Modeling mortality with pandemics: a Bayesian transitory vanishing jump approach

Julius Goes [1] Anne Leucht [1] Karim Barigou [2]

[1] Institute of Statistics – University Bamberg

[2] Ecole d'actuariat – Université Laval

Longevity 18
08. Sep.23

Institut für Statistik
Universität Bamberg

UNIVERSITÉ LAVAL

Table of Contents

1. Introduction
2. Model Proposition
3. Estimation Procedure
4. Empirical Results
5. Outlook

Motivation I

- The Lee-Carter approach by Lee and Carter (1992) is among the earliest model for stochastic mortality forecasting. With many extensions thereof in recent years.
- Lee-Carter model is effective in capturing a general mortality trend
  $\rightarrow$ does not account for large, unexpected jumps caused by e.g. pandemics or wars.
- Lee and Carter (1992) treat the 1918 influenza (spanish flu) as an anomaly and use an intervention model, effectively removing its influence.

Motivation II

- Earliest methods relied on outlier techniques (e.g. Li & Chan, 2005) and remove potential data points.
- H. Chen and Cox (2009) incorporate a jump processes into the Lee-Carter model.
  $\rightarrow$ either permanent or transitory mortality shocks.
- Adaptation by Liu and Li (2015) to account for age dependent shock.
- More recently, other approaches haven been introduced (e.g. F.-Y. Chen et al., 2022; Robben & Antonio, 2023; van Berkum et al., 2022).
  $\rightarrow$ All assume that the shock completely disappears the next period
- We introduce a model, where the shock effect slowly vanishes over time.

Vanishing Effects

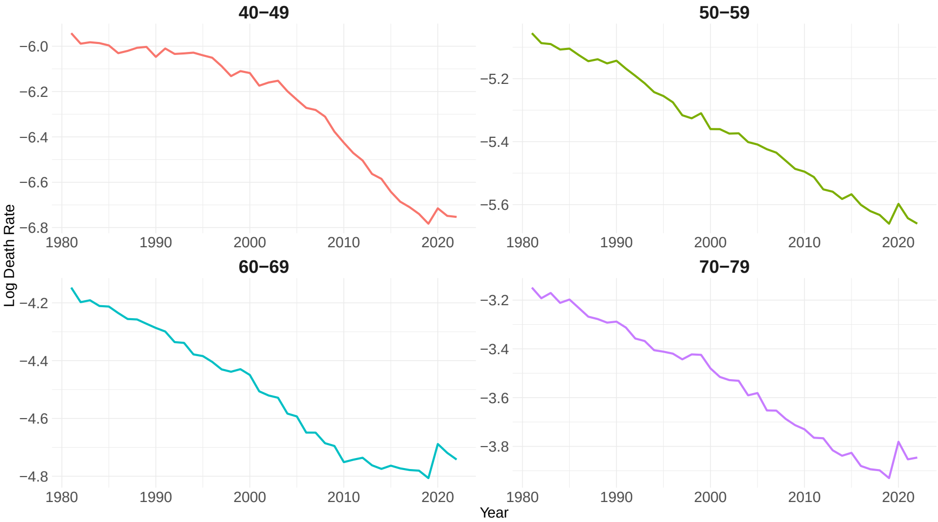

Figure: Log Death Rates for Spain

Lee-Carter Model

- Let $D_{x,t}$ denote the total number of deaths and $E_{x,t}$ central exposure to risk, with $x = x_1, x_2, \ldots, x_A$ and $t = t_1, t_2, \ldots, t_T$.
- We denote $m_{x,t}$ to be the central mortality rate given by

$$m_{x,t} = \frac{D_{xt}}{E_{xt}}.$$

- The Lee-Carter model including a centered error term is as follows

$$\ln\left(m_{x,t}\right) = \alpha_x + \beta_x \kappa_t + e_{x,t},$$

where $e_{x,t} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_e^2\right)$.

Some Shock Extensions

- Extension of κ_t by H. Chen and Cox (2009) that allows for transitory effects.
- Adaption of Chen and Cox by Liu and Li (2015) to account for an age dependent shock.
- Liu and Li (2015) defined three different variants of their shock model. The one closest to our adaptation is given by

$$\ln(m_{x,t}) = \alpha_x + \beta_x \kappa_t + \qquad\qquad + e_{x,t}. \tag{1}$$

Some Shock Extensions

- Extension of κ_t by H. Chen and Cox (2009) that allows for transitory effects.
- Adaption of Chen and Cox by Liu and Li (2015) to account for an age dependent shock.
- Liu and Li (2015) defined three different variants of their shock model. The one closest to our adaptation is given by

$$\ln(m_{x,t}) = \alpha_x + \beta_x \kappa_t + \beta_x^{(J)} N_t Y_t + e_{x,t}. \tag{1}$$

- $N_t \in \{0, 1\}$ represents a binary variable that equals one if a mortality jump occurs in year t and zero otherwise.
- Y_t denotes a jump severity variable, measuring the effect of the mortality jump.
- The parameters $\beta_x^{(J)}$ denote a new age pattern of shocks

Own Model Proposition

- Our model extents the variant of Liu and Li (2015) by introducing a vanishing effect J_t controlled by the parameter a

$$\ln\left(m_{x,t}\right) = \alpha_x + \beta_x \kappa_t + \qquad\qquad + e_{x,t} \tag{2}$$

Own Model Proposition

- Our model extents the variant of Liu and Li (2015) by introducing a vanishing effect J_t controlled by the parameter a

$$\ln\left(m_{x,t}\right) = \alpha_x + \beta_x \kappa_t + \beta_x^{(J)} J_t + e_{x,t} \tag{2}$$

Own Model Proposition

- Our model extents the variant of Liu and Li (2015) by introducing a vanishing effect J_t controlled by the parameter a

$$\begin{aligned} \ln\left(m_{x,t}\right) &= \alpha_x + \beta_x \kappa_t + \beta_x^{(J)} J_t + e_{x,t} \\ J_t &= a J_{t-1} + N_t Y_t, \end{aligned} \tag{2}$$

where $a \in [0, 1)$ controls the rate of decay.

Own Model Proposition

- Our model extents the variant of Liu and Li (2015) by introducing a vanishing effect J_t controlled by the parameter a

$$\begin{aligned} \ln\left(m_{x,t}\right) &= \alpha_x + \beta_x \kappa_t + \beta_x^{(J)} J_t + e_{x,t} \\ J_t &= a J_{t-1} + N_t Y_t, \end{aligned} \tag{2}$$

  where $a \in [0, 1)$ controls the rate of decay.
- Own model is a generalization of Eq. (1) that is recovered if $a = 0$.
- The time effect κ_t is assumed to be a random walk with drift

$$\kappa_t = \kappa_{t-1} + d + \xi_t, \qquad \xi_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\xi^2\right).$$

Mortality Improvements

- Two common routes of estimation called Route I and Route II in the terminology of Haberman and Renshaw (2012).
- The traditional approach (Route I) estimates mortality rates directly.
- Route II models the first differences of log mortality rates, called mortality improvements.
- Mortality improvements are given by

$$Z_{x,t} := \ln\left(m_{x,t+1}\right) - \ln\left(m_{x,t}\right).$$

- Own model specification of Eq. (2) is then

$$\begin{aligned} Z_{x,t} &= \beta_x\left(\kappa_{t+1} - \kappa_t\right) + \beta_x^{(J)}\left(J_{t+1} - J_t\right) + \varepsilon_{x,t} \\ Z_{x,t} &= \beta_x\left(d + \xi_{t+1}\right) + \beta_x^{(J)}\Delta J_t + \varepsilon_{x,t}, \end{aligned} \tag{3}$$

where $\varepsilon_{x,t} = e_{x,t+1} - e_{x,t}$ with $\varepsilon_{x,t} \overset{iid}{\sim} \mathcal{N}(0, 2\sigma_e^2)$.

Identifiability Constraints I

- The Lee Carter model suffers from non-identifiability.
- Own approach of Eq. (3) can be thought an extension of the Lee Carter model.
- According to Hunt and Blake (2020, Theorem 1.) the typical Lee Carter transformations can be generalized into higher dimensions.
- Let $\boldsymbol{b}_x = (\beta_x, \beta_x^J)^\intercal$ and $\boldsymbol{v}_t = (\Delta\kappa_t, \Delta J_t)^\intercal$, then Eq. (2) can written as

$$Z_{x,t} = \boldsymbol{b}_x^\intercal \boldsymbol{v}_t + \varepsilon_{x,t}.$$

- Let there be a (2×2) invertible matrix $\mathbf{A}$, then the parameters can be transformed using

$$\{\tilde{\boldsymbol{b}}_x, \tilde{\boldsymbol{v}}_t\} = \left\{\boldsymbol{b}_x \mathbf{A}^{-1}, \mathbf{A}\boldsymbol{v}_t\right\}.$$

- Since $\mathbf{A} \in \mathbb{R}^{2\times 2}$ there are a total of four free parameters
$\rightarrow$ we need to impose four identifiability constraints !

Identifiability Constraints II

- To make the model identifiable we impose the standard sum-to-one constraints

$$\sum_{x=1}^{A} \beta_x = 1 \quad \text{and} \quad \sum_{x=1}^{A} \beta_x^{(J)} = 1.$$

- Other constraints are possible, i.e. like using the Euclidean norm.
- In addition, we apply a corner constraint and set $\Delta J_{t_1} = 0$ as well as $\Delta k_{t_1} = d$.
- Constraints on the age parameters are the same as Liu and Li (2015).

Parameter Estimation

- Parameter estimation done in a Bayesian setting.
- NIMBLE (Numerical Inference for statistical Models for Bayesian and Likelihood Estimation) (de Valpine et al., 2017)
- NIMBLE is a framework for writing statistical models and algorithms.
- Can be accessed in R via the *nimble* package (de Valpine et al., 2023)
- Allows to sample discrete parameters and lets you choose which sampler to select for each parameter.

Prior Overview

Parameter	Prior Distribution	Hyperprior 1	Hyperprior 2
Age Parameter			
$(\beta_1, \ldots, \beta_A)$	Dirichlet$(1, \ldots, 1)$	-	-
$\left(\beta_1^{(J)}, \ldots, \beta_A^{(J)}\right)$	Dirichlet$(1, \ldots, 1)$	-	-
Time Parameter			
$\Delta\kappa_t$	$\Delta\kappa_t \overset{iid}{\sim} \mathcal{N}(d, \sigma_\xi^2)$	$d \sim \mathcal{N}(0, 5)$	$\sigma_\xi \sim \mathcal{N}_+(0, 2)$
N_t	$N_t \overset{iid}{\sim}$ Bern(p)	$p \sim$ Beta$(1, 20)$	-
Y_t	$Y_t \overset{iid}{\sim} \mathcal{N}(\mu_Y, \sigma_Y^2)$	$\mu_Y \sim \mathcal{N}_+(0, 5)$	$\sigma_Y \sim \mathcal{N}_+(0, 2)$
a	$a \sim$ Beta$(1, 5)$	-	-
Other Parameters			
σ_ε	$\sigma_\varepsilon \sim \mathcal{N}_+(0, 2)$	-	-

Table: Selection of Priors

Data

- We obtained death rates for countries which were heavily affected by the COVID-19 Pandemic, according to deaths per 100 K people (https://coronavirus.jhu.edu/data/mortality).
- Data from two sources: Human Mortality Database (HMD) and Eurostat.

Country	Year	Source
Italy	1980 - 2021	Eurostat (demo_magec)
	2022	Eurostat (demo_r_mkw_05)
Spain	1980 - 2021	Eurostat (demo_magec)
	2022	Eurostat (demo_r_mkw_05)
United States	1980 - 2021	HMD

Table: Sources of Data

- Total of $A = 10$ age groups with $x \in [0, 10), [10, 20), \ldots, [90, \infty)$.
- $T = 41$ years, respectively $T = 42$ (for european countries)

Validation Approach

- Compare own modelling approach with that of Liu and Li (2015) for all three countries.
- Classical out-of-sample evaluation not feasible due to pandemic's recent occurrence.
- Comparison of in-sample fit only, via widely applicable or Watanabe-Akaike information criterion (WAIC; Watanabe, 2010) as well as the leave-one-out cross validated predictive density (LOO-CV).
- The LOO-CV gets multiplied by -2 to be on the deviance scale (lower is better).

Results I

Model	Italy	Spain	US
WAIC			
Own	-1568.39	**-1534.34**	**-1898.16**
Liu-Li	**-1569.27**	-1526.26	-1896.61
LOO-CV			
Own	-1563.38	**-1527.30**	**-1890.85**
Liu-Li	**-1564.92**	-1517.83	-1889.60

Table: In-Sample Fit comparison.

Results II

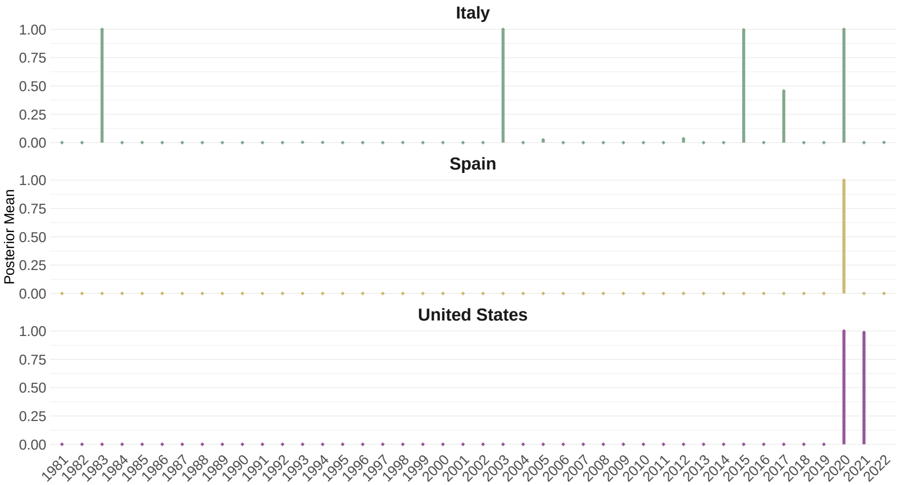

Figure: Posterior mean estimate of N_t (Jump occurrence) for each country.

Results III

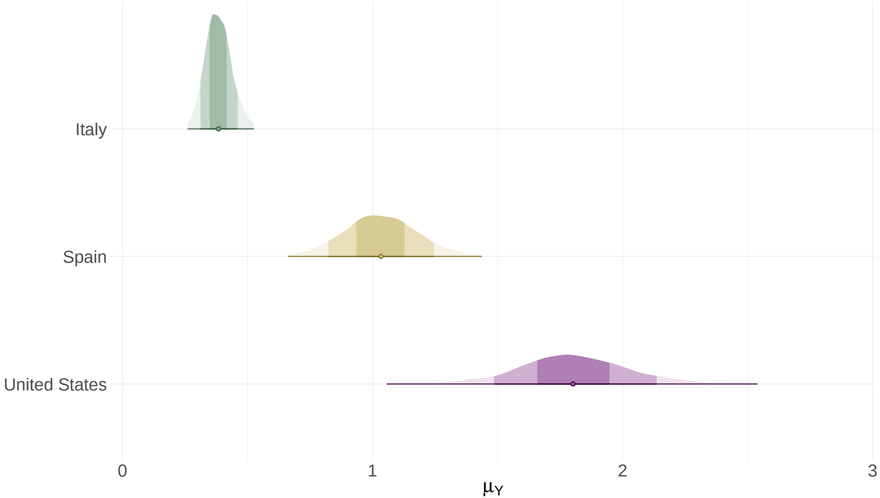

Figure: Posterior estimate of μ_Y (mean of jump effect) for each country.

Results IV

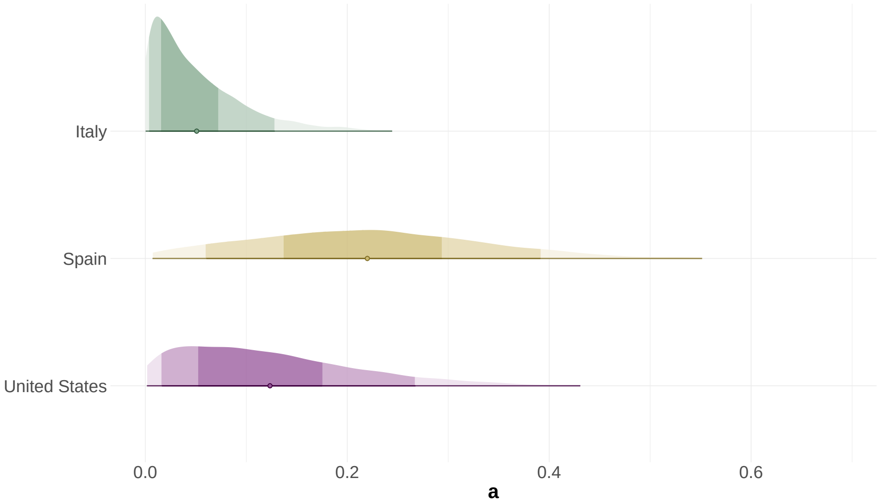

Figure: Posterior estimate of the vanishing parameter a for each country.

Results V

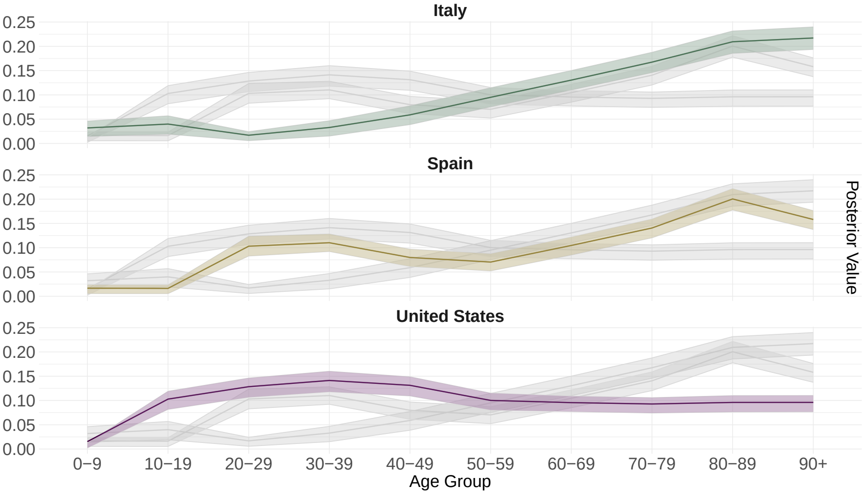

Figure: Age Pattern of $\beta_x^{(J)}$. Thick line denotes posterior mean, colored shades denote 50%- posterior intervals. Light gray shades show posterior estimates of the other countries.

Outlook

- Analyse how jump models affect forecasts of mortality rates
- Use a hierarchical structure to estimate the vanishing parameter a as well as jump parameters μ_Y and σ_Y for multiple countries jointly.
- Use some regularization on the vanishing parameter a, e.g. the horseshoe prior.

Thank you for your attention!

Bibliography I

Chen, F.-Y., Yang, S. S., & Huang, H.-C. (2022).Modeling pandemic mortality risk and its application to mortality-linked security pricing. *Insurance: Mathematics and Economics*, *106*, 341–363. https://doi.org/10.1016/j.insmatheco.2022.06.002

Chen, H., & Cox, S. H. (2009).Modeling Mortality With Jumps: Applications to Mortality Securitization. *Journal of Risk and Insurance*, *76*(3), 727–751. https://doi.org/10.1111/j.1539-6975.2009.01313.x

de Valpine, P., Turek, D., Paciorek, C. J., Anderson-Bergman, C., Lang, D. T., & Bodik, R. (2017).Programming with models: Writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics*, *26*(2), 403–413.

de Valpine, P., Paciorek, C., Turek, D., Michaud, N., Anderson-Bergman, C., Obermeyer, F., Cortes, C. W., Rodríguez, A., Lang, D. T., Zhang, W., Paganin, S., Hug, J., Zhang, W., Babu, J., Ponisio, L., & Sujan, P. (2023, June). Nimble: MCMC, Particle Filtering, and Programmable Hierarchical Modeling.

Haberman, S., & Renshaw, A. (2012).Parametric mortality improvement rate modelling and projecting. *Insurance: Mathematics and Economics*, *50*(3), 309–333. https://doi.org/10.1016/j.insmatheco.2011.11.005

Hunt, A., & Blake, D. (2020).Identifiability in age/period mortality models. *Annals of Actuarial Science*, *14*(2), 461–499. https://doi.org/10.1017/S1748499520000111

Lee, R. D., & Carter, L. R. (1992).Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, *87*(419), 659–671. https://doi.org/10.1080/01621459.1992.10475265

Bibliography II

Li, S.-H., & Chan, W.-S. (2005).Outlier analysis and mortality forecasting: The United Kingdom and Scandinavian countries. *Scandinavian Actuarial Journal*, *2005*(3), 187–211. https://doi.org/10.1080/03461230510006973

Liu, Y., & Li, J. S.-H. (2015).The age pattern of transitory mortality jumps and its impact on the pricing of catastrophic mortality bonds. *Insurance: Mathematics and Economics*, *64*, 135–150. https://doi.org/10.1016/j.insmatheco.2015.05.005

Robben, J., & Antonio, K. (2023).Catastrophe risk in a stochastic multi-population mortality model. *arXiv*.

van Berkum, F., Melenberg, B., & Vellekoop, M. (2022, September). Estimating the impact of the COVID-19 pandemic using granular mortality data.

Vehtari, A., Gelman, A., & Gabry, J. (2017).Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4

Watanabe, S. (2010).Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory.. *Journal of machine learning research*, *11*(12), 3571–3594.

In Sample Criteria I

- The WAIC is based on the log pointwise predictive density (lppd) defined as

$$\text{lppd} = \sum_{i=1}^{N} \ln \int p(y_i|\theta)p(\theta|y_i)d\theta,$$

  where $y_1, \ldots, y_N$ denotes the data at hand.

- The above quantity is computed using draws from the posterior where $\theta^{(s)}$ denote the usual simulations with $s = 1, \ldots S$, then

$$\widehat{\text{lppd}} = \sum_{i=1}^{N} \ln \left(\frac{1}{S} \sum_{s=1}^{S} p(y_i|\theta^{(s)}) \right)$$

In Sample Criteria II

- The WAIC is given by

$$\text{WAIC} = -2\text{lppd} + 2p_{waic}$$
$$p_{waic} = \sum_{i=1}^{N} V_{s=1}^{S} \left(\ln p(y_i|\theta^{(s)}) \right),$$

where $V_{s=1}^{S}$ denotes the sample variance over all s posterior draws.

In Sample Criteria III

- The leave-one-out predictive density (lppd_{loo}) in a Bayesian setting is calculated by

$$\text{lppd}_{\text{loo}} = \sum_{i=1}^{N} \ln p(y_i | y_{-i}),$$

where

$$p(y_i | y_{-i}) = \int p(y_i | \theta) p(\theta | y_{-i}) d\theta$$

- Instead of estimating the posterior $p(\theta | y_{-i})$ for each of the N cross validation points, the density can be approximated using importance sampling (Vehtari et al., 2017).
- For the lppd_{loo} to be on the deviance scale we calculate loo-cv $= -2\text{lppd}_{\text{loo}}$

Sampler Overview

Sampler	Parameter
AF Slice Sampler	
	$(\beta_1, \ldots, \beta_A)$
	$\left(\beta_1^{(J)}, \ldots, \beta_A^{(J)}\right)$
Binary Sampler	
	$N_t \quad \forall t \in \{1, \ldots, T\}$
Gibbs	
	d
	p
Random Walk Metropolis	
	σ_ε
	σ_ξ
Slice Sampler	
	$\Delta\kappa_t \quad \forall t \in \{1, \ldots, T\}$
	μ_ξ
	σ_Y

Table: Overview of Samplers

Death Rates Italy I

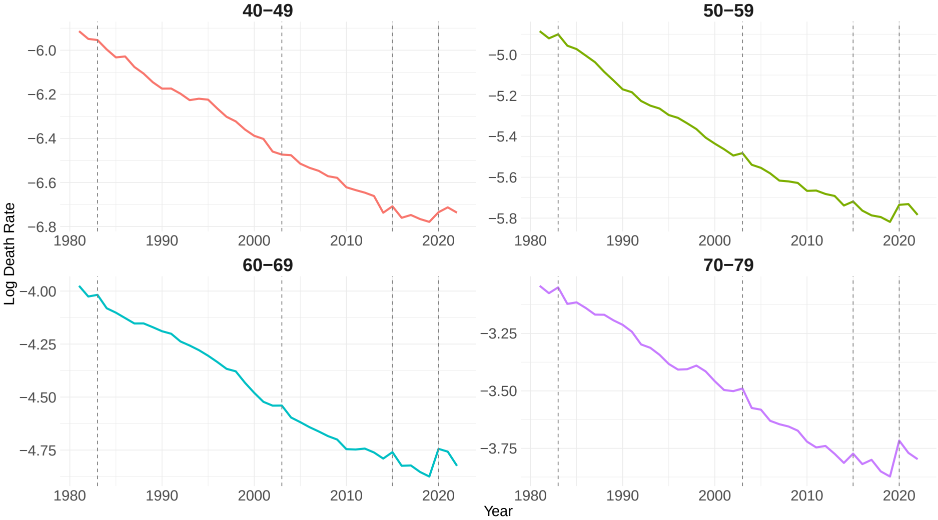

Figure: Log Death Rates for Italy

Death Rates Italy II

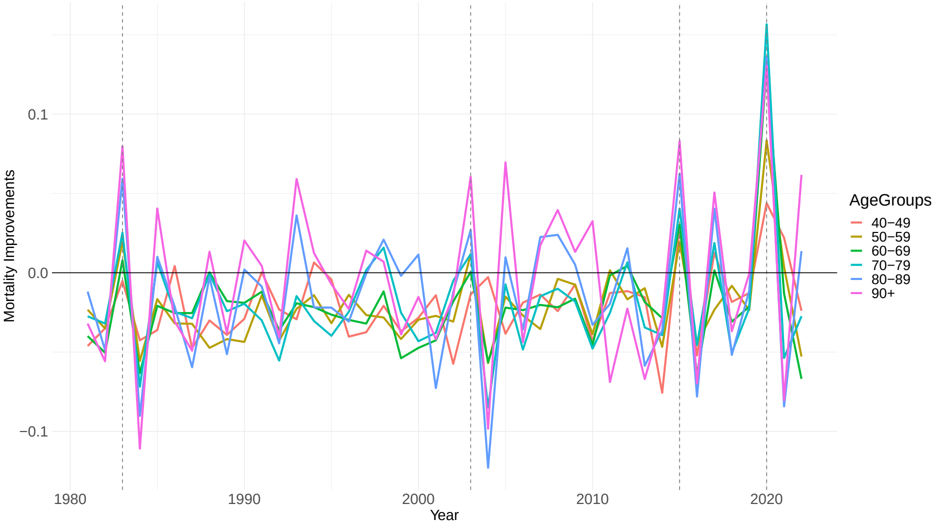

Figure: Mortality Improvements Italy