A Flexible Two-Dimensional Mortality Model for Use in Indirect Estimation

John Wilmoth *Univ. of California, Berkeley*
Sarah Zureick *Univ. of Wisconsin, Madison*
Vladimir Canudas-Romo *Johns Hopkins University*
Mie Inoue *World Health Organization*
Cheryl Sawyer *UN Population Division*

CEDEPLAR, Belo Horizonte, Brazil
13 June 2011

Background

- This study began in 2006 when I was working for the Population Division of the United Nations – inspired by comments from Hania Zlotnik about how to exploit better the information collected in the Human Mortality Database (HMD)
- At that time, Vladimir Canudas-Romo was working at the University of California in Berkeley as part of the HMD team

Today's presentation …

- Is based mostly on an article that will be published in one of the next issues of the journal *Population Studies*
- Includes some additional analysis of an interesting case: Costa Rica

Model Life Tables

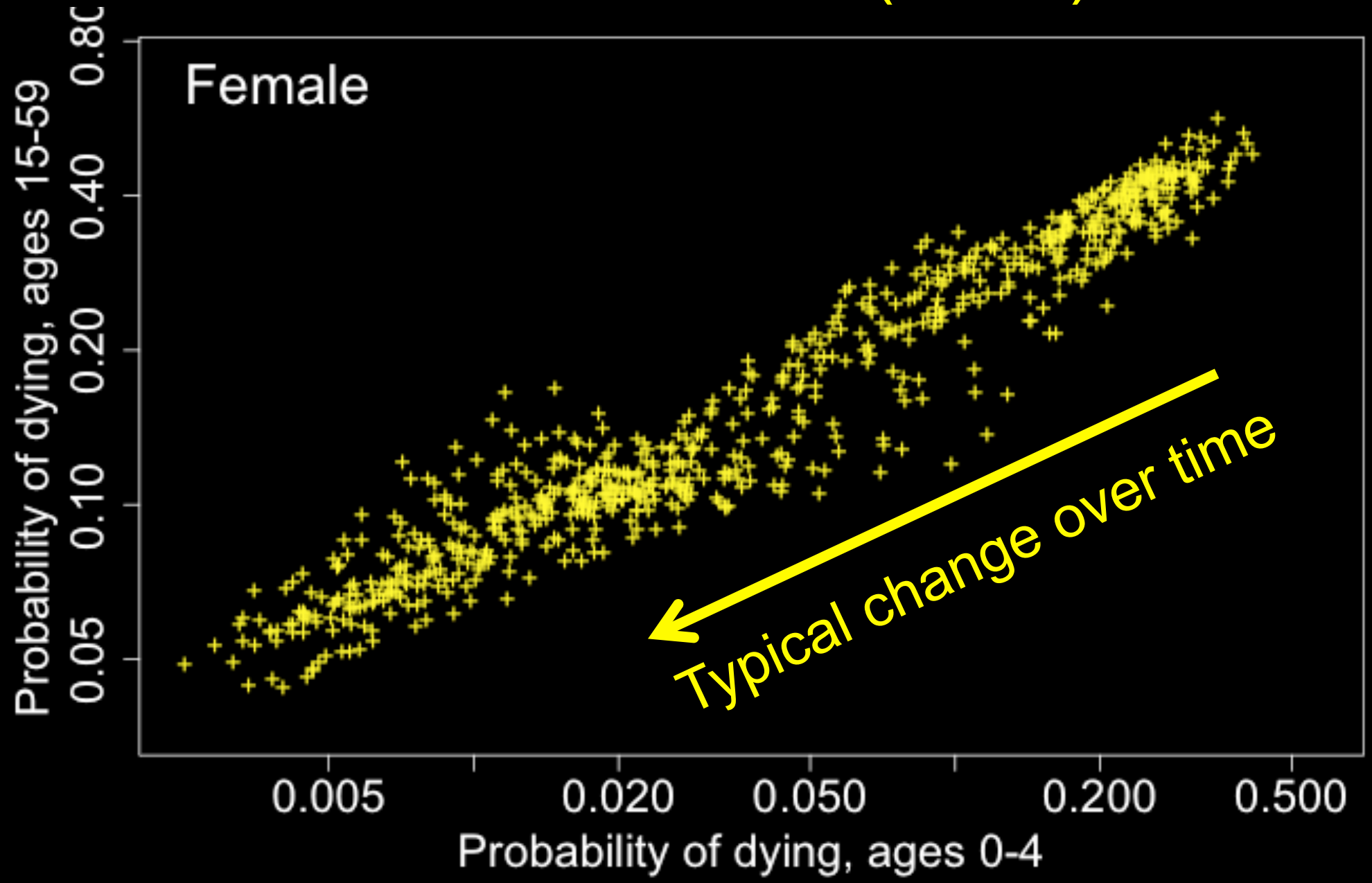
- Used for mortality estimation in populations that lack reliable age-specific data
- Each model life table system is based on an empirical model of typical variations in the age pattern of mortality, estimated using some collection of observed life tables
- For practical applications, these models should allow 1 or 2 “entry” parameters

Model Life Table Systems

- Three are currently used to produce internationally comparable estimates:
  - Coale-Demeny (1966, 1983)
  - United Nations (1982)
  - Modified Brass logit (2003)
- Each system implies a particular set of relationships between mortality risks at different ages – for example, ${}_{45}q_{15}$ vs. ${}_{5}q_{0}$

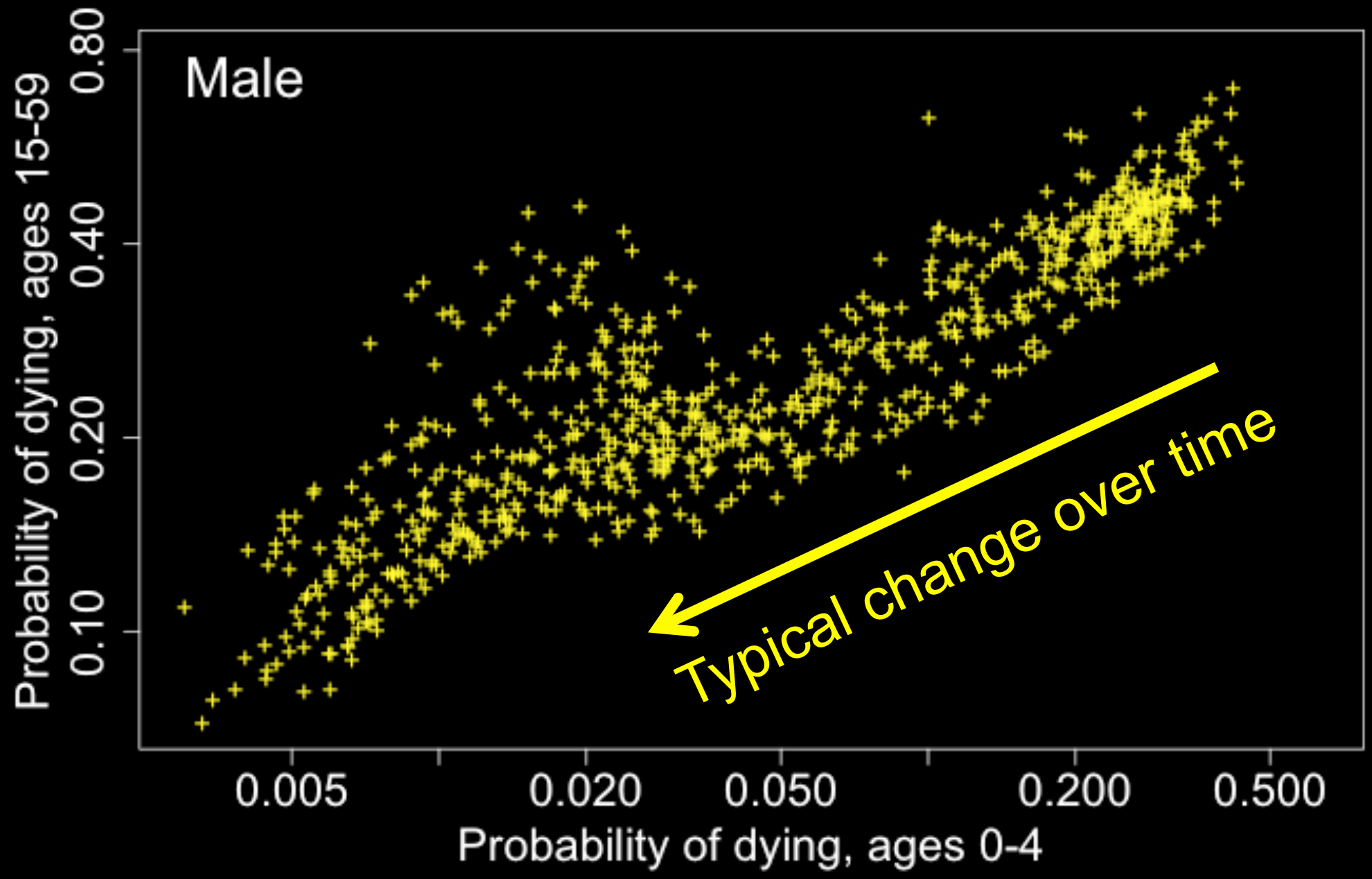
Adult ($_{45}q_{15}$) vs. Child ($_5q_0$) Mortality

Historical Data (HMD)



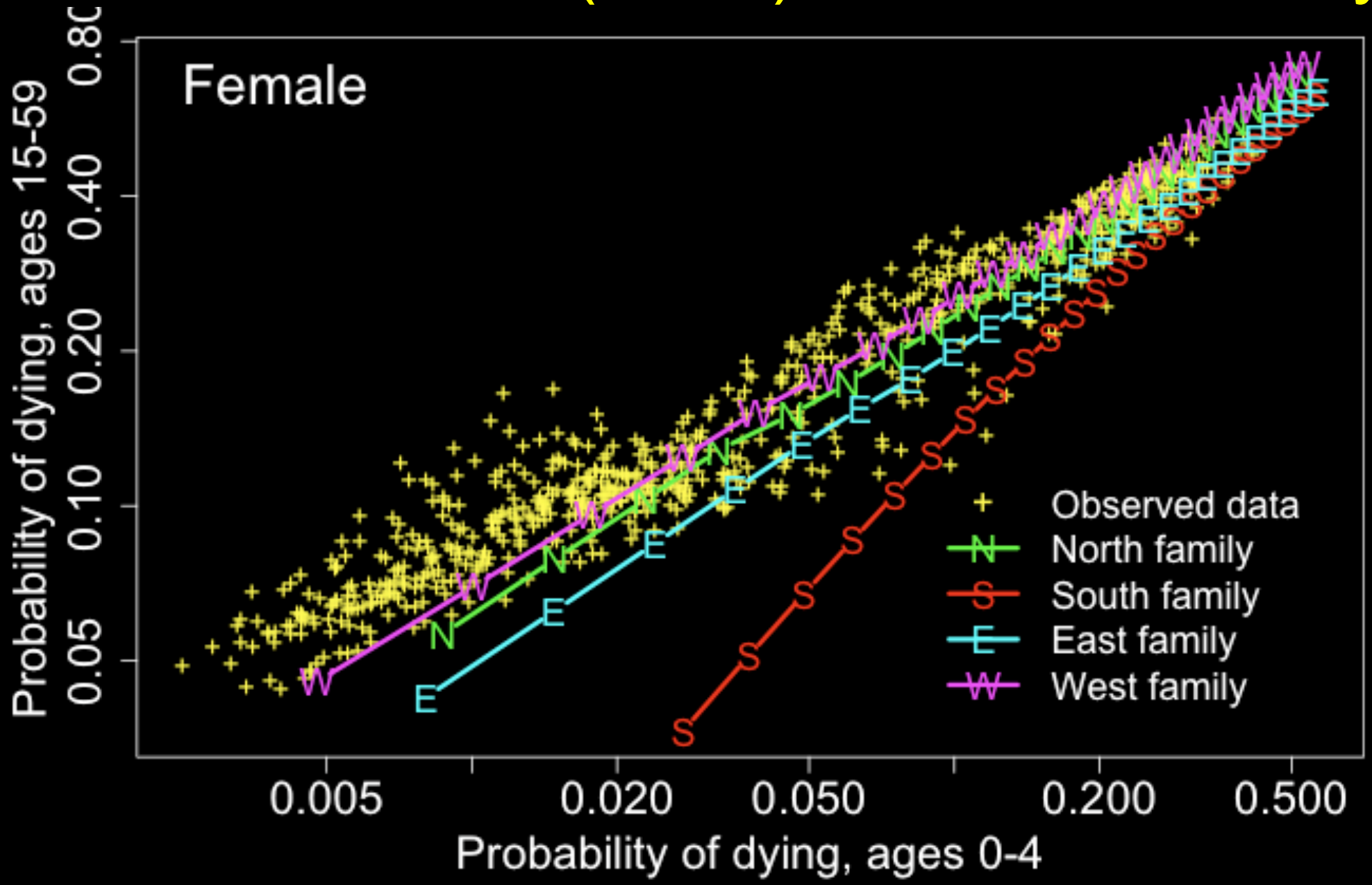
Adult ($_{45}q_{15}$) vs. Child ($_5q_0$) Mortality

Historical Data (HMD)



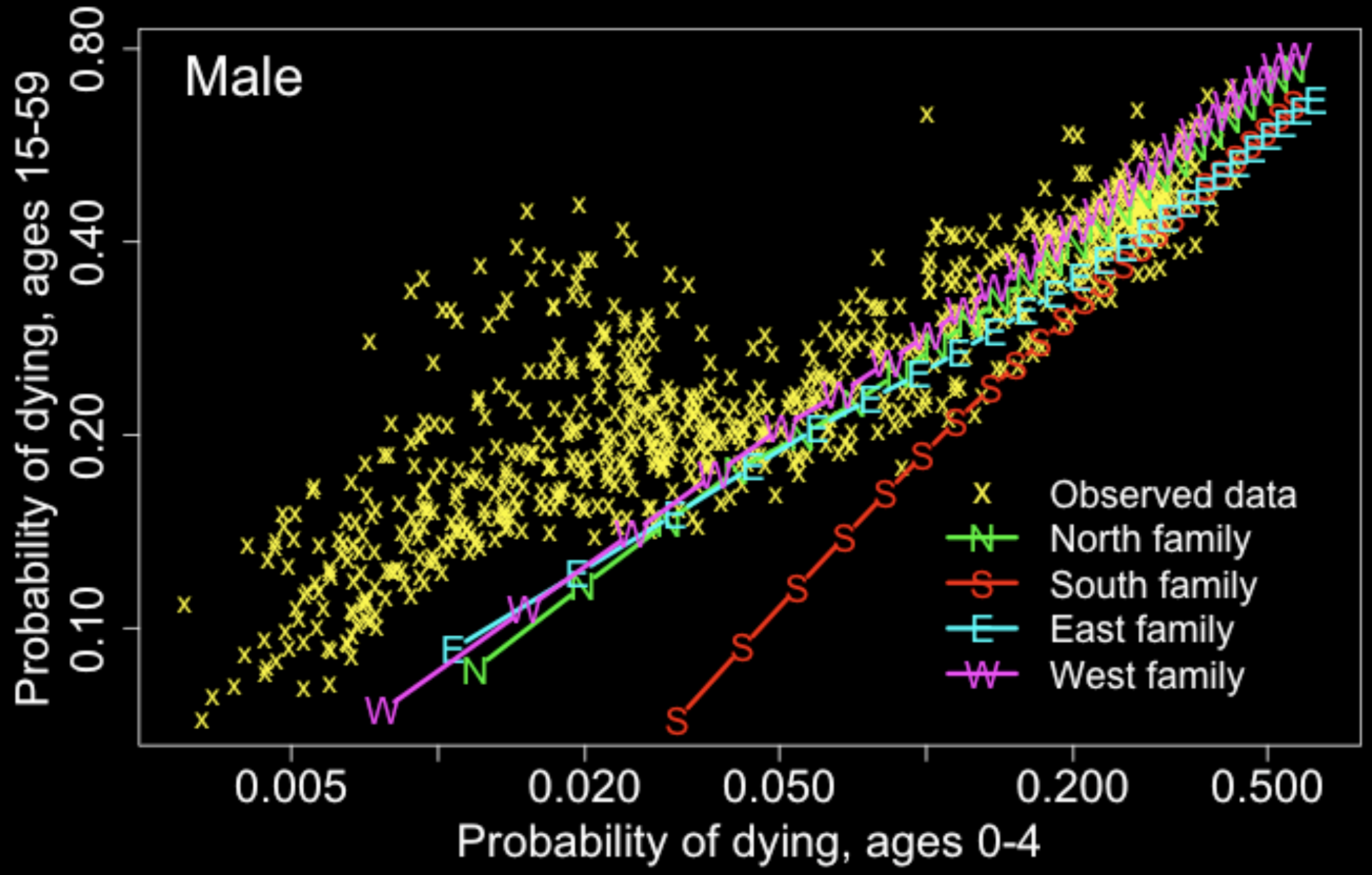
Adult ($_{45}q_{15}$) vs. Child ($_5q_0$) Mortality

Historical Data (HMD) + Coale-Demeny



Adult ($_{45}q_{15}$) vs. Child ($_5q_0$) Mortality

Historical Data (HMD) + Coale-Demeny



Motivations for this Analysis

- The need to recalibrate earlier model life table systems to better reflect the full body of empirical evidence now available
- The desire to create a model life table system that is as reliable, as flexible, and yet as simple as possible

Log-Quadratic Mortality Model

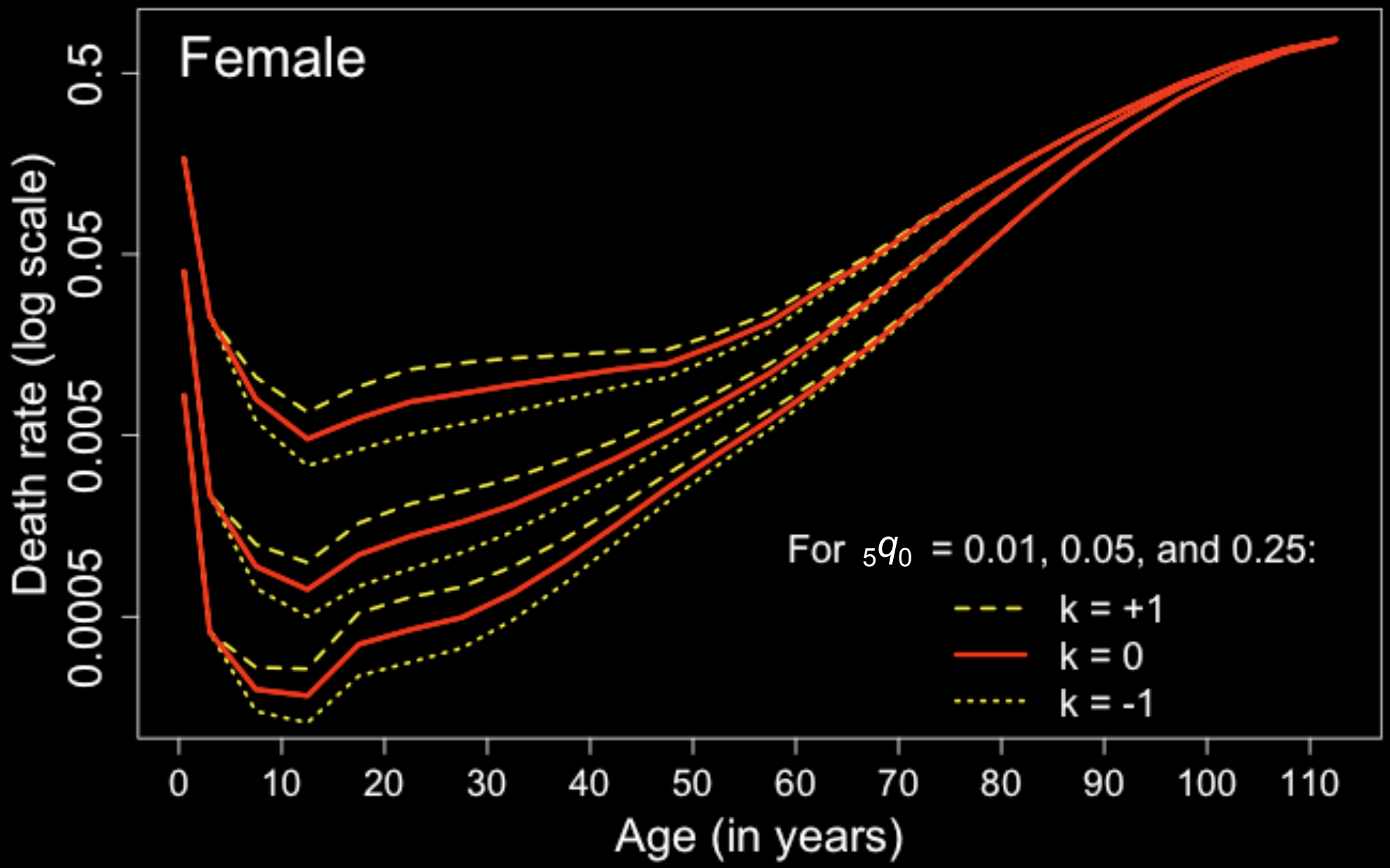
$$\log(\, m_x \,) = a_x + b_x h + c_x h^2 + v_x k + \varepsilon_x$$

Thus, there two entry parameters:

- $h = \log(\, {}_5q_0 \,)$ reflects the level of child mortality
- k reflects the level of excess adult mortality and is chosen to match ${}_{45}q_{15}$ or other global measure of adult mortality (if available)

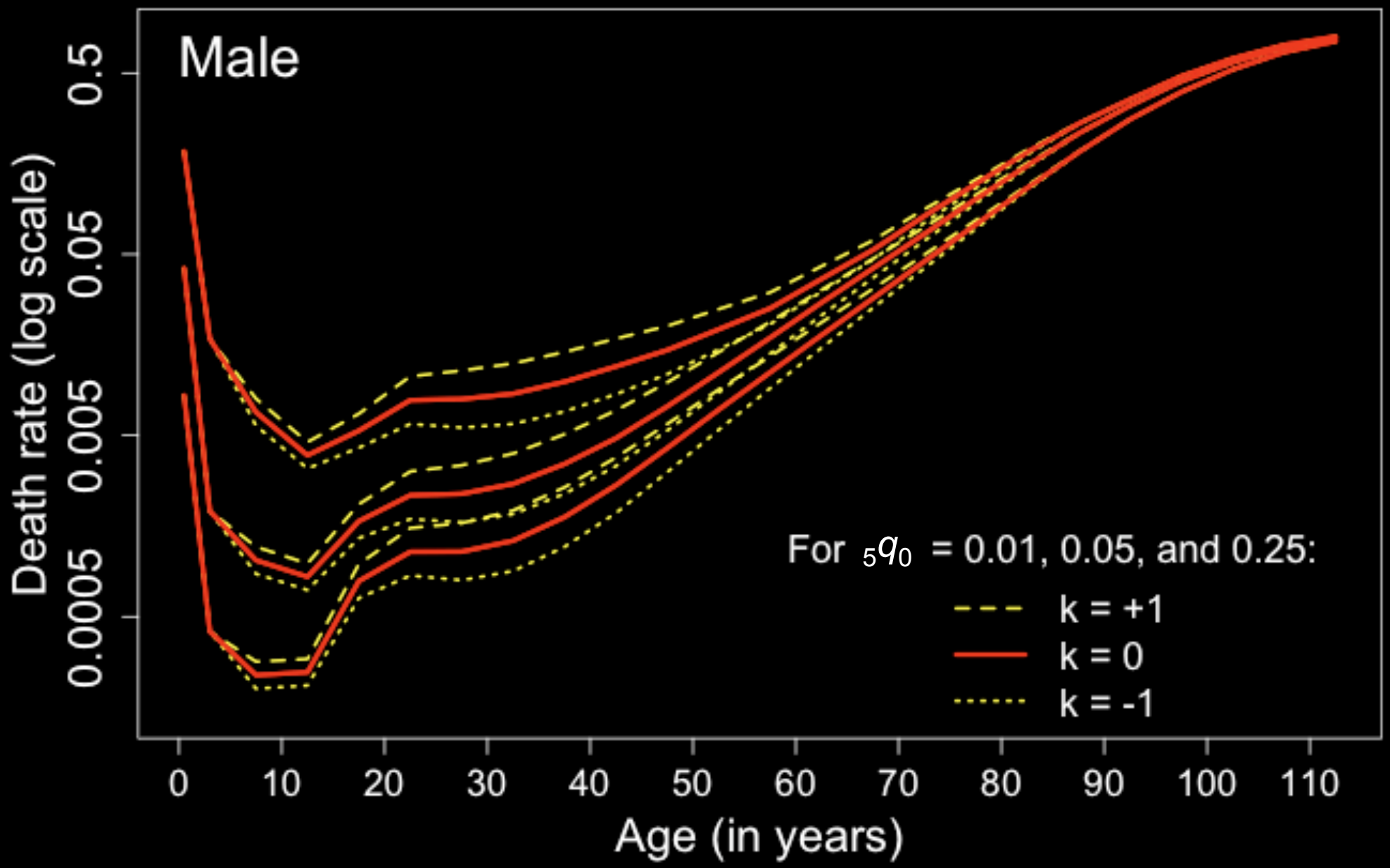
Log-Quadratic Model: Age Patterns

$$\log(m_x) = a_x + b_x h + c_x h^2 + v_x k$$



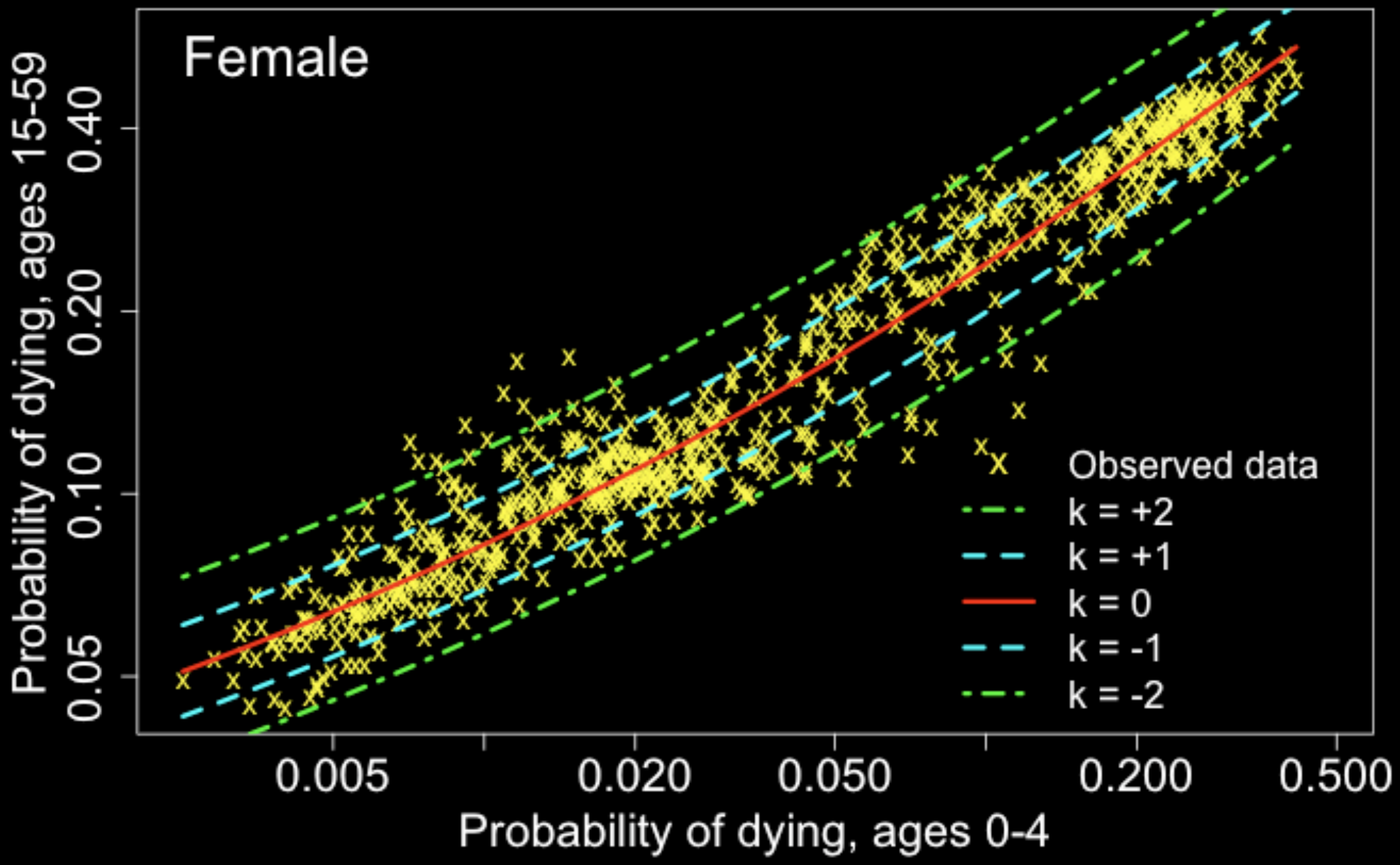
Log-Quadratic Model: Age Patterns

$$\log(m_x) = a_x + b_x h + c_x h^2 + v_x k$$



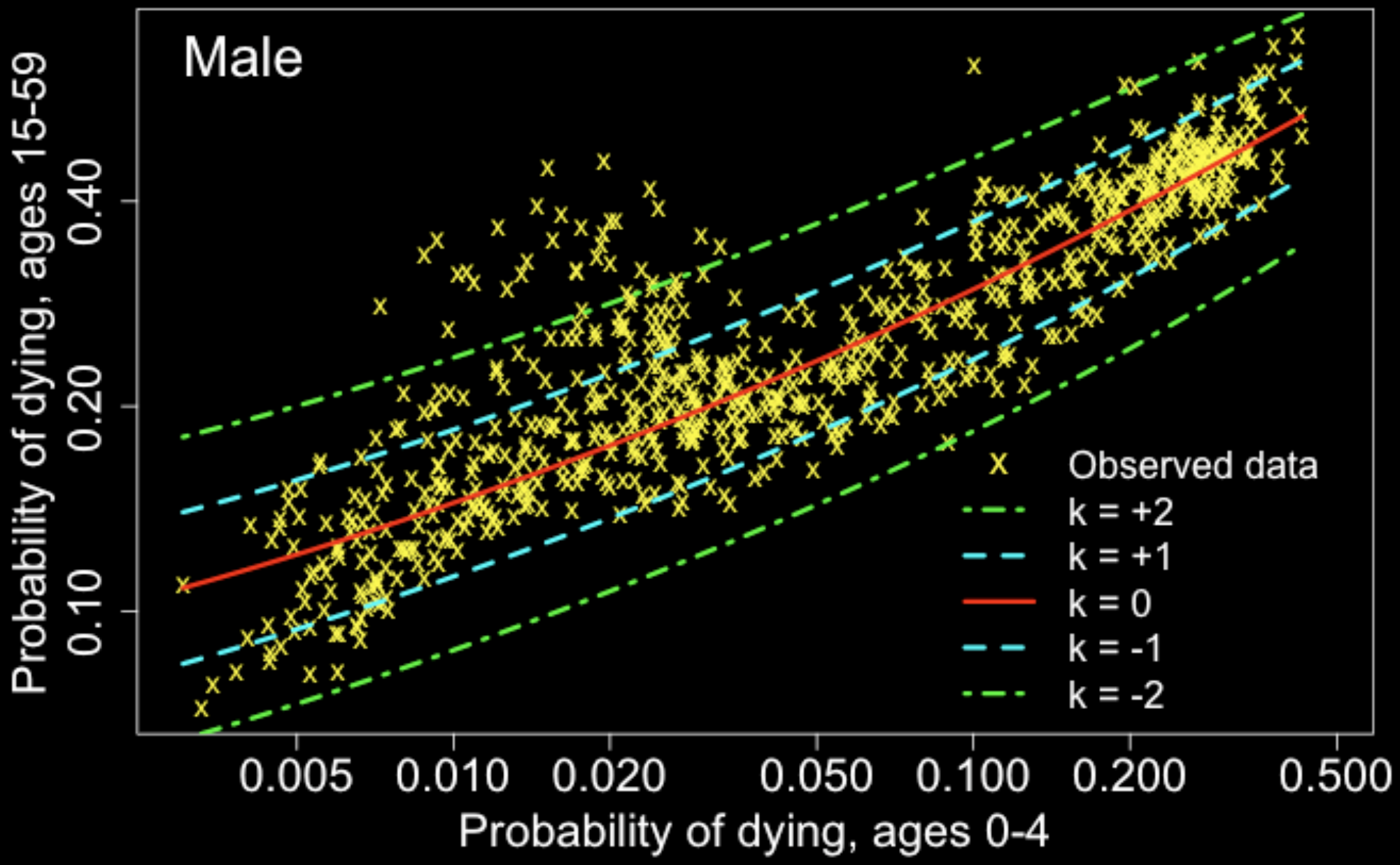
Adult ($_{45}q_{15}$) vs. Child ($_5q_0$) Mortality

Historical Data (HMD) + Log-Quadratic



Adult ($_{45}q_{15}$) vs. Child ($_5q_0$) Mortality

Historical Data (HMD) + Log-Quadratic



Fitting and Testing the Model

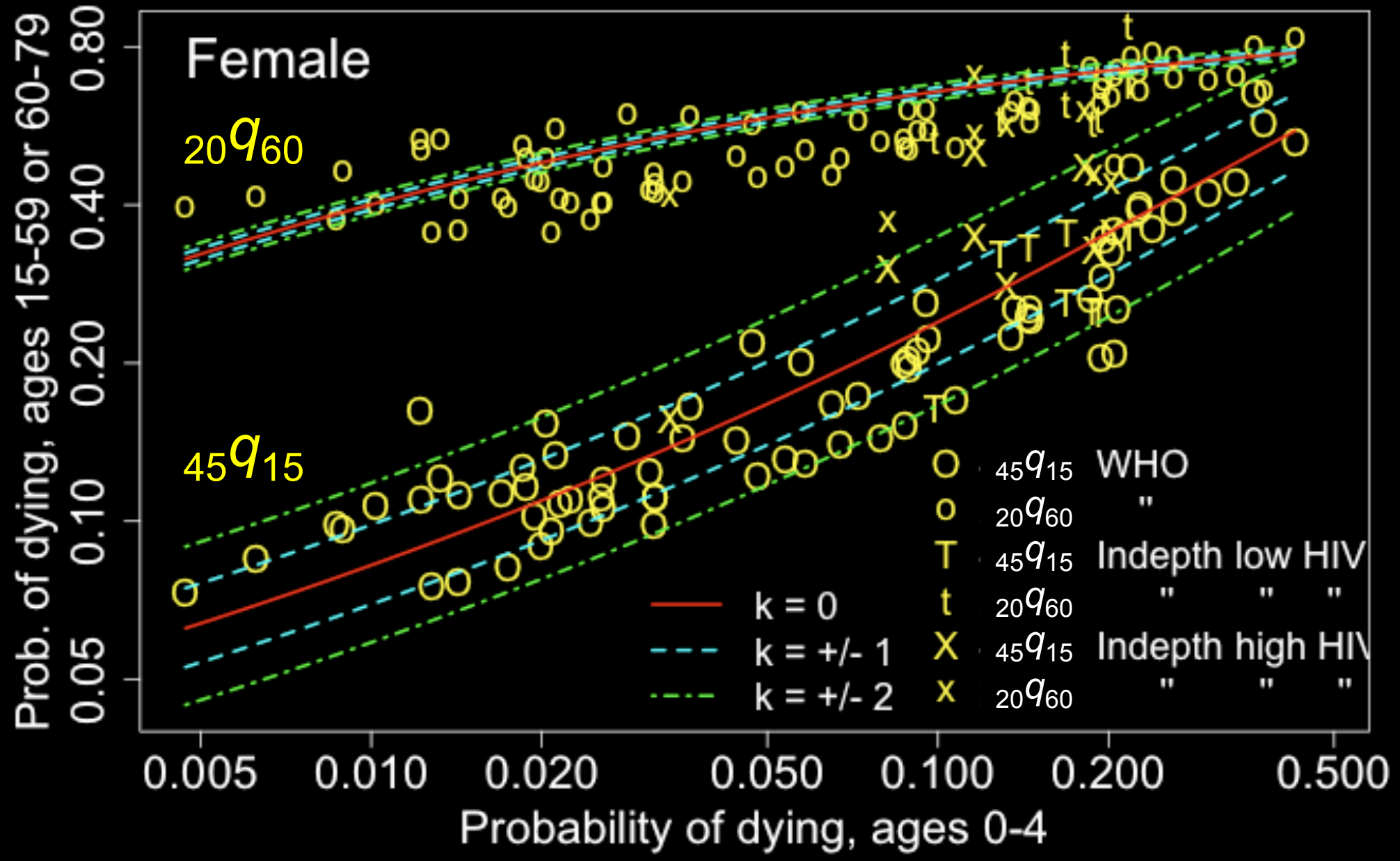
- Fitted using all HMD 5x5 life tables:
  - After eliminating some overlapping series
  - Using mortality estimates for civilian populations during wartime (when available)
- Testing has involved:
  - Internal tests using HMD data
  - External tests using other collections of empirical life tables (WHO, INDEPTH, HLD)

HMD Data Used for Fitting the Model

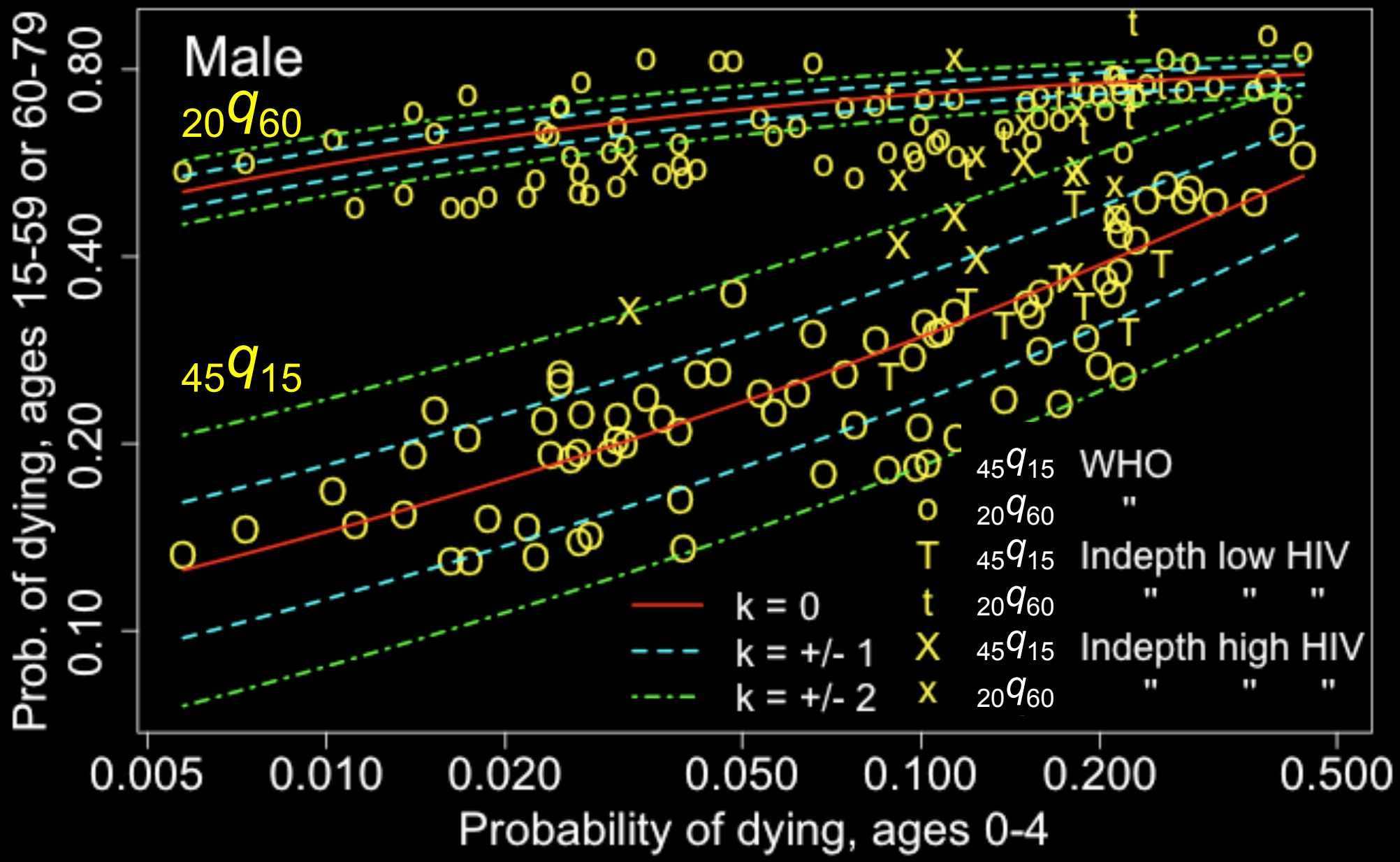
- Data from 5 continents (# of populations):
  - Europe (31)
  - North America (2)
  - South America (1)
  - Asia (2)
  - Oceania (2)
- And from 4 centuries (starting date):

  Sweden (1751), France (1816), Denmark (1835)
  … Russia (1959), Taiwan (1970), Chile (1992)

Fitting the Log-Quadratic Model
Data from Developing Countries



Fitting the Log-Quadratic Model
Data from Developing Countries



Summary of Test Results

- Tests based on root mean square error for estimates of e_0, ${}_1q_0$, ${}_{45}q_{15}$, and ${}_{20}q_{60}$
- Log-quadratic model yields consistently smaller estimation errors compared to Coale-Demeny or UN model life tables
- Log-quadratic model and modified logit models have similar levels of accuracy (in this case the main advantages of the log-quadratic are flexibility and simplicity)

Practical Applications of the Model for Estimating Mortality

- For some populations, the only available empirical data refer to child mortality (${}_5q_0$)
- For other populations, there may also be data on adult mortality (${}_5q_0$ and ${}_{45}q_{15}$)
- By design, the log-quadratic model can accommodate both situations

Two Versions of the Log-Quadratic Model

One-parameter ($k = 0$) :

$$\log(m_x) = a_x + b_x h + c_x h^2$$

Two-parameter (k to match ${}_{45}q_{15}$) :

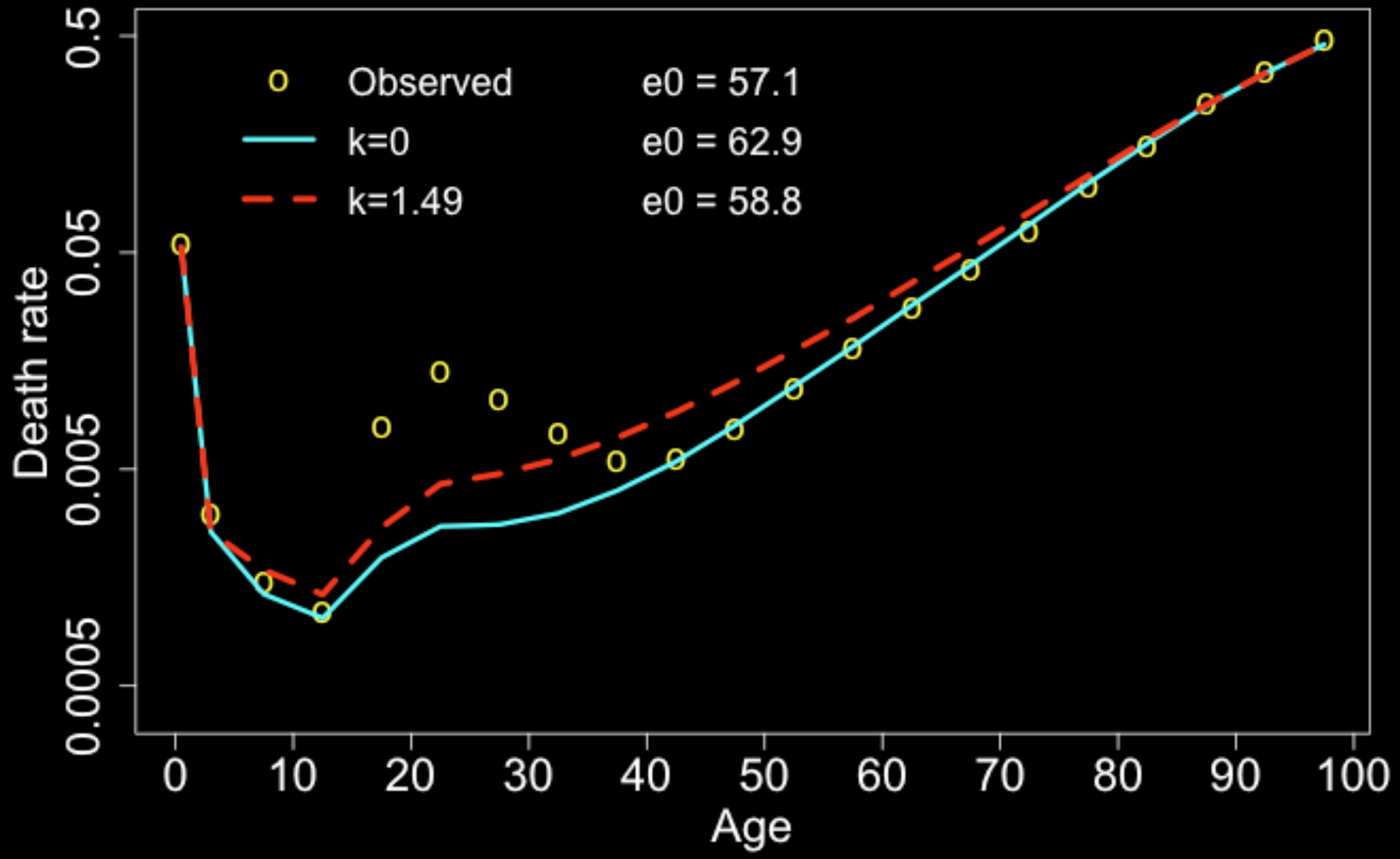
$$\log(m_x) = a_x + b_x h + c_x h^2 + v_x k$$

where $h = \log({}_5q_0)$

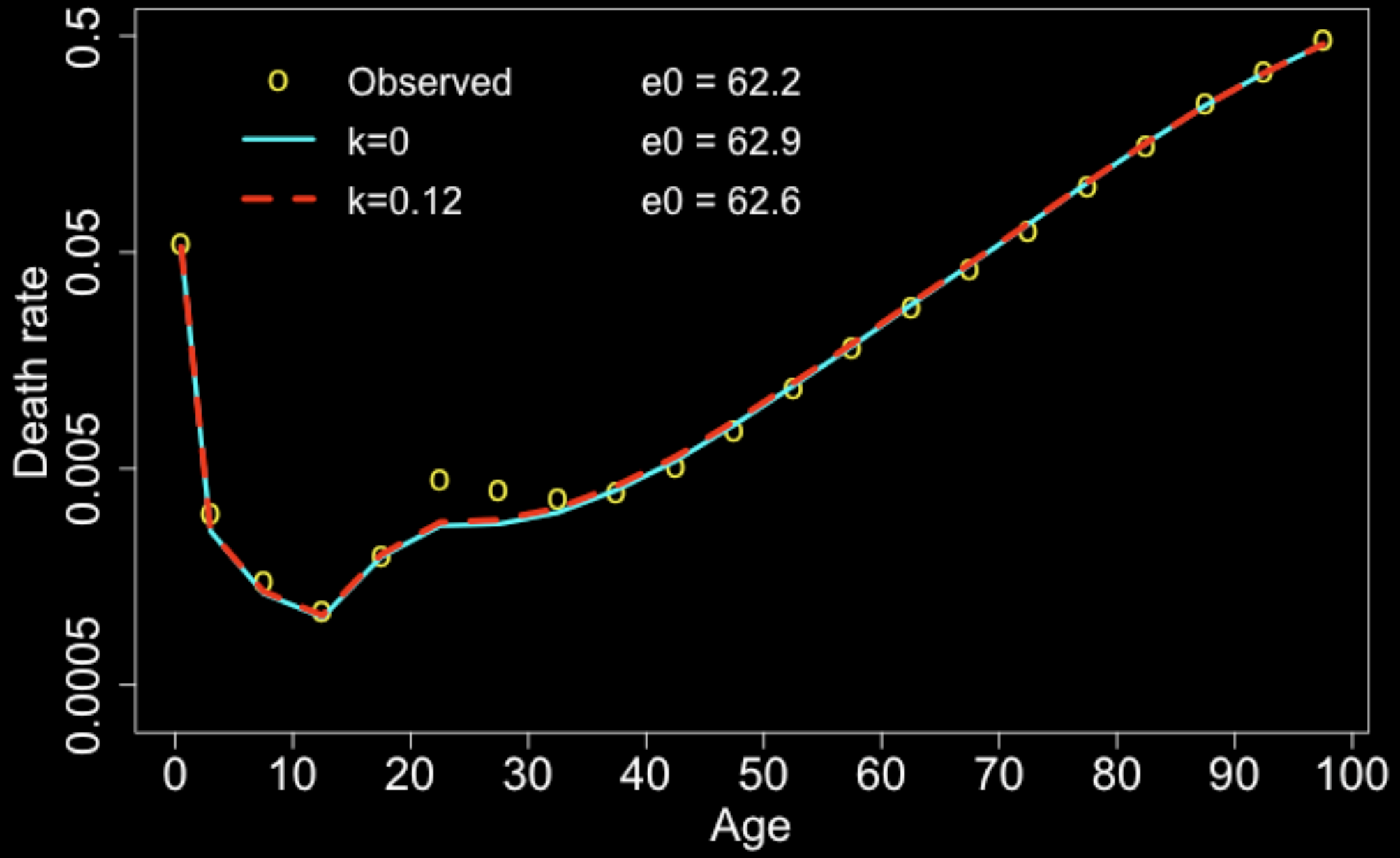
Some Not-so-typical Examples

- England & Wales during 2nd World War
- Denmark during the Spanish flu epidemic
- Spain during the Civil War
- Portugal in the 1960s
- Russia in recent years

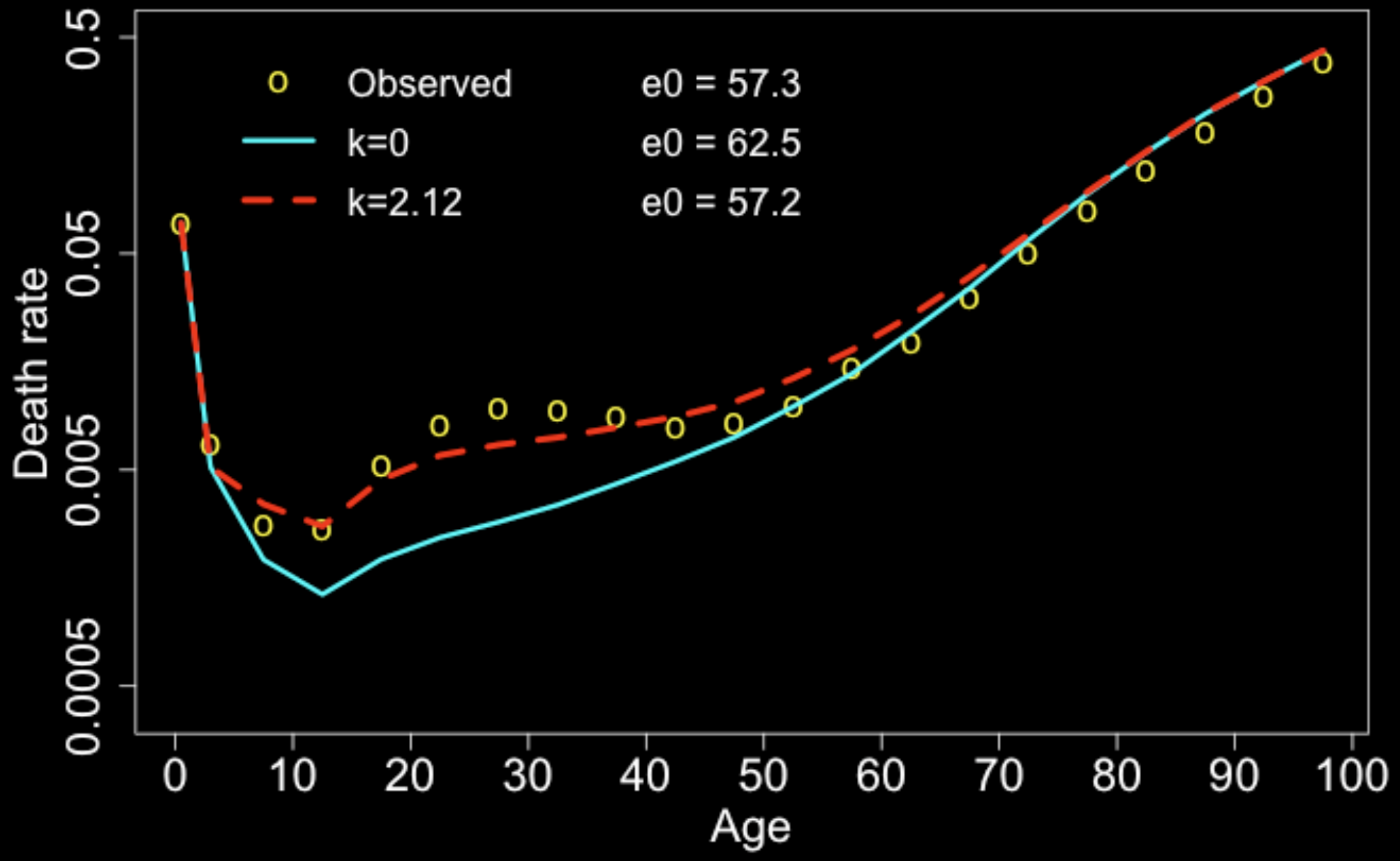
England & Wales, 1943-45, Men
Total population



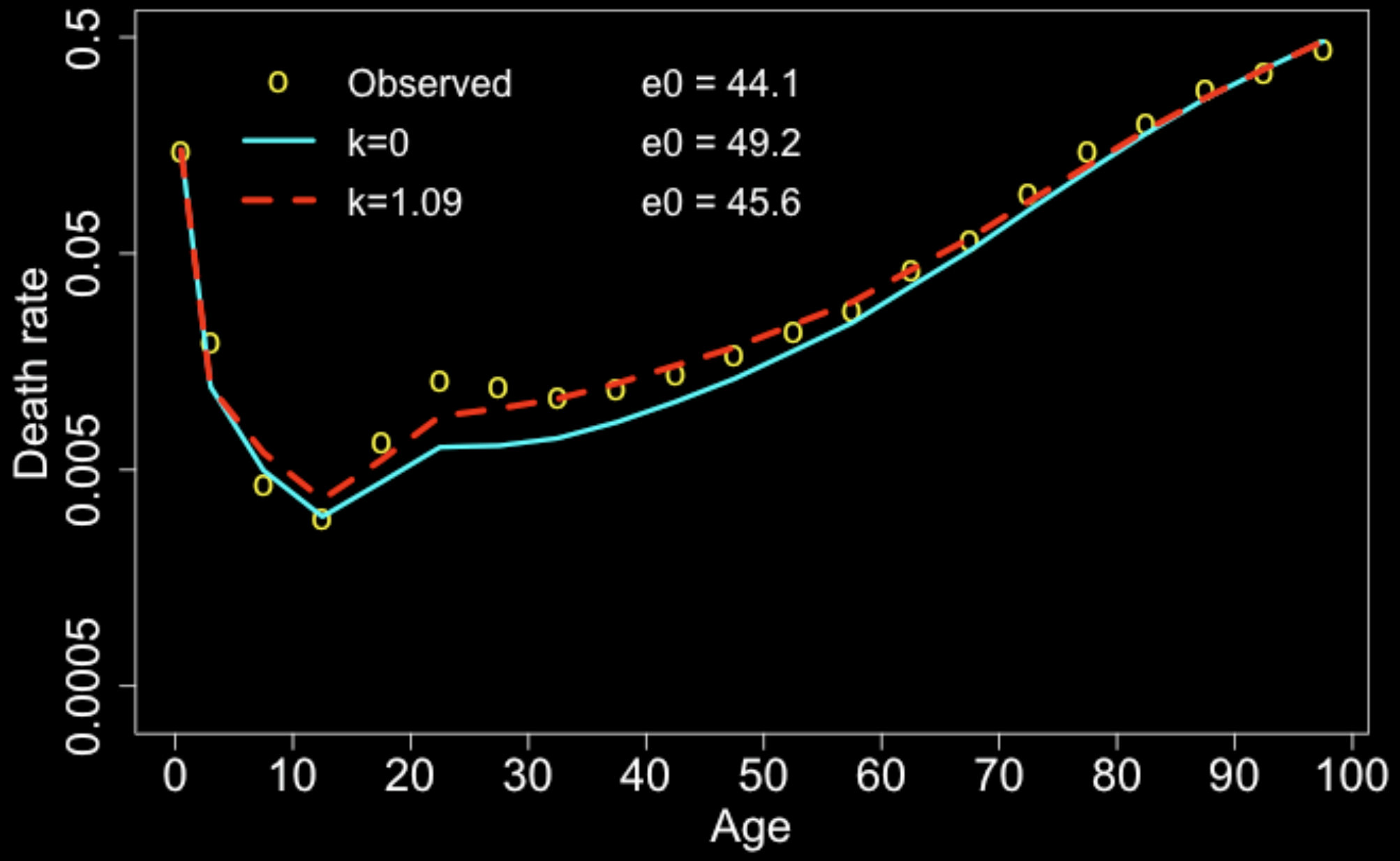
England & Wales, 1943-45, Men
Civilian population only



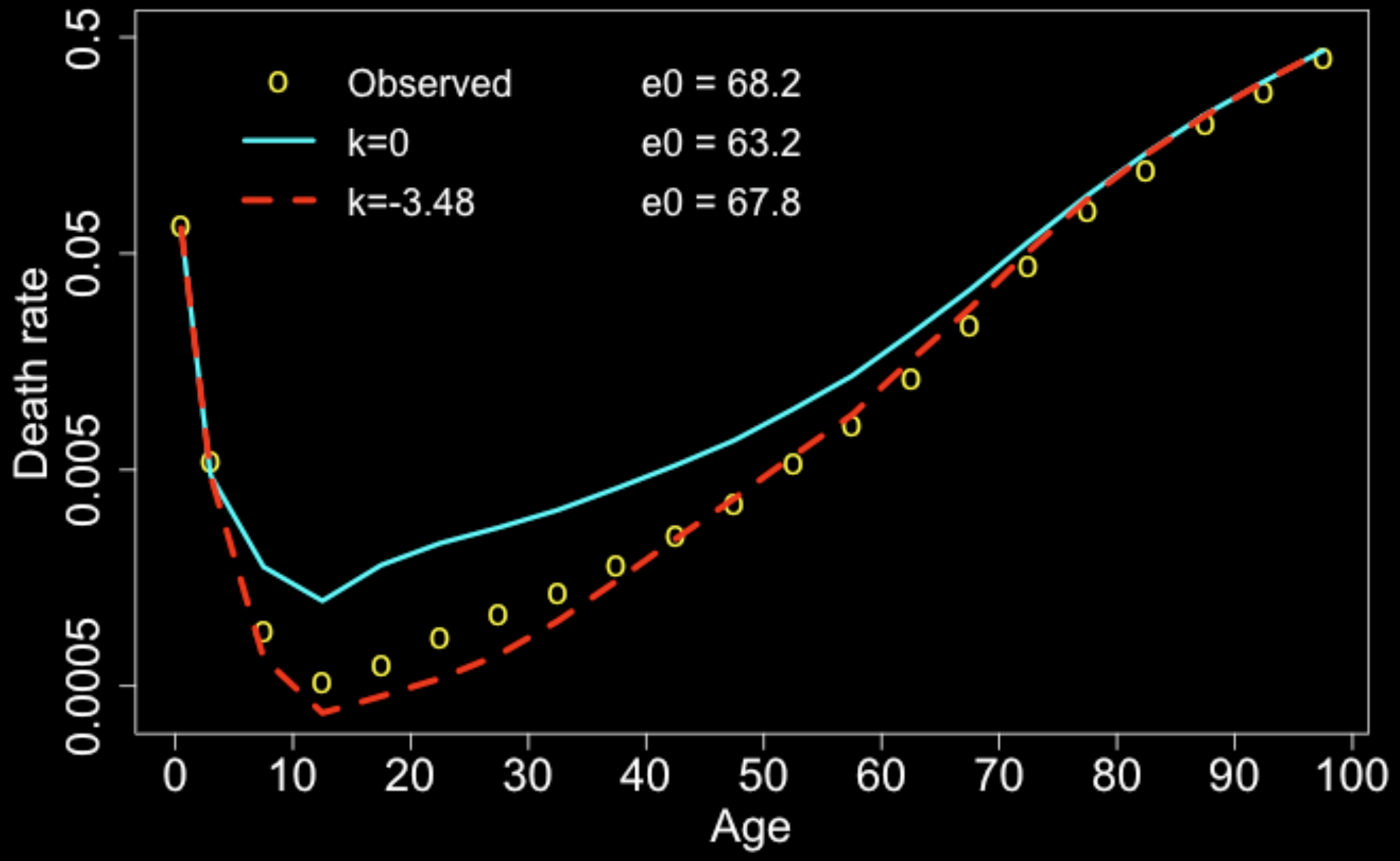
Denmark, 1918, Women



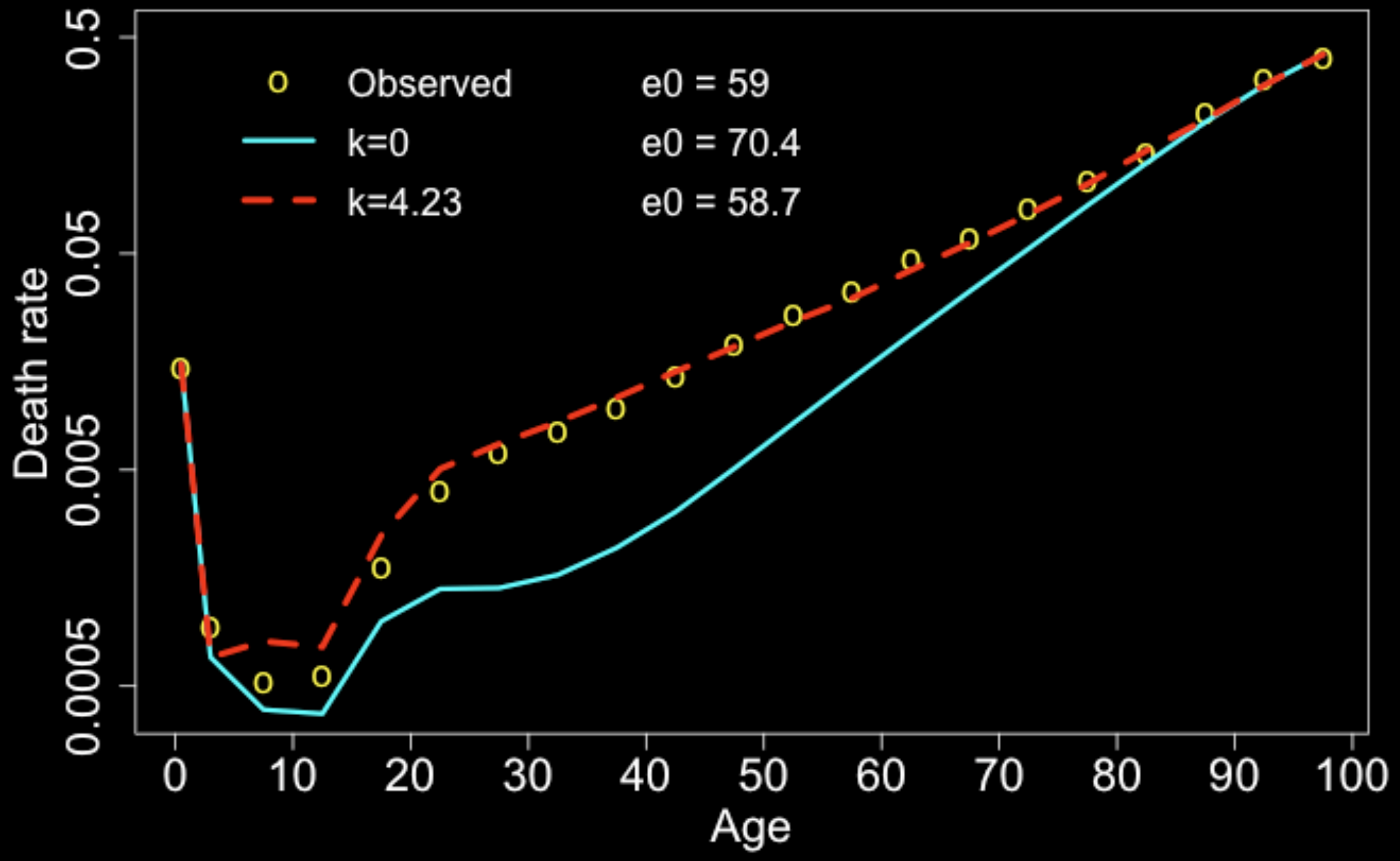
Spain, 1936-39, Men



Portugal, 1960-69, Women



Russia, 2000-06, Men

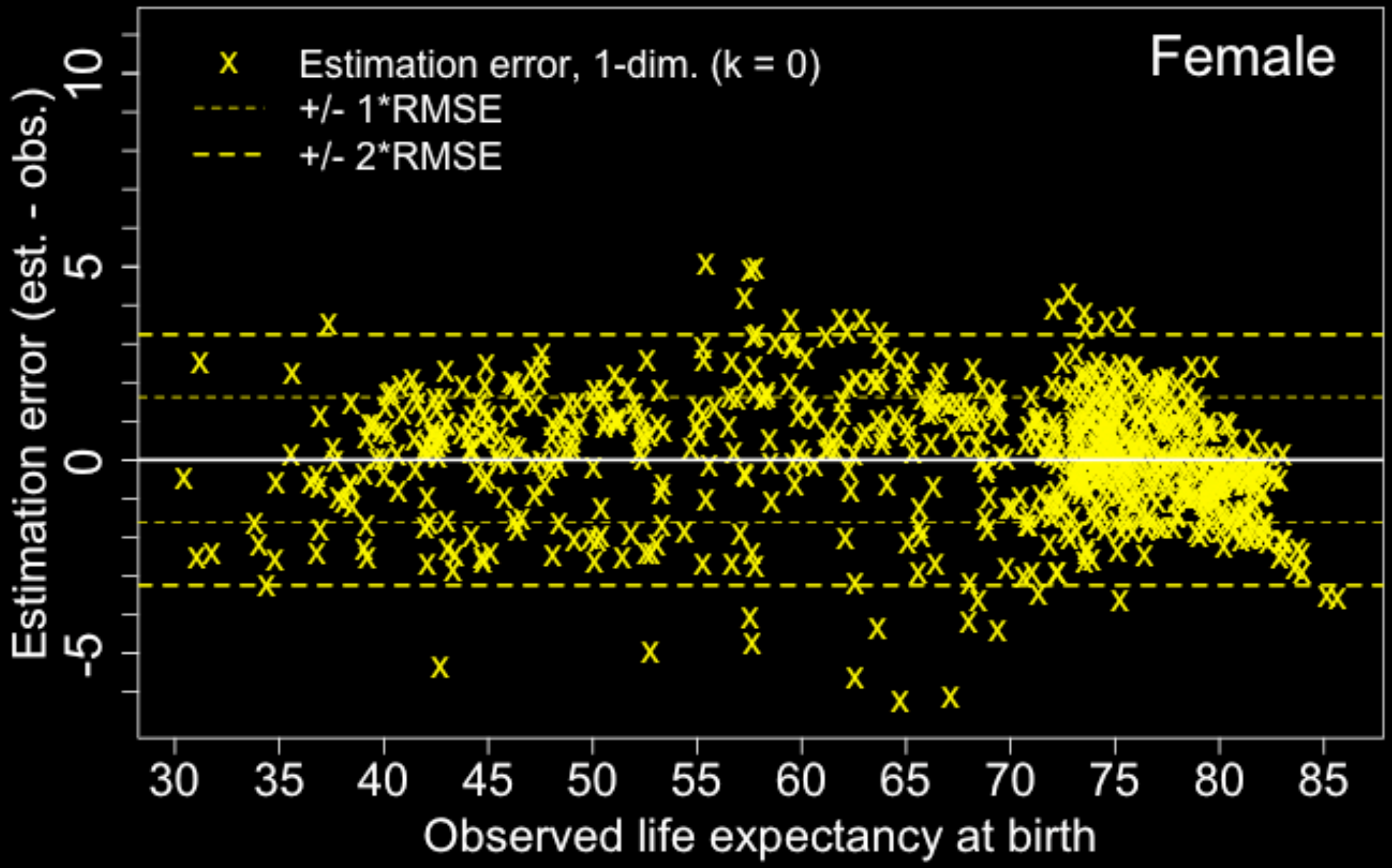


Typical Estimation Errors: Some Questions

- How big are in-sample estimation errors?
- When using model life tables, how much uncertainty is associated with estimates of e_0 (or other summary measures)?
- The in-sample estimation errors provide lower-bound estimates of uncertainty

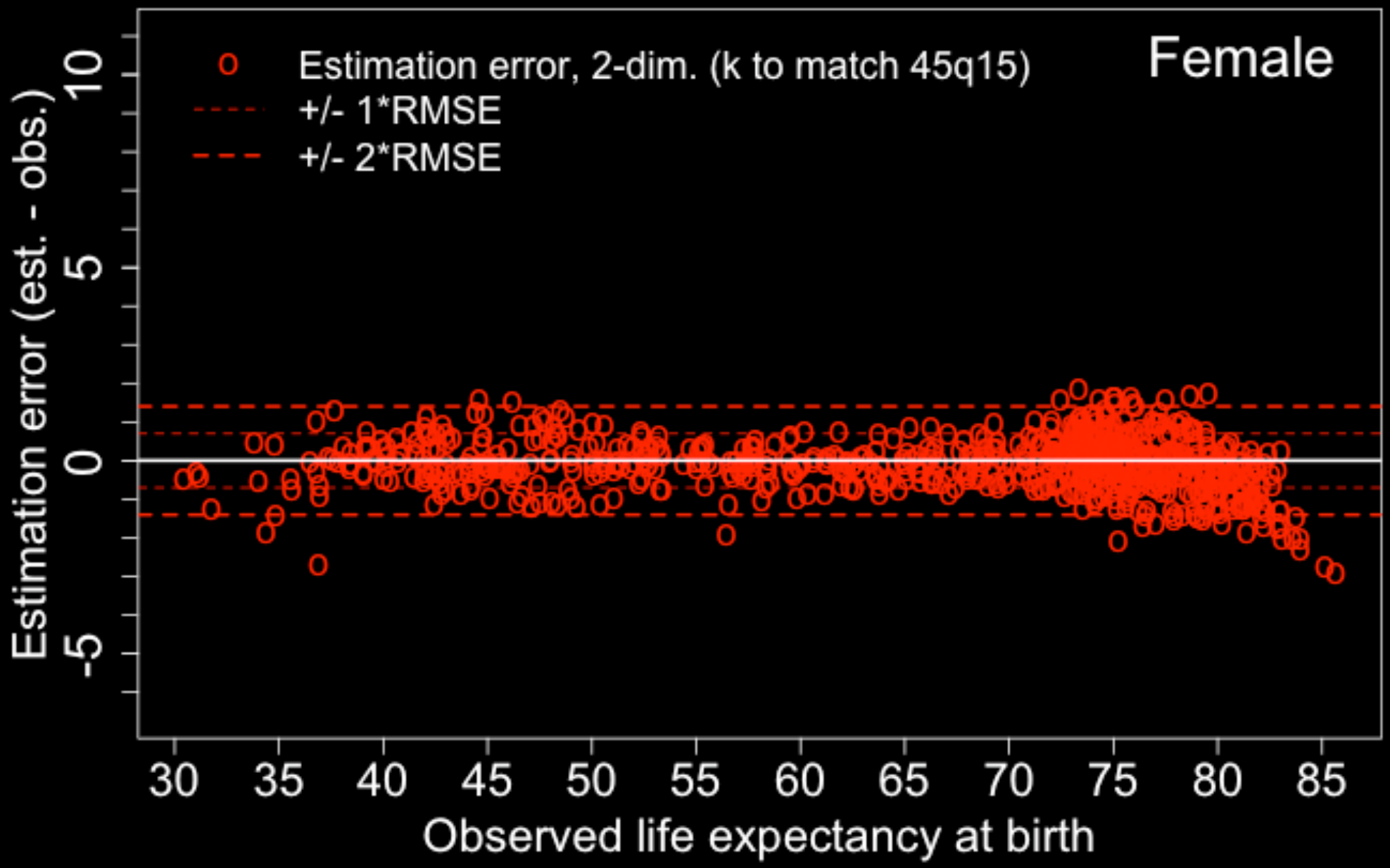
Estimation Errors in Female e_0

1-parameter model: ± 1 or 2 RMSEs



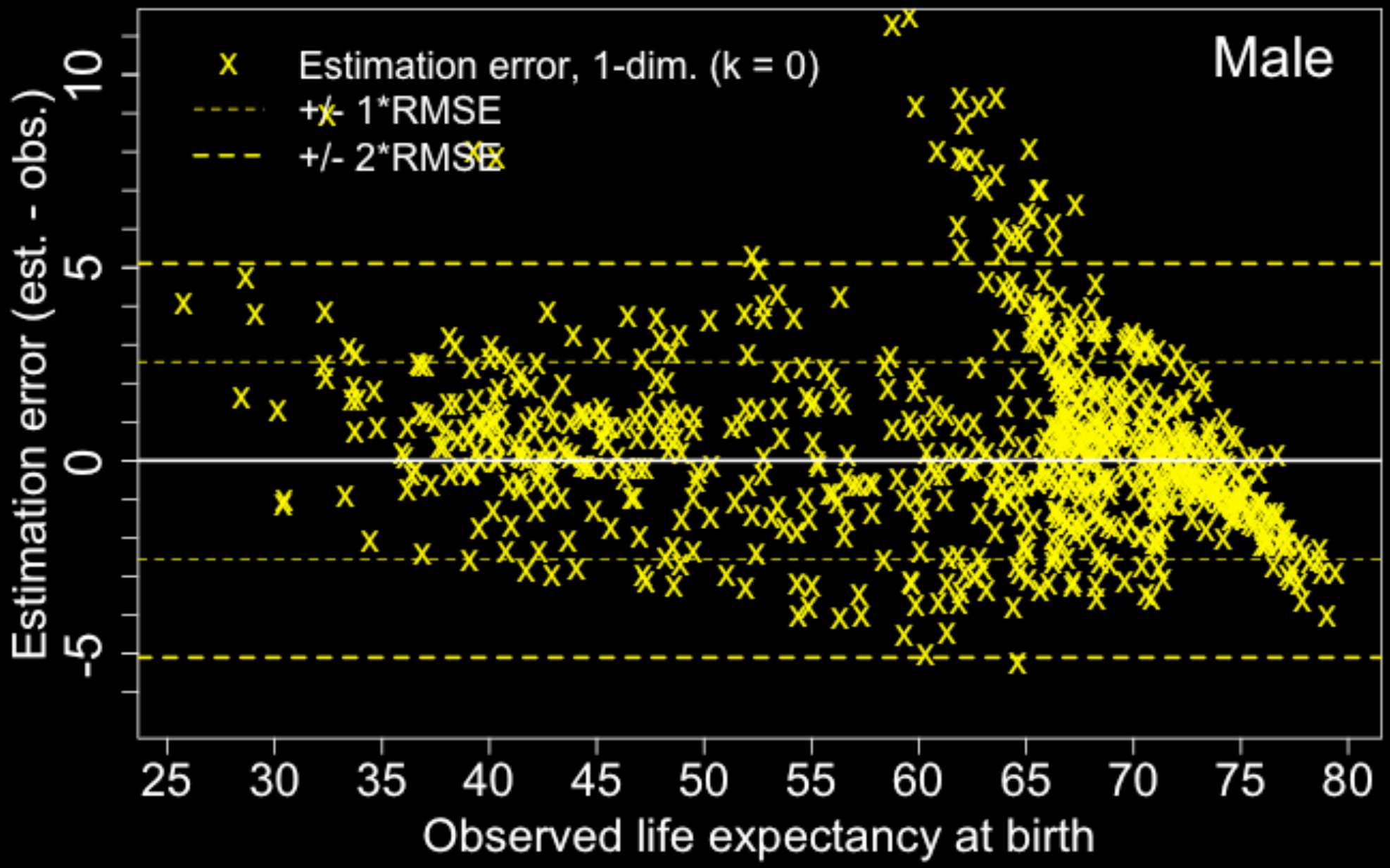
Estimation Errors in Female e_0

2-parameter model: ± 1 or 2 RMSEs



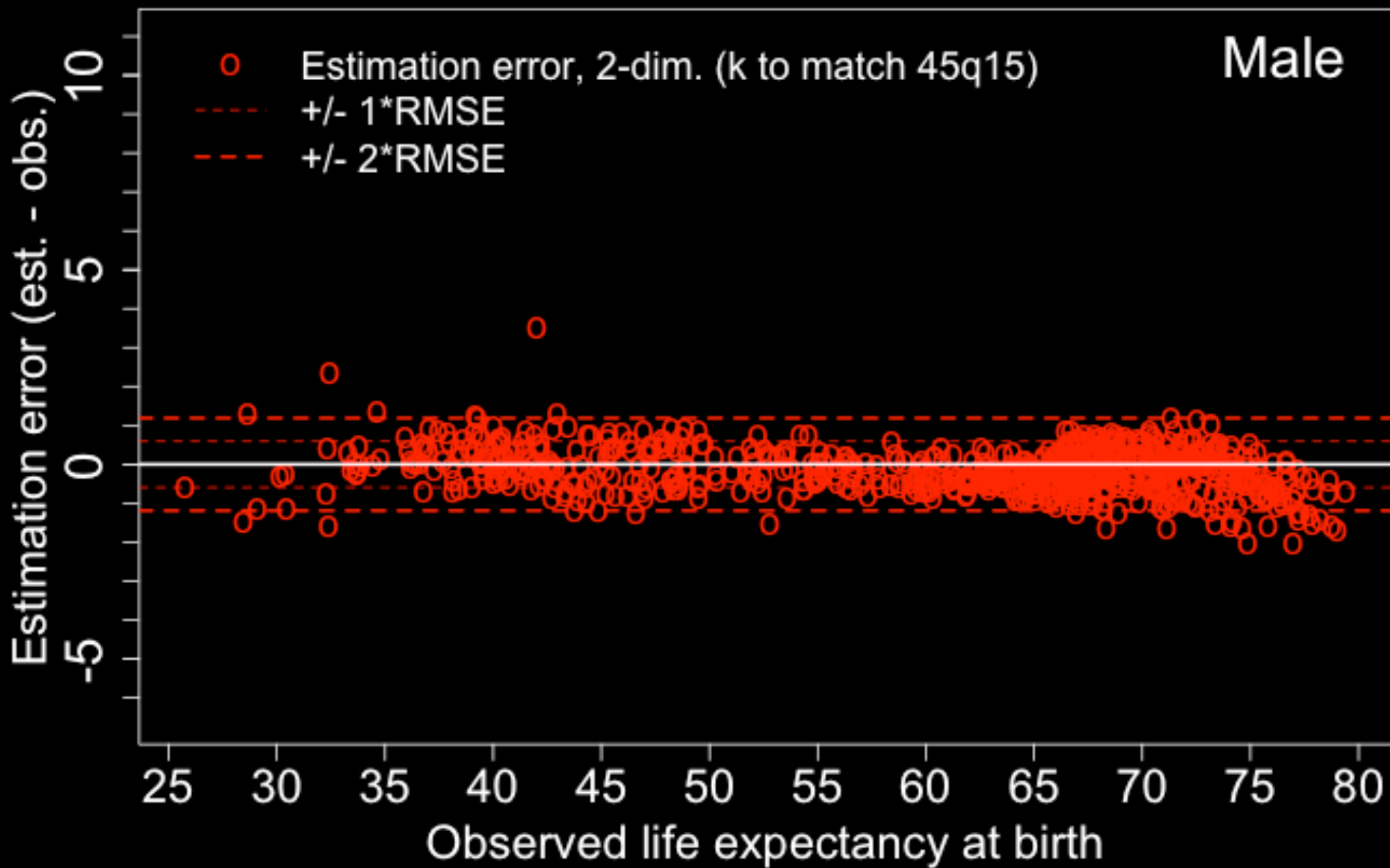
Estimation Errors in Male e_0

1-parameter model: ± 1 or 2 RMSEs



Estimation Errors in Male e_0

2-parameter model: ± 1 or 2 RMSEs



Typical Estimation Errors: Some Answers

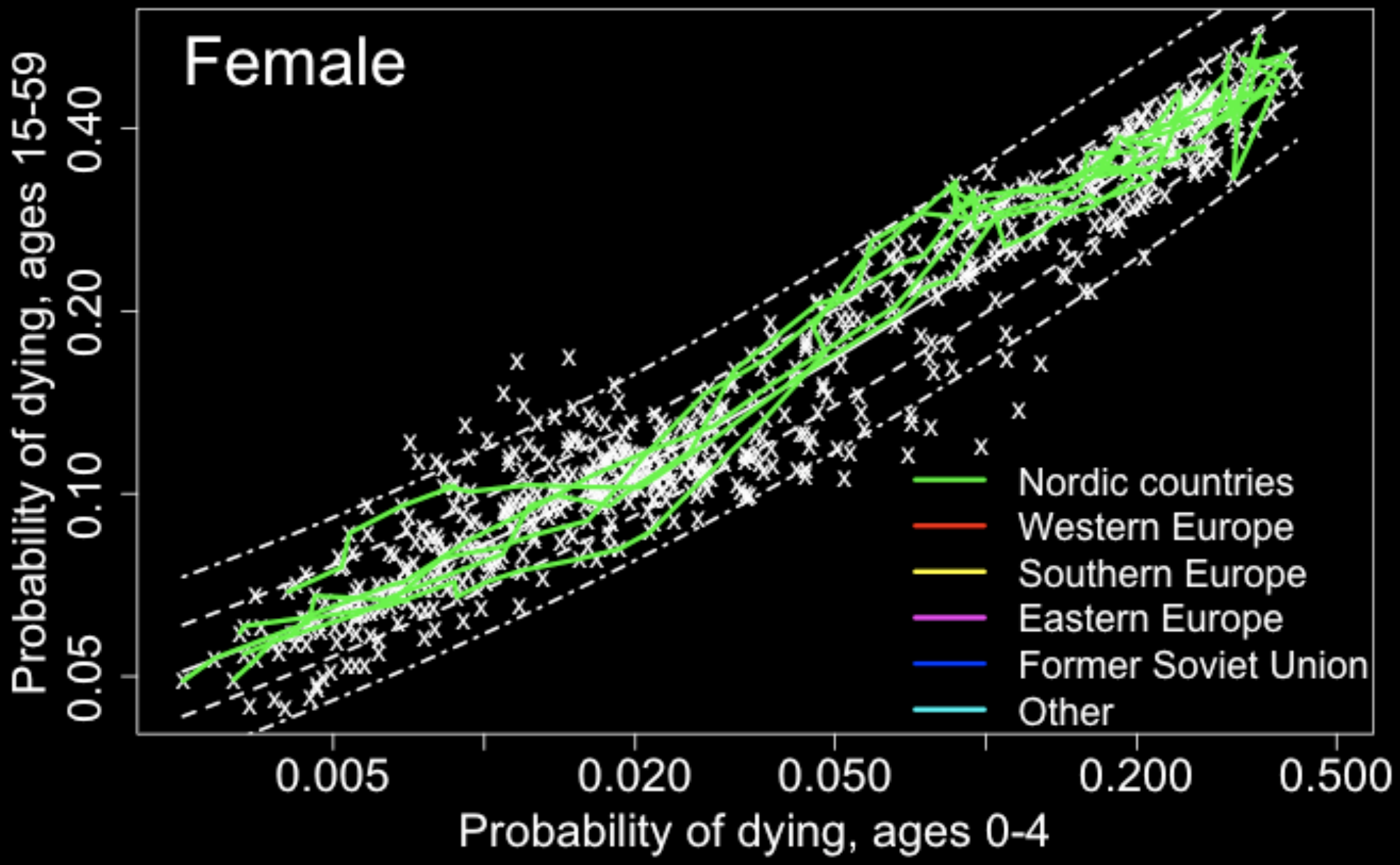
- When we know both ${}_{5}q_{0}$ and ${}_{45}q_{15}$ accurately, estimation errors in e_0 are around 1-1.5 years for women and men
- When we know only ${}_{5}q_{0}$, estimation errors in e_0 are roughly ± 3 years for women and ± 5 years for men
- Remember: these are lower bounds

Changes in Age Patterns of Mortality by Region

- Most countries included in the HMD conform closely to the log-quadratric model
- Three regions have distinctive patterns:
  - Southern Europe, 1960s and earlier
  - Former Soviet Union, 1970s and later
  - Eastern Europe, mostly in 1970s and 1980s

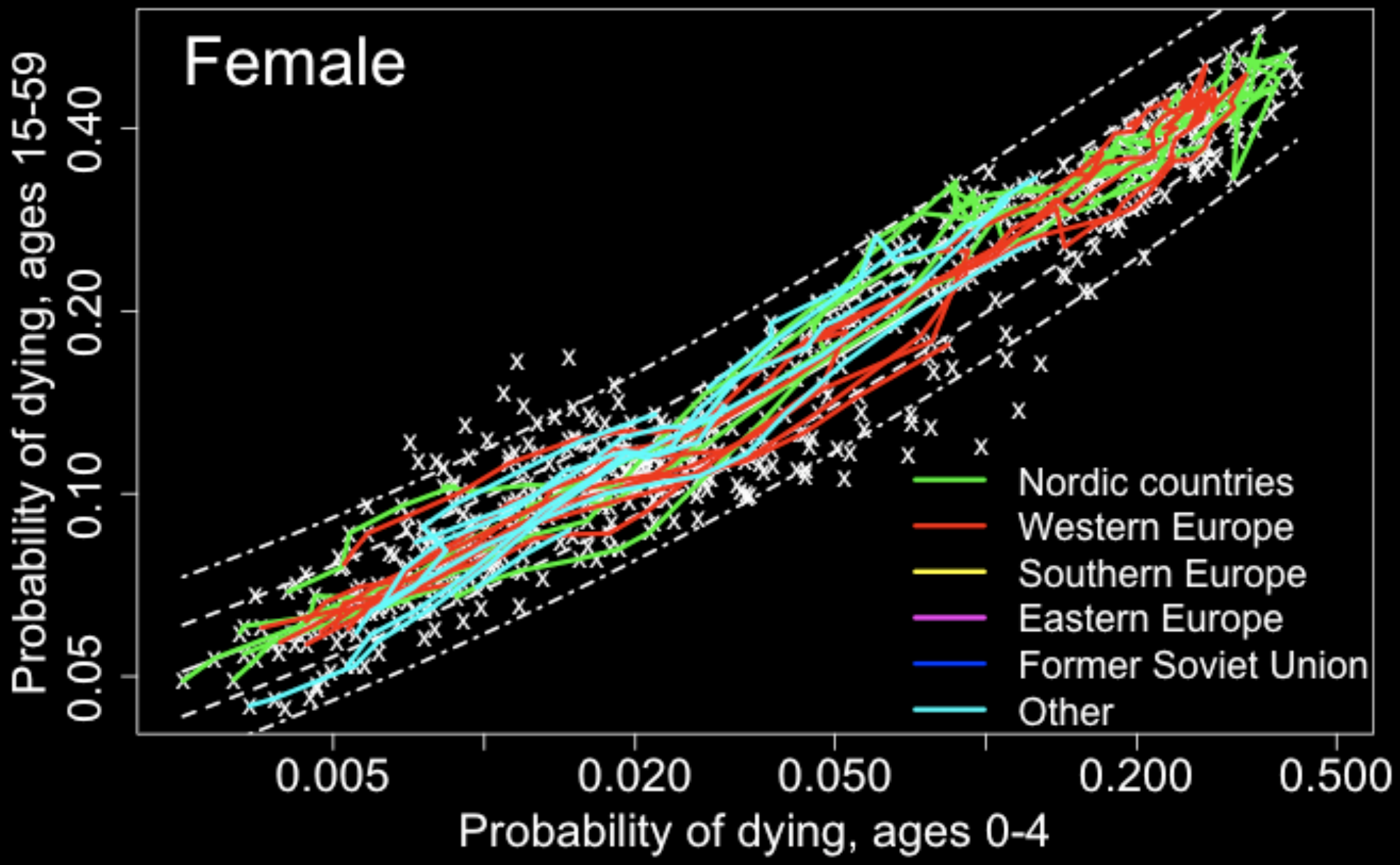
Country Trends + Log-Quadratic Model

Nordic countries only



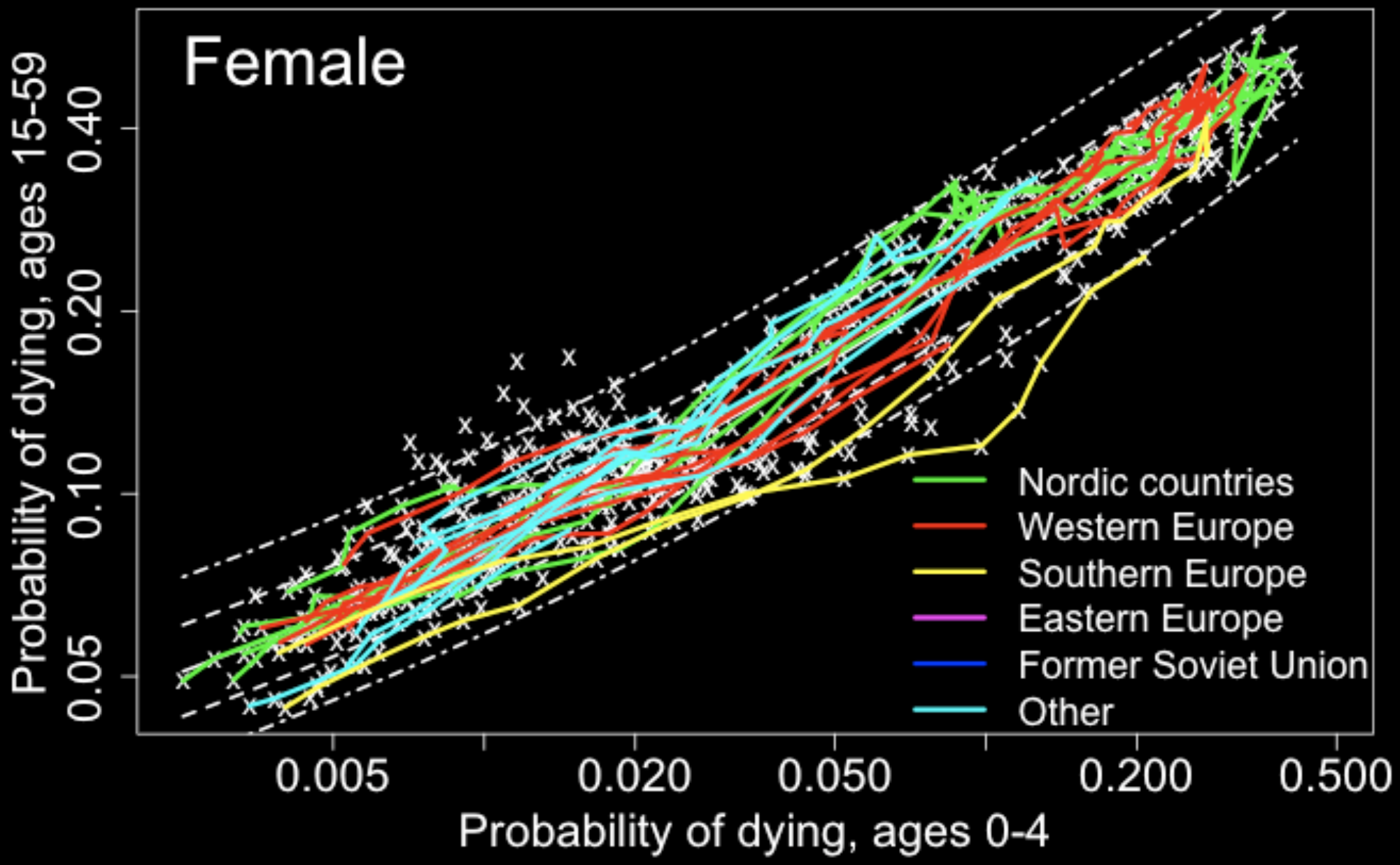
Country Trends + Log-Quadratic Model

Add Western Europe and Other



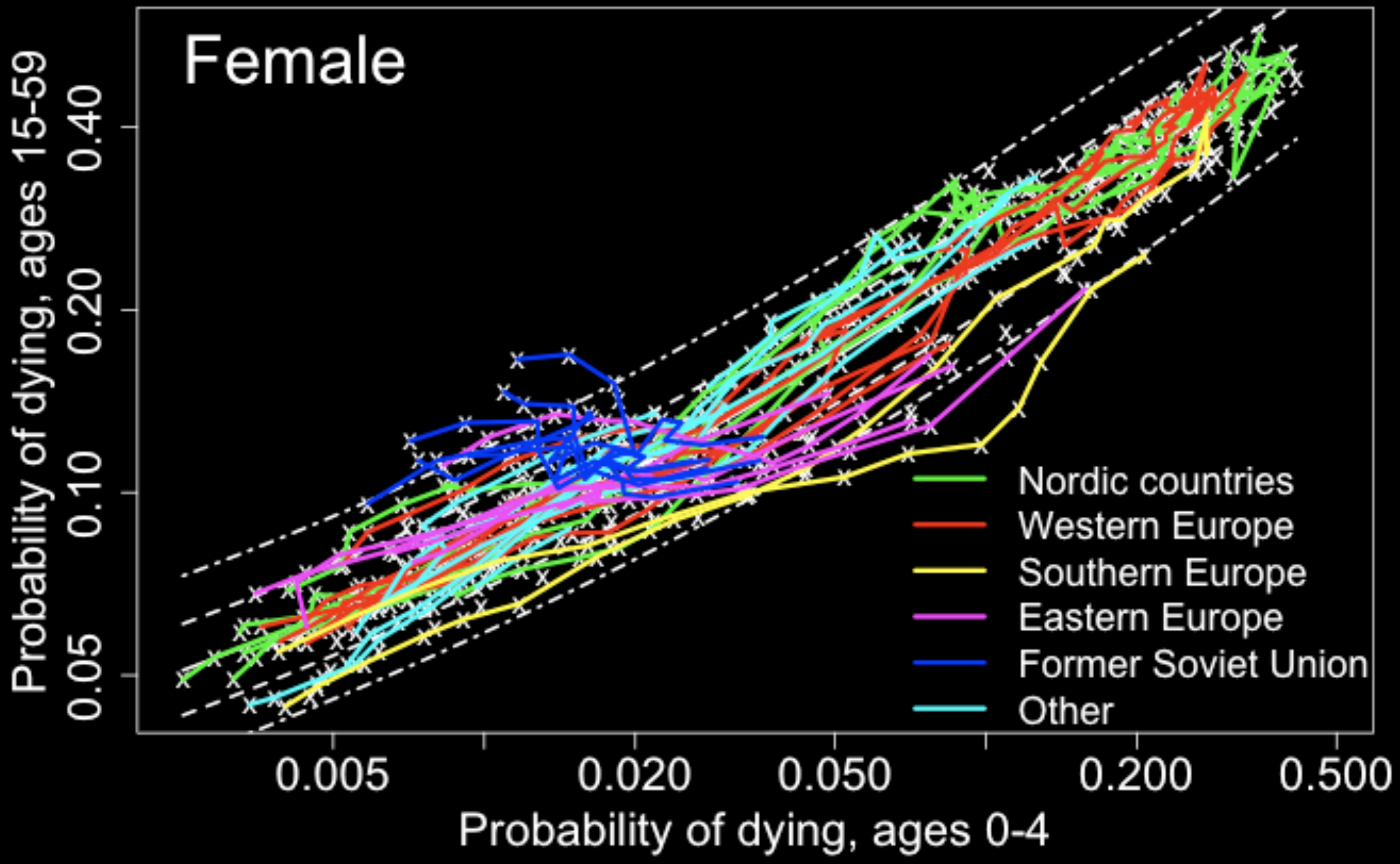
Country Trends + Log-Quadratic Model

Add Southern Europe



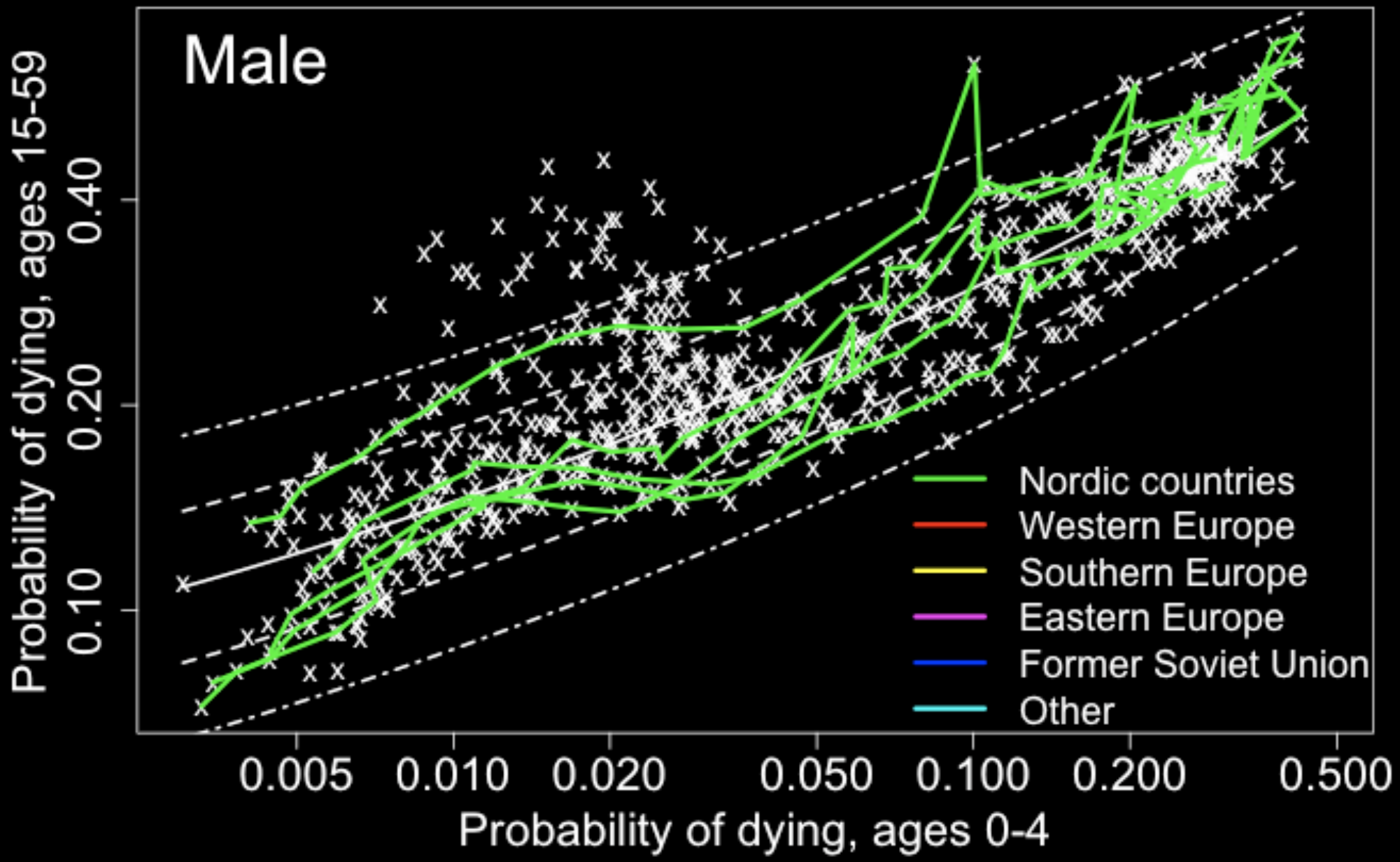
Country Trends + Log-Quadratic Model

Add Eastern Europe and former USSR



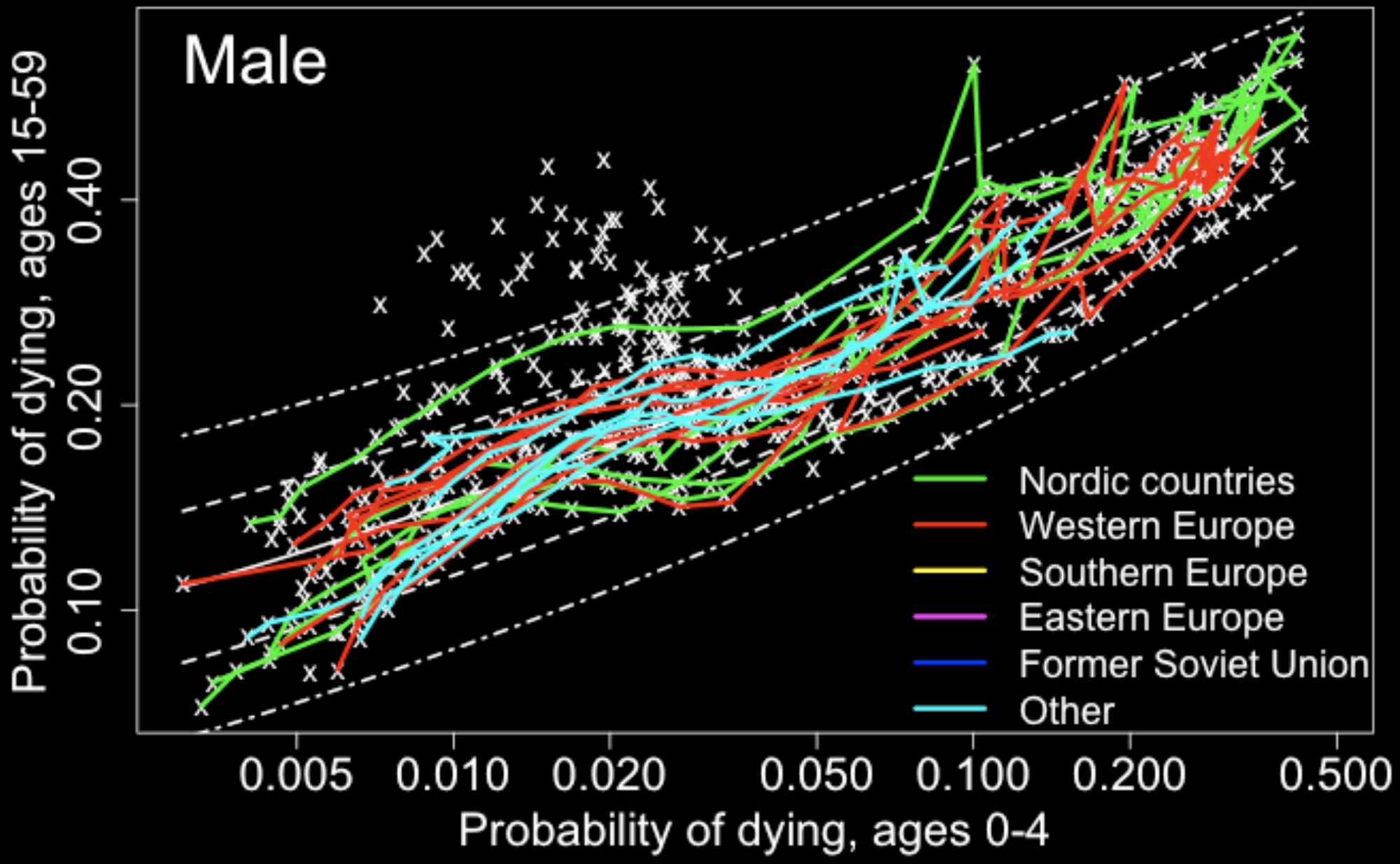
Country Trends + Log-Quadratic Model

Nordic countries only



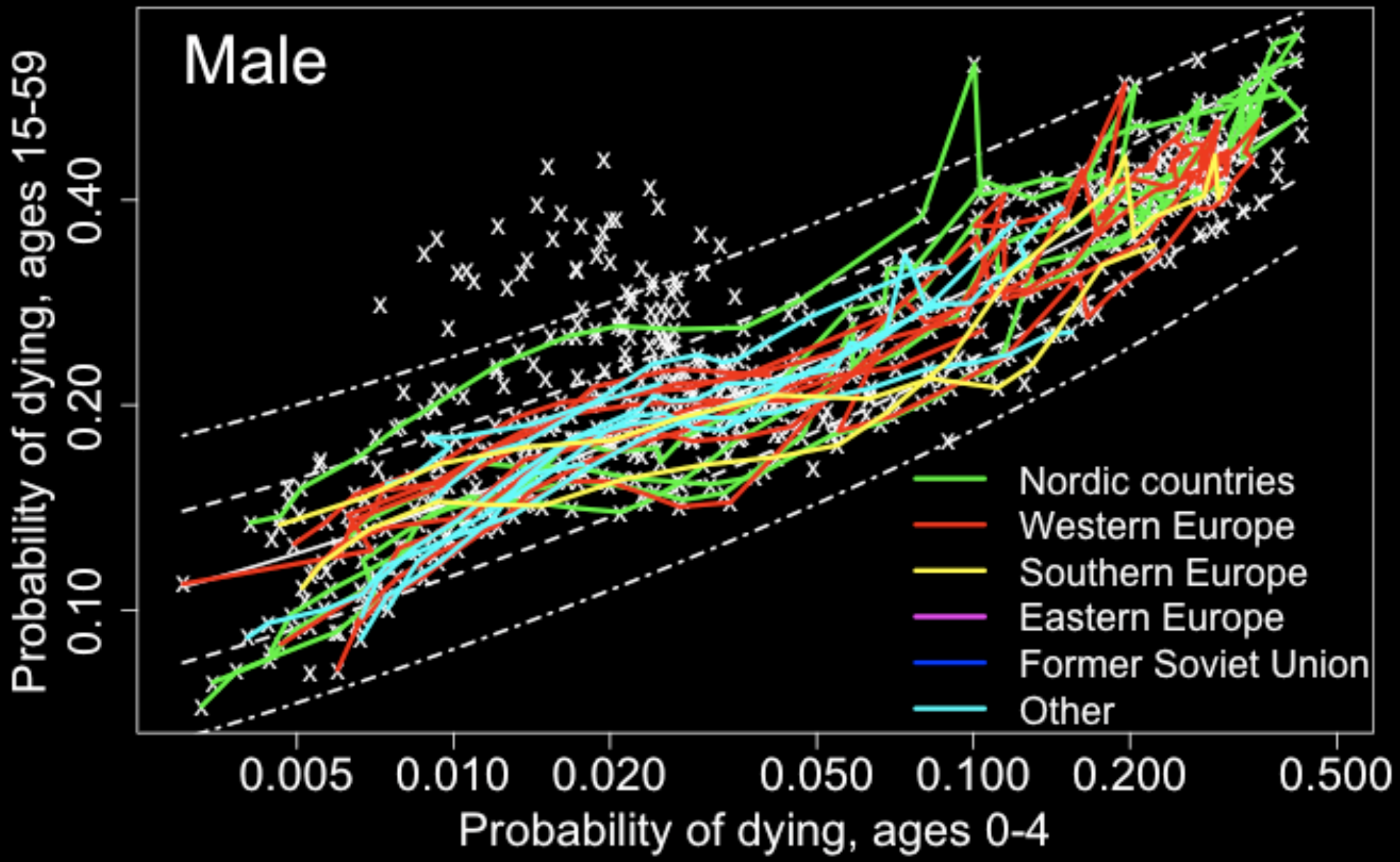
Country Trends + Log-Quadratic Model

Add Western Europe and Other



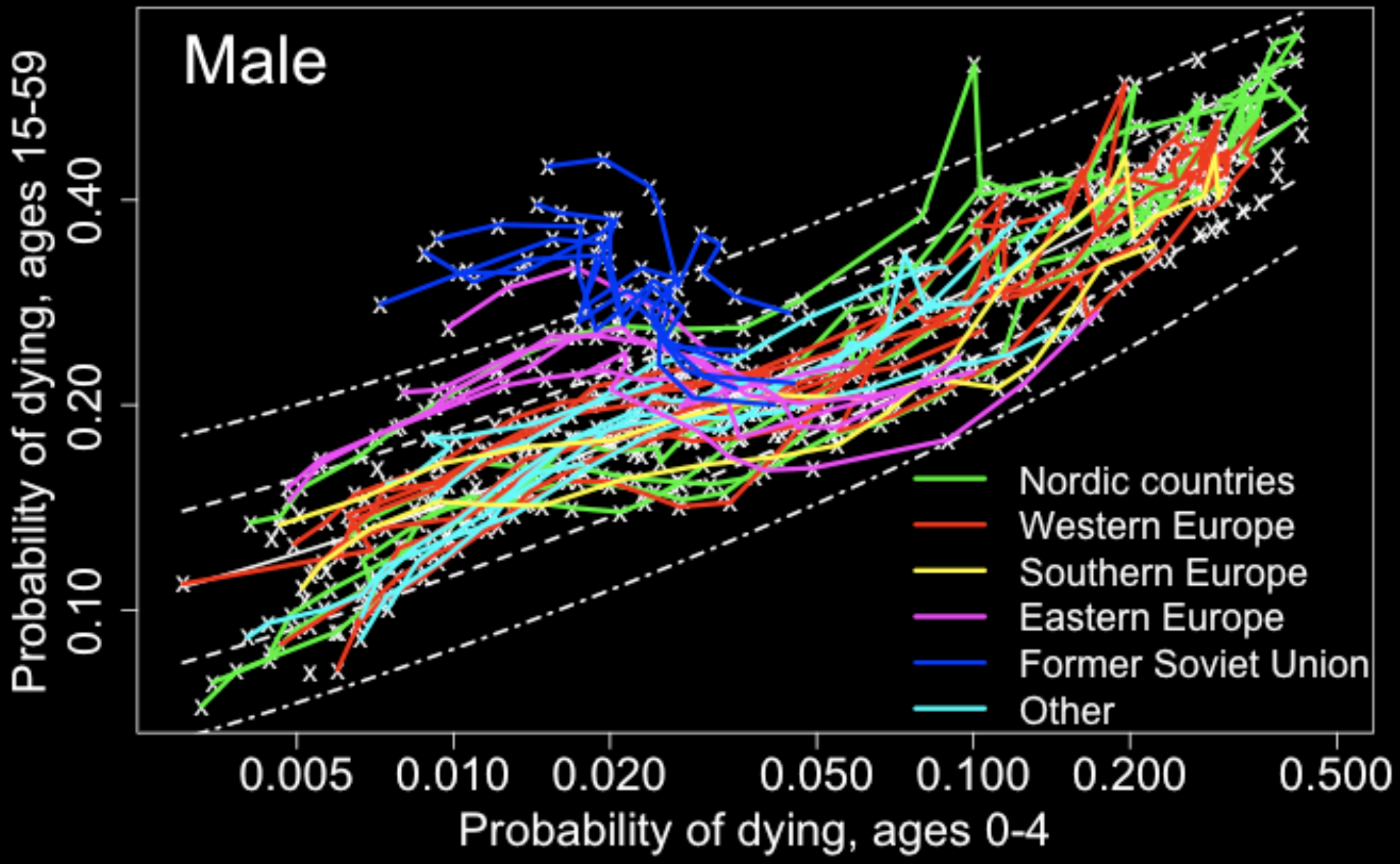
Country Trends + Log-Quadratic Model

Add Southern Europe



Country Trends + Log-Quadratic Model

Add Eastern Europe and former USSR



Interesting Example: Costa Rica

Historical Exception or Bad Data?

- Article in *Demography* by Rosero-Bixby
- Comparison to historical data

THE EXCEPTIONALLY HIGH LIFE EXPECTANCY OF COSTA RICAN NONAGENARIANS*

LUIS ROSERO-BIXBY

Robust data from a voter registry show that Costa Rican nonagenarians have an exceptionally high live expectancy. Mortality at age 90 in Costa Rica is at least 14% lower than an average of 13 high-income countries. This advantage increases with age by 1% per year. Males have an additional 12% advantage. Age-90 life expectancy for males is 4.4 years, one-half year more than any other country in the world. These estimates do not use problematic data on reported ages, but ages are computed from birth dates in the Costa Rican birth-registration ledgers. Census data confirm the exceptionally high survival of elderly Costa Ricans, especially males. Comparisons with the United States and Sweden show that the Costa Rican advantage comes mostly from reduced incidence of cardiovascular diseases, coupled with a low prevalence of obesity, as the only available explanatory risk factor. Costa Rican nonagenarians are survivors of cohorts that underwent extremely harsh health conditions when young, and their advantage might be just a heterogeneity in frailty effect that might disappear in more recent cohorts. The availability of reliable estimates for the oldest-old in low-income populations is extremely rare. These results may enlighten the debate over how harsh early-life health conditions affect older-age mortality.

Two key findings have emerged from recent studies of old-age mortality in humans (Vaupel et al. 1998): (1) mortality rates are declining substantially, and (2) the increase of death rates with age decelerates among the oldest-old. In the words of Vaupel et al. (1998), these findings are perplexing and hard to reconcile: according to evolutionary biology, there is no possible selection against mutations occurring after reproduction and nurturing have ceased. A possible explanation of the old-age deceleration is *heterogeneity in frailty*; that is, as the frail die at early ages, the old tend to be a select subpopulation of the fittest (Barbi, Caselli, and Vallin 2003; Horiuchi and Wilmoth 1998; Vaupel et al. 1998). In turn, a possible explanation of the mortality decline at old ages is a cohort effect of past improvements in health conditions at early ages; that is, recent improvements in health status among the elderly would echo events that happened decades ago when cohorts were young. These two explanations are somehow contradictory: does high, early-life mortality make a cohort stronger by eliminating the frail, or does the cohort become weaker because of accumulated injuries? An important scientific debate is taking place in this regard (Barbi and Vaupel 2005; Finch and Crimmins 2004, 2005).

The *heterogeneity in frailty* argument has been mostly supported by mathematical and simulation models (Vaupel, Manton, and Stallard 1979); by indirect evidence from genetic homogeneous populations such as twins (Yashin and Iachine 1997); and by observations in other species, such as the Mediterranean fruit fly (Vaupel and Carey 1993). Indirect methods have been developed to determine the existence of heterogeneity from cohort mortality patterns (Manton, Stallard, and Vaupel 1981). Data showing low death rates at old ages in low-income populations that saw harsh health conditions at young ages might support the *heterogeneity in frailty* argument, given the prejudice that the poor cannot be

*Luis Rosero-Bixby, Centro Centroamericano de Población, Universidad de Costa Rica, Apartado 2060, San Jose, Costa Rica; e-mail: Lrosero@ccp.ucr.ac.cr. The Wellcome Trust Foundation (Grant No. 072406/Z/03/Z) and the Florida Ice and Farm Co. of Costa Rica provided support for this and other studies on aging in Costa Rica. The Costa Rican *Tribunal Supremo de Elecciones* provided the databases. Daniel Antich from the Universidad de Costa Rica provided assistance in processing the databases. Albert I. Hermalin from the University of Michigan provided suggestions and encouragement to improve this manuscript. German Rodríguez from Princeton University provided statistical advice to estimate the model.

Luis Rosero-Bixby

Demography
August 2008

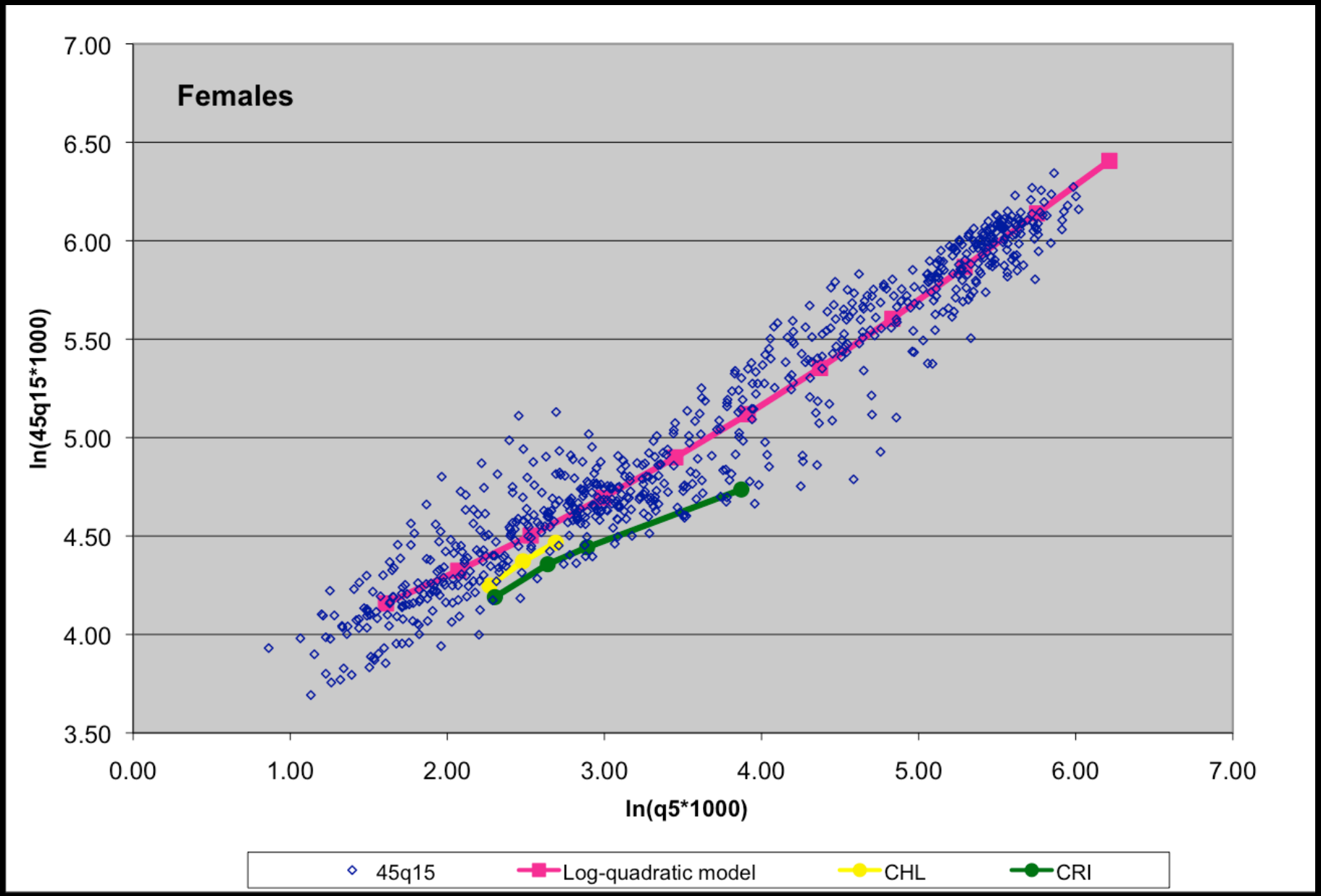
45(1): 673-691

Rosero-Bixby's Main Points

- Age data obtained from voter registry, in theory not affected by age reporting errors
- His estimates of mortality above age 90 match closely those from vital statistics
- Calculated death rates above 90 are the lowest in the world: at least 14% lower than average of 13 high-income countries

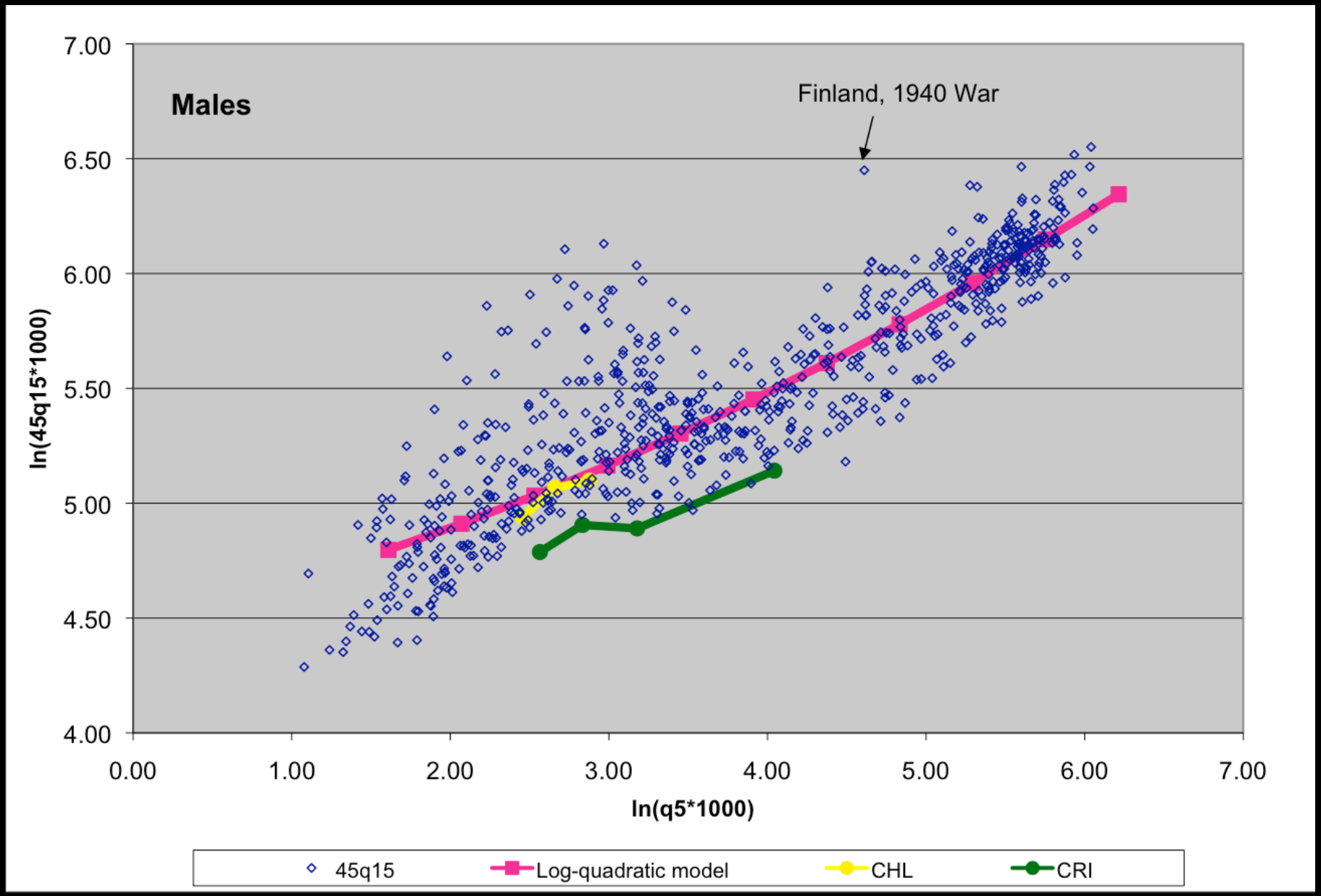
Adult ($_{45}q_{15}$) vs. Child ($_5q_0$) Mortality

Costa Rica + HMD + Log-Quadratic



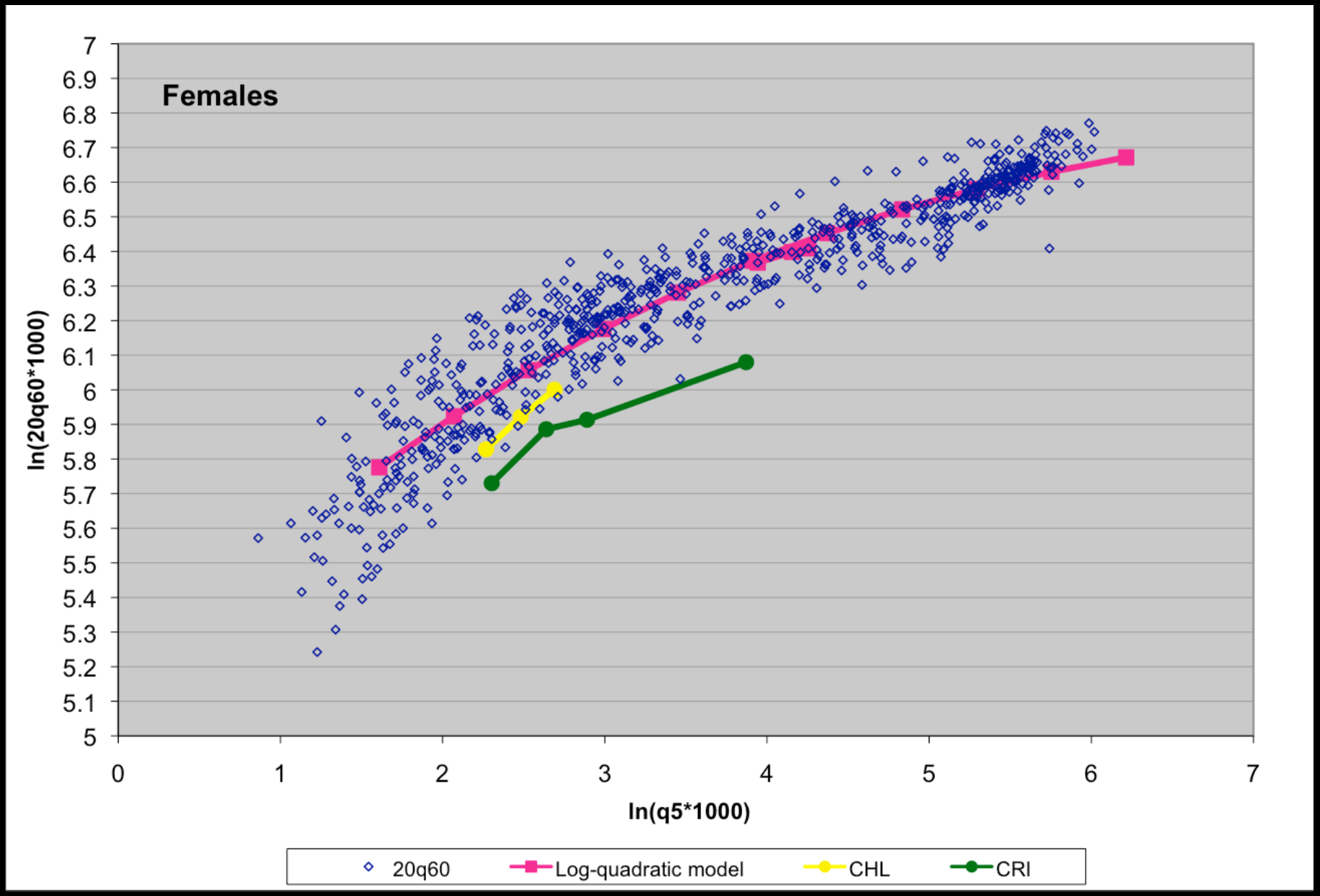
Adult ($_{45}q_{15}$) vs. Child ($_5q_0$) Mortality

Costa Rica + HMD + Log-Quadratic



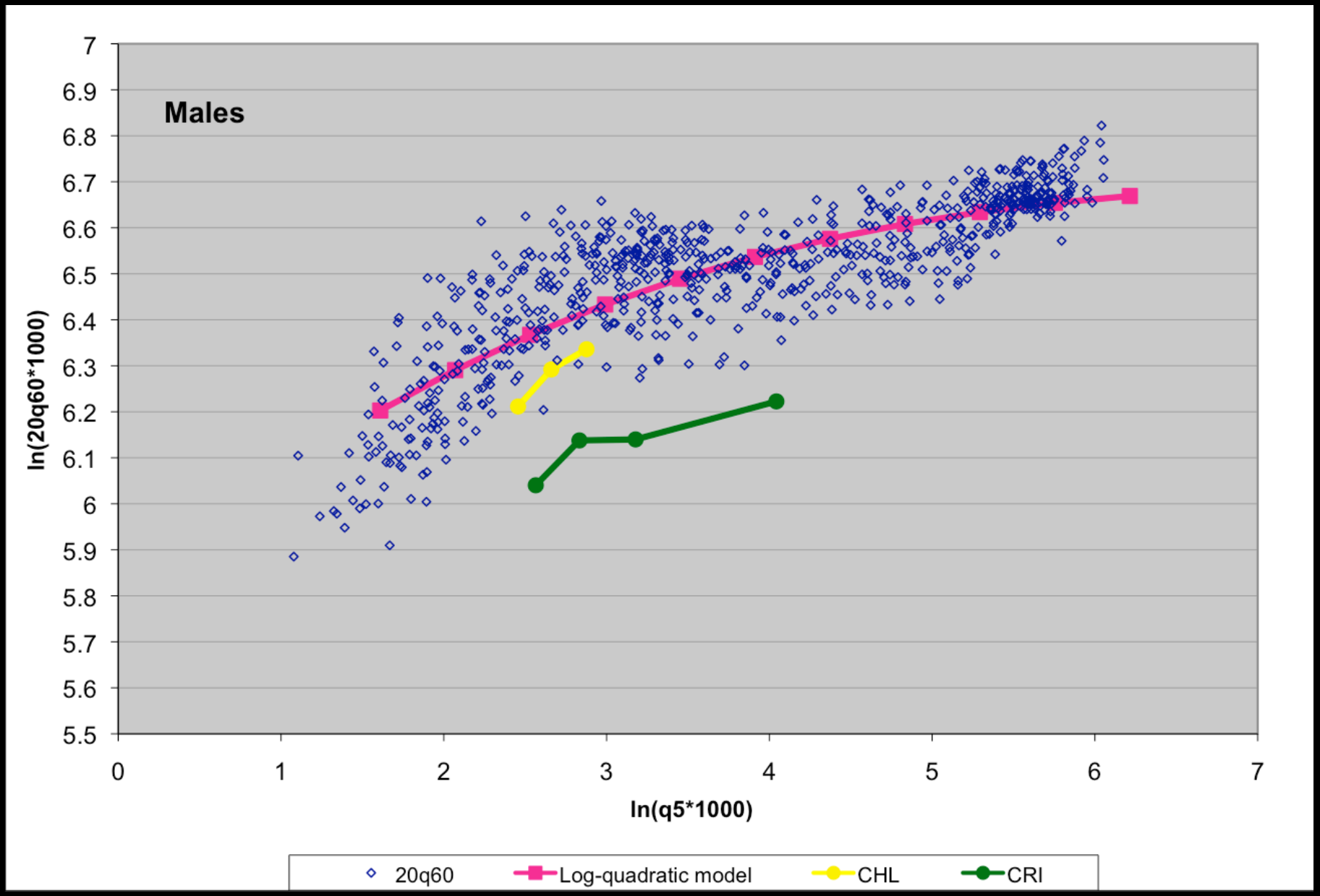
Older adult ($_{20}q_{60}$) vs. Child ($_5q_0$) Mortality

Costa Rica + HMD + Log-Quadratic



Older adult ($_{20}q_{60}$) vs. Child ($_5q_0$) Mortality

Costa Rica + HMD + Log-Quadratic



Main Conclusions

- Coale-Demeny and UN model life table systems yield biased estimates at low or even moderately low levels of mortality
- Modified logit model yields more accurate results but is somewhat awkward to apply and unnecessarily complicated
- Log-quadratic model is as reliable as any alternative method and offers significant advantages in terms of flexibility and simplicity

Main Conclusions (cont.)

- Estimation uncertainty for values of e_0 derived from child mortality is *at least* ± 3 years for women and ± 5 years for men
- Some regions have distinctive patterns in certain time periods, but these can be well represented by a single model (no need for discrete "families" of model life tables)
- Limitations: war mortality, AIDS?
- Estimates for Costa Rica are implausible

The End