# An Improved Method of Computing Multistate Survivorship Proportions for the Terminal Age Groups 

Just, P. and Liaw, K.-L.

IIASA Working Paper
WP-83-065

July 1983

Just, P. and Liaw, K.-L. (1983) An Improved Method of Computing Multistate Survivorship Proportions for the Terminal Age Groups. IIASA Working Paper. WP-83-065 Copyright © 1983 by the author(s). http://pure.iiasa.ac.at/2247/

Working Papers on work of the International Institute for Applied Systems Analysis receive only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work. All rights reserved. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage. All copies must bear this notice and the full citation on the first page. For other purposes, to republish, to post on servers or to redistribute to lists, permission must be sought by contacting repository@iiasa.ac.at

# Working Paper 

```
AN IMPROVED METHOD OF COMPUTING
MULTISTATE SURVIVORSHIP PROPORTIONS
FOR THE TERMINAL AGE GROUPS
Peer Just
Kao-Lee Liaw
July 1983
WP-83-65
```


# AN IMPROVED METHOD OF COMPUTING <br> MULTISTATE SURVIVORSHIP PROPORTIONS <br> FOR THE TERMINAL AGE GROUPS 

Peer Just
Kao-Lee Liaw

July 1983
WP-83-65

Working Papers are interim reports on work of the International Institute for Applied Systems Analysis and have received only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute or of its National Member Organizations.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS A-2361 Laxenburg, Austria

## FOREWORD

Low fertility levels in IIASA countries are creating aging populations whose demands for health care and income maintenance (social security) will increase to unprecedented levels, thereby calling forth policies that will seek to promote increased family care and worklife flexibility. The new Population Program will examine current patterns of population aging and changing lifestyles in IIASA countries, project the needs for health and income support that such patterns are likely to generate during the next several decades, and consider alternative family and employment policies that might reduce the social costs of meeting these needs.

Multiregional and multistate demographic methods are being increasingly adopted in applied population research. This has stimulated a reexamination of a number of aspects of the basic methodology. One of these is the proper survivorship of the last two age groups in a population and an appropriate calculation of the corresponding survivorship proportions. This problem is particularly relevant in studies of the future age composition of the elderly population. This paper suggests several alternatives to improve the projection of the population in the last two age groups. The so-called preferred approach recommended by the authors may be readily integrated into the standard framework of multistate projections.

Andrei Rogers
Leader, Population Program

## ABSTRACT

The aging of populations is a phenomenon which has become an important research topic. Demographers, however, have given inadequate attention to the projection of the number of old people and their future age composition. This paper shows that the conventional method for estimating the survivorship proportions of the very old tends to produce misleading results with respect to the size and composition of the aged. Several alternatives are suggested here to overcome these problems. An empirical example is used to point out the problems of the conventional approach and to evaluate the suggested improvements.

CONTENTS

INTRODUCTION

1. THE PROBLEM OF THE CONVENTIONAL APPROACH
2. THE DISAGGREGATION APPROACH
3. THE PREFERRED APPROACH
3.1 Linear Approach
3.2 Exponential Approach
4. COMPARISON OF THE PROJECTION RESULTS

CONCLUSION
APPENDIX A: DATA FOR SWEDEN 1974, FEMALES, STOCKHOLM AND THE REST OF THE COUNTRY

APPENDIX B: PROJECTED GROWTH RATES OF SWEDISH FEMALES 1975-2024

REFERENCES

AN IMPROVED METHOD OF COMPUTING MULTISTATE SURVIVORSHIP PROPORTIONS<br>FOR THE TERMINAL AGE GROUPS

## INTRODUCTION

In multistate demography, an arbitrary convention has been used to obtain the submatrices of survivorship (and outmigration) proportions for the two oldest age groups from the submatrices representing the age compositions of the multiregional life table (stationary) population. The convention mismatches two subpopulations and causes problems that have become particularly serious when one tries to construct a single-year projection model for a population with a high expectation of life. Since the convention was inherited from the common practice in single-state mathematical demography, the problem has been hidden or ignored for a long time. As population aging has become an important research topic in recent years, we find it timely to focus in this paper on pointing out the problems of this convention and proposing alternative solutions. To illustrate our arguments we shall draw on a multiregional example using Swedish data.

In section 1 , we identify the problems of the conventional approach. Section 2 presents an alternative approach that reduces the severity of mismatching subpopulations by further splitting up the last age group into smaller ones. Section 3 shows our preferred alternative approach that yields sensible and interpretable results. Section 4 uses the 1974 Swedish data shown in Appendix A to evaluate the quality of the projection results that are generated by the different approaches. The last section summarizes the main points.

1. THE PROBLEM OF THE CONVENTIONAL APPROACH

Consider a population with $N(\geq 1)$ regions and $w+1$ age groups: $(0, h),(h, 2 h), \ldots,(x, x+h), \ldots,(w h, \infty)$. Let $h \stackrel{L}{L}$ be an $\mathrm{N} \times \mathrm{N}$ matrix representing the place-of-residence-by-place-of-birth population distribution of the $x$ age group (i.e., between exact ages $x$ and $x+h$ ) in the multiregional life table population. Specifically, the element in the ith row and the jth column is the number of residents at exact ages between $x$ and $x+h$ in region $j$ who were born in region $i$. Also let $\underset{\sim}{S} \underset{\sim}{S}$ be the $N \times N$ submatrix of survivorship (and outmigration) proportions, where the element in row $j$ and column $i$ represents the proportion of the individuals in the ith region and the $x+h$ age group who will reside in the jth region $h$ years later. The diagonal elements of $h_{\sim}^{S}$ are surviving stayer proportions, whereas the off-diagonal elements of ${ }_{h \sim x}^{S}$ are surviving outmigration proportions. Without foreign migration, the constraint

$$
\begin{equation*}
0 \leq h^{i j} \leq 1 \tag{1}
\end{equation*}
$$

must be satisfied for each element $h^{S_{x}^{i j}}$ of $h S_{x}$
By definition, we have

$$
\begin{equation*}
h \underset{\sim}{L} \underset{\sim}{\mathrm{~L}} \mathrm{~h} \underset{\sim}{\mathrm{~S}} \underset{\mathrm{x}}{\mathrm{~h}} \underset{\sim}{\mathrm{~L}} \mathrm{x} \quad \text { for } \mathrm{x}=0, h, 2 h, \ldots,(w-1) h \tag{2}
\end{equation*}
$$

and

The conventional approach is (i) to solve equation (2) for $h \underset{\sim}{S}$ for the first $w-1$ age groups, (ii) to solve for $h \underset{\sim}{S}(w-1) h$ from the inappropriate equation

$$
\begin{equation*}
\infty_{\sim}^{L}{ }_{\sim}^{L} h(w-1) h(w) h(w-1) h \tag{4}
\end{equation*}
$$

and (iii) to set ${ }_{h}{\underset{\sim}{w h}}$ to zero. The convention is found in most books that deal with uniregional as well as multiregional population projections (e.g., Keyfitz (1968 and 1977) and Rogers (1975)).

A problem of the conventional approach is that the last two $h_{\sim}^{S}{ }_{w h}$ do not contain any demographically meaningful quantities. For example, in a single-region life table based on the 1974 Swedish female mortality data for 86 age groups ( $0,1,2, \ldots, 85+$ ), we found that ${ }_{1} \mathrm{~L}_{84}=0.36424$ and ${ }_{\infty} \mathrm{L}_{85}=1.79784$. According to the conventional approach, we get ${ }_{1} S_{84}={ }_{\infty} L_{85} 1_{1}^{L_{84}^{-1}}=4.93592$ and $1^{S}{ }_{85}=0$. These are, of course, meaningless survivorship proportions. By disaggregating the female population into two regions (Stockholm and the rest of Sweden), and incorporating interregional migration information (see Appendix A), the conventional approach yields

$$
{ }_{1} S_{\sim 84}=\left[\begin{array}{ll}
5.193331 & 0.016234  \tag{5}\\
0.057824 & 4.856811
\end{array}\right]
$$

and

$$
{ }_{1}{\underset{\sim}{\sim}}_{85}=\left[\begin{array}{ll}
0 & 0  \tag{6}\\
0 & 0
\end{array}\right]
$$

Though they happen to be between zero and unity, the off-diagonal elements of ${ }_{1 \sim 84}$ are much too large to represent the surviving outmigration proportions of the relevant subpopulations. Furthermore, the diagonal elements of $1 \mathrm{~S}_{84}$ fall far above the upper bound of the constraint that must hold for the definition of surviving stayer proportions. In general, these elements tend to be farther above the upper bound, when the expectation of life at age wh is high, or when both $w$ and $h$ are small. Of course, one must not be serious about interpreting equation (6). With mortality and migration information available for all age groups, it is regrettable that the conventional approach ends up with meaningless submatrices of survivorship and outmigration proportions for the last two age groups. If the last age group of the raw data was $65+$, then the $h \underset{\sim}{S}$ submatrices would contain no useful information about the post-retirement subpopulations.

One may argue that the submatrices $h_{\sim}^{S}$ x are constructed solely for the purpose of projection. Therefore, as long as the resulting projections are reasonably good, one need not care about interpretations. Unfortunately, the projections can turn out to be truly bad, particularly when a single-year projection model is used. Consider again the aforementioned biregional Swedish population. Applying the submatrix $1_{\sim}^{S_{84}}$ in equation (5) to the 1974 subpopulation in the 84 th age group gives

$$
{ }_{\infty}^{\mathrm{K}} 85(1975)={ }_{1 \sim 84}^{S_{1}} \mathcal{L}_{\sim 84}^{\mathrm{K}}(1974)=\left[\begin{array}{ll}
5.193331 & 0.016234  \tag{7}\\
0.057824 & 4.856811
\end{array}\right]\left[\begin{array}{r}
2132 \\
10169
\end{array}\right]=\left[\begin{array}{l}
11237 \\
49512
\end{array}\right]
$$

Comparing this vector with the observed initial subpopulation in the 85+ age group

$$
\infty_{\sim}^{K}{ }_{\sim 5}(1974)=\left[\begin{array}{r}
8789 \\
40714
\end{array}\right]
$$

we see that the conventional approach implies an annual growth rate for the last age group of $28 \%$ in the Stockholm region and $22 \%$ in the rest of Sweden. Lumping the two regions together gives an annual growth rate for this age group of $23 \%$. These phenomenally high growth rates mean that the projection is totally misleading as far as the last age group is concerned. We will show later that the growth rate of the whole population generated by the conventional approach is also distorted quite badly.

In short, the conventionally constructed ${ }_{h}{\underset{\sim}{X}}^{x}$ submatrices for the last two age groups are suitable neither for interpretation nor for projection. Therefore, we shall consider a few alternative approaches that may produce more satisfactory results.

## 2. THE DISAGGREGATION APPROACH

One can partly overcome the aforementioned problems by disaggregating the last open-ended age group into an arbitrary number (J) of age groups: $\quad h_{\sim}^{\hat{K}}{ }_{w h} h_{\sim}^{\hat{K}}(w+1) h, \cdots{ }_{\infty} \hat{K}_{(w+J)}^{\hat{K}} h^{\prime}$. Any reasonable interpolation procedure such as a third-degree spline can be applied, as long as the constraint

$$
\begin{equation*}
\infty_{\sim}^{K} w h=\sum_{j=0}^{J-1} \underset{\sim}{\hat{K}}(w+j) h+\infty_{\sim}^{\hat{K}}(w+J) h \tag{8}
\end{equation*}
$$

is satisfied. Note that the age-composition of these new groups will not have a lasting effect on the projected population because the people in these age groups at the initial time will be totally replaced by the incoming cohorts after Jh years.

Under the assumption that the observed mortality and mobility rates of the last age group apply to all additional age groups, the submatrices of survivorship proportions can be computed as follows. The submatrix $h \underset{\sim}{S}(w-1) h$ in the conventional approach, expressed in terms of occurrence/exposure rates, is given by (Ledent 1978)

$$
\begin{equation*}
h \underset{\sim}{S}(w-1) h=\frac{1}{h} \underset{\sim}{M_{w h}^{-1}}\left[\underset{\sim}{I}-\frac{h}{2} \underset{\sim}{M}(w-1) h\right] \tag{9}
\end{equation*}
$$

where $\underset{\sim}{M}$ wh denotes the matrix of observed rates as set out in Willekens and Rogers (1978). This matrix is now replaced by a matrix $h \underset{\sim}{S}(w-1) h$ calculated by

$$
\begin{equation*}
h_{\sim}^{\hat{S}}(w-1) h=\left[\underset{\sim}{I}+\frac{h}{2} \underset{\sim}{M}{ }_{w h}\right]^{-1}\left[\underset{\sim}{I}-\frac{h}{2} \underset{\sim}{M}(w-1) h\right] \tag{10}
\end{equation*}
$$

Assuming constant rates for all ages $w+j$ ( $j=0, \ldots, J-2$ ), the corresponding submatrices of survivorship proportions are given by

$$
\begin{align*}
\mathrm{h}_{\sim}^{\mathrm{S}}(\mathrm{w}+j) \mathrm{h} & =\left[\underset{\sim}{I}+\frac{\mathrm{h}}{2} \underset{\sim}{\mathrm{M}} \mathrm{wh}\right]^{-1}\left[\underset{\sim}{I}-\frac{h}{2} \underset{\sim}{\mathrm{M}} \mathrm{wh}\right] \quad j=0 \text { to } \mathrm{J}-2 \\
& =\mathrm{h}_{\sim}^{P} \mathrm{wh} \tag{11}
\end{align*}
$$

The troublesome submatrix of survivorship proportions $h \sim(w+J+1) h$ for the open-ended age group can now be calculated in two ways. Either by applying the same procedure as for the other additional age groups and setting

$$
\begin{equation*}
h \underset{\sim}{S}(w+J+1) h \stackrel{P}{h} \underset{\sim}{P} h \tag{12}
\end{equation*}
$$

or using the conventional formula for the last survivorship matrix but with identical rates for the last and next to the last age groups. From equation (9) we then get

$$
\begin{align*}
\underset{\sim}{S}(w+J-1) h & =\frac{1}{h} \underset{\sim}{M_{W h}^{-1}}\left[\underset{\sim}{I}-\frac{h}{2} \underset{\sim}{M}\right. \\
& =\frac{1}{h} \underset{\sim}{M} \underset{W h}{M}-\frac{1}{2} \underset{\sim}{I} \tag{13}
\end{align*}
$$

The computation based on equation (12) will further on be referred to as disaggregation approach (I), and the alternative involving equation (13) as disaggregation approach (II). For both alternatives, $h \sim(w+J) h$ is set to zero. The population can now be projected forward by the expanded growth matrix. The structure of this matrix is the same as that for the conventional approach: the $N \times N$ submatrices in the diagonal are all zero.

For the Swedish data in Appendix A, we have disaggregated the $85+$ age group into 11 age groups ( $85,86, \ldots, 94,95+$ ). According to equation (10), we now have

$$
{ }_{1} S_{84}=\left[\begin{array}{ll}
0.862548 & 0.000403  \tag{14}\\
0.002809 & 0.858338
\end{array}\right]
$$

Although it looks much more sensible than the corresponding submatrix generated by the conventional approach (see equation (5)), this submatrix tends to understate the level of survivorship proportions, because in equation (10) the average mortality level of the $85+$ age group is assumed to be applicable to the singleyear age 85.

The submatrices computed according to equation (11) are

$$
{\underset{\sim}{\sim}}_{\sim}^{S}{ }_{85}={ }_{1}^{S}{\underset{\sim}{S}}_{86}=\ldots={ }_{1}{\underset{\sim}{S}}_{93}=\left[\begin{array}{ll}
0.833907 & 0.000472  \tag{15}\\
0.001427 & 0.823267
\end{array}\right]
$$

These submatrices also look sensible, except that the survivorship proportions tend to be understated for ${ }_{1}{\underset{\sim}{S}}_{85}$ and overstated for ${ }_{1}{\underset{\sim}{S}}_{93^{\circ}}$ For the next to the last age group, the disaggregation approach (I) yields

$$
{ }_{1}{\underset{\sim}{\sim}}_{94}=\left[\begin{array}{ll}
0.833907 & 0.000472  \tag{16}\\
0.001427 & 0.823267
\end{array}\right]
$$

whereas the disaggregation approach (II) gives

$$
{ }_{1} \mathrm{~S}_{\sim 94}=\left[\begin{array}{ll}
5.020865 & 0.016090  \tag{17}\\
0.048612 & 4.658370
\end{array}\right]
$$

While equation (16) does not give any information in addition to that of equation (15), equation (17) does not permit any meaningful interpretation. Therefore, the justification for the disaggregation approach lies in the hope of producing less distorted projections only.

## 3. THE PREFERRED APPROACH

Our preferred approach uses equation (2) to find the first w submatrices of survivorship and outmigration proportions and then uses the basic equation (3) to find ${ }_{h}{ }_{\sim}^{S}{ }_{W h}$. The solutions are

$$
\begin{equation*}
\underset{\sim}{S} \underset{\sim}{S}=\sim_{\sim}^{L} x+h \underset{\sim}{L}{ }_{\sim}^{L} \quad \text { for } x=0, h, 2 h, \ldots,(w-1) h \tag{18}
\end{equation*}
$$

and

$$
\begin{equation*}
h S_{\sim}^{S}{ }_{w h}=\left(\infty_{\sim}^{L} w h^{L}-h \underset{\sim}{S}(w-1) h h \underset{\sim}{L}(w-1) h\right) \infty_{\sim}^{L}{ }_{\sim}^{L}-1 \tag{19}
\end{equation*}
$$

To compute the submatrix $h_{\sim}^{S}(w-1) h$ according to equation (18), we need to know the value of $h_{\sim}^{L}{ }_{\sim}^{w h}$, which is different from the submatrix $\infty_{\sim}^{L} \mathrm{~L}_{\mathrm{wh}}$ that is computed in the conventional multiregional life table. To make the distinction perfectly clear, we write

$$
\begin{equation*}
h \underset{\sim}{L} x_{0}^{h} \int_{\sim}^{\ell} \underset{x+\lambda}{ } d x \tag{20}
\end{equation*}
$$

and

$$
\begin{equation*}
\infty_{\sim x}^{L}=\int_{0}^{\infty}{\underset{\sim x}{ }}^{x+\lambda} d x \tag{21}
\end{equation*}
$$

 place-of-birth population distribution at exact age $x+\lambda$ among the survivals of the arbitrarily specified birth cohort in ${\underset{\sim}{~}}_{0} 0^{*}$ *
 us back to the conventional approach.

To compute $h \underset{\sim}{L} x^{\text {a }}$ and $\infty_{\sim}^{L} \underset{\sim}{L}$ from the matrices of the observed mobility and mortality rates $(\underset{\sim}{( })$ ), an assumption about the underlying mathematical model must be made so that the integrals in equations (20) and (21) can be conveniently evaluated. Two well-known alternative assumptions are (i) that $\ell_{\sim}$ is piece-wise linear within the interval $h$, and (ii) that $\underset{\sim}{M}$ is a good approximation of the matrix of instantaneous mobility and mortality

[^0]rates ( $\underset{\sim}{\mu}$ ) which is in turn assumed to be piece-wise constant within the individual age groups. The approaches based on these two assumptions are called the linear and the exponential approaches, respectively. The former approach occasionally generates some nonsensical quantities like a negative outmigration proportion, whereas the latter does not have such a problem. Since it is widely used and does generate satisfactory results when singleyear age groups are used, the linear approach shall be discussed first.

### 3.1 Linear Approach

Assuming a piece-wise linear $\underset{\sim}{\ell}$ x over an integral $h$, equation (18) can be expressed in terms of observed occurrence/exposure rates as shown by equation (10). In particular, the survivorship proportions from the second last to the last age group is given by

$$
\begin{equation*}
h \underset{\sim}{S}(w-1) h=\left[\underset{\sim}{I}+\frac{h}{2} \underset{\sim}{M} w h\right]^{-1}\left[\underset{\sim}{I}-\frac{h}{2} \underset{\sim}{M}(w-1) h\right] \tag{22}
\end{equation*}
$$

In contrast, the submatrix $h_{\sim}^{S}(w-1) h$ in the conventional approach was calculated according to equation (9).

Finally, the proportion surviving within the last age group has to be derived from formula (19). Replacing $h \underset{\sim}{S}(w-1) h$ in (19) by

$$
\begin{equation*}
h \stackrel{S}{\sim}(w-1) h=h \stackrel{L}{\sim}{ }_{\sim}^{L} h \underset{\sim}{\sim}{ }_{\sim}^{L}(w-1) h \tag{23}
\end{equation*}
$$

leads to

$$
\begin{equation*}
h \underset{\sim}{S}{ }_{w h}=\underset{\sim}{I}-h_{\sim}^{L}{ }_{w h} \infty_{\sim}^{L}{\underset{\sim h}{L}}^{-1} \tag{24}
\end{equation*}
$$

Following Willekens and Rogers (1978), we can substitute $\infty_{\infty}{ }_{\sim}^{L} w h$ in (24) by

$$
\begin{equation*}
\infty_{\sim}^{L}{ }_{w h}=\underset{\sim}{M_{w h}^{-1}} \quad \underset{\sim}{-1}{ }^{\ell} \tag{25}
\end{equation*}
$$

and finally get

Some further manipulations* show that the additional submatrix of survivorship proportions $\underset{\sim}{\mathrm{S}} \underset{\mathrm{wh}}{ }$ is given by

$$
\begin{equation*}
\mathrm{h}_{\sim \mathrm{Sh}}^{\mathrm{S}}=\underset{\sim}{I}-\frac{\mathrm{h}}{2}\left[{\underset{\sim}{P}}_{\sim}^{P} \underset{\sim}{ }+\underset{\sim}{I}\right] \underset{\sim}{M} \underset{\sim}{P}{ }_{\sim}^{P} \tag{26}
\end{equation*}
$$

with $\underset{\sim}{\sim}{\underset{\sim}{w h}}$ to be calculated by

$$
\begin{equation*}
\underset{\mathrm{h}}{\mathrm{P}} \mathrm{wh}=\left[\underset{\sim}{I}+\frac{\mathrm{h}}{2} \underset{\sim}{\mathrm{M}} \mathrm{wh}^{-1}\left[\underset{\sim}{I}-\frac{\mathrm{h}}{2} \underset{\sim}{\mathrm{M}} \mathrm{wh}\right]\right. \tag{27}
\end{equation*}
$$

Note that in the conventional approach, the zero matrix $\infty_{\infty} \mathrm{P}$ wh is given but ${ }_{h}{ }_{\sim}^{\sim} W_{h}$ is not computed. It can be shown that $\frac{h}{2}[h \underset{\sim}{P} w h+\underset{\sim}{I}] \stackrel{h}{2}(\underset{\sim}{\ell}(w+1) h+\underset{\sim}{\ell} \underset{\sim}{\ell}) \underset{\sim}{\ell}{ }_{\sim}^{-1}$, which represents the average numbers of person-years lived in various regions during the $h$ years beyond age wh by region of residence at exact age wh. This interpretation makes equation (26) intuitively clear. The survivorship proportions computed by formulas (22) and (26) have now to be placed into the last row of the growth matrix. Since an additional element $h_{\sim}^{S} \underset{\sim}{S}$ is used, the growth matrix is rewritten as
*First replacing $h{\underset{\sim}{w h}}^{\underset{W}{L}}$ by

$$
\mathrm{h} \underset{\sim}{\mathrm{~L}} \mathrm{wh}=\frac{\mathrm{h}}{2}\left[\mathrm{~h}_{\sim}^{\mathrm{P}} \underset{\mathrm{wh}}{ }+\underset{\sim}{I}\right] \underset{\sim}{\ell} \mathrm{wh}
$$

and then using the property (Ledent 1978)

$$
\begin{aligned}
\mathrm{h}_{\sim}^{\mathrm{P}} \mathrm{~Wh} & =\left[\underset{\sim}{I}-\frac{\mathrm{h}}{2} \underset{\sim}{\mathrm{M}}\right. \\
& \left.=\left[\underset{\sim}{I}+\frac{\mathrm{h}}{2} \underset{\sim}{\mathrm{I}} \underset{\sim}{\mathrm{M}}+\frac{\mathrm{h}}{2}\right]_{\sim}^{\mathrm{M}}\right]^{-1}\left[\underset{\sim}{I}-\frac{\mathrm{h}}{2}\right]_{\sim}^{\mathrm{M}} \\
& \left.\mathrm{Mh}^{-1}\right]
\end{aligned}
$$

Applying the linear approach to the Swedish data in Appendix A yields

$$
1_{\sim}^{S}{ }_{\sim}{ }_{84}=\left[\begin{array}{ll}
0.862548 & 0.000403  \tag{29}\\
0.002809 & 0.858338
\end{array}\right]
$$

and

$$
1_{\sim}^{S} 85=\left[\begin{array}{ll}
0.833907 & 0.000472  \tag{30}\\
0.001427 & 0.823267
\end{array}\right]
$$

Both these matrices appear much more sensible than those of the conventional approach. Note that the ${ }_{1}{\underset{\sim}{S}}_{84}$ submatrix does not differ between the preferred linear approach and the disaggregation approaches. The value of the submatrix $1_{\sim}^{S} 85$ of the preferred linear approach is identical to those of $1 \underset{\sim}{S} 85^{\prime}$ 1 ${\underset{\sim}{S}}_{86^{\prime}}$ ..., ${\underset{\sim}{S}}_{93}$ of both disaggregation approaches.

### 3.2 Exponential Approach

The assumptions that the observed matrix of occurrence/ exposure rates $(\underset{\sim}{M})$ is a good approximation of the matrix of the corresponding instantaneous rates $(\underset{\sim}{\mu})$ and that $\mu_{x}$ is piecewise constant within individual age groups imply that the model of the multiregional life table can be written as the differential equation

$$
\begin{equation*}
\dot{\imath}_{\sim+\lambda}=-\underset{\sim}{M}{\underset{\sim}{x}+\lambda}_{\ell}^{l} \quad \text { for } x=0, h, 2 h, \ldots, w h \tag{31}
\end{equation*}
$$

where ${\underset{\sim}{l}}_{x+\lambda}$ is the derivative of ${\underset{\sim}{l}}_{x+\lambda}$ with respect to age, and $\lambda$ is constrained by $0 \leq \lambda<h$ for the first $w$ age groups and by $0 \leq \lambda$ for the last age group. The solution of the differential equation is

$$
\begin{equation*}
{\underset{\sim}{\ell}+\lambda}=e^{-\lambda \underset{\sim}{M}} \underset{\sim}{\ell} \underset{x}{ } \quad \text { for } x=0, h, 2 h, \ldots, w h \tag{32}
\end{equation*}
$$

In other words, the matrix of probabilities of surviving between exact ages $x$ and $x+h$ is simply

$$
\begin{equation*}
\lambda_{\sim}^{P} \underset{\sim}{P}=e^{-\lambda \underset{\sim}{M}} \quad \text { for } x=0, h, 2 h, \ldots, w h \tag{33}
\end{equation*}
$$

Substituting equation (32) and (33) into equations (20) and (21), expanding $e^{-\lambda M} \sim$ into a Taylor series, and then integrating, we get

$$
\begin{aligned}
& =\left(\underset{\sim}{I}-h_{\sim}^{P} X_{\sim}\right) \underset{\sim}{M} \underset{\sim}{\ell}
\end{aligned}
$$

and

Substituting equations (34) and (35) into equations (18) and (19), we get

$$
\begin{align*}
& \text { for } x=0, h, 2 h, \ldots,(w-1) h \tag{36}
\end{align*}
$$

and

The exponential approach starts with computing all $h \underset{\sim}{P}$ from ${\underset{\sim}{x}}^{x}$ and then using equations (36) and (37) to compute all $h_{\sim}^{S} X_{x}$. To obtain a highly accurate result for $\underset{\sim}{P} \underset{x}{ }$, we recommend the following computation formula

$$
\begin{align*}
\underset{\mathrm{h}}{\mathrm{P}}= & {\left.\left[\underset{\sim}{\mathrm{I}}+\frac{1}{2} \underset{\sim}{\mathrm{~h}} \mathrm{~m}_{\mathrm{x}}+\frac{5}{44}\left(\mathrm{hM}_{\sim}\right)^{2}+\frac{1}{66}\left(\mathrm{hM}_{\sim}\right)^{3}+\frac{1}{792}(\underset{\sim}{\mathrm{~h}})^{4}\right)^{4}+\frac{1}{15840}\left(\mathrm{hM}_{\sim \mathrm{x}}\right)^{5}+\frac{1}{665280}\left(\mathrm{hM}_{\sim}\right)^{6}\right]^{-1} } \\
& {\left[\underset{\sim}{I}-\frac{1}{2} \mathrm{hM}_{\sim}+\frac{5}{44}\left(\mathrm{hM}_{\sim}\right)^{2}-\frac{1}{66}\left(\mathrm{hM}_{\sim}\right)^{3}+\frac{1}{792}\left(\mathrm{hM}_{\sim}\right)^{4}-\frac{1}{15840}\left(\mathrm{hM}_{\sim}\right)^{5}+\frac{1}{665280}\left(\mathrm{hM}_{\sim}\right)^{6}\right] } \tag{38}
\end{align*}
$$

This formula was derived by the matrix continued fraction method of electrical engineers (Shieh et al. 1978). For an explanation of the logic underlying the method, see Liaw and Ledent (1980).

Applying the exponential approach to the Swedish data in Appendix A yields

$$
1_{\sim}^{S_{\sim 4}}=\left[\begin{array}{ll}
0.861465 & 0.000408  \tag{39}\\
0.002758 & 0.856931
\end{array}\right]
$$

and

$$
1_{\sim}^{S_{85}}=\left[\begin{array}{ll}
0.834322 & 0.000468  \tag{40}\\
0.001415 & 0.823769
\end{array}\right]
$$

Comparing equations (39) and (40) with equations (29) and (30) suggests that for single-year models, the linear and exponential approaches tend to yield similar results. By comparing these four equations with equations (5) and (6), our preferred approach is clearly better than the conventional approach, as far as interpretability is concerned. However, one must remember that the survivorship proportions for the second to last age group tend to understate their true values, since the average mortality of the open-ended age group tends to be higher than the true mortality of the first corresponding closed interval (wh, wh+h).

## 4. COMPARISON OF THE PROJECTION RESULTS

The survivorship submatrices generated by the conventional approach, the two versions of the disaggregation approach, and the two versions of our preferred approach are combined with reproduction submatrices to form alternative growth matrices for the 1974 Swedish data shown in Appendix A. Here we want to evaluate the projections generated by these growth matrices in terms of their effects on the changing population size of the 85+ age group and on the entire system's annual growth rate.

Right from the first projected year (1975), the conventional approach separates itself from the other approaches by forcing the $85+$ age group to grow by an unrealistically high growth rate of $20.47 \% . *$ In contrast, the two versions of the preferred approach both imply a growth rate of $3.87 \%$, whereas the first version of the disaggregation approach results in a growth rate of $3.44 \%$ and the second version a growth rate of $8.66 \%$. Figure 1 shows that among all the approaches, the conventional one tends to produce most erratic projection for the last open-ended age group, because it amplifies the irregularities in the changing sizes of the incoming cohorts. The first version of the disaggregation approach shows a strong tendency to underproject the population size in the $85+$ age group, whereas the second version of the disaggregation approach is moderately sensitive to the irregularities in the initial age composition but does not show a strong tendency to under- or over-project the size of the oldest age group. The two versions of our preferred approach produce practically identical results that neither amplify the irregularities of the incoming cohorts nor exhibit any tendency of overor under-projection, assuming the age-specific rates are timeinvariant.

The conventional approach not only yields unrealistic projections for the $85+$ age group but also distorts significantly the growth rate of the entire population, although the proportion

[^1]

1 - conventional approach
2 - disaggregation approach I
3 - disaggregation approach II
4 - preferred linear and exponential approach

Figure 1. Size index of population aged 85 and over (1974 = 100).
of the population in the last age group is usually relatively small. Table 1 shows that between 1974 and 1975 the growth rate of the entire population is projected to be $0.56 \%$ by the conventional approach and $0.34 \%$ by our preferred approaches. This difference is completely due to the different treatments of the open-ended age group. We also see from Figure 2 that the growth rate of the entire population generated by the conventional approach tends to fluctuate significantly in response to the changing sizes of the cohorts entering the last age group, whereas the corresponding growth rates generated by our preferred approaches exhibit a more regular pattern. The more irregular the initial age composition, the worse the conventional approach will perform. For countries that have more irregular age compositions than Sweden (e.g., the Soviet Union and the Federal Republic of Germany), the differences between the approaches would be more dramatic than what is revealed in our example. The patterns of population growth rates generated by the two disaggregation approaches do not differ much from those of our preferred approaches, as shown in Appendix B.

Table 1. Annual growth rate in percent, 1974-1975.

| Approach | Population 85+ | Total population |
| :--- | :---: | :--- |
| Conventional | 20.47 | 0.56 |
| Disaggregation I | 3.44 | 0.33 |
| Disaggregation II | 8.66 | 0.40 |
| Preferred linear | 3.87 | 0.34 |
| Preferred |  |  |
| exponential | 3.88 | 0.34 |



KEY
1 - conventional approach
2 - preferred linear and exponential approach

Figure 2. Annual growth rate (in percent): 1974-2024.

## CONCLUSION

We have pointed out the problems of the conventional approach to population projection, arising from an inappropriate specification of the survivorship proportions for the last two age groups. After examining several alternative approaches that may remedy these problems, we are convinced that our preferred approaches described in section 3 are superior to the other alternatives in terms of (1) the interpretability of the survivorship proportions, (2) the reliability of the projection results, and (3) the computational effect involved.

# APPENDIX A: DATA FOR SWEDEN 1974, FEMALES, STOCKHOLM AND REST OF THE COUNTRY 





－その
内人

















49
49
59
51
52
53
54
55
55
57
58
59
60
61
62
63
64
65
66
67
68
69
73
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
tot
gross
crude (x1000)



> 0.00000 .00000000000000000000000000000
9991. 100.00

00.00
27.53

## 


6639. 100.00
73.06

|  |  |
| :---: | :---: |
| 45. | 0.35 |
| 60. | 0.47 |
| 43. | 0.38 |
| 47. | 0.36 |
| 45. | 0.35 |
| 41. | 0.32 |
| 44. | 0.34 |
| 39. | 0.30 |
| 37. | 0.29 |
| 29. | 6.23 |
| 38. | 0.29 |
| 42. | 6.33 |
| 33. | 0.30 |
| 33. | 0.26 |
| 40. | 0.31 |
| 46. | 0.36 |
| 59. | 0.46 |
| 44. | 0.34 |
| 33. | 0.26 |
| 26. | 0.20 |
| 39. | 0.30 |
| $2 i$. | 0.16 |
| 21. | 0.16 |
| 24. | 0.19 |
| 21. | 0.16 |
| 16. | 0.12 |
| 13. | 0.10 |
| 16. | 0.12 |
| 10. | 0.08 |
| 14. | 0.11 |
| 20. | 0.16 |
| 9. | 0.477 |
| 9. | 0.07 |
| 7. | 0.65 |
| 4. | 0.063 |
| 4. | 0.133 |
| 23. | 0.18 |

12884. 100.00
25.32


|  |  |
| :--- | :--- |
| 5. | 0.58 |
| 78. | 0.53 |
| 63. | 0.43 |
| 85. | 0.60 |
| 73. | 0.50 |
| 78. | 0.53 |
| 87. | 0.59 |
| 81. | 0.55 |
| 63. | 0.43 |
| 73. | 0.50 |
| 69. | 0.47 |
| 73. | 0.50 |
| 79. | 0.54 |
| 89. | 0.69 |
| 56. | 0.58 |
| 93. | 0.67 |
| 74. | 0.59 |
| 66. | 0.45 |
| 55. | 0.37 |
| 57. | 0.39 |
| 37. | 0.25 |
| 41. | 0.28 |
| 32. | 0.22 |
| 24. | 0.16 |
| 28. | 0.19 |
| 23. | 0.16 |
| 30. | 0.20 |
| 16. | 0.11 |
| 27. | 0.18 |
| 15. | 9.10 |
| 10. | 0.67 |
| 13. | 0.69 |
| 6. | 0.64 |
| 12. | 0.68 |
| 10. | 0.07 |
| 10. | 0.07 |
| 15. | 0.10 |
|  |  |


0000000000000000000000000000000000000

1.221
7.661
78.26


1.431
9.214
28.35

$-2.407$

 $\stackrel{\square}{a}$







ヘining－ ダオ




 number




|  |  <br> A－ g－vworond <br>  |
| :---: | :---: |
|  | －00000000000000ーーーーーーーーーーーーーーーーーーーーーー <br>  |
| ث <br> © | 0000000000000000000000000000000000000 |
| $\begin{array}{ll}N & \text {－} \\ 0 & 8 \\ \& & \dot{8}\end{array}$ | 0000000000000000000000000000000000000 |
| $\stackrel{N}{N}$ |  <br>  |
|  | N <br>  <br>  |
| A 0 0 |  |
| $\begin{array}{ll} \text { N } & \overline{8} \\ \text { on } & 8 \\ - & 8 \end{array}$ | 0000000000000000000000000000000000000 － |
| N C ＋ | N． |
|  | 0000000000000000000000000000000000000 <br>  |
|  | 0000000000000000000000000000000000000 |
|  |  <br>  |
| $\begin{aligned} & N A Q \\ & \infty \\ & N A N \\ & N=0.0 \end{aligned}$ |  |
|  | $000000-000000000-00-1-100-0000-00$ ： <br>  <br>  |
| 0 in in | －Q00 dodo －ininining giv inginnuw wivioog ogiciso <br>  |

APPENDIX B: PROJECTED GROWTH RATES OF SWEDISH FEMALES, 1975-2024

| Year | 1 | 2 | 3 | 4 | 5 |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1975 | 0.564200 | 0.333100 | 0.399800 | 0.338200 | 0.338500 |
| 1976 | 0.318300 | 0.298000 | 0.330900 | 0.308300 | 0.308600 |
| 1977 | 0.282100 | 0.261000 | 0.289200 | 0.275100 | 0.275400 |
| 1978 | 0.235100 | 0.223300 | 0.245800 | 0.239900 | 0.240100 |
| 1979 | 0.206700 | 0.186300 | 0.203100 | 0.204000 | 0.204100 |
| 1980 | 0.173100 | 0.151000 | 0.162300 | 0.168500 | 0.168600 |
| 1981 | 0.134300 | 0.118200 | 0.124600 | 0.134600 | 0.134700 |
| 1982 | 0.089500 | 0.088900 | 0.091000 | 0.103600 | 0.103600 |
| 1983 | 0.084500 | 0.063700 | 0.062400 | 0.076500 | 0.076400 |
| 1984 | 0.052600 | 0.043000 | 0.038700 | 0.053200 | 0.053100 |
| 1985 | 0.046400 | 0.025700 | 0.021800 | 0.033500 | 0.033400 |
| 1986 | 0.039000 | 0.010700 | 0.018000 | 0.016500 | 0.016400 |
| 1987 | -0.006400 | -0.004300 | 0.002900 | 0.002000 | 0.001800 |
| 1988 | -0.046100 | -0.017000 | -0.011500 | -0.010600 | -0.010800 |
| 1989 | -0.046200 | -0.027800 | -0.021400 | -0.021500 | -0.021700 |
| 1990 | -0.021300 | -0.037900 | -0.031100 | -0.031500 | -0.031700 |
| 1991 | -0.036900 | -0.047500 | -0.041100 | -0.040900 | -0.041200 |
| 1992 | -0.051000 | -0.056500 | -0.052200 | -0.050000 | -0.050200 |
| 1993 | -0.056200 | -0.065000 | -0.057700 | -0.058900 | -0.059000 |
| 1994 | -0.063700 | -0.074200 | -0.068000 | -0.067900 | -0.068000 |
| 1995 | -0.118300 | -0.083200 | -0.075000 | -0.076900 | -0.077000 |
| 1996 | -0.135300 | -0.092100 | -0.082000 | -0.085300 | -0.085400 |
| 1997 | -0.093700 | -0.100900 | -0.094800 | -0.093500 | -0.093600 |
| 1998 | -0.120000 | -0.108800 | -0.106900 | -0.101600 | -0.101700 |
| 1999 | -0.142400 | -0.115600 | -0.113000 | -0.109300 | -0.109400 |
| 2000 | -0. 169600 | -0.122200 | -0.115000 | -0.116500 | -0.116600 |
| 2001 | -0.173500 | -0.129100 | -0.122400 | -0.123100 | -0.123200 |
| 2002 | -0. 121300 | -0.135700 | -0.129700 | -0.129600 | -0.129600 |
| 2003 | -0.142600 | -0.141900 | -0.135400 | -0.135800 | -0.135900 |
| 2004 | -0.166800 | -0.147800 | -0.141000 | -0.141500 | -0.141600 |
| 2005 | 0.052800 | -0.154300 | -0.154100 | -0.147900 | -0.148000 |
| 2006 | -0.051700 | -0.160600 | -0.162800 | -0.155200 | -0.155300 |
| 2007 | -0.350300 | -0.164400 | -0.160400 | -0.160300 | -0.160400 |
| 2008 | -0.270200 | -0.166400 | -0.165100 | -0.162200 | -0.162400 |
| 2009 | -0.197800 | -0.166400 | -0.168100 | -0.162700 | -0.162900 |
| 2010 | -0.191900 | -0. 165200 | -0.170500 | -0.162400 | -0.162600 |
| 2011 | -0.200400 | -0.163300 | -0.169600 | -0.161700 | -0.162000 |
| 2012 | -0.200600 | -0.161200 | -0.160000 | -0.161000 | -0.161200 |
| 2013 | -0.147600 | -0.161300 | -0.162000 | -0.160900 | -0.161100 |
| 2014 | -0.158700 | -0.162000 | -0. 165800 | -0.161800 | -0.162000 |
| 2015 | -0.160700 | -0.163300 | -0.133700 | -0.163800 | -0.163900 |
| 2016 | -0.152000 | -0.172000 | -0.151700 | -0.167200 | -0.167400 |
| 2017 | -0.181000 | -0.179700 | -0.201000 | -0.172200 | -0.172300 |
| 2018 | -0.205700 | -0.181300 | -0.195200 | -0.178800 | -0.178800 |
| 2019 | -0.193000 | -0.186900 | -0.192500 | -0.187200 | -0.187300 |
| 2020 | -0.155200 | -0.196600 | -0.202400 | -0.197900 | -0.197900 |
| 2021 | -0.148000 | -0.209000 | -0.216900 | -0.211100 | -0.211000 |
| 2022 | -0.165300 | -0.223400 | -0.232400 | -0.226500 | -0.226400 |
| 2023 | -0.191600 | -0.239300 | -0.241300 | -0.243400 | -0.243300 |
| 2024 | -0.196500 | -0.257900 | -0.261000 | -0.261700 | -0.261500 |

KEY

```
1 - conventional approach
2 - disaggregation approach I
3 - disaggregation approach II
4 - preferred approach linear
5 - preferred approach exponential
```


## REFERENCES

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

Keyfitz, N. (1977) Applied Mathematical Demography. New York: John Wiley.

Liaw, K.-I., and J. Ledent (1980) Discrete Approximation of a Continuous Model of Multistate Demography. PP-80-14. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Ledent, J. (1978) Some Methodological and Empirical Considerations in the Construction of Increment-Decrement Life Tables. RM-78-25. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A. (1975) Introduction to Multiregional Mathematical Demography. New York: John Wiley.

Shieh, L.S., R. Yates, and J. Navarro (1978) Representation of Continuous-Time State Equations by Discrete-Time State Equations. IEEE 8(6):485-492.

Willekens, F., and A. Rogers (1978) Spatial Population Analysis: Methods and Computer Programs. RR-78-18. Laxenburg, Austria: International Institute for Applied Systems Analysis.


[^0]:    *For convenience, we usually let $\underset{\sim}{\ell}$ be such that most $\underset{\sim}{\ell}$
    are invertible.

[^1]:    *All the growth rates mentioned in this section are instantaneous rates per year.

