

Robust Dynamic mortality modelling under Structural Shocks

Karen Y. Barrera

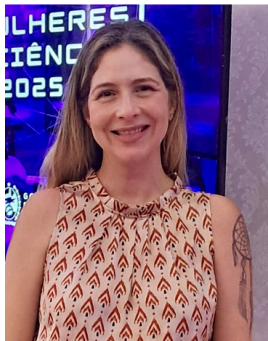
Joint with Thaís C. Fonseca and Viviana G. R. Lobo

Institute of Mathematics, Universidade Federal do Rio de Janeiro

September 08, 2026



Joint work with:

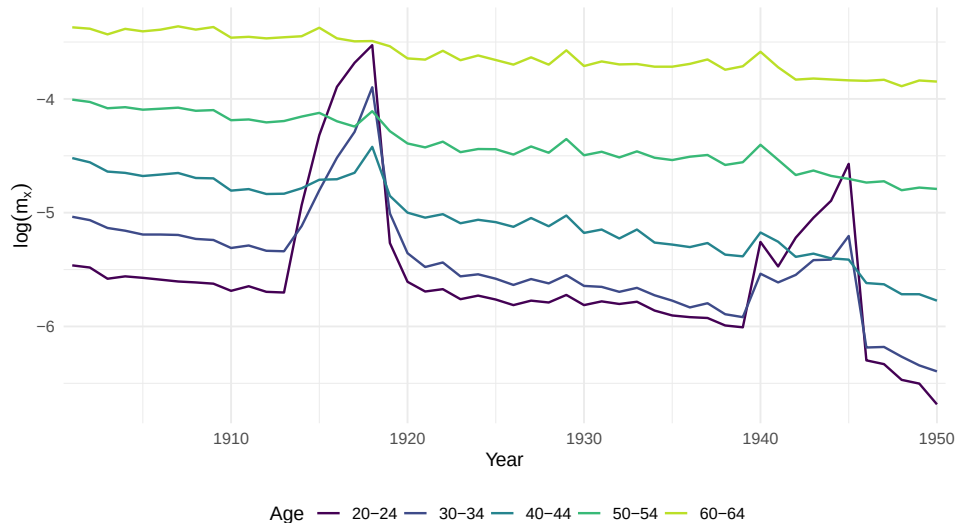


Thais C. Fonseca (IM/UFRJ)

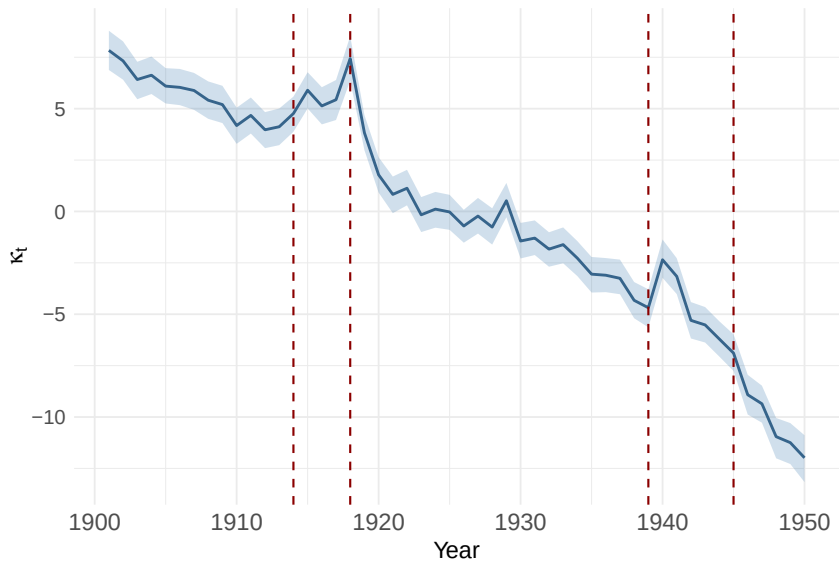


Viviana Lobo (IM/UFRJ)

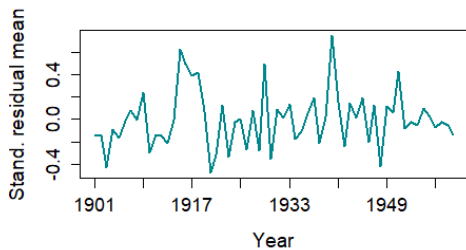
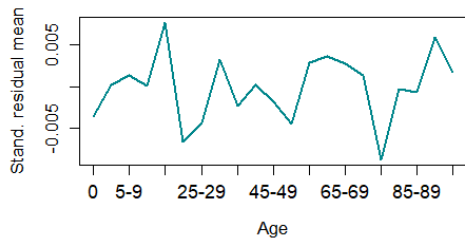
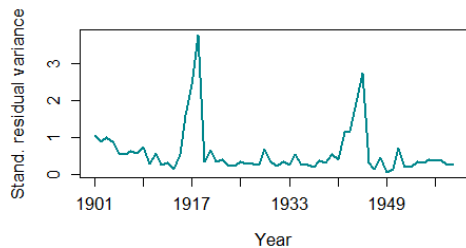
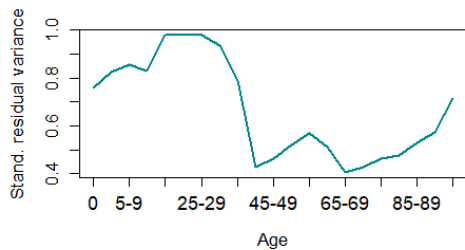
Motivation: Mortality shocks are not ordinary forecast errors



Posterior mean of the mortality index



Standardized Residuals of a mortality model



Literature: Existing models address different parts of the problem

- ▶ **Liu–Li(2015)**: separates occurrence from severity, but π is **fixed** and N_t is serially independent.
- ▶ **Goes et al.(2025)**: adds AR/MA structure so a *realised* shock persists and decays, but the **probability of a new shock** is still constant.
- ▶ **Scale-mixture** literature: handles unusual residual volatility, but has no shock/jump mechanism at all.

How can we robustify mortality forecasts?

1. Estimate the variance robustly
 - 1.1 Scale-mixing parameter
2. Remove distortion in parameter estimates
 - 2.1 Include a mechanism to capture and incorporate mortality shocks.
 - 2.2 Dynamic probability of mortality shock occurrence

$$\text{observational mortality} = \text{trend} + \text{shock} + \text{variability}$$

Lee-Carter with Age Heterogeneity

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\alpha} + \boldsymbol{\beta}\kappa_t + \boldsymbol{\varepsilon}_t, & \boldsymbol{\varepsilon}_t &\sim N_A(\mathbf{0}, \boldsymbol{\Sigma}) \\ \kappa_t &= \kappa_{t-1} + \theta + w_{1t}, & w_{1t} &\sim N(0, \sigma_w^2) \end{aligned}$$

Based on: Pedroza (2006)

where:

- ▶ $\mathbf{y}_t = (y_{1t}, y_{2t}, \dots, y_{At})'$: Observed log-mortality rates,
- ▶ $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_A)'$: Age-specific intercept,
- ▶ $\boldsymbol{\beta} = (\beta_1, \dots, \beta_A)'$: Age sensitivity to changes in κ_t ,
- ▶ $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_A^2)$: Age-specific variance

Limitation: Still assumes **constant variance over time**

The Liu-Li model

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\alpha} + \boldsymbol{\beta} \kappa_t + \boldsymbol{\beta}^{(J)} N_t Y_t + \boldsymbol{\varepsilon}_t, & \boldsymbol{\varepsilon}_t &\sim N_A(\mathbf{0}, \boldsymbol{\Sigma}) \\ \kappa_t &= \kappa_{t-1} + \theta + w_{1t}, & w_{1t} &\sim N(0, \sigma_w^2) \end{aligned}$$

where:

- ▶ $N_t \sim \text{Bernoulli}(\pi)$: Jump indicator,
- ▶ $Y_t \sim N(\mu_Y, \sigma_Y^2)$: Jump severity,
- ▶ $\boldsymbol{\beta}^{(J)} = (\beta_1^{(J)}, \dots, \beta_A^{(J)})'$: Age-specific impact to jump events.

Limitation: variance and a probability of jump occurrence remain **constant** over time.

Two Bayesian proposals

1. **Time-varying shock probability** π_t : dynamic Beta evolution (Fonseca & Ferreira, 2017)
 - ▶ Induces **memory** in the filtered probability: a shock in the present raises the *propensity* for a shock in the years that follow
 - ▶ Nests Liu & Li (2015) exactly as the boundary case $\delta = 1$
2. **Dynamic observational precision** λ_t : scale-mixture formulation capturing temporal heteroscedasticity (Fonseca, Lobo and Schmidt, 2023)
 - ▶ Lets **periods** of unusual volatility be recognized as such, without forcing that variation into the jump component

Proposal 1: dynamic shock probability

$$\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}\kappa_t + \boxed{\boldsymbol{\beta}^{(J)} N_t Y_t} + \boldsymbol{\varepsilon}_t,$$

$$\kappa_t = \kappa_{t-1} + \theta + w_t,$$

$$N_t \mid \pi_t \sim \text{Bernoulli}(\boxed{\pi_t}),$$

$$\boldsymbol{\varepsilon}_t \sim \mathcal{N}_A(\mathbf{0}, \boxed{\boldsymbol{\Sigma}})$$

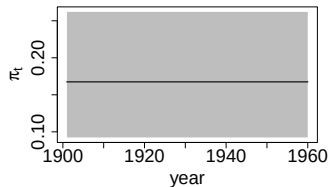
$$w_t \sim \mathcal{N}(0, \sigma_w^2)$$

$$\pi_t = \frac{\phi_{1,t}\pi_{t-1}}{\phi_{1,t}\pi_{t-1} + \phi_{2,t}(1 - \pi_{t-1})}$$

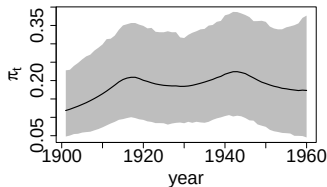
- ▶ $\phi_{1,t} \sim \text{Beta}(\delta a_{t-1}, (1 - \delta)a_{t-1})$, $\phi_{2,t} \sim \text{Beta}(\delta b_{t-1}, (1 - \delta)b_{t-1})$
- ▶ $\delta \in (0, 1]$ is a discount factor - **varying smoothness**

How does δ work?

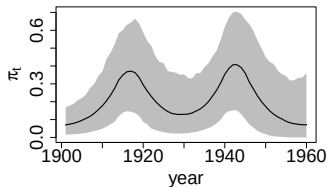
$$\delta = 1$$



$$\delta = 0.95$$



$$\delta = 0.85$$



- ▶ **Small δ :** π_t swings quickly with recent shocks or their absence
- ▶ **Large δ :** π_t moves slowly and smoothly
- ▶ $\delta = 1$: every year contributes equally, and we recover the **constant-probability** Liu & Li (2015) model exactly.

Proposal 2: dynamic stochastic variance

$$\mathbf{y}_t = \boldsymbol{\alpha} + \beta \kappa_t + \boxed{\beta^{(J)} N_t Y_t} + \frac{\boldsymbol{\varepsilon}_t}{\boxed{\sqrt{\lambda_t}}}, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}_A(\mathbf{0}, \boldsymbol{\Sigma}),$$

$$\kappa_t = \kappa_{t-1} + \theta + w_{1t},$$

$$w_{1t} \sim \mathcal{N}(0, \sigma_w^2),$$

$$N_t | \pi_t \sim \text{Bernoulli}(\boxed{\pi_t}),$$

$$\pi_t = \frac{\phi_{1,t} \pi_{t-1}}{\phi_{1,t} \pi_{t-1} + \phi_{2,t} (1 - \pi_{t-1})}$$

$$\ln(\lambda_t) = \eta_t + v_{1t},$$

$$v_{1t} \sim \mathcal{N}\left(-\frac{\nu}{2}, \nu\right),$$

$$\eta_t = \eta_{t-1} + w_{2t},$$

$$w_{2t} \sim \mathcal{N}(0, W_\eta).$$

How does λ_t work?

- ▶ If a year looks **unusually noisy**, λ_t **inflates** the observational variance
- ▶ If a year looks **normal**, λ_t stays **near one**, and the model behaves just like the standard Gaussian case
- ▶ By letting η_t evolve smoothly, the model recognizes that **volatility itself has regimes**, a noisy year raises the expected volatility of the years that follow

What about ν ?

- ▶ ν controls **how far λ_t is allowed to wander** from one.
- ▶ Small $\nu \Rightarrow \lambda_t$ stays tightly clustered around 1 $\Rightarrow$ model behaves almost exactly like a standard Gaussian
- ▶ Large $\nu \Rightarrow \lambda_t$ can swing much further from 1 $\Rightarrow$ model tolerates (and can detect) much more extreme deviation from baseline variance.

Bayesian Estimation: Two-Block MCMC

Block 1: Conditional on $\beta^{(j)(j-1)}, N_t^{(j-1)}, Y_t^{(j-1)}$ define:

$$\tilde{\mathbf{y}}_t = \mathbf{y}_t - \beta^{(j)(j-1)} N_t^{(j-1)} Y_t^{(j-1)}$$

- ▶ Sample $\kappa_{1:T}^{(j)}$ and $\eta_{1:T}^{(j)}$ via FFBS.
- ▶ Sample $\lambda_t^{(j)}$ via Metropolis-Hastings
- ▶ Sample static parameters $(\alpha^{(j)}, \beta^{(j)}, \Sigma^{(j)}, \theta^{(j)}, \sigma_w^2^{(j)}, \nu^{(j)})$

Block 2: Conditional on $\alpha^{(j)}, \beta^{(j)}, \kappa_t^{(j)}$ define:

$$\bar{\mathbf{y}}_t = \mathbf{y}_t - \alpha^{(j)} - \beta^{(j)} \kappa_t^{(j)}$$

- ▶ Sample jump indicator $N_{1:T}^{(j)}$ and magnitude $Y_{1:T}^{(j)}$
- ▶ Sample $\pi_{1:T}^{(j)}$ via FFBS
- ▶ Sample $\beta^{(j)(j)}$

Application

1. Data:

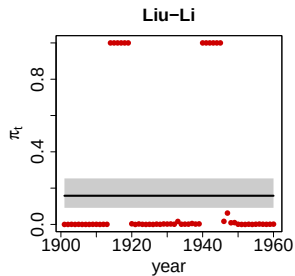
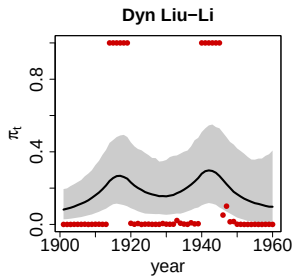
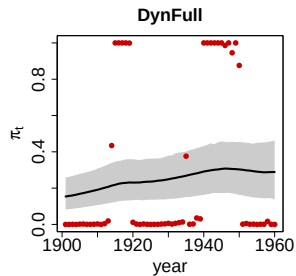
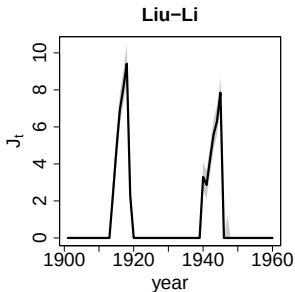
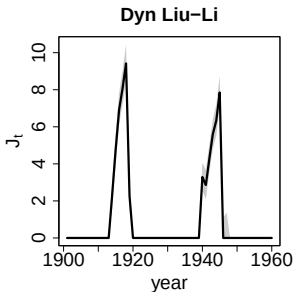
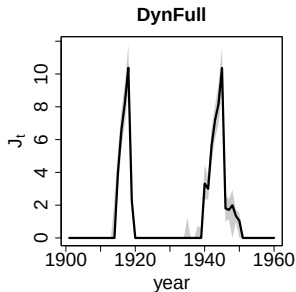
1.1 England and Wales mortality, Human Mortality Database (2026)

- ▶ 21 age groups (0-1 to 95-99)
- ▶ Years 1901-1960
- ▶ Covers WWI (1914-1918) and WWII (1939-1945)

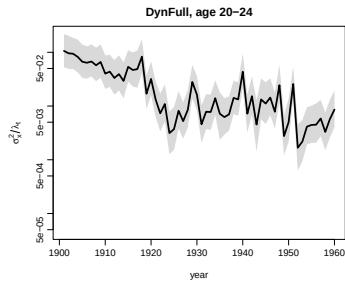
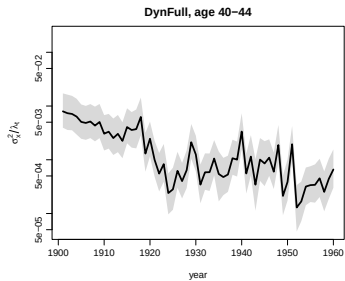
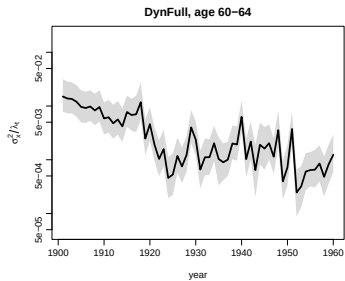
1.2 USA mortality, Human Mortality Database (2026)

- ▶ 21 age groups (0-1 to 95-99)
- ▶ Years 1970-2020
- ▶ Covers the COVID-19 pandemic

Results: England and Wales (1901-1960)

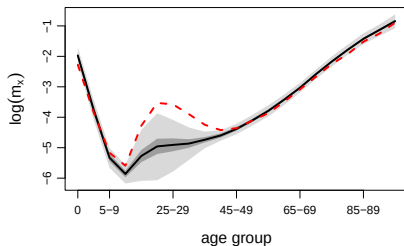


Variance σ_x^2/λ_t

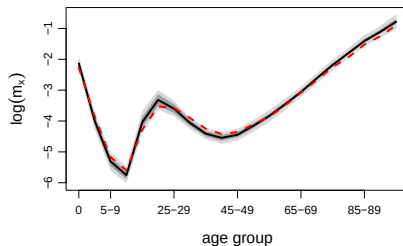


Fitted Mortality Curves (War year)

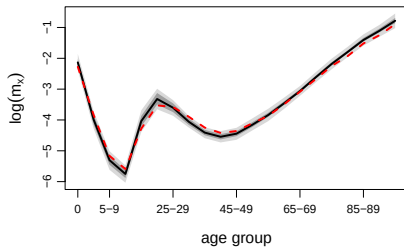
LC (1918)



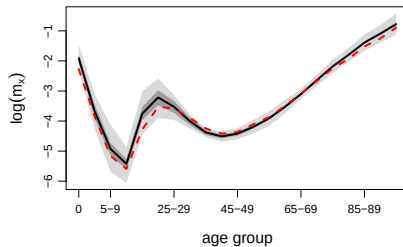
Liu-Li (1918)



Dyn Liu-Li (1918)

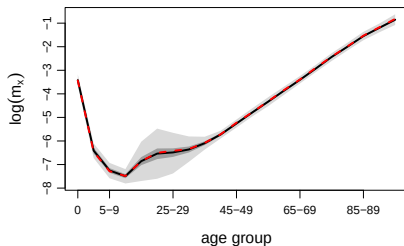


DynFull (1918)

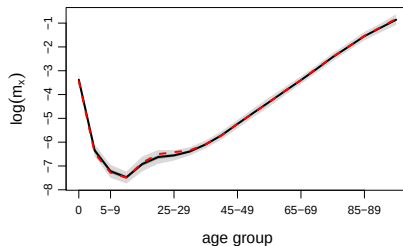


Fitted Mortality Curves (Normal year)

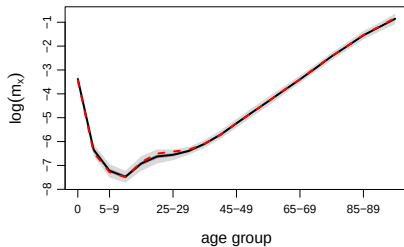
LC (1949)



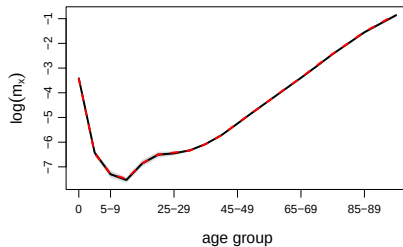
Liu-Li (1949)



Dyn Liu-Li (1949)



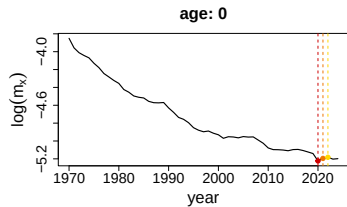
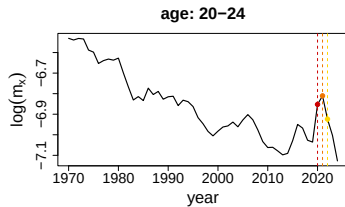
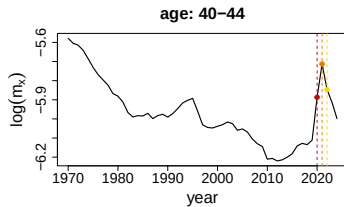
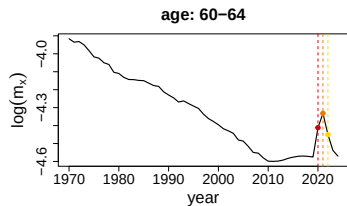
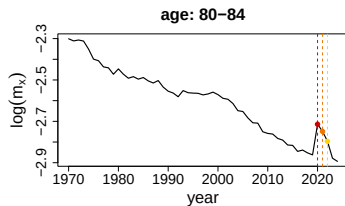
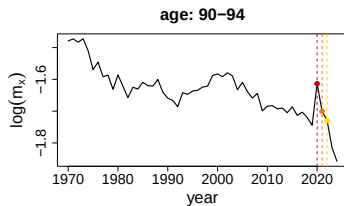
DynFull (1949)



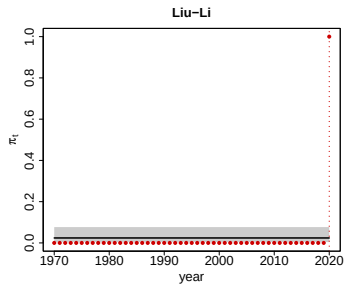
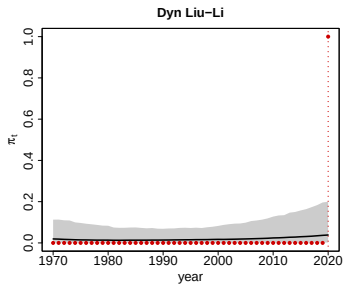
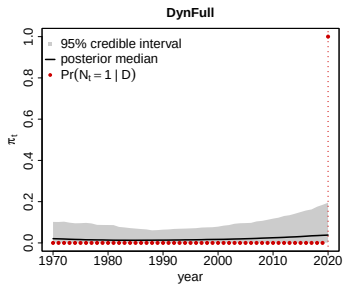
In-sample goodness-of-fit metrics

Model	RMSE	MAE	WAIC
LC	0.19176	0.10078	-1925.463
Liu-Li	0.07771	0.06250	-2688.169
Dyn Liu-Li	0.07763	0.06244	-2687.931
DynFull	0.10310	0.06612	-3352.381

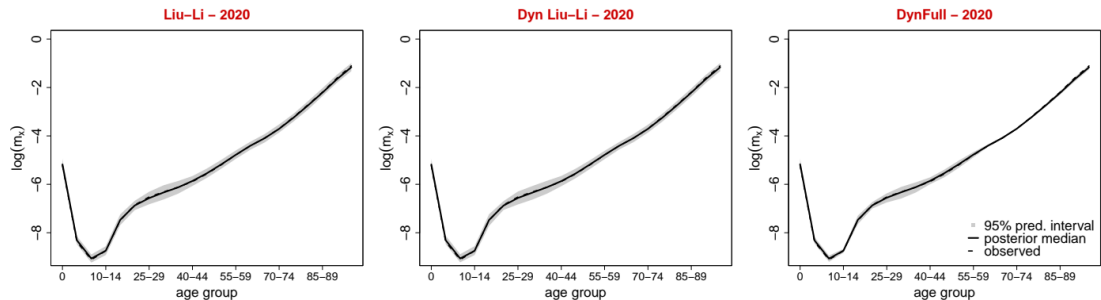
Results: USA (1970-2020)



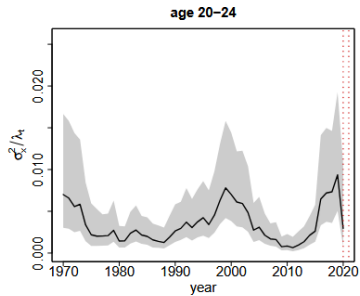
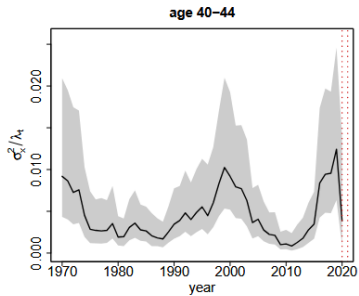
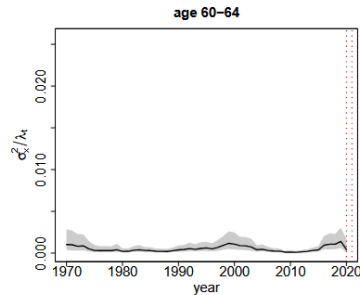
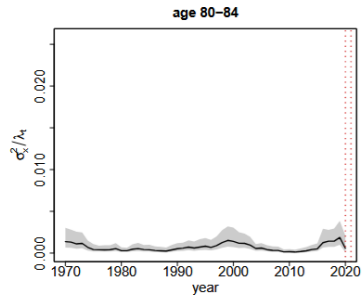
Probability of mortality shock occurrence



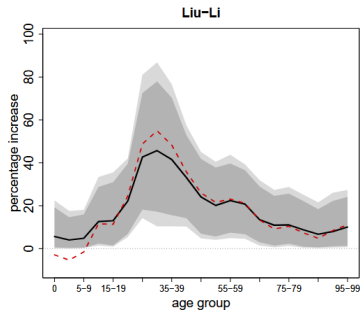
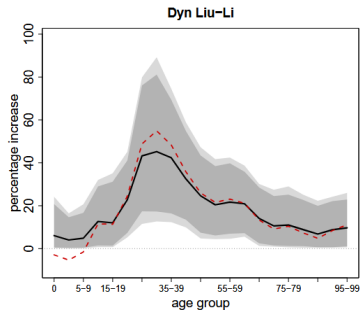
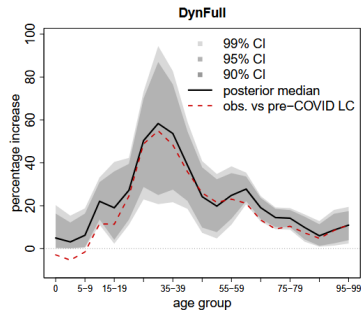
Fitted Mortality Curve



Variance σ_x^2/λ_t

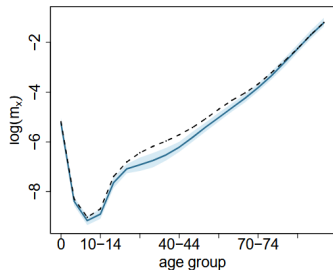


Observed percentage increase in death rates from the 1970 to 2020

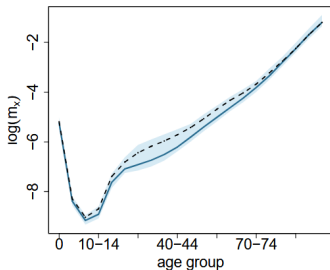


Forecast two years ahead

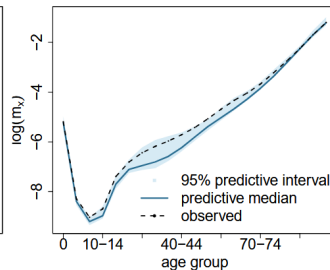
Liu-Li - 2021



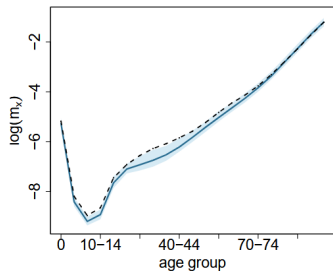
Dyn Liu-Li - 2021



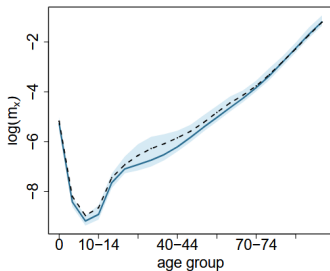
DynFull - 2021



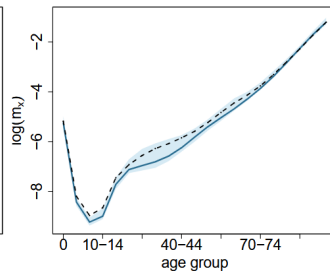
Liu-Li - 2022



Dyn Liu-Li - 2022



DynFull - 2022



Goodness-of-fit comparison, scenario 1970–2020

Model	WAIC	LOO
Lee–Carter	−2799.688	−2798.907
Liu and Li	−3076.436	−2880.279
Dyn Liu and Li	−3073.772	−2881.703
DynFull	− 3617.359	− 3603.601

Thank you!

