ON THE SPEED OF CONVERGENCE TO STABILITY

Young J. Kim

Department of Population Dynamics
The Johns Hopkins University
Baltimore, Maryland 21205

Running Title: SPEED OF CONVERGENCE

ABSTRACT

Two alternative representations of the dynamics of populations in n age groups are presented using column-stochastic and row-stochastic matrices that are similar to the normed population projection matrix. One representation describes a process for the reciprocal of age-specific momentum, while the other describes a Markov chain for the age distribution of total population reproductive value. Both processes are characterized by the stable net maternity function, and the limiting distribution for both processes is given by the age distribution of total reproductive value in the stable population. Using populations in two age groups, it is shown explicitly how the speed of convergence is determined by the shape of the stable net maternity function. Finally, a large number of observed populations in 10 age groups are used to find measures that best describe the speed of convergence. Population entropy, and the skewness of the age distribution of stable population reproductive value are found to have rank correlations of around 0.98.

Introduction

Stable population theory states that when age-specific vital rates remain unchanged over time, a population with an arbitrary initial age distribution comes to have a fixed age distribution. The resultant age distribution depends only on the specified vital rates, and is independent of the initial age distribution. The speed of convergence to stability is generally known to depend on the shape of the net maternity function.

In pioneering work on convergence to stability, Coale (1968; 1972) related the speed of convergence to the moments of net maternity function. By adjusting the area under the net maternity function, he considered a relative difference between the real root and the real part of the complex root that has the largest real part among all complex roots. With this rescaling of net maternity function, he concluded that convergence is faster for a smaller mean, a larger variance, and a larger skewness. Keyfitz (1972; 1985) also gave qualitatively the same results as those obtained by Coale . Le Bras (1969), on the other hand, rescaled the net maternity function by transforming it to the stable net maternity function. Since the maximal root of the stable net maternity function is 1, he considered only the complex roots of the stable net maternity function in terms of moments.

Using alternative representations of the population dynamics, Sykes (1973) demonstrated that convergence to stability is governed by the stable net maternity function. Arthur (1981, 1982) also showed that the dynamics of the growth adjusted birth sequence is determined by the stable net maternity function. How the speed of convergence is determined by the stable net maternity function was discussed by Tuljapurkar (1982). He argued that population entropy determines the rate at which the Kullback distance

decreases.

In this paper, we use the discrete time formulation of the dynamics. We first present two alternative parameterizations of the dynamics of populations in n age groups, following Sykes (1973), in much detail than presented by Tuljapurkar (1982). It is shown that the dynamics of an initial population to its stable limit is equivalent to a renewal process and/or a Markov chain. The stable net maternity function characterizes both these processes. Then populations in two age groups are considered to gain insight into the way a population converges to stability. Finally, an empirical study of correlates of the speed of convergence to stability is presented using a large number of actual observed human populations with 10 age groups. Two measures that correlate slightly better than population entropy are found. The paper ends with a brief discussion.

The Dynamics of a Population in n Age Groups

We follow the usual restriction to a closed female population, and define the population in n age group at time t as a column vector of length n, $x_t = (x_{t1}, x_{t2}, \cdots, x_{tn})'$ [a prime denotes a transpose of a vector or a matrix]. The dynamics of the population are then described as

$$x_t = A \, x_{t-1} \quad , \tag{1}$$

where $A = \{a_{ij}\}$ is an n x n population projection matrix (PPM) that has elements

$$a_{ij} = \begin{cases} b_j, & i=1 \\ s_j, & i=2,3,\cdots,n; \; j=i-1 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

In equation (2), b_j is the "birth" rate that represents the number of persons in the first age group at time t+1 per person in the jth age group at time t, and s_j is the probability that persons in the jth age group survive to age group j+1. From equation (1), the population at time t in terms of an initial population x_0 can be written as

$$x_t = A^t x_0 \quad .$$
(3)

We let λ_j represent eigenvalues of A, and u_j and v_j denote the right and left eigenvectors associated with λ_j, j=1, 2, $\cdots$,n. If we assume the PPM A to be irreducible and primitive [mild conditions satisfied by all human populations (Sykes, 1969)], then there exists a dominant eigenvalue of multiplicity one. Let that eigenvalue be λ, and let the right and left positive eigenvectors associated with λ be denoted as u and v. With this notation, equation (3) can be written, using the spectral decomposition of A (Leslie, 1945; Keyfitz, 1968, p. 59-62),

$$x_t/\lambda^t = (v,x_0)u + \Sigma_{j=2}^{n} (\lambda_j/\lambda)^t(v_j,x_0)u_j \quad ,$$
(4)

where (v,x) denotes the inner product, $\Sigma_{i=1}^{n} v_i x_i$, and hence is a scalar. Thus the speed of convergence to stability is determined by how fast the difference

$$x_t/\lambda^t - (v,x_0)u - \Sigma_{j=2}^{n} (\lambda_j/\lambda)^t(v_j,x_0)u_j$$
(5)

converges to a zero vector. In equations (4) and (5), the eigenvalues λ_j are roots of the characteristic equation

$$\Sigma_{i=1}^{n} \lambda^{-i}p_i b_i = 1 \quad ,$$
(6)

and u_j and v_j are the corresponding right and left eigenvectors. Elements of the positive eigenvectors $u=\{u_i\}$ and $v=\{v_i\}$ associated with λ are given by

$$u_i = \lambda^{-(i-1)}p_i$$

$$v_i = \Sigma_{j=i}^{n} \lambda^{-j}p_jb_j \ /[u_i \ \Sigma_{j=1}^{n}\Sigma_{k=j}^{n} \ \lambda^{-k}p_kb_k] \quad , \tag{7}$$

where $p_i = \Pi_{j=1}^{i-1} s_j$ is the probability of survival from the first age group to the $\underline{i\text{th}}$ age group, with p_1 defined as 1. The eigenvectors in equation (7) are normalized in such a way that $u_1 = 1$ and $(u,v)=1$. The double sum in the denominator of equation (7) gives the mean age μ of the stable net maternity function, i.e.,

$$\mu = \Sigma_{j=1}^{n}\Sigma_{k=j}^{n} \ \lambda^{-k}p_kb_k \quad . \tag{8}$$

For an alternative representation of the dynamics represented in equation (1), define the n x n diagonal matrix, $U = \text{diag}(u_i)$, where u_i are components of u. [This part heavily draws on Sykes, 1973; see also Tuljapurkar, 1982]. By premultiplying equation (1) by U^{-1} and dividing by λ^t, we obtain

$$U^{-1}x_t/\lambda^t = [U^{-1}AU/\lambda] \ U^{-1}x_{t-1}/\lambda^{t-1} \quad . \tag{9}$$

By defining a column vector of unit length y_t as

$$y_t = U^{-1}x_t/[\lambda^t(v,x_0)] \tag{10}$$

and a matrix F as

$$F = U^{-1}AU/\lambda \quad , \tag{11}$$

equation (9) can be written

$$y_t - F\ y_{t-1} \quad . \tag{12}$$

The matrix $F - \{f_{ij}\}$ in equations (11) and (12) has the same pattern as the PPM A, and its elements are given by

$$f_{ij} - \begin{bmatrix} f_j, & i-1 \\ 1, & i-2,3,\cdots,n; \ j-i-1 \\ 0, & \text{otherwise} \end{bmatrix} \tag{13}$$

where

$$f_j - \lambda^{-j} p_j b_j \tag{14}$$

is the stable net maternity function, or the age distribution of the characteristic equation, and hence F is a row-stochastic matrix. Although F is a row-stochastic matrix, equation (12) does not describe a Markov chain because it operates on a column vector. Instead, F defines a renewal process as will be shown below. This obviously lead Tuljapurkar (1982) to state that the initial state of the Markov chain associated with F is not naturally identified.

Equation (12) gives y_t in terms of an initial population y_0,

$$y_t - F^t\ y_0 \quad , \tag{15}$$

where F^t converges to

$$\lim_{t\to\infty} F^t - 1\ v(F)' \tag{16}$$

with 1 denoting a column vector all of whose elements are 1. The $v(F)'$, the column proportionality of the limiting matrix, has its ith element

$$v_i(F) - u_i v_i \quad , \tag{17}$$

which is the age distribution of total reproductive value in the stable population, and sums to 1. With equation (16), equation (15) then yields, in the limit,

$$\lim_{t \to \infty} y_{ti} = 1 \ , \qquad i=1,2,\cdots,n \qquad . \tag{18}$$

In this process, the <u>ith</u> element of y_t, $y_{ti} = x_{ti}/[\lambda^t(v,x_0)u_i]$, is the ratio of the number of persons in age group i to the corresponding stable equivalent population at time t (Keyfitz, 1971). That is, y_{ti} is the reciprocal of the age-specific momentum for the age group i (Schoen and Kim, 1991), which converges to unity as the population evolves toward stability . Equation (18) describes this result.

Finally, we note that the dynamics of y_t in equation (12) reduce to a one-dimensional process for the reciprocal of the birth momentum, $y_t(1)$,

$$y_t(1) = \Sigma_{k=1}^{n} f_k \ y_{t-k}(1) \qquad . \tag{19}$$

Equation (19) represents a renewal process (repeated averaging) of the growth adjusted birth sequence (reciprocal of birth momentum) governed by the stable net maternity function f. (cf. Arthur, 1981 and 1982).

Alternatively, we can premultiply the original process in equation (1) by a diagonal matrix, $V = \text{diag}(v_i)$, where v_i are components of v, and obtain

$$Vx_t/\lambda^t = [VAV^{-1}/\lambda] \ Vx_{t-1}/\lambda^{t-1} \qquad . \tag{20}$$

We define a vector of unit length, z_t,

$$z_t = Vx_t/[\lambda^t(v,x_0)] \ , \tag{21}$$

and the matrix G,

$$G = VAV^{-1}/\lambda \qquad . \tag{22}$$

The matrix G has the same pattern as the PPM **A**, and its elements are given by

$$g_{ij} = \begin{cases} g_j, & i=1 \\ 1-g_j, & i=2,3,\cdots,n; \quad j=i-1 \\ 0, & \text{otherwise} \end{cases} \qquad (23)$$

where

$$g_j = f_j \ / \ \Sigma_{i=j}^{n} \ f_i \quad , \qquad (24)$$

which represents the probability of reproduction conditioanl on not having reproduced by that age. By transposing equation (20), we obtain

$$z_t' = z_{t-1}' \ G' \quad , \qquad (25)$$

where G' is now a row-stochastic matrix, and thus equation (25) defines a Markov chain for z_t. The vector z_t defined in equation (21) is the age distribution of the population reproductive value. The limiting (stationary) distribution of the Markov chain is given by the right eigenvector **u**(G) of G, i.e.,

$$\lim_{t\to\infty} G^t = \ u(G) \ 1' \quad , \qquad (26)$$

where **u**(G) is again given by

$$u_i(G) = u_i v_i \quad . \qquad (27)$$

Thus from equations (25)-(27), the age distribution of reproductive values, z_t, converges to

$$\lim_{t\to\infty} z_t = \{u_i z_i\} \quad . \qquad (28)$$

In sum, the dynamics of a population described in equation (1) are equivalent to a renewal process for the reciprocal of birth momentum, and to a Markov chain for the age distribution of population reproductive value. Convergence of both processes is determined by the stable net maternity distribution, $f-\{f_i\}$.

While the above two representations show that convergence to stability is determined by the stable net maternity function $\{f_i\}$, they do not provide an answer to what characteristics of the distribution determine the speed of convergence. That the speed of convergence is determined by the relative sizes of the eigenvalues of the stochastic matrices F or G, where the maximal eigenvalue is unity, is of no use unless we know how the eigenvalues are determined by the distribution $\{f_j\}$. We now turn to this question in the context of a population with 2 age groups.

Convergence of Populations in Two Age Groups

By considering populations in two age groups, we can examine the convergence process more concretely without invoking the eigenstructure. Let 2 x 2 PPM be

$$A = \begin{bmatrix} a & b \\ p & 0 \end{bmatrix} , \qquad (29)$$

where $\underline{a}$ and $\underline{b}$ represent the "birth" rates in the age groups 1 and 2, respectively [b_1 and b_2 in equation (2) are written as $\underline{a}$ and $\underline{b}$ here to avoid subscripts], and p is the survival probability between age groups 1 and 2. Then as a special case of 2 x 2 model considered in Kim (1985), the tth power of the PPM, A^t, can be written as

$$A^t = (\Pi_{j=1}^{t} \lambda_j) \begin{bmatrix} 1 & b/\lambda_t \\ p/\lambda_t & bp/\lambda_{t-1}\lambda_t \end{bmatrix} , \tag{30}$$

where λ_t satisfies the generalized characteristic equation,

$$1 = a/\lambda_t + bp/\lambda_t\lambda_{t-1} . \tag{31}$$

Equations (30) and (31) hold for all values of $t = 1, 2, \cdots$ if we set initial conditions $\lambda_0=\infty$ and $\lambda_1=a$. The representation of A^t in equation (32) shows that convergence of $A^t/(\Pi\lambda_j)$ is completely determined by the convergence of λ_t to its limit λ, which is the positive real root of the characteristic equation,

$$1 = a/\lambda + bp/\lambda^2 . \tag{32}$$

In examining how λ_t converges to λ, it is more informative to rewrite equation (32) as

$$\lambda_t = a + bp/\lambda_{t-1} . \tag{33}$$

Then the signed difference between two successive λ_t's satisfies

$$\lambda_t - \lambda_{t-1} = bp(1/\lambda_{t-1} - 1/\lambda_{t-2}) , \tag{34}$$

which can be rewritten, using equation (33), as

$$\lambda_t - \lambda_{t-1} = - (\lambda_{t-1}-\lambda_{t-2})/[1 + (a/bp)\lambda_{t-2}] . \tag{35}$$

Repeated applications of equation (35) yield

$$\lambda_t - \lambda_{t-1} = (-1)^t\{1/\{[a/bp][1+(a/bp)\lambda_1]\cdots[1+(a/bp)\lambda_{t-2}]\}\} . \tag{36}$$

Since the $(a/bp)\lambda_j$ in the denominator of equation (36) is positive for all j, it shows that $|\lambda_t - \lambda_{t-1}|$ decreases monotonically over t, and that the magnitude of $(a/bp)\lambda_j$ determines the speed of convergence. That is, the larger the value of $(a/bp)\lambda_j$, the faster is the convergence.

Notice that the factor $(a/bp)\lambda_j$ is, in the limit, the ratio of a/λ to bp/λ^2, the elements of the stable net maternity function, or the characteristic equation. It thus shows that 1) the speed of convergence to stability is determined by the stable net maternity function, and more specifically, 2) the larger the first element (hence the smaller the second element since the two elements sum to one) the faster is the convergence.

The convergence in the 2 x 2 case can be seen more easily if we consider the row-stochastic matrix F,

$$F = \begin{bmatrix} 1-f & f \\ 1 & 0 \end{bmatrix} , \qquad (37)$$

instead of A in equation (29). Equation (37) can be considered as describing a population x_t that has a net reproduction rate (NRR) of 1, with no mortality between age groups 1 and 2, or more generally as describing arbitrary fertility and mortality schedules in the context of equation (12). At every time point t, F^t is given by

$$F^t = \begin{bmatrix} 1-f+f^2-\cdots+(-1)^{t-1}f^t & f-f^2+\cdots+(-1)^t f^t \\ 1-f+f^2-\cdots+(-1)^t f^{t-1} & f-f^2+\cdots+(-1)^{t-1}f^{t-1} \end{bmatrix} , \qquad (38)$$

which can be considered as a special case of that considered in Kim and Sykes (1978) for changing vital rates. Equation (38) converges to

$$\lim_{t \to \infty} F^t = (1/(1+f)) \begin{bmatrix} 1 & f \\ 1 & f \end{bmatrix} \qquad (39)$$

where $1+f = \mu$ is the mean of the stable net maternity function, defined in equation (8). Since convergence of $1-f+f^2-\cdots+(-1)^{t-1}f^t$ to $1/(1+f)$ is faster if f is smaller, a faster convergence to stability is achieved when the second element f is small. Thus we obtain the same result as that obtained for the general 2 x 2 case.

While the results in this section are suggestive, generalizations are difficult to make because the 2 x 2 case offers only one degree of freedom. The general n x n case, however, is not analytically tractable so we now turn to an empirical study.

Empirical Study of the Determinants of Convergence

Data compiled in two volumes of Keyfitz and Flieger (1968; 1971) were used in an empirical search for determinants of the speed of convergence to stability. There are 90 populations that have PPM tabulated in Keyfitz and Flieger (1971), and 177 in Keyfitz and Flieger (1968), totaling 267 populations.

From equation (5), the dominant factor that determines the speed at which a population converges to stability is the ratio $|\lambda_2|/\lambda$, where the eigenvalues are ordered $|\lambda_2|>|\lambda_3|>\cdots>|\lambda_n|$. Thus, the ratio $|\lambda_2|/\lambda$ was used as a reference for the speed of convergence to stability, and rank correlations were calculated for various measures. Since the moments of the net maternity function were traditionally considered as the determinants of the speed of convergence to stability (Coale, 1972, p.61-116; Keyfitz, 1972,

p.7-13; Keyfitz, 1985, p.255-262), we first examined the first three central moments of the net maternity function and stable net maternity function. Results are poor, except that the third moment of the stable net maternity function has correlation of around 0.9. These are shown in Appendix.

To illustrate the effect of the level of fertility on the stable net maternity function, reproductive patterns of Togo 1961 and Japan 1963 are presented in Figure 1. The top figure presents the net maternity function of Togo and Japan, which have similar age patterns, except for the level. The middle figure shows the stable net maternity function f, which shows a marked difference between the two populations, as a result of their different growth rates. The bottom figure shows the stable reproductive function $\{u_i v_i\}$.

Since Tuljapurkar (1982) showed that the population entropy H defined as

$$H - -(1/\mu) \; \Sigma \; f_i \; \ln(f_i) \quad , \qquad\qquad (40)$$

where $\{f_i\}$ is the stable net maternity function [equation (14)], and μ is the mean of that distribution [equation (8)], determines the asymptotic rate at which the Kullback distance K_t converges, H was rank correlated with $|\lambda_2|/\lambda$. The rank correlation is found to be .974, .980 and .979, respectively, in the 90, 177 and 267 populations. These correlations are high. When various composite measures of the first three moments were examined, the coefficient of variation (CV) of the net maternity function was found to correlate comparably with the population entropy.

Next, in view of the essential role $\{V_i\}-\{u_i v_i\}$ plays in the limiting distribution of both the Markov chain and renewal process, moments of $\{V_i\}$ were examined. Two measures that correlate as high as (or better than) population entropy H were found. They involve measures of skewness,

$$\beta_1(V) - (\mu_3)^2/(\sigma^2)^3 \ , \tag{41}$$

and

$$sk(V) - \mu_3/\mu^3 \ , \tag{42}$$

where μ, σ^2, and μ_3 are the first three moments of $\{V_i\}$, the age distribution of the stable population reproductive value. Their correlation coefficients are shown in Table 1.

Summary and Discussion

We have presented two alternative representations of the dynamics of a population x_t in n age groups. One of the two matrices involved is row-stochastic, and defines a process for the reciprocal of age-specific momentum of population growth. The other matrix is column-stochastic, and defines a Markov chain for the age-specific population reproductive value. Both matrices are completely characterized by the stable net maternity function, which is the age component of the characteristic equation, $\{f_j\} - \{\lambda^{-j}p_j b_j\}$. Thus, the net maternity function enters only through the stable net maternity function f. Therefore, populations with the same shape of net maternity function, but at different levels, converge at varying speeds (as do Togo and Japan), while populations with different shapes of net maternity function, but with a common stable net maternity function converge at the same speed.

Search for the determinants of the speed of convergence to stability was approached in two ways. First, an analytical result was obtained by using populations in two age groups. It was shown that convergence is faster when the first element of the stable net maternity function is larger, thus if the mean is smaller and the skewness is larger. However, with only two age

groups, all moments change simultaneously when one element of the distribution changes, so that more specific answers cannot be obtained.

We then turned to an experimental study. Using a large number of observed human populations in ten age groups, the rank order of populations by $|\lambda_2|/\lambda$ was correlated with rank orders of various "determinants". The population entropy H correlates highly, as do two measures of the skewness of the stable population reproductive value function V. The coefficient of variation of the stable net maternity function f also correlates highly.

In summary, a larger value of H, and more right-skewed stable population reproductive value distribution bring about faster convergence to stability. The measures of the skewness of the stable population reproductive value function correlates slightly better than population entropy H in all subsets analyzed. The coefficient of variation of the stable net maternity function also correlates highly, though slightly less than the others. That this measure correlates highly is expected because both CV and H measure the spread of the distribution adjusted for the mean (of the stable net maternity function).

References

Arthur, W.B. (1981). Why a population converges to stability. American Mathematical Monthly 88 8:557-563.

Arthur, W.B. (1982). The ergodic theorems of demography: A simple proof. Demography 19: 439-445.

Coale, A.J. (1968). Convergence of a human population to a stable form. JASA 63: 395-435.

Coale, A.J. (1972). The Growth and Structure of Human Populations: A Mathematical Investigation. Princeton: Princeton University Press.

Keyfitz, N. (1968). Introduction to the Mathematics of Population. Reading: Addison-wesley.

Keyfitz, N. (1971). On the momentum of population growth. Demography 8: 71-80.

Keyfitz, N. (1972). Population waves, in Population Dynamics. T.N.E. Greville (ed.) New york: Academic Press.

Keyfitz, N. (1985). Applied Mathematical Demography. (2nd Ed.) New York: Springer-Verlag.

Keyfitz, N. and W. Flieger (1968). World Population: An Analysis of Vital Data. Chicago: University of Chicago Press.

Keyfitz, N. and W. Flieger (1971). Population. San Francisco: Freeman.

Kim, Y.J. (1985). On the dynamics of populations with two age groups. Demography 22(3): 455-468.

Kim, Y.J. and Z.M. Sykes (1978). Dynamics of some special populations with NRR=1. Demography 15(4): 559-569.

Le Bras, H. (1969). Retour d'une population à l'état stable après une "catastrophe". <u>Population</u> <u>26</u> : 525-572.

Leslie, P.H. (1945). On the use of matrices in population mathematics. <u>Biometrika</u> <u>33</u> : 183-212.

Schoen, R. and Y.J. Kim (1991). Movement toward stability is a fundamental principle of population dynamics. <u>Demography</u> (in press).

Sykes, Z.M. (1969). On discrete population theory. <u>Biometrics</u> <u>25</u>: 285-293.

Sykes, Z.M. (1973). On discrete population theory. II: An alternative parameterization. (unpublished manuscript).

Tuljapurkar, S.C. (1982). Why use entropy? It determines the rate of convergence. <u>Journal of Mathematical Biology</u> <u>13</u>: 325-337.
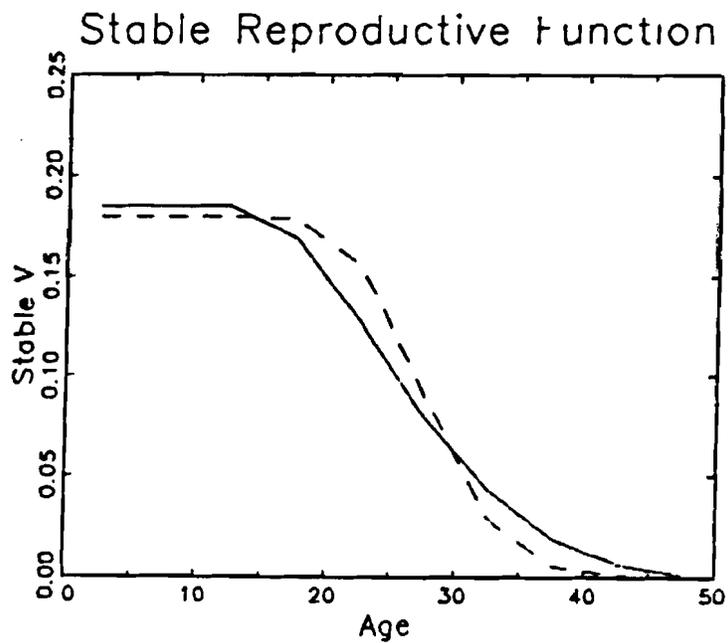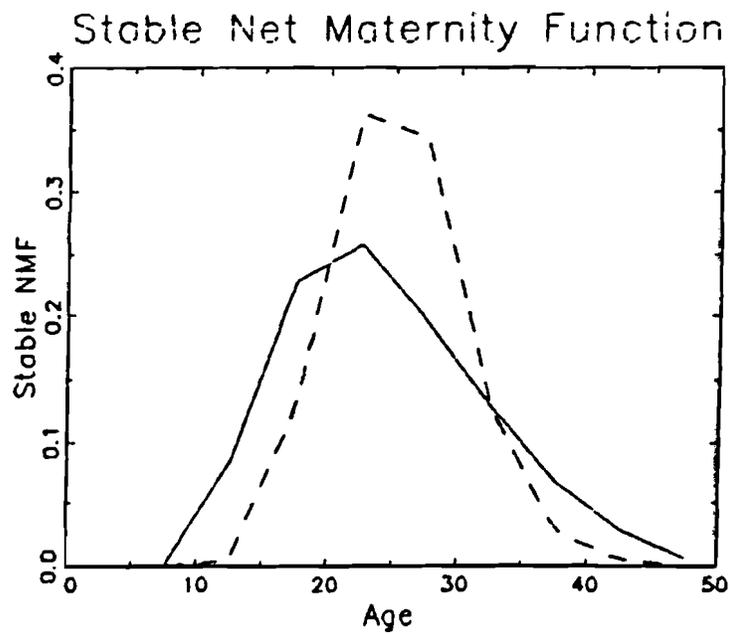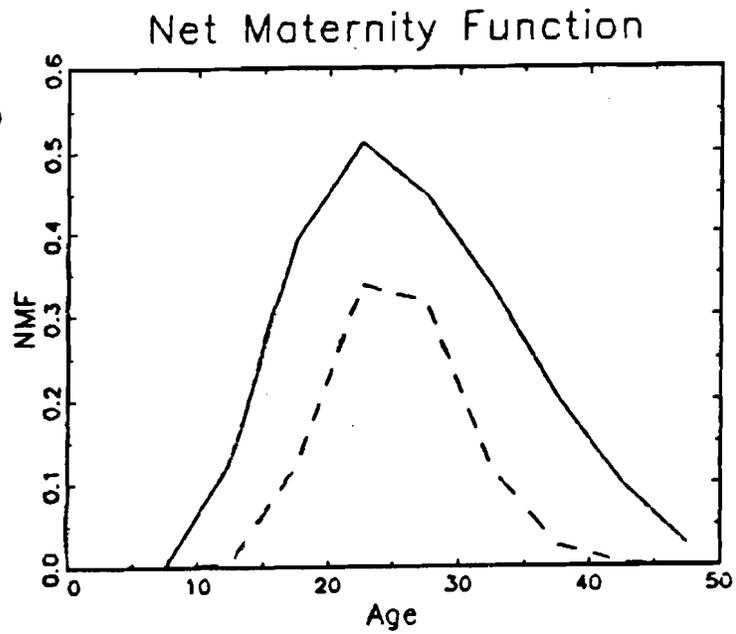
Table 1. Rank correlation (absolute value) coefficients between $|\lambda_2|/\lambda$ and various "determinants" of the speed of convergence to stability

	90 populations	177 populations	Total 267 populations
H(f)	0.974	0.980	0.979
CV(f)	0.975	0.955	0.961
β_1(V)	0.980	0.986	0.985
sk(V)	0.979	0.985	0.984

Note: See the text for notation.

Figure 1.

Reproductive patterns
for Togo 1961 (solid line)
and Japan 1963 (broken
line).

Net Maternity Function

Stable Net Maternity Function

Stable Reproductive Function

Table A. Rank correlation (absolute value) coefficients between $|\lambda_2|/\lambda$
and the first three moments of net maternity and stable net
maternity functions

	90 populations	177 populations	Total 267 populations
Maternity function			
μ	0.015	0.494	0.329
σ^2	0.800	0.522	0.651
μ_3	0.744	0.795	0.780
Net Maternity function			
μ	0.458	0.753	0.664
σ^2	0.773	0.411	0.548
μ_3	0.900	0.935	0.925