DEMOGRAPHY© Volume 10, Number 2 May 1973

ESTIMATING INTERNAL MIGRATION FROM INCOMPLETE DATA USING MODEL MULTIREGIONAL LIFE TABLES

Andrei Rogers
The Technological Institute, Northwestern University, Evanston, Illinois 60201

Abstract—A principal feature of current methods of estimating demographic measures from incomplete data is the use of model life tables that approximate the mortality of a region for which reliable mortality data are unavailable. Observed decennial rates of survivorship may be used to identify out of a set of such model life tables one that best matches the observed data. This paper introduces the concept of a model *multiregional* life table and outlines a procedure for selecting an appropriate one using place-of-birth-by-residence data.

Introduction

In most developing countries of the world today, studies that seek to assess the severity of population-related problems or that strive to evaluate the impacts of programs and policies on population trends are seriously impaired by the general absence of reliable data on mortality, fertility, and migration. At the same time, the need for accurate population data in such countries is urgent, for without this information it is virtually impossible to formulate intelligent plans for social and economic development.

Frequently the only reliable sources of demographic data are the decennial censuses of population. These offer indirect measures of mortality, fertility, and migration, for the age distribution of a population manifests the past behavior of these fundamental components of population growth and change. For example, if a population undisturbed by migration is enumerated in two consecutive decennial censuses, and if each census contains tabulations of the population by five-year age groups, then it is a simple matter to compute the proportion of each cohort that survived the decade. Thus, in Table 1, we find that the proportion of persons 10 to 14 in the 1960 Brazilian Census to those 0 to 4 in the 1950 Census is 0.9236, which is an approximation of the survival ratio ${}_{10}s(0) = {}_5L_{10}/{}_5L_0$ in a life table representing the schedule of mortality during the intercensal decade. (We have also assumed that the Brazilian population was stable. This, of course, is not a necessary assumption and merely reflects the absence of place-of-residence-by-place-of-birth data in the 1950 Census. Since such data for two consecutive decennial censuses are needed later in this paper, we have been forced to use the assumption of stability in order to illustrate our method.)

Two fundamental problems need to be resolved in order to construct a life table from survival ratios such as appear in column 3 of Table 1. First, to complete a life table constructed on the basis of 10-year survival ratios, we must either have available adequate records of births during the decade or be able to estimate infant and child mortality indirectly, for example, by associating the mortality rates of persons under age five to those over five. Second, as is illustrated by the survival ratios in column 3 of Table 1,

TABLE 1.—Population of Brazil, by Age, and Age-Specific Survival Proportions

Age at Last Birthday in 1950	Population in 1950[a] w (1950)	Population in 1960[b] w (1960)	Initial Estimate of 10-Year Survival Proportion[c] $_{10}\hat{s}$	Model Life Table 5-Year Survival Probability[d] $_5s$
x	(1)	(2)	(3)	(4)
0	8,864,612	8,187,797	0.9236	0.949
5	7,295,864	7,379,287	1.0114	0.983
10	5,993,054	6,131,768	1.0231	0.982
15	5,401,265	5,176,448	0.9583	0.975
20	4,488,145	4,555,825	1.0152	0.970
25	3,788,899	3,915,675	1.0334	0.966
30	3,334,634	3,302,842	0.9904	0.961
35	2,866,077	2,758,709	0.9625	0.955
40	2,417,513	2,297,136	0.9502	0.946
45	2,019,236	1,822,006	0.9023	0.931
50	1,681,387	1,348,394	0.8019	0.907
55	1,333,616	936,095	0.7019	0.871
60	986,956	1,076,529	0.4375	0.818
65	684,174	--	--	0.744
70	787,965	--	--	0.640
75+		--	--	0.395

a-Obtained by multiplying each corresponding element in the 1960 population, in the last column of Table 4, by 0.73195, the reciprocal of the 10-year stable growth ratio.
b-Source: Table 4.
c-Equation (3). With these survival proportions we enter Table 1.3 of United Nations (1967), p. 94 and obtain, by interpolation, estimates of the expectation of life at birth.
d-Interpolated using model life tables in United Nations (1967), pp. 87-88 (e(0) = 50.00).

the considerable age misreporting and differential omission by age that occurs in such censuses often leads to survival ratios that are obviously too low or that exceed unity. Most of the available methods for adjustment are essentially arbitrary and may significantly affect the age pattern of mortality that is being estimated.

To assist developing countries in their quest for better methods of estimating demographic measures from incomplete data, the United Nations recently published a manual on this subject (United Nations, 1967). The principal feature of the procedures outlined there is the use of *model life tables* which utilize the regularities exhibited by available mortality data collected in countries with accurate registration systems to systematically approximate the mortality of a region for which such data are unavailable. Entering these tables with $_{10}\hat{s}(0) = 0.9236$, say, we find that the associated expectation of life at birth is approximately 46 years. Repeating this procedure with $_{10}\hat{s}(15) = 0.9583$ and $_{10}\hat{s}(55) = 0.7019$, say, we obtain expectations of life at birth of approximately 56 and 48 years, respectively. Thus an average value of about 50 years seems to be indicated by the data. The *five-year* survival probabilities that are associated with a United Nations model life table with an expectation of life at birth of 50 years appear in column 4 of Table 1.

The above procedure, with a few refinements, forms the essence of the "U.N. method" of estimating a single-region life table from incomplete data. The purpose of this paper is to illustrate how the same method may be generalized to include the estimation of multiregional life tables (see Appendix) which describe the mortality and mobility schedules of an interregional population system that is open to internal migration. We begin by introducing the concept of *model multiregional life tables* and then describe how we may enter such tables with a set of initial estimates of survival ratios and out-migration proportions, obtained by applying the PRPB method (Rogers and von Rabenau, 1971) to two consecutive place-of-birth-by-place-of-residence census age distributions.

Model Multiregional Life Tables

A model single-region life table approximates the mortality schedule of a region that is closed to migration by resorting to the mortality experience in other regions with populations and conditions that may be presumed to be similar to the one being studied. The most common method for incorporating the mortality experience of more than one region is by means of regression, and several studies have used this method to develop model life tables for use in situations where accurate observed data are unavailable (Coale and Demeny, 1966; United Nations, 1967). The usual procedure in such studies is to associate the regional *probabilities of dying at each age*, $q_i(x)$, with the regional *expectation of life at a given age*, ${}_ie(x)$. For example, consider the following regional male expectations of life at birth and probabilities of dying within the first five years of age:

Region of Birth (i)	${}_ie(0)$	$q_i(0)$
1. San Francisco	67.62	0.029863
2. Los Angeles	67.50	0.032562
3. San Diego	67.44	0.033425
4. Rest of Calif.	67.35	0.032782
5. Rest of the U.S.	66.74	0.036100

A simple linear regression of the five $q_i(0)$'s on the associated five ${}_ie(0)$'s yields

$$q_i(0) = 0.42487 - 0.00582\,{}_ie(0) \qquad (1)$$

and a coefficient of determination (r^2) of 0.84. Thus, the probability of dying within the first five years of age for males in a region in the United States may be approximated by inserting the appropriate expectation of life at birth into (1) and then solving for $q_i(0)$. Comparable equations may be derived for all age groups $x = 0, 5, 10, \cdots$, and the results then may be used to compute a model single-region life table.

A model multiregional life table approximates the mortality and mobility schedule of a system of regions that are open to migration by resorting to the mortality and mobility experience in other regions with populations and conditions that may be presumed to be similar to the ones being studied. As in the case of model single-region life tables, we may summarize the generalized experience of several regions by means of regression. Specifically, we may associate the regional *probabilities of out-migrating at each age*, $p_{ij}(x)$, with the *fraction of the expectation of life at birth of persons born in region i that is expected to be lived in region j*, ${}_ie_j(0)/{}_ie(0)$. For example, consider the regional male expectations of life at birth and the probabilities of out-migrating within the first five years of age that appear in Tables 2 and 3. A simple linear regression of

TABLE 2.—Regional Expectations of Life at Birth, by Place of Residence: U. S. Males, 1955–1960

Region of Birth	Region of Residence 1	2	3	4	5	Total
1. San Francisco	32.51	5.50	1.10	5.59	22.92	67.62
2. Los Angeles	4.11	36.06	1.56	3.62	22.16	67.50
3. San Diego	3.64	7.67	21.72	2.46	31.95	67.44
4. Rest of Calif.	8.81	7.39	1.27	27.09	22.78	67.35
5. Rest of the U.S.	1.34	2.69	0.58	0.87	61.26	66.74

TABLE 3.—Regional Probabilities of Out-migrating within the First Five Years of Age: U. S. Males, 1955–1960

Region of Birth	Region of Destination 1	2	3	4	5
1. San Francisco	--	0.019455	0.002767	0.025164	0.043671
2. Los Angeles	0.015556	--	0.005049	0.014432	0.044633
3. San Diego	0.011699	0.030910	--	0.006736	0.060250
4. Rest of Calif.	0.038998	0.027956	0.002855	--	0.035587
5. Rest of the U.S.	0.002062	0.005028	0.000693	0.001074	--

the 20 $p_{ij}(0)$'s on the associated 20 fractions ${}_ie_j(0)/{}_ie(0)$ yields

$$p_{ij}(0) = 0.00586 + 0.11837[{}_ie_j(0)/{}_ie(0)] \tag{2}$$

and a coefficient of determination of 0.84. Thus, the probability of out-migrating within the first five years of age for males in a region in the United States may be approximated by inserting the appropriate pair of expectations of life at birth into (2) and then solving for $p_{ij}(0)$. Comparable equations may be obtained for all age groups $x = 0, 5, 10, \cdots$ and all pairs of interregional flows $i, j = 1, 2, \cdots$ $(i \neq j)$. The results then may be combined with those found for mortality in order to compute a model multiregional life table, such as the one that appears in abbreviated form in Table 4.

Selecting the Appropriate Model Multiregional Life Table

The United Nations method of using two consecutive decennial census-enumerated age distributions to obtain the initial age-specific estimates of the 10-year survival proportions may be generalized to multiregional population systems if age-specific place-of-residence-by-place-of-birth data are available for both census years. (If the population is stable, then, as in the single-region case, one census distribution and the stable growth rate are sufficient, for if the multiregional population is stable, then so are its place-of-residence-by-place-of-birth age distributions [c.f. equation (12) in Rogers and von Rabenau, 1971].) This may be easily demonstrated by defining the single-region procedure in algebraic form and then reverting to

TABLE 4.—Model Multiregional (Two-Region) Life Table: $_ie(0) = 48.00$, $_ie_j(0) = 3.00$; $_je(0) = 56.00$, $_je_i(0) = 1.00$

Age at Last Birthday x	Survival Probability $_5\underline{s}_{ii}$	Out-migration Probability $_5\underline{s}_{ij}$	Death Rate $\underline{d}_i$	Out-migration Rate $\underline{o}_i$	Survival Probability $_5\underline{s}_{jj}$	Out-migration Probability $_5\underline{s}_{ji}$	Death Rate $\underline{d}_j$	Out-migration Rate $\underline{o}_j$
0	.903001	.010998	.028128	.001278	.943252	.003646	.017147	.000415
5	.955623	.014249	.006880	.002065	.976262	.004720	.004318	.000684
10	.957233	.012191	.005353	.001677	.976369	.004039	.003346	.000556
15	.945145	.015512	.007116	.001627	.969130	.005150	.004574	.000539
20	.931015	.022434	.009000	.003155	.961810	.007414	.005857	.001044
25	.927682	.019835	.010147	.002966	.958600	.006540	.006640	.000980
30	.925674	.015665	.011508	.002276	.955565	.005164	.007548	.000751
35	.922744	.012881	.012752	.001953	.951804	.004240	.008478	.000644
40	.919444	.009750	.013937	.001449	.946752	.003210	.009509	.000478
45	.906495	.007957	.015515	.001174	.934887	.002615	.011049	.000387
50	.882233	.006978	.020488	.001014	.914633	.002290	.014869	.000334
55	.841617	.006614	.026833	.000974	.881251	.002160	.020014	.000319
60	.783157	.006422	.039988	.000990	.831045	.002087	.030119	.000323
65	.704240	.005889	.056022	.001001	.758342	.001898	.043944	.000325
70	.595842	.004894	.084484	.000916	.654997	.001568	.068318	.000295
75	.417788	.003646	.125338	.000849	.487352	.001133	.105143	.000272
80	.345166	.003761	.252055	.000797	.462207	.001161	.203532	.000246

matrix algebra to define the corresponding multiregional method. First, observe that the single-region procedure for estimating $s_i(x)$ may be expressed as follows:

$$\begin{aligned} {}_{10}\hat{s}_i(x) &= \frac{w_i^{(t+10,x+10)}}{w_i^{(t,x)}} \\ &= \{w_i^{(t,x)}\}^{-1} w_i^{(t+10,x+10)} \qquad (3) \end{aligned}$$

where $w_i^{(t,x)}$ denotes the number of persons at age x in region i at time t. Next, consider the two-region model of population growth, say, which distinguishes persons by place of residence and by place of birth (Rogers and von Rabenau, 1971):

$$\begin{aligned} {}_iw_i^{(t+10,x+10)} &= {}_{10}s_{ii}(x)\,{}_iw_i^{(t,x)} + {}_{10}s_{ji}(x)\,{}_iw_j^{(t,x)} \\ {}_jw_i^{(t+10,x+10)} &= {}_{10}s_{ii}(x)\,{}_jw_i^{(t,x)} + {}_{10}s_{ji}(x)\,{}_jw_j^{(t,x)}, \end{aligned}$$

where

${}_hw_k^{(t,x)}$ = the number of persons at age x in region k at time t who were born in region h; and

${}_{10}s_{hk}(x)$ = the proportion of persons at age x in region h at time t who are at age $x + 10$ in region k at time $t + 10$.

Expressing the above in matrix form, we have that

$$w_i^{(t+10,x+10)} = \{W^{(t,x)}\}\,{}_{10}s_i(x); \qquad (4)$$

whence

$${}_{10}\hat{s}_i(x) = \{W^{(t,x)}\}^{-1} w_i^{(t+10,x+10)}. \qquad (5)$$

Observe that (5) is the matrix expression of (3).

To illustrate the application of (5), we present in Table 5 the 1960 total population of Brazil disaggregated by age, place of residence, and place of birth. Since such data for 1950 are unavailable, we assume that the population is stable and obtain comparable 1950 data by multiplying each element in Table 5 by 0.73195, the reciprocal of the observed ten-year growth ratio. Applying (5) to each region and every age group, we obtain the initial age-specific estimates of the 10-year survival and out-migration proportions that are set out in Table 6. Observe that, as in Table 1, a few of the survival proportions exceed unity. And note that several out-migration proportions are negative.

In the single-region case we "adjusted" the initial estimates of survival proportions that appear in column 3 of Table 1 by resorting to model single-region life tables. (The adjusted estimates were set out in column 4 of Table 1.) Analogously, in the multiregional case, we may adjust the initial estimates of survival and out-migration proportions that appear in Table 6 by interpolating in an appropriate set of model multiregional life tables. (The adjusted estimates are the proportions that appear in Table 4.) The interpolation may be carried out by regressing expectations of life at birth, by place of residence, on age-specific survival and out-migration proportions, as follows:

$${}_he_k(0) = f[{}_{10}s_{hh}(x),\, {}_{10}s_{hk}(x),\, {}_{10}s_{kk}(x),\, {}_{10}s_{kh}(x)]. \qquad (6)$$

Conclusion

The purpose of this paper has been to describe a multiregional generalization of what by now is probably the "conventional" method of estimating mortality from incomplete data for single regions that are closed to migration. The generalization draws on the concept of a model multiregional life table and the observation that the PRPB method of estimating interregional migration streams from place-of-residence-by-place-of-birth data is simply the multiregional counterpart of the single-region procedure for obtaining initial "unadjusted" estimates of survival proportions.

The calculations presented in this

Table 5.—Population of Brazil by Age, Region of Residence, and Region of Birth: 1960

Age at Last Birthday	Population of the North-Center-West Region, (i), by Age and Place of Birth			Population of the Rest of Brazil Region (j), by Age and Place of Birth			Total Population of Brazil
x	W_i	$_iW_i$	$_jW_i$	W_j	$_jW_j$	$_iW_j$	W
0-4	964,648	936,120	28,528	11,146,312	11,136,797	9,515	12,110,960
5-9	865,376	807,936	57,440	9,102,338	9,087,044	15,294	9,967,714
10-14	709,357	638,075	71,282	7,478,441	7,458,394	20,047	8,187,797
15-19	587,401	511,044	76,357	6,791,886	6,763,832	28,054	7,379,287
20-24	489,284	392,641	96,643	5,642,485	5,613,903	28,582	6,131,768
25-29	403,929	319,037	84,892	4,772,519	4,744,507	28,012	5,176,448
30-34	321,959	251,139	70,820	4,233,866	4,210,924	22,942	4,555,825
35-39	303,966	226,379	77,587	3,611,709	3,591,533	20,176	3,915,675
40-44	238,143	181,395	56,748	3,064,697	3,046,101	18,596	3,302,842
45-49	191,930	145,870	46,060	2,566,779	2,551,176	15,603	2,758,709
50-54	143,057	101,854	41,203	2,154,079	2,143,139	10,940	2,297,136
55-59	97,242	71,029	26,213	1,724,765	1,716,802	7,963	1,822,006
60-64	84,969	57,133	27,836	1,263,425	1,256,218	7,207	1,348,394
65-69	45,445	27,298	18,147	890,650	888,118	2,532	936,095
70+	62,592	37,804	24,788	1,013,937	1,008,938	4,999	1,076,529
Total	5,509,297	4,704,753	804,544	65,457,888	65,217,426	240,462	70,967,185

Source: Special tabulation for the Brazilian Ministry of Planning expanded proportionately from a 1 1/2 percent sample to the total population reported in the Brazilian Census of 1960.

TABLE 6.—Initial Estimates of the Age-Specific 10-Year Survival and Out-migration Proportions: North-Center-West Region (i) and the Rest of Brazil Region (j)

Age at Last Birthday x	$_{10}\hat{s}_{ii}$	$_{10}\hat{s}_{ij}$	$_{10}\hat{s}_{jj}$	$_{10}\hat{s}_{ji}$
0	0.9312	0.0200	0.9149	0.0064
5	0.8641	0.0282	1.0167	0.0060
10	0.8404	0.0289	1.0281	0.0097
15	0.8525	0.0223	0.9581	0.0075
20	0.8737	0.0052	1.0247	0.0022
25	0.9690	-0.0044	1.0343	0.0050
30	0.9866	0.0109	0.9881	0.0018
35	0.8805	0.0077	0.9703	-0.0015
40	0.7667	-0.0162	0.9615	0.0042
45	0.6650	-0.0238	0.9198	0.0020
50	0.7660	0.0107	0.8006	0.0030
55	0.5243	-0.0306	0.7072	0.0064
60+	0.4224	0.0032	0.4371	0.0013

Source: Computed using the data in Table 4.

study have been included only in order to illustrate the methodology, and no claim is made regarding their accuracy or applicability to other study areas. Moreover, much theoretical and empirical work still remains to be done before the procedures set out in this paper can become practically useful. In concluding this paper, therefore, we shall mention a few directions of research that warrant further study.

First, the theoretical foundation of multiregional life tables is incomplete and should receive the attention of demography's more mathematically sophisticated scholars. Second, empirical efforts to construct sets of model multiregional life tables should be encouraged, for it is only through such studies that we will be able to determine whether regularities similar to those found in single-region life tables also may be identified in multiregional life tables. Third, the use of model multiregional life tables to estimate other demographic measures from incomplete data should be explored. For example, as in the single-region case, a region's annual birth rate can be estimated by adding the annual rate of increase to the estimated annual death and net migration rates which may be found by applying the life table values for these rates to the average of the two consecutive decennial age distributions. Finally, alternative methods for selecting the appropriate model multiregional life table should be investigated. For example, instead of the PRPB initial estimates of survival and out-migration proportions, we could, in the absence of place-of-birth data, use estimates of these proportions obtained from a time series of population distributions (Rogers, 1967, 1971a). With a two-region population system, we would need at least three decennial censuses:

$$w_i^{(t+10,x+10)} = {}_{10}s_{ii}(x)w_i^{(t,x)} + {}_{10}s_{ji}(x)w_j^{(t,x)}$$

$$w_i^{(t+20,x+10)} = {}_{10}s_{ii}(x)w_i^{(t+10,x)} + {}_{10}s_{ji}(x)w_j^{(t+10,x)},$$

which in matrix form is

$$w_i^{(x+10)} = \{W^{(x)}\}_{10}s_i(x); \qquad (7)$$

whence

$${}_{10}\hat{s}_i(x) = \{W^{(x)}\}^{-1}w_i^{(x+10)} \qquad (8)$$

Note the similarity of (8) and (5).

Appendix

The Multiregional Life Table

Imagine a population system consisting of two-regions, i and j, say, and let δ denote the state of being dead. An individual in region i at time t may, during the interval $(t, t+1)$, migrate continually between i and j or enter the state of death, δ. The central problem in going from observed data on deaths and migration, in such a two-region population system, to a two-region life table describing the mortality and mobility regime for this population lies in the estimation of the age-specific migration and death probabilities $p_{ij}(x)$ and $q_i(x)$, where

$p_{ij}(x)$ = probability that an individual at aged x and present in region i at time t will survive and be present in region j at time $t+1$;

$q_i(x)$ = probability that an individual at aged x and present in region i at time t will die before time $t+1$; and

$$p_{ii}(x) = 1 - p_{ij}(x) - q_i(x). \qquad (1)$$

In our study, we have adopted a method of estimation that generalizes Chiang's (1968) single-region concept of fraction of person-years lived (see Rogers, 1971b).

With estimates of the age-specific migration and death probabilities for each region, we may compute a two-region life table in a manner that is analogous to the procedures used to obtain a single-region life table. First, let

${}_{ix}l_{ii}(y)$ = expected number of survivors alive at age y in region j among the $l_i(x)$ individuals of age x now alive in region i;

${}_{ix}l_{ij}(y)$ = expected number of survivors alive at age $y+1$ in region j among the ${}_{ix}l_i(y)$ individuals of age y now alive in region i and previously living in region i at age x;

${}_{ix}d_i(y)$ = expected number of deaths between ages y and $y+1$ among the ${}_{ix}l_i(y)$ individuals of age y now alive in region i and previously living in region i at age x.

Next, observe that

$${}_{ix}l_i(x) = l_i(x), \qquad (2)$$

and define

$${}_{ix}l_i(y) = 0, \quad \text{for} \quad y < x \qquad (3)$$

$${}_{ix}l_j(y) = 0, \quad \text{for} \quad y \leq x. \qquad (4)$$

Finally, note that

$${}_{ix}l_i(y) = {}_{ix}l_{ii}(y) + {}_{ix}l_{ij}(y) + {}_{ix}d_i(y), \qquad (5)$$

$${}_{ix}l_i(y+1) = {}_{ix}l_{ii}(y) + {}_{ix}l_{ji}(y), \qquad (6)$$

$${}_{ix}l_{ij}(y) = {}_{ix}l_i(y)p_{ij}(y), \qquad (7)$$

$${}_{ix}d_i(y) = {}_{ix}l_i(y)q_i(y). \qquad (8)$$

To obtain the two-region life table, we start our cycle of computations with $l_i(x)$ and apply (7) and (2) to find ${}_{ix}l_{ii}(x)$:

$${}_{ix}l_{ii}(x) = {}_{ix}l_i(x)p_{ii}(x) = l_i(x)p_{ii}(x). \qquad (9)$$

Next, we draw on (7) and (4) to define

$${}_{ix}l_{ji}(x) = {}_{ix}l_j(x)p_{ji}(x) = 0. \qquad (10)$$

Therefore, by (6),

$${}_{ix}l_i(x+1) = {}_{ix}l_{ii}(x) + {}_{ix}l_{ji}(x) = {}_{ix}l_{ii}(x). \qquad (11)$$

Setting $y = x$ in (7) yields ${}_{ix}l_{ij}(x)$, and (7) and (4) may be used to define

$$ {}_{ix}l_{ji}(x) = {}_{ix}l_j(x)p_{ji}(x) = 0. \qquad (12)$$

Thus, by (6),

$$ {}_{ix}l_j(x+1) = {}_{ix}l_{jj}(x) + {}_{ix}l_{ij}(x) = {}_{ix}l_{ij}(x). \qquad (13)$$

Given $l_i(x)$, ${}_{ix}l_{ii}(x)$, and ${}_{ix}l_{ij}(x)$, we may apply the appropriate fractions of person-years lived to obtain the number of person-years lived by the $l_i(x)$ individuals in region i and in region j, respectively:

$$ {}_{ix}L_{ii}(x) = a_{ii}\ {}_{ix}l_{ii}(x) + a_{ij}\ {}_{ix}l_{ij}(x) + a_{i\delta}\ {}_{ix}d_i(x) \qquad (14)$$

and

$$ {}_{ix}L_{ij}(x) = (1 - a_{ii})\,{}_{ix}l_{ii}(x) + (1 - a_{ij})\,{}_{ix}l_{ij}(x), \qquad (15)$$

where, for example, a_{ij} denotes the fraction of a year lived in region i by persons who were in i at time t and in j at time $t + 1$.

Analogously, recalling (4), (10), and (12), we conclude that

$$ {}_{ix}L_{jj}(x) = 0, \qquad (16)$$

and

$$ {}_{ix}L_{ji}(x) = 0. \qquad (17)$$

We begin the next cycle of computations by using ${}_{ix}l_{ii}(x)$ to obtain ${}_{ix}l_i(x+1)$ and ${}_{ix}l_{ij}(x)$ to find ${}_{ix}l_j(x+1)$, following (11) and (13). Applying (7), with the appropriate subscript modifications, twice to ${}_{ix}l_i(x+1)$ and twice to ${}_{ix}l_j(x+1)$, we can derive ${}_{ix}l_{ii}(x+1)$, ${}_{ix}l_{ij}(x+1)$, ${}_{ix}l_{jj}(x+1)$, and ${}_{ix}l_{ji}(x+1)$. Using these quantities, we may obtain ${}_{ix}L_{ii}(x+1)$, ${}_{ix}L_{ij}(x+1)$, ${}_{ix}L_{jj}(x+1)$, and ${}_{ix}L_{ji}(x+1)$.

The remainder of the life table specific to the "cohort" $l_i(x)$ now follows directly. First, we complete the survivorship and migration history of this particular group of individuals. The resulting set of ${}_{ix}L_{ii}(y)$, ${}_{ix}L_{ij}(y)$, ${}_{ix}L_{jj}(y)$, and ${}_{ix}L_{ji}(y)$ values permits us to compute the total person-years in prospect beyond age $k (k \geq x)$ for the "cohort" $l_i(x)$, by region of residence. For example,

$$ {}_{ix}T_i(k) = \sum_{y=k}^{\omega} [{}_{ix}L_{ii}(y) + {}_{ix}L_{ji}(y)], \qquad (18)$$

where ω denotes the maximum age to which anyone can live. The expectation of life beyond age k for the "cohort" $l_i(x)$, by region of residence, follows directly. For example,

$$ {}_{ix}e_i(k) = \frac{{}_{ix}T_i(k)}{{}_{ix}l_i(k)}. \qquad (19)$$

Denoting the average expected remaining lifetime beyond age k of an individual alive at age x in region i by ${}_{ix}e(k)$, we have that

$$ {}_{ix}e(k) = \frac{{}_{ix}T_i(k) + {}_{ix}T_j(k)}{{}_{ix}l_i(k)} = {}_{ix}e_i(k) + {}_{ix}e_j(k). \qquad (20)$$

That is, an individual currently at age x in region i can expect to live a total of ${}_{ix}e(k)$ years beyond age $k (k \geq x)$, of which ${}_{ix}e_i(k)$ years will be spent in region i and ${}_{ix}e_j(k)$ years will be spent in region j.

The above described process for computing a life table specific to the "cohort" $l_i(x)$ may be applied to generate such a table for each $l_i(y)$, $j = 1, 2, \cdots, n$. The entire process than may be repeated for each $l_j(y)$, $y = 0, 1, 2, \cdots, \omega$. The mass of information that is the end product of all these computations we shall call *the multiregional life table.*

One of the most useful applications of a multiregional life table is in population projection. The expected number of persons between ages x and $x + 5$ surviving five years is

$$\int_0^5 w_i(x+t)\,\frac{l_i(x+5+t)}{l_i(x+t)}\,dt, \qquad (21)$$

where $w_i(x)$ denotes the number of persons at age x.

The conventional approximation used in integrating (21) over the five-year interval is the substitution for (21) of the following product of two integrals divided by a third (Keyfitz, 1968, p. 247):

$$\frac{\int_0^5 w_i(x+t)\,dt \int_0^5 l_i(x+5+t)\,dt}{\int_0^5 l_i(x+t)\,dt}$$

$$= \frac{W_i(x)L_i(x+5)}{L_i(x)} = W_i(x)s_i(x), \text{ say.} \qquad (22)$$

Because, in our two-region population system, each regional population is exposed to migration, its survivors five years hence may be found in both of the two regions: $W_i(x)s_{ii}(x)$ in region i and $W_i(x)s_{ij}(x)$ in region j, say, where

$$s_i(x) = s_{ii}(x) + s_{ij}(x), \qquad (23)$$

with $s_{ij}(x)$ denoting the probability that an individual in region i at age x will be in region j at age $x + 5$. But

$$s_i(x) = \frac{L_{i\cdot}(x+5)}{L_{\cdot i}(x)}$$

$$= \frac{{}_{ix}L_{ii}(x+5) + {}_{jx}L_{ii}(x+5) + {}_{ix}L_{ij}(x+5) + {}_{jx}L_{ij}(x+5)}{{}_{ix}L_{ii}(x) + {}_{jx}L_{ji}(x)}. \qquad (24)$$

Thus, to obtain $s_{ii}(x)$ and $s_{ij}(x)$, we equate the right-hand sides of (23) and (24) and partition the numerator in (24) into total stationary regional populations in region i and in region j, respectively:

$$s_{ii}(x) + s_{ij}(x) = \frac{[{}_{ix}L_{ii}(x+5) + {}_{jx}L_{ii}(x+5)] + [{}_{ix}L_{ij}(x+5) + {}_{jx}L_{ij}(x+5)]}{{}_{ix}L_{ii}(x) + {}_{jx}L_{ji}(x)},$$

whence

$$s_{ii}(x) = \frac{{}_{ix}L_{ii}(x+5) + {}_{jx}L_{ii}(x+5)}{{}_{ix}L_{ii}(x) + {}_{jx}L_{ji}(x)}, \qquad (25)$$

and

$$s_{ij}(x) = \frac{{}_{ix}L_{ij}(x+5) + {}_{jx}L_{ij}(x+5)}{{}_{ix}L_{ii}(x) + {}_{jx}L_{ji}(x)}. \qquad (26)$$

Acknowledgments

This research has been supported by a grant from the Ford and Rockefeller Foundations Program in Social Science, Law, and Population Policy. The author gratefully acknowledges the research assistance of Jacques Ledent and Antonio Rocha.

References

Chiang, Chin Long. 1968. Introduction to Stochastic Processes in Biostatistics. New York: John Wiley and Sons.

Coale, Ansley J., and Paul Demeny. 1966. Regional Model Life Tables and Stable Populations. Princeton: Princeton University Press.

Keyfitz, Nathan. 1968. Introduction to the Mathematics of Population. Reading, Massachusetts: Addison-Wesley.

Rogers, Andrei. 1967. Estimating Interregional Population and Migration Operators from Interregional Population Distributions. Demography 4:515–531.

———. 1971a. Matrix Methods in Urban and Regional Analysis. San Francisco: Holden-Day.

———. 1971b. The Multiregional Life Table. Multiregional Growth Models Project, Technological Institute, Working Paper, No. 2. Evanston, Ill.: Northwestern University.

———, and Burkhard von Rabenau. 1971. Estimation of Interregional Migration Streams from Place-of-Birth-by-Residence Data. Demography 8:185–194.

United Nations. 1967. Methods of Estimating Basic Demographic Measures from Incomplete Data. Manual IV, Manuals on Methods of Estimating Population. Series A, Population Studies, No. 42. New York: United Nations.