Mortality Measurement Beyond Age 100

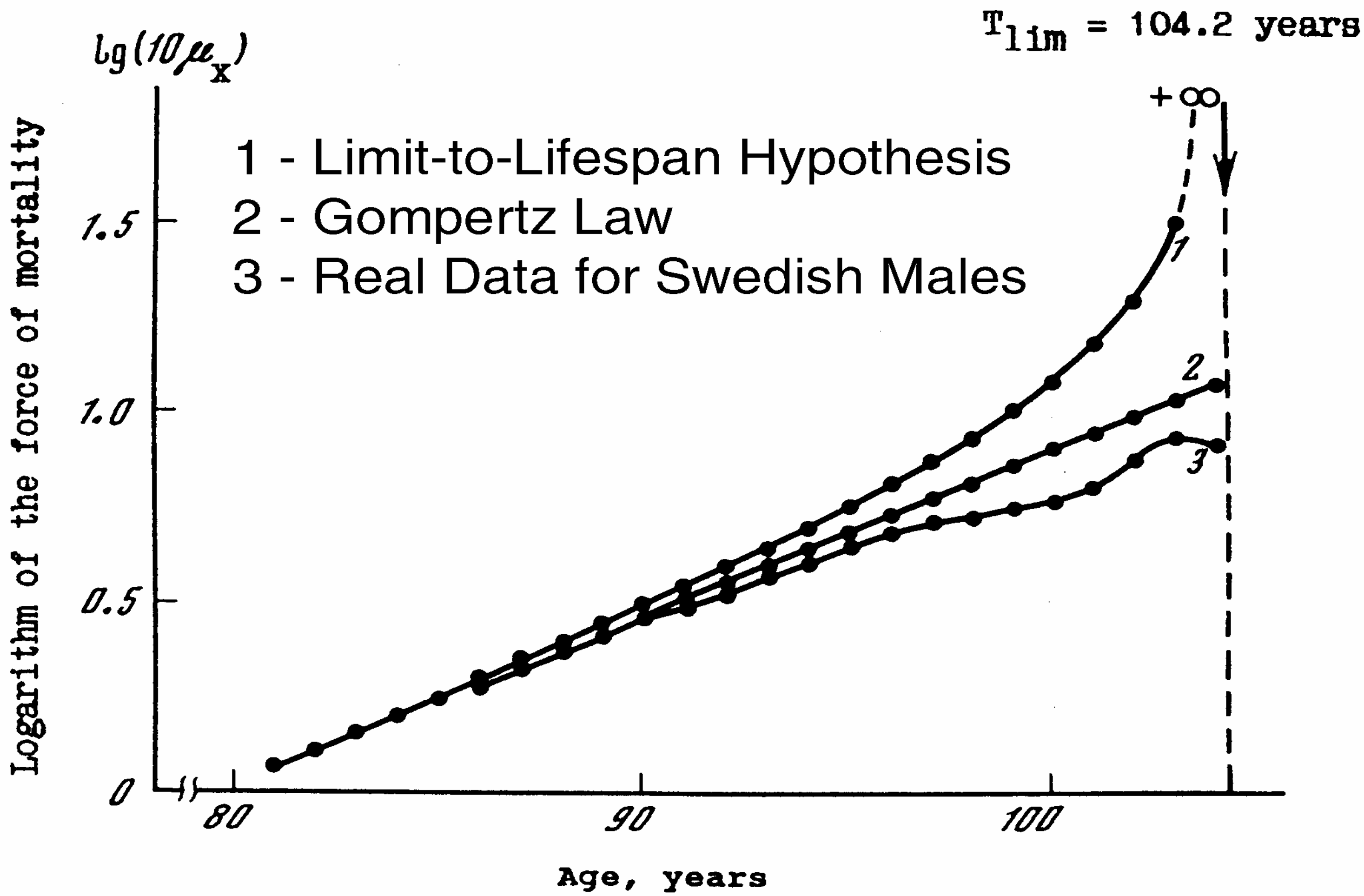
Dr. Natalia S. Gavrilova, Ph.D.
Dr. Leonid A. Gavrilov, Ph.D.

Center on Aging
NORC and The University of Chicago
Chicago, Illinois, USA

A growing number of persons living beyond age 100 emphasizes the need for accurate measurement and modeling of mortality at advanced ages.

What do we know about late-life mortality?

Mortality at Advanced Ages



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

Mortality Deceleration in Other Species

Invertebrates:

- Nematodes, shrimps, bdelloid rotifers, degenerate medusae (Economos, 1979)
- Drosophila melanogaster (Economos, 1979; Curtsinger et al., 1992)
- Medfly (Carey et al., 1992)
- Housefly, blowfly (Gavrilov, 1980)
- Fruit flies, parasitoid wasp (Vaupel et al., 1998)
- Bruchid beetle (Tatar et al., 1993)

Mammals:

- Mice (Lindop, 1961; Sacher, 1966; Economos, 1979)
- Rats (Sacher, 1966)
- Horse, Sheep, Guinea pig (Economos, 1979; 1980)

However no mortality deceleration is reported for

- Rodents (Austad, 2001)
- Baboons (Bronikowski et al., 2002)

Existing Explanations of Mortality Deceleration

- **Population Heterogeneity (Beard, 1959; Sacher, 1966).** *"... sub-populations with the higher injury levels die out more rapidly, resulting in progressive selection for vigour in the surviving populations*" **(Sacher, 1966)**
- **Exhaustion of organism's redundancy (reserves) at extremely old ages so that every random hit results in death (Gavrilov, Gavrilova, 1991; 2001)**
- **Lower risks of death for older people due to less risky behavior (Greenwood, Irwin, 1939)**
- **Evolutionary explanations (Mueller, Rose, 1996; Charlesworth, 2001)**

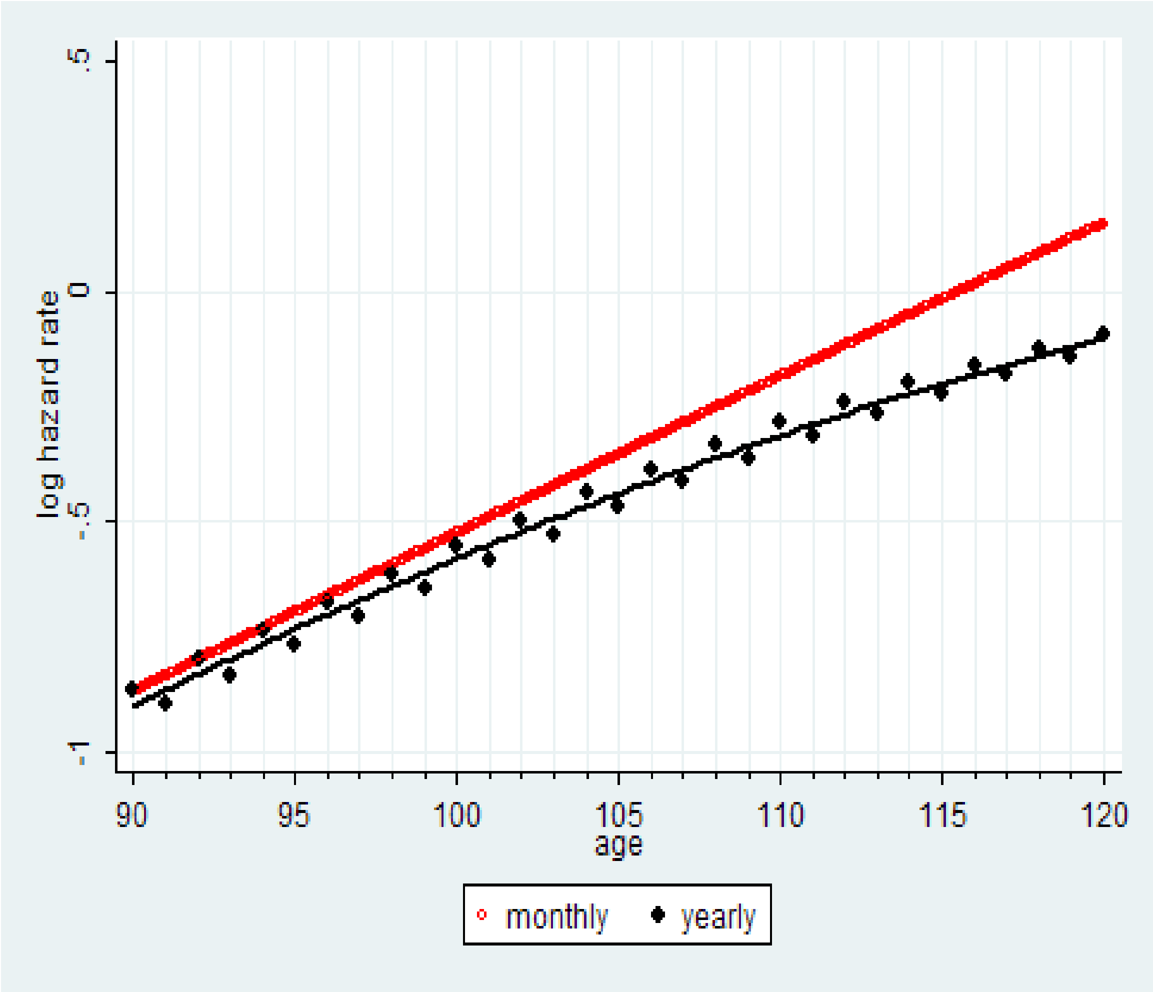
Mortality force (hazard rate) is the best indicator to study mortality at advanced ages

$$\mu_x = -\frac{dN_x}{N_x dx} = -\frac{d\ln(N_x)}{dx} \approx -\frac{\Delta \ln(N_x)}{\Delta x}$$

- **Does not depend on the length of age interval**
- **Has no upper boundary and theoretically can grow unlimitedly**
- **Famous Gompertz law was proposed for fitting age-specific mortality force function (Gompertz, 1825)**

Monthly Estimates of Mortality are More Accurate

Simulation assuming Gompertz law for hazard rate



Stata package uses the Nelson-Aalen estimate of hazard rate:

$$\mu_x = \mathrm{H}(x) - \mathrm{H}(x-1) = \frac{d_x}{n_x}$$

H(x) is a cumulative hazard function, d_x is the number of deaths occurring at time x and n_x is the number at risk at time x before the occurrence of the deaths. This method is equivalent to calculation of probabilities of death:

$$q_x = \frac{d_x}{l_x}$$

Problems in Hazard Rate Estimation At Extremely Old Ages

1. Mortality deceleration in humans may be 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
3. Ages of very old people may be highly exaggerated

Social Security Administration Death Master File Helps to Alleviate the First Two Problems

- Allows to study mortality in large, more homogeneous single-year or even single-month birth cohorts
- Allows to estimate mortality in one-month age intervals narrowing the interval of hazard rates estimation

What Is SSA DMF ?

- SSA DMF is a publicly available data resource (available at Rootsweb.com)
- Covers 93-96 percent deaths of persons 65+ occurred in the United States in the period 1937-2010
- Some birth cohorts covered by DMF could be studied by the method of extinct generations
- Considered superior in data quality compared to vital statistics records by some researchers

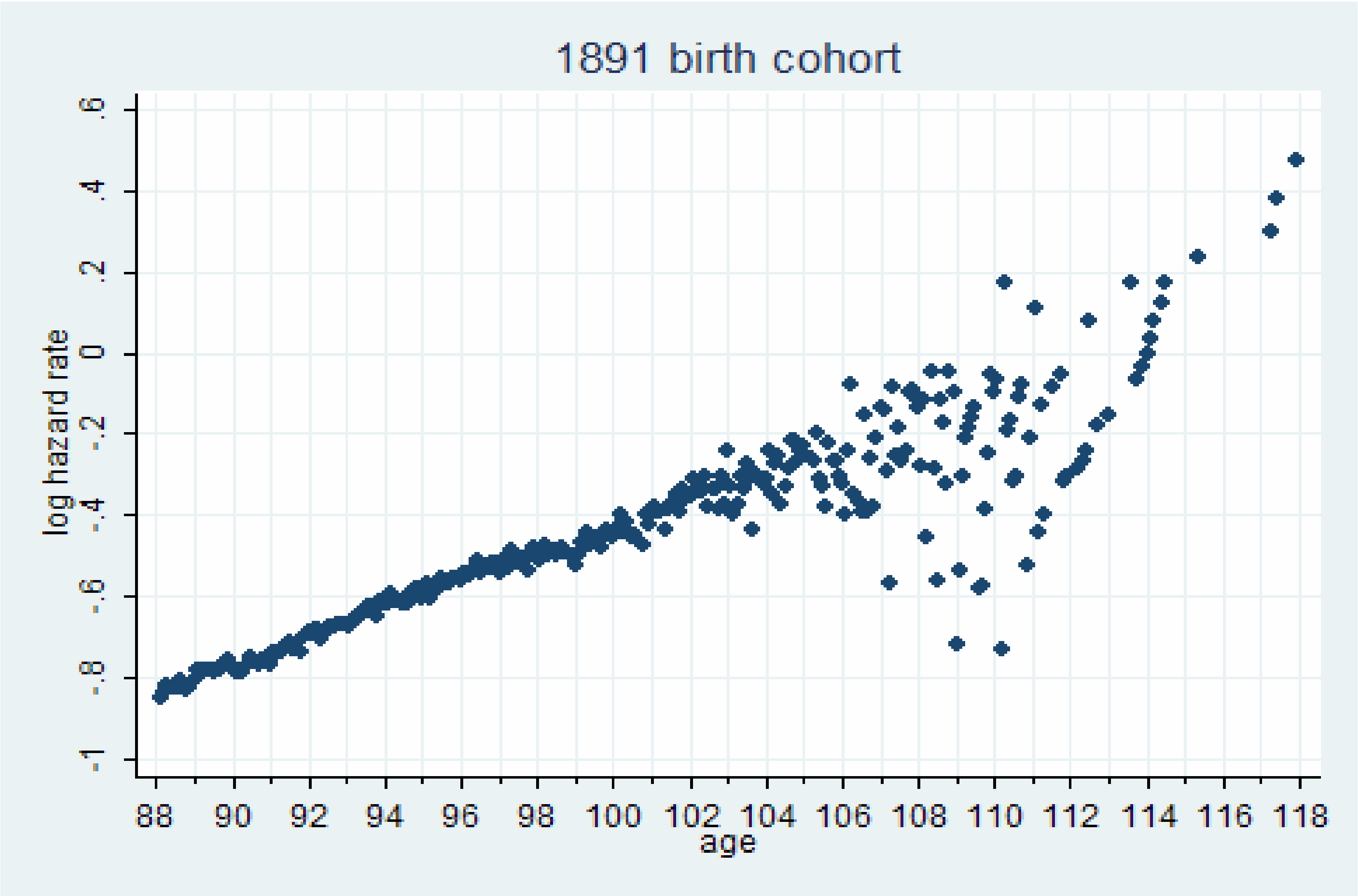
Social Security Administration Death Master File (DMF) Was Used in This Study:

To estimate hazard rates for relatively homogeneous single-year extinct birth cohorts (1881-1895)

To obtain monthly rather than traditional annual estimates of hazard rates

To identify the age interval and cohort with reasonably good data quality and compare mortality models

Hazard rate estimates at advanced ages based on DMF



Nelson-Aalen estimates of hazard rates using Stata 11

Hypothesis

Mortality deceleration at advanced ages among DMF cohorts may be caused by poor data quality (age exaggeration) at very advanced ages

If this hypothesis is correct then mortality deceleration at advanced ages should be less expressed for data with better quality

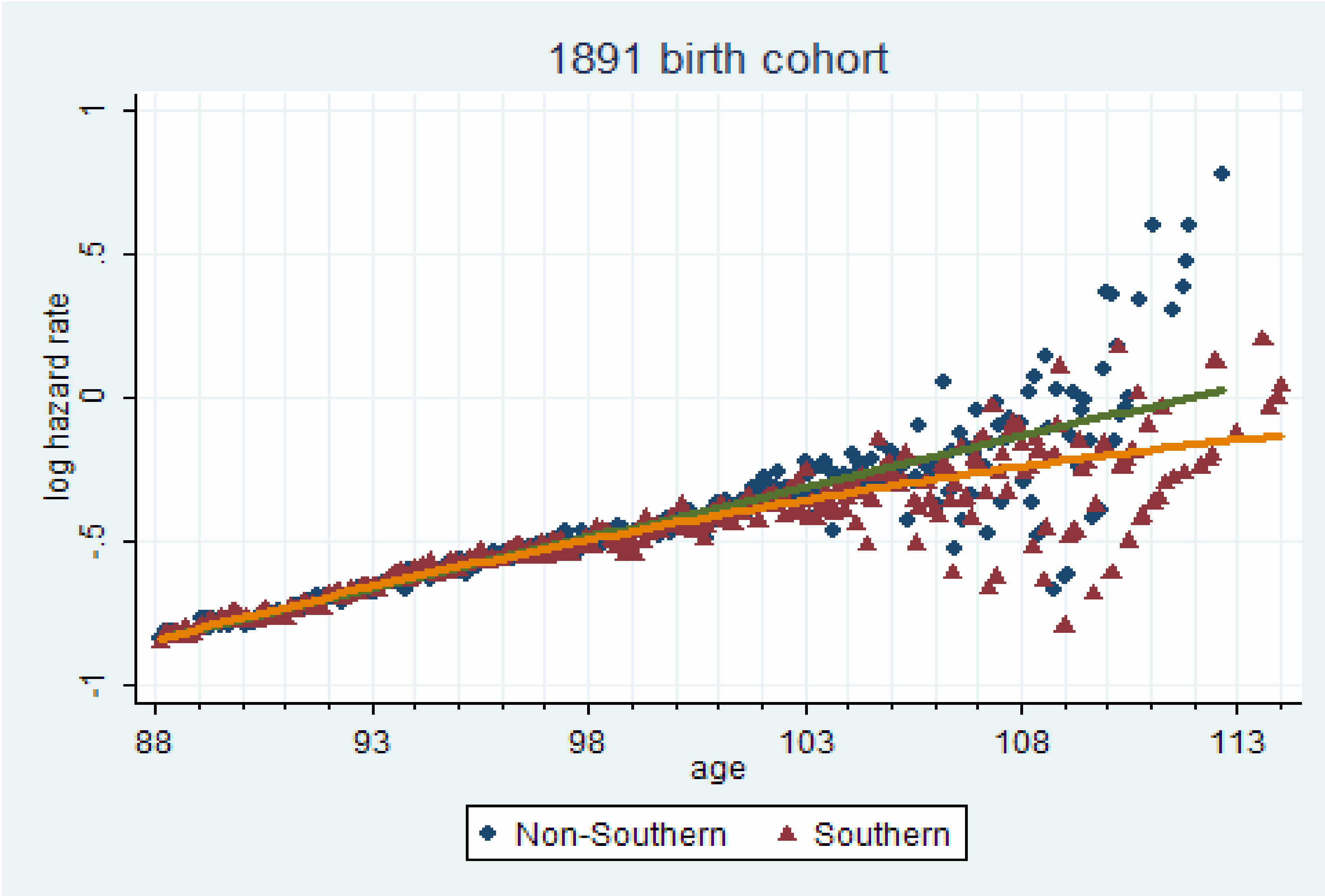
Quality Control (1)

Study of mortality in the states with different quality of age reporting:

Records for persons applied to SSN in the Southern states were found to be of lower quality (Rosenwaike, Stone, 2003)

We compared mortality of persons applied to SSN in Southern states, Hawaii, Puerto Rico, CA and NY with mortality of persons applied in the Northern states (the remainder)

Mortality for data with presumably different quality



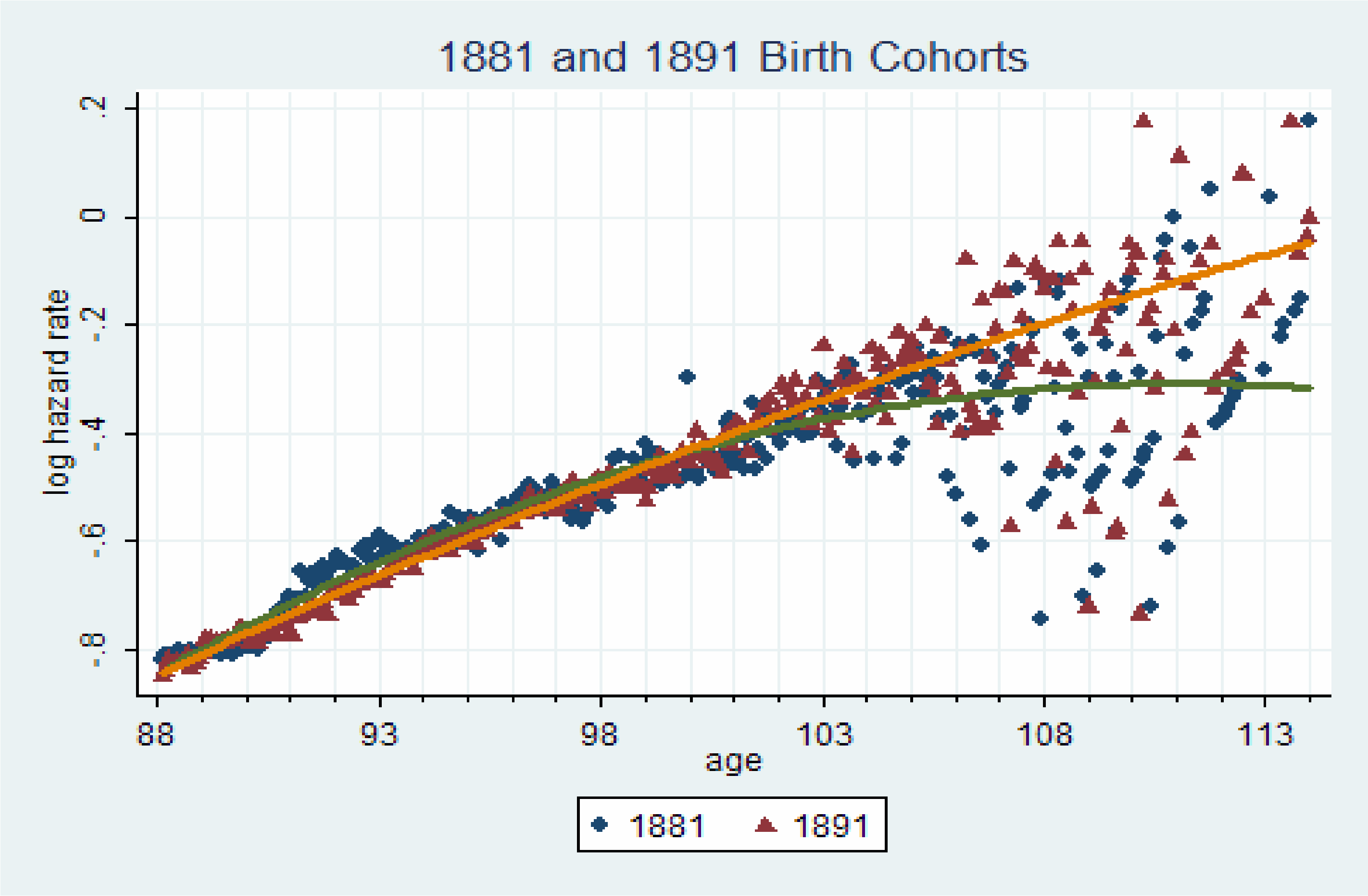
The degree of deceleration was evaluated using quadratic model

Quality Control (2)

Study of mortality for earlier and later single-year extinct birth cohorts:

Records for later born persons are supposed to be of better quality due to improvement of age reporting over time.

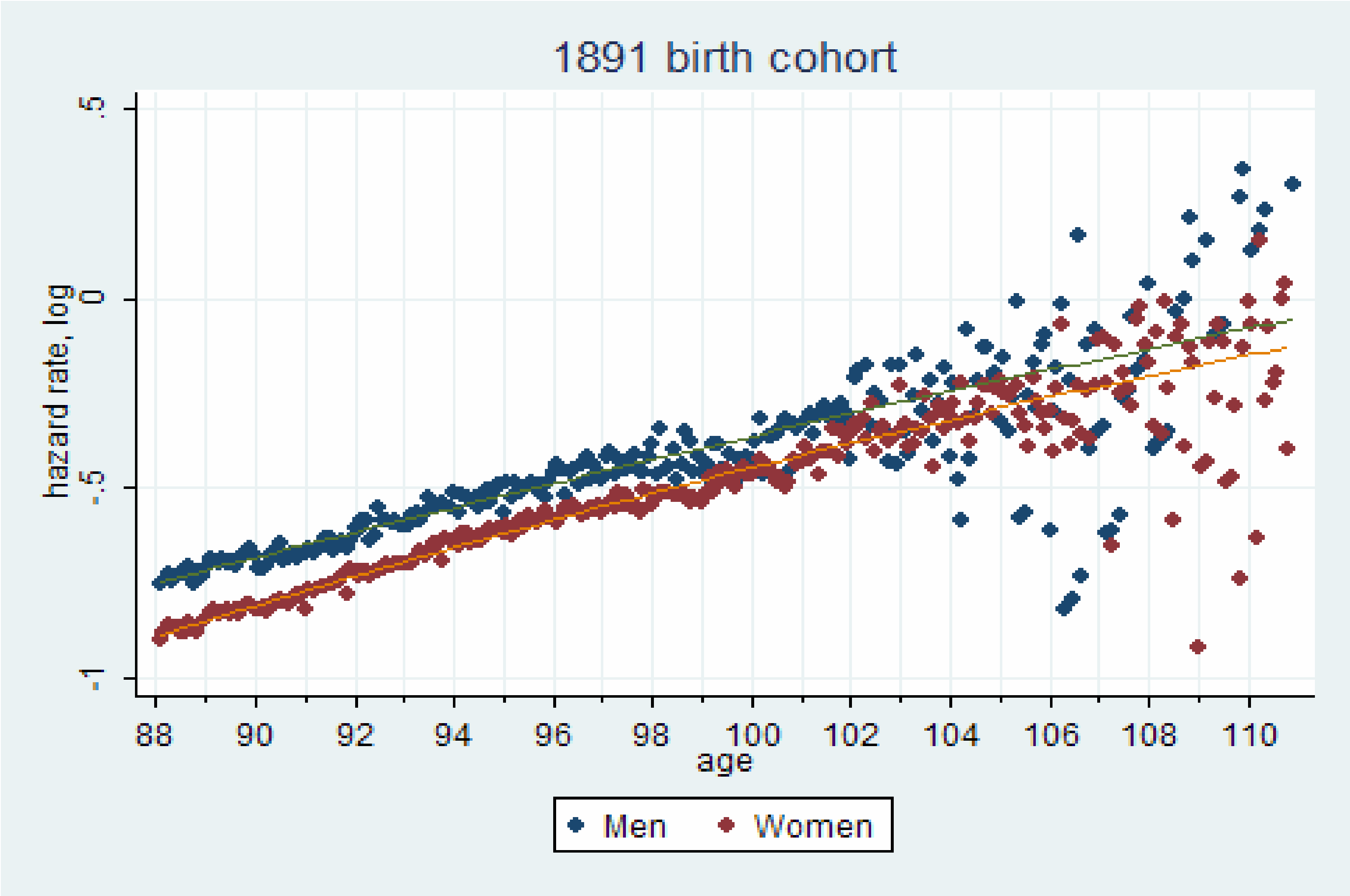
Mortality for data with presumably different quality



At what age interval data have reasonably good quality?

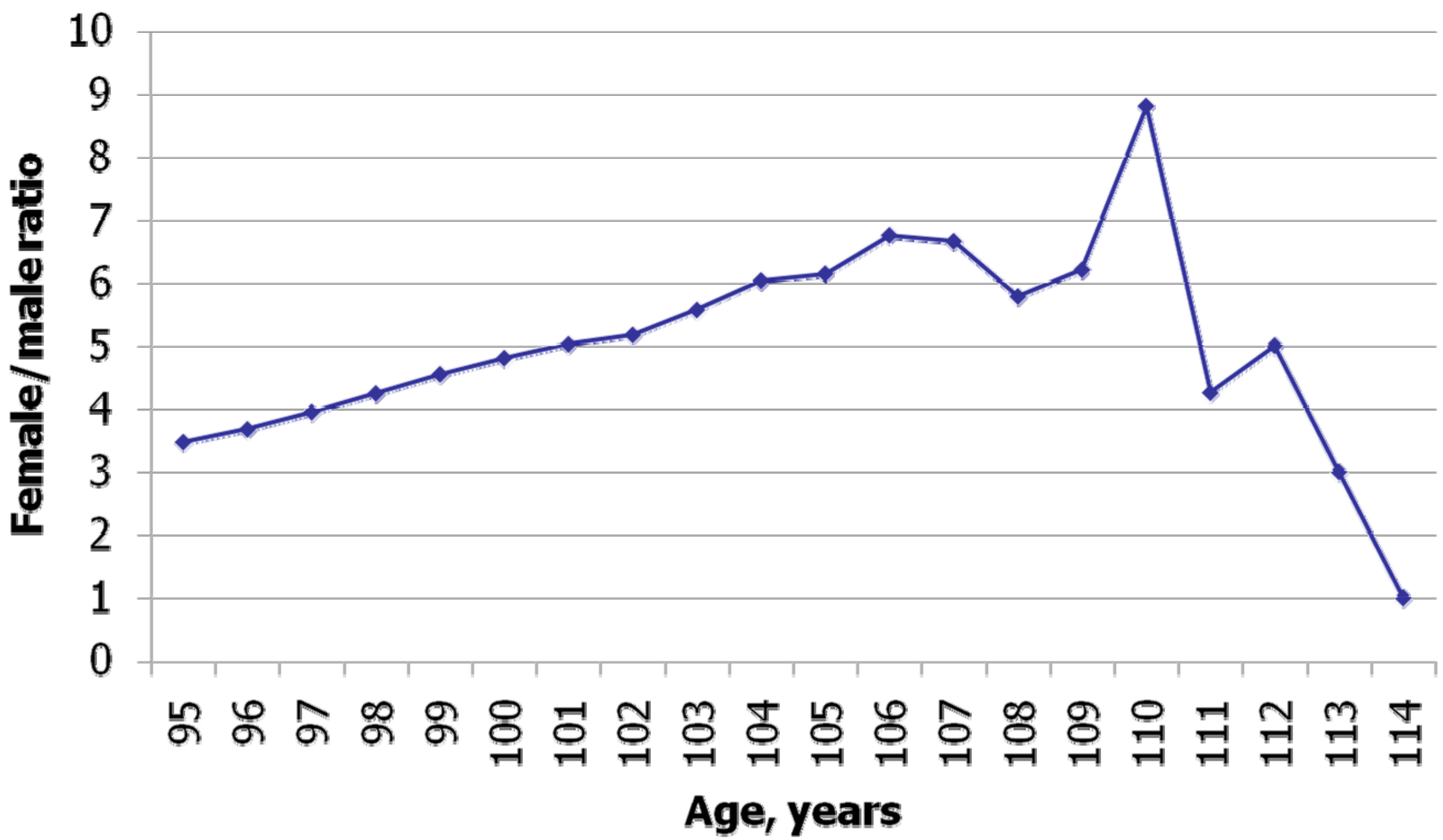
A study of age-specific mortality by gender

Women have lower mortality at all ages

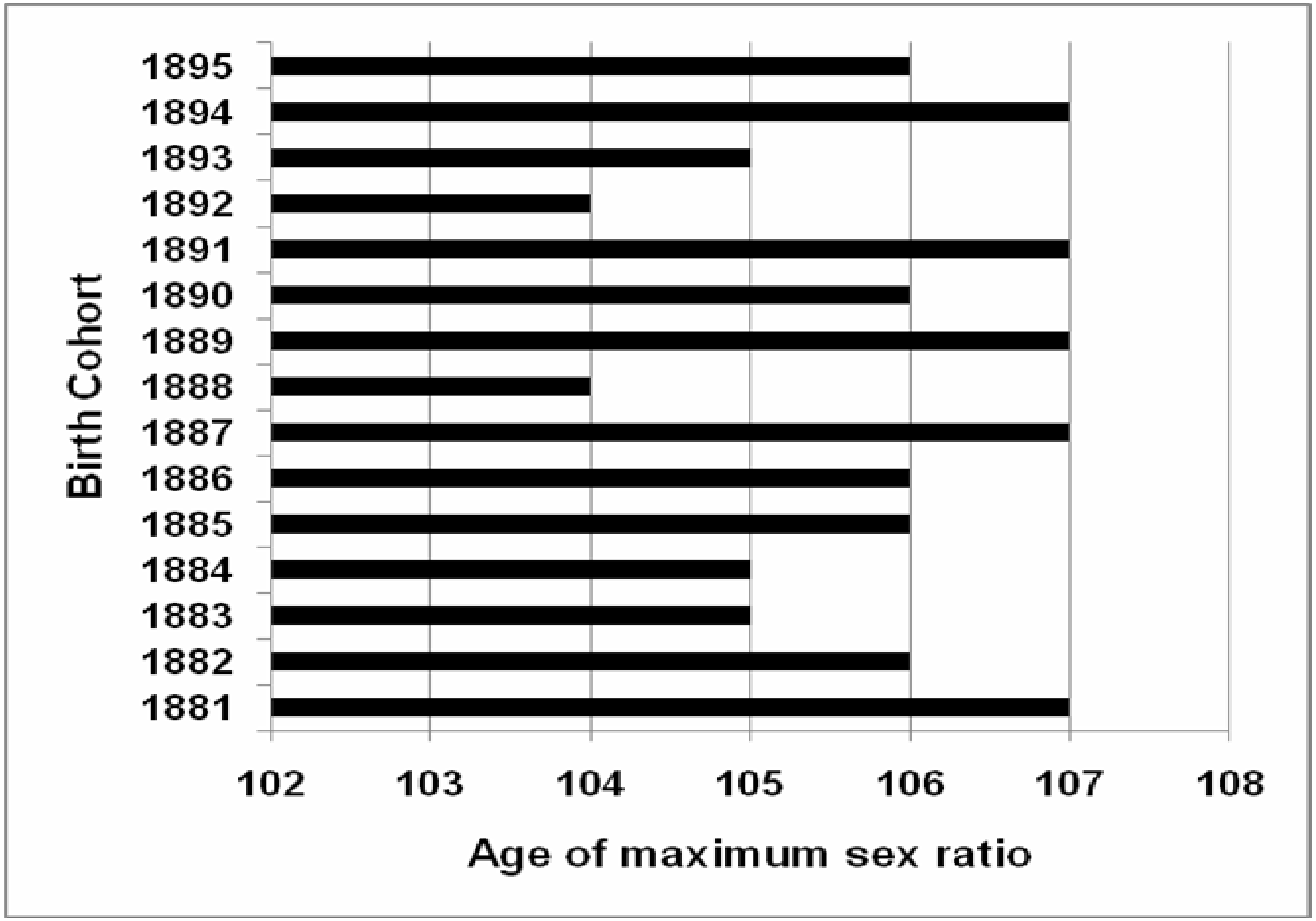


Hence number of females to number of males ratio should grow with age

Observed female to male ratio at advanced ages for combined 1887-1892 birth cohort



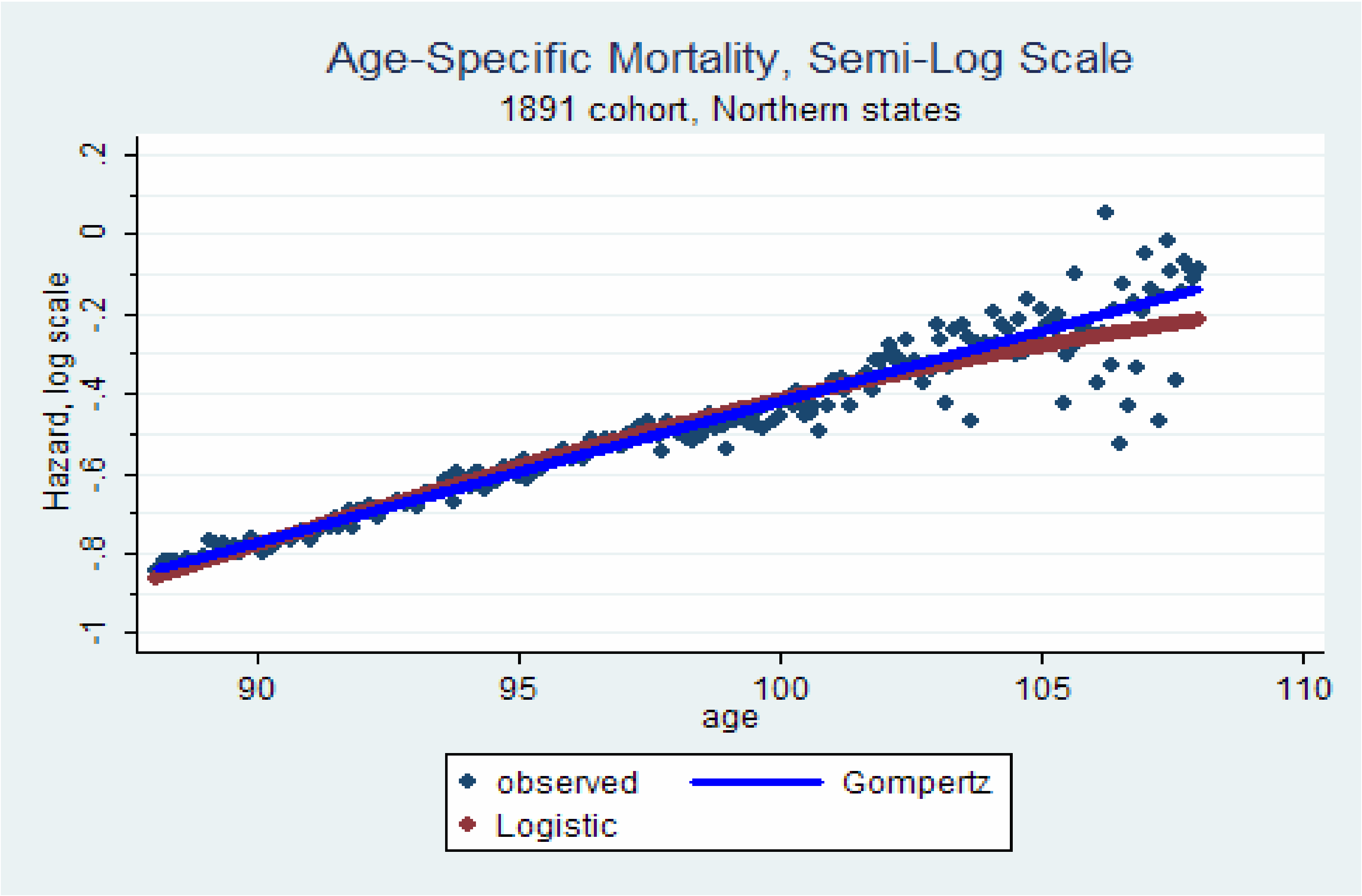
Age of maximum female to male ratio by birth cohort



Modeling mortality at advanced ages

- Data with reasonably good quality were used: Northern states and 88-106 years age interval
- Gompertz and logistic (Kannisto) models were compared
- Nonlinear regression model for parameter estimates (Stata 11)
- Model goodness-of-fit was estimated using AIC and BIC

Fitting mortality with logistic and Gompertz models



Bayesian information criterion (BIC) to compare logistic and Gompertz models, by birth cohort

Birth cohort	1886	1887	1888	1889	1890	1891	1892	1893	1894	1895
Cohort size at 88 years	111657	114469	128768	131778	135393	143138	152058	156189	160835	165294
Gompertz	-594776.2	-625303	-709620.7	-710871.1	-724731	-767138.3	-831555.3	-890022.6	-946219	-921650.3
logistic	-588049.5	-618721.4	-712575.5	-715356.6	-722939.6	-739727.6	-810951.8	-862135.9	-905787.1	-863246.6

Better fit (lower BIC) is highlighted in red

Conclusion: In 8 out of 10 cases Gompertz model demonstrates better fit than logistic model for age interval 88-106 years

Comparison to mortality data from the Actuarial Study No.116

- 1900 birth cohort in Actuarial Study is the earliest available cohort for comparison
- Used 1894 birth cohort for comparison because later birth cohorts are less likely to be extinct
- Historical studies suggest that adult life expectancy in the U.S. did not experience substantial changes during the period 1890-1900 (Haines, 1998)

In Actuarial Study death rates at ages 95 and older were extrapolated

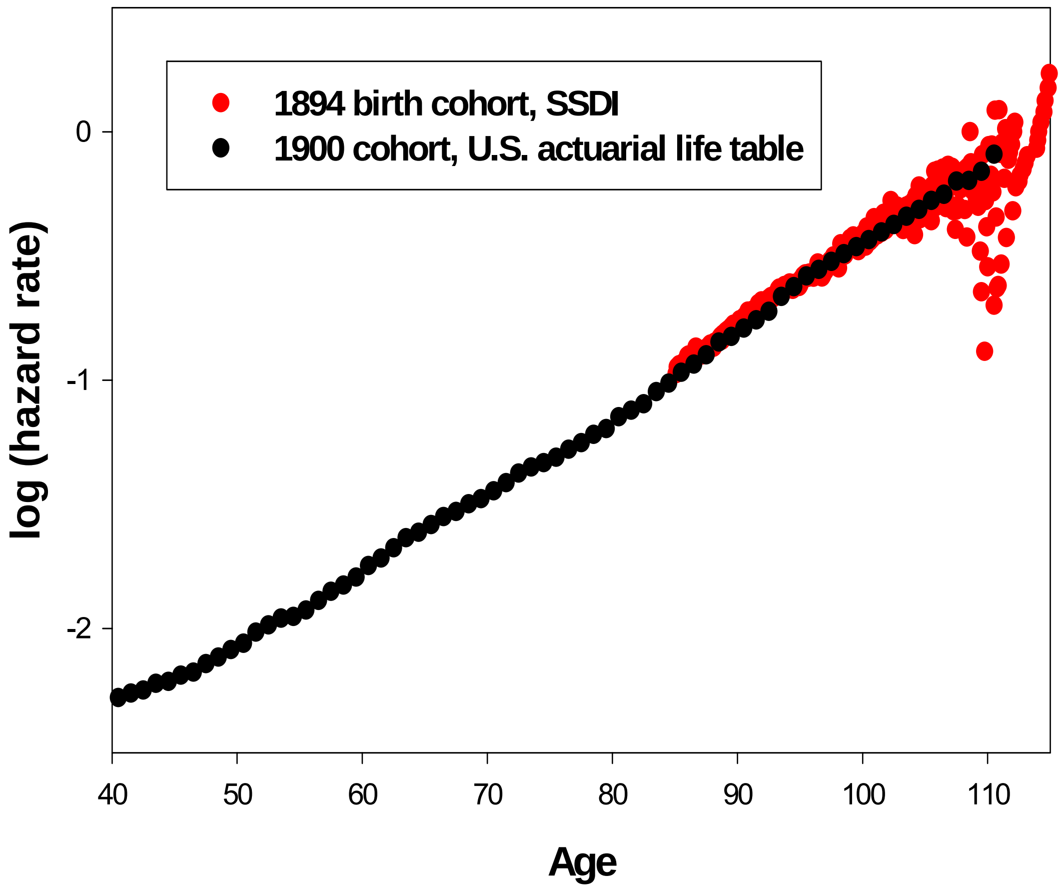
$$q_x = q_{x-1} \bullet \left(\frac{q_{94}}{q_{93}} \bullet \frac{99-x}{5} + 1.05 \bullet \frac{x-94}{5} \right) \qquad x = 95, 96, 97, 98, 99$$

$$q_x = 1.05 \bullet q_{x-1} \qquad x = 100, 101, 102, \ldots$$

- **We used conversion formula (Gehan, 1969) to calculate hazard rate from life table values of probability of death:**

$$\mu_x = -\ln(1-q_x)$$

Mortality at advanced ages:
Actuarial 1900 cohort life table and SSDI 1894 birth cohort

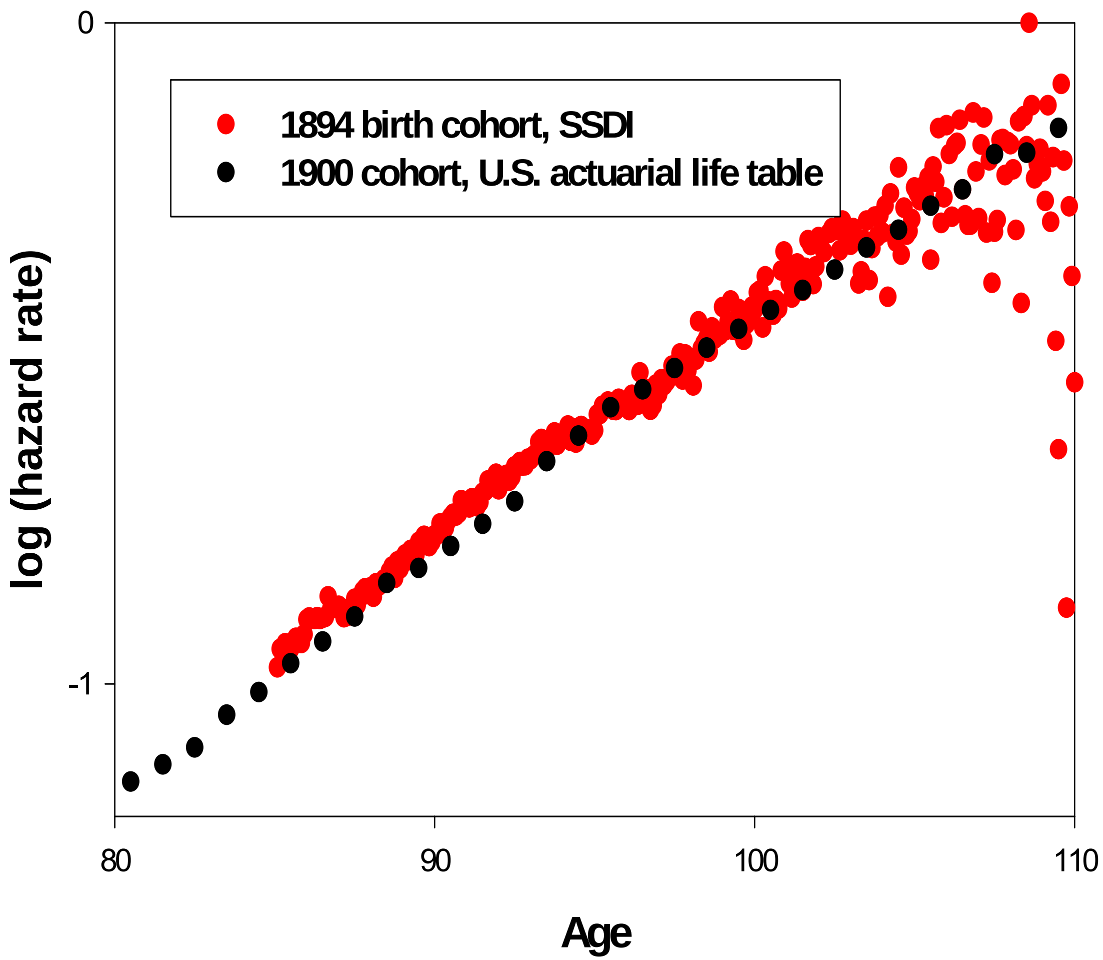


Source for actuarial life table:
Bell, F.C., Miller, M.L. Life Tables for the United States Social Security Area 1900-2100
Actuarial Study No. 116

Hazard rates for 1900 cohort are estimated by Sacher formula

Estimating Gompertz slope parameter

Actuarial cohort life table and SSDI 1894 cohort



1900 cohort, age interval 40-104
alpha (95% CI):
0.0785 (0.0772,0.0797)

1894 cohort, age interval 88-106
alpha (95% CI):
0.0786 (0.0786,0.0787)

Hypothesis about two-stage Gompertz model is not supported by real data

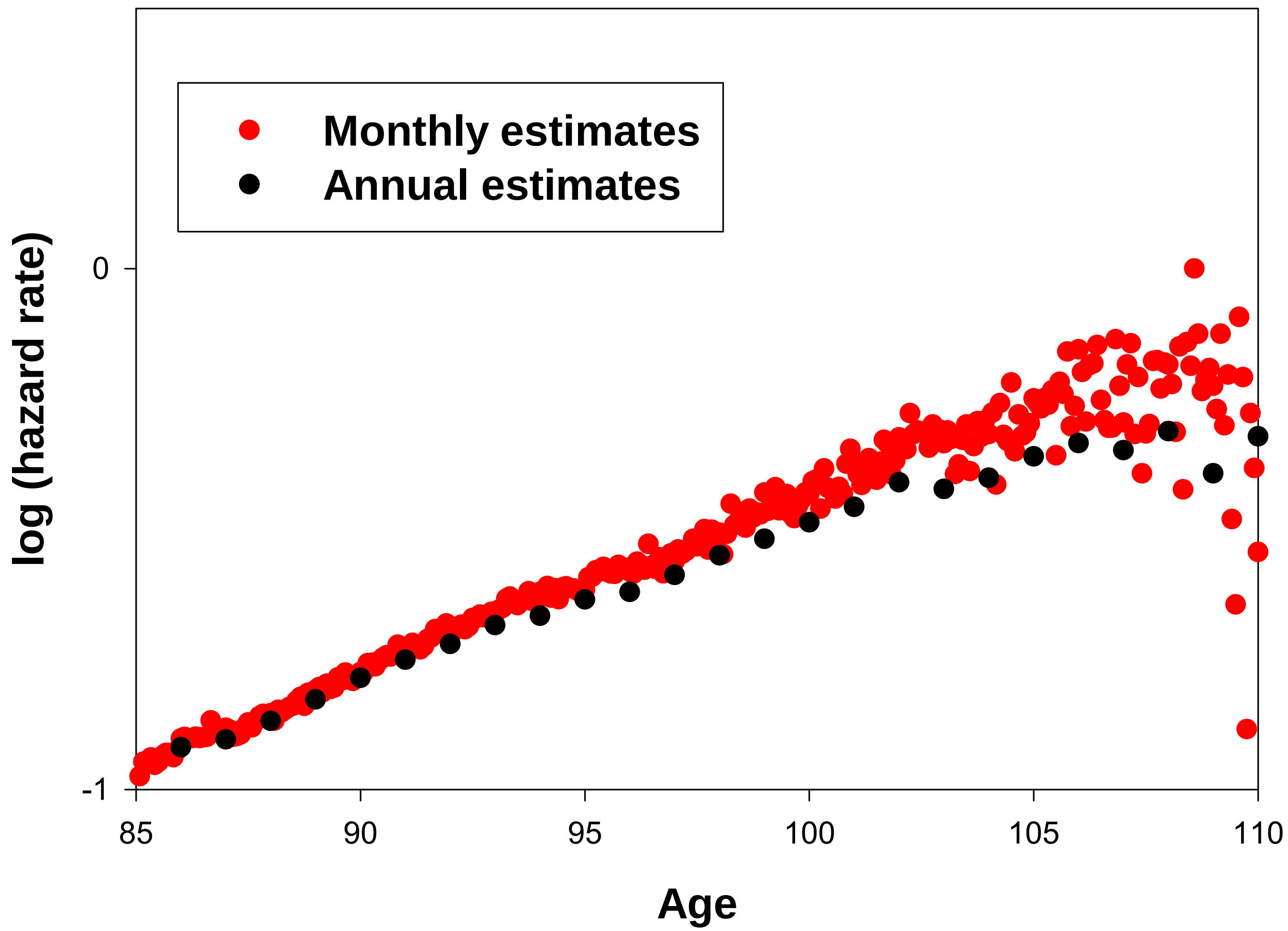
What estimate of hazard rate is the most accurate?

Simulation study comparing several existing estimates:

- Nelson-Aalen estimate available in Stata
- Sacher estimate (Sacher, 1956)
- Gehan (pseudo-Sacher) estimate (Gehan, 1969)
- Actuarial estimate (Kimball, 1960)

Mortality of 1894 birth cohort

Monthly and Annual Estimates of Hazard Rates using Nelson-Aalen formula (Stata)



Sacher formula for hazard rate estimation (Sacher, 1956; 1966)

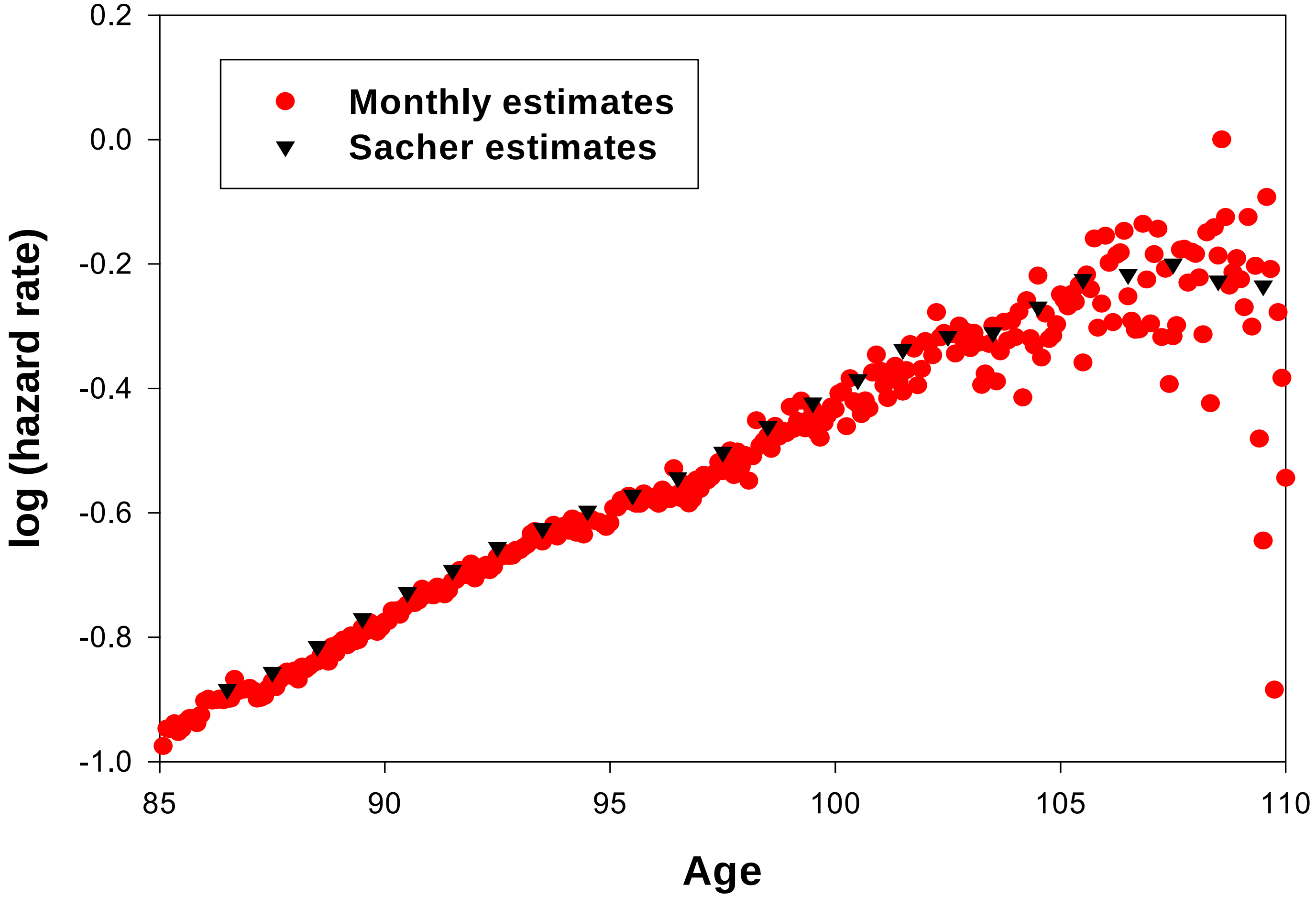
$$\mu_x = \frac{1}{\Delta x}\left(\ln l_{x-\frac{\Delta x}{2}} - \ln l_{x+\frac{\Delta x}{2}}\right) = \frac{1}{2\Delta x}\ln\frac{l_{x-\Delta x}}{l_{x+\Delta x}}$$

Hazard rate

l_x - survivor function at age x; Δx – age interval

Simplified version suggested by Gehan (1977):

Mortality of 1894 birth cohort
Sacher formula for annual estimates of hazard rates

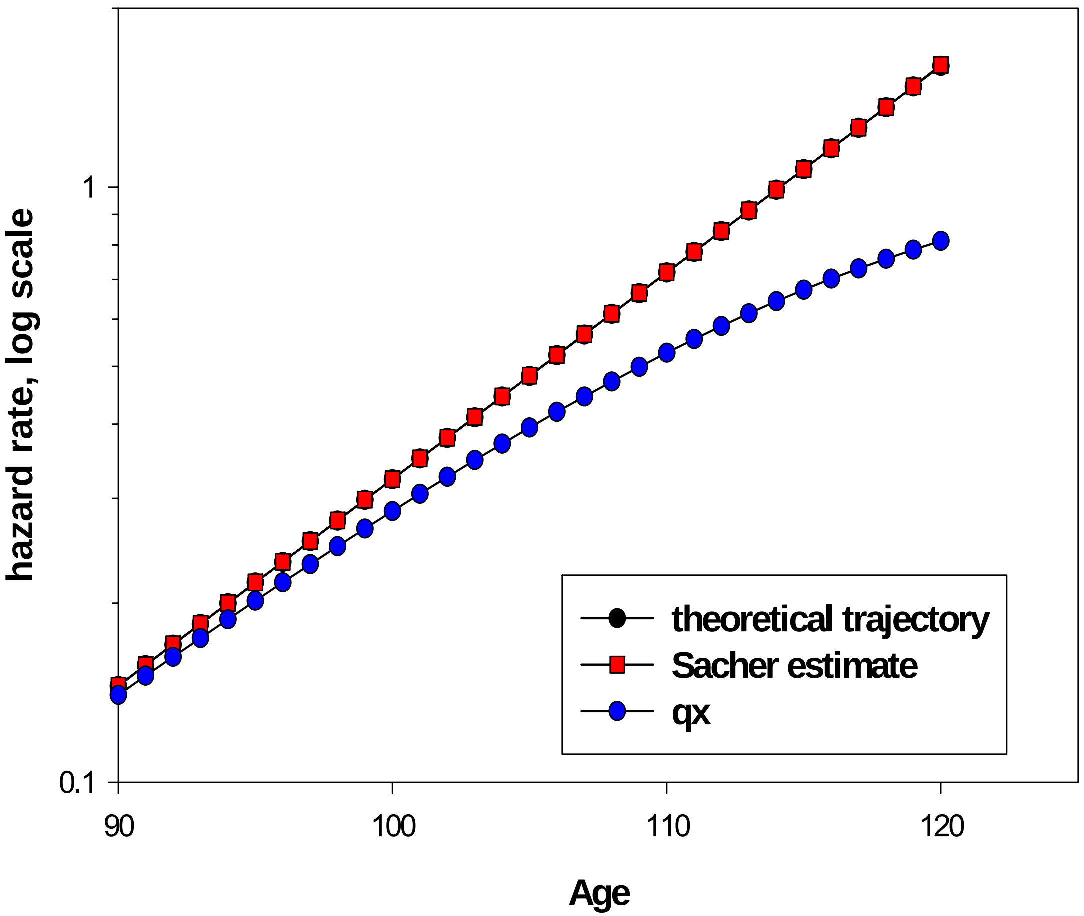


Simulation study to identify the most accurate mortality indicator

- Simulate yearly l_x numbers assuming Gompertz function for hazard rate in the entire age interval and initial cohort size equal to 10^{11} individuals
- Gompertz parameters are typical for the U.S. birth cohorts: slope coefficient (alpha) = 0.08 year^{-1}; R_0= 0.0001 year^{-1}
- Focus on ages beyond 90 years
- Accuracy of various hazard rate estimates (Sacher, Gehan, and actuarial estimates) and probability of death is compared at ages 100-110

Simulation study of Gompertz mortality

Compare Sacher hazard rate estimate and probability of death in a yearly age interval



Sacher estimates practically coincide with theoretical mortality trajectory

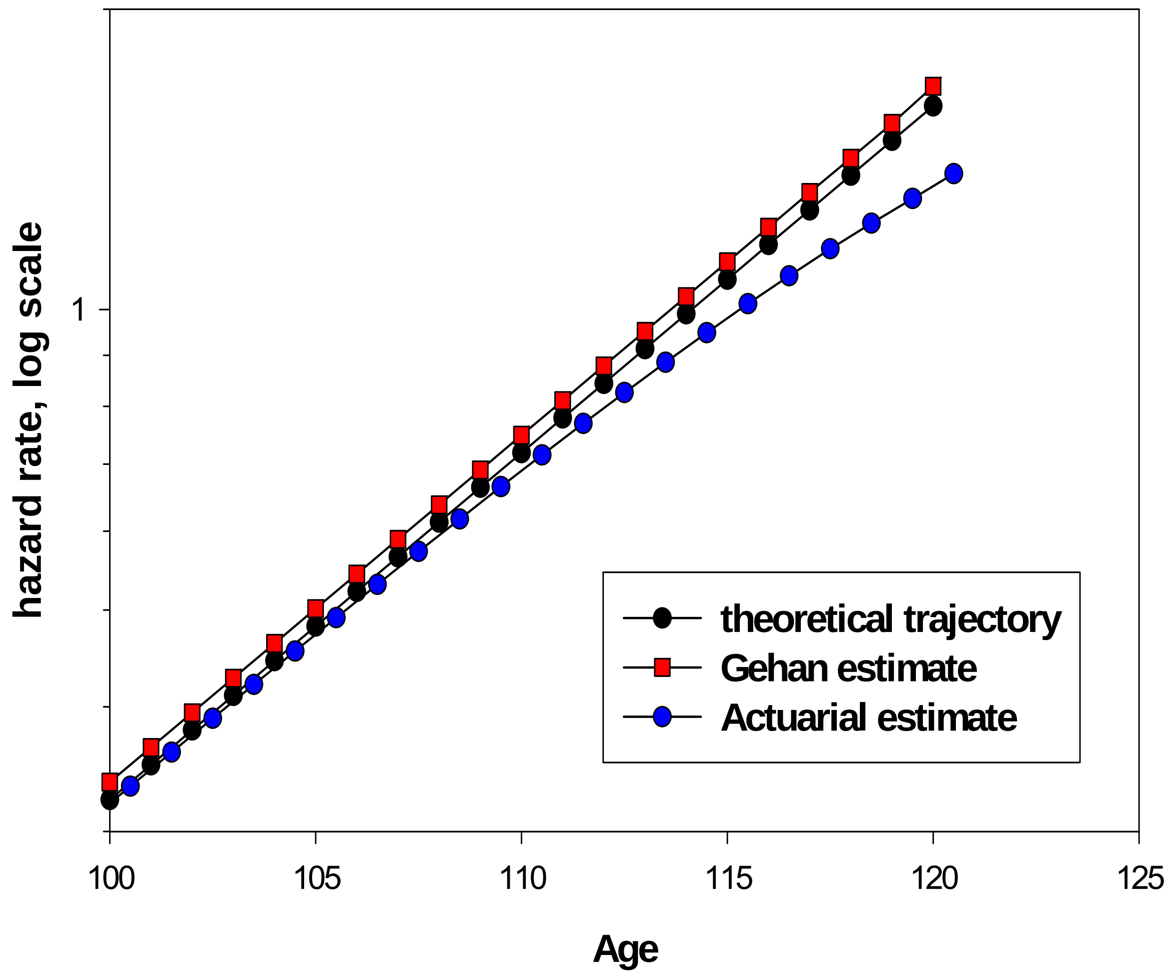
$$\mu_x = \frac{1}{2\Delta x} ln \frac{l_{x-\Delta x}}{l_{x+\Delta x}}$$

Probability of death values strongly undeestimate mortality after age 100

$$q_x = \frac{d_x}{l_x}$$

Simulation study of Gompertz mortality

Compare Gehan and actuarial hazard rate estimates



Gehan estimates slightly overestimate hazard rate because of its half-year shift to earlier ages

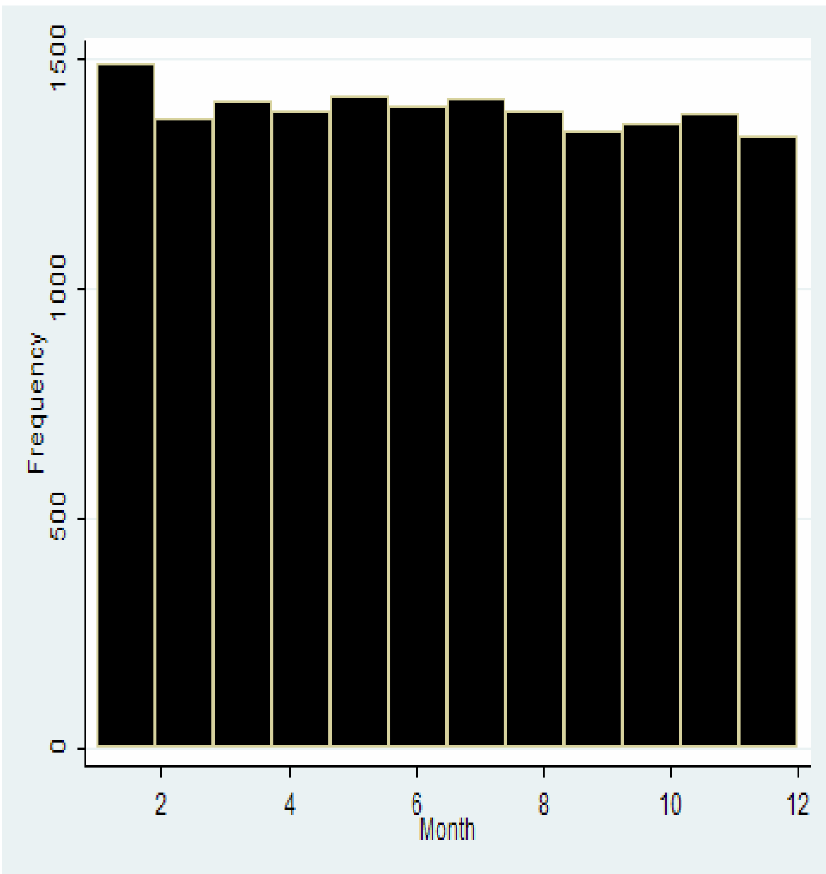
$$\mu_x = -\ln(1 - q_x)$$

Actuarial estimates undeestimate mortality after age 100

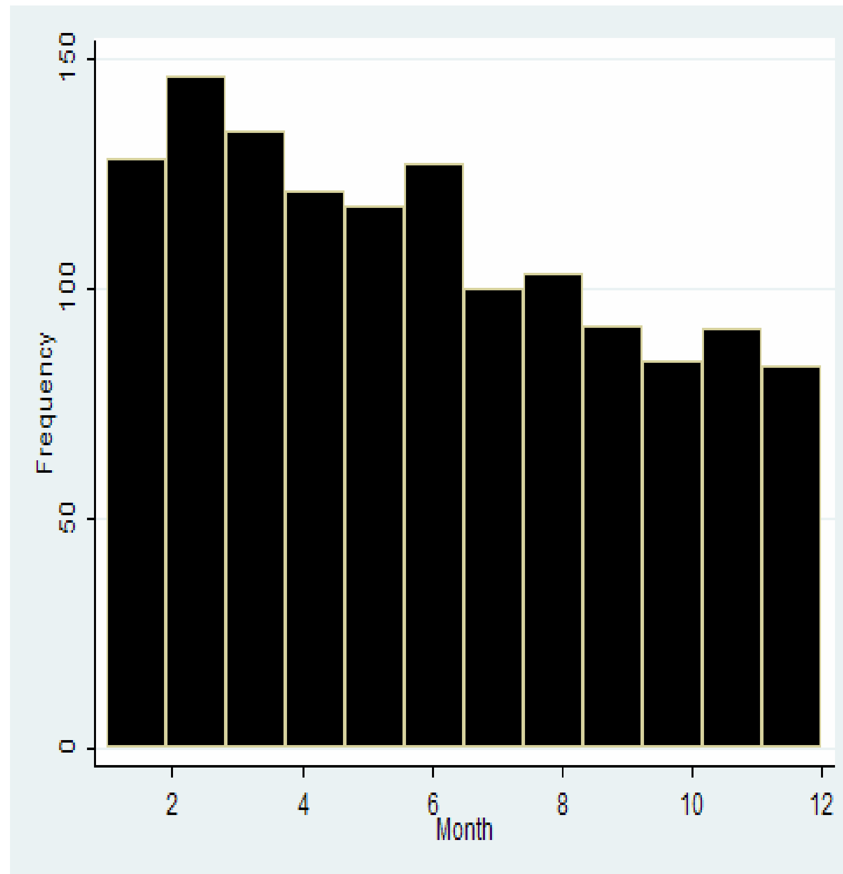
$$\mu_{x+\frac{\Delta x}{2}} = \frac{2}{\Delta x}\,\frac{l_x - l_{x+\Delta x}}{l_x + l_{x+\Delta x}}$$

Deaths at extreme ages are not distributed uniformly over one-year interval

90-year olds



102-year olds



1891 birth cohort from the Social Security Death Index

Accuracy of hazard rate estimates

Relative difference between theoretical and observed values, %		
Estimate	**100 years**	**110 years**
Probability of death	11.6%, understate	26.7%, understate
Sacher estimate	0.1%, overstate	0.1%, overstate
Gehan estimate	4.1%, overstate	4.1%, overstate
Actuarial estimate	1.0%, understate	4.5%, understate

Conclusions

- Deceleration of mortality in later life is more expressed for data with lower quality. Quality of age reporting in DMF becomes poor beyond the age of 107 years
- Below age 107 years and for data of reasonably good quality the Gompertz model fits mortality better than the logistic model (no mortality deceleration)
- Sacher estimate of hazard rate turns out to be the most accurate and most useful estimate to study mortality at advanced ages

Acknowledgments

This study was made possible thanks to:

generous support from the

- National Institute on Aging
- Stimulating working environment at the Center on Aging, NORC/University of Chicago

For More Information and Updates Please Visit Our Scientific and Educational Website on Human Longevity:

- http://longevity-science.org

And Please Post Your Comments at our Scientific Discussion Blog:

- http://longevity-science.blogspot.com/