ERGO Center of Excellence in Insurance

Eine Einrichtung der TUM gefördert von der ERGO Group

Estimation and Projection of Affine Mortality Models: A Multi-Country Comparison

Francesco Ungolo (TUM)
Joint work with M. Sherris and Y. Zhou (UNSW)

Longevity 16 conference
13th August 2021

Agenda

Introduction

- Stochastic mortality models attracted considerable attention in the literature, following the seminal work of Lee and Carter (1992), particularly in discrete time (Cairns et al. (2006), Renshaw and Haberman (2006), Cairns et al. (2009));
- Initially focused on age-period data, several extensions have been proposed to account for the cohort effect (Willets (2004), Cairns et al. (2009), Gallop (2008));
- Continuous time models have been recently proposed, drawing on the mathematical framework of interest rate models (Milevsky and Promislow (2001), Jevtic et al. (2013), Blackburn and Sherris (2013)).

Age-cohort data

- The analysis of longevity-linked cash flows requires the use of cohort survival functions;
- The use of models based on age-period and age-period-cohort models may not help in the identification of cohort effects;
- The cohort effect is deemed to be more stable and persistent across adult life (McCarthy (2020))

Continuous time mortality models

- Analytical tractability - closed form survival curves for affine class;
- Consistency between mortality dynamics and functional form of the survival curve;
- Stability of parameter estimates;
- Use of mathematical finance techniques relying on arbitrage-free pricing;
- Natural extensions to multi-factor models, capturing different features of the mortality surface.

The affine framework (Duffie and Kan (1996)) - (1)

The intensity process for a cohort born in year t is driven by the vector of latent factors $X(t)$ with dynamics:

$$\mathrm{d}X(t) = \Delta\left[\theta^Q - X(t)\right]\mathrm{d}t + \Sigma D(X(t), t)\,\mathrm{d}W^Q(t) \tag{1}$$

where

- $\Delta \in \mathbb{R}^{M\times M}$ is the mean reversion matrix;
- $\theta^Q \in \mathbb{R}^M$ is the long term mean of the process;
- $\Sigma \in \mathbb{R}^{M\times M}$ is the volatility matrix;
- $W^Q(t)$ is a standard Brownian motion;
- $D(X(t), t)$ is a diagonal matrix;

The affine framework (Duffie and Kan (1996)) - (2)

- The survival probability of newborn in year t until time T, $S(t,T)$, is modelled as an exponentially affine function of $X(t)$:

$$\begin{aligned} S(t,T) &= \mathbb{E}\left[\exp\left(-\int_t^T \mu(t,s)\,\mathrm{d}s\right) \mid \mathscr{F}_t\right] \\ &= \exp\left[-\bar{\mu}(t,T)(T-t)\right] \\ &= \exp\left[A(t,T) + B(t,T)' X(t)\right] \end{aligned} \tag{2}$$

- The factor loading $B(t,T)$ and $A(t,T)$ depend on the mortality dynamics specified for $X(t)$;
- $\bar{\mu}(t,T)$ is the average force of mortality between t and T:

$$\bar{\mu}(t,T) = -\frac{1}{T-t}\log S(t,T) = -\frac{B(t,T)'}{T-t}X(t) - \frac{A(t,T)}{T-t} \tag{3}$$

Mortality models

- Blackburn-Sherris model with three factors (BS, dependent and independent factors, Blackburn and Sherris (2013) and Huang et al. (2020));
- Arbitrage-Free Nelson-Siegel with Level, Slope and Curvature factors (AFNS, dependent and independent factors, Christensen et al. (2011));
- Cox-Ingersoll-Ross with three independent factors (CIR, Cox et al. (1985))

Data

We approximate the force of mortality by means of the central death rates $m_{x,t}$ for the male population aged 50-99 of the following countries, sourced from the HMD:

- USA (cohorts born from 1883-1915);
- Australia (1872-1916);
- Denmark (1790-2014);
- England and Wales (1795-1914);
- Japan (1897-1916).

The probability of an individual aged x at time t to survive until age $x + T - t$ is:

$$S_x(t, T) = \prod_{j=1}^{T-t} e^{-m_{x+j-1,t-x}} \tag{4}$$

The corresponding average force of mortality is:

$$\bar{\mu}_x(t, T) = \frac{1}{T-t} \sum_{j=1}^{T-t} m_{x+j-1,t-x}, \tag{5}$$

Example and univariate formulation

For the cohort born in year $t = 1, \ldots, L$ we have:

$$\begin{pmatrix} \bar{\mu}(t,t+1) \\ \bar{\mu}(t,t+2) \\ \vdots \\ \bar{\mu}(t,t+N) \end{pmatrix} = - \begin{pmatrix} \frac{A(t,t+1)}{1} \\ \frac{A(t,t+2)}{2} \\ \vdots \\ \frac{A(t,t+N)}{N} \end{pmatrix} - \begin{pmatrix} \frac{B_1(t,t+1)}{1} & \frac{B_2(t,t+1)}{1} & \frac{B_3(t,t+1)}{1} \\ \frac{B_1(t,t+2)}{2} & \frac{B_2(t,t+2)}{2} & \frac{B_3(t,t+2)}{2} \\ \vdots & \vdots & \vdots \\ \frac{B_1(t,t+N)}{N} & \frac{B_2(t,t+N)}{N} & \frac{B_3(t,t+N)}{N} \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} + \begin{pmatrix} \varepsilon_{t,1} \\ \varepsilon_{t,2} \\ \vdots \\ \varepsilon_{t,N} \end{pmatrix} \tag{6}$$

$$= \bar{\mu}_t = A + BX(t) + \varepsilon_t \tag{7}$$

Let:

- $\bar{\mu}_t = [\bar{\mu}_{t,1}, \ldots, \bar{\mu}_{t,N}]'$
- $A = [a_1, \ldots, a_N]'$;
- $B = [b_1, \ldots, b_N]'$; (b_i is a 3-dimensional vector).

Inference: State-space formulation of the system

State-space formulation with univariate measurement formulation (Koopman and Durbin (2000)):

$$\text{State equation:} \quad X(t) = e^{-K} X(t-1) + \eta_t \tag{8}$$

$$\text{Measurement equation:} \quad \bar{\mu}_{t,i} = a_i + b_i X(t) + \varepsilon_{t,i} \tag{9}$$

where:

- K is the mean reversion term under the real-world probability measure P;
- $\eta_t \sim N(0, R)$;
- $\varepsilon_{t,i} \sim N\left(0, r_c + r_1 \sum_{k=1}^{i} e^{r_2 k}/i\right)$;
- η_t and ε_t are independently distributed.

Inference: Kalman-Filter Maximum Likelihood (1)

The Kalman filter provides the least square estimator of the distribution of $X(t)$, conditional on the measurement $\bar{\mu}$. It consists of the following two recursive steps for $t = 1, \dots, L$

1. Forecasting ($i = 1$ only):

$$\widehat{X}(t \mid t-1) = \mathbb{E}(X(t) \mid \bar{\mu}_{1:t-1}) = e^{-K'}\widehat{X}(t-1) \tag{10}$$
$$\widehat{\Sigma}(t \mid t-1) = \mathbb{V}(X(t) \mid \bar{\mu}_{1:t-1}) = e^{-K}\widehat{\Sigma}(t-1)\, e^{-K'} + R$$

2. Time-update ($i = 1, \dots, N-1$ on the right-hand side):

$$\widehat{X}(t) = \mathbb{E}(X(t) \mid \bar{\mu}_{1:t-1}, \bar{\mu}_{t,1:i-1}) = \widehat{X}(t \mid t-1) + K_{t,i} v_{t,i} \tag{11}$$
$$\widehat{\Sigma}(t) = \mathbb{V}(X(t) \mid \bar{\mu}_{1:t-1}, \bar{\mu}_{t,1:i-1}) = \widehat{\Sigma}(t \mid t-1) - K_{t,i} F_{t,i} K'_{t,i}$$
$$= (I - K_{t,i} b_i)\widehat{\Sigma}(t \mid t-1)(I - K_{t,i} b_i)' + K_{t,i}\omega^2_{t,i} K'_{t,i}$$

Inference: Kalman-Filter Maximum Likelihood (2)

where

$$
\begin{aligned}
\omega_{t,i}^2 &= r_c + r_1 \sum_{k=1}^{i} e^{r_2 k} / i \\
v_{t,i} &= \bar{\mu}_{t,i} - a_i - b_i \widehat{X}(t \mid t-1) \\
F_{t,i} &= b_i \widehat{\Sigma}(t \mid t-1) b_i' + \omega_i^2 \\
K_{t,i} &= \widehat{\Sigma}(t \mid t-1) b_i' F_{t,i}^{-1}
\end{aligned}
$$

The log-likelihood function is readily obtained:

$$
\log L(\psi \mid \bar{\mu}_{1:L}) = -\frac{TL}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{L} \sum_{i=1}^{N} \left(\log F_{t,i} + v_{t,i}^2 F_{t,i}^{-1} \right) \tag{12}
$$

Resources

- The Github repository `affine_mortality` (Ungolo, Sherris and Zhou (2021b)) contains the code implementing the maximum likelihood with univariate Kalman Filter estimation and analysis of the model, see

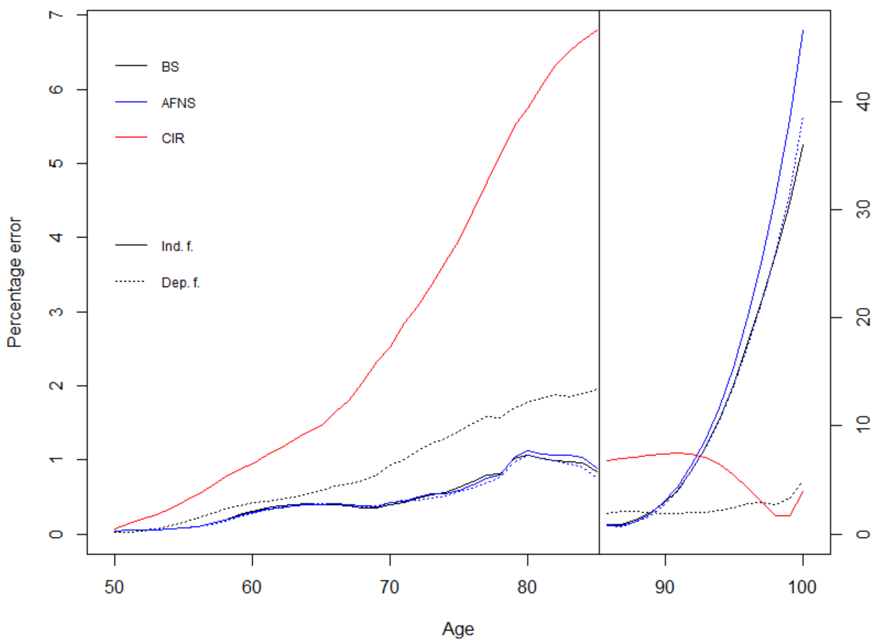
  `https://github.com/ungolof/affine_mortality.git`

- It allows to:
  - Estimate the model parameters (different optimization methods);
  - Estimate parameters uncertainty by bootstrap and multiple imputation;
  - Compute goodness of fit measures, and residuals (and plotting);
  - Project survival curves for future cohorts.

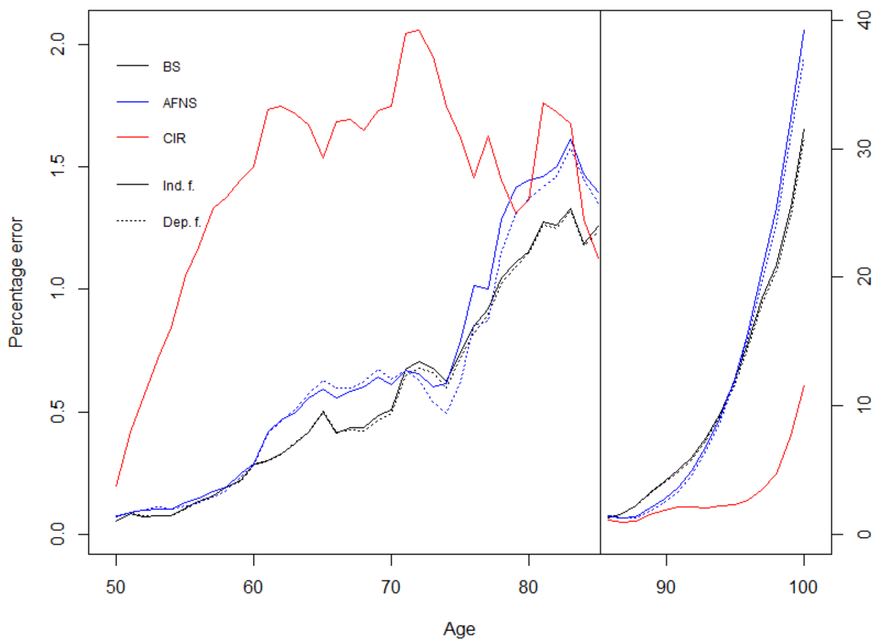
Goodness of fit - Quantitative comparison

	Model	USA	Aus	Dnk	E&W	Jap
AIC	BS ind.	−19848.08	−26397.85	−71916.27	−69179.36	−13153.40
	BS dep.	−19865.51	−26529.58	−72405.65	−69541.13	−13165.64
	AFNS ind.	−19683.26	−25999.51	−71883.57	−69144.38	−13162.81
	AFNS dep.	−20003.19	−26245.26	−72196.84	−69481.86	−13175.81
	CIR	−20106.66	−27109.87	−72430.60	−69516.39	−13366.66
BIC	BS ind.	−19783.18	−26329.22	−71835.38	−69098.97	−13094.50
	BS dep.	−19768.16	−26426.65	−72284.33	−69420.54	−13077.30
	AFNS ind.	−19629.17	−25942.32	−71816.16	−69077.39	−13113.74
	AFNS dep.	−19932.88	−26170.92	−72109.22	−69394.75	−13112.01
	CIR	−20009.31	−27006.93	−72309.28	−69395.80	−13278.32
RMSE	BS ind.	0.00184	0.00187	0.00253	0.00159	0.00029
	BS dep.	0.00059	0.00183	0.00251	0.00161	0.00029
	AFNS ind.	0.00227	0.00207	0.00261	0.00163	0.00030
	AFNS dep.	0.00205	0.00197	0.00248	0.00158	0.00030
	CIR	0.00178	0.00128	0.00253	0.00133	0.00027

Goodness of fit - Mean Absolute Percentage Error

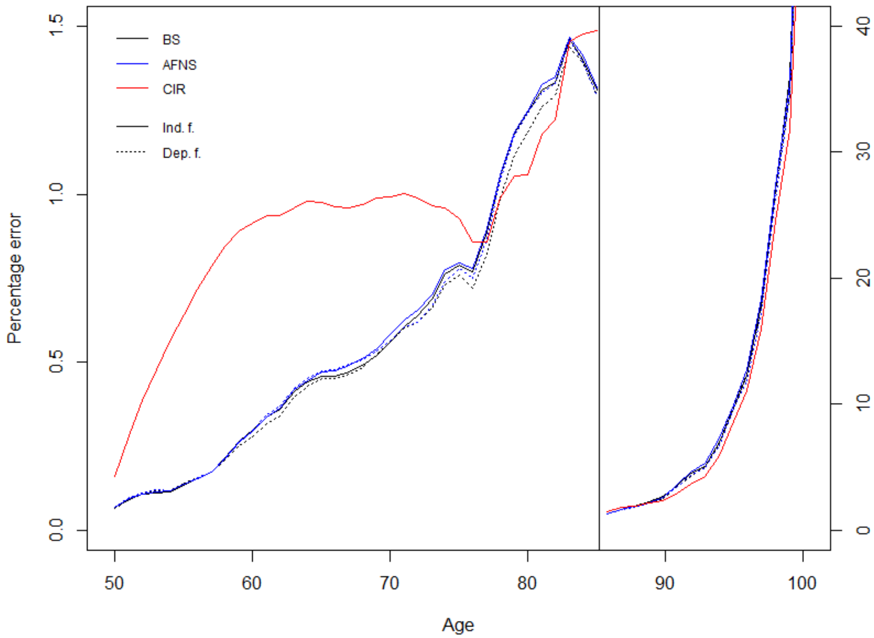


(a) USA

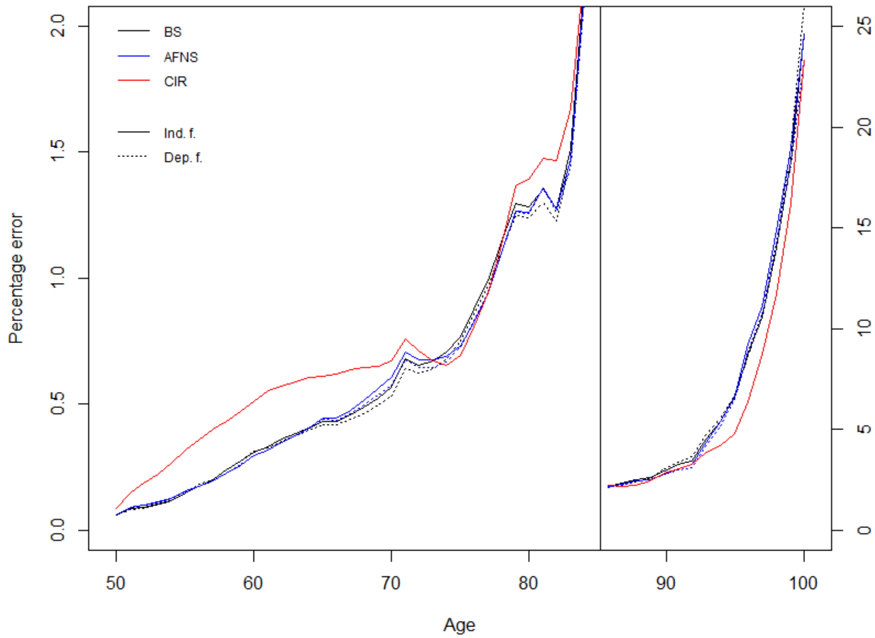


(b) Australia

Goodness of fit - Mean Absolute Percentage Error



(a) Denmark



(b) England & Wales

Goodness of fit - Mean Absolute Percentage Error

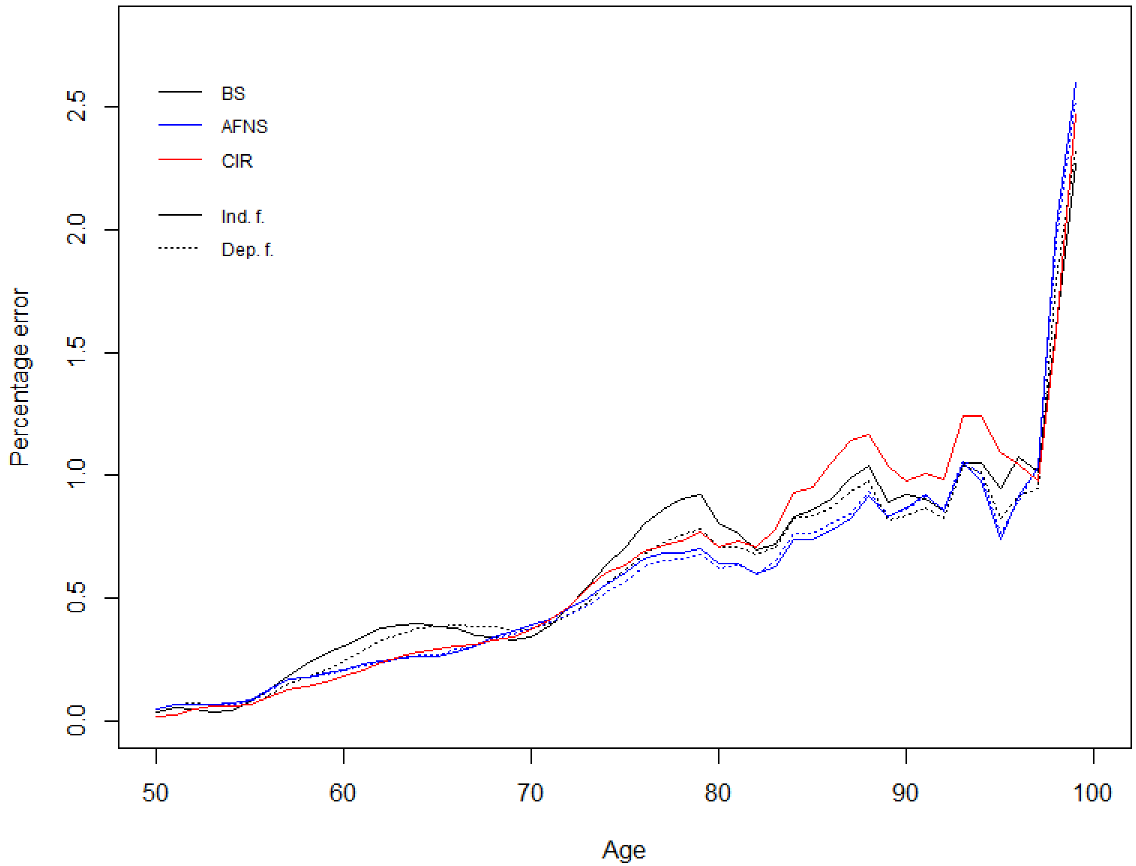
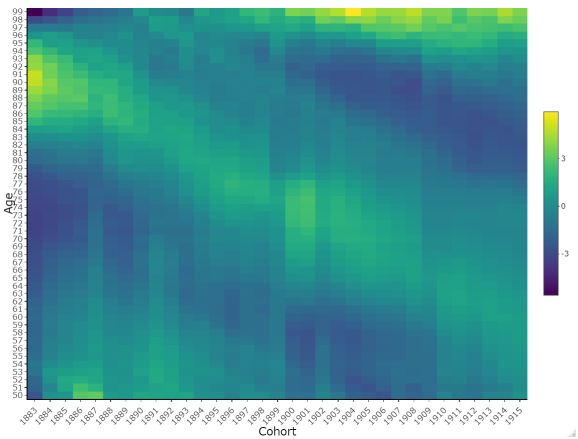
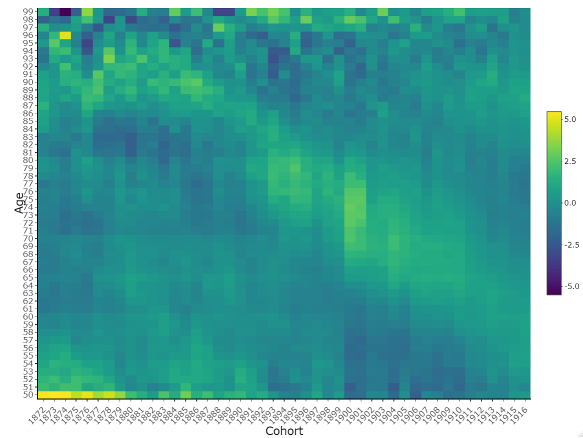


Abbildung: Japan

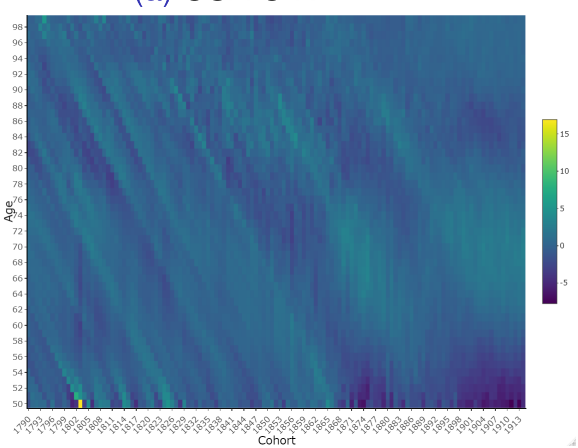
Standardized residuals



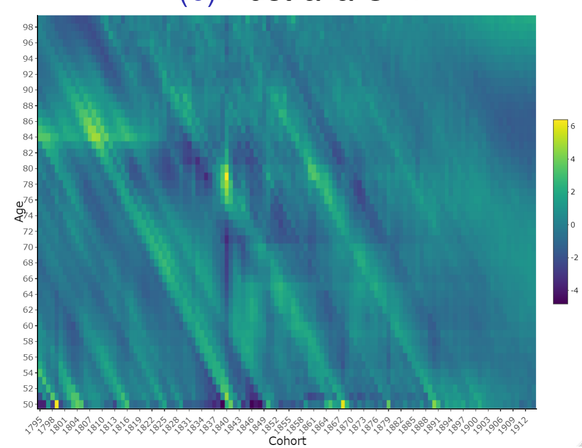
(a) USA CIR



(b) Australia CIR



(c) Denmark CIR



(d) E& W BS dep.

ERGO Center of Excellence in Insurance

Eine Einrichtung der TUM gefördert von der ERGO Group

Projection - Root Mean Squared Error

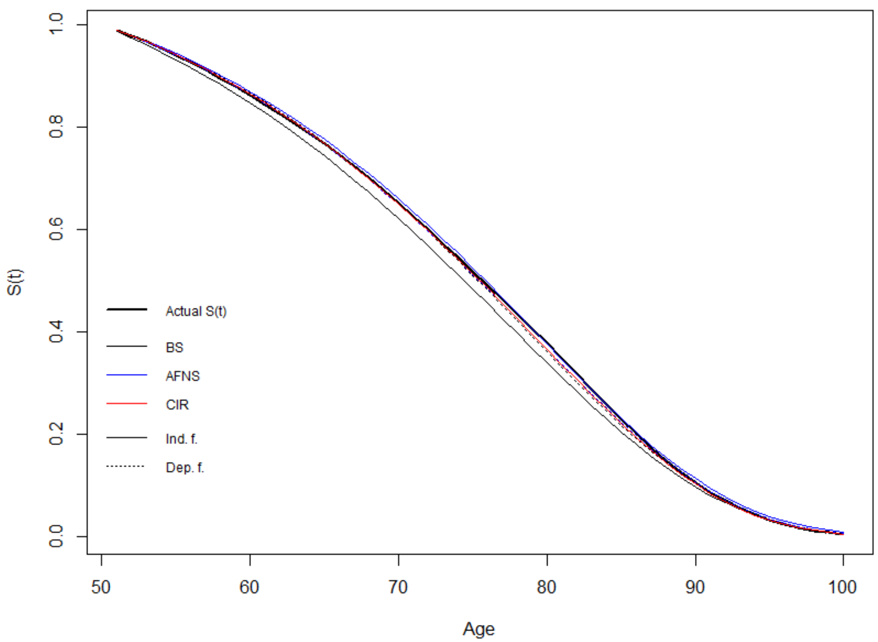
Optimal forecast of the cohort survival curve:

$$S_x(t+1, T+1) = \exp\left(B(t,T)'\mathbb{E}[X(t+1) \mid X(t)] + A(t,T)\right) \tag{13}$$

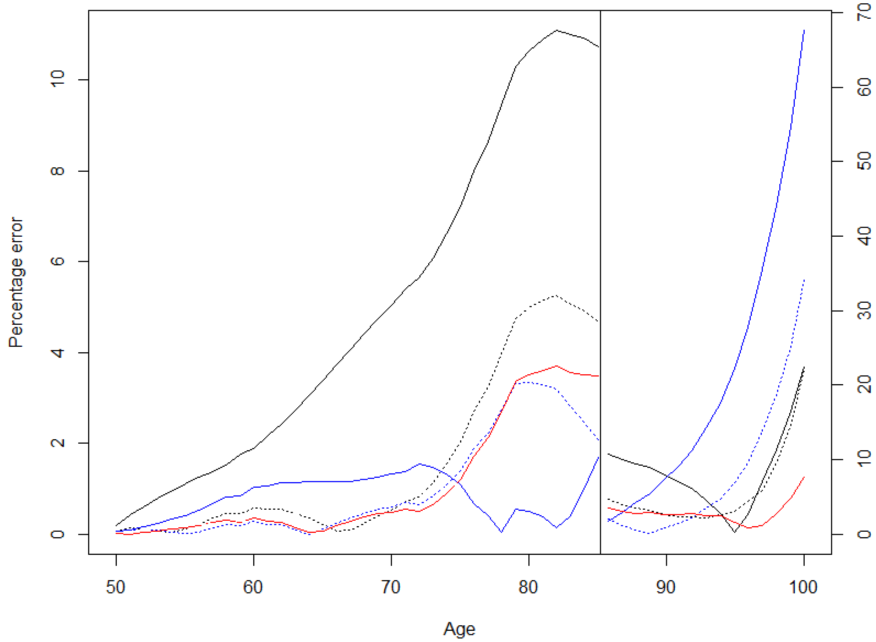
Tabelle: RMSE for comparing Actual and Best-Estimate average mortality rates for projected cohorts.

Country	**Cohort**	**Blackburn-Sherris**		**AFNS**		**CIR**
		Indepedent	**Dependent**	**Indepedent**	**Dependent**	
USA	1916	0.00246	0.00121	0.00279	0.00137	0.00062
Australia	1917	0.00366	0.00360	0.00409	0.00314	0.00545
Denmark	1915	0.00258	0.00194	0.00203	0.00158	0.00345
E&W	1915	0.00304	0.00311	0.00313	0.00251	0.00271
Japan	1917	0.00074	0.00090	0.00073	0.00077	0.00034

Projection - USA

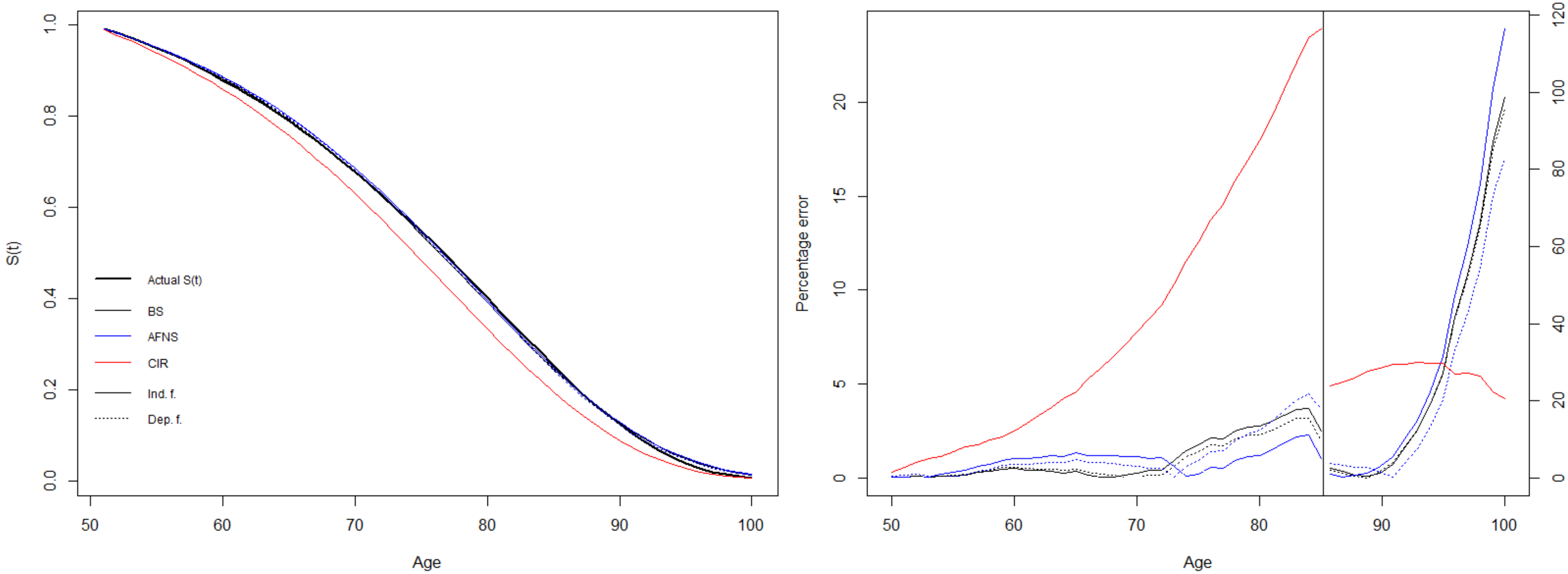


(a) S(t) 1916 cohort



(b) Mean Absolute Percentage Error

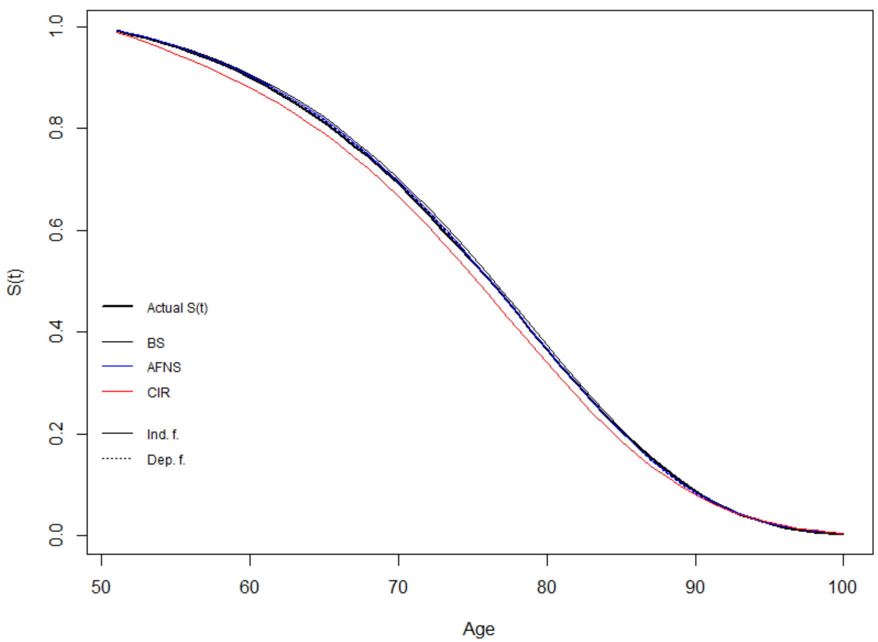
Projection - Australia



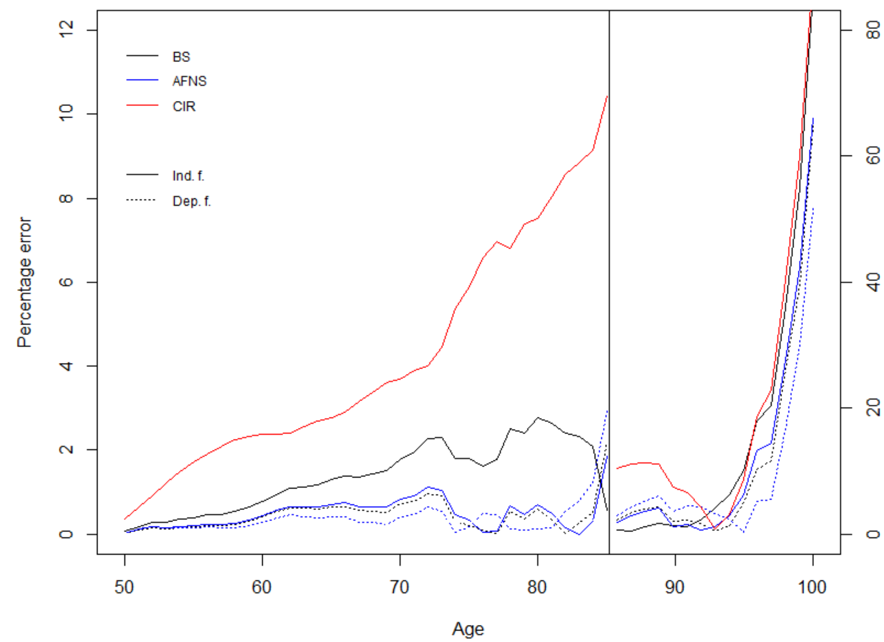
(a) S(t) 1917 cohort

(b) Mean Absolute Percentage Error

Projection - Denmark

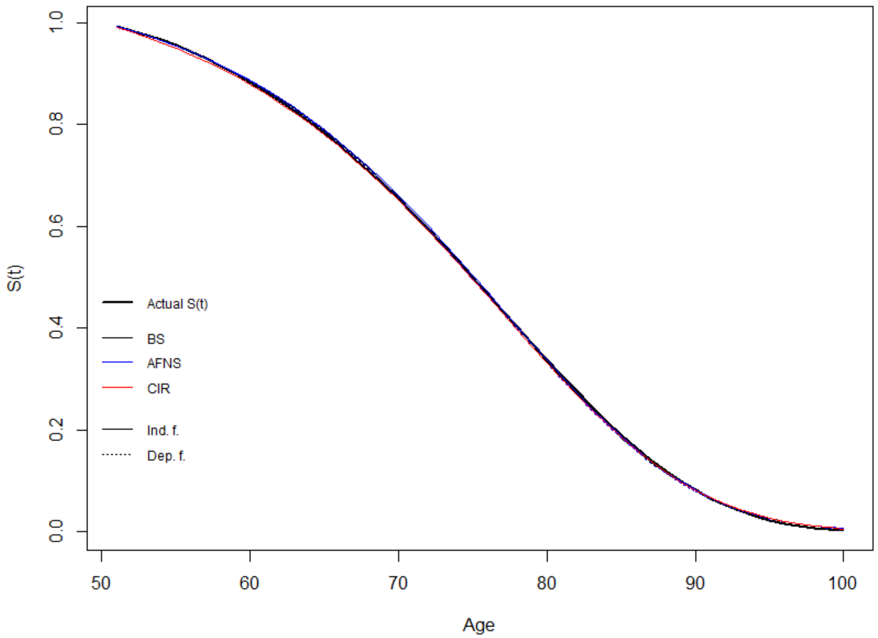


(a) S(t) 1915 cohort

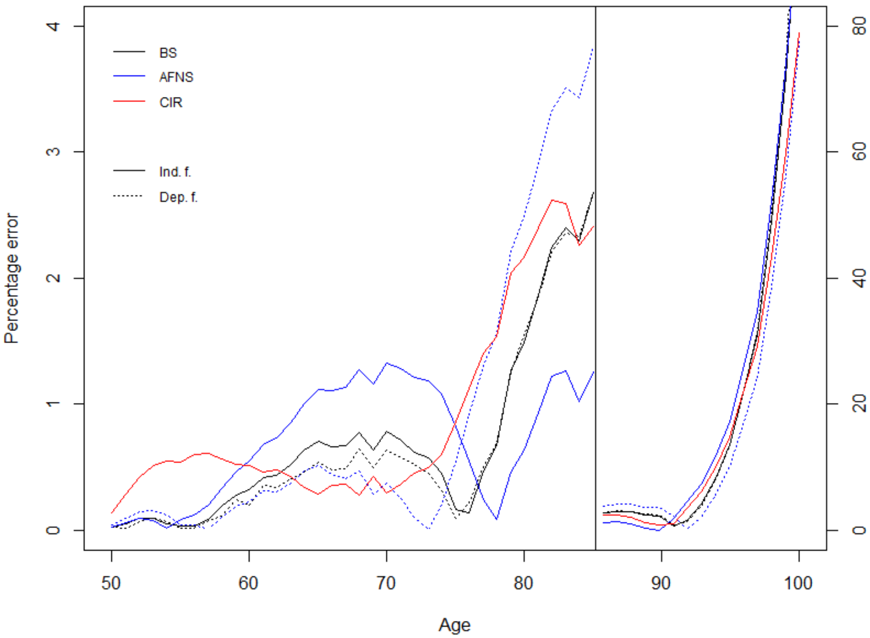


(b) Mean Absolute Percentage Error

Projection - England & Wales

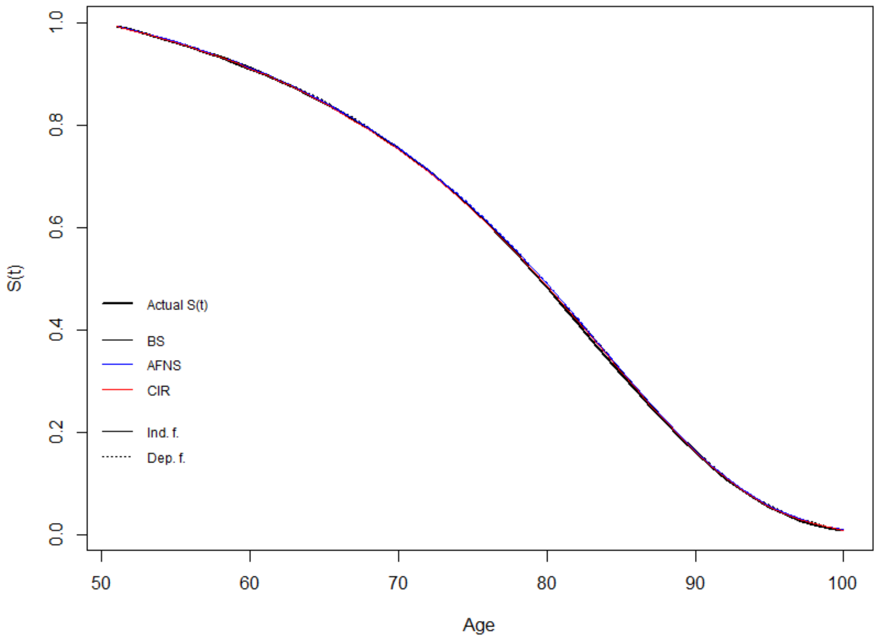


(a) S(t) 1915 cohort

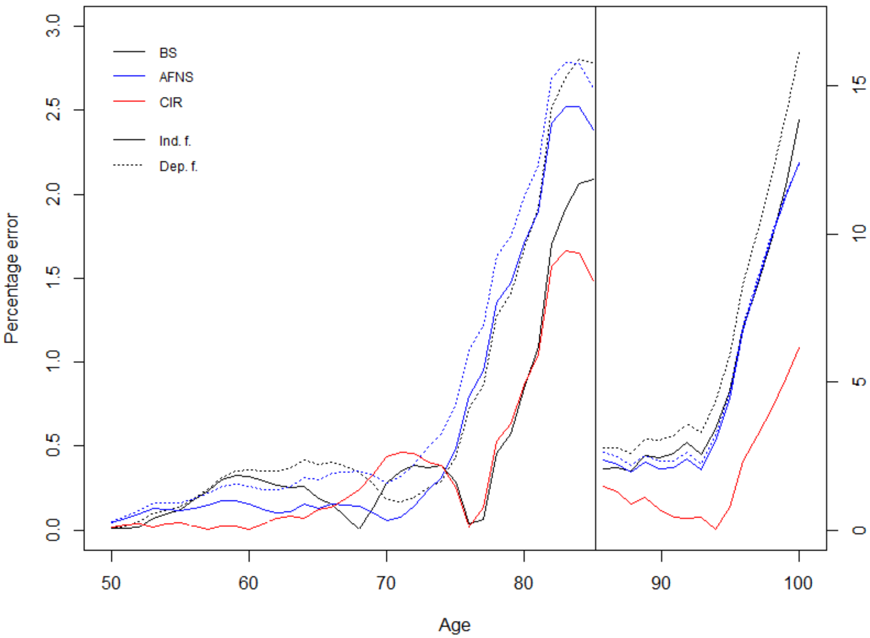


(b) Mean Absolute Percentage Error

Projection - Japan



(a) S(t) 1917 cohort



(b) Mean Absolute Percentage Error

Conclusions

- The analysis of age-cohort data allows to directly obtain cohort survival curves useful for pricing longevity-linked cash flows;
- The maximum likelihood estimator using the univariate Kalman Filter provides a robust method for parameter estimation;
- The CIR mortality model is capable to best fit historical data for most countries, especially at older ages;
- The analysis of standardized residuals showed the presence of a period effect;
- Simpler Gaussian models (particularly those with dependent factors) provide a reasonable means to project future survival curves;
- All models analysed proved to be robust with respect to the range of cohorts used for estimation;

References

- Blackburn, C. and Sherris, M. (2013). Consistent dynamic affine mortality models for longevity risk applications. *Insurance: Mathematics and Economics, 53:64-73*.
- Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., and Balevich, I. (2009). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal, 13(1):1-35*;
- Cairns, A. J., Blake, D., and Dowd, K. (2006b). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance, 73(4):687-718*;
- Christensen, J. H., Diebold, F. X., and Rudebusch, G. D. (2011). The affine arbitrage-free class of Nelson-Siegel term structure models. *Journal of Econometrics, 164(1):4-20*;
- Cox, J.C., J.E. Ingersoll and S.A. Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica. 53 (2): 385–407*;
- Duffie, D., Kan, R., A yield-factor model of interest rates, *Mathematical Finance, Vol. 6, Issue 4, 379-406*;
- Gallop, A. (2008). Mortality projections in the United Kingdom. *Paper for conference: Living to 100 and Beyond Symposium. Society of Actuaries*;
- Huang, Z., Sherris, M., Villegas, A., and Ziveyi, J. (2019). The application of affine processes in cohort mortality risk models. *Working paper, CEPAR Working Paper*;

ERGO Center of Excellence in Insurance

Eine Einrichtung der TUM gefördert von der ERGO Group

References

- Jevtic, P., Luciano, E., and Vigna, E. (2013). Mortality surface by means of continuous time cohort models. *Insurance: Mathematics and Economics, 53:122-133*;
- Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association, 87(419):659-671*;
- McCarthy, D. (2020), 80 will be the new 70: Old-age mortality postponement in the United States and its likely effect on the finances of the OASI program, *Journal of Risk and Insurance, 88, 381-412*;
- Milevsky, M. A. and Promislow, S. D. (2001). Mortality derivatives and the option to annuitise. *Insurance: Mathematics and Economics, 29(3):299-318*
- Renshaw, A. and Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics, 38(3):556 - 570*;
- Ungolo, F., Sherris M., and Zhou Y. (2021a), Estimation and Projection of Affine Mortality Models: A Multi-Country Comparison, *Working paper*;
- Ungolo, F., Sherris M., and Zhou Y. (2021b), `affine_mortality`: A Github repository for estimation and analysis of affine mortality models, *Working paper*;
- Willets, R. (2004). The Cohort Effect: Insights and Explanations. *British Actuarial Journal*

ERGO Center of
Excellence in Insurance

Eine Einrichtung der TUM gefördert von der ERGO Group

Thank you for your attention!