Extension, Compression, and Beyond – A Unique Classification System for Mortality Evolution Patterns

Matthias Börger
Institut für Finanz- und Aktuarwissenschaften (ifa)
Lise-Meitner-Straße 14
89081 Ulm, Germany

Martin Genz (corresponding author)
Institut für Finanz- und Aktuarwissenschaften (ifa) & Institut für Versicherungswissenschaften, Universität Ulm
Lise-Meitner-Straße 14
89081 Ulm, Germany
phone: +49 731 20 644 264
fax: +49 731 20 644 299
email: m.genz@ifa-ulm.de

Jochen Ruß
Institut für Finanz- und Aktuarwissenschaften (ifa) & Institut für Versicherungswissenschaften, Universität Ulm
Lise-Meitner-Straße 14
89081 Ulm, Germany

Abstract

There exists a variety of literature on the question how the age distribution of deaths changes over time as life expectancy increases. However, corresponding terms like extension, compression, or rectangularization are sometimes defined only vaguely, and statistics used to detect certain scenarios can be misleading. The matter is further complicated since often mixed scenarios prevail and the considered age range can have an impact on observed mortality patterns.

In this paper, we establish a unique classification framework for realized mortality scenarios which allows for the detection of both, pure and mixed scenarios. It determines whether changes of the deaths curve over time show elements of extension or contraction, compression or decompression, left or right shifting mortality, and concentration or diffusion. The framework can not only test the presence of a particular scenario, but also assign a unique scenario to any observed mortality evolution. Furthermore, it can detect different mortality scenarios for different age ranges in the same population. We also present a methodology for the implementation of our classification framework and apply it to mortality data for US females.

Key words

Mortality scenario classification, longevity, rectangularization, shifting mortality, extension, compression

1. Introduction

Mortality evolutions, i.e. realized changes in mortality rates, have been analyzed extensively in the last decades. These analyses typically deal with changes in the distribution of lifetimes and hence go far beyond determining trends in the evolution of life expectancy. In this sense, changes in aggregated statistics like life expectancy are simply a consequence of the underlying change of the age distribution of deaths.

There is a wide range of literature on the question how realized mortality changed over time and how patterns of past developments (which we also call mortality scenarios) can be described and classified. In this context, different terms have been created, e.g. rectangularization, compression, extension, expansion, and shifting mortality. These terms have been helpful in the analysis of historical mortality evolution patterns. Their definitions are, however, mostly intuitive which can lead to ambiguity. For instance, Fries (1980) defines rectangularization as the convergence of the survival curve to a theoretical but not completely reachable final state, where everybody dies at the same age. Many authors have adopted this definition (see e.g. Cheung et al. 2005; Kannisto 2000; Manton and Tolley 1991). However, as we show in Section 2, this definition can be misleading. Similarly intuitive but difficult to verify from observed mortality patterns is the definition of compression in Debón et al. (2011) as a “state in which mortality from exogenous causes is eliminated and the remaining variability in the age at death is caused by genetic factors.” Thus, a precise and feasible definition for each mortality scenario is necessary to test its occurrence in practice.

Furthermore, different authors have defined certain scenarios in different ways. In contrast to the aforementioned intuitive definition in Debón et al. (2011), many authors use certain statistics of the deaths curve, i.e. the age distribution of deaths, to define compression. According to Kannisto (2001), (old-age) compression can be observed if the modal age at death M, i.e. the age with the largest number of deaths, increases and $SD(M+)$, i.e. the standard deviation of the distribution of deaths above M, decreases at the same time. This definition is (at least implicitly) applied by other authors, e.g. Cheung and Robine (2007) or Ouellette and Bourbeau (2010). Wilmoth and Horiuchi (1999), on the other hand, identify compression by a shrinking inter-quartile range (IQR), i.e. the length of the age range between the 25th and the 75th percentile of the distribution of deaths. Analogously, Kannisto (2000) uses the so called $C\alpha$-statistics, i.e. the shortest age range in which $\alpha\%$ of all deaths occur. Thatcher et al. (2010) observe compression if the slope parameter in a logistic mortality model increases with time. We will show in Section 2 that different definitions of compression will not always yield the same results.

Scenario definitions can also be critical when they only rely on observations for a rather small age range. For instance, when analyzing the evolutions of M and $SD(M+)$, one completely ignores the mortality evolution for all ages below M. As we show in an example in Section 2, if M increases and $SD(M+)$ decreases, compression need not be present for the whole age range under consideration.

The distinction between different scenarios is also not always clear. For instance, Wilmoth (2000) states that rectangularization should be "best thought of as 'compression of mortality'". Also for Myers and Manton (1984), compression and rectangularization seem to be the same scenario. Others like Nusselder and Mackenbach (1996) see rectangularization as a special case of compression in which the life expectancy increases with time. A similar issue exists for definitions of extension, expansion, and shifting mortality. For example, Debón et al. (2011) use the terms expansion and shifting mortality but do not explain the differences between them. Others define the three terms in different ways: Wilmoth and Horiuchi (1999) use the term expansion if the force of mortality decreases faster for older ages than for younger ages; Bongaarts (2005), on the other hand, explains the scenario of shifting mortality as a result of "delays in the timing of deaths", i.e. the force of mortality curve exhibits simply a shift in age; in Cheung et al. (2005) the term "longevity extension" is used for a scenario where longevity beyond the modal age at death increases.

Sometimes, one scenario is defined by the absence of some other scenario. For instance, Canudas-Romo (2008) regards shifting mortality as a scenario “where the compression of mortality has stopped”. Obviously, such a definition implies that these scenarios, i.e. shifting mortality and compression, are mutually exclusive. This rules out mixed scenarios by definition. As we will see later, elements of different mortality scenarios can often be present at the same time. Therefore, analyses which solely focus on testing for one particular scenario, e.g. compression, can never provide a comprehensive insight into the mortality evolution.

In this paper, we address these issues and establish a unique classification framework for mortality scenarios. The framework is based on observed changes in the deaths curve for the age range under consideration. We build on existing concepts like shifting mortality, extension, and compression and combine these concepts to a framework which particularly allows for the detection of mixed scenarios of mortality change. We provide precise definitions of scenarios and show how they can be identified. Furthermore, our framework is applicable to any age range from some starting age to the age at which everybody has died. Thus, the age range can be chosen such that it suits best the question at hand. We show that different scenarios might prevail for different age ranges and that our framework can identify this. For instance, sometimes scenarios can be observed where more and more deaths get shifted from younger to older ages, but where deaths become more and more evenly spread within the older ages. Such a scenario might be thought of as compression on the age range starting at 0, but quite the opposite on the age range starting at 60 (see Section 2). We also provide a possible methodology for the implementation of our framework and show its practical applicability in an example.

The remainder of this paper is organized as follows: Section 2 illustrates different issues in identifying mortality scenarios and sets out the requirements for a new classification framework. We establish a framework that allows for a unique classification of mortality scenarios and a clear definition of the different patterns in Section 3. The framework is based on changes in the shape of the deaths curve over time. Section 4 discusses the implementation of the framework. In particular, statistics for measuring relevant changes in the shape of the deaths curve are discussed, and a method for detecting trend changes is introduced. For illustrative purposes, we then present an application to the mortality evolution of US females in Section 5. Finally, Section 6 concludes.

2. Typical Issues with Scenario Definitions and Statistics

In this section, we identify and discuss some shortcomings of existing approaches for the classification of mortality scenarios. These shortcomings motivate a need for a new classification framework that will be developed in Section 3.

2.1. Imprecise Mortality Scenario Definitions

Mortality scenarios describe patterns in the evolution of mortality over time, i.e. a process of change. However, in the literature we find several imprecise mortality scenario definitions. One example is the attempt to define the process of change solely by some (only theoretical and hence unreachable) final stage. This is the case when rectangularization is defined as a process where the survival curve approaches a rectangular form. However, a rectangular form can be reached "on different routes".

This is illustrated by the left panel of Fig. 1 which shows a hypothetical, but not unrealistic evolution of deaths curves $d(x)$ over time.[1] Assume that at some point in time, mortality in some population follows the curve labeled "State 1". At some later point in time, it follows "State 2", etc., until it reaches "State 5". Without using any formal definition, one would intuitively conclude that some scenario of compression takes place between States 3 and 5. Between States 1 and 3, however, a scenario which is somewhat the opposite of compression can be observed.

[1] All deaths curves in this paper are scaled such that the areas underneath the curves each integrate to 1. Thus, the corresponding survival curves start with a radix of 1. Also note that all examples in Sections 2 and 3 are based on hypothetical, illustrative curves which are, however, reasonable since overall mortality improves and life expectancy increases.

If, however, one looks at the corresponding survival curves $l(x)$ (right panel of Fig. 1), one might intuitively conclude that with every step the shape becomes more rectangular. Therefore, one might identify the whole transition from State 1 to State 5 as rectangularization which is sometimes seen as a special case of compression. This clearly contradicts the observation that between States 1 and 3 the opposite of compression prevails.

So, one can conclude that the definition of a mortality scenario by some theoretical final state that is being approached will not always lead to a correct result.

2.2. Misleading or Insufficient Statistics

Of course, a reduction of complexity by looking at some key statistics of deaths or survival curves rather than at the whole curves is desirable. On the other hand this always leads to a loss of information. Therefore, one should very carefully identify statistics that preserve that part of the information one is interested in. Unfortunately, for some statistics that are frequently being used to describe patterns of mortality changes, this is not the case (at least if they are not analyzed together with additional statistics). In this subsection, we will explain this point.

Getting back to Fig. 1, we can observe that the modal age at death M increases from state to state starting with 83 years in State 1 and reaching 104 years in State 5. At the same time, $SD(M+)$ decreases from state to state starting at 7.62 in State 1 and ending at 2.71 in State 5. Following, for example, Robine et al. (2008), this would mean that compression prevails throughout the process and in particular also between States 1 and 3, which is inconsistent with the intuition from looking at the left panel of Fig. 1.

Sometimes, different statistics that were designed to measure the same thing can lead to contradicting results. For example, compression is often defined by a reduction of the IQR and/or a $C\alpha$-statistic (see Wilmoth and Horiuchi 1999; Kannisto 2000). Figure 2 shows two scenarios of mortality evolution where the structures of the mortality distributions have changed considerably from State 1 to State 2, with clear characteristics of mortality improvement and compression.

However, in the left panel the IQR remains unchanged, whereas in the right panel, $C50$ remains at the same value. Thus, neither IQR nor $C50$ alone are always able to identify compression.

Such issues can always occur when statistics that only take into account parts or certain points of the deaths curve are used to identify changes of the entire deaths curve.

2.3. Ignoring Mixed Scenarios

Next, we show that it may not be appropriate to define a certain mortality scenario as the opposite of some other scenario, or more generally that mixed scenarios should be allowed for and hence more than one "dimension" is required to get a full picture of a mortality scenario.

A classical example is the relationship between shifting mortality[2] (or alternatively extension) and compression. The left panel of Fig. 3 shows a mixed scenario where (in the transition from State 1 to State 2) shifting mortality and compression seem to coexist. Therefore, identification of one scenario should not rule out the other. Analogously, in the right panel of Fig. 3 neither shifting mortality nor compression can be observed. Thus, rejection of one scenario does not imply that the other scenario prevails. So clearly it is not suitable to consider compression and shifting mortality as disjoint categories. This again shows the need for a more sophisticated classification system which combines different concepts of compression, shifting mortality, etc. in the form of mixed scenarios.

2.4. Impact of Age Range

Sometimes, different types of mortality evolution occur in different age ranges. Myers and Manton (1984) compare the survival curve starting at age 0 with the survival curve starting at age 65 for US females and males between 1962 and 1979. They observe a clear tendency toward rectangularization for the entire age range but not in the older ages. If one is interested primarily

[2] In the introduction we have pointed out that the terms expansion, extension, and shifting mortality co-exist in the literature. We consider expansion and extension to be the same and use the term extension for that. We consider shifting mortality to be a different phenomenon (see Section 3).

in a certain age range (e.g. old age mortality) one should therefore only consider the corresponding part of the mortality curve.

However, when restricting the age range, undesired effects may occur whenever statistics are being used which depend on the number of people being alive at the beginning of the considered age range, for example $d(M)$, i.e. the number of deaths at age M. Assume one is interested in the age range starting at age 65. If between two points in time younger age mortality decreased, then more people would reach age 65. Even if older age mortality did not change at all, $d(M)$ would increase (with M remaining unchanged), suggesting a change in old age mortality. And if a change in old age mortality actually occurred, the change in $d(M)$ would be affected by both, the change in old age mortality that one is interested in and a change in younger age mortality that one is not interested in. These undesired effects can be eliminated by "normalizing" the population sizes such that at all considered points in time the number of people alive at the beginning of the considered age range is the same (e.g. $l(65) = 1$).

The left panel of Fig. 4 shows some mortality evolution over the entire age range. Here, clearly compression towards higher ages can be observed. If one is only interested in the age range 65+, one might intuitively look at the respective age range of the left panel of Fig. 4 (i.e. without normalizing) which displays signs of compression. However, in the normalized curves (right panel of Fig. 4) the deaths curve of State 2 looks less dense than for State 1 which is an indication against compression.

3. A New Classification Framework for Mortality Scenarios

In the previous section, we have identified shortcomings of existing approaches for the classification of mortality evolutions. We will now propose a new framework where unique mortality scenarios are defined based on observable changes in the shape of the deaths curve. Note that in this section, we introduce the "intellectual concept" of the framework whereas in the next section, we describe a methodology that can be applied to estimate the statistics used in our framework and to identify trends and trend changes in these statistics.

Our framework combines and uniquely defines four concepts for the change of mortality over time which are well known from the literature: the concept of shifting mortality (see e.g. Canudas-Romo 2008), the concept of longevity extension (see e.g. Rossi et al. 2013), the concept of compression of mortality (see e.g. Myers and Manton 1984), and the concept of concentration of mortality (see e.g. Kannisto 2001). As we will show, only a combined look at all four dimensions (which automatically allows us to consider both pure and mixed scenarios) gives a full picture of the considered mortality evolutions.

Our classification framework can be applied for any age range which includes the right tail of the deaths curve. Depending on the question at hand, the age range could start e.g. at zero, some juvenile age, or the retirement age. In particular, it is possible that the classification framework identifies different mortality scenarios for different age ranges (see Section 5 for an example).

For any given age range, we will use four key characteristics of the deaths curve each corresponding to one of the concepts mentioned above. Significant changes in one or several characteristics over time mean that the deaths curve has changed. Conversely, if these four characteristics remain unaltered, changes in a deaths curve are regarded as immaterial. We will show that these four characteristics are sufficient to distinguish between a great variety of different deaths curves and to uniquely classify mortality scenarios. The four characteristics are:

The position of a deaths curve's peak is measured by the modal age at death M and describes general shifts in the distribution of deaths. Since the shape of a deaths curve typically changes over time, a pure shift of the entire deaths curve will rarely occur, and therefore we consider its "center" M as a reference point. An increase in M indicates *right shifting mortality,* while a decrease in M implies *left shifting mortality*. Note that in this section, we assume that the modal age at death can be determined uniquely.[3]

The support of a deaths curve is determined by its upper bound, which we refer to as UB.[4] We denote the respective changes of UB as *extension* (if UB increases over time) and *contraction* (if UB decreases over time). Estimating UB in practice involves some ambiguity (see Section 4 for more details).

The degree of inequality in the distribution of deaths, which we denote by DoI, is the least obvious of the four key characteristics. However, Fig. 5 shows two deaths curves which are significantly different although the other three statistics of our framework coincide. Therefore, an additional statistic is required which is related to the shape of the curve. The deaths curve of State 2 is almost zero up to age 50, while State 1 shows a somewhat more balanced distribution of deaths over all ages. DoI is designed to pick up such differences by measuring the equality/inequality of the distribution of deaths over the whole age range. Intuitively, a low value of DoI indicates that deaths are rather equally distributed over the whole considered age range and vice versa. We use the terms *compression*/*decompression* if DoI increases/decreases and refer to Section 4 for more details.

Finally, the height of the peak of a deaths curve is given by $d(M)$. This component addresses the evolution of a deaths curve at and close to its "center" M. An increase in $d(M)$ is referred to as *concentration* and indicates that the distribution of deaths becomes more concentrated around M. The counterpart to *concentration* is what we refer to as *diffusion*, and it is observed if $d(M)$ decreases. Note that similar to DoI, $d(M)$ can also be seen as an indicator for the

[3] Only in rather theoretical scenarios, the peak might not be unique, e.g. because of multiple peaks of the same height or a plateau. In such a case, one might use a suitable alternative to M or modify the framework to include additional statistics.

[4] In theory, UB can only exist if the probability of death reaches one for some age. If the probability of death remains below one for all ages, any age could be reached in principle. Research by several authors, see e.g. Gampe (2010), indicates that probabilities of death typically flatten out at very old ages, possibly somewhere around 0.5. Thus, the population surviving up to such ages would get halved every year, but if the initial population was large enough, there would be a few survivors up to any age. Therefore, one could argue that UB does not exist in theory, which is, however, irrelevant for our application.

equality/inequality of the distribution of deaths: A large $d(M)$ implies that many deaths are concentrated at and around M. However, $d(M)$ is a more local measure for a small region around M, whereas DoI measures the equality/inequality of the distribution of deaths over the whole age range.

Of course each of the four components mentioned above can remain unchanged over time. In this case, the respective component is referred to as *neutral*. Thus, every component can attain three states.[5]

Two of the four statistics explained above (UB and M) primarily determine the "position" of the deaths curve, while the other two ($d(M)$ and DoI) primarily describe its shape. We believe that these four characteristics provide a good trade-off between granularity and complexity. The four components are summarized in Table 1. In principle, any combination of the three different states for each component is possible. This implies that we can classify both pure and mixed scenarios, which was one of the requirements from Section 2. In a pure scenario, only one component of the "scenario vector" is different from neutral. For instance, the vector (*neutral, extension, neutral, neutral*) denotes a pure extension scenario. On the other hand, a vector like (*neutral, extension, compression, neutral*) describes a mixed scenario which contains elements of both extension and compression. In total, there are $3^4 = 81$ possible mortality scenarios which might seem unfeasible at first glance. However, many scenarios will hardly be observed in practice, e.g. (*left shifting mortality, extension, compression, diffusion*). Those scenarios are nevertheless part of our classification framework to make sure that there are no unclassifiable evolutions and that classifications are unique.

[5] If a distinction between different intensities of increase or decrease is desired, more than three states can be considered or additional information about the slope of the respective trend line (see Section 4) can be added.

4. Methodology for the Implementation of the Classification Framework

The application of the classification framework introduced in Section 3 involves two main steps: First, the four statistics need to be estimated from deaths curves for each year in the observation period. A reasonable estimator for each of the statistics is proposed in the following subsection. Thereafter, trends in the resulting time series need to be analyzed in order to determine the prevailing states in each of the four scenario components. This is addressed in Subsection 4.2. Obviously, various different estimators and methods could be used in both steps, and thus the specific estimators and methods described in this section are only one possible implementation.

4.1. Estimation of Statistics

We now explain how we calculate the four statistics from the deaths curve in any given year. Both, raw or smoothed deaths curves can be used in principle. In our application in Section 5 we explain why we prefer using smoothed data.

For the *position of a deaths curve's peak* measured by M, we use the following estimator by Kannisto (2001):

$$M = x_{d\_max} + \frac{d(x_{d\_max}) - d(x_{d\_max} - 1)}{\left(d(x_{d\_max}) - d(x_{d\_max} - 1)\right) + \left(d(x_{d\_max}) - d(x_{d\_max} + 1)\right)},$$

where x_{d_max} is the age for which the largest number of deaths is observed. As a byproduct, the *height of a deaths curve's peak*, can then be estimated by the number of deaths at age x_{d_max}:

$$d(M) = d(x_{d\_max}).$$

For the *upper bound of* a *deaths curve* UB, we use the age at the α percentile of the distribution of deaths, x_α, plus an estimate for the remaining life expectancy at that age. Thus, the estimator for UB is

$$UB = x_\alpha + e_{x_\alpha}.$$

This approach builds on Rossi et al. (2013) who propose using the 90th percentile of the distribution of deaths as an approximation for the highest attainable ages. We prefer our combined estimator as it is considerably less biased. In our application (see Section 5) we set $\alpha = 99\%$. For the populations we have analyzed, this choice provides a reasonable compromise between only cutting off a small part of the distribution of deaths and stability in the statistic's evolution over time. For smaller (sub-)populations, however, smaller values for α might be more appropriate.

The statistic measuring the *degree of inequality* DoI in the distribution of deaths needs to take into account the whole age range. Therefore, statistics like $SD(M+)$, IQR, or $C\alpha$ which are commonly used to measure compression (see Subsection 2.2) are not feasible. An intuitive alternative is the area between the actual deaths curve and a hypothetical flat deaths curve $d_{flat}(x)$ as illustrated in Fig. 6. Using discrete data, this area can be approximated by adding up the absolute differences in the numbers of deaths between the two deaths curves. Thus, we estimate DoI as

$$DoI = c \cdot \sum_{x=x_0}^{\lfloor UB \rfloor} \left| d(x) - d_{flat}(x) \right| = c \cdot \sum_{x=x_0}^{\lfloor UB \rfloor} \left| d(x) - \frac{l_{x_0}}{(UB - x_0 + 1)} \right|,$$

where x_0 is the starting age of the deaths curve and $c = \frac{\lfloor UB \rfloor - x_0 + 1}{2 \cdot l_{x_0}(\lfloor UB \rfloor - x_0)}$ is a scaling factor such that DoI assumes its maximum value of 1 in case all people die at the same age. The minimum value of DoI is 0 in case deaths are uniformly distributed over all ages, i.e. $d(x) = d_{flat}(x)$ holds for all x.

Note that the dependence of DoI on UB is uncritical in our framework since we are only interested in changes of DoI over time. A potential misestimation/bias of UB would affect DoI in the same way for each point in time. Further, changes in UB over time do not automatically imply changes in DoI. For instance, if UB increases while the deaths curve's shape does not change materially, the slight changes of $d(x)$, $d_{flat}(x)$, and the scaling factor c would basically cancel each other.

As mentioned above, alternative estimators could be used for the four statistics. In particular, there is an extensive literature on measuring UB which is sometimes referred to as "maximum lifespan" (see e.g. Finch and Pike 1996) or "finite lifespan" (see Fries 1980). Alternative estimators for UB can, amongst others, be found in Cheung and Robine (2007), Fries (1980), or Wilmoth (1997). As alternative measures for DoI, one could consider the variance in the number of deaths, the Gini-Index as proposed by Debón et al. (2011), or the entropy as originally proposed by Demetrius (1974) and adopted by Keyfitz (1985) and Wilmoth and Horiuchi (1999). These statistics also consider the whole age range as required. However, the Gini-Index and the entropy are defined on the survival curve which makes them less intuitive in our deaths curve based framework.

4.2. Determination of Prevailing States

After estimating the four statistics for each year in the observation period (an example of the resulting time series is shown in Fig. 7), the trends prevailing at each point in time need to be determined. We will now introduce a possible methodology which we found to be suitable for all datasets we have analyzed. However, a different methodology or modifications of our methodology, e.g. with respect to the significance levels in the different tests could be used and might be advisable for certain applications.

Elimination of Outliers

Potential outliers should be eliminated as they are irrelevant with respect to long-term trends, but can significantly blur the trend analysis. Such outliers are typically caused by extreme events like the Spanish Flu. In order to detect whether a data point is an outlier, we fit a linear regression to the 10 adjacent data points. The sample variance of the residuals (assumed to be normally distributed) can then be used to derive a 99% prediction interval for the data point under consideration. If the data point lies outside the prediction interval, it is considered an outlier.

Determination of Trends, Trend Changes, and Jumps

In order to determine trends in the four statistics, we fit piecewise linear trends to the respective time series. Most of the time, mortality evolves rather steadily over time, and hence the piecewise linear trends should connect continuously. However, jumps can occur in case of extreme changes, e.g. the fall of the Soviet Union or wars like WW II, or changes in data processing methods. Thus, at every data point of a time series under consideration, the previous linear trend can persist, a new trend can commence starting at the end point of the previous trend (change in slope) or a new trend can commence at some other level (jump and change in slope). The following methodology first determines which of the three possibilities is the most likely one for each data point and then analyzes how many changes in slope and jumps are most suitable to describe the structural patterns in the entire time series and where they should occur.

In order to identify "candidate" data points for trend changes, i.e. changes in slope with or without jumps, we first perform a *preliminary analysis:* We carry out three fits for every possible combination of three data points:[6] a straight regression line to the data from the first to the third data point, a continuous regression line to the data from the first to the third data point with a change in slope at the second data point, and two straight regression lines (to the data from the first to the second and from the second to the third data point, respectively) that allow for a jump at the second data point. A set of Chow tests (see Chow 1960) is used to determine which trend evolution is most likely for the second data point, under the assumption that adjacent trend changes are located at the first and third data point or that these data points are the first or last data points of the entire time series. In the first Chow test (significance level of 1%), the null hypothesis of one persistent trend, i.e. no change at the second data point, is tested against a continuous change in slope. The result of the test (the new null hypothesis) is then tested versus a

[6] If the time series has *k* data points, this means that we consider al *k**(*k*-1)*(*k*-2)/6 possible triples.

jump in a second Chow test. Note that the results of the Chow tests usually depend on the choice for the first and third data point. Thus, whether a data point is a candidate for a trend change (and if so, of which kind) depends on the position of the neighboring trend changes.

After the preliminary analysis, we use the following main algorithm to identify the number and locations of trend changes that result in an optimal fit:[7]

- *Step 1:* We commence by fitting a straight regression line to the entire time series. This is the case of no trend change at all, i.e. the number of possible trend changes n is zero.
- *Step 2:* The number of possible trend changes is increased from n to $n+1$.
- *Step 3:* We determine the sample variances of the residuals from the fit with n trend changes. They will be required as variance estimators in Step 5. The sample variances are to be computed separately for each period with constant trend. We use a regime switch argument here to justify that the variance can change when the trend changes and thereby allow for heteroscedasticity as it can e.g. be observed in Fig. 7.
- *Step 4:* Building on the preliminary analysis, we determine all feasible combinations of $n+1$ candidate data points for trend changes. The preliminary analysis also indicates for each candidate data point whether the trend change would be a change in slope with or without jumps. If there is no feasible combination, the fit with n trend changes is the overall optimal fit and the algorithm terminates.
- *Step 5:* For each feasible combination of trend changes from Step 4 we fit a piecewise linear trend curve to the data (and allow for discontinuities only where the type of the potential trend change is a change with jump). In order to account for heteroscedasticity, we use the sample variances from Step 3 as weights.
- *Step 6:* The optimal trend change positions (and thus also the trend change types) for $n+1$ trend changes are determined by comparing the fits from Step 5 by the Akaike

[7] Note that the presentation of the algorithm aims for a clear presentation of and distinction between the steps involved and does not pay attention to computing efficiency.

Information Criterion (AIC; Akaike 1973). The number of parameters is two (initial intercept and slope) plus $n+1$ for the trend change positions plus $n+1$ for the changes in slope plus one for every jump (i.e. the new intercepts after the jumps).

- *Step 7:* Finally, we compare the optimal fits with n and with $n+1$ trend changes to assess the contribution of the additional trend change to the time series representation. To this end, we use another Chow test (again with significance level 1%). Since the original test by Chow only considers one trend change versus none, we use an extended version of the test: The test statistic remains unchanged, but the number of parameters increases (one for each trend change position, each intercept, and each slope). Note that for $n \geq 2$, we can account for heteroscedasticity also in this test by applying variance estimates from the optimal fit with $n-1$ trend changes as weights. The null hypothesis in the Chow test is the case of n trend changes. Thus, the additional trend change is only accepted if it significantly improves the fit, which is in line with our intention of determining long-term trends. If the null hypothesis stands, the time series can be adequately described by n trend changes, and the algorithm terminates. If the additional trend change significantly improves the fit, we return to Step 2.

Testing for Increasing, Decreasing, or Neutral Statistics

Finally, we have to determine if the resulting trend curve (see the lines in Fig. 7) should be considered increasing, decreasing, or neutral in the context of our framework. For each period with constant trend, we use an F-test with a significance level of 10% to analyze whether the slope of the trend is significantly different from zero. If the slope is not significantly different from zero, the state *neutral* is assigned. Otherwise, we consider the statistic as *increasing* (*decreasing*) if the slope is positive (negative) during the corresponding period of time. This definition implies that the state *neutral* is not only assigned if the slope is clearly close to zero but also if the uncertainty in the underlying data is too large to identify a significant trend.

5. Application of the Classification Framework

In this section, we apply our classification framework to the mortality evolution of females in the USA.[8] We derive log mortality rates $\ln(m(x,t))$ for ages 0 to 109 from the deaths and exposure data in the Human Mortality Database (HMD) for years 1933 to 2013. For each calendar year, these log mortality rates are then smoothed and extrapolated using P-splines. This allows us to derive normalized (see Subsection 2.4) and smoothed deaths curves. We prefer this approach over using the raw deaths curves from the HMD for several reasons: Potential disturbing effects resulting from birth cohorts of different sizes are eliminated; random effects in the data which might, e.g., lead to double peaks in the deaths curve are significantly reduced; the potential impact of age misspecifications in the raw data, in particular with respect to estimating UB, is reduced; and the time series for the four statistics exhibit less random fluctuations and are thus easier to analyze.

We consider deaths curves covering different age ranges as discussed in Subsection 2.4: The curves $d_{10}(x,t)$ start with a fixed radix at age 10 and thus exclude effects from infant mortality, whereas the $d_{60}(x,t)$ curves allow for an analysis of mortality at typical retirement ages. Figure 7 displays the four components of our classification framework for both starting ages along with the respective piecewise linear trend lines.[9]

By definition, the curves for the modal age at death M coincide for both starting ages. From a theoretical perspective, the same holds for UB. However, the chosen estimator yields slight differences for the different starting ages. Since the two sets of data points would be difficult to distinguish and the resulting scenarios for this component are the same for both starting ages, we only display UB for starting age 10.

---

[8] We have also applied the framework to several other populations, e.g. Sweden, Japan, and West Germany. In all cases, the framework yielded reasonable and informative results. For the sake of brevity, however, we only show the results for one population. We chose US females for illustration because the variety of different observed scenarios was the largest. We refer to Genz (2017) for an application of our framework to a larger number of countries and a comparison of the respective mortality patterns.

[9] We have also considered the starting ages 0, i.e. the complete age range, and 30 in order to exclude effects of young adult's mortality like accidents, etc. It turned out that the observed scenarios for starting ages 0, 10, and 30 are quite similar.

From Fig. 7 we can see that our framework identifies several trend changes for each of the statistics and both starting ages. Such trend changes can either mean that the direction of the trend changes (e.g. from increasing to neutral or decreasing, etc.), or that only the intensity of the trend (i.e. the slope of the trend line) changes significantly while its direction remains unchanged. For example, the first two trend changes in M are changes in the direction of the trend, whereas the subsequent trend changes (except the last one) only concern the intensity of the increase (i.e. the pace of the right shift in mortality). Thus, a trend change does not inevitably lead to a change in the scenario vector. Moreover, as mentioned in Section 4, trends can change with or without a jump in the absolute level of the statistic. In our example, such jumps occur for all statistics except $d(M)_{60}$.

The direction of each trend as well as the position of the trend changes and their types (i.e. a change in slope with or without an upward/downward jump) are summarized in Fig. 8. This representation allows for an easy visual assessment of the scenario vector at each point in time. For instance, in the year 2010, for both starting ages, the scenario vector is (0, +, +, -), i.e. the scenario is neutral with respect to shifting mortality and exhibits extension, compression, and diffusion at the same time.

By comparing Fig. 7 to Fig. 8 we find some periods with seemingly increasing (decreasing) trends in Fig. 7, but a classification as "neutral" in Fig. 8. One such example is the first trend for UB. Here the underlying data has a relatively strong variance, and therefore the seemingly increasing trend is not significant (as explained at the very end of Subsection 4.2).

The results of our analysis particularly underline the need for combining different concepts of mortality change in one framework (see Subsection 2.3) as we observe mixed scenarios over almost the whole observation period. There are even periods where all four indicators change, e.g. between 1973 and 1982 for starting age 10. During this period we simultaneously observe right shifting mortality and extension (i.e. both the mode and the upper bound of the deaths curve move to the right) combined with a compression of the whole curve and an increase of the concentration around the mode. In contrast, pure scenarios seem to be very rare: Only for starting age 60 and years 1941 to 1948 we find a scenario of pure diffusion.

Furthermore, the results show that each of our four components is relevant in the sense that no component can be explained by the others. For instance, as one would expect, M and UB increase over the observation period in general, i.e. we observe right shifting mortality and extension. However, particularly for UB there are some periods (see for example the 1990s) where we observe the opposite trend, i.e. contraction, and thus these two statistics do not move in the same direction throughout the entire observation period. This also holds for DoI and $d(M)$ although they also frequently follow the same trend. For example, after 2006 $d(M)$ decreases for both starting ages, while DoI increases for both starting ages, i.e. we observe diffusion and compression at the same time.

The results also highlight the importance of choosing a suitable age range. For both, DoI and $d(M)$, we find several time periods where the trends differ by starting age. For instance, between 1975 and 1990 we observe compression for starting age 10, but decompression for starting age 60.

6. Conclusion

In this article we explain why many existing approaches to classify patterns of mortality evolution have four major shortcomings: Mortality scenario definitions are often imprecise and intuitive rather than rigorous; some frequently used statistics are not always sufficient to identify the respective scenarios; mixed scenarios are usually not accounted for; often, the impact of the considered age range is being ignored.

We propose a new framework for classifying patterns of mortality evolution. Our approach is based on changes of the deaths curve and uses four statistics that should be considered simultaneously. Each mortality scenario then consists of four components: (1) the deaths curve can exhibit a right shift or a left shift or be neutral in that respect; (2) the deaths curve can exhibit extension or contraction or be neutral in that respect; (3) the deaths curve can exhibit compression or decompression or be neutral in that respect; (4) the deaths curve can exhibit concentration or diffusion or be neutral in that respect. This approach overcomes the shortcomings of previous approaches: Each mortality evolution is uniquely and precisely classified; by considering all four components simultaneously, mixed scenarios are automatically detected; the framework is applicable to different age ranges.

For some of the statistics used, the estimation is not straightforward. Beyond an introduction of the intellectual concept of the framework, we therefore also introduce a methodology that can be used to estimate the statistics and determine trends and trend changes in the data. Also, we apply our approach to data for US females, illustrating that the structure of the change in mortality can be quickly assessed and well understood. We further demonstrate empirically that none of the four components can be explained by the other three and that results can significantly differ for different age ranges.

Note that the purpose of our framework is a classification of realized mortality evolutions. In this sense it is purely descriptive, i.e. it does not provide explanations for observed trends and trend changes. It seems obvious that any research that intends to provide such explanations or seeks to explore a link between determinants of mortality and observed patterns of mortality change needs as a prerequisite a common understanding which pattern of mortality change has been observed in which situation. Our methodology can provide this and hence serves as a basis for such research. In particular, the detected trend changes can be an indication when and how demographic changes have occurred. Similarly, by applying our framework to different populations, time and structure of differences in their demographic evolutions can be detected, which again can serve as a basis for research on the causes.

If a mortality model is to be calibrated to historical data, our framework can also be used to identify suitable time spans (e.g. without major trend breaks). Further, the framework can be applied for testing whether existing mortality projections are consistent with observed trends in the most recent history.

References

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. *2nd International Symposium of Information Theory*, B.N. Petrov and F. Csaki (eds.), Akademiai Kiado, Budapest, 267-281.

Bongaarts, J. (2005). Long-Range Trends in Adult Mortality: Models and Projection Methods. *Demography , 42* (1), 23-49.

Canudas-Romo, V. (2008). The Modal Age at Death and the Shifting Mortality Hypothesis. *Demographic Research , 19* (30), 1179-1204.

Cheung, S. L., Robine, J.-M., Tu, E. J.-C., & Caselli, G. (2005). Three Dimensions of the Survival Curve: Horizontalization, Verticalization, and Longevity Extension. *Demography , 42* (2), 243-258.

Cheung, S. L., & Robine, J.-M. (2007). Increase in Common Longevity and the Compression of Mortality: The Case of Japan. *Population Studies , 61* (1), 85-97.

Chow, G.C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica, 28(*3), 591-605.

Debón, A., Martínez-Ruiz, F., & Montes, F. (2011). *Temporal evolution of some mortality indicators. Application to Spanish data.* Society of Actuaries. Life Monographs. https://soa.org/library/monographs/life/living-to-100/2011/mono-li11-1b-debon.pdf. Accessed 16 January 2015.

Demetrius, L. (1974). Demographic Parameters and Natural Selection. *Proceedings of the National Academy of Sciences of the United States of America, 12*, 4645-4647

Finch, C. E., & Pike, M. C. (1996). Maximum Life Span Predictions From the Gompertz Mortality Model. *Journal of Gerontology: Biological Sciences , 51A* (3), B183-B194.

Fries, J. F. (1980). Aging, Natural Death, and the Compression of Morbidity. *The New England Journal of Medicine , 303* (3), 130-135.

Gampe, J. (2010). Human Mortality Beyond Age 110. In H. Maier, J. Gampe, B. Jeune, J.-M. Robine, & J. W. Vaupel (Ed.), *Demographic Research Monographs – Supercentenarians,* Heidelberg: Springer Verlag, 219-230.

Genz, M. (2017). *A Comprehensive Analysis of the Patterns of Worldwide Mortality Evolution.* Society of Actuaries. Life Monographs. https://www.soa.org/Library/Monographs/Life/Living-To-100/2017/2017-living-100-monograph-genz-paper.pdf. Accessed 15 December 2017.

Human Mortality Database. (n.d.). *Univerity of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany).* www.mortality.org. Accessed 16 August 2014.

Kannisto, V. (2000). Measuring the Compression of Mortality. *Demographic Research,* doi: 10.4054/DemRes.2000.3.6

Kannisto, V. (2001). Mode and Dispersion of the Length of Life. *Population: An English Selection, 13* (1), 159-172.

Keyfitz, N. (1985). *Applied Mathematical Demography,* New York: Springer Verlag.

Manton, K. G., & Tolley, H. D. (1991). Rectangularization of the Survival Curve - Implications of an Ill-Posed Question. *Journal of Ageing and Health , 3* (2), 172-193.

Myers, G. C., & Manton, K. G. (1984). Compression of Mortality: Myth or Reality? *The Gerontologist , 24* (4), 346-353.

Nusselder, W. J., & Mackenbach, J. P. (1996). Rectangularization of the Survival Curve in the Netherlands, 1950-1992. *The Gerontologist , 36* (6), 773-782.

Ouellette, N., & Bourbeau, R. (2010). *Recent adult mortality trends in Canada, the United States and other low mortality countries.* Society of Actuaries. Life Monographs. https://soa.org/library/monographs/life/living-to-100/2011/mono-li11-4b-ouellette.pdf. Accessed 16 January 2015.

Robine, J.-M., Cheung, S. L. K., Horiuchi, S., & Thatcher, A. R. (2008). *Is There a Limit to the Compression of Mortality?* Society of Actuaries. Life Monographs. https://www.soa.org/essays-monographs/2008-living-to-100/mono-li08-03-cheung.pdf. Accessed 15 December 2017.

Rossi, I. A., Rousson, V., & Paccaud, F. (2013). The Contribution of Rectangularization to the Secular Increase of Life Expectancy: An Empirical Study. *International Journal of Epidemiology , 42*, 250-258.

Thatcher, A. R., Cheung, S. L., Horiuchi, S., & Robine, J.-M. (2010). The Compression of Deaths Above the Mode. *Demographic Research , 22* (17), 505-538.

Wilmoth, J. R. (1997). In Search of Limits. In K. W. Wachter, & C. E. Finch (Ed.), *Between Zeus and the Salmon - The Biodemography of Longevity* (pp. 38-64). Washington D.C.: The National Academies Press.

Wilmoth, J. R., & Horiuchi, S. (1999). Rectangularization Revisited: Variability of Age at Death within Human Populations. *Demography , 36* (4), 475-495.

Wilmoth, J. R. (2000). Demography of Longevity: Past, Present, and Future Trends. *Experimental Gerontology , 35*, 1111-1129.

Tables

Table 1 Scenario components, attainable states, and criteria

Scenario component	Attainable states	Criterion (in terms of deaths curve characteristic and statistic to be computed)
1	Right shifting mortality Left shifting mortality Neutral	Peak shifts to the right; M increases Peak shifts to the left; M decreases Peak does not move; M constant
2	Extension Contraction Neutral	Support is prolonged; UB increases Support shrinks; UB decreases Support remains unchanged; UB constant
3	Compression Decompression Neutral	Distribution of deaths less balanced; DoI increases Distribution of deaths more balanced; DoI decreases Distribution of deaths equally balanced, DoI constant
4	Concentration Diffusion Neutral	More deaths at/around M; $d(M)$ increases Less deaths at/around M; $d(M)$ decreases Number of deaths at/around M unchanged; $d(M)$ constant

Figures

Fig. 1 Mortality evolution in a hypothetical example. Left: deaths curves; right: survival curves

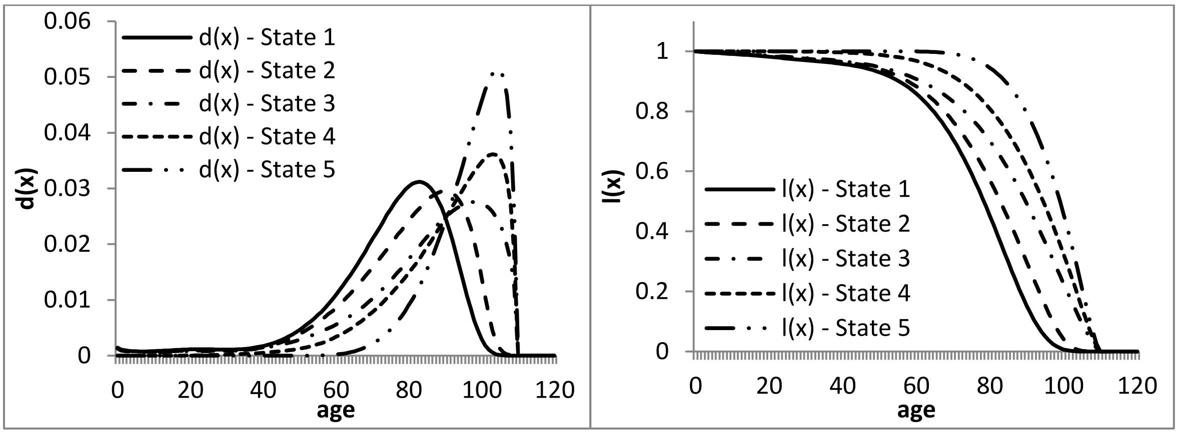

Fig. 2 Mortality evolutions with compression. Left: unchanged IQR; right: unchanged $C50$

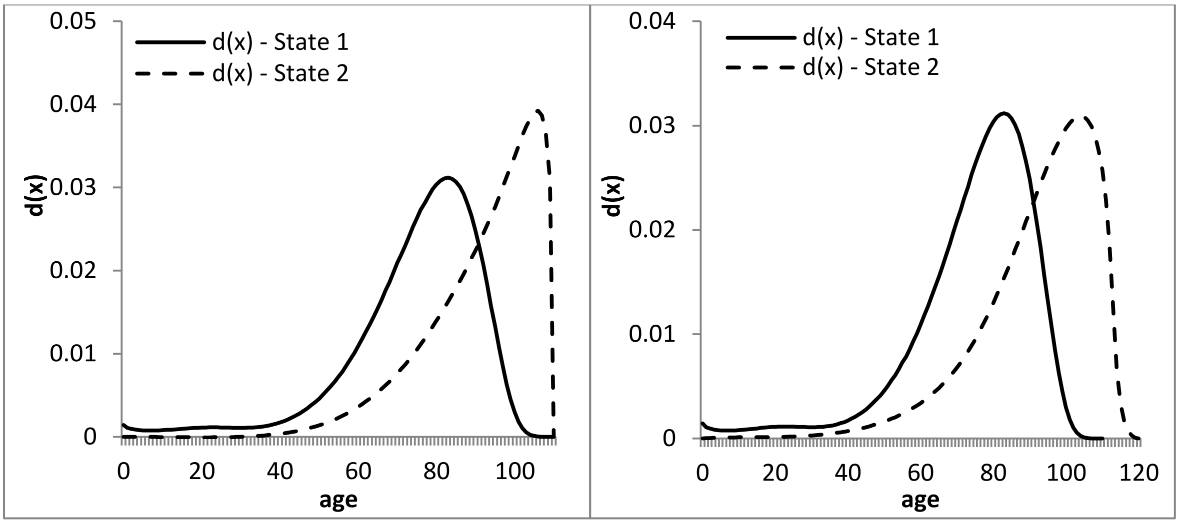

Fig. 3 Two hypothetical examples. Left: shifting mortality and compression coexist; right: neither shifting mortality nor compression exists

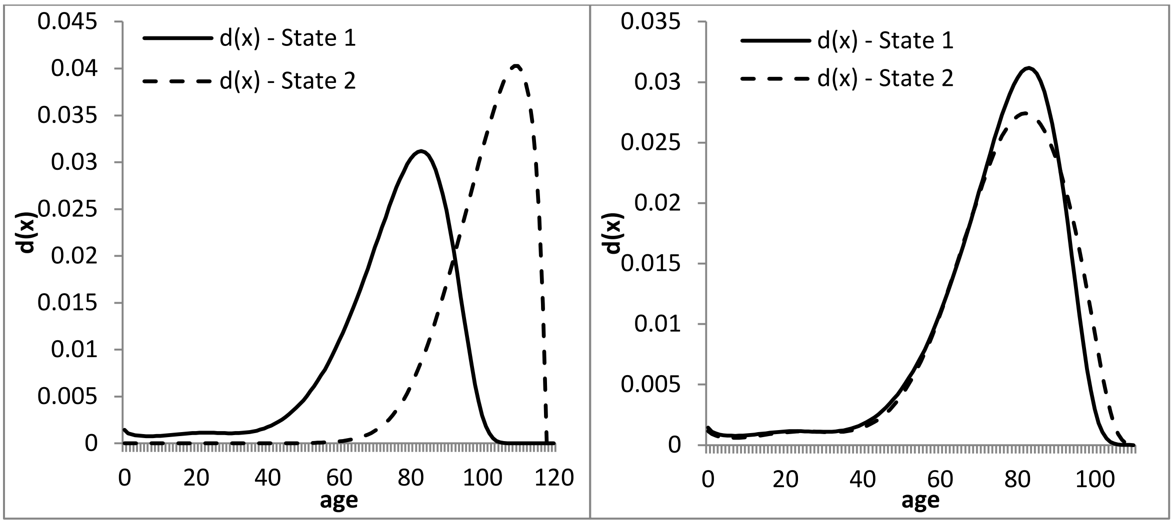

Fig. 4 Mortality evolution with increasing number of survivors to age 65. Left: complete age range; right: starting at age 65 with normalized l(65)

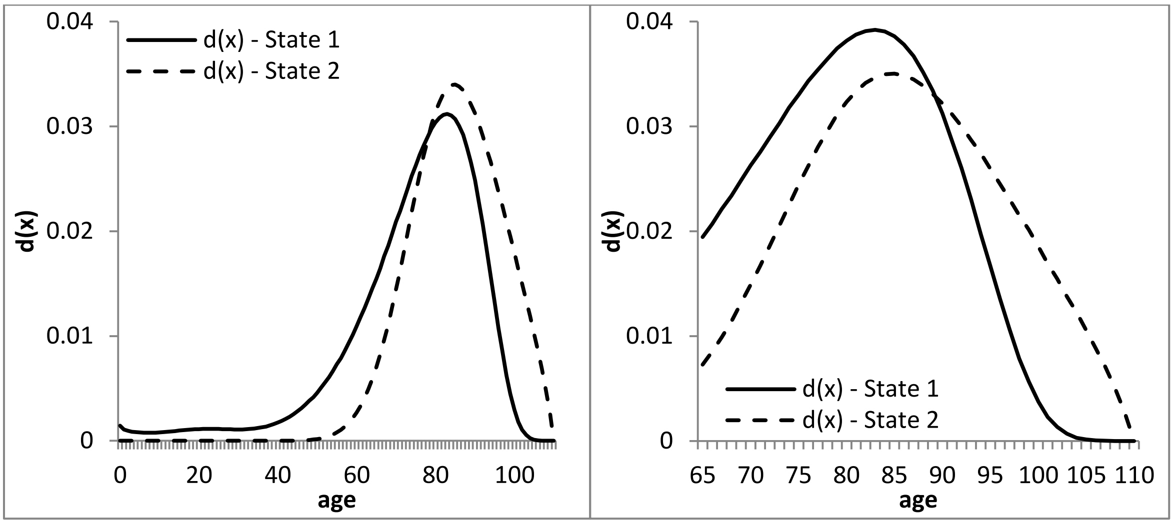

Fig. 5 Mortality evolution with constant M, $d(M)$, and UB, but changing DoI

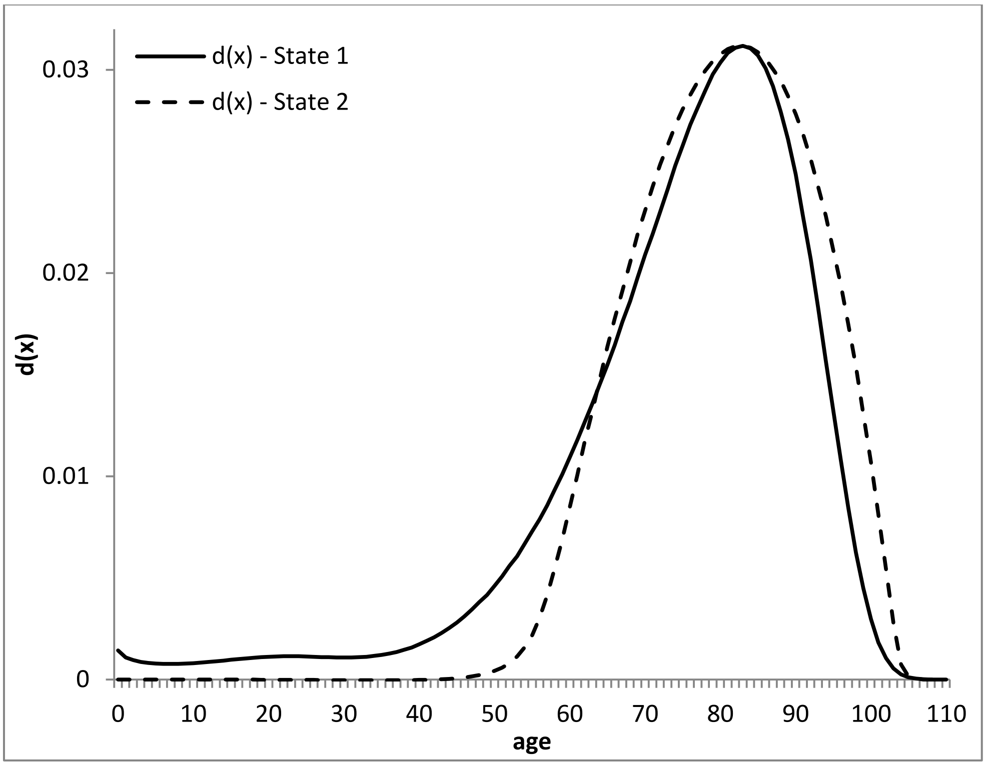

Fig. 6 DoI as the area between observed deaths curve $d(x)$ and hypothetical flat deaths curve $d_{flat}(x)$

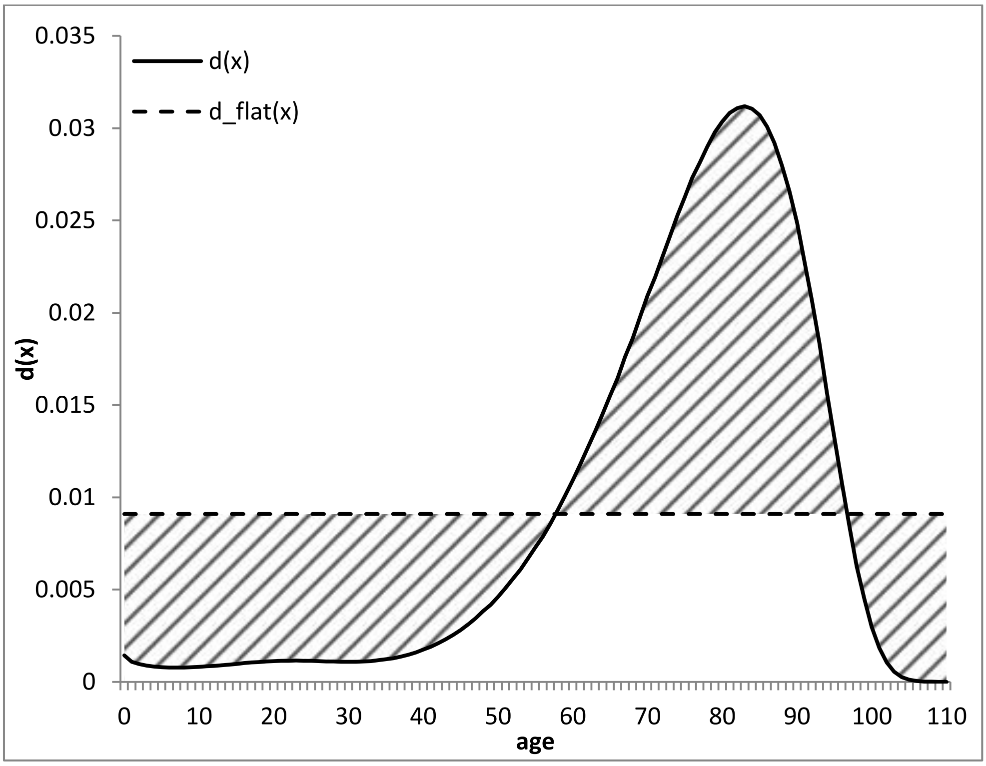

Fig. 7 Development of the four components of our classification framework for US females from 1947 to 2013. Upper left panel: M; upper right panel: UB; lower left panel: DoI; lower right panel: $d(M)$

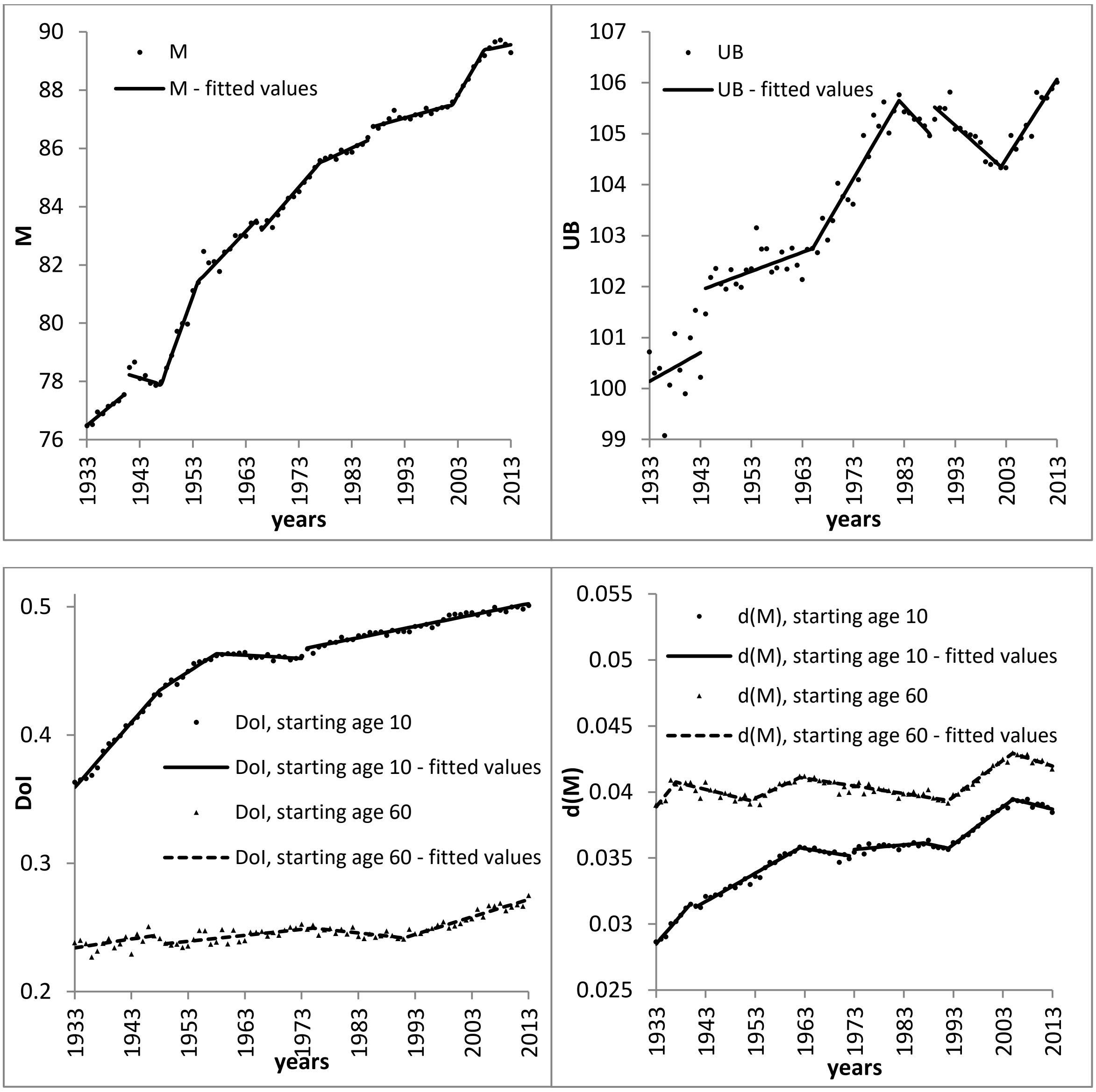

Fig. 8 Time bars of mortality evolution for US females, each statistic, and both starting ages