Extension, Compression, and Beyond*

A Unique Classification System for Mortality Evolution Patterns

* And selected aspects of related papers.

- ICA 2018, Berlin, June, 8th 2018
- Martin Genz
- Joint work with Matthias Börger, Jan Feifel, Markus Pauly, and Jochen Ruß
- Institute for Finance and Actuarial Sciences and University of Ulm, Germany

About the speaker

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)

Martin Genz

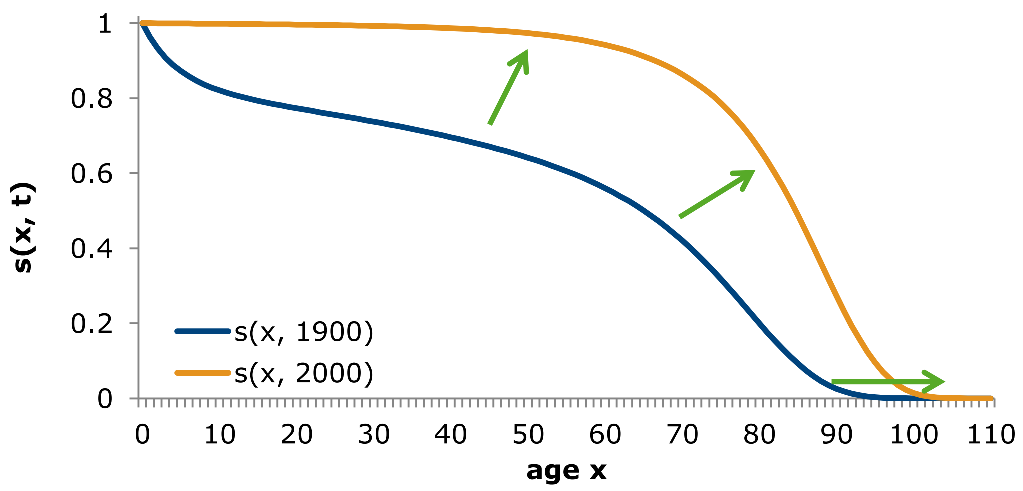
- joined ifa in 2013
- Master of Science (Mathematics and Management, University of Ulm, 2013)
- Ph.D. student since 2013
- Junior member of the DGVFM since 2014

Agenda

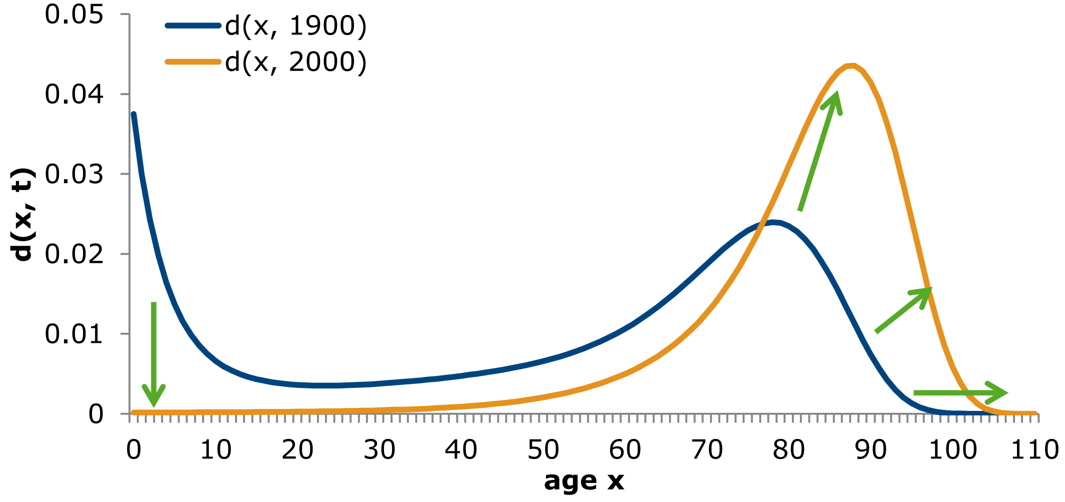
Key question

- Life expectancy increases in many countries.
- But changes in life expectancy (and other typically used statistics) are only a **consequence of the underlying change of the age distribution of deaths**.

Survival curve of Swedish females 1900 and 2000



Deaths curve of Swedish females 1900 and 2000



- Key question: **How does the shape of these curves change over time?**

Classification of mortality evolutions in the past

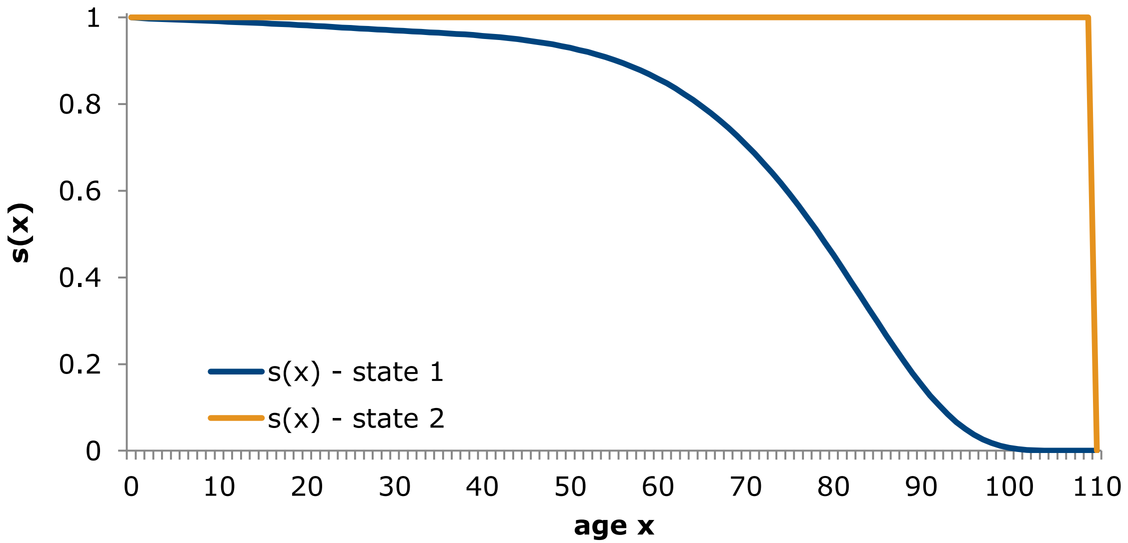
Shortcomings

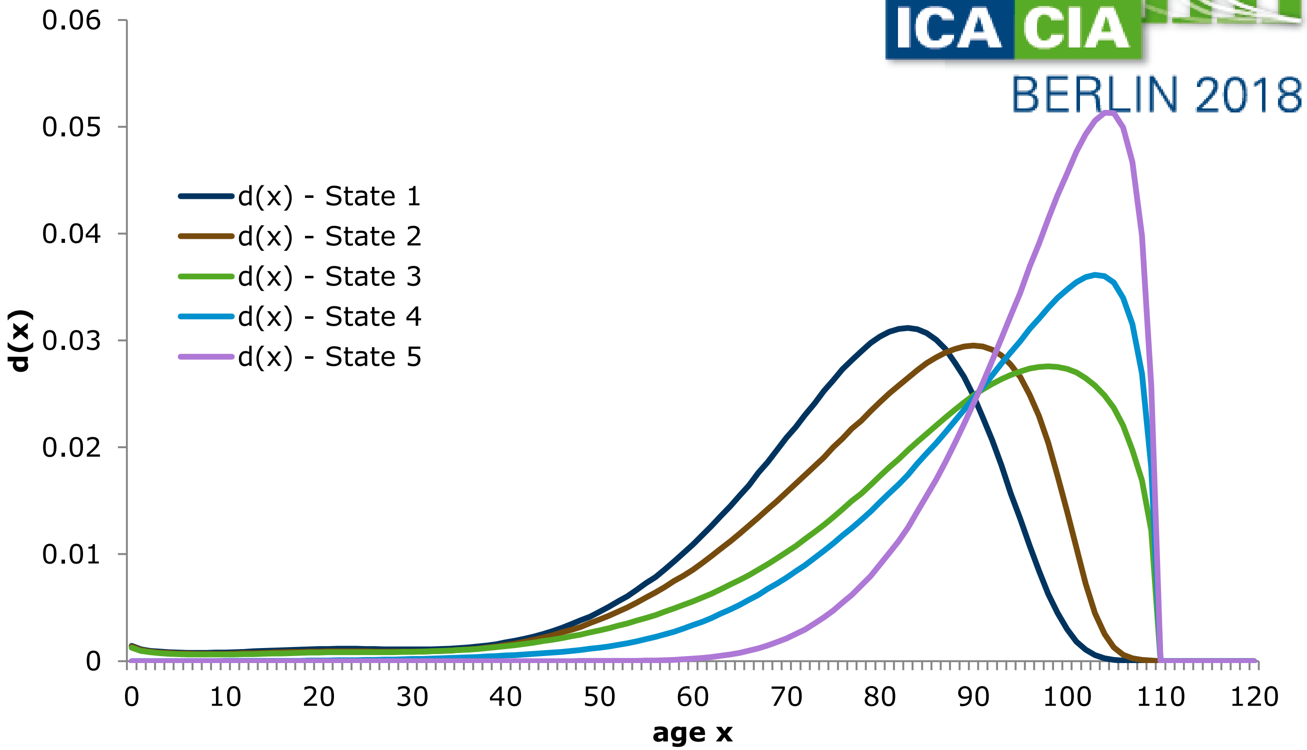
- There is a variety of literature on the question how the age distribution of deaths changes over time and we have identified some shortcomings there.
- In what follows, we discuss examples for four shortcomings:
  - Different notions for certain observations have been established (e.g. compression, extension, rectangularization, etc.) but sometimes these **scenarios were defined imprecisely**.
  - Some of these scenarios were supposed to be **mutually exclusive** but there are counterexamples.
  - Several often used **statistics are insufficient or even misleading**.
  - Often effects caused by the **choice of a certain age range** under observation were not considered.

Classification of mortality evolutions in the past

Shortcomings

- **Imprecise scenario definitions:**
  - E.g., **rectangularization** is defined by a final state.





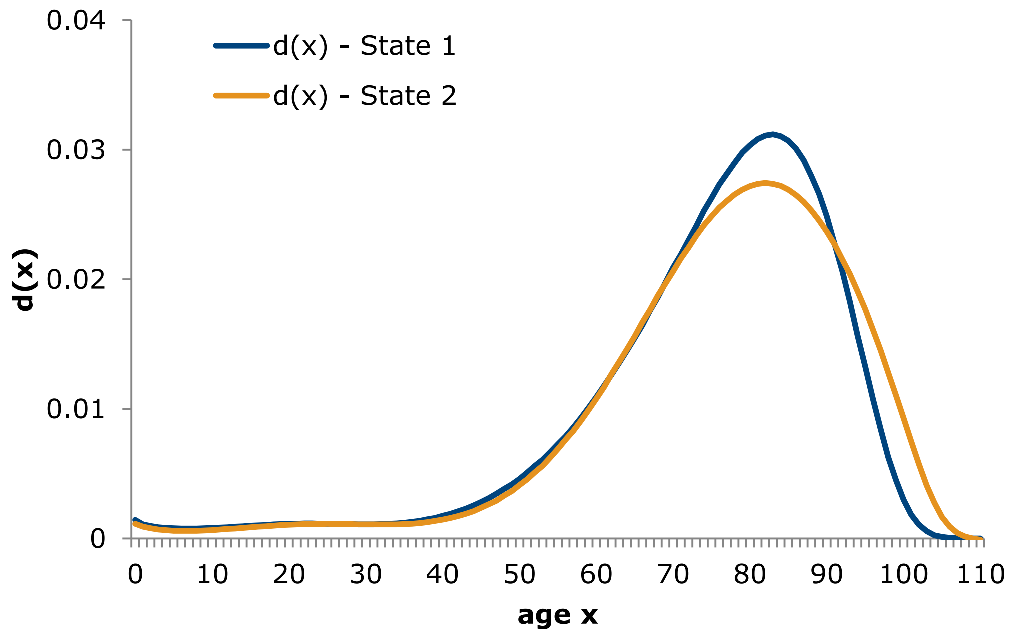
- Rectangularization: Present state (blue line) and final state (yellow line)

- Overall rectangularization…
- … but two different evolutions between State 1 and 3 and State 3 and 5, respectively.

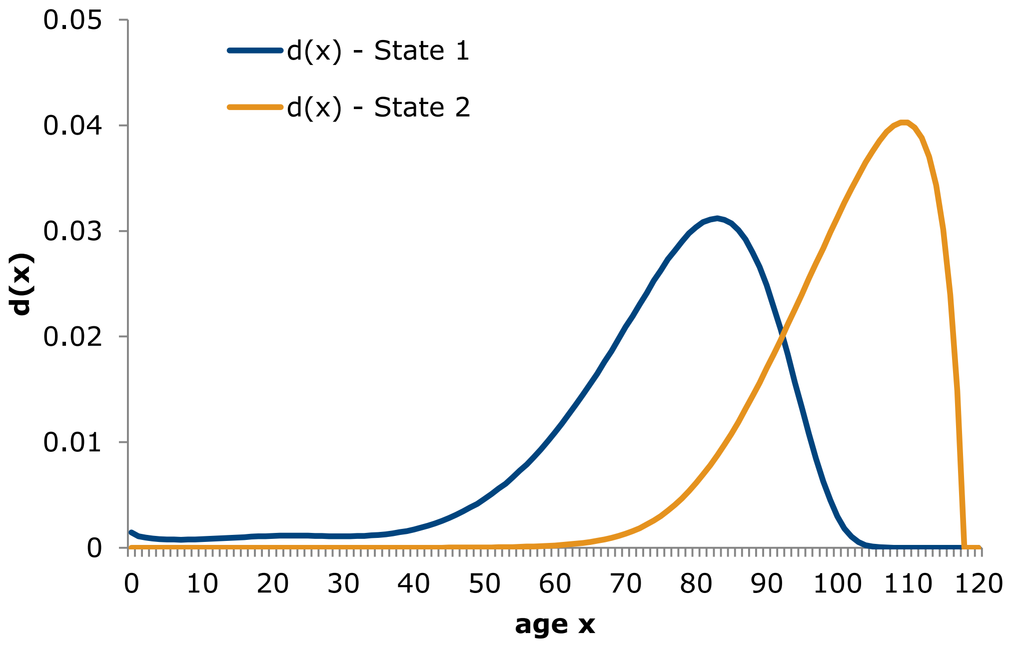
Classification of mortality evolutions in the past

Shortcomings

- **Exclusiveness of scenarios:**
  - E.g., **compression** and **shifting mortality** are assumed to be opposing scenarios.



- Neither compression nor shifting mortality prevail.



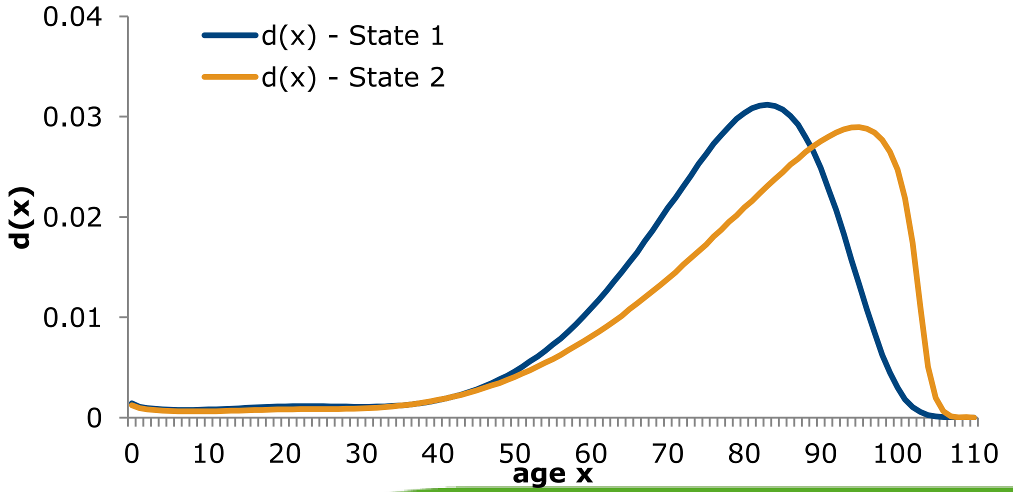
- Compression and shifting mortality coexist.

Classification of mortality evolutions in the past

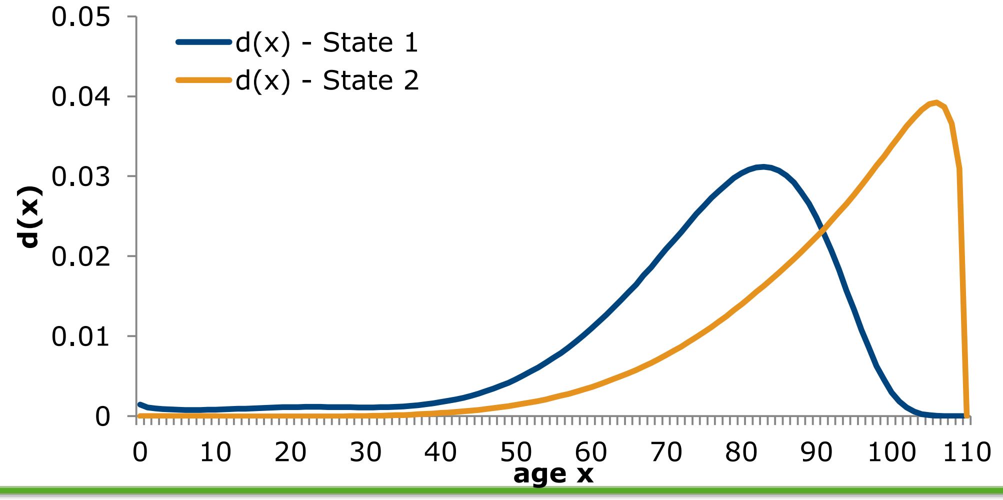
Shortcomings

- **Insufficient or misleading statistics:**
  - For example, in literature **compression** is linked with…
    - increasing M and decreasing SD(M+),
    - decreasing IQR,
    - decreasing Cα (e.g. C50).

– Example 1: **M increases and SD(M+) decreases**, but there is no compression:



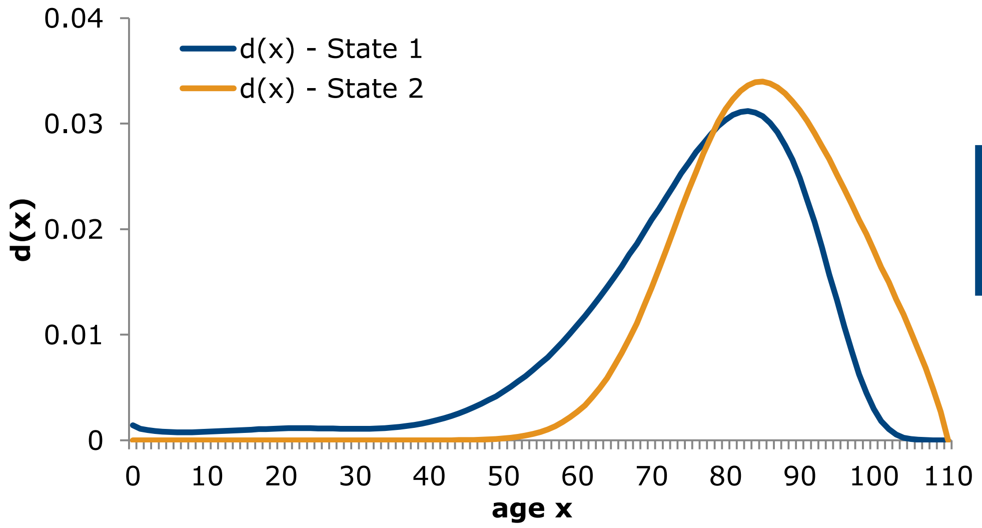
– Example 2: Considerable process of compression, but **IQR stays constant:**



Classification of mortality evolutions in the past

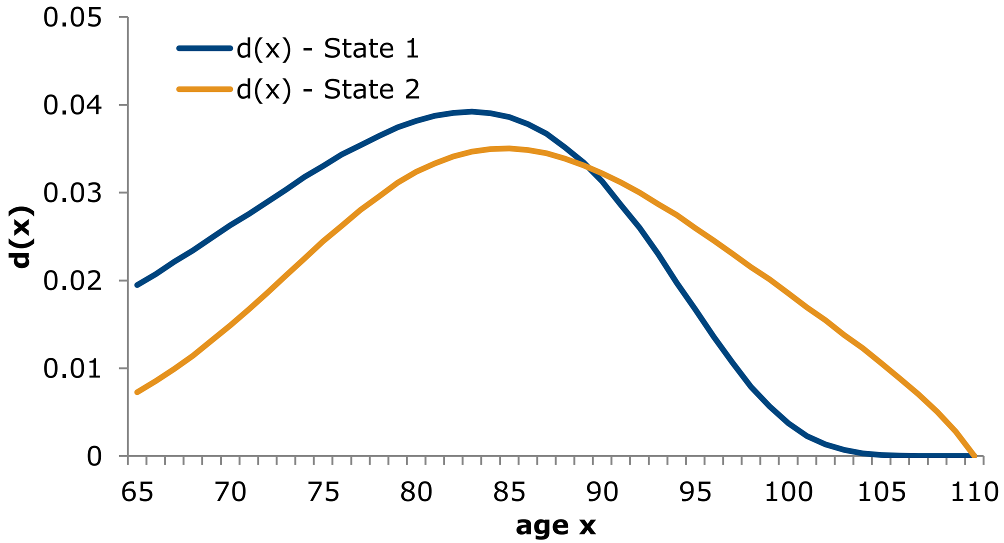
Shortcomings

- The choice of the **age range matters:**
  - The age range should be chosen depending on the question at hand.



- Compression into the higher ages

s(65)=1



- No compression within the higher ages.

A new classification framework

Requirements

In light of these shortcomings of previous approaches, a new classification system should…

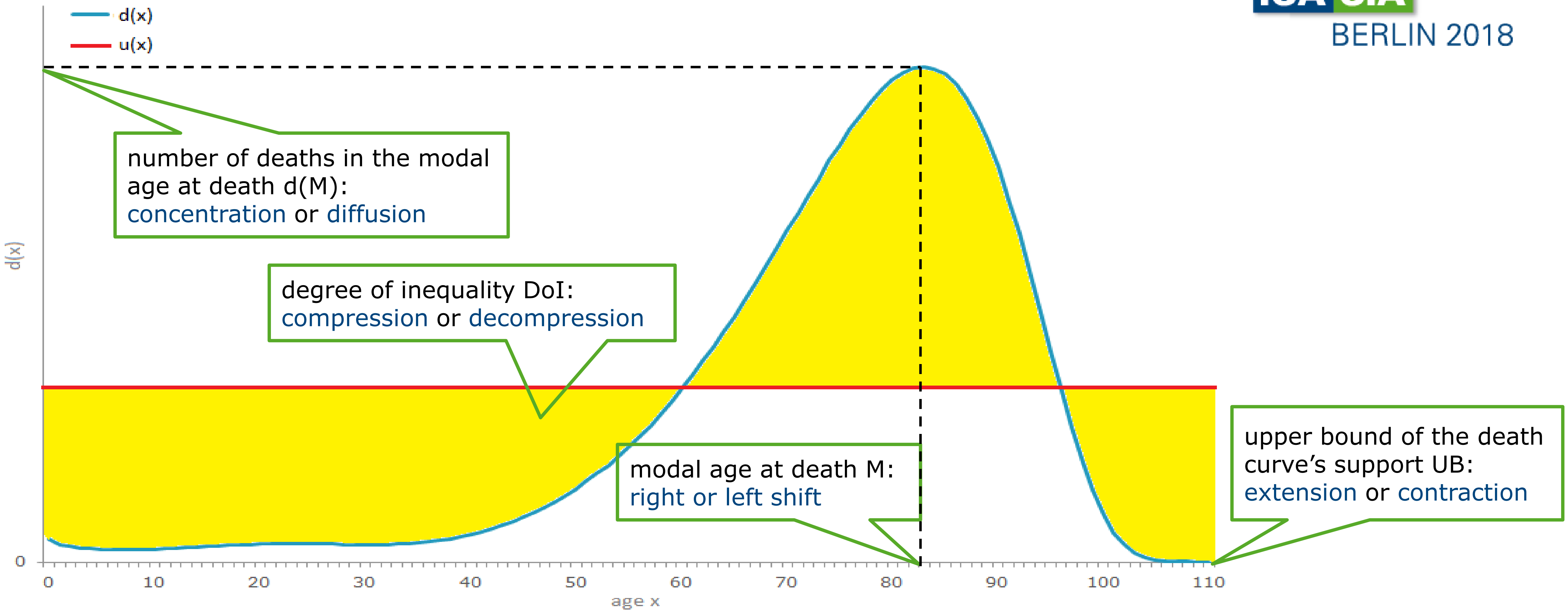
- … uniquely categorize every material change in mortality patterns,
- … allow for mixed scenarios,
- … be applicable to different age ranges,
- … build on statistics that can be feasibly calculated and easily interpreted,
- … be extendable by additional components if needed.

Our new approach:

- We use the deaths curve as basis for the framework.
- We define 4 characteristics of the deaths curve for a unique classification of observed mortality evolutions.

A new classification framework

Details



A new classification framework

Details

Each scenario is defined by a **4-dimensional vector** where each component can have three specifications:

component	attainable states	indicates changes in…
M	right shift / neutral / left shift	… the position of the deaths curve
UB	extension / neutral / contraction	
DoI	compression / neutral / decompression	… the shape of the deaths curve
d(M)	concentration / neutral / diffusion	

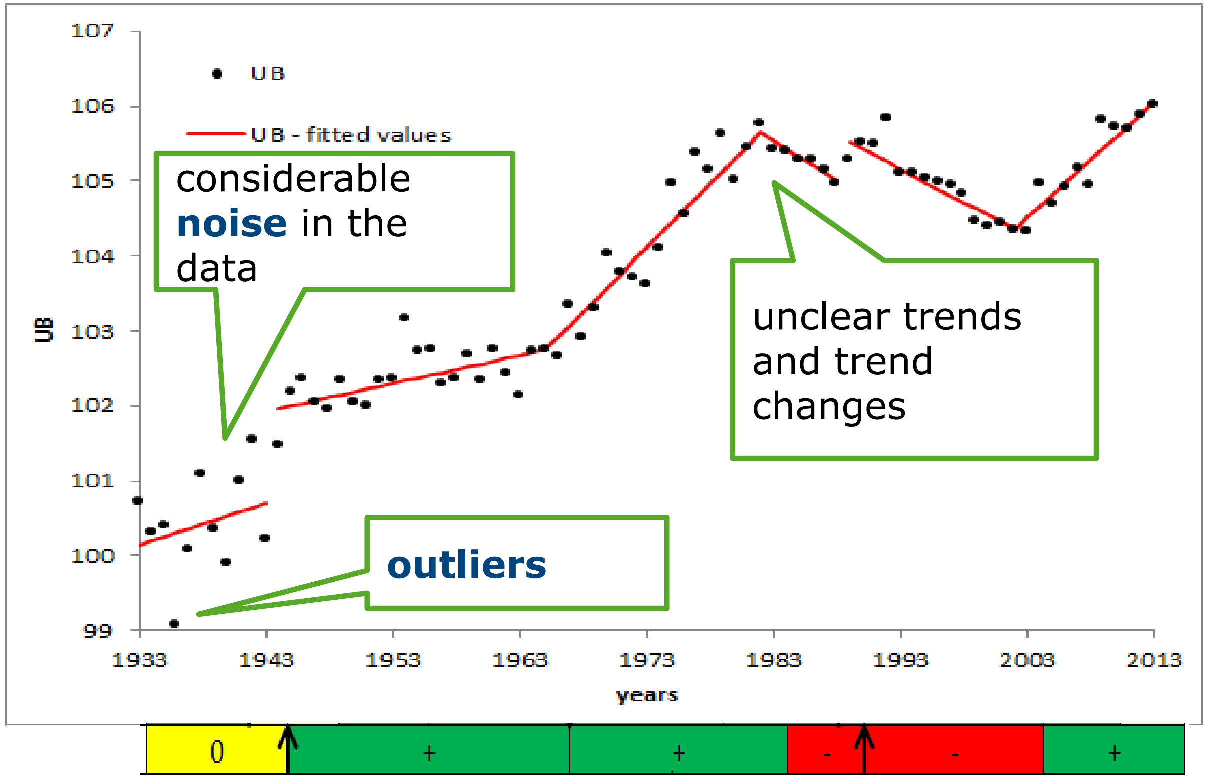
- This allows for $3^4=81$ different scenarios (some of which might not be relevant in practice)
- The framework satisfies the requirements:
  - Each observed mortality evolution can uniquely be classified in one of those scenarios.
  - Pure and mixed scenarios are included.
  - The framework can be applied to age ranges starting at any given age up to UB.
  - Feasible and easily interpretable statistics are used.
  - The framework is extendable by additional statistics if needed.

In the paper, we discuss different issues in estimating these statistics, e.g. how to estimate UB, and methods for the detection of trends.

A new classification framework

Application: UB for US females, starting age 10

- For each calendar year, we estimate the four statistics. This gives us four time series.

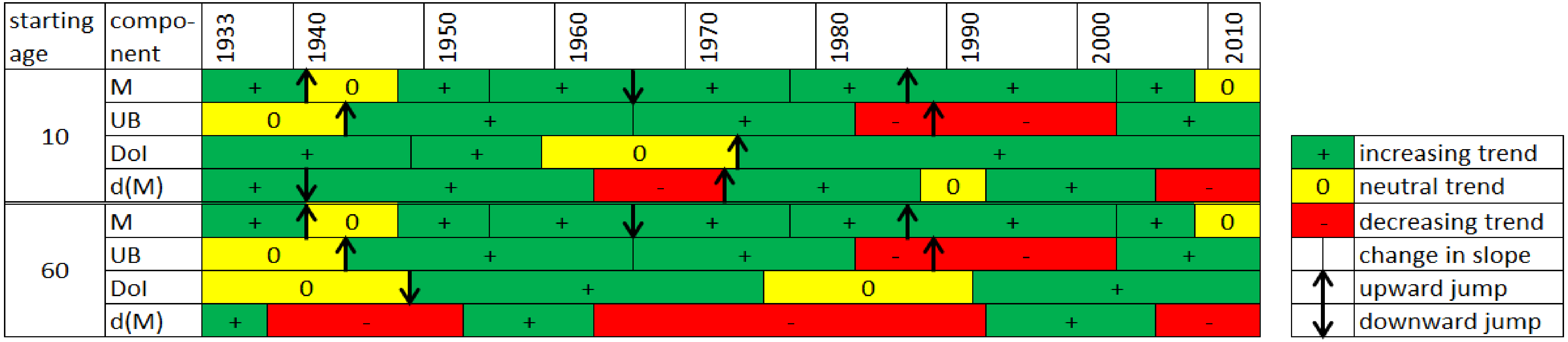


(Source: Human Mortality Database (HMD) and own calculation)

- The time series for each statistic…
  - … are "**noisy**" → We need to eliminate the noise.
  - … have **outliers** → We need to eliminate the outliers.
  - … in parts have **unclear trends** → We need a method to identify periods of stable trends.
- We fit a polygonal curve to the data and allow for **changes in slope with or without jumps**. Thus we achieve a decomposition of the time range into periods, where the time series follows a linear trend.
- For each period we detect whether a trend is significantly **increasing**, **neutral**, or **decreasing** using a statistical test.
- This can be illustrated by a time bar.

A new classification framework

Application: The mortality evolution of US females



What does this mean for the classification framework?

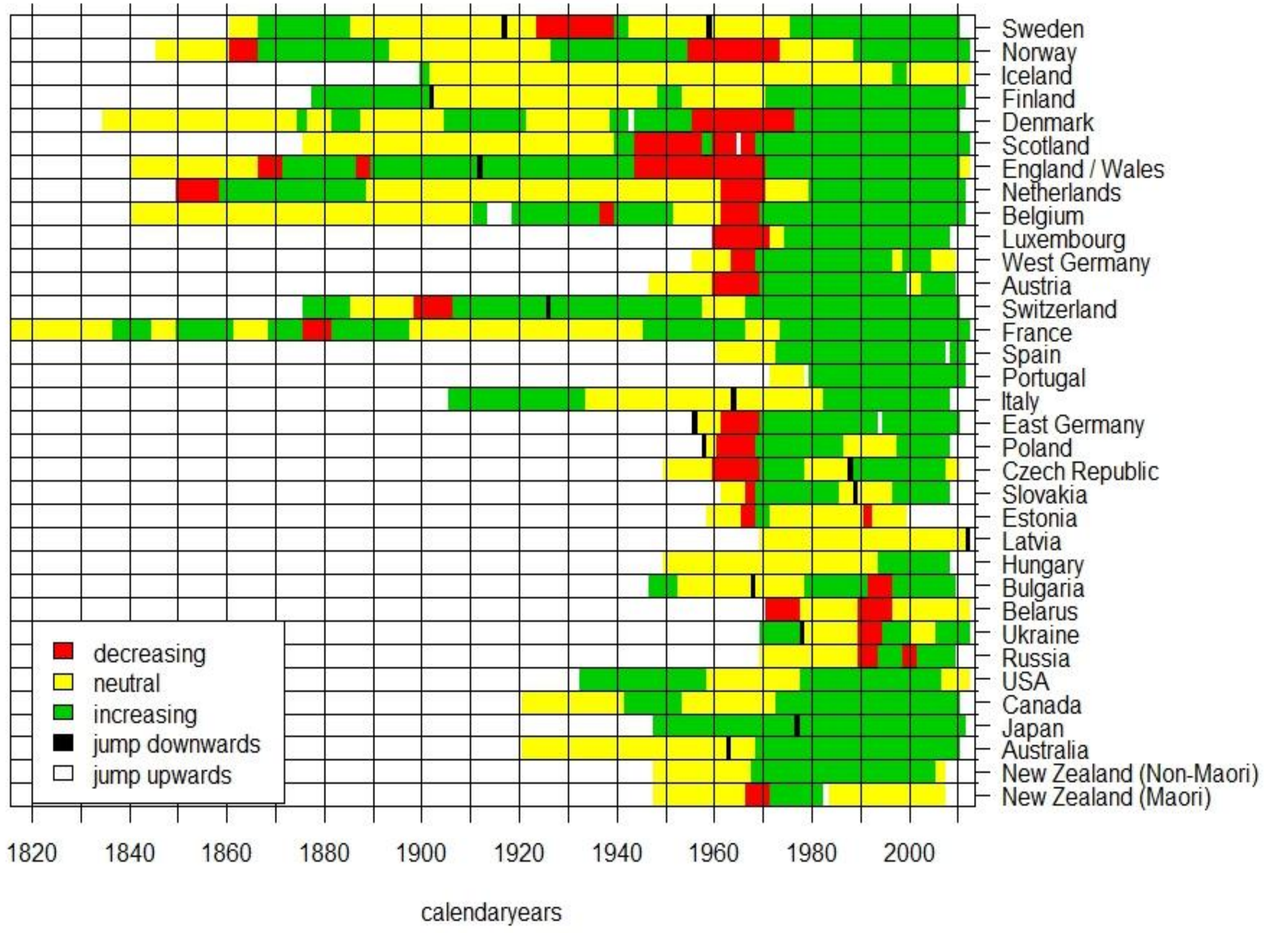
We see that…

- … each component of the vector develops independently from the others (no redundant information).
- … there are mixed scenarios (rather the rule than an exception).
- … there are different scenarios for different age ranges (age range matters).

Related Papers – 1: A Comprehensive Analysis of the Patterns of Worldwide Mortality Evolution

Typical Results

- We applied the classification framework on a comprehensive dataset from the HMD and analyzed differences in the trends of the four statistics between…
  - … neighboring countries, …
  - … males and females, and …
  - … starting ages 0 and 60.
- We found:
  - Dependencies in the trends among neighboring countries (see figure).
  - Differences in the trends between Eastern European populations an the rest of the world.
  - Differences in the trends between males and females for both starting ages, which are smaller for starting age 60.
  - Greater differences between the starting ages for males than for females.



Related Papers – 2: Who wants to live forever? An analysis of the maximum lifespan in the US

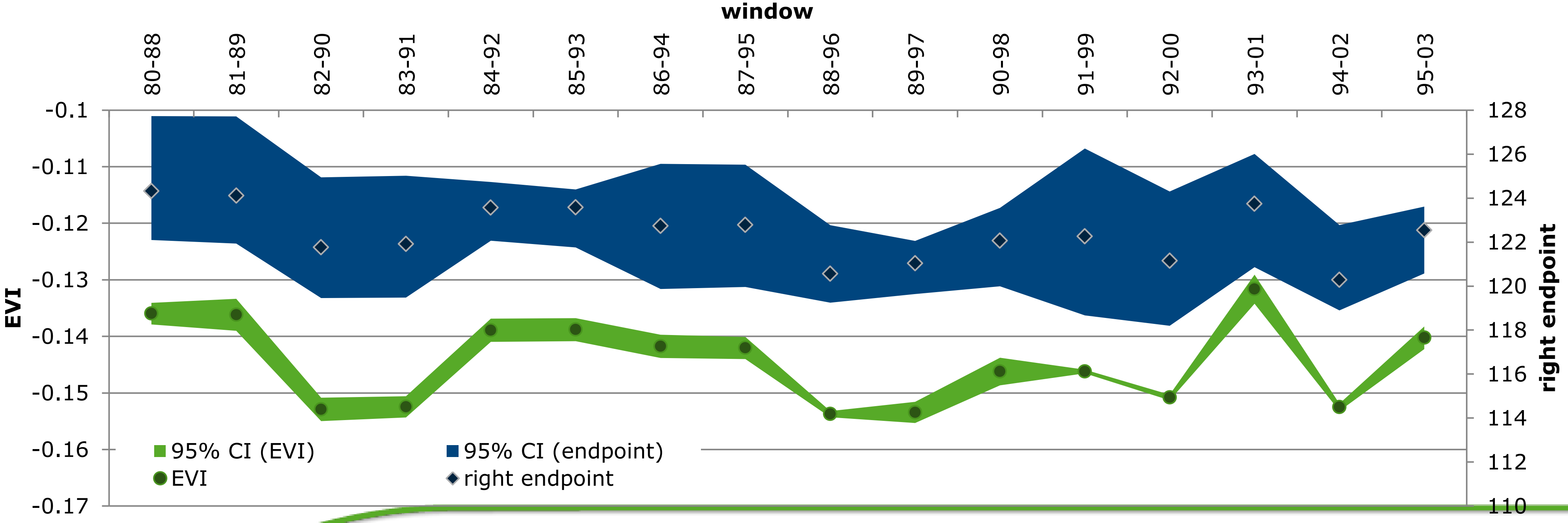
The Estimation of the Right Endpoint - Idea

- So far, we assumed that there are sufficiently good estimators for each of the four statistics of the classification framework. However, for example the **estimation of UB** is difficult.
- We have estimated this statistic using a **high quantile plus the remaining life expectancy** in this age. But this estimator highly depends on the quality / availability of old age mortality data.
- This is a typical problem of **extreme value theory (EVT)**. Still, two challenges remain:
  - Old age mortality **data** typically is sparse and often censored.
  - Solution: Combine (parts of) the HMD-data with the (complete) data from the International Database on Longevity (IDL) → **combined dataset** (CDS).
- Resulting data is randomly right censored. Thus "classical" EVT-methods are not sufficient.
  - Solution: Einmahl et al. (2008) introduced estimators for the extreme value index γ (EVI) and the right endpoint of the deaths curve for randomly right censored data.
  - For the existence of a right endpoint x_F of a distribution function F, it holds:
    - If $\boldsymbol{\gamma < 0}$ → $\boldsymbol{x_F < \infty}$, i.e. a right endpoint of F does exist.
    - If $\boldsymbol{\gamma > 0}$ → $\boldsymbol{x_F = \infty}$, i.e. a right endpoint of F does not exist.

Related Papers – 2: Who wants to live forever? An analysis of the maximum lifespan in the US

The Estimation of the Right Endpoint - Results

- We have used different estimation methods and different parameters of the estimation algorithm to estimate the EVI.
  - Key question: is there a finite right endpoint of the deaths curve?
  - If so, we determine a best estimate for it.
- Below a typical example for our results indicating that no one will live forever:



(source: HMD, IDL, and own calculations)

Summary

ICA CIA BERLIN 2018

In the paper, we have…

- … identified **shortcomings** of previous approaches for classification of mortality scenarios,
- … derived **requirements** for a new framework,
- … identified 4 central **characteristics** of the deaths curve,
- … derived a **new classification framework** based on these characteristics, which
  - … builds on clear scenario definitions,
  - … provides a unique classification for each mortality evolution,
  - … allows for mixed scenarios,
  - … is applicable for different age ranges,
- … discussed **methods** we suggest for the determination of prevailing states, and
- … **applied the framework** to concrete data for US females.

In two related papers, we have…

- … applied the framework on a comprehensive dataset from the HMD, and …
- … **estimated the right endpoint of the deaths curve** using censored EVT and a combined dataset.

Thank you for your attention!

References

Börger, M., Genz, M., & Ruß, J. (2016). Extension, Compression, and Beyond – A Unique Classification System for Mortality Evolution Patterns. https://www.ifa-ulm.de/fileadmin/user_upload/download/forschung/2016_ifa_Boerger-etal_Extension-Compression-and-Beyond.pdf

Einmahl, J.H., Fils-Villetard, A., and Guillou, A. (2008). Statistics of Extremes Under Random Censoring. *Bernoulli*, 14(1): 207-227.

Feifel, J., Genz, M., Pauly, M. (2017). Who wants to live forever? – An Analysis of the Maximum Lifespan in the US. https://www.ifa-ulm.de/fileadmin/user_upload/download/forschung/2017_ifa_Feifel-etal_Who-wants-to-live-forever-An-analysis-of-the-maximum-lifespan-in-the-US.pdf

Genz, M., 2017. A Comprehensive Analysis of the Patterns of Worldwide Mortality Evolution. https://www.soa.org/Library/Monographs/Life/Living-To-100/2017/2017-living-100-monograph-genz-paper.pdf

HMD (2015). Human Mortality Database (http://www.mortality.org)

IDL (2015). International Database on Longevity (http://www.supercentenarians.org)