DEMOGRAPHY© Volume 18, Number 3 August 1981

METHODS FOR COMPARING THE MORTALITY EXPERIENCE OF HETEROGENEOUS POPULATIONS

Kenneth G. Manton
Center for Demographic Studies, and Department of Community and Family Medicine, Duke University, Durham, North Carolina 27706

Eric Stallard
Center for Demographic Studies, Duke University, Durham, North Carolina 27706

James W. Vaupel
Institute of Policy Sciences and Public Affairs, Duke University, Durham, North Carolina 27706

Abstract—Methods are presented which produce Maximum Likelihood Estimates (MLE) of the degree of heterogeneity in individual mortality risks under a variety of assumptions about the age trajectory of those mortality risks. With these estimates of the degree of population heterogeneity it is possible to adjust comparisons of mortality risks across populations for the effects of population heterogeneity, differential mortality selection, and different age trajectories of the force of mortality. These methods are demonstrated by applying a variety of standard assumptions about the age trajectory of the force of mortality to the analysis of a broad range of cohort mortality data for the U.S. and Swedish populations. The estimates of the degree of heterogeneity, produced under all of the selected force of mortality models, consistently indicated a considerable degree of heterogeneity in mortality risks.

INTRODUCTION

In a previous paper (Vaupel et al., 1979a) a model was presented which illustrated the bias produced in life table parameters by the operation, over age, of systematic mortality selection on a heterogeneous population. The effects of such bias on a variety of different types of mortality analyses were discussed. Two important types of mortality analyses—comparison of the mortality experience of different populations, or of different birth cohorts for the same population—are particularly subject to the bias of population heterogeneity since the comparisons could involve populations under very different mortality conditions. Maximum likelihood (ML) estimation procedures which deal with this bias are presented in this paper. These procedures are designed to permit comparisons of the mortality experience across cohorts within the same population by explicitly estimating the magnitude of heterogeneity in that population while simultaneously adjusting the comparisons for that level of heterogeneity.

These procedures are applied to comparisons of the mortality experience across successive cohorts of males and females in the U.S. and Swedish populations. Estimates of the magnitude of population heterogeneity in these populations are produced under a variety of assumptions about the age trajectory of the force of mortality for individuals. The estimates of heterogeneity under all of the models selected were consistent in indicating a fairly high degree of population heterogeneity.

The remainder of this paper is organized into three sections. In the methods section we (a) present a mathematical

model of population heterogeneity and mortality selection, (b) discuss procedures for estimating the degree of heterogeneity and (c) present a variety of models of the age trajectory of individual mortality risks that will be employed in the analysis. In the second section we discuss the preparation of the cohort mortality data. Finally, in the results section we both review the estimation of the degree of population heterogeneity from the cohort data and evaluate that degree of heterogeneity for our perception of individual mortality risks.

METHODS

This section is organized into three parts. First, we briefly review the model presented in Vaupel et al. (1979a), generalize the results presented in that paper, and establish necessary definitions and mathematical notation. Second, the likelihood function for the heterogeneity model presented in Vaupel et al. (1979) is derived so that maximum likelihood procedures (see Appendix) can be used to generate (a) parameter estimates, (b) asymptotic confidence intervals for our parameter estimates, and (c) likelihood chi-squared tests for hierarchially structured models. Third, the models of the age trajectory of individual mortality risks employed in generating estimates of population heterogeneity are presented.

Heterogeneous Population Mortality Model

The heterogeneous population mortality model presented in Vaupel et al. (1979a) introduced the concept of individual "frailty" into life table computations. This was operationalized by defining an individual's frailty to represent the proportional increase or decrease in the force of mortality operating on that individual as compared to some standard force of mortality. For convenience, this standard force of mortality was scaled so that it was equal to the average force of mortality faced by the cohort at birth. It was further assumed that each individual's relative frailty remained constant throughout his or her life. These assumptions led directly to the concept of the "standard individual," defined to be an individual who, at each age x, is subjected to the standard force of mortality specific to age x. Thus, the standard individual is assumed to have a relative frailty of unity or one (1.0).

To deal with individual differences in the force of mortality, we define $\mu_i(x, y, z)$ to be the force of mortality for an individual in population group i, at exact age x, at some instant in time y, and with a relative frailty z. Under the assumption that an individual's frailty represents the proportional increase or decrease in the standard force of mortality, it follows that

$$\mu_i(x, y, z) = z\mu_i(x, y), \qquad (1a)$$

where $\mu_i(x, y)$ is the standard force of mortality in population group i, at exact age x, at time y. For simplicity, subscripts and arguments will be suppressed throughout this paper whenever they are not essential to convey meaning. Specifically, the symbols i (population), x (age), and y (time) will be left implicit in most of the following. Thus equation (1a) is taken to be equivalent to:

$$\mu(z) = z\mu, \qquad (1b)$$

where μ is the standard force of mortality.

Because different individuals are assumed to have different levels of frailty, it follows from (1b) that the members of each cohort will, at any given age, face differential mortality risks. This implies that the distribution of frailty in the cohort will change over time as mortality selectively removes the more frail cohort members. This process of mortality selection also implies that the cohort survivors at any given age will *not* represent a random sample of the cohort survivors at some earlier age (see Figures 1 and 2). Instead, these cohort survivors represent a sample of the earlier survivors with systematically lower frailty.

The extent of the selection for lower frailty can be determined if the propor-

tion of survivors from the initial observation time is known and if it is assumed that frailty at the initial observation time is gamma distributed (Vaupel, 1979a). While this latter assumption is convenient because it simplifies the mathematics, these estimates can be obtained for any distribution which has a moment generating function (MGF). To see this, it is necessary to examine the relationship between the force of mortality $\mu(z)$ faced by an individual at some given level of frailty and the probability of survival to each age $s_i(x, y, z)$ or, more simply, $s(z)$. The reader may recall that the force of mortality is often defined as:

$$\mu(z) = \frac{\partial}{\partial x}\{-\ln[s(z)]\}, \tag{2a}$$

which yields:

$$s(z) = e^{-H(z)}, \tag{2b}$$

where

$$H(z) = H(x, y, z) \tag{2c}$$

$$= \int_{x_0}^{x} \mu(t,(y - x) + t, z)\, dt. \tag{2d}$$

The symbol $H(z)$ denotes the cumulative force of mortality faced by an individual over the age interval (x_0, x). Note that we may set x_0 equal to the age of initial observation. For example, if we follow the survivorship from birth of a given cohort, then we set $x_0 = 0$.

The distribution function of the cumulative forces of mortality for cohort members, $F[H(z)]$, is defined by the relationship:

$$dF[H(z)] = f[H(z)]\, dH(z), \tag{3}$$

where $f[H(z)]$ is the probability density function (pdf) of $H(z)$. Note that $dF[H(z)]$ is the probability that a given individual's cumulative force of mortality is in the interval $(H(z), H(z) + dH(z))$. Thus, using (2b), we can write the expected value of the survivorship proportion as:

$$\bar{s} = E[s(z)] \tag{4a}$$

$$= \int_0^{\infty} e^{-H(z)} f[H(z)] dH(z) \tag{4b}$$

$$= \text{MGF}[-H(z)]. \tag{4c}$$

In writing (4) we assume that individuals are stochastically independent. Thus, the proportion surviving is given by evaluating the MGF of the negative cumula tive force of mortality. Note that equation (4b) is actually a special case of the MGF with the auxiliary variable t (in $e^{tH(z)}$) set to the value (-1). Thus, the MGF in (4c) is written as a function of the random variable $(-H(z))$. Obviously, this does not require any assumption concerning the distribution of frailty, z, at the initial observation time x_0. However, if $H(z)$ is assumed to be a fixed (i.e., non-random) function of z, then the initial distribution of frailty may be introduced in (4) by appropriate change of variables in $dF(H(z))$.

The average force of mortality at age x can be analytically determined to be a negative cumulant function (C) by analogy to (2a), i.e.,

$$\bar{\mu} = \frac{\partial}{\partial x}\{-\ln[\bar{s}]\} \tag{5a}$$

$$= \frac{\partial}{\partial x}\{-\ln[\text{MGF}[-H(z)]]\} \tag{5b}$$

$$= \frac{\partial}{\partial x}\{-C[-H(z)]\} \tag{5c}$$

$$= \frac{\partial}{\partial x}\left\{-C\left[-z\int_{x_0}^{x} \mu(t)\, dt\right]\right\} \tag{5d}$$

where C denotes the cumulant function (Kendall and Stuart, 1969, p. 67). From equation (5c) we see that the cohort force of mortality $\bar{\mu}$ may be equated to the partial derivative of the negative cumulant function of the negative cumulative force of mortality over the age interval (x_0, x). This implies (5d) that the cohort force of mortality $\bar{\mu}$ and the standard force of mortality μ will be different in heterogeneous

populations. The extent of this difference determines the degree of bias to be expected in mortality analyses which do not take into account the relationship in (5). In other words, comparisons of the mortality experience of heterogeneous populations which employ ratios or differences of the age specific cohort forces of mortality, $\bar{\mu}_i$'s, are subject to the difficulty that for heterogeneous populations the observed $\bar{\mu}_i(x)$ represents a function of both the standard force of mortality at age x and the proportion of the cohort surviving to that age. If the proportions surviving to age x differ between the populations being compared, the comparison will not represent the differential risks faced by individuals in the two populations.

Likelihood Function for Gamma Distributed Frailty

In this section we present the derivation of the likelihood function. The likelihood derivation is accomplished in three steps. First, we will show that, if frailty is gamma distributed at the initial observation time, then equations (4) and (5) imply that $\bar{\mu}$ is a simple function of μ and $\bar{s}$. Second, drawing upon results presented in Vaupel et al. (1979a), we will show that the force of mortality among those who die at any given age is gamma distributed with $\bar{\mu}$ as one parameter. Third, with $\bar{\mu}$ expressed as a function of μ and $\bar{s}$, we employ this gamma distribution to derive the sampling distribution of the average (mean) force of mortality in the cohort. Thus, the observed cohort force of mortality will be treated as a random variable sampled from a fully specified gamma distribution with mean $\bar{\mu}$.

i. Functional relationship between $\bar{\mu}$ and μ, $\bar{s}$.

Following Vaupel et al. (1979a), we assume that the initial distribution of individual frailty is the gamma distribution. This allows us to write the probability that a given individual's frailty is in the interval $(z, z + dz)$ as:

$$dF(z) = f(z)\, dz, \tag{6}$$

where

$$f(z) = \lambda^k z^{k-1} e^{-\lambda z}/\Gamma(k) \tag{7a}$$

is the gamma pdf. As shown in (7a), the gamma pdf has two parameters—a scale parameter λ and a shape parameter k. These parameters are combined to determine the mean and variance of z as follows:

$$\begin{aligned} \bar{z} &= E(z) \\ &= k/\lambda \end{aligned} \tag{8a}$$

$$\begin{aligned} \text{var}(z) &= E\{[z - \bar{z}]^2\} \\ &= \bar{z}^2/k. \end{aligned} \tag{8b}$$

Equation (8b) shows that the variance of z is inversely proportional to k. Thus, the higher the value of k, the lower will be the degree of heterogeneity. Indeed, from (8a) and (8b) we find the coefficient of variation to be the inverse square root of k, i.e.,

$$CV(z) = 1/k^{1/2}. \tag{8c}$$

This suggests that the degree of heterogeneity can be summarized by the single parameter k.

By solving (8a) for λ and substituting the result in (7a) we find that:

$$f(z) = \left(\frac{k}{\bar{z}}\right)^k z^{k-1}\, e^{-kz/\bar{z}}/\Gamma(k), \tag{7b}$$

which shows that the gamma pdf may be conveniently reparameterized in terms of the mean $\bar{z}$ and shape k. We also have, by definition, the relationship

$$\bar{\mu}(x_0) = \bar{z}\, \mu(x_0) \tag{9}$$

which, in view of our assumption that the standard force of mortality and the cohort force of mortality are equal at birth, implies that

$$\bar{z} = 1,\ x_0 = 0. \tag{10}$$

Equation (10) implies that the initial gamma pdf is a function of only one unknown parameter, k. This reduction to one parameter is achieved because a gamma variable may be rescaled to produce another gamma variable without changing the shape parameter. We will now employ this property to solve equations (4) and (5).

By substituting equation (1) in (2d), we find that:

$$H(z) = z\,H, \tag{11}$$

where

$$H = \int_{x_0}^{x} \mu(t)\,dt$$

is the cumulative standard force of mortality over the age interval (x_0, x). Substitution of (10) and (11) in (7b) allows (3) to be written as:

$$dF[H(z)] = \left(\frac{k}{H}\right)^k [H(z)]^{k-1} \frac{e^{-kH(z)/H}}{\Gamma(k)}\,dH(z). \tag{12}$$

Equation (12) shows that $H(z)$ is also a gamma variable with mean H and shape k. Reference to any standard statistical text (e.g., Hastings and Peacock, 1975) provides the moment generating function required to solve (4) and (5):

$$\bar{s} = (1 + H/k)^{-k} \tag{13}$$

$$\bar{\mu} = \mu/(1 + H/k). \tag{14}$$

From equations (13) and (14), it is obvious that:

$$\bar{\mu} = \mu\,\bar{s}^{1/k}. \tag{15}$$

Equation (15) shows that, if frailty is gamma distributed, the cohort force of mortality is the product of the standard force of mortality and the $(1/k)th$ power of the survivorship proportion.

ii. Distribution of force of mortality among those who die at age x.

It is important to note that equation (15) is the basic functional relationship in Vaupel et al. (1979a). This permits us to draw upon two results presented in that paper: (1) frailty among the survivors at any age x is gamma distributed with the same value of the shape parameter k as at birth. This means that the force of mortality among the survivors at age x will be gamma distributed with mean $\bar{\mu}$ and shape k. (2) Frailty among those who die at any age x is also gamma distributed with the same age dependent scale parameter $\lambda(x)$ as among those surviving to age x but with shape parameter $k + 1$. This means that the force of mortality among those who die at age x will be gamma distributed with shape $k + 1$ and mean $\bar{\mu}^*$, where

$$\bar{\mu}^* = \bar{\mu}(k + 1)/k. \tag{16}$$

Thus, the probability that the force of mortality among those who die at age x is in the range $(\mu^*(z), \mu^*(z) + d\mu^*(z))$ may be written in the form:

$$dF[\mu^*(z)] = \left(\frac{k}{\bar{\mu}}\right)^{k+1} [\mu^*(z)]^k \frac{e^{-k\,\mu^*(z)/\bar{\mu}}}{\Gamma(k+1)}\,d\mu^*(z). \tag{17}$$

iii. Sampling distribution of the average force of mortality.

In order to derive the sampling distribution of $\mathring{\mu}$, the observed cohort force of mortality, we need to do two things. First, using $d_i(x, y)$, or more simply d, to denote the number of deaths in population group i at age x at time y, we obtain the distribution of the sum of μ^* for the d persons (i.e., $\sum_{m=1}^{d} \mu_m^*$) from the d-fold convolution of (17). Second, we rescale this distribution by the factor $k/[d \cdot (k + 1)]$, yielding:

$$dF[\mathring{\mu}] = \left[\frac{d \cdot (k+1)}{\mu \cdot \bar{s}^{1/k}}\right]^{d\cdot(k+1)} \cdot \frac{\mathring{\mu}^{d\cdot(k+1)-1} \cdot e^{-d\cdot(k+1)\cdot \mathring{\mu}/(\mu\cdot\bar{s}^{1/k})}}{\Gamma[d \cdot (k+1)]} \tag{18}$$

with mean and variance

$$E(\mathring{\mu}) = \mu\,\bar{s}^{1/k} = \bar{\mu} \tag{19}$$

$$\text{var}(\mathring{\mu}) = [E(\mathring{\mu})]^2/[d \cdot (k + 1)]. \tag{20}$$

The parameters we wish to estimate are k and μ. Equation (18) can be used to form the likelihood of k and μ if empirical estimates of $d_i(x, y)$, $\mathring{\mu}_i(x, y)$ and $\bar{s}_i(x, y)$ are available. These empirical estimates can be obtained from cohort life tables as de-

scribed below, thereby permitting k and μ to be estimated using the ML procedures described in the Appendix. However, the fact that there are two parameters, k and μ, associated with each observed mortality count d suggests that to obtain estimates constraints will have to be imposed on the variation of these parameters over observed data points. These constraints will be discussed below.

Models of the Standard Force of Mortality

In the previous section, it was shown that, if the initial population distribution of frailty were the gamma distribution, then the likelihood function could be expressed as a function of two parameters per observed data point. Additionally, the gamma shape parameter which does not vary over age is a characteristic of the population and, therefore, may be reasonably assumed not to vary over successive cohorts of the same population. Actually, this assumption is reasonable if mortality selection is the dominant factor in the age changes in the heterogeneity distribution for each cohort. This would be the case if the effects of modern medicine on life expectancy over the past century can be represented by reductions in the age specific standard forces of mortality over successive cohorts—but not by a reduction in the coefficient of variation and, hence, the shape parameter k (see equation 8c), of the reduced mortality risk levels. This suggests that, considering only one population group, if we have observations for n_a ages for n_c successive cohorts, we need to deal with $n + 1$ parameters from n observed data points, where $n = n_a \cdot n_c$. This condition makes it clear that we have to make some assumption about the age trajectory of individual mortality risks in order to obtain estimates of the parameter k. We will consider four such models of the age trajectory of individual mortality risks. These four models were selected because they are flexible and can represent an extremely broad range of age trajectories of individual mortality risks. This flexibility is required because we assume that the effects of modern medicine and various public health measures can be adequately represented in the model of the standard force of mortality. The range of estimates of heterogeneity produced under these various models reflects the magnitude of population heterogeneity if the "true" individual age trajectory of mortality risks were represented by one of the models selected.

The first model selected represents the standard force of mortality as a Gompertz function of age. This model was selected because the Gompertz function is often selected to represent the age trajectory of mortality risks in a population or cohort. Consequently, the estimate of population heterogeneity produced under this model would be the appropriate estimate of heterogeneity if the Gompertz function was indeed representative of the age change of mortality risks for individuals. The Gompertz function is:

$$\mu(x, y) = \alpha_{y_0} e^{\beta x}, \qquad (21a)$$

where

$$y_0 = y - x$$

represents the year the cohort was born. In this model each cohort is permitted to have a different initial mortality constant, α. The parameter β, the Gompertz rate parameter which governs the trajectory of the age increase in the standard force of mortality, is assumed constant over successive cohorts. For estimation purposes, we find it more convenient to deal with a reparameterization of (21a):

$$\mu(x, y) = \exp[\ln(\alpha) + \beta x + c_{y_0}], \qquad (21b)$$

where c_{y_0} represents a contrast between the cohort born at y_0 and some other specified cohort (in our example, this will be the 1885 birth cohort) to which the α in (21b) applies directly.

The second model was selected because it does not require the individual force of mortality to follow any parametric form, i.e., an age specific force of mortality is not constrained to be related to other age

specific forces of mortality within the cohort. What is assumed, however, is that each age specific force of mortality is proportional to the force of mortality at the same age in other cohorts. This model can be written:

$$\mu(x, y) = \exp[\ln\mu(x)] + c_{y_0}] \quad (22)$$

where c_{y_0} is a contrast between the force of mortality for standard individuals at age x in the cohort born in year y_0 and the force of mortality at age x in a cohort born in some other year, e.g., given the data described in the next section, 1885. Since the contrast parameter is in the exponent, it represents the proportionality factor between the force of mortality at the same age in different cohorts.

The third model extends the second model by allowing the proportionality factor c_{y_0} to increase or decrease, exponentially, across age. This model may be represented as:

$$\mu(x, y) = \exp[\ln[\mu(x)] + c_{y_0} + \gamma_x (1885 - y_0)] \quad (23a)$$

where γ_x is the change in the proportionality factor c_{y_0} for age x. The fourth model extends the second even further by allowing the proportionality factor to vary exponentially across cohorts as well as across ages. This model may be represented as:

$$\mu(x, y) = \exp[\ln[\mu(x)] + c_{y_0} + \gamma_x (1885 - y_0) + \beta_{y_0} x] \quad (23b)$$

where β_{y_0} is the change in the proportionality factor c_{y_0} for a given cohort.

DATA

Estimation of the parameters of (18) requires that we obtain both the cohort survival variables, $\bar{s}_i(x, y)$, and the cohort mortality counts, $d_i(x, y)$, as well as the observed cohort forces of mortality, $\mathring{\mu}_i(x, y)$ for the populations to be compared. Mortality data for U.S. white males and females and Swedish males and females for the years 1850 to 1975 were obtained from a number of published sources. Swedish data were used because of the high quality of the Swedish vital statistics system as well as the unusually long historical time series for this country. This latter aspect is particularly important to our strategy of estimating the range of k by analyzing a series of cohort mortality experiences.

The Swedish data for 1850 to 1880 were derived from Keyfitz and Flieger (1968), while the data for 1881 to 1975 were derived from the Statistical Abstract of Sweden (1976). The Keyfitz and Flieger data were in the form of abridged life table mortality probabilities (i.e., ${}_1q_0$, ${}_4q_1$, and ${}_5q_x$, $5 \leq x \leq 80$) for five-year time periods. In contrast, the Statistical Abstract of Sweden data were in the form of complete life table mortality probabilities (i.e., ${}_1q_x$, $0 \leq x \leq 89$) for either ten-year time periods (1881 to 1930) or five-year time periods (1931 to 1975). Neither of these data forms is appropriate for the construction of the complete cohort life tables required for our analysis. Consequently, interpolation techniques (described in Vaupel et al., 1979b) were employed to provide period and cohort life tables.

For the United States, the data problems were more severe. First, there is the difficulty that the death registration system was not completed in the United States until 1933. Second, there are questions about the reliability of mortality data, though the problems seem to be greater for the nonwhite population. Third, census enumeration errors affect the denominators of the death rates, though, as for the mortality data, these problems seem to be greater for the nonwhite population. These difficulties suggested that we employ alternatives to the "official" life tables provided by the vital statistics system. Indeed, a great deal of effort has been expended by demographers in attempting to provide such alternate life tables. For U.S. whites in 1850, we employed the life table data (i.e., ${}_1q_x$, $x = 0,1,2,3,4,5,10,\ldots 85$) from Jacobson (1957). For 1901 and the decennial years 1910 to 1950, U.S. white life table data

(i.e., l_x, $x = 0,1,5,10,\ldots 90$) which were adjusted for census enumeration errors were taken from Coale and Zelnik (1963). For 1960 and 1970, the Siegel (1974) estimates of census errors were employed to adjust the "official" complete life tables for those years. Here it should be noted that the Siegel estimates involved the Coale and Zelnik methodology, so that it is unlikely that bias would be introduced into the cohort life tables at the point of crossing over from the Coale and Zelnik data to the Siegel adjusted data. For 1975 we constructed complete life tables from the mortality counts provided by the National Center for Health Statistics and the 1975 resident population projections provided by the U.S. Bureau of the Census. The interpolation methods applied to the Swedish data (Vaupel et al., 1979b) were also applied to U.S. data.

With these data, it was possible to construct complete life tables up to age 89 for each cohort born in the 35-year period 1850 to 1885. The cohort survivorship variables $\bar{s}_i(x, y)$ were estimated from the l_x-column of the cohort life tables for the cohort born in year $(y\text{-}x)$ in population group i as the ratio l_x/l_0, where l_0 is the cohort life table radix and l_x is the cohort life table number of survivors at exact age x. Additionally, with the cohort survivorship variables $\bar{s}_i(x, y)$ available, we were able to normalize the cohort radix for each of 8 birth cohorts in such fashion that the number of survivors at age 65 matched the published population figures at age 65 for those cohorts. The 8 cohorts were the 1850, 1855, 1860, 1865, 1870, 1875, 1880, and 1885 birth cohorts. By normalizing the cohort radices in this fashion, it was possible to estimate the $d_i(x, y)$'s required for use in our likelihood function.

The observed cohort force of mortality is not contained in the cohort life tables. Consequently, to estimate these quantities, we used the negative logarithmic transformation of the age specific probability of survival, i.e.,

$$\mathring{\mu}_i(x, y) = -\ln[1 - q_i(x, y)], \quad (24)$$

where $q_i(x, y)$ is the life table probability of death at age x in year y in population i.

Finally, because we were forced to employ interpolated data and because the Gompertz function is appropriate only for adult mortality, we restricted our analysis to every fifth year of age in the age range 35 to 85 and age 89. This restriction will have the effect of minimizing bias due to interpolation. Thus, the analysis is concerned with $n_c = 8$ cohorts and $n_a = 12$ ages, for a total of $n = 96$ data points for each of four populations.

RESULTS

In this section we will present the parameter estimates for the distribution of individual frailty derived under four different sets of model assumptions concerning the changes over both age and cohort of the standard force of mortality. These estimates were based on independent analyses of four population groups—U.S. white males and females and Swedish males and females. For each population group, the parameter estimates were based on data from 8 successive cohorts born in 1850, 1855, 1860, 1865, 1870, 1875, 1880, and 1885, respectively. For each cohort, data from 12 ages were analyzed—ages 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, and 89. Hence, the parameter estimates for each population group were based on a total of 96 data points.

This section is in three parts. We first examine the level of fit of each of the four different models, selecting a model which best represents the data by organizing the models so that they can be tested as a series of hierarchical hypotheses with the Gompertz model representing the simplest or baseline model. We then present the parameters and evaluate the fit of the Gompertz model since the Gompertz model is a commonly used actuarial function. In the third part of this section we present and discuss the best fitting of the more general model forms.

Evaluation of Models of Population Heterogeneity

As we have indicated it is necessary to select a particular model of the age trajectory of individual mortality risks before estimates of heterogeneity can be produced. In a previous section we selected four such models with the standard Gompertz function being the basis for the most restrictive of them. To evaluate the fit of these models it is necessary to realize that a likelihood chi-squared value may be calculated for each model (see Appendix). To test between the fits of the four models we shall calculate a series of F-statistics by

$$F = \frac{\chi_{B,A}^2/(D.F._A - D.F._B)}{\chi_B^2/D.F._B} \qquad (25)$$

where $\chi_{B,A}^2$ is the likelihood chi-squared value obtained from the comparison of the likelihood function values in the simpler (fewer parameters) model A with degrees of freedom $D.F._A$ and the corresponding value for the more complex model B with degrees of freedom $D.F._B$; and χ_B^2 is the "goodness of fit" chi-squared value for model B with k fixed at the estimated value (see Appendix). The F-statistic then represents a test, with degrees of freedom $D.F._A - D.F._B$ and $D.F._B$, of the significance of the average improvement in fit produced by the additional parameters added in the more complex model. By making the tests in this way there is less danger of over-fitting the model (i.e., of adding parameters that explain only interpolation and other data bias) since the test of the efficacy of the additional parameters is made against the residual mean square error for the more complex model. Let us designate the four models as Model I (the Gompertz model; equation 21b), Model II (the age specific proportional hazard model; equation 22), Model III (the age specific, proportional hazard model with exponential changes, equation 23a) and Model IV (the age specific proportional hazard model with exponential age and cohort changes, equation 23b). The models may be tested in a hierarchical sequence from Model I to Model IV. The F-statistics for these hierarchial tests are presented, for each of the four population groups, in Table 1.

The results in Table 1 are clear. The additional parameters in Model III improve the fit to the data markedly—χ^2 change going from Model I to Model III was 217,056, 23,326, 270,011 and 27,439 for U.S. white males, Swedish males, U.S. white females and Swedish females respectively. The improvement in going from Model II to Model III was also highly significant. However, the improvement in going from Model III to Model IV was insignificant ($p > 0.25$) for three of the population groups, being significant ($p < 0.001$) only for U.S. white females. As a consequence we select Model III as our best fitting model.

Model I: The Gompertz Model

As we have indicated we will discuss the analysis produced under Model I, or the Gompertz specification, since it is a familiar model to most demographers and because it is a model that has commonly been used to model mortality.

This model was parameterized so that the parameters α and β applied directly to the Gompertz function for the 1885 birth cohort. Seven parameters representing contrasts of the other seven birth cohorts with the 1885 birth cohorts were also estimated. Finally, a single parameter k representing the initial population frailty distribution was estimated. Hence, for each population group, the Gompertz model required estimation of 10 parameters. These parameters, estimated via ML procedures, and their standard errors, are presented in Table 2.

Examination of Table 2 shows that, for all four population groups, all 10 parameters are estimated with a high degree of precision. The Gompertz constant $\ln(\alpha)$ is largest for U.S. white males (−8.23) and smallest for Swedish females (−9.11) with

Table 1.—F-Statistics for Tests of the Fit of a Hierarchical Sequence of Models

Comparisons	U.S. White Males	Swedish Males	U.S. White Females	Swedish Females
Model I versus Model II				
$F_{10,76}$	139.8	636.3	233.7	451.4
Model II versus Model III				
$F_{11,65}$	27.1	8.4	53.0	11.9
Model III versus Model IV				
$F_{6,59}$	1.1	0.7	4.7	1.3

Swedish males (−8.89) and U.S. white females (−8.71) ranking third and second, respectively. Thus, Swedes have a lower initial mortality constant than have Americans and, within country, females have a lower initial mortality constant than males. Interestingly, the Gompertz rate parameters exhibit the reverse pattern with U.S. white males (0.080) and females (0.082) ranking fourth and third, respectively, while Swedish males (0.088) and females (0.089) rank second and first respectively. Thus, there is a negative correlation between the $\ln(\alpha)$ and β parameters in the Gompertz function model for the standard force of mortality. The existence of a negative correlation was anticipated from the results of the application of the Gompertz function to cohort mortality rates (Strehler, 1977).

The Gompertz rate parameters have a simple interpretation as the proportional increases in the standard force of mortality over single years of age. To see this, consider the expression for the ratio of the standard forces of mortality at ages $(x + 1)$ and x. From equation (21a), this ratio is found to be simply e^{β} which for small values of β is approximately equal to $(1 + \beta)$. For example, for U.S. white males the estimated β value (0.080) is interpreted as an approximate 8 percent increase in the standard force of mortality per single year of age. The actual value computed using the ratio e^{β} is only slightly higher at 8.3 percent.

Also included in Table 2 are the contrast parameters (c_{y_0}'s) for the seven cohorts born in 1850, 1855, 1860, 1865, 1870, 1875, and 1880. For all four population groups, the c_{y_0}'s indicate a monotonic decline in the standard force of mortality over successive cohorts. The c_{y_0}'s may be interpreted in any of three ways. First, they may be added to the $\ln(\alpha)$ parameter to produce an estimate of $\ln(\alpha)$ as in equation (21b). For example, to compute $\ln(\alpha_{1850})$ we simply add $\ln(\alpha)$ and c_{1850}, yielding −7.92 (U.S. white males), −8.61 (Swedish males), −8.14 (U.S. white females), and −8.71 (Swedish females). Here we see that, although the Gompertz constants maintain the same relative ranking as for the 1885 cohort, the difference between Swedish males and U.S. white females is now much larger (rising

Table 2.—Parameter Estimates for Gompertz Model

	U.S. White Males	Swedish Males	U.S. White Females	Swedish Females
$\ln(\alpha_{1885})$	-8.23 (3.58×10^{-3})	-8.89 (1.61×10^{-2})	-8.71 (4.44×10^{-3})	-9.11 (1.66×10^{-2})
β	8.00×10^{-2} (6.80×10^{-5})	8.85×10^{-2} (2.84×10^{-4})	8.22×10^{-2} (7.96×10^{-5})	8.90×10^{-2} (2.84×10^{-4})
c_{1850}	3.12×10^{-1} (2.03×10^{-3})	2.81×10^{-1} (8.82×10^{-3})	5.74×10^{-1} (2.62×10^{-3})	4.00×10^{-1} (9.02×10^{-3})
c_{1855}	2.79×10^{-1} (1.96×10^{-3})	2.50×10^{-1} (8.66×10^{-3})	5.18×10^{-1} (2.51×10^{-3})	3.69×10^{-1} (8.85×10^{-3})
c_{1860}	2.41×10^{-1} (1.90×10^{-3})	2.35×10^{-1} (8.46×10^{-3})	4.48×10^{-1} (2.39×10^{-3})	3.60×10^{-1} (8.65×10^{-3})
c_{1865}	2.11×10^{-1} (1.83×10^{-3})	1.93×10^{-1} (8.48×10^{-3})	3.79×10^{-1} (2.30×10^{-3})	3.11×10^{-1} (8.71×10^{-3})
c_{1870}	1.72×10^{-1} (1.76×10^{-3})	1.51×10^{-1} (8.39×10^{-3})	2.97×10^{-1} (2.20×10^{-3})	2.75×10^{-1} (8.69×10^{-3})
c_{1875}	1.14×10^{-1} (1.71×10^{-3})	1.14×10^{-1} (8.13×10^{-3})	2.09×10^{-1} (2.11×10^{-3})	2.37×10^{-1} (8.46×10^{-3})
c_{1880}	5.44×10^{-2} (1.65×10^{-3})	8.34×10^{-2} (7.95×10^{-3})	1.24×10^{-1} (2.02×10^{-3})	1.61×10^{-1} (8.28×10^{-3})
k	3.93 (1.91×10^{-2})	3.20 (5.35×10^{-2})	2.84 (1.40×10^{-2})	2.79 (4.38×10^{-2})

Note: Standard errors are in parentheses.

from 0.18 to 0.47). Second the c_{y_0}'s may be exponentiated (as in equation (21a)) to produce a measure of the proportional increase in risk experienced by the earlier cohorts as compared to the 1885 birth cohort. For example, the 1850 cohorts had an increased risk of 37, 32, 78, and 49 percent respectively, in the U.S. white male, Swedish male, U.S. white female and Swedish female populations as compared to the 1885 birth cohorts in these four population groups. Third, differences in the c_{y_0}'s between successive cohorts may be divided by the time interval length to estimate the proportional decrease in the standard force of mortality over single years of time. For example, as a rough measure of the overall rate of decrease, we may take the difference between c_{1850} and c_{1885} (which, by definition is zero) and divide by 35 years to find net decreases of 0.9, 0.8, 1.6 and 1.1 percent per year, re-

spectively, in the U.S. white male, Swedish male, U.S. white female, and Swedish female population groups.

The final parameter in Table 2 is the heterogeneity parameter k. This parameter describes the shape of the heterogeneity distribution so that larger values of k are indicative of a reduced level of heterogeneity. That is, as shown in equation (8c), the coefficient of variation is obtained from the inverse square root of k so that large k's translate into small coefficients of variation. Here it is seen that under the assumption that the standard force of mortality can be modeled as a Gompertz function of age, U.S. white males have the largest k (3.93), Swedish females have the smallest k (2.79), while Swedish males (3.20) and U.S. white females (2.84) rank second and third respectively. Thus, the Gompertz model implies that Swedish females have the largest coefficient of variation (0.60), U.S. white females the second largest (0.59), Swedish males the third largest (0.56), while the smallest coefficient of variation is obtained for U.S. white males (0.50). A second way of interpreting these shape parameter estimates relies on the property that a gamma distribution with shape parameter k may be rescaled to produce another gamma distribution with the same shape parameter k. Consequently, k represents the shape parameter of the distribution of either the force of mortality at any given age or the individual frailty levels among survivors at any given age. In Figures 1 and 2 we have graphed the frailty distribution for the 1885 birth cohorts in the U.S. white male and Swedish female population groups using the values of k reported in Table 2.

The four plots in Figure 1 are the gamma distributions of frailty in the 1885 U.S. white male birth cohort at birth and for survivors to ages 45, 65, and 85. Each plot is normalized to a probability mass equal to the survivorship proportion $\bar{s}(x)$, i.e.,

$$f_x(x, y, z) = f(x_0, y_0, z) \cdot s(x, y, z), \quad (26)$$

where $f(x_0, y_0, z)$ is given in (7b); $s(x, y, z)$ is formed by solving (13) for H and substituting the result in (2b); and $f_x(x, y, z)$ is the function plotted. By integrating both sides of (26) with respect to z, we find that:

$$\int_0^\infty f_x(x, y, z)\, dz = E[s(x, y, z)] = \bar{s}(x, y), \quad (27)$$

so that the effects of mortality selection on the cohort can be seen as a "shrinking" of the distribution. Note the rapid removal of persons with high frailty levels ($z > 1.0$) at each age: by age 85, only those persons with very low frailty values survive.

The four plots in Figure 2 are the gamma distributions of frailty in the 1885 Swedish female birth cohort at birth and for survivors to ages 45, 65 and 85. As in Figure 1, these plots have been normalized to a probability mass of $\bar{s}(x)$. Here, we find that Swedish females, with a much lower k than U.S. white males (2.79 versus 3.93) have greater variance in the initial distribution and, because the shape parameter k remains constant under the operation of mortality selection, they also exhibit greater relative variance at each of the ages 45, 65, and 85.

In the above two figures we have examined the distribution of frailty at birth and among survivors to three ages—45, 65, and 85. A third way of interpreting the estimated shape parameter values is based on the fact that the force of mortality among those who die at any given age is gamma distributed with shape parameter $k + 1$ and mean $\bar{\mu}^*$ where from equation (16) we have:

$$\bar{\mu}^*/\bar{\mu} = (k + 1)/k. \quad (28)$$

Thus, the quantity $(k + 1)/k$ is a measure of relative risk, giving the proportional increase in average risk for those who die at any given age compared to the average risk for the survivors to that age. Using the values of k from Table 2, we find that Swedish females have the largest relative risk (1.36) and U.S. white males have the

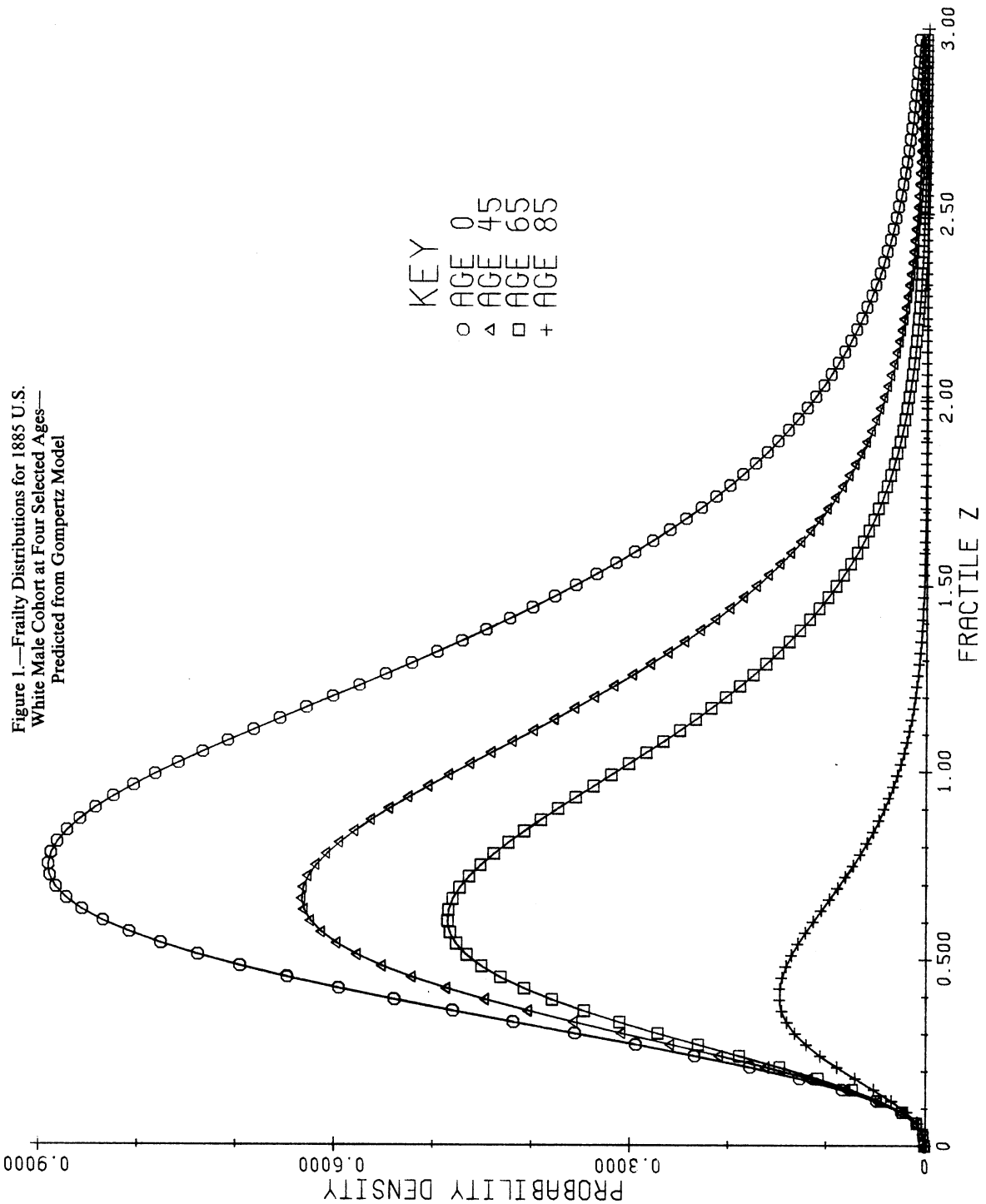

Figure 1.—Frailty Distributions for 1885 U.S. White Male Cohort at Four Selected Ages—Predicted from Gompertz Model

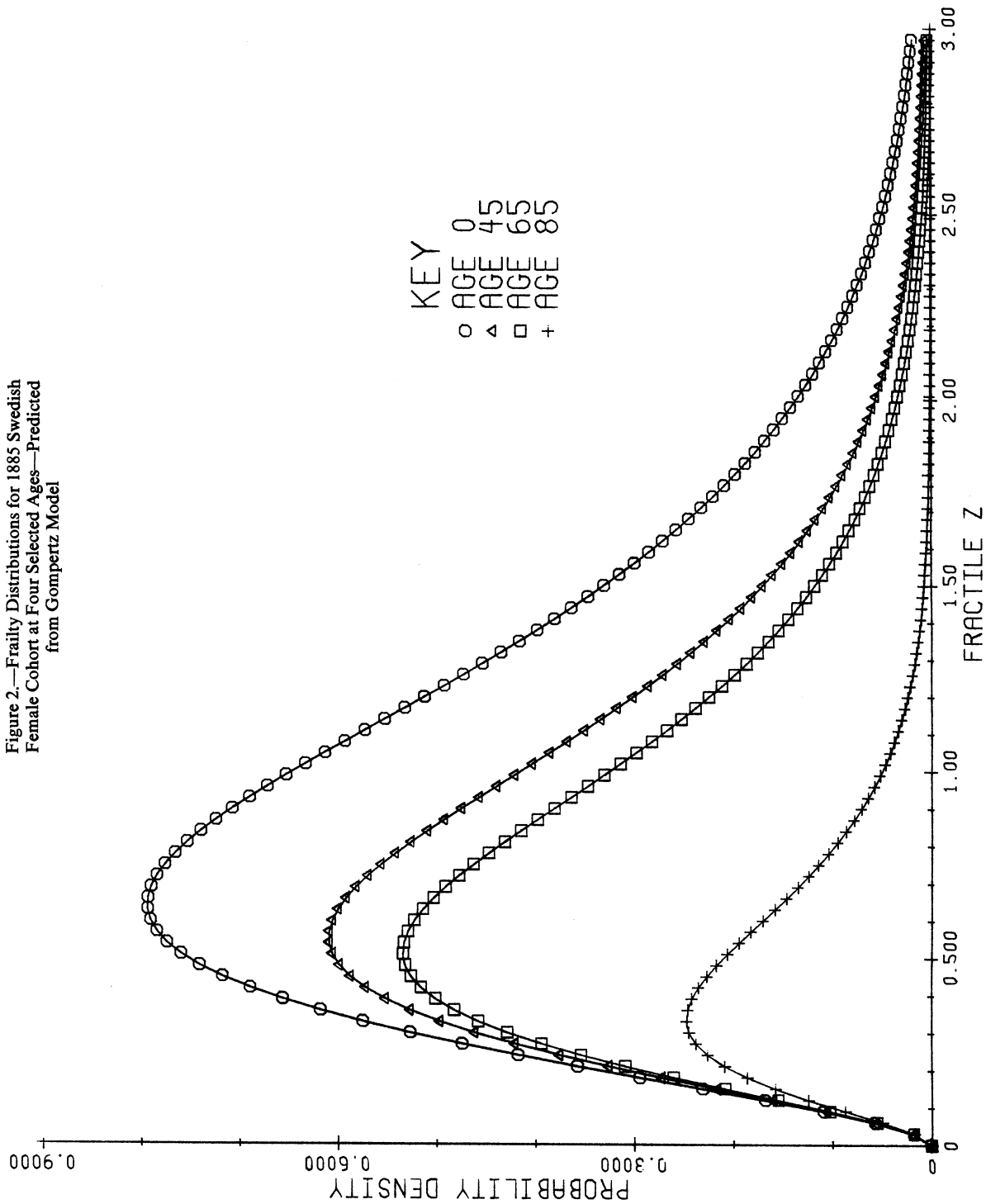

Figure 2.—Frailty Distributions for 1885 Swedish Female Cohort at Four Selected Ages—Predicted from Gompertz Model

smallest (1.25), while Swedish males (1.31) and U.S. white females (1.35) rank third and second, respectively. It is interesting to note that the female population groups exhibit virtually identical relative risks, especially given the fact that the k values were determined from independent analyses. Males, on the other hand, exhibit about a 6 percent difference between Sweden and the U.S. population groups.

Model III: Age Specific Standard Force of Mortality Model

In the third model we tested, the Gompertz function for the 1885 birth cohort was replaced by 12 μ parameters representing the 12 age specific standard forces of mortality in the 1885 cohort and 12 γ parameters representing the change of the age specific forces of mortality over cohorts. One of the seven contrast parameters was constrained to eliminate a collinearity with the γ parameters. Hence this third model was estimated with 31 parameters which included 6 c_{y_0}'s and k from the Gompertz model but replaced the $\ln(\alpha)$ and β parameters with 12 $\ln(\mu(x))$ parameters and 12 γ_x parameters. The estimated parameter values and their standard errors are presented in Table 3.

To illustrate the improvement of Model III over the other two models we present, in Figure 3, the cohort forces of mortality predicted from the three models, and the observed forces of mortality, for the 8 U.S. white male cohorts. In Figure 3 the observed values are indicated by the symbol ○, the values predicted from Model I by □, the values predicted from Model II by ◇, and the values predicted from Model III by +. The cohort forces of mortality, both predicted and observed, are plotted against "date" so that to determine age from the plot the birth date of any given cohort has to be subtracted from date, i.e., the values at 1885 are the first age points for the 1850 cohorts so that 1885−1850 =35.

In Figure 3 we see that Model I (the Gompertz) systematically deviates from the observed values at both younger ages (below age 75, over prediction) and at later ages (under prediction) for all 8 cohorts. For Model II the deviations are not as large and they change sign over cohorts. That is, the pattern of deviation from the observed for Model II is the same as the Gompertz up to the 1870 cohort. For the 1875 and later cohorts the pattern reverses with under prediction below age 75 and over prediction after age 75. The deviations of Model III from the data are very small and show no regular pattern. Thus the plots for the U.S. white male cohorts clearly indicate that Model III is the best model with no systematic pattern of deviations.

The most important changes in parameter values between Models I and III are the declines in the heterogeneity parameter k for U.S. white males (0.60 versus 3.93) and Swedish males (1.59 versus 3.20), U.S. white females (0.82 versus 2.84) and Swedish females (2.40 versus 2.79). These reductions in k imply substantially greater heterogeneity, especially for the two U.S. groups, than implied under Model I. For example, these values of k yield coefficients of variation of 1.29 and 1.10, respectiyely, for U.S. white males and females which are substantially larger than the prior estimates of 0.50 and 0.59. For Swedish males, the coefficients of variation increases from 0.56 to 0.79 while for Swedish females the increase is more modest, rising from 0.60 to 0.65.

A comparison of the 6 c_{y_0}'s common to the two models shows that in Model III the cohort changes are much larger than under the Gompertz for males and smaller in Model III than in the Gompertz for females. For example, c_{1850} increases for U.S. white males (0.34 versus 0.31), for Swedish males (0.36 versus 0.28), but decreases for U.S. white females (0.43 versus 0.57) and Swedish females (0.33 versus 0.40). As discussed we may divide these c_{1850}'s by 35 years to produce a rough measure of the overall rate of decrease in the standard force of mortality in the four population groups. These net decreases are 1.0, 1.0, 1.2, and

Table 3.—Parameter Estimates for Age Specific Standard Force of Mortality Model with Changes in Age Specific Standard Forces of Mortality over Cohorts

	U.S. White Males	Swedish Males	U.S. White Females	Swedish Females
$\ln \mu_{35}$	-4.37 (3.48×10^{-2})	-4.75 (5.76×10^{-2})	-4.73 (2.90×10^{-2})	-5.00 (3.21×10^{-2})
$\ln \mu_{40}$	-4.26 (3.79×10^{-2})	-4.87 (6.19×10^{-2})	-4.72 (3.14×10^{-2})	-5.04 (3.39×10^{-2})
$\ln \mu_{45}$	-4.01 (4.11×10^{-2})	-4.75 (6.50×10^{-2})	-4.56 (3.38×10^{-2})	-5.02 (3.56×10^{-2})
$\ln \mu_{50}$	-3.61 (4.52×10^{-2})	-4.49 (6.84×10^{-2})	-4.28 (3.66×10^{-2})	-4.77 (3.60×10^{-2})
$\ln \mu_{55}$	-3.17 (5.10×10^{-2})	-4.15 (7.35×10^{-2})	-3.98 (4.03×10^{-2})	-4.49 (3.73×10^{-2})
$\ln \mu_{60}$	-2.63 (5.93×10^{-2})	-3.74 (8.13×10^{-2})	-3.58 (4.51×10^{-2})	-4.13 (3.95×10^{-2})
$\ln \mu_{65}$	-1.98 (7.20×10^{-2})	-3.25 (9.36×10^{-2})	-3.10 (5.21×10^{-2})	-3.66 (4.37×10^{-2})
$\ln \mu_{70}$	-1.24 (9.00×10^{-2})	-2.68 (1.13×10^{-1})	-2.53 (6.26×10^{-2})	-3.13 (5.16×10^{-2})
$\ln \mu_{75}$	-0.38 (1.18×10^{-1})	-2.04 (1.46×10^{-1})	-1.89 (7.97×10^{-2})	-2.55 (6.58×10^{-2})
$\ln \mu_{80}$	0.75 (1.57×10^{-1})	-1.28 (2.00×10^{-1})	-1.07 (1.03×10^{-1})	-1.90 (8.94×10^{-2})
$\ln \mu_{85}$	2.19 (2.16×10^{-1})	-0.41 (2.88×10^{-1})	-0.10 (1.41×10^{-1})	-1.22 (1.28×10^{-1})
$\ln \mu_{89}$	3.51 (2.76×10^{-1})	0.41 (3.92×10^{-1})	0.80 (1.84×10^{-1})	-0.63 (1.74×10^{-1})
c_{1850}	0.34 (1.98×10^{-2})	0.36 (6.68×10^{-2})	0.43 (2.15×10^{-2})	0.33 (6.08×10^{-2})
c_{1855}	0.29 (1.68×10^{-2})	0.32 (5.67×10^{-2})	0.40 (1.80×10^{-2})	0.32 (5.09×10^{-2})
c_{1860}	0.26 (1.38×10^{-2})	0.27 (4.67×10^{-2})	0.36 (1.46×10^{-2})	0.31 (4.18×10^{-2})
c_{1865}	0.24 (1.06×10^{-2})	0.25 (3.68×10^{-2})	0.33 (1.11×10^{-2})	0.28 (3.25×10^{-2})

Table 3.—(Continued)

c_{1870}	0.21 (7.79×10^{-3})	0.20 (2.73×10^{-2})	0.28 (7.94×10^{-3})	0.26 (2.34×10^{-2})
c_{1875}	0.14 (5.13×10^{-3})	0.14 (1.83×10^{-2})	0.21 (5.19×10^{-3})	0.22 (1.54×10^{-2})
γ_{35}	9.57×10^{-3} (8.84×10^{-4})	-7.16×10^{-3} (2.48×10^{-3})	5.72×10^{-3} (9.62×10^{-4})	-3.32×10^{-3} (2.21×10^{-3})
γ_{40}	1.22×10^{-2} (9.12×10^{-4})	3.54×10^{-3} (2.49×10^{-3})	6.23×10^{-3} (9.97×10^{-4})	1.79×10^{-3} (2.22×10^{-3})
γ_{45}	1.16×10^{-2} (9.34×10^{-4})	4.78×10^{-3} (2.46×10^{-3})	5.55×10^{-3} (1.01×10^{-3})	3.15×10^{-3} (2.22×10^{-3})
γ_{50}	8.65×10^{-3} (9.62×10^{-4})	3.28×10^{-3} (2.43×10^{-3})	4.19×10^{-3} (1.03×10^{-3})	1.78×10^{-4} (2.19×10^{-3})
γ_{55}	6.14×10^{-3} (9.84×10^{-4})	1.66×10^{-3} (2.40×10^{-3})	4.67×10^{-3} (1.04×10^{-3})	-1.16×10^{-4} (2.15×10^{-3})
γ_{60}	4.53×10^{-3} (9.96×10^{-4})	4.11×10^{-4} (2.38×10^{-3})	5.34×10^{-3} (1.08×10^{-3})	-6.86×10^{-4} (2.11×10^{-3})
γ_{65}	2.77×10^{-3} (9.99×10^{-4})	-9.97×10^{-4} (2.36×10^{-3})	7.26×10^{-3} (1.13×10^{-3})	-8.09×10^{-4} (2.08×10^{-3})
γ_{70}	4.50×10^{-3} (1.01×10^{-3})	-2.11×10^{-3} (2.37×10^{-3})	8.33×10^{-3} (1.23×10^{-3})	-5.74×10^{-4} (2.07×10^{-3})
γ_{75}	9.66×10^{-3} (1.08×10^{-3})	-1.96×10^{-3} (2.41×10^{-3})	1.42×10^{-2} (1.38×10^{-3})	1.14×10^{-3} (2.10×10^{-3})
γ_{80}	1.54×10^{-2} (1.30×10^{-3})	-2.89×10^{-4} (2.51×10^{-3})	1.94×10^{-2} (1.74×10^{-3})	2.76×10^{-3} (2.20×10^{-3})
γ_{85}	2.58×10^{-2} (1.70×10^{-3})	2.80×10^{-3} (2.80×10^{-3})	3.01×10^{-2} (2.28×10^{-3})	6.41×10^{-3} (2.44×10^{-3})
γ_{90}	4.24×10^{-2} (2.42×10^{-3})	7.23×10^{-3} (3.31×10^{-3})	4.24×10^{-2} (3.12×10^{-3})	1.11×10^{-2} (2.83×10^{-3})
k	0.60 (3.21×10^{-2})	1.59 (3.37×10^{-1})	0.82 (5.65×10^{-2})	2.40 (4.01×10^{-1})

Note: Standard errors are in parentheses.

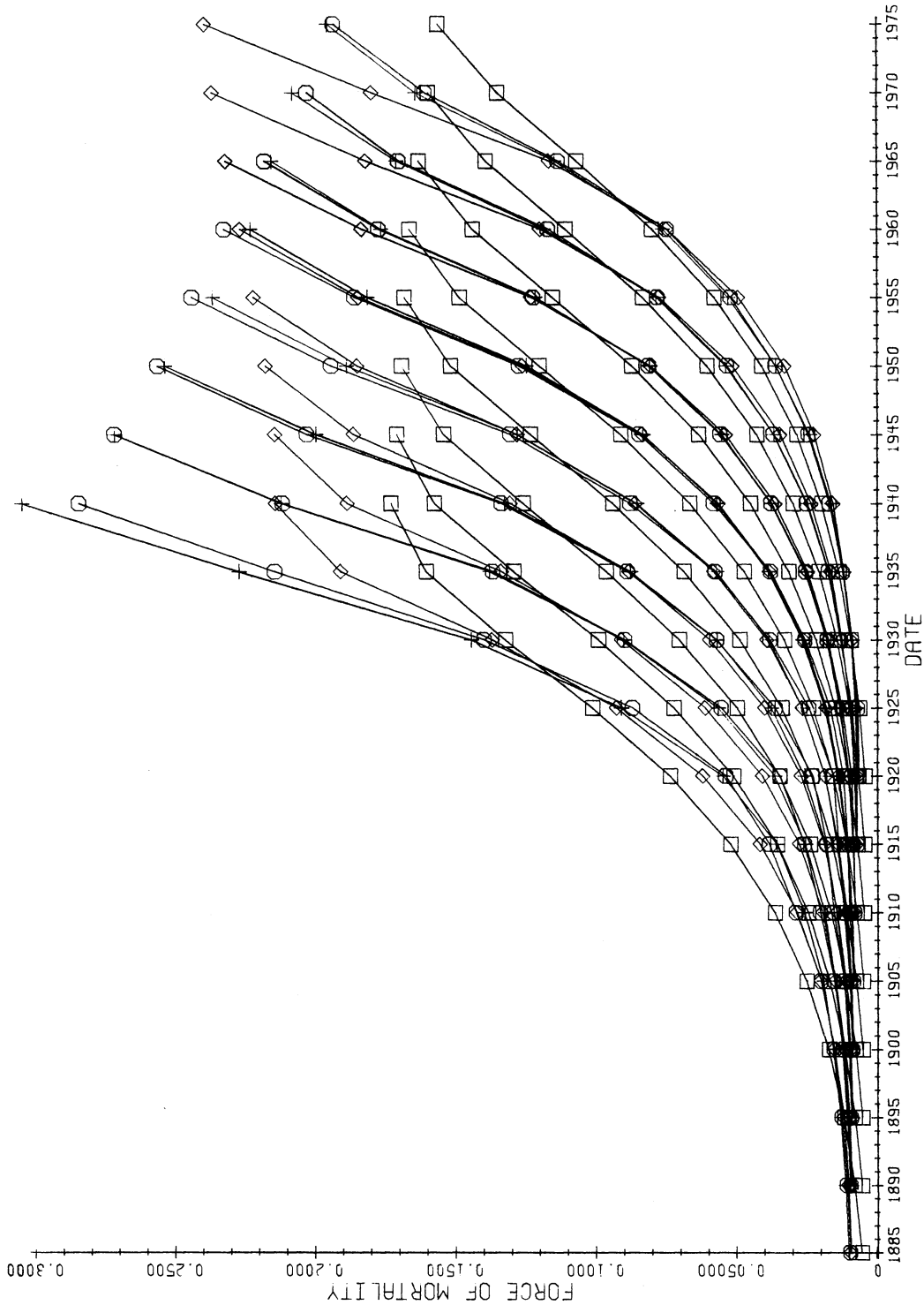

Figure 3.—Observed and Predicted Cohort Forces of Mortality for Eight U.S. White Male Cohorts

0.9 percent per year, respectively, in the U.S. white male, Swedish male, U.S. white female and Swedish female population groups (compare with 0.9, 0.8, 1.6 and 1.1 percent under the Gompertz model). Clearly, the age specific model leads to more similar rates of change.

The next parameters to be examined are the $\ln(\mu(x))$'s. As indicated, these parameters are direct estimates of the standard forces of mortality in the four population groups for the 1885 cohort at the specified ages. From Table 3 it can be seen that these parameters generally form a monotonic increasing sequence over age in all four population groups. Additionally, it is seen that the standard force of mortality at all ages is highest for U.S. white males, second highest for U.S. white females, third highest for Swedish males, and lowest for Swedish females. The age trajectory of individual mortality risks for all four population groups is far more rapid than in the observed data or under the Gompertz model.

The final set of parameters in the model, the γ_x's, represent the change in the age specific standard force of mortality over cohorts. As discussed, these parameters estimates are conditional on the contrast parameter c_{1880} which in Model III was constrained to the values obtained in Model II (i.e., c_{1880} was fixed at 0.07, 0.07, 0.12 and 0.13 respectively for U.S. white males and females and Swedish males and females). Thus, these parameter estimates cannot be unique and must be interpreted jointly with the contrast parameters. In general the estimates of γ_x are larger for both U.S. males and females than for Swedish males and females. The parameters show that for both U.S. groups there was an age specific decrease for all ages in the individual force of mortality. For Sweden a number of the γ_x are insignificant. To determine the total magnitude of change the γ_x have to be multiplied by $(1885 - y_0)$ and added to the appropriate c_{y_0}. For example, for the 1850 cohort each γ_x parameter would be multiplied by 35 and would represent the value by which, at age x, the 1850 cohort had higher logarithmically transformed standard forces of mortality than in the 1885 cohort.

DISCUSSION

The purpose of this paper is to illustrate methods for comparing the mortality experience of different human populations by adjusting the comparisons for the degree of individual heterogeneity to mortality risks within each population. This required using human mortality data for national populations and four different models for the age change of individual mortality risks to generate a range of estimates of the parameter k. This was done because the estimates of heterogeneity vary with the assumptions made about age change of the force of mortality for individuals. The estimates of k varied from 3.93, 2.84, 3.20 and 2.79 for U.S. males and females and Swedish males and females respectively, if the age trajectory of individual mortality risks followed the Gompertz model, to 0.60, 0.82, 1.59 and 2.40 under a model which did not assume a specific parametric form for the age increase in individual mortality risks.

The estimates of k between roughly 0.6 to 3.9 suggests that substantial population heterogeneity is present in all populations and that, consequently, the effects of mortality selection will have important implications for analyses of elderly populations. To illustrate, for the 1975 Swedish female period life table, the currently calculated life expectancy at birth is 78.15 years. In contrast, for $k_{SF} = 3$, the adjusted life expectancy would be 1.79 years less; for $k_{SF} = 4$, it would be 1.11 years less. Even for $k_{SF} = 8$, a value considerably larger than the values estimated in this paper, the adjusted life expectancy would be 0.34 years less than the currently calculated value of 78.15 years. To gain some perspective on the magnitude of these discrepancies, recall that Keyfitz (1977) reports that the total elimination of cancer death would affect life expectancy by about 2.3 years. Thus, with our esti-

mates of k in the range of 0.6 to 3.9 the substantive implications of the heterogeneity suggested by our estimates for life table computations are of the same order of magnitude as the elimination of major causes of death.

Given that the methods proposed in this paper can be applied to generate numerical estimates of the degree of heterogeneity in a given population, this raises the more fundamental issue of how we shall interpret such estimates. A simplistic interpretation would deal with frailty as a biological constant determined solely by genetic factors. However, even a brief review of the epidemiologic literature suggests that numerous risk factors, which are modifiable throughout at least part of the lifespan, contribute to significant risk heterogeneity to specific diseases and hence to total mortality. Since (a) our estimates of the degree of risk heterogeneity are based on a model of mortality selection and (b) the process of mortality selection requires only that the mortality risks are distributed over individuals, regardless of the source of this heterogeneity, it is clear that these estimates cannot be simply decomposed into genetic and nongenetic components of heterogeneity. Thus, the frailty variable z must be interpreted as a measure of the combined effects of both factors operating to systematically increase or decrease a given individual's mortality risk vis-à-vis the standard force of mortality. In any event, the challenge for future researchers is (a) to adequately account for such heterogeneity in their analyses of mortality, especially when analyzing mortality in elderly populations, and (b) to attempt to unconfound the various sources of heterogeneity so that those which are correlated with an individual's probability of survival, and which follow the age trajectory of the standard forces of mortality rather than the cohort force of mortality, are clearly identified as being subject to the effects of mortality selection.

APPENDIX

Presented here are the details of the ML procedure employed in estimating the parameters in equation (18). By employing the Stirling approximation to the gamma function (Γ) in (18), is possible to express the log likelihood for k and μ in the form:

$$L(k, \mu|d, \bar{s}, \mathring{\mu}) = \tfrac{1}{2} \ln (k + 1) + d(k + 1)\left\{\ln\left[\frac{\mathring{\mu}}{\mu \bar{s}^{1/k}}\right] + \left[1 - \frac{\mathring{\mu}}{\mu \bar{s}^{1/k}}\right]\right\}, \quad \text{(A.1)}$$

where a constant term $\ln[(d/(2\pi))^{1/2}/\mathring{\mu}]$ need not be considered, and hence, is omitted from (A.1). Our objective is to find the parameter values for k and μ which maximize the sum, over all observed data points, of functions of the form (A.1). To achieve this objective, we employed the Newton-Raphson algorithm which, starting with some arbitrary initial parameter values, iteratively improves the parameter estimates until the maximum of the sum of L's in (A.1) is found. Although estimation is conducted separately for each model, differences in the sums of L's may be transformed to approximate χ^2 variables for hierarchical model testing.

Estimation

In the Newton-Raphson procedure, a parameter vector $\mathbf{b}$ is replaced, at each iteration, with a "better" estimate of $\mathbf{b}$, i.e.,

$$\mathbf{b} \leftarrow \mathbf{b} + \mathbf{J}^{-1}\mathbf{g}, \quad \text{(A.2)}$$

where $\mathbf{g}$ is the vector of first-order partial derivatives of the sum of L's with respect to the $\mathbf{b}$ vector. At convergence, $\mathbf{J}$ is the covariance matrix of the first-order partial derivatives, while $\mathbf{J}^{-1}$ is the covariance matrix of the parameter estimates.

To simplify notation, recall that the models of the standard force of mortality

which we have considered can be written in the general form:

$$\mu = \exp(\mathbf{b}^{*T}\mathbf{w}), \tag{A.3}$$

where $\mathbf{w}$ is the vector of independent variables specific to a given model and $\mathbf{b}^*$ is the vector of linear coefficients. We also need the following definition:

$$\bar{s} = \exp(-\bar{H}). \tag{A.4}$$

From (A.3) and (A.4), it is apparent that the denominator terms in (A.1) can be expressed as:

$$\mu\,\bar{s}^{1/k} = \exp(\mathbf{b}^{*T}\mathbf{w} - \bar{H}/k), \tag{A.5}$$

which suggests that the vector $\mathbf{b}$ in (A.2) is of the form:

$$\mathbf{b}^T = (\mathbf{b}^{*T}, k). \tag{A.6}$$

By substituting (A.5) in (A.1), we find that the components of $\mathbf{g}$ associated with the log likelihood for each element of $\mathbf{b}$ are of the form:

$$\frac{\partial L}{\partial b_i^*}(\mathbf{b}|d, \bar{s}, \mathring{\mu}) = -d(k+1)\,w_i\left[1 - \frac{\mathring{\mu}}{\mu\,\bar{s}^{1/k}}\right] \tag{A.7}$$

$$\frac{\partial L}{\partial k}(\mathbf{b}|d, \bar{s}, \mathring{\mu}) = \frac{1}{2(k+1)} + d\left\{\ln\left(\frac{\mathring{\mu}}{\mu\,\bar{s}^{1/k}}\right) + \left[1 - \frac{\mathring{\mu}}{\mu\,\bar{s}^{1/k}}\right]\right\} - d(k+1)\frac{\bar{H}}{k^2}\left[1 - \frac{\mathring{\mu}}{\mu\,\bar{s}^{1/k}}\right]. \tag{A.8}$$

Similarly, the components of $\mathbf{J}$ associated with the log likelihood for pairs of elements of $\mathbf{b}$ are of the form:

$$E\left[\frac{\partial L(\ )}{\partial b_i^*}\cdot\frac{\partial L(\ )}{\partial b_j^*}\right] = d(k+1)\,w_i\,w_j \tag{A.9}$$

$$E\left[\frac{\partial L(\ )}{\partial b_i^*}\cdot\frac{\partial L(\ )}{\partial k}\right] = d(k+1)\,w_i\,\bar{H}/k^2 \tag{A.10}$$

$$E\left[\frac{\partial L(\ )}{\partial k}\cdot\frac{\partial L(\ )}{\partial k}\right] = \frac{1}{2(k+1)^2} + d(k+1)\frac{\bar{H}^2}{k^4}. \tag{A.11}$$

From (A.6), we see that there are at least two parameters associated with the log likelihood for each observation. Hence, constraints must be imposed on the parameters in order to achieve unique parameter estimates. This implies that the elements of $\mathbf{g}$ corresponding to a given parameter will be the sum, over those observations for which the parameter is constrained to be equal, of terms of the form (A.7) or (A.8). Similarly, the elements of $\mathbf{J}$ corresponding to a given pair of parameters will be the sum over those observations to which both parameters apply, of terms of the form (A.9), (A.10) or (A.11).

Model Testing

When dealing with alternate models describing the same phenomena, one usually needs to have some statistic by which both the relative performance of each model and the absolute level of fit may be assessed. For hierarchical models, we can employ the likelihood ratio χ^2 approximation:

$$\chi^2 = 2\cdot\sum_{j=1}^{n}\left[L_j(\mathbf{b}_2|d, \bar{s}, \mathring{\mu}) - L_j(\mathbf{b}_1|d, \bar{s}, \mathring{\mu})\right], \tag{A.12}$$

where n is the number of observations and where $\mathbf{b}_1$ and $\mathbf{b}_2$ are the parameter vectors for the two models being compared. For this test to be appropriate, $\mathbf{b}_1$ must represent a "subset" of $\mathbf{b}_2$, i.e., imposition of constraints on $\mathbf{b}_2$ must lead to $\mathbf{b}_1$. For example, the age specific standard force of mortality model may be constrained to reproduce the Gompertz model. When the test can be performed, the degrees of freedom are given by the number of constraints imposed on $\mathbf{b}_2$ to reproduce $\mathbf{b}_1$.

ACKNOWLEDGMENT

The research reported in this paper was conducted under NIA Grant AG01159-03 and 04.

REFERENCES

Coale, A. J. and M. Zelnik. 1963. New Estimates of Fertility and Population in the United States. Princeton: Princeton University Press.

Hastings, N. A. and J. B. Peacock. 1975. Statistical Distributions. London: Butterworths.

Jacobson, P. H. 1957. An Estimate of the Expectation of Life in the United States in 1850. Milbank Memorial Fund Quarterly 35:197–201.

Kendall, M. G. and A. Stuart. 1969. The Advanced Theory of Statistics. Vol. I, Distribution Theory. New York: Hafner.

Keyfitz, N. 1977. What Difference Would It Make If Cancer Were Eradicated? An Examination of the Taeuber Paradox. Demography 14:411–418.

——— and W. Flieger. 1968. World Population. Chicago: University of Chicago Press.

Siegel, J. S. 1974. Estimates of Coverage of the Population by Sex, Race, and Age in the 1970 Census. Demography 11:1–23.

Statistical Abstract of Sweden. 1976. Stockholm: National Center of Statistics.

Strehler, B. L. 1977. Time, Cells, and Aging. New York: Academic.

Vaupel, J. W., K. G. Manton and E. Stallard. 1979a. The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. Demography 16:439–454.

Vaupel, J. W., K. G. Manton and E. Stallard. 1979b. What Demographic Difference Does It Make That The Frail Die First? A Model Of Mortality And Some Suggestive Results Based On Swedish Mortality Since 1778. Center For The Study Of Health Policy, Institute of Policy Sciences and Public Affairs, Working Paper No. 1079, Duke University, Durham, N. C.