STABLE MIGRATION RATES FROM THE MULTIREGIONAL GROWTH MATRIX OPERATOR

LEROY O. STONE*

RESUMEN

La distribución de población entre numerosos grupos de edades y localizaciones regionales es transformada desde un punto del tiempo hacia otro por medio de una matriz de operaciones que fue recientemente formulada por Rogers en esta revista. Rogers discute la estabilidad de la distribución de edades por regiones, y este artículo continua esa discusión mediante la indicación de fórmulas para tasas estables de migración para edad específica y probabilidades para cada región. Esas formulas deben ser de algún interes desde que la mayor parte de las discusiones previas en teoría de poblaciones estables tratan con poblaciones cerradas.

SUMMARY

The distribution of population among several age groups and regional locations is transformed from one time point to another by means of a matrix operator that was recently formulated by Rogers in this journal. Rogers discussed the stability of the regional age distributions, and this article continues the discussion by indicating the formulas for stable age-specific migration rates and the probabilities for each region. These formulas should be of some interest since most of the previous discussion in stable population theory deals with closed populations.

In a paper recently published in this journal,[1] Andrei Rogers used a matrix formulation to demonstrate again[2] the eventual stability of a regional age distribution which is subject to constant age-specific mortality, fertility, and migration rates. He also indicated the stable growth rate of population (the dominant characteristic root of the matrix operator).

It might be of some interest to add to this discussion by indicating formulas for stable age-specific migration ratios and the probabilities in each region. Classical work in stable population theory has dealt largely with closed populations, thus ignoring migration.[3] The purpose of this note is, then, to indicate these formulas.

It is appropriate to begin by recapitulating the basic definitions in Rogers' paper, but extending them to incorporate a desirable improvement not mentioned there.[4]

Let w_{iat} refer to the number of persons in the ath age group and the ith region at time t. The set of such numbers for all $n \cdot m$ groups ($i = 1, 2. \ldots, n$; $a = 1, 2, \ldots, m$) is put in a column vector W_t. The value of this vector at time $t + 1$ is generated by a square matrix operator G_t, so that

$$W_{t+1} = G_t \cdot W_t \,. \qquad (1)$$

Aside from zeroes, G_t has three kinds of elements in Rogers' formulation.[5] The first is "S_{iat}" which represents the proportion who will be found alive in region i at time $t + 1$ among those who were alive in *the same* region at time t. Over this

* Dominion Bureau of Statistics, Canada. The writer is grateful to Andrei Rogers, and is solely responsible for any errors in this paper.

[1] Andrei Rogers, "The Multiregional Matrix Growth Operator and the Stable Interregional Age Structure," *Demography*, III (1966) 537–44.

[2] Cf. Hannes Hyrennius, "Population Growth and Replacement," in Philip M. Hauser and Otis Dudley Duncan (eds.), *The Study of Population* (Chicago: University of Chicago Press, 1959), pp. 472–85; Hannes Hyrennius, "La mesure de la reproduction et de l'accroissement naturel," *Population*, III (1948), 271–92.

[3] The relevant literature is becoming voluminous. Typical works are Alvaro Lopez, *Problems of Stable Population Theory* (Princeton: Princeton University Press, 1961), and Nathan Keyfitz, "The Population Projection as a Matrix Operator," *Demography*, I (1964), 56–73.

[4] Rogers, *loc. cit.*

[5] *Ibid.* The reader should also refer to Rogers' paper for the exact arrangement of the elements in G_t, which is a bit too cumbersome to be repeated again here.

time interval these persons have aged from a to $a + 1$.

The second type of element of G_t reflects inter-regional migration. Let "$p_{j(t+1)|iat}$" represent the proportion who will migrate to region j and survive there till time $t + 1$ among those who were aged a years in region i at time t. Again these persons are aged $a + 1$ at time $t + 1$.

The third element of G_t reflects births between times t and $t + 1$. This is "b_{ia}," which Rogers defines as the number of surviving births per person of ath childbearing age in region i. These are the births which are found alive at time $t + 1$, and the rate is multiplied by w_{iat} to generate the contribution of persons aged a at time t to the survivors of births in region i.

It is not clear how Rogers' formulation of b_{ia} generally allows for the inter-regional migration between t and $t + 1$ of the births that are within this time interval. For an interval of as much as ten years (which is used in Rogers' examples), it is a safe assumption that the parents who are in-migrating to region i between t and $t + 1$ bring with them their offspring of the same time interval, and have more children before time $t + 1$. It would appear that Rogers' formulation does not generally allow for this potentially important impact of migration on the number of survivors of recent births (since time t) on the child population of each region.

In order to correct this apparent deficiency we may redefine b_{ia} and provide a fourth type of element for the matrix operator G_t. Let "$b_{iat,i}$" represent the rate at which persons aged a at time t in region i contribute surviving births to the population of the *same* region at time $t + 1$. Thus, the actual number of survivors so contributed is $w_{iat} \cdot b_{iat,i}$. Further let "$b_{iat,j}$" be the rate at which persons aged a at time t in region i contribute survivors of births (between t and $t + 1$) to the population of region j at time $t + 1$; so that the actual number so contributed is given by $w_{iat} \cdot b_{iat,j}$. This contribution is mainly a result of the migration of parents (between t and $t + 1$) from region i to region j. The element $b_{iat,i}$ replaces Rogers' b_{ia} in the matrix operator, while the element $b_{iat,j}$ is placed in the iath row and jth column. As first approximations, we may use

$$b_{iat\ i} \cong b_{ia} \cdot S_{iat} , \tag{2}$$

and

$$b_{iat.j} \cong b_{ia} \cdot p_{j(t+1)|iat} . \tag{3}$$

These suggested modifications of Rogers' formulation do not affect the discussion that follows.

THE STABLE MIGRATION RATES

The total number of in-migrants (between t and $t + 1$) to region i, who were aged a at time t is given by formula (4).

$$I_{iat} = \sum_{j=1}^{n} p_{i(t+1)|jat} \cdot w_{jat} , \qquad i \neq j . \tag{4}$$

The ratio of this number to a selected denominator defines the in-migration rate. As the denominator, an arbitrary choice is usually made among w_{iat}, $w_{i(a+1)(t+1)}$, and some linear combination of these two terms. Because of the simplicity of the results, w_{iat} is used here. Thus the in-migration rate is *defined* here as in formula (5).

$$\gamma(I_{iat}) = \sum_{j=1}^{n} p_{i(t+1)|jat} \cdot (w_{jat}/w_{iat}) . \tag{5}$$

Since G_t is a square non-negative matrix (of dimension $m \cdot n \times m \cdot n$), its dominant characteristic root is non-negative and unique. Associated with this root is a vector W with a unique non-negative direction in the $m \cdot n$-dimensional space.[6] As is well known, the dominant root is the largest positive value of λ that satisfies the determinantal equation (6).

$$|G_t - \lambda I| = 0 , \tag{6}$$

and it defines the latent growth rate of population. The associated latent vector W is such that

$$G_t \cdot W = \lambda W . \tag{7}$$

[6] Rogers, *loc. cit.*

Having solved (6) for λ, we may substitute it into formula (7) and solve the following system of linear equations

$$(G_t - \lambda I)W = 0 \,, \qquad (7a)$$

to get W. The ratio of any two components of W remains constant over all nontrivial solutions of W. Thus, in the stable population (defined as the latent vector W) the ratio on the right-hand side of equation (5) is a constant with respect to j, i, and a. That is

$$w_{jat}/w_{iat} = c_{ija} \text{ (constant for the stable population).} \qquad (8)$$

Thus, equation (5) defines the stable age-specific in-migration ratio.

$$\gamma(I_{iat}) = \sum_{j=1}^{n} p_{i(t+1)|jat} \cdot c_{ija} \,, \; j \neq i \,. \qquad (5a)$$

We may also define the age-specific in-migration probability (the out-migration probability for all areas outside of region i). To do this, observe that in the stable population

$$w_{jat}/\sum_{j=1}^{n} w_{jat} = h_{ja} \,, \qquad j \neq i \quad (9)$$

is a constant with respect to j, a, and i. The stable age-specific in-migration probability for region i is defined in equation (10).

$$P^r(I_{iat}) = \sum_{j=1}^{n} p_{i(t+1)|jat} \cdot h_{ja} \,, \; i \neq j \,. \qquad (10)$$

For the puposes of computation the following alternatives for equations (5a) and (10) can be shown to be correct.

$$\gamma(I_{iat}) = \lambda \cdot c_{ia} - S_{iat} \,, \qquad (5b)$$

where

$$w_{i(a+1)t}/w_{iat} = c_{ia} \text{ (constant).}$$

$$P\gamma(I_{iat}) = (\lambda \cdot c_{ia} - S_{iat}) \times (h_{ja}/c_{ija}) \,. \qquad (10a)$$

Table 1.—STABLE DECENNIAL MIGRATION RATIOS AND PROBABILITIES[(a)] FOR CALIFORNIA AND THE REST OF THE UNITED STATES

(Based on the parameters in Roger's matrix operator aud referring to persons who were alive at the end of the decade)

Areas	Age group 10-19 years at the beginning of the ten-year interval				
	Ratios (percent)			Probabilities	
	In-migration	Out-migration (b)	Net migration	In-migration (c)	Out-migration (c)
California.............	8.5	15.0	-6.5[(d)]	0.04	0.15
Rest of the United States...............	7.0	4.0	3.0[(d)]	0.15	0.04

[(a)] Figures are calculated from equations (5a), (5b), (10), (10b), (11) and (12).

[(b)] Out-migration is probability expressed as a percentage.

[(c)] Figures are accounted for by the fact that there are only two regions.

[(d)] These figures reflect possible errors in the relevant source data. With an out-migration probability of 0.15 for California, the proportion who survive in California should not have exceeded 0.85; but Rogers' figure for this proportion is 0.93. It is not likely that this represents the survival rate independent of migration, for the implied 7 percent of deaths is much too high for the age group 10-19 years (at the beginning of the decade). The crude decennial death rate would be of this order.

SOURCE: A. Rogers, op. cit., p. 542.

The *time-independent* out-migration probability for the same age group and region is already defined in setting up G_t. It is as shown in equation (11).

$$\sum_{j=1}^{n} p_{j(t+1)|iat}\,, \qquad j \neq i\,. \quad (11)$$

By the definition of "$p_{j(t+1)|iat}$," the stable *net* migration ratio for the same age group and region is as shown in equation (12).

$$\gamma(N_{iat}) = \gamma(I_{iat}) - \sum_{=1}^{n} p_{j(t+1)|iat}\,, \qquad i \neq j\,. \quad (12)$$

Of course all the rates and probabilities pertain to migrants who survive to the end of the unit time interval of migration. Calculation of the foregoing ratios and probabilities may be illustrated from Rogers' example of two regions: California and the rest of the United States. (See Table 1.)

Table 1 indicates a stable decennial net migration ratio of *minus* six percent for California among persons aged 10–19 years at the beginning of each decade. This result is based on parameters estimated from data for the 1950–60 decade;[7] and it differs very sharply from the 1950–60 net migration ratio for California in the age group 10–19 years. According to Eldridge's estimates,[8] this ratio was positive and at least 40 percent. As noted in the footnotes to Table 1, Rogers' 'Survival-in-California' ratio for the 10–19 age group seems much too high.[9]

[7] *Ibid.*, 541.

[8] Hope T. Eldridge, *Net Intercensal Migration for States and Geographic Divisions of the United States, 1950–1960,* (Philadelphia: Population Studies Center, University of Pennsylvania, 1965) p. 155.

[9] There are several possible sources of the apparent discrepancy, including the adjustment of the matrix operator to ensure that it predicts exactly the enumerated 1960 population (which was subject to enumeration error). While the apparent discrepancy does not reflect negatively on the value of Rogers' formulation, it does indicate the need to exercise much care in the estimation of the matrix operator. The specific issue here is the extent to which errors in the basic population statistics should be allowed to influence the elements in the matrix operator.