Biodemographic Study of Mortality Trajectories at Advanced Old Ages

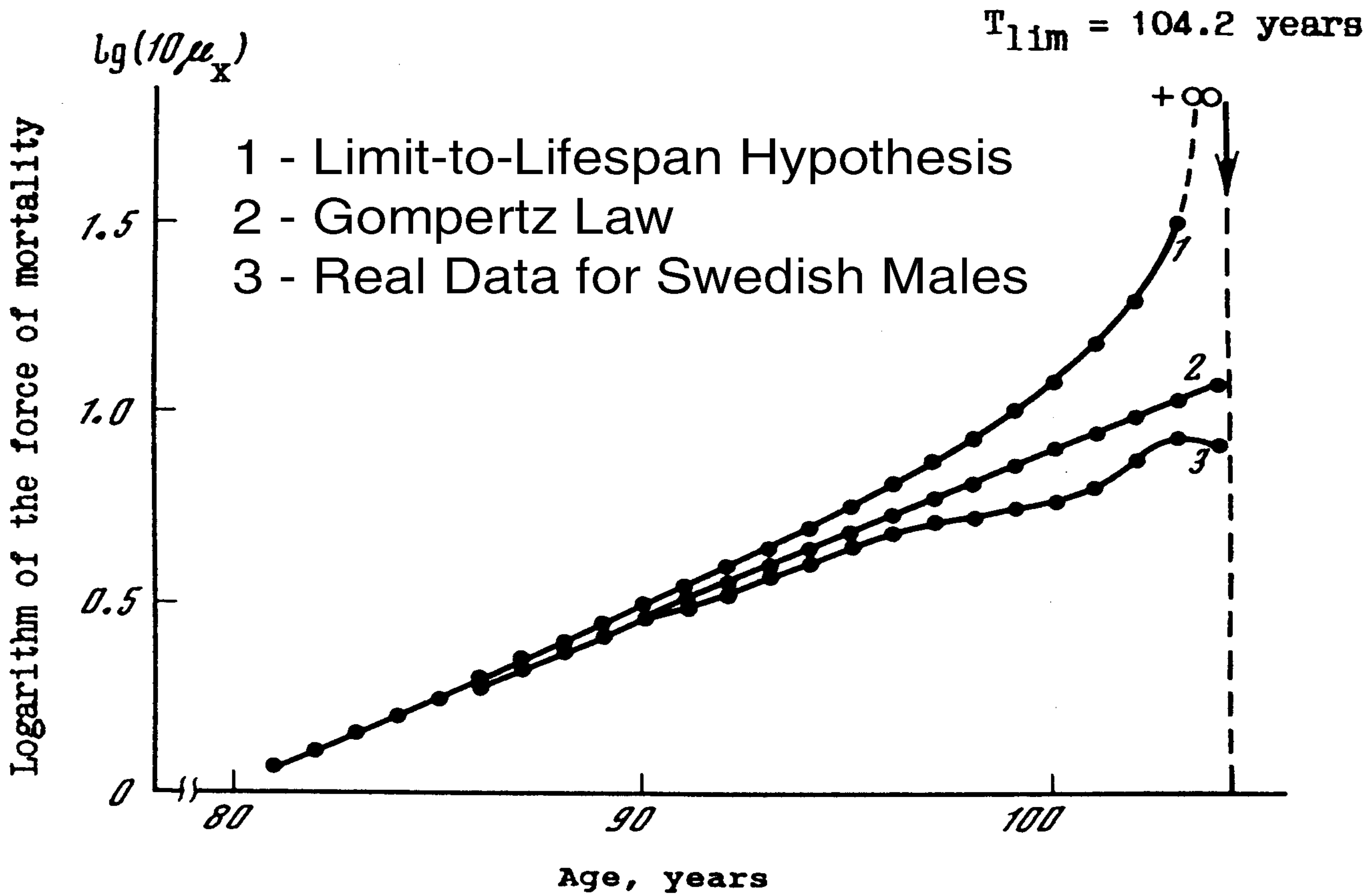
Natalia S. Gavrilova, Ph.D.
Leonid A. Gavrilov, Ph.D.

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

The growing number of persons living beyond age 80 underscores the need for accurate measurement of mortality at advanced ages.

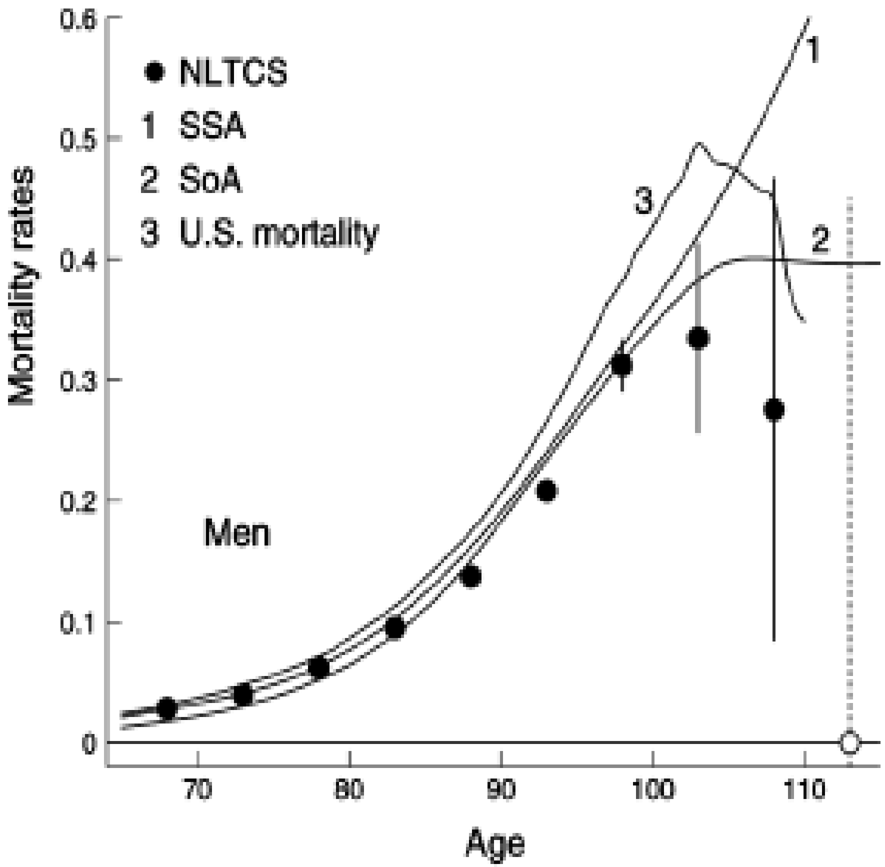
Earlier studies suggested that the exponential growth of mortality with age (Gompertz law) is followed by a period of deceleration, with slower rates of mortality increase.

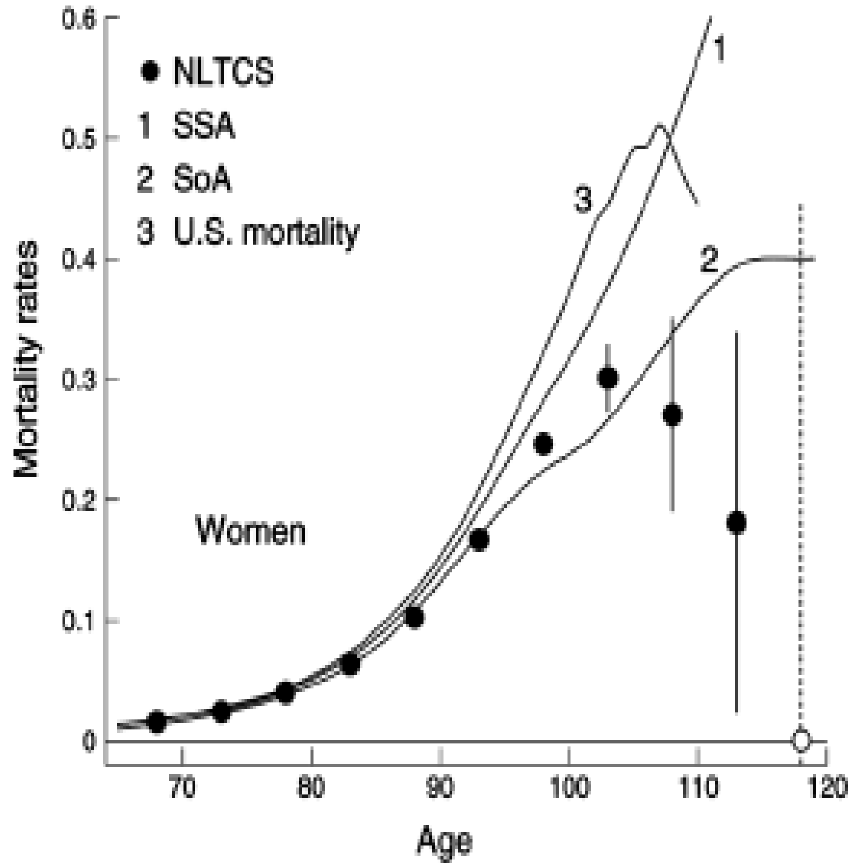
Mortality at Advanced Ages – more than 20 years ago



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

Mortality at Advanced Ages, Recent Study





Source: **Manton et al. (2008). Human Mortality at Extreme Ages: Data from the NLTCS and Linked Medicare Records.** *Math.Pop.Studies*

Problems with 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

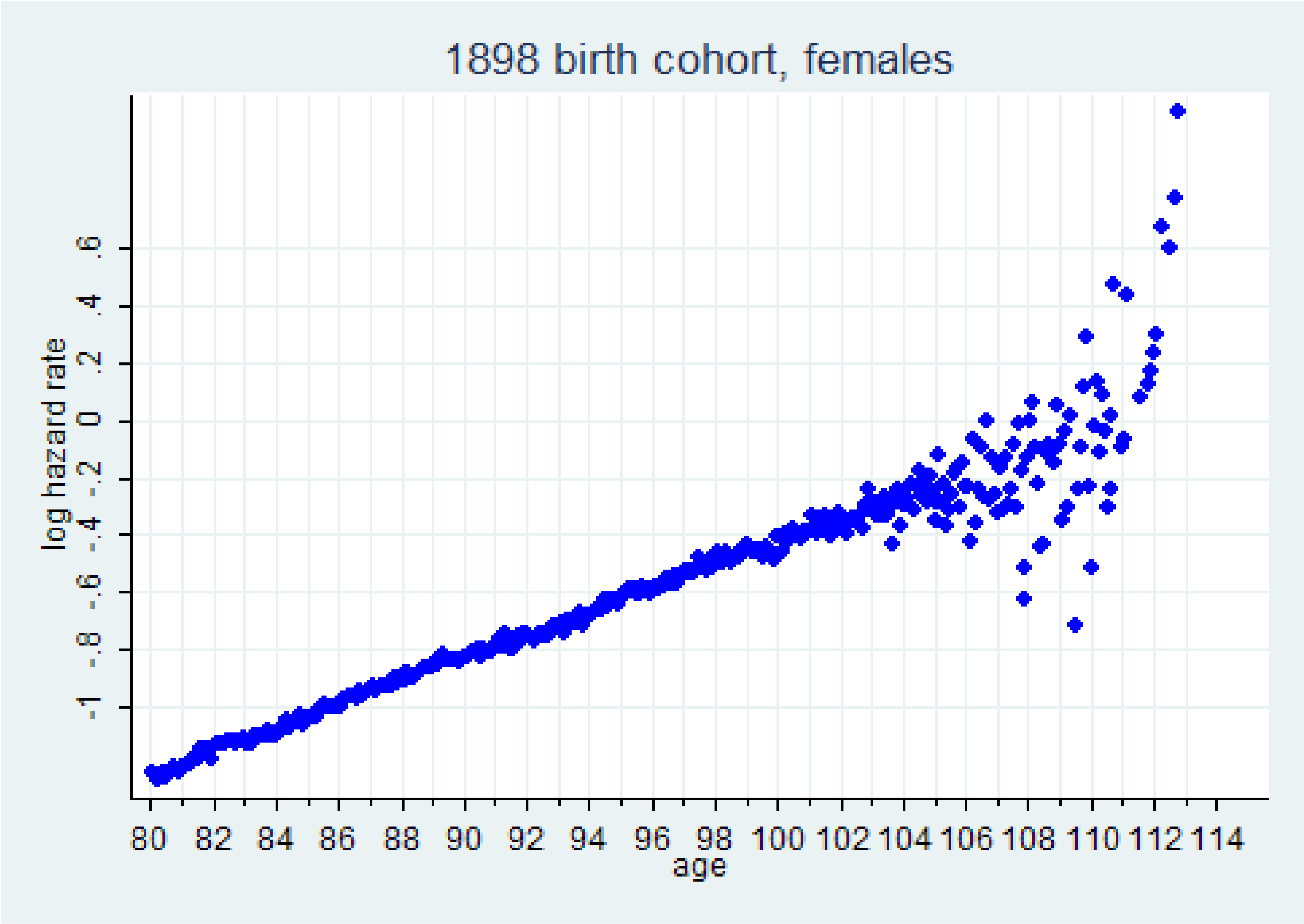
Social Security Administration's Death Master File (SSA's DMF) Helps to Alleviate the First Two Problems

- 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's DMF ?

- As a result of a court case under the Freedom of Information Act, SSA is required to release its death information to the public. SSA's DMF contains the complete and official SSA database extract, as well as updates to the full file of persons reported to SSA as being deceased.
- SSA DMF is no longer a publicly available data resource (now is available from Ancestry.com for fee)
- We used DMF full file obtained from the National Technical Information Service (NTIS). Last deaths occurred in September 2011.

SSA DMF birth cohort mortality

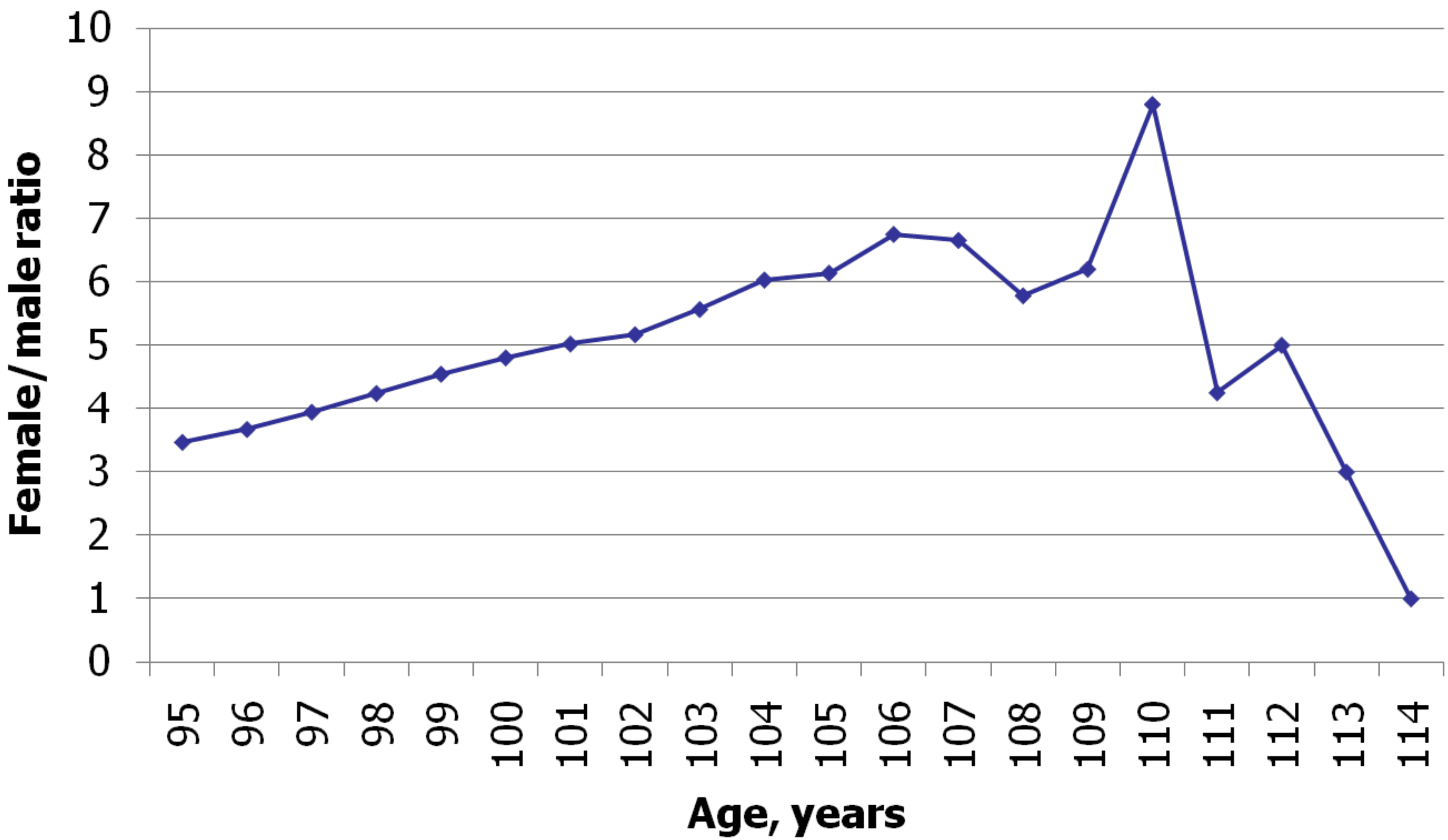


Nelson-Aalen monthly estimates of hazard rates using Stata 11

Conclusions from our earlier study of SSA DMF

- **Mortality deceleration at advanced ages among DMF cohorts is more expressed for data of lower quality**
- **Mortality data beyond ages 106-107 years have unacceptably poor quality (as shown using female-to-male ratio test). The study by other authors also showed that beyond age 110 years the age of individuals in DMF cohorts can be validated for less than 30% cases (Young et al., 2010)**
- **Source: Gavrilov, Gavrilova, *North American Actuarial Journal*, 2011, 15(3):432-447**

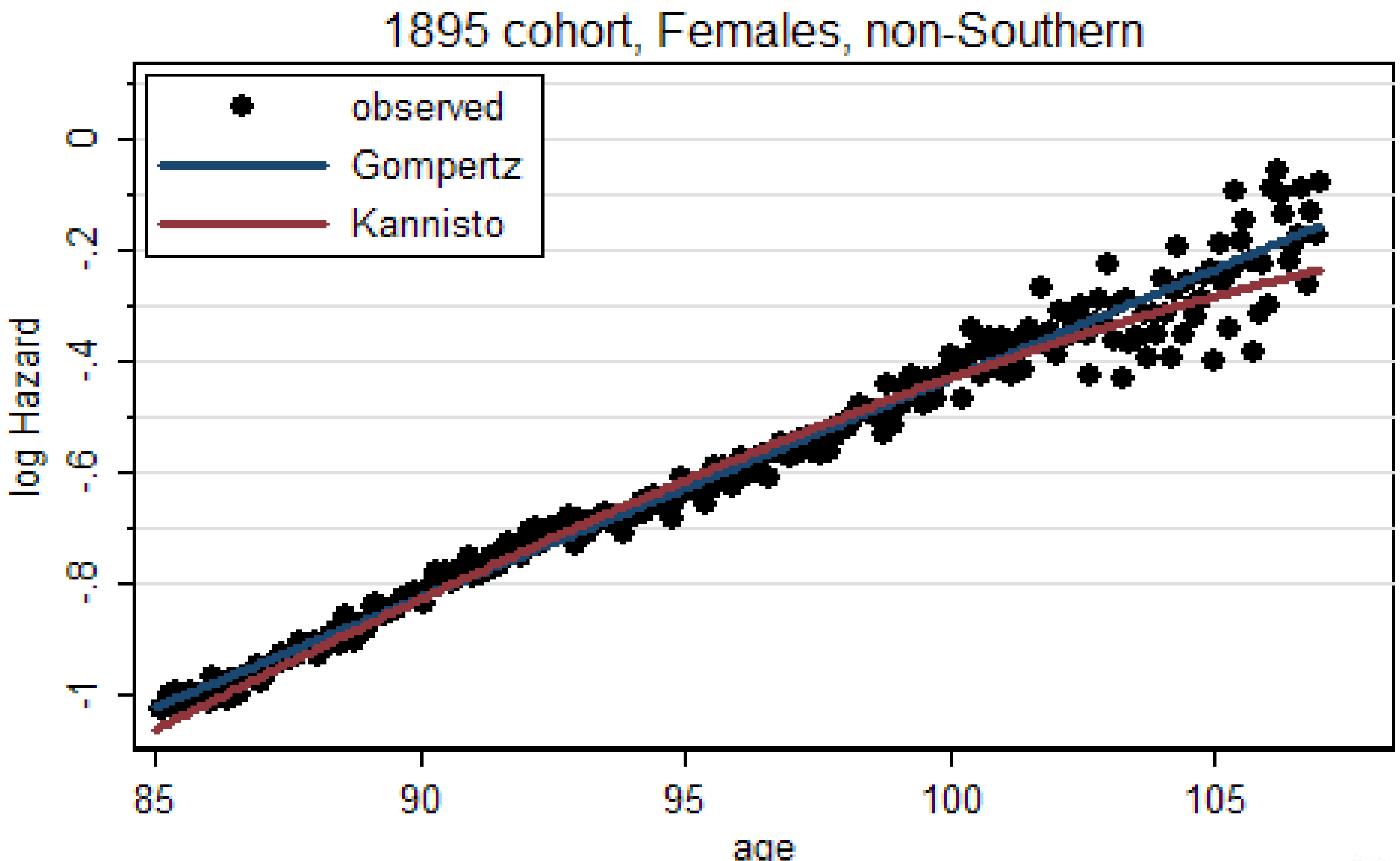
Observed female to male ratio at advanced ages for combined 1887-1892 birth cohort



Selection of competing mortality models using DMF data

- Data with reasonably good quality were used: non-Southern states and 85-106 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

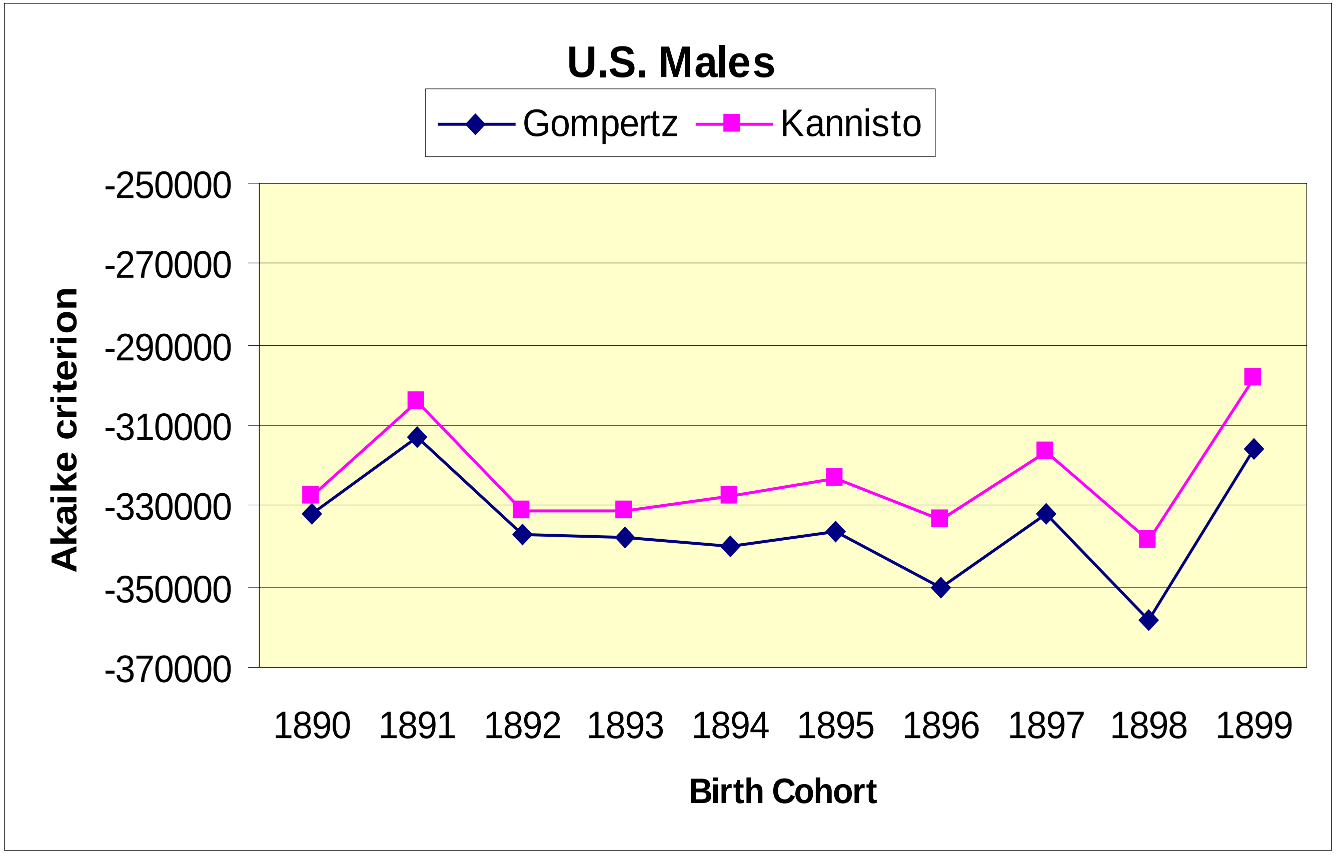
Fitting mortality with Kannisto and Gompertz models



Gompertz model $\mu_x = ae^{bx}$.

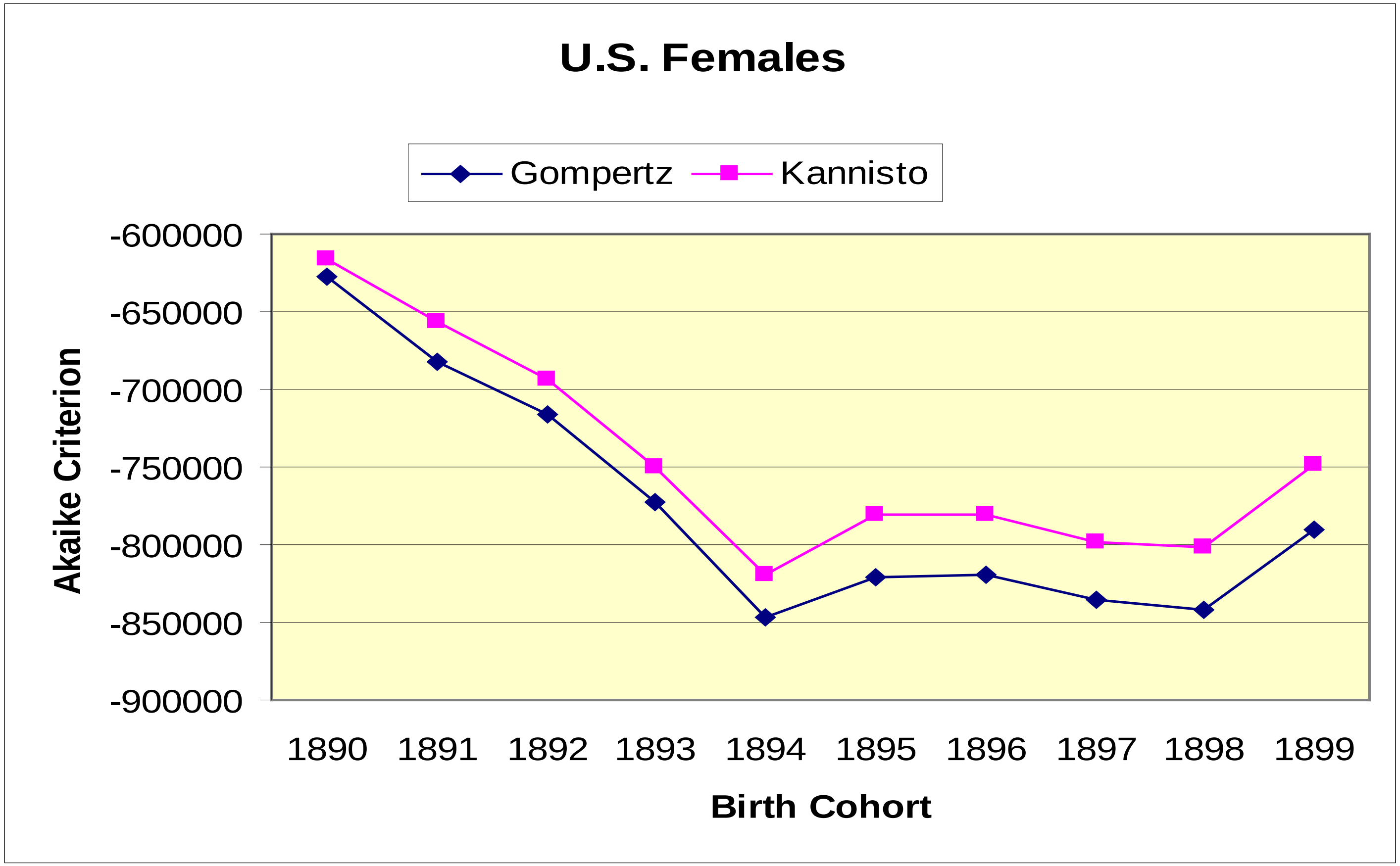
Kannisto model $\mu_x = \frac{ae^{bx}}{1+ae^{bx}}$

Akaike information criterion (AIC) to compare Kannisto and Gompertz models, **men**, by birth cohort (non-Southern states)



Conclusion: In all ten cases Gompertz model demonstrates better fit than Kannisto model for men in age interval 85-106 years

Akaike information criterion (AIC) to compare Kannisto and Gompertz models, **women**, by birth cohort (non-Southern states)



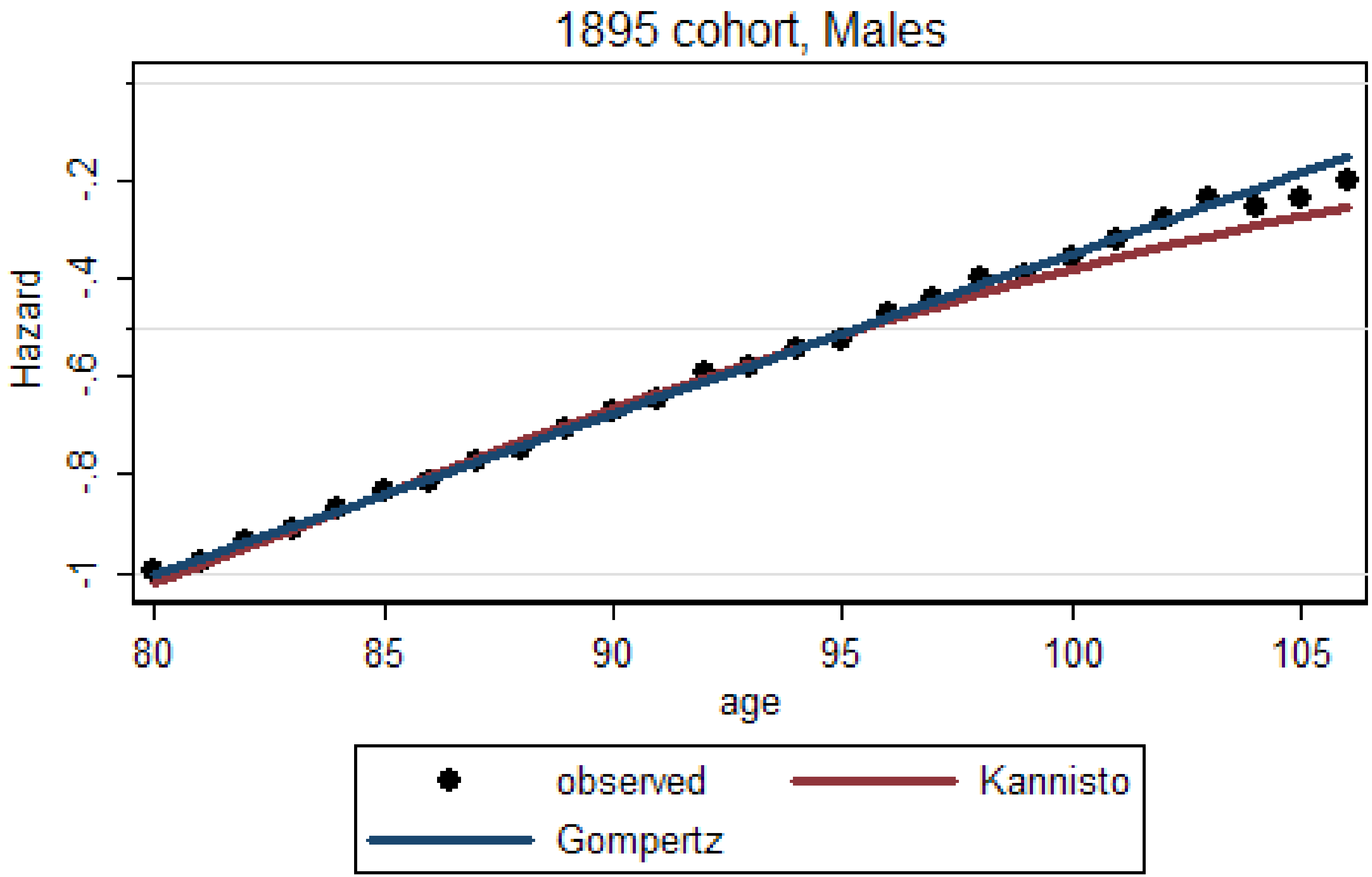
Conclusion: In all ten cases Gompertz model demonstrates better fit than Kannisto model for women in age interval 85-106 years

The second studied dataset: U.S. cohort death rates taken from the Human Mortality Database

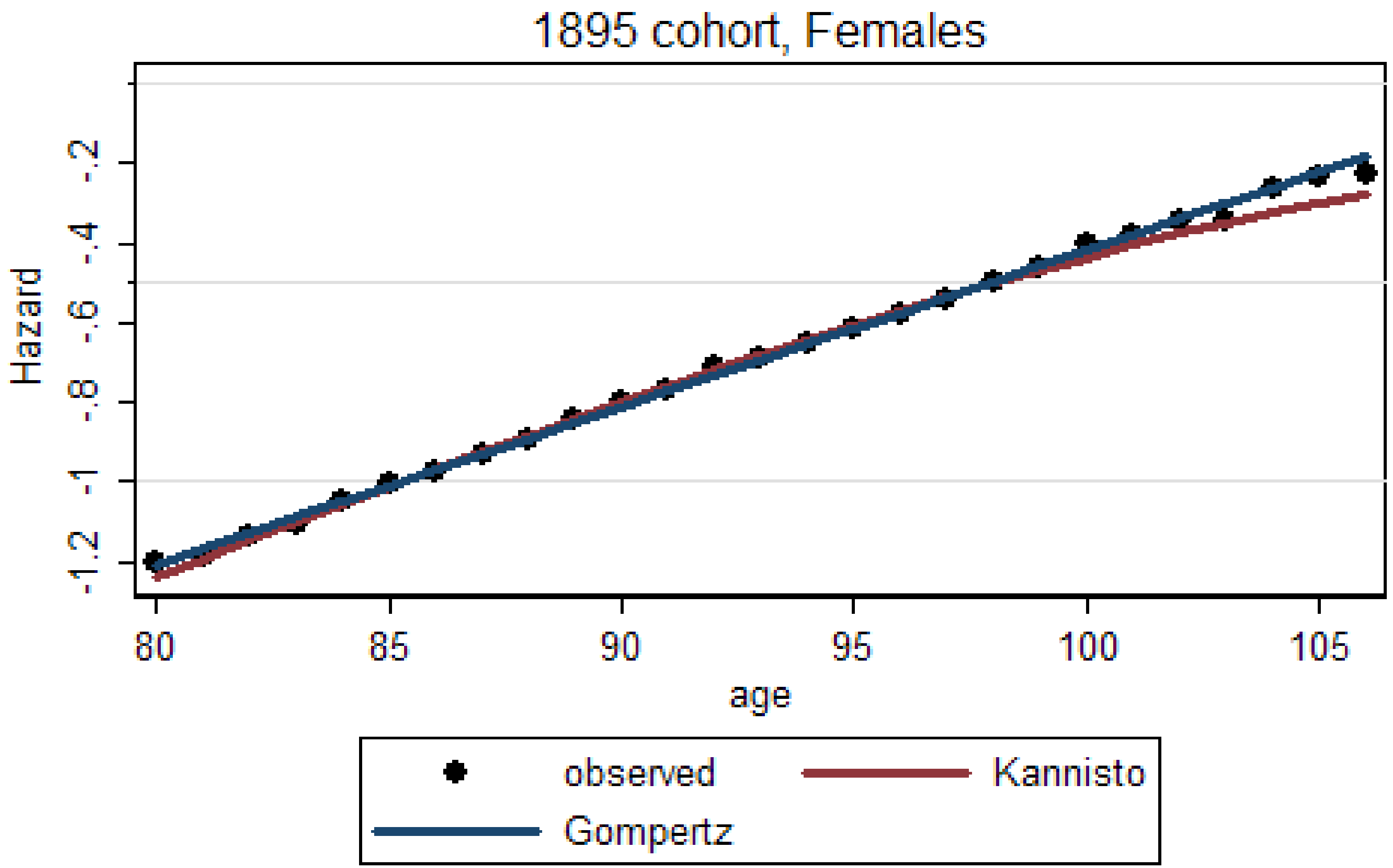
Selection of competing mortality models using HMD data

- Data with reasonably good quality were used: 80-106 years age interval
- Gompertz and logistic (Kannisto) models were compared
- Nonlinear weighted regression model for parameter estimates (Stata 11)
- Age-specific exposure values were used as weights (Muller at al., Biometrika, 1997)
- Model goodness-of-fit was estimated using AIC and BIC

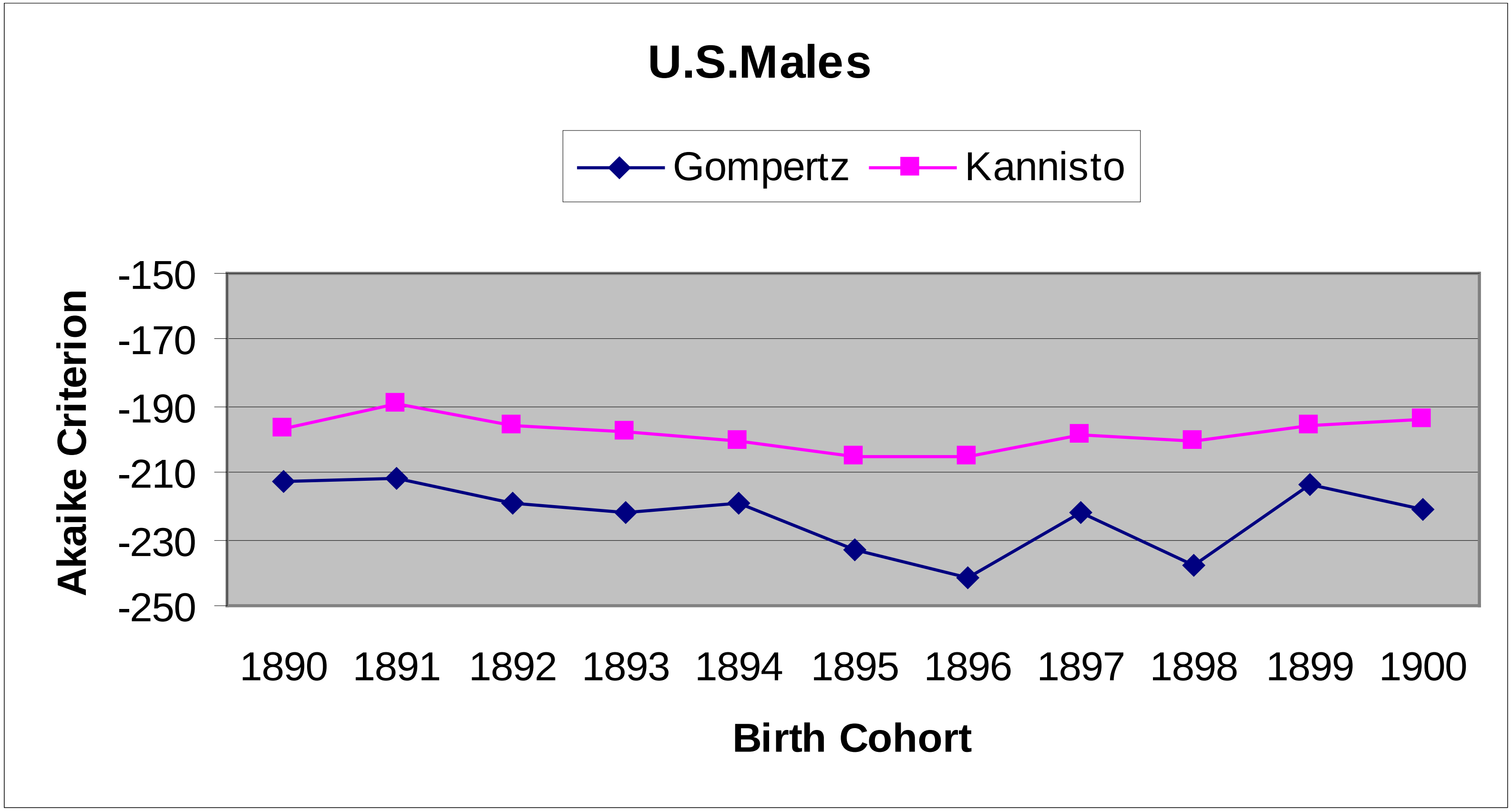
Fitting mortality with Kannisto and Gompertz models, HMD U.S. data



Fitting mortality with Kannisto and Gompertz models, HMD U.S. data

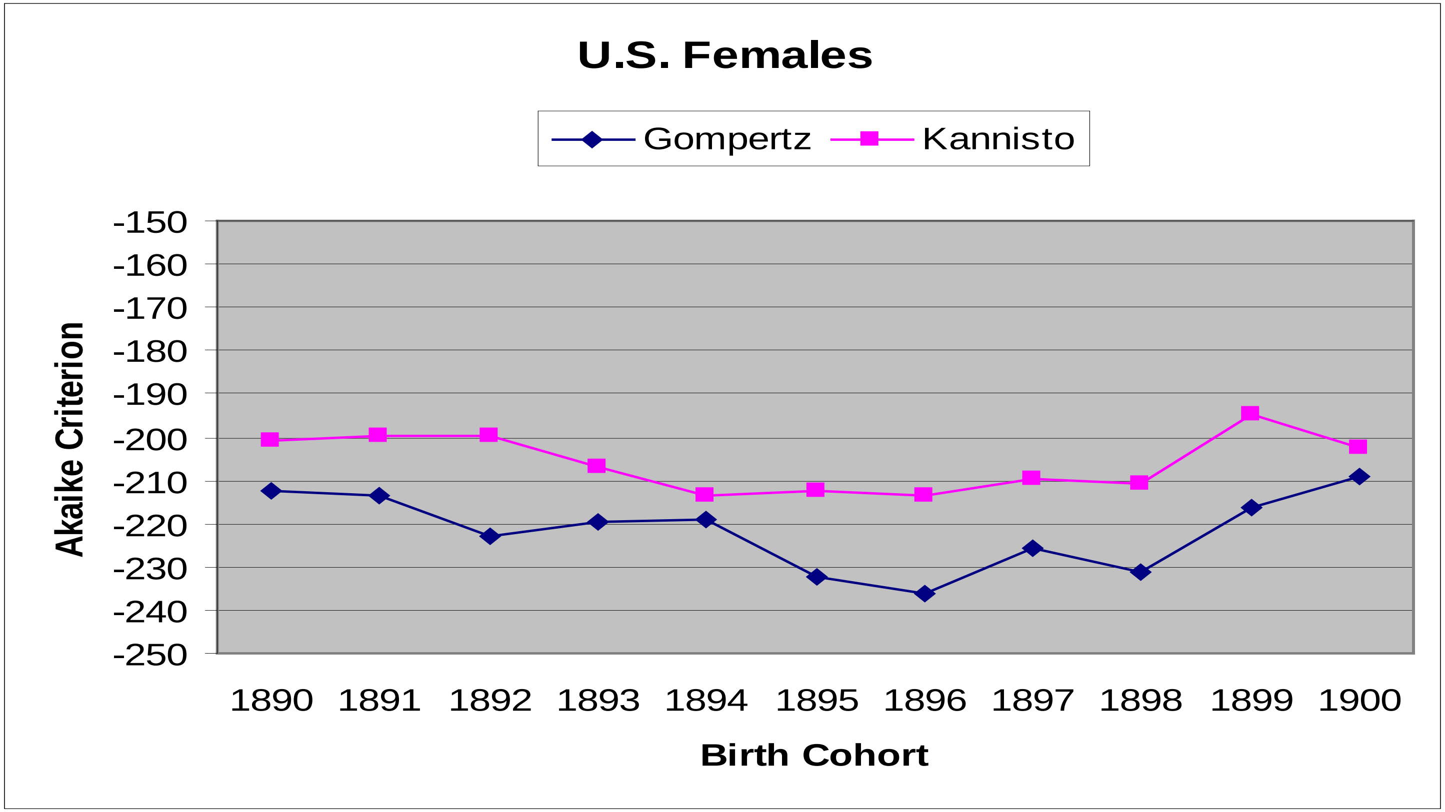


Akaike information criterion (AIC) to compare Kannisto and Gompertz models, **men**, by birth cohort (HMD U.S. data)



Conclusion: In all ten cases Gompertz model demonstrates better fit than Kannisto model for men in age interval 80-106 years

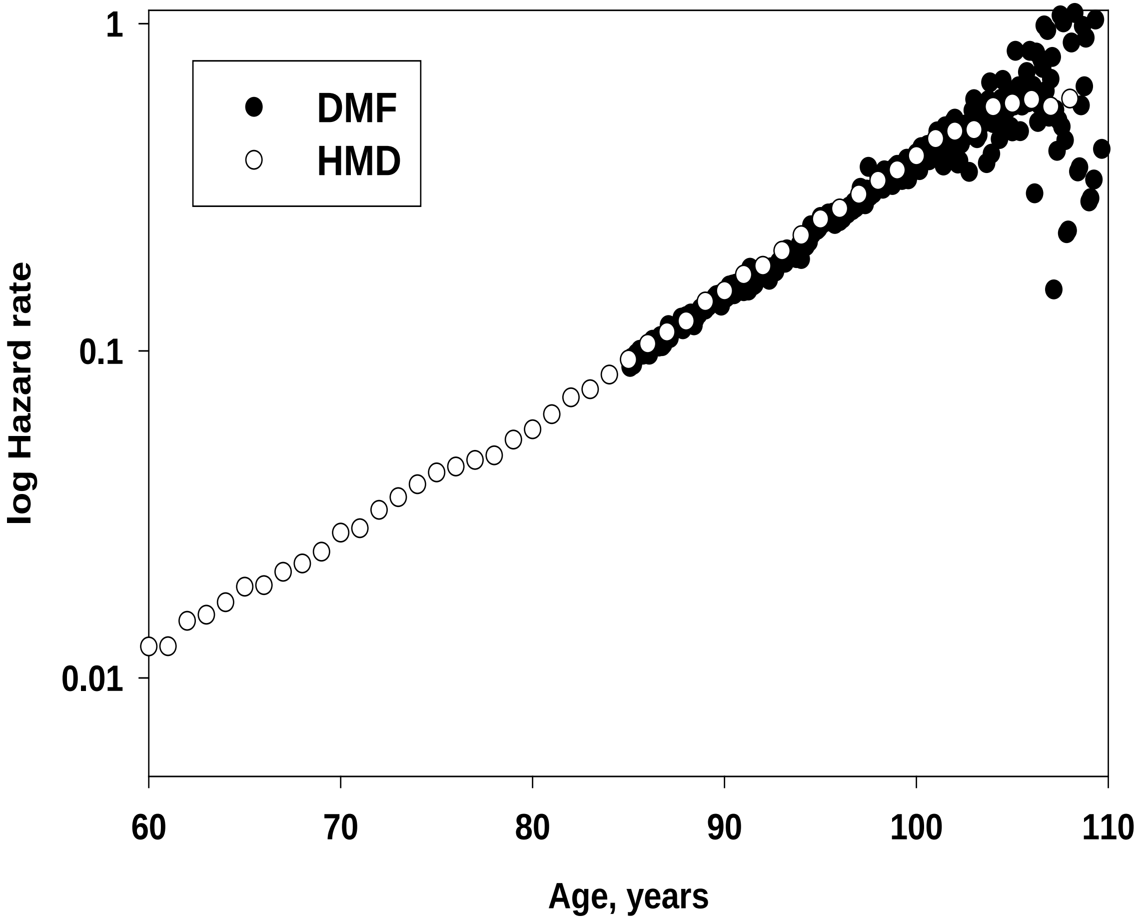
Akaike information criterion (AIC) to compare Kannisto and Gompertz models, **women**, by birth cohort (HMD U.S. data)



Conclusion: In all ten cases Gompertz model demonstrates better fit than Kannisto model for women in age interval 80-106 years

Compare DMF and HMD data

Females, 1898 birth cohort



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

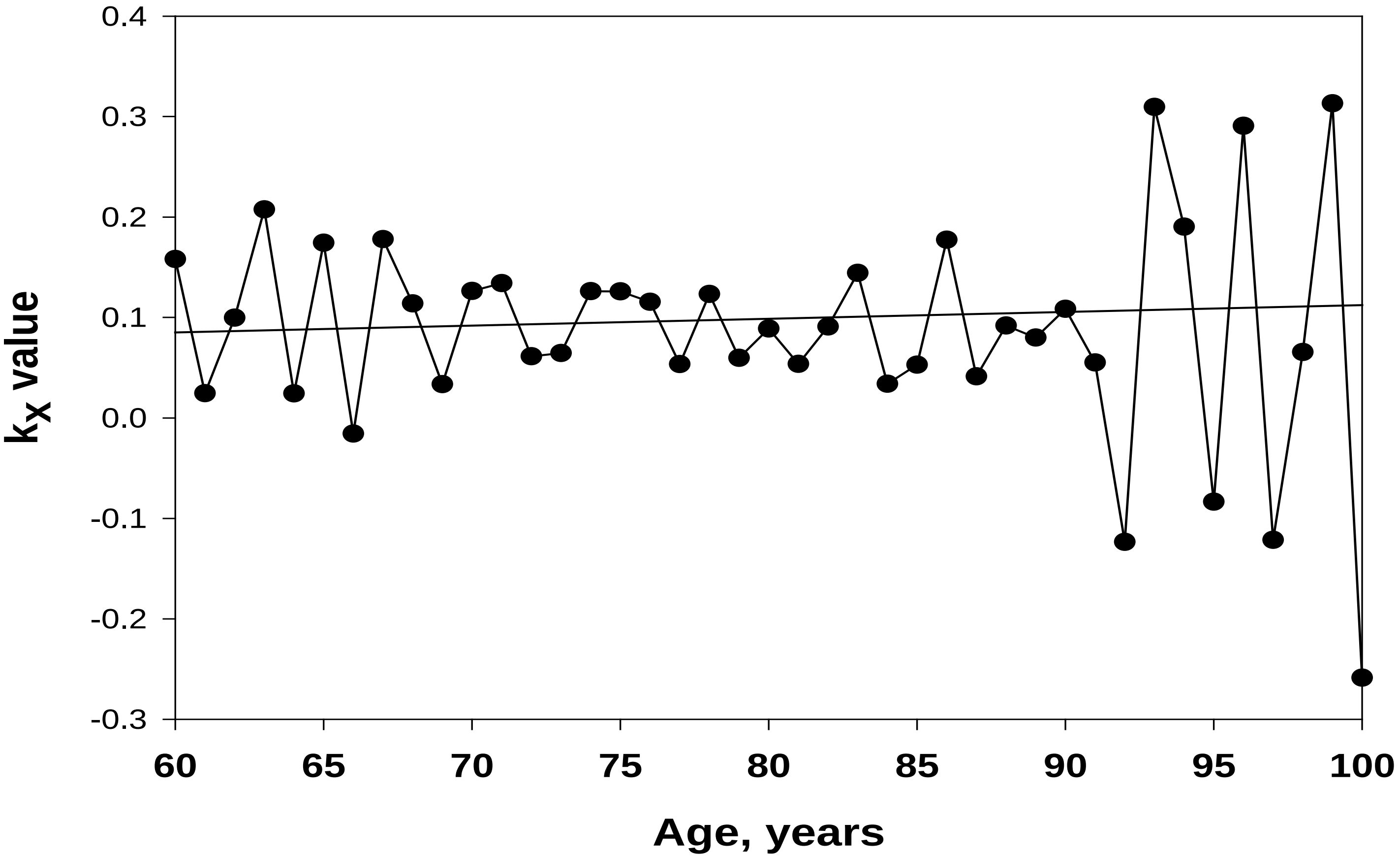
Alternative way to study mortality trajectories at advanced ages: Age-specific rate of mortality change

Suggested by Horiuchi and Coale (1990), Coale and Kisker (1990), Horiuchi and Wilmoth (1998) and later called 'life table aging rate (LAR)'

$$k(x) = d \ln \mu(x)/dx$$

- **Constant k(x) suggests that mortality follows the Gompertz model.**
- **Earlier studies found that k(x) declines in the age interval 80-100 years suggesting mortality deceleration.**

Age-specific rate of mortality change
Swedish males, 1896 birth cohort



Flat k(x) suggests that mortality follows the Gompertz law

Study of age-specific rate of mortality change using cohort data

- Age-specific cohort death rates taken from the Human Mortality Database
- Studied countries: Canada, France, Sweden, United States
- Studied birth cohorts: 1894, 1896, 1898
- k(x) calculated in the age interval 80-100 years
- k(x) calculated using one-year mortality rates

Slope coefficients (with p-values) for linear regression models of k(x) on age

Country	Sex	Birth cohort					
		1894		1896		1898	
		slope	p-value	slope	p-value	slope	p-value
Canada	F	-0.00023	0.914	0.00004	0.984	0.00066	0.583
	M	0.00112	0.778	0.00235	0.499	0.00109	0.678
France	F	-0.00070	0.681	-0.00179	0.169	-0.00165	0.181
	M	0.00035	0.907	-0.00048	0.808	0.00207	0.369
Sweden	F	0.00060	0.879	-0.00357	0.240	-0.00044	0.857
	M	0.00191	0.742	-0.00253	0.635	0.00165	0.792
USA	F	0.00016	0.884	0.00009	0.918	0.000006	0.994
	M	0.00006	0.965	0.00007	0.946	0.00048	0.610

All regressions were run in the age interval 80-100 years.

In previous studies mortality rates were calculated for five-year age intervals

$$k_x = \frac{\ln(m_x) - \ln(m_{x-5})}{5}$$

- **Five-year age interval is very wide for mortality estimation at advanced ages.**
- **Assumption about uniform distribution of deaths in the age interval does not work for 5-year interval**
- **Mortality rates at advanced ages are biased downward**

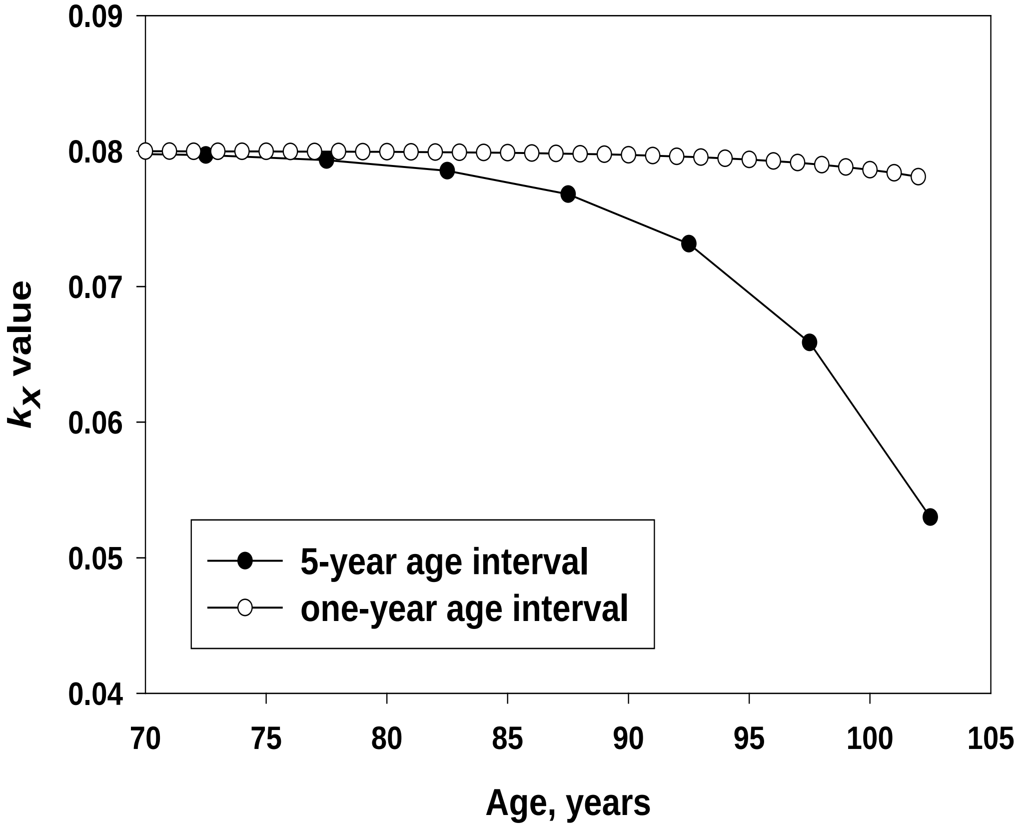
Simulation study of mortality following the Gompertz law

- 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 $year^{-1}$
- Numbers of survivors were calculated using formula (Gavrilov et al., 1983):

$$\frac{N_x}{N_0} = \frac{N_{x0}}{N_0} \exp\left(\left(-\frac{a}{b}\right)(e^{bx} - e^{bx0})\right)$$

where Nx/N0 is the probability of survival to age x, i.e. the number of hypothetical cohort at age x divided by its initial number N0. *a* and *b (slope)* are parameters of Gompertz equation

Age-specific rate of mortality change with age, kx, by age interval for mortality calculation
Simulation study of Gompertz mortality



Taking into account that underlying mortality follows the Gompertz law, the dependence of k(x) on age should be flat

Conclusions

- Below age 107 years and for data of reasonably good quality the Gompertz model fits mortality better than the Kannisto model (no mortality deceleration) for 20 studied single-year U.S. birth cohorts
- Age-specific rate of mortality change remains flat in the age interval 80-100 years for 24 studied single-year birth cohorts of Canada, France, Sweden and United States suggesting that mortality follows the Gompertz law

Acknowledgments

This study was made possible thanks to:

generous support from the

- National Institute on Aging (R01 AG028620)
- 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/

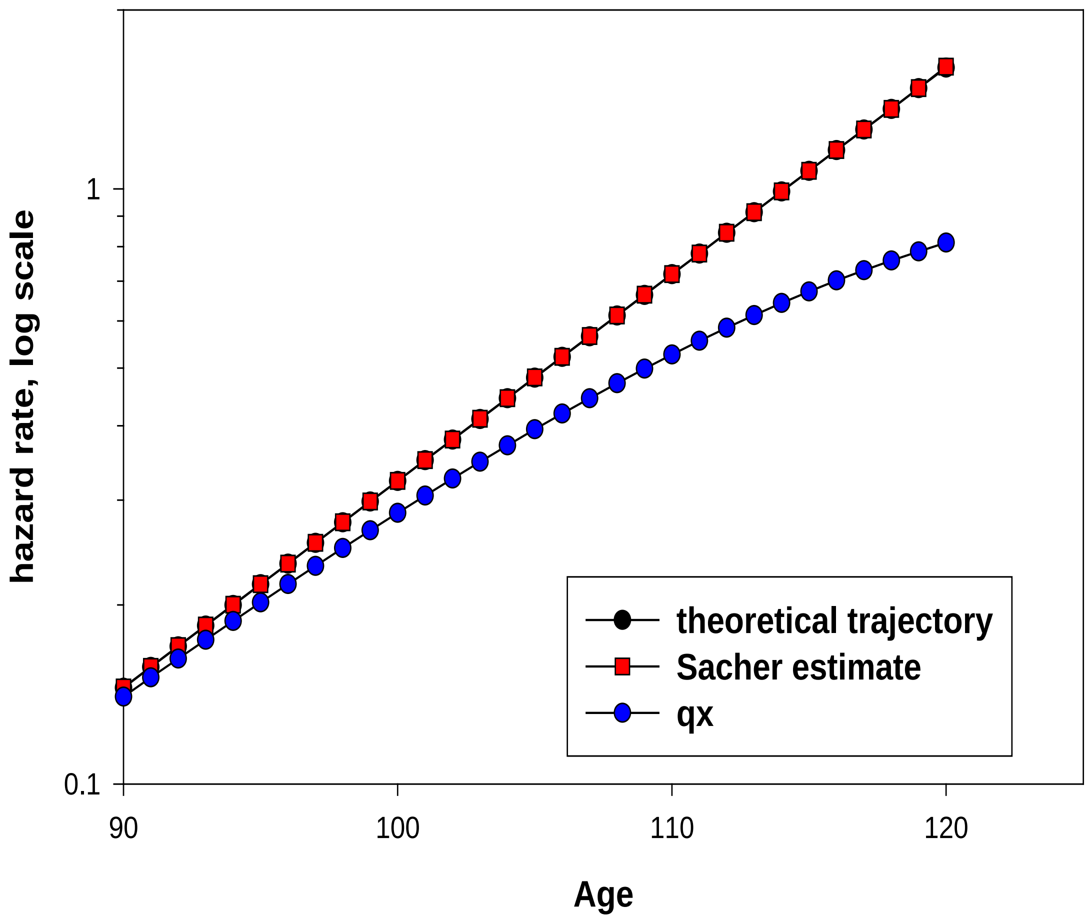
Which estimate of hazard rate is the most accurate?

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)**

Simulation study of Gompertz mortality

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



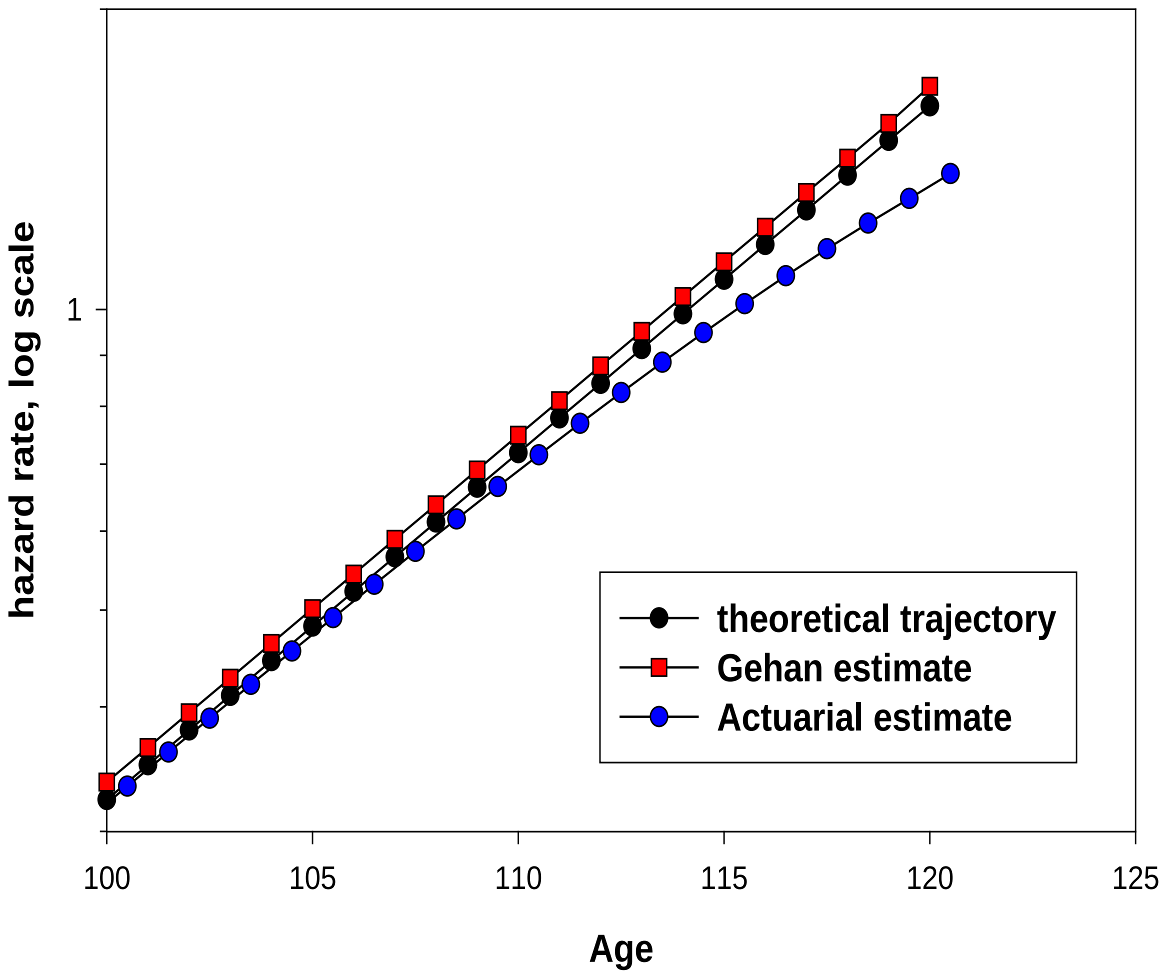
Sacher estimates practically coincide with theoretical mortality trajectory

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

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

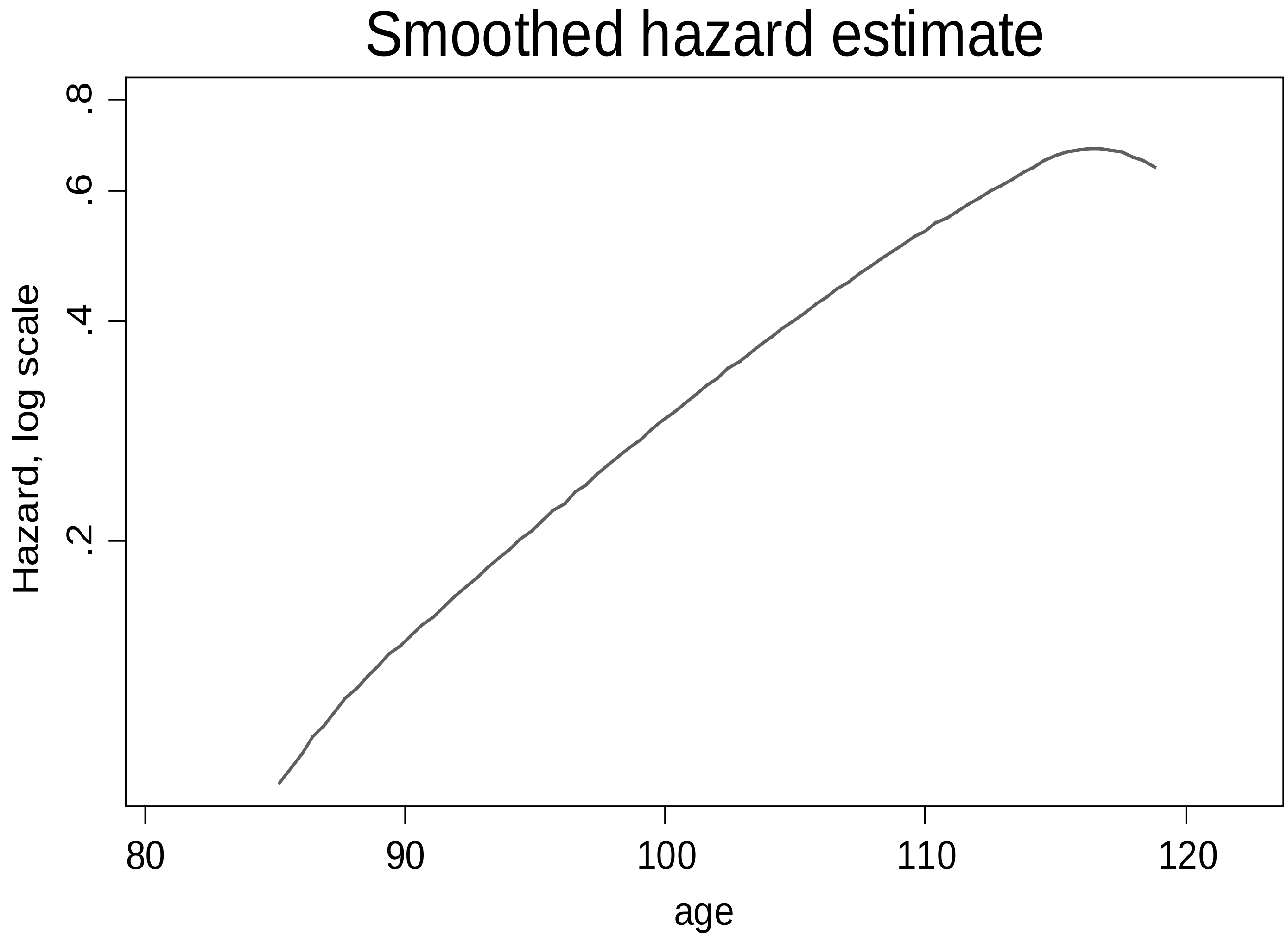
$$\mu_x = -\ln(1 - q_x)$$

Actuarial estimates undeestimate mortality after age 100

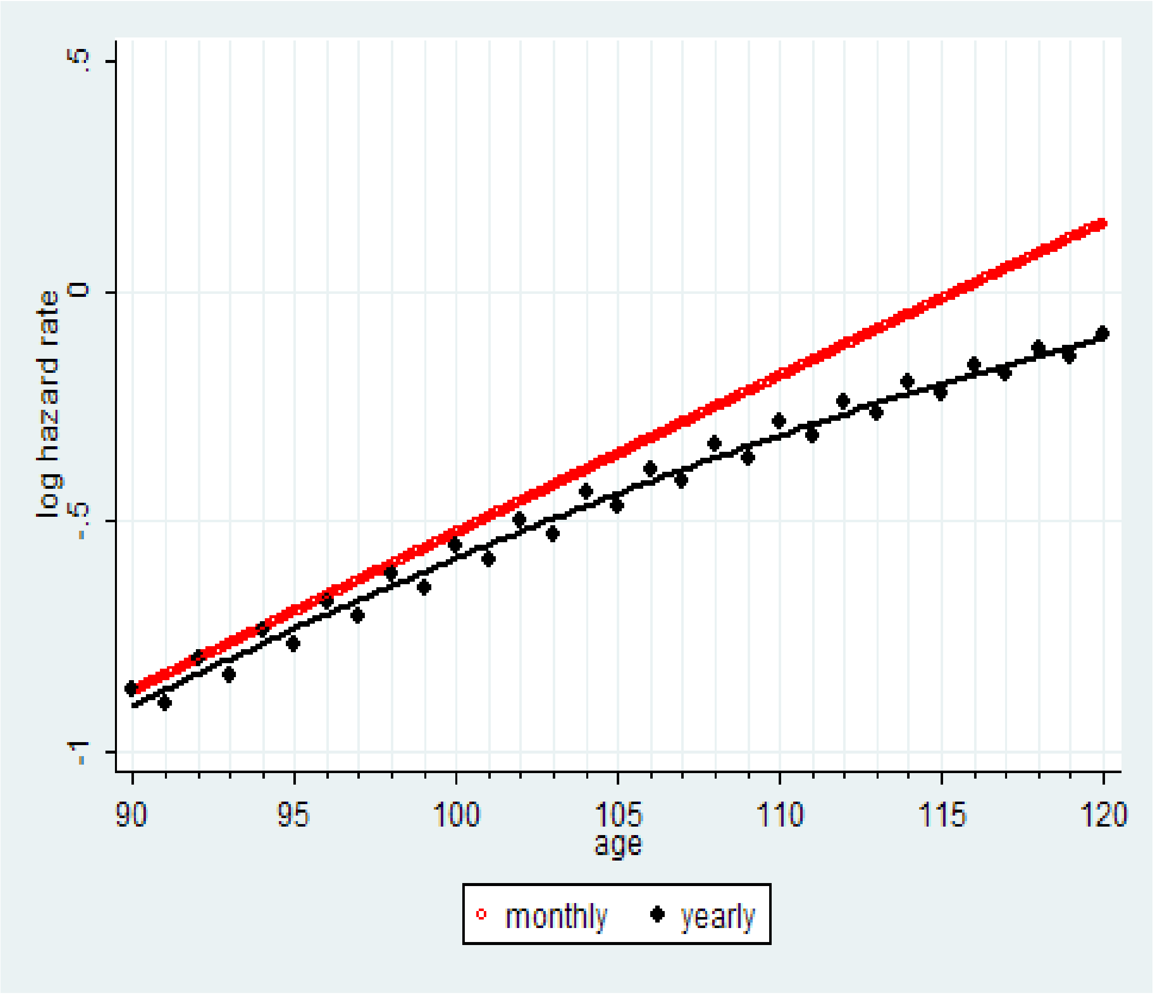
$$\mu_{x+\frac{\Delta x}{2}} = \frac{2}{\Delta x}\frac{l_x - l_{x+\Delta x}}{l_x + l_{x+\Delta x}}$$

Simulation study of the Gompertz mortality

Kernel smoothing of hazard rates



Monthly Estimates of Mortality are More Accurate
Simulation assuming Gompertz law for hazard rate



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

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

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:

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

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

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

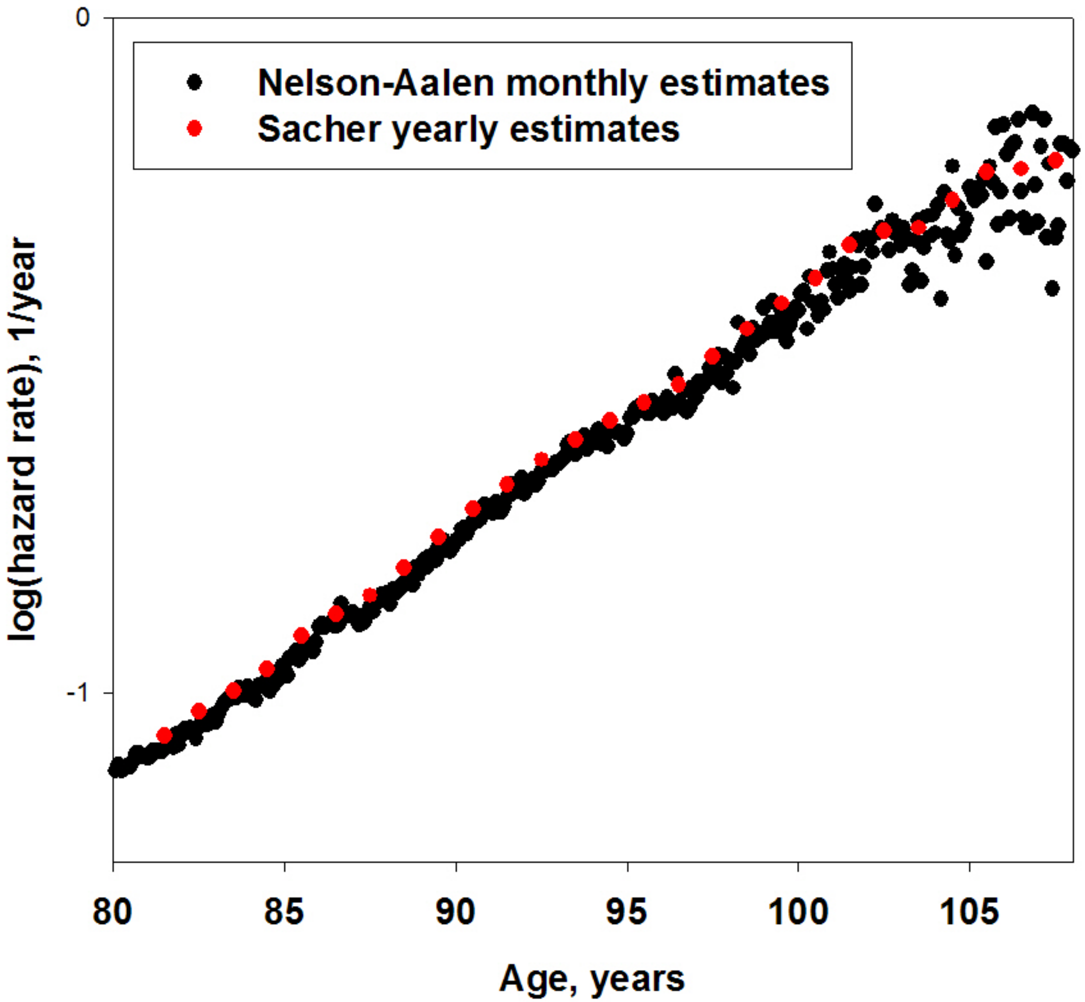
Hazard rate

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

Simplified version suggested by Gehan (1969):

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

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

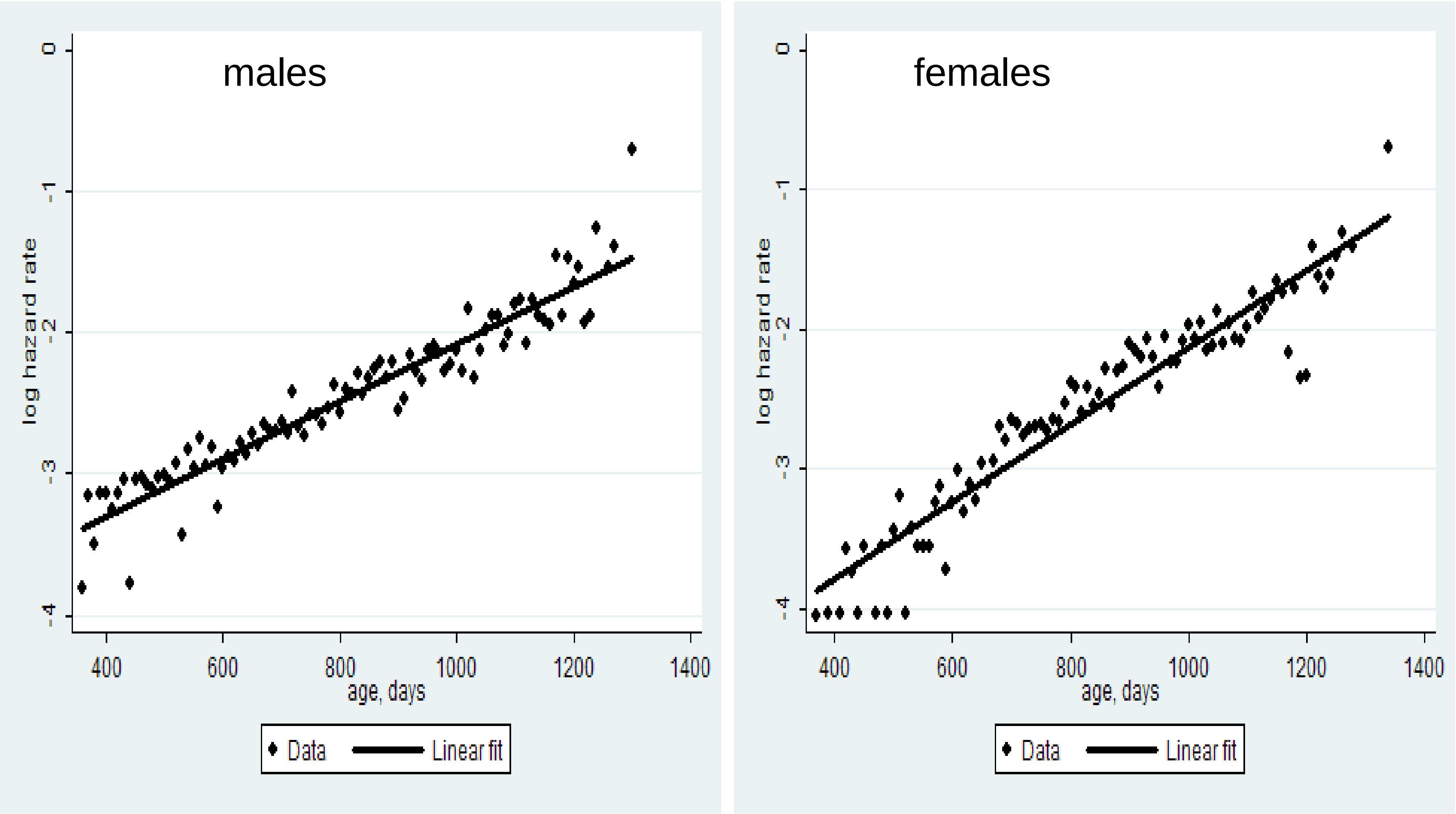


What about other mammals?

Mortality data for mice:

- Data from the NIH Interventions Testing Program, courtesy of Richard Miller (U of Michigan)
- Argonne National Laboratory data, courtesy of Bruce Carnes (U of Oklahoma)

Mortality of mice (log scale)
Miller data



- **Actuarial estimate of hazard rate with 10-day age intervals**

Bayesian information criterion (BIC) to compare the Gompertz and Kannisto models, mice data

Dataset	Miller data Controls		Miller data Exp., no life extension		Carnes data Early controls		Carnes data Late controls	
Sex	M	F	M	F	M	F	M	F
Cohort size at age one year	1281	1104	2181	1911	364	431	487	510
Gompertz	**-597.5**	**-496.4**	**-660.4**	**-580.6**	**-585.0**	**-566.3**	**-639.5**	**-549.6**
Kannisto	-565.6	-495.4	-571.3	-577.2	-556.3	-558.4	-638.7	-548.0

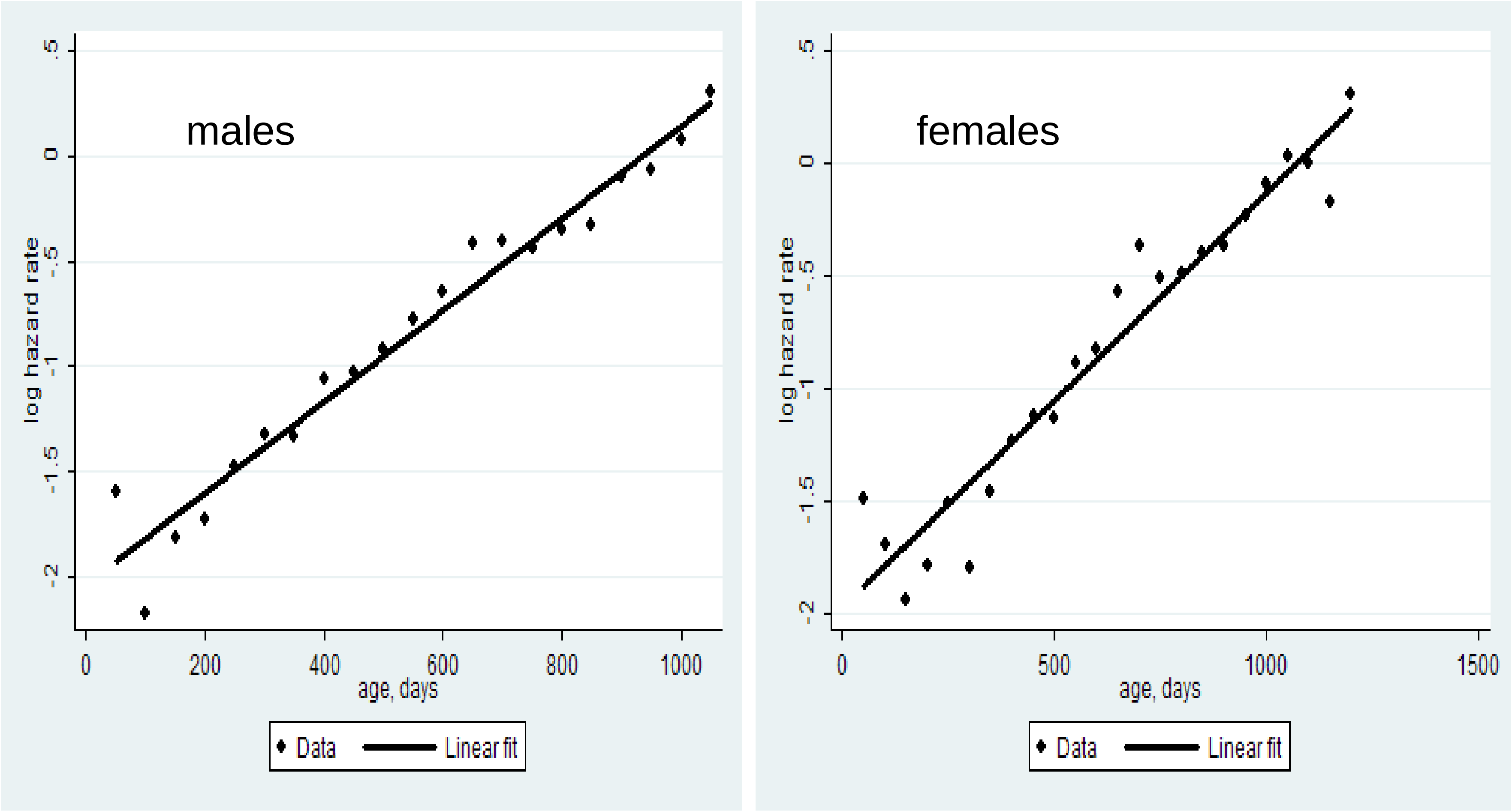
Better fit (lower BIC) is highlighted in red

Conclusion: In all cases Gompertz model demonstrates better fit than Kannisto model for mortality of mice after one year of age

Laboratory rats

- **Data sources: Dunning, Curtis (1946); Weisner, Sheard (1935), Schlettwein-Gsell (1970)**

Mortality of Wistar rats



- **Actuarial estimate of hazard rate with 50-day age intervals**
- **Data source: Weisner, Sheard, 1935**

Bayesian information criterion (BIC) to compare Gompertz and Kannisto models, rat data

Line	Wistar (1935)		Wistar (1970)		Copenhagen		Fisher		Backcrosses	
Sex	M	F	M	F	M	F	M	F	M	F
Cohort size	1372	1407	1372	2035	1328	1474	1076	2030	585	672
Gompertz	**-34.3**	**-10.9**	**-34.3**	**-53.7**	**-11.8**	**-46.3**	**-17.0**	**-13.5**	**-18.4**	**-38.6**
Kannisto	7.5	5.6	7.5	1.6	2.3	-3.7	6.9	9.4	2.48	-2.75

Better fit (lower BIC) is highlighted in red

Conclusion: In all cases Gompertz model demonstrates better fit than Kannisto model for mortality of laboratory rats