CHOICES IN THE CONSTRUCTION OF MULTIREGIONAL LIFE TABLES

Jacques Ledent
Philip Rees

November 1980
WP-80-173

Appearing simultaneously as WP 289,
School of Geography, University of Leeds,
Leeds LS2 9JT, U.K.

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

FOREWORD

Interest in human settlement and systems policies has been a central part of urban-related work at the International Institute for Applied Systems Analysis (IIASA) from the outset. From 1975 through 1978 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 work that might apply the newly developed mathematical methodology to other research topics.

This paper is a result of the continuing collaborative work being carried out by IIASA scholars and the Migration and Settlement network. In it, Jacques Ledent of HSS and Philip Rees of the University of Leeds, U.K., consider several important issues connected with the construction of multiregional life tables. They focus, in particular, on problems revolving around the three approaches proposed to date for calculating the probability matrices that are the starting point for all applied life-table analyses: the movement, the transition, and the hybrid methods of estimation. They conclude that transition data should be used rather than movement data, where available, and that transition methods rather than hybrid methods should be applied to the transition data.

Selected papers summarizing previous work on migration and settlement at IIASA are listed at the back of this paper.

Andrei Rogers
Chairman
Human Settlements
and Services Area

ACKNOWLEDGMENTS

The authors wish to thank Andrei Rogers for bringing them together to work on a topic of mutual concern in which he has played such an important and pioneering role. Our thanks also go to Paul Drewe, author of the Migration and Settlement research report on the Netherlands (Drewe 1980) for allowing us to use some of the data contained in the report. Philip Rees wishes to thank the staff at IIASA for providing accommodation, services, and hospitality so conducive to thinking and writing. Finally, we both are grateful for Susanne Stock's skill and effort in typing the paper.

ABSTRACT

The methods of multiregional life table construction are explored through an investigation of a tree of choices with respect to approach, data, rates definition, probability definition, and stationary population or life-years-lived calculation. Two principal approaches are discussed: the movement approach and the transition approach, although a third label "hybrid approach" is used to characterize many of the developments in the field to date. Methods are applied to two population systems, that of the Netherlands, where moves data are available, and that of Great Britain, where the data on migration come in the transition form. The discussion of methods and the computer runs of the life table model lead to a clear set of recommended choices for the would-be life table constructor. The fairly simple and direct transition approach using migrant data measured by a five-year question is our preferred choice. The paper ends by speculating on how solutions to the unresolved problems of multiregional population analysis might be sought.

CONTENTS

CHOICES IN THE CONSTRUCTION OF MULTIREGIONAL LIFE TABLES

1. PURPOSE OF THE PAPER

Over the past five years or so (1975-79) more than a dozen scholars from national member organizations of the International Institute for Applied Systems Analysis have been engaged in collaboration with IIASA researchers in a multinational project to describe and analyze the pattern of population movement and change among the regions of their countries (Rogers 1976; Willekens 1978). One of the analytic tools used in these studies has been the multiregional life table (Rogers 1973a, 1975a; Willekens and Rogers 1978). Many different types of migration data have been employed as input to the multiregional life table program and a variety of different methods of converting these migration data and associated mortality data into the probabilities needed in the life table have been suggested (Rogers 1973a, 1975a; Rees and Wilson 1975, 1977; Rogers and Ledent 1976; Ledent 1978, 1980a; Rees 1978, 1980a).

The purpose of this paper is to review systematically the consequences of using different data types, different rate definitions, and different probability estimation methods. We do this by applying different methods to the same data set and examining the resulting effects; we adopt different data sets and the same methods and examine the results. Out of this experimentation with Dutch and British data emerge recommendations as to the most reliable methods to use given available data, the most reliable data to seek out, and the data set which should ideally be collected and the methods which should ideally be applied.

In a sense our search for optimum data and an optimum methodology resembles the search of the surfer for the perfect wave. There is always a better wave to be found on the next beach on the next day. However, given the waves and beaches available today choices have to be made, and our series of experiments are intended to inform future constructors of multiregional life tables.

However, although our experiments are with multiregional population systems, most of our methods and conclusions carry over into other multistate population systems: the movement of people between marital status states, the transfers of workers into and out of the labor force, the movement of pupils and students through an educational system.

We first outline by way of reminder the principal steps involved in constructing a multiregional regional life table and identify the steps about which choices can be made.

Two general approaches to multiregional life table construction can be distinguished; on the one hand, the *movement* approach, initially proposed in the context of marital status analysis by Schoen and Nelson (1974) and Schoen (1975) and further developed by Rogers and Ledent (1976) and Schoen and Land (1979) and, on the other hand, the *transition* approach suggested by Rogers (1975a) and Rees and Wilson (1975, 1977) and developed and implemented by Ledent (1978, 1980a, 1980b). However, both these approaches were predated by and developed in part from the original work of Rogers (1973a, 1973b, 1975a) in which rate and probability definition equations follow, by and large, the logic of the movement approach but in which the transfer data used consist of changes of residence (numbers of migrants) as in the transition approach. Table 1 shows the concepts and data associated with each approach, and why we have elected to call the earlier stream of work the *hybrid* approach.

Table 1. The approaches to multiregional life table construction and their origin.

Concepts \ Data	Moves	Transitions
Movement	MOVEMENT APPROACH Schoen and Nelson (1974) Schoen (1975) Rogers and Ledent (1976) Schoen and Land (1979)	HYBRID APPROACH Rogers (1973a, 1973b, 1975a)
Transition		TRANSITION APPROACH Rogers (1975a) Rees and Wilson (1975, 1977) Ledent (1978, 1980a, 1980b)

Within these three broad approaches a variety of subapproaches or choices are explored (as in Schoen and Land 1979), and in the final section of the paper, recommendations are made about selecting the best methods and best data for multiregional life table construction, within the constraints that face the researcher.

Throughout the paper we consider national population systems (Holland, Great Britain) closed off from the rest of the world though clearly such systems are unrealistic. None of the methodological points made are seriously affected, we believe, by this action though ideally flows of people to and from the rest of the world should have been included.

2. MULTIREGIONAL LIFE TABLES: A REMINDER

2.1 The General Steps

The steps in constructing a life table that computes from current migration and mortality behavior the likely life history of people born in different states are outlined in Rogers (1975a) and in Willekens and Rogers (1978). They are as follows.

(1) DATA: The necessary data on populations migration, regional mortality and regional fertility* are assembled and/or estimated in one of a selection of ways (see Willekens and Rogers 1978; Willekens, Por and Raquillet 1979; Rees 1980).

(2) RATES: Migration, mortality or survivorship rates are computed from the assembled data in one of a variety of ways.

(3) PROBABILITIES: Probabilities of interregional transition are computed from the observed rates in one of a diversity of methods. The transitions are from one exact age to another (age x to age x + n), and involve "staying and survival" or "migration and survival" or "non-survival". In some cases, the "non-survival" transition may be broken down into "staying and non-survival" or "migration and non-survival". In most instances, the formulae expressing the transition probabilities in terms of the input rates follow from a specific assumption relating to the derivation of the number of years lived between two consecutive ages. There are however exceptions in the case of the transition approach in which the linear and cubic spline interpolation methods for estimating the age-specific probabilities do not preclude any method for calculating the aforementioned numbers of person-years lived.

---

*Not strictly speaking necessary for life table computation per se.

(4) BASIC LIFE TABLE STATISTICS: The following life table statistics may then be generated (Willekens and Rogers 1978:21-23):

1. life history of a regional birth cohort
2. number of survivors at exact age x
3. number of years lived between two consecutive ages, or, the age composition of (the) stationary population.

(5) DERIVED LIFE TABLE STATISTICS: From the first three sets of statistics a further five follow (Willekens and Rogers 1978:23):

4. number of years lived beyond age x
5. life expectancies by region of birth
6. life expectancies by region of residence
7. migraproduction rates
8. survivorship rates.

At each of these steps except the last, step (5), these are alternative methods or choices, and we will analyze the consequence of choices within a step on the results in subsequent steps, particularly on the life table statistics produced at step (5).

2.2 Definitions

The variables employed in the multiregional life table model are defined where they are introduced. However, since our notation differs somewhat from that of Rogers (1975a) or Willekens and Rogers (1978), the principles behind the notation are explained in Appendix 1, to which the reader can turn when further clarification is required. In particular, the role of that "jack of all trades", the subscript x for age, is explained.

2.3 The Multiregional Life Table Functions

These are stated here to remind the reader about how multiregional life tables are constructed. However, it must be stressed that the equations are only one of a family of alternatives that will be discussed.

First, rates of mortality and mobility are computed

$$M_x^{\delta(i)} = D_x^{\delta(i)}/K_x^i \tag{1}$$

$$M_x^{ij} = D_x^{ij}/K_x^i \tag{2}$$

where $M_x^{\delta(i)}$ is the observed rate for a T-year period at which persons die in region i between age x and x+n, $D_x^{\delta(i)}$ is the number of deaths in region i to persons aged x to x+n at death, K_x^i is the population risk in region i aged x to x+n (normally the mid-period population or average of initial and final populations in a period). In equation (2), M_x^{ij} is the observed rate at which people move between regions i and j while aged x to x+n, and D_x^{ij} is the number of such moves or displacements. Then, probabilities of transition are worked out

$$p_x^{ij} = \frac{\frac{n}{T} M_x^{ij}}{1 + \frac{n}{2T} M_x^{\delta(i)} + \frac{n}{2T} \sum\limits_{k \neq i} M_x^{ik}} \tag{3}$$

The variable p_x^{ij} is the probability that a person attaining age x in region i will survive at age x + n in region j, where n is the age interval adopted in the life table, T is the length of the time period to which the data refer and $\frac{n}{T}$ is the ratio of age interval of the model to time period of the data required to make the two equivalent in the model. Equation (3) is referred to as the "Option 1" method in Rogers (1975a:82) and Willekens and Rogers (1978:51-52).

Probabilities of non-survival for persons in region i are given by

$$p_x^{i\delta} = \frac{\frac{n}{T} M_x^{\delta(i)}}{1 + \frac{n}{2T} M_x^{\delta(i)} + \frac{n}{2T} \sum\limits_{k \neq i} M_x^{ik}} \tag{4}$$

where $p_x^{i\delta}$ is the probability that a person attaining age x in region i will die before attaining age x+n.

The probabilities of survival and staying are found by subtraction

$$p_x^{ii} = 1 - \sum_{k \neq i} p_x^{ik} - p_x^{i\delta} \tag{5}$$

These probabilities of transition are then applied to the regional cohorts at birth, ${}_0^i\ell$ (age 0 in region i or birth in region i) which may be allocated values of 1, so that subsequent statistics have a "per unit born" or "probability-like" interpretation, or values of say 100,000, when subsequent statistics will have the interpretation of "numbers in a hypothetical cohort". For convenience, we adopt here the "probability" interpretation.

In general, the transition history of the hypothetical cohort is traced out as follows

$${}_0^j\ell_x^{ik} = {}_0^j\ell_x^i \; p_x^{ik} \tag{6}$$

$${}_0^j\ell_{x+n}^k = \sum_i {}_0^j\ell_x^{ik} \tag{7}$$

$${}_0^j\ell_x^{i\delta} = {}_0^j\ell_x^i \; p_x^{i\delta} \tag{8}$$

where ${}_0^j\ell_x^i$ is the probability that persons born in region j attain age x in region i and ${}_0^j\ell_x^{ik}$ is the probability that those persons make a transition to region k at age x+n (that is, are in region k at age x+n, n years after birthday x). Equation (7) gathers together these transitions to obtain the probability that persons born in region j will be in region k at age x+n. The probability of dying before reaching the x+n-th birthday, given you were born in region j and were present age x in region i, is estimated through equation (8). Note that equation (6) contains a key assumption: that the probability of making an interregional transition between regions i and k is independent of region of birth. This is the normal assumption that has to be made to construct a multiregional life table but Ledent (1980c) has shown that if we can replace equation (6) by

$$ {}^{j}_{0}\ell^{ik}_{x} = {}^{j}_{0}\ell^{i}_{x}\ {}^{j}_{0}p^{ik}_{x} \tag{6'} $$

where ${}^{j}_{0}p^{ik}_{x}$ is the probability that a person born in region j will make a region i to region k transition in age interval x to x+n, then we must obtain a better estimate of people's life histories.

It is convenient at this point in the exposition to move to a matrix notation and rewrite (6) and (7) as

$$ \underset{\sim}{\ell}_{x+n} = \underset{\sim}{p}_{x}\ \underset{\sim}{\ell}_{x} \tag{9} $$

where

$$ \underset{\sim}{\ell}_{x} = [{}^{j}_{0}\ell^{i}_{x}] \quad \text{and} \quad \underset{\sim}{p}_{x} = [p^{ik}_{x}] $$

Then, the matrix $L_x = [{}^{ij}_{0}L^{\cdot}_{x}]$ of number of life years spent in region i by persons born in region j, is obtained from

$$ \underset{\sim}{L}_{x} = \tfrac{1}{2}[\underset{\sim}{\ell}_{x} + \underset{\sim}{\ell}_{x+n}] \tag{10} $$

The total number of years lived beyond age x by persons newly born, $\underset{\sim}{T}_{x}$ in matrix form, is defined as

$$\underset{\sim}{T}_x = \sum_{y=x}^{z} \underset{\sim}{L}_y \tag{11}$$

where z denotes the last age group (z years and over). Life expectancy can be defined (Willekens and Rogers 1978:40,42) either by place of residence

$${}_x\underset{\sim}{e}_x = \underset{\sim}{T}_x(\underset{\sim}{\ell}_x)^{-1} \tag{12}$$

or by place of birth

$${}_0\underset{\sim}{e}_x = \underset{\sim}{T}_x[\underset{\sim}{\ell}_x]_d^{-1} \tag{12'}$$

where ${}_x\underset{\sim}{e}_x$ refers to a matrix of life expectancies beyond age x conditional on region of residence at age x, where ${}_0\underset{\sim}{e}_x$ is a matrix of life expectancies beyond age x conditional on region of birth; where $[\underset{\sim}{\ell}_x]_d$ is a diagonal matrix with elements of the vector $\{\underset{\sim}{i}\}'\underset{\sim}{\ell}_x$ in the diagonal, $\{i\}'$ being a row vector of ones.

A set of products of the multiregional life table model alternative to the statistics associated with the expected number of years lived in various regions by regional cohorts are the statistics on numbers of migrations that particular regional cohorts expect to make. Among the statistics that are produced are the gross migraproduction rate, the generalized net mobility function, and the net migraproduction rate. These rates (defined in Rogers 1975b and incorporated in the multiregional life table program in Willekens and Rogers 1978:99-117) measure in various ways the numbers of migrations out of particular regions expected to be made by regional birth cohorts.

The gross migraproduction rate matrix, $\underset{\sim}{\mathrm{GMR}}$, is defined as

$$\underset{\sim}{\mathrm{GMR}} = n \sum_{x=0}^{z} [\underset{\sim}{M}_x^0]_d \tag{13}$$

where $[\underset{\sim}{M}_x^0]_d$ is a matrix of total outmigration rates from region 1 to r arranged along the principal diagonal with zeroes elsewhere.

The general mobility function is computed as

$$\underset{\sim}{Y}_x = [\underset{\sim}{M}_x^0]_d \, \underset{\sim}{L}_x \tag{14}$$

and the net migraproduction rate matrix is simply a sum of the age-specific mobility matrices over all ages

$$\underset{\sim}{\mathrm{NMR}} = \sum_{x=0}^{z} \underset{\sim}{Y}_x \tag{15}$$

The distinction between the expectation-of-life and migraproduction-rate statistics is an important one to keep in mind as the choices in life table construction are reviewed in the paper. Certain choices will be preferred if good time spent or life expectancy statistics are what is sought; other choices will be recommended if good migraproduction statistics are the goal.

The final product of the multiregional life table model, which we consider here, are the survivorship rates (also called survivorship proportions in the literature). Normally, these are computed once the life table stationary populations $\underset{\sim}{L}_x$ are known (Rogers and Ledent 1975; Rogers 1975a)

$$\underset{\sim}{S}_x = \underset{\sim}{L}_{x+n} \, [\underset{\sim}{L}_x]^{-1} \tag{16}$$

although we shall show later that it is possible to change the order of calculation here and compute the $\underset{\sim}{L}_x$ variables given observed values for the $\underset{\sim}{S}_x$ variables. The survivorship rates are used as part of the input to population projections.

2.4 Choices: Alternative Approaches to the Estimation of Multiregional Life Tables

What choices face the person who wishes to construct a multiregional life table? The alternatives can be conveniently viewed in the form of a table, the columns of which refer to the steps in the construction of life tables, and the rows to the three approaches we distinguish (Figure 1).

The first choice is between data types. Either data on events taking place at given points in time may be used or data on changes in a person's characteristics between two points in time. The first type of data we term "movements" data, the second "transitions" data. Movements data include deaths, births and migrations; transitions data include non-survivors, newly-born infants and migrants. Occasionally, transition data may be classified by region of birth, although we know only of one country (United States) in which the necessary tabulations have been produced.

The second choice is between periods over which the data are collected. Movements data are collected usually in registration systems for annual periods. They can be used directly with a single year of age classification of the population or if this is too fine, the one year data can be either aggregated over a large period or suitably multiplied by $\frac{n}{T}$ where n is the age interval used in the life table model and T the period for which data have been collected.

Transitions data derive from periodic national censuses in which the two points in time are a fixed interval apart, usually one to ten years apart. In the U.K. census of 1961 (and 1981), a one year question only was employed; in the Soviet census of 1970 a two year question was utilized; in the Australian

Figure 1. Choices in the construction of multiregional life tables.

Type of Data	Approach	Data Period and Age Interval	Rates Derived from Data: Mortality	Rates Derived from Data: Mobility	Probabilities of Survival and/or Stationary Population
Moves	Movement	not important	conventional	conventional	• linear ("Option 3") • exponential • cubic • iterative-interpolative (with linear or exponential methods)
Transitions	Hybrid	$T < n$	conventional	assuming direct equivalence of age group data	• linear ("Option 1") • exponential • movement-based linear • movement-based exponential
		$T = n$		averaging over successive age group data	
	Transition	$T < n$	directly estimated (from census data or accounts)		• linear/direct ("Option 2") • estimation of $\underset{\sim}{p}_x$: linear or cubic spline interpolation method; estimation of $\underset{\sim}{L}_x$: linear or direct method
		$T = n$	conditionally estimated		

Notes to Figure 1

1. "Option 1", "Option 2", and "Option 3" are the titles given to various probability estimation equations in Rogers (1975a) and Willekens and Rogers (1978).

2. T = length of time period to which mortality and migration data refer
   n = the age interval used in the life table model

3. ⟷ indicates that the two choices under separate columns are linked of necessity.

4. The choices are defined in the following sections of the paper.

<table>
<tr><th rowspan="2">Type of Data</th><th rowspan="2">Approach</th><th rowspan="2">Period and Interval</th><th colspan="2">Rates</th><th rowspan="2">Probab-ilities</th><th rowspan="2">Life Years</th></tr>
<tr><th>Mort.</th><th>Mob.</th></tr>
<tr><td>3.</td><td>3.</td><td>--</td><td>--</td><td>--</td><td colspan="2">3.3.1
3.3.2
3.3.3
3.3.4 | 3.3.5</td></tr>
<tr><td rowspan="2">4. & 5.</td><td>4.</td><td>4.4
4.5</td><td>--</td><td>4.3</td><td colspan="2">4.2.2
4.2.4
4.2.1
4.2.2</td></tr>
<tr><td>5.</td><td>5.8</td><td colspan="2">5.2.4
5.2.5</td><td>5.3.1
5.3.2
5.3.3</td><td>5.4</td></tr>
</table>

census of 1971, the Canadian census of 1971 and the U.S. censuses of 1940, 1960, and 1970 a five year question was used; in the French censuses of 1962, 1968 and 1975 eight, six and seven year questions were posed, respectively, in order to link together successive censuses. In certain censuses information about more than one period of observation is available: in the U.K. censuses of 1966 and 1971, in the Japanese census of 1970, and in the Australian census of 1976 both one year and five year migrant data are tabulated. In the majority of situations, therefore, the researcher has no choice of period length. However, when transition data are employed, period length does turn out to have a considerable effect on our estimate of the regional distribution of the life expectancies of babies born in the various regions, whereas when movement data are employed period length has no effect as long as migration patterns remain stable.

The third set of choices involves the definition of rates: mortality and mobility rates in the movement and hybrid approaches, survivorship and non-survivorship rates in the transition approach. In the movement approach conventional definitions of mortality and mobility rates are used, and so no choices are involved, unless the movement data are classified by both age and birth cohort, when improvements to conventional methods are possible. In the hybrid and transition approaches there are alternative ways of computing rates, and we make a careful analysis of the alternatives.

The fourth choice that needs to be made concerns the probability definition equation. The choices involve the "options" defined by Rogers (1975a), Rogers and Ledent (1976) and Willekens and Rogers (1978), the modifications developed by Ledent (1978, 1980a, 1980b) and suggestions made in Rees and Wilson (1975, 1977) and Rees (1980a).

Each method has its advantages and disadvantages: there is a fair measure of agreement between different methods for normal, low mobility and mortality situations; where mobility is very high, probabilities do differ significantly.

Finally, there are a set of choices which carry over from the conventional life table concerning the function assumed for the life years lived/stationary population. In the case of the movement approach, these generally follow directly from the choices made for the probabilities, except that a choice is possible if the interpolative-iterative method is used. In the case of the hybrid approach the life years or stationary population equations are all directly linked with choices of probability estimation equation. This is also true for the first choice of probability equation in the transition approach but with the second and third alternatives either of two stationary population equations can be used.

3. THE MOVEMENT APPROACH

3.1 General Characteristics

The movement approach to multistate life tables grows naturally out of conventional life table methods. The movements considered in the conventional life table are those into life-births and those out of life-deaths. In the multiregional life table model further movements, into and out of regions, are incorporated. However, this introduces, as Ledent (1980) points out, a severe methodological problem:

> *The fact is that any model of the life-table type is a transition model: that is,... moves have to be transformed into transitions.* (Ledent 1980a:548)

However, none of the methods we discuss or suggest effect this transformation satisfactorily, since we would need statistics such as the average number of interregional moves per inter-regional transition to use as divisors of moves. Such statistics are unavailable unless a country measures both movements through a registration system and asks a retrospective migration question of the right kind in the national periodic census.

Why has this problem not been encountered before in life table work? The reason is fairly simple. In conventional life tables the number of movements always equals the number of transitions, that is:

the number of births (moves from preexistence to existence)	=	the number of transitions from preexistence to existence or the number of persons making those transitions (infants born)

and

the number of deaths (moves from existence to post-existence)	=	the number of transitions from existence to post-existence or the number of persons making those transitions (the non-survivors)

with one or two significant individual exceptions.

This problem of moves-transitions inequality undoubtedly occurs in other multistate population systems such as marital status classified populations but hithertofore has not come to light because over short periods (say, of a year) the inequality was probably small. However, legal changes in divorce law in many countries have increased the possibilities of multiple moves within short time periods. Such an inequality is so obvious and serious in the case of employment classified populations that substantial surveys have been mounted to measure the numbers of persons continuously in unemployment, and to test the claim that rising unemployment consists of a larger number of short spells rather than the lengthening of the spells of unemployment of persons already unemployed.

If interest is focussed on assessing the numbers of moves that are likely to occur in a population, then counter-arguments in favor of the transition approach can be mounted, although in as much as the stationary population statistics involved in migraproduction functions will be biased, so will those statistics themselves.

3.2 General Equations

Formally, the movement approach is characterized by two sets of equations--flow and orientation equations--originally proposed by Schoen and Nelson (1974) and Schoen (1975) in the context of marital status analysis and later generalized by Rogers and Ledent (1976); the generalization pertains to the introduction a second subscript referring to the place-of-birth.

The flow equations may be defined as

$$ {}_0^j\ell_{x+n}^i = {}_0^j\ell_x^i - \sum_{\substack{k=1 \\ k\neq i}}^{r} {}_0^jd_x^{ik} - {}_0^jd_x^{\delta(i)} + \sum_{\substack{k=1 \\ k\neq i}}^{r} {}_0^jd_x^{ki} \qquad (17) $$

where ${}_0^jd_x^{ik}$ terms are the life table displacements from region i to region k that occur to members of the born-in-region j cohort in the age interval x to x+n, and the ${}^{j}_{0}d^{\delta(i)}_{x}$ are the numbers of deaths that occur in region i to persons in the region j birth cohort between ages x and x+n. So the first set of decrements from ${}^{j}_{0}\ell^{i}_{x}$ are the estimated out-migrations, the second set of decrements are the deaths and the set of increments are the estimated in-migrations all occurring to the life table cohort born in region j. This equation is the life table equivalent to the simple components-of-growth equation frequently used to estimate regional populations.

The orientation equations may be defined as

$$ {}^{j}_{0}d^{ik}_{x} = m^{ik}_{x} \; {}^{ij}_{0}L^{\cdot}_{x} \tag{18}$$

$$ {}^{j}_{0}d^{\delta(i)}_{x} = m^{\delta(i)}_{x} \; {}^{ij}_{0}L^{\cdot}_{x} \tag{19}$$

which contain the assumption that the life table rates of moving between region i and region k at age x, m^{ik}_{x}, and the life table rates of dying in region i are independent of place of birth. If movement statistics are classified by place of birth of the person making the move, then this restriction can be relaxed. Normally, however, we assume

$$ m^{ik}_{x} = M^{ik}_{x} \tag{20}$$

$$ m^{\delta(i)}_{x} = M^{\delta(i)}_{x} \tag{21}$$

where the capital M's refer to the observed rates corresponding to the left-hand side life table rates. The observed rates were defined earlier in equations (1) and (2).

Let us now substitute the right-hand sides of equations (18) and (19) for the displacements variables in the right-hand side of (17), rearranging the ℓ-variables at the same time and multiplying both sides by -1

$$ {}^{j}_{0}\ell^{i}_{x} - {}^{j}_{0}\ell^{i}_{x+n} = \sum_{\substack{k=1 \\ k\neq i}}^{r} m^{ik}_{x} \; {}^{ij}_{0}L^{\cdot}_{x} + m^{\delta(i)}_{x} \; {}^{ij}_{0}L^{\cdot}_{x} $$

$$ - \sum_{\substack{k=1 \\ k\neq i}}^{r} m^{ki}_{x} \; {}^{kj}_{0}L^{\cdot}_{x} \tag{22} $$

Now by using the device invented by Rogers and Ledent (1976) of arranging all the m rates in a matrix thus:

$$ \underset{\sim}{m}'_{x} = \begin{bmatrix} m^{\delta(1)}_{x} + \sum\limits_{k\neq 1} m^{1k}_{x} & -m^{12}_{x} & \cdots & -m^{1r}_{x} \\ -m^{21}_{x} & m^{\delta(2)}_{x} + \sum\limits_{k\neq 2} m^{2k} & \cdots & -m^{2r}_{x} \\ \vdots & \vdots & \ddots & \vdots \\ -m^{r1}_{x} & -m^{r2}_{x} & & m^{\delta(k)}_{x} + \sum\limits_{k\neq r} m^{rk} \end{bmatrix} \tag{23} $$

we can re-express equation (22) as a matrix equation

$$ \underset{\sim}{\ell}_{x} - \underset{\sim}{\ell}_{x+n} = \underset{\sim}{m}_{x} \underset{\sim}{L}_{x} \tag{24} $$

where

$$ \underset{\sim}{\ell}'_{x} = \begin{bmatrix} {}^{1}\ell^{1}_{x} & \cdots & {}^{1}\ell^{r}_{x} \\ \vdots & & \vdots \\ {}^{r}\ell^{1}_{x} & \cdots & {}^{r}\ell^{r}_{x} \end{bmatrix} \tag{25} $$

and

$$\underset{\sim}{L}_x' = \begin{bmatrix} {}^{11}L_x & \cdots\cdots & {}^{1r}L_x \\ \vdots & & \vdots \\ {}^{r1}L_x & \cdots\cdots & {}^{rr}L_x \end{bmatrix} \tag{26}$$

The transpositions are used to preserve the original superscript ordering used in our algebraic equations.

To compute the $\underset{\sim}{L}_x$ matrix the equation (24) is rearranged

$$\underset{\sim}{L}_x = \underset{\sim}{m}_x^{-1}(\underset{\sim}{\ell}_x - \underset{\sim}{\ell}_{x+n}) \tag{27}$$

Note that the above reasoning also applies to the last age group (whose length is w - z where w is the maximal age one can reach); equation (24) still holds but, since the second term on the left-hand side is zero, it simply becomes (Ledent 1978)

$$\underset{\sim}{\ell}_z = \underset{\sim}{m}_z \underset{\sim}{L}_z \tag{28}$$

so that $\underset{\sim}{L}_z$ can be obtained from

$$\underset{\sim}{L}_z = \underset{\sim}{m}_z^{-1} \underset{\sim}{\ell}_z \tag{29}$$

The sequence of multiregional life table statistics then follows once $\underset{\sim}{L}_x$ is known, as outlined in an earlier section.

3.3 Applied Calculation

To derive the values for the $\underset{\sim}{p}_x$, $\underset{\sim}{\ell}_x$ and $\underset{\sim}{L}_x$ variables, some four methods have been suggested

1) the linear method
2) the exponential method
3) the cubic method
4) the iterative-interpolative method

Each of these is discussed in turn.

3.3.1 The Linear Method ("Option 3")

Here the assumption is made that the stationary population is a simple average (linear integration) of $\underset{\sim}{\ell}_x$ and $\underset{\sim}{\ell}_{x+n}$ matrices:

$$L_x = \frac{n}{2}[\underset{\sim}{\ell}_x + \underset{\sim}{\ell}_{x+n}] \tag{30}$$

where n is the age group interval (equal to the period length). This is simply a restatement of equation (9). Now we can substitute for $\underset{\sim}{L}_x$ in equation (24) from the right-hand side of equation (30) and also replace $\underset{\sim}{m}_x$ by the observed $\underset{\sim}{M}_x$ [the assumptions of equations (20) and (21)] to yield

$$\underset{\sim}{\ell}_x - \underset{\sim}{\ell}_{x+n} = \underset{\sim}{M}_x \frac{n}{2}[\underset{\sim}{\ell}_x + \underset{\sim}{\ell}_{x+n}] \tag{31}$$

Then if we multiply out the right-hand side thus,

$$\underset{\sim}{\ell}_x - \underset{\sim}{\ell}_{x+n} = \underset{\sim}{M}_x \frac{n}{2} \underset{\sim}{\ell}_x + \underset{\sim}{M}_x \frac{n}{2} \underset{\sim}{\ell}_{x+n} \tag{32}$$

regroup all terms involving $\underset{\sim}{\ell}_x$ on the left-hand side and terms involving $\underset{\sim}{\ell}_{x+n}$ on the right-hand side, we obtain

$$\underset{\sim}{\ell}_x - \underset{\sim}{M}_x \frac{n}{2} \underset{\sim}{\ell}_x = \underset{\sim}{M}_x \frac{n}{2} \underset{\sim}{\ell}_{x+n} + \underset{\sim}{\ell}_{x+n}$$

$$(\underset{\sim}{I} - \frac{n}{2} \underset{\sim}{M}_x)\underset{\sim}{\ell}_x = (\underset{\sim}{I} + \frac{n}{2} \underset{\sim}{M}_x)\underset{\sim}{\ell}_{x+n}$$

so that (as given in Rogers and Ledent 1976)

$$\underset{\sim}{\ell}_{x+n} = (\underset{\sim}{I} + \frac{n}{2} \underset{\sim}{M}_x)^{-1}(\underset{\sim}{I} - \frac{n}{2} \underset{\sim}{M}_x)\underset{\sim}{\ell}_x \tag{33}$$

Because $(\underset{\sim}{I} + \frac{n}{2} \underset{\sim}{M}_x)$ and $(\underset{\sim}{I} - \frac{n}{2} \underset{\sim}{M}_x)$ are commutative, we can also write equation (33) as

$$\underset{\sim}{\ell}_{x+n} = (\underset{\sim}{I} - \frac{n}{2} \underset{\sim}{M}_x)(\underset{\sim}{I} + \frac{n}{2} \underset{\sim}{M}_x)^{-1} \underset{\sim}{\ell}_x \tag{34}$$

as derived in Ledent (1978), where it is shown that the linear formula for $\underset{\sim}{p}_x$ is equivalent to assuming that movements out of or into a state are evenly distributed over an age/time interval (Ledent 1978:42fn). So that, we can replace the probabilities matrix in equation (25) by a matrix expression involved the observed rates of mortality and mobility:

$$\underset{\sim}{p}_x = (\underset{\sim}{I} + \frac{n}{2} \underset{\sim}{M}_x)^{-1}(\underset{\sim}{I} - \frac{n}{2} \underset{\sim}{M}_x) \tag{35}$$

or if the rates are not annual rates we should modify this to

$$\underset{\sim}{p}_x = (\underset{\sim}{I} + \frac{n}{2T} \underset{\sim}{M}_x)^{-1}(\underset{\sim}{I} - \frac{n}{2T} \underset{\sim}{M}_x) \tag{36}$$

In the methods and programs monograph by Willekens and Rogers (1978), this is referred to as the "Option 3" method, and this is its title in the p-definition column, movement method row in Figure 1.

As for the probabilities of non-survival $p_x^{i\delta}$, they can be assembled in a vector $\{p_x^{\delta}\}$ which can be obtained as a residual from

$$\{p_x^{\delta}\} = \{i\} - \underset{\sim}{p}_x'\{i\} \tag{37}$$

where $\{i\}$ is a column vector of ones

$\underset{\sim}{p}_x'$ is the transpose matrix of $\underset{\sim}{p}_x$

In fact, these probabilities can be further disaggregated to account for the place in which the deaths actually occur. It has been shown (Ledent 1978) that the matrix $\underset{\sim}{p}_x^{\delta}$--whose (i,j)-th element is the probability of dying in region i for an individual aged x in region j--can be derived from

$$\underset{\sim}{p}_x^{\delta} = n\, \underset{\sim}{M}_x^{\delta}[\underset{\sim}{I} + \frac{n}{2}\, \underset{\sim}{M}_x]^{-1} \tag{38}$$

where $\underset{\sim}{M}_x^{\delta}$ is a diagonal matrix of the regional death rates.

3.3.2 The Exponential Method

An alternative starting point for the calculation of $\underset{\sim}{p}_x$'s and $\underset{\sim}{L}_x$'s is that of assuming that the instantaneous forces of mortality and mobility are equal to the observed discrete counterparts over an age interval: that is,

$$\underset{\sim}{\mu}(y) = \underset{\sim}{M}_x \tag{39}$$

for all y such that $x \le y < x+n$ where $\underset{\sim}{\mu}(y)$ is a matrix of instantaneous rates of mortality $\mu^{i\delta}(y)$ and mobility $\mu^{ij}(y)$, functions of continuous age y, arranged in the same fashion as $\underset{\sim}{M}_x$ [whose transpose was defined in equation (23)]. Krishnamoorthy (1979) and Schoen and Land (1979) have shown that this assumption leads to the following expression for the $\underset{\sim}{p}_x$ probabilities

$$\underset{\sim}{p}_x = \exp(-n\ \underset{\sim}{M}_x) \tag{40}$$

The right side expression is evaluated by using matrix equivalent of the Taylor expansion for computing e^{-x}

$$\underset{\sim}{p}_x = \underset{\sim}{I} - n\ \underset{\sim}{M}_x + \frac{n^2}{2!}\ \underset{\sim}{M}_x^2 - \frac{n^3}{3!}\ \underset{\sim}{M}_x^3 + \frac{n^4}{4!}\ \underset{\sim}{M}_x^4 - \ldots \tag{41}$$

with as many terms being used in the computing algorithm to give $\underset{\sim}{p}_x$ probabilities accurate to the sixth decimal place or 10^{-6}. Note that

$$\underset{\sim}{p}_x^{\delta} = \underset{\sim}{M}_x^{\delta}\ \underset{\sim}{M}_x^{-1}[\underset{\sim}{I} - \exp(-n\ \underset{\sim}{M}_x)] \tag{42}$$

To estimate the $\underset{\sim}{L}_x$ matrix in this method, the $\underset{\sim}{m}_x$ terms in equation (24) are replaced by the observed rates

$$\underset{\sim}{L}_x = \underset{\sim}{M}_x^{-1}(\underset{\sim}{\ell}_x - \underset{\sim}{\ell}_{x+n}) \tag{43}$$

although occasionally computation of the inverse of $\underset{\sim}{M}_x$ does give problems when the mortality rates are low and need to be carefully checked.

3.3.3 The Cubic Method

As an alternative to the linear and exponential methods, one can use the cubic method proposed by Schoen and Nelson (1974) and further extended by Ledent (1978); again the extension pertains to the introduction of a second subscript relating to the place of birth. The integration of $\underset{\sim}{L}_x$ is carried out by fitting a curve of degree three through four successive values (the conventional life table version is explained and derived in Keyfitz 1968) $\underset{\sim}{\ell}_{x-n}$, $\underset{\sim}{\ell}_x$, $\underset{\sim}{\ell}_{x+n}$, and $\underset{\sim}{\ell}_{x+2n}$:

$$\underset{\sim}{L}_x = \frac{13n}{24}[\underset{\sim}{\ell}_x + \underset{\sim}{\ell}_{x+n}] - \frac{n}{24}[\underset{\sim}{\ell}_{x-n} + \underset{\sim}{\ell}_{x+2n}] \tag{44}$$

with slight modifications of this formula for the first, second, and last but one last ages:

$$\underset{\sim}{L}_0 = \frac{n}{2}\underset{\sim}{\ell}_0 + \frac{13n}{24}\underset{\sim}{\ell}_n - \frac{n}{24}\underset{\sim}{\ell}_{2n} \tag{45}$$

$$\underset{\sim}{L}_n = \frac{n}{2}(\underset{\sim}{\ell}_n + \underset{\sim}{\ell}_{2n}) \tag{46}$$

$$\underset{\sim}{L}_{z-n} = \frac{n}{2}(\underset{\sim}{\ell}_{z-n} + \underset{\sim}{\ell}_z) \tag{47}$$

Note that the special treatment of these age groups follows the procedure used in the application of the analogous method used in the construction of an ordinary life table (Keyfitz 1968). To derive the $\underset{\sim}{\ell}_x$ values either the linear or the exponential method can be used to give initial values which are input to equation (44) and then used in a rearrangement of equation (24)

$$\underset{\sim}{\ell}_{x+n} = \underset{\sim}{\ell}_x - \underset{\sim}{M}_x \underset{\sim}{L}_x \tag{48}$$

with $\underset{\sim}{M}_x$ substituted for $\underset{\sim}{m}_x$ for computational purposes, to give fresh $\underset{\sim}{\ell}_x$ estimates. The procedure is repeated until satisfactory $\underset{\sim}{\ell}_x$ values are obtained (that is, when the sixth decimal place value does not change with successive iterations).

If the probability matrix values are required they may be obtained from

$$\underset{\sim}{p}_x = \underset{\sim}{\ell}_{x+n}\, \underset{\sim}{\ell}_x^{-1} \tag{49}$$

using the $\underset{\sim}{\ell}_x$ values that have been generated from equations (44) through (48). If some of the states employed are initially empty (which will virtually never be the case in multiregional applications), the assuming radices of 1 for those states will ensure that the computations (inversion of $\underset{\sim}{\ell}_x$) can be carried out.

3.3.4 The Iterative-Interpolative Method

The linear, exponential and cubic methods rely on the assumption that the life table rates are equal to the observed rates. This assumption is not tested. However, Keyfitz (1966, 1968) has developed "life table that iterates to the data"; Oechsli (1972, 1975) has outlined the principles upon which this method might be developed for a multistate system. In the context of interregional migration, such a method was first developed by Ledent and Rogers (1972) and later improved by Ledent (1978:54-57). A further improvement of this method is presented below.

The calculation comprises two steps. First, the mortality and mobility curves are graduated to small intervals and then the rates for the small age intervals are adjusted so that, aggregated to the larger age intervals normally employed in multiregional life tables (5 years), they match the observed rates.

The method adopted here fits a cubic spline function to the observed mortality and mobility rates to carry out the required interpolation (using a similar method to that described by McNeil, Trussell and Turner 1977). The interpolation was to rates applicable to one-year age groups (rather than 0.2 year age groups) because of computing constraints.

Once the values of the mortality and mobility rates for the single year age groups have been interpolated, one uses either the linear or the exponential or the cubic method to derive the ${}_h\underset{\sim}{p}_x$, ${}_h\underset{\sim}{\ell}_x$, and ${}_h\underset{\sim}{L}_x$ values for single years of age where h is the small age group interval (one in this case). From these values, the implied five years of age rates of mortality and mobility can be derived from equation (24) rewritten as

$$\underset{\sim}{m}_x^* = (\underset{\sim}{\ell}_x - \underset{\sim}{\ell}_{x+n})\underset{\sim}{L}_x^{-1} \qquad (50)$$

where the $\underset{\sim}{\ell}_x$ and $\underset{\sim}{\ell}_{x+n}$ matrices are selected from the more detailed ${}_h\underset{\sim}{\ell}_x, {}_h\underset{\sim}{\ell}_{x+h}, {}_h\underset{\sim}{\ell}_{x+2h}, \ldots, {}_h\underset{\sim}{\ell}_{x+n}$ series and

$$\underset{\sim}{L}_x = \sum_{y=x}^{x+n-h} {}_h\underset{\sim}{L}_y \qquad (51)$$

The $\underset{\sim}{m}_x^*$ estimate thus obtained will generally not agree with the observed $\underset{\sim}{M}_x$. Improved estimates of the single year mortality and mobility rates are obtained by adjusting the initial estimates using

$${}_hm_y^{ij}(k + 1) = {}_hm_y^{ij}(k) \frac{M_x^{ij}}{m_x^{ij*}(k)} \qquad (52)$$

and

$${}_hm_y^{\delta(i)}(k + 1) = {}_hm_y^{\delta(i)}(k) \frac{M_x^{\delta(i)}}{m_x^{\delta(i)*}(k)} \qquad (53)$$

for all i from 1 to r, all j from 1 to r, and all y such that $x \le y \le x+n-h$, where k and k + 1 refer to successive iterations of the procedure. The procedures involving equations (50) through (53) are repeated until convergence is achieved, that is, until

$$|m_x^{ij*}(k) - M_x^{ij}| < .000001 \qquad (54)$$

$$\left| m_x^{\delta(i)*}(k) - M_x^{\delta(i)} \right| < .000001 \tag{55}$$

for all i, j and x.

3.3.5 About the Last Age Group

The treatment of the last age group is the same regardless of the method used. One simply substitutes the observed rates matrix $\underset{\sim}{M}_z$ for the life table rates matrix $\underset{\sim}{m}_z$ in equation (28).

3.4 The Effects of Alternative Methods

Schoen (1979) has looked at the effect of the first three of the probability estimation methods in a marital status classified population system. Here we extend the comparison to a multiregional population system and include the additional interpolative-iterative method in the comparison. To ascertain how important the choice of method of probability and life years estimation is we apply each method to data from Drewe (1980) on the Netherlands case in the IIASA Migration and Settlement series of country studies. Migration data collected are of the movement type as is clear from Drewe's description

> As regards migration data, movers (migrating families or single persons) receive a special card ("verhuiskaart") from the municipality of origin, which they are requested to fill in and hand over to the municipality of destination. After registration, the card is returned to the municipality of origin, and from there it is passed on to the Central Bureau of Statistics (Drewe 1980:16).

Drewe's data have been aggregated for convenience of presentation to a four-region system by combining the South-West and South regions into one. The detailed makeup of each region in terms of Dutch provinces is spelled out in the footnote to Table 2.

Selected multiregional life table statistics are presented in Tables 2 through 4, and in Figure 2. Table 2 shows the probabilities ($\underset{\sim}{p}_x$'s) for the age 20 to age 25 transition. The columns of the table refer to a separate computer run employing the method indicated at the top of the column: the last two

Table 2. Netherlands: transition probabilities, ages 20 to 25 ($\underset{\sim}{p}_{20}$).

Regions	Linear	Exponential	Cubic	Interpolative-iterative Linear	Interpolative-iterative Exponential
DEATH PROBABILITIES					
North	.00376	.00375	.00376	.00376	.00376
East	.00378	.00378	.00378	.00378	.00378
West	.00299	.00299	.00299	.00299	.00299
South	.00385	.00385	.00385	.00385	.00385
SURVIVAL AND STAYING PROBABILITIES					
North	.80886	.80876	.80892	.80929	.80933
East	.76554	.76747	.76593	.76682	.76689
West	.86852	.86925	.86835	.86851	.86854
South	.84295	.84357	.84321	.84357	.84360
OUT-MIGRATION PROBABILITIES FROM NORTH TO:					
East	.07472	.07392	.07460	.07426	.07423
West	.09180	.09153	.09175	.09162	.09161
South	.02086	.02103	.02097	.02107	.02107
OUT-MIGRATION PROBABILITIES FROM EAST TO:					
North	.04218	.04175	.04205	.04186	.04184
West	.12994	.12887	.12963	.12911	.12907
South	.05856	.05814	.05859	.05844	.05844
OUT-MIGRATION PROBABILITIES FROM WEST TO:					
North	.02438	.02430	.02440	.02438	.02438
East	.05722	.05674	.05724	.05711	.05709
South	.04689	.04671	.04702	.04701	.04701
OUT-MIGRATION PROBABILITIES FROM SOUTH TO:					
North	.01036	.01045	.01040	.01044	.01044
East	.04774	.04742	.04762	.04745	.04743
West	.09510	.09472	.09492	.09469	.09468

Method of probability estimation: Linear, Exponential, Cubic, Interpolative-iterative (Linear, Exponential).

Note: The probabilities in the table are given to five decimal places. They were computed to six.

SOURCE: The population, deaths and migrations data for the Netherlands were abstracted from Appendix A in Drewe (1980:44-48). The five regions Drewe used were aggregated to four by combining his South and South-West regions. The regions are aggregations of the Dutch provinces and are composed as follows: North is made up of Groningen, Friesland and Drenthe; East is made up of Overijssel and Gelderland; West is an aggregation of Utrecht, Noord-Holland, and Zuid-Holland; South (in this paper) is an amalgam of Noord-Brabant, Limburg and Zeeland.

columns show two alternative versions of the interpolative-iterative method, one in which the linear method is used with the (converged) one year mortality and mobility rates, and the other in which the exponential method is used. We could also have used the cubic method. A glance at the table and Table 3 shows that this sub-choice makes very little difference to the resulting life table statistics.

The arrangement of statistics in Table 2 is as follows. The probabilities are most usually shown in the form of a matrix or table of migration probabilities, such as in the linear case

$$
\begin{array}{l c}
\underline{\text{From}} & \underline{\text{Migration to}} \\
 & \begin{array}{cccc} \text{North} & \text{East} & \text{West} & \text{South} \end{array} \\
\begin{array}{l} \text{North} \\ \text{East} \\ \text{West} \\ \text{South} \end{array} &
\begin{pmatrix}
.80886 & .07472 & .09180 & .02086 \\
.04218 & .76554 & .12994 & .05856 \\
.02438 & .05722 & .86852 & .04689 \\
.01036 & .04774 & .09510 & .84295
\end{pmatrix}
\end{array}
$$

and a vector of death probabilities

$$
\begin{array}{l c}
\underline{\text{From}} & \text{Death} \\
\begin{array}{l} \text{North} \\ \text{East} \\ \text{West} \\ \text{South} \end{array} &
\begin{pmatrix}
.00376 \\
.00378 \\
.00299 \\
.00385
\end{pmatrix}
\end{array}
$$

In Table 2 the migration probabilities and death probabilities have been rearranged in one continuous column, the death probabilities first, then the retention probabilities (the diagonal elements of the matrix), followed by the remaining elements in the matrix, row by row. The same rearrangement has been affected in Table 3 for life expectancies by region of birth and the <u>percentage distribution</u> of life expectancies for each region of birth by the regions in which life is spent.

Table 3. Netherlands: regional life expectancies at birth (in years) and their distribution by region of residence (in percentages).

	Method of probability estimation				
				Interpolative-iterative	
Regions	Linear	Exponential	Cubic	Linear	Exponential
LIFE EXPECTANCIES BY REGION AT BIRTH					
North	74.60	74.57	74.65	74.64	74.64
East	74.53	74.49	74.57	74.56	74.56
West	74.81	74.78	74.86	74.86	74.85
South	74.44	74.40	74.49	74.48	74.47
PERCENT OF LIFE SPENT IN REGION OF BIRTH					
North	60.00	60.01	59.99	59.99	59.99
East	56.47	56.50	56.48	56.49	56.49
West	64.78	64.79	64.76	64.76	64.76
South	69.98	69.99	69.98	69.99	69.99
PERCENT OF LIFE SPENT IN OTHER REGIONS: BORN IN NORTH					
East	14.75	14.74	14.75	14.75	14.75
West	17.44	17.43	17.43	17.43	17.43
South	7.81	7.81	7.82	7.83	7.83
PERCENT OF LIFE SPENT IN OTHER REGIONS: BORN IN EAST					
North	8.75	8.74	8.75	8.75	8.75
West	21.06	21.04	21.04	21.03	21.03
South	13.72	13.72	13.73	13.73	13.73
PERCENT OF LIFE SPENT IN OTHER REGIONS: BORN IN WEST					
North	7.01	7.01	7.01	7.01	7.01
East	13.80	13.80	13.81	13.81	13.81
South	14.40	14.40	14.42	14.42	14.42
PERCENT OF LIFE SPENT IN OTHER REGIONS: BORN IN SOUTH					
North	3.45	3.45	3.46	3.46	3.46
East	10.01	10.00	10.01	10.01	10.01
West	16.56	16.55	16.55	16.54	16.54

The principal feature of the results of these five experiments with different alternatives of the movement approach is that they all are in very close agreement. The maximum disagreement between methods in life expectancies at birth is 0.09 of a year (South exponential result compared with South cubic) or just a little over a month, and between retention percentages is 0.03 of a percent (Table 3). The probabilities of transition from age 20 to 25 (Table 2) show greater differences but these are compensated for by differences in the opposite direction at other ages.

We have no "true" life table values against which to evaluate the methods properly, but we can select the interpolative-iterative method as a standard, and theoretically the best alternative, against which to compare the others. In Figure 2a are plotted the differences between the retention probabilities for the North region and those for the linear, cubic and exponential methods, and in Figure 2b the corresponding differences for the probabilities of migration from the South region to the North are plotted.

The cubic method turns out to be the closest to the interpolative-iterative method (that is, closest to the horizontal line plotted at the zero level) in terms of retention and migration probabilities. The linear method fares better than the exponential for the retention probabilities shown in Figure 2a but slightly worse for the migration probabilities plotted in Figure 2b. Overall, the linear and exponential methods perform at much the same "error" level in distributing life across the Dutch regions, but the exponential method gives slightly worse total life expectancies. As might be expected the linear, exponential and cubic methods fare worst compared with the interpolative-iterative method between ages 15 and 25 when the migration propensity curves are altering shape and direction dramatically, and at the last two ages transitions where rather unsatisfactory assumptions are made.

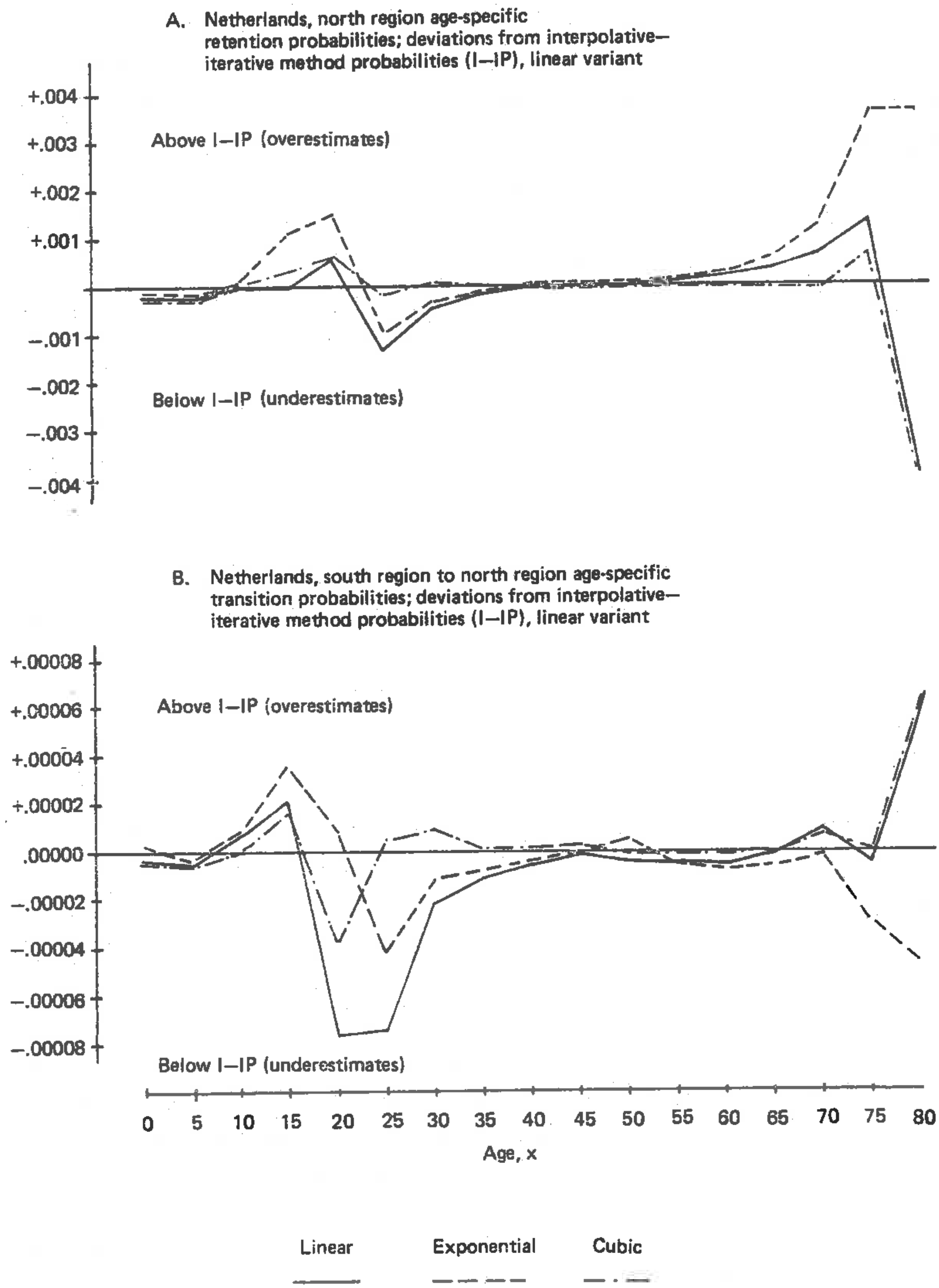


Figure 2. Comparisons of selected probabilities from linear, exponential, and cubic methods with those from the iterative-interpolative method.

A final set of statistics which can be used to evaluate the movement approach alternatives are the mean durations to transfer (within an age interval)--shown in Table 4--obtained for each method by applying the equation given by Ledent (1978: 39-40, equation 48)

$$[\underset{\sim}{a}_x \otimes \underset{\sim}{m}_x] = (\underset{\sim}{L}_x - n\, \underset{\sim}{\ell}_{x+n})\underset{\sim}{L}_x^{-1} \tag{56}$$

where $[\underset{\sim}{a}_x \otimes \underset{\sim}{m}_x]$ is a matrix each element of which consists of a term in a matrix $\underset{\sim}{a}_x$ multiplied by the corresponding term in the $\underset{\sim}{m}_x$ matrix, so that

$$[\underset{\sim}{a}_x \otimes \underset{\sim}{m}_x]' = \begin{bmatrix} [a_x^{\delta(1)}m_x^{\delta(1)} + \sum\limits_{k\neq 1} a_x^{1k}m_x^{1k}] & -a_x^{12}m_x^{12} & \cdots & -a_x^{1r}m_x^{1r} \\ -a_x^{21}m_x^{21} & [a_x^{\delta(2)}m_x^{\delta(2)} + \sum\limits_{k\neq 2} a_x^{2k}m_x^{2k}] & \cdots & -a_x^{2r}m_x^{2r} \\ \vdots & \vdots & \ddots & \vdots \\ -a_x^{r1}m_x^{r1} & -a_x^{r2}m_x^{r2} & \cdots & [a_x^{\delta(r)}m_x^{\delta(r)} + \sum\limits_{k\neq r} a_x^{rk}m_x^{rk}] \end{bmatrix} \tag{57}$$

The a_x^{ik} variables are the average time spent in region i by persons making moves to region k, and the $a_x^{\delta(k)}$ are the average times spent in k before death, all in the age transition x to x+n. The a_x^{ik}'s are computed by dividing each off-diagonal element in the $[\underset{\sim}{a}_x \otimes \underset{\sim}{m}_x]$ matrix by the corresponding element in the $\underset{\sim}{m}_x$ matrix; the off-diagonal values can then be inserted in the diagonal and the $a_x^{\delta(k)}$'s worked out. Numerically, the application of equation (56) gives rise to slight problems in the case of younger ages when the determinant of $\underset{\sim}{L}_x$ is close to zero.

Table 4. Netherlands, age transition 70 to 75: mean duration to transfer.

		Migration to			
From	Death	North	East	West	South
	LINEAR				
North	2.50	0	2.50	2.50	2.50
East	2.50	2.50	0	2.50	2.50
West	2.50	2.50	2.50	0	2.50
South	2.50	2.50	2.50	2.50	0
	EXPONENTIAL				
North	2.42	0	2.32	2.33	2.35
East	2.42	2.32	0	2.32	2.32
West	2.42	2.33	2.32	0	2.32
South	2.42	2.36	2.32	2.32	0
	CUBIC				
North	2.62	0	2.37	2.46	2.44
East	2.62	2.31	0	2.42	2.37
West	2.61	2.22	2.25	0	2.27
South	2.60	2.35	2.31	2.32	0
	INTERPOLATIVE-ITERATIVE (LINEAR VARIANT)				
North	2.61	0	2.39	2.50	2.48
East	2.61	2.31	0	2.42	2.37
West	2.60	2.22	2.25	0	2.28
South	2.59	2.30	2.39	2.45	0

From the results shown in Table 4 we can see that the cubic and interpolative-iterative are necessarily more accurate than the alternative linear and exponential approaches. Since mortality rate functions are steeply increasing between ages 70 and 75, we would expect the mean duration of transfer to death to exceed 2.5; since mobility rate functions are declining between ages 70 to 75, we would expect the mean duration of transfer not to exceed 2.5. These expectations are fulfilled in the cubic and interpolative-iterative cases but not in the linear or exponential.

3.5 The Negativity Problem and Related Issues

The observation has been made (Rees 1978) that the linear method can produce negative probability estimates, typically on the diagonal, if high annual migration rates are observed. The reason for this result is that no non-negativity constraints are imposed on the probabilities, and that the linear method does not effect a "conceptual" conversion of movements to transitions. To ensure non-negativity would require that only one move per transition occur (unlikely)* and that the conversion from a one-year to five-year period be affected by powering the one-year probabilities:

$$\underset{\sim}{p}_x = [(\underset{\sim}{I} + \tfrac{1}{2}\,\underset{\sim}{M}_x)^{-1}(\underset{\sim}{I} - \tfrac{1}{2}\,\underset{\sim}{M}_x)]^n \tag{58}$$

The counter-observation, however, has been made (Rees 1979b) that migration rates sufficiently high for the results of equation (58) and equation (35) to differ significantly are rarely observed: they do not, for example, occur in interregional migration between regions in Great Britain (Rees 1979a).

*This could be achieved if the number of moves per transition were known: each movement flow could be divided by this number.

To test the likelihood of such negative probability estimates, a simple hypothetical two-region system has been designed:

$$\underset{\sim}{M} = v\begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix} + w\begin{bmatrix} 0.005 & -0.010 \\ -0.005 & 0.010 \end{bmatrix} \qquad (59)$$

where v is the mortality level and w the mobility level. The mortality levels were allowed to vary from 1 to 100, yielding mortality rates of from 1 per thousand to 100 per thousand; and the mobility levels were allowed to vary from 1 to 200, generating movement rates of from 5 per thousand to 2000 per thousand. What equation (58) does is to generate hypothetical $\underset{\sim}{M}$ matrices which then input into the linear probability equation (35), the modified power version of the linear probability equation, equation (58), or the exponential equation (40).

It turns out that the second region exhibits negative retention probabilities (p^{22}) at high mortality and mobility levels, so that in Figure 3 the region 2 retention probabilities have been plotted on graphs for each of the three estimation equations. Only the linear equation produces negative values; the power-linear and exponential equations always give positive results, very close together, which approach an asymptote result as the mobility level increases. Increasing the mortality level has only a moderate effect on the outcome.

However, although negative retention probabilities only occur with the linear equation at mortality levels not observed in real world regional systems, a worrying discrepancy between the results of the linear equation and the other two appears at lower mobility levels between 10 and 15 (or between 100/1000 and 150/1000 total outmigration rate for region 2). One year mobility rates are observed at this level in many countries, but not for interregional migration, merely for migration from residences in general. In Rees (1979b) results of using power-linear were very little different from those of the linear equation for a three-region, British population system.

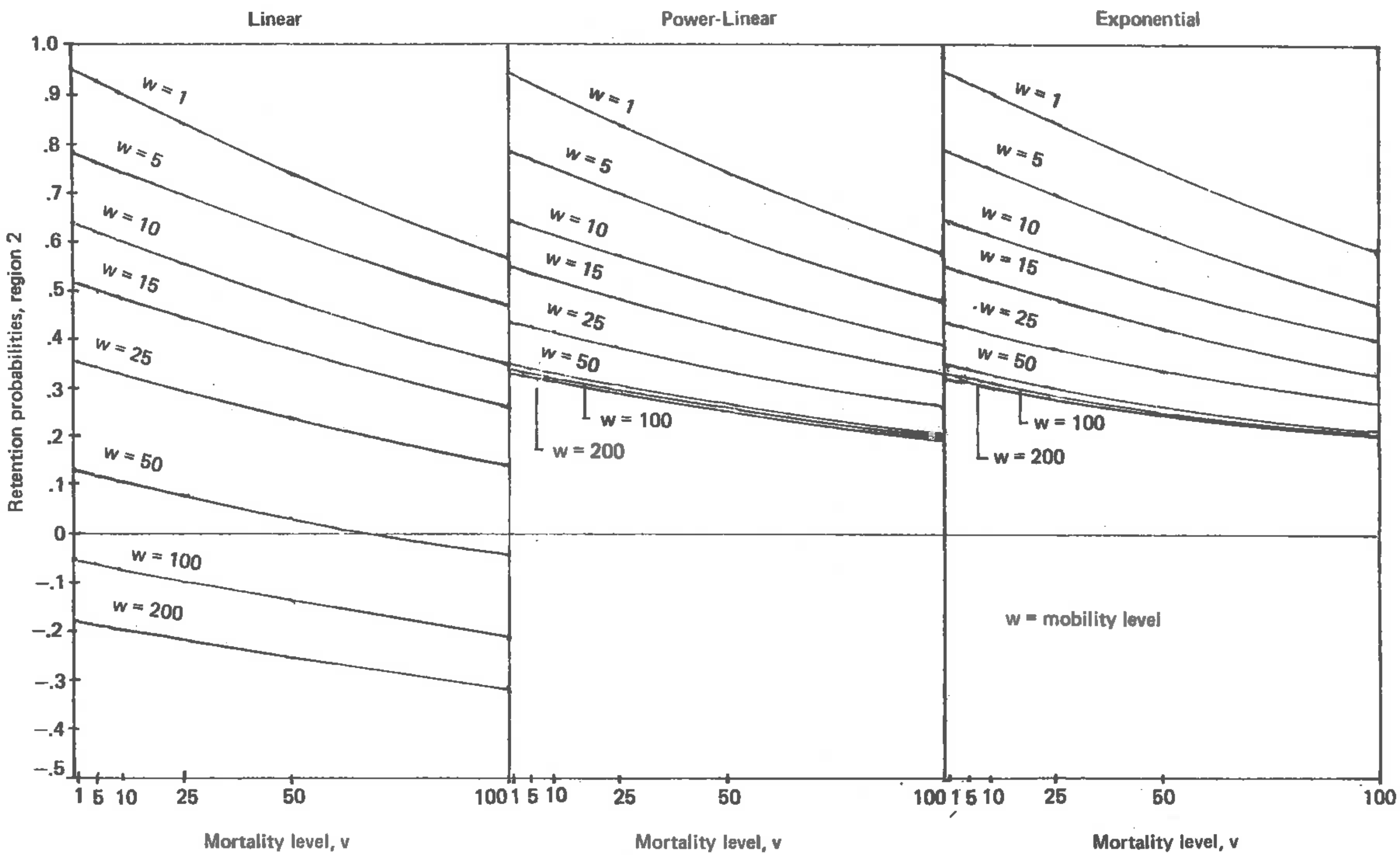


Figure 3. The influence of mortality and mobility levels on estimated probabilities, movement approach.

Our general conclusion must be therefore that the negativity problem is not a serious one but that it would be as well to keep in mind the problem when new and highly mobile multistate systems are being investigated.

3.6 Conclusions About the Movement Approach

Let us review now the choices facing someone wishing to construct a multiregional life table and having available only movement type data (Figure 1).

Firstly, if these movement data are classified by place of birth, they should be employed to construct r separate multiregional life tables, one per region of birth. Although we can not present empirical evidence in favour of this choice (although Ledent 1980 gives the evidence in the transition case), it is clearly justified on theoretical grounds since it relaxes a key assumption of the model. However, normally this choice is not available.

The second set of choices concern probability and stationary population estimation equations. Again we would argue on theoretical grounds that the interpolative-iterative method is the best, although rather more demanding in programming and computer-time terms. Of the non-iterative methods the cubic is closest to the interpolative-iterative, followed by the linear, and then the exponential. There are arguments in favour of the exponential or a power modification of the linear if mobility levels are very high.

However, the differences between the various methods are probably not large enough to be of empirical interest, and the extensive use to which the linear (Option 3) method has been put (in IIASA's Migration and Settlement Task) would seem justified.

4. THE HYBRID APPROACH: USING MORTALITY AND MOBILITY RATES WITH TRANSITION DATA

4.1 General Observations

It is clear from the preceding section that the movement approach to the calculation of a multiregional life table constitutes a generalization of the classical method of ordinary life table construction: in most instances, matrices are simply substituted for scalars. A characteristic of this approach is the observation and calculation of mobility rates in age-time arrangements identical to the common observation and calculation of mortality rates. It thus requires data in the form of moves or migrations that are to be established from a population register.

In fact, many countries do not maintain such registers and geographical mobility data are then obtained from a population census in the form of migrants or changes of residence derived from questions about residence one or five years earlier. The proper approach for taking care of this data type is the transition approach which attempts to generalize the methods of ordinary life table construction based on survivorship rates.

Historically, however, the case of mobility data from census information was dealt with by Rogers (1973a, 1973b, 1975) in yet another approach in which the concepts were borrowed from the classical or movement approach. Because of these two ingredients--methods from the movement approach and inputs being transition data--we have characterized this approach as the *hybrid* approach, to contrast it with the transition approach. These distinctions were recognized in part in Rogers (1975a) in the discussion of probability estimation options and in the description of those options (Options 1, 2, and 3) in Willekens and Rogers (1978). We hope in this section to make the distinctions as clear as we can.

4.2 Probability Estimation and Associated Equations

Formally, the transition approach is characterized by a set of flow and orientation equations pertaining to the evolution between ages x and x+n of the closed group of people ${}^{i}_{x}\ell^{i}_{x}$ present at age x in region i, regardless of their place of birth. Note that

$$ {}^{i}_{x}\ell^{i}_{x} = \sum_{j=1}^{r} {}^{j}_{0}\ell^{i}_{x} \tag{60} $$

and observe the difference with the movement approach.

The basic flow equation may be defined as (Rogers 1973a, 1975a)

$$ {}^{i}_{x}\ell^{i}_{x+n} = {}^{i}_{x}\ell^{i}_{x} - \sum_{\substack{k=1 \\ k \neq i}}^{r} {}^{i}_{x}\ell^{ik}_{x} - {}^{i}_{x}\ell^{i\delta(.)}_{x} \tag{61} $$

where ${}^{i}_{x}\ell^{ik}_{x}$ terms are the number of persons in the life table population who, present at age x in region i, survive to age x+n in region k and ${}^{i}_{x}\ell^{i\delta(.)}_{x}$ is the number of persons in the same closed group who die before reaching age x+n.

The definition of the orientation equations here is not as straightforward as in the movement case. Rogers himself (Rogers 1973a, 1975a) does not present any clear specification of such equations but his approach of the problem (Rogers 1975a:82-84) can be viewed as consisting of the following:

$$ {}^{i}_{x}\ell^{ik}_{x} = {}^{i}_{x}m^{ik}_{x} \; {}^{ii}_{x}L^{i}_{x} \qquad \text{(note that } k \neq i\text{)} \tag{62} $$

and

$$ {}^{i}_{x}\ell^{i\delta(.)}_{x} = {}^{i}_{x}m^{i\delta(i)}_{x} \; {}^{ii}_{x}L^{i}_{x} \tag{63} $$

where the mobility and mortality rates thus defined are--in spite of the identical notation used--different from those defined earlier in the movement approach: ${}^{ii}_{x}L^{i}$ are the life years lived in region i by members of the region i, age x cohort over the x to x+n age interval. The interpretation of such a set of

The interpretation of such a set of equations is rather simple from the point of view of classical mathematical demography: any move by a member of the closed group ${}^{i}_{x}\ell^{ik}$ out of region i is regarded as a death, at least as long as the focus is on the age/time interval x to x+n. [The provision for return migration within such an interval which Rogers (1975a) evoked in theoretical terms but disregards at the applied level is, however, inconsistent with such a view.] In other words, the specification of the above orientation equations implicitly rules out the possibility of multiple moves per age interval. Therefore, an individual who has just moved out from a certain region is not exposed to the risk of dying until he or she reaches the next exact age in the series 0,n,2n,...! We will see later on an important consequence of such an assumption.

Then, substituting the orientation equations into the flow equations leads to

$$ {}^{i}_{x}\ell^{i}_{x+n} = {}^{i}_{x}\ell^{i}_{x} - \sum_{\substack{k=1 \\ k\neq i}}^{r} {}^{i}_{x}m^{ik}_{x}\ {}^{ii}_{x}L^{i}_{x} - {}^{i}_{x}m^{i\delta(i)}_{x}\ {}^{ii}_{x}L^{i}_{x} \qquad (64) $$

which together with (62) constitutes the fundamental system of equations leading to the derivation of formulae for estimating the survival probabilities. Unfortunately, this system cannot be simply summarized into a matrix equation similar to equation (24) of the movement approach.

In practice, the derivation of such formulae requires the adoption of a method for calculating the number of person-years lived as well as a linkage of the life table with the observed rates. With regard to the latter, it is customary to assume--because of the format in which mortality and census migration data are generally collected and tabulated--that the mortality and mobility rates are independent of the place of residence at age x

$$ {}^{i}_{x}m^{ik}_{x} = M^{ik}_{x} \tag{65}$$

and

$$ {}^{i}_{x}m^{i\delta(i)}_{x} = M^{\cdot\delta(i)}_{x} \tag{66}$$

How does one observe such rates in practice? In the case of mortality, the observed rate can be simply assimilated with the conventional mortality rate $M^{\cdot\delta(i)}_{x}$ pertaining to region i. The case of the mobility rate is more complicated and will be dealt with later on.

4.2.1 The Linear Approach ("Option 1")

Rogers (1973a, 1975a) simply assumes that the number of person-years lived by the members of the group ${}^{i}_{x}\ell^{i}_{x}$ can be obtained linearly from the ℓ-statistics as

$$ {}^{ki}_{x}L^{i}_{x} = \frac{n}{2}[{}^{i}_{x}\ell^{k}_{x} + {}^{i}_{x}\ell^{ik}_{x}] \qquad \text{for all k's} \tag{67}$$

including k = i. (Note that ${}^{i}_{x}\ell^{ii}_{x} = {}^{i}_{x}\ell^{i}_{x+n}$ and ${}^{i}_{x}\ell^{ik}_{x} = 0$ if $k \neq i$.) Then, it can be established (Rogers 1975a:82-84) that

$$ p^{ij}_{x} = \frac{\frac{n}{T} M^{ij}_{x}}{1 + \frac{n}{2T} \sum\limits_{k\neq i}^{T} M^{ik}_{x} + \frac{n}{2T} M^{\delta(i)}_{x}} \qquad \text{for all } j \neq i \tag{68}$$

$$ p^{i\delta}_{x} = \frac{\frac{n}{T} M^{\delta(i)}_{x}}{1 + \frac{n}{2T} \sum\limits_{k\neq i}^{r} M^{ik}_{x} + \frac{n}{2T} M^{\delta(i)}_{x}} \tag{69}$$

and

$$p_x^{ii} = \frac{1 - \frac{n}{2T} \sum_{k \neq i}^{r} M_x^{ik} - \frac{n}{2T} M_x^{\delta(i)}}{1 + \frac{n}{2T} \sum_{k \neq i}^{r} M_x^{ik} + \frac{n}{2T} M_x^{\delta(i)}} \tag{70}$$

$$\text{or} \qquad 1 - \sum_{k \neq i} p_x^{ik} - p_x^{i\delta} \tag{71}$$

These equations have already appeared as equations (3), (4), and (5) but are repeated here for convenience.

4.2.2 The Exponential Method

As an alternative to the linear approach, it is possible to assume that the mortality and mobility propensities are constant over the whole interval x to x+n. Then we have

$$p_x^{ii} = e^{-n(\sum_{k \neq i}^{r} M_x^{ik} + M_x^{\delta(i)})} \tag{72}$$

and

$$p_x^{ij} = \frac{M_x^{ij}}{\left(\sum_{k \neq i}^{r} M_x^{ik} + M_x^{\delta(i)}\right)} \left(1 - e^{-n(\sum_{k \neq i}^{r} M_x^{ik} + M_x^{\delta(i)})}\right) \tag{73}$$

and

$$p_x^{i\delta} = \frac{M_x^{\delta(i)}}{\left(\sum_{k \neq i}^{r} M_x^{ik} + M_x^{\delta(i)}\right)} \left(1 - e^{-n(\sum_{k \neq i}^{r} M_x^{ik} + M_x^{\delta(i)})}\right) \tag{74}$$

In such circumstances, the number ${}^{ii}_{x}L^{i}$ of person-years lived in region i is obtained by substituting $p^{ii}_{x}\ {}^{i}_{x}\ell^{i}_{x}$ [where p^{ii}_{x} is given by (72)] into (c)

$$ {}^{ii}_{x}L^{i}_{x} = \frac{1 - e^{-n(\sum_{k\neq i} M^{ik}_{x} + M^{\cdot\delta(i)}_{x})}}{(\sum_{k\neq i} M^{ik}_{x} + M^{\cdot\delta(i)}_{x})}\ {}^{i}_{x}\ell^{i}_{x} \tag{75} $$

and finally we have

$$ {}^{ji}_{x}L^{i}_{x} = \frac{M^{ij}_{x}}{(\sum_{k\neq i} M^{ik}_{x} + M^{\cdot\delta(i)}_{x})}\left[n - \frac{1 - e^{-n(\sum_{k\neq i} M^{ik}_{x} + M^{\cdot\delta(i)}_{x})}}{(\sum_{k\neq i} M^{ik}_{x} + M^{\cdot\delta(i)}_{x})}\right]{}^{i}_{x}\ell^{i}_{x} \tag{76} $$

4.2.3 The Linear and Exponential Methods from the Movement Approach

As seen earlier, a fundamental problem with the hybrid approach as exposed above is the ruling out of the possibility of multiple moves which leads to an artificial reducing of the exposure to the risk of dying.

Ledent (1978:67) indicates that the multiple move assumption can be somewhat attenuated to allow for the possibility of death after a migration (two migrations within the same interval are still not allowed). This does not lead to a much greater complexity for the derivation of the survival probabilities for which explicit formulae have been derived by Ledent (1978: 71-73) in the case of a linear calculation of the L-statistics.

An alternative possibility of achieving the same result has been proposed by Willekens and Rogers (1978); it simply consists of substituting the movement-based linear formulae ("Option 3") for the above linear formulae ("Option 1"). In the same vein, it is also reasonable to substitute the movement-based exponential formulae for the above exponential formulae.

4.2.4 About the Last Age Group

The treatment of the last age group is the same for all methods. The equation

$$\underset{\sim}{L}_z = \underset{\sim}{M}_z^{-1} \underset{\sim}{\ell}_z \tag{77}$$

is used (the migration rates are assumed to be zero).

4.3 Rate Estimation Equations

4.3.1 The Approach So Far: Equivalence of Data and Model Age Labels

At this point it is useful to retrace our steps in Figure 1 and go back to the column labeled "Rates - Definition". We concentrate our discussion on migration rates, since the mortality rates are as conventionally computed.

In defining rates in the movement approach (and mortality rates in all approaches) the match between the definition of the events by the data collecting agency and in the model are usually identical. This is the situation shown in Figure 4a for a five-year age interval and one-year time interval, although the same principle applies no matter what the intervals. The dots in the data Lexis diagram refer to movements at the point in age-time space that they occur.

In the case of the hybrid approach in which transition data are used, difficulties occur. We cannot in this case assume an equivalence of the age-time space in which the data are gathered, and that used in the model, as is clear from Figure 4b for a one-year time interval and from Figure 4c for a five-year time interval. The transition data are recorded in an end of interval census (at time t+1 in Figure 4b or time t+5 in Figure 4c) and are classified by age at the end of the time interval. Their age at the start of the interval can be worked out using the Lexis diagram and will be x-1 to x+n-1 in the one-year case when n = 5 and x-n to x in the five-year case when n = 5.

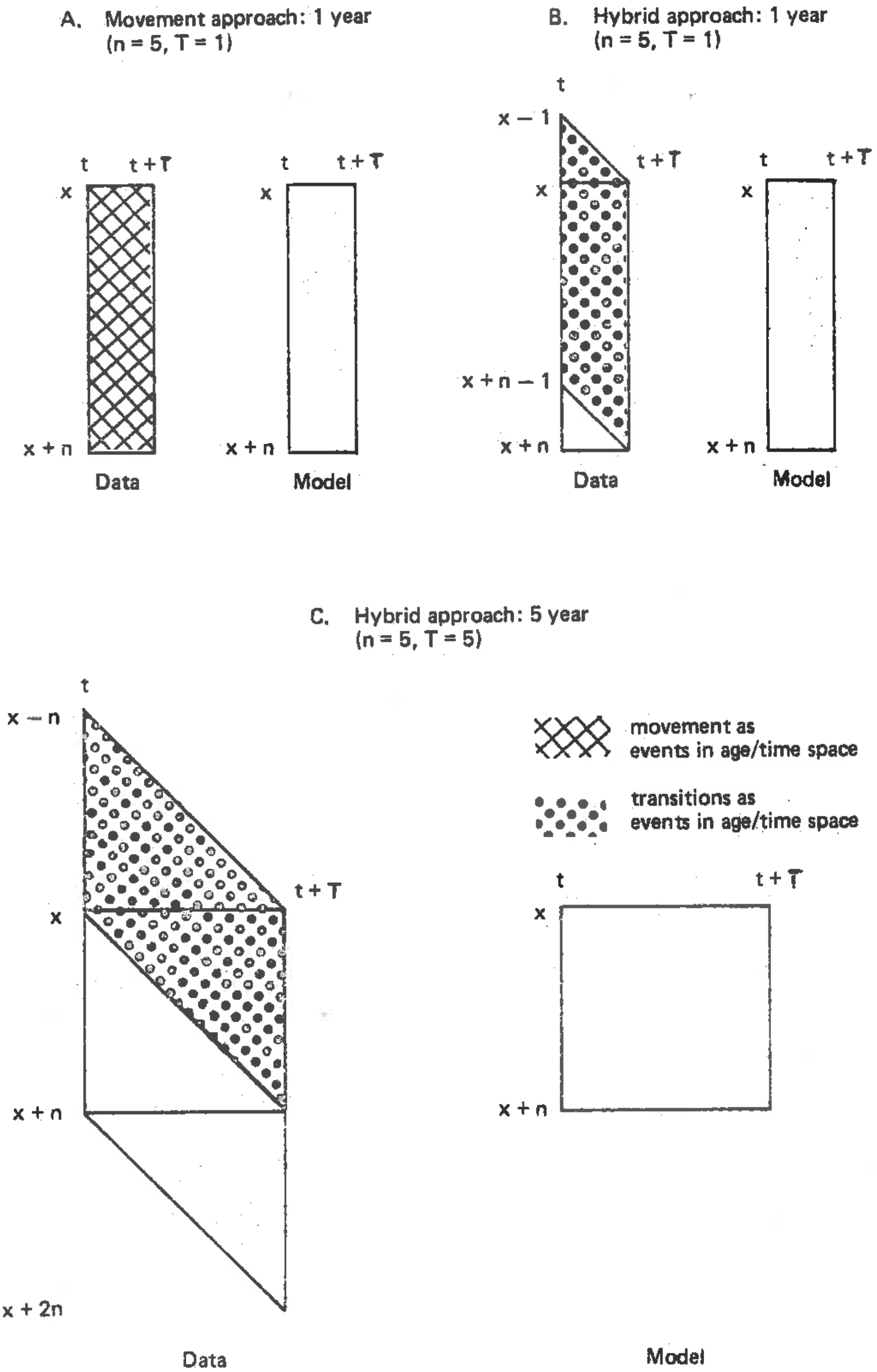


Figure 4. The data and model matching problem in the hybrid approach.

It is therefore incorrect to define mobility rates as

$$M_x^{ij} = \frac{K_{.x}^{ij}}{K_x^i} \tag{78}$$

or

$$M_x^{ij} = \frac{K_{.x+n}^{ij}}{K_x^i} \tag{79}$$

equivalent when T = n to $M_x^{ij} = \dfrac{K_{x.}^{ij}}{K_x^i}$ (80)

where $K_{.x}^{ij}$ refers to migrants classified by age at their region j location at the end of the measurement period, at the time of the census.

Both alternatives will be equally incorrect when n = T but the definition in equation (78) will be better if T < n. If n = 5 and T = 1, equation (78) will give a reasonable approximation on the assumption that events in the bottom triangle in Figure 4b are the same as events in the top triangle. Unfortunately, neither in Rogers (1975a) nor in Willekens and Rogers (1978) is the method of rate computation made explicit, but looking at the empirical work on the United States population (Rogers 1975a), normally equation (78) is employed with one-year period data and equation (79) with five-year data. For the first age group, it is assumed, in the one-year period case,

$$M_0^{ij} = K_{.1}^{ij} \;/\; K_0^i \tag{81}$$

where $K_{.1}^{ij}$ are the migrants from region i to region j aged one to four at the end of the period, and K_0^i is the mid-period population aged one to four. In the five-year period case, the equation adopted is

$$M_0^{ij} = K_{.5}^{ij} \ / \ K_0^{i} \tag{82}$$

where $K_{.5}^{ij}$ are the i,j migrants aged five to nine at the end of the period. For the last age group in the one- and five-year cases, it is assumed

$$M_z^{ij} = K_{.z}^{ij} \ / \ K_z^{i} \tag{83}$$

where z refers to the last age group z years and beyond.

The problem with the first age group equations is that they neglect to take into account infant migrants--$K_{.0}^{ij}$--and the problem with the last age group equations is that again there is a mismatch between the migrants' age and those of the denominator.

4.3.2 An Averaging Procedure

The correct procedure, following Rees (1979a), is to estimate the necessary model inputs from adjacent age group data, thus*:

$$M_x^{ij} = \frac{(1 - \frac{T}{2n})K_{.x}^{ij} + (\frac{T}{2n})K_{.x+n}^{ij}}{K_x^{i}} \tag{84}$$

with the n = 5, T = 1 equation being

$$\begin{aligned} M_x^{ij} &= \left[(1 - \tfrac{1}{10})K_{.x}^{ij} + (\tfrac{1}{10})K_{.x+n}^{ij}\right] / \ K_x^{i} \\ &= \left[(\tfrac{9}{10})K_{.x}^{ij} + (\tfrac{1}{10})K_{.x+n}^{ij}\right] / \ K_x^{i} \end{aligned} \tag{85}$$

*This assumes $T \leq n$. When $T > n$, then more age groups become relevant.

and the n = 5, T = 5 equation being

$$M_x^{ij} = \left[(1 - \tfrac{5}{10})K_{.x}^{ij} + (\tfrac{5}{10})K_{.x+n}^{ij}\right] / K_x^i$$
$$= \left[(\tfrac{1}{2})K_{.x}^{ij} + (\tfrac{1}{2})K_{.x+n}^{ij}\right] / K_x^i \qquad (86)$$

Modifications of these equations are needed for the first age group and the last. The general estimation equation for the 0 to n transition where $T \leq n$ is:

$$M_0^{ij} = \left[K_{.0}^{\beta(i)j} + K_{.T}^{ij} + (\tfrac{T}{2n})K_{.n}^{ij}\right] / K_0^i \qquad (87)$$

where $K_{.0}^{\beta(i)j}$ is the number of persons born in region i who migrate to region j and are aged 0 to T at the end of the period; $K_{.T}^{ij}$ are the region i to region j migrants who are aged T to n at the end of the period. When the age group interval equals the time period length, equation (87) reduces to

$$M_0^{ij} = \left[K_{.0}^{\beta(i)j} + (\tfrac{1}{2})K_{.n}^{ij}\right] / K_0^i \qquad (88)$$

With the last but one and last age groups, there is the difficulty that the Lexis diagram "geometric" weights are poor estimators of the distribution of migrants. A simple, although inadequate, suggestion would be to use population weights

$$M_{z-n}^{ij} = \left[\left(1 - \frac{T}{2n}\right)K_{.z-n}^{ij} + \left(\frac{K_z^j}{K_z^j + K_{z+n}^j}\right)K_{.z}^{ij}\right] / K_{z-n}^i \qquad (89)$$

and

$$M_{z-n}^{ij} = \left[\left(\frac{K_{z+n}^j}{K_z^j + K_{z+n}}\right)K_{.z}^{ij}\right] / K_z^i \qquad (90)$$

where z-n is the age which starts the last but one age group in the migrant data, and z is the age which starts the last age group; K_z^j is the population of region j aged w to w+n; and K_{z+n}^j is the population aged z+n and over. This method requires knowledge of population stocks to an age one interval greater than the migrant data, which is usually the case.

These rate estimation methods would, of course, be much improved through the adoption of an interpolative-iterative method of the type outlined in section 3.

4.4 Period of Data Measurement

A final element of choice in the hybrid approach concerns the period over which the migration data have been measured. The observation has been made by several researchers (Rees 1977a; Long and Boertlein 1975) that

$$K_x^{ij}(5) \neq 5\ K_x^{ij}(1) \tag{91}$$

or more generally that

$$K_x^{ij}(T_2) \neq \frac{T_2}{T_1}\ K_x^{ij}(T_1) \tag{92}$$

where the term in brackets indicates the time period over which the transitions were measured.

Three possible explanations for this observation have been proposed, and these are briefly considered here:

(i) It is a result of changes in the migration propensities over the different time periods of measurement --from t-5 to t compared with from t-1 to t, for example.

(ii) If the migration rates are transition rates then a multiplicative, Markov process should link the two observations. That is,

$$\underset{\sim}{H}(T_2) = \underset{\sim}{H}(T_1)^{(T_2/T_1)} \tag{93}$$

is a better description of the process, where $\underset{\sim}{H}$ is a matrix of transition rates defined as

$$h^{ij} = K^{ij} / K^{i}(t) \tag{94}$$

neglecting for the moment the treatment of age which is more complicated. The multiplicative process should create sequences of shorter period transitions which will be consolidated into one longer period transition with the short period transitions that "disappear" in the consolidation numbering $\left[\frac{T_2}{T_1} K_x^{ij}(T_1) - K_x^{ij}(T_2)\right]$.

(iii) Or it could be that, although the first and second explanations hold, they are insufficient to account for the differences observed. Instead, if there was to be a substantially higher probability of a person returning to his previous region of residence or region of birth than of a person making a new migration there, this might account for the discrepancy.

Results reported by Rees (1977) for the one year periods 1965-66 and 1970-71 which exhibited very little difference in their migration levels and patterns, and for the five-year period 1966-71 suggest that the first explanation has been a minor one. The second explanation appears to count for some of the five-year - one-year differences, particularly if migration propensities are high (cf. Figure 4 results), but a good deal is left over to be explained. Ledent (1980c) has shown for the United States between 1965 and 1970 that return migration was a very important phenomenon, which might well account for the one-year - five-year phenomenon.

Our conclusion from this discussion is that if equation (91) is true and equation (93) fails to hold, then the choice of period length to which the migration data apply could have a serious influence on the estimation of multiregional life table.

Speculations by Ledent (1978:133-137) were confirmed by analysis by Rees (1979b) which suggested for Great Britain that a substantial difference exists between the one-year based and five-year based life tables. This analysis is repeated in the present wider context.

4.5 The Data Sets Used and the Experiments Performed

The one-year data used consists of the original data assembled by Rees (1979a) on population, births, deaths, and migrants for the ten-region Great Britain country study in the IIASA Migration and Settlement Task. The data have been aggregate from ten to three regions (East Anglia, South East, Rest of Britain), and the migrant data manipulations, equivalent to the denominator on the right-hand side of equation (85), have bee left undone.

The five-year data are for the three-region system East Anglia, South East, and Rest of Britain and derive from the population accounts presented in Rees (1980a) which derive from an earlier study (Rees 1977b).

A third data set is also employed that combines the one-year deaths data for 1970 with the five-year migration data for 1966-71 for the three-region system. Comparisons can then be made between the first and third sets, holding the mortality information constant.

The data have been used with assumption that Great Britain constitutes a closed population system. Although this assumption is incorrect, it has to serve for present purposes either until work on developing a life table including the rest of the world is carried out (the second author's intention) or until a method of constructing life tables for an open system is developed. As a result the retention probabilities will tend to be inflated by the level of the emigration probabilities, as Rees (1980a) has pointed out.

Table 5 sets out the experiments carried out in order to analyze the effect that different method and data set choices have on the Great Britain multiregional life table. Obviously, only a selection of the results can be presented here: reference back to Table 5 will help clarify the nature of each analysis as it is presented.

We begin first by considering the effect of choice of probability and associated estimation equations.

4.6 The Effect of Choice of Probability and Associated Estimation Equations

Before beginning the presentation of the effect of choosing different probability estimation equations, we set out in Table 5 the region 1 (East Anglia) probabilities and in Table 6 the Great Britain life expectancy values and their distribution for the three-region system that are equivalent to those published in the UK Migration and Settlement country study (Rees 1979a, 1979b, Table 11, p. 48). One year data is employed, rates are computed by averaging successive census age groups for transition data, and the "Option 3" probability estimation method (movement-based/linear method) is utilized. Tables 6 and 7 represent therefore the best published estimates to date of the Great Britain multiregional life table.

We can now ask the questions: What would have been the effect of using different methods on these estimates? Table 8 collects together the probability vectors that refer to death or to outmigration from the East Anglia population for the four probability estimation methods. Table 6 values, our best estimates to date, appear in the third column under "Run 7". When the exponential version of the movement-based method is employed, the death probabilities are lowered somewhat (compare run 7 with run 8, and run 3 with run 4) and the migration probabilities are lowered a little. These were the differences

Table 5. A list of the experiments (computer runs) carried out for the hybrid approach on the Great Britain data sets.

Experiment (computer run)	Period choice: deaths	Period choice: migrants	Migration rate choice	Probability choice	Stationary population choice on L-choice
DATA SET ONE (gb1)					
1	one year	one year	equivalent	transition-based	/linear
2	"	"	"	"	/exponential
3	"	"	averaged	"	/linear
4	"	"	"	"	/exponential
5	"	"	equivalent	movement-based	/linear
6	"	"	"	"	/exponential
7	"	"	averaged	"	/linear
8	"	"	"	"	/exponential
DATA SET TWO (gb5)					
9	five years	five years	equivalent	transition-based	/linear
10	"	"	"	"	/exponential
11	"	"	averaged	"	/linear
12	"	"	"	"	/exponential
13	"	"	equivalent	movement-based	/linear
14	"	"	"	"	/exponential
15	"	"	averaged	"	/linear
16	"	"	"	"	/exponential
DATA SET THREE (gb 5 two)					
17	one year	one year	equivalent	transition-based	/linear
18	"	"	"	"	/exponential
19	"	"	averaged	"	/linear
20	"	"	"	"	/exponential
21	"	"	equivalent	movement-based	/linear
22	"	"	"	"	/exponential
23	"	"	averaged	"	/linear
24	"	"	"	"	/exponential

NOTES For discussion, see text section specified.

Data sets: section 4.5
Period choice: section 4.4
Rate choice: equivalent - section 4.3.1 averaged - section 4.3.2
Probability choice: transition-based/linear - section 4.2.1
transition-based/exponential - section 4.2.3
movement-based/linear - section 4.2.2
movement based/exponential - section 4.2.4

Table 6. Great Britain, hybrid approach: transition probabilities art of East Anglia, averaged rates, movement-based/ linear method with one year (1970) data (run 7).

Age, x	Death probability $p_x^{1\delta}$	Probabilities of surviving in: East Anglia p_x^{11}	South East p_x^{12}	Rest of Britain p_x^{13}
0	.01894	.81079	.08019	.09008
5	.00184	.87637	.05846	.06333
10	.00166	.90383	.04617	.04834
15	.00359	.83890	.08405	.07346
20	.00403	.75125	.12771	.11701
25	.00272	.78867	.10869	.09992
30	.00343	.85372	.06679	.07606
35	.00610	.88411	.05273	.05706
40	.01128	.91518	.03579	.03775
45	.01810	.92525	.03019	.02646
50	.02924	.90665	.03425	.02985
55	.04784	.92024	.01711	.01481
60	.07744	.89055	.01773	.01429
65	.13213	.83799	.01535	.01453
70	.19736	.76556	.01944	.01765
75	.28768	.68812	.01264	.01156
80	.41545	.56364	.01116	.00974
85	1.00000	.00000	.00000	.00000

Table 7. Great Britain, hybrid approach: life expectancies and their regional distribution, averaged rates, movement-based/linear method with one year (1970) data (run 7).

Region of birth	Life expectancy (years)	Percent spent in: East Anglia	South East	Rest of Britain
East Anglia	72.43	41.02	25.78	33.19
South East	72.45	4.11	64.55	31.35
Rest of Britain	71.49	2.31	16.31	81.38

Table 8. Great Britain, hybrid approach: transition probabilities out of East Anglia based on one year deaths and migrants data, 1970.

	Method of probability estimation			
	Transition-based		Movement-based	
Age	Linear Run 3	Exponential Run 4	Linear Run 7	Exponential Run 8
PROBABILITY OF DYING (p_x^{14})				
0	.01719	.01714	.01894	.01894
5	.00174	.00174	.00184	.00184
10	.00159	.00159	.00166	.00166
15	.00334	.00333	.00359	.00359
20	.00359	.00357	.00403	.00403
25	.00235	.00234	.00272	.00272
30	.00310	.00310	.00343	.00344
35	.00572	.00571	.00610	.00610
40	.01082	.01081	.01128	.01128
45	.01748	.01747	.01810	.01810
50	.02811	.02809	.02924	.02925
55	.04691	.04688	.04784	.04784
60	.07593	.07584	.07744	.07741
65	.12982	.12951	.13213	.13194
70	.19294	.19195	.19736	.19669
75	.28331	.28062	.28768	.28543
80	.40952	.40142	.41545	.40817
85	1.00000	1.00000	1.00000	1.00000
PROB. EAST ANGLIA → SOUTH EAST (p_x^{12})				
0	.08303	.08275	.08019	.07977
5	.05954	.05946	.05846	.05834
10	.04685	.04682	.04617	.04611
15	.08588	.08567	.08405	.08376
20	.13294	.13215	.12771	.12651
25	.11308	.11260	.10869	.10793
30	.06849	.06835	.06679	.06658
35	.05373	.05367	.05273	.05264
40	.03643	.03640	.03579	.03575
45	.03090	.03089	.03019	.03016
50	.03525	.03523	.03425	.03420
55	.01780	.01779	.01711	.01709
60	.01879	.01877	.01773	.01767
65	.01672	.01668	.01535	.01524
70	.02179	.02168	.01944	.01915
75	.01488	.01474	.01264	.01228
80	.01423	.01394	.01116	.01046
85	.00000	.00000	.00000	.00000
PROB. MIGRATION OUT OF EAST ANGLIA ($p_x^{12} + p_x^{13}$)				
0	.17282	.17225	.17027	.16958
5	.12228	.12211	.12179	.12160
10	.09482	.09475	.09451	.09442
15	.15839	.15802	.15751	.15710
20	.24673	.24527	.24472	.24311
25	.21018	.20930	.20861	.20763
30	.14375	.14347	.13685	.14253
35	.11050	.11037	.10979	.10964
40	.07417	.07412	.07354	.07348
45	.05738	.05737	.05665	.05661
50	.06536	.06531	.06410	.06403
55	.03291	.03289	.03192	.03188
60	.03361	.03357	.03202	.03192
65	.03225	.03217	.02988	.02968
70	.04155	.04134	.03709	.03654
75	.02860	.02833	.02420	.02348
80	.02685	.02632	.02090	.01954
85	.00000	.00000	.00000	.00000

	Method of probability estimation			
	Transition-based		Movement-based	
Age	Linear Run 3	Exponential Run 4	Linear Run 7	Exponential Run 8
PROB. EAST ANGLIA → EAST ANGLIA (p_x^{11})				
0	.80998	.81061	.81079	.81148
5	.87598	.87615	.87637	.87656
10	.90359	.90367	.90383	.90392
15	.83827	.83865	.83890	.83931
20	.74968	.75116	.75125	.75287
25	.78747	.78837	.78867	.78965
30	.85315	.85343	.85372	.85403
35	.88378	.88392	.88411	.88426
40	.91501	.91507	.91518	.91524
45	.92514	.92518	.92525	.92529
50	.90653	.90660	.90665	.90673
55	.92018	.92023	.92024	.92029
60	.89047	.89059	.89055	.89067
65	.83793	.83832	.83799	.83838
70	.76551	.76672	.76556	.76677
75	.68809	.69105	.68812	.69108
80	.56362	.57226	.56364	.57229
85	.00000	.00000	.00000	.00000
PROB. EAST ANGLIA → REST OF BRITAIN (p_x^{13})				
0	.08980	.08950	.09008	.08981
5	.06274	.06265	.06333	.06326
10	.04797	.04793	.04834	.04831
15	.07252	.07235	.07346	.07334
20	.11379	.11312	.11701	.11659
25	.09711	.09670	.09992	.09970
30	.07526	.07512	.07606	.07595
35	.05677	.05670	.05706	.05700
40	.03774	.03772	.03775	.03773
45	.02648	.02646	.02646	.02645
50	.03010	.03008	.02985	.02983
55	.01511	.01510	.01481	.01479
60	.01482	.01480	.01429	.01426
65	.01553	.01549	.01453	.01445
70	.01976	.01966	.01765	.01739
75	.01372	.01359	.01156	.01120
80	.01263	.01238	.00974	.00909
85	.00000	.00000	.00000	.00000

NOTE: All these runs employ rates computed by the averaging method.

observed in the movement approach. Compared with the linear approach, however, the life expectancies are lower in all regions (Table 9). This is intuitively unreasonable since lower mortality rates in all regions should not lead to lower life expectancy in all regions. The retention probabilities have been raised, however, and this raises the populations at risk of dying and thus, the number dying in earlier ages, leading to the lower life expectancies.

Another element in our comparison of exponential and linear methods is the non-negativity problem. This has already been discussed for the movement-based probability estimation methods (see section 3.5). The same toy model (equation 59) can be applied in the case of the transition-based probability estimation formulae (equations 68 and 70 for the linear case and equations 72 and 73 for the exponential case). Figure 5 presents the results for the region 2 retention probabilities where negative probabilities occur. The migration level is taken only up to 100 because at this level, the transition rate out of region 2 is 1, the maximum that can occur (and highly unlikely since some of the population will die). The problem is worse than in the movement-based case since in the formulae transitions are still being treated in part as if they were movements by being multiplied by n. Significant differences between linear and exponential formulae begin at lower mobility levels (between 5 and 10 rather than 10 and 15). The exponential formulae still ensure non-negativity.

The differences between linear and exponential formulae are very small for the transition based equations--only .03, .03, and .04 of a year for East Anglia, the South East, and the Rest of Britain, respectively (or about 12 days). These negligible differences are comparable to those observed in the movement approach analysis of the Dutch population system. However, the differences between movement-based linear and exponential formulae for the British population system are larger--.09, .07, and .20 of a year for the three regions (or about 44 days or one and a half months). This difference between the British and Dutch observations is probably due to the higher levels of migration in the U.K. compared with Holland.

Table 9. Great Britain, hybrid approach: regional life expectancies at birth (in years) and their distribution by region of residence (in percentages) based on one year deaths and migrant data, 1970. Probability estimation comparison.

	Method of probability and L-function estimation			
	Transition-based		Movement-based	
	Linear	Exponential	Linear	Exponential
Region	Run 3	Run 4	Run 7	Run 8
LIFE EXPECTANCIES BY REGION AT BIRTH				
East Anglia	72.74	72.71	72.43	72.34
South East	72.66	72.63	72.45	72.32
Rest of Britain	71.83	71.79	71.69	71.49
PERCENT OF LIFE SPENT BY PERSONS BORN IN EAST ANGLIA				
East Anglia	40.82	40.84	41.02	41.02
South East	26.27	26.27	25.78	25.97
Rest of Britain	32.91	32.89	33.19	33.01
PERCENT OF LIFE SPENT BY PERSONS BORN IN SOUTH EAST				
East Anglia	4.36	4.36	4.11	4.28
South East	63.99	63.99	64.55	63.98
Rest of Britain	31.65	31.64	31.35	31.74
PERCENT OF LIFE SPENT BY PERSONS BORN IN REST OF BRITAIN				
East Anglia	2.42	2.42	2.31	2.50
South East	16.87	16.87	16.31	17.51
Rest of Britain	80.71	80.70	81.38	79.99

Note: All runs presented in this table are based on the averaging rates method.

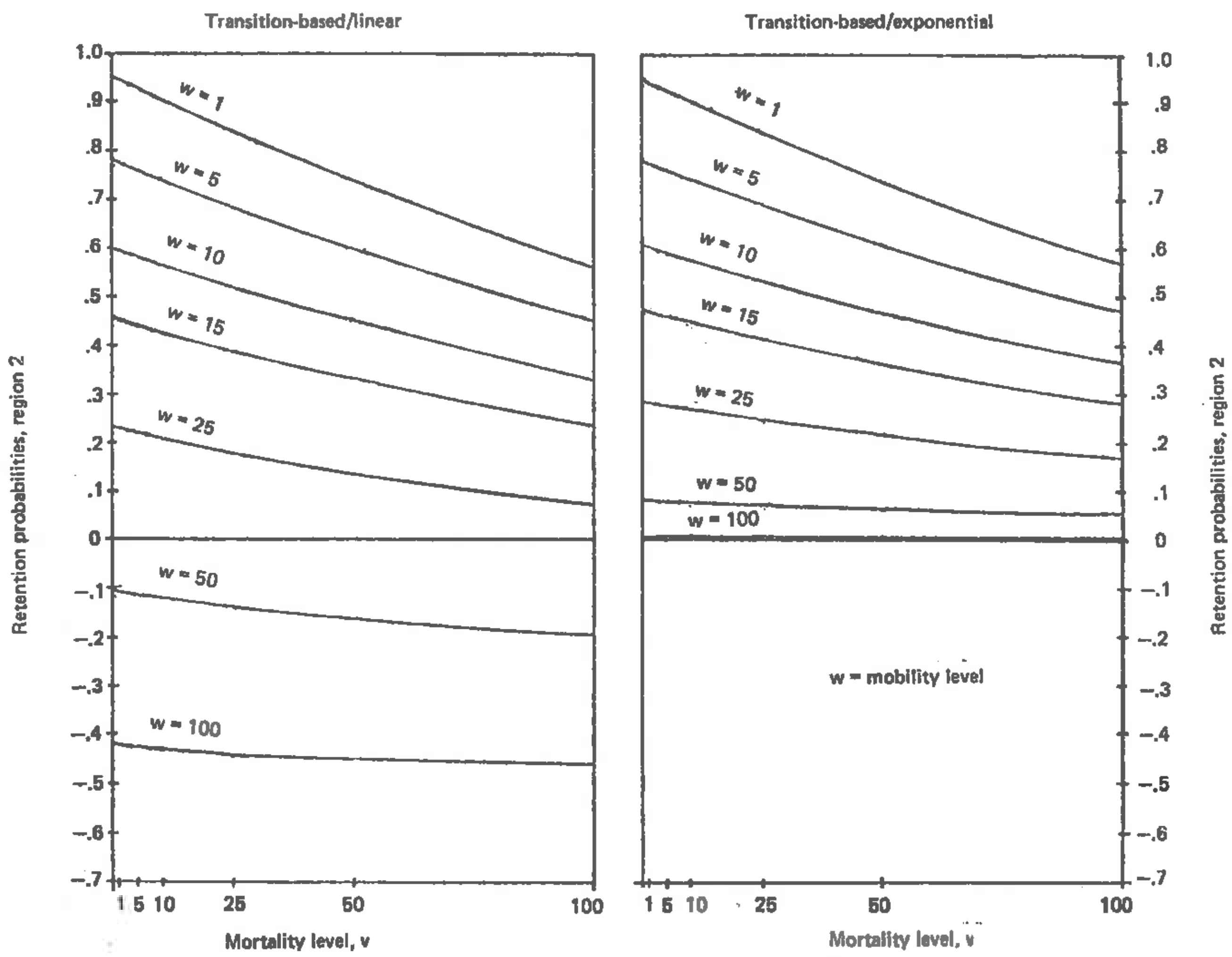


Figure 5. The influence of mortality and mobility levels on estimated probabilities, hybrid approach.

The second comparison that should be made is between the transition-based and movement-based formulae. Here the differences are larger. Using the transition formulae persons appear to live about .27 years or 100 days longer than when movement approach formulae are used (see Table 9), and slightly lower proportions of this time are spent in the region of birth [although there are exceptions (Rest of Britain), exponential method, comparing run 4 with run 8]. Our earlier (and Appendix A.3) analysis of the transition-based formulae suggest that both mortality and migration probabilities are underestimated: this means that the life expectancies should be lower in reality, and nearer the movement-based numbers, but that the percentage of lives spent outside the region of birth should be higher.

Are our probability estimation method comparisons confirmed for the other data sets, with their different migrant data and deaths data (in the case of data set two)? Table 10 extracts some selected probabilities from the data set two and three runs corresponding to those discussed in Table 8. The same comparisons can be made.

4.7 The Effect of Choice of Migration Rate Estimators

Two methods of computing migration rates were outlined earlier in the paper. Although our recommendation would always be to adopt the averaged rates method (section 4.3.2) where interpolative-iterative methods were not used, it is instructive to compare the results of the two alternatives in order to gauge how much error has resulted from the past use of the equivalent rate method.

The annual or annual equivalent migration rates from the South East to the Rest of Great Britain are plotted against age in Figure 6. For one-year data and annual rates the two rate schedules are very close except for ages 0 to 4 (plotted at age 0 on the graph) where the rates assuming equivalence are below those using the averaging method. The reason is that infant migrants are ignored in the equivalent method. The differences in rates estimation methods decrease the years spent outside the region of birth by an average of 0.73 percent over the three regions (Table 11).

Table 10. Great Britain, hybrid approach: transition probabilities out of East Anglia for ages 20 to 25 based on the five-year deaths and migrants data, 1966-71, and the combined one-year deaths and five-year migrants data.

Ages 20 to 25	Method of probability estimation			
	Transition-based		Movement-based	
	Linear	Exponential	Linear	Exponential
Data set two →	Run 11	Run 12	Run 15	Run 16
Data set three →	Run 19	Run 20	Run 23	Run 24
PROBABILITY OF DYING IN EAST ANGLIA ($p_{20}^{1\delta}$)				
Data set two	.00378	.00378	.00401	.00401
Data set three	.00382	.00382	.00406	.00406
PROBABILITY OF STAYING IN EAST ANGLIA (p_{20}^{11})				
Data set two	.86605	.86627	.86651	.86674
Data set three	.86602	.86623	.86647	.86670
PROBABILITY OF MIGRATING FROM EAST ANGLIA TO SOUTHEAST (p_{20}^{12})				
Data set two	.06819	.06808	.06704	.06689
Data set three	.06819	.06808	.06703	.06687
PROBABILITY OF MIGRATING FROM EAST ANGLIA TO REST OF BRITAIN (p_{20}^{13})				
Data set two	.06197	.06188	.06244	.06236
Data set three	.06197	.06187	.06245	.06237

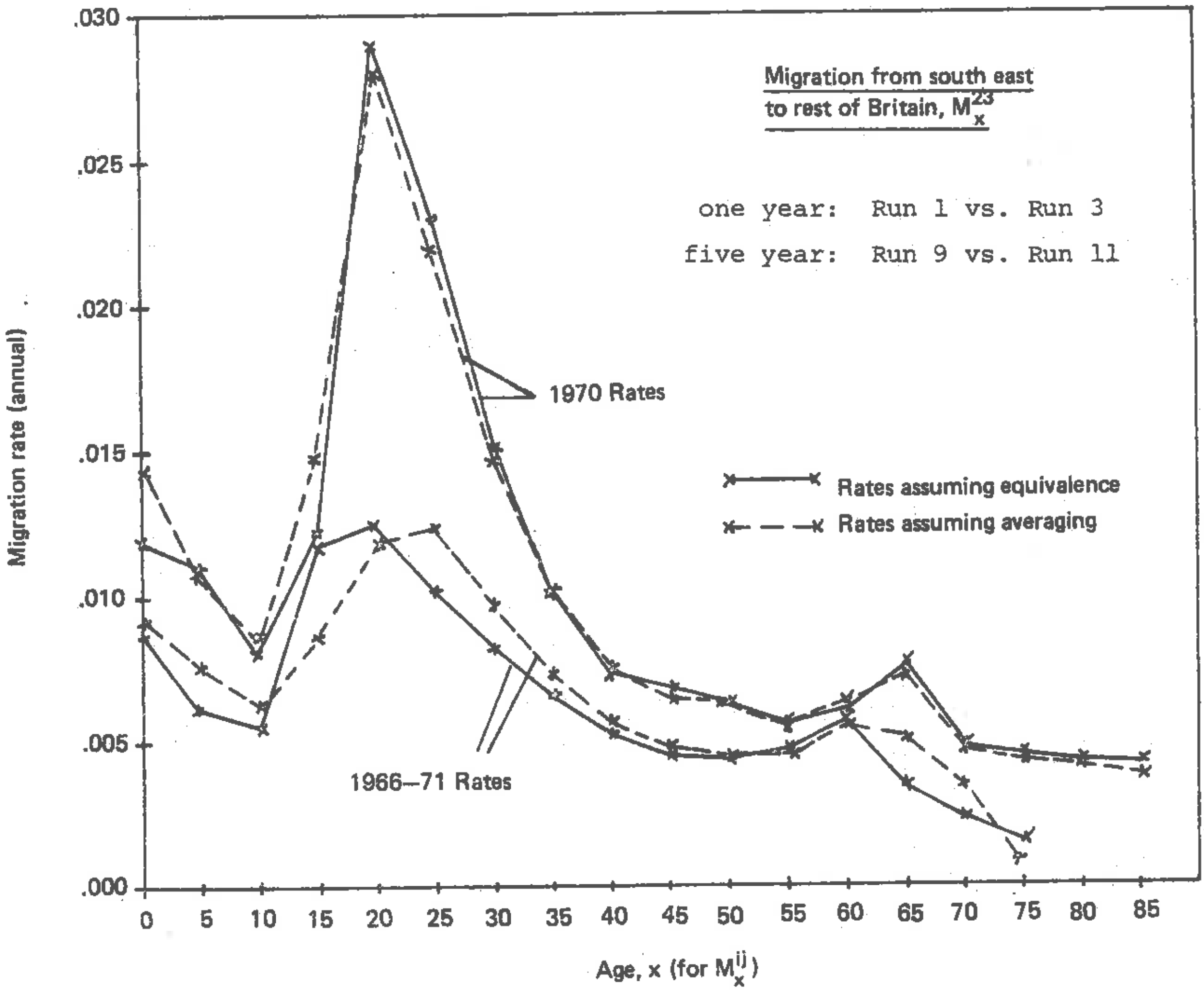


Figure 6. The two methods of rates estimation contrasted.

Table 11. Great Britain, hybrid approach: regional life expectancies at birth (in years) and their distribution by region of residence (in percentages) based on one-year deaths and migrant data and five-year deaths and migrant data, 1970. Rates estimation comparison.

	Migrant data period and method of rate calculation			
	One-year migrants data		Five-year migrants data	
	Equivalent rates	Averaged rates	Equivalent rates	Averaged rates
Region	Run 1	Run 3	Run 9	Run 11
LIFE EXPECTANCIES BY REGION AT BIRTH				
East Anglia	72.74	72.74	73.02	73.00
South East	72.66	72.66	72.66	72.66
Rest of Britain	71.83	71.83	71.76	71.78
PERCENT OF LIFE SPENT BY PERSONS BORN IN EAST ANGLIA				
East Anglia	41.98	40.82	57.78	56.02
South East	25.82	26.27	19.14	20.05
Rest of Britain	32.20	32.91	23.08	23.93
PERCENT OF LIFE SPENT BY PERSONS BORN IN SOUTH EAST				
East Anglia	4.28	4.36	3.57	3.64
South East	64.70	63.99	74.40	73.63
Rest of Britain	31.03	31.65	22.03	22.73
PERCENT OF LIFE SPENT BY PERSONS BORN IN REST OF BRITAIN				
East Anglia	2.38	2.42	1.75	1.88
South East	16.59	16.87	11.52	11.85
Rest of Britain	81.04	80.71	86.72	86.27

Note: All runs presented in this table are based on transition-based/linear method of probability and stationary population estimation.

The differences are much greater for rates computed from five-year migrant data (the annual equivalent rates shown in Figure 6 are derived by dividing the five-year migration rates by five). The effect of adopting the equivalent method is to shift the migration rate curve to the left by an average of about half the age interval (two and a half years). Infant migrations are again ignored but this has less of an effect in the five-year case because of a probable underestimation of infant migrants in the Great Britain accounts from which the five-year data derive (Rees 1980a). The years spent outside the region of birth are decreased by an average of just under one percent through the adoption of equivalent rates compared with averaged rates.

Rate estimation method does not, however, appear to have an effect on the computation of total life expectancy at birth.

4.8 The Effect of Choice of Alternative Period Length Data

Much more dramatic differences than those between rates estimation methods are apparent in the Figure 6 schedules and the Table 11 statistics, however. The five-year annual equivalent migration rates are lower at all ages than the one-year migration rates. (It is better to look at the averaged rates in this case because different equations were used in the one-year and five-year cases for the equivalent rates.) These differences are particularly dramatic for the ages at which migration rates are highest (15-30).

The differences in the migration rate schedules feed through to the probability matrices and so the rest of the life table. Tables 11 and 12 show that independent of rates estimation method or probability estimation method, the five-year based life tables exhibit far lower expectations of life outside regions of birth than the one-year based life tables. The comparisons in Table 11 show 10.3 percent less of regional life expectancies spent outside the region of birth in the five-year tables, and 10.1 percent less in the Table 11 comparisons. These results confirm the premonitions of Ledent (1978:133-137), and the findings of Rees (1979b:46-51).

Table 12. Great Britain, hybrid approach: regional life expectancies at birth (in years) and their distribution by region of residence (in percentages), using averaged rates and the exponential method. A comparison of one-year (1970) and five-year (1966-71) rates.

	Method or probability estimation and data period			
	Transition-based		Movement-based	
	One-year	Five-year	One-year	Five-year
	Run 4	Run 12	Run 8	Run 16
LIFE EXPECTANCIES BY REGION AT BIRTH				
East Anglia	72.41	72.98	72.34	72.74
South East	72.63	72.64	72.32	72.43
Rest of Britain	71.69	71.77	71.49	71.55
PERCENT OF LIFE SPENT BY PERSONS BORN IN EAST ANGLIA				
East Anglia	40.84	56.02	41.02	56.13
South East	26.27	20.05	25.97	19.87
Rest of Britain	32.89	23.93	33.01	24.00
PERCENT OF LIFE SPENT BY PERSONS BORN IN SOUTH EAST				
East Anglia	4.36	3.65	4.28	3.62
South East	63.99	73.62	63.98	73.50
Rest of Britain	31.64	22.73	31.74	22.88
PERCENT OF LIFE SPENT BY PERSONS BORN IN REST OF BRITAIN				
East Anglia	2.42	1.88	2.50	1.93
South East	16.87	11.85	17.51	12.29
Rest of Britain	80.70	86.27	79.99	85.77

Note: All runs presented in this table use averaged rates and the exponential method of probability/stationary population estimation.

By way of completeness one can note that the difference in mortality rates (1970 vs. 1966-71) or in number of age groups (18 vs. 16) does not affect the one-year - five-year migrant data comparison. Table 13 assembles life expectancies and distribution percentages in the by now familiar format for a comparison of similar data set two (five-year deaths data and five-year migrants data) with a synthetic data set three (one-year deaths data for 1970 and five-year migrants data for 1966-71 with 16 age groups). The mortality data have no effect on the regional distribution of life statistics. The differences in the total life expectations between the one-year and five-year cases are due to the differences in the migration schedules rather than to variations in the mortality pattern. Ideally, this should not be the case but in the hybrid approach this coupling unfortunately occurs.

4.9 Conclusions about the Hybrid Approach

To summarize, the elements of choice in calculating a multi-regional life table using the hybrid approach can be ranked in order of increasing impact on the life table statistics as follows.

(1) The choice of linear or exponential formulae has a very small impact. From Table 14 we can see that the average difference between linear and exponential $_0e$ values is .07 of a year (25 days) and retention percentages are shifted by an average of .25 of a percent.

(2) Using transition-based rather than movement-based equations for probability estimation shifts life expectancies by .21 of a year (77 days) but changes retention probabilities by only .13 of a percent.

(3) Employment of equivalent migration rates as opposed to averaged rates makes virtually no difference to life expectancies but does on average shift retention probabilities .46 of a percent.

Table 13. Great Britain, hybrid approach: regional life expectancies at birth (in years) and their distribution by region of residence (in percentages), using averaged rates, the exponential method, and five-year migrant data. A comparison of results using one-year (1970) and five-year (1966-71) deaths data.

	Method of probability estimation and deaths data period			
	Transition-based		Movement-based	
	Five-year deaths	One-year deaths	Five-year deaths	One-year deaths
Region	Run 12	Run 20	Run 16	Run 24
LIFE EXPECTANCIES BY REGION AT BIRTH				
East Anglia	72.98	72.86	72.74	72.63
South East	72.64	72.75	72.43	72.54
Rest of Britain	71.77	71.66	71.55	71.45
PERCENT OF LIFE SPENT BY PERSONS BORN IN EAST ANGLIA				
East Anglia	56.02	56.01	56.13	56.15
South East	20.05	20.10	19.87	19.92
Rest of Britain	23.93	23.89	24.00	23.93
PERCENT OF LIFE SPENT BY PERSONS BORN IN SOUTH EAST				
East Anglia	3.65	3.64	3.62	3.61
South East	73.62	73.70	73.50	73.60
Rest of Britain	22.73	22.67	22.88	22.79
PERCENT OF LIFE SPENT BY PERSONS BORN IN REST OF BRITAIN				
East Anglia	1.88	1.88	1.93	1.93
South East	11.85	11.89	12.29	12.33
Rest of Britain	86.27	86.23	85.77	85.73

Note: All runs presented in this table use averaged rates and the exponential method of probability/stationary population estimations. The migrant data all refer to the 1966-71 five-year period.

Table 14. Great Britain, hybrid approach; summary of the effects of the choices in terms of differences in life expectancy (years) and retention (percent).

Region	Linear versus Exponential	Transition-based versus Movement-based	Equivalent versus Averaged	One-year migrant versus Five-year migrant
LIFE EXPECTANCY (years) $({}_{0}^{i}e)$				
East Anglia	+.05	+.25	+.01	-.28
South East	+.06	+.20	+.00	-.11
Rest of Britain	+.09	+.19	+.01	+.02
Average, GB	+.07	+.21	+.01	-.14
RETENTION (percents) $({}^{i}\theta^{i} \times 100)$				
East Anglia	+.02	-.16	+1.22	-15.46
South East	+.24	-.16	+.74	-9.62
Rest of Britain	+.50	+.07	+.41	-5.57
Average, GB	+.25	-.13	+.77	-10.22

Note: Each figure in the table is the value of the difference between the statistic for all 24 runs of the life table program adopting the upper alternative of the pair of alternatives which head each column.

e.g.: Average life expectancy computed for East Anglia using the linear formulae = 72.77 years

Average life expectancy computed for East Anglia using the exponential formulae = 72.72 years

Difference = +0.05 of a year

(4) Use of five-year migrant data rather than one-year has a slight effect on life expectancy values of .14 of a year (51 days). However, all other effects are completely overshadowed by the dramatic effect of period of migration measurement on the retention statistics. When a switch is made in computing the three-region multiregional life table for Great Britain from one-year to five-year migrant data, the percentage of life estimated to be spent in the region of birth increases by an average of 10.22 percent or about 7.4 years.

It is difficult to make recommendations as to the best choice among the alternatives described in this section but nevertheless advice should be offered.

(1) It is a matter of indifference whether linear of exponential equations are chosen, except when migration rates are extremely high when the latter should be used.

(2) Although logically transition-based formulae should be used with transition data, the ones used here are less satisfactory than the equivalent movement formulae.

(3) Soundly computed migration rates employing the averaging principle should be employed.

(4) Two reasons can be suggested for using five-year transition data when they are available. First, if the life table age interval is five years, then these are the data which match the model transitions most closely. Second, there is evidence of much return migration over a five-year period not reflected in the model when one-year migrant data are used. This issue will be returned to later in the paper. There we will discuss whether this recommendation should be made less specific--that is, we should recommend only that T should equal n for transition data--or whether we should, in fact, prefer say a T = n = 5 system to a T = n = 10 or T = n = 1 system.

Given the difficulties and tensions inherent in a hybrid model that mixes impact transition data and movement methods, we turn to the third set of methods involving the transition approach.

5. THE TRANSITION APPROACH: USING SURVIVORSHIP RATES

5.1 General Discussion

The transition approach is one which takes into account both the "transition" nature of the multiregional life table and the definition of much migration data in transition terms. In contrast to the movement and hybrid approaches which are based on the assumption of equality of mortality rates and mobility rates in the life table and the observed populations, the transition approach instead assumes equality of life table and observed survivorship rates or proportions. The advantage of beginning the life table model with survivorship rates is that one begins with rates defined in a way very close to that of the probabilities of survival. Survivorship rates together with non-survivorship rates should, if properly defined, add up to unity (but never exceed it) and measure changes in state between fixed points in time (such as t and t+T) just as life table survival probabilities measures changes in state between fixed points in age (such as x and x+n).

Demographers have in the past attempted to measure such rates in countries where mortality data were lacking by comparing the populations in successive age groups in two consecutive censuses. Rogers (1975a) has adopted these methods to multi-regional populations and shown how survival probabilities can be estimated from survivorship rates. This he called the "Option 2" method. However, the method has been relatively neglected because it was proven to produce probability estimates that were negative or greater than one, and because it was felt that estimation of reliable survivorship rates from data was difficult.

In the next part of the paper, subsection 5.2, we show that reliable estimates of survivorship rates can be generated either through the prior estimation of multiregional population accounts (Rees and Wilson 1977) or through use of census migration data and start of period population data or through use of census migration, mortality, and population data. Care must be taken in handling "closure of the system" when survivorship rates and we show that this is best analyzed in a multiregional accounting framework even if multiregional population accounts are not themselves prepared. To date in the paper we have ignored the closure issue but in our discussion we show that it has a hidden influence on multiregional life tables estimated via the movement or hybrid approaches.

The discussion then leads on in subsection 5.3 to a consideration of ways of estimating survival probabilities from survivorship rates. We examine the slight modification proposed by Ledent (1978) to the original "Option 2" method, which appears to fail to overcome its disadvantages fully. Then an averaging method, suggested by Rees and Wilson (1977), is shown to work well and to avoid the negative or greater than one probabilities of the "Option 2" method. Finally, the refinement proposed by Ledent (1980a, 1980b) of the Rees and Wilson method in which a cubic spline function is used instead of averaging for interpolation.

As for the stationary population (L-statistics), we first consider a linear method identical to the one used in the movement and hybrid approaches. Such a linear approach explicitly appears in the "Option 2" method but, in this case, it implicitly involves more than the simple use of a linear formula. Such an observation led Ledent (1978) to devise a direct method for calculating the L_x-statistics which can be substituted for the linear approach when the survival probabilities are calculated independently from these statistics (as with the linear and cubic spline interpolation methods). Below, the linear and direct methods are compared.

After a detailed discussion of the way the British three-region data were handled, and the life table experiments set up, we outline in the latter half of this section of the paper the results of our experiments again working from right to left in the choice table (Figure 1).

5.2 Survivorship Rates: Methods of Estimation

5.2.1 An Accounts Framework

Although it will be unusual to compute survivorship rates directly from accounts, an accounts framework is essential for understanding the problems involved in handling less-than-accounts data.

Figure 7 sets out the accounts framework for the Great Britain regional system we are experimenting with. In this system transitions within the rest of the world are ignored so that it is not possible to close the system for life table purposes by adopting the rest of the world as a fourth region to add to three we consider within Great Britain. For population projection purposes this does not matter as we can close the system using the immigrant and emigrant information (see Rees 1979c). However, for life table purposes adjustments are made, either implicitly by default or explicitly by design.

Let us first define the "true" survivorship and non-survivorship rates in terms of the accounts variable. The survivorship rate, S_x^{ij}, as previously defined, is expressed in terms of accounts variables as follows

$$S_x^{ij} = K_x^{\varepsilon(i)\sigma(j)} \ / \ K_x^{\varepsilon(i).(.)}(t) \qquad (95)$$

where $K_x^{\varepsilon(i)\sigma(j)}$ is the number of persons in existence at the start of the period in region i in the age group with the age range x to x+n who survive in region j at the end of the period in the age group with ages ranging from x+n to x+2n, and $K_x^{\varepsilon(i).(.)}$ is the population initially resident at time t in region i in

Final state / Initial state		Survival σ in:				Non-survival δ in:				Total
		I			E	I			E	
		East Anglia 1	South East 2	Rest of Britain 3	Rest of World 4	East Anglia 1	South East 2	Rest of Britain 3	Rest of World 4	
Existence ε in: I	East Anglia, 1	$K_x^{\epsilon(1)\sigma(1)}$	$K_x^{\epsilon(1)\sigma(2)}$	$K_x^{\epsilon(1)\sigma(3)}$	$K_x^{\epsilon(1)\sigma(4)}$	$K_x^{\epsilon(1)\delta(1)}$	$K_x^{\epsilon(1)\delta(2)}$	$K_x^{\epsilon(1)\delta(3)}$	$K_x^{\epsilon(1)\delta(4)}$	$K_x^{\epsilon(1)*(*)}$
	South East, 2	$K_x^{\epsilon(2)\sigma(1)}$	$K_x^{\epsilon(2)\sigma(2)}$	$K_x^{\epsilon(2)\sigma(3)}$	$K_x^{\epsilon(2)\sigma(4)}$	$K_x^{\epsilon(2)\delta(1)}$	$K_x^{\epsilon(2)\delta(2)}$	$K_x^{\epsilon(2)\delta(3)}$	$K_x^{\epsilon(2)\delta(4)}$	$K_x^{\epsilon(2)*(*)}$
	Rest of Britain, 3	$K_x^{\epsilon(3)\sigma(1)}$	$K_x^{\epsilon(3)\sigma(2)}$	$K_x^{\epsilon(3)\sigma(3)}$	$K_x^{\epsilon(3)\sigma(4)}$	$K_x^{\epsilon(3)\delta(1)}$	$K_x^{\epsilon(3)\delta(2)}$	$K_x^{\epsilon(3)\delta(3)}$	$K_x^{\epsilon(3)\delta(4)}$	$K_x^{\epsilon(3)*(*)}$
E	Rest of World, 4	$K_x^{\epsilon(4)\sigma(1)}$	$K_x^{\epsilon(4)\sigma(2)}$	$K_x^{\epsilon(4)\sigma(3)}$	$K_x^{\epsilon(4)\sigma(4)}$	$K_x^{\epsilon(4)\delta(1)}$	$K_x^{\epsilon(4)\delta(2)}$	$K_x^{\epsilon(4)\delta(3)}$	$K_x^{\epsilon(4)\delta(4)}$	$K_x^{\epsilon(4)*(*)}$
Total		$K_x^{\epsilon(*)\sigma(1)}$	$K_x^{\epsilon(*)\sigma(2)}$	$K_x^{\epsilon(*)\sigma(3)}$	$K_x^{\epsilon(*)\sigma(4)}$	$K_x^{\epsilon(*)\delta(1)}$	$K_x^{\epsilon(*)\delta(2)}$	$K_x^{\epsilon(*)\delta(3)}$	$K_x^{\epsilon(*)\delta(4)}$	$K_x^{\epsilon(*)*(*)}$

<u>Notes</u>

K = count of people making transitions indicated by subscripts
ϵ = existence (lifestate)
σ = survival (lifestate)
δ = death (lifestate)
x = age which starts age group at time t for transitions

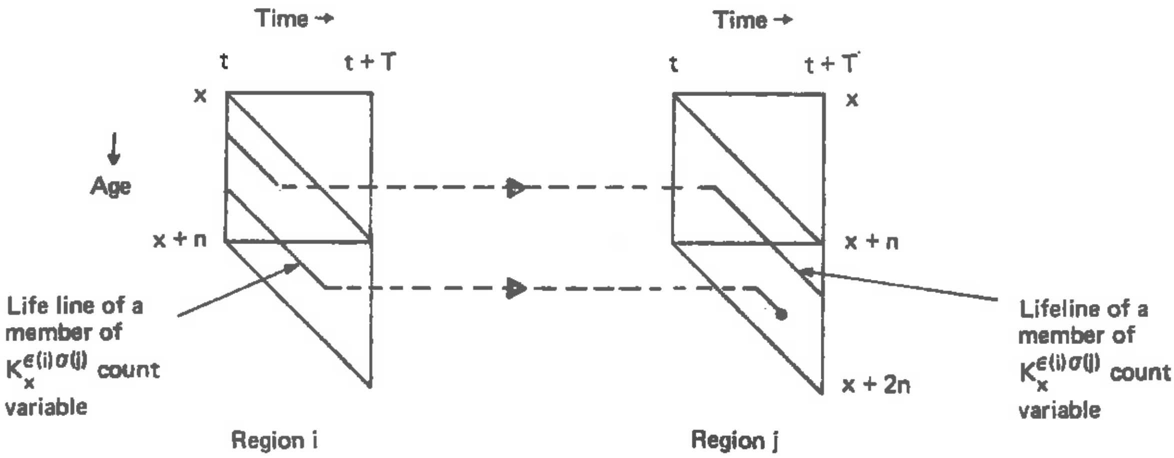


Figure 7. An accounts framework for a Great Britain, four-region system (after Rees 1978).

the age group with age range x to x+n. We assume in these definitions that the age group interval, n, is equal to the period length, T.

The non-survivorship rate, $S_x^{i\delta}$, is a sum of more disaggregated rates:

$$S_x^{i\delta} = \sum_k S_x^{i\delta(k)} \tag{96}$$

where

$$S_x^{i\delta(j)} = K_x^{\varepsilon(i)\delta(j)} \;/\; K_x^{\varepsilon(i)\cdot(\cdot)}(t) \tag{97}$$

The numerator $K_x^{\varepsilon(i)\delta(j)}$ is defined as the number of persons in existence at the start of the period, dying in either age group x to x+n or age group x+n to 2x+n (see Figure 7's Lexis diagram). Note that since

$$K_x^{\varepsilon(i)\cdot(\cdot)}(t) = \sum_k K_x^{\varepsilon(i)\delta(k)} + \sum_k K^{\varepsilon(i)\delta(k)} \tag{98}$$

so

$$\sum_k S_x^{ik} + S_x^{i\delta} = 1 \tag{99}$$

Examples of an accounts table and associated survivorship rates for x = 20 are given in Tables 15 and 16, respectively.

The definitions for the first and last age groups are simple adaptations of equations (95) and (97) for the first age group:

$$S_{-n}^{ij} = K_{-n}^{\beta(i)\sigma(j)} \;/\; K_{-n}^{\beta(i)\cdot(\cdot)}(t) \tag{100}$$

$$S_{-n}^{i\delta(j)} = K_{-n}^{\beta(i)\delta(j)} \;/\; K_{-n}^{\beta(i)\cdot(\cdot)}(t) \tag{101}$$

Table 15. An accounts table for x = 20, Great Britain, 1966-71.

Final state / Initial state			Survival in: Aged 25-29				Death in: Aged 20-24 or 25-29				Totals
			EA 1	SE 2	RB 3	RW 4	EA 1	SE 2	RB 3	RW 4	
Existence in: Aged 20-24	East Anglia	1	79,620	7,578	8,141	5,296	334	13	15	10	101,007
	South East	2	14,408	984,059	79,740	101,985	26	3,691	148	175	1,184,032
	Rest of Britain	3	9,440	84,510	2,074,448	107,473	18	145	8,056	208	2,284,298
	Rest of World	4	7,790	88,045	70,123	0	14	149	134	0	166,255
	Totals		111,258	1,164,192	2,232,452	214,554	392	3,998	8,353	393	3,735,592

SOURCE: Rees (1980a).

Table 16. Survivorship and non-survivorship rates for x = 20, Great Britain, 1966-71.

Final state / Initial state			Survival in: Aged 25-29				Death in: Aged 20-24 or 25-29				Totals
			EA 1	SE 2	RB 3	RW 4	EA 1	SE 2	RB 3	RW 4	
Existence in: Aged 20-24	East Anglia	1	.78826	.07502	.08060	.05243	.00331	.00013	.00015	.00010	1.00000
	South East	2	.01217	.83111	.06735	.08596	.00002	.00312	.00012	.00015	1.00000
	Rest of Britain	3	.00413	.03700	.90813	.04703	.00001	.00006	.00353	.00009	1.00000
	Rest of World	4	-	-	-	-	-	-	-	-	-
	Totals		-	-	-	-	-	-	-	-	

SOURCE: Rees (1980a).

so that

$$S_{-n}^{i\delta} = \sum_{k} S_{-n}^{i\delta(k)} \tag{102}$$

where the -n label indicates birth during the period and survival to or death before reaching the first age group, ages 0 to n, at the end of the period.

For the last age group, the corresponding equations are

$$S_{2}^{ij} = K_{2}^{\varepsilon(i)\sigma(j)} \;/\; K_{2}^{\varepsilon(i).(.)}(t) \tag{103}$$

$$S_{2}^{i\ (j)} = K_{2}^{\varepsilon(i)\sigma(j)} \;/\; K_{2}^{\varepsilon(i).(.)}(t) \tag{104}$$

and

$$S_{z}^{i} = \sum_{k} S_{z}^{i\delta(k)} \tag{105}$$

where persons are aged z and over at the start of the period and z+n and over at the end of the period, if they survive.

5.2.2 Closing the System by Default

In this subsection we define survivorship rates in a fashion that simulates the treatment of the rest of the world in the movement and hybrid approaches. This is not to suggest that this is the way survivorship rates are computed in those approaches. By constructing these alternate ways of closing the system we can get a feel for some of the biases introduced by the treatment of closure in the three approaches to life table construction.

The survivorship rates involving migration within the nation are defined as in equations (95) and (100). The non-survivorship rate is implicitly defined as in equations (96) and (97) although if full accounts were not available, the following approximation might be used

$$S_x^{i\delta} = S_x^{\delta(i)} = K_x^{\varepsilon(.)\delta(i)} \;/\; K_x^{\varepsilon(i).(.)} \tag{106}$$

In other words, observed deaths in a region are substituted for deaths to persons initially in that region at the start of the period. Then the survivorship and staying rates (retention rates) are defined implicitly as residuals

$$S_x^{ii} = 1 - \sum_{\substack{k \neq i \\ k \varepsilon I}} S_x^{ik} - S_x^{i\delta} \qquad \text{for } i \;\varepsilon\; I \tag{107}$$

where $j \;\varepsilon\; I$ means that region j belongs to the set of regions I making up the internal or national system. The equivalent accounts definition of retention rates as residuals would be

$$S_x^{ii} = 1 - \sum_{\substack{k \neq i \\ k \varepsilon I,E}} S_x^{ik} - S_x^{i\delta} \qquad \text{for } i \;\varepsilon\; I \tag{108}$$

where E refers to the external set of regions.

If we label accounts-based survivorship rates with a label (0), and subsequent definitions with (1), (2) and so on, we can show explicitly how the various definitions are connected. The retention rate from equation (107), $S_x^{ii}(1)$ is therefore

$$S_x^{ii}(1) = S_x^{ii}(0) + \sum_{k \varepsilon E} S_x^{ik}(0) \qquad \text{for } i \;\varepsilon\; I \tag{109}$$

In other words retention rates will, if the system is closed by default, be greater than their time values by emigration rate. In other words, it is assumed that people who otherwise leave the country remain in their region of birth.

5.2.3 Closing the System by Regarding Emigration as Death

In the life table model itself, normally only survivorship rates are defined and non-survivorship rates are obtained, if needed, as a residual

$$s_x^{i\delta} = 1 - \sum_k s_x^{ik} \tag{110}$$

Therefore, one method of closing the system would be to use equation (95) to define the survivorship rates, correctly, for internal transfers, and by implication to alter equation (110) to

$$s_x^{i\delta}(2) = 1 - \sum_{k \varepsilon I} s_x^{ik}(0) \tag{111}$$

so that in this case

$$s_x^{i\delta}(2) = s_x^{i\delta}(0) + \sum_{k \varepsilon E} s_x^{ik}(0) \tag{112}$$

This method assumes that when people emigrate out of the country, their lives are lost to the system, and they are treated as dead.

The problem with this method of closure is that the life expectancies for regional cohorts will behave with age in a way very different from those of any other life tables, as can be seen from comparing the $\underset{\sim}{s}_{20}^{\delta}(2)$ rates with those for $\underset{\sim}{s}_{20}^{\delta}(1)$ shown in Figure 8.

5.2.4 Closing the System by Reducing the Population at Risk

If the data to hand are in form of population accounts, it is possible to close the system by redefining the population at risk: emigrants (both survivors and non-survivors) are abstracted from the initial population, so that

By default

$$\underset{\sim}{S}_{20}(1) = \begin{pmatrix} .84050 & .07502 & .08060 \\ .01217 & .91710 & .06735 \\ .00413 & .03700 & .95521 \end{pmatrix} \qquad \underset{\sim}{S}^{\delta}_{20}(1) = \begin{pmatrix} .00388 \\ .00338 \\ .00366 \end{pmatrix}$$

By combining emigration and death

$$\underset{\sim}{S}_{20}(2) = \begin{pmatrix} .78826 & .07502 & .08060 \\ .01217 & .83111 & .06735 \\ .00413 & .03700 & .90813 \end{pmatrix} \qquad \underset{\sim}{S}^{\delta}_{20}(2) = \begin{pmatrix} .05611 \\ .08937 \\ .05072 \end{pmatrix}$$

By reducing the population at risk

$$\underset{\sim}{S}_{20}(3) = \begin{pmatrix} .83197 & .07918 & .08507 \\ .01332 & .90942 & .07369 \\ .00434 & .03883 & .95306 \end{pmatrix} \qquad \underset{\sim}{S}^{\delta}_{20}(3) = \begin{pmatrix} .00378 \\ .00357 \\ .00377 \end{pmatrix}$$

By using conditional survivorship rates

$$\underset{\sim}{S}_{20}(4) = \begin{pmatrix} .83205 & .07919 & .08508 \\ .01332 & .90957 & .07370 \\ .00434 & .03883 & .95314 \end{pmatrix} \qquad \underset{\sim}{S}^{\delta}_{20}(4) = \begin{pmatrix} .00368 \\ .00341 \\ .00369 \end{pmatrix}$$

By expanding to include the rest of the world

$$\underset{\sim}{S}_{20}(5) = \begin{pmatrix} .78826 & .07502 & .08060 & .05243 \\ .01217 & .83111 & .06735 & .08596 \\ .00413 & .03700 & .90813 & .04703 \\ .00003 & .00037 & .00030 & .99230 \end{pmatrix} \qquad \underset{\sim}{S}^{\delta}_{20}(5) = \begin{pmatrix} .00368 \\ .00341 \\ .00369 \\ .00700 \end{pmatrix}$$

Rows = origins 1 = East Anglia 2 = South East 3 = Rest of Britain
Columns = destinations 1 = East Anglia 2 = South East 3 = Rest of Britain

Figure 8. $\underset{\sim}{S}_{20}$ matrices for Great Britain consequent on various methods of closing the system.

$$S_x^{ij}(3) = K_x^{\varepsilon(i)\sigma(j)} \Big/ \left(K_x^{\varepsilon(i).(0)}(t) - \sum_{k\varepsilon E} K_x^{\varepsilon(i)\sigma(k)} - \sum_{k\varepsilon E} K_x^{\varepsilon(i)\delta(k)}\right) \qquad i,j \ \varepsilon \ I \tag{113}$$

$$S_x^{i\delta}(3) = \sum_{j\varepsilon I} K^{\varepsilon(i)\delta(j)} \Big/ \left(K_x^{\varepsilon(i).(.)}(t) - \sum_{k\varepsilon E} K_x^{\varepsilon(i)\sigma(k)} - \sum_{k\varepsilon E} K^{\varepsilon(i)\delta(k)}\right) \qquad i \ \varepsilon \ I \tag{114}$$

The survivorship and non-survivorship rates are then conditional on staying within the internal set of regions and they will still have the property

$$\sum_{k\varepsilon I} S_x^{ik}(3) + S_x^{i\delta}(3) = 1 \tag{115}$$

Compared with the first definition (by default) the diagonal elements of $\underset{\sim}{S}_x$ are lower and the off-diagonal elements are higher, as can be seen in the sample calculations for x = 20 (Figure 8). Spreading the adjustments over all terms in the $\underset{\sim}{S}_x$ matrices in the fashion suggested here preserves the pattern of population redistribution contained in the accounts whereas the first method reduces the amount of such redistribution by overestimating the diagonal terms.

Survivorship rates computed by this method will be inflated compared with those computed directly from the accounts by the ratios of their populations at risk:

$$S_x^{ij}(3) = S_x^{ij}(0) \frac{K_x^{\varepsilon(i).(.)}(t)}{\left(K_x^{\varepsilon(i).(.)}(t) - \sum_{k\varepsilon E} K_x^{\varepsilon(i)\sigma(k)} - \sum_{k\varepsilon E} K^{\varepsilon(i)\delta(k)}\right)} \tag{116}$$

$$i,j \ \varepsilon \ I$$

5.2.5 Closing the System Using Conditional Survivorship Rates

Frequently, the measurement of survivorship rates cannot start from an accounts table either because the necessary emigration data are unavailable or poorly estimated or because of problems in implementing the necessary computer programs (Plessis-Fraissard and Rees 1976, 1978 and Plessis-Fraissard 1980). Ledent (1980a, 1980b) has proposed a method which uses census migration data only to compute conditional survivorship rates. Overall survivorship rates are computed independently using mortality and mid-period population data, and these then are multiplied by the conditional survivorship rates to yield estimates of the survivorship rates.

The conditional survivorship rates are defined as

$$\bar{S}_x^{ij}(4) = K_x^{\varepsilon(i)\sigma(j)} \Big/ \sum_{k\varepsilon I} K_x^{\varepsilon(i)\sigma(k)} \qquad i,j\ \varepsilon\ I \qquad (117)$$

where $K_x^{\varepsilon(i)\sigma(k)}$ are entries in census migration tables. Sometimes the diagonal values, $K_x^{\varepsilon(i)\sigma(i)}$, are not published, but they can usually be computed using the census population figures and data from the immigrant tables:

$$K_{.x+n}^{\varepsilon(i)\sigma(i)} = K_{.x+n}^{\varepsilon(.)\sigma(i)} - \sum_{\substack{k\varepsilon I \\ k\neq i}} K_{.x+n}^{\varepsilon(k)\sigma(i)} - \sum_{k\varepsilon E} K_{.x+n}^{\varepsilon(k)\sigma(i)} \qquad i\ \varepsilon\ I \qquad (118)$$

assuming that

$$K_{x.}^{\varepsilon(i)\sigma(i)} = K_{.x+n}^{\varepsilon(i)\sigma(i)} \qquad (119)$$

which will be true when T = n. This assumption we have made throughout section 5 to date in order to keep the presentation straightforward, although the assumption is not strictly necessary.

We now show how s_x^{ij} estimates might be computed from estimates of overall survivorship rates, $s_x^{i\sigma}$. However, the procedure actually used is to introduce these as overall survival probabilities and we describe this method afterwards.

Given good estimates of the $\underset{\sim}{s}_x^{\sigma}$ matrices computed from the accounts or from conventional mortality data, we can then compute the survivorship rates as

$$s_x^{ij}(4) = \bar{s}_x^{ij}(4) \cdot s_x^{i\sigma}(4) \tag{120}$$

with the overall survivorship rates being equivalent to $s_x^{i\sigma}(0)$ where

$$\begin{aligned} s_x^{i\sigma}(0) &= \sum_k s_x^{ik}(0) \\ &= 1 - \sum_k s_x^{i\delta(k)}(0) \end{aligned} \tag{121}$$

How then do the $s_x^{ij}(4)$ compare with our original accounts based definitions, $s_x^{ij}(0)$? We can express equation (120) as

$$s_x^{ij}(4) = \frac{K_x^{\varepsilon(i)\delta(j)}}{\sum\limits_{k\varepsilon I} K_x^{\varepsilon(i)\sigma(k)}} \cdot \frac{\sum\limits_k K_x^{\varepsilon(i)\sigma(k)}}{K_x^{\varepsilon(i).(.)}(t)} \tag{122}$$

Using this definition together with that of $s_x^{ij}(0)$ in equation (95), we obtain

$$s_x^{ij}(4) = s_x^{ij}(0) \frac{\sum\limits_k K_x^{\varepsilon(i)\sigma(k)}}{\sum\limits_{k\varepsilon I} K_x^{\varepsilon(i)\sigma(k)}} \qquad \text{for } i,j \ \varepsilon \ I \tag{123}$$

Thus, the survivorship rates are inflated, in comparison to those in the original accounts by the ratio of all survivors (including emigrants) of the region i, age x to x+n population to just those surviving within the internal set of regions. However, the estimates should be very close if the $S_x^{i\sigma}(4)$ rates are close to the $S_x^{i\sigma}(0)$ rates. In Figure 8, this has been assumed and the $\mathbf{S}_{20}(4)$ estimates are very close to the $\mathbf{S}_{20}(3)$ estimates.

In practice, for many population systems, overall survivorship rates cannot be computed directly from population accounts because those accounts have not been estimated. So, instead we need a method for estimating such overall survivorship rates, $\mathbf{S}_x$, from conventional mortality data, so that we can compute the survivorship rates

$$\mathbf{S}_x = \overline{\mathbf{S}}_x \, \mathbf{S}_x^{\sigma} \tag{124}$$

For the purposes of this task, note that if the Markovian assumption holds, we have the following relationship between the basic inputs of the movement and transition approaches, assuming further the use of a linear integration formula for calculating $\mathbf{L}_x$ (Ledent 1978, equation 67, page 48):

$$\mathbf{S}_x = (\mathbf{I} + \tfrac{n}{2}\,\mathbf{M}_{x+n})^{-1}\,(\mathbf{I} - \tfrac{n}{2}\,\mathbf{M}_x) \tag{125}$$

and therefore that

$$\overline{\mathbf{S}}_x \, \mathbf{S}_x^{\sigma} = (\mathbf{I} + \tfrac{n}{2}\,\mathbf{M}_{x+n})^{-1}\,(\mathbf{I} - \tfrac{n}{2}\,\mathbf{M}_x) \tag{126}$$

Premultiplying both sides of equation (126) by $(\mathbf{I} + \tfrac{n}{2}\,\mathbf{M}_{x+n})$, we obtain

$$(\mathbf{I} + \tfrac{n}{2}\,\mathbf{M}_{x+n})\overline{\mathbf{S}}_x \, \mathbf{S}_x^{\delta} = (\mathbf{I} - \tfrac{n}{2}\,\mathbf{M}_x) \tag{127}$$

Next we premultiply this equality by a row vector of ones, $\{i\}'$. Since $\{i\}'\underset{\sim}{M}_x$ and $\{i\}'\underset{\sim}{M}_{x+n}$ are row vectors of conventional mortality rates, $\{M_x^\delta\}'$ and $\{M_{x+n}^\delta\}'$, respectively, and since $\{i\}'\underset{\sim}{\overline{S}}_x \underset{\sim}{S}_x^\sigma$ is a row vector, $\{S_x^\sigma\}'$, whose typical element is $s_x^{i\sigma}$, we can write equation (127) as

$$\{S_x^\sigma\}' + \frac{n}{2}\{M_{x+n}^\delta\}' \underset{\sim}{\overline{S}}_x \underset{\sim}{S}_x^\sigma = \{i\}' - \frac{n}{2}\{M_x^\delta\}' \tag{128}$$

which becomes, after transposition,

$$\{S_x^\sigma\} + \frac{n}{2} \underset{\sim}{S}_x^\sigma \underset{\sim}{\overline{S}}_x'\{M_{x+n}^\delta\} = \{i\} - \frac{n}{2}\{M_x^\delta\} \tag{129}$$

Then, if we premultiply both sides of (129) by $[\underset{\sim}{S}_x^\sigma]^{-1}$, we obtain

$$\{i\} + \frac{n}{2} \underset{\sim}{\overline{S}}_x'\{M_{x+n}^\delta\} = \{S_x^\sigma\}^{-1}\{i\} - \frac{n}{2}\{S_x^\sigma\}^{-1}\{M_x^\delta\} \tag{130}$$

so that

$$\{S_x^\sigma\}^{-1} = (\underset{\sim}{I} - \frac{n}{2} \underset{\sim}{M}_x^\delta)^{-1} (\underset{\sim}{I} + \frac{S}{2} \underset{\sim}{\overline{S}}_x' \underset{\sim}{M}_{x+n}^\delta)\{i\} \tag{131}$$

where $\{S_x^\sigma\}^{-1}$ denotes a vector whose elements are the reciprocals of the elements of $\{S_x^\sigma\}$.

We can estimate $\underset{\sim}{M}_x^\delta$ and $\underset{\sim}{M}_{x+n}^\delta$ in the normal way from conventional mortality and can measure $\underset{\sim}{\overline{S}}_x$ from census migration and population data, and so find the S_x^σ terms. Equation (131) ensures the conversion of place of death rates into place of initial residence rates through the use of the $\underset{\sim}{\overline{S}}_x$ matrix. This equation operates in a manner analogous to the migration and death equations in an accounts-based model, and is based on the hypothesis that people die at the rate of the region they move to.

Note that the above derivation of the survivorship matrices $\underset{\sim}{S}_x$ from the knowledge of the conditional survival matrices $\underset{\sim}{\bar{S}}_x$ and the conventional mortality rates $\underset{\sim}{M}_x$ and $\underset{\sim}{M}_{x+n}$ is only valid when $n = T$. In case $n \neq T$, one will then adopt an alternative closure method or, if the same type of closure is required, one will follow the multiregional life table construction method proposed by Ledent (1980b): see section 5.8 below.

5.2.6 Closing the System by Including the Rest of the World

A final method of closing the system would, of course, be to include the flows within the rest of the world that were set to zero in Table 15. Approximate survivorship rate values have been computed for the rest of the world row and are displayed in Figure 8 to show that the numbers can be estimated.

The usual objection to such a procedure is that the rates of migration from the rest of the world to the internal regions do not reflect the time rate of return migration by persons born within the country. However, this applies to any of the interregional flows, and ideally we should begin all multiregional life tables with migration data cross-classified by region of birth as Ledent (1980c) points out.

5.2.7 Summary of Survivorship Rate Choices

Some five alternative methods of survivorship rate calculation have been presented. The second and fifth choices are not considered as they lead to life tables of very different nature than the ones we have been considering so far. The first choice, although easy to put into operation, leads to overestimates of retention rates and to overestimates of the amount of life spent in the region of birth. We therefore concentrate on the third and fourth methods, the third which presupposes that a set of population accounts is available and the fourth which can be used with census migration data and mortality statistics.

5.3 Methods of Converting Survivorship Rates into Survival Probabilities

5.3.1 The "Option 2" Method

A first method, denoted originally as the "Option 2" method, was proposed by Rogers (1975:85-88). The following linkage between survivorship rate matrices, $\underset{\sim}{s}_x$, and survival probability matrices, $\underset{\sim}{p}_x$, was derived assuming the linear integration formula for $\underset{\sim}{L}_x$ (Rogers 1975, equation A.14, p. 85)

$$\underset{\sim}{s}_x = [\underset{\sim}{I} + \underset{\sim}{p}_{x+n}]\ \underset{\sim}{p}_x\ [\underset{\sim}{I} + \underset{\sim}{p}_x]^{-1} \qquad (132)$$

Equation (132) can be rearranged to yield either a formula for $\underset{\sim}{p}_x$ or one for $\underset{\sim}{p}_{x+n}$. Postmultiplying both sides of equation (132) by $[\underset{\sim}{I} + p_x]$ and rearranging we obtain

$$\underset{\sim}{p}_x = [\underset{\sim}{I} + \underset{\sim}{p}_{x+n} - \underset{\sim}{s}_x]^{-1}\ \underset{\sim}{s}_x \qquad (133)$$

which suggests $\underset{\sim}{p}_x$ can be obtained if $\underset{\sim}{p}_{x+n}$ and $\underset{\sim}{s}_x$ are known. Alternatively, one can rewrite equation (132) as

$$\underset{\sim}{p}_{x+n} = \underset{\sim}{s}_x[\underset{\sim}{I} + \underset{\sim}{p}_x]\underset{\sim}{p}_x^{-1} - \underset{\sim}{I} \qquad (134)$$

which suggests that $\underset{\sim}{p}_{x+n}$ can be derived if $\underset{\sim}{s}_x$ and $\underset{\sim}{p}_x$ are known.

In the first case (the case originally considered by Rogers 1975a), if an estimate of $\underset{\sim}{p}_{z-n}$ is available then the matrices $\underset{\sim}{p}_x$ (for x = z-2n,z-3n,...,n,0) can be obtained. In the second case (suggested by Ledent 1978), if $\underset{\sim}{p}_0$ is known, then the series of matrices $\underset{\sim}{p}_x$ (for x = n,2n,...,2-n) can be obtained from knowledge of the survivorship matrices $\underset{\sim}{s}_x$ (for x = 0,n,2n, ...,z-2n).

Here we use the second way and compute the $\underset{\sim}{p}_{x+n}$'s proceeding from younger to older ages. The first $\underset{\sim}{p}_x$, $\underset{\sim}{p}_0$, can be found from the relations assumed in the linear integration of $\underset{\sim}{L}_x$

$$\underset{\sim}{s}_{-n} = \underset{\sim}{L}_0 \, \underset{\sim}{\ell}_0^{-1} \tag{135}$$

$$\underset{\sim}{s}_{-n} = \frac{1}{2}(\underset{\sim}{\ell}_0 + \underset{\sim}{\ell}_n)\underset{\sim}{\ell}_0^{-1} \tag{136}$$

which, if $\underset{\sim}{\ell}_0 = \underset{\sim}{I}$ and $\underset{\sim}{\ell}_n = \underset{\sim}{p}_0 \, \underset{\sim}{\ell}_0$, becomes

$$\underset{\sim}{s}_{-n} = \frac{1}{2}(\underset{\sim}{I} + \underset{\sim}{p}_0) \tag{137}$$

so that

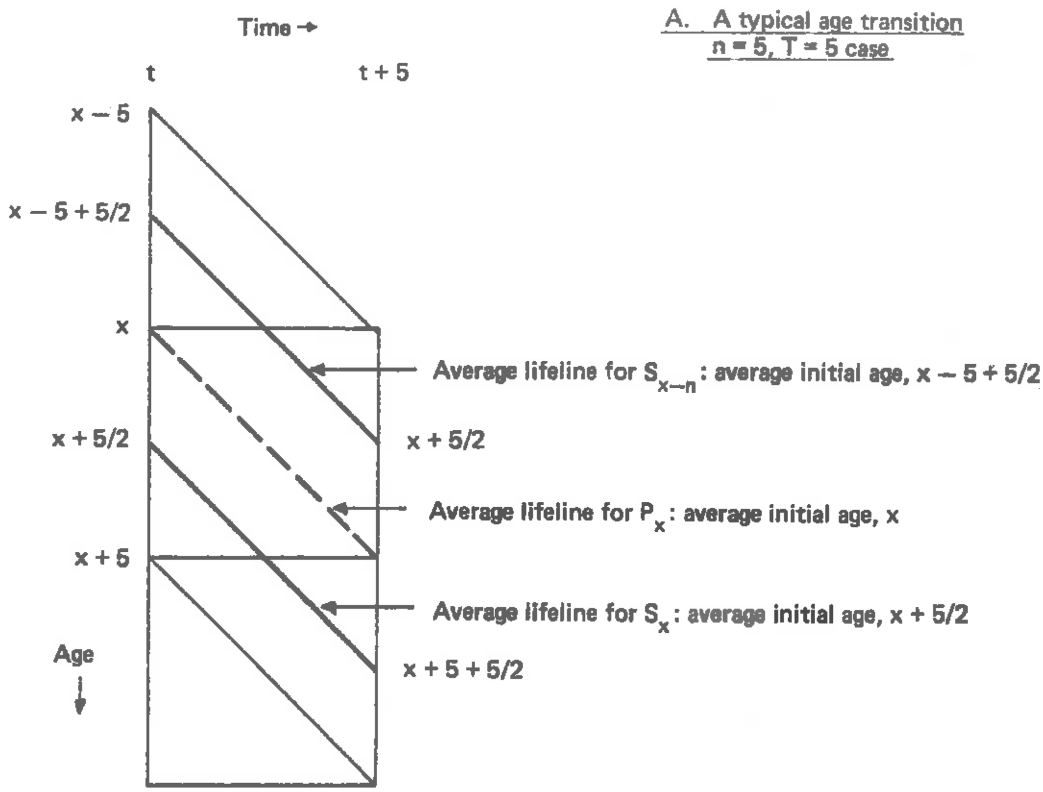
$$\underset{\sim}{p}_0 = 2\underset{\sim}{s}_{-n} - \underset{\sim}{I} \tag{138}$$

when $\underset{\sim}{p}_0$ can be estimated by setting $\underset{\sim}{s}_{-n}$ equal to the observed $\underset{\sim}{s}_{-n}$.

However, the "Option 2" approach has its drawbacks. Because the method relies on a formula linking statistics of two consecutive age intervals, estimation errors made for a given age interval are passed on to the next and so they cumulate. Also, it is critically dependent on the initial $\underset{\sim}{p}_0$ or $\underset{\sim}{p}_{z-n}$ estimate. Equation (138) turns out to be a poor estimate for $\underset{\sim}{p}_0$, and undesirable estimates outside the 0, 1 range are generally obtained even if the observed survivorship matrices are free from any measurement error. In addition, the method assumes that the age interval, n, is equal to the time period length, T.

5.3.2 The Linear Interpolation Method

As an alternative to the "Option 2" method, Rees and Wilson (1977) suggest that a simple solution to the problem of deriving survival probabilities from survivorship rates is to average the latter (see Figure 9a)



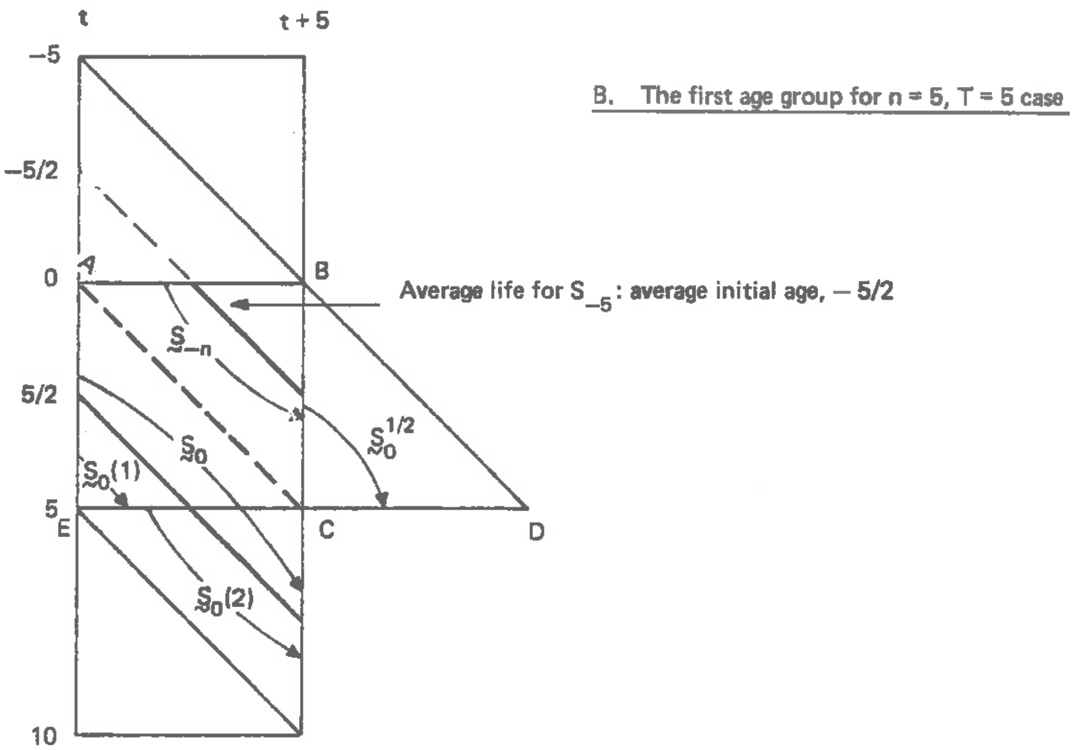
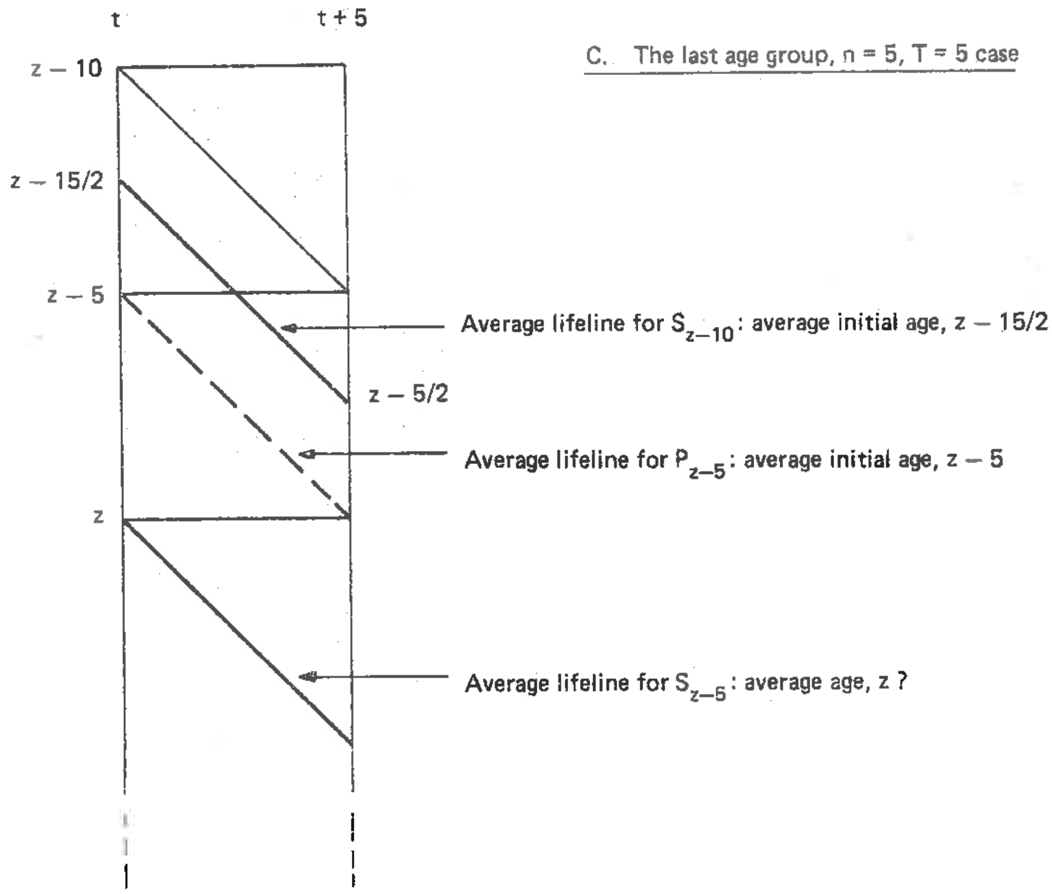


Figure 9. Lexis diagrams illustrating the linear interpolation method applied to survivorship rates.



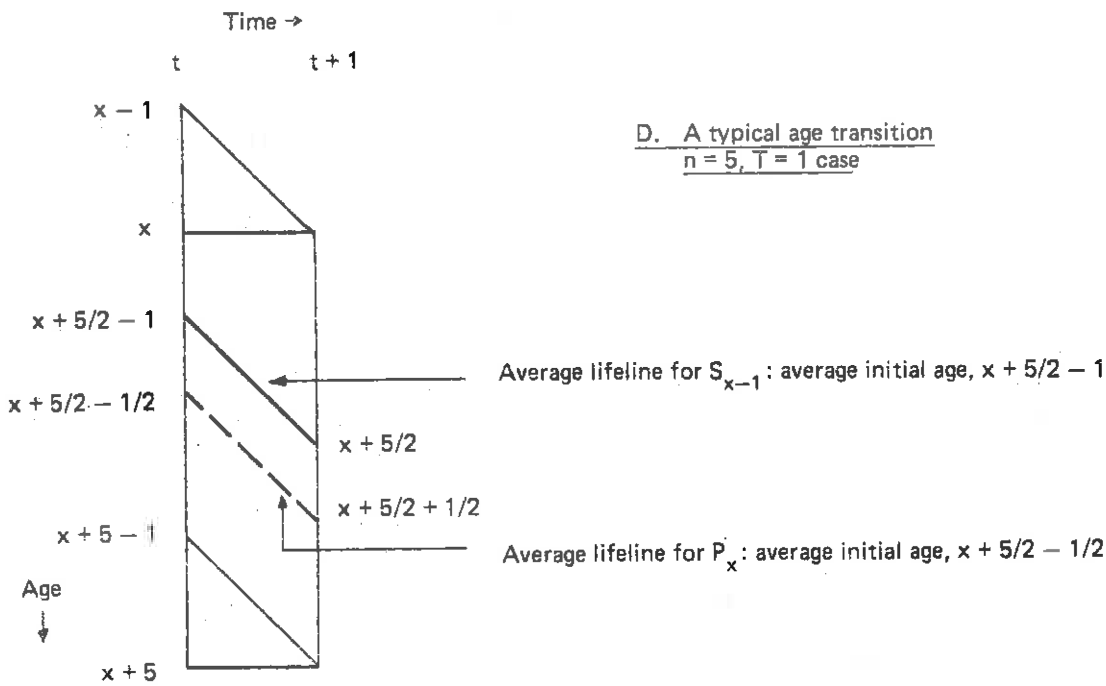


Figure 9. Continued.

$$\underset{\sim}{p}_x = \frac{1}{2}[\underset{\sim}{S}_{x-n} + \underset{\sim}{S}_x] \tag{139}$$

The logic behind such a procedure is that to interpolate all one has to do is assume average ages at the start of the time interval (for the populations involved in the transitions) for the $\underset{\sim}{S}_{x-n}$, $\underset{\sim}{S}_x$, and $\underset{\sim}{p}_x$ rates, and then to assume a linear function for the rate between the average age for $\underset{\sim}{S}_{x-n}$ and $\underset{\sim}{S}_x$, from which the $\underset{\sim}{p}_x$ rate can be read using its assumed average age.

Indeed (139) does not apply to the first age group. In practice, we take

$$\underset{\sim}{p}_0 = \frac{1}{2}(\underset{\sim}{S}^2_{-n} + \underset{\sim}{S}_0) \tag{140}$$

The $\underset{\sim}{S}_{-n}$ rates matrix is squared so that it refers to a T-year period equivalent rather than a $\frac{T}{2}$-year period, which the average infant born in a time interval will experience (see Figure 9b).

The last age transition is also problematical. It is difficult to know what average age to assume for the $\underset{\sim}{S}_{z-n}$ rates. This will depend on whether we use the $\underset{\sim}{S}_{z-n}$ rates as referring only to age group z-n to z (at the start of the time interval of survivorship) or whether we use $\underset{\sim}{S}_{z-n}$ rates as referring to age group z-n and above. In the former case, an average of z can be assumed and equation (139) used; in the latter case, the average should be higher and in Figure 9d, an average age of $z+\frac{5}{2}$ is assumed. Then, we would obtain p_{z-n} from

$$\underset{\sim}{p}_{z-n} = \frac{1}{3}(2\underset{\sim}{S}_{z-2n} + \underset{\sim}{S}_{z-n}) \tag{141}$$

All of the above reasoning assumes, as in the "Option 2" method, that n = T. However, the above "averaging" procedure can be easily generalized into a linear interpolation procedure (see Figure 9d drawn for a situation where n = 5 and T = 1).

Survivorship rates can be measured only by age groups at the end of the period (at time t+1 when a census has been taken). The survivorship rates can then be plotted at the average ages so determined and the probability read off the graph (Figure 10).

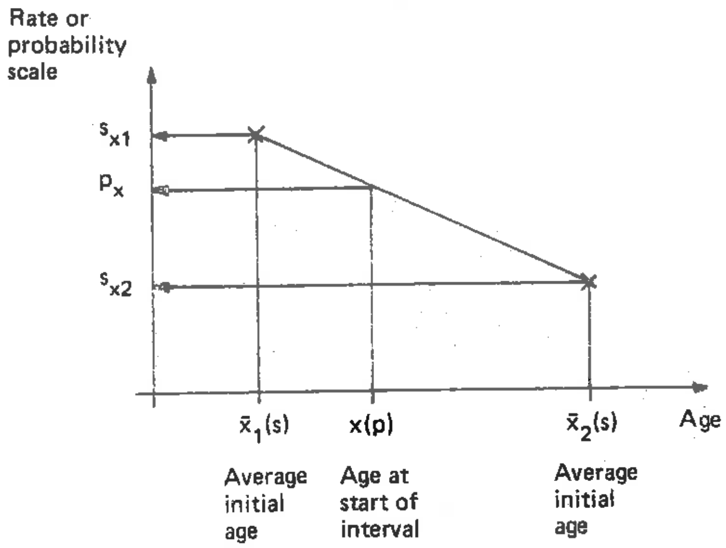


Figure 10. A graph illustrating how the linear interpolation for estimating p_x survival probabilities from S_x survivorship rates works.

Normally, we will wish to carry out such interpolation numerically rather than graphically. Let $\bar{x}_1(s)$ be the average age associated with $\underset{\sim}{S}_{x_1}$, the survivorship rates "younger" than $\underset{\sim}{p}_x$; let $\bar{x}_2(s)$ be the average associated with $\underset{\sim}{S}_{x_2}$, the "older" survivorship rates, and let x(p) be the associated with the $\underset{\sim}{p}_x$ probabilities. Then, if we assume a linear function between $\underset{\sim}{S}_{x_1}$ and $\underset{\sim}{S}_{x_2}$, $\underset{\sim}{p}_x$ is given by

$$p_x = S_{x_1} + \left(\frac{S_{x_2} - S_{x_1}}{\bar{x}_2(s) - \bar{x}_1(2)} \right) [x(p) - \bar{x}_1(s)] \tag{142}$$

where the first term is the initial constant, the second term the slope and the third term the age interval to which the slope is applied. Note that, by substituting $x(p) = \frac{\bar{x}_1(s) + \bar{x}_2(s)}{2}$ into (142), we simply rederive equation (139).

Note that the averaging method proposed in (139) relies on an arithmetic average. Alternatively, the averaging could be geometric rather than arithmetic

$$p_x = [S_x \; S_{x-n}]^{\frac{1}{2}} \tag{143}$$

In the case of the first age group we would have

$$p_0 = [S_0 \; S_{-n}^2]^{\frac{1}{2}} = S_0^{\frac{1}{2}} \; S_{-n} \tag{144}$$

This estimate is better than that of (140) since $S_0^{\frac{1}{2}}$ is a reasonable estimate of the transitions that occur in ΔBCD in Figure 9c taking infants from the t+5 line to their fifth birthdays, although it will tend to overestimate survival chances of the infant from 0 to n since $S_0^{\delta}(2)$ is undoubted less than $S_0^{\delta}(1)$ (see Figure 9c). However, the true p_0 is quite close to the equation (144) expression and we adopt it in our linear interpolation method in preference to equation (140).

Again in the case $n \neq T$, such a method could be generalized into an interpolation method based on

$$p_x = S_{x_1} \, e^{\left[\ln\left(\frac{S_{x_2}}{S_{x_1}}\right) / (x_2(s) - x_1(s)) \right] [\bar{x}(p) - x_1(s)]} \tag{145}$$

It must be admitted that the linear interpolation method is crude, particularly in its assumptions about average ages. An additional real world complication is that different average ages may be applicable to different transitions. If this refinement were to be added it would be necessary to interpolate between non-survivorship rates as well and then constrain the resulting probabilities so that they add to 1. However, crude although this linear interpolation, including the variant just described, is, it does have the considerable advantage over the "Option 2" method in guaranteeing that if the $\underset{\sim}{S}_x$ rates have been properly measured and lie between 0 and 1, then the $\underset{\sim}{p}_x$ probabilities will do so also. The method is clearly more satisfactory when n is small (say 1) than when large (say 5). It is also fairly logical to improve the method by fitting a function to the observed survivorship rates that is better than the point to point linear one which the averaging method assumes.

5.3.3 Cubic Spline Interpolation Method

Such a method has been developed by Ledent (1980b) using a cubic spline function similar to that employed by McNeil, Trussell and Turner (1977). Figure 11 shows how the technique works in principle. Plotted on a graph of survivorship rate versus age are the values for S_{-5}^{12}, S_{0}^{12},..., S_{20}^{12} (East Anglia to South East migratory transition).

The age scale used on the graph refers to age at the beginning of an x to x+5 transition associated with a p_x probability. Survivorship rates, S_x, are plotted at x+2.5 years on this scale because the "S_x persons" making a transition are initially 2.5 years older than the "p_x persons". An additional age scale has been added at the bottom of the diagram to indicate the average age at which transitions are assumed to take place either for "S_x persons" or "p_x persons". The straightlines linking these data points represent the function used by implication when the arithmetic averaging approach is employed. This works fairly well when the survivorship function is either decreasing fairly smoothly as it is between age 25 and 75 or

Notes

(1) Age at start of age interval in which p-transitions take place

(2) S_x rates plotted at age indicated above at x + 2.5

(3) p_x rates plotted at age indicated above

(4) The S_{-5} matrix has been squared before this value is plotted

(5) This is the observed rate not the conventional rate

(6) Alternative age scale indicating average age at which transitions are assumed to take place

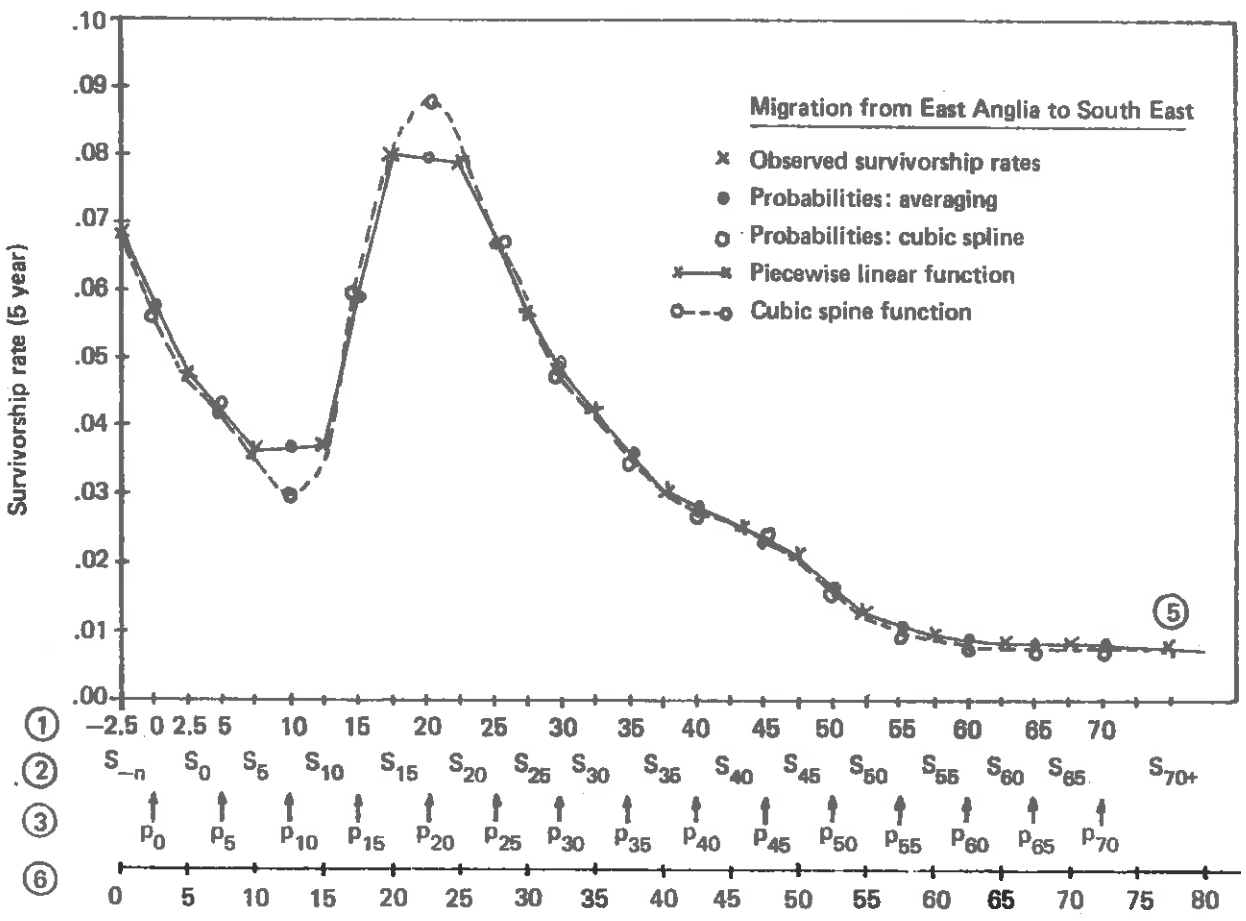


Figure 11. A survivorship rate profile showing the difference between piecewise linear interpolation and cubic spline interpolation.

increasing smoothly as it is between 15 and 20. At the turning points of the function, however, the piecewise linear function behaves less well.

The alternative is to fit a smoothing function using the cubic spline method. Such a function is indicated by the pecked line on Figure 11. The required p_x probabilities may then be found as the value of the ordinate at ages $x+\frac{n}{2}$ for $x = 0,n,2n,\ldots,z-n$. One difficulty that sometimes arises in the case of low mortality levels for the younger age groups is that application of the cubic spline interpolation procedure may yield estimates of the survivorship rates and survival probabilities such that $\sum_j s_x^{ij} > 1$ or $\sum_j p_x^{ij} > 1$.

To get round this difficulty one could either constrain the p estimates to add to one

$$p_x^{ij}(b) = p_x^{ij}(a)\ k \tag{146}$$

where

$$\sum_j p_x^{ij}(a)k + p_x^{i\delta}(a) = 1 \tag{147}$$

and

$$k = 1 \ / \ \sum_j p_x^{ij}(a) + p_x^{i\delta} \tag{148}$$

where $p_x^{ij}(a)$ and $p_x^{i\delta}(a)$ are the initial estimates and $p_x^{ij}(b)$ the adjusted estimates. Or, the error could be loaded onto the retention probabilities by using

$$p_x^{ii} = 1 - \sum_{j\neq i} p_x^{ij} - p_x^{i\delta} \tag{149}$$

instead of the interpolation method.

A comparison of the survival probabilities resulting from the two interpolation methods is given in Table 17. Figure 11 plots the full age range for the East Anglia to South East transition. The major differences occur, as one would expect, at the major turning points of the survivorship/probability schedules: at the p_{10} local minimum and the p_{20} maximum. The cubic spline values are lower for p_{10} migration terms and higher for p_{20} migration terms than the corresponding averaging method values. The cubic spline retention values for p_{10} are higher and for p_{20} are lower than the values for the other method. The values for the other age intervals are fairly close and successive differences tend to cancel out, being different in sign. Thus, although the cubic spline interpolated values are clearly closer to the "truth", we should not expect the two different interpolation methods to give very different life expectancy results. If a good cubic spline smoothing subroutine is available, it should be used; otherwise the averaging method will be perfectly adequate. Note that special care should be taken in using the best values and best "plotting ages" for the first and last survivorship rates: the best values are those directly observed.

5.4 Stationary Population Estimators

5.4.1 The Linear Method

The "Option 2" method for computing the $\underset{\sim}{p}_x$'s involves the assumption that $\underset{\sim}{L}_x$ is computed linearly from

$$\underset{\sim}{L}_x = \frac{n}{2}[\underset{\sim}{\ell}_x + \underset{\sim}{\ell}_{x+n}] \tag{150}$$

with the last value computed using the (life table) survivorship rate

$$\underset{\sim}{L}_z = \underset{\sim}{S}_{z-n} \underset{\sim}{L}_{z-n} \tag{151}$$

Table 17. Survival probabilities, p_x, for x = 5 to 35; for East Anglia estimated from survivorship rates using the averaging and the cubic spline interpolation methods.

A. Linear interpolation (run 26)

		Transition from East Anglia to		
Age	Death	East Anglia	South East	Rest of Britain
5	.00290	.90582	.04190	.04939
10	.02208	.92246	.03687	.03850
15	.00324	.88879	.05905	.04892
20	.00387	.84225	.08001	.07387
25	.00383	.85369	.06780	.07448
30	.00453	.88824	.04969	.05754
35	.00682	.91321	.03681	.04317

B. Cubic spline interpolation (run 28)

		Transition from East Anglia to		
Age	Death	East Anglia	South East	Rest of Britain
5	.00077	.90709	.04164	.05050
10	.00223	.93080	.03185	.03512
15	.00326	.89189	.05898	.04587
20	.00403	.83112	.08625	.07860
25	.00370	.85159	.06783	.07688
30	.00433	.89044	.04887	.05636
35	.00655	.91370	.03621	.04354

However, because of the methodology inherent to the "Option 2" method (the linear assumption is used twice, once when deriving the survival probabilities from the survivorship rates and again when deriving the L-statistics), the estimates of the L-statistics are such that for all age groups

$$\underset{\sim}{L}_{x+n} = \underset{\sim}{S}_x \underset{\sim}{L}_x \tag{152}$$

This indeed suggests the use of a direct method for calculating $\underset{\sim}{L}_x$, in the case the estimation of the survival probabilities do not have any stringent consequences for the calculation of $\underset{\sim}{L}_x$.

5.4.2 The Direct Method

The method we propose here (originally suggested by Ledent 1978, 1980a) takes the normal life table definitions of survivorship rates

$$\underset{\sim}{s}_{-n} = \frac{1}{n} \underset{\sim}{L}_0 \underset{\sim}{\ell}_0^{-1} \tag{153}$$

$$\underset{\sim}{s}_x = \underset{\sim}{L}_{x+n} \underset{\sim}{L}_x^{-1} \tag{154}$$

and

$$\underset{\sim}{s}_z = \underset{\sim}{L}_z \underset{\sim}{L}_{z-n}^{-1} \tag{155}$$

and reverses them, substituting the observed $\underset{\sim}{S}_x$ for the life table $\underset{\sim}{s}_x$, starting with

$$\underset{\sim}{L}_0 = n \underset{\sim}{S}_{-n} \underset{\sim}{\ell}_0 \tag{156}$$

and continuing with

$$\underset{\sim}{L}_n = \underset{\sim}{S}_0 \underset{\sim}{L}_0 \tag{157}$$

and more generally with

$$\underset{\sim}{L}_{x+n} = \underset{\sim}{S}_x \underset{\sim}{L}_x \tag{158}$$

where $\underset{\sim}{S}_x$ $(n = -n, 0, n, 2n, \ldots, z-n, z)$ are the observed values of the survivorship rates. Note that this series of equations makes possible a solution to the last age group problem so that

$$\underset{\sim}{L}_z = \underset{\sim}{S}_z \underset{\sim}{L}_z + \underset{\sim}{S}_{z-n} \underset{\sim}{L}_{z-n} \tag{159}$$

where $\underset{\sim}{S}_z$ are the observed survivorship rates in the last, open-ended age group. Equation (159) leads to

$$\underset{\sim}{L}_z = (\underset{\sim}{I} - \underset{\sim}{S}_z)^{-1} \underset{\sim}{S}_{z-n} \underset{\sim}{L}_{z-n} \tag{160}$$

This gives us an expression for the life table survivorship rate $\underset{\sim}{s}_{z-n}$

$$\underset{\sim}{s}_{z-n} = (\underset{\sim}{I} - \underset{\sim}{S}_z)^{-1} \underset{\sim}{S}_{z-n} \tag{161}$$

One interesting implication of using this direct method for computing the life years lived variable ($\underset{\sim}{L}_x$'s other interpretation) is that life expectancies at birth ${}_0\underset{\sim}{\ell}_0$ will depend only on the survivorship rates.

Given that

$$\begin{aligned} \underset{\sim}{L}_0 &= \underset{\sim}{s}_{-n}(n\ \underset{\sim}{\ell}_0) \\ \underset{\sim}{L}_n &= \underset{\sim}{s}_0 \underset{\sim}{L}_0 = \underset{\sim}{s}_0 \underset{\sim}{s}_{-n}(n\ \underset{\sim}{\ell}_0) \\ &\vdots \end{aligned} \tag{162}$$

$$\underset{\sim}{L}_x = \underset{\sim}{s}_{x-n}\, \underset{\sim}{L}_{x-n} = \underset{\sim}{s}_{x-n}\, \underset{\sim}{s}_{x-2n} \cdots \underset{\sim}{s}_0\, \underset{\sim}{s}_{-n}(n\, \underset{\sim}{\ell}_0)$$
$$\vdots \tag{162}$$
$$\underset{\sim}{L}_z = \underset{\sim}{s}_{z-n}\, \underset{\sim}{L}_{z-n} = \underset{\sim}{s}_{z-n}\, \underset{\sim}{s}_{z-2n} \cdots \underset{\sim}{s}_0\, \underset{\sim}{s}_{-n}(n\, \underset{\sim}{\ell}_0)$$

Typically

$$\underset{\sim}{L}_x = \left(\prod_{y=-n}^{x-n} \underset{\sim}{s}_y\right)(n\, \underset{\sim}{\ell}_0) \tag{163}$$

so that

$$\sum_{x=0}^{z} \underset{\sim}{L}_x = \sum_{x=0}^{z} \left(\prod_{y=-n}^{x-n} \underset{\sim}{s}_y\right)(n\, \underset{\sim}{\ell}_0) \tag{164}$$

Since life expectancies by region of birth at birth are computed as

$$_0\underset{\sim}{e}_0 = \left[\sum_{x=0}^{z} (\underset{\sim}{L}_x)\right] \underset{\sim}{\ell}_0^{-1} \tag{165}$$

they can be re-expressed as

$$_0\underset{\sim}{e}_0 = \left[\sum_{x=0}^{z} \left(\prod_{y=-n}^{x-n} \underset{\sim}{s}_y\right)\right](n\, \underset{\sim}{\ell}_0)\, \underset{\sim}{\ell}_0^{-1}$$
$$= n \sum_{x=0}^{z} \left(\prod_{y=-n}^{x-n} \underset{\sim}{s}_y\right) \tag{166}$$

Life expectancies beyond age u are

$$_0\underset{\sim}{\ell}_u = \sum_{x=u}^{z} \underset{\sim}{L}_u\, \underset{\sim}{\ell}_u^{-1} \tag{167}$$

so that

$$ {}_0\underset{\sim}{\ell}_u = n \sum_{x=n}^{z} \left(\prod_{y=-n}^{x-n} \underset{\sim}{s}_y \right) \underset{\sim}{\ell}_0 \; \underset{\sim}{\ell}_u^{-1} \tag{168} $$

5.4.3 About the Last Age Group

Before going on to discuss the empirical implementation of the transition approach, it will be useful to clarify how survivorship rates for the last age groups are treated in the life table and accounts based projection models.

In life table based projection models the bottom right-hand corner of the growth operator matrix, $\underset{\sim}{G}$, arranged like a Leslie matrix with each term an n × n interregional submatrix

$$ \begin{bmatrix} \ddots & \vdots & \vdots & \vdots \\ \cdots & \underset{\sim}{s}_{z-2n} & 0 & 0 \\ \cdots & 0 & \underset{\sim}{s}_{z-n} & 0 \end{bmatrix} $$

whereas in accounts based projection models the equivalent corner looks like this

$$ \begin{bmatrix} \ddots & \vdots & \vdots & \vdots \\ \cdots & \underset{\sim}{S}_{z-2n} & 0 & 0 \\ \cdots & 0 & \underset{\sim}{S}_{z-n} & \underset{\sim}{S}_z \end{bmatrix} $$

The lower case s letters refer to life table derived survivorship rates and the upper case S letters to observed (accounts-based) survivorship rates.

The life table survivorship submatrices are defined as follows:

$$\underset{\sim}{s}_{z-2n} = \underset{\sim}{L}_{z-n}\ \underset{\sim}{L}^{-1}_{z-2n} \tag{169}$$

and

$$\underset{\sim}{s}_{z-n} = \underset{\sim}{L}_{z}\ \underset{\sim}{L}^{-1}_{z-n} \tag{170}$$

where $\underset{\sim}{L}_z$ are the life years lived beyond age z or the stationary population in ages z and over. The accounts based definitions are

$$\underset{\sim}{S}_{z-2n} = (\underset{\sim}{K}_{z-2n})'\ [\underset{\sim}{k}_{z-2n}]_d^{-1} \tag{171}$$

$$\underset{\sim}{S}_{z-n} = (\underset{\sim}{K}_{z-n})'\ [\underset{\sim}{k}_{z-n}]_d^{-1} \tag{172}$$

$$\underset{\sim}{S}_{z} = (\underset{\sim}{K}_{z+})'\ [\underset{\sim}{k}_{z+}]_d^{-1} \tag{173}$$

where K refers to an n × n accounts matrix of survivors in the age group starting with age z-2n, z-n or z and extending to z-n, z or all ages beyond z respectively, and $[k]_d$ is a diagonal matrix of initial populations in the age group indicated by the age subscript.

Now $\underset{\sim}{s}_z$ in the life table is set to zero because it has a similar meaning as $\underset{\sim}{p}_z$. The $\underset{\sim}{p}_z$ probability refers to the probability of surviving from exact age z to exact age w by which no one remains alive ($\ell_w = 0$). In effect, we assume $\underset{\sim}{p}_{z\ z+n} \cdot \underset{\sim}{p}_{z+n\ z+2n} \cdots \underset{\sim}{p}_{w-n}\ \underset{\sim}{p}_w = 0$. Similarly, $\underset{\sim}{s}_z \cdot \underset{\sim}{s}_{z+n} \cdot \underset{\sim}{S}_{z+2n} \cdots \underset{\sim}{S}_{w-n} = 0$ since $\underset{\sim}{L}_w = 0$. In the observed projection matrix, however, we only want to carry things forward for T years and therefore clearly the $\underset{\sim}{S}_z$ survivorship rate will be non-zero.

5.5 The Data Sets Used and Experiments Carried Out

For computing multiregional life tables using the transition approach additional data are required. For the transition approach we require data on the stayers and survivors for each of data sets one, two and three (see section 4.5). These are set out in Appendix sections A.2.4 and A.2.5. The five-year figures come from the accounts computed in Rees (1977b) and presented in part in Rees (1980a). The one-year figures were computed from the 1971 census migration and population tables using equation (118) and an infant equivalent.

For the direct survivorship rate method [$S_x^{ij}(4)$'s - see section 5.2.1] information on the initial populations for the 1966-71 and on surviving and non-surviving emigrants are needed. These come from Rees (1977b) via Rees (1980a), and are set out in Appendix section A.2.6.

Table 18 sets out in tabular form the details of the numerical experiments carried out on the Great Britain multiregional population system. The experiments are concentrated on those data sets (two and three) that use the five-year migration data, following the conclusions that we came to about period choice as a result of the hybrid approach experiments. Just a single run (run 41) is carried out with the one-year period migration data.

We consider first the consequences of the various methods used to convert survivorship rates into survival probabilities and subsequent life table statistics, since this is where the crux of previous objections to the transition has resided.

5.6 The Effect of Different Methods of Probability Estimation from Survivorship Rates

5.6.1 "Option 2": Continued Problems

As Rogers (1975a) found, the probability estimates derived using the Option 2 method are badly estimated as Table 19 shows. Despite well-behaved survivorship rates, correctly estimated, the survival probabilities, the $\underset{\sim}{p}_x$'s, are often negative

Table 18. A list of experiments (computer runs) carried out for the transition approach on the Great Britain data sets.

Experiment (Computer run)	Period choice: Deaths	Period choice: Migrants	Survivorship rate choice	Probability choice	Stationary population choice
DATA SET TWO (gb5)					
25	five years	five years	direct	option 2	linear
26	"	"	"	averaging	linear
27	"	"	"	"	direct
28	"	"	"	cubic spline	linear
29	"	"	"	"	direct
30	"	"	conditional	option 2	linear
31	"	"	"	averaging	linear
32	"	"	"	"	direct
33	"	"	"	cubic spline	linear
34	"	"	"	"	direct
DATA SET THREE (gb5 two)					
35	one year	five years	conditional	option 2	linear
36	"	"	"	averaging	linear
37	"	"	"	"	direct
38	"	"	"	cubic spline	linear
39	"	"	"	"	direct
DATA SET ONE (gb1)					
40	one year	one year	conditional	cubic spline	linear

For discussion, see text section specified.

NOTES

Data sets: section 5.5
Survivorship rate choice: section 5.2
Probability choice: section 5.3
Stationary population choice: section 5.4

The computer run numbers follow on from those of Table 4.

Table 19. Survivorship rates and survival probabilities using the transition approach: direct rates, "option 2", linear method (run 25).

Age, x	Death	Retention	Migration		Death	Retention	Migration	
	$s_x^{1\delta}$	s_x^{11}	s_x^{12}	s_x^{13}	$p_x^{1\delta}$	p_x^{11}	p_x^{12}	p_x^{13}
-5	.01612	.91968	.03432	.02988				
0	.00415	.89176	.04734	.05675	.03225	.83936	.06863	.05976
5	.00164	.91988	.03646	.04202	-0.2417	.95462	.01857	.05098
10	.00251	.92504	.03728	.03517	.02650	.88332	.05577	.03441
15	.00397	.85253	.08083	.06268	-.02157	.97251	.01379	.03527
20	.00378	.83197	.07918	.08507	.02868	.72889	.14993	.09250
25	.00388	.87542	.05681	.06389	-.01986	.97557	-.02767	.07196
30	.00518	.90107	.04257	.05118	.02676	.77254	.13993	.06078
35	.00845	.92536	.03105	.03515	-.01472	1.06976	-.09656	.04152
40	.00141	.93449	.02686	.02458	.03181	.79080	.14197	.03543
45	.02320	.93602	.02149	.01929	-.00184	1.11865	-.13438	.01757
50	.03783	.93815	.01327	.01074	.04978	.77384	.14836	.02803
55	.06218	.91733	.01043	.01006	.02828	1.15370	-.18028	-.00170
60	.10234	.87421	.00896	.01449	.09991	.71469	.15566	.02974
65	.15819	.81816	.00878	.01487	.10942	1.10316	-.22014	.00756
70	.24879	.72703	.00869	.01548	.21761	.56381	.18351	.03507
75	.45647	.52738	.00681	.00935	1.00000	.00000	.00000	.00000

NOTES

Regions: 1 = East Anglia 2 = South East 3 = Rest of Britain

($p_x^{1\delta}$'s, p_x^{12}'s) or exceed one (p_x^{11}'s). The principal reason for these results are that the $\underset{\sim}{p}_0$ probabilities are initially poorly estimated by equation (105). The linear assumption is a very poor one given the mortality and migration schedules over years 0 to 5. The initial errors will get passed on to successively higher age intervals. If better values are substituted for $\underset{\sim}{p}_0$ and $\underset{\sim}{p}_5$ then the errors generated by the Option 2 approach are a good deal less but they do not disappear.

Because the $\underset{\sim}{p}_x$'s are ill-estimated, the life expectancy statistics are unacceptable in general, except those of life expectancy at birth. These are acceptable because they rely entirely on the information content of the $\underset{\sim}{S}_x$'s above and are thus identical to those obtained when the direct method of life years estimation is used, although the $\underset{\sim}{p}_x$'s will be very different (see Table 20).

5.6.2 A Quick Comparison of Transition and Hybrid Approach Methods

Since the distribution of life expectancies is relatively similar across all transition approach runs, it is probably be useful at this stage to compare results with those of the hybrid approach. In the fourth column of statistics in Table 20 is reproduced run 15 in which conventional mortality rates, averaged migration rates, the Option 3 probability estimation method and the linear method for $\underset{\sim}{L}_x$ were used. The transition approach runs employing the direct measurement of survivorship rates and their direct use in the life years lived calculation result in life expectancies at birth about 2 months less than those in the hybrid approach. Where conditional survivorship rates are computed, on the basis initially of the same mortality information, the life expectancies are very close, only some 10 days different. Since the survivorship and non-survivorship rates in the case of runs 25 and 27 are directly observed, whereas the overall survivorship rates in the conditional run and in the hybrid approach run are the product of an equation based on the movement concept, our judgement must be in favor of the more directly derived values.

Table 20. Great Britain, transition approach: regional life expectancies at birth (years) and their distribution by region of residence (in percentages) based on five-year migrant data. Option 2, averaging and Option 3 methods compared.

Regions		(1)	(2)	(3)	(4)
	Approach →	Transition Approach			Hybrid Approach
	Rate method →	Direct	Direct	Conditional	Averaged
	Prob. method →	Option 2	Linear Int.	Linear Int.	Option 3
	$\underset{\sim}{L}_x$ method →	Linear	Direct	Direct	Linear
	Experiment →	Run 25	Run 27	Run 32	Run 15
	column →	(1)	(2)	(3)	(4)
LIFE EXPECTANCIES BY REGION AT BIRTH					
East Anglia		72.57	72.57	72.72	72.79
South East		72.29	72.29	72.48	72.51
Rest of Britain		71.57	71.57	71.72	71.70
PERCENT OF LIFE SPENT BY PERSONS BORN IN EAST ANGLIA					
East Anglia		52.19	52.19	52.18	56.19
South East		21.47	21.47	21.47	19.78
Rest of Britain		26.34	26.34	26.35	24.03
PERCENT OF LIFE SPENT BY PERSONS BORN IN SOUTH EAST					
East Anglia		3.75	3.75	3.76	3.51
South East		71.93	71.93	71.91	22.54
Rest of Britain		24.32	24.32	24.33	22.54
PERCENT OF LIFE SPENT BY PERSONS BORN IN REST OF BRITAIN					
East Anglia		1.91	1.91	1.91	1.83
South East		12.10	12.10	12.10	11.59
Rest of Britain		86.00	86.00	85.99	86.58

Note: All runs in this table use the linear method for $\underset{\sim}{L}_x$, and data set two.

When we come to compare the distribution of life expectancies at birth amongst the four runs in Table 20, the grouping of results is different. All the transition approach runs in the table (and the others not reported are close to these results) give retention percentages of circa 52, 72, and 86 percent for East Anglia, the South East, and the Rest of Britain, respectively, whereas the hybrid approach runs yield the rather different figures of circa 56, 74, and 86 percent. The reason for this difference is clear. The transition approach computations are based on closing the population system by distributing the emigration rate over all survivorship rate terms (see sections 5.2.4 and 5.2.5) whereas in the hybrid approach the population system is closed by default and the emigration rate is added only to the diagonal element in the probability matrix. [Compare the $\underset{\sim}{S}_{20}(3)$, $\underset{\sim}{S}_{20}(4)$, and $\underset{\sim}{S}_{20}(1)$ matrices in Figure 9.] The average difference between equivalent period transition and hybrid runs in retention percentages must be circa 2.5 percent, which is an effect much larger than any other within the hybrid approach, for example, except for the period of migration differences.

5.6.3 The Linear Interpolation and Cubic Spline Interpolation Methods for Computing Survival Probabilities

In Table 21 are displayed in columns (2) and (4) the survival probabilities, $\underset{\sim}{p}_x$'s, that result from respectively averaging and fitting a cubic spline smoothing function to the survivorship rates observed in column (1). The estimates differ at the turning points of the migration function as we have already observed (Figure 12). What we can now do, however, is go further and compare the observed survivorship rates with those computed in the life table from the survival probabilities via the linear integration of the $\underset{\sim}{\ell}_x$'s used to generate the $\underset{\sim}{L}_x$'s for input to equation (125). In virtually all age groups the cubic spline derived life table survivorship rates are closer to the observed than the averaging method derived $\underset{\sim}{s}_x$'s, as one might expect. However, both methods produce reasonable results.

Table 21. A comparison of observed and life table survivorship rates for the averaging and cubic spline methods for generating survival probabilities, linear method for stationary population.

Age, x	Runs 26 and 28 Survivorship rates Direct method s_x^{12} Column (1)	Run 26 Survival probabilities Linear interlation method p_x^{12} Column (2)	Run 26 Survivorship rates (life table) Linear method s_x^{12} Column (3)	Run 28 Survival probabilities cubic spline p_x^{12} Column (4)	Run 28 Survivorsip rates (life table) Linear method s_x^{12} Column (5)
-5	.03432		.02796		.02761
0	.04734	.05593	.04976	.04947	.04924
5	.03646	.04190	.03962	.04302	.03708
10	.03728	.03687	.04752	.02985	.04497
15	.08083	.05905	.06862	.05906	.07145
20	.07918	.08001	.07433	.08822	.07766
25	.05681	.06800	.05969	.06778	.05930
30	.04257	.04969	.04371	.04860	.04299
35	.03105	.03681	.03307	.03603	.03249
40	.02686	.02895	.02661	.02819	.02659
45	.02149	.02417	.02081	.02494	.02099
50	.01327	.01738	.01460	.01706	.01422
55	.01043	.01185	.01070	.01113	.01037
60	.08896	.00970	.00918	.00959	.00909
65	.00878	.00887	.00864	.00872	.00859
70	.00869	.00874	.04102	.00876	.04102
75	.00681	.00000		.00000	

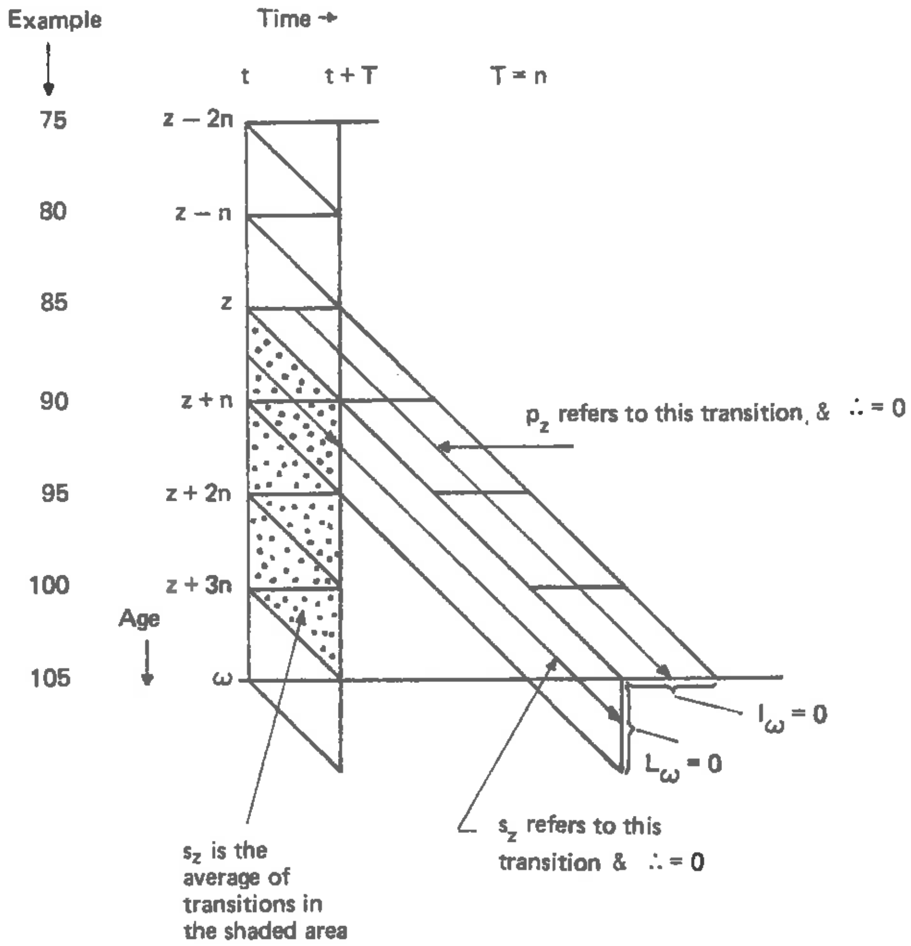


Figure 12. A Lexis diagram illustrating the last age group problem.

Table 22 presents the life expectancies for the runs whose probabilities were presented above together with those for the corresponding direct method for estimating life years lived. In Table 23 are shown the total survivorship rates ($s_x^{i\delta}$) for East Anglia associated with the four runs. The differences between runs in both expectation of life at birth and its distribution among the regions are due very largely to differences in the estimation of the $\underset{\sim}{S}_{-5}$ and $\underset{\sim}{S}_0$ life table survivorship rates. The run 26 values in Table 23 show that the linear method for deriving $\underset{\sim}{L}_x$ and hence, in this case, $\underset{\sim}{S}_x$, gives poor

Table 22. Great Britain, transition approach: regional life expectancies at birth (years) and their distribution by region of residence (in percentages) based on five-year data. Averaging and cubic spline methods compared.

	Approach, rate →	Transition approach - direct rates method			
	Prob. method →	Linear interpolation	Linear interpolation	Cubic spline	Cubic spline
	$\underset{\sim}{L}_x$ method →	Linear	Direct	Linear	Direct
Regions	Experiment →	Run 26	Run 27	Run 28	Run 29
LIFE EXPECTANCIES BY REGION AT BIRTH					
East Anglia		72.30	72.57	72.78	72.57
South East		72.03	72.29	72.55	72.29
Rest of Britain		71.32	71.57	71.84	71.57
PERCENT OF LIFE SPENT BY PERSONS BORN IN EAST ANGLIA					
East Anglia		52.50	52.19	52.50	52.19
South East		21.32	21.47	21.23	21.47
Rest of Britain		26.18	26.34	26.27	26.34
PERCENT OF LIFE SPENT BY PERSONS BORN IN SOUTH EAST					
East Anglia		3.72	3.75	3.73	3.75
South East		72.11	71.93	72.07	71.93
Rest of Britain		24.17	24.32	24.20	24.32
PERCENT OF LIFE SPENT BY PERSONS BORN IN REST OF BRITAIN					
East Anglia		1.89	1.91	1.87	1.91
South East		12.04	12.10	12.04	12.10
Rest of Britain		86.07	86.00	86.09	86.00

results for these first two age groups. If we square the total survivorship S_{-n} rate we obtain rate 0.98194 for a comparable length period that is very little lower than the S_0 value, whereas in reality it should be substantially below it, given the behavior of the mortality function at very young ages. Since the original accounts estimates (Rees 1977) upon which the directly measured survivorship rates (runs 27 and 29 in Table 23) have information about mortality for not only for single years but for short periods within the first year built into them, these are the rates that must be believed. Note that choice of the linear or direct methods for computing $\underset{\sim}{L}_x$ has small effects on the distribution of expected life.

Table 23. Total survivorship rates (life table) for East Anglia associated with selected transition approach runs.

Age, x	Linear interpolation Linear Run 26	Linear interpolation Direct Run 27[1]	Cubic spline Linear Run 28	Cubic spline Direct Run 29
-5	.99093	.98388	.99217	.98388
0	.98948	.99585	.99179	.99585
5	.99752	.99836	.99850	.99836
10	.99735	.99749	.99726	.99749
15	.99644	.99603	.99636	.99603
20	.99613	.99622	.99612	.99622
25	.99581	.99612	.99597	.99612
30	.99432	.99482	.99455	.99482
35	.99095	.99155	.99126	.99155
40	.98505	.98593	.98550	.98593
45	.97545	.97680	.97615	.97680
50	.95987	.96217	.96117	.96217
55	.93426	.93782	.93595	.93782
60	.89475	.89766	.89726	.89766
65	.83567	.84181	.83895	.84181
70	1.63964	1.63964	1.63964	1.63964

Notes

1. These are also the rates which are observed.

2. The runs all use the direct method of survivorship rate measurement.

Our conclusions concerning the choice of method for deriving survival probabilities from survivorship rates are as follows.

The cubic spline interpolation method gives the best results, but those of the linear interpolation method are not significantly worse. Thus, if a cubic spline computing routine is unavailable, the averaging method may be used with a fair degree of confidence. If single year of age data are available, then very good results should be obtained using the linear interpolation method in adapted form

$$\underset{\sim}{p}_x = \frac{1}{2}(\underset{\sim}{S}_{x-1} + \underset{\sim}{S}_x) \tag{174}$$

The Option 2 method gives unsatisfactory results, although these stem in major part from the poor estimation of the initial probabilities by the linear method.

5.7 The Effect of Measuring Survivorship Rates via the Direct or Conditional Methods

The differences between the direct and conditional methods of survivorship rate estimation have two sources: the first is the way in which mortality is handled, either implicitly through observed non-survivorship rates or explicitly through computing mortality rates using linear assumption formulae derived originally from the movement approach. The second is in the way the system is closed as explained in section 5.2.

The first difference has a small effect on the life expectancies: the life expectancies in the conditional method column (run 32) of Table 24 are 0.16 years older on average than those in the direct method column (run 27), as we had earlier observed in comparing runs 32 and 25 in Table 20. The second difference has a very small effect: the distribution percentages differ for retention by an average of only .01 of a percent between the direct rate and conditional rate runs. The third column in Table 24 simply shows the effect of switching to the 1970 deaths data which lowers life expectancy for the East Anglia

Table 24. Great Britain, transition approach: regional life expectancies at birth (years) and their distribution by region of residence (in percentages): a comparison of the direct and conditional rates methods.

Regions		(1)	(2)	(3)	(4)	(5)
	Data set →	Data set two		Data set three	Data set three	Data set one
	Rate method →	Direct	Conditional	Conditional	Conditional	Conditional
	p_x method →	Linear interpolation	Linear interpolation	Linear interpolation	Cubic spline	Cubic spline
	$L_{\sim x}$ method →	Direct	Direct	Direct	Linear	Linear
	Experiment →	Run 27	Run 32	Run 37	Run 38	Run 40
LIFE EXPECTANCIES BY REGION OF BIRTH						
East Anglia		72.57	72.72	73.54	73.53	72.43
South East		72.29	72.48	73.51	73.50	72.44
Rest of Britain		71.57	71.72	72.40	72.41	71.70
PERCENT OF LIFE SPENT BY PERSONS BORN IN EAST ANGLIA						
East Anglia		52.19	52.18	51.99	52.28	40.63
South East		21.47	21.47	21.64	21.41	25.78
Rest of Britain		26.34	26.35	26.37	26.31	33.59
PERCENT OF LIFE SPENT BY PERSONS BORN IN SOUTH EAST						
East Anglia		3.75	3.76	3.79	3.77	4.29
South East		71.93	71.91	71.85	71.98	63.66
Rest of Britain		24.32	24.33	24.35	24.25	32.05
PERCENT OF LIFE SPENT BY PERSONS BORN IN REST OF BRITAIN						
East Anglia		1.91	1.91	1.94	1.90	2.41
South East		12.10	12.10	12.25	12.20	16.76
Rest of Britain		86.00	85.99	85.81	85.90	80.83

Rest of Britain born and raises it for the South East born. The retention percentages for East Anglia, the South East, and the Rest of Britain are all lowered a little as a result.

5.8 The Period of Migration Measurement Effect

The last two columns of Table 24 enable us to confirm the major finding of the hybrid approach section: that of period choice. Runs 38 and 40 differ only in the period to which the migration data refer and in the greater number of age groups in the one year (deaths and migrants) data set (18 compared with 16).

Because the data set one used to perform run 40 refers to unequal time and age intervals, the methodology used was slightly different than in run 38. The life table survivorship rates $\underset{\sim}{s}_x$ are not to be calculated from $\bar{\underset{\sim}{s}}_x$ as suggested in subsection 5.2.5. Instead, we use an equivalent method described in Ledent (1980a, 1980b). A set of $\bar{\underset{\sim}{p}}_x$ is obtained from the set of $\bar{\underset{\sim}{s}}_x$ using the cubic spline interpolation and then $\bar{\underset{\sim}{p}}_x$ is transformed into $\underset{\sim}{p}_x$ using a method similar to the one used to transform $\bar{\underset{\sim}{s}}_x$ into $\underset{\sim}{s}_x$.

This latter difference reduces the life expectancies a little in the one year (deaths and migrants) run compared with the one year deaths and five year migrants run. The major effect, however, is the shift in the distribution of life due to the different measurement period for the migrant data. The retention percentages are 8.6 percent lower on average for the one year period run, and are even lower than in comparable hybrid approach runs because of the different method of system closure used in the transition approach.

5.9 Conclusions about the Transition Approach

Our conclusions for the transition approach choices are fairly straightforward, even if some of the comparisons have been a little complex. We would argue that a method based on the direct measurement of survivorship rates from population accounts, on the averaging of these rates to give probability estimates and on the direct use of survivorship rates for calculating the life years lived/stationary population statistics gives good results in a simple and elegant fashion. Where population accounts (really, good emigration estimates) are not available, a conditional method based just on census migration data and a method for estimating non-survivorship rates from conventional mortality rates may be substitued.

Added precision may be obtained through use of cubic spline interpolation for probability estimation or through more finely disaggregated age data. However, more important than such added sophistication is careful attention to measuring or estimating rates or probabilities for the first age group, where age disaggregation will clearly improve life expectancy estimates considerably. Although less influential is its effect on life expectancies, the proper treatment of the last age group is also important, and we have shown how the last age group can be treated consistently and correctly in both the life table and population projection models.

6. COMPARING THE ALTERNATIVE APPROACHES

6.1 Movement, Hybrid, and Transition Approaches

Ideally, a comparison of these approaches should be based on the construction of multiregional life tables for the same population system using the different techniques. Unfortunately, movement data were not available for the Great Britain regional system since the U.K. does not maintain a population registration system. Such a comprehensive comparison will have to await the analysis of large samples of individual migration histories, available from countries such as Sweden. Here we confine our remarks to the hybrid and transition approaches, noting that our results for the Great Britain regional system for 1970 using the movement-based method of probability estimation ("Option 3") will not be far removed from analyses with movement data, had they been available, since at the interregional scale relatively few "surplus" moves are made within one year.

We have chosen to review the principal results of the hybrid and transition approaches by looking at that most sensitive of summary statistics--${}^{1}\theta^{1}$--the proportion of the life predicted for an East Anglia born infant that he or she will expect to spend in his or her native region. In Figure 13, these statistics are plotted for all 40 "experiments" with the Great Britain three-region population system. The hybrid approach results are plotted on the left-hand scale and transition approach values on the right-hand scale. The alphabetically labelled vertical bars pick out typical differences between particular analyses that differ only in one choice of data or method. Table 25 lists these effects in order of magnitude and shows the numerical shift in the retention percentage that each choice effects.

Dwarfing all other effects are those labelled (A-h) and (A-t) on Figure 13: if we use five-year data for migrants rather than one we increase the retention percentage by 15.5 in the hybrid approach or 11.7 percent in the transition case. Or put another way, if we use one-year data rather than five, we are overestimating the amount of life expected to be spent outside East Anglia by East Anglian natives by 15.5 or 11.7 percent. The second most important effect, labelled (B), less than one the size of the first, results when we use the transition approach with five-year data rather than the hybrid approach. This can most probably be interpreted as a difference due principally to the different way in which the system is closed in each approach. This effect varies with the importance of international outmigration and for many countries will not be serious. The third ranked effect, (C), is the shift retention percentages of circa 1.2 percent that takes place when migration rates are measured (incorrectly) using the equivalent approach rather than the average (correctly).

The other effects contained in Figure 13 and listed in Table 25 are rather small in magnitude, despite the considerable theoretical interest attached to the choices involved. The really important choices are those concerned with the selection and handling of migration data. Our results thus confirm the preliminary experiments and speculations of Ledent (1978:126-137).

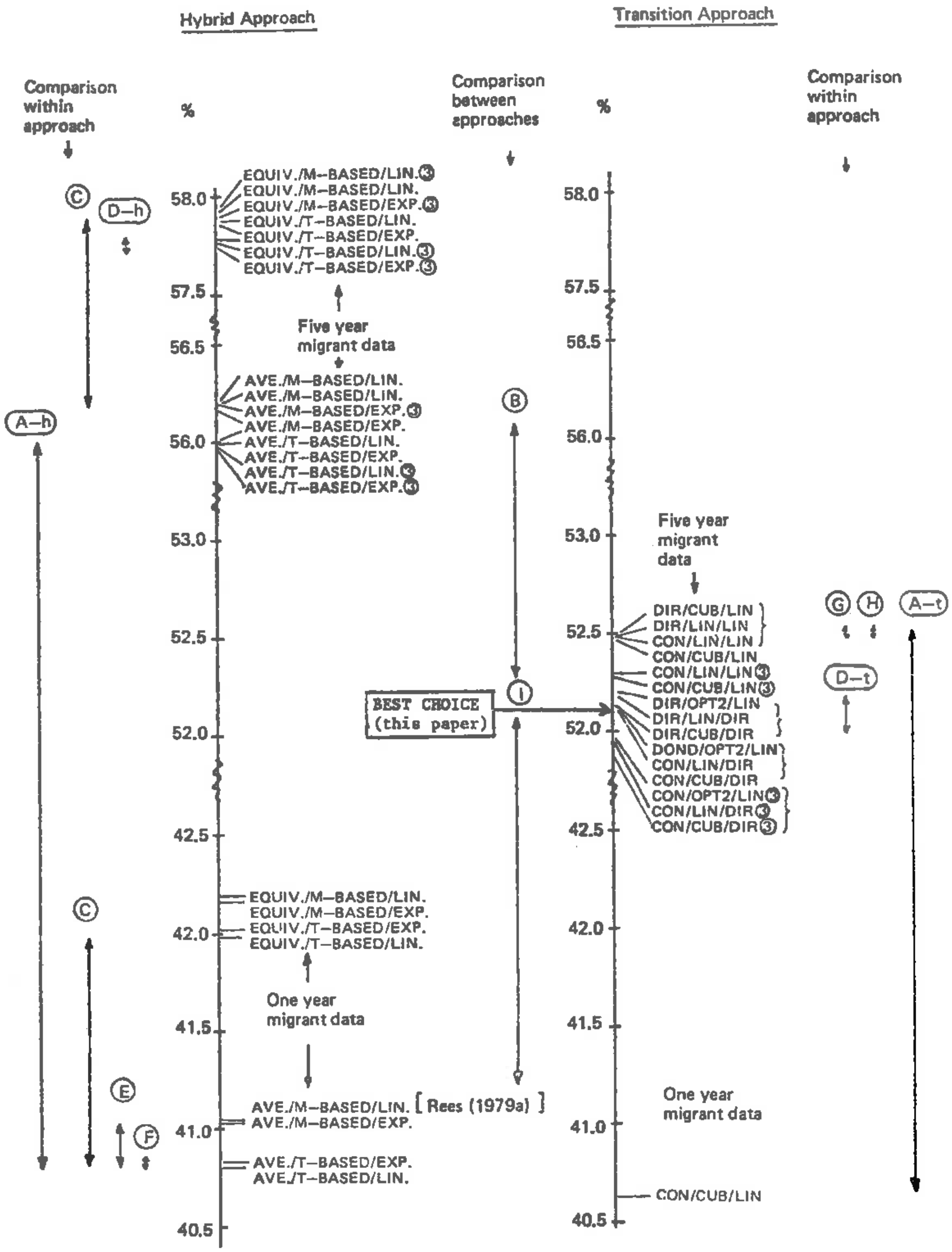


Figure 13. Percentage of expected life of East Anglia born spent in East Anglia: hybrid and transition approaches.

③ - refers to data set three

} brackets same value runs

see tables 15 and 18 for detailed definitions of the run title abbreviations

Table 25. Shifts in retention percentages for East Anglia due to changes in data or method.

Effect	Less preferred alternative		More preferred alternative	Applicable approach
A-h	one-year migrant data -15.46%	vs	five-year migrant data +15.46%	hybrid
A-t	one-year migrant data -11.65%	vs	five-year migrant data +11.65%	transition
B	closure by default +4.50%	vs	explicit closure -4.50%	hybrid vs transition
C	equivalent mobility rates +1.22%	vs	averaged mobility rates -1.22%	hybrid
D-t	one-year deaths data -0.20%	vs	five-year deaths data +0.20%	transition
E	movement-based probabilities +0.16%	vs	transition-based probabilities -0.16%	hybrid
F	linear p_x and L_x formulae +0.02%	vs	exponential p_x and L_x formulae -0.02%	hybrid
D-h	one-year deaths data -0.02%	vs	five-year deaths data +0.02%	hybrid
G	linear interpolation of probabilities -0.00%	vs	cubic spline interpolation of probabilities +0.00%	transition
H	conditional survivorship rates -0.00%	vs	direct survivorship rates transition +0.00%	transition
I	currently published figure (Rees 1979a) [41.02%] -11.17%	vs	best set of choices in this paper (run 29) [52.19%] +11.17%	

Our recommendations to future multistate modelers are contained in Table 25. The rightmost alternative has been labeled "more preferred", and should where possible be chosen. Clearly, the further down the table one goes the more indifferent about the choice one becomes, and from effect (E) downwards adoption of the preferred alternative will depend on whether, in general, the resources for the extra-programming or computational effort are available. (The authors are currently revising their programs to make the preferred choices more accessible.) These recommendations assume that the order of preference for approaches, given a choice: first ranked, transition; second ranked, hybrid; and third ranked, movement. We try in the remainder of the paper to review the arguments behind this ranking.

6.2 The Markovian Assumption of Independence of Previous Region of Residence

The multiregional life table model assumes a Markovian world in which the probability of migration depends only on the state occupied at the start of a time/age interval. The discrepancy between results with one-year and five-year period migration and the very different life tables that result from using place-of-birth migration data (Ledent 1980c) suggest that the real world of people living in and moving amongst regions is not a Markovian one. Return migration (Vanderkamp 1972; Long and Boertlein 1975; Ledent 1980c) is a real world process not captured in the multiregional population model.

This Markovian assumption (there are others) is used differently in the movement and transition approaches. It is used in the movement approach to obtain the transition probabilities ($\underset{\sim}{p}_x$'s) from the movement data (the $\underset{\sim}{M}_x$'s) within an age/time interval. In the transition approach the transition probabilities are measured more directly, and within an age/time interval the Markov assumption is not used.

In both approaches this Markov assumption is used to proceed from one age to the next. The degree of error so introduced can be gauged by comparing the one-year and five-year results. In the one-year runs this Markov assumption has been used five times as frequently as in the five-year runs. Over all the age intervals, the one-year model (irrespective of whether the model has $T = 1$, $n = 5$ or $T = 1$, $n = 1$) will use this Markov assumption 86 times whereas the five-year model will use it 18 times. Thus, the larger the time/age interval, the smaller the number of intervals and the lesser the impact of this Markovian assumption.

Of course, if we take this argument to its logical conclusion, the best measurement of people's expected life histories is of course to average their actual migration histories over their lifetime. Such an exercise would involve observation over periods of a lifetime and so would be impractible. Note that if we simply reduce the period of observation to five or one years while maintaining long age intervals all sorts of other aggregation errors are liable to be introduced.

Our conclusion is that a five-year age interval, five-year time period based model is a reasonable compromise between the desire to decrease the reliance of the life table model on the Markovian assumption and the desire to reduce the error introduced by age aggregation (particularly important for total life expectancy computations). The main competitor to the five-year age interval, five-year time interval model is the single year of age, annual period model. This has undoubted advantages in planning contexts since, for educational and social service planning, the client groups do not fall into neat quinquennial age groups, and in planning the time horizon is short and the model outputs must be revised earlier. Fortunately, the effect of non-applicability of the Markovian assumption we have discussed is less serious for the population projection products of multiregional population models as Philipov and Rogers (1980) have shown.

6.3 Solving the "One Year-Five Year Problem"

From what we have said previously, it is clear that the calculation of a multiregional life table from movement data or transition data is not adequate and that some procedure is necessary to correct the transition probabilities obtained by a strict application of the available methods to such data.

We note first that the "one year/five year problem" has been recognized in more general form in mobility studies for some time (Blumen, Kogan and McCarthy 1955; Bartholomew 1973), particularly in manpower studies. The probability of staying in a job n years is observed to be much higher than a one-year probability matrix applied n times would suggest.

In manpower studies models which partition the population into movers and stayers have proved useful. Kitsul and Philipov (1980) have constructed a model that partitions the population into a low mobility group and a high mobility group. Although the mobility rates computed by the model for the two groups cannot be given a behavioral interpretation, the model does reproduce fairly closely the observed five-year transition probabilities matrix from the one-year matrix. Determination of model parameters depends on the availability of five-year and one-year migration matrices. Kitsul and Philipov use the same Great Britain system and data as we do in this paper (originating in Rees 1977b), so that, in principle, the equations which they determined, could be applied to the Great Britain migration data to be produced from the 1981 Census. The Office of Population Censuses and Surveys have, in their wisdom, decided to drop the five-year migration question asked in 1971 and 1966 of a sample of households, and concentrate their efforts on asking all households a one-year question (a much more costly exercise than asking both one- and five-year questions of 10% of households, as was the case in 1971 and 1966). Alternatively, the actual empirical ratios of one-year to five-year rates could be used to inflate the 1981 one-year observations.

An alternative approach to the problem has been taken by Castro (1980), who generalized the simple model suggested by Rees (1977a) to a full binomial probability model. Although this model works well in predicting five-year migration rates from one-year rates for all migrations, it does not work well for migration between aggregated spatial units since it uses the Markovian assumption of independence of prior state.

What the previous attempts at the problem fail to do, in our opinion, is to build in to the one-year to five-year connecting model, the key behavior that produces the "problem", namely return migration. The corollary is that progress with the one-year/five-year problem can be made only given good individual migration history data with which the true causes of the one year-five year problem can be especially identified and then modeled.

If return migration (rather than just multiple migration *per se*) is the explanation for the one year-five year problem, then this suggests a reason why a Markovian model appeared to work well for all migrations between residences (Rees 1977a; Castro 1980) but not for interregional migration. Clearly, return migration is a fairly diffuse process. A migrant may move to another area, fail to make a success of it, and return to his previous district. It is unlikely that he will be able to return to the very dwelling he left since someone else will now be living there! The desire to return will be one of wanting to live in the same milieu, the same environment as before, perhaps within reach of former friends and family. But such desires can be satisfied in many spatial locations around the previous residence, some quite far away. For an international migrant the ambience which he wishes to capture may include the whole of his native country.

The implications of these arguments is that, as we move from a multistate population model made operational at the regional scale (say 10 regions of 5.5 million people each on average) where the Markovian assumption of independence is violated to one operational at the household scale (say 17 million "regions" of 3.2 persons each on average) where the assumption holds, we may find a happy medium at a fine spatial scale (say 130 countries and districts of 0.42 million people each) where we can "live with" the Markovian assumption. Gibberd (1980) has shown that, although a Markovian process may hold for a disaggregated spatial system, an aggregated system may exhibit non-Markovian behavior, just as a result of aggregation let alone a change in behavior which we are suggesting here. It should be possible to examine migration matrices at various spatial scales in a number of countries and determine whether these speculations have any validity.

6.4 Concluding Remarks

This paper has dealt with the construction of multiregional life tables from an applied point of view. New methods were developed for a variety of life table calculations, building on previous work, and old and new methods were applied in as comparable a fashion as possible to a common multiregional population system. Our principal conclusions were that transition data should be used rather than movement data, where available, and that transition methods rather than hybrid methods should be applied to the transition data.

These proposals are based, almost exclusively, on examination of the life tables generated by multiregional population models in just one country. We believe our results are applicable in other developed countries. Whether our conclusions would be the same had we examined other, rather different, outputs of multiregional population models, namely, migraproduction rates and projected populations is another question that deserves to be considered in further work.

REFERENCES

Bartholomew, D. (1973) *Stochastic Models for Social Processes*. Chichester: Wiley.

Blumen, I., F. Kogan, and P. McMarthy (1955) *The Industrial Mobility of Labor as a Probability Process*. Cornell Studies of Industry and Labor Relations, 4. Ithaca: New York.

Castro, L. (1980) Aggregated data transformation: transitions and moves. Draft of unpublished paper (personal communication).

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

Gibberd, R. (1980) Aggregation of population projection models. Draft of unpublished paper (personal communication).

Keyfitz, N. (1966) A life table that agrees with the data. *Journal of the American Statistical Association* 61(1): 305-312.

Keyfitz, N. (1968) *An Introduction to the Mathematics of Population*. Reading, Mass.: Addison-Wesley.

Kitsul, P., and D. Philipov (1980) *The One-Year - Five-Year Migration Problem*. WP-80-81. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Krishnamoorthy, S. (1979) Classical approach to increment-decrement life tables: an application to the study of marital status of United States females, 1970. *Mathematical Biosciences* 44(1-2): 139-154.

Ledent, J. (1978) *Some Methodological and Empirical Considerations in the Construction of Increment-Decrement Life Tables*. RM-78-25. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Ledent, J. (1980a) Multistate life tables: movement versus transition perspectives. *Environment and Planning A* 12: 533-562.

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

Ledent, J. (1980c) *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 A. Rogers (1972) *An Interpolative-Iterative Procedure for Constructing a Multiregional Life Table*. Multiregional Growth Models Project, Working Paper WP 9/72, The Technological Institute, Northwestern University, Evanston, Illinois.

Long, L.H., and C.G. Boertlein (1975) *The Geographical Mobility of Americans: an International Comparison*. Current Population Reports, Special Studies, Series P-23, No. 64. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census.

McNeil, D.R., T.J. Trussell, and J.C. Turner (1977) Spline interpolation of demographic data. *Demography* 14:245-252.

Oechsli, F.W. (1972) The parity and nuptiality problem in demography. Unpublished Ph.D. dissertation, Graduate Division, University of California, Berkeley.

Oechsli, F.W. (1975) A population model based on a life table that includes marriage and parity. *Theoretical Population Biology* 7:229-245.

Office of Population Censuses and Surveys (1974a) *Census 1971, Great Britain, Migration Tables, Part I: 10% Sample*. London: H.M.S.O.

Office of Population Censuses and Surveys (1974b) *Census 1971, Great Britain, Age, Marital Condition and General Tables*. London: H.M.S.O.

Philipov, D., and A. Rogers (1980) *Multistate Population Projections*. WP-80-57. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Plessis-Fraissard, M., and P.H. Rees (1976) *Computer Program for Constructing Age-Sex Disaggregated Multiregional Population Accounts (DAME): Users' Notes*. Working Paper 170. School of Geography, University of Leeds.

Plessis-Fraissard, M., and P.H. Rees (1978) Computer program 1: a program for constructing age-sex disaggregated multiregional accounts (DAME). *Environment and Planning A* 10 (2):225-226.

Plessis-Fraissard, M. (1980) Migrants, spouses and households: an investigation. Unpublished Ph.D. thesis in preparation, School of Geography, University of Leeds.

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

Rees, P.H. (1977b) The Future Population of East Anglia and Its Constituent Counties (Cambridge, Norfolk and Suffolk). A report prepared for the East Anglia Economic Planning Council by the University of Leeds Industrial Services Ltd., under Department of the Environment Contract Number DGR/461/23.

Rees, P.H. (1978) Increment-decrement life tables: some further comments for an accounting point of view. *Environment and Planning A* 10(705-726).

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

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

Rees, P.H. (1979c) Regional population projection models and accounting methods. *Journal of the Royal Statistical Society*, Series A, 142(2):223-255.

Rees, P.H. (1980a) Problems of multiregional population analysis: data collection and demographic accounting. *Sistemi Urbani*, forthcoming. (Originally produced as Working Paper 229, School of Geography, University of Leeds.)

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

Rees, P.H., and A.G. Wilson (1975) Accounts and models for spatial demographic analysis 3: rates and life tables. *Environment and Planning A* 7:199-231.

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

Rogers, A. (1973a) The multiregional life table. *Journal of Mathematical Sociology* 3:127-137.

Rogers, A. (1973b) The mathematics of multiregional demographic growth. *Environment and Planning* 5:3-29.

Rogers, A. (1975a) *Introduction to Multiregional Mathematical Demography*. New York: Wiley.

Rogers, A. (1975b) *Spatial Migration Expectancies*. RM-75-77. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A. (1976) *The Comparative Migration and Settlement Study: A Summary of Workshop Proceedings and Conclusions*. RM-76-01. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A., and J. Ledent (1975) Multiregional Population Projections, DP-128, Center for Mathematical Studies in Economics and Management Science, Northwestern University, Evanston, Illinois; published in revised form in pages 31-58, *Optimization Techniques: Modeling and Optimization in the Service of Man, Part I*, edited by J. Cea, 1976. Springer, Berlin.

Rogers, A., and J. Ledent (1976) Increment-decrement life tables: a comment. *Demography* 13(2):287-290.

Schoen, R. (1975) Constructing increment-decrement life tables. *Demography* 12(2):313-324.

Schoen, R. (1979) Calculating increment-decrement life tables by estimating mean duration of transfer from observed rates. *Mathematical Biosciences* 47:255-259.

Schoen, R., and K.C. Land (1979) A general algorithm for estimating a Markov-generated increment-decrement life table with applications to marital status patterns. *Journal of the American Statistical Association* 74:761-776.

Schoen, R., and V.E. Nelson (1974) Marriage, divorce and mortality: a life table analysis. *Demography* 11:267-290.

Vanderkamp, J. (1972) Return migration: its significance and behavior. *Western Economic Journal* 10:460-465.

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

Willekens, F., A. Por, 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 and Computer Programs.* RR-78-18. Laxenburg, Austria: International Institute for Applied Systems Analysis.

APPENDIX 1: THE MEANINGS OF X

A.1.1 The Three Meanings

For convenience and following conventional practice, we use one subscript, x, to denote age in all discrete data and life table variables. However, there are three pairs of distinct and different meanings that can be attached to x, and only one meaning applies to each variable in general. These are set out below and in Figure A.1.

	meaning	stock variables	flow variables
(i)	exact age x	all lifelines crossing the (x, t) to (x, t+T) line	all flows in the (x, t), (x, t+T), (x+n, t+2T) and (x+n, t+T) age-time space
(ii)	age cohort x	all lifelines crossing the (x, t) to (x+n, t) line	all flows in the (x, t), (x+n, t+T), (x+2n, t+T) and (x+n, t) age-time space
(iii)	age group x	all lifelines crossing the $(x, t+\frac{T}{2})$ to $(x+n, t+\frac{T}{2})$ line	all flows in the (x, t), (x, t+T), (x+n, t+n) and (x+n, t) age-time space

Figure A.2 illustrates these meanings with respect to data variables used in the analysis.

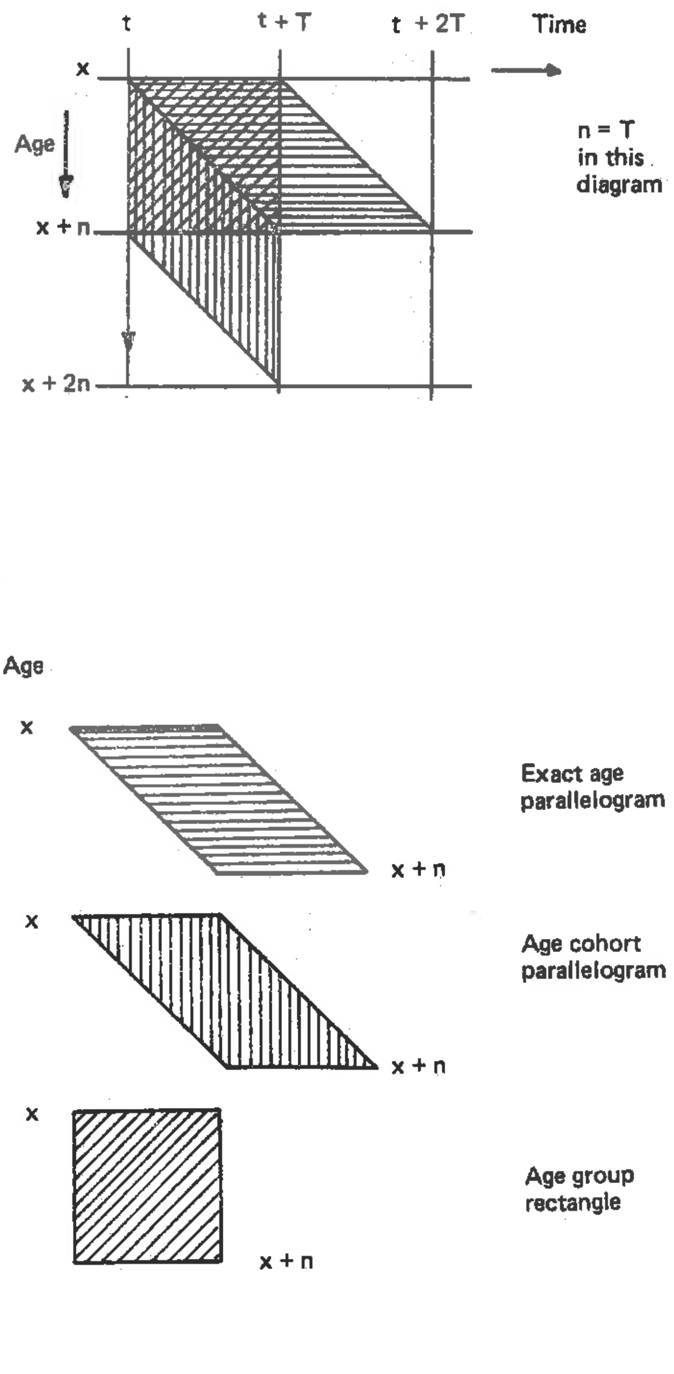


Figure A.1 Alternative meanings of x.

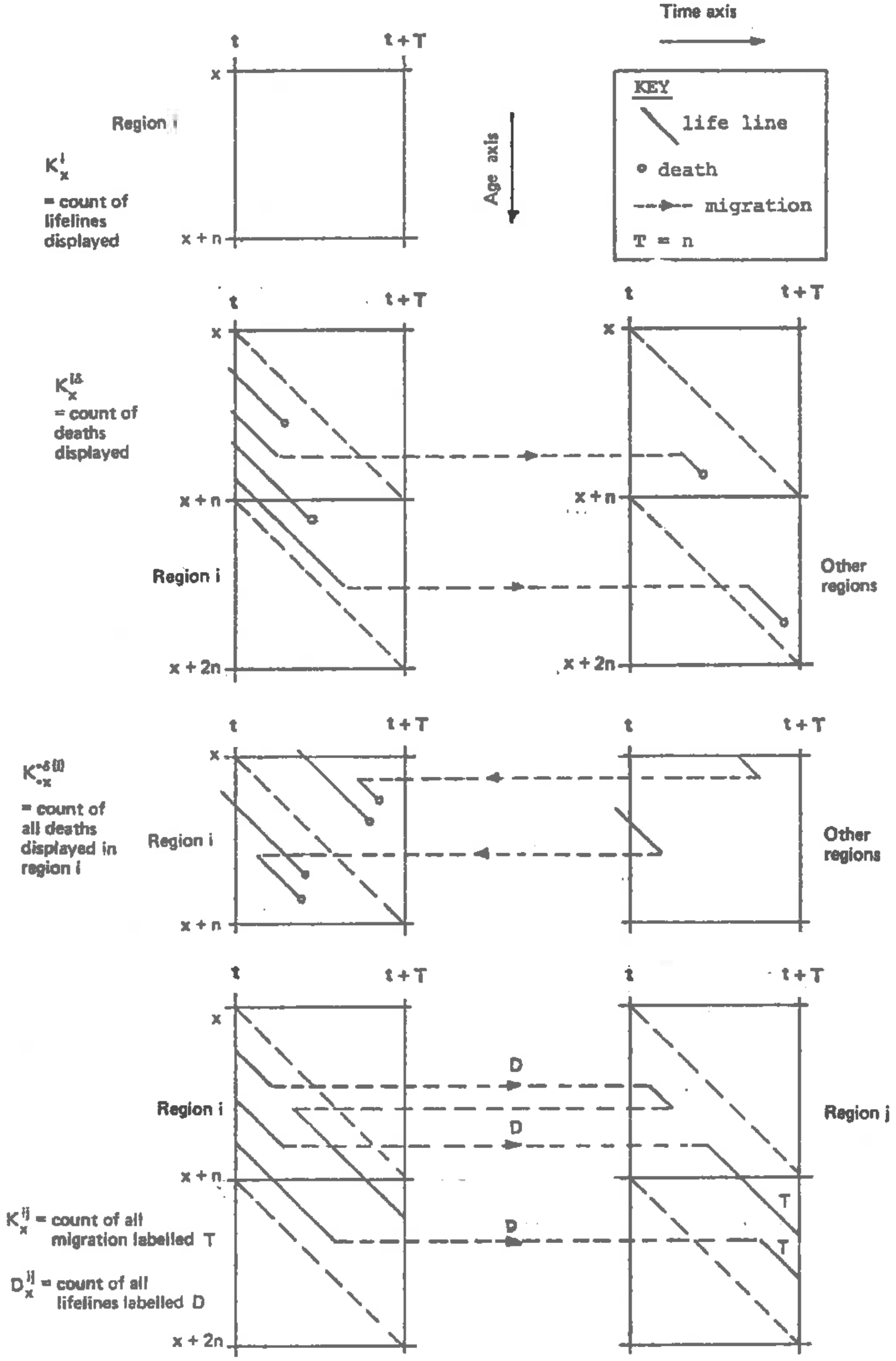


Figure A.2 The meaning of data items clarified by means of a Lexis diagram.

A.1.2 Examples of Variables Associated with Three Meanings

	meaning	stock variables	flow variables
(i)	exact age x	ℓ_x^i, ${}_0^j\ell_x^i$	p_x^{ij}, p_x^i, $\underset{\sim}{p}_x$
		$\underset{\sim}{\ell}_x$	${}_0^j\ell_x^{ik}$, ${}_0^j\ell_x^{i\delta}$
(ii)	age cohort x	$K_x^i(t)$, $\underset{\sim}{k}_x$	S_x^{ij}, $S_x^{i\delta}$, $\bar{S}_x^{ij}$, $S_x^{i\sigma}$
			s_x^{ij}, $s_x^{i\delta}$
(iii)	age group x	$K_x^i(t+\frac{T}{2})$	$K_x^{\cdot(\cdot)\delta(i)}$, D_x^{ij}
			$M_x^{i\delta}$, $M_x^{\delta(i)}$, M_x^{ij}

APPENDIX 2. THE DATA SETS USED

A.2.1 Data Set One (gb1)

This data set refers to the year 1970 and is (together with additional data--see A.2.4) used in runs 1 to 8 of the hybrid approach and run 41 of the transition approach. The data are aggregated in part from those presented in Rees (1979a, Appendix C) and in part are the original data gathered but before manipulation (Rees 1979a:74-79).

EAST ANGLIA

Age group		Population (m.y., 1970)	Deaths (1970)	Age group (initial)	Migrants (1970-71) to East Anglia	South East	Rest of Britain	Age group (final)
x	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Births		(26,018)*	--	Birth	0	243	263	0
0	0-4	128,700	489	0-3	0	1,940	2,103	1-4
5	5-9	134,700	50	4-8	0	1,785	1,880	5-9
10	10-14	113,900	38	9-13	0	1,036	1,099	10-14
15	15-19	114,200	83	14-18	0	1,890	1,590	15-19
20	20-24	137,700	113	19-23	0	4,330	3,710	20-24
25	25-29	108,600	57	24-28	0	2,880	2,431	25-29
30	30-34	103,100	69	29-33	0	1,560	1,720	30-34
35	35-39	101,300	123	34-38	0	1,201	1,269	35-39
40	40-44	97,800	221	39-43	0	749	790	40-44
45	45-49	109,300	397	44-48	0	700	600	45-49
50	50-54	92,400	545	49-53	0	718	613	50-54
55	55-59	100,800	985	54-58	0	372	319	55-59
60	60-64	95,800	1,539	59-63	0	389	302	60-64
65	65-69	84,700	2,393	64-68	0	307	286	65-69
70	70-74	61,400	2,684	69-73	0	319	289	70-74
75	75-79	43,795	2,940	74-78	0	160	148	75-79
80	80-84	27,672	2,899	79-83	0	104	92	80-84
85	85+	17,633	3,396	84+	0	71	66	85+
Totals		1,673,500	19,021		0	20,753	19,570	

* The population column total in this and subsequent tables does not include the element in the births row. Brackets around the birth row element are used in the tables that follow to indicate this.

SOUTH EAST

Age group x	(1)	Population (m.y., 1970) (2)	Deaths (1970) (3)	Age group (initial) (4)	Migrants (1970-71) to East Anglia (5)	South East (6)	Rest of Britain (7)	Age group (final) (8)
	Births	(267,522)	--	Birth	480	0	2,087	0
0	0-4	1,392,600	5,192	0-3	3,840	0	16,696	1-4
5	5-9	1,366,500	417	4-8	2,829	0	14,953	5-9
10	10-14	1,197,800	330	9-13	1,831	0	9,683	10-14
15	15-19	1,139,500	725	14-18	2,200	0	14,300	15-19
20	20-24	1,382,200	972	19-23	5,841	0	39,810	20-24
25	25-29	1,152,200	805	24-28	4,349	0	26,280	25-29
30	30-34	1,059,800	949	29-33	2,869	0	16,189	30-34
35	35-39	1,027,800	1,261	34-38	2,006	0	10,676	35-39
40	40-44	1,031,000	2,363	39-43	1,434	0	7,603	40-44
45	45-49	1,151,100	4,463	44-48	1,490	0	7,712	45-49
50	50-54	1,019,400	6,455	49-53	1,241	0	6,362	50-54
55	55-59	1,097,800	11,738	54-58	1,250	0	6,156	55-59
60	60-64	998,300	17,427	59-63	1,669	0	6,174	60-64
65	65-69	824,600	23,619	64-68	1,259	0	6,307	65-69
70	70-74	614,500	27,208	69-73	579	0	2,945	70-74
75	75-79	410,500	27,512	74-78	349	0	1,828	75-79
80	80-84	268,275	28,533	79-83	221	0	1,157	80-84
85	85+	181,872	36,202	84+	142	0	761	85+
Totals		17,315,502	196,171	Totals	35,879	0	197,679	

REST OF BRITAIN

Age group		Population	Deaths	Age group	Migrants (1970-71) to			Age group
		(m.y., 1970)	(1970)	(initial)	East Anglia	South East	Rest of Britain	(final)
x	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Births	(578,281)	--	Birth	303	1,933	0	0
0	0-4	2,936,400	12,941	0-3	2,422	15,465	0	1-4
5	5-9	3,040,300	1,096	4-8	2,576	15,636	0	5-9
10	10-14	2,630,100	820	9-13	1,553	9,064	0	10-14
15	15-19	2,441,200	1,549	14-18	2,630	21,980	0	15-19
20	20-24	2,678,000	1,926	19-23	4,510	48,860	0	20-24
25	25-29	2,242,700	1,642	24-28	3,250	25,010	0	25-29
30	30-34	2,091,000	2,016	29-33	1,956	15,141	0	30-34
35	35-39	2,059,700	3,106	34-38	1,515	10,841	0	35-39
40	40-44	2,143,600	5,880	39-43	1,118	8,167	0	40-44
45	45-49	2,329,300	11,258	44-48	799	6,103	0	45-49
50	50-54	2,000,900	15,744	49-53	636	4,823	0	50-54
55	55-59	2,150,500	27,445	54-58	507	3,784	0	55-59
60	60-64	2,007,300	41,573	59-63	559	3,281	0	60-64
65	65-69	1,692,500	56,217	64-68	496	3,015	0	65-69
70	70-74	1,225,600	62,936	69-73	293	1,852	0	70-74
75	75-79	805,980	62,784	74-78	197	1,322	0	75-79
80	80-84	462,000	56,097	79-83	116	790	0	80-84
85	85+	260,720	58,612	84+	70	461	0	85+
Totals		35,776,081	423,642	Totals	25,506	198,068	0	

A.2.2 Data Set Two (gb5)

This data set refers to the five-year intercensal period 1966-71 and is used in runs 9 to 16 of the hybrid approach and in runs 25 to 34 of the transition approach (with additional data--see A.2.5). The data are derived from the accounts computed in Rees (1977b) and presented in part in Rees (1980a).

EAST ANGLIA

x	Age group (1)	Population (ave. for 1966-71) (2)	Deaths (1966-71) (3)	Age group (initial) (4)	Migrants (1966-71) to: East Anglia (5)	South East (6)	Rest of Britain (7)	Age group (final) (8)
	Births	(124,770)	--	Birth	0	4,201	3,658	0-4
0	0-4	128,516	2,458	0-4	0	5,739	6,880	5-9
5	5-9	124,005	223	5-9	0	3,972	4,578	10-14
10	10-14	113,736	177	10-14	0	3,893	3,673	15-19
15	15-19	116,506	422	15-19	0	9,396	7,286	20-24
20	20-24	116,128	471	20-24	0	7,578	8,141	25-29
25	25-29	100,491	336	25-29	0	4,895	5,505	30-34
30	30-34	94,820	580	35-39	0	3,697	4,445	40-44
35	35-39	94,820	580	35-39	0	2,859	3,237	45-49
40	40-44	99,107	1,050	40-44	0	2,640	2,416	45-49
45	45-49	98,376	1,835	45-49	0	1,964	1,763	50-54
50	50-54	96,507	2,746	50-54	0	1,274	1,031	55-59
55	55-59	98,317	4,880	55-59	0	1,003	967	60-64
60	60-64	92,849	7,627	60-64	0	794	1,284	65-59
65	65-69	79,798	10,996	65-69	0	652	1,105	70-74
70	70-74	61,738	13,058	70-74	0	492	876	75-79
75	75+	86,104	45,240	75+	0	549	754	80+
Totals		1,725,814	92,504		0	55,598	57,599	

SOUTH EAST

Age group x	(1)	Population (ave. for 1966-71) (2)	Deaths (1966-71) (3)	Age group (initial) (4)	Migrants (1966-71) to: East Anglia (5)	South East (6)	Rest of Britain (7)	Age group (final) (8)
	Births	(1,352,048)	--	Birth	6,519	0	33,213	0-4
0	0-4	1,355,979	28,558	0-4	11,171	0	58,934	5-9
5	5-9	1,297,371	2,182	5-9	7,513	0	39,793	10-14
10	10-14	1,159,903	1,689	10-14	5,710	0	31,756	15-19
15	15-19	1,212,193	3,842	15-19	12,590	0	71,702	20-24
20	20-24	1,273,144	4,124	20-24	14,408	0	79,740	25-29
25	25-29	1,115,041	3,863	25-29	9,799	0	57,623	30-34
30	30-34	1,019,232	4,683	30-34	7,366	0	41,016	35-39
35	35-39	1,017,379	6,882	40-44	6,039	0	32,917	40-44
40	40-44	1,066,841	12,569	40-44	5,512	0	27,662	45-49
45	45-49	1,064,446	21,435	45-49	4,737	0	23,445	50-54
50	50-54	1,060,723	34,380	50-54	4,612	0	23,250	55-59
55	55-59	1,069,806	58,179	55-59	6,428	0	25,097	60-64
60	60-64	970,123	86,011	60-64	4,638	0	27,368	65-69
65	65-69	779,252	111,490	65-69	3,379	0	13,168	70-74
70	70-74	584,392	129,949	70-74	2,008	0	6,800	75-79
75	75+	823,520	443,177	75+	2,123	0	6,384	80+
Totals		18,221,393	953,013		114,552		599,868	

REST OF BRITAIN

Age group		Population (ave. for 1966-71)	Deaths (1966-71)	Age group (initial)	Migrants (1966-71) to:			Age group (final)
					East Anglia	South East	Rest of Britain	
x	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Births	(2,975,636)	--	Birth	8,889	30,178	0	0-4
0	0-4	2,980,333	66,769	0-4	8,869	52,833	0	5-9
5	5-9	2,899,912	5,435	5-9	6,129	37,861	0	10-14
10	10-14	2,634,849	4,309	10-14	5,814	46,077	0	15-19
15	15-19	2,598,980	8,464	15-19	10,725	119,097	0	20-24
20	20-24	2,456,332	9,379	20-24	9,440	84,510	0	25-29
25	25-29	2,159,700	8,185	25-29	7,129	51,816	0	30-34
30	30-34	2,044,905	10,060	30-34	4,864	37,201	0	35-39
35	35-39	2,085,387	16,192	35-39	4,422	31,037	0	40-44
40	40-44	2,228,875	30,183	40-44	5,550	26,522	0	45-49
45	45-49	2,219,553	52,850	45-49	5,017	19,553	0	50-54
50	50-54	2,167,280	80,139	50-54	4,800	14,992	0	55-59
55	55-59	2,168,242	137,235	55-59	1,819	12,220	0	60-64
60	60-64	2,003,826	206,321	60-64	2,824	11,886	0	65-69
65	65-69	1,655,347	272,473	65-69	2,480	7,995	0	70-74
70	70-74	1,218,643	305,976	70-74	1,710	5,950	0	75-79
75	75+	1,500,375	874,751	75+	1,729	6,083	0	80+
Totals		37,998,175	2,088,721		92,210	595,811	0	

A.2.3 Data Set Three (gb5 two)

These are the data additional to or which replace those in data set two. They are used in hybrid approach runs 17 to 24.

Age group		Population (average for 1966-71)		
		East Anglia	South East	Rest of Britain
x	(1)	(2)	(3)	(4)
75	75-79	42,322	392,668	791,046
80	80-84	26,742	256,776	453,440
85	85+	17,040	174,076	255,889
Totals		86,104	823,520	1,500,375

Age group		Deaths (1970 deaths × 5)		
		East Anglia	South East	Rest of Britain
x	(1)	(2)	(3)	(4)
0	0-4	2,442	25,277	65,673
5	5-9	230	1,980	5,227
10	10-14	190	1,598	2,549
15	15-19	423	3,856	8,246
20	20-24	476	4,477	8,833
25	25-29	264	3,895	7,906
30	30-34	315	4,563	9,858
35	35-39	576	6,241	15,724
40	40-44	1,120	12,226	30,570
45	45-49	1,787	20,635	53,638
50	50-54	2,846	33,583	85,266
55	55-59	4,804	57,193	138,357
60	60-64	7,458	84,676	207,505
65	65-69	11,273	111,600	274,915
70	70-74	13,494	129,375	312,894
75	75-79	14,206	131,664	308,103
80	80-84	14,008	136,550	275,288
85	85+	16,409	173,251	287,630
Totals		92,321	942,640	2,098,182

Note:

The additional migrant numbers are generated for age groups 75-79, 80-84, and 85+, by assuming $\underset{\sim}{S}_{75} = \underset{\sim}{S}_{75+}$, $\underset{\sim}{S}_{80} = \underset{\sim}{S}_{75+}$, and $\underset{\sim}{S}_{80+} = \underset{\sim}{S}_{75+}$.

A.2.4 Data Set One (gb1): Additional Data for the Transition Approach

Age group (initial)	Surviving stayers (1970-71)			Age group (final)
	East Anglia	South East	Rest of Britain	
(1)	(2)	(3)	(4)	(5)
Birth	25,041	259,556	565,329	0
0-3	97,137	1,007,688	2,297,635	1-4
4-8	127,238	1,323,340	2,981,784	5-9
9-13	114,819	1,176,186	2,725,149	10-14
14-18	105,670	1,074,007	2,450,063	15-19
19-23	116,418	1,263,286	2,554,826	20-24
24-28	100,678	1,104,032	2,177,691	25-29
29-33	91,693	988,667	2,015,049	30-34
34-38	90,928	959,182	1,984,661	35-39
39-43	94,928	1,005,003	2,018,277	40-44
44-48	101,930	1,068,182	2,262,936	45-49
49-53	94,075	998,956	2,081,521	50-54
54-58	98,074	1,044,470	2,125,216	55-59
59-63	94,370	988,735	2,032,129	60-64
64-68	83,092	818,596	1,729,241	65-69
69-73	65,902	603,175	1,282,008	70-74
74-78	44,449	414,329	837,948	75-79
79-83	28,083	254,296	459,861	80-84
84+	17,898	172,500	259,363	85+
Totals	1,591,942	16,524,186	34,930,687	

SOURCE: These surviving stayer numbers are computed by subtracting total internal in-migrants and total external in-migrants (immigrants) to each region (as given in Table 3A of O.P.C.S., 1974a:61,78 and 79) from the adjusted census data, 1971 populations of the regions as given in Rees, 1980a (derived originally from O.P.C.S., 1974b).

A.2.5 Data Set Two (gb5) and Data Set Three (gb5 two): Additional Data for the Transition Approach

EAST ANGLIA

Age group (initial)	Population (at c.d. 1966)	Surviving emigrants (1966-71)	Non-surviving emigrants (1966-71)	Reduced population at risk	Surviving stayers (1966-71)	Age group (final)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Birth	124,770	2,330	18	122,422	112,592	0-4
0-4	125,605	4,366	9	121,230	108,108	5-9
5-9	112,961	4,011	4	108,946	100,217	10-14
10-14	107,773	3,334	5	110,434	96,606	15-19
15-19	121,081	4,827	8	116,246	99,103	20-24
20-24	101,007	5,296	10	95,701	79,620	25-29
25-29	89,723	3,557	7	86,159	75,425	30-34
30-34	89,399	2,544	6	86,849	78,257	35-39
35-39	94,233	2,133	9	92,091	85,217	40-44
40-44	99,805	1,497	11	98,297	91,858	45-49
45-49	92,391	991	12	91,388	85,541	50-54
50-54	96,639	649	12	95,978	90,042	55-59
55-59	96,590	435	14	96,141	88,193	60-64
60-64	88,949	325	17	88,606	77,460	65-69
65-69	74,484	177	15	74,292	60,783	70-74
70-74	56,693	87	13	56,593	41,145	75-79
75+	80,788	86	29	80,673	42,545	80+
Totals	1,652,890	36,645	199	1,616,046	1,412,712	Totals

SOUTH EAST

Age group (initial)	Population (at c.d. 1966)	Surviving emigrants (1966-71)	Non-surviving emigrants (1966-71)	Reduced population at risk	Surviving stayers (1966-71)	Age group (final)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Birth	1,352,048	44,327	402	1,307,319	1,243,694	0-4
0-4	1,409,274	84,055	171	1,325,048	1,249,290	5-9
5-9	1,238,611	77,256	56	1,161,299	1,112,234	10-14
10-14	1,119,789	63,870	72	1,055,847	1,015,938	15-19
15-19	1,309,088	91,847	148	1,217,093	1,128,627	20-24
20-24	1,184,032	101,785	175	1,082,072	984,059	25-29
25-29	1,065,889	81,742	164	983,983	912,460	30-34
30-34	1,014,606	55,390	154	959,062	905,165	35-39
35-39	1,050,777	41,027	189	1,009,561	961,140	40-44
40-44	1,110,358	28,717	212	1,081,429	1,031,932	45-49
45-49	1,047,619	19,024	249	1,028,346	973,153	50-54
50-54	1,112,749	12,506	258	1,099,985	1,026,101	55-59
55-59	1,088,262	8,539	276	1,079,447	975,196	60-64
60-64	945,861	6,014	333	939,514	806,332	65-69
65-69	735,148	3,183	286	731,679	595,172	70-74
70-74	562,468	1,385	219	560,864	408,765	75-79
75+	801,795	1,782	615	799,398	420,930	80+
Totals	18,148,374	722,449	3,979	17,421,946	15,750,188	Totals

REST OF BRITAIN

Age group (initial)	Population (at c.d. 1966)	Surviving emigrants (1966-71)	Non-surviving emigrants (1966-71)	Reduced population at risk	Surviving stayers (1966-71)	Age group (final)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Birth	2,975,636	46,972	444	2,928,220	2,834,255	0-4
0-4	3,058,068	89,577	215	2,968,276	2,892,220	5-9
5-9	2,783,506	83,011	69	2,700,426	2,651,944	10-14
10-14	2,520,237	68,014	83	2,452,140	2,394,405	15-19
15-19	2,719,156	94,074	177	2,624,905	2,485,220	20-24
20-24	2,284,298	107,473	208	2,176,617	2,074,448	25-29
25-29	2,086,948	87,173	191	1,999,584	1,931,915	30-34
30-34	2,040,381	59,043	184	1,981,154	1,926,934	35-39
35-39	2,163,027	43,673	231	2,119,123	2,061,548	40-44
40-44	2,333,200	30,207	286	2,302,707	2,227,808	45-49
45-49	2,162,408	19,954	305	2,142,148	2,054,382	50-54
50-54	2,243,063	13,023	328	2,229,712	2,101,598	55-59
55-59	2,202,582	8,685	343	2,193,554	2,008,920	60-64
60-64	1,967,267	6,178	422	1,960,667	1,703,828	65-69
65-69	1,573,370	3,238	352	1,569,780	1,268,768	70-74
70-74	1,151,224	1,661	283	1,149,280	831,111	75-59
75+	1,438,635	1,501	593	1,436,541	713,197	80+
Totals	37,703,006	763,457	4,715	36,934,834	34,162,501	Totals

SOURCE: Rees (1980a).

APPENDIX 3. A SUMMARY OF HYBRID APPROACH AND TRANSITION APPROACH RESULTS

A.3.1 Hybrid Approach Results

Choice: Period choice deaths	migrants	Migrant rate choice	Probability/ stationary population choice	Run	East Anglia life (years) $^1_0\ell$	Percent $^1\theta^1$	$^1\theta^2$	$^1\theta^3$	South East life (years) $^2_0\ell$	Percent $^2\theta^1$	$^2\theta^2$	$^2\theta^3$	Rest of Britain life (years) $^3_0\ell$	Percent $^3\theta^1$	$^3\theta^2$	$^3\theta^3$
DATA SET ONE (gb1)																
1	1	equiv.	t-based/lin.	1	72.74	41.98	25.82	32.20	72.66	4.28	64.70	31.03	71.83	2.38	16.59	81.04
"	"	equiv.	" /exp.	2	72.70	42.00	25.81	32.18	72.62	4.28	64.70	31.02	71.79	2.38	16.59	81.03
"	"	ave.	" /lin.	3	72.74	40.82	26.27	32.91	72.66	4.36	63.99	31.65	71.83	2.42	16.87	80.71
"	"	ave.	" /exp.	4	72.71	40.84	26.27	32.89	72.63	4.36	63.99	31.64	71.79	2.42	16.87	80.70
"	"	equiv.	m-based/lin.	5	72.44	42.18	25.34	32.48	72.45	4.03	65.25	30.72	71.68	2.27	16.04	81.70
"	"	equiv.	" /exp.	6	72.35	42.17	25.53	32.31	72.33	4.21	64.68	31.13	71.49	2.45	17.22	80.33
"	"	ave.	" /lin.	7	72.43	41.02	25.78	33.19	72.45	4.11	64.55	31.35	71.49	2.31	16.31	81.38
"	"	ave.	" /exp.	8	72.34	41.02	25.97	33.01	72.32	4.28	63.98	31.74	71.49	2.50	17.51	79.99
DATA SET TWO (gb 5)																
5	5	equiv.	t-based/lin.	9	73.02	57.78	19.14	23.08	72.66	3.57	74.40	22.03	71.76	1.75	11.52	86.72
"	"	equiv.	" /exp.	10	73.00	57.78	19.14	23.08	72.65	3.57	74.40	22.04	71.74	1.75	11.53	86.72
"	"	ave.	" /lin.	11	73.00	56.02	20.05	23.93	72.66	3.64	73.63	22.73	71.78	1.88	11.85	86.27
"	"	ave.	" /exp.	12	72.98	56.02	20.05	23.93	72.64	3.65	73.62	22.73	71.77	1.88	11.85	86.27
"	"	equiv.	m-based/lin.	13	72.83	57.91	18.90	23.19	72.53	3.45	74.69	21.86	71.69	1.72	11.28	87.00
"	"	equiv.	" /exp.	14	72.78	57.85	18.99	23.16	72.45	3.55	74.25	22.20	71.55	1.82	11.94	86.24
"	"	ave.	" /lin.	15	72.79	56.19	19.78	24.03	72.51	3.51	73.95	22.54	71.70	1.83	11.59	86.58
"	"	ave.	" /exp.	16	72.74	56.13	19.87	24.00	72.43	3.62	73.50	22.88	71.55	1.93	12.29	85.77
DATA SET THREE (gb5 two)																
1	5	equiv.	t-based/lin.	17	72.92	57.75	19.19	23.06	72.80	3.57	74.46	21.99	71.69	1.76	11.57	86.67
"	"	equiv.	" /exp.	18	72.88	57.75	19.19	23.06	72.76	3.57	74.45	21.99	71.64	1.76	11.57	86.67
"	"	ave.	" /lin.	19	72.90	56.01	20.10	23.90	72.79	3.64	73.70	22.67	71.70	1.88	11.89	86.23
"	"	ave.	" /exp.	20	72.86	56.01	20.10	23.89	72.75	3.64	73.70	22.67	71.66	1.88	11.89	86.23
"	"	equiv.	m-based/lin.	21	72.73	57.93	18.94	23.13	72.66	3.44	74.79	21.78	71.61	1.71	11.31	86.98
"	"	equiv.	" /exp.	22	72.66	57.88	19.03	23.09	72.56	3.54	74.35	22.11	71.44	1.81	11.98	86.21
"	"	ave.	" /lin.	23	72.70	56.20	19.83	23.97	72.64	3.50	74.04	22.45	71.62	1.82	11.63	86.55
"	"	ave.	" /exp.	24	72.63	56.15	19.92	23.93	72.54	3.61	73.60	22.79	71.45	1.93	12.33	85.73

Note: See Table 5 for a fuller version of the choices.

A.3.2 Transition Approach Results

Choice						East Anglia				South East				Rest of Britain			
Period choice		Surv. rate choice	Prob. choice	Stationary population choice	Run	life (years)	Percent			life (years)	Percent			life (years)	Percent		
deaths	migrants					${}^{1}_{0}\ell$	${}^{1}\theta^{1}$	${}^{1}\theta^{2}$	${}^{1}\theta^{3}$	${}^{2}_{0}\ell$	${}^{2}\theta^{1}$	${}^{2}\theta^{2}$	${}^{2}\theta^{3}$	${}^{3}_{0}\ell$	${}^{3}\theta^{1}$	${}^{3}\theta^{2}$	${}^{3}\theta^{3}$
DATA SET TWO (gb5)																	
5	5	Direct	Opt. 2	Lin.	25	72.57	52.19	21.47	26.34	72.29	3.75	71.93	24.32	71.57	1.91	12.10	86.00
"	"	"	Lin. Int.	Lin.	26	72.30	52.50	21.32	26.18	72.03	3.72	72.11	24.17	71.32	1.89	12.04	86.07
"	"	"	"	Direct	27	72.57	52.19	21.47	26.34	72.29	3.75	71.93	24.32	71.57	1.91	12.10	86.00
"	"	"	Cubic S.	Lin.	28	72.78	52.50	21.23	26.27	72.55	3.73	72.07	24.20	71.84	1.87	12.04	86.09
"	"	"	"	Direct	29	72.57	52.19	21.47	26.34	72.29	3.75	71.93	24.32	71.57	1.91	12.10	86.00
"	"	Cond.	Opt. 2	Lin.	30	72.72	52.18	21.47	26.35	72.48	3.76	71.91	24.33	71.72	1.91	12.10	85.99
"	"	"	Lin. Int.	Lin.	31	72.33	52.50	21.32	26.18	72.09	3.73	72.11	24.16	71.34	1.89	12.04	86.07
"	"	"	"	Direct	32	72.72	52.18	21.47	26.35	72.48	3.76	71.91	24.33	71.72	1.91	12.10	85.99
"	"	"	Cubic S.	Lin.	33	72.61	52.50	21.23	26.27	72.38	3.73	72.07	24.20	71.63	1.87	12.04	86.08
"	"	"	"	Direct	34	72.72	52.18	21.47	26.35	72.48	3.76	71.91	24.33	71.12	1.91	12.10	85.99
DATA SET THREE (gb5 two)																	
1	5	Cond.	Opt. 2	Lin.	35	73.54	51.99	21.64	26.37	73.51	3.79	71.85	24.35	72.40	1.94	12.25	85.81
"	"	"	Lin. Int.	Lin.	36	73.24	52.29	21.50	26.22	73.20	3.77	72.02	24.21	72.11	1.92	12.19	85.89
"	"	"	"	Direct	37	73.54	51.99	21.64	26.37	73.51	3.79	71.85	24.35	72.40	1.94	12.25	85.81
"	"	"	Cubic S.	Lin.	38	73.53	52.28	21.41	26.31	73.50	3.77	71.98	24.25	72.41	1.90	12.20	85.90
"	"	"	"	Direct	39	73.54	51.99	21.64	26.37	73.51	3.79	71.85	24.35	72.40	1.94	12.25	85.81
DATA SET ONE (gb1)																	
1	1	Cond.	Cubic S.	Lin.	40	72.43	40.63	25.78	33.59	72.44	4.29	63.66	32.05	71.70	2.41	16.76	80.83

Note: See Table 18 for a fuller version of the choices.

Abbreviations have the following meaning:

Surv. = survivorship   Prob. = probability   Stat. population = Stationary population   Cond. = conditional

Opt. 2 = Option 2   Lin. Int. = linear interpolation   Cubic S. = cubic spline

SELECTED PAPERS ON MIGRATION AND
SETTLEMENT AT IIASA

Andrei Rogers, editor, *Migration and Settlement: Selected Essays*, RR-78-6, reprinted from *Environment and Planning A* 10(5):469-617.

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* 9(1979):275-310.

Andrei Rogers, editor, *Essays in Multistate Mathematical Demography*, RR-80-10, reprinted from *Environment and Planning A* 12(5):485-622.

Nathan Keyfitz, *Multidimensionality in Population Analysis*, RR-80-33, reprinted from *Sociological Methodology 1980*, K. Schuessler, ed., pp. 191-218.