International Institute for
Applied Systems Analysis
www.iiasa.ac.at

Parametrized Multistate Population Projections

Rogers, A. and Planck, F.

IIASA Collaborative Paper
August 1984

PARAMETRIZED MULTISTATE POPULATION PROJECTIONS

Andrei Rogers

Friedrich Planck

August 1984

CP-84-35

ABOUT THE AUTHORS

Andrei Rogers is Professor of Geography and Director of the
Population Program, Institute of Behavioral Science,
University of Colorado, Boulder, Colorado, 80309, USA.

Friedrich Planck is Research Associate at the Institute of
Social Sciences, University of Erlangen, Nurnberg, D-85,
Federal Republic of Germany.

ABSTRACT

This paper reports progress on the development of a population

projection process that emphasizes model selection over demographic

accounting. Transparent multiregional/multistate population projections

that rely on parametrized model schedules are illustrated, together with

simple techniques that extrapolate the recent trends exhibited by the

parameters of such schedules. The parametrized schedules condense the

amount of demographic information, expressing it in a language and

variables that are more readily understood by the users of the

projections. In addition, they permit a concise specification of the

expected temporal patterns of variation among these variables, and they

allow a disaggregated focus on demographic change that otherwise would not

be feasible.

PARAMETRIZED MULTISTATE POPULATION PROJECTIONS

It has been argued that the population projection process should be
formulated as one of model selection rather than of demographic accounting
(Brass, 1974 and 1977; Keyfitz 1972). This paper reports progress on the
development of such a projection process. It describes methods for
generating multiregional/multistate population projections that rely on
parametrized model schedules and simple techniques that extrapolate trends
for the parameters of such schedules. The parametrized schedules condense
the amount of information to be specified as assumptions, simplifying and
making more transparent what is being modeled; they express this condensed
information in a language and variables that are readily understood by the
users of the projections; they permit a more concise specification of the
expected temporal patterns of variation among these variables; and they
allow a finer disaggregation of demographic change components than would
otherwise be feasible.

1. INTRODUCTION

Multistate generalizations of the classical single state projection
models widely used in applied demography today assess the numerical
consequences, to an observed or hypothetical (single-sex) multistate
population, of a particular set of assumptions regarding future patterns
of mortality, fertility, and interstate transfers. The multistate model
of demographic growth and change expresses the population projection
process by means of a simple matrix operation in which a population set
out as a vector is multiplied by a growth matrix that survives the
population forward over time. The projection computes the state- and
age-specific survivors of a given sex and adds to this total the
corresponding surviving new births.

Multistate demographic projections incorporate two important aspects of population dynamics that lead to greater consistency among projected outputs: 1) accounting identities that interconnect changes in events and flows with changes in population stocks, and 2) interstate transition probabilities that reflect the influences of past events and flows through the current age and status distributions of the aggregate population. For example, the number of widowings in a given region will be influenced by the number of married women residing there, a number that in turn is influenced by the number of marriages in previous periods in all regions and the number of married women inmigrating to and outmigrating from the region of interest.

To ensure that accounting identities connecting events, flows, and stocks are respected, the multistate projection model traces the evolution of each status-specific category of individuals by adjusting an initial stock to take into account the number of events and flows that are expected to occur during a projection period. In this way, changes in the number of events and flows are reflected in the projected age- and status-specific distribution of the population.

The influences running in the reverse direction are also included. Changes in age- and status-specific population stocks influence future events and flows. For example, increases in the number of marriages at a particular age in a given region will lead to increases in the number of married persons and thereby produce a rise in nuptial births there in the future.

Multistate population projections generally need to keep track of enormous amounts of data. The disaggregations incorporated in such projections are introduced either because forecasts of the specified

population subgroups are important in their own right, or because it is believed that simple and regular trends are more likely to be discovered at relatively higher levels of disaggregation.

High levels of disaggregation permit a greater flexibility in the use of the projections by a wide variety of users; they also often lead to a detection of greater consistency in patterns of behavior among more homogeneous population subgroups. But greater disaggregation requires the estimation of even greater numbers of data points, both those describing initial population stocks and those defining the future rates of events and flows that are expected to occur. The practical difficulties of obtaining and interpreting such data soon outstrip the benefits of disaggregation.

Mathematical descriptions of schedules of demographic rates, here called parametrized model schedules, offer a means for condensing the amount of information to be specified as assumptions. They also express this condensed information in a language and variables that are more readily understood by the users of the projections, and they provide a convenient way of associating the variables to one another, extrapolating them over time, and relating them to variables describing the economic environment that underlies the projections.

The use of parametrized model schedules in the population projection process allows one to develop an effective description of how the components of demographic change (e.g. mortality, fertility, and migration) are assumed to vary over time in terms of a relatively few parameters. Insofar as the assumptions correctly anticipate the future, the projection foretells what indeed comes to pass. Insofar as the parameters are readily interpretable by non-demographer users of the

projection, they make possible the assessment of the reasonableness of a
set of assumptions instead of a set of projected population totals.

As Keyfitz (1972) correctly observes, a trend extrapolation of each
age-specific rate in a population projection is an excessive concession to
flexibility that can readily produce erratic results. On the other hand,
to assume that change in a set of rates occurs uniformly at all ages is to
go against experience. Parametrized model schedules offer a way of
introducing flexibility, while at the same time retaining the
interdependence between the rates of a particular schedule.

The aim of this paper is to illustrate a procedure for multistate
population projection that requires the specification of future trends for
a number of significant parameters defining a collection of model
schedules. The intent of such a procedure is, in the words of William
Brass (1977, p. 15):

> ...to sketch out a procedure for population projection which
> requires the estimation of future trends for a minimum of
> significant parameters.
> ...to shift as far as practicable from the appearance of a
> bookkeeping, accounting system to one in which the somewhat crude
> model elements are apparent and, thus, their inescapable lack of
> certainty displayed.

The illustration considers a two region-four state description of the
Swedish female population in 1974 and examines alternative projections of
that population into the future. We begin with a description of
parametrized model schedules and the input data, continue with a
discussion of the associated multistate life tables and constant
coefficient projections, and conclude with an exposition of simple
variable coefficient projections that are driven by assumed patterns of
change in fertility, marital status transitions, and internal migration.

2. PARAMETRIZED MODEL SCHEDULES AND INPUT DATA [1]

The use of mathematical functions, expressed in terms of a small set of parameters, to smooth and describe parsimoniously schedules of age-specific rates is a common practice in demography. Such functions have been fitted to mortality and fertility data, for example, and the results have been widely used for data smoothing, interpolation, comparative analysis, data inference, and forecasting (Brass 1971, Coale and Demeny 1966 and 1983, Coale and Trussell 1974, Heligman and Pollard 1980, Hoem et al. 1981, and United Nations 1967 and 1983.)

More recently, the range of parametrized schedules has been expanded to include interstate transfers such as migration (Rogers, Raquillet, and Castro 1978; Rogers and Castro 1981) and changes in marital status other than first marriage (Rogers and Williams 1982, and Williams 1981). Thus it is now possible to define a model (hypothetical) multistate dynamics that describes the evolution of a single-sex population exposed to parametrized schedules of mortality, fertility, migration, and several forms of marital status change (that is, first marriage, divorce, widowhood, and remarriage).

Parametrized model schedules describe the remarkably persistent regularities in age pattern that are exhibited by many empirical schedules of age-specific rates. Mortality schedules, for example, normally show a moderately high death rate following birth, after which the rates drop to a minimum between ages 10 to 15, then increase slowly until about age 50, and thereafter rise at an increasing pace until the last years of life.

---

[1] This section is drawn from Rogers (1982).

Fertility rates generally start to take on nonzero values at about age 15 and attain a maximum somewhere between ages 20 and 30; the curve is unimodal and declines to zero once again at some age close to 50. Similar unimodal profiles may be found in schedules of first marriage, divorce, and remarriage. The most prominent regularity in age-specific schedules of migration is the high concentration of migration among young adults; rates of migration also are high among children, starting with a peak during the first year of life, dropping to a low point at about age 16, turning sharply upward to a peak near ages 20 to 22, and declining regularly thereafter except for a possible slight hump or upward slope at the onset of the principal ages of retirement. Although data on rates of labor force entry and exit are very scarce, the few published studies that are available indicate that regularities in age pattern also may be found in such schedules. Figure 1 illustrates a number of typical age profiles exhibited by schedules of rates in multistate demography.

The shape or _profile_ of a schedule of age-specific rates is a feature that may be usefully examined independently of its intensity or _level_. This is because there are considerable empirical data showing that although the latter tends to vary significantly from place to place, the former remains remarkably similar.

The level at which occurrences of an event or a flow take place in a multistate population system may be represented by the area under the curve of the particular schedule of rates. In fertility studies, for example, this area is called the gross reproduction rate if the rates refer to parents and babies of a single sex. By analogy, therefore, we shall refer to areas under all schedules of rates as _gross transition rates_ (GTRs), inserting the appropriate modifier when dealing with a
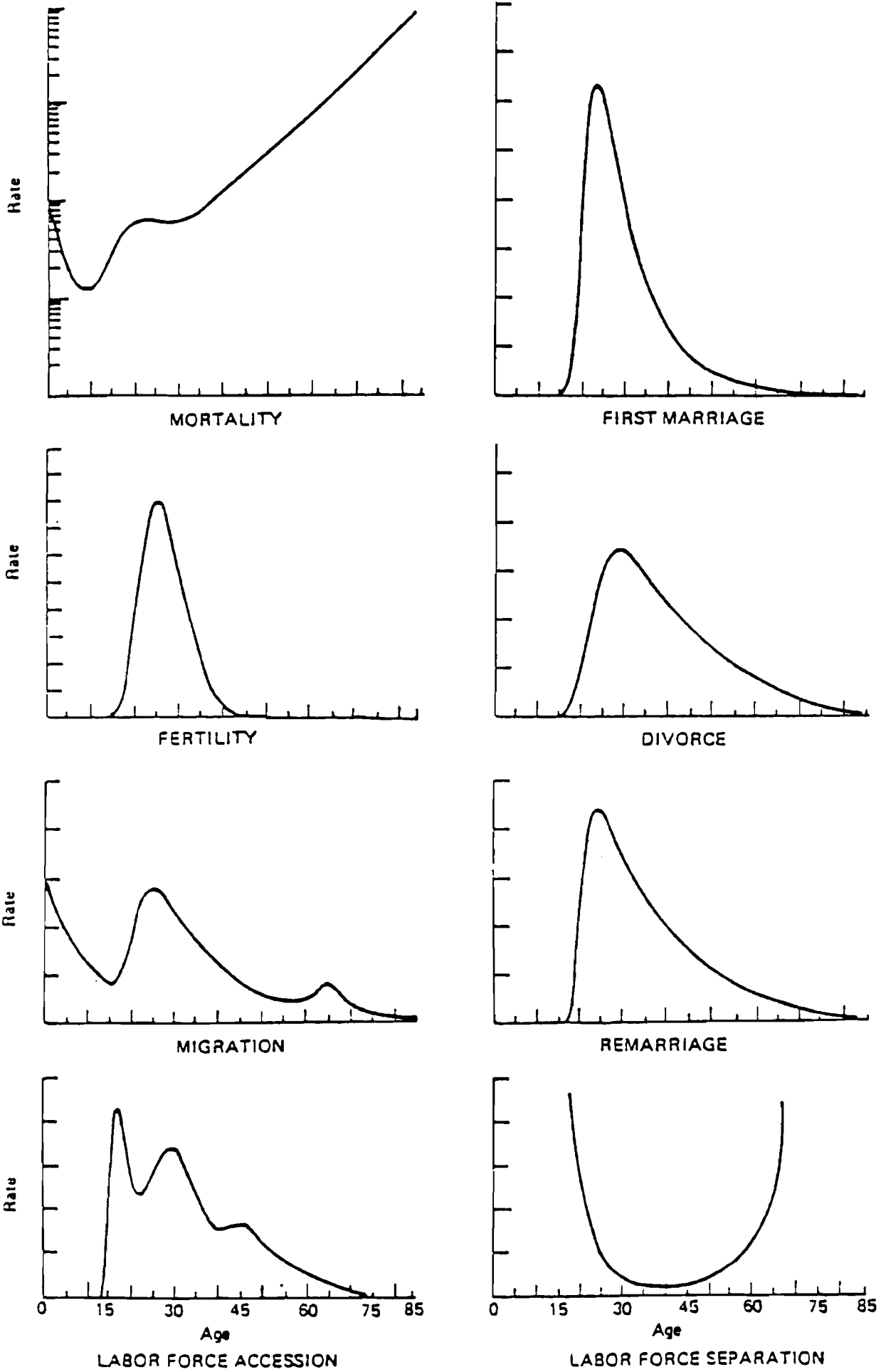
Figure 1. Multistate schedules.
Source: Rogers (1982).

particular event or flow--for example, gross mortality transition rate and gross accession transition rate. The term "transition" is introduced throughout in order to distinguish this aggregate measure of level from the other more common gross rates used in demography, such as the directional gross (instead of net) rate of migration.

The gross transition rate measures the intensity of particular events within a state population or of flows between state populations during a given interval of time. The index, therefore, is a cross-sectional measure and should not be confused with the net transition rate (such as the net reproduction rate), which is a cohort-related index that measures the intensity of such events or flows over a lifetime. Moreover, in a multistate framework, where return flows such as remarriages play an important role, gross and net rates can give widely differing indications of interstate movement intensities.

2.1 Mortality

Three principal approaches have been advanced for summarizing age patterns of mortality: functional descriptions in the form of mathematical expressions with a few parameters (Benjamin and Pollard 1980), numerical tabulations generated from statistical summaries of large data sets (Coale and Demeny 1966 and 1983), and relational procedures associating observed patterns with those found in a standard schedule (Brass 1971).

The search for a "mathematical law" of mortality has, until recently, produced mathematical functions that were successful in capturing empirical regularities in only parts of the age range, and numerical tabulations have proven to be somewhat cumbersome and inflexible for

computer-based applied analysis. Consequently, the relational methods
first proposed by William Brass have become widely adopted. With two
parameters and a standard life table, it has become possible to describe
and analyze a large variety of mortality regimes parsimoniously.

Recently, Heligman and Pollard (1980) described a mathematical model
that appears to provide satisfactory representations of a wide variety of
age patterns of mortality across the entire age range. Their function
defines the variable $q(x)$, the probability of dying within one year for an
individual at age x. We have found it more convenient to focus instead on
$d(x)$, the annual death rate at age x, and to adopt the slightly modified
Heligman-Pollard formula, suggested by Brooks et al. (1980) of the IMPACT
Project, that appears as Equation 1 in Figure 2. The three terms in that
equation represent infant and childhood mortality (I), mortality due to
accidents (A), and a senescent mortality component (S) which reflects
mortality due to aging. Figure 3 exhibits those three components and
their sum, drawing on Australian data for 1950.

Death rates differ markedly not only between ages, but also between
sexes, marital states, and occasionally regions. At the IMPACT Project,
model schedules based on Equation 1 of Figure 2 have been successfully
fitted to Australian age-specific data for the death rates of persons of
each sex and marital status (Figure 4). Not all components of the
Heligman-Pollard curve were used; the first component was omitted for
married males and females , as well as for divorced and widowed females;
the first and second components were both omitted for divorced and widowed
males.

MORTALITY

$$d(x) = d_I(x) + d_A(x) + d_S(x) \tag{1}$$

WHERE

$$d_I(x) = \begin{cases} Q_0 & \text{for } x = 0 \\ Q_1 x^\gamma & \text{for } x > 0 \end{cases}$$

$$d_A(x) = Q_A \, e^{-\left(\frac{\ln x - \ln X_A}{\sigma}\right)^2} \qquad \text{for } x \geq 0$$

$$d_S(x) = Q_S \, \frac{e^{x/X_S}}{1 + Q_S \, e^{x/X_S}} \qquad \text{for } x \geq 0$$

WIDOWHOOD

$$d(x) = d_A(x) + d_S(x) \tag{2}$$

FERTILITY, MARRIAGE, AND DIVORCE

$$f(x) = a \, e^{-\alpha(x-\mu) - e^{-\lambda(x-\mu)}} \tag{3}$$

Figure 2. Model Schedules.
Source: Rogers (1982)

MIGRATiON

$$m(x) = a_1 e^{-\alpha_1 x} + a_2 e^{-\alpha_2 (x-\mu_2) - e^{-\lambda_2 (x-\mu_2)}} + R + c \qquad (4)$$

WHERE

$$R = a_3 e^{-\alpha_3 (x-\mu_3) - e^{-\lambda_3 (x-\mu_3)}} \qquad (5)$$

OR

$$R = a_3 e^{\alpha_3 x} \qquad (6,$$

OR

$$R = 0$$

NOTE: $a_1 = \alpha_1 = 0$ FOR MIGRATION OF MARRIED, WIDOWED,
                          AND DIVORCED PERSONS
      AND
         $R = c = 0$
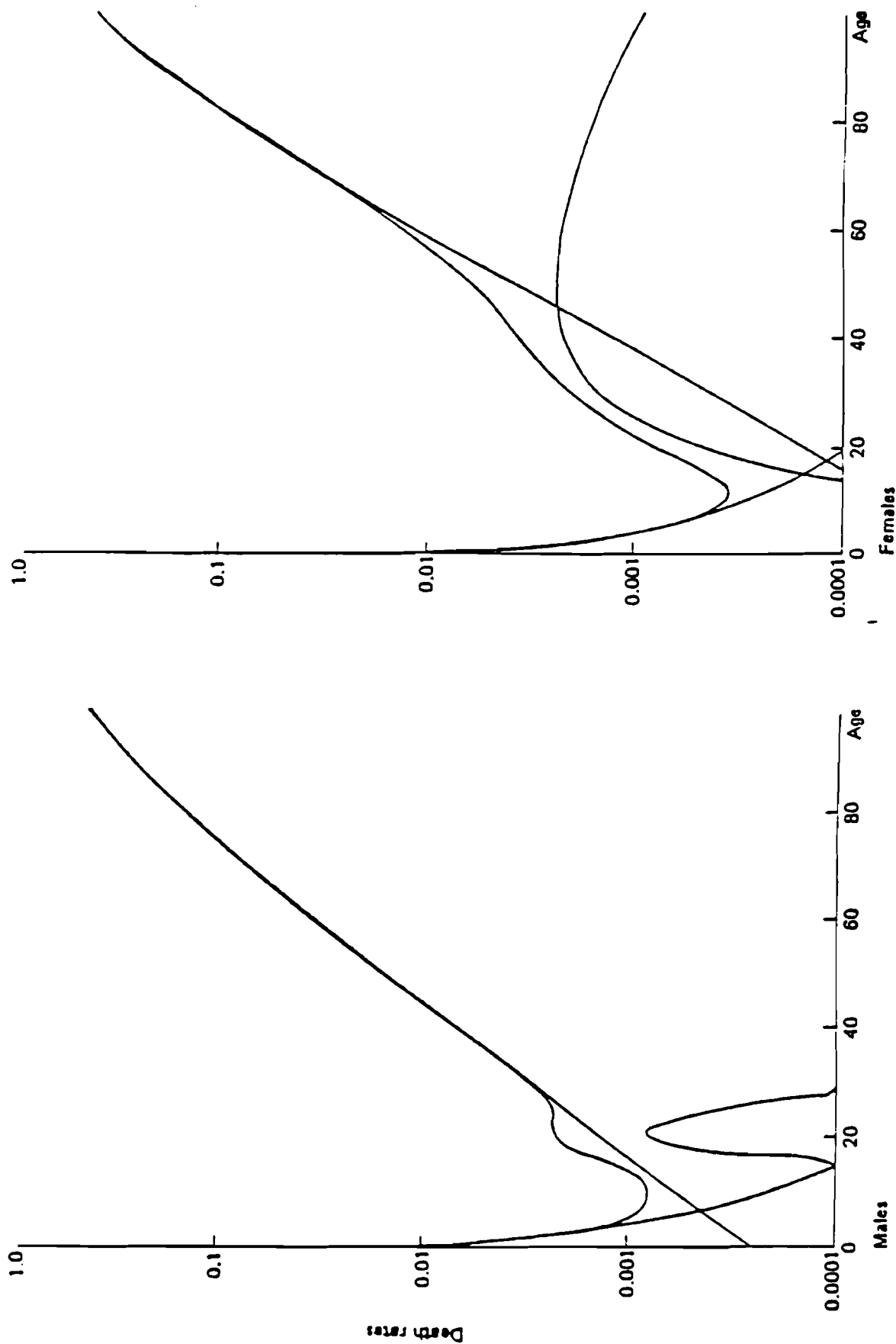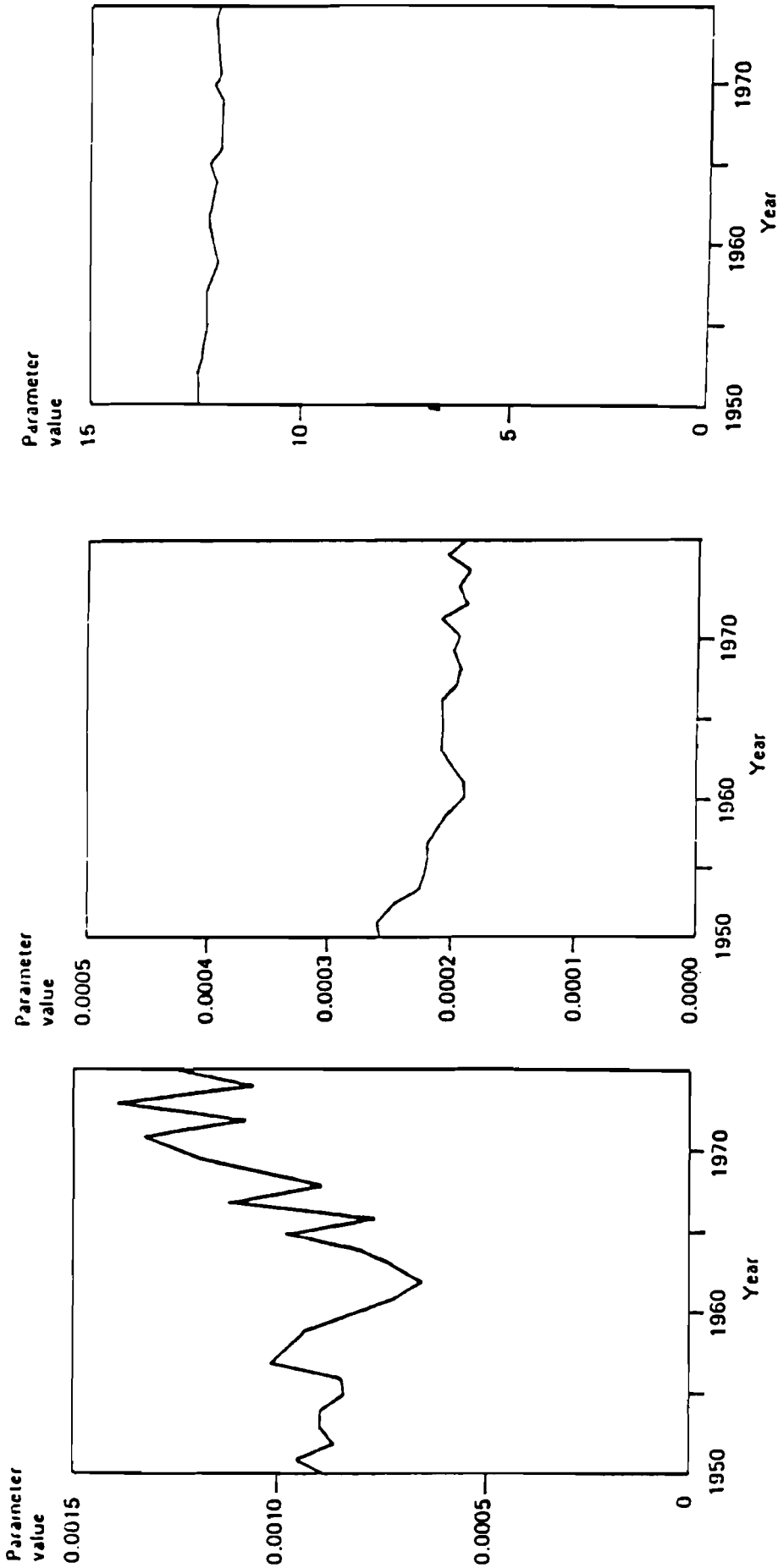
Figure 2 (continued) Model Schedules.

Figure 3.   The components and total death rates for the Heligman-Pollard functions for never married Australian males and females for 1950.

Source: Brooks et al. (1980).

Parameter value

The value of the parameter Q_A for never married males for the years 1950–1975

Parameter value

The values of the parameter Q_S for never married males for 1950–1975

Parameter value

The values of the parameter X_S for never married males for 1950–1975

Figure 4.   The values of the parameters for never married Australian males for the years 1950–1975.

Source: Brooks et al. (1980).

After fitting such model schedules in each region of a multiregional system, movements over time in their parameters could then be analyzed and used for projecting future mortality by age, sex, marital status, and region. For example, linear regression equations could be fitted to the trajectories set out in Figure 4, and short extrapolations of those trends could produce the needed projected future regimes of mortality. The relatively large number of parameters, however, suggests the desirability of extrapolating some function of the parameters instead--for example, the two-parameter Brass (1971) logit transformation of the mortality schedule.

2.2 Fertility

Among the relatively large number of different parametric functions that have been proposed recently for representing schedules of age-specific fertility, the formula put forward by Coale and Trussell (1974) has assumed a certain pre-eminence. This formula can be viewed as the product of two component schedules: a model nuptiality schedule and a model marital fertility schedule. The former adopts the double-exponential first marriage function of Coale and McNeil (1972):

$$g(x) = \frac{0.19465}{k} \, e^{-\frac{0.174}{k}(x - x_0 - 6.06k) \, - \, e^{-\frac{0.2881}{k}(x - x_0 - 6.06k)}}$$

where x_0 is the age at which a consequential number of first marriages begin to occur, and k is the number of years in the observed population into which one year of marriage in the standard population is transformed. Integrating, one finds

$$G(x) = \int_0^x g(a)\,da$$

which when multiplied by the proportion who will ever marry represents the proportion married at each age.

Coale and Trussell (1974) argue that marital fertility either follows a pattern that Henry (1961) called natural fertility or deviates from it in a regular manner that increases with age, such that the ratio of marital fertility to natural fertility can be expressed by

$$\frac{r(x)}{n(x)} = M\ e^{mv(x)}$$

where M is a scaling factor that sets the ratio r(x)/n(x) equal to unity at some fixed age, m indicates the degree of control of marital fertility, and v(x) and n(x) are fixed values that are assumed to remain invariant across populations and over time.

Multiplying the two-parameter model schedule of proportions ever married at each age by the one-parameter model schedule of marital fertility, Coale and Trussell (1974) generated an extensive set of model schedules that describe empirical fertility rates with surprising accuracy. Their representation as

$$f(x) = G(x) \cdot r(x) = G(x)n(x)e^{mv(x)}$$

allows one to obtain fertility age profiles (but not levels) that depend only on the fixed single-year values of the functions n(x) and v(x), and on estimates for x_0, k, and m.

If the populations to be projected are already disaggregated by marital status, such that the proportions married, never married, and previously married at each age are known, appropriate model schedules for

the age-specific fertility rates of women of each marital status may be developed. This allows one to consider separately marital and non-marital fertility, each of which may be influenced by different demographic and economic factors. In the illustrative projection developed later in this paper, a double-exponential function (set out as Equation 3 in Figure 2) is used to describe fertility rates at age x for women of each marital status in each region. Figure 5 illustrates the fit of this function to the 1962-1971 age-specific fertility rates of Denmark analyzed by Hoem et al. (1981).

The shape of the double-exponential curve is defined by the three parameters, α, μ, and λ, and the level of the curve is defined by the scaling parameter a. Although these parameters are not readily interpretable, it is possible to derive the propensity, mean, variance, and mode of the double-exponential function in terms of them (Coale and McNeil 1972; Rogers and Castro 1981; and Sams 1981).

2.3 Migration

A recent study of age patterns in migration schedules (Rogers and Castro 1981) has shown that such patterns exhibit an age profile that can be adequately described by the mathematical expression appearing as Equation 4 in Figure 2. The four terms in the equation represent childhood migration, labor force age migration, retirement migration, and a constant level of migration across all ages.

The shape of the second term, the labor force component of the curve, is the double exponential formula put forward by Coale and McNeil (1972). The first term, a simple negative exponential curve, describes the migration age profile of children and adolescents. Finally, the post-labor force component is a constant, another double-exponential, or
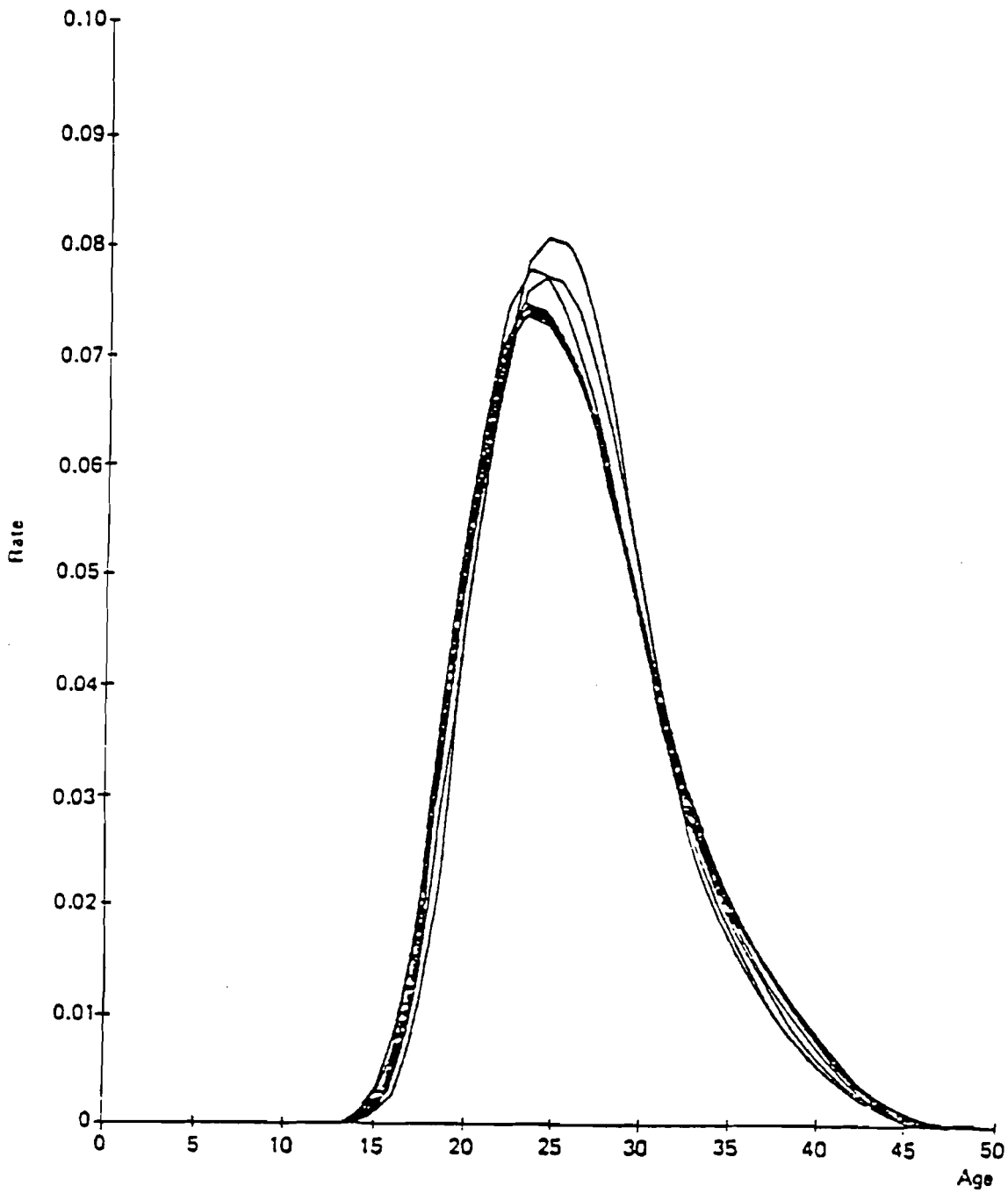
Figure 5.   The double exponential model fertility schedule:
            Denmark, 1962-71.

Source: Hoem et al.   (1981) and Rogers   (1982).

an upward sloping positive exponential. The fourth term decribes a
constant level of migration across all ages. The migration rate, m(x),
therefore, depends on values taken on by anywhere from 7 to 11
parameters. Figure 6 illustrates the fit of the nine parameter model
schedule to intercommunal migration in the Netherlands.

2.4 Marital Status

Coale and McNeil's (1972) double-exponential model schedule of first
marriages was introduced a decade ago. Parametrized schedules of other
changes in marital status, however, seem to have been first used only
recently, in a study carried out by the IMPACT Project in Australia
(Powell 1977). Working with a detailed demographic data bank produced by
Brown and Hall (1978), Williams (1981) fitted gamma distributions to
Australian rates of first marriage, divorce, remarriage of divorcees, and
remarriage of widows, for each year from 1921 to 1976. These model
schedules provided adequate descriptions of Australian marital status
changes, although some difficulties arose with age distributions that
exhibited steep rises in early ages; in particular, the age distributions
of first marriages. This difficulty was overcome by the addition of a
second time-invariant gamma distribution.

Functions based on the Coale-McNeil double-exponential distribution
seem better able to cope with the problem of steeply rising age
distributions than the gamma distribution. Figure 7 illustrates the
goodness-of-fit of the double-exponential distribution to data on
Australian males in 1976. Although the parameters of both functions can be
expressed in terms of the propensity, mean age and variance in age, the
double-exponential function requires a further parameter--the modal
age--whose movements over time may be more difficult to model and project.

19
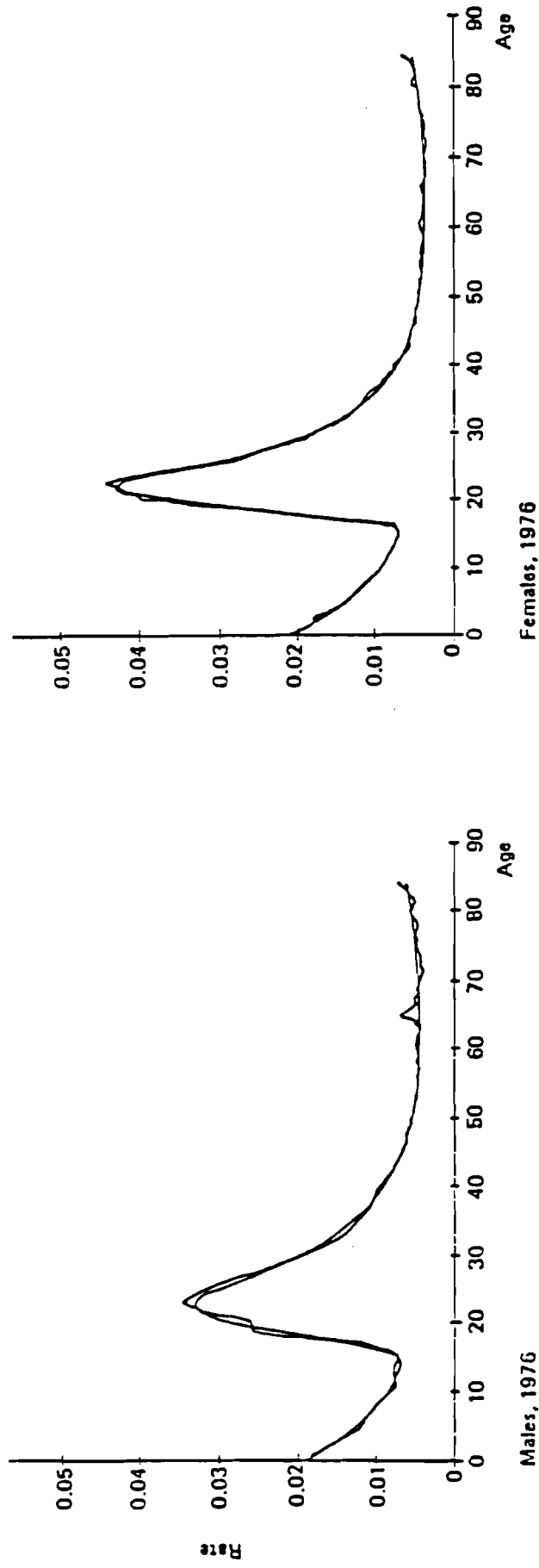
Figure 6. Model migration schedules for the Netherlands.

Source: Drewe (1980) and Rogers (1982).

First marriage

Divorce

Remarriage of divorced

Remarriage of widowed

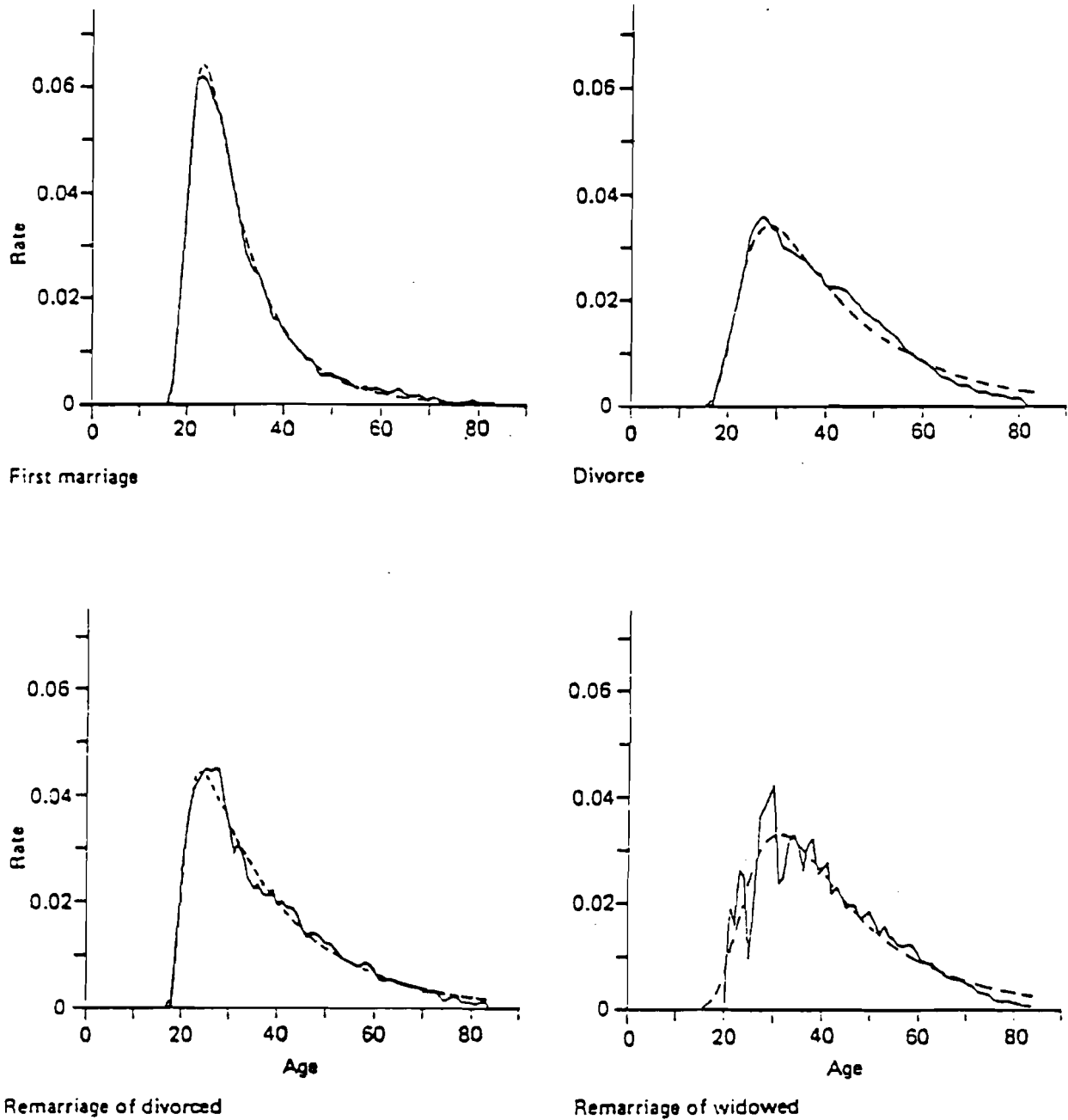Figure 7.   Double exponential model schedules of marital
             status change (--- model schedule, ——— observed
             data): Australian males 1976.

Source: Brown and Hall (1978) and Rogers (1982).

2.5 Other Transitions

The notion of model schedules may be used to describe a wide range of demographic transitions. We have considered mortality, fertility, migration, marriage, divorce, and remarriage. We could as easily have focused on flows between different states of, for instance, income, education, health, and labor force activity.
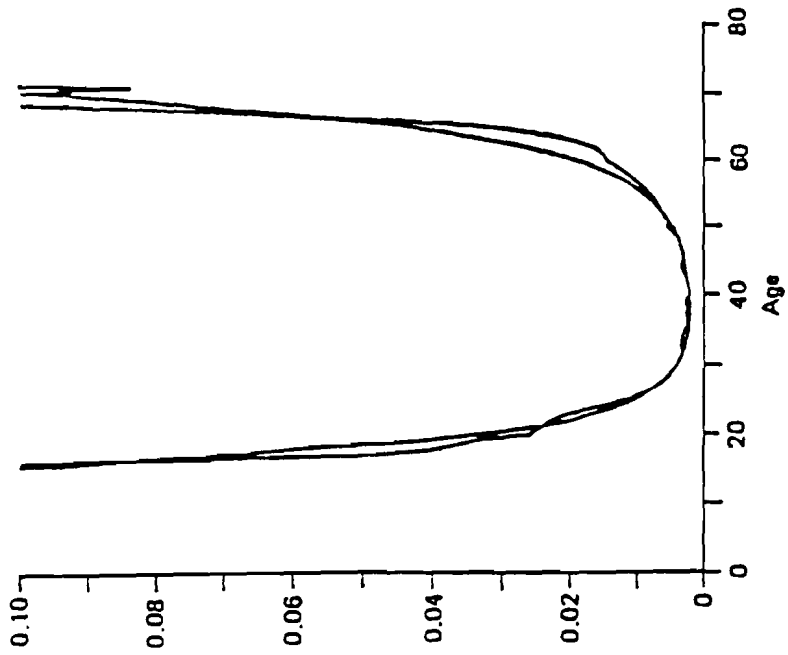
Consider, for example, the flows between active and inactive statuses in studies of labor force participation. Rates of entry into the labor force, called accession rates, exhibit an age profile that can be described as the sum of three double exponential distributions. Rates of exit from the labor force, called separation rates, may be described by a U-shaped curve defined as

$$h(x) = a_1 e^{-\alpha_1 x} + a_3 e^{\alpha_3 x} + c$$

Figure 8 illustrates the fit of these two curves to accession and separation rates, respectively, of Danish males in 1972-74 (Hoem and Fong 1976).

2.6 Input Data: Swedish Females, 1974

To illustrate the process of carrying out a parametrized multistate population projection, we have brought together data that describe the mortality, fertility, migration, and marital status change patterns of the Swedish female population in 1974. Data describing the first three components of change were provided by Arne Arvidsson of the Swedish Central Bureau of Statistics for a study of Sweden's migration and settlement structure (Andersson and Holmberg 1980). Data on marital status change flows were unavailable in the detail required and had to be

A. Model labor force accession schedule

B. Model labor force separation schedule

Figure 8. Model schedules of male labor force accession and separation:
Denmark 1972-1974.

Source: Hoem and Fong (1976) and Roqers (1982).

inferred by borrowing the age profiles observed in Norway in 1977-78
(Brunborg et al. 1981). Table 1 sets out the resulting crude rates of
events and flows in the two region system of Stockholm and the rest of
Sweden, and Table 2 presents the parameters that define the corresponding
model schedules of age-specific rates. Figures 9 through 12 illustrate
the fits of the model schedules to observed data, including a number of
male schedules for purposes of comparison.

Our experience with fitting the Heligman-Pollard function to Swedish
data suggests that the model schedule is over-parametrized. (A similar
observation is made by Brooks et al. 1980.) With so many variables to
estimate, very similar distributions can be obtained with significantly
different combinations of values for the parameters. The net result of
this is the creation of relatively large fluctuations in parameter
estimates over time, as changes in the values of one parameter produce
compensating shifts in those of another. To dampen such fluctuations we
follow the suggestion of Brooks et al. (1980) and fix the values of X_A
and σ. This establishes the position and shape of the accident component
but permits its level Q_A to change from year to year.

Except for mortality, the level parameters of all model schedules have
values scaled to produce a unit area under the curve (i.e., a gross
transition rate of unity). When used for projection purposes, these
parameters need to be multiplied by the appropriate observed or forecasted
gross transition rates.

3. MULTISTATE LIFE TABLES

The simplest life tables recognize only one class of decrement, e.g.,
death, and their construction is normally initiated by estimating a set of

TABLE 1. OBSERVED CRUDE RATES OF MULTISTATE TRANSFERS (PER THOUSAND): SWEDEN, FEMALES, 1974

TO FROM	STOCKHOLM				REST OF SWEDEN				DEATH	BIRTH
	NEVER MARRIED	MARRIED	WIDOWED	DIVORCED	NEVER MARRIED	MARRIED	WIDOWED	DIVORCED		
STOCKHOLM										
NEVER MARRIED	--	23.5	--	--	25.2	1.2	--	--	4.2	9.3
MARRIED	--	--	12.1	25.6	--	15.8	0.1	0.1	5.8	18.9
WIDOWED	--	1.1	--	--	--	0.0	5.8	--	43.7	0.3
DIVORCED	--	39.8	--	--	--	2.0	--	13.6	8.6	9.1
REST OF SWEDEN										
NEVER MARRIED	5.0	0.9	--	--	--	21.5	--	--	4.6	9.2
MARRIED	--	2.5	0.1	0.1	--	--	13.3	12.1	6.4	19.4
WIDOWED	--	0.0	0.6	--	--	1.02	--	--	46.8	0.3
DIVORCED	--	1.5	--	1.9	--	36.5	--	--	9.6	7.7
SWEDEN TOTAL									9.5	13.0

24

TABLE 2. MODEL SCHEDULE PARAMETERS: SWEDEN, FEMALES, 1974

PARAMETERS	MORTALITY			WIDOWHOOD	PARAMETERS	MIGRATION	
	SWEDEN	STOCKHOLM	REST OF SWEDEN	SWEDEN		STOCKHOLM	REST OF SWEDEN
Q_1	0.00173	0.00158	0.00193	--	a_1	0.028	0.018
γ	0.167	0.178	0.170	--	α_1	0.092	0.104
Q_S	0.00001	0.00002	0.00001	0.00002	a_2	0.047	0.092
x_S	9.06	9.93	9.24	9.75	μ_2	19.32	17.92
Q_A	0.00024	0.00044	0.00026	0.00036	α_2	0.094	0.140
x_A	22.0	22.0	22.0	22.5	λ_2	0.369	0.662
σ	0.4	0.4	0.4	0.5	a_3	0.00008	--
Q_0	0.00714	0.00649	0.00729	--	μ_3	85.84	--
					α_3	0.392	--
					λ_3	0.072	--
					c	0.002	0.003

TABLE 2 (CONTINUED)   MODEL SCHEDULE PARAMETERS: SWEDEN, FEMALES, 1974

| PARAMETERS	FERTILITY			PARAMETERS	MARITAL STATUS			
	SWEDEN	STOCKHOLM	REST OF SWEDEN		FIRST MARRIAGE	DIVORCE	REMARRIAGE OF WIDOWED	REMARRIAGE OF DIVORCED
a	0.095	0.052	0.145	a	0.157	0.081	0.129	0.084
μ	32.79	35.89	30.21	μ	20.29	20.72	20.64	17.23
α	0.328	0.361	0.303	α	0.140	0.072	0.115	0.080
λ	0.130	0.115	0.148	λ	0.343	0.224	0.333	1.0
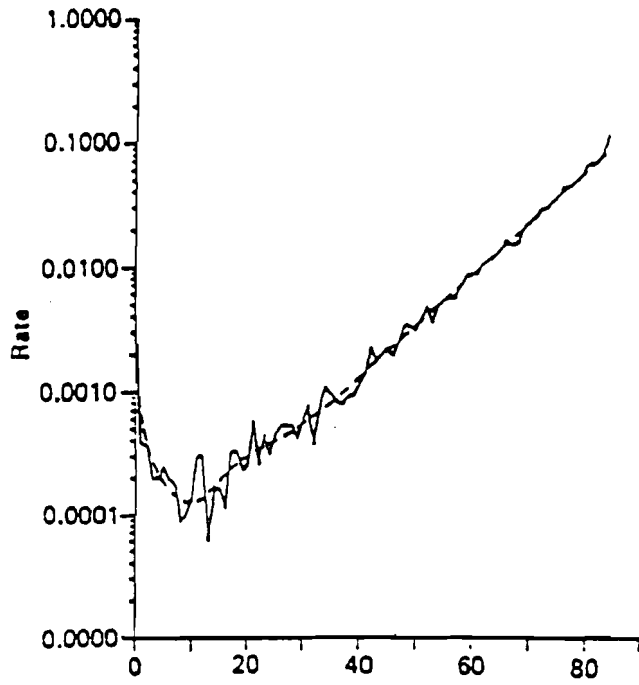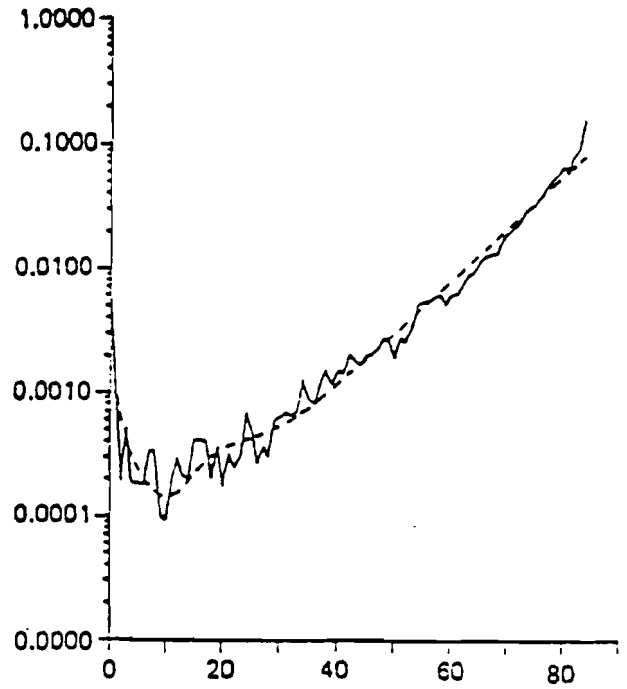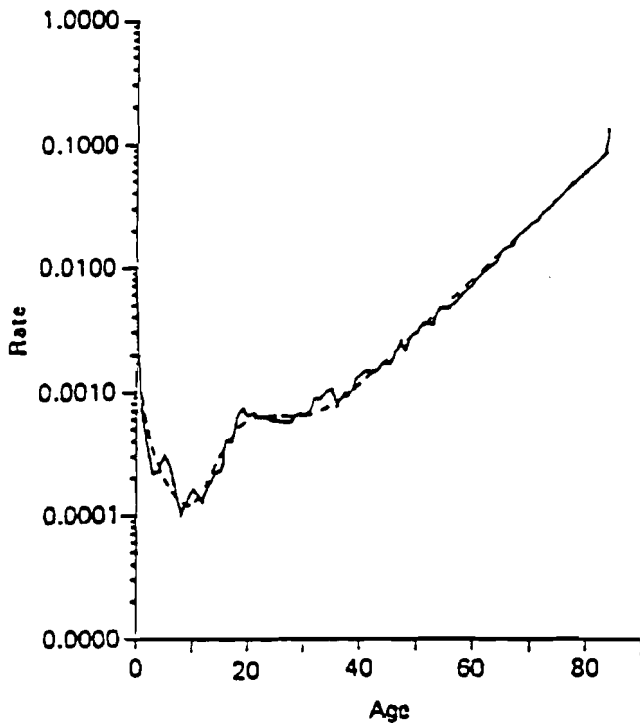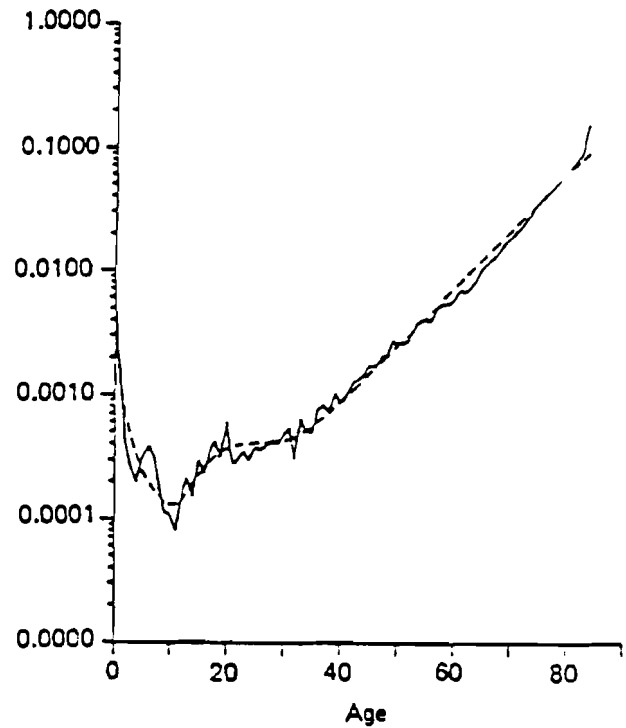
Figure 9. Model mortality schedules for Swedish data (--- model schedule, — observed data), 1974.
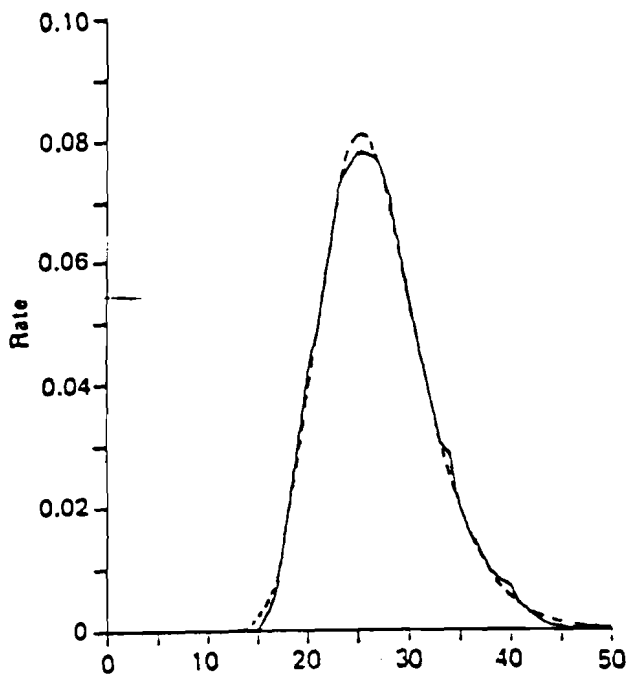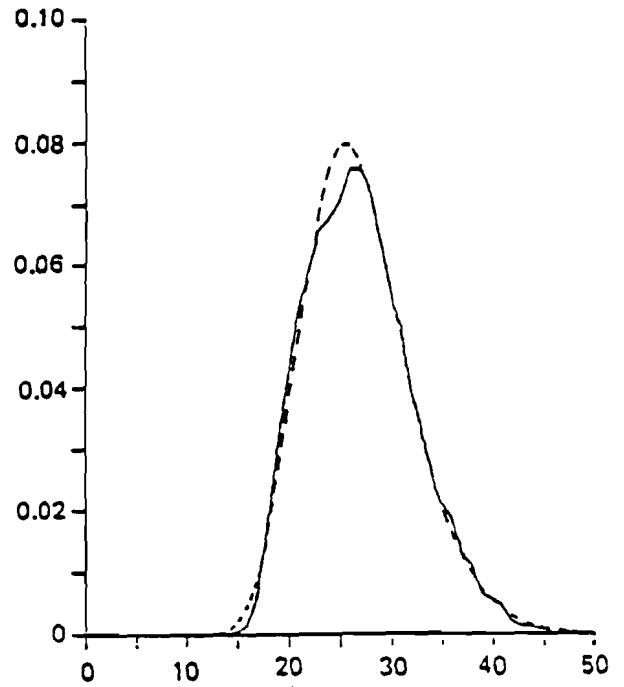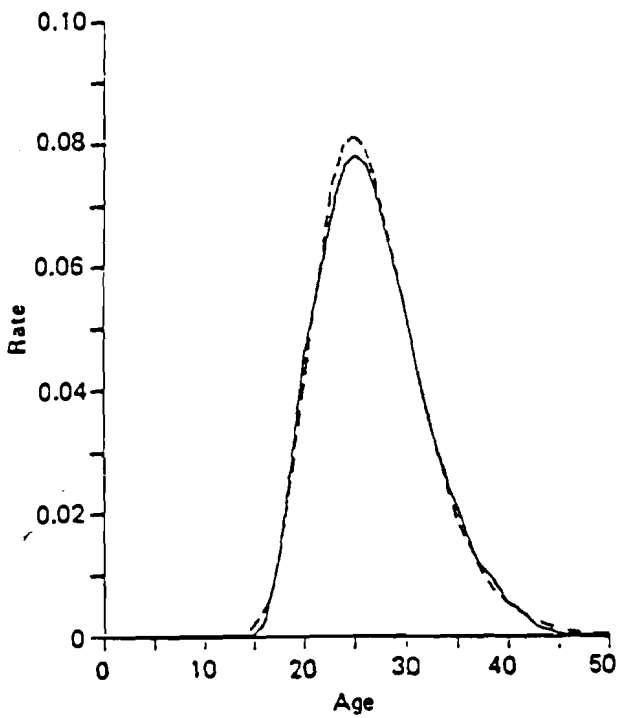
Source: Andersson and Holmberg (1980) and Rogers (1982).

Figure 10. Model fertility schedules for Swedish data (--- model schedule, — observed data), 1974.

Source: Andersson and Holmberg (1980) and Rogers (1982).

Figure 11.   Model migration schedules for Swedish
             data (--- model schedule, —— observed data),
             1974.

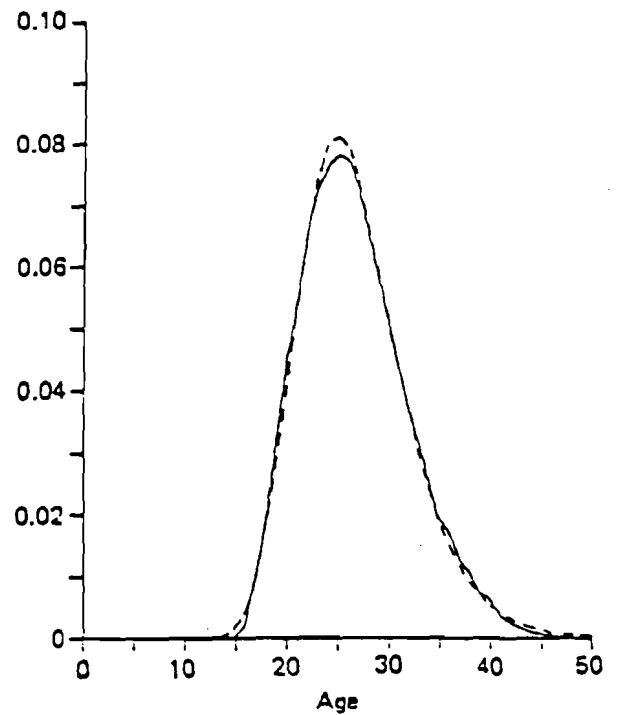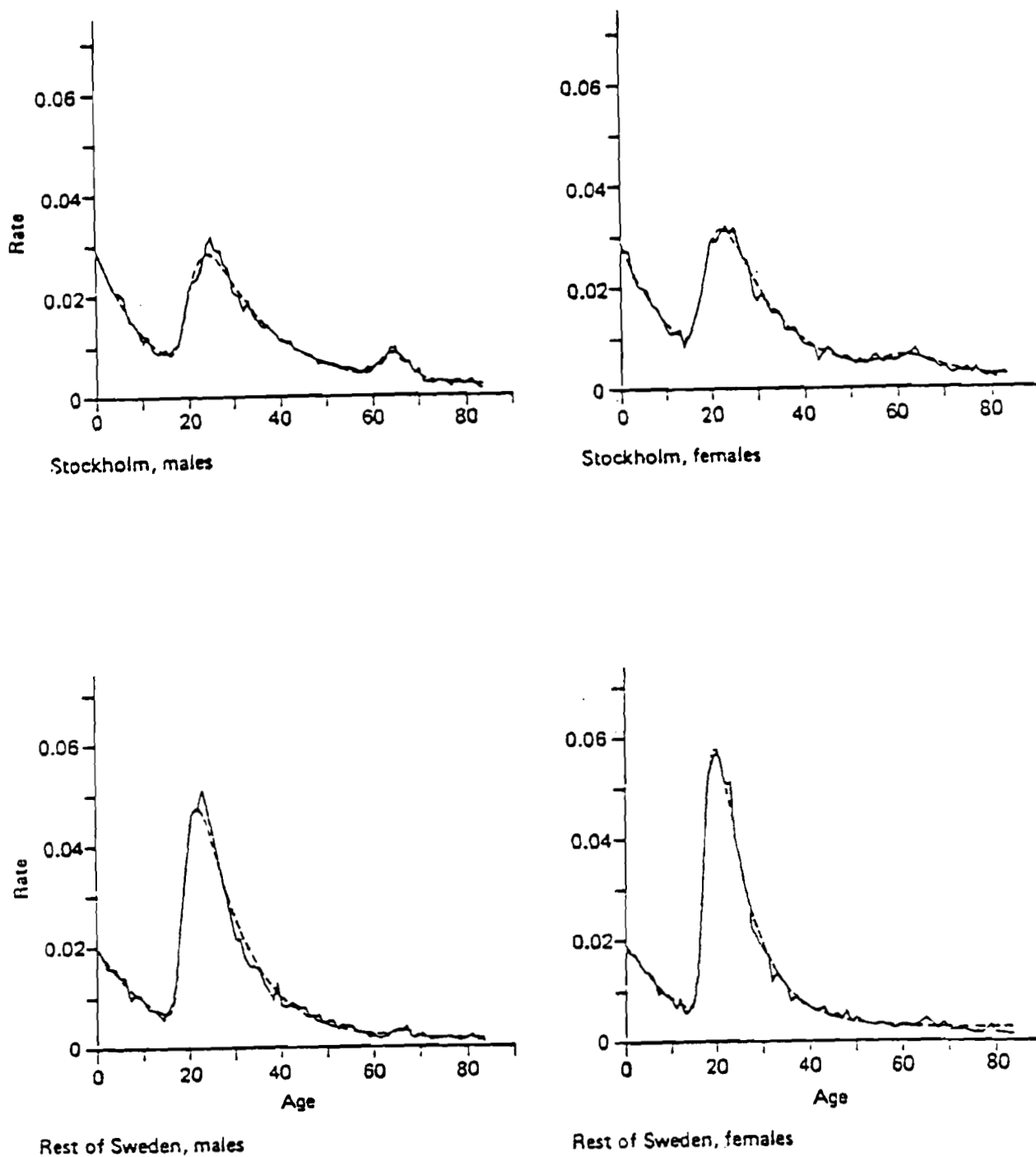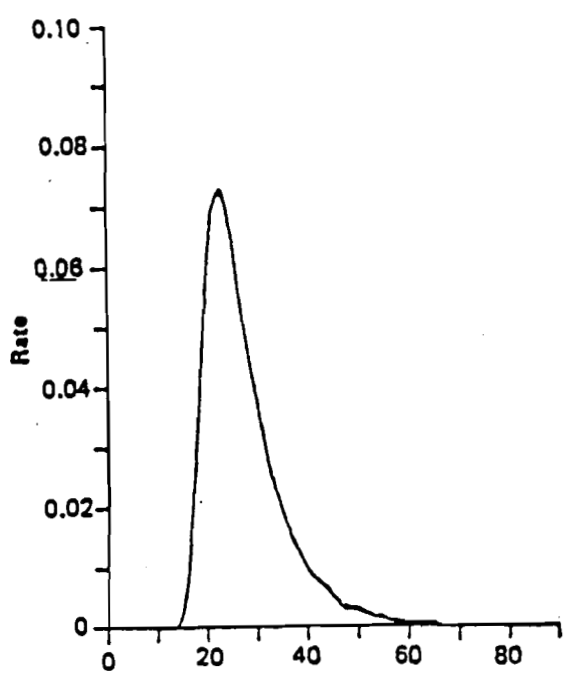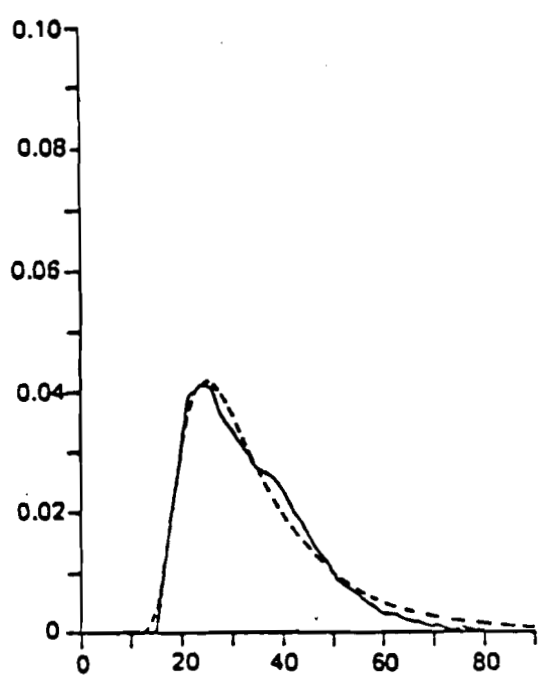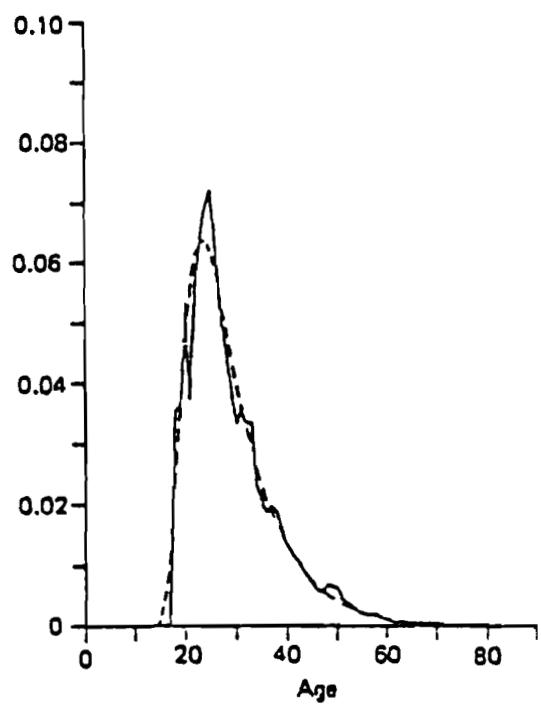Source: Andersson and Holmberg (1980) and Rogers (1982).
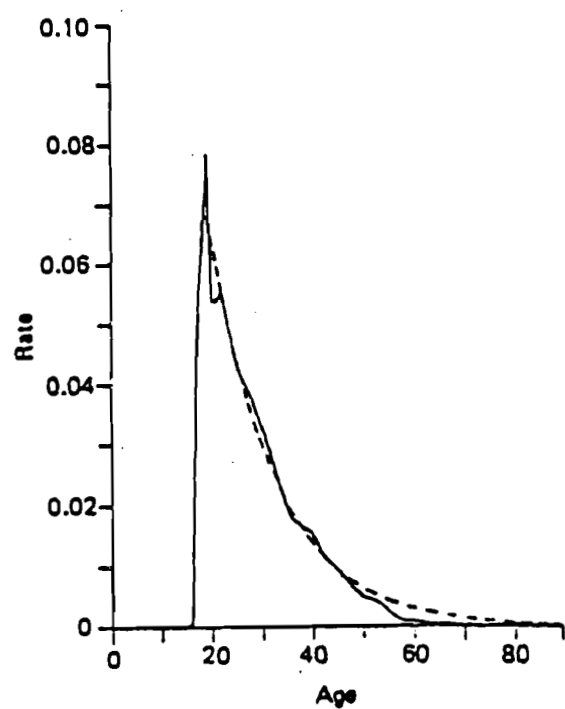
Figure 12. Model schedules of marital status change: Norwegian
females (--- model schedule, — observed data), 1977-1978.

Source: Brunborg et al. (1981) and Rogers (1982).

age-specific probabilities of leaving the population, e.g., dying, within each interval of age from observed data on age-specific exit rates.

Extending simple life tables to recognize several modes of exit from the population gives rise to multiple-decrement life tables. A further generalization of the life table concept arises with the recognition of entries as well as exits. Such increment-decrement life tables allow for multiple movements between several states, for example, transitions between marital statuses and death (single, married, divorced, widowed, dead), or between labor force statuses and death (employed, unemployed, retired, dead).

Multiple radix increment-decrement life tables that recognize several regional populations, each with a region-specific schedule of mortality and several destination-specific schedules of internal migration are called multiregional life tables. They represent the most general class of life tables and were originally developed for the study of interregional migration between interacting multiple regional populations. Their construction is usually initiated by estimating a matrix of age-specific death and migration rates.

One of the most useful statistics provided by a life table is the average expectation of life beyond age x, calculated by applying age-specific probabilities of survival to a hypothetical cohort of babies and then observing at each age their average length of remaining life in each state.

Table 3 presents four sets of expectations of life at birth, associated with our illustration focusing on Swedish females in 1974. The first is for the total population; the second is for a two-region disaggregation of this total into the populations of Stockholm and the

TABLE 3. A COMPARISON OF MODEL-BASED AND DATA-BASED MULTISTATE LIFE TABLES: EXPECTATIONS OF LIFE AT BIRTH, BY REGION OF RESIDENCE AND STATE OF EXISTENCE

1-STATE

BORN IN SWEDEN
77.9 (78.2)

2-STATES

LIVING IN	BORN IN	
	STOCKHOLM	REST OF SWEDEN
STOCKH.	37.9 (38.1)	8.2 (8.4)
REST OF SWEDEN	39.8 (40.1)	69.5 (69.8)
TOTAL	77.7 (78.2)	77.7 (78.2)

4-STATES

LIVING AS	BORN IN SWEDEN
NEVER MARRIED	37.2 (37.4)
MARRIED	26.6 (26.3)
WIDOWED	6.0 (6.1)
DIVORCED	8.0 (8.4)
TOTAL	77.8 (78.2)

8-STATES

LIVING IN/AS	BORN IN	
	STOCKHOLM	REST OF SWEDEN
STOCKHOLM		
NEVER MARRIED	22.7 (22.7)	2.5 (2.6)
MARRIED	9.2 (9.2)	3.3 (3.3)
WIDOWED	1.7 (1.7)	0.8 (0.8)
DIVORCED	4.1 (4.4)	1.5 (1.7)
REST OF SWEDEN		
NEVER MARRIED	14.4 (14.6)	34.7 (34.7)
MARRIED	16.5 (16.3)	23.5 (23.3)
WIDOWED	4.2 (4.2)	5.3 (5.4)
DIVORCED	4.8 (5.0)	6.1 (6.3)
TOTAL	77.7 (78.2)	77.7 (78.2)

rest of Sweden; the third is for a four-state disaggregation of the Swedish total into the never married, married, widowed, and divorced categories; and the fourth is for an eight-state disaggregation that combines the two regional states with the four marital states.

Two sets of life expectancies are distinguished in Table 3. Those set out in parentheses were obtained using the observed data; those without parentheses were calculated on the basis of rates defined by the model schedules presented in Table 2. The differences are insignificant in all instances, with no deviation exceeding six months.

According to Table 3, a baby girl exposed to the 1974 Swedish mortality regime could expect to live about 78 years. Of this total, she could expect to live approximately 26 years in the married state and 8 years as a divorcee. If the girl was born in Stockholm, she could expect to live just over a half of her total life, and almost two-thirds of her married life, outside her region of birth.

4. MULTISTATE PROJECTION WITH CONSTANT COEFFICIENTS

Multistate generalizations of the classical projection model of mathematical demography typically involve three basic steps. The first ascertains the starting age-by-state distribution and the age-state-specific schedules of fertility, mortality, and interstate flows to which the multistate population has been subject during a past period; the second adopts a set of assumptions regarding the future behavior of such schedules; and the third derives the consequences of applying these schedules to the initial population.

A multistate population projection calculates the state- and age-specific survivors of a population of a given sex and adds to this total the new births that survive to the end of the unit time interval. Given appropriate data, survivorship proportions can be obtained as part of the calculations carried out in developing a multistate life table or from the observed data, and they then can be applied to the initial population. For example, it is possible to simultaneously determine the projected male or female population and its age/marital status/regional distribution from the observed age/marital status/region-specific flows of marital status changes, regional migrant inflows and outflows, deaths, and fertility. The projected population so derived should then be augmented by the numbers of international migrant arrivals and departures (disaggregated by age, marital status, and region of arrival or departure) to give the projected male or female population by age, marital status, and region of residence.

The asymptotic properties of multistate population projections have been extensively studied in mathematical demography. This body of theory draws on the properties of matrices with non-negative elements and establishes the existence of a unique, real, positive, dominant characteristic root and an associated positive characteristic vector to which the population converges as it approaches its stable distribution.

As with most population projection models in the demographic literature, the multistate projection model deals only with a single sex at a time. However, the separate projection of the evolution of the male and female populations generally leads to inconsistencies, such as the

number of married males not coinciding with the number of married females for a given year, the total number of new widows during a year not coinciding with the total number of deaths among married men that year, and so on. Thus it is somewhat unrealistic to project the transitions among individuals of one sex without taking into account parallel transitions among individuals of the other sex. Methods for coping with this inconsistency and incorporating it into a multistate projection process are not yet well developed, but they are discussed, for example, by Sanderson (1981). Such methods are not considered in this paper.

Tables 4 and 5 present four sets of illustrative projections of the 1974 Swedish female population corresponding to the four sets of multistate life expectancies listed earlier in Table 3. All projections were carried out with unchanging age-specific rates. Once again, the numbers in parentheses refer to results obtained with observed data and those without parentheses refer to figures derived by means of model-schedule based computations. And once again the differences between the two are relatively minor.

The projections set out in Tables 4 and 5 show Sweden's female population to be relatively stationary over the next 30 years with the one- and two-state projections showing a very slight increase and the four- and eight-state results indicating a very small decrease. The annual rate of growth in the year 2004 is negative in all instances, however.

TABLE 4. A COMPARISON OF MODEL-SCHEDULE-BASED AND DATA-BASED MULTISTATE POPULATION PROJECTIONS

POPULATION PROJECTIONS

VARIABLE	1-STATE SWEDEN TOTAL	2-STATES STOCKHOLM	2-STATES R.SWEDEN	2-STATES SWEDEN TOTAL	4-STATES NEVER MARRIED	4-STATES MARRIED	4-STATES WIDOWED	4-STATES DIVORCED	4-STATES SWEDEN TOTAL
1974									
POP (000)	4,099	767	3,332	4,099	1,659	1,890	385	164	4,099
M. AGE	38.6	38.2	38.7	38.6	20.5	47.0	70.5	49.7	38.6
SHARE	100.0	18.7	81.3	100.0	40.5	46.1	9.4	4.0	100.0
2004									
POP (000)	4,111 (4,120)	732 (740)	3,369 (3,381)	4,101 (4,121)	1,653 (1,659)	1,541 (1,528)	408 (417)	416 (434)	4,018 (4,038)
M. AGE	41.0 (41.1)	41.4 (41.5)	40.9 (41.0)	41.0 (41.1)	23.8 (23.8)	48.6 (48.5)	74.1 (73.8)	54.3 (54.8)	41.6 (41.6)
SHARE	100.0 (100.0)	17.8 (18.0)	82.2 (82.0)	100.0 (100.0)	41.1 (41.1)	38.4 (37.8)	10.2 (10.3)	10.3 (10.7)	100.0 (100.0)
G. RATE	-0.0014 (-0.0014)	-0.0030 (-0.0028)	-0.0011 (-0.0011)	-0.0014 (-0.0014)	0.0001 (0.0001)	-0.0064 (-0.0067)	-0.0075 (-0.0085)	0.0074 (0.0088)	-0.0024 (-0.0024)
STABLE									
M. AGE	41.7 (41.8)	42.9 (43.0)	42.7 (42.8)	42.7 (42.8)	30.2 (30.3)	48.8 (48.6)	74.1 (73.7)	58.0 (58.7)	44.1 (44.2)
SHARE	100.0 (100.0)	16.9 (17.1)	83.1 (82.9)	100.0 (100.0)	43.7 (43.7)	35.5 (34.9)	9.5 (9.5)	11.3 (11.9)	100.0 (100.0)
G. RATE	-0.0017 (-0.0016)	—	—	-0.0041 (-0.0041)	—	—	-0.0065 (-0.0064)	—	—

TABLE 5. A COMPARISON OF MODEL-SCHEDULE BASED AND DATA-BASED MULTISTATE POPULATION PROJECTIONS

8 - STATE POPULATION PROJECTION

VARIABLE	STOCKHOLM				REST OF SWEDEN				SWEDEN TOTAL
	NEVER MARRIED	MARRIED	WIDOWED	DIVORCED	NEVER MARRIED	MARRIED	WIDOWED	DIVORCED	
1974									
POP (000)	304	364	68	31	1,355	1,527	317	132	4,099
M. AGE	20.6	46.0	70.2	48.8	20.5	47.2	70.6	49.9	38.6
SHARE	7.4	8.9	1.7	0.8	33.1	37.3	7.7	3.2	100.0
2004									
POP (000)	277 (282)	258 (260)	71 (72)	107 (114)	1,379 (1,379)	1,281 (1,274)	340 (346)	304 (313)	4,017 (4,041)
M. AGE	24.4 (24.2)	47.5 (47.4)	74.3 (74.1)	53.0 (53.5)	23.6 (23.7)	48.8 (48.7)	73.9 (73.7)	54.8 (55.4)	41.5 (41.6)
SHARE	6.9 (7.0)	6.4 (6.4)	1.8 (1.8)	2.7 (2.8)	34.3 (34.1)	31.9 (31.5)	8.5 (8.6)	7.6 (7.8)	100.0 (100.0)
G. RATE	-0.0020 (-0.0016)	-0.0086 (-0.0087)	-0.0109 (-0.0125)	0.0061 (0.0086)	0.0006 (0.0007)	-0.0059 (-0.0061)	-0.0066 (-0.0075)	0.0074 (0.0086)	-0.0024 (-0.0023)
STABLE									
M. AGE	30.2 (29.9)	47.4 (47.2)	74.0 (73.5)	56.7 (57.6)	30.1 (30.2)	49.0 (48.9)	74.0 (73.7)	58.3 (59.1)	44.0 (44.1)
SHARE	6.8 (7.0)	5.6 (5.7)	1.5 (1.5)	2.7 (3.0)	37.0 (36.7)	29.9 (29.4)	8.1 (8.1)	8.3 (8.6)	100.0 (100.0)
G. RATE	——	——	——	——	-0.0064 (-0.0063)				——

5. MULTISTATE PROJECTION WITH VARIABLE COEFFICIENTS

The population projections summarized in Tables 4 and 5 indicate that parametrized projections provide a reasonably accurate alternative to ones carried out with observed rates. The principal advantage of such model-schedule-based projections, however, lies not in their ability to replicate the results of data-based constant coefficient (CC) projections but rather in the transparency and flexibility that they bring to variable coefficient (VC) projections.

It is widely recognized that the age-specific rates of a demographic event (e.g. death) or flow (e.g. migration) are interdependent and vary differentially with changes in levels. Thus to project them individually, age by age, or to assume that they will change uniformly at all ages, is to invite potentially unreasonable results. An important feature of parametrized projection procedures is the ease with which they permit the introduction of changes in age profiles that are associated with changes in levels.

The relational approach proposed by Brass (1971) is a convenient method for altering both the level and the age profile of each parametrized model schedule used in the projection process. Two parameters are associated with every schedule, and extrapolating their values into the future produces the needed projected values of various age-specific components of change.

An alternative to Brass's method is the regression approach, used by Coale and Demeny (1966 and 1983) and others, which embodies a correlational perspective that associates rates at different ages to an index of level. We adopt this approach and use the data on Swedish females to illustrate the transparency and flexibility of parametrized projection methods.

Perhaps the simplest and most straightforward way of projecting the values to be assumed by a model schedule's parameters is to associate such values with the schedule's level as measured by its gross transition rate. Forecasts of levels, then, lead to corresponding forecasts of profiles. To implement this procedure, regression equations were developed that set each model schedule parameter as a dependent variable with the appropriate gross transition rate as the independent variable. Ideally all parameters should have been estimated simultaneously with a system of equations. However, for the purposes of our illustration, it was felt that single-equation ordinary least squares estimates were adequate.

5.1 Estimation of the Parameters

In the Swedish illustrative VC projection, the age profiles of the mortality, widowhood, and remarriage model schedules were held constant; each of the four fertility model schedule parameters was regressed on the gross reproduction rate, of the seven migration model schedule parameters on the gross migraproduction rate (the retirement peak was treated exogenously and held fixed), of the four first-marriage model schedule parameters on the gross first-marriage transition rate, and of the four divorce model schedule parameters on the gross divorce transition rate.

Fifty-three regional model fertility schedules from seven countries, 330 model migration schedules from eight countries, and a time series of 30 model first-marriage and 12 model divorce schedules from Australia constituted the data bank used to estimate the various regression equations. All intercept terms were rescaled to produce exact fits to initial conditions, i.e. such that when 1974 gross transition rates were entered into the regression equations, 1974 model schedule parameters were obtained.

An example of the unrescaled estimated regression results is provided by the four fertility equations:

$$a = 0.063 + 0.028(GRR) \qquad (r = 0.921)$$

$$\mu = 28.39 + 4.21(GRR) \qquad (r = 0.899)$$

$$\alpha = 0.244 + 0.027(GRR) \qquad (r = 0.519)$$

$$\lambda = 0.137 - 0.022(GRR) \qquad (r = -0.924)$$

With the exception of λ, all parameters increase with increases in fertility level. The "height" parameter a rises, the "position" parameter μ shifts to the right on the age axis, and the "descent" parameter α becomes larger, making the double exponential curve steeper at ages past the mode.

5.2 Assumptions About Levels

To develop the illustrative projection, it was assumed that the gross reproduction rate would increase linearly from 0.82 to 0.90 in the Stockholm region by 1984 and from 0.93 to unity in the rest of Sweden. Moreover, it was assumed that this rise in fertility would be totally due to increases in non-marital fertility, which would converge to equal levels in both regions. Post-1984 fertility levels were assumed to remain fixed.

Migration out of the Stockholm region was assumed to increase slightly from 1974 to 1984, growing linearly from an initial gross migraproduction rate (GMR) of 1.43 to 1.50. The migration level in the reverse direction was held constant at a GMR of 0.29. Differentials in this rate among marital statuses were assumed to diminish, such that by 1984 the same rate would be exhibited by married, widowed, and divorced females. The GMRs of the never-married population, however, were set at higher levels over the decade. All levels beyond 1984 were assumed to remain fixed.

The level of first-marriage was assumed to continue its pattern of decrease, with an expected 10 percent fall in the gross transition rate. Divorce, on the other hand, was assumed to exhibit a 20 percent increase in its gross transition rate. Whereas changes in first marriage rates were introduced in both regions, changes in divorce rates were permitted only in the non-Stockholm (rest of Sweden) region. Once again all changes were assumed to occur in a linear pattern over the decade 1974 to 1984 and were held fixed thereafter.

Table 6 sets out the initial (1974) and final (1984 onwards) gross transition rates that produced the illustrative projection. Figure 13 shows graphically the changes in age profiles produced by the assumed changes in levels.

5.3 Results

Figures 14 and 15 summarize some of the aggregate results produced by the 50-year illustrative projections. The constant coefficient (CC) and variable coefficient (VC) projections are contrasted in Figure 14, for all four levels of disaggregation. Figure 15 focuses on the expected elderly populations in the 8-state CC and VC projections.

Figure 14 vividly demonstrates the extent to which aggregation bias can alter projection totals. Over a period of 50 years an over-projection of some 280,000 people is introduced by aggregating the 8-state model into a 1-state one in the CC projection, growing to 360,000 in the case of the VC projection. The major impact is apparently produced by the aggregation over the four marital states.

All projections show a gradual increase in Sweden's future population, peaking to a maximum either in 1984 or a decade later, and in all but one instance (the single-region VC projection) declining immediately

42

TABLE 6. INITIAL (AND FINAL) GROSS TRANSITION RATES: SWEDEN, FEMALES, 1974

| TO → FROM ↓	STOCKHOLM				REST OF SWEDEN				DEATH	BIRTH
	NEVER MARRIED	MARRIED	WIDOWED	DIVORCED	NEVER MARRIED	MARRIED	WIDOWED	DIVORCED		
STOCKHOLM										
NEVER MARRIED	--	1.38 (1.14)	--	--	1.38 (1.50)	0.06	--	--	1.22	0.61 (0.73)
MARRIED	--	--	2.24	1.43	--	1.04 (1.20)	0.02	0.01	1.21	0.97
WIDOWED	--	0.89	--	--	--	0.0	1.08 (1.20)	--	1.21	0.61 (0.73)
DIVORCED	--	3.47	--	--	--	0.17	--	1.00 (1.20)	1.21	0.61 (0.73)
REST OF SWEDEN										
NEVER MARRIED	0.25	0.06	--	--	--	1.38 (1.14)	--	--	1.27	0.64 (0.73)
MARRIED	--	0.21	0.02	0.01	--	--	2.24	0.71 (0.92)	1.26	1.18
WIDOWED	--	0.0	0.21	--	--	0.86	--	--	1.26	0.64 (0.73)
DIVORCED	--	0.14	--	0.20	--	3.20	--	--	1.26	0.64 (0.73)
SWEDEN TOTAL									1.26	0.91 (0.99)
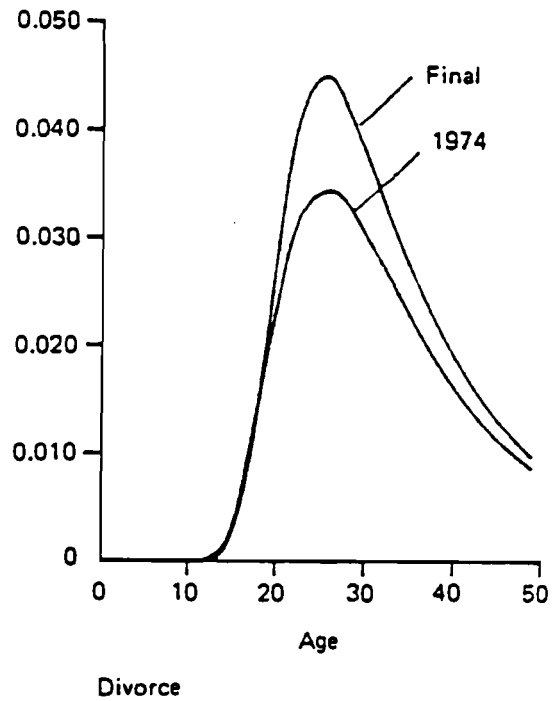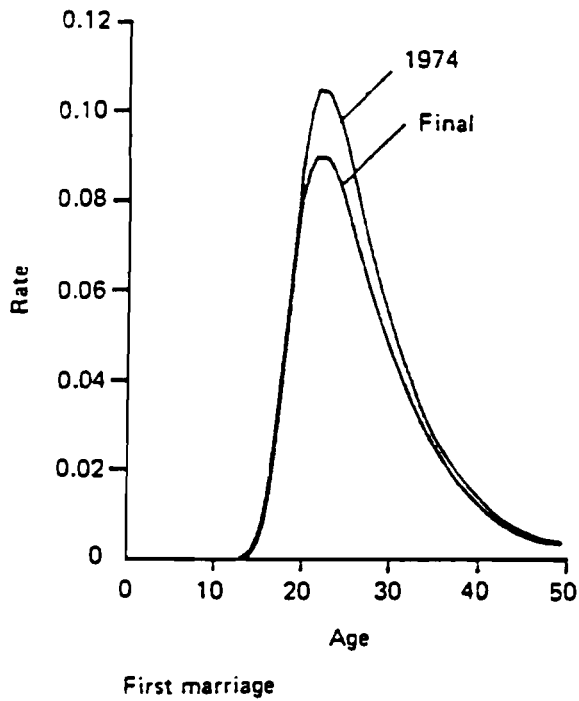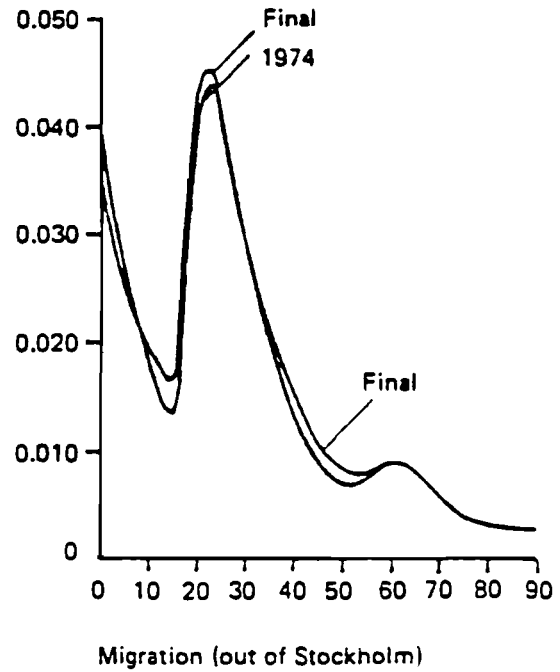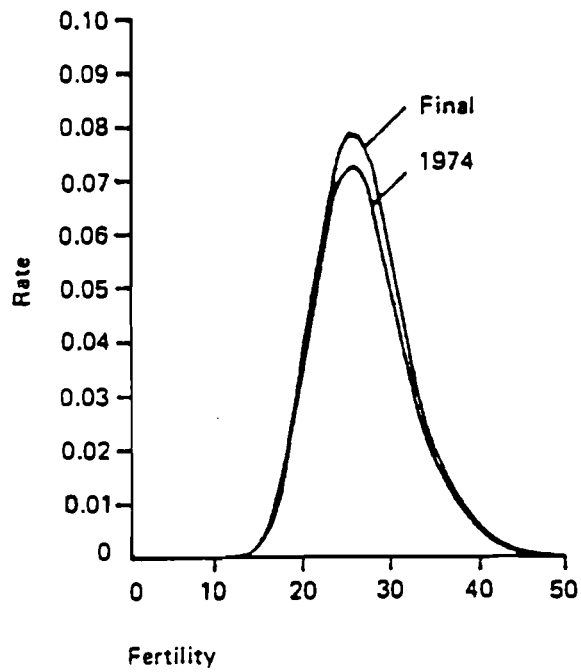
Figure 13. Initial and final model schedules: Swedish females, 1974 and projected.
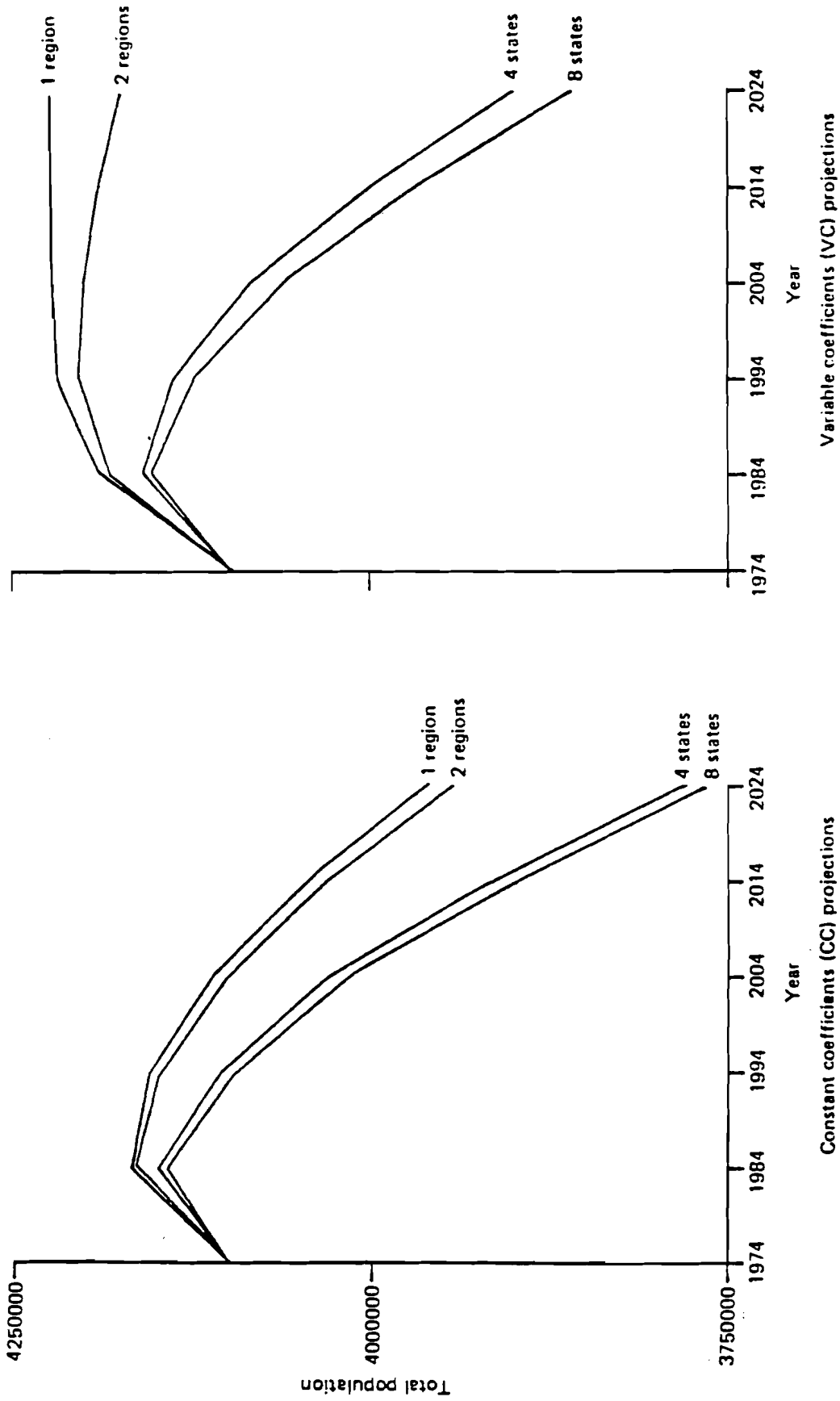
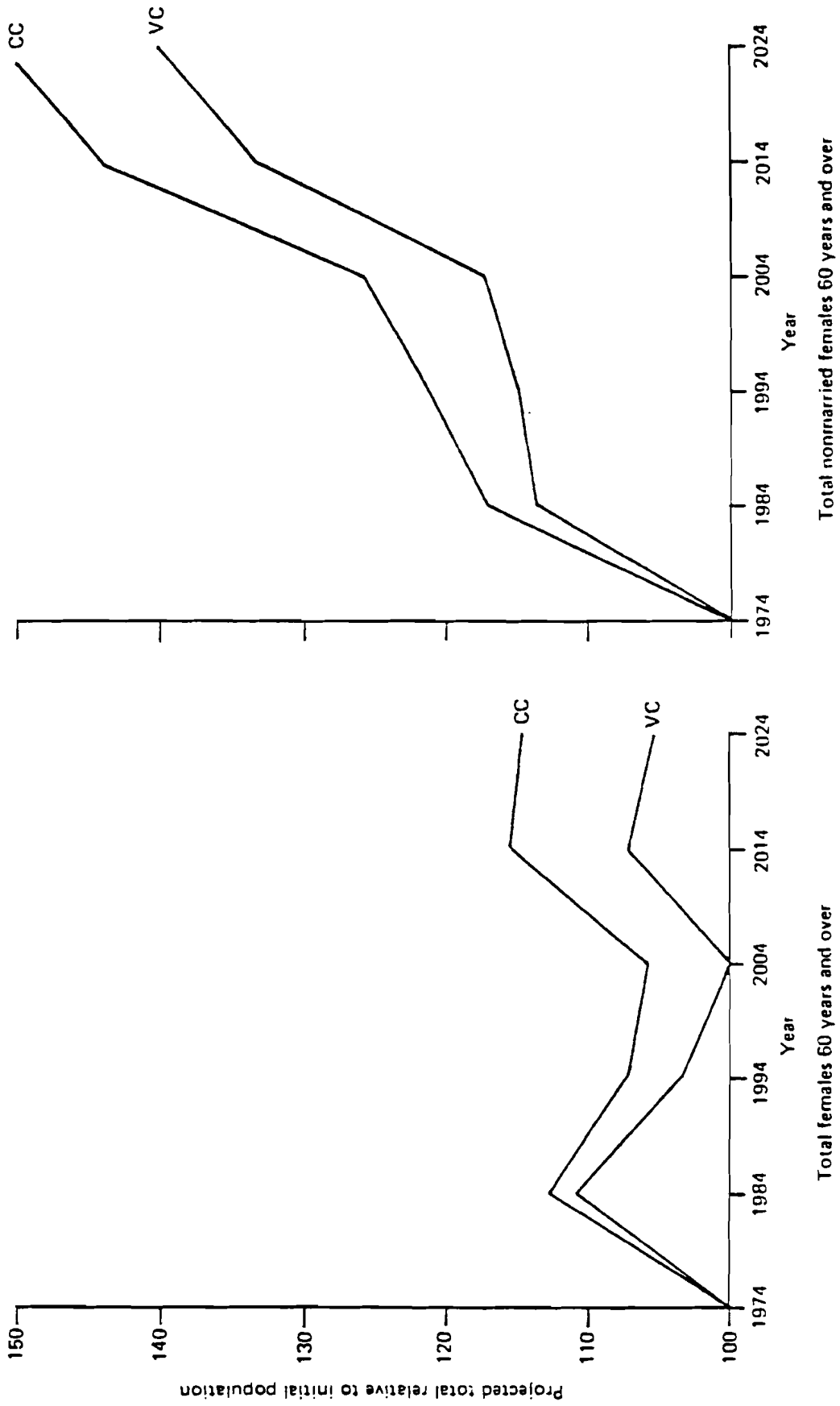Figure 14. Aggregation bias in projections: Swedish females, 1974-2024.

45

Figure 15. Projected elderly population: Stockholm, females, 60 years and over.

thereafter. The lowest total populations in the year 2024 are projected by the CC models and the highest by the corresponding VC models. For example, the eight-state CC projection gives a total population of 3.76 million individuals; the corresponding VC projection increases this sum by a hundred thousand.

Figure 15 indicates that although only modest increases are to be expected in the elderly population over the next decades, the population of the nonmarried elderly should increase dramatically. The CC and the VC projections show a growth of between 40 to 50 percent in this subgroup over the 50-year projection period.

6. CONCLUSION

The research summarized in this paper demonstrates that it is possible to carry out multistate population projections of considerable generality and levels of disaggregation using a parametrized modeling approach that emphasizes model selection in place of demographic accounting. The replacement of observed schedules by model schedules brings both economy and transparency. In the eight-state model, for example, 44 observed schedules containing over 2,000 age-specific rates were replaced by the same number of schedules defined in terms of a total of 227 parameters. This more compact representation of the input data identifies the broad patterns exhibited by the demographic components of growth and change, thereby making more transparent the ways in which model schedule levels and age profiles influence population stocks and flows.

Although the model schedule parameters are not always demographically interpretable, future research is likely to link them to variables that are. Moreover the parameters can be readily extrapolated into the future to produce reasonable age patterns of rates, and their adoption also

permits the introduction of changes in schedules of age-specific rates
that alternative methods do not (e.g., increased mortality due to a rise
in the relative number of deaths attributable to accidents).

Finally, it seems likely that econometric explanatory models, with
parameters instead of rates as dependent variables, will produce results
that are at least as effective as current models that focus on observed
flows or rates.

REFERENCES

Andersson, A. and I. Holmberg (1980) Migration and Settlement: 3.
Sweden. RR-80-5. Laxenburg, Austria: International Institute for
Applied Systems Analysis.

Benjamin, B. and J. H. Pollard (1980) The Analysis of Mortality and Other
Actuarial Statistics. London: Heinemann.

Brass, W. (1971) On the scale of mortality. Pages 86-110 in Biological
Aspects of Demography, edited by W. Brass. London: Taylor and
Francis.

Brass, W. (1974) Perspectives in population prediction: illustrated by
the statistics of England and Wales. Journal of the Statistical
Society, Series A, 134(4): 532-570.

Brass, W. (1977) Population projections for planning and policy. Paper
delivered at the Eighth Annual Summer Seminar in Population, East-West
Center, University of Hawaii, June 1977.

Brooks, C., D. Sams and P. Williams (1980) A Time Series of Smooth
Approximations for Age, Sex and Marital Status Specific Death Rates in
Australia, 1950/51 to 1975/76 with Projections to the Year 2000.
Research Memorandum. Melbourne, Australia: Impact Project Research
Centre.

Brown, H.P. and A.R. Hall (1978) Australian Demographic Data Bank, Volume
I: Recorded Vital Statistics 1921-1976. Canberra: Research School of
Social Sciences, Australian National University.

Brunborg, H., J. Monnesland and R. Selmer (1981) Framskriving av
Folkemengden etter. Ekteskapelig Status 1979-2025 (Projections of the
Population by Marital Status 1979-2025). Rapporter 81/12. Oslo:
Central Bureau of Statistics.

Coale, A.J. and P. Demeny (1966) Regional Model Life Tables and Stable
Populations. Princeton, New Jersey: Princeton University Press.

Coale, A.J. and P. Demeny (1983) Regional Model Life Tables and Stable
Populations. Second edition. New York: Academic Press.

Coale, A.J. and D.R. McNeil (1972) The distribution by age of the
frequency of first marriage in a female cohort. Journal of the
American Statistical Association 67:743-749.

Coale, A.J. and J. Trussell (1974) Model fertility schedules: variations
in the age structure of childbearing in human populations. Population
Index 40(2): 185-206.

Heligman, L. and J.H. Pollard (1980) The age pattern of mortality.
Journal of the Institute of Actuaries, (1980) 107:49-80.

Henry, L. (1961) Some data on natural fertility. Eugenics Quarterly
8(2):81-91.

Hoem, J.M. and M.S. Fong (1976) A Markov Chain Model of Working Life Tables. Working Paper No. 2. Copenhagen: Laboratory of Actuarial Mathematics, University of Copenhagen.

Hoem, J.M., D. Madsen, J.L. Nelsen, E.M. Ohlsen, H.O. Hansen and B. Rennermahn (1981) Experiments in modelling recent Danish fertility curves. Demography 18(2):231-244.

Keyfitz, N. (1972) On future population. Journal of the American Statistical Association 67:347-363.

Powell, A. A. (1977) The Impact Project: An Overview. First progress report of the Impact Project, volume 1. Canberra: Australian Government Printing Service.

Rogers, A. (1982) Parametrized Multistate Population Dynamics. WP-82-125. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A. and P. Williams (1982) A Framework for Multistate Demoeconomic Modeling and Projection, with an Illustrative Application. WP-82-69. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A. and L.J. Castro (1981) Model Migration Schedules. RR-81-30. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A., R. Raquillet and L.J. Castro (1978) Model migration schedules and their applications. Environment and Planning A 10(5):475-502.

Sams, D. (1981) The Double Exponential Function. Research Memorandum. Melbourne, Australia: Impact Project Research Centre.

Sanderson, W.C. (1981) An Analytically Based Two-Sex Marriage Model and Maximum Likelihood Estimates of its Parameters: Austria, 1979. WP-81-130. Laxenburg, Austria: International Institute for Applied Systems Analysis.

United Nations (1967) Methods of Estimating Basic Demographic Measures from Incomplete Data. Manual IV. New York: Department of International Economic and Social Affairs, United Nations.

United Nations (1983) Indirect Techniques for Demographic Estimation. Manual X. New York: Department of International Economic and Social Affairs, United Nations.

Williams, P. (1981) Marriage and Divorce in Australia: A Time Series of Fitted Distributions, 1921/22 to 1975/76. Working Paper B-20. Melbourne, Australia: Impact Project Research Centre.