

# **Modelling and quantifying mortality and longevity risk**

## **Module B2 : Projection for Multi-population Models**

Michel Vellekoop

Actuarial Summer School  
Warsaw, Sept 18-19, 2025

Please do not copy or redistribute without permission

# Overview



In this module:

- Projection of multi-population models
- Performance criteria for projections

# Projection of multi-population models

---

# Projections for AG2022

- Simulation of time series: Cholesky factorization of  $C = H^T H$  gives

$$\begin{aligned} Y_t &= X_t \Theta + Z_t, \quad Z_t \sim N(0_4, C) \\ &= X_t \Theta + H \tilde{Z}_t, \quad \tilde{Z}_t \sim N(0_4, I_4) \end{aligned}$$

for

$$Y_t = \begin{bmatrix} K_{t+1}^m - K_t^m \\ K_{t+1}^v - K_t^v \\ \kappa_{t+1}^m \\ \kappa_{t+1}^v \end{bmatrix}, \quad X_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \kappa_t^m & 0 \\ 0 & 0 & 0 & \kappa_t^v \end{bmatrix}, \quad \Theta = \begin{bmatrix} \theta^m \\ \theta^v \\ a^m \\ a^v \\ c^m \\ c^v \end{bmatrix}, \quad Z_t = \begin{bmatrix} \epsilon_{t+1}^m \\ \epsilon_{t+1}^v \\ \delta_{t+1}^m \\ \delta_{t+1}^v \end{bmatrix}.$$

- Combination with  $(A_x^g, B_x^g, K_t^g, \alpha_x^g, \beta_x^g, \kappa_t^g)$  gives  $\mu_{xtg}^{\text{NL}}$  and thus  $q_{xtg}^{\text{NL}}$ .

## Simulating Scenarios for Mortality

Model does not (just) generate a table, but stochastic scenario generator (which complements those for stock prices, interest rates etcetera).

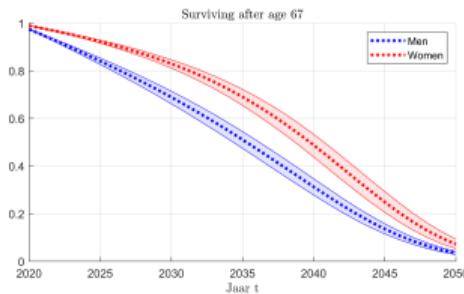
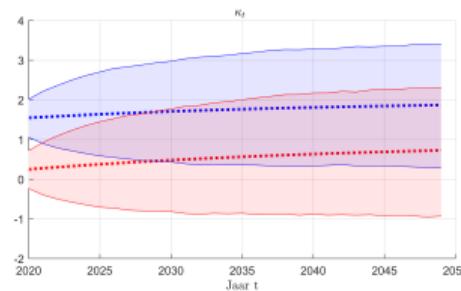
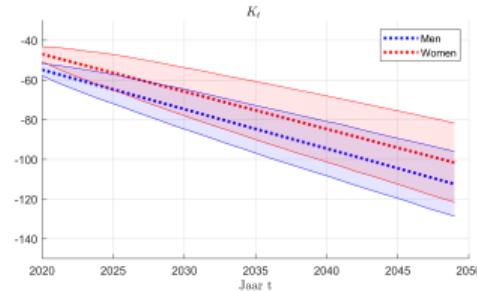
- To start simulating we also need, apart from blue parameters and entries of covariance matrix  $C$ , the starting values  $K_{t_0}$  and  $\kappa_{t_0}$ .
- Every new draw of the sequence of normal random variables generates a new stochastic life table.
- By simulating many times one can find expected values, variances, quantiles etcetera of relevant quantities such as life expectancies, annuity values or the NPV for a large portfolio of insurance products.
- If one just wants to have a best estimate of the log mortality rates (to create a single table) one can simply put all the values of  $\epsilon_t$  and  $\delta_t$  equal to zero. This gives the most likely path for the stochastic hazard rates.

# Determining Quantiles

- Simulation of distribution of relevant quantity generates quantiles. This can usually **not** be done by
  - determining the quantiles of the life table (i.e. by determining a 2.5% life table and 97.5% life table of dying probabilities),
  - and then using these tables to find the quantiles for the values of liabilities in a certain insurance portfolio.
- Compare: if  $Z$  is standard Gaussian, its quantiles for 2.5% and 97.5% are at  $-1.96$  and  $1.96$ . But the quantiles of  $Z^2$  for 2.5% and 97.5% are **not** at  $3.84$  and  $3.84$ . That would mean 95% of probability mass would be at  $3.84$ !
- Quantiles cannot be calculated by substitution under nonlinear transformations. Instead, simulate scenarios and apply them to portfolio values, this can never go wrong!

# Nonlinear Transformations

Model AG2022 Pre-COVID: Time series and Projections  
EU/POL Data 1970-2019



Best estimates and 95% confidence intervals.

# Period versus Cohort Tables

Actuarieel Genootschap & Actuarieel Instituut  
AG&AI

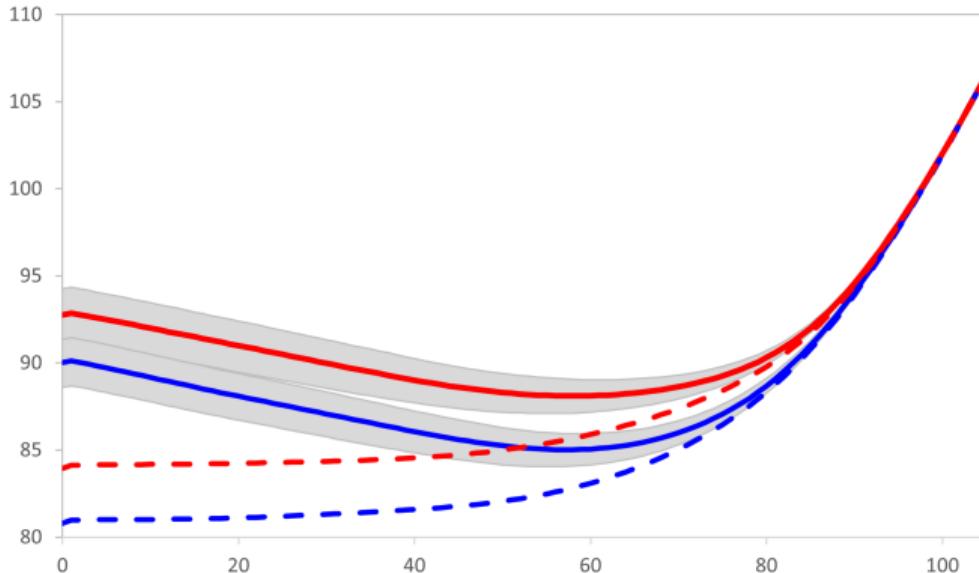
		C	D	E	F	G	H	I	J	K	L
	0	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
2	0	0,002221	0,002147	0,002075	0,002004	0,001935	0,001868	0,001803	0,001739	0,001677	0,001617
3	1	0,000198	0,000192	0,000185	0,000179	0,000172	0,000166	0,000161	0,000155	0,000150	0,000144
4	2	0,000125	0,000121	0,000118	0,000114	0,000110	0,000107	0,000103	0,000100	0,000096	0,000093
5	3	0,000087	0,000084	0,000081	0,000078	0,000075	0,000073	0,000070	0,000067	0,000065	0,000062
6	4	0,000067	0,000065	0,000062	0,000060	0,000058	0,000054	0,000052	0,000050	0,000048	0,000046
7	5	0,000061	0,000059	0,000057	0,000055	0,000053	0,000051	0,000049	0,000047	0,000045	0,000043
8	6	0,000055	0,000052	0,000050	0,000048	0,000046	0,000044	0,000042	0,000041	0,000039	0,000037
9	7	0,000049	0,000048	0,000046	0,000044	0,000043	0,000041	0,000040	0,000038	0,000037	0,000036
10	8	0,000041	0,000040	0,000047	0,000046	0,000044	0,000042	0,000041	0,000039	0,000038	0,000037
11	9	0,000047	0,000044	0,000043	0,000041	0,000040	0,000038	0,000037	0,000036	0,000034	0,000033
12	10	0,000054	0,000053	0,000051	0,000049	0,000048	0,000046	0,000044	0,000043	0,000041	0,000040
13	11	0,000059	0,000057	0,000055	0,000053	0,000051	0,000050	0,000048	0,000046	0,000045	0,000043
14	12	0,000071	0,000069	0,000068	0,000066	0,000065	0,000063	0,000062	0,000060	0,000059	0,000057
15	13	0,000085	0,000083	0,000081	0,000079	0,000077	0,000075	0,000073	0,000071	0,000069	0,000067
16	14	0,000100	0,000097	0,000095	0,000092	0,000090	0,000088	0,000085	0,000083	0,000081	0,000079
17	15	0,000125	0,000122	0,000119	0,000116	0,000113	0,000110	0,000107	0,000105	0,000102	0,000099
18	16	0,000178	0,000174	0,000170	0,000166	0,000162	0,000158	0,000154	0,000150	0,000146	0,000142
19	17	0,000197	0,000193	0,000188	0,000183	0,000179	0,000174	0,000170	0,000165	0,000161	0,000156
20	18	0,000244	0,000239	0,000234	0,000229	0,000224	0,000219	0,000214	0,000210	0,000205	0,000200
21	19	0,000294	0,000288	0,000283	0,000277	0,000272	0,000266	0,000260	0,000255	0,000249	0,000243
22	20	0,000321	0,000315	0,000309	0,000303	0,000297	0,000291	0,000285	0,000279	0,000273	0,000267

When calculating probability of survival over  $k$  years in 2024 do we incorporate **current** or **projected** survival probabilities in later years?

$$e_x^{\text{per}}(2024) = \frac{1}{2} + \sum_{k=0}^{\infty} \prod_{i=0}^k (1 - q_{x+i}(2024)),$$

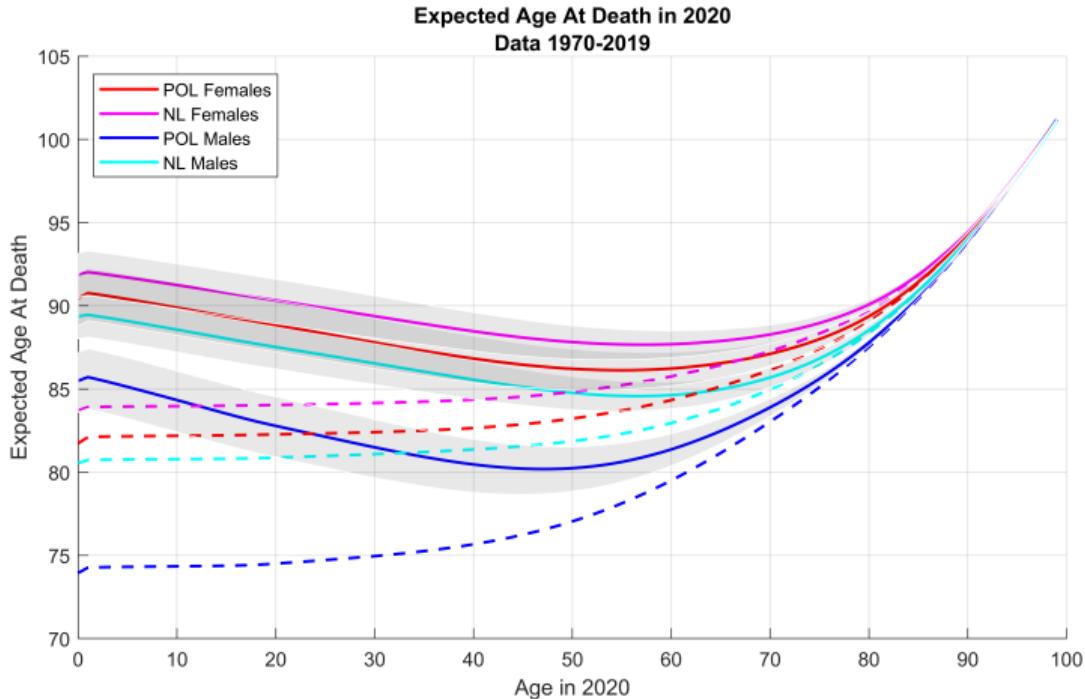
$$e_x^{\text{coh}}(2024) = \frac{1}{2} + \sum_{k=0}^{\infty} \prod_{i=0}^k (1 - q_{x+i}(2024 + i)).$$

# Prognosis and Uncertainty



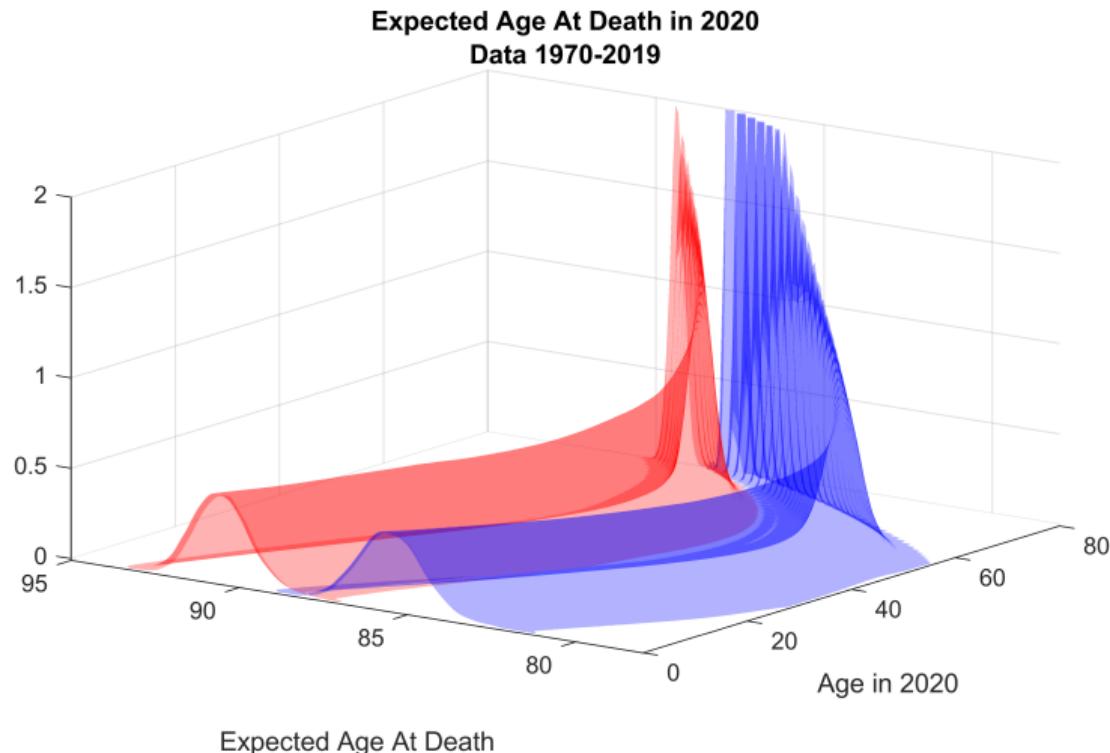
Expected age at death given age (in 2022)  
Dutch Females (red) and Males (blue),      Period (dashed line) and Cohort (line).

# Prognosis and Uncertainty



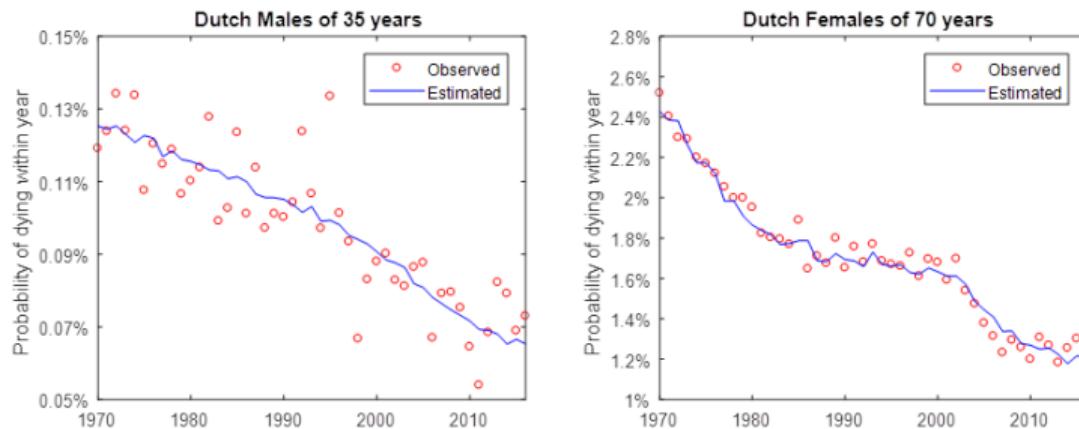
Best estimate age at death given age **in 2020** and 95% confidence intervals.

# Prognosis and Uncertainty



Distribution of age at death given age in 2022 (POL females/males).

# Filtering and Smoothing



More noise under 35-year olds (uncertainty in mortality **observations** does not imply more uncertainty in death **probabilities**).

Notice that **degree of smoothing is age-dependent**, since information from different ages is combined with weights to extract best statistical information from data.

# Filtering and Smoothing

- More fluctuations at lower ages: in the Netherlands in 2017
  - 13 of the  $\pm 100.000$  girls with age 15 died, while
  - 909 of the  $\pm 101.000$  women with age 67 died.
- One extra death would thus amount to
  - a relative change of 7.7% among girls with age 15, and
  - a relative change of 0.11% among women with age 67.
- We meet **death frequencies**, not **death probabilities** and measurement noise varies with age.
- Maximal likelihood methods incorporate this information in the estimation procedure, by assigning different weights to information from different age groups.

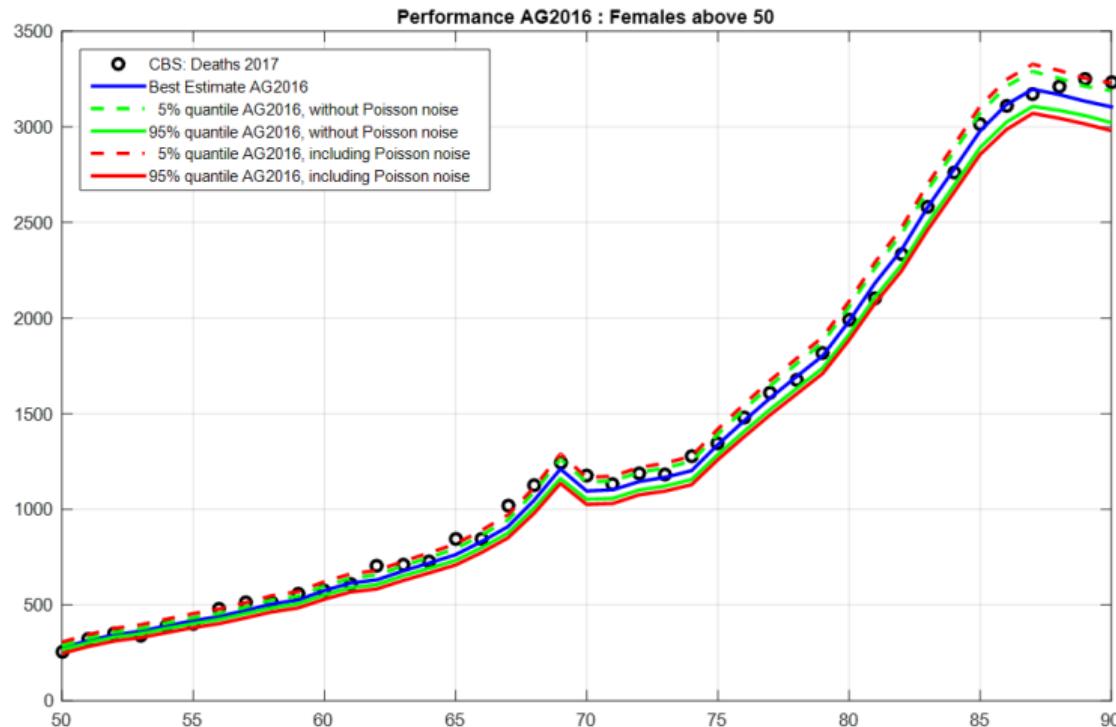
# Performance criteria

---

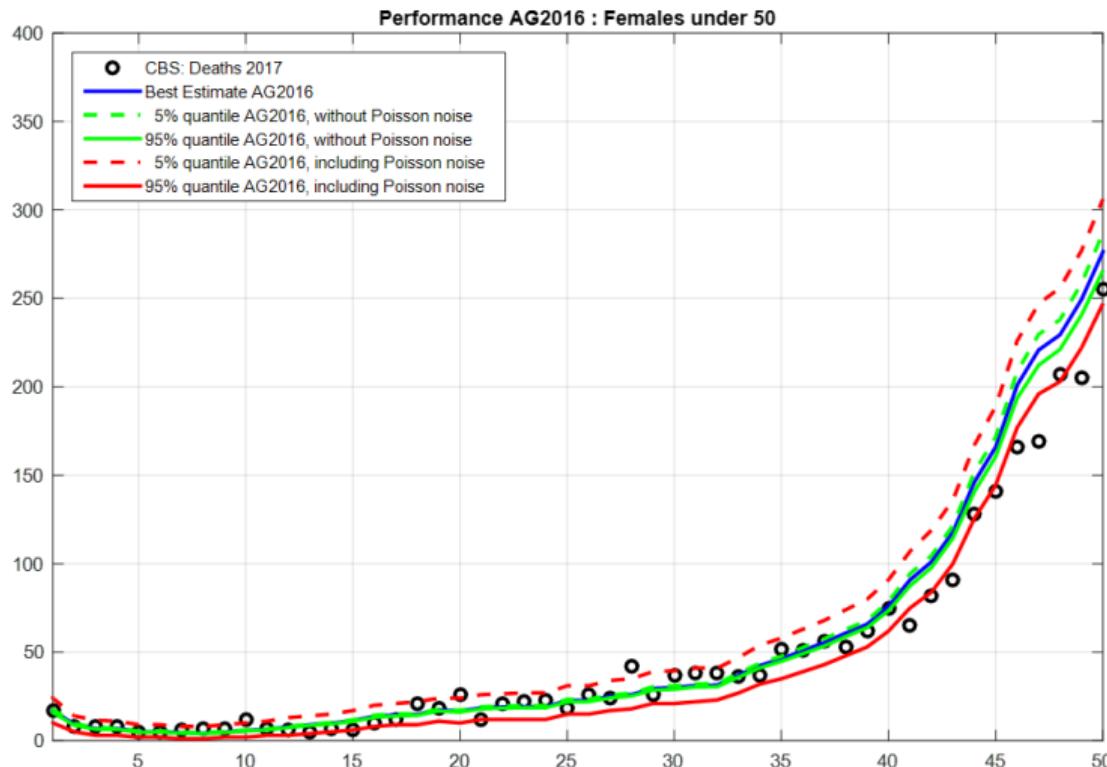
## Considerations when choosing a model

- Goodness-of-fit (in-sample)
- Backtest (out-of-sample fit)

# Results AG2016 in 2017



# Results AG2016 in 2017



## Considerations when choosing a model

- Goodness-of-fit (in-sample)
- Backtest (out-of-sample fit)
- Parsimony
- Time consistency

## Time Consistency

- Maximal likelihood estimate for parameters  $\hat{\Theta}$ ,  $\hat{C}$ :

$$(\hat{\Theta}, \hat{C}) = \arg \max_{(\Theta, C)} \left( -\frac{1}{2} \text{tr}[C^{-1} \sum_{t=1}^{T-1} (Y_t - X_t \Theta)(Y_t - X_t \Theta)'] - \frac{T-1}{2} \ln(2\pi|C|) \right).$$

- If realised time series observation for  $t = T$  matches estimated value,  $Y_T = X_T \hat{\Theta}$ , and we leave the estimated covariance matrix unchanged,  $C = \hat{C}$ , then updated estimate for parameter vector is found as

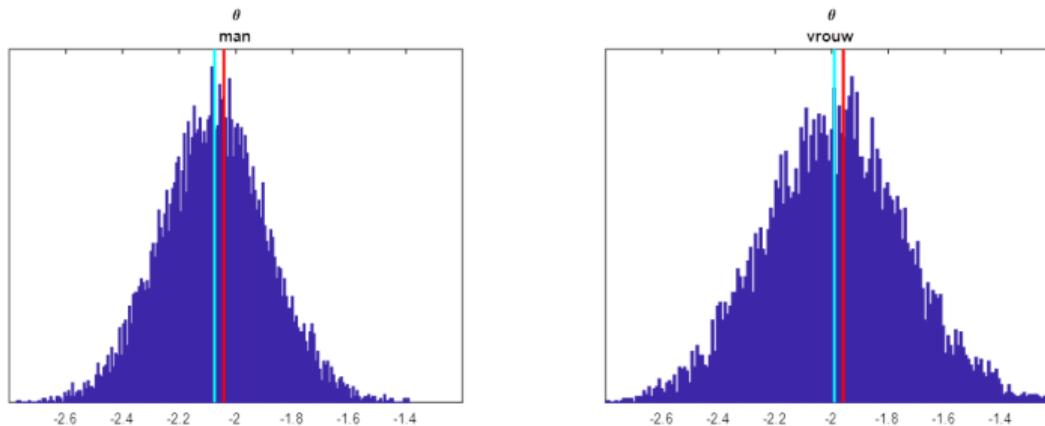
$$\begin{aligned} \hat{\Theta}_{\text{new}} &= \arg \max_{\Theta} \left( -\frac{1}{2} \text{tr}[\hat{C}^{-1} \sum_{t=1}^{T-1} (Y_t - X_t \Theta)(Y_t - X_t \Theta)'] \right. \\ &\quad \left. - \frac{1}{2} \text{tr}[X_T' \hat{C}^{-1} X_T (\hat{\Theta} - \Theta)(\hat{\Theta} - \Theta)'] \right). \end{aligned}$$

- But since  $\Theta = \hat{\Theta}$  maximizes both terms in the brackets, we find that  $\hat{\Theta}_{\text{new}} = \hat{\Theta}$ .

## Considerations when choosing a model

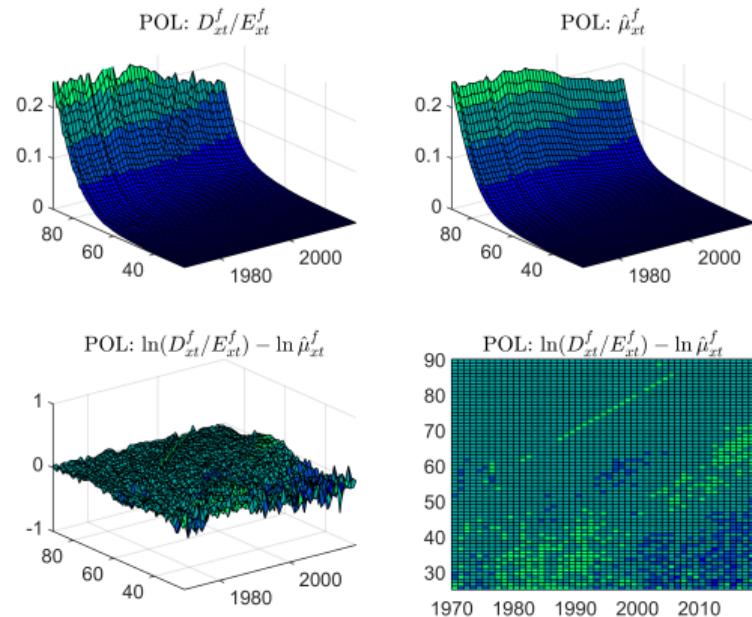
- Goodness-of-fit (in-sample)
- Backtest (out-of-sample fit)
- Parsimony
- Time consistency
- Visual inspection dynamics of dying probabilities
- Robustness

# Parameter Uncertainty: Resampling Methods



- Many possible numbers of deaths were simulated for model (here: AG2018), using the calibrated parameter values.
- For each of these samples, the parameters are recorded that would be found if that sample had been used to calibrate the model.

## Other approaches : Cohort Effects Poland



**Advantage:** Specific effect per generation taken into account.

**Disadvantage:** May be hard to project.

## Other approaches : Dynamic Models in continuous time

- Time inhomogeneous **Markov jump process**: finite states  $\{1, 2, \dots, p\}$  plus absorbing death state  $\{p + 1\}$ , initial distribution  $\pi$  on  $p$  first states, and intensity a function  $\lambda(t)$  times fixed  $p \times p$  matrix  $\mathbf{T}$ . Distribution remaining time  $\tau$  until we reach death state:

$$\mathbb{P}(\tau \geq T | \tau \geq t) = \pi' e^{-\int_t^T \lambda(u) du \mathbf{T}} \mathbf{e}$$

with  $\mathbf{e}$  a vector of  $p$  ones.

- Choice

$$\int_t^T \lambda(u) du = \int_t^T e^{\beta(s-t)} ds = \beta^{-1}(e^{\beta(T-t)} - 1)$$

gives **matrix-Gompertz** distribution (Albrecher et al. 2022).

- **Advantage**: exploits structure of multiple states before death.  
**Disadvantage**: interpretation 'states' may be difficult.

## Other approaches : Dynamic Models in continuous time

- **Affine stochastic models** for interest rate can be relabeled to model intensity.
- On  $\mathcal{G}_t$ , the information set containing all actuarial information apart from the time of death, we may take  $\mu_t = a_0 + a'_1 X_t$  for affine process

$$dX_t = (\zeta(t) - LX_t)dt + \Sigma(G_0 + \sum_i G_i(X_t)_i)^{\frac{1}{2}} dW_t,$$

with  $W$  Brownian motion,  $L$ ,  $\Sigma$ ,  $G_i$ ,  $G_0$  matrices and  $\zeta$  a deterministic function.

- Remaining lifetime  $\tau$  for intensity process  $\mu_t$  satisfies for  $T \geq t \geq 0$

$$\mathbb{P}(\tau \geq T | \tau \geq t) = \mathbb{E}[e^{-\int_t^T \mu_s ds} | \mathcal{G}_t] = \mathbb{E}[e^{-a_0(T-t) - a'_1 \int_t^T X_s ds} | \mathcal{G}_t] = e^{A(t, T) + B(t, T)' X_t}$$

with

$$\partial_t B(t, T)_i = -\frac{1}{2} B(t, T)' \Sigma G_i \Sigma' B(t, T) + (L' B(t, T))_i - (a_1)_i, \quad B(T, T) = 0,$$

$$A(t, T) = \int_t^T (B(s, T)' \zeta(s) + \frac{1}{2} B(s, T)' \Sigma G_0 \Sigma' B(s, T) - a_0) ds.$$

## Other approaches : Dynamic Models in continuous time

- Log-linear characteristic functions for (integrated) process results in

$$\mathbb{P}(\tau \geq T | \tau \geq t) = e^{A(t, T) + B(t, T)' X_t}$$

and thus in closed form expressions (in terms of solutions of ordinary differential equations).

- Different structures proposed, such as multivariate autoregressive process, CIR square root processes etc.: (Milevsky & Promislow, 2001), (Blackburn & Sherris, 2013), (Jetvić & Regis, 2019), (Ungolo et al. 2023).
- **Advantage:** flexible structure, Kalman filtering for calibration.  
**Disadvantage:** information on mortality may not arrive continuously.

# Challenges

- Implicit assumption is always that order of magnitude for future shocks is **comparable to those in the past**
- Realizations can deviate in future due to
  - Behavioral, social-economic or ethical developments which differ from what has been the case historically
  - Medical developments and successes in fighting of diseases
  - Possible effects of as yet unknown viruses and bacteria, such as **COVID-19**.

## Key References

- Albrecher H., Bladt, M., Blad, M. and Yslas, J.. (2022), Mortality modeling and regression with matrix distributions. *Insurance: Mathematics and Economics* 107, 68-87.
- Blackburn, C. and Sherris, M. (2013), Consistent dynamic affine mortality models for longevity risk applications. *Insurance: Mathematics and Economics* 53(1), 64-73.
- Koninklijk Actuarieel Genootschap (2022) Projection Table AG2022.
- Jevtić, P. and Regis, L. (2019) A continuous-time stochastic model for the mortality surface of multiple populations. *Insurance: Mathematics and Economics* 88, 299-318.
- Milevsky, M.A. and Promoslow, M.D. (2001). Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics* 29(3), 181-195.
- Ungolo, F., Garces, L. P. D. M., Sherris, M., and Zhou, (2023). Estimation, Comparison, and Projection of Multifactor Age–Cohort Affine Mortality Models. *North American Actuarial Journal*, 1-23.