Modelling and quantifying mortality and longevity risk

Module B1 : Multi-population Mortality Models

Michel Vellekoop

Actuarial Summer School
Warsaw, Sept 18-19, 2025

Overview

In this module:

- Motivation for combining models for multiple populations
- Estimation of multi-population models

Motivation to combine populations

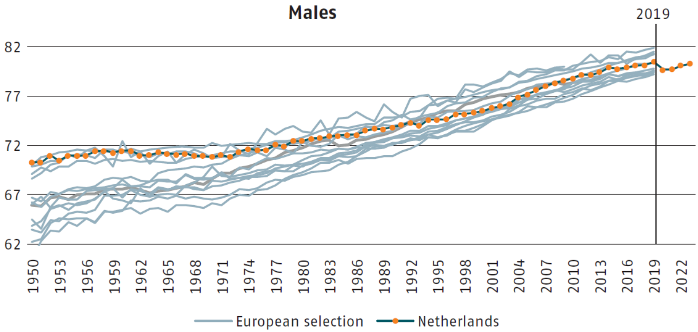
Micro vs Macro Risk

- In mortality models, we can distinguish (at least) two sources of uncertainty:
  - Macro Longevity Risk
    The changes over time in the probabilities of dying
    (due to random innovations in time effect κ_t), and
  - Micro Longevity Risk
    The Poisson-distributed number of deaths *given* these probabilities
    (i.e. the Grim Reaper throwing his coins for any individual).
- Second risk becomes relatively smaller if we increase the size of our portfolio.
- The first one does not. This is of fundamental importance for the risk management of life insurance products.

Correlation structure

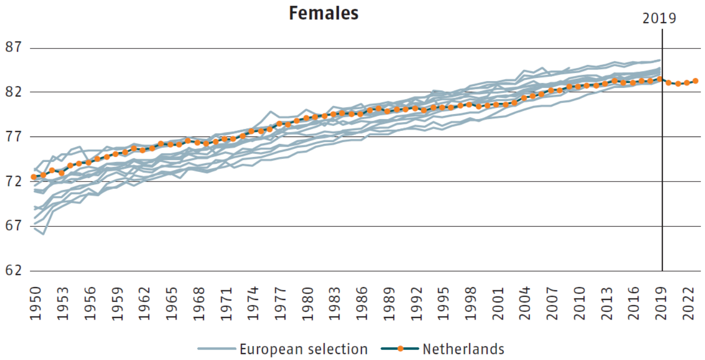
- The single factor Lee-Carter model creates perfect dependence between changes for different ages.
- Correlation between changes in $\ln \mu_{x_1,t}$ and $\ln \mu_{x_2,t}$ equals one for all ages x_1, x_2.
- In this module we discuss multi-population models such as the Dutch Actuarial Society AG2020/AG2022 based on (Li and Lee, 2005).
- Introducring multiple stochastic factors and data from different countries allows us to distinguish between country-specific and more general trends in mortality changes.

Period Life Expectancies at birth ♂



European peers and the Netherlands

Period Life Expectancies at birth ♀

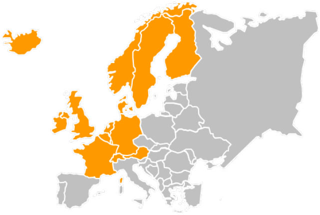


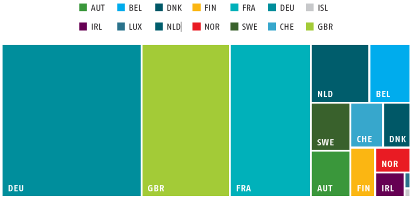
European Peers and the Netherlands

Looking beyond the border

- KAG model structure (pre-pandemic !) distinguishes
  - Long-term trend which is common among Western-European countries (medical improvements will not stop at a border)
  - Short term deviations from this trend for any single country (effects such as start/end of smoking epidemic may vary per population)
- We only take countries which are comparable in wealth: the top half of the European Union in terms of GDP (per capita). Using more data, we aim to enhance stability of forecasts.
- Main assumption that mortality in countries will not diverge from common European trend can be tested over time by studying the development of the country-specific time series.

Datsets for Netherlands and European peer group

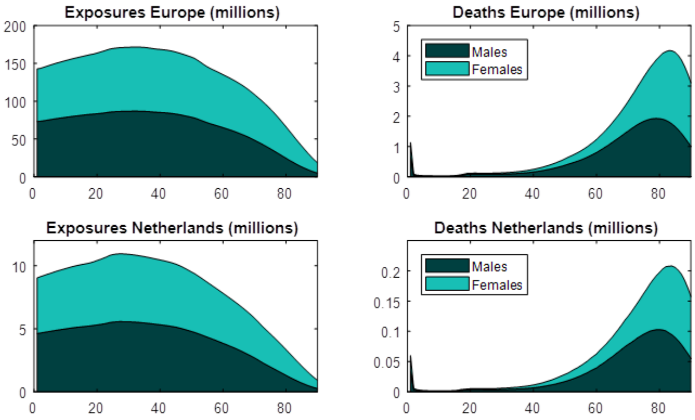




Left: Countries included in analysis
(European countries with GDP per person above the European average),

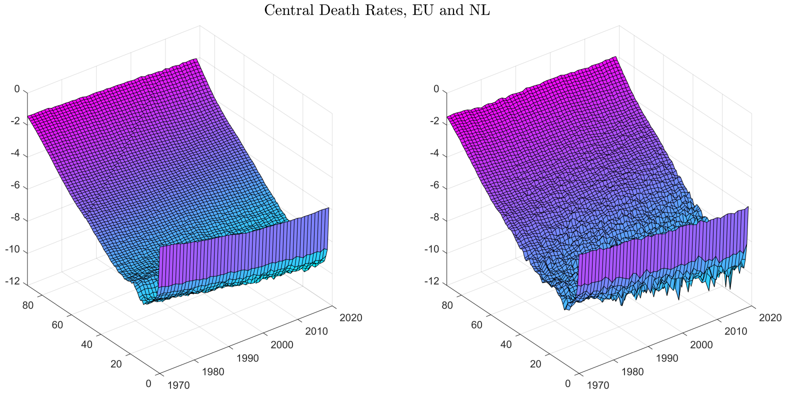
Right: Relative sizes for death counts (2019).
(Germany 787.320, Netherlands 126.995)

Datsets for Netherlands and European peer group



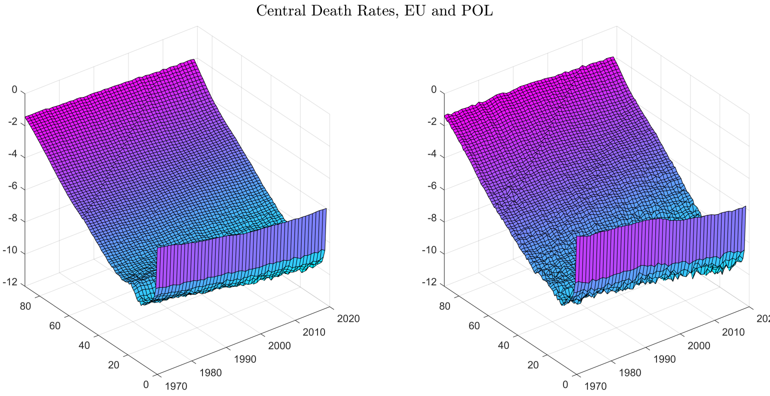
Dataset sizes for exposures and death counts.

Raw data



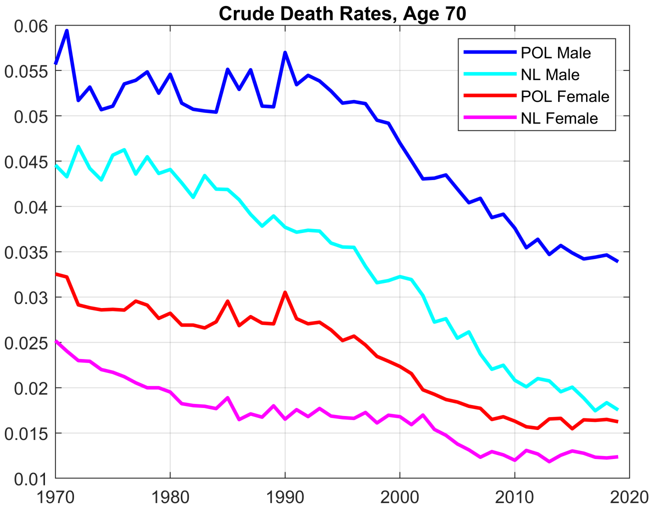
Log central death rates, $\ln(D_{xt}/E_{xt})$
for females in Europe (left) and the Netherlands (right).

Raw data



Log central death rates, $\ln(D_{xt}/E_{xt})$
for females in Europe (left) and Poland (right).
Emphasizes importance of proper model for observation noise!

Raw data



Crude Death Rates, Age 70
0.06
0.055
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
1970
1980
1990
2000
2010
2020
POL Male
NL Male
POL Female
NL Female

Stochastic Model for the Dutch Population

- AG2020 (model of the Dutch Actuarial Society)

$$
\begin{aligned}
D_{xtg} &\sim \mathrm{Poisson}(E_{xtg}\,\mu_{xtg}) \\
\ln \mu_{xtg} &= \ln \mu^{\mathrm{EU}}_{xtg} + \ln \mu^{\mathrm{NL}}_{xtg}
\end{aligned}
$$

- with long term trend fitted using countries in Europe with comparable GDP

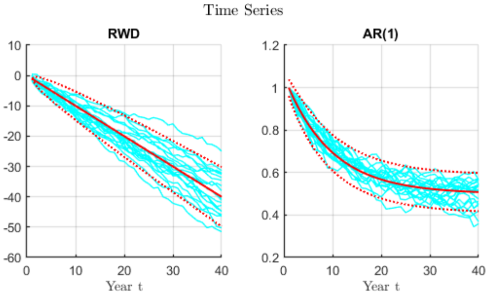
$$
\begin{aligned}
\ln \mu^{\mathrm{EU}}_{xtg} &= A^g_x + B^g_x\,K^g_t \\
K^g_{t+1} &= K^g_t + \theta^g + \epsilon^g_{t+1}
\end{aligned}
$$

- and short term dynamics which describe difference with this European trend

$$
\begin{aligned}
\ln \mu^{\mathrm{NL}}_{xtg} &= \alpha^g_x + \beta^g_x\,\kappa^g_t \\
\kappa^g_{t+1} &= a^g\,\kappa^g_t + \delta^g_{t+1} + c^g
\end{aligned}
$$

with $|a^g| < 1$, and $Z_t := (\epsilon^m_t, \delta^m_t, \epsilon^f_t, \delta^f_t)$ i.i.d. Gaussian with mean $(0,0,0,0)$ and a given 4×4 covariance matrix C which does not change over time.

Dynamics in the long run



Common European trend for all in peer group (Left, RWD)

$$K^g_{t+1} = K^g_t + \theta^g + \epsilon^g_{t+1},$$

combined with transient effect (Right, AR(1)) that describes individual country's deviation from the common trend:

$$\kappa^g_{t+1} = a^g \, \kappa^g_t + \delta^g_{t+1} + c^g .$$

Estimation of multi-population models

Maximum Likelihood estimation

- First we fit European data using $\ln \mu_{xtg}^{\mathrm{EU}} = A_x^g + B_x^g K_t^g$

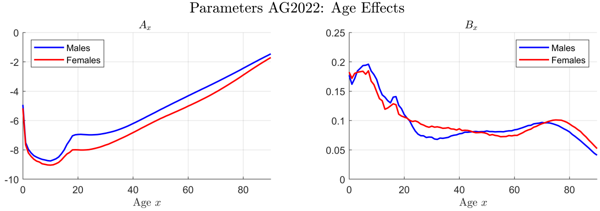
$$\max_{(A_x^g, B_x^g, K_t^g)} \prod_{x \in \mathcal{X}^{\mathrm{EU}}} \prod_{t \in \mathcal{T}^{\mathrm{EU}}} \frac{(E_{xtg}^{\mathrm{EU}} e^{A_x^g + B_x^g K_t^g})^{D_{xtg}^{\mathrm{EU}}} \exp(-E_{xtg}^{\mathrm{EU}} e^{A_x^g + B_x^g K_t^g})}{D_{xtg}^{\mathrm{EU}}!},$$

- then the Dutch deviation $\ln \mu_{xtg}^{\mathrm{NL}} = \ln \mu_{xtg}^{\mathrm{EU}} + \alpha_x^g + \beta_x^g \kappa_t^g$

$$\max_{(\alpha_x^g, \beta_x^g, \kappa_t^g)} \prod_{x \in \mathcal{X}^{\mathrm{NL}}} \prod_{t \in \mathcal{T}^{\mathrm{NL}}} \frac{(E_{xtg}^{\mathrm{NL}} e^{\hat{A}_x^g + \hat{B}_x^g \hat{K}_t^g + \alpha_x^g + \beta_x^g \kappa_t^g})^{D_{xtg}^{\mathrm{NL}}} \exp(-E_{xtg}^{\mathrm{NL}} e^{\hat{A}_x^g + \hat{B}_x^g \hat{K}_t^g + \alpha_x^g + \beta_x^g \kappa_t^g})}{D_{xtg}^{\mathrm{NL}}!}$$

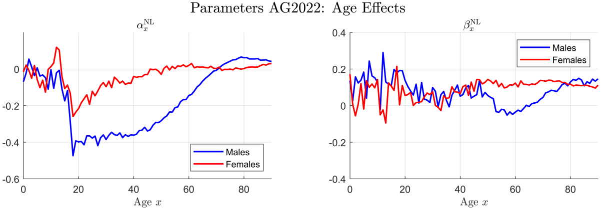
  using estimates of first step.
- Note that we treat time series K and κ as parameters here.
- In numerical procedures we maximize the logarithm of the likelihood.

Age Parameters for the Dutch Population, AG2022

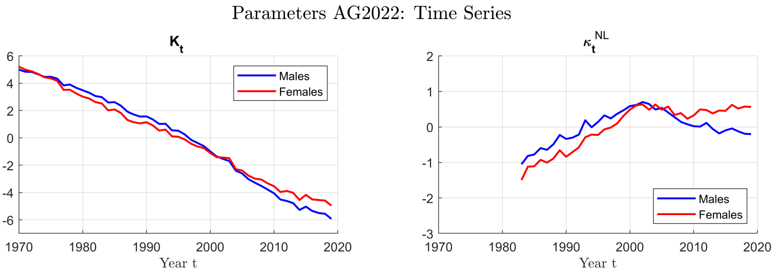


$$\ln \mu_{xtg}^{\mathrm{EU}} = A_x^g + B_x^g K_t^g$$

$$\ln \mu_{xtg}^{\mathrm{NL}} = \alpha_x^g + \beta_x^g \kappa_t^g$$



Time Series for the Dutch Population, AG2022

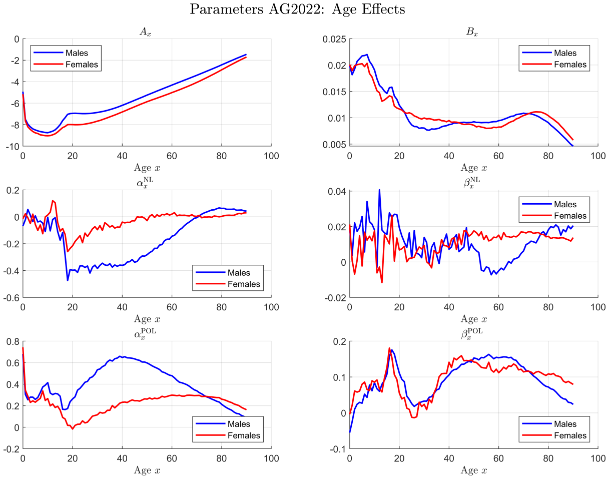


$$\ln \mu_{xtg}^{\mathrm{EU}} = A_x^g + B_x^g \, K_t^g , \qquad K_{t+1}^g = K_t^g + \theta^g + \epsilon_{t+1}^g$$

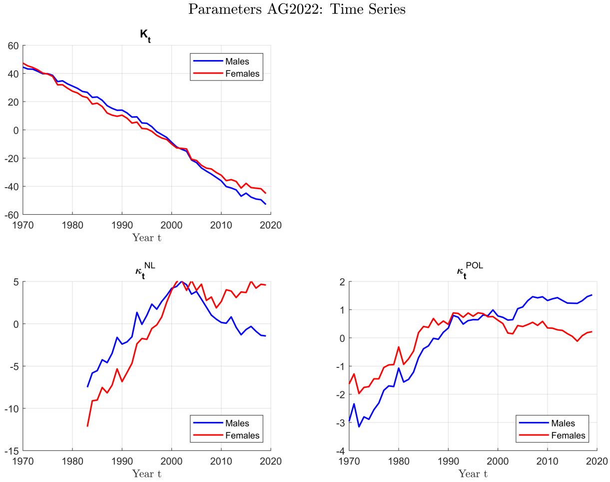
$$\ln \mu_{xtg}^{\mathrm{NL}} = \alpha_x^g + \beta_x^g \, \kappa_t^g , \qquad \kappa_{t+1}^g = a^g \, \kappa_t^g + \delta_{t+1}^g + c^g$$

Notice that we (only) use recent data for the Dutch deviation series κ_t^g.

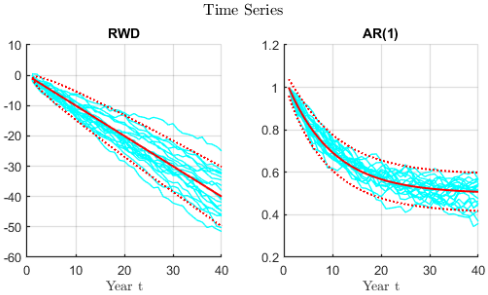
Comparing European countries, Netherlands and Poland



Comparing European countries, Netherlands and Poland



Dynamics in the long run



Common European trend for all in peer group (Left, RWD)

$$K^g_{t+1} = K^g_t + \theta^g + \epsilon^g_{t+1},$$

combined with transient effect (Right, AR(1)) that describes individual country's deviation from the common trend:

$$\kappa^g_{t+1} = a^g \, \kappa^g_t + \delta^g_{t+1} + c^g.$$

Random Walk with Drift

- Dynamics for K are assumed to follow (we suppress superscripts g):

$$K_t = K_{t-1} + \theta + \epsilon_t$$

  with i.i.d. Gaussian $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ and $\theta < 0$.

- This implies that

$$K_t = K_0 + t\theta + \sum_{s=0}^{t-1} \epsilon_s$$

  so

$$\mathbb{E}K_t = K_0 + t\theta$$

  and

$$\mathrm{Var}(K_t) = t\sigma_\epsilon^2.$$

First Order Autoregressive Model

- For both genders, dynamics for κ are assumed to follow

$$\kappa_t = a\kappa_{t-1} + \delta_t + c$$

with i.i.d. Gaussian sequence $\delta_t \sim N(0, \sigma_\delta^2)$.
- This implies that

$$\kappa_t \;=\; \tfrac{c}{1-a} + (\kappa_0 - \tfrac{c}{1-a})a^t + \sum_{s=0}^{t-1} a^s \delta_{t-s}$$

- We thus find that mean, variance and autocorrelation converge to fixed values

$$\begin{aligned}
\mathbb{E}\kappa_t &= \tfrac{c}{1-a} + (\kappa_0 - \tfrac{c}{1-a})a^t \quad \overset{t\uparrow\infty}{\to} \quad \tfrac{c}{1-a} \\
\mathrm{Var}\,\kappa_t &= \sigma_\delta^2 \sum_{s=0}^{t-1} a^{2s} = \sigma_\delta^2 \frac{1-a^{2t}}{1-a^2} \quad \overset{t\uparrow\infty}{\to} \quad \frac{\sigma_\delta^2}{1-a^2} \\
\rho_{t,t+s} &= a^s \sqrt{\frac{1-a^{2t}}{1-a^{2(t+s)}}} \quad \overset{t\uparrow\infty}{\to} \quad a^s
\end{aligned}$$

Prediction of future values: RWD

- If today is time t, value at a future time $t+h$ can be written as

$$K_{t+h} = K_t + h\theta + \sum_{s=0}^{h-1} \epsilon_{t+s}$$

which has the Gaussian distribution $N(K_t + h\theta, \sigma_\epsilon^2 h)$.

- So our best point estimate for the future value at time $t+h$ is

$$K_t + h\theta$$

and the 95% confidence bounds are

$$\pm 1.96 \sigma_\epsilon \sqrt{h}$$

on top of that value.

Prediction of future values: AR(1)

- If today is time t, value at a future time $t+h$ can be written as

$$\kappa_{t+h} = \frac{c}{1-a} + (\kappa_t - \frac{c}{1-a})a^h + \sum_{s=0}^{h-1} a^s \delta_{t+h-s}$$

  with distribution $N(\frac{c}{1-a} + (\kappa_t - \frac{c}{1-a})a^h, \ \frac{\sigma_\delta^2(1-a^{2h})}{1-a^2})$.
- So our best point estimate for the future value h years from now is

$$\frac{c}{1-a} + (\kappa_t - \frac{c}{1-a})a^h$$

  i.e. the current value κ_t times an exponentially decreasing weight plus a limit value to which it converges.
- The 95% confidence bounds are

$$\pm 1.96\sigma_\delta \frac{\sqrt{1-a^{2h}}}{\sqrt{1-a^2}}$$

  on top of that value. This equals $1.96\sigma_\delta$ for $h=1$ but converges to a fixed strictly positive value $1.96\sigma_\delta/\sqrt{1-a^2}$ for $h \to \infty$.

Time Series for the Dutch Population, AG2022

- Calibration of time series

$$
\begin{aligned}
K^g_{t+1} &= K^g_t + \theta^g + \epsilon^g_{t+1}, \\
\kappa^g_{t+1} &= a^g\,\kappa^g_t + c^g + \delta^g_{t+1},
\end{aligned}
$$

uses $Y_t = X_t\Theta + Z_t$, $Z_t \sim N(0_4, C)$ iid, for

$$
Y_t = \begin{bmatrix} K^m_{t+1} - K^m_t \\ K^v_{t+1} - K^v_t \\ \kappa^m_{t+1} \\ \kappa^v_{t+1} \end{bmatrix}, \quad X_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \kappa^m_t & 0 & 1 & 0 \\ 0 & 0 & 0 & \kappa^v_t & 0 & 1 \end{bmatrix}, \quad \Theta = \begin{bmatrix} \theta^m \\ \theta^v \\ a^m \\ a^v \\ c^m \\ c^v \end{bmatrix}, \quad Z_t = \begin{bmatrix} \epsilon^m_{t+1} \\ \epsilon^v_{t+1} \\ \delta^m_{t+1} \\ \delta^v_{t+1} \end{bmatrix}.
$$

- Maximal likelihood estimate for parameters $(\hat{\Theta},\ \hat{C})$ based on observations $\{(Y_t, X_t) :\ 1 \leq t \leq T-1\}$ is

$$
(\hat{\Theta}, \hat{C}) = \underset{(\Theta, C)}{\arg\max}\ \ln \prod_{t=1}^{T-1} \left((2\pi|C|)^{-1/2} \exp[-\tfrac{1}{2}(Y_t - X_t\Theta)' C^{-1} (Y_t - X_t\Theta)] \right).
$$

Key References

- Barrieu et al. (2010), Understanding, modelling and managing longevity risk: key issues and main challenges, *Scandinavian Actuarial Journal* 3, 203-231.
- Brouhns, N., Denuit, M., and Vermunt, J.K. (2002) A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics* 31(3), 373-393.
- Koninklijk Actuarieel Genootschap (2022) Projection Table AG2022.
- Lee, R.D. and Carter, L.R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* 87(419), 659-675.
- Li, N. and Lee, R. (2005) Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography* 42(3), 575-594.
- Pitacco et al. (2009) Modelling Longevity Dynamics for Pensions and Annuity Business, Oxford University Press.