Sensitivity analysis in multiregional demographic models

F J Willekens

International Institute for Applied Systems Analysis, Schloss Laxenburg, 2361 Laxenburg, Austria
Received 7 January 1977, in revised form 16 March 1977

Abstract. The theory of multiregional mathematical demography investigates how fertility, mortality, and migration combine to shape the growth of multiregional population systems. Population dynamics have been studied for cases where the structural parameters, namely the age-specific rates of fertility, mortality, and migration, are fixed. This paper addresses the question of how the system behaves under changing structural parameters. By applying the technique of matrix differentiation, sensitivity functions are derived which link changes in multiregional life-table statistics and in population projections to changes in the age-specific rates. A review of the technique, which may be used for the sensitivity analysis of any matrix model, is given in the appendix.

The field of mathematical demography is concerned with the mathematical description of how fertility and mortality combine to determine the characteristics of populations and to shape their growth. Traditionally, demographers have restricted their attention to fertility and mortality (Coale, 1972; Keyfitz, 1968), and have in fact assumed that populations are 'closed' to migration, that is, populations are undisturbed by in- and out-migration. This is an unrealistic assumption, especially in population analysis at the subnational level. The introduction of migration into mathematical demography has been pioneered by Rogers (1975). He describes, in analytical terms, how fertility, mortality, and migration combine to determine the features and the growth of multiregional population systems. The basic tool used is matrix algebra.

Mathematical demography demonstrates how various population characteristics may be expressed in terms of observed age-specific fertility, mortality, and migration rates. The fundamental assumption underlying the models is that the age-specific rates, that is, the structural parameters, are known exactly and that they remain fixed over time. The implications of this are expressed by Keyfitz (1968, page 27): "The object [of population projection] is to understand the past rather than to predict the future; apparently the way to think effectively about an observed set of birth and death rates is to ask what it would lead to if continued".

No one truly believes that fertility, mortality, and migration schedules are measured without observation error and that they will remain unchanged for a prolonged period of time. However, variations in structural parameters have not been considered until recently (Goodman, 1969; 1971; Keyfitz, 1971; Preston, 1974).

It is the purpose of this paper to improve understanding of the impact on the population system of changes in its structural parameters. The system considered is the multiregional demographic one described in Rogers (1975). The parameters are the age-specific fertility, mortality, and migration rates. In general terms, the problem is to find how sensitive stationary population characteristics and population projections are to changes in age-specific rates. In another paper I have investigated some long-run impacts of the changes, namely the sensitivity of the stable population characteristics (Willekens, 1976).

There are two approaches to impact analysis. The first is the *simulation approach*, or the arithmetic approach as Keyfitz (1971, page 275) calls it. It is simply the computation of the life table and population projection under the old and the new rates. The differences in the output give the impact of changing the rates. Suitable

tools for the simulation approach are provided by model life tables, such as those developed by Coale and Demeny (1966) for a single-region demographic system and by Rogers (1975, chapter 6) for a multiregional system. For an illustration of this approach see Rogers (1975, pages 169–172), Rogers and Willekens (1976, pages 28–30). Besides its demanding character in terms of computer time, the approach tells us nothing about the complete set of parameters on which the changes in the final results depend. It will be found useful, however, for verification of the results of the second approach, which is the *analytical approach*. This procedure derives a general formula for assessing the impact of a particular change in terms of well-known population variables. Such a formula will be designated a *sensitivity function*. Partial differentiation will be seen to be the basic ingredient in the analysis of such functions.

In this paper, impact analysis is performed using the analytical approach. It is assumed that all the functions are differentiable with respect to the variables in which the changes occur. Since multiregional demographic models are formulated in matrix terms, matrix differentiation techniques are applied; and because not much work has been done in the area of matrix calculus, the appendix reviews several relevant topics of such a calculus⁽¹⁾.

1 Impact of changes in age-specific rates on life-table functions

The concept of a multiregional life table as developed by Rogers (1973; 1975) is a device for exhibiting the mortality and migration history of a set of regional cohorts as they age. It is assumed that the age-specific rates describing the mortality and mobility experience of an actual population remain constant, and that the system of regions is undisturbed by external migration.

The first part of this section sets out the life-table functions. The cohorts that will be considered are birth cohorts or *radices*. Their life histories are of special interest because they provide the information required by population projection models. The life-table statistics are given by place of birth. In the second part I combine the life-table functions with the matrix differentiation techniques described in the appendix. This enables one to develop life-table sensitivity functions.

1.1 The multiregional life table

All the life-table functions are derived from a set of age-specific death and out-migration rates. Let $M(x)$ denote the matrix of observed annual rates for the persons in the age interval from x to $x+h$. The length of the interval h is arbitrary. Without loss of generality I shall consider age intervals of five years. For an N -region system, $M(x)$ is

$$M(x) = \begin{bmatrix} M_{11}(x) + \sum_{j \neq 1}^N M_{1j}(x) & -M_{21}(x) & -M_{31}(x) & \dots \\ -M_{12}(x) & M_{22}(x) + \sum_{j \neq 2}^N M_{2j}(x) & -M_{32}(x) & \dots \\ -M_{13}(x) & -M_{23}(x) & M_{33}(x) + \sum_{j \neq 3}^N M_{3j}(x) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad (1)$$

(1) All major textbooks on matrix algebra that have been consulted lack a chapter on matrix calculus, although some scattered treatment may occur. The only unified treatment of matrix differentiation that I have found is by Dwyer and MacPhail (1948). A simplified and extended version appeared twenty years later in Dwyer (1967). The formulae given there are general enough to handle differentiation problems that occur in life-table functions and in the analysis of population projections over a finite time horizon.

where $M_{i8}(x)$ is the age-specific annual death rate in region i , and $M_{ij}(x)$ is the age-specific annual out-migration rate from region i to region j . This out-migration rate is estimated by dividing the annual number of out-migrants to j by the midyear population of i .

Let $P(x)$ be the matrix of age-specific probabilities of dying and out-migrating:

$$P(x) = \begin{bmatrix} p_{11}(x) & p_{21}(x) & \dots \\ p_{12}(x) & p_{22}(x) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}, \quad (2)$$

with $p_{ij}(x)$ being the probability that an individual in region i at exact age x will survive and be in region j at exact age $x+5$. The diagonal element $p_{ii}(x)$ is the probability that an individual will survive and be in region i at the end of the interval. If $q_i(x)$ is the probability that an individual in region i at age x will die before reaching age $x+5$, then the following relationship follows

$$p_{ii}(x) = 1 - q_i(x) - \sum_{j \neq i}^N p_{ij}(x). \quad (3)$$

The matrix of probabilities is derived from the matrix of vital rates (Rogers and Ledent, 1976; Schoen, 1975):

$$P(x) = [I + \frac{5}{2}M(x)]^{-1} [I - \frac{5}{2}M(x)]. \quad (4)$$

Note that $P(x)$ is the transition matrix of an *absorbing* Markov chain.

The probability that an individual starting out in region j , that is, born in j , will be in region i at exact age x is denoted by ${}_j\hat{l}_i(x)$. The matrix containing those probabilities is

$$\hat{l}(x) = \begin{bmatrix} {}_1\hat{l}_1(x) & {}_2\hat{l}_1(x) & \dots \\ {}_1\hat{l}_2(x) & {}_2\hat{l}_2(x) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}. \quad (5)$$

The matrix $\hat{l}(x)$ describes the state of the system at time x . If one assumes that the state at time x depends only on the previous state $\hat{l}(x-5)$ (that is, the Markov property), then equation (5) may be written as

$$\hat{l}(x) = P(x-5)P(x-10) \dots P(0) = P(x-5)\hat{l}(x-5). \quad (6)$$

Define

$$l(x) = \hat{l}(x)l(0), \quad (7)$$

where $l(0)$ is a diagonal matrix of the cohorts of babies born in the N regions at a given instant in time. Typically, ${}_i l_i(0)$ is called the radix of region i and is set equal to some arbitrary constant such as 100000. Then $l(x)$ is the matrix of the number of people at exact age x by place of residence and by place of birth.

Another life-table function is the total number of people of age-group x , that is, aged x to $x+4$ years, in each region by place of birth:

$$L(x) = \begin{bmatrix} {}_1L_1(x) & {}_2L_1(x) & \dots \\ {}_1L_2(x) & {}_2L_2(x) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}, \quad (8)$$

with ${}_jL_i(x)$ being the number of people in region i in age-group x who were born in region j . In addition the element ${}_jL_i(x)$ can be thought of as the total person-years lived in region i between ages x and $x+5$, by the people born in region j (Rogers, 1975, page 65). The matrix $L(x)$ is given by

$$L(x) = \int_0^5 l(x+t) dt = \left[\int_0^5 \hat{l}(x+t) dt \right] l(0). \quad (9)$$

With the assumption of a uniform distribution of out-migrations and deaths over the five-year age interval, we may obtain numerical values for $L(x)$ by the linear interpolation

$$L(x) = \frac{5}{2} [l(x) + l(x+5)],$$

or

$$L(x) = \frac{5}{2} [I + P(x)] l(x). \quad (10)$$

Another useful expression for $L(x)$ is found by noting that

$$[I + P(x)] = [I + \frac{5}{2}M(x)]^{-1} \left[[I + \frac{5}{2}M(x)] + [I - \frac{5}{2}M(x)] \right] = 2[I + \frac{5}{2}M(x)]^{-1}. \quad (11)$$

Equation (10) may therefore be rewritten as

$$L(x) = 5[I + \frac{5}{2}M(x)]^{-1} l(x). \quad (12)$$

If $L(x)$ is aggregated over all age groups beyond age x , the expected total number of person-years remaining to the people at exact age x can be defined as

$$T(x) = \sum_{y=x}^z L(y), \quad (13)$$

where z is the terminal age group.

By expressing $T(x)$ per individual, we get the matrix of expectation of life of an individual at exact age x :

$$e(x) = T(x)l^{-1}(x) = \left[\sum_{y=x}^z L(y) \right] l^{-1}(x). \quad (14)$$

The matrix $e(x)$ expresses the life expectancies by place of residence. An element ${}_ie_j(x)$ denotes the expected number of years lived in region j beyond age x by an individual now residing in region i and x years of age. The expectation-of-life matrix by place of birth is given by

$$\bar{e}(x) = T(x)\bar{l}^{-1}(x), \quad (15)$$

where

$$\bar{l}(x) = \begin{bmatrix} {}_1l(x) & 0 & \dots \\ 0 & {}_2l(x) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad (16)$$

is a diagonal matrix. The element ${}_il(x)$ denotes the total number of people of exact age x who are born in region i . An element ${}_ie_j(x)$ gives the expected number of years lived in region j beyond age x by an individual born in region i and now x years of age (his current place of residence is not considered).

A very useful life-table function is the survivorship matrix. It is an essential component of the population projection matrix. Rogers (1975, page 79) has shown

that the survivorship matrix

$$\mathbf{S}(x) = \begin{bmatrix} s_{11}(x) & s_{21}(x) & \dots \\ s_{12}(x) & s_{22}(x) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad (17)$$

is given by

$$\mathbf{S}(x) = \mathbf{L}(x+5)\mathbf{L}^{-1}(x). \quad (18)$$

The element $s_{ij}(x)$ denotes the proportion of individuals aged x to $x+4$ in region i who survive to be $x+5$ to $x+9$ years old five years later, and are then in region j .

The survivorship matrix may be expressed directly in terms of the observed age-specific rates⁽²⁾. Substitution of equation (12) into equation (18) gives

$$\mathbf{S}(x) = [\mathbf{I} + \frac{5}{2}\mathbf{M}(x+5)]^{-1}\mathbf{P}(x)\mathbf{I}(x)\mathbf{I}^{-1}(x)[\mathbf{I} + \frac{5}{2}\mathbf{M}(x)],$$

which yields, after replacing equation (4) for $\mathbf{P}(x)$ and reworking,

$$\mathbf{S}(x) = [\mathbf{I} + \frac{5}{2}\mathbf{M}(x+5)]^{-1}[\mathbf{I} - \frac{5}{2}\mathbf{M}(x)]. \quad (19)$$

I have now set up the important life-table functions, and can proceed to the analysis of their sensitivities to changes in the underlying rates, that is, in $\mathbf{M}(x)$.

1.2 Sensitivity analysis of life-table functions

The fundamental question posed in this section is: what is the effect on the various life-table statistics of a change in the observed age-specific rates? To resolve this question the life-table functions are combined with the matrix differentiation techniques given in the appendix.

This section is divided into five parts, each of which starts out with a specific life-table function. The derivative of this function with respect to an element of the matrix of age-specific rates yields the corresponding sensitivity function.

1.2.1 Sensitivity of the probabilities of dying and out-migrating. Recall the estimating formula set out in equation (4). The probability matrix $\mathbf{P}(x)$ depends only on $\mathbf{M}(x)$. Therefore $\mathbf{P}(a)$ is not affected by a change in $\mathbf{M}(x)$ for $a \neq x$.

The derivative of $\mathbf{P}(x)$ with respect to an arbitrary element of $\mathbf{M}(x)$ is, by virtue of formulae (A13) and (A25) of the appendix,

$$\begin{aligned} \frac{\partial \mathbf{P}(x)}{\partial \langle \mathbf{M}(x) \rangle} &= \frac{\partial [\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1}}{\partial \langle \mathbf{M}(x) \rangle} [\mathbf{I} - \frac{5}{2}\mathbf{M}(x)] + [\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1} \frac{\partial [\mathbf{I} - \frac{5}{2}\mathbf{M}(x)]}{\partial \langle \mathbf{M}(x) \rangle} \\ &= -[\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1} \frac{\partial [\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]}{\partial \langle \mathbf{M}(x) \rangle} [\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1} [\mathbf{I} - \frac{5}{2}\mathbf{M}(x)] \\ &\quad + [\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1} \frac{\partial [\mathbf{I} - \frac{5}{2}\mathbf{M}(x)]}{\partial \langle \mathbf{M}(x) \rangle} = -[\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1} [\frac{5}{2}\mathbf{J}\mathbf{P}(x) + \frac{5}{2}\mathbf{J}], \end{aligned}$$

where $\mathbf{J}$ is a matrix of the same dimension as $\mathbf{M}(x)$, with all elements zero except for a unit element in the position of the arbitrary element $\langle \mathbf{M}(x) \rangle$. (This notation is further explained in the appendix.) Therefore the sensitivity function for $\mathbf{P}(x)$ is

$$\frac{\partial \mathbf{P}(x)}{\partial \langle \mathbf{M}(x) \rangle} = -\frac{5}{2}[\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1}\mathbf{J}[\mathbf{P}(x) + \mathbf{I}]. \quad (20)$$

Substitution from equation (11) for $[\mathbf{P}(x) + \mathbf{I}]$ yields

$$\frac{\partial \mathbf{P}(x)}{\partial \langle \mathbf{M}(x) \rangle} = -5[\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1}\mathbf{J}[\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1}. \quad (21)$$

⁽²⁾ The author is grateful to Jacques Ledent for the derivation.

1.2.2 *Sensitivity of the number of people at exact age a .* A change in $\mathbf{M}(x)$ does not affect $\mathbf{l}(a)$ for $a \leq x$. Therefore we look only at the case $a > x$. Note that because of equation (6) $\mathbf{l}(a)$ may be written as

$$\mathbf{l}(a) = \mathbf{P}(a-5)\mathbf{P}(a-10) \dots \mathbf{P}(x)\mathbf{l}(x) .$$

On recalling that $\mathbf{M}(x)$ affects only $\mathbf{P}(x)$, one may write

$$\begin{aligned} \frac{\partial \mathbf{l}(a)}{\partial \langle \mathbf{M}(x) \rangle} &= \mathbf{P}(a-5)\mathbf{P}(a-10) \dots \frac{\partial \mathbf{P}(x)}{\partial \langle \mathbf{M}(x) \rangle} \mathbf{l}(x) \\ &= -\frac{5}{2} \left[\mathbf{P}(a-5)\mathbf{P}(a-10) \dots \mathbf{P}(x+5) [\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1} \mathbf{J} [\mathbf{P}(x) + \mathbf{I}] \mathbf{l}(x) \right] . \end{aligned} \quad (22)$$

By inserting

$$\mathbf{I} = [\mathbf{I} - \frac{5}{2}\mathbf{M}(x)]\mathbf{l}(x)\mathbf{l}^{-1}(x) [\mathbf{I} - \frac{5}{2}\mathbf{M}(x)]^{-1}$$

into equation (22), substituting for $\mathbf{P}(x)$, and using equation (10), one obtains

$$\frac{\partial \mathbf{l}(a)}{\partial \langle \mathbf{M}(x) \rangle} = -\mathbf{l}(a)\mathbf{l}^{-1}(x) [\mathbf{I} - \frac{5}{2}\mathbf{M}(x)]^{-1} \mathbf{J} \mathbf{L}(x) . \quad (23)$$

Another expression is

$$\frac{\partial \mathbf{l}(a)}{\partial \langle \mathbf{M}(x) \rangle} = -\mathbf{l}(a)\mathbf{l}^{-1}(x+5) [\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1} \mathbf{J} \mathbf{L}(x) . \quad (24)$$

For $a = x+5$, we have

$$\frac{\partial \mathbf{l}(x+5)}{\partial \langle \mathbf{M}(x) \rangle} = -\mathbf{P}(x) [\mathbf{I} - \frac{5}{2}\mathbf{M}(x)]^{-1} \mathbf{J} \mathbf{L}(x) = -[\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1} \mathbf{J} \mathbf{L}(x) . \quad (25)$$

An interesting formulation of the sensitivity function is

$$\frac{\partial \mathbf{l}(a)}{\partial \langle \mathbf{M}(x) \rangle} = \mathbf{l}(a)\mathbf{l}^{-1}(x)\mathbf{P}^{-1}(x) \frac{\partial \mathbf{P}(x)}{\partial \langle \mathbf{M}(x) \rangle} \mathbf{l}(x) ,$$

or

$$\mathbf{l}^{-1}(a) \frac{\partial \mathbf{l}(a)}{\partial \langle \mathbf{M}(x) \rangle} = \mathbf{l}^{-1}(x+5) \frac{\partial \mathbf{P}(x)}{\partial \langle \mathbf{M}(x) \rangle} \mathbf{l}(x) . \quad (26)$$

This shows that the relative sensitivity of $\mathbf{l}(a)$ to changes in $\mathbf{M}(x)$ is a weighted average of the relative sensitivity of $\mathbf{P}(x)$, and is independent of a . Consider the first age group and suppose that all regions have the same radices, that is, $\mathbf{l}(0)$ is a scalar matrix—a diagonal matrix with the same diagonal elements. The relative sensitivity of any $\mathbf{l}(a)$ is then proportional to the relative sensitivity of $\mathbf{P}(0)$.

1.2.3 *Sensitivity of the number of people in age-group $(a, a+4)$.* What is the impact of a change in $\mathbf{M}(x)$ on the number of people in age-group $(a, a+4)$ and on their spatial distribution? It is clear that $\mathbf{M}(x)$ does not affect $\mathbf{L}(a)$ for $a < x$. Therefore I shall consider here the case of $a \geq x$. Recall from equation (10) that

$$\mathbf{L}(a) = \frac{5}{2} [\mathbf{l}(a+5) + \mathbf{l}(a)] .$$

Differentiating both sides gives

$$\frac{\partial \mathbf{L}(a)}{\partial \langle \mathbf{M}(x) \rangle} = \frac{5}{2} \left[\frac{\partial \mathbf{l}(a+5)}{\partial \langle \mathbf{M}(x) \rangle} + \frac{\partial \mathbf{l}(a)}{\partial \langle \mathbf{M}(x) \rangle} \right] .$$

If $a = x$, then $\partial \mathbf{l}(a)/\partial \langle \mathbf{M}(x) \rangle = 0$ and we have

$$\frac{\partial \mathbf{L}(x)}{\partial \langle \mathbf{M}(x) \rangle} = \frac{5}{2} \frac{\partial \mathbf{l}(x+5)}{\partial \langle \mathbf{M}(x) \rangle} = -\frac{5}{2} [\mathbf{I} + \frac{5}{2}\mathbf{M}(x)]^{-1} \mathbf{J} \mathbf{L}(x) \quad (27)$$

which has the following alternative expression:

$$\frac{\partial L(x)}{\partial \langle M(x) \rangle} = -\frac{5}{4} [P(x) + I] J L(x) . \quad (28)$$

If $a > x$, we know that $P(a)$ is independent of $M(x)$ and therefore

$$\frac{\partial L(a)}{\partial \langle M(x) \rangle} = \frac{5}{2} [P(a) + I] \frac{\partial l(a)}{\partial \langle M(x) \rangle} = -L(a) l^{-1}(x) [I - \frac{5}{2} M(x)]^{-1} J L(x) \quad (29)$$

$$= -L(a) l^{-1}(x+5) [I + \frac{5}{2} M(x)]^{-1} J L(x) , \quad (30)$$

which may also be written as

$$\frac{\partial L(a)}{\partial \langle M(x) \rangle} = \frac{5}{2} [P(a) + I] l(a) l^{-1}(a) \frac{\partial l(a)}{\partial \langle M(x) \rangle} ,$$

whence, since $\frac{5}{2} [P(a) + I] l(a) = L(a)$,

$$L^{-1}(a) \frac{\partial L(a)}{\partial \langle M(x) \rangle} = l^{-1}(a) \frac{\partial l(a)}{\partial \langle M(x) \rangle} . \quad (31)$$

Equation (31) indicates that the relative sensitivity of the number of people in age-group $(a, a+4)$ is equal to the relative sensitivity of the number of people at exact age a for $a > x$.

1.2.4 Sensitivity of the expectation of life at age a . I now derive the sensitivity function of the most important life-table statistic, namely the expectation of life. I shall distinguish between the expectation of life by place of residence and the expectation of life by place of birth.

1.2.4.1 Expectation of life by place of residence. Consider first the sensitivity of $e(x)$. Differentiating both sides of equation (14) yields

$$\frac{\partial e(x)}{\partial \langle M(x) \rangle} = \frac{\partial \left[\sum_{y=x}^z L(y) \right]}{\partial \langle M(x) \rangle} l^{-1}(x) + \left[\sum_{y=x}^z L(y) \right] \frac{\partial l^{-1}(x)}{\partial \langle M(x) \rangle} . \quad (32)$$

From equations (27) and (29) we see that

$$\begin{aligned} \frac{\partial \left[\sum_{y=x}^z L(y) \right]}{\partial \langle M(x) \rangle} &= - \left[\sum_{y=x+5}^z L(y) \right] l^{-1}(x) [I - \frac{5}{2} M(x)]^{-1} J L(x) - \frac{5}{2} [I + \frac{5}{2} M(x)]^{-1} J L(x) \\ &= - \left[\sum_{y=x+5}^z L(y) l^{-1}(x) + \frac{5}{2} P(x) + \frac{5}{2} I - \frac{5}{2} I \right] [I - \frac{5}{2} M(x)]^{-1} J L(x) \\ &= - \left[\sum_{y=x+5}^z L(y) l^{-1}(x) + L(x) l^{-1}(x) - \frac{5}{2} I \right] [I - \frac{5}{2} M(x)]^{-1} J L(x) \\ &= -[e(x) - \frac{5}{2} I] [I - \frac{5}{2} M(x)]^{-1} J L(x) . \end{aligned} \quad (33)$$

Since $l(x)$ is independent of $M(x)$, equation (32) may be written as follows:

$$\frac{\partial e(x)}{\partial \langle M(x) \rangle} = -[e(x) - \frac{5}{2} I] [I - \frac{5}{2} M(x)]^{-1} J L(x) l^{-1}(x) \quad (34)$$

or

$$\frac{\partial e(x)}{\partial \langle M(x) \rangle} = -5[e(x) - \frac{5}{2} I] [I - \frac{5}{2} M(x)]^{-1} J [I + \frac{5}{2} M(x)]^{-1} . \quad (35)$$

For $a < x$ we have

$$\frac{\partial e(a)}{\partial \langle M(x) \rangle} = \frac{\partial \left[\sum_{y=x+5}^z L(y) + L(x) + \sum_{y=a}^{x-5} L(y) \right] I^{-1}(a)}{\partial \langle M(x) \rangle}.$$

We know that

$$\frac{\partial L(y)}{\partial \langle M(x) \rangle} = 0, \quad \text{for } y < x,$$

and

$$\frac{\partial l(a)}{\partial \langle M(x) \rangle} = 0, \quad \text{for } a < x.$$

Therefore

$$\frac{\partial e(a)}{\partial \langle M(x) \rangle} = \frac{\partial \left[\sum_{y=x}^z L(y) \right] I^{-1}(a)}{\partial \langle M(x) \rangle},$$

$$\frac{\partial e(a)}{\partial \langle M(x) \rangle} = \frac{\partial e(x)}{\partial \langle M(x) \rangle} l(x) I^{-1}(a), \quad (36)$$

$$\frac{\partial e(a)}{\partial \langle M(x) \rangle} = -[e(x) - \tfrac{5}{2}I] [I - \tfrac{5}{2}M(x)]^{-1} J L(x) I^{-1}(a) \quad (37)$$

$$= -e(x) [I - \tfrac{5}{2}M(x)]^{-1} J L(x) I^{-1}(a) + \tfrac{5}{2} [I - \tfrac{5}{2}M(x)]^{-1} J L(x) I^{-1}(a). \quad (38)$$

The second component of the sensitivity function follows from the linear approximation (10) of the continuous relationship (9).

Consider the continuous definition of $e(a)$

$$e(a) = \left[\int_a^\omega l(t) dt \right] I^{-1}(a)$$

where ω is the terminal age. Differentiation, for $a \leq x$, yields

$$\begin{aligned} \frac{\partial e(a)}{\partial \langle M(x) \rangle} &= \left[\int_a^\omega \frac{\partial l(t)}{\partial \langle M(x) \rangle} dt \right] I^{-1}(a) \\ &= \left[\int_x^\omega -l(t) I^{-1}(x) [I - \tfrac{5}{2}M(x)]^{-1} J l(x) dt \right] I^{-1}(a). \end{aligned}$$

Since $l(t)$ is independent of $M(x)$ for $t < x$, we have

$$\begin{aligned} \frac{\partial e(a)}{\partial \langle M(x) \rangle} &= - \left[\int_x^\omega l(t) dt \right] I^{-1}(x) [I - \tfrac{5}{2}M(x)]^{-1} J l(x) I^{-1}(a) \\ &= -e(x) [I - \tfrac{5}{2}M(x)]^{-1} J l(x) I^{-1}(a), \end{aligned} \quad (39)$$

which is equivalent to the first term of equation (38) with $l(x)$ replaced by $L(x)$ in the discrete case. The expression (39), written in terms of differentials, is similar to the sensitivity function of the expectation of life, given by Keyfitz (1971, page 276) for the single-region case as

$$de(a) = -e(x) [dM(x)] l(x) I^{-1}(a),$$

where $e(\cdot)$, $l(\cdot)$, and $M(\cdot)$ are scalars. The term $[I - \tfrac{5}{2}M(x)]^{-1}$ in equation (39) occurs because I have considered observed rates where Keyfitz derived the formula using instantaneous rates. If $M(x)$ contained instantaneous rates, then $M(x) \doteq 0$ and $[I - \tfrac{5}{2}M(x)] \doteq I$.

For $a > x$, we have

$$\frac{\partial e(a)}{\partial \langle M(x) \rangle} = \frac{\partial \left[\sum_{y=a}^z L(y) \right]}{\partial \langle M(x) \rangle} I^{-1}(a) + \left[\sum_{y=a}^z L(y) \right] \frac{\partial I^{-1}(a)}{\partial \langle M(x) \rangle}.$$

Applying equations (23), (29), and (A25) gives

$$\begin{aligned} \frac{\partial e(a)}{\partial \langle M(x) \rangle} &= -e(a) I^{-1}(x) [I - \frac{1}{2} M(x)]^{-1} J L(x) I^{-1}(a) \\ &\quad + e(a) I^{-1}(a) l(a) I^{-1}(x) [I - \frac{1}{2} M(x)]^{-1} J L(x) I^{-1}(a) = 0. \end{aligned} \quad (40)$$

Therefore a change in $M(x)$ does not affect $e(a)$ for $a > x$.

1.2.4.2 Expectation of life by place of birth. The life expectancy by place of birth is given by formula (15). By making use of derivations analogous to the ones in section 1.2.4.1, it can be shown that for $a \leq x$ the sensitivity functions of $\bar{e}(a)$ and $e(a)$ are very similar:

$$\frac{\partial \bar{e}(a)}{\partial \langle M(x) \rangle} = -[e(x) - \frac{1}{2} I] [I - \frac{1}{2} M(x)]^{-1} J L(x) \bar{I}^{-1}(a). \quad (41)$$

Contrary to the finding of the previous section, a change in $M(x)$ affects $\bar{e}(a)$ for $a > x$. For $a > x$, we may write

$$\frac{\partial \bar{e}(a)}{\partial \langle M(x) \rangle} = \frac{\partial \left[\sum_{y=a}^z L(y) \right]}{\partial \langle M(x) \rangle} \bar{I}^{-1}(a) + \left[\sum_{y=a}^z L(y) \right] \frac{\partial \bar{I}^{-1}(a)}{\partial \langle M(x) \rangle}. \quad (42)$$

By equation (29) we have

$$\frac{\partial \left[\sum_{y=a}^z L(y) \right]}{\partial \langle M(x) \rangle} = -e(a) I^{-1}(x) [I - \frac{1}{2} M(x)]^{-1} J L(x).$$

On the other hand

$$\frac{\partial \bar{I}^{-1}(a)}{\partial \langle M(x) \rangle} = -I^{-1}(a) \frac{\partial \bar{I}(a)}{\partial \langle M(x) \rangle} I^{-1}(a),$$

where

$$\frac{\partial \bar{I}(a)}{\partial \langle M(x) \rangle} = Q \quad (43)$$

is a diagonal matrix. The diagonal consists of the elements of the vector

$$\{1\}' \frac{\partial l(a)}{\partial \langle M(x) \rangle} = -\{1\}' l(a) I^{-1}(x) [I - \frac{1}{2} M(x)]^{-1} J L(x).$$

Introducing these results into equation (42) gives

$$\begin{aligned} \frac{\partial \bar{e}(a)}{\partial \langle M(x) \rangle} &= -e(a) I^{-1}(x) [I - \frac{1}{2} M(x)]^{-1} J L(x) \bar{I}^{-1}(a) - \bar{I}^{-1}(a) Q \bar{I}^{-1}(a) \\ &= e(a) I^{-1}(a) \frac{\partial l(a)}{\partial \langle M(x) \rangle} \bar{I}^{-1}(a) - \bar{I}^{-1}(a) \frac{\partial \bar{I}(a)}{\partial \langle M(x) \rangle} \bar{I}^{-1}(a). \end{aligned} \quad (44)$$

1.2.5 Sensitivity of the survivorship proportions. As in the preceding sections, we treat $S(a)$ for $a = x$ and for $a \neq x$ separately. The survivorship matrix is given by

equation (19). Differentiating with respect to $\langle M(x) \rangle$ yields

$$\frac{\partial S(x)}{\partial \langle M(x) \rangle} = -\frac{5}{2} [I + \frac{5}{2} M(x+5)]^{-1} J, \quad (45)$$

or

$$\frac{\partial S(x)}{\partial \langle M(x) \rangle} = -\frac{5}{2} S(x) [I - \frac{5}{2} M(x)]^{-1} J. \quad (46)$$

To illustrate the dynamic relationship between the life-table statistics, one may express the sensitivity of $S(x)$ in relation to the sensitivity of other statistics. For example, combination of equation (46) with equation (29) yields

$$S^{-1}(x) \frac{\partial S(x)}{\partial \langle M(x) \rangle} = P^{-1}(x) \frac{\partial L(x)}{\partial \langle M(x) \rangle} L^{-1}(x),$$

and combination of equation (46) with equation (23) gives

$$S^{-1}(x) \frac{\partial S(x)}{\partial \langle M(x) \rangle} = -\frac{5}{2} P^{-1}(x) \frac{\partial I(x+5)}{\partial \langle M(x) \rangle} L^{-1}(x).$$

The relative sensitivity of $S(x)$ may be regarded as a weighted measure of the sensitivities of other statistics.

I now turn to the sensitivity of $S(a)$ to changes in $M(x)$ for $a \neq x$. For $a > x$ and for $a < x-5$, $S(a)$ is independent of a change in $M(x)$. The sensitivity of $S(x-5)$ to a change in $M(x)$ is derived easily. Write equation (19) for $x-5$:

$$S(x-5) = [I + \frac{5}{2} M(x)]^{-1} [I - \frac{5}{2} M(x-5)].$$

Then

$$\begin{aligned} \frac{\partial S(x-5)}{\partial \langle M(x) \rangle} &= -\frac{5}{2} [I + \frac{5}{2} M(x)]^{-1} J [I + \frac{5}{2} M(x)]^{-1} [I - \frac{5}{2} M(x-5)] \\ &= -\frac{5}{2} [I + \frac{5}{2} M(x)]^{-1} J S(x-5). \end{aligned} \quad (47)$$

The relationship between the sensitivity of $S(x)$ and $S(x-5)$ is thus

$$\frac{\partial S(x)}{\partial \langle M(x) \rangle} = S(x) P^{-1}(x) \frac{\partial S(x-5)}{\partial \langle M(x) \rangle} S^{-1}(x-5), \quad (48)$$

and

$$\frac{\partial S(x-5)}{\partial \langle M(x) \rangle} = P(x) S^{-1}(x) \frac{\partial S(x)}{\partial \langle M(x) \rangle} S(x-5). \quad (49)$$

2 Impact of changes in age-specific rates on the population projection

Population projection is often carried out under the assumption that an observed population-growth regime will remain constant. This implies that the observed age-specific rates will not change over the projection period. (This is a crude assumption and no demographer or planner considers it to be realistic. Nevertheless it produces a useful benchmark against which to compare other alternative projections.) In this section I deal with the question of how sensitive are population projections to changes in age-specific rates. These variations may occur at any point in time. If they occur in the base year, they can be related to observation errors. The sensitivity functions developed remain exactly the same no matter what the causes of the variations are.

In the first part the population-growth model is set out as a system of first-order linear homogeneous difference equations with constant coefficients, as in Rogers

(1975, chapter 5). The second part studies the sensitivity of population growth to changes in observed age-specific rates.

2.1 The discrete model of multiregional demographic growth

Population growth may be expressed in terms of the changing level of population or in terms of the variation of the number of births over time. In demography it has been a custom to formulate the discrete model of population growth in terms of total population, whereas the continuous version describes the birth trajectory (Keyfitz, 1968; Rogers, 1975). In this paper I shall consider only the discrete population-growth model.

A multiregional growth process may be described as a matrix multiplication (Rogers, 1975, page 123):

$$\{K^{(t+1)}\} = G\{K^{(t)}\} \quad (50)$$

where the vector $\{K^{(t)}\}$ describes the regional age-specific population distribution at time t , with

$$\{K^{(t)}\} = \begin{bmatrix} \{K^{(t)}(0)\} \\ \{K^{(t)}(5)\} \\ \vdots \\ \{K^{(t)}(z)\} \end{bmatrix}, \quad \text{and} \quad \{K^{(t)}(x)\} = \begin{bmatrix} K_1^{(t)}(x) \\ K_2^{(t)}(x) \\ \vdots \\ K_N^{(t)}(x) \end{bmatrix}, \quad (51)$$

z being the terminal age interval and N the number of regions.

Each element $K_i^{(t)}(x)$ denotes the number of people in region i at time t who are x to $x+4$ years old. Note that $t+1$ represents the next moment in time, that is, five years later than t ; age-groups and time intervals of five years are being considered.

The operator G is the generalized Leslie matrix

$$G = \begin{bmatrix} 0 & 0 & B(\alpha-5) & \dots & B(\beta-5) & \dots & 0 & 0 \\ S(0) & 0 & & & & & & \\ 0 & S(5) & \dots & \dots & \dots & \dots & & \\ \vdots & \vdots & & & & & & \\ 0 & 0 & & & & & S(z-5) & 0 \end{bmatrix}, \quad (52)$$

with $S(x)$, the matrix of survivorship proportions, set out in section 1.1. The first and last ages of childbearing may be denoted by α and β , respectively, and

$$B(x) = \begin{bmatrix} b_{11}(x) & b_{21}(x) & \dots \\ b_{12}(x) & b_{22}(x) & \dots \\ \vdots & \vdots & \end{bmatrix},$$

where an element $b_{ij}(x)$ denotes the average number of babies born during the unit time interval in region i and alive in region j at the end of that interval, per individual living in region i at the beginning of the interval, and aged x to $x+4$ years old.

The matrix $B(x)$ may be estimated by the following expression (Rogers, 1975, pages 120–121):

$$B(x) = \frac{1}{2} L(0) I^{-1}(0) [F(x) + F(x+5)S(x)].$$

Since

$$L(0) = \frac{5}{2} [I(5) + I(0)] = \frac{5}{2} [P(0) + I] I(0) ,$$

we may write

$$B(x) = \frac{5}{4} [P(0) + I] [F(x) + F(x+5) S(x)] . \quad (53)$$

The quantities $L(0)$, $I(0)$, $P(0)$, and $S(x)$ have been defined in section 1.1. $P(0)$ and $S(x)$ are given by the life table, and $F(x)$ is a diagonal matrix containing the annual regional birthrates of people aged x to $x+4$ years old. The number of births in year t from people aged x to $x+4$ years old at t is $F(x)\{K^{(t)}(x)\}$. The number of births during a five-year period starting at t , from people aged x to $x+4$ years old at t , is

$$\frac{5}{2} [F(x)\{K^{(t)}(x)\} + F(x+5)\{K^{(t+1)}(x+5)\}] = \frac{5}{2} [F(x) + F(x+5)S(x)]\{K^{(t)}(x)\} .$$

Of these births, a proportion $L(0)[5I(0)]^{-1}$ will be surviving in the various regions at the end of the time interval. Note that this model of the survivorship and migratory behavior of zero- to four-year old children does not incorporate the observed dependence of the migration pattern of children on that of their parents. For a critique and possible extension see Rees and Wilson (1977, pages 146–150).

Because of the particular structure of the generalized Leslie matrix, equation (50) may be written as two equation systems:

$$\{K^{(t+1)}(0)\} = \sum_{\alpha=5}^{\beta-5} B(x)\{K^{(t)}(x)\} , \quad (54)$$

$$\{K^{(t+1)}(x+5)\} = S(x)\{K^{(t)}(x)\} , \quad \text{for } 5 \leq x \leq z-5 . \quad (55)$$

The vector $\{K^{(t)}(x)\}$ may be expressed in the form

$$\begin{aligned} \{K^{(t+x/5)}(x)\} &= [S(x-5) S(x-10) \dots S(5) S(0)]\{K^{(t)}(0)\} \\ &= A(x)\{K^{(t)}(0)\} , \quad \text{say,} \end{aligned} \quad (56)$$

where we define

$$A(x) = \begin{cases} I , & \text{for } x = 0 \\ \prod_{y=x-5}^0 S(y) , & \text{for } x = 5, 10, \dots, z \end{cases} = L(x) L^{-1}(0) , \quad (57)$$

with

$$\prod_{y=x-5}^0 S(y) = S(x-5) S(x-10) \dots S(5) S(0) . \quad (58)$$

The element $a_{ij}(x)$ of $A(x)$ is the proportion of individuals aged 0 to 4 years old in region i who will survive to be x to $x+4$ years old exactly x years later, and will at that time be in region j .

2.2 Sensitivity analysis of the population projection

Recall the population growth model defined in equation (50):

$$\{K^{(t+1)}\} = G\{K^{(t)}\} .$$

The assessment of the sensitivity of $\{K^{(t+1)}\}$ to changes in age-specific rates $M(x)$ may be analyzed by means of a two-step process. The first step considers the sensitivity of the growth matrix to changes in age-specific rates. The second step derives a sensitivity function which describes the impact on the population distribution of a change in the growth matrix. In the sensitivity analysis of life-table statistics, we

were not concerned with the time when the change in $M(x)$ occurred. The time consideration was irrelevant since the life table is a static model. For the sensitivity analysis of the population growth, however, it is important to know not only the age group *where* a change in $M(x)$ occurs, but also the time *when* the change occurs. I shall denote this time by t_0 . The time at which the change in the population distribution is measured will be denoted by t_1 .

Besides the change in $\{K^{(t_1)}\}$ caused by a change in the age-specific rates at t_0 , one may also consider the problem of how a unique change in $\{K^{(t_0)}\}$ affects $\{K^{(t_1)}\}$. These are two separate sensitivity problems. In the first, the parameter changes at t_0 and then remains at this new level. The second problem, however, is equivalent to a parameter change at the period preceding t_0 only. These two sensitivity problems will be treated separately.

2.2.1 Sensitivity of the growth matrix. The growth matrix G is composed of two types of submatrices, $S(x)$ and $B(x)$. It has been shown in section 1.2, equations (45) and (47), that a change in $M(x)$ affects only $S(x)$ and $S(x-5)$:

$$\begin{aligned}\frac{\partial S(x)}{\partial \langle M(x) \rangle} &= -\frac{5}{2} [I + \frac{5}{2} M(x+5)]^{-1} J, \\ \frac{\partial S(x-5)}{\partial \langle M(x) \rangle} &= -\frac{5}{2} [I + \frac{5}{2} M(x)]^{-1} J S(x-5).\end{aligned}$$

The sensitivity function of $B(x)$ remains to be derived. In equation (53) $B(x)$ depends on the age-specific death rates and out-migration rates through $S(x)$ and $P(0)$, and on the age-specific fertility rates $F(x)$ and $F(x+5)$. Consider the partial derivative of $B(x)$ with respect to $M(x)$:

$$\begin{aligned}\frac{\partial B(x)}{\partial \langle M(x) \rangle} &= \frac{5}{4} \frac{\partial [P(0) + I]}{\partial \langle M(x) \rangle} F(x) + \frac{5}{4} \frac{\partial [P(0) + I]}{\partial \langle M(x) \rangle} F(x+5) S(x) \\ &\quad + \frac{5}{4} [P(0) + I] F(x+5) \frac{\partial S(x)}{\partial \langle M(x) \rangle}.\end{aligned}\quad (59)$$

Since $P(0)$ is affected by a change in $M(x)$ only if $x = 0$, and because for this case $F(x)$ and $F(x+5)$ are 0, equation (59) reduces to

$$\frac{\partial B(x)}{\partial \langle M(x) \rangle} = \frac{5}{4} [P(0) + I] F(x+5) \frac{\partial S(x)}{\partial \langle M(x) \rangle}, \quad (60)$$

which, by equation (45), is

$$\frac{\partial B(x)}{\partial \langle M(x) \rangle} = -\frac{25}{8} [P(0) + I] F(x+5) [I + \frac{5}{2} M(x+5)]^{-1} J. \quad (61)$$

Since a change in $M(x)$ affects $S(x-5)$, it also affects $B(x-5)$:

$$\frac{\partial B(x-5)}{\partial \langle M(x) \rangle} = -\frac{25}{8} [P(0) + I] F(x) [I + \frac{5}{2} M(x)]^{-1} J S(x-5). \quad (62)$$

The sensitivity of $B(x)$ with respect to $F(x)$ and $F(x+5)$ may also be derived easily:

$$\frac{\partial B(x)}{\partial \langle F(x) \rangle} = \frac{5}{4} [P(0) + I] J = \frac{5}{2} L(0) [5I(0)]^{-1} J, \quad (63)$$

and

$$\frac{\partial B(x-5)}{\partial \langle F(x) \rangle} = \frac{5}{4} [P(0) + I] J S(x-5). \quad (64)$$

Thus the impact of a unit change in the fertility matrix $F(x)$ on the element $B(x)$ is $\frac{5}{2}$ times the proportion of newborn babies that will be alive at the end of the time interval.

Having derived sensitivity functions for the elements of the growth matrix, we can now proceed to the question of how changes in the growth matrix affect the growth of the population. This is sometimes called trajectory sensitivity.

2.2.2 Sensitivity of the population trajectory. Recall the population growth equation (50):

$$\{K^{(t+1)}\} = G\{K^{(t)}\}.$$

Since G is assumed to be constant over time, the population distribution at time t_1 is given by

$$\{K^{(t_1)}\} = G^{t_1 - t_0} \{K^{(t_0)}\}.$$

I shall assume that the change in the growth matrix occurs at t_0 . Without loss of generality, t_0 may be set equal to zero, and t_1 equal to t . Then

$$\{K^{(t)}\} = G^t \{K^{(0)}\}.$$

The sensitivity of $\{K^{(t)}\}$ to a change in G is

$$\frac{\partial \{K^{(t)}\}}{\partial \langle G \rangle} = \frac{\partial [G^t]}{\partial \langle G \rangle} \{K^{(0)}\}.$$

The sensitivity of G^t to a change in $\langle G \rangle$ is given by equation (A24). Applying this result yields

$$\frac{\partial \{K^{(t)}\}}{\partial \langle G \rangle} = \sum_{i=0}^{t-1} G^i J G^{t-1-i} \{K^{(0)}\}. \quad (65)$$

A problem which is related to the computation of the sensitivity of $\{K^{(t)}\}$ to changes in G is to find out under what conditions variations in G do not affect $\{K^{(t)}\}$. Such specific conditions have been derived by Tomović and Vukobratović (1972); they will not be discussed here. This and similar problems of trajectory insensitivity or invariance are receiving increasing attention in system theory and optimal-control theory. For a review of some applications in the social sciences see Erickson and Norton (1973).

The next section addresses the topic of the sensitivity of population growth to changes in the population distribution at a certain point in time. This will be called the analysis of small perturbations around the growth path.

2.2.3 Perturbations around the population-growth path. The impact on $\{K^{(t)}\}$ of a change in $\{K^{(0)}\}$ is very simple in the time-invariant equation system (50). Applying the results of vector differentiation set out in the appendix gives

$$\frac{\partial \{K^{(t)}\}}{\partial \{K^{(0)}\}'} = \frac{\partial [G^t]}{\partial \{K^{(0)}\}'} = G^t, \quad (66)$$

where $\{K^{(0)}\}'$ is the transpose of $\{K^{(0)}\}$.

Equation (66) relates changes in the state vector at time t to changes in the state vector at time zero. If the growth matrix is time-dependent this problem cannot be solved analytically and one must rely on simulation. An illustration of such a situation is when the model incorporates a feedback loop, that is, the growth matrix at time t depends on the state vector at time t . For an application of feedback models to urban analysis see Forrester (1969). Nelson and Kern (1971) have simulated the impact of small perturbations around the trajectory for a Forrester-type of urban model.

3 Conclusion

This paper has been devoted to the problem of sensitivity analysis in multiregional demographic systems. From mathematical demography we know that demographic change may be traced back to changes in age-specific fertility, mortality, and migration rates. To show how the mechanism works has been the subject of this paper.

A set of sensitivity functions have been derived which relate a change in demographic characteristics to a change in the vital rates. A summary of the important functions is given in table 1. The primary purpose of this study was to contribute to the knowledge of spatial population dynamics by presenting a unifying technique of impact assessments. In single-region mathematical demography, ordinary differential calculus is used to perform sensitivity analysis. In multiregional demography, where one is dealing with matrix and vector functions, the application of ordinary calculus is very complicated. Instead, matrix differentiation techniques prove to be very useful. A review of these techniques is given in the appendix. These mathematical tools have been applied to derive analytical expressions for multiregional demographic features, such as life-table statistics and population projection, which represent the impacts of changes in vital rates. The sensitivity functions reveal how each spatial demographic characteristic depends on the age-specific rates and how it reacts to changes in those rates. *Matrix differentiation techniques form a powerful tool for the analysis of structural change in multiregional systems.*

Table 1. Table of sensitivity functions.

Function $A(a)$	Sensitivity function $\frac{\partial A(a)}{\partial \langle M(x) \rangle}$	
$M(a)$	$\begin{cases} 0 & \text{for } a \neq x \\ J & \text{for } a = x \end{cases}$	
$P(a) = [I + \frac{5}{2}M(a)]^{-1}[I - \frac{5}{2}M(a)]$	$\begin{cases} 0 & \text{for } a \neq x \\ -5[I + \frac{5}{2}M(x)]^{-1}J[I + \frac{5}{2}M(x)]^{-1} & \text{for } a = x \end{cases}$	
$l(a) = P(a-5)l(a-5)$	$\begin{cases} 0 & \text{for } a \leq x \\ -l(a)l^{-1}(x+5)[I + \frac{5}{2}M(x)]^{-1}JL(x) & \text{for } a > x \end{cases}$	
$L(a) = \frac{5}{2}[I + P(a)]l(a)$	$\begin{cases} \frac{5}{2}[I + \frac{5}{2}M(x)]^{-1}JL(x) & \text{for } a = x \\ -L(a)l^{-1}(x+5)[I + \frac{5}{2}M(x)]^{-1}JL(x) & \text{for } a > x \end{cases}$	
$e(a) = \left[\sum_{y=a}^z L(y) \right] l^{-1}(a)$	$\begin{cases} 0 & \text{for } a > x \\ -[e(x) - \frac{5}{2}I][I - \frac{5}{2}M(x)]^{-1}JL(x)l^{-1}(a) & \text{for } a \leq x \end{cases}$	
$\bar{e}(a) = \left[\sum_{y=a}^z L(y) \right] \bar{l}^{-1}(a)$	$\begin{cases} -e(a)l^{-1}(x)[I - \frac{5}{2}M(x)]^{-1}JL(x)\bar{l}^{-1}(a) & \text{for } a > x \\ -\bar{l}^{-1}(a)Q\bar{l}^{-1}(a) & \text{for } a = x \\ -[e(x) - \frac{5}{2}I][I - \frac{5}{2}M(x)]^{-1}JL(x)\bar{l}^{-1}(a) & \text{for } a < x \end{cases}$	
$S(x) = [I + \frac{5}{2}M(x+5)]^{-1}[I - \frac{5}{2}M(x)]$	$\begin{cases} -\frac{5}{2}[I + \frac{5}{2}M(x+5)]^{-1}J & \text{for } a = x \\ -\frac{5}{2}[I + \frac{5}{2}M(x)]^{-1}JS(x-5) & \text{for } a = x-5 \\ 0 & \text{for } a > x \text{ or } a < x-5 \end{cases}$	

Acknowledgement. This paper is based on my PhD dissertation entitled *Analytics of Multiregional Population Distribution Policy* (1976) submitted to Northwestern University, Evanston, Illinois. I am most grateful to Professor A Rogers for his guidance and close cooperation.

References

- Bellman R, 1970 *Introduction to Matrix Analysis, Second Edition* (McGraw-Hill, New York)
- Coale A J, 1972 *The Growth and Structure of Human Populations: A Mathematical Investigation* (Princeton University Press, Princeton, NJ)
- Coale A J, Demeny P, 1966 *Regional Model Life Tables and Stable Populations* (Princeton University Press, Princeton, NJ)
- Dwyer P S, 1967 "Some applications of matrix derivatives in multivariate analysis" *Journal of the American Statistical Association* 62 607-625
- Dwyer P S, MacPhail M S, 1948 "Symbolic matrix derivatives" *Annals of Mathematical Studies* 16 517-534
- Emre E, Hüseyin Ö, 1975 "Generalization of Leverrier's algorithm to polynomial matrices of arbitrary degree" *IEEE Transactions on Automatic Control* 20 136
- Erickson D L, Norton F E, 1973 "Application of sensitivity constrained optimal control to national economic policy formulation" in *Control and Dynamic Systems: Advances in Theory and Applications* Ed. C T Leondes (Academic Press, New York) pp 132-237
- Forrester J W, 1969 *Urban Dynamics* (MIT Press, Cambridge, Mass)
- Goodman L A, 1969 "The analysis of population growth when the birth and death rates depend upon several factors" *Biometrics* 25 659-681
- Goodman L A, 1971 "On the sensitivity of the intrinsic growth rate to changes in the age-specific birth and death rates" *Theoretical Population Biology* 2 339-354
- Keyfitz N, 1968 *Introduction to the Mathematics of Population* (Addison-Wesley, Reading, Mass)
- Keyfitz N, 1971 "Linkages of intrinsic to age-specific rates" *Journal of the American Statistical Association* 66 275-281
- Lancaster P, 1969 *Theory of Matrices* (Academic Press, New York)
- Nelson N, Kern F, 1971 "Perturbation and sensitivity analysis of an urban model" in *Conference on Decision and Control* Ed. J Tou, Department of Electrical Engineering, University of Florida, Gainesville, Fla, pp 283-288
- Neudecker N, 1969 "Some theorems on matrix differentiation with special reference to Kronecker matrix products" *Journal of the American Statistical Association* 64 953-963
- Newbery A C, 1974 "Numerical analysis" in *Handbook of Applied Mathematics* Ed. C E Pearson (Van Nostrand Reinhold, New York) pp 1002-1057
- Preston S H, 1974 "Effect of mortality change on stable population parameters" *Demography* 11 119-130
- Rees P H, Wilson A G, 1977 *Spatial Population Analysis* (Edward Arnold, London)
- Rogers A, 1973 "The mathematics of multiregional demographic growth" *Environment and Planning* 5 3-29
- Rogers A, 1971 *Matrix Methods in Urban and Regional Analysis* (Holden-Day, San Francisco)
- Rogers A, 1975 *Introduction to Multiregional Mathematical Demography* (John Wiley, New York)
- Rogers A, Ledent J, 1976 "Increment-decrement life tables: a comment" *Demography* 13 287-290
- Rogers A, Willekens F, 1976 "Spatial population dynamics" *Papers of the Regional Science Association* 36 3-34
- Schoen R, 1975 "Constructing increment-decrement life tables" *Demography* 12 313-324
- Tomović R, Vukobratović M, 1972 *General Sensitivity Theory* (American Elsevier, New York)
- Willekens F J, 1976 "Sensitivity analysis" *IIASA Research Memorandum RM-76-49* International Institute for Applied Systems Analysis, Schloss Laxenburg, Laxenburg, Austria

APPENDIX

Matrix differentiation techniques

The purpose of this appendix is to provide the necessary mathematical tools to perform sensitivity analysis of structural change in multiregional demographic systems. The basic notion is that of matrix differentiation. Neudecker (1969, page 953) defines matrix differentiation as the procedure of finding partial derivatives of the elements of a matrix function with respect to the elements of the argument matrix. Although not much has been written on matrix differentiation and the technique is not covered in most textbooks on matrix algebra, this appendix is not intended to be complete. It covers only the techniques applied in this paper and some direct extensions. It is mainly based on the papers by Dwyer and MacPhail (1948) and Dwyer (1967).

Let Y be a $P \times Q$ matrix with elements y_{ij} and let X be an $M \times N$ matrix with elements x_{kl} . Dwyer (1967) makes a distinction between the position of an element in the matrix and its value. The symbol $\langle X \rangle_{kl}$ is used to indicate a specific k, l -element of X . Its scalar value is x_{kl} . Less formally, $\langle X \rangle_{kl}$ may be replaced by $\langle X \rangle$. Therefore $\langle X \rangle$ is an arbitrary element of the matrix X . As in conventional notation, X' denotes the transpose of X , and X^{-1} is the inverse of X .

The relevant results of matrix calculus are given below. To introduce some notation, I start out with the differentiation of a matrix with respect to its elements. This is followed by the differentiation of a matrix with respect to a scalar, and the differentiation of a scalar function with respect to a matrix. The most important scalar function is the determinant. The tools provided in the section on the differentiation of matrix products are frequently used in performing sensitivity analysis of multiregional systems. Also of great importance is the derivative of the inverse. The next section gives some chain rules of matrix differentiation. Vector calculus and matrix calculus are closely related, since a vector is a matrix with only one row or one column. The formulae for vector differentiation, however, have a different appearance and are less complex. Therefore a separate section will be devoted to vector differentiation.

A1 *Differentiation of a matrix with respect to its elements*

The derivative of a matrix X with respect to the element $\langle X \rangle_{kl}$ is

$$\frac{\partial X}{\partial \langle X \rangle_{kl}} = J_{kl}, \quad (A1)$$

where J_{kl} denotes an $M \times N$ matrix with zero elements everywhere except for a unit element in the k th row and l th column.

Similarly

$$\frac{\partial X'}{\partial \langle X \rangle_{kl}} = J'_{lk}, \quad (A2)$$

where J'_{lk} is an $N \times M$ matrix with all elements zero except for a unit element in the l th row and k th column.

Instead of considering the derivative of a matrix with respect to an element, one may also consider the derivative of a matrix element with respect to the matrix:

$$\frac{\partial \langle Y \rangle_{ij}}{\partial Y} = K_{ij}, \quad (A3)$$

where K_{ij} is a $P \times Q$ matrix with zeros everywhere except for a unit element in the i th row and j th column.

Similarly

$$\frac{\partial \langle Y \rangle_{ij}}{\partial Y'} = K'_{ji} . \quad (\text{A4})$$

For convenience the subscripts will be dropped. For example, $\langle X \rangle$ will denote an arbitrary element of X , and J will denote a matrix with all elements zero except for a unit element in the appropriate place, as determined by the location of $\langle X \rangle$.

A2 Differentiation of a matrix with respect to a scalar and of a scalar with respect to a matrix

Let $Y(a)$ be a matrix function of the scalar a . The derivative

$$\frac{\partial Y(a)}{\partial a} \quad (\text{A5})$$

is a matrix with elements $\partial y_{ij}/\partial a$. Each element of $Y(a)$ is differentiated.

The derivative of a scalar function with respect to a matrix is denoted by

$$\frac{\partial f(X)}{\partial X} , \quad (\text{A6})$$

and is a matrix with elements

$$\frac{\partial f(X)}{\partial \langle X \rangle_{ij}} . \quad (\text{A7})$$

Two important scalar functions are considered: the determinant and the trace. I begin by making the assumption that X is a square matrix.

A2.1 Determinant. The determinant of the square matrix X can be evaluated in terms of the cofactors of the elements of the i th row (Rogers, 1971, page 81):

$$|X| = x_{i1} X_{i1}^c + x_{i2} X_{i2}^c + \dots + x_{iN} X_{iN}^c .$$

It can easily be seen that

$$\frac{\partial |X|}{\partial \langle X \rangle_{ij}} = X_{ij}^c , \quad (\text{A8})$$

where X_{ij}^c is the cofactor of the element x_{ij} , and

$$\frac{\partial |X|}{\partial X} = \text{cof} X = [\text{adj} X]' ,$$

where $\text{cof} X$ is the matrix of cofactors, and $\text{adj} X$ is the adjoint matrix of the matrix X . But if X is nonsingular,

$$\text{cof} X = |X| [X']^{-1} . \quad (\text{A9})$$

Equation (A8) may be written as

$$\frac{\partial |X|}{\partial X} = |X| [X']^{-1} . \quad (\text{A10})$$

This formula is well-known in matrix theory and can also be found in Bellman (1970, page 182).

It should be noted that if X is symmetric, then

$$\frac{\partial |X|}{\partial \langle X \rangle_{ij}} = \begin{cases} 2X_{ij}^c & \text{for } i \neq j , \\ X_{ij}^c & \text{for } i = j . \end{cases} \quad (\text{A11})$$

A2.2 Trace. The trace of the square matrix X is the sum of its diagonal elements, and

$$\frac{\partial \text{tr}(X)}{\partial \langle X \rangle_{ij}} = \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j, \end{cases} \quad (\text{A12})$$

with

$$\frac{\partial \text{tr}(X)}{\partial X} = I,$$

where I is the identity matrix.

A3 Differentiation of matrix products

Let U and V be two matrix functions of the matrix X . The derivative of their product, $Y = UV$, with respect to $\langle X \rangle$ is

$$\frac{\partial Y}{\partial \langle X \rangle} = \frac{\partial [UV]}{\partial \langle X \rangle} = \frac{\partial U}{\partial \langle X \rangle} V + U \frac{\partial V}{\partial \langle X \rangle}. \quad (\text{A13})$$

The derivative of a product of three matrices is

$$\frac{\partial Y}{\partial \langle X \rangle} = \frac{\partial [UVW]}{\partial \langle X \rangle} = \frac{\partial U}{\partial \langle X \rangle} VW + U \frac{\partial V}{\partial \langle X \rangle} W + UV \frac{\partial W}{\partial \langle X \rangle}. \quad (\text{A14})$$

These general formulae may be applied to various cases—some cases of interest are listed below. The matrices A and B are constant, that is, independent of X . The matrices J and K are as defined in section A1.

$$Y = \begin{cases} AX, \\ XB, \\ X'B, \\ XX, \\ X'X, \\ AXB, \\ XXX, \\ AXA^{-1}, \end{cases} \quad \text{then } \frac{\partial Y}{\partial \langle X \rangle} = \begin{cases} AJ, \\ JB, \\ J'B, \\ JX + XJ, \\ J'X + X'J, \\ AJB, \\ JXX + XJX + XXJ, \\ AJA^{-1}. \end{cases} \quad (\text{A15})$$

$$XB, \quad (\text{A16})$$

$$X'B, \quad (\text{A17})$$

$$XX, \quad (\text{A18})$$

$$X'X, \quad (\text{A19})$$

$$AXB, \quad (\text{A20})$$

$$XXX, \quad (\text{A21})$$

$$AXA^{-1}, \quad (\text{A22})$$

The derivative of the power of a square matrix can readily be computed using these formulae, thus

$$\frac{\partial [X^n]}{\partial \langle X \rangle} = JX^{n-1} + \sum_{s=1}^{n-2} X^s JX^{n-1-s} + X^{n-1} J, \quad (\text{A23})$$

or, if we write $X^0 = I$, then

$$\frac{\partial [X^n]}{\partial \langle X \rangle} = \sum_{s=0}^{n-1} X^s JX^{n-1-s}. \quad (\text{A24})$$

The derivative of an inverse follows. By definition

$$XX^{-1} = I.$$

Therefore

$$\frac{\partial [XX^{-1}]}{\partial \langle X \rangle} = \frac{\partial I}{\partial \langle X \rangle} = 0,$$

but also

$$\frac{\partial[XX^{-1}]}{\partial\langle X \rangle} = \frac{\partial X}{\partial\langle X \rangle} X^{-1} + X \frac{\partial[X^{-1}]}{\partial\langle X \rangle}.$$

It follows therefore that

$$\frac{\partial[X^{-1}]}{\partial\langle X \rangle} = -X^{-1} J X^{-1}. \quad (\text{A25})$$

One application of this result is

$$\frac{\partial X A X^{-1}}{\partial\langle X \rangle} = J A X^{-1} - X A X^{-1} J X^{-1}. \quad (\text{A26})$$

So far I have considered the derivative $\partial Y / \partial \langle X \rangle$, where Y is a matrix product and $\langle X \rangle$ is an arbitrary element of X . The result is a matrix of partial derivatives. But what is the formula for $\partial Y / \partial X$, where X represents the full matrix? This question has been studied by Neudecker (1969). Its solution involves the transformation of a matrix into a vector and the use of Kronecker products. For example, let $Y = AXB$ and suppose that one is interested in the derivative of Y with respect to X .

If Y is of order $P \times Q$ define the PQ column vector, $\text{vec} Y$ (denoted this way to distinguish it from the vector $\{Y\}$), by

$$\text{vec} Y = \begin{bmatrix} \{Y_{\cdot 1}\} \\ \{Y_{\cdot 2}\} \\ \vdots \\ \{Y_{\cdot Q}\} \end{bmatrix}.$$

In a similar way one can construct $\text{vec} X$. Neudecker shows that

$$\text{vec}(AXB) = [B' \otimes A] \text{vec} X, \quad (\text{A27})$$

where $\otimes$ denotes the Kronecker product. Equation (A27) may be differentiated by making use of the formulae for vector differentiation:

$$\frac{\partial \text{vec}[AXB]}{\partial \text{vec} X} = [B' \otimes A]'$$

Since the transpose of a Kronecker product is the Kronecker product of the transposes, we have⁽³⁾

$$\frac{\partial \text{vec}[AXB]}{\partial \text{vec} X} = B \otimes A'. \quad (\text{A28})$$

I shall not explore the various formulae for $\partial Y / \partial X$ further since they are not explicitly used in this paper.

A4 Chain rules of differentiation

Let $f(Y)$ be a scalar function of Y and let Y be a matrix function of X . Then

$$\frac{\partial f(Y)}{\partial\langle X \rangle} = \sum_{kl} \frac{\partial f(Y)}{\partial\langle Y \rangle_{kl}} \frac{\partial\langle Y \rangle_{kl}}{\partial\langle X \rangle}, \quad (\text{A29})$$

$$\frac{\partial f(Y)}{\partial\langle X \rangle} = \text{tr} \left[\frac{\partial f(Y)}{\partial Y} \frac{\partial Y'}{\partial\langle X \rangle} \right]. \quad (\text{A30})$$

⁽³⁾ For an exposition of the properties of Kronecker products or direct products see Lancaster (1969, pages 256–259).

If $\mathbf{Y}$ is a matrix function of a scalar a , that is, $\mathbf{Y}(a)$, the formula becomes

$$\frac{\partial f(\mathbf{Y})}{\partial a} = \text{tr} \left[\frac{\partial f(\mathbf{Y})}{\partial \mathbf{Y}} \frac{\partial \mathbf{Y}'}{\partial a} \right]. \quad (\text{A31})$$

Consider also the derivative

$$\frac{\partial f(\mathbf{Y})}{\partial \mathbf{X}} = \sum_{kl} \frac{\partial f(\mathbf{Y})}{\partial \langle \mathbf{Y} \rangle_{kl}} \frac{\partial \langle \mathbf{Y} \rangle_{kl}}{\partial \mathbf{X}}. \quad (\text{A32})$$

Several interesting applications arise. For example, let $f(\mathbf{Y}) = |\mathbf{X} - \lambda \mathbf{I}|$, where $\mathbf{X}$ may be the population-growth matrix. Then

$$\begin{aligned} \frac{\partial |\mathbf{X} - \lambda \mathbf{I}|}{\partial \langle \mathbf{X} \rangle} &= \text{tr} \left[\frac{\partial |\mathbf{X} - \lambda \mathbf{I}|}{\partial [\mathbf{X} - \lambda \mathbf{I}]} \frac{\partial [\mathbf{X} - \lambda \mathbf{I}']}{\partial \langle \mathbf{X} \rangle} \right], \\ \frac{\partial |\mathbf{X} - \lambda \mathbf{I}|}{\partial \langle \mathbf{X} \rangle} &= |\mathbf{X} - \lambda \mathbf{I}| \text{tr} \left[[(\mathbf{X} - \lambda \mathbf{I})']^{-1} \mathbf{J}' \right] \\ &= \text{tr} \left[[\text{cof}(\mathbf{X} - \lambda \mathbf{I})] \mathbf{J}' \right], \end{aligned} \quad (\text{A33})$$

and

$$\frac{\partial |\mathbf{X} - \lambda \mathbf{I}|}{\partial \mathbf{X}} = \sum_{kl} \frac{\partial |\mathbf{X} - \lambda \mathbf{I}|}{\partial \langle [\mathbf{X} - \lambda \mathbf{I}] \rangle_{kl}} \frac{\partial \langle [\mathbf{X} - \lambda \mathbf{I}] \rangle_{kl}}{\partial \mathbf{X}} = \sum_{kl} |\mathbf{X} - \lambda \mathbf{I}| [\mathbf{X} - \lambda \mathbf{I}]_{kl}'^{-1} \mathbf{J}'_{kl}, \quad (\text{A34})$$

$$\frac{\partial |\mathbf{X} - \lambda \mathbf{I}|}{\partial \mathbf{X}} = |\mathbf{X} - \lambda \mathbf{I}| [\mathbf{X} - \lambda \mathbf{I}]'^{-1} = \text{cof}[\mathbf{X} - \lambda \mathbf{I}], \quad (\text{A35})$$

where $\text{cof}[\mathbf{X} - \lambda \mathbf{I}]$ is the cofactor matrix of $[\mathbf{X} - \lambda \mathbf{I}]$.

If $\mathbf{Y}(r)$ is a function of the scalar r , then

$$\frac{\partial |\mathbf{Y}(r)|}{\partial r} = \text{tr} \left[\frac{\partial |\mathbf{Y}(r)|}{\partial [\mathbf{Y}(r)]} \frac{\partial [\mathbf{Y}(r)]'}{\partial r} \right] = \text{tr} \left[|\mathbf{Y}(r)| [\mathbf{Y}(r)]'^{-1} \frac{\partial [\mathbf{Y}(r)]'}{\partial r} \right],$$

and, since $\text{tr} \mathbf{AB} = \text{tr} [\mathbf{AB}]' = \text{tr} \mathbf{B}' \mathbf{A}'$, then

$$\frac{\partial |\mathbf{Y}(r)|}{\partial r} = |\mathbf{Y}(r)| \text{tr} \left[\frac{\partial [\mathbf{Y}(r)]}{\partial r} [\mathbf{Y}(r)]^{-1} \right]. \quad (\text{A36})$$

Formula (A36) is not only of interest in a study of the sensitivity of the determinant of a polynomial matrix, but is also useful in the computation of the determinant, as shown by Emre and Hüseyin (1975, page 136). An application of equation (A36) that is relevant is

$$\frac{\partial |\mathbf{A} - \lambda \mathbf{I}|}{\partial \lambda} = -|\mathbf{A} - \lambda \mathbf{I}| \text{tr} [\mathbf{A} - \lambda \mathbf{I}]^{-1}. \quad (\text{A37})$$

This formula can also be found in Newbery (1974, page 1016). Finally consider the application where $f(\mathbf{Y}) = \text{tr} [\mathbf{AXB}]$, whence

$$\frac{\partial f(\mathbf{Y})}{\partial \mathbf{X}} = \frac{\partial \text{tr} [\mathbf{AXB}]}{\partial \mathbf{X}} = \mathbf{A}' \mathbf{B}'. \quad (\text{A38})$$

A5 Vector differentiation

Vectors may be considered as matrices with only one row or one column, and the rules for matrix differentiation may be applied. But the derivative of a vector or of a vector equation has a simpler form than the matrix analogue. It is therefore worthwhile to list the formulae for vector differentiation separately. Two cases are considered: the derivative of a scalar function with respect to a vector and the derivative of a vector function with respect to a vector.

A5.1 Differentiation of a scalar function with respect to a vector. Consider the general scalar function $f(\{x\})$, where $\{x\}$ is the argument vector. Some relevant formulations of $f(\{x\})$ and their derivatives are listed below.

$$f(\{x\}) = \begin{cases} \{a\}'\{x\}, & \text{then } \frac{\partial f(\{x\})}{\partial \{x\}} = \{a\}, & \text{(A39)} \\ \{x\}'\{x\}, & 2\{x\}, & \text{(A40)} \\ \{x\}'A\{x\}, & A\{x\} + A'\{x\}. & \text{(A41)} \end{cases}$$

A5.2 Differentiation of a vector function with respect to a vector. Let $\{f(\{x\})\}$ denote a column vector of scalar functions $f_i(\{x\})$, where $\{x\}$ is the argument vector and $\{f(\{x\})\}$ represents a system of equations. For example, let $\{f(\{x\})\}$ be a system of linear equations in $\{x\}$, that is, $\{f(\{x\})\} = A\{x\}$, then

$$\frac{\partial A\{x\}}{\partial \{x\}_i} = \{a_i\}, \quad \text{(A42)}$$

where $\{a_i\}$ is the i th column of A .

The derivatives of $\{f(\{x\})\}$ with respect to all the elements of the argument vector form a matrix if the argument vector is a row vector. For example

$$\frac{\partial A\{x\}}{\partial \{x\}'} = A. \quad \text{(A43)}$$

The determinant $|\partial\{f(\{x\})\}/\partial\{x\}'|$ is known as the Jacobian or functional determinant.

Corresponding to the chain rule of matrix differentiation, one may formulate the chain rule of vector differentiation. Let $\{x\}$, $\{y\}$, and $\{z\}$ be vectors. It can be shown that

$$\frac{\partial \{y\}}{\partial \{x\}'} = \frac{\partial \{y\}}{\partial \{z\}'} \frac{\partial \{z\}}{\partial \{x\}'}. \quad \text{(A44)}$$