# Deriving age-specific death rates from life expectancy forecasts

Marius Pascariu and Vladimir Canudas-Romo

Max-Planck Odense Center on the Biodemography of Aging
University of Southern Denmark

mpascariu@health.sdu.dk

12th International Longevity Risk
and Capital Markets Solutions Conference,
Chicago, USA

September 30, 2016

# Motivation

1. Constructing a life table starting from $m_x$'s is straightforward, the way back from $e_x$ to $m_x$ gives an infinity of solutions;
2. Some people believe in expert forecasts;
3. Forecasting $e_x$ may be 'easier';
4. The method can be used for validating the forecasts given by traditional models.

# Previous work on death rates estimation

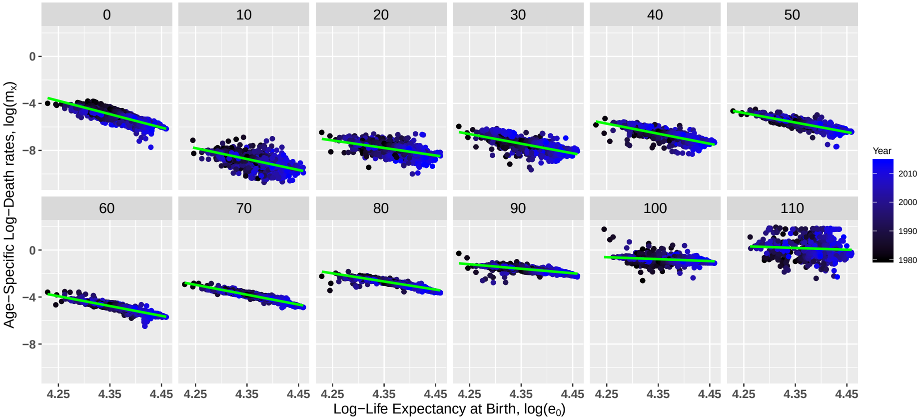
Methods based on:

- records of population growth and distribution by age;
- responses to questions about fertility and mortality;
- reported age distributions or reported child survival;
- Lee-Carter model.

| **Source** | **Date** | **Description** |
|---|---|---|
| United Nations | 1955 | 1-parameter system indexed on infant mortality |
| Gabriel & Ronen | 1958 | 1-parameter... |
| Ledermann & Breas | 1959 | 1 or 2 |
| Coale & Demeny | 1966 | + Regions |
| Brass | 1971 | Logit life table system |
| United Nations | 1981 | Revised UN + Regions |
| WHO | 2000 | Improved Brass |
| Murray et al. | 2003 | Modified Logit |
| Wilmoth et al. | 2013 | Log-quadratic Model |
| Ševčíková et al. | 2016 | Algorithm based on Lee-Carter |

# Life expectancy vs. Death-rates

*Linear relation between life expectancy at birth and mortality rates, by age.*
*Based on HMD mortality data starting from 1980 for 43 countries and regions.*




Data source: www.mortality.org

# Linear-link model

$$\log(m_x) = \beta_x \log e_0 + \nu_x k + \epsilon_x, \quad (1)$$

where

- $\beta_x$ is age-specific pattern of human mortality;
- $\nu_x$ is the rate of mortality improvement over ages;
- $k$ is the parameter to be optimize;
- $\epsilon_x \backsim N(0, \sigma^2)$.

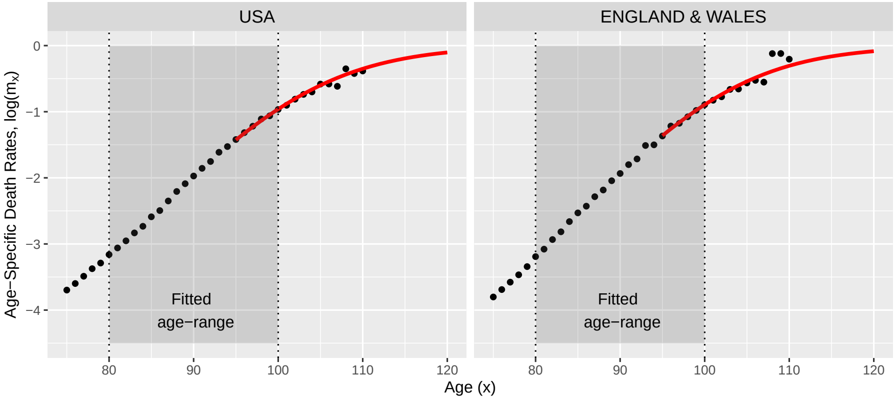
# Algorithm

1. Extend historical $m_x(t)$ to higher age groups up to age 120 for all $t$. Kannisto, Logistic, Gompertz method (?)
2. Estimate $\beta_x$ by using the least squares approach by minimizing the sum of squared residuals given by $\sum_i [\log(m_x) - \beta_x log(e_0)]^2 = r_{xi}$ ;
3. Estimate parameter $\nu_x$ by computing a singular value decomposition $SVD[r_{xi}]$ of the matrix of regression residuals;
4. Smooth the $\beta_x$ and $\nu_x$ parameters using splines.
5. Compute the mortality rates by $\log(m_x) = \beta_x \log e_x + \nu_x k$, where $k = 0$.
6. Optimize the mortality curve by finding the value of $k$ where the difference between $e_x^*(\tau)$ and $e_x(\tau)$ is below tolerance level.

# Kannisto extension of old age mortality (2014)

## The Kannisto model

$$logit(m_x) = ln(\alpha) + \beta\chi + \epsilon_\chi, \quad (2)$$

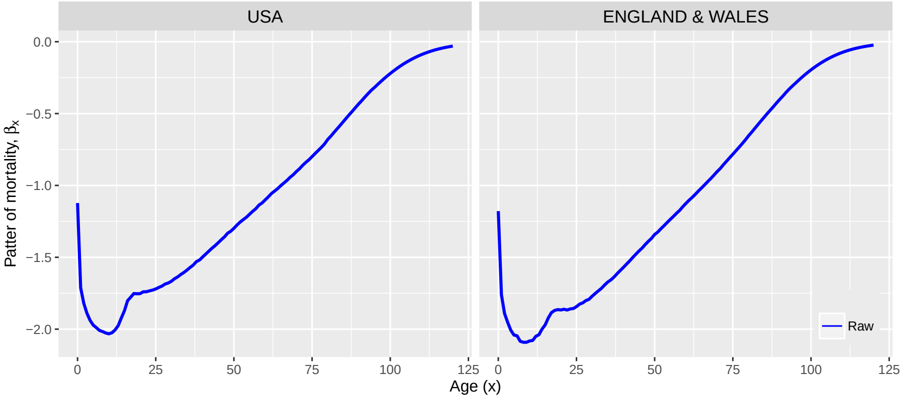


# Algorithm

1. Extend historical $m_x(t)$ to higher age groups up to age 120 for all $t$. Kannisto, Logistic, Gompertz method (?)
2. Estimate $\beta_x$ by using the least squares approach by minimizing the sum of squared residuals given by $\sum_i \left[\log(m_x) - \beta_x log(e_0)\right]^2 = r_{xi}$ ;
3. Estimate parameter $\nu_x$ by computing a singular value decomposition $SVD\left[r_{xi}\right]$ of the matrix of regression residuals;
4. Smooth the $\beta_x$ and $\nu_x$ parameters using splines.
5. Compute the mortality rates by $\log(m_x) = \beta_x \log e_x + \nu_x k$, where $k = 0$.
6. Optimize the mortality curve by finding the value of $k$ where the difference between $e^*_x(\tau)$ and $e_x(\tau)$ is below tolerance level.

# General pattern of human mortality, $\beta_x$

*The general pattern of human mortality in female population.*



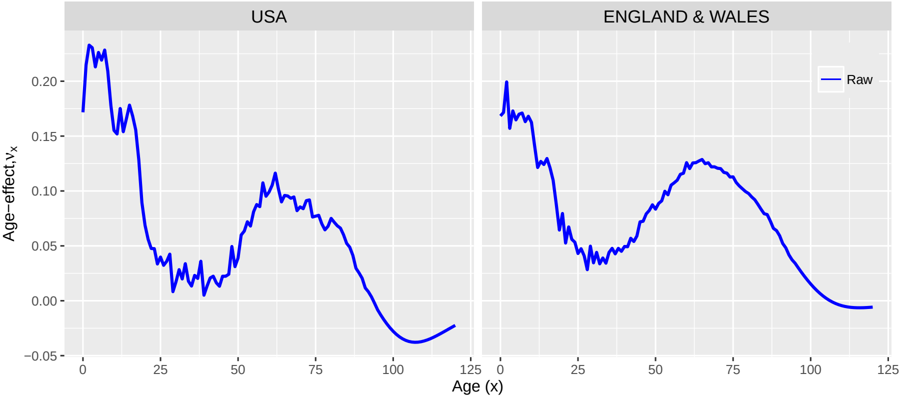

*Model fitted for 1980-2013 period (HMD)*

# Algorithm

1. Extend historical $m_x(t)$ to higher age groups up to age 120 for all $t$. Kannisto, Logistic, Gompertz method (?)
2. Estimate $\beta_x$ by using the least squares approach by minimizing the sum of squared residuals given by $\sum_i [\log(m_x) - \beta_x log(e_0)]^2 = r_{xi}$ ;
3. Estimate parameter $\nu_x$ by computing a singular value decomposition $SVD[r_{xi}]$ of the matrix of regression residuals;
4. Smooth the $\beta_x$ and $\nu_x$ parameters using splines.
5. Compute the mortality rates by $\log(m_x) = \beta_x \log e_x + \nu_x k$, where $k = 0$.
6. Optimize the mortality curve by finding the value of $k$ where the difference between $e_x^*(\tau)$ and $e_x(\tau)$ is below tolerance level.

# The age-effect, $\nu_x$

*The $\nu_x$ parameter gives the speed of improvement in $m_x$ in the analyzed period.*



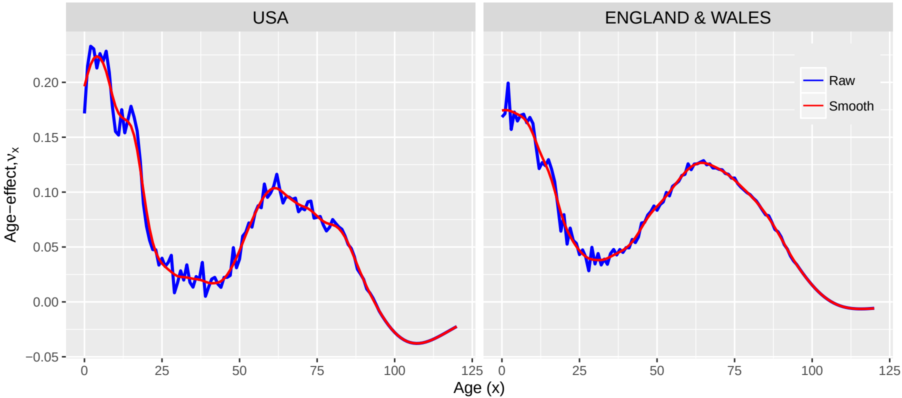

*Model fitted for 1980-2013 period (HMD)*

# Algorithm

1. Extend historical $m_x(t)$ to higher age groups up to age 120 for all $t$. Kannisto, Logistic, Gompertz method (?)
2. Estimate $\beta_x$ by using the least squares approach by minimizing the sum of squared residuals given by $\sum_i \left[\log(m_x) - \beta_x log(e_0)\right]^2 = r_{xi}$ ;
3. Estimate parameter $\nu_x$ by computing a singular value decomposition $SVD\,[r_{xi}]$ of the matrix of regression residuals;
4. Smooth the $\beta_x$ and $\nu_x$ parameters using splines.
5. Compute the mortality rates by $\log(m_x) = \beta_x \log e_x + \nu_x k$, where $k = 0$.
6. Optimize the mortality curve by finding the value of $k$ where the difference between $e_x^*(\tau)$ and $e_x(\tau)$ is below tolerance level.

# The age-effect, $\nu_x$

*Smoothing is important in obtaining graduated mortality curves and to avoid projecting age-specific noise in the jump-off life table*



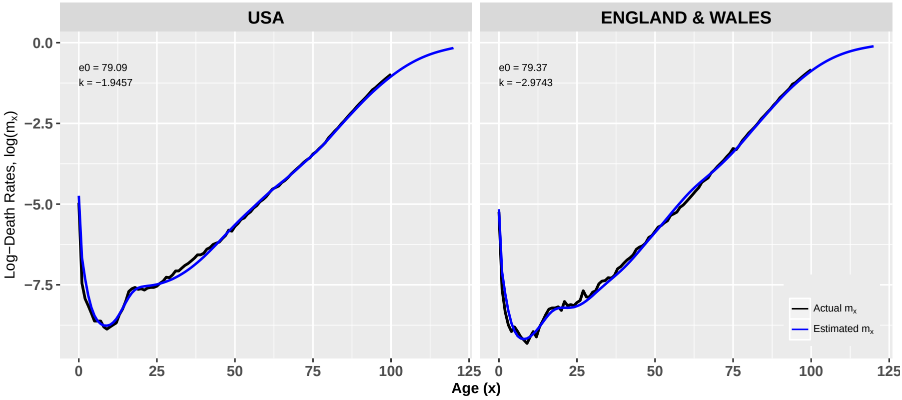

*Model fitted for 1980-2013 period (HMD)*

# Algorithm

1. Extend historical $m_x(t)$ to higher age groups up to age 120 for all $t$. Kannisto, Logistic, Gompertz method (?)
2. Estimate $\beta_x$ by using the least squares approach by minimizing the sum of squared residuals given by $\sum_i [\log(m_x) - \beta_x log(e_0)]^2 = r_{xi}$ ;
3. Estimate parameter $\nu_x$ by computing a singular value decomposition $SVD[r_{xi}]$ of the matrix of regression residuals;
4. Smooth the $\beta_x$ and $\nu_x$ parameters using splines.
5. Compute the mortality rates by $\log(m_x) = \beta_x \log e_x + \nu_x k$, where $k = 0$.
6. Optimize the mortality curve by finding the value of $k$ where the difference between $e^*_x(\tau)$ and $e_x(\tau)$ is below tolerance level.

# Reconstructed Mortality Curve, $m_x$ (1995)

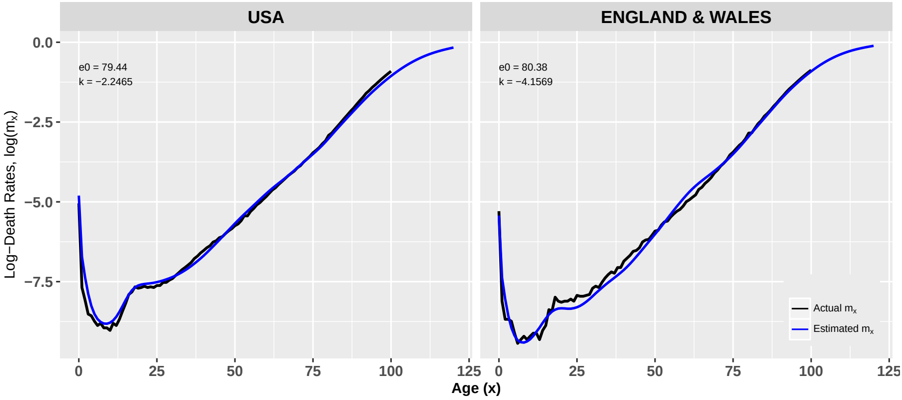
*Observed and estimated death rates for female populations in 1995.*
*High degree of accuracy.*




*Model fitted for 1965-1990 period (HMD)*

# Reconstructed Mortality Curve, $m_x$ (2000)

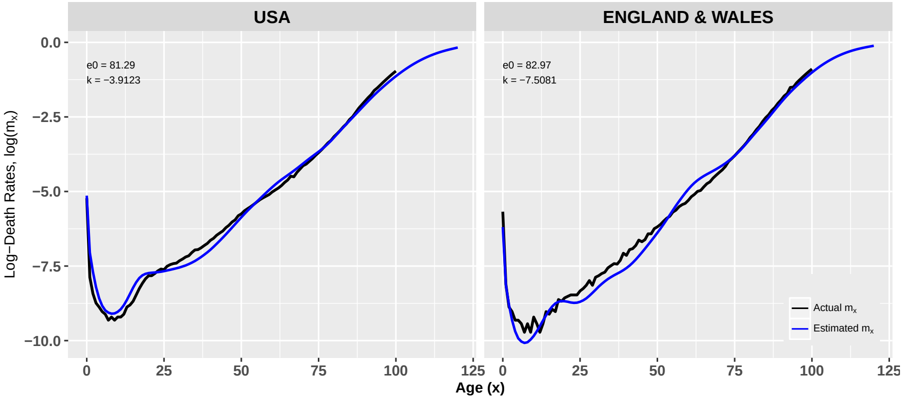
*Observed and estimated death rates 10 years in the future.*
*The estimated mortality curve maintain its accuracy.*




*Model fitted for 1965-1990 period (HMD)*

# Reconstructed Mortality Curve, $m_x$ (2013)

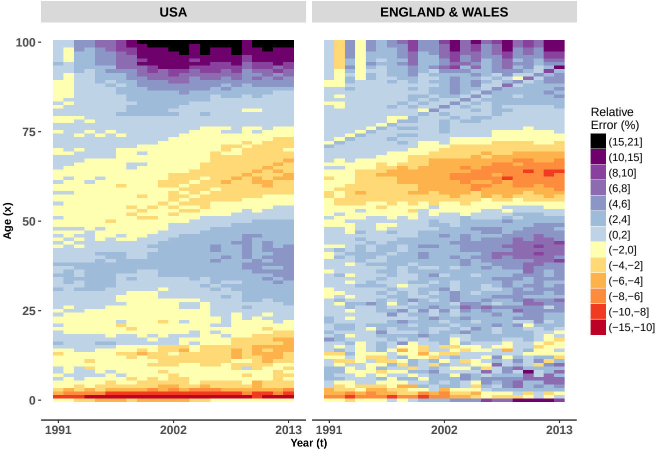
*Observed and estimated age-specific death rates in 2013.*
*The two mortality curves give exactly the same $e_0$.*




*Model fitted for 1965-1990 period (HMD)*

# Age specific reconstruction error

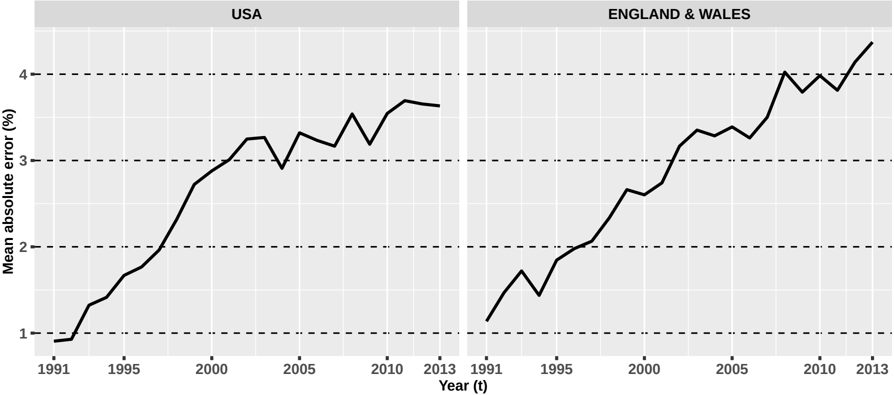
*The relative errors are usually in the range of $\pm 4\%$. Larger errors can be expected at advanced ages or at infancy.*



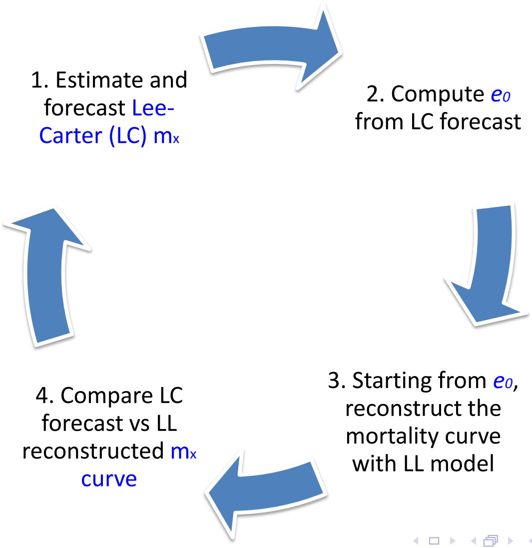

*Model fitted for 1965-1990 period (HMD)*

# Accuracy: 1991-2014

*The mean absolute error in a 24 year forecasting window is less than 4.2%!*

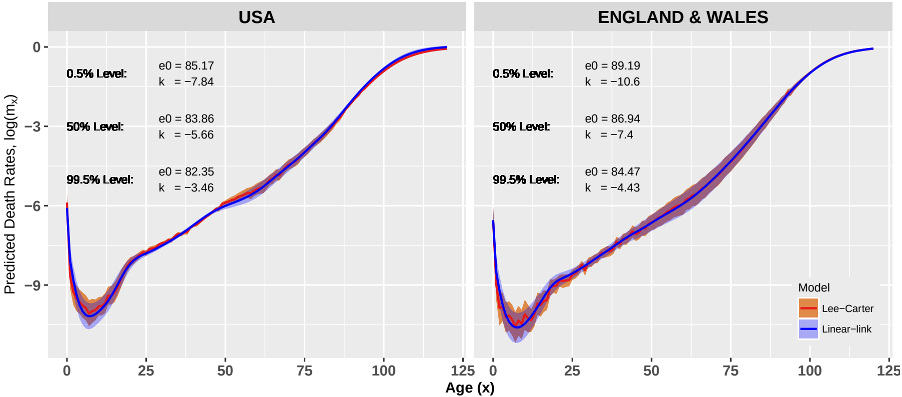



*Model fitted for 1965-1990 period (HMD)*


1. Estimate and forecast Lee-Carter (LC) $m_x$
2. Compute $e_0$ from LC forecast
3. Starting from $e_0$, reconstruct the mortality curve with LL model
4. Compare LC forecast vs LL reconstructed $m_x$ curve

# Reconstruction of Lee-Carter forecast (2040)

*Comparison of the mortality curves predicted by Lee-Carter and Linear-link models in 2040 from female population.*




*Model fitted for 1980-2013 period (HMD)*

# Conclusion & Discussion

- Linear model to derive the entire schedule of age-specific death rates based on a single value of life expectancy and prior knowledge of human mortality pattern ($m_x$ vs. $e_x$);
- The optimal length of data to be used in the fitting of the model is between 30 and 35 years. A larger time interval would only make the parameter estimates to lose their relevance;
- The model can accurately reconstruct a Lee-Carter forecast starting from a single value of life expectancy at birth.

# Thank you!