# Modelling the Age Patterns of Demographic Events

Guy Abel

ESRC Centre for Population Change,
University of Southampton, United Kingdom

6th July 2010

# Outline

1. Motivation for using model schedules and examples.
2. Mathematical Models.
3. Relational Models.
4. Model Life Tables.
5. Schedules over time.

# Motivation

- Intensity of demographic events varies sharply with age.
- Large sets of data can be cumbersome or unavailable.
- Demographers have searched for more parsimonious representations (a.k.a. model schedules) of age variation.
- Model schedules allow:
    1. Identification of irregularities in the data.
    2. Simplify assumptions for population projections.
    3. Estimation of incomplete data.
    4. Explanation of behavioral patterns of populations.

# Model Schedules

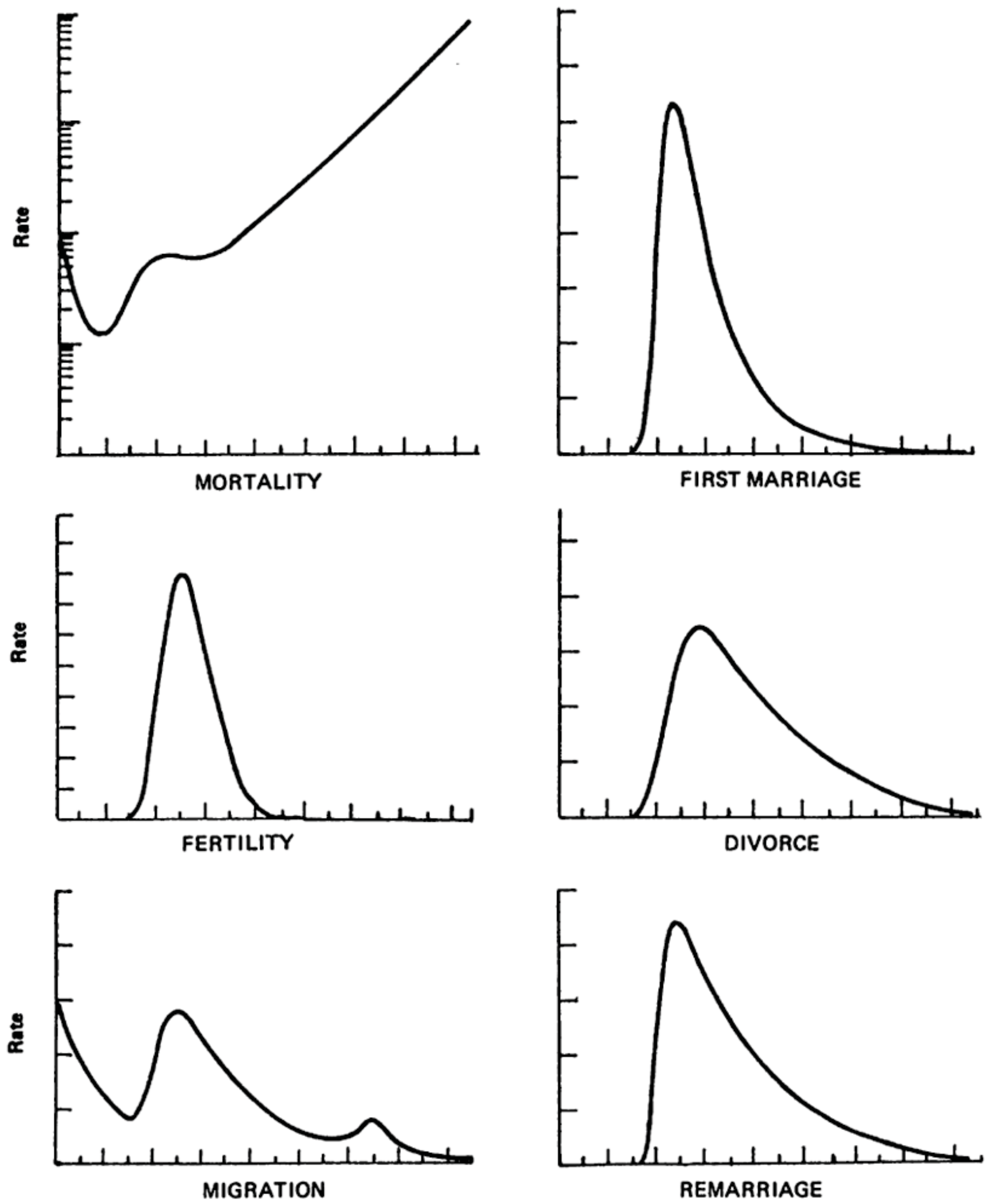
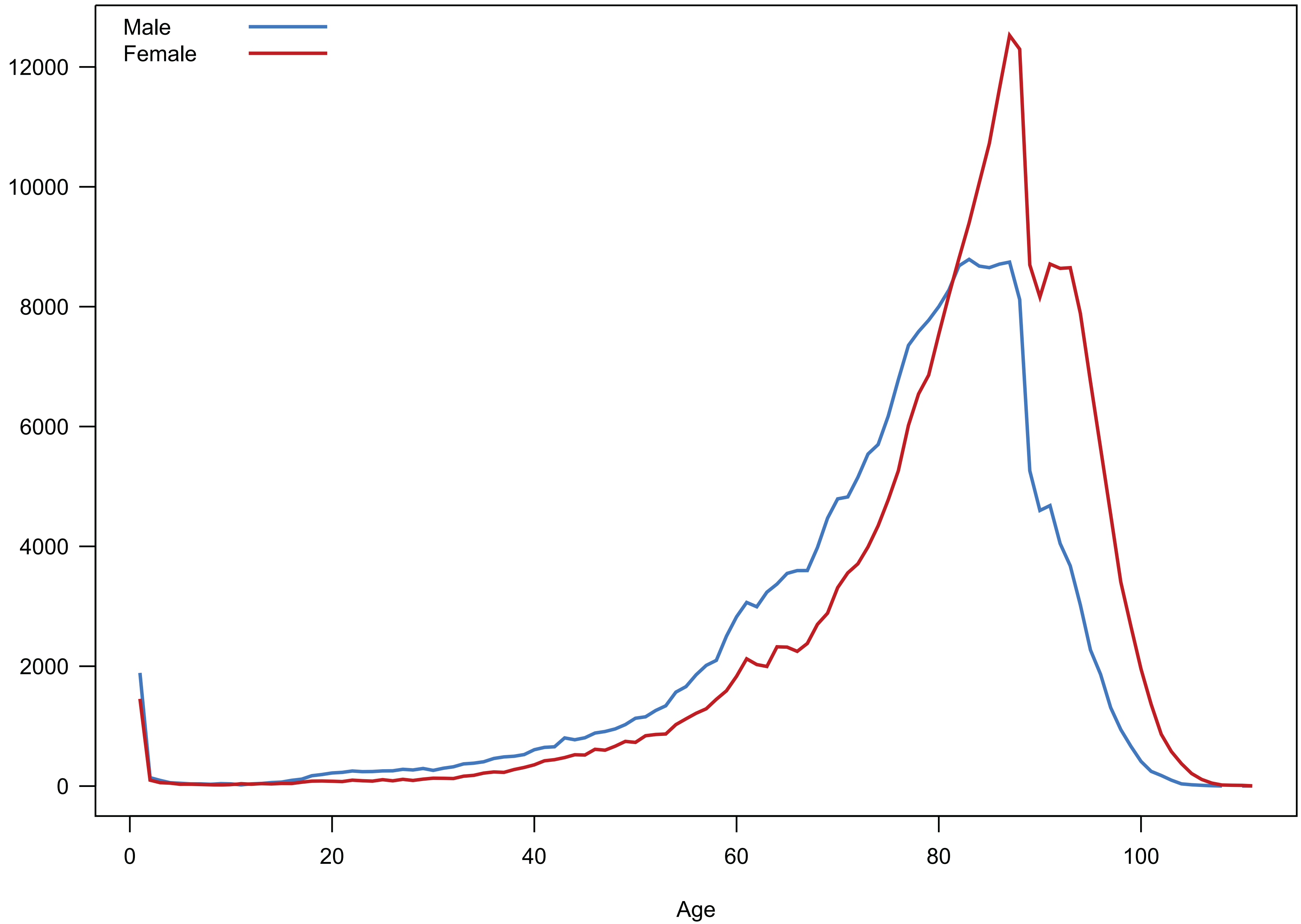
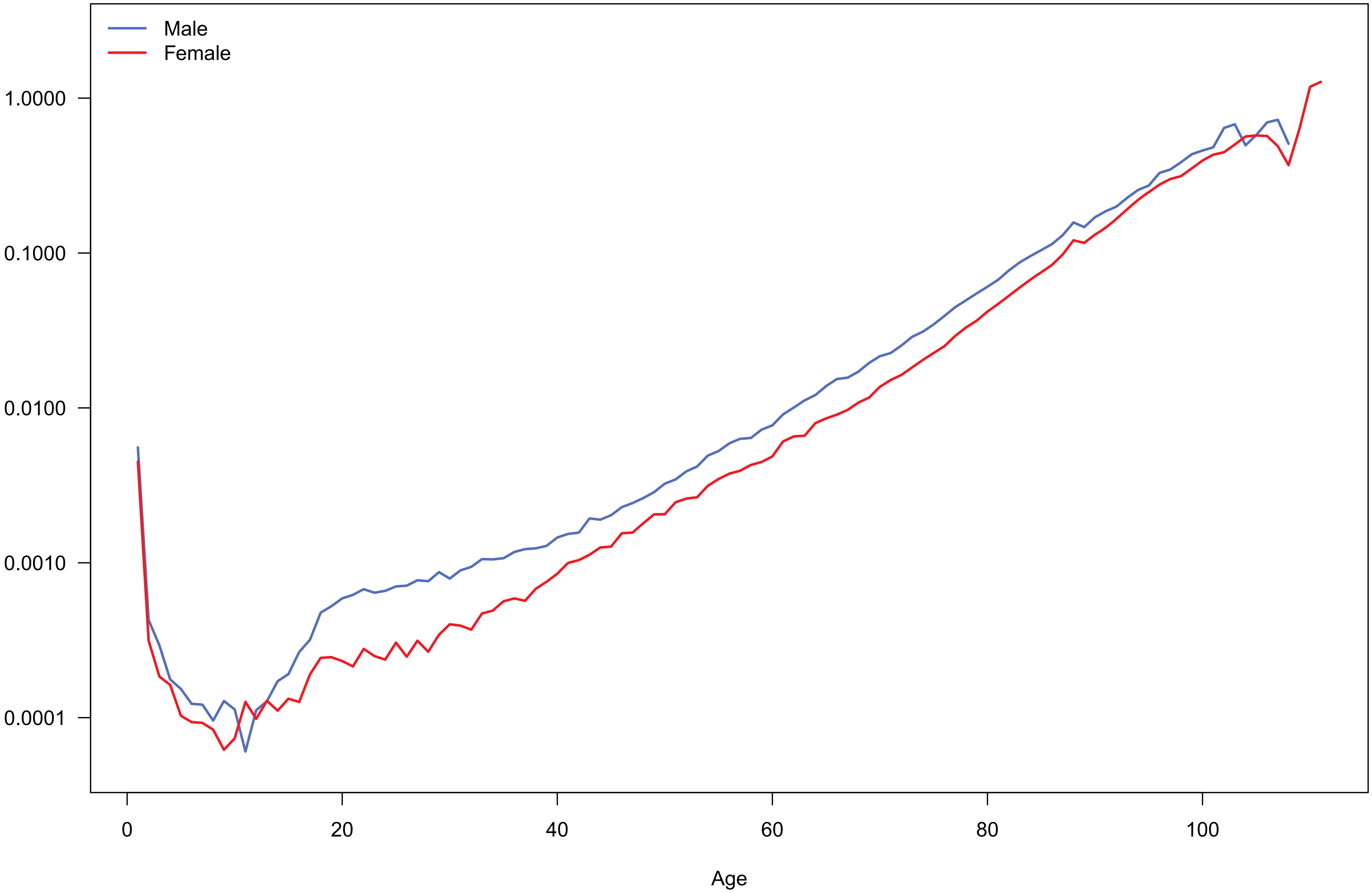



*Figure 1. Multistate Schedules. Source: Rogers (1982).*

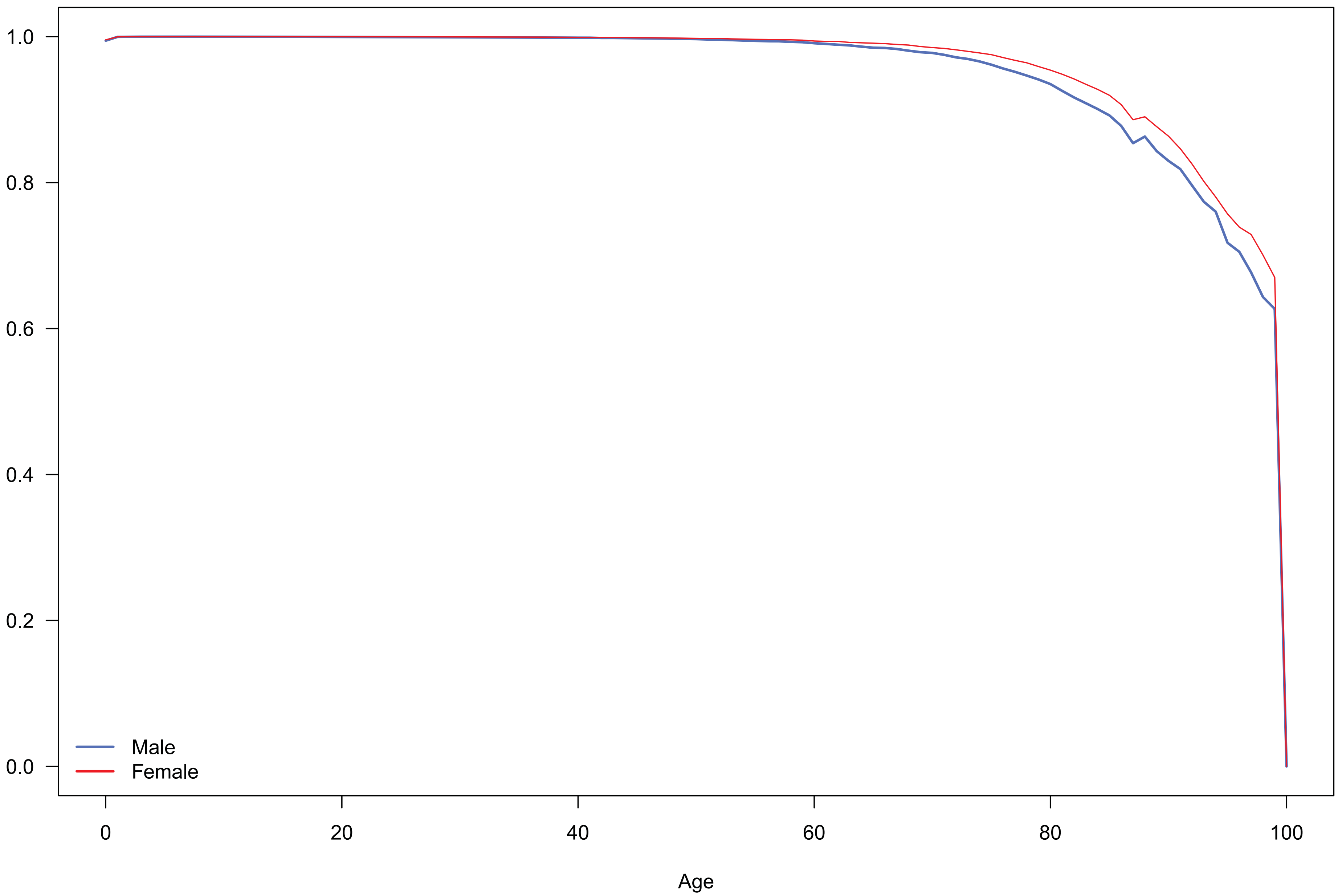
# Number of Deaths ($D_x$) in England and Wales, 2007



# Mortality Rate ($m_x$) in England and Wales, 2007



# Survivorship ($l_x$) for England and Wales, 2007



# Gompertz

- First demographic mathematical models developed for mortality.
- Mortality functions link the age-specific probability of death, $m_x$ to age $(x)$ using a function, $f(x)$
- Gompertz (1825) suggested that a *law of geometric progression pervades* in mortality after a certain age.

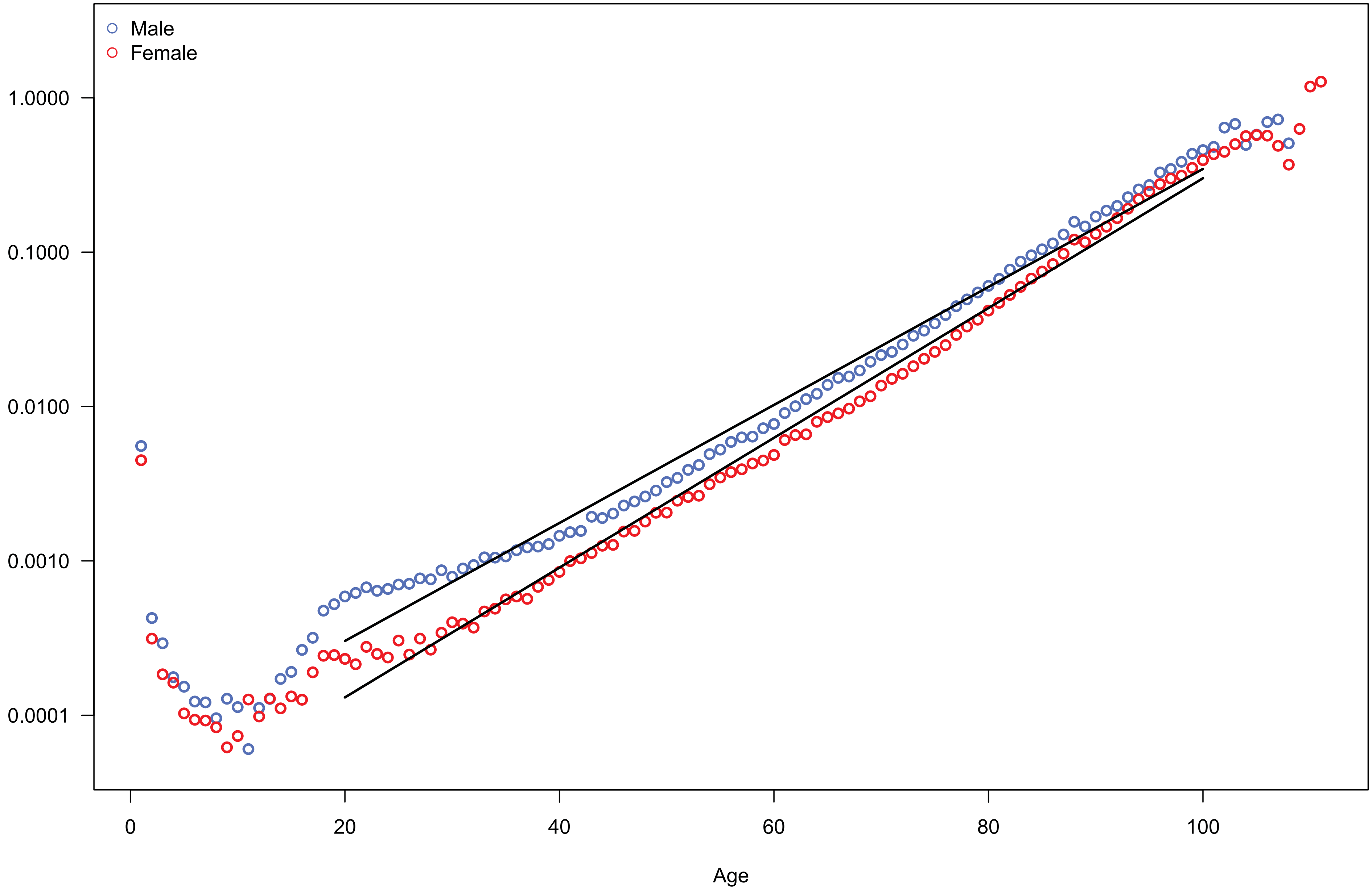
$$f(x) = \alpha e^{\beta x} \tag{1}$$

- For log-mortality this is a linear function of age

$$\log f(x) = \log(\alpha) + \beta x \tag{2}$$

Can use standard regression approach to calculate $\alpha$ and $\beta$

# Fits of Gompertz Models for England and Wales, 2007

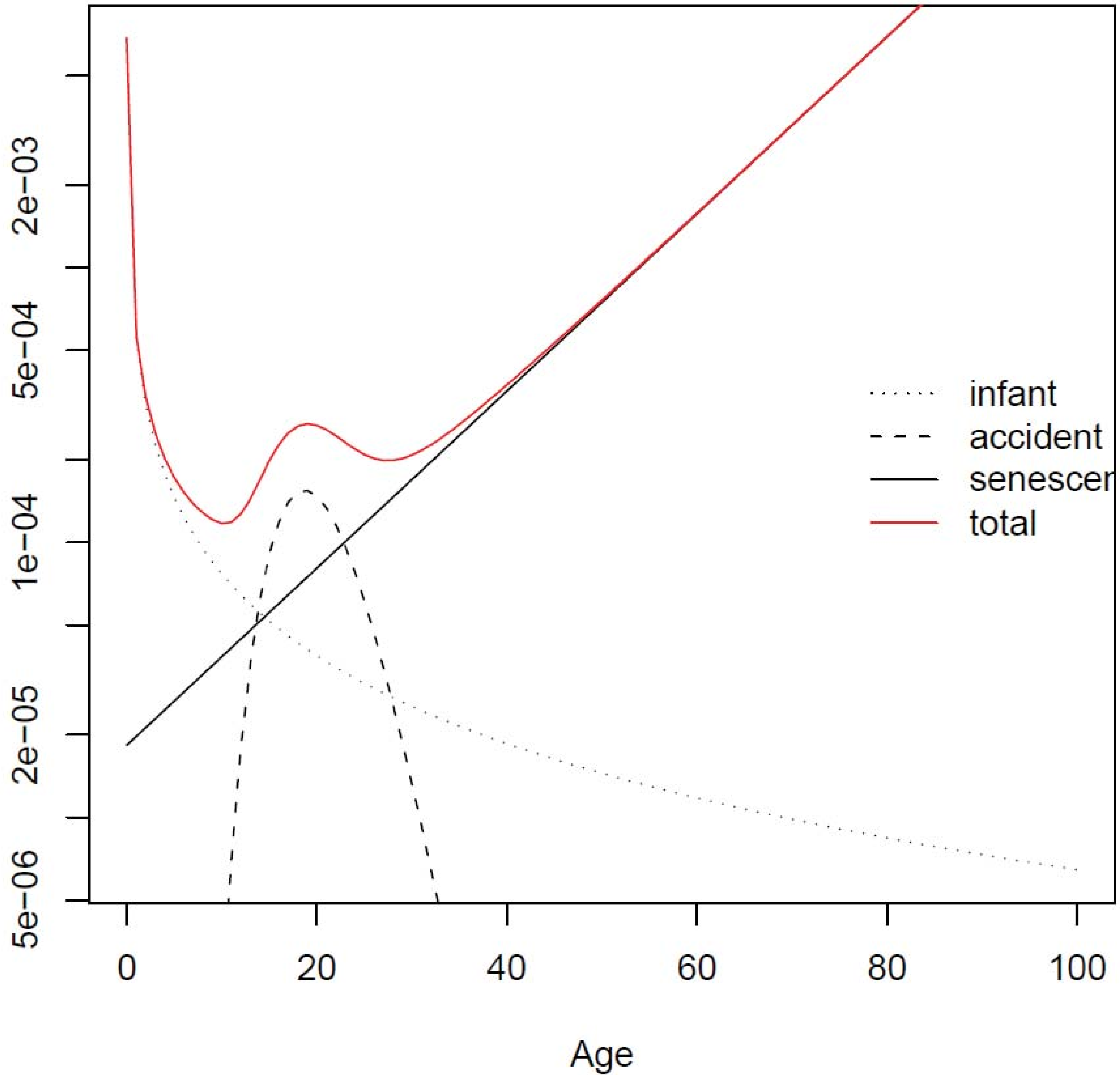


# Competing Components Models

- Far more complex mathematical models have been developed.
- Most build together lines and curves for different components of the mortality (or fertility or migration) schedules.
- Heligman-Pollard is one such example for mortality:

$$f(x) = A^{(x+B)^C} + Dexp[-E\{\log \frac{x}{F}\}^2] + \frac{GH^x}{1+GH^x} \tag{3}$$

# Heligman-Pollard for 1995 UK mortality, source: Jones (2005)



# Competing Components Models

- Double exponential model used for fertility (amongst others)

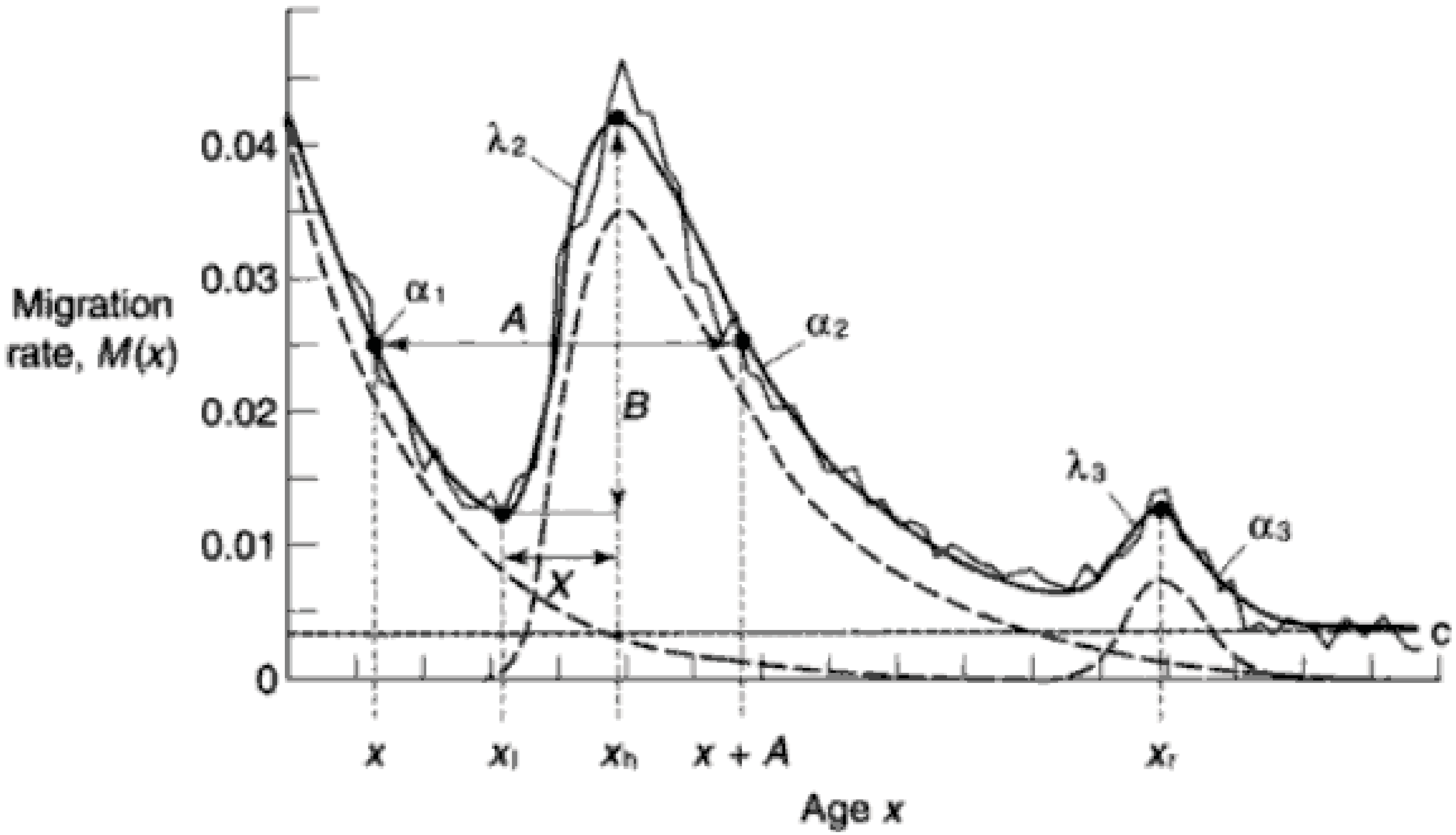
$$f(x) = a \exp(-\alpha(x-\mu) - \exp(-\lambda(x-\mu)))$$

- Coale-Trussel formula (and other formulations) also popular. These interact nuptiality and fertility schedules.
- Rogers and Castro (1981) propose a 7 or 11 parameter model for migration

$$m(x) = a_1 \exp(-\alpha_1 x) + a_2 \exp(-\alpha_2(x-\mu_2) - exp(-\lambda_2(x-\mu_2))) + R$$

where $R = a_3 \exp(\alpha_3(x-\mu_3) - exp(-\lambda_3(x-\mu_3)))$ if retirement peak.

# Migration Models, source: Rogers and Castro (1981)




$\alpha_1$ = rate of descent of pre-labor force component
$\lambda_2$ = rate of ascent of labor force component
$\alpha_2$ = rate of descent of labor force component
$\lambda_3$ = rate of ascent of post-labor force component
$\alpha_3$ = rate of descent of post-labor force component
c = constant

$x_l$ = low point
$x_h$ = high peak
$x_r$ = retirement peak
$x$ = labor force shift
$A$ = parental shift
$B$ = jump

# Mathematical Models

Advantages:

- Reduce dimensions, small number of parameters relative to number of data points.
- Parameters can have nice interpretations.
- Good for comparisons.

Disadvantages

- Hard to fit.
- Starting values difficult to find.
- Estimates frequently unstable.
- Might not adapt to new sources of demographic change.

# Relational Models

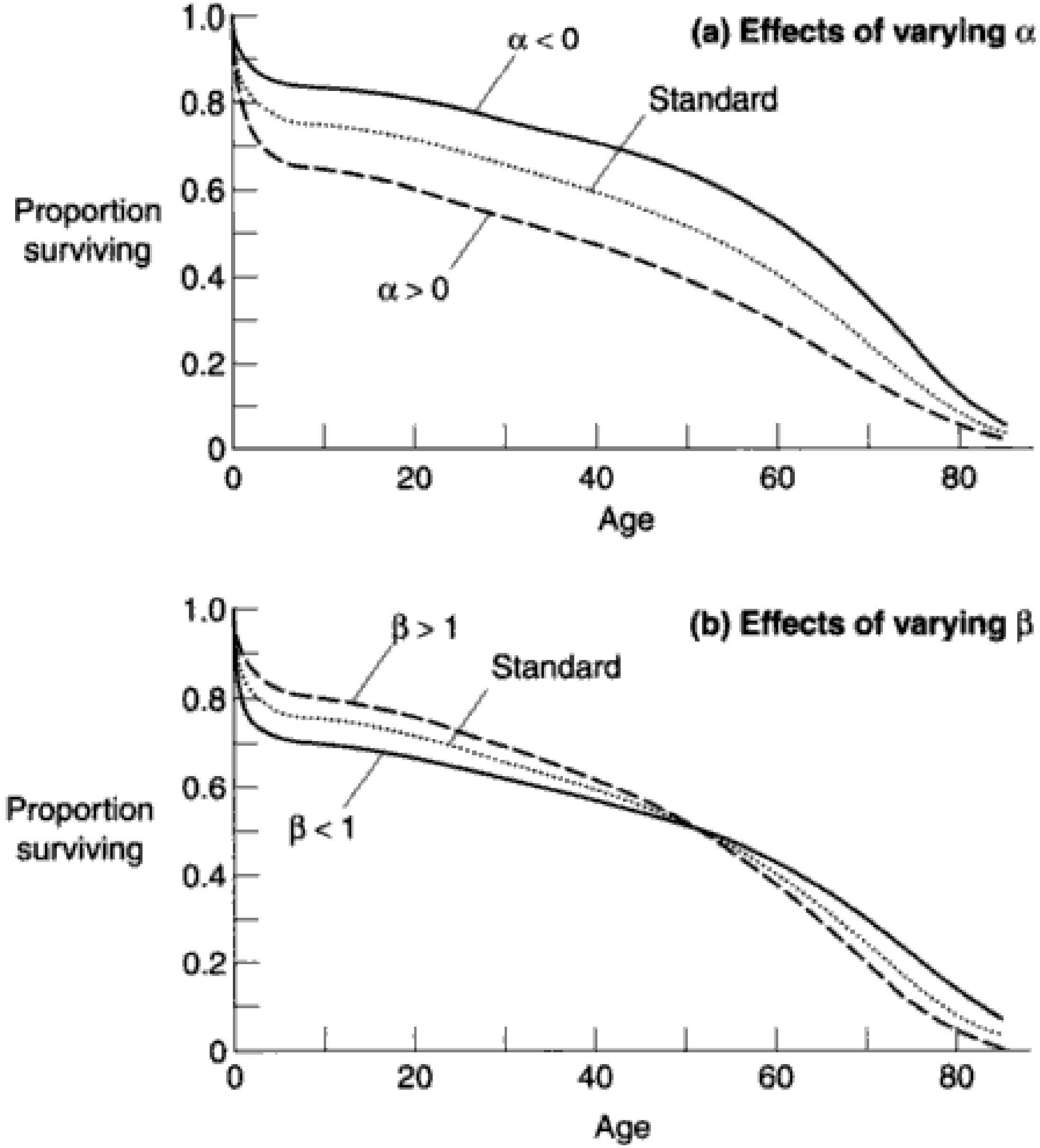
- Brass (1971) proposed a model to relate observed mortality data (which might be incomplete) to a standard schedule, $X^{(s)}$
- Requires logit transformation on our survivorship probabilities

$$\widehat{Y} = logit(x) = \alpha + \beta X^{(s)}$$

  $\alpha$ is interpreted as the level of mortality
  $\beta$ is the shape of mortality

- Can apply to model any standard schedule of mortality, fertility or migration to adjusted using estimated parameters

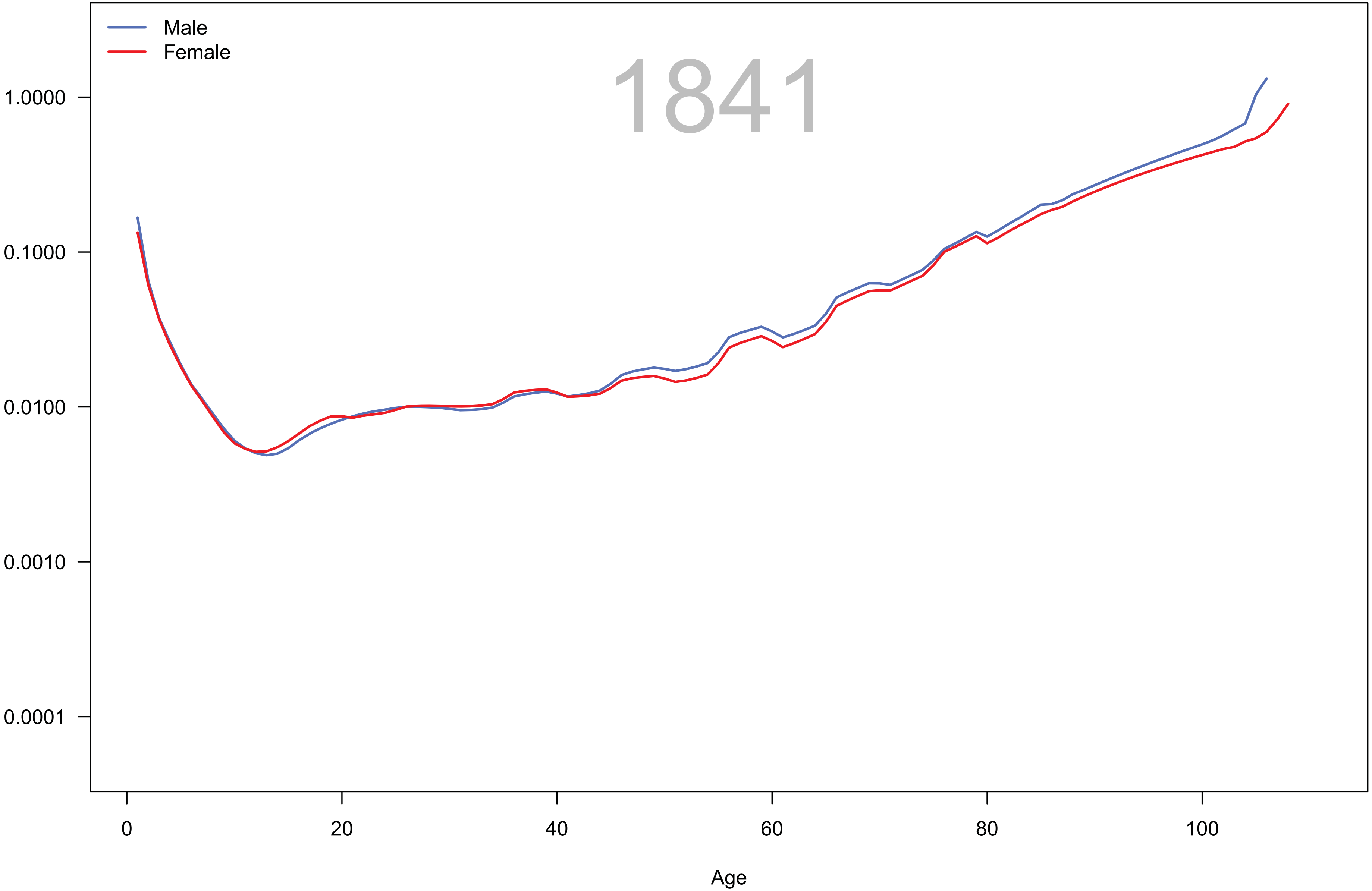
# Effect of Changing Brass Model Parameters, source: Preston et. al. (2001)



# Model Life Tables

- Variety of life tables systems
- Based on a collection of real life tables, usually of high quality
- Use a multivariate analysis technique to find clusters
- Combine them to provide a standard
- e.g. Coale-Demeny (1966, 1983)
  - Divided countries into 4 regional models (North, South, East, West) based on mortality rates at different ages
  - Use available data to estimate life expectancy
  - Use available data to determine mortality pattern
  - Use standard table that fits your assumptions

# Logarithm of $m_x$ 1841 to 2007



# UK $F_x$ (000's) 1974 to 2007

# Parametric Models

- Lee Carter (1992) mortality model used in U.S. Census Bureau population forecasts

$$\log(m(x,t)) = a_x + K(t)b_x$$

  $a_x$ average of log mortality at age $x$
  $b_x$ age pattern of mortality change
  $K(t)$ the level of mortality at time $t$
  - Special case of log-linear model
  - Special case of PCA, using the first principal component of the log mortality matrix $m_x t$
  - Age profiles evolve in implausible ways.
- ARIMA on Model Parameters, see for example McNown and Rogers (1989)

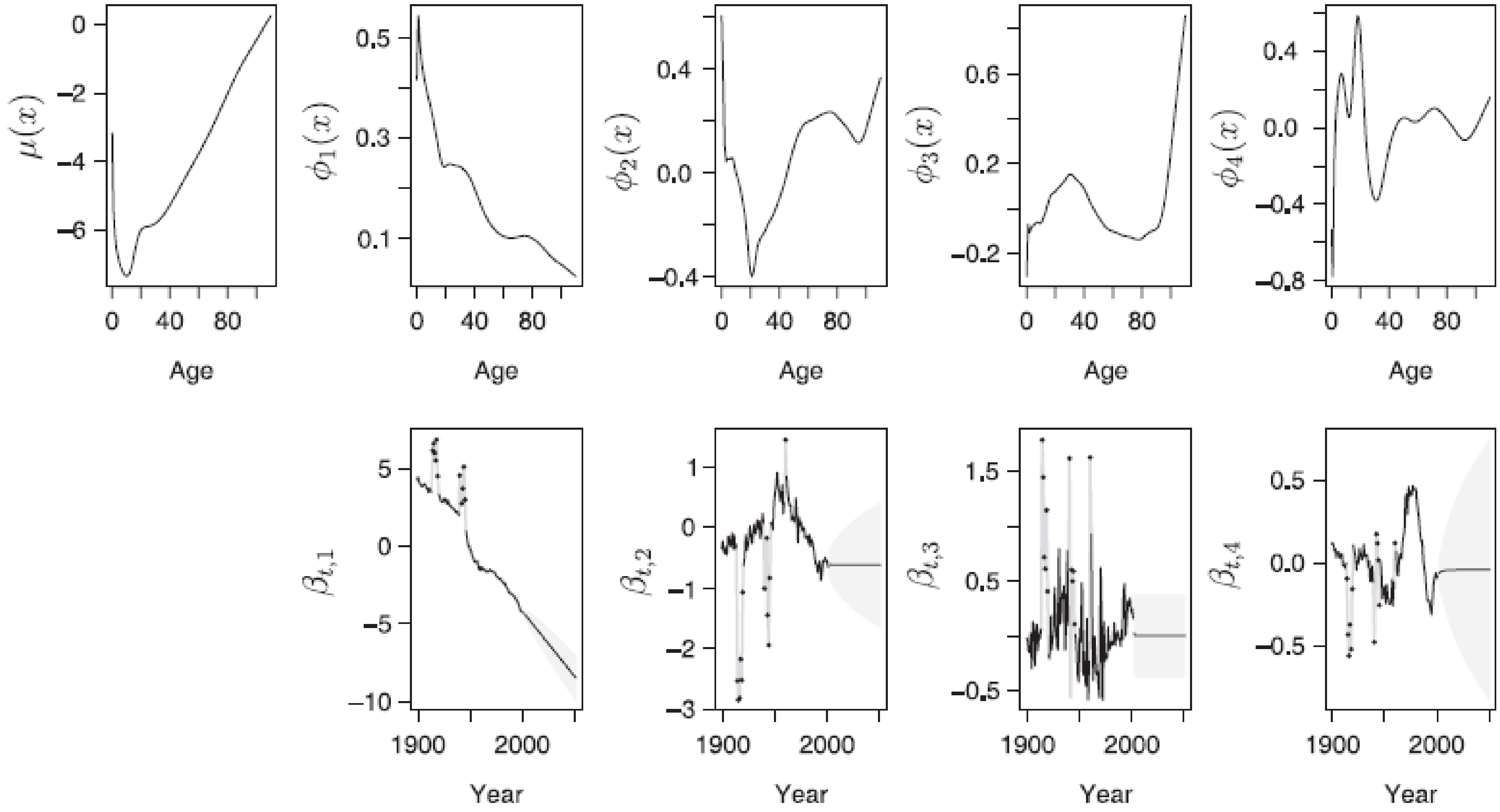
# Functionals

- Recently authors have started to frame the modelling of profiles within the functional data paradigm.
- Hyndman and Ullah (2007) (below), Girosi and King (2008) and Ramsey et.al. (2009)
  1. Nonparametric smoothing of the each curve to obtain $f_t(x)$.
  2. Fit functional model to $f_t(x)$ to decompose curves into principal components

$$f_t(x) = \mu(x) + \sum_{k=1}^{K} \beta_{t,k}\phi_k(x) + e_t(x)$$

     $\phi_k(x)$ is the $k$th principal component (basis) functions
     $\beta_{t,k}$ are corresponding scores
     $\mu(x)$ mean function
  3. Univariate time series for each $\beta_{t,k}$
  4. Forecasts of principal components scores multiplied with principal components to obtain estimated future curves.

# Basis functions and coefficients for French Mortality, source: Hyndman and Ullah (2007)



# Thank You

Thank you for listening!

Research prepared within the Centre for Population Change (CPC) funded by the ESRC Grant number RES-625-28-0001. All opinions are those of the author only.