Mortality Measurement at Advanced Ages: A Study of the Social Security Administration Death Master File

Leonid A. Gavrilov* and Natalia S. Gavrilova†

Presented at the Living to 100 and Beyond Symposium

Orlando, Fla.

January 7-9, 2008

Copyright 2008 by the Society of Actuaries.
All rights reserved by the Society of Actuaries. Permission is granted to make brief excerpts for a published review. Permission is also granted to make limited numbers of copies of items in this monograph for personal, internal, classroom or other instructional use, on condition that the foregoing copyright notice is used so as to give reasonable notice of the Society's copyright. This consent for free limited copying without prior consent of the Society does not extend to making copies for general distribution, for advertising or promotional purposes, for inclusion in new collective works or for resale.

* Center on Aging, NORC and The University of Chicago

† Center on Aging, NORC and The University of Chicago

Abstract

Accurate estimates of mortality at advanced ages are essential to improving forecasts of mortality and the population size of the oldest old age group. However, estimation of hazard rates at extremely old ages poses serious challenges to researchers: (1) The observed mortality deceleration may be at least partially an artifact of mixing different birth cohorts with different mortality (heterogeneity effect); (2) Standard assumptions of hazard rate estimates may be invalid when risk of death is extremely high at old ages; (3) Ages of very old people may be exaggerated. One way of obtaining estimates of mortality at extreme ages is to pool together international records of persons surviving to extreme ages with subsequent efforts of strict age validation. This approach helps researchers to resolve the third of the above-mentioned problems but does not resolve the first two problems because of inevitable data heterogeneity when data for people belonging to different birth cohorts and countries are pooled together.

In this paper we propose an alternative approach, which gives an opportunity to resolve the first two problems by compiling data for more homogeneous single-year birth cohorts with hazard rates measured at narrow (monthly) age intervals. Possible ways of resolving the third problem of hazard rate estimation are elaborated. This approach is based on data from the Social Security Administration Death Master File (DMF). Some birth cohorts covered by DMF could be studied by the method of extinct generations. Availability of month of birth and month of death information provides a unique opportunity to obtain hazard rate estimates for every month of age. Study of several single-year extinct birth cohorts shows that mortality trajectory at advanced ages follows the Gompertz law up to the ages 102-105 years without a noticeable deceleration. Earlier reports of mortality deceleration (deviation of mortality from the Gompertz law) at ages below 100 appear to be artifacts of mixing together several birth cohorts with different mortality levels and using cross-sectional instead of cohort data. Age exaggeration and crude assumptions applied to mortality estimates at advanced ages may also contribute to mortality underestimation at very advanced ages.

1. Introduction

Accurate estimates of mortality at advanced ages are essential to improving forecasts of mortality and the population size of the oldest old age group. It is now considered as an established fact that mortality at advanced ages has a tendency to deviate from the Gompertz law, so that the logistic model often is used to fit human mortality (Horiuchi and Wilmoth, 1998). Estimation of mortality at extreme ages is difficult because data for extremely long-lived individuals are scarce and subjected to age exaggeration. Traditional demographic estimates of mortality based on period data encounter the well known denominator problem. More accurate estimates are obtained using the method of extinct generations (Vincent, 1951). In order to obtain good quality estimates of mortality at advanced ages, researchers are forced to pool together data for several calendar periods. Single-year life tables for many countries have very small numbers of survivors to age 100 that makes estimates of mortality at advanced ages unreliable. On the other hand, aggregation of deaths for several calendar periods creates a heterogeneous mixture of cases from different birth cohorts. Mortality deceleration observed in these data might be an artifact of data heterogeneity. In addition to that, many assumptions about distribution of deaths in the age/time interval used in mortality estimation are not valid for extreme old ages when mortality is particularly high and grows very rapidly.

Thus the estimation of hazard rates at extremely old ages poses several serious challenges to researchers:

(1) The observed mortality deceleration may be at least partially an artifact of mixing different birth cohorts with different mortality (heterogeneity effect);
(2) Standard assumptions of hazard rate estimates may be invalid when risk of death is extremely high at old ages;
(3) Ages of very old people may be exaggerated.

One way of obtaining estimates of mortality at extreme ages is to pool together international records of persons surviving to extreme ages with subsequent efforts of strict age validation (Robine and Vaupel, 2001; Robine, Cournil et al. 2005). This approach helps to resolve the third problem mentioned above but does not allow researchers to resolve the first two problems because of inevitable data heterogeneity when data for people belonging to different birth cohorts and countries are pooled together. In this project, we propose an alternative approach, which allows us to resolve partially the first two problems. This approach is based on data from the Social Security Administration Death Master File (DMF), which allows researchers to compile data for large single-year birth cohorts. Possible ways of resolving the third problem of hazard rate estimation are also elaborated.

2. Mortality at Advanced Ages: A Historical Review

The history of mortality studies at extreme ages is very rich in ideas and findings. Early studies starting with Gompertz (1825) himself suggested that the Gompertz law of mortality is not applicable to extreme old ages, and that mortality deceleration and leveling-off takes place at advanced ages (for an excellent historical review of studies on mortality deceleration at extreme old ages, see Olshansky, 1998). In 1939 two British researchers, Greenwood and Irwin, published an article "Biostatistics of Senility," with an intriguing finding that mortality force stops increasing with age at extreme old ages and becomes constant (Greenwood and Irwin, 1939). Their study and findings were considered to be so important that they were featured at the front page of academic journal "Human Biology," where their study was published. This study, accomplished by the famous British statistician and epidemiologist, Major Greenwood, may be interesting to discuss here because it correctly describes the mortality pattern at advanced ages for humans.

The first important finding was formulated by Greenwood and Irwin in the following way: "***…the increase of mortality rate with age advances at a slackening rate, that nearly all, perhaps all, methods of graduation of the type of Gompertz's formula over-state senile mortality***" (Greenwood and Irwin, 1939, p. 14). This observation was

confirmed later by many authors (see review in Gavrilov and Gavrilova, 1991), and it is known as the "late-life mortality deceleration."

The authors also suggested "***the possibility that with advancing age the rate of mortality asymptotes to a finite value***" (Greenwood and Irwin, 1939, p. 14). Their conclusion that mortality at exceptionally high ages follows a first order kinetics (also known as the law of radioactive decay with exponential decline in survival probabilities) was confirmed later by other researchers, including A.C. Economos (1979; 1980), who demonstrated the correctness of this law for humans and laboratory animals (linear decrease for the logarithm of the number of survivors). This observation is known now as the "mortality leveling-off" at advanced ages, and as the "late-life mortality plateau" (Curtsinger et al., 2006).

Moreover, Greenwood and Irwin made the first estimates for the asymptotic value of human mortality (one-year probability of death, q_x) at extreme ages using data from the life insurance company. According to their estimates, "***… the limiting values of q_x are 0.439 for women and 0.544 for men***" (Greenwood and Irwin, 1939, p. 21). It is interesting that these first estimates are very close to estimates obtained later using more numerous and accurate human data including recent data on supercentenarians (Robine and Vaupel, 2001). The authors also proposed an explanation of this phenomenon. According to Greenwood and Irwin (1939), centenarians live in a more protected environment than younger age groups and hence have lower risk of death than is predicted by the Gompertz formula.

The actuaries including Gompertz himself first noted this phenomenon of mortality deceleration. They also proposed a logistic formula for fitting mortality growth with age in order to account for mortality fall-off at advanced ages (Perks, 1932; Beard, 1959, 1971). Robert Eric Beard (1959) introduced a model of population heterogeneity with gamma distributed individual risk in order to explain mortality deceleration at older ages. This explanation is the most common explanation of mortality deceleration now.

The same phenomenon of "almost non-aging" survival dynamics at extreme old ages is detected in many other biological species. In some species mortality plateau can occupy a sizable part of their life (Carey et al., 1992; Gavrilov and Gavrilova, 2006).

Biologists were well aware of mortality leveling-off since the 1960s. For example, Lindop (1961) applied Perks formula in order to account for mortality deceleration at older ages in mice. George Sacher believed that the observed mortality deceleration in mice and rats can be explained by population heterogeneity: "one effect of such residual heterogeneity is to bring about a decreased slope of the Gompertzian at advanced ages. This occurs because sub-populations with the higher injury levels die out more rapidly, resulting in progressive selection for vigour in the surviving populations" (Sacher, 1966, p. 435). Strehler and Mildvan (1960) considered mortality deceleration at advanced ages as a prerequisite for all mathematical models of aging to explain.

Later Economos published a series of articles claiming a priority in the discovery of a "non-Gompertzian paradigm of mortality" (Economos, 1979; 1980; 1983; 1985). He found that mortality leveling-off is observed in rodents (guinea pigs, rats, mice) and invertebrates (nematodes, shrimps, bdelloid rotifers, fruit flies, degenerate medusae *Campanularia Flexuosa*). In the 1990s, the phenomenon of mortality deceleration and leveling-off became widely known after publications, which demonstrated mortality leveling-off in large samples of *Drosophila melanogaster* (Curtsinger et al., 1992; 2006) and medflies *Ceratitis capitata* (Carey et al., 1992), including isogenic strains of Drosophila (Curtsinger et al., 1992; Fukui et al., 1993; 1996). Mortality plateaus at advanced ages are observed for some other insects: house fly *Musca vicina*, blowfly *Calliphora erythrocephala* (Gavrilov, 1980), house fly *Musca domestica* (Gavrilov and Gavrilova, 2006), fruit flies *Anastrepha ludens*, *Anastrepha obliqua*, *Anastrepha serpentine*, parasitoid wasp *Diachasmimorpha longiacaudtis* (Vaupel et al., 1998), and bruchid beetle *Callosobruchus maculates* (Tatar et al., 1993). Interestingly, the failure kinetics of manufactured products (steel samples, industrial relays, and motor heat insulators) also demonstrates the same "non-aging" pattern at the end of their "lifespan" (Economos, 1979).

The existence of mortality plateaus is well established for a number of lower organisms, mostly insects. In the case of mammals, data are much more controversial. Although Lindop (1961) and Sacher (1966) reported short-term periods of mortality deceleration in mice at advanced ages and even used Perks formula in their analyses, Austad (2001) recently argued that rodents do not demonstrate mortality deceleration even in the case of large samples. Study of baboons found no mortality deceleration at advanced ages (Bronikowski et al., 2002). In the case of humans this problem is not yet resolved completely, because of scarceness of data and/or their low reliability. Thus, more studies on larger human birth cohorts are required to establish with certainty the true mortality trajectory at advanced ages.

The phenomenon of late-life mortality leveling-off presents a theoretical challenge to many models and theories of aging. One interesting corollary from these intriguing observations is that there seems to be no fixed upper limit for individual life span (Gavrilov, 1984; Gavrilov and Gavrilova, 1991; Wilmoth, 1997).

Population heterogeneity is probably the most common explanation of mortality deceleration proposed by British actuary Eric Beard in 1959. Another explanation of this phenomenon comes from the reliability theory of aging, which explains mortality leveling-off by an exhaustion of organism's redundancy (reserves) at extremely old ages so that every random hit results in death (Gavrilov and Gavrilova, 1991; 2001; 2006). There is also an opinion that lower risks of death for older people are due to their less risky behavior (Greenwood and Irwin, 1939). Finally, some researchers suggest evolutionary theory explanations for the phenomenon of mortality leveling-off (Mueller and Rose, 1996; Charlesworth, 2001).

3. Hazard Rate Estimation at Advanced Ages

A conventional way to obtain estimates of mortality at advanced ages is a construction of demographic life table with probability of death (q_x) as one of the important life table functions. Although probability of death is a useful indicator for mortality studies, it may be not the most convenient one for studies of mortality at

advanced ages. First, the values of q_x depend on the length of the age interval Δx for which it is calculated. This hampers both analyses and interpretation. For example, if one-day probability of death follows the Gompertz law of mortality, probability of death calculated for other age intervals does not follow this law (see Gavrilov and Gavrilova, 1991; Le Bras, 1976). Thus it turns out that the shape of age-dependence for q_x depends on the arbitrary choice of age interval. Also, by definition q_x is bounded by unity, which makes it difficult to study mortality at advanced ages.

A more useful indicator for mortality studies at advanced age is instantaneous mortality rate, mortality force or hazard rate, μ_x which is defined as follows:

$$\mu_x = -\frac{dN_x}{N_x dx}$$

where N_x is a number of living individuals at age x.

Hazard rate does not depend on the length of age interval (it is measured at the instant of time x), has no upper boundary and has a dimension of rate (time^{-1}). It should also be noted that the famous law of mortality, the Gompertz law, was proposed for fitting the hazard rate rather than probability of death (Gompertz, 1825).

The empirical estimates of hazard rates are often based on suggestion that age-specific mortality rate or death rate (number of deaths divided by exposure) is a good estimate of theoretical hazard rate. One of the first empirical estimates of hazard rate was proposed by George Sacher (Sacher, 1956; 1966):

$$\mu_x = \frac{1}{\Delta x}(ln\ l_{x-\frac{\Delta x}{2}} - ln\ l_{x+\frac{\Delta x}{2}}) = \frac{1}{2\Delta x}\ ln\frac{l_{x-\Delta x}}{l_{x+\Delta x}}$$

This estimate is unbiased for slow changes in hazard rate if $\Delta x \Delta \mu_x << 1$ (Sacher, 1966).

A simplified version of Sacher estimate (for small age intervals equal to unity) often is used in biological studies of mortality: $\mu_x = -\ln(1-q_x)$. This estimate is based on the assumption that hazard rate is constant over the age interval.

At advanced ages when death rates are high and grow very rapidly, the assumptions about small changes in hazard rate or a constant hazard rate within the age interval become questionable. Violation of these assumptions may lead to biased estimates of hazard rates calculated on the annual basis. Narrowing the age interval for hazard rate estimation from one-year to one-month may help to alleviate this problem.

4. Social Security Administration Death Master File as a Source of Mortality Data for Advanced Ages.

The Social Security Administration Death Master File (DMF) is a publicly available data source that allows a search for individuals using various search criteria: birth date, death date, first and last names, Social Security number, place of last residence, etc. (see Table 1). This resource covers individual deaths that occurred in the period 1937-2007, although coverage is not complete for deaths that occurred before 1972 (see Faig, 2002 for more details). Many researchers suggest that the quality of SSA/Medicare data is superior to vital statistics records because of strict evidentiary requirements in application for Medicare while age reporting in death certificates is made by proxy informant (Kestenbaum, 1992; Kestenbaum and Ferguson, 2002; Rosenwaike et al., 1998; Rosenwaike and Stone, 2003).

Social Security Administration Death Master File (DMF) was used in the study of mortality kinetics after ages 85-90 years. The advantage of this data source is that some birth cohorts covered by DMF could be studied by the method of extinct generations (Vincent, 1951; Kannisto, 1988; 1994). Availability of month of birth and month of death information provides a unique opportunity to obtain hazard rate estimates for every month of age, which is important given extremely high mortality after age 100 years.

The information from the DMF was collected for individuals who lived 85 years and over and died before 2008. The DMF database is unique because it represents mortality experience for very large birth cohorts of the oldest-old persons. In this study mortality measurements were made for cohorts, which are more homogeneous in respect to the period of birth and historical life course experiences.

TABLE 1

Information Available in

the SSA Death Master File

1. first, last names, SSN
2. date, month, year of birth
3. month, year of death
4. state of the SSN issuance
5. town, county, state, zip code of the last residence
6. death date verification code

The DMF collects deaths for persons who receive SSA benefits and currently covers over 90 percent of deaths occurring in the United States (Faig, 2002) and 93 to 96 percent of deaths of individuals aged 65 or older (Hill, Rosenwaike, 2001). Despite certain limitations, this data source allows researchers to obtain detailed estimates of

mortality at advanced ages. We already used this data resource for centenarians' age validation in the study of centenarian genealogies (Gavrilova, Gavrilov, 2007). This data resource is also useful for mortality estimates for several extinct or almost extinct birth cohorts in the United States.

5. Hazard Rate Estimates at Advanced Ages Using Data from the Social Security Death Master File

In this study we collected information from the DMF publicly available at Rootsweb.com. The total number of collected records is over 9 million with more than 900,000 records belonging to persons who lived 100 years and longer. Several birth cohorts (those born before 1890) may be considered extinct or almost extinct, so it is possible to apply the method of extinct generations (Vincent, 1951) and estimate mortality kinetics at very advanced ages up to 115-120 years.

The last deaths in the DMF available at the Rootsweb Web site occurred in 2007 (when the present study was conducted). We obtained data for persons who died before 2008 and were born in 1875-1894. There were no persons born before 1882 and only a few persons born in 1883-1887 who died in 2006 or 2007. Thus, the 1875-1887 birth cohorts in this sample may be considered extinct or almost extinct. Later birth cohorts are close to extinction but not completely extinct. Assuming that the number of living persons belonging to 1875-87 birth cohorts in 2007 is close to zero, it is possible to construct a cohort life table using the method of extinct generations, which was suggested and explained by Vincent (1951) and developed further by Kannisto (1994). In the first stage of our analyses we calculated an individual life span in completed months:

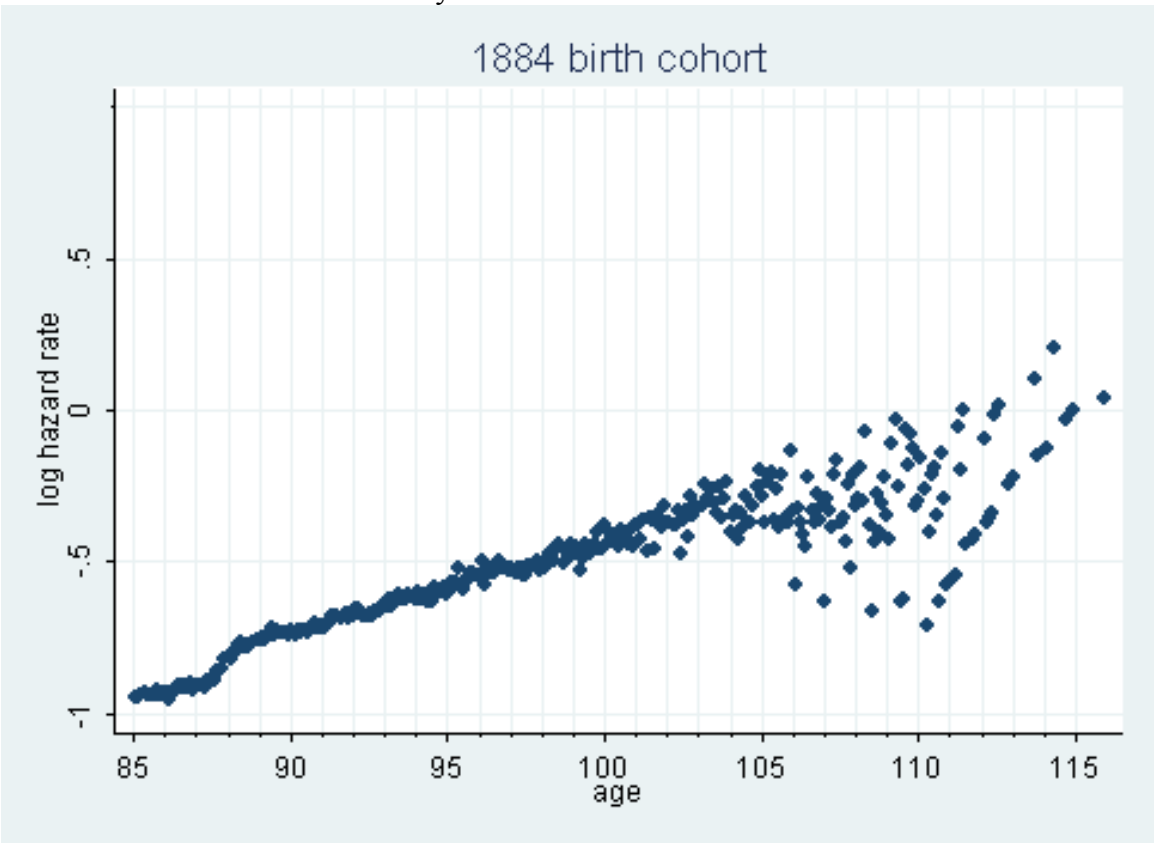
Lifespan in months = (death year – birth year) x 12 + death month – birth month

Then it is possible to estimate the hazard rate at each month of age using standard methods of survival analysis. For cohorts born after 1889, hazard rates may be overestimated at very advanced ages (over 110 years), although mortality trajectory at earlier ages should not be significantly affected by life span truncation. All calculations

were done using Stata statistical package, procedures “stset” and “sts” (Stata Corp, 2005). This software provides nonparametric estimates of major survival functions including the Nelson-Aalen estimates of hazard rate (force of mortality). Note that hazard rate, in contrast to probability of death, q(x), has a dimension of time frequency, because of the time interval in the denominator (reciprocal time, $time^{-1}$). Thus the values of hazard rates depend on the chosen units of time measurement (day^{-1}, $month^{-1}$ or $year^{-1}$). In this study survival times were measured in months, so the estimates of hazard rates initially had a dimension of $month^{-1}$. For the purpose of comparability with other published studies, which typically use the $year^{-1}$ time scale, we have transformed monthly hazard rates to the more conventional units of $year^{-1}$, by multiplying these estimates by a factor of 12 (one month in the denominator of hazard rate formula is equal to 1/12 year). Also note that hazard rate, in contrast to probability of death can be greater than 1, and therefore its logarithm can be greater than 0 (and we indeed observed this at extreme old ages in some rare cases as will be described later). We estimated hazard rates for several single-year birth cohorts—those born in 1879, 1884, 1887, 1889 and 1894.

The results of hazard rate estimates for the 1884 birth cohort are presented in Figure 1. Note the drop of mortality at ages 85-88, which corresponds to poor death coverage in the Social Security Death Master File for deaths occurring before 1972. Note that from age 90 up to ages 102-105 years, mortality grows steadily without obvious deceleration. Only after age 105 years does mortality tend to decelerate, although high statistical noise makes mortality estimates beyond age 105 years less reliable.

FIGURE 1
Hazard Rate (per year) for 1884 Birth Cohort.
Data from the Social Security Administration Death Master File



A recent study of age validation among supercentenarians (Rosenwaike and Stone, 2003) showed that age reporting among supercentenarians in the SSA database is rather accurate with the exception of persons born in the Southern states. In order to improve the quality of our dataset regarding the correct age reporting, we excluded records for those persons who applied for a Social Security number in the Southeast (AR, AL, GA, MS, LA, TN, FL, KY, SC, NC, VA and WV), Southwest (AZ, NM, TX, OK), Puerto Rico and Hawaii. This step of data cleaning, however, did not change significantly the overall trajectory of mortality at advanced ages, but decreased the number of too low mortality estimates after age 105 years (see Figures 2-3).

FIGURE 2
Hazard Rate (per year) for 1889 Birth Cohort.
Data from the Social Security Administration Death Master File

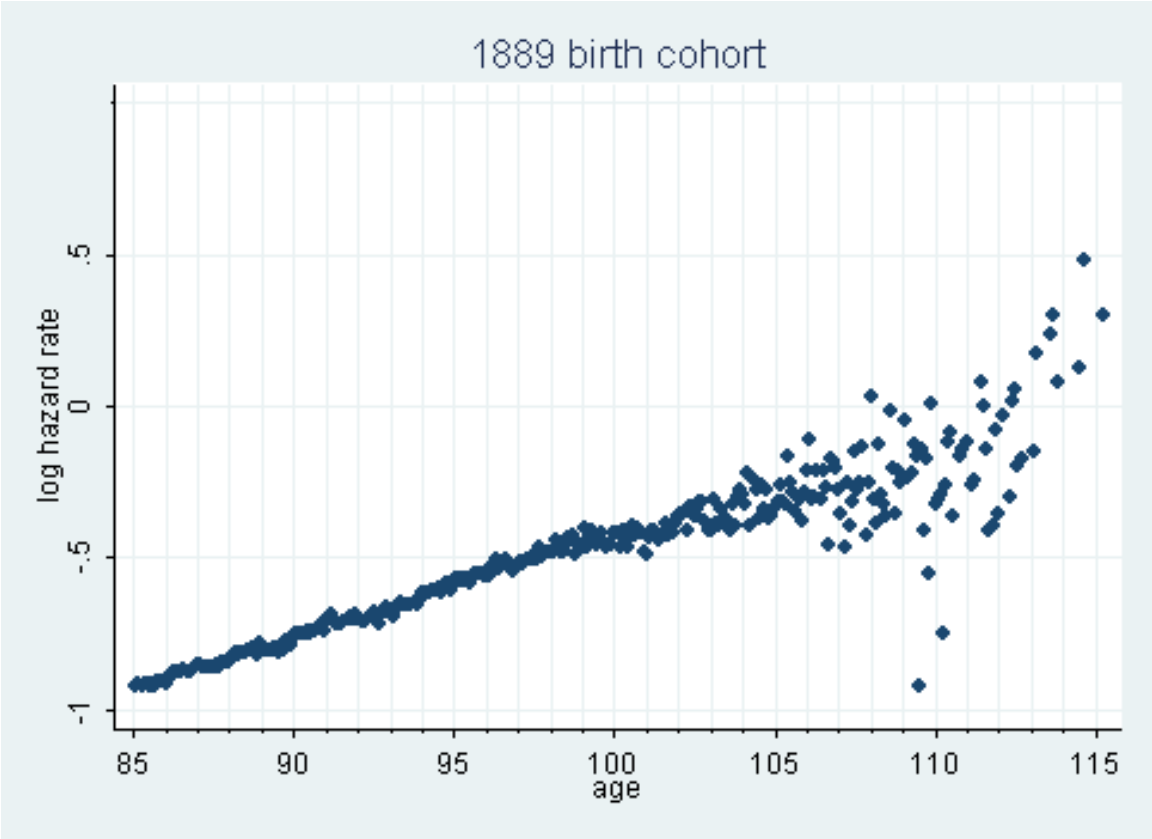
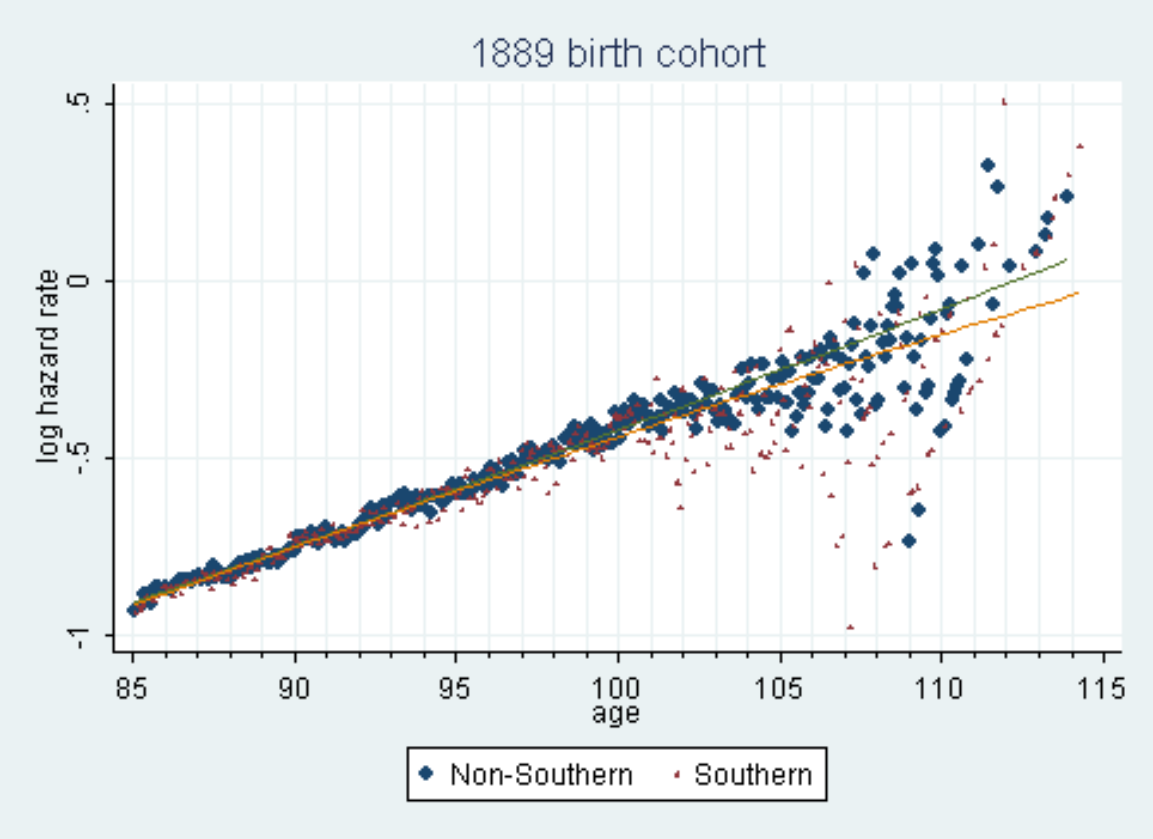


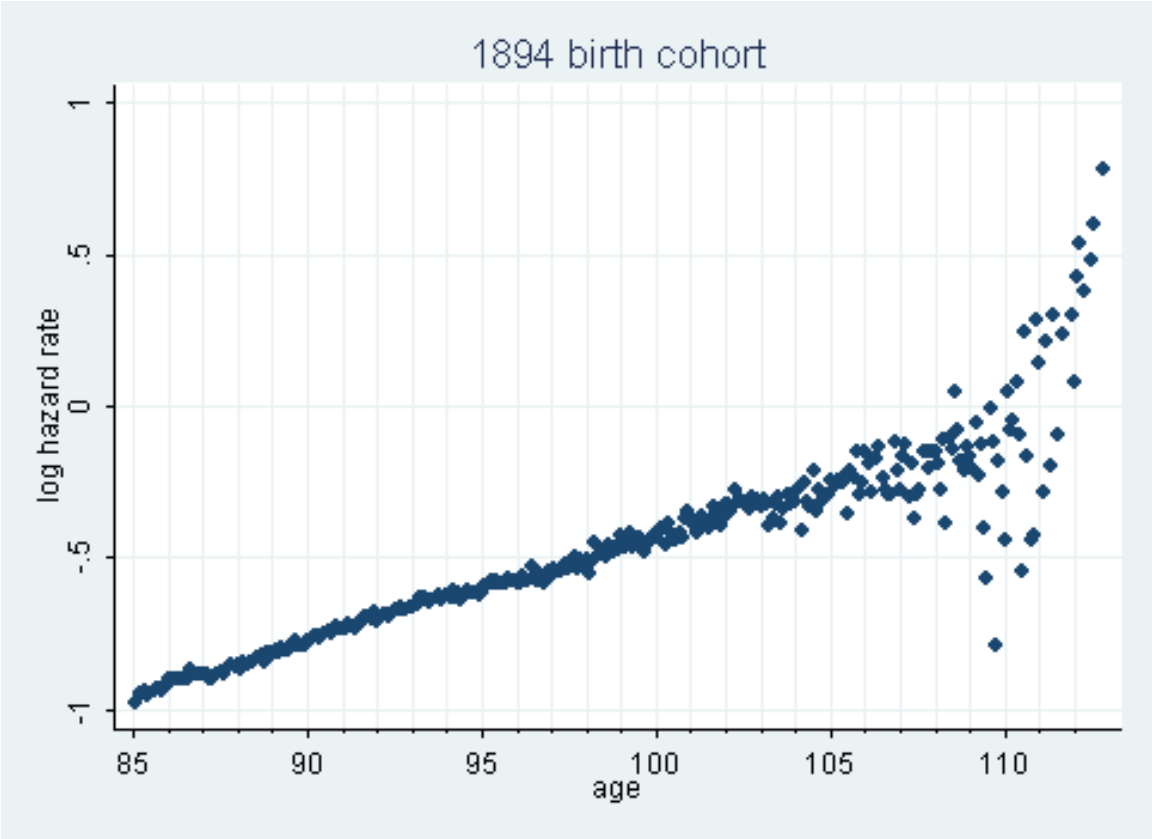
FIGURE 3
Hazard Rate (per year) for 1889 Birth Cohort.
Data for Southern and non-Southern states are split. Data are fitted using quadratic regression. Data from the Social Security Administration Death Master File



Note that quadratic fit of mortality for the non-Southern population is straight and shows good agreement with the Gompertz model, while mortality of Southern states has a tendency for mortality deceleration. Indeed, quadratic fit of mortality showed that mortality of non-Southern states has slightly positive quadratic term, i.e., mortality in these states grows slightly faster with age than predicted by the Gompertz model (possibly due to effects of data truncation in non-extinct birth cohorts). At the same time, mortality in Southern states with presumably lower quality of age reporting had slightly negative quadratic term demonstrating mortality deceleration. This example shows that lower quality of age reporting may indeed result in mortality deceleration.

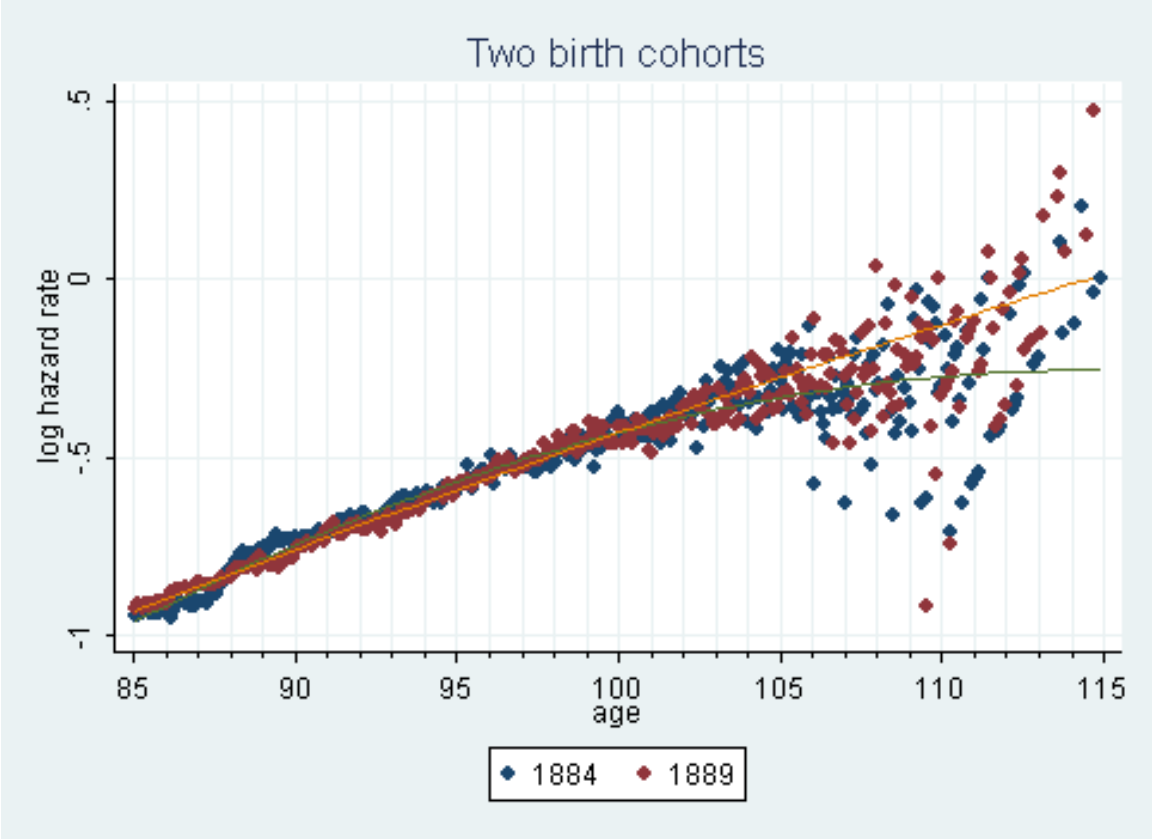
Mortality after age 110 years for cohorts born after 1890 is affected by data truncation (Figure 4). These examples (Figures 1-4) demonstrate that for a single-year birth cohort mortality agrees well with the Gompertz law up to very advanced ages. Previous studies of mortality at advanced ages used aggregated data combining several birth cohorts with different mortality, and this aggregation apparently resulted in early mortality deceleration and subsequent leveling-off as it was demonstrated by heterogeneity model (Beard, 1971). Mortality deceleration and even decline of mortality often is observed for data with low quality. On the other hand, improvement of data quality results in straighter mortality trajectory in semi-log scale (Kestenbaum and Ferguson, 2002). In our study, the more recent 1894 birth cohort demonstrates straighter trajectory and lower statistical noise after age 105 than the older 1884 one (see Figures 1 and 4). Thus, we may expect that cohorts born after 1894 would demonstrate even better fit by the Gompertz model than the older ones because of improved quality of age reporting. Testing this hypothesis now is hampered by the problem of data truncation for non-extinct birth cohorts.

FIGURE 4
Hazard Rate (per year) for 1894 Birth Cohort.
Data from the Social Security Administration Death Master File



Comparison of mortality rates between two birth cohorts (1884 and 1889) found no significant difference in mortality levels between individuals born in 1884 and 1889 (Figure 5). This finding is expected in view of small historical changes in mortality at old ages (80 and over) until the 1960s. Taking into account that individuals born in 1884 and 1889 were already octogenarians in the 1960s, we cannot expect significant differences in mortality between these two birth cohorts. Note, however, that mortality of the later birth cohort (with presumably better age reporting) is also closer to the Gompertz model.

FIGURE 5
Comparison of Hazard Rates for Two Birth Cohorts (1884 and 1889). Data are fitted using quadratic regression. Data from the Social Security Administration Death Master File



We already noted that the period of mortality deceleration in mammals is very short compared to lower organisms. It appears to be relatively short in humans too. This observation agrees well with the prediction of the reliability theory of aging that more complex living systems/organisms with many vital subsystems (like mammals) may experience very short or no period of mortality plateau at advance ages in contrast to more simple living organisms (Gavrilov and Gavrilova, 1991; 2001; 2003a; 2004; 2006).

6. Gender-Specific Differences in Mortality after Age 100

The SSA DMF does not provide information about the sex of the deceased. To avoid this limitation of the data sample, we conducted a procedure of sex identification

using information about the 1,000 most commonly used baby names in the 1900s provided by the Social Security Administration (http://www.ssa.gov/OACT/babynames). These data come from a sample of 5 percent of all Social Security cards issued to individuals who were born during the 1900s in the United States. From the lists of male and female first names, we removed first names consisting of initials and names for which sex was unclear (like Jessie or Lonnie). It is interesting to note that the SSA male list contains explicitly female first names (Mary, Elizabeth) and the same problem was observed for the female list, which indicates that SSA data apparently contain many sex misidentifications. These female names were removed from the male list and the same procedure was done for the female list. Using the final lists of male and female first names, we identified sex in approximately 90 percent of cases for the studied birth cohorts of persons aged 85 years and over. The remaining 10 percent of persons with unknown sex had the same mean lifespan as the remaining 90 percent of individuals with identified sex pooled together, so the existence of possible sex bias in our sex identification looks unlikely. Data samples for 1886, 1887 and 1889 birth cohorts were used for a more detailed study of gender-specific mortality at advanced ages.

The result of hazard rate estimation for males and females is presented in Figure 6. Note that male mortality continues to exceed female mortality up to very advanced ages, and this difference narrows very slowly with age. At age 110 years the number of remaining males (9 persons) and females (44 persons) is too small for accurate estimates of hazard rate after this age (data for 1886 birth cohort).

FIGURE 6
Hazard Rate (per year) for Males and Females from 1887 Birth Cohort. Data are fitted using quadratic regression.
Data from the Social Security Administration Death Master File

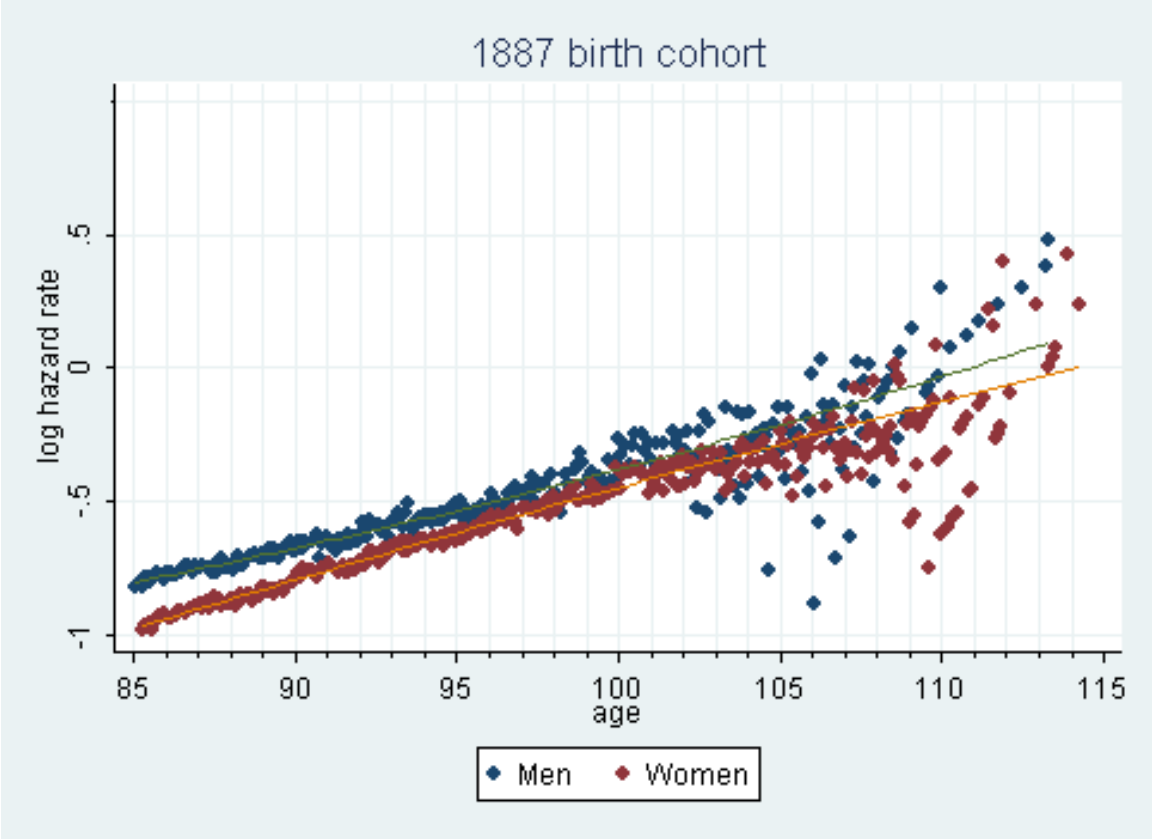
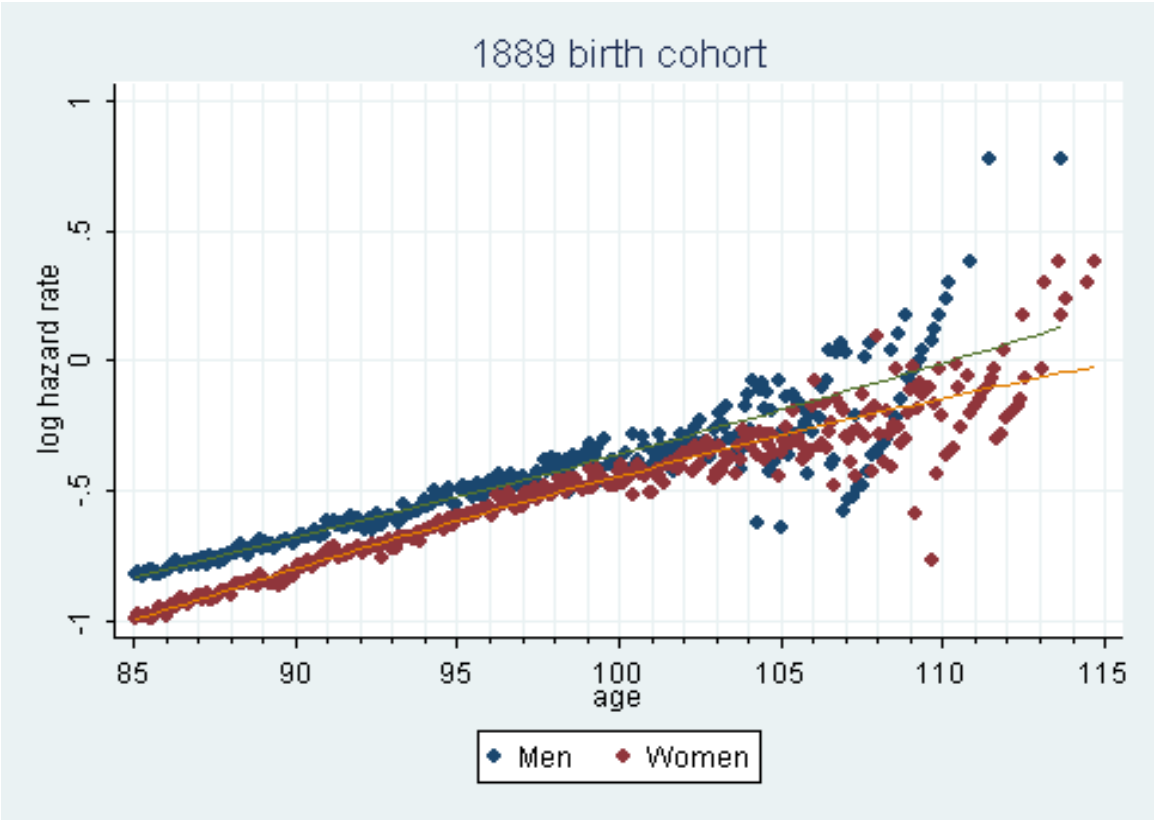


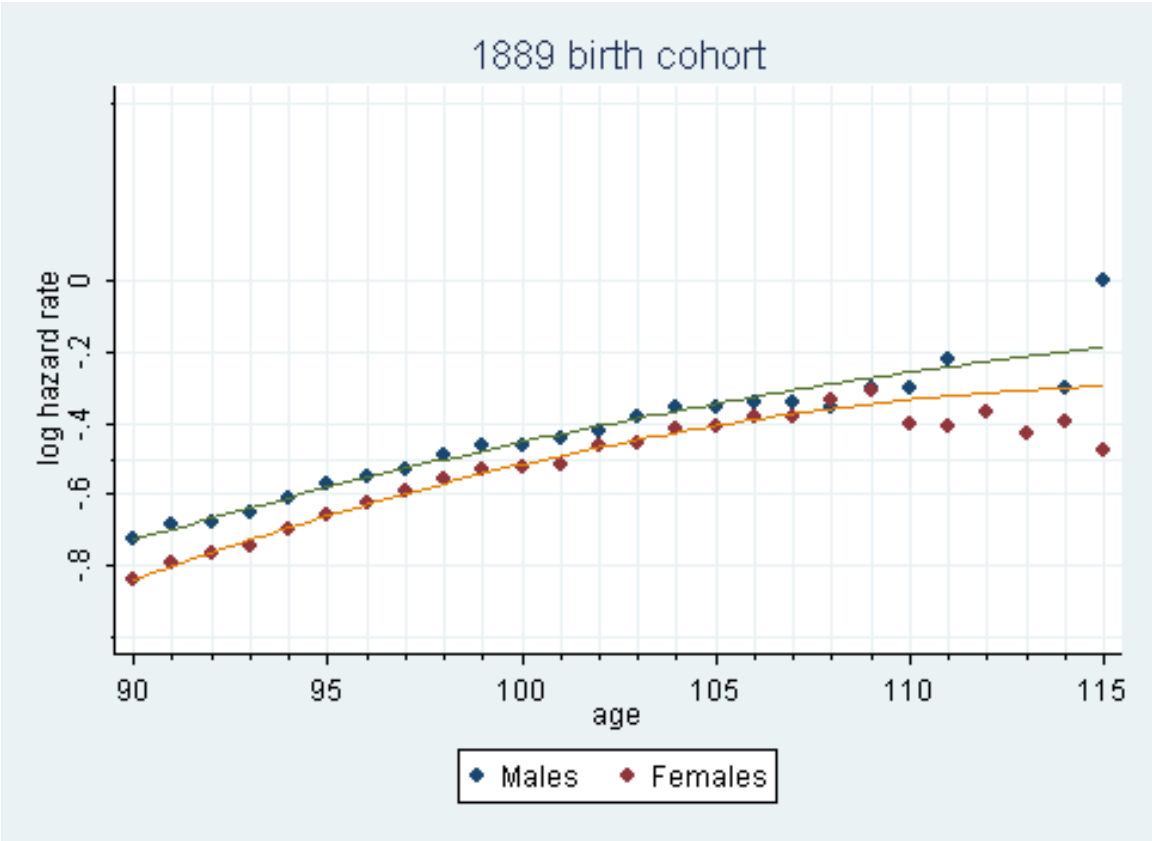
Figure 7 shows mortality of men and women for the 1889 birth cohort, which is essentially the same as for the 1887 birth cohort (see Figure 6).

FIGURE 7
Hazard Rate (per year) for Males and Females from 1889 birth cohort. Data are fitted using quadratic regression.
Data from the Social Security Administration Death Master File



Interestingly, the hazard rate trajectory based on crude estimates of lifespan measured in whole years (like in ordinary demographic life tables) creates an impression of more pronounced mortality deceleration (Figure 8) than estimates obtained for every month of life (Figure 7).

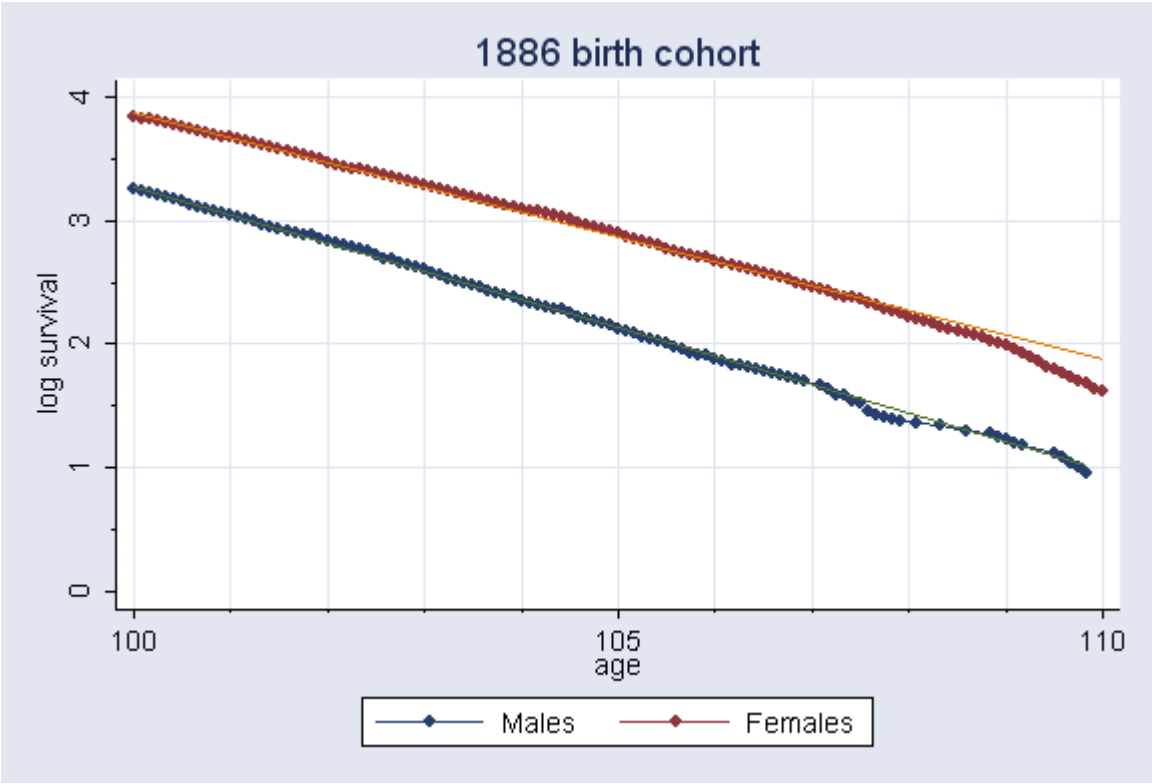
FIGURE 8
Hazard Rate (per year) for Males and Females from 1889 Birth Cohort. Lifespan is estimated in whole years. Data are fitted using quadratic regression.
Data from the Social Security Administration Death Master File



7. Other Ways to Check for Mortality Leveling-Off at Advanced Ages

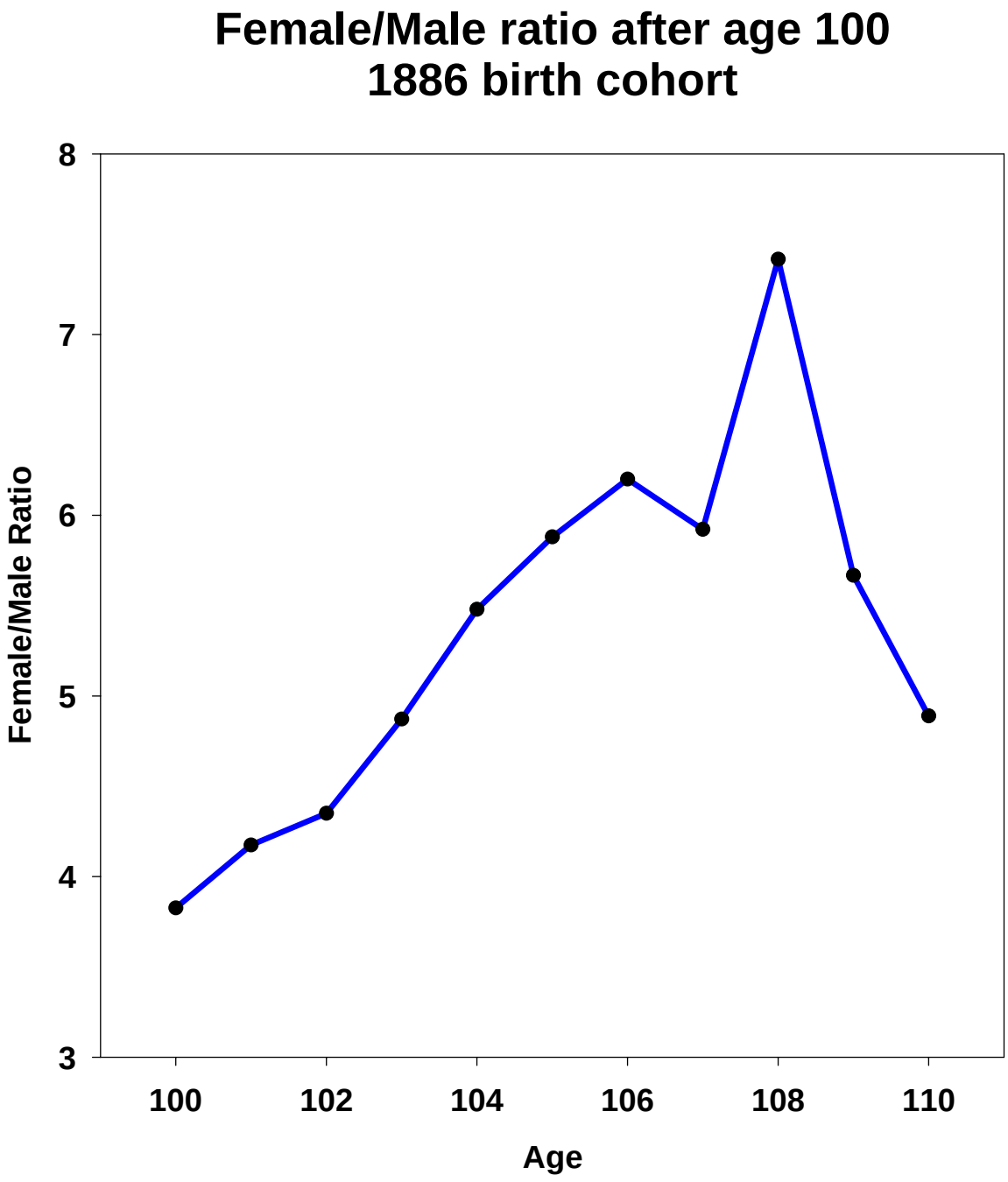
Another way to check the leveling-off of mortality at advanced ages is to draw a plot of survival curve in semilog coordinates. This plot should be linear if mortality at advanced ages is constant (leveled off). The plot presented in Figure 9 shows some deviations from the linear dependence.

FIGURE 9
Logarithm of Survival as a Function of Age for Males and Females from 1886 Birth Cohort. Data are fitted using liner regression.
Data from the Social Security Administration Death Master File



Male mortality after age 100 is higher than female mortality, leading to a rapid increase in the proportion of females among the extremely old (see Figure 10). Note the decline of female proportion after age 108 years, which may suggest a possibility of age exaggeration by some very old males.

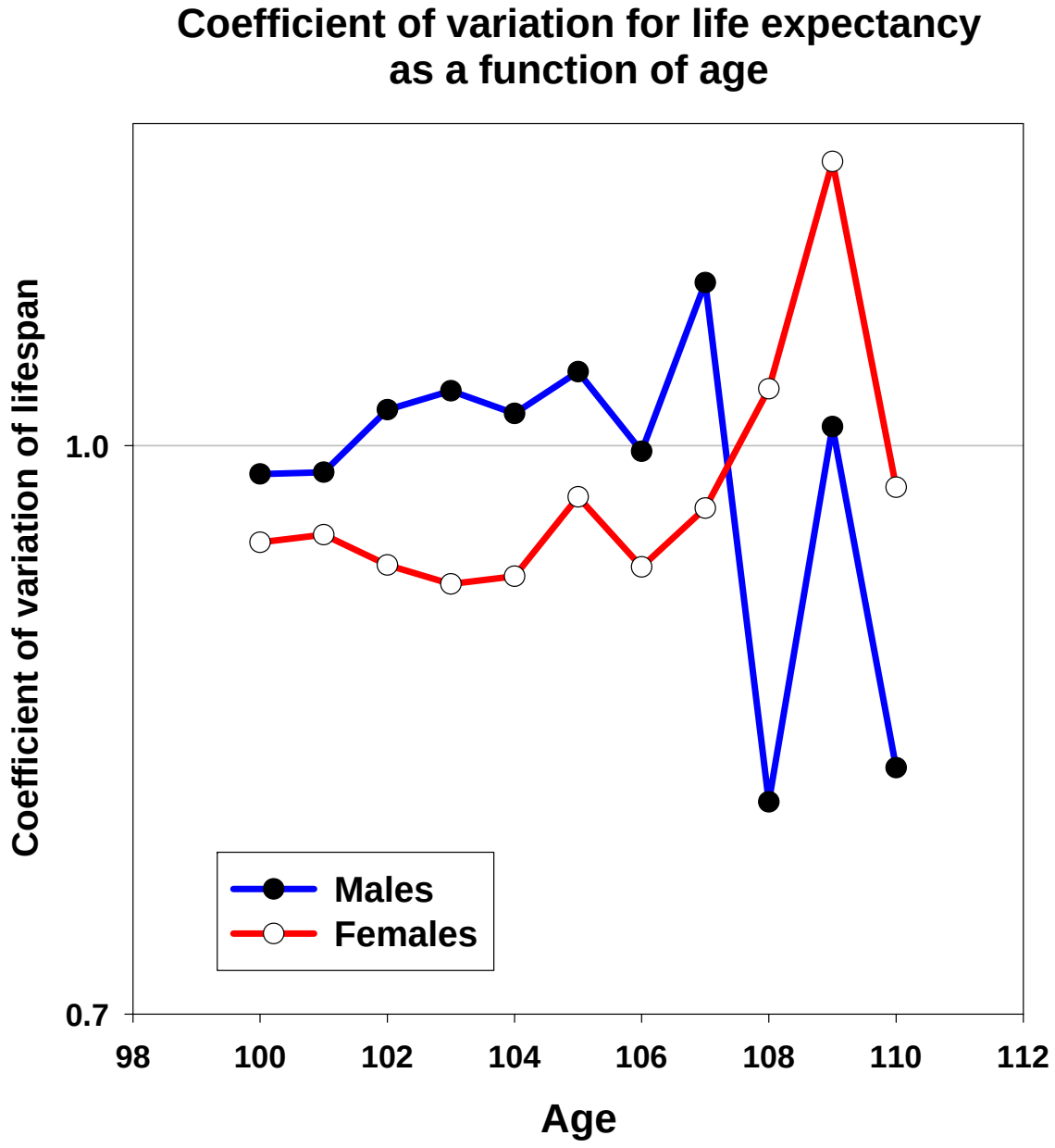
FIGURE 10
Female/Male Ratio as a Function of Age for the 1886 Birth Cohort.
Data from the Social Security Administration Death Master File



Visual inspection may be sometimes misleading, so we used nonparametric measures of mortality leveling-off. One such measure is a coefficient of variation for life expectancy at different ages:
$CV = \sigma/\mu$, where σ is a standard deviation and μ is mean.

This is a dimensionless characteristic and often is reported as a percentage. For non-aging systems with constant mortality over age (mortality kinetics of radioactive decay) the coefficient of variation is equal to 100 percent, while for aging systems it is lower. On the opposite, "rejuvenating" systems with declining mortality have coefficient of variation for lifespan higher than 100 percent. We estimated coefficient of variation for life expectancy separately for males and females at different ages, starting with age 100 years. The results of our estimations are presented in Figure 11. Note that males demonstrate a non-aging mortality kinetics rather early—after age 102 years while females continue to demonstrate an aging kinetics up to age 106 years.

FIGURE 11
Coefficient of Variation for Life Expectancy as a Function of Age for the 1886 Birth Cohort. Data from the Social Security Administration Death Master File



These results do not show an increase of the coefficient of variation for life expectancy over 100 percent at advanced ages, so the decline of mortality at extremely high ages seems unlikely. We should also keep in mind that age exaggeration is more common for old men than for old women, so that the observed early non-aging kinetics among men may be an artifact of age overestimation.

8. Conclusions

The SSA Death Master File (DMF) covers deaths that occurred in the period 1937-2007 and is considered by some researchers superior in quality to the official U.S. vital statistics. Some birth cohorts in DMF may be considered as extinct or almost extinct. Detailed information about birth and death dates of decedents allowed us to estimate hazard rates of older persons at each month of age. Study of several extinct birth cohorts showed that mortality grows steadily without deceleration from 85 to 102-105 years of age. Then statistical noise rapidly increases and mortality tends to decelerate. Thus the study of large and relatively homogeneous birth cohorts demonstrates that late-life mortality deceleration appears to be not that strong—cohort mortality at advanced ages continues to grow up to age 105 years. On the other hand, non-parametric approach demonstrated that the late-life mortality plateau is likely not an artifact and is expressed earlier in males than in females. Age exaggeration and crude assumptions applied to mortality estimates at advanced ages may also contribute to mortality underestimation at very advanced ages.

Acknowledgments

This study was made possible thanks to a generous support from the Society of Actuaries as well as the Center on the Demography and Economics of Aging pilot grant (P30 AG012857), National Institute on Aging NIA (K02 AG00976). We are grateful to Thomas Edwalds, Kenneth Faig and Ward Kinkade for helpful comments and suggestions on the earlier version of this manuscript as well as to the participants of the International Symposium “Living to 100” in Orlando, FL (January 2008) for stimulating discussion.

References

Allison, P. 1995. *Survival Analysis Using the SAS ® System: A Practical Guide*. SAS Institute.

Austad, S.N. 2001. "Concepts and Theories of Aging." In E.J. Masoro and S.N. Austad. *Handbook of the Biology of Aging,* pp. 3-22. San Diego, CA: Academic Press.

Beard, R.E. 1959. "Note on Some Mathematical Mortality Models. In: G.E.W. Wolstenholme and M. O'Connor (eds.). *The lifespan of animals,* pp. 302-311. Boston: Little, Brown.

Beard, R.E. 1971. "Some Aspects of Theories of Mortality, Cause of Death Analysis, Forecasting and Stochastic Processes." In W. Brass (ed.). *Biological Aspects of Demography,* pp. 57-68. London: Taylor & Francis.

Bronikowski, A.M., Alberts, S.C., Altmann, J., Packer, C., Carey, K.D., and Tatar, M. 2002. "The Aging Baboon: Comparative Demography in a Non-Human Primate." Proc. Natl. Acad. Sci. USA 99: 9591-9595.

Carey, J.R., Liedo, P., Orozco, D., and Vaupel, J.W. 1992. "Slowing of Mortality Rates at Older Ages in Large Medfly Cohorts." *Science* 258: 457-461.

Costa, D.L., and Lahey, J. 2003. Becoming Oldest-Old: Evidence from Historical U.S. Data. NBER Working Paper No. W9933. http://ssrn.com/abstract=439614.

Curtsinger, J.W., Fukui, H., Townsend, D., and Vaupel, J.W. 1992. Demography of Genotypes: Failure of the Limited Life-Span Paradigm in *Drosophila melanogaster*." *Science* 258 : 461-463.

Curtsinger, J.W., Gavrilova, N.S., and Gavrilov, L.A. 2006. "Biodemography of Aging and Age-Specific Mortality in Drosophila melanogaster." In E.J. Masoro and S.N. Austad (eds.), *Handbook of the Biology of Aging*, pp. 261-288. San Diego: Academic Press.

Depoid, F. 1973. "La Mortalite des Grands Viellards." *Population* 28: 755-92.

Doblhammer, G. 1999. "Longevity and Month of Birth: Evidence from Austria and Denmark." *Demographic Research* [Online] 1, 1-22. Available: http://www.demographic-research.org/Volumes/Vol1/3/default.htm

_____. 2003. "The Late Life Legacy of Very Early Life." Rostock, *MPIDR Working Paper WP-2003-030*.

Doblhammer G, Vaupel, J.W. 2001. "Lifespan Depends on Month of Birth." Proc. Natl. Acad. USA 98: 2934-2939.

Economos, A.C. 1979. "A Non-Gompertzian Paradigm for Mortality Kinetics of Metazoan Animals and Failure Kinetics of Manufactured Products. AGE 2: 74-76.

_____. 1980. Kinetics of metazoan mortality. *J. Social Biol. Struct.* 3: 317-329.

_____. 1983. "Rate of Aging, Rate of Dying and the Mechanism of Mortality." Arch. Gerontol. and Geriatrics 1: 3-27.

_____. 1985. "Rate of Aging, Rate of Dying and Non-Gompertzian Mortality—Encore. *Gerontology* 31: 106-111.

Faig, K. 2002. "Reported Deaths of Centenarians and Near-Centenarians in the U.S. Social Security Administration's Death Master File." In: Proceedings of the Society of Actuaries "Living to 100 and Beyond International Symposium," Orlando, FL.

Fukui, H.H., Ackert, L., and Curtsinger, J.W. 1996. "Deceleration of Age-Specific Mortality Rates in Chromosomal Homozygotes and Heterozygotes of Drosophila melanogaster." Experimental Gerontology 31: 517-531.

Fukui, H.H., Xiu, L., and Curtsinger, J.W. 1993. "Slowing of Age-Specific Mortality Rates in Drosophila melanogaster." Experimental Gerontology 28: 585-599.

Gavrilov, L.A. 1980. "Study of Life Span Genetics Using the Kinetic Analysis." Ph.D. Thesis, Moscow, Russia: Moscow State University.

_____. 1984. "Does a Limit of the Life Span Really Exist?" *Biofizika* 29: 908-911.

Gavrilov, L.A., Gavrilova, N.S. 1991. *The Biology of Life Span: A Quantitative Approach*. New York: Harwood Academic Publisher.

_____. 1999. "Season of Birth and Human Longevity." Journal of Anti-Aging Medicine 2: 365-366.

_____. 2001. "The Reliability Theory of Aging and Longevity." J. Theor. Biol. 213: 527-545.

_____. 2002. "Early-Life Seasonal Programming of Adult Lifespan: Evidence from the 19th Century Birth Cohorts." Annual Meeting of the Social Science History Association, St. Louis, 24-27 October 2002. Available at: http://www.ssha.org/abstract2002/abs348.html.

_____. 2003a. "The Quest for a General Theory of Aging and Longevity." Science's SAGE KE (Science of Aging Knowledge Environment) for 16 July 2003; Vol. 2003, No. 28, 1-10.

_____. 2003b. “Early-Life Factors Modulating Lifespan.” In Rattan, S.I.S. (ed.). *Modulating Aging and Longevity*, pp. 27-50. Dordrecht, The Netherlands: Kluwer Academic Publishers.

_____. 2004. “Early-Life Programming of Aging and Longevity: The Idea of High Initial Damage Load (the HIDL Hypothesis).” Annals of the New York Academy of Sciences 1019: 496-501.

_____. 2006. “Reliability Theory of Aging and Longevity.” In E.J. Masoro and S.N. Austad (eds.). *Handbook of the Biology of Aging*, pp. 3-42. San Diego: Academic Press.

_____. 2005. “Search for Predictors of Exceptional Human Longevity: Using Computerized Genealogies and Internet Resources for Human Longevity Studies.” In: Proceedings of the Society of Actuaries “Living to 100 and Beyond International Symposium”, Orlando, FL.

Gavrilova, N.S., Gavrilov, L.A., Evdokushkina, G.N., and Semyonova, V.G. 2003. “Early-Life Predictors of Human Longevity: Analysis of the 19th Century Birth Cohorts.” Annales de Demographie Historique 2: 177-198.

Gompertz, B. 1825. “On the Nature of the Function Expressive of the Law of Human Mortality and on a New Mode of Determining Life Contingencies.” Philos.Trans.Roy.Soc.London A 115: 513-585.

Greenwood, M., and Irwin, J.O. 1939. “The Biostatistics of Senility.” Hum. Biol. 11: 1-23.

Hill, M.E., and Rosenwaike, I. 2001. “The Social Security Administration's Death Master File: The Completeness of Death Reporting at Older Ages.” Soc Secur Bull. 64: 45-51.

Horiuchi, S, and Wilmoth, J.R. 1998.“Deceleration in the Age Pattern of Mortality at Older Ages.” *Demography* 35: 391-412.

Kestenbaum, B. 1992. “A Description of the Extreme Aged Population Based on Improved Medicare Enrollment Data.” Demography 29: 565-80.

Kestenbaum B., and Ferguson, B.R. 2002. “Mortality of the Extreme Aged in the United States in the 1990s, Based on Improved Medicare Data.” In Proceedings of the Society of Actuaries “Living to 100 and Beyond International Symposium”, Orlando, FL.

Laake, K., and Sverre, J.M. 1996. “Winter Excess Mortality: A Comparison between Norway and England plus Wales.” Age and Ageing 25: 343-348.

Le Bras, H. 1976. “Lois de Mortalité et Age Limité.” Population 31: 655-692.

Lindop, P.J. 1961. “Growth Rate, Lifespan and Causes of Death in SAS/4 Mice.” Gerontologia 5: 193-208.

Olshansky, S.J. 1998. “On the Biodemography of Aging: A Review Essay.” Population and Development Review 24: 381-393.

Perks, W. 1932. “On Some Experiments in the Graduation of Mortality Statistics.” Journal of the Institute of Actuaries 63: 12-57.

Robine, J-M, and Vaupel J.W. 2001. “Supercentenarians: Slower Ageing Individuals or Senile Elderly?” Experimental Gerontology 36: 915-930.

Rosenwaike, I., and Logue, B. 1983. “Accuracy of Death Certificate Ages for the Extreme Aged." Demography 20: 569-85.

Rosenwaike, I., Hill, M., Preston, S., and Elo, I. 1998. **“**Linking Death Certificates to Early Census Records: The African American Matched Records Sample.” Historical Methods 31: 65-74.

Rosenwaike, I., and Stone, L.F. 2003. “Verification of the Ages of Supercentenarians in the United States: Results of a Matching Study.” Demography 40: 727-739.

Sacher, G.A. 1956. “On the Statistical Nature of Mortality, with Especial Reference to Chronic Radiation Mortality. Radiology 67: 250-257.

Sacher, G.A. 1966. “The Gompertz Transformation in the Study of the Injury-Mortality Relationship: Application to Late Radiation Effects and Ageing.” In P.J. Lindop and G.A. Sacher (eds.). *Radiation and Ageing,* pp. 411-441. London: Taylor and Francis.

Shrestha, L.B., and Preston, S.H. 1995. “Consistency of Census and Vital Registration Data on Older Americans: 1970-1990.” Survey Methodology 21: 167-177.

Tatar, M., Carey, J.R., and Vaupel, J.W. 1993. “Long-Term Cost of Reproduction with and without Accelerated Senescence in *Callosobruchus maculatus*: Analysis of Age-Specific Mortality.” Evolution 47: 1302-1312.

Thatcher, A.R. 1999. “The Long-Term Pattern of Adult Mortality and the Highest Attained Age.” J. R. Statist. Soc. A 162(1): 5-43.

Vaupel, J.W., Carey, J.R., Christensen, K., Johnson, T., Yashin, A.I., Holm, N.V., Iachine, I.A., Kannisto, V., Khazaeli, A.A., Liedo, P., Longo, V.D., Zeng, Y., Manton, K., and Curtsinger, J.W. 1998. “Biodemographic Trajectories of Longevity.“ *Science* 280 : 855-860.

Vincent, P. 1951. “La Mortalite des Viellards.” Population 6 : 181-204.

Wilkinson, P., Pattenden, S., Armstrong, B., Fletcher, A., Kovats, R.S., Mangtani, P., and McMichael, A.J. 2004. "Vulnerability to Winter Mortality in Elderly People in Britain: Population-Based Study." BMJ 329: 647.

Wilmoth, J.R. 1997. "In Search of Limits." In K.W. Wachter and C.E. Finch (eds.). *Between Zeus and the Salmon. The Biodemography of Longevity,* pp. 38-64. Washington, DC: National Academy Press.