Demographic Research a free, expedited, online journal
of peer-reviewed research and commentary
in the population sciences published by the
Max Planck Institute for Demographic Research
Konrad-Zuse Str. 1, D-18057 Rostock · GERMANY
www.demographic-research.org

DEMOGRAPHIC RESEARCH

VOLUME 18, ARTICLE 14, PAGES 409-436
PUBLISHED 27 MAY 2008
http://www.demographic-research.org/Volumes/Vol18/14/

Research Article

Constant global population with demographic heterogeneity

Joel E. Cohen

Table of Contents

1	Introduction	410
2	The classical stationary population model	410
3	Heterogeneous stationary populations	412
3.1	Heterogeneous stationary populations: discrete case	412
3.2	Heterogeneous stationary populations: continuous case	416
3.3	Analytically soluble example	417
4	Constant global population with migration between countries	418
4.1	Alternative projections of a constant global population with international migration	422
5	Demographic data	427
6	Acknowledgements	430
	References	431
	Appendix 1	433
	Appendix 2	434
	Appendix 3	436

Demographic Research – Volume 18, Article 14
research article

Constant global population with demographic heterogeneity

Joel E. Cohen [1]

Abstract

To understand better a possible future constant global population that is demographically heterogeneous, this paper analyzes several models. Classical theory of stationary populations generally fails to apply. However, if constant global population size $P(global)$ is the sum of all country population sizes, and if constant global annual number of births $B(global)$ is the sum of the annual number of births of all countries, and if constant global life expectancy at birth $\mathbf{e}(global)$ is the population-weighted mean of the life expectancy at birth of all countries, then $B(global) \cdot \mathbf{e}(global)$ always exceeds $P(global)$ unless all countries have the same life expectancy at birth, in which case $B(global) \cdot \mathbf{e}(global) = P(global)$.

[1] Laboratory of Populations Rockefeller University & Columbia University 1230 York Avenue, Box 20, New York, NY 10065-6399 USA, 2008-02-20. E-mail: cohen@rockefeller.edu. Lab web page http://www.rockefeller.edu/labheads/cohenje/cohenje.html

1. Introduction

In most countries of the world, human birth rates have been dropping and human life expectancies have been increasing (apart from countries severely afflicted by HIV-AIDS, economic disruption, and violent disorder). The end of global population growth lies within the range of plausible scenarios for the coming half-century or century (United Nations Population Division 2005). For example, in its low variant prepared in 2004, the UN Population Division projected that global population would peak in 2040 and then decline (until 2050, the horizon of the projection). In its medium variant, the projected global net rate of reproduction fell to 1.00 by 2025-2030 and thereafter continued to decline, implying a future cessation of global population growth. Yet no demographic projections assume that all countries will have identical demographic characteristics.

It is therefore timely and useful to model a global population with unchanging total size and demographic variation among countries. This paper presents models of constant global populations and an apparently new inequality that connects population size, birth rates and life expectancy in a heterogeneous stationary population.

These models differ from graded models of stationary populations (e.g., Seal 1945; Vajda 1947; Bartholomew 1963; Keyfitz 1973), which assume that individuals advance through a linear succession of grades. Graded models have been used to study promotion in hierarchical organizations.

Coale (1972) emphasized that any earthbound population must have an average rate of growth that approaches zero as the time interval over which it is observed increases without limit, because infinite increase (implied by an average rate of increase that is positive, no matter how small) is impossible on a finite Earth. Numerous models have been proposed to investigate how populations approach stationary size, including models with and without age structure, with and without migration among subpopulations, and with rates of birth, death and migration that are constant or changing in time (Espenshade 1978; Feeney 1971; Kim and Schoen 1996; Land and Rogers 1982; Le Bras 1971; Rogers 1968, 1990, 1995; Rogers and Castro 1981; Rogers and Henning 1999; Rogers et al. 2004; Romaniuc 2005; Schoen 1988, 2002; Schoen and Kim 1993, 1998-not an exhaustive list). These thoughtful and useful studies do not appear to have attained the results presented here.

2. The classical stationary population model

The classical stationary population model describes an age-structured, single-sex population, closed to migration, with a constant number of births B per year equal to the constant number of deaths D per year so that the total population size P remains constant. The life table or survival function is the probability that a newborn individual survives to any given age or older, and is the complement of (i.e., one minus) the cumulative distribution function of length of life. The life table $l(x)$ depends on age only; it is assumed constant over time and applicable to all individuals. These attributes define a homogeneous stationary population.

A homogeneous stationary population arises in stable population theory from six conditions (Ryder 1975, p. 3): a fixed life table for females; a fixed maternity function for females; a net reproduction rate equal to 1; a fixed ratio of male to female births; a fixed survival function for males; and closure of the population to migration (this last condition can be relaxed). In the present paper, the assumptions about males will be ignored and a homogeneous stationary population will be treated as a single-sex model.

Three of the many well known properties of a homogeneous stationary population will be examined here.

First, in a homogeneous stationary population, if the life expectancy **e** is defined as the average number of years an individual lives from birth to death, then $P = B\mathbf{e}$ (e.g., Keyfitz 1968, p. 7, Pollard 1973, p. 11). If the per capita crude birth rate $b = \frac{B}{P}$ is the number of births per year per individual in the population, then $P = B\mathbf{e}$ is equivalent to $1 = \frac{B}{P}\mathbf{e} = b\mathbf{e}$ or $b = \frac{1}{\mathbf{e}}$.

Second, in a stationary population, let a^* be the age at which all individuals begin to make a constant contribution of c units of money (e.g., dollars or Euros) per year to a retirement fund, let b^* be the age at which all individuals stop contributing c units of money per year to a retirement fund and start drawing a constant annuity of one unit of money per year from it, and let w^* be the maximum length of life in this population. Also let r be the interest or discount rate (assumed fixed and positive) so that one unit of money invested today is worth e^{rt} units of money after t years, or one unit of money to be received t years in the future has discounted present value e^{-rt}. Then for a retirement system in balance (Alho and Spencer 2005, p. 84), the annual contribution every individual must make between age a^* and b^* to keep the retirement system in balance is $c = \frac{A}{D}$, where

$$A = \int_{b^*}^{w^*} e^{-rx}\, l(x)\, dx, \quad D = \int_{a^*}^{b^*} e^{-rx}\, l(x)\, dx.$$

Here $l(x)$ is the life table, A is the expected present value at the birth of an individual of his or her retirement annuity, and cD is the expected present value at the birth of an individual of his or her lifetime contributions to the retirement system. Balance requires that $cD = A$, hence $c = \frac{A}{D}$.

Third, in a stationary population, let $\overline{A}$ be the average age, σ^2 the variance of the length of life (from birth to death), and (as before) **e** the life expectancy (average length of life). Then $\overline{A} = \frac{\mathbf{e}+\frac{\sigma^2}{\mathbf{e}}}{2}$ (e.g., Ryder 1975, p. 8; Preston et al. 2001, p. 112).

3. Heterogeneous stationary populations

In a possible future, suppose each country has a homogeneous stationary population, but different countries have different life tables and birth rates, and migration between countries is assumed to be negligible or zero. I define this situation to be a heterogeneous stationary population. In this model, the world has constant population size but the life expectancy of an individual depends on the country of birth.

A substantively different but formally equivalent heterogeneous stationary population is an honorific society with different classes of members (corresponding to different fields of learning or achievement) in which each class is permitted to elect a fixed number of new members per year, and members remain, until death, in the society and in the same class within the society. In the absence of changes in survival and in the absence of changes in the age distribution at election, each class will eventually become a stationary population. Different classes may elect different numbers of new members each year (each such election counts as a "birth" to the class) and may have different life expectancies (from election to death).

Given that the three identities above ($P = B\mathbf{e}$, $c = \frac{A}{D}$, $\overline{A} = \frac{\mathbf{e}+\frac{\sigma^2}{\mathbf{e}}}{2}$) apply to each country's population (or each class's membership) separately, do they also apply to the global population if each quantity on both sides of each equation is correctly aggregated from the corresponding quantities for each country? The answer established below is, somewhat unexpectedly, no. None of the three identities holds in general for the global population if each quantity on both sides of each equation is correctly aggregated from the corresponding quantities for each country. For the identities $c = \frac{A}{D}$ and $\overline{A} = \frac{\mathbf{e}+\frac{\sigma^2}{\mathbf{e}}}{2}$, inequalities may hold in either direction, depending on the example.

The main new positive finding below is that the first identity $P = B\mathbf{e}$ is a special case of an inequality $P \leq B\mathbf{e}$ in which equality holds if and only if all countries have identical life expectancies. In greater detail, suppose the population size $P(global)$ is the sum of $P(country)$ for all countries, and the annual number of births $B(global)$ is the sum of $B(country)$ for all countries, and the life expectancy at birth $\mathbf{e}(global)$ is the population-weighted mean of $\mathbf{e}(country)$ for all countries. Then $B(global) \cdot \mathbf{e}(global)$ always exceeds or overstates $P(global)$ unless $\mathbf{e}(country)$ is the same for all countries.

3.1 Heterogeneous stationary populations: discrete case

In a heterogeneous stationary population, the mathematically most elementary case assumes a fixed, finite number m of countries, each indexed by $j = 1, \ldots, m$. Assume that country j has an annual number of births $B(j) > 0$, life table $l(j, x)$ (where j is the country, x is the age in years, $l(j,0)=1$ and $l(j,w^*)=0$), and life expectancy at birth

$$\mathbf{e}(j)=\int_0^{w^*} l(j,x)\,dx.$$

Then the stationary population size of country j is $P(j)=B(j)\mathbf{e}(j)$, the global population size is $\sum_{j=1}^m P(j)$ and the fraction of the global population in country j is $p(j)$=$\frac{P(j)}{P}$. The global number of births per year is $B=\sum_{j=1}^m B(j)$. The global population's life table $l(x)$ is the population-weighted mean of the country life tables

$$l(x)=\sum_{j=1}^m p(j)\,l(j,x);$$

Appendix 1 gives a detailed derivation. The global life expectancy is the population-weighted mean of the country life expectancies $\mathbf{e}=\sum_{j=1}^m p(j)\,\mathbf{e}(j)$.

We now prove that in a heterogeneous stationary population, $B\mathbf{e}\geq P$, and $B\mathbf{e}=P$ holds if and only if all countries have the same life expectancy.
Proof.

$$\begin{aligned}
B\mathbf{e} &= [B(1)+\ldots+B(m)][p(1)\mathbf{e}(1)+\ldots+p(m)\mathbf{e}(m)]\\
&= \frac{1}{P}[B(1)+\ldots+B(m)][P(1)\mathbf{e}(1)+\ldots+P(m)\mathbf{e}(m)]\\
&= \frac{1}{P}[B(1)+\ldots+B(m)][B(1)\mathbf{e}(1)^2+\ldots+B(m)\mathbf{e}(m)^2]\\
&\geq \frac{1}{P}[B(1)\mathbf{e}(1)+\ldots+B(m)\mathbf{e}(m)]^2\\
&= \frac{1}{P}[P(1)+\ldots+P(m)]^2\\
&= P.
\end{aligned}$$

The inequality follows from Cauchy's inequality (Pólya and Szegö 1972, p. 68), stated in Appendix 2 below, with $u(j)=\sqrt{B(j)}$ and $v(j)=\sqrt{B(j)}\mathbf{e}(j)$ for $j=1,\ldots,m$. Then equality holds if and only if $\mathbf{e}(j)$ is a constant independent of j, i.e., if and only if all countries have the same life expectancy. QED

The life expectancy at birth of a randomly chosen person, $\mathbf{e}=\sum_{j=1}^m p(j)\mathbf{e}(j)$, differs in general from the life expectancy of a randomly chosen newborn $e_b=\sum_{j=1}^m b(j)\mathbf{e}(j)$, where $b(j)=\frac{B(j)}{B}$ is the share of global births in country $j=1,\cdots,m$. A reader of a prior draft pointed out that, in a heterogeneous stationary population, $Be_b=P$. Should the weights in calculating global life expectancy be countries' shares of global population or countries' shares of global births? The following hypothetical example suggests that countries' shares of population are more appropriate weights. Suppose the world has $m = 2$ countries. Suppose each country is a stationary homogeneous population, and each has half the world's births each year. In country 1, life expectancy at birth is 100 years, while in country 2, life expectancy at birth is 10 years. Consequently, country 1 has 10 times the population size of country 2. The life expectancy of a randomly chosen newborn is $\frac{100+10}{2} = 55$ years, but the life expectancy of a randomly chosen person is $\frac{10}{11} \cdot 100 + \frac{1}{11} \cdot 10 = 91.8$ years as most of the world's people are living in country 1. In computing the global life expectancy, weighting each country's life expectancy at birth by its share of population is intuitively as well as analytically the preferable alternative.

To examine the pension contributions and annuities in a heterogeneous stationary population, we assume that all countries $j = 1, \dots, m$ have the same first age of contribution a^*, the same last age of contribution b^*, the same maximum length of life w^*, and the same discount rate r. For country j, the contribution per individual per year required to balance country j's retirement system is $c(j) = \frac{A(j)}{D(j)}$, where

$$A(j) = \int_{b^*}^{w^*} e^{-rx}\, l(j,x)\, dx \quad \text{and} \quad D(j) = \int_{a^*}^{b^*} e^{-rx}\, l(j,x)\, dx.$$

Therefore the average annual contribution per person if each country balances its own retirement system is

$$\overline{c} = \sum_{j=1}^{m} p(j)\, c(j) = \sum_{j=1}^{m} p(j)\, \frac{A(j)}{D(j)}.$$

The contribution c required per person using the global life table satisfies $cD = A$ where

$$A = \int_{b^*}^{w^*} e^{-rx}\, l(x)\, dx = \sum_{j=1}^{m} p(j)\, A(j)$$

and

$$D = \int_{a^*}^{b^*} e^{-rx}\, l(x)\, dx = \sum_{j=1}^{m} p(j)\, D(j).$$

Hence

$$c = \frac{\sum_{j=1}^{m} p(j)\, A(j)}{\sum_{j=1}^{m} p(j)\, D(j)}.$$

It is easy to generate numerical values of $p(j)$, $A(j)$ and $D(j)$ such that $\overline{c} < c$ and other numerical values such that $\overline{c} > c$.

Similarly, let $\overline{A}(j)$ be the average age, $\sigma^2(j)$ be the variance of the length of life from birth to death, and (as before) $\mathbf{e}(j)$ the life expectancy of country $j = 1, \ldots, m$. The variance of the length of life is the mean squared length of life minus the squared life expectancy. Then $\overline{A}(j) = \frac{1}{2}(\mathbf{e}(j) + \frac{\sigma^2(j)}{\mathbf{e}(j)})$. As before, the global life expectancy is $\mathbf{e} = \sum_{j=1}^{m} p(j)\, \mathbf{e}(j)$ and the global average age is

$$\begin{aligned}\overline{A} &= \sum_{j=1}^{m} p(j)\, \overline{A}(j) \\ &= \frac{\sum\limits_{j=1}^{m} p(j)\, \mathbf{e}(j) + \sum\limits_{j=1}^{m} p(j)\, \frac{\sigma^2(j)}{\mathbf{e}(j)}}{2} \\ &= \frac{\mathbf{e} + \sum\limits_{j=1}^{m} p(j)\, \frac{\sigma^2(j)}{\mathbf{e}(j)}}{2}.\end{aligned}$$

The global population's variance of the length of life is the population-weighted mean of the countries' mean squared length of life, minus the square of the population-weighted mean of the countries' life expectancies, or

$$\sigma^2 = \sum_{j=1}^{m} p(j)\, (\sigma^2(j) + \mathbf{e}^2(j)) - (\sum_{j=1}^{m} p(j)\, \mathbf{e}(j))^2.$$

Thus we are interested in the relation between

$$\overline{A} = \frac{\mathbf{e} + \sum\limits_{j=1}^{m} p(j)\, \frac{\sigma^2(j)}{\mathbf{e}(j)}}{2} \quad \text{and} \quad \frac{\mathbf{e} + \frac{\sigma^2}{\mathbf{e}}}{2},$$

or equivalently (after cancelling the factor of $\frac{1}{2}$ and removing the first term $\mathbf{e}$ found in both expressions) in the relation between

$$\sum_{j=1}^{m} p(j)\, \frac{\sigma^2(j)}{\mathbf{e}(j)} \quad \text{and} \quad \frac{\sigma^2}{\mathbf{e}}.$$

As in the previous case, it is easy to generate numerical values of $p(j)$, $\sigma^2(j)$ and $\mathbf{e}(j)$ such that $\sum\limits_{j=1}^{m} p(j)\, \frac{\sigma^2(j)}{\mathbf{e}(j)} < \frac{\sigma^2}{\mathbf{e}}$ and other numerical values such that $\sum\limits_{j=1}^{m} p(j)\, \frac{\sigma^2(j)}{\mathbf{e}(j)} > \frac{\sigma^2}{\mathbf{e}}$.

While the three identities give useful information about a homogeneous stationary population, none of them necessarily holds for a heterogeneous stationary population. Two of the three identities may fail to hold in either direction, but the fundamental identity $P = B\mathbf{e}$ for a homogeneous stationary population can be replaced by the inequality $P \leq B\mathbf{e}$ for a heterogeneous stationary population, with equality if and only if all countries have the same life expectancy.

3.2 Heterogeneous stationary populations: continuous case

A frequency histogram of the values of life expectancy **e** where the unit of observation is a country summarizes international variation in **e**: the horizontal axis is **e** (discretized), and the height of the vertical bars represents the number of countries with life expectancy in the corresponding interval of **e**. The total area under the frequency histogram (i.e., the sum of the heights of all the bars) is m, the number of countries. Now suppose the number m of units of observation becomes large while the width of the bins used to calculate the frequency histogram becomes small in such a way that the shape of the frequency histogram, divided by its total area m, approaches a limiting probability density function $f(\mathbf{e})$ on the non-negative half line $[0, \infty)$. The probability density function $f(\mathbf{e})$ satisfies $\int_0^\infty f(\mathbf{e})\, d\mathbf{e} = 1$. For example, instead of using a country as the unit of measurement, one could measure the life expectancy of each province (primary administrative subunit) or of each county (secondary administrative subunit) worldwide. As m increases, it will be assumed that the number of units of observation with life expectancy in the interval from e_1 to $e_2 > e_1$ is increasingly well approximated by $m \int_{e_1}^{e_2} f(\mathbf{e})\, d\mathbf{e}$ while the total size of the heterogeneous stationary population remains constant at P. It will be assumed that the probability density function $f(\mathbf{e})$ and all other functions occurring here are properly integrable (as defined by Pólya and Szegö 1972, p. 46). For each life expectancy **e**, the average number of annual births per country with life expectancy **e** will be represented by $B(\mathbf{e}) \geq 0$. It will be assumed further that the mean and variance of $B(\mathbf{e})$ with respect to $f(\mathbf{e})$ are finite. All of these assumptions are automatically satisfied in the discrete case and are reasonable assumptions for a continuous approximation to a demographically realistic situation.

Under the additional assumption that each unit of observation is a homogeneous stationary population, the population size $P(\mathbf{e})$ of each unit with life expectancy **e** is $P(\mathbf{e}) = B(\mathbf{e})\, \mathbf{e}$. Then the stationary global population size is the weighted sum over all **e** of the units' population sizes,

$$P = m \int_0^\infty B(\mathbf{e})\, \mathbf{e}\, f(\mathbf{e})\, d\mathbf{e}.$$

The global number of births per year is

$$B = m \int_0^\infty B(\mathbf{e})\, f(\mathbf{e})\, d\mathbf{e}.$$

As in the discrete case, the global life expectancy is the population-weighted mean of the life expectancies of each country. To avoid confusion with the dummy variable **e** for life expectancy, the symbol $\overline{e}$ will be used to denote the global life expectancy:

$$\overline{e} = \frac{m}{P} \int_0^\infty \mathbf{e}\, P(\mathbf{e})\, f(\mathbf{e})\, d\mathbf{e} = \frac{m}{P} \int_0^\infty \mathbf{e}^2\, B(\mathbf{e})\, f(\mathbf{e})\, d\mathbf{e}.$$

By Schwarz's inequality (Pólya and Szegö 1972, p. 68), which is the continuous version of Cauchy's inequality, $P \leq B\overline{e}$. In detail, Schwarz's inequality states that if $u(x)$ and $v(x)$ are two functions that are properly integrable in the interval $[a, b]$, then

$$(\int_a^b u(x)\, v(x)\, dx)^2 \leq \int_a^b [u(x)]^2\, dx \int_a^b [v(x)]^2\, dx.$$

Let $u(\mathbf{e}) = \sqrt{B(\mathbf{e})\, f(\mathbf{e})}$ and $v(\mathbf{e}) = \mathbf{e}\, \sqrt{B(\mathbf{e})\, f(\mathbf{e})}$. Then Schwarz's inequality becomes

$$(\int_0^\infty B(\mathbf{e})\, \mathbf{e}\, f(\mathbf{e})\, d\mathbf{e})^2 \leq \int_0^\infty B(\mathbf{e})\, f(\mathbf{e})\, d\mathbf{e} \int_0^\infty \mathbf{e}^2\, B(\mathbf{e})\, f(\mathbf{e})\, d\mathbf{e}$$

or $(\frac{P}{m})^2 \leq \frac{B}{m} \cdot \frac{P\,\overline{e}}{m}$ which simplifies to $P \leq B\,\overline{e}$. QED

As in the discrete case, the inequality does not depend on the number m of countries or other units of observation.

3.3 Analytically soluble example

To confirm and quantify the inequality $P \leq B\,\overline{e}$ in an analytically soluble, hypothetical example of the continuous case of a heterogeneous stationary population, suppose that the probability density function (pdf) of the country life expectancy **e** is the gamma distribution with parameters ρ (a positive real number) and r (a positive integer),

$$f(y) = \frac{\rho^r\, e^{-\rho y}\, y^{r-1}}{(r-1)!},$$

for all $y \geq 0$. Initially assume $r > 1$. The case $r = 1$ will be considered separately. If Y is a random variable with gamma pdf $f(y)$, then for $n = 0, 1, 2, \ldots$,

$$E(Y^n) = \int_0^\infty y^n\, f(y)\, dy = \frac{(r+n-1)^{(n)}}{\rho^n},$$

where $(a)^{(n)} = a\,(a-1)\cdots(a-n+1)$ is the falling factorial, e.g., $(a)^{(0)} = 1$, $(a)^{(1)} = a$, $(a)^{(2)} = a\,(a-1)$. Suppose $B(\mathbf{e}) = \frac{1}{\mathbf{e}}$ (here **e** represents life expectancy, not the base of natural logarithms) so that the births per year in each country are (in some arbitrary units) inversely proportional to the life expectancy of that country. Then $B(\mathbf{e})\,\mathbf{e} = 1$ so $P = m$ and $\overline{e} = \frac{m}{P}\,E(Y) = \frac{r}{\rho}$. Also

$$\begin{aligned} B &= m \int_0^\infty \frac{1}{y}\, f(y)\, dy \\ &= m \int_0^\infty \frac{\rho^r\, e^{-\rho y}\, y^{r-2}}{(r-1)!}\, dy \\ &= m \frac{\rho}{r-1} \int_0^\infty \frac{\rho^{r-1}\, e^{-\rho y}\, y^{r-2}}{(r-2)!}\, dy \\ &= m \frac{\rho}{r-1} \end{aligned}$$

since the last integrand on the right is the gamma pdf with parameters ρ and $r-1$. Therefore $B\,\overline{e} = m\frac{r}{r-1} > m$. The inequality is expected from the general inequality $P \leq B\,\overline{e}$.

When $r = 1$, the gamma distribution becomes the exponential distribution so that $P = m$ and $\overline{e} = \frac{m}{P}E(Y) = \frac{1}{\rho}$, but

$$B = m \int_0^\infty \frac{1}{y} f(y)\, dy = m \int_0^\infty \rho\, e^{-\rho y}\, y^{-1}\, dy$$

which diverges to infinity, so the inequality $P < B\,\overline{e}$ holds.

At the other extreme, if $r \to \infty$, then $\overline{e} \to \infty$ and $B \to 0$ while $B\,\overline{e} \downarrow m$ so that the inequality approaches equality. In the gamma pdf with parameters ρ and r, the coefficient of variation, i.e., the standard deviation divided by the mean, is $r^{-\frac{1}{2}}$, which approaches 0 as $r \to \infty$. Thus the demographic interpretation of the case when $r \to \infty$ is that global life expectancy grows longer on average and, relative to this increasing average, also grows less variable from country to country.

4. Constant global population with migration between countries

In a heterogeneous stationary population, each country has constant size and no individuals migrate from one country to another, by assumption. We now consider a discrete-time model of a constant global population with migration among countries. In this model, countries may change size in time. Later we will give a hypothetical numerical example.

Let the letters h, i, j index countries $1, \ldots, m$. Let $P(i,t) \geq 0$ be the population size of country i at t. Let $P(t)$ be the global population size (sum of country population sizes) at t. For the country of origin i and the country of destination j, where $i \neq j$, let $M(i,j,t) \geq 0$ be the number of migrants from country i to country j between t and $t+1$, defined as the number of people who were in country i at t and in country j (different from i) at $t+1$. (This *de facto* definition does not coincide with the legal definition of a migrant used by many countries.) Let $M(i,i,t) = 0$ for all i. The $m \times m$ matrix $M(t)$ with elements $M(i,j,t)$ is called the migration matrix.

The number of emigrants $E(i,t)$ from country i between t and $t+1$ is the number of people who were in country i at t and were elsewhere at $t+1$, namely, the sum of row i of the migration matrix: $E(i,t) = \sum_j M(i,j,t) \geq 0$. The number of immigrants $I(i,t)$ to country i between t and $t+1$ is the number of people who were elsewhere than country i at t and were in country i at $t+1$, namely, the sum of column i of the migration matrix: $I(i,t) = \sum_h M(h,i,t) \geq 0$.

The number of net migrants to country i between t and $t+1$ is

$$N(i,t) = I(i,t) - E(i,t) = \sum_h M(h,i,t) - \sum_j M(i,j,t). \tag{1}$$

Net migration is positive if immigration exceeds emigration, zero if immigration equals emigration and negative if emigration exceeds immigration. Then $N(t)$, the sum of net migrants over all countries, is guaranteed to be zero (with no further assumptions) because

$$N(t) = \sum_i N(i,t) = \sum_i I(i,t) - \sum_i E(i,t) = \sum_{h,i} M(h,i,t) - \sum_{i,j} M(i,j,t) = 0.$$

Each country's population may change as a result of migration and vital events (i.e., births and deaths): for $i = 1,\ldots,m$,

$$P(i,t+1) = P(i,t) + N(i,t) + B(i,t) - D(i,t), \tag{2}$$

where $B(i,t)$ is the number of births and $D(i,t)$ is the number of deaths between t and $t+1$ in country i. To assure that $P(i,t+1) \geq 0$, it will be required that

$$P(i,t) \geq D(i,t) - B(i,t) - N(i,t).$$

If each term in equation (2) is summed over all countries $i = 1,\ldots,m$, and if $B(t)$ is the sum of births and $D(t)$ is the sum of deaths in all countries between t and $t+1$, then because $N(t) = 0$, we have

$$P(t+1) = P(t) + B(t) - D(t).$$

Therefore $P(t+1) = P(t)$ if and only if $B(t) = D(t)$. We now make the additional important assumption that $B(t) = D(t)$. Migration, births and deaths may vary in time while keeping a constant total population size.

The assumption that $B(t) = D(t)$ is a significant constraint, as the following hypothetical scenario illustrates. If all migrants were people past the ages of childbearing and if they all moved from a high-mortality country to a low-mortality country and immediately acquired that lower rate of mortality, as will be further assumed below, then $D(t{+}1) < D(t)$ and global births would also have to decrease, even though the migration involved only people past childbearing. This problem arises only if the model has age structure.

If each country were modeled as having its own age structure and its own life table, then it would be necessary to specify the age structure of migrants (e.g., Rogers 1968, 1995; Rogers and Castro 1981; Rogers and Henning 1999; Rogers, Castro and Lea 2004). To avoid tracking ages, we now make the further simplifying assumption that, for each country j, for each interval from t to $t+1$, and for every age x, country j has a force of mortality $\mu(j,t)$ (also called the per capita instantaneous death rate or hazard of mortality) that applies to every individual while that individual lives in country j during the interval from t to $t+1$. The force of mortality applies equally to natives and to immigrants, assuming that, upon migration, migrants change instantaneously and completely from the life table of the country of origin to the life table of the country of destination. The period life expectancy $\mathbf{e}(j,t)$ of country j is then $\frac{1}{\mu(j,t)}$ during the time interval from t to $t+1$.

A simple numerical example illustrates some of the main features and limitations of this model (Table 1). In this example, all countries are directly linked by migration in both directions and no country has zero natural increase (births minus deaths), yet the total population size is constant from $t = 0$ to $t = 1$. Because no country has unchanging size, this global population of constant size with migration is not a heterogeneous stationary population.

Table 1: Numerical example of a constant global population with international migration

time	**t=0**								**t=1**
country i	**P(i,0)**	**M(i,1,0)**	**M(i,2,0)**	**M(i,3,0)**	**E(i,0)**	**N(i,0)**	**B(i,0)**	**D(i,0)**	**P(i,1)**
country 1	20	0	2	3	5	5	1	2	24
country 2	30	4	0	5	9	0	6	3	33
country 3	50	6	7	0	13	-5	5	7	43
total	100	10	9	8	27	0	12	12	100
		=I(1,0)	=I(2,0)	=I(3,0)					

Note: t=time (years), $P(i,t)$= population of country i at start of year t, $M(i,j,t)$=number of individuals in country i at t who were in country j at $t+1$, $E(i,t)$=number of emigrants from country i between t and $t+1$, $I(i,t)$=number of immigrants to country i between t and $t+1$, $N(i,t)$=net migration of country i from t to $t+1$, $B(i,t)$=births in country i in year t, $D(i,t)$=deaths in country i in year t.

Table 2: **Numerical example of births, life expectancy and population size in a stationary population with migration**

time country i	t=0 P(i,0)	t=1 P(i,1)	t=1/2 P(i,1/2)	D(i,0)	t=1/2 e(i,0)	B(i,0)	t=1/2 b(i,1/2)	t=1/2 d(i,1/2)
country 1	20	24	22	2	$\frac{22}{2}$= 11	1	$\frac{1}{22}$	$\frac{2}{22}$
country 2	30	33	31.5	3	$\frac{31.5}{3}$= 10.5	6	$\frac{6}{31.5}$	$\frac{3}{31.5}$
country 3	50	43	46.5	7	$\frac{46.5}{7}$= 6.64...	5	$\frac{5}{46.5}$	$\frac{7}{46.5}$
total	100	100	100	12	$\frac{100}{12}$= 8.33...	12	$\frac{12}{100}$	$\frac{12}{100}$

Symbols are defined in Table 1 note.

Table 2 continues the numerical example from Table 1. Because $\mathbf{e}(j,t) = \frac{1}{\mu(j,t)}$ and $\mu(j,t) = \frac{D(j,t)}{\text{person-years lived}}$, the life expectancy $\mathbf{e}(j,t)$ of country j is the person-years lived between t and $t+1$ divided by the number $D(j,t)$ of deaths between t and $t+1$. The person-years lived are determined by the exact timing of events of migration, birth and death between t and $t+1$. A reasonable linear approximation to person-years lived is the average of the population in the country at t and at $t+1$, $\frac{P(j,t)+P(j,t+1)}{2}$, denoted $P(j,t+\frac{1}{2})$ for brevity. Hence $\mathbf{e}(j,t) = \frac{P(j,t+\frac{1}{2})}{D(j,t)}$ is the (approximate) life expectancy of country j between t and $t+1$. Then the life expectancy $\mathbf{e}(t)$ of the global population between t and $t+1$ is the weighted average of the life expectancy in each country. The weights are person-years lived in each country divided by the sum of the person-years lived over all countries. The period life expectancy $\mathbf{e}(t) > 0$ between t and $t+1$ reflects births, deaths and migration because all three factors affect the person-years lived between t and $t+1$. In this example, the global life expectancy is

$$\frac{22 \cdot 11 + 31.5 \cdot 10.5 + 46.5 \cdot \frac{46.5}{7}}{100} = \frac{\frac{22^2}{2} + \frac{31.5^2}{3} + \frac{46.5^2}{7}}{100} = 8.816$$
$$> \frac{P(t)}{B(t)} = \frac{100}{12} = 8.333,$$

which is the estimate of the global life expectancy that would be obtained assuming a *homogeneous* stationary population and ignoring differences among countries.

In general, in a constant global population with migration between countries, for each year $[t, t+1]$ considered separately, $B(t)\,\mathbf{e}(t) \geq P(t+\frac{1}{2}) = P(t) = P(t+1)$, and

$B(t)\,\mathbf{e}(t) = P(t)$ holds if and only if all countries have the same life expectancy. The proof (in Appendix 2) parallels the proof in a heterogeneous stationary population, except that here the Cauchy-Schwarz inequality is applied to the person-years lived instead of to the population at t or $t+1$. These results depend on the assumption that $B(t) = D(t)$ at each time.

A similar inequality holds when averaging over time as well as over countries. The time-average birth rate and the time-average life expectancy of the global population from 0 to T are

$$\overline{B} = \frac{1}{T}\sum_{0}^{T-1} B(t), \quad \overline{e} = \frac{1}{T}\sum_{0}^{T-1} \mathbf{e}(t).$$

Then $P \le \overline{B}\,\overline{e}$, and $P < \overline{B}\,\overline{e}$ unless all countries have equal life expectancies at all times. These results depend on the assumption that $B(t) = D(t)$ at each time. The proof is in Appendix 3.

4.1 Alternative projections of a constant global population with international migration

Suppose we have complete data for each country in the world on each country's population size at times t and $t+1$ and its immigrants and emigrants and births and deaths from t to $t+1$, and suppose global population size $P(t) = P(t+1)$ is constant. How shall we project future population sizes globally and by country? For a recent review of some projection techniques in use, see Howe and Jackson (2005).

The remainder of this subsection offers three linear ways of projecting the future of a constant global population with international migration and shows by numerical example that all three approaches have problems. The three approaches assume that

(i) the change in the population of each country is constant over time, or
(ii) demographic rates relative to the population of the sending country are constant (a sender-controlled linear representation), or
(iii) demographic rates relative to the population of the receiving country are constant (a receiver-controlled linear representation).

All three methods can give exactly the same population sizes of all countries at $t = 0$ and $t = 1$. The three methods may give very different projections of the future from $t = 2$ onward. Only the projection based on constant linear change keeps the global population size constant after $t = 1$. However, in our numerical example, the first method leads to a projection of a negative population size for country 3 by year 8, while the third method leads to a projection of a negative population size for country 3 by year 5. The limitations of each of the three methods of projection illustrated by this numerical example indicate that further work is needed to develop a useful projection method for a constant global population with international migration.

We now present the analysis and numerical example in detail. Equations (1) and (2) and the remainder of this section are independent of any assumption that the life table is exponential or otherwise. We may rewrite the definition of net migration in equation (1) as

$$N(i,t) = \sum_h M(h,i,t) - E(i,t) = \sum_{h \neq i} P(h,t)\, \frac{M(h,i,t)}{P(h,t)} - P(i,t)\, \frac{E(i,t)}{P(i,t)}.$$

Let $b(i,t) = \frac{B(i,t)}{P(i,t)}$ be the crude birth rate per capita and let $d(i,t) = \frac{D(i,t)}{P(i,t)}$ be the crude death rate per capita of country i from t to $t+1$. (The denominators are intentionally $P(i,t)$ and not $P(i, t+\frac{1}{2})$.) Then

$$P(i,t+1) = P(i,t)\,[1 - \frac{E(i,t)}{P(i,t)} + b(i,t) - d(i,t)] + \sum_{h \neq i} P(h,t) \frac{M(h,i,t)}{P(h,t)}.$$

Define, for $i = 1, \ldots, m$, $a(i,i,t) = 1 - \frac{E(i,t)}{P(i,t)} + b(i,t) - d(i,t)$, and for $h \neq i$, $h = 1, \ldots, m$, $a(h,i,t) = \frac{M(h,i,t)}{P(h,t)}$. The off-diagonal elements of the matrix a are thus nonnegative. (In the numerical example, the matrix a is computed in Table 3.) Then equation (2) is equivalent to

$$P(i,t+1) = \sum_{h=1}^{m} P(h,t)\, a(h,i,t), \qquad i = 1, \ldots, m \tag{3}$$

which is a donor-controlled linear representation of a constant global population with international migration. Then the global population size is the same at t and $t+1$:

$$\begin{aligned}
\sum_{i=1}^{m} P(i,t+1) &= \sum_{h=1}^{m} P(h,t) \sum_{i=1}^{m} a(h,i,t) \\
&= \sum_{h=1}^{m} P(h,t)\,[a(h,h,t) + \sum_{i \neq h} a(h,i,t)] \\
&= \sum_{h=1}^{m} [P(h,t) - E(h,t) + B(h,t) - D(h,t) + \sum_{i \neq h} M(h,i,t)] \\
&= \sum_{h=1}^{m} [P(h,t) - E(h,t) + B(h,t) - D(h,t) + E(h,t)]
\end{aligned}$$

$$= \sum_{h=1}^{m} [P(h,t) + B(h,t) - D(h,t)]$$
$$= \sum_{h=1}^{m} P(h,t).$$

Dividing both sides by the global population P and letting $p(i,t) = \frac{P(i,t)}{P}$ gives

$$p(i,t+1) = \sum_{h=1}^{m} p(h,t)\, a(h,i,t), \qquad i = 1,\ldots,m.$$

Unlike a time-inhomogeneous m-state Markov chain (e.g., Iosifescu 1980), here neither the row sums nor the column sums of the transition matrix a are required to be 1, and the elements of a may fail to be nonnegative, and the eigenvalue of a largest in modulus may fail to be 1.

Table 3: Numerical example of sender-controlled linear representation of a stationary population with migration

time country i	t=0 p(i,0)	a(i,1,0)	a(i,2,0)	a(i,3,0)	E(i,0)	b(i,0)	d(i,0)	t=1 p(i,1)
country 1	0.2	$1-\frac{5}{20}+\frac{1}{20}-\frac{2}{20}=\frac{14}{20}$	$\frac{2}{20}$	$\frac{3}{20}$	5	$\frac{1}{20}$	$\frac{2}{20}$	0.24
country 2	0.3	$\frac{4}{30}$	$1-\frac{9}{30}+\frac{6}{30}-\frac{3}{30}=\frac{24}{30}$	$\frac{5}{30}$	9	$\frac{6}{30}$	$\frac{3}{30}$	0.33
country 3	0.5	$\frac{6}{50}$	$\frac{7}{50}$	$1-\frac{13}{50}+\frac{5}{50}-\frac{7}{50}=\frac{35}{50}$	13	$\frac{5}{50}$	$\frac{7}{50}$	0.43
total	1.0	0.9533...	1.04	1.0166...	27	12	12	1

Note: $p(i,t)$= fraction of global population in country i at t, $a(i,j,t)$=sender-controlled migration matrix at t (see text for definition)

Whenever the matrix a is nonnegative and some power of it has strictly positive elements (as in the numerical example in Table 3), the Perron-Frobenius theorem (e.g., Gantmacher 1960) assures that the proportion of global population found in each country will converge to a constant fraction and that the size of the global population and of each country will eventually change by a constant factor each year. That factor of annual change is given by the eigenvalue of the a matrix with maximal absolute value (the so-called spectral radius of the a matrix), which is guaranteed to be a positive real number. The long-run proportions of global population in each country are given by the elements of the eigenvector corresponding to the spectral radius normalized to sum to 1. Although the matrix a is constructed to assure that global population remains constant between $t = 0$ and $t = 1$, there is no guarantee that when the matrix a is assumed constant the global population size will remain constant beyond $t = 1$ in the short term or long term. Unless the spectral radius of the matrix a happens to be 1, the global population and each country's population will not be stationary in the long run.

In the numerical example in Table 3, the three eigenvalues of a are approximately 1.0086, 0.6298 and 0.5616, hence the spectral radius of the a matrix is 1.0086. Consequently after projecting each country's population a long time assuming a constant a matrix, the population of each country and of the world will increase by about 0.86% per year. From the last two lines of Table 5, in the projection assuming a constant a matrix, the annual increase of country 1's population is $100 \cdot (\frac{31.12}{30.83} - 1) = 0.93\%$ (using the original projection before rounding), of country 2's population 0.90%, and of country 3's population 0.75%. The rates of increase for all three countries from $t = 99$ to $t = 100$ are all 0.86%, as expected from the spectral radius of a.

Rewriting (2) in the form (3) represents the number of migrants as proportional to the size of the population of the sending country. An obvious alternative is to represent the number of migrants as proportional to the size of the population of the receiving country. Rewrite (1) as

$$N(i,t) = I(i,t) - \sum_j M(i,j,t) = P(i,t)\frac{I(i,t)}{P(i,t)} - \sum_{j \neq i} P(j,t)\frac{M(i,j,t)}{P(j,t)}$$

so that

$$P(i,t+1) = P(i,t)\,[1 + \frac{I(i,t)}{P(i,t)} + b(i,t) - d(i,t)] - \sum_{j \neq i} P(j,t)\frac{M(i,j,t)}{P(j,t)}.$$

Define, for $i = 1, \ldots, m, \quad z(i,i,t) = 1 + \frac{I(i,t)}{P(i,t)} + b(i,t) - d(i,t)$, and for $j \neq i$, $j = 1, \ldots, m, \quad z(i,j,t) = -\frac{M(i,j,t)}{P(j,t)}$, so that all off-diagonal elements of the matrix z will be 0 or negative. (In the numerical example, the matrix z is computed in Table 4.)

Then (2) is equivalent to

$$P(i,t+1) = \sum_{j=1}^{m} z(i,j,t)\,P(j,t), \qquad i = 1,\ldots,m, \tag{4}$$

which is a receiver-controlled linear representation of a constant global population with international migration. Dividing both sides by the total population P and letting $p(i,t) = \frac{P(i,t)}{P}$ gives

$$p(i,t+1) = \sum_{j=1}^{m} z(i,j,t)\,p(j,t), \qquad i = 1,\ldots,m.$$

In this example, no row sum, column sum or eigenvalue equals 1.

Table 4: Numerical example of receiver-controlled linear representation of a stationary population with migration

time country i	t=0 p(i,0)	z(i,1,0)	z(i,2,0)	z(i,3,0)	row sum	b(i,0)	d(i,0)	t=1 p(i,1)
country 1	0.2	$1+\frac{10}{20}+\frac{1}{20}-\frac{2}{20}=\frac{29}{20}$	$-\frac{2}{30}$	$-\frac{3}{50}$	1.3233	$\frac{1}{20}$	$\frac{2}{20}$	0.24
country 2	0.3	$-\frac{4}{20}$	$1+\frac{9}{30}+\frac{6}{30}-\frac{3}{30}=\frac{42}{30}$	$-\frac{5}{50}$	1.1000	$\frac{6}{30}$	$\frac{3}{30}$	0.33
country 3	0.5	$-\frac{6}{20}$	$-\frac{7}{30}$	$1+\frac{8}{50}+\frac{5}{50}-\frac{7}{50}=\frac{56}{50}$	0.5867	$\frac{5}{50}$	$\frac{7}{50}$	0.43
total	1.0	10 =I(1,0)	9 =I(2,0)	8 =I(3,0)	3.0100	12	12	1

Note: $p(i,t)$= fraction of global population in country i at t, $z(i,j,t)$=receiver-controlled migration matrix at t (see text for definition)

None of these three linear approaches to projecting a constant global population with international migration is satisfactory. If one assumes the change $N(i,t)$+$B(i,t)$-$D(i,t)$

in the population of country i is constant for all t, equation (2) gives a projection technique. If one assumes the matrix a is constant for all t, equation (3) gives a sender-controlled projection technique. If one assumes the matrix z is constant for all t, equation (4) gives a receiver-controlled projection technique. When the matrices are related as above, all three methods give exactly the same population sizes of all countries at $t = 0$ and $t = 1$. In this numerical example, the three methods give very different projections of the future $t = 2, \ldots$ (Table 5).

Table 5: Numerical example of projections of a stationary population with migration by 3 methods: constant linear change, constant a matrix, and constant z matrix

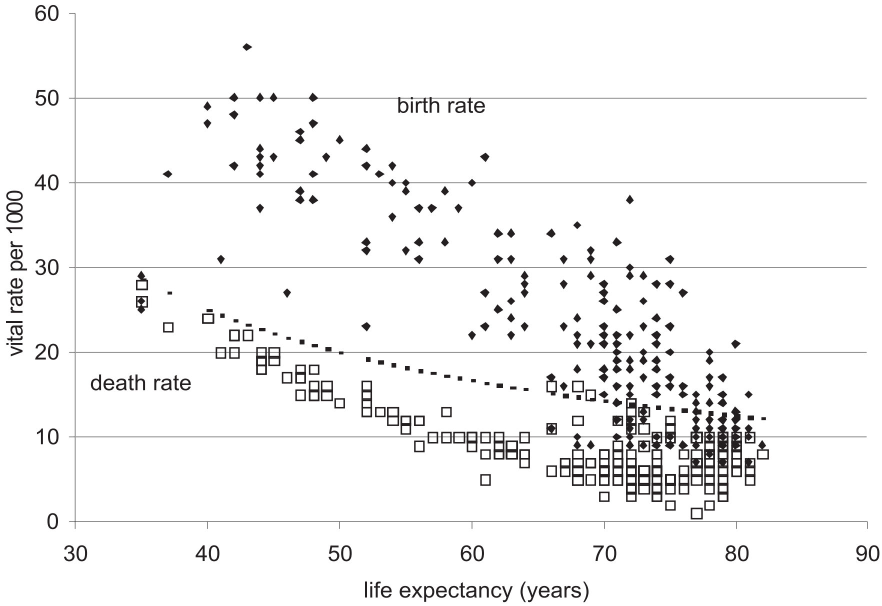
	constant linear change			constant *a* matrix			constant *z* matrix		
year t	P(1,t)	P(2,t)	P(3,t)	P(1,t)	P(2,t)	P(3,t)	P(1,t)	P(2,t)	P(3,t)
0	20	30	50	20	30	50	20	30	50
1	24	33	43	24	33	43	24	33	43
2	28	36	36	26.36	34.82	39.20	30.02	37.10	33.26
3	32	39	29	27.80	35.98	37.20	39.06	42.61	19.59
4	36	42	22	28.72	36.77	36.21	52.62	49.88	0.28
5	40	45	15	29.35	37.36	35.78	72.96	59.28	−27.11
6	44	48	8	29.82	37.83	35.68	103.46	71.12	−66.09
7	48	51	1	30.20	38.24	35.75	149.25	85.48	−121.65
8	52	54	−6	30.53	38.62	35.93	218.01	101.99	−200.97
9	56	57	−13	30.83	38.98	36.17	321.37	119.28	−314.29
10	60	60	−20	31.12	39.33	36.44	476.89	134.15	−476.24

Note: $P(i, t)$= population of country i at start of year t. Matrices a and z are defined in text.

5. Demographic data

Today's global population remains far from any of the models considered here. To examine the present empirical relevance of the heterogeneous stationary population model, Figure 1 shows the crude birth rate, the crude death rate, and the reciprocal of the life expectancy (which would be equal to the crude birth rate and to the crude death rate in a homogeneous stationary population), as a function of life expectancy for both sexes combined, for 222 countries or territories using data available in 2006. For most countries, the birth rate exceeds the death rate, so (ignoring international migration) the population is growing and has a younger age structure than it would if it were stationary with the same life table. The crude death rate falls below the reciprocal of the life expectancy. Most countries are not approximately stationary.

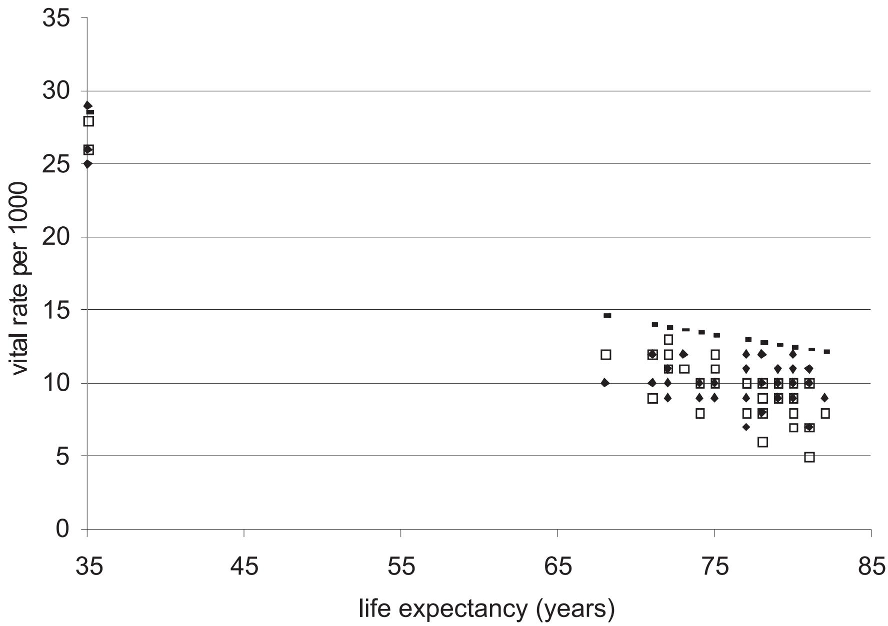
Figure 1: Crude birth rate (filled diamonds) and crude death rate (open squares) (annual, per 1000) as a function of life expectancy (years) for both sexes combined in 222 countries or territories



Note: Apparently filled squares represent the coincidence of a crude birth rate and crude death rate. Countries with missing data are excluded. Dashed line shows 1000/life expectancy, the theoretical prediction for both birth rate and death rate (per 1000) in a homogeneous stationary population.
Source of data: Reference Bureau DataFinder, http://www.prb.org/datafind/datafinder7.htm, accessed 2006-08-11

If a country is well described by the homogeneous stationary population model, its crude birth rate equals its crude death rate, and both equal the reciprocal of the life expectancy. Figure 2 examines 37 countries or territories, selected from the data in Figure 1, with rates of natural increase (crude birth rate minus crude death rate) that are small in magnitude, specifically, in the range from -3 to +3 per thousand. Does the reciprocal of the life expectancy predict well the crude birth rate and the crude death rate in countries where these rates are nearly equal, as the homogeneous stationary population model predicts?

Figure 2: Crude birth rate (filled diamonds) and crude death rate (open squares) (annual, per 1000) as a function of life expectancy (years) for both sexes combined in 37 countries or territories with birth rate minus death rate in the range from -3 to +3 per thousand



Note: Apparently filled squares represent the coincidence of a crude birth rate and crude death rate. Countries with missing data are excluded. Dashed line shows 1000/life expectancy, the theoretical prediction for both birth rate and death rate (per 1000) in a homogeneous stationary population.
Source of data: Reference Bureau DataFinder, http://www.prb.org/datafind/datafinder7.htm, accessed 2006-08-11

These countries fall into two groups: unfortunate countries with very low life expectancy, where the reciprocal of life expectancy does predict vital rates as expected in a homogeneous stationary population; and countries with life expectancy above the average estimated for the world as a whole, where crude birth rates and crude death rates fall below the reciprocal of life expectancy. Ryder (1975, p. 4) called the first group countries with "inefficient" replacement and the second group countries with "efficient" replacement. It remains to be seen whether countries with efficient replacement will be better described by the stationary population model in the future.

6. Acknowledgements

I acknowledge with thanks the helpful comments of anonymous referees, the support of U.S. National Science Foundation grant DMS-0443803, the assistance of Priscilla K. Rogerson, and the hospitality of Mr. and Mrs. William T. Golden and family during this work.

References

Alho, J. and Spencer, B. (2005). *Statistical Demography and Forecasting.* Springer, New York.

Bartholomew, D. (1963). A multi-stage renewal process. *J. R. Statist. Soc.*, B 25:150–168.

Coale, A. (1972). Alternative paths to a stationary population. In Westoff, C. and Parke, R., editors, *Demographic and Social Aspects of Population Growth. U.S. Commission on Population Growth and the American Future, research reports vol. 1.*, pages 589–603. Government Printing Office, Washington, DC.

Espenshade, T. (1978). Zero population growth and the economies of developed nations. *Popul. Devel. Rev.*, 4:645–680.

Feeney, G. (1971). Comment on a proposition of H. Le Bras. *Theor. Popul. Biol.*, 2:122–123.

Gantmacher, F. (1960). *Theory of Matrices.* Chelsea, New York.

Howe, N. and Jackson, R. (2005). Projecting immigration: a survey of the current state of practice and theory. A Report of the CSIS Global Aging Initiative, with contributions by Rebecca Strauss and Keisuke Nakashima. Center for Strategic and International Studies, Washington, DC. http://ideas.repec.org/p/crr/crrwps/2004-32.html, last accessed 2008-02-20.

Iosifescu, M. (1980). *Finite Markov Processes and Their Applications.* John Wiley, New York, Editura Tehnica, Bucharest.

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

Keyfitz, N. (1973). Individual mobility in a stationary population. *Popul. Stud.*, 27:335–352.

Kim, Y. and Schoen, R. (1996). Populations with sinusoidal birth trajectories. *Theor. Popul. Biol.*, 50:105–123.

Land, K. and Rogers, A., editors (1982). *Multidimensional Mathematical Demography.* Academic, New York. pages 477-503.

Le Bras, H. (1971). Équilibre et croissance de populations soumises à des migrations. *Theor. Popul. Biol.*, 2:100–121.

Marshall, A. and Olkin, I. (1979). *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York.

Pollard, J. (1973). *Mathematical Models for the Growth of Human Populations.* Cambridge University Press, Cambridge.

Pólya, G. and Szegö, G. (1972). *Problems and Theorems in Analysis, Vol. I: Series, Integral Calculus, Theory of Functions.* Springer-Verlag, New York, Heidelberg, Berlin.

Preston, S., Heuveline, P., and Guillot, M. (2001). *Demography: Measuring and Modeling Population Processes.* Blackwell Publishing, Oxford.

Rogers, A. (1968). *Matrix analysis of interregional population growth and distribution.* University of California Press, Berkeley.

Rogers, A. (1990). Requiem for the net migrant. *Geogr. Anal.*, 22:283–300.

Rogers, A. (1995). *Multiregional demography: principles, methods and extensions.* John Wiley, New York.

Rogers, A. and Castro, L. (1981). Model migration schedules. Research Report-81-30. International Institute for Applied Systems Analysis, Laxenburg.

Rogers, A., Castro, L., and Lea, M. (2004). Model migration schedules: three alternative linear parameter estimation methods. Institute of Behavioral Science, Research Program on Population Processes, University of Colorado at Boulder, Working Paper POP2004-0004.

Rogers, A. and Henning, S. (1999). The internal migration patterns of the foreign-born and native-born populations in the United States: 1975-80 and 1985-90. *Int. Migr. Rev.*, 33:403–429.

Romaniuc, A. (2005). Stationary population as theoretical concept and as policy vision for Canada. Presented at workshop on "Population Changes and Public Policy", London, Ontario, Canada, February 3-4, 2005. sociology.uwo.ca/popchange/Romaniuc,%20 Stationary%20population.pdf.

Ryder, N. (1975). Notes on stationary populations. *Popul. Index*, 41:3–28.

Schoen, R. (1988). *Modeling Multigroup Populations.* Plenum Press, New York.

Schoen, R. (2002). On the impact of spatial momentum. *Demographic Res.*, 6:49–66. www.demographic-research.org/Volumes/Vol6/3/.

Schoen, R. and Kim, Y. (1993). Two-state spatial dynamics in the absence of age. *Theor. Popul. Biol.*, 44:67–79.

Schoen, R. and Kim, Y. (1998). Momentum under a gradual approach to zero growth. *Popul. Stud.*, 52:295–299.

Seal, H. (1945). The mathematics of a population composed of k stationary strata each recruited from the stratum below and supported at the lowest level by a uniform annual number of entrants. *Biometrika*, 33:226–230.

United Nations Population Division (2005). World Population Prospects: The 2004 Revision, Highlights, ESA/P/WP.193, 24 February 2005. Department of Economic and Social Affairs. United Nations, New York.

Vajda, S. (1947). The stratified semi-stationary population. *Biometrika*, 34:243–254.

Appendix 1

The global population's life table is the population-weighted mean of the country life tables, i.e., $l(x) = \sum_{j=1}^{m} p(j)\, l(j, x)$.
Proof.

Let X be a continuous, real nonnegative-valued random variable that represents an individual's length of life or exact age at death. Then the global probability of surviving from birth to age x or longer is $l(x) = P[X \geq x]$. The conditional probability of surviving from birth to age x or longer, given residence in country j, is $l(j, x) = P[X \geq x \,|\, j]$. Everybody in the global population lives in some country (by assumption). Each country's life table applies to all individuals in that country regardless of their *current* age, by definition. The decomposition formula for conditional probabilities is $P[X \geq x] = \sum_{j=1}^{m} P[X \geq x \,|\, j]\, p(j)$. This is the same as $l(x) = \sum_{j=1}^{m} p(j)\, l(j, x)$ in other notation.

Appendix 2

Cauchy's inequality: Let $u(1), \ldots, u(m),\ v(1), \ldots, v(m)$ be arbitrary real numbers. Then

$$[u(1)\, v(1) + \ldots + u(m)\, v(m)]^2 \leq [u(1)^2 + \ldots + u(m)^2]\, [v(1)^2 + \ldots + v(m)^2].$$

Equality holds if and only if for some real constants λ and μ with $\lambda^2 + \mu^2 > 0$ we have

$$\lambda\, u(j) + \mu\, v(j) = 0 \quad \text{for } j = 1, \ldots, m.$$

Jensen's inequality (e.g. Marshall and Olkin 1979, p. 454): For any real interval (a, b) and any convex function f defined on (a, b) and any real numbers $u(1), \ldots, u(m)$ each in (a, b) and any nonnegative real numbers $v(1), \ldots, v(m)$ such that $\sum_{j=1}^{m} v(j) = 1$, one has

$$f(\sum u(j)\, v(j)) \leq \sum v(j)\, f(u(j)).$$

Jensen established an integral analog using Lebesgue measure.

In a constant global population of size $P(t)$ with migration between countries, if $B(t) = D(t)$ for all t, then for each year $[t, t+1]$ considered separately,

$$B(t)\, \boldsymbol{e}(t) \geq P(t + \frac{1}{2}) = P(t) = P(t+1),$$

and $B(t)\, \boldsymbol{e}(t) = P(t)$ holds if and only if all countries have the same life expectancy.
Proof.

The indices i, j run over all countries from 1 to m. Write $t' = t + \frac{1}{2}$ so that $P(j, t') = \frac{P(j,t)+P(j,t+1)}{2}$ is the person-years lived in country j from t to $t+1$. Then the global life expectancy is

$$\begin{aligned}
\mathbf{e}(t) &= \frac{\sum_j P(j, t')\, \mathbf{e}(j, t)}{\sum_i P(i, t')} \\
&= \frac{\sum_j \frac{P(j,t')^2}{D(j,t)}}{P(t')} \\
&= \frac{\sum_j (\frac{P(j,t')}{\sqrt{D(j,t)}})^2}{P(t)}
\end{aligned}$$

Hence

$$
\begin{aligned}
B(t)\,\mathbf{e}(t) &= D(t)\frac{\sum_j (\frac{P(j,t')}{\sqrt{D(j,t)}})^2}{P(t)} \\
&= \sum_i D(i,t)\,\frac{\sum_j (\frac{P(j,t')}{\sqrt{D(j,t)}})^2}{P(t)} \\
&= \sum_i \sqrt{D(i,t)}^2 \frac{\sum_j (\frac{P(j,t')}{\sqrt{D(j,t)}})^2}{P(t)} \\
&\geq \frac{(\sum_j \sqrt{D(j,t)}\,\frac{P(j,t')}{\sqrt{D(j,t)}})^2}{P(t)} \\
&= \frac{P(t)^2}{P(t)} \\
&= P(t)
\end{aligned}
$$

where the inequality is Cauchy-Schwarz.

Appendix 3

In a constant global population of size $P(t)$ with migration between countries, if $B(t) = D(t)$ for all t, then the time-average birth rate $\overline{B} = \frac{1}{T}\sum_0^{T-1} B(t)$ and the time-average life expectancy $\overline{e} = \frac{1}{T}\sum_0^{T-1} \mathbf{e}(t)$ of the global population from 0 to T satisfy $P \leq \overline{B}\,\overline{e}$, and $P < \overline{B}\,\overline{e}$ unless all countries have equal life expectancies at all times.
Proof.

By the result in Appendix 2, the fixed global population size $P = P(0) > 0$ satisfies, for every $t = 0, 1, \ldots, T, \quad B(t)\,\mathbf{e}(t) \geq P$, with equality if and only if all countries have the same life expectancy between t and $t+1$.

Henceforth the lower and upper limits of summation are 0 and $T-1$, respectively. Since $B(t) \geq \frac{P}{\mathbf{e}(t)}$ for every $t = 0, 1, \ldots, T-1$,

$$\overline{B} = \frac{1}{T}\sum B(t) \geq \frac{P}{T}\sum \frac{1}{\mathbf{e}(t)} \geq \frac{P}{\frac{1}{T}\sum \mathbf{e}(t)} = \frac{P}{\overline{e}}.$$

The first inequality above follows from Cauchy-Schwarz and is strict if and only if, for at least one t, at least two countries have different life expectancies. The second inequality follows from Jensen's inequality (Appendix 2) and is strict unless $\mathbf{e}(t)$ is constant for t from 0 to $T-1$.

In detail, on the positive real line, $\frac{1}{y}$ is a strictly convex function of y. By Jensen's inequality, the average of the reciprocal of y is greater than or equal to the reciprocal of the average of y, and the inequality is strict if and only if y takes at least two distinct values each with positive measure. In particular, taking $y(t) = \mathbf{e}(t)$ gives

$$\sum \frac{1}{\mathbf{e}(t)} \geq \frac{1}{\sum \mathbf{e}(t)}$$

with strict inequality unless $\mathbf{e}(t)$ is almost always constant.