A Set of New Stochastic Trend Models

- Johannes Schupp
- ICA Berlin, 8th June 2018

About the speaker

Johannes Schupp

Institut für Finanz- und Aktuarwissenschaften (ifa)

- ifa is an independent actuarial consulting firm.
- Our consulting services in all lines of insurance business include:
  - typical actuarial tasks and actuarial modelling
  - insurance product development
  - risk management, Solvency II, asset liability management
  - data analytics
  - market entries (cross-border business, setup of new insurance companies, Fintechs)
  - professional education
  - academic research on actuarial topics of practical relevance
- located in Ulm, Germany
- currently about 30 consultants
- academic cooperation with the University of Ulm (offering the largest actuarial program in Germany)

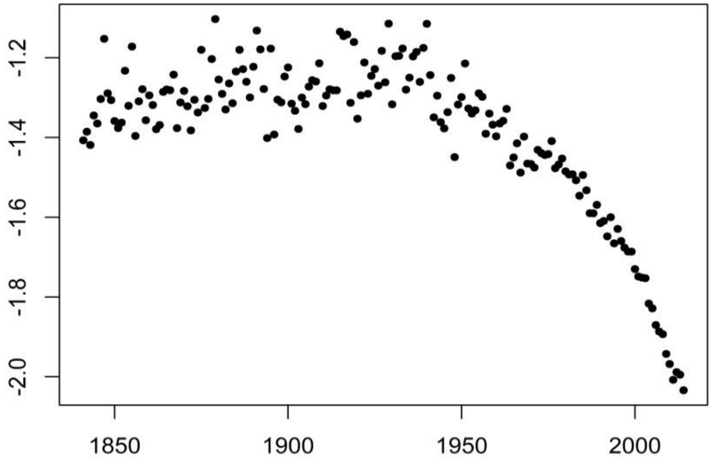
- joined ifa in 2014
- qualified actuary (German Association of Actuaries DAV, 2018)
- Master of Science (Mathematics and Management, University of Ulm, 2014)

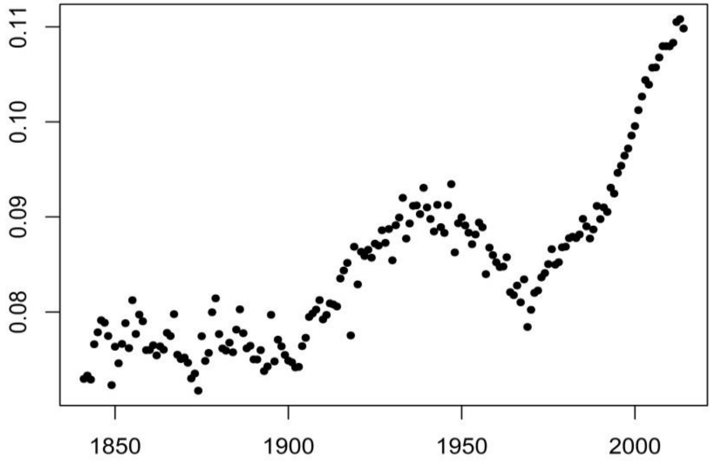
Introduction

- Traditional actuarial mathematics
  - Deterministic future mortality rates
- Reality: Uncertainty about the evolution of mortality
  - Measure longevity risk in pension or annuity portfolios with stochastic mortality models
  - Estimate a distribution of future mortality
- Parametric mortality models: Lee-Carter model, Cairns-Blake-Dowd model, APC model, etc. reduce the information about exposures and deaths to a few parameters:
  - CBD: Two time dependent period effects (Cairns et al. (2006)): $\log\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = \kappa_t^1 + \kappa_t^2 \cdot (x - \bar{x})$
  - Estimate period effects with $L(\kappa_t^1, \kappa_t^2) \rightarrow max$ with the assumption of $D_{x,t} \sim Poi(E_{x,t} \cdot \widehat{m}_{x,t})$ or $D_{x,t} \sim Bin\left(E_{x,t}, \hat{q}_{x,t}\right)$

Introduction

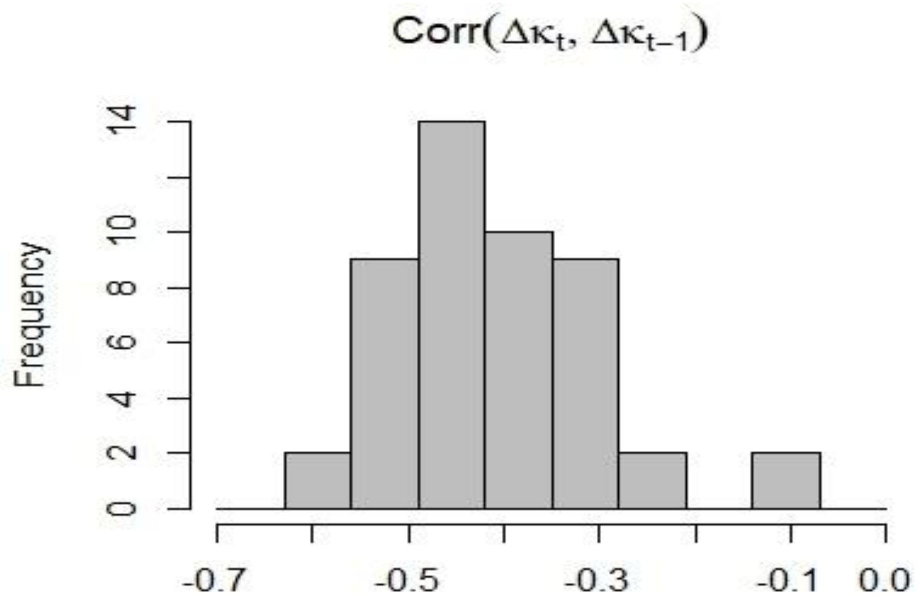
- Period effects calibrated for English and Welsh males older than 65
- Popular choice for stochastic forecasts: a (multivariate) random walk with constant drift (RWD)
- Extrapolating only the most recent trend, systematically underestimates future uncertainty, see e.g. Sweeting (2011), Li et al. (2011), Börger et al. (2014)
- The historic paths indicate a several times changing mortality trend
- Decreasing annual volatility
- In principle, our approaches can be applied to any parametric mortality model with period effects





Introduction

- Trend stationary or Difference stationary mortality?
  - Without any prior knowledge of the trend/drift we can analyze the $Corr(\Delta\kappa_t, \Delta\kappa_{t-1})$
  - Difference Stationary: $Corr(\Delta\kappa_t, \Delta\kappa_{t-1}) \approx 0$
  - Trend Stationary: $Corr(\Delta\kappa_t, \Delta\kappa_{t-1}) \approx -\frac{1}{2}$
  - Correlation of 48 period effects (standard CBD model κ_t^1, κ_t^2, males and females, two age spans (50-89, >65) for England & Wales, Sweden, Switzerland, Australia, Italy, Netherlands)
- Period effects appear **trend stationary** and the historic trend **changed** once in a while
  - Only a piecewise linear trend
  - Random changes in the trends slope
  - Random fluctuation around the prevailing trend



Introduction

Problems in existing trend estimation

- Little number of historic trend changes observed (see Börger and Schupp (2018))
  - unlikely to be generated by the trend change density in some cases
  - inconsistent prediction possible as the distribution for simulations unlikely generates the historic trends
- Choice of information criteria (see Burnham and Anderson (2002)):
  - Pure Likelihood would be biased if we compare models with different numbers of parameters
  - AIC-type should be used to select a best model in a set of candidates
  - BIC-type only for $n \to \infty$ and 'truth' must be a candidate
  - Many applicants of Information Criteria in the context of mortality modeling simply use the BIC because it's more robust

Agenda

- Specification of a Stochastic Trend Process
- Parameter Estimation
  - Existing approach: On the basis of historic trends
  - Forecast-oriented approach based on likelihood
  - Combination of both: Improve the set of candidates
- Parameter Uncertainty
  - Uncertainty in the starting values
- Conclusion

Stochastic Trend Process

- Continuous piecewise linear, changing trend and random fluctuation around the prevailing trend
  - Model the trend process with random noise → $\kappa_t = \hat{\kappa}_t + \epsilon_t;\ \epsilon_t \sim f$
  - Extrapolate the most recent actual mortality trend → $\hat{\kappa}_t = \hat{\kappa}_{t-1} + \beta_t$
  - In every year, there is a possible change in the mortality trend with probability p
  - In the case of a trend change $\lambda_t = M_t \cdot S_t$
    - With absolute magnitude of the trend change $M_t \sim h$
    - Sign of the trend change S_t bernoulli distributed with values -1, 1 each with probability $\frac{1}{2}$

→ $\beta_t = \beta_{t-1} + \lambda_t$, where $\lambda_t = \begin{Bmatrix} 0 & \text{with probability } 1-p \\ M_t \cdot S_t & \text{with probability } p \end{Bmatrix} \sim g$

- We propose to use: $\mathbf{f} = \mathcal{N}(0, \sigma^2_{\epsilon,t})$, $\mathbf{h} = \mathcal{LN}(\mu, \sigma^2)$. In principle, also other distributions are possible (Pareto, Normal, t-distribution,…)

Parameters to be estimated for projections starting in t=0: $p, \sigma^2_{\epsilon,t}, \mu, \sigma^2, \beta_0, \hat{\kappa}_0$

Parameter Estimation

Calibration based on historic trends

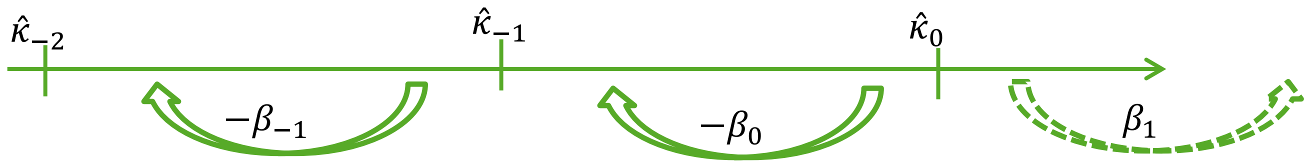
- Use historic trends to estimate parameters (see Sweeting (2011), Hunt and Blake (2015), Börger and Schupp (2018)).
- For $k \in 0, \dots, m$ estimate optimal trend process $\beta_0, \hat{\kappa}_0, \lambda_{-N+2}, \dots, \lambda_0$
  - Estimate the k trend change points $\mathrm{t}_1, \dots, t_k$
  - Estimate the k changes in slope $\lambda_{t_1} \dots, \lambda_{t_k}$
  - Estimate starting trend and value $\beta_0, \hat{\kappa}_0$
- Calibration based on an iterative method proposed by Muggeo (2003). Update $\sigma^2_{\epsilon,t}$ iteratively
  - Find trend, s.th. $f_{\mathcal{N}}\left(\kappa^i_{-N,\dots,0} \middle| trend\right) = f_{\mathcal{N}}\left(\kappa^i_{-N,\dots,0} \middle| \sigma^2_{\epsilon,t}, \beta_0, \hat{\kappa}_0, \lambda_{-N+2}, \dots, \lambda_0\right) \to max$
- Choose optimal trend process with Modified – BIC
  - Requirement of very robust Information Criteria, recommended by Muggeo and Adelfio (2011)

Parameter Estimation

Alternative I

Convoluted Likelihood

- Stochastic forecasts require: $\mu, \sigma^2, p, \sigma^2_{\epsilon,t}, \hat{\kappa}_0, \beta_0$. Not necessarily a historic trend. The focus here will be solely on forecasts.
- Idea: Classic MLE: $L(\mu, \sigma^2, p, \sigma^2_{\epsilon,t}, \hat{\kappa}_0, \beta_0 \, | \kappa^i) \rightarrow \max$
- Example: Consider last three years:



$$\hat{\kappa}_t = \hat{\kappa}_{t-1} + \beta_t$$
$$\beta_t = \beta_{t-1} + \lambda_t$$

  - Known trend in 0, unknown trend in -1 (possible trend change λ_0)
  - $L(\mu, \sigma^2, p, \sigma^2_{\epsilon,t}, \hat{\kappa}_0, \beta_0 | \kappa_{-2}, \kappa_{-1}, \kappa_0)$
  - $= f_{\mathcal{N}}(\kappa_0 - \hat{\kappa}_0 | \sigma^2_{\epsilon,t}, \widehat{\kappa_0}) \cdot f_{\mathcal{N}}(\kappa_{-1} - (\hat{\kappa}_0 - \beta_0) | \sigma^2_{\epsilon,t}, \hat{\kappa}_0, \beta_0) \cdot (f_{\mathcal{N}} * g)(\kappa_{-2} | \mu, \sigma^2, p, \sigma^2_{\epsilon,t}, \hat{\kappa}_0, \beta_0)$
  - $= f_{\mathcal{N}}(\epsilon_0 | \sigma^2_{\epsilon,t}, \hat{\kappa}_0) \cdot f_{\mathcal{N}}(\epsilon_{-1} | \sigma^2_{\epsilon,t}, \hat{\kappa}_0, \beta_0) \cdot \int_{\mathbb{R}} g(\lambda_0 | \mu, \sigma^2, p) \cdot f_{\mathcal{N}}(\kappa_{-2} - (\hat{\kappa}_0 - \beta_0 - (\beta_0 - \lambda_0)) | \sigma^2_{\epsilon,t}, \hat{\kappa}_0, \beta_0) d\lambda_0 \rightarrow \max$
- Knowing $\mu, \sigma^2, p, \sigma^2_{\epsilon}, \hat{\kappa}_0, \beta_0$, we can give a likelihood for the historic data→ Maximize it!

Parameter Estimation

Alternative I

- Consider the complete history: $L(\theta|\kappa_{-N,\dots,0}) \to \max$ with $\theta \coloneqq \mu, \sigma^2, p, \sigma^2_{\epsilon,t}, \hat{\kappa}_0, \beta_0$
- We can calculate the trend process recursively $\hat{\kappa}_{-s} = \hat{\kappa}_0 - s\beta_0 + \sum_{l=1}^{s-1} l \cdot \lambda_{-(s-1-l)}$, $0 \leq s$
  - $L(\mu, \sigma^2, p, \sigma^2_{\epsilon,t}, \hat{\kappa}_0, \beta_0 \,|\kappa_{-N,\dots,0}) = f_{\mathcal{N}}(\epsilon_0|\sigma^2_{\epsilon,t}, \hat{\kappa}_0) \cdot f_{\mathcal{N}}(\epsilon_{-1}|\sigma^2_{\epsilon,t}, \hat{\kappa}_0, \beta_0)$

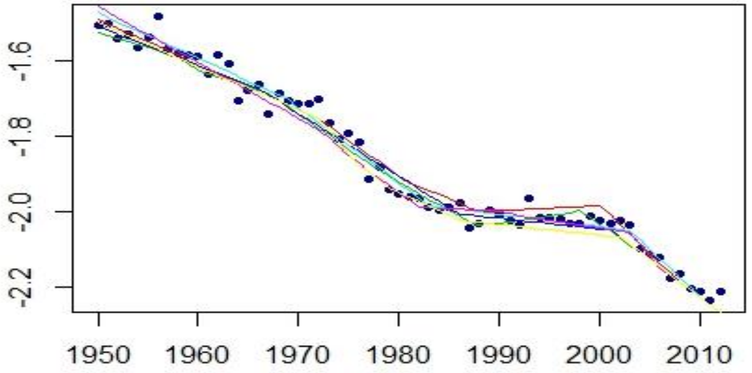
$$\cdot \int_{\mathbb{R}^{N-1}} \prod_{s=2}^{N} g(\lambda_{-(s-2)}|\theta) \cdot f_{\mathcal{N}}\left(\kappa_{-s} - (\hat{\kappa}_0 - s\beta_0 + \sum_{l=1}^{s-1} l \cdot \lambda_{-(s-1-l)})|\theta\right) d\lambda_{-(N-2),\dots,0} \to max$$

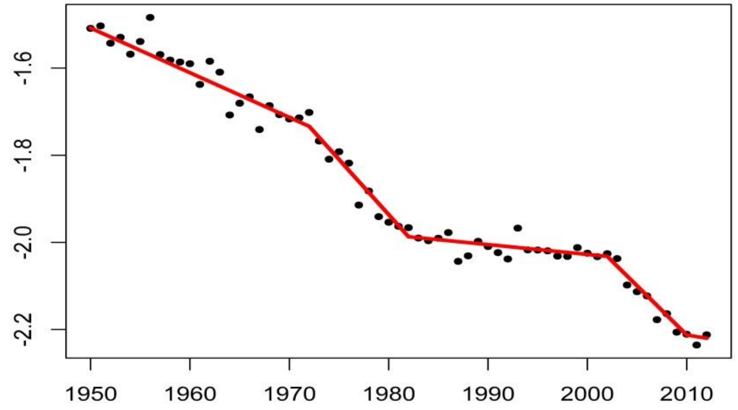
- Challenge: In parameter calibration, we need to solve and optimize this N-1 dimensional integral
- Possible Solution: Use Monte-Carlo integration
  - Basic idea: $\mathrm{I} = \int f(x)g(x)dx$, simulate $x^1, \dots, x^m$ with $x^i \sim g$, estimate $\hat{I} = \frac{1}{m}\sum_{i=1}^{m} f(x^i)$.
  - Here: Simulate $x^1, \dots, x^m$ trends according to $x^l = (\lambda_{-N+2}, \dots, \lambda_0)^l$ with $\lambda_j \sim g$
  - Calculate $\hat{I} = \frac{1}{m}\sum_{l=1}^{m} \prod_{j=-N}^{0} f_{\mathcal{N}}(\epsilon_j^i|\theta, x^l)$ for $i = 1,2$
  - Starting in $\mathrm{t} = 0$ we simulate historic trends. The estimated parameters can be used for projections directly.

Parameter Estimation

Alternative I

- Advantages:
  - Maximum of consistency in forecasts
  - Flexibility on distributional assumptions
- Disadvantages and open issues:
  - No historic trends
  - Trends x^l with a high likelihood ($\prod_{j=-N}^{0} f_{\mathcal{N}}\left(\epsilon_j \middle| \theta, x^l\right)$) are extremely rare
  - Full complexity of trend model not achieved
  - Numeric approaches failed due to the pole of g around zero
- However, with the available methods, convergence can only be achieved by very strong, critical interventions
- NLD-females ($\sigma_\epsilon^2 = \sigma_{\epsilon,t}^2, \beta_0, \hat{\kappa}_0$ fixed)
  - $\mu = -5.0, \sigma^2 = 0.03125,\ p = 0.1$ with MLE
  - $\mu = -4.0, \sigma^2 = 0.024,\ p = 0.0634$ with historic trend estimation (Alternative II)





Parameter Estimation

Alternative II

Combined likelihood-like approach

- Include the distribution of the trend changes used for simulations within the calibration
- Simplified model, which has the essential advantages of the previous model and can be calibrated efficiently
- For $k \in 1, \dots, m$ estimate optimal trend process $\beta_0, \hat{\kappa}_0, \lambda_{-N+2}, \dots, \lambda_0$

  $f_{\mathcal{N}}\left(\kappa^i_{-N,\dots,0} \middle| \sigma^2_{\epsilon,t}, \beta_0, \hat{\kappa}_0, \lambda_{-N+2}, \dots, \lambda_0\right) \cdot g(\lambda_{-N+2}, \dots, \lambda_0 | \mu, \sigma^2, p)$, where exactly k of $\lambda_{-N+2}, \dots, \lambda_0$ are unequal to zero (trend curve with k trend changes). Update $\sigma^2_{\epsilon,t}, \mu, \sigma^2, p$ iteratively.

  - Estimate k trend change points $t_1, \dots, t_k$
  - Estimate k changes in slope $\lambda_{t_1} \dots, \lambda_{t_k}$
  - Estimate starting trend and value $\beta_0, \hat{\kappa}_0$
  - →2k+2 parameters to estimate

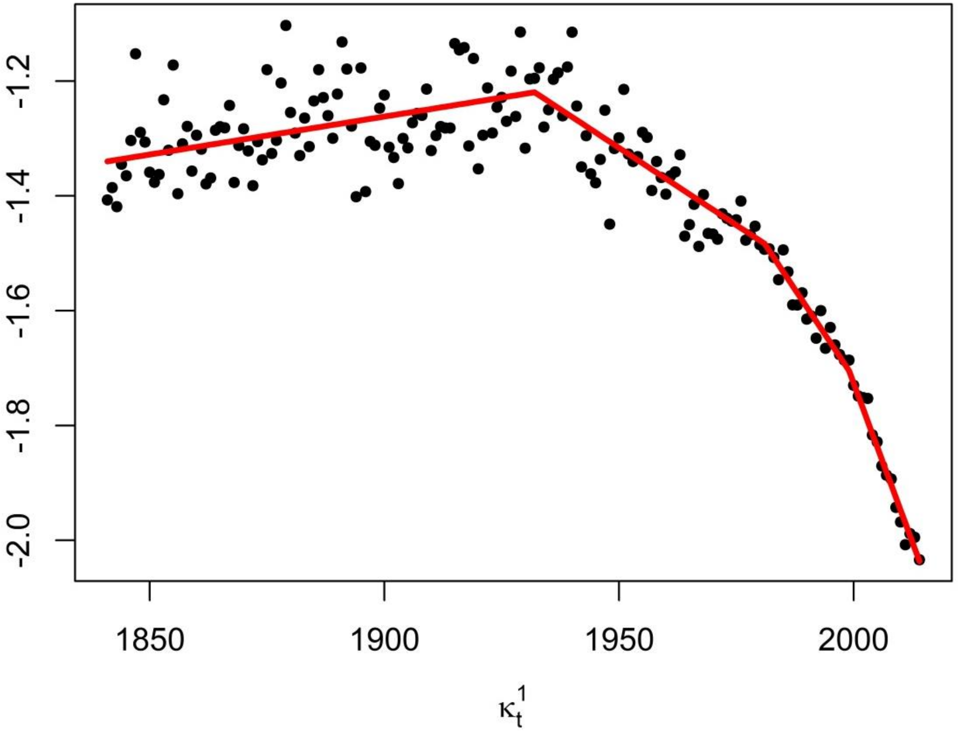
BEFORE: Update $\sigma^2_{\epsilon,t}$ iteratively

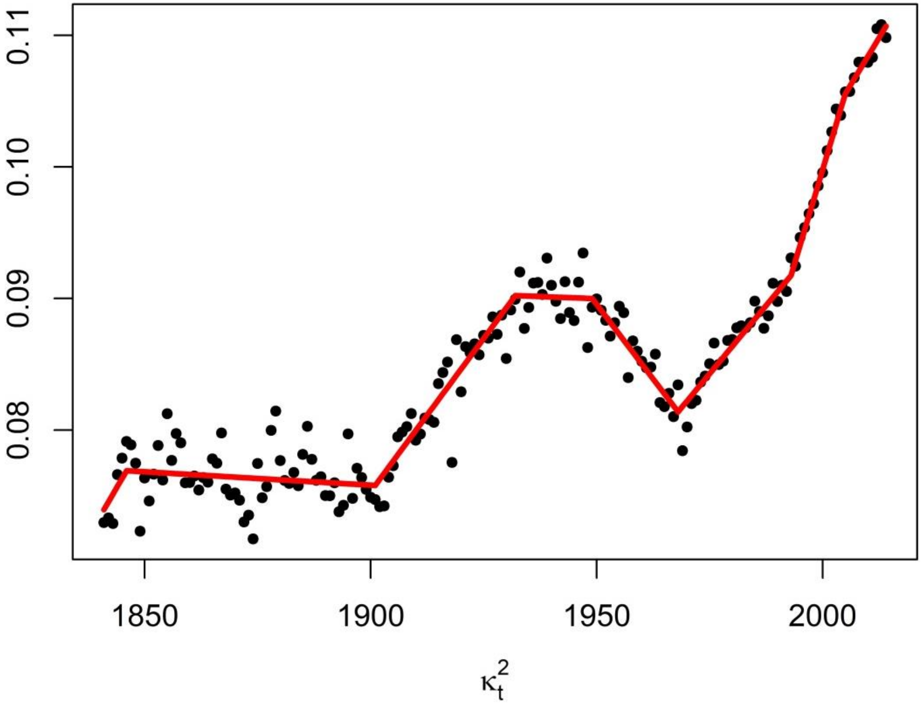
- Find trend s.th. $f_{\mathcal{N}}\left(\kappa^i_{-N,\dots,0} \middle| \sigma^2_{\epsilon,t}, \beta_0, \hat{\kappa}_0, \lambda_{-N+2}, \dots, \lambda_0\right) \cdot g(\lambda_{-N+2}, \dots, \lambda_0 | \mu, \sigma^2, p) \rightarrow max$
  - Change the set of candidates
  - Choose best model in the set of candidates with AIC

Parameter Estimation

Alternative II

Optimal historic trend models





- Consistency between historic trends and stochastic simulation
  - One step approach: All parameters required for stochastic forecasts are part of the calibration: $p, \sigma_{\epsilon,t}^2, \mu, \sigma^2, \beta_0, \hat{\kappa}_0$

Parameter Uncertainty

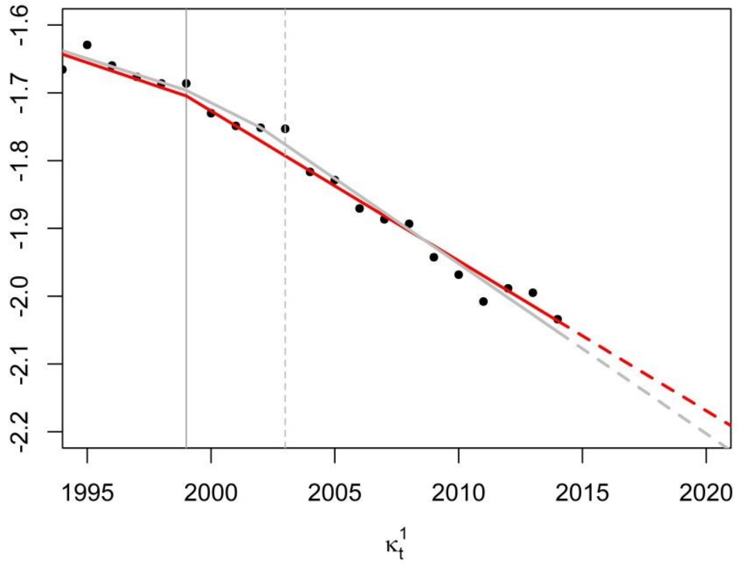
Using Akaike weights

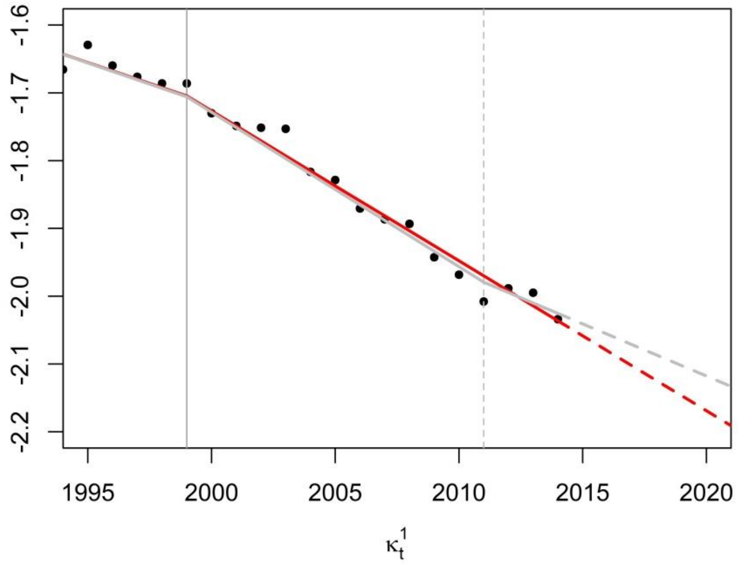
Parameters to be estimated for a projection: $\boldsymbol{p}, \boldsymbol{\sigma^2_{\epsilon,t}}, \boldsymbol{\mu}, \boldsymbol{\sigma^2}, \boldsymbol{\beta_0}, \boldsymbol{\hat{\kappa}_0}$

- Exemplarily: Uncertainty in starting values $\beta_0, \hat{\kappa}_0$:
  - Chance for another trend change since the last detected change due to limited data after that potential trend change.
  - Fix calibrated trend changes $(\beta_{0,0}, \hat{\kappa}_{0,0})$ and add an additional trend change in each year after the last trend change s=1,…,N-1 where N is the length of data after the last trend change. Denote the resulting starting values as $\beta_{0,s}, \hat{\kappa}_{0,s}$. Choose local optima of IC_s as further candidates $\coloneqq s^*$. Use $\beta_{0,s^*}, \hat{\kappa}_{0,s^*}$ in a simulation with weight:

$$\mathbb{P}\left(\beta_{0,s^*}, \hat{\kappa}_{0,s^*}\right) = \frac{\exp(-\frac{1}{2}(IC_s - IC_0))}{\sum_{i \in s^*} \exp(-\frac{1}{2}(IC_i - IC_0))}$$

  - Three scenarios:
    - Basis (red trend) with 78%
    - Steeper mortality trend with 8%
    - Flatter mortality trend with 14%





Conclusion - Summary

ICA CIA BERLIN 2018

- Period Effects of Parametric Mortality Models seem to be trend stationary with random changes in the trends slope.
- Model Risk with respect to the trend process is highly significant
- Parameter estimation of a trend model based on
  - Historic trends within a two step approach
- Complete MLE with some critical limitations
- One step approach that combines the advantages of both approaches
  - Consistent projections of the trend
  - Interpretable historic trends
- Parameter Uncertainty included with Akaike weights

Literature

- Börger, M., Fleischer, D., Kuksin, N., 2014. Modeling Mortality Trend under Modern Solvency Regimes. ASTIN Bulletin, 44: 1–38.
- Börger, M. and Schupp J., 2018. Modeling Trend Processes in Parametric mortality models. Insurance: Mathematics and Economics, 78: 369-380.
- Burnham, K.P. and Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer Science & Business Media.
- Cairns, A., Blake, D., Dowd, K., 2006. A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration. Journal of Risk and Insurance, 73: 687–718.
- Hunt, A. and Blake, D. (2015). Modelling mortality and longevity risk management – chapter 4: Consistent mortality projections allowing for trend changes and cohort effects. Doctoral thesis, Cass Business School.
- Li, J. S.-H., Chan, W.-S., and Cheung, S.-H. (2011). Structural changes in the Lee-Carter indexes: detection and implications. North American Actuarial Journal, 15(1): 13–31.
- Muggeo, V. M., 2003. Estimating regression models with unknown breakpoints. Statistics in medicine, 22.19: 3055–3071.

Literature

- Muggeo, V.M. and Adelfio, G., 2011. Efficient change point detection for genomic sequences of continuous measurements. Bioinformatics, 27 (2):161-166.
- Sweeting, P., 2011. A Trend-Change Extension of the Cairns-Blake-Dowd Model. Annals of Actuarial Science, 5: 143–162.

Contact

Johannes Schupp(M.Sc.)
+49 (731) 20 644-241
j.schupp@ifa-ulm.de