Solvency 2: Longevity Stress and the Danish Longevity Benchmark

Working Group, The Danish Society of Actuaries

September 18, 2012

Abstract. This paper proposes a partial internal model for longevity risk within the Solvency 2 framework. The model is closely linked to the mechanisms associated with the Danish longevity benchmark, where the underlying mortality intensity and the trend is estimated yearly based on mortality experience from the Danish life and pension insurance sector and on current data from the entire Danish population. Within this model, we derive a new estimate for the 99.5% quantile for longevity risk, which differs from the longevity stress of 20% from the standard model. The new stress explicitly reflects the risk associated with unexpected changes in the underlying population mortality intensity on a one year horizon and with a 99.5% confidence level. In addition, the model contains a component, which quantifies the non-systematic longevity risk. This last component depends on the size of the portfolio.

Contents

1 Excecutive summary

This paper, which was prepared by a working group[1] under the Danish Society of Actuaries during the Summer 2012, proposes a new model for quantifying the longevity risk within Solvency 2 for Danish life and pension insurance companies. The proposed model is closely linked to the so-called longevity benchmark that was introduced by the Danish FSA in 2010 and subsequently used by all Danish life and pension insurance companies since 2011. The new longevity stress contains three components. The first two components are closely related to the construction of the longevity benchmark and applies for any company which uses the Danish longevity benchmark. These components reflect the risk associated with changes in the underlying, current mortality and the risk associated with changes in the assumptions concerning the future improvement of the mortality intensity. The third and last component quantifies the company's non-systematic risk associated with the yearly analysis of the deviation from the benchmark index for the underlying current mortality intensity. In contrast to the first two parts, this last part is company-specific and depends on the size of the company's insurance portfolio.

1.1 Organization of the paper

The paper presents the methodological arguments and theoretical background for the model. It is the intention that the document can serve as part of the documentation in an application for the use of a (partial) internal model for longevity risk within Solvency 2. The paper is organized as follows. Section 2 contains an executive summary in Danish, written for Danish decision makers. Section 3 provides background information and gives an overview of the proposed model to quantity longevity risk within the Danish regulatory regime. A detailed technical description of the model can be found in Section 4 and Appendix A.

1.2 The Danish longevity benchmark and Solvency 2

In December 2010, the Danish financial supervisory authorities (The Danish FSA) introduced a so-called longevity benchmark, which should be used as best estimate by all Danish life and pensions insurance companies for market based valuation of life insurance liabilities. This index is composed of a current mortality intensity and a set of assumptions concerning the expected future mortality improvement. The current mortality intensity is estimated from 5 years of data for a portfolio of insured lives from several large Danish life insurance companies, and hence this mortality intensity will in general differ from the mortality intensity estimated from the total Danish population. We shall also refer to these mortality intensities as the *sector mortality intensity* for the population of insured lives, and the *population mortality intensity* for the average mortality intensity for the total Danish population.

The life and pension insurance companies are now obliged to perform a yearly estimation procedure prescribed by the Danish FSA in order to determine whether the companies' mortality experiences deviate significantly from the underlying sector mortality. This estimation is performed within a Poisson regression model, where the underlying

[1]The working group consisted of Kristoffer Andre Bork (chairman), Ivan Toftegaard Carlsen, Gregers Østervig Frank, Kristian Hasløv, Søren Fiig Jarner, Lars Sommer Hansen, Aage Møller Holst, Nichlas Abel Korsgaard, Liselotte Milting, Thomas Møller and Merete Lykke Rasmussen.

sector mortality is fixed. More precisely, the model specifies a parametrization for the company's mortality intensity in terms of the underlying sector mortality via a set of fixed regressors. If the company's current mortality intensity deviates significantly from the benchmark, the company must use an adjusted intensity for the current mortality, where the adjustment is estimated in the regression model. If the difference is not significant, the benchmark intensity is used as best estimate for the current mortality. With this Danish benchmark model, the life and pension companies' current mortality assumptions are adapted to the observed sector mortality intensity via the yearly estimation and test procedures. In particular, this ensures that the mortality assumptions for the current mortality are based on the most recent data.

The assumptions about the expected future mortality improvements in the benchmark are estimated from mortality data for the entire Danish population during the last 30 years. This leads to age- and gender-dependent improvement factors that reflect the average improvements during the last 30 years in the Danish population. The benchmark assumptions for the mortality improvements are updated on a yearly basis. The companies are not required to perform an analysis of the trend in their own population since it is assumed that the portfolio of a single company will in general be too small to establish a significant difference from the benchmark trend. Therefore, the Danish life and pension companies typically use the country specific trend when determining the best estimate for the life insurance liabilities. In particular, this ensures that the companies use assumptions concerning future mortality improvements that are similar to the observed historical improvements.

The Solvency 2 standard model prescribes a pragmatic longevity stress of 20% for all ages and all time points. Since the mortality assumptions for Danish life and pension insurance companies are revised yearly by including recent mortality experience and by following a relatively advanced model, it seems reasonable to propose a more refined longevity stress within this model which is more consistent with a 99.5% confidence level on a one year time horizon. Therefore, the Danish Society of Actuaries has established a working group in order to propose a new model for the longevity stress, which measures the systematic risk associated with the yearly updating procedure for the longevity benchmark and the non-systematic risk associated with the analysis performed by the individual companies.

1.3 The proposed longevity stress

The proposed longevity stress consists of three parts: A stress for the current mortality, a stress for the trend (the systematic mortality risk), and a stress for the non-systematic risk. Whereas the two first components are directly applicable for any Danish life and pension insurance company, the third stress is company-specific and depends crucially on the size of the insurance portfolio, or, more specifically, on the expected number of deaths in the insurance portfolio. Table 1 shows these three components which can be used together with the Danish FSA Benchmark model.

Hence the market value of life insurance liabilities can be determined with the proposed longevity stress by first reducing the current mortality intensity by 6%. This age-independent stress may be interpreted as a way of modeling the uncertainty associated with the yearly updating procedure for the underlying benchmark of the current mortality (the sector mortality). Similarly, the updating procedure for the assumptions concerning the future mortality improvements may lead to changes from year to year.

Stress, current mortality	6%
Stress, future improvements (trend)	6%
Stress, non-systematic	$2.6/\sqrt{5H}$

Table 1: Combined 99.5% longevity stress for the current underlying mortality, the trend and the non-systematic risk. The parameter H is the expected number of deaths in the insurance portfolio during a period of five years.

This risk is quantified by increasing the improvement rates by 6% for all ages. As mentioned above, these two corrections, which may be interpreted as a model for the systematic risk associated with changes in the longevity benchmark, can be used by all life and pension insurance companies and do not depend on the size of the companies' insurance portfolios. The third and last component of the proposed longevity stress deals with the non-systematic risk and depends on the expected number of deaths in the company's insurance portfolio. In Table 1 this dependence is parameterized via the quantity H, which is the expected number of deaths in the portfolio during a period of five years.

Expected number of deaths, one year	1	10	100	1,000	10,000
Expected number of deaths, 5 years (H)	5	50	500	5,000	50,000
Non-systematic stress	52.0%	16.4%	5.2%	1.6%	0.5%

Table 2: The non-systematic stress for various portfolio sizes. The size of the portfolio is measured by the expected number of deaths, calculated by using the Danish FSA longevity benchmark.

Table 2 contains examples of this non-systematic longevity stress for various portfolio sizes. For a portfolio with 1,000 expected deaths each year, i.e. H=5,000, the stress is 1.6%. However, the stress can be considerably larger for small portfolios. Hence, the stress associated with the non-systematic risk is linked to the size of the data which is used for the company's statistical analysis. This non-systematic risk is calculated for the portfolio as a whole and for both genders combined.

The two stress components for the current mortality and the trend may be interpreted as the 99.5% quantile in the distribution for the remaining life time associated with the yearly updating procedure for the longevity benchmark. Similarly, the non-systematic stress reflects the 99.5% quantile in the deviation of the company's own mortality intensity from the benchmark. Hence, the model quantifies the non-systematic risk more precisely by taking into account both the size of the portfolio and the underlying mechanism, where the mortality assumptions are updated yearly. The model is not able to quantify the risk in a portfolio where the expected number of deaths is very small, e.g. if only one death is assumed in a period of five years; in this situation, the non-systematic stress presented above will exceed 100%. For such portfolios, we propose to apply alternative methods and introduce a priori information about the portfolio, for example by comparing to the mortality experience from similar portfolios.

Table 3 presents the best estimate for the remaining life times at various ages, calculated by using the longevity benchmark. In addition, the table contains the increase in the expected life time associated with the stress of 20% from the Solvency 2 standard model. For a person aged 60 years, we see that the stress of the standard model leads to an increase of 1.7 years for males and 1.8 years for females. The last column of the table contains the increase in the expected life time which corresponds to the new systematic stress, where the current mortality is reduced by 6%, and where the future mortality improvements are increased by 6%. This can be interpreted as the stress for a portfolio, which is sufficiently large to eliminate the non-systematic risk. For a 60-year old, this stress leads to an increase of 0.5–0.6 years, which is hence the lower bound for the increase in the expected remaining life time in this model for a 60-year old.

		Remaining life time	Increase of life time			
Gender	Age	Benchmark	20%	H=500	H=5,000	Systematic
Males	20	66.3	1.7	1.1	0.8	0.7
	40	45.1	1.8	1.1	0.8	0.7
	60	24.7	1.7	0.9	0.7	0.5
	80	8.5	1.1	0.6	0.4	0.3
Females	20	68.6	1.8	1.2	0.9	0.7
	40	47.5	1.9	1.1	0.8	0.7
	60	27.4	1.8	1.0	0.7	0.6
	80	10.5	1.3	0.7	0.5	0.4

Table 3: Expected remaining life time for males and females at various ages with the Danish longevity benchmark, and the increase with various stresses: The standard model, the systematic stress of 6% for the current mortality and trend, and the combined systematic and non-systematic stress with number of expected deaths H equal to 500 and 5,000, respectively.

Finally, the table contains the increase in the expected remaining life time which arises when we include the components for both the systematic and the unsystematic risk for a portfolio with 500 and 5,000 expected deaths during a period of five years. For the small portfolio with 500 expected deaths, the new stress leads to a total increase of 0.9 years for 60-year old males and 1.0 years for females, whereas the increase is 0.7 years for males and females aged 60 in a portfolio with 5,000 expected deaths during a period of five years.

One can alternatively compare the proposed stress with the stress in the standard model by translating the proposed stress into an age- and time-dependent stress in the underlying mortality intensity. This is done in Tables 4 and 5 below. In Table 4, we have listed the company-independent stress for the systematic risks, whereas Table 5 contains the total stress for a 60-year old, where the non-systematic risk is included for portfolios with various values for the number of expected deaths H.

It follows from Table 5 that the stress for a 60-year old is approximately 11–12% for a portfolio of size $H = 500$. This stress is obtained by increasing the stress from Table 4 by $2.6/\sqrt{5 \cdot 500}$. For a portfolio of size $H = 50$, the corresponding stress is 23–24%. This can be compared to the Solvency 2 standard model, where the current and future mortality are reduced by 20% for all times and all ages.

Time	Age (males)				Age (females)			
	20	40	60	80	20	40	60	80
0	6%	6%	6%	6%	6%	6%	6%	6%
10	8%	7%	7%	6%	8%	7%	6%	6%
20	8%	8%	7%	6%	9%	8%	7%	6%
30	8%	9%	7%	6%	9%	7%	7%	6%
40	11%	9%	6%		10%	8%	6%	
50	11%	7%	6%		8%	8%	6%	
60	10%	6%			9%	6%		

Table 4: Age- and time-dependent stress in percent for the systematic risk at various start ages for males and females.

Time	Deaths H (males)				Deaths H (females)			
	50	500	5,000	50,000	50	500	5,000	50,000
0	23%	11%	8%	7%	23%	11%	8%	7%
10	23%	12%	8%	7%	23%	12%	8%	7%
20	23%	12%	9%	7%	23%	12%	9%	7%
30	24%	12%	9%	8%	24%	12%	9%	8%
40	23%	11%	8%	7%	23%	11%	8%	7%
50	22%	11%	8%	7%	22%	11%	8%	7%

Table 5: Total stress in percent for the systematic and non-systematic risk at age 60 with varying H for males and females.

2 Executive summary in Danish

Finanstilsynet introducerede i december 2010 et såkaldt levetidsbenchmark, der skal anvendes til fastlæggelse af bedste skøn for dødeligheden i den danske pensionsbranche. Levetidsbenchmarket består af to dele: Et benchmark for den nuværende observerede dødelighed i form af en årlig dødsintensitet (svarende til sandsynligheden for at dø inden for et år) for hver alder og for hvert køn, samt et benchmark for den forventede fremtidige levetidsforbedring (den nedadgående trend i den årlige dødsintensitet), ligeledes for hver alder og køn.

Det enkelte pensionsselskab skal gennemføre en af Finanstilsynet fastlagt statistisk analyse af, om den observerede dødsintensitet i selskabet afviger fra benchmarket for den nuværende observerede dødelighed. Hvis der er en signifikant afvigelse, skal selskabet tage højde for denne afvigelse ved fastsættelse af bedste skøn. Hvis der derimod ikke kan påvises en signifikant afvigelse, skal Finanstilsynets benchmark benyttes. Det er antagelsen, at det enkelte selskab ikke har en tilstrækkelig stor kundebestand, som tillader, at selskabet kan estimere afvigelser i forhold til benchmarket for den årlige trend. Derfor er der ikke krav om, at der skal gennemføres en tilsvarende analyse i forhold til trenden. Finanstilsynet opdaterer levetidsbenchmarket årligt, ligesom de danske selskaber skal foretage den statistiske analyse årligt og på baggrund heraf anmelde nye dødelighedsforudsætninger.

Da dødelighedsforudsætningerne i de danske selskaber således fastsættes efter en avanceret model og opdateres årligt, synes Solvens 2's stress på 20% for levetidsrisiko i standardmodellen umiddelbart for simpel i sin tilgang. Med udgangspunkt i den danske levetidsmodel kan der fastlægges stresstests, der for de fleste danske selskaber vil være mere retvisende i forhold til risikoniveauet i Solvens 2.

Den danske Aktuarforening har derfor nedsat en arbejdsgruppe med henblik på at udvikle en model for levetidsstresset. Målene for arbejdsgruppen har været:

- at levere input, som danske pensionsselskaber vil kunne anvende til en ansøgning om anvendelse af en (partiel) intern model for levetidsrisiko,
- at foreslå risikostød, som knytter sig til risikoen ved den årlige opdatering af Finanstilsynets benchmark for den nuværende dødelighed og trenden,
- at foreslå et yderligere risikostød, der afspejler den selskabsspecifikke realisationsrisiko, og
- at de foreslåede risikostød er forholdsvist simple og forklarlige for de beslutningstagere, der skal godkende de anvendte (partielle) interne modeller.

Resultatet af analysen består af et samlet forslag til risikostød for henholdsvis den aktuelle dødelighed og trenden (den systematiske risiko) samt et særligt risikostød for realisationsrisikoen (den ikke-systematiske risiko). Stødene vedrørende de systematiske risici er selskabsuafhængige, mens stødet for realisationsrisikoen er selskabsafhængigt. Tabel 6 viser disse tre komponenter, som skal anvendes sammen med Finanstilsynets levetidsbenchmark.

Risikostød, nuværende dødelighed	6%
Risikostød, trend	6%
Risikostød, realisationsrisiko	$2{,}6/\sqrt{5H}$

Table 6: Risikostød vedrørende den nuværende dødelighed, trend og realisationsrisiko. Størrelsen H angiver det forventede antal dødsfald over en periode på 5 år.

Ved opgørelse af livsforsikringshensættelserne med de foreslåede risikostød skal den nuværende dødelighed først reduceres med 6%. Dette aldersuafhængige stød kan fortolkes som en modellering af usikkerheden i forbindelse med den årlige opdatering af det fælles danske benchmark for den nuværende dødelighed. Den årlige opdatering af de aldersafhængige satser for de forventede fremtidige levetidsforudsætninger fra levetidsbenchmarket er tilsvarende forbundet med en risiko, som beskrives ved at øge forbedringsraterne med 6% for alle aldre. Disse to korrektioner, der kan fortolkes som en beskrivelse af den systematiske risiko forbundet med opdatering af levetidsbenchmarket, kan anvendes for alle pensionsselskaber og afhænger ikke af selskabernes størrelse eller sammensætningen af selskabernes bestande. Risikostødets tredje og sidste komponent er som nævnt ovenfor selskabsspecifik og afhænger af det forventede antal dødsfald i selskabet. I tabel 6 er dette parametriseret via størrelsen H, som angiver det forventede antal dødsfald over en femårig periode.

Forventet antal dødsfald pr. år	1	10	100	1.000	10.000
Forventet antal dødsfald på 5 år (H)	5	50	500	5.000	50.000
Usystematisk dødelighedsstress	52,0%	16,4%	5,2%	1,6%	0,5%

Table 7: Usystematisk stress for varierende porteføljestørrelser. Porteføljens størrelse måles ud fra det forventede antal dødsfald, beregnet under anvendelse af Finanstilsynets levetidsbenchmark.

Tabel 7 viser størrelsen af stresset ved varierende porteføljestørrelser. Det fremgår, at dette stress udgør 1,6% ved et forventet antal dødsfald på 1.000 pr. år, svarende til 5.000 dødsfald over en periode på 5 år. Ved mindre porteføljer vil det usystematiske stress være større. Realisationsrisikoen tager således udgangspunkt i størrelsen af det datamateriale, som har været anvendt i selskabets statistiske analyse. Det bemærkes, at realisationsrisikoen beregnes samlet for mænd og kvinder, hvilket vurderes at give det mest retvisende billede, idet mænd og kvinder sædvanligvis indgår i et risikofællesskab.

Stødene i tabel 6 vedrørende den aktuelle dødelighed og trenden kan fortolkes som 99,5% fraktilen i sandsynlighedsfordelingen for restlevetiden knyttet til den årlige opdatering af levetidsbenchmarket. Risikostødet vedrørende realisationsrisikoen afspejler 99,5% fraktilen i sandsynlighedsfordelingen for selskabets estimat på afvigelsen i forhold til benchmarket for den aktuelle dødelighed. Realisationsrisiko kvantificeres dermed mere præcist, idet der eksplicit tages højde for forsikringsbestandens størrelse og hvorledes dødelighedsforudsætningerne rent faktisk opdateres hvert år.

I tabel 8 findes bedste skøn for de forventede levetider med levetidsbenchmarket for udvalgte aldre, samt forøgelsen af de forventede levetider beregnet under anvendelse af standardmodellens levetidsstress på 20%. For en 60-årig ses standardmodellens stress på 20% at føre til en forøgelse af restlevetiden på 1,7 år for mænd og 1,8 år for kvinder. Den sidste søjle viser den tilsvarende forøgelse af levetiderne ved det nye systematiske stress, hvor den nuværende dødelighed reduceres med 6% i alle aldre, og hvor de årlige forbedringsrater forøges med 6%. Disse tal svarer i princippet til situationen, hvor bestanden er så stor, at den ikke-systematiske risiko er elimineret. Det fremgår af tabellen, at dette stress fører til en levetidsforøgelse på 0,5-0,6 år for en 60-årig, hvilket udgør nedre grænse for levetidsforøgelsen i modellen for en 60-årig.

		Restlevetid	Forøgelse af restlevetid			
Køn	Alder	Benchmark	20%	H=500	H=5.000	Systematisk stød
Mænd	20	66,3	1,7	1,1	0,8	0,7
	40	45,1	1,8	1,1	0,8	0,7
	60	24,7	1,7	0,9	0,7	0,5
	80	8,5	1,1	0,6	0,4	0,3
Kvinder	20	68,6	1,8	1,2	0,9	0,7
	40	47,5	1,9	1,1	0,8	0,7
	60	27,4	1,8	1,0	0,7	0,6
	80	10,5	1,3	0,7	0,5	0,4

Table 8: Forventet restlevetid for mænd og kvinder i udvalgte aldre under anvendelse af Finanstilsynets levetidsbenchmark og forøgelse heraf ved standardmodellens stress på 20%, ved de nye stress for $H = 500$, $H = 5.000$ samt den rene systematiske risiko.

Endelig viser tabellen forøgelsen af restlevetiderne under indregning af både den systematiske og den usystematiske risiko for en bestand med henholdsvis 500 og 5.000 forventede dødsfald over en femårig periode. Ved en portefølje med 500 forventede dødsfald over en periode på 5 år vil det nye levetidsstress føre til en forøgelse af restlevetiden på 0,9 år for mænd og 1,0 år for kvinder, mens levetidsforøgelsen forbundet med levetidsstresset udgør 0,7 år for både mænd og kvinder i alder 60 ved en bestand med i alt 5.000 forventede dødsfald.

Det samlede stød til dødsintensiteterne, som er sammenligneligt med det 20% stød, der gælder i standardmodellen i Solvens 2, er vist for udvalgte aldre i nedenstående tabeller. Tabel 9 viser de samlede selskabsuafhængige stød for de systematiske risici, mens tabel 10 viser de samlede stød for en 60-årig, hvor realisationsrisikoen er medregnet for forskellige værdier for H.

	Alder (mænd)				Alder (kvinder)			
Tid	20	40	60	80	20	40	60	80
0	6%	6%	6%	6%	6%	6%	6%	6%
10	8%	7%	7%	6%	8%	7%	6%	6%
20	8%	8%	7%	6%	9%	8%	7%	6%
30	8%	9%	7%	6%	9%	7%	7%	6%
40	11%	9%	6%		10%	8%	6%	
50	11%	7%	6%		8%	8%	6%	
60	10%	6%			9%	6%		

Table 9: Alders- og tidafhængige stress i procent for den systematiske risiko ved varierende startalder for mænd og kvinder.

	Dødsfald H (mænd)				Dødsfald H (kvinder)			
Tid	50	500	5.000	50.000	50	500	5.000	50.000
0	23%	11%	8%	7%	23%	11%	8%	7%
10	23%	12%	8%	7%	23%	12%	8%	7%
20	23%	12%	9%	7%	23%	12%	9%	7%
30	24%	12%	9%	8%	24%	12%	9%	8%
40	23%	11%	8%	7%	23%	11%	8%	7%
50	22%	11%	8%	7%	22%	11%	8%	7%

Table 10: Samlet stress i procent for den systematiske og usystematiske risiko i alder 60 med varierende forventet antal dødsfald (værdi af H) for mænd og kvinder.

Det fremgår af tabel 10, at dødelighedsstresset for en 60-årig udgør 11-12% ved en bestand med et forventet antal dødsfald på 500. Dette stress er beregnet ved at øge stressene for 60-årige fra tabel 9 med $2{,}6/\sqrt{5 \cdot 500}$. For en bestand med blot 50 forventede dødsfald ses det tilsvarende stress for en 60-årig at udgøre 23-24%. Under Solvens 2 ville den aktuelle og fremtidige dødsintensitet skulle reduceres med 20% for alle aldre og alle tidspunkter.

3 Background: Solvency 2, longevity risk and the Danish longevity benchmark

3.1 Longevity risk in the standard model

With Solvency 2, life and pension insurance companies must calculate the risk associated with longevity. In the Solvency 2 directive, Article 105, 2(b), this risk is described as *"the risk of loss, or of adverse change in the value of insurance liabilities, resulting from changes in the level, trend, or volatility of mortality rates, where a decrease in the mortality rate leads to an increase in the value of insurance liabilities (longevity risk)"*.

In the standard model, the longevity risk is quantified by reducing the mortality intensity by 20% regardless of age, trend or size of the portfolio. This follows e.g. from "Draft Implementing measures Solvency 2 October 2011 Article 107 LUR3" (not published): *"The capital requirement for longevity risk referred to in point (b) of Article 105(3) of Directive 2009/138/EC shall be equal to the loss in basic own funds of insurance and reinsurance undertakings that would result from an instantaneous permanent decrease of 20% in the mortality rates used for the calculation of technical provisions."*

Hence, the risk is defined as the economic loss stemming from an instantaneous, but permanent, decrease in the mortality intensity used for calculating the technical provisions. Following the general framework under Solvency 2, the risk is calibrated using a 99.5% confidence level (Value-at-Risk measure) in a one year time horizon.

3.2 Longevity risk in the Danish longevity model

The longevity risk can be split into (at least) two parts: A systematic part, which is related to the risk for a change in the current level of mortality and the risk for a change in the future mortality improvements (the trend in mortality), and a non-systematic part related to the uncertainty/randomness in the estimation of the company's mortality, which depends on the size of the portfolio.

The approach in the standard model with a reduction of 20% should be viewed as a simple approximation of the aggregate impact of the systematic and the non-systematic risk parts. Within the framework of the Danish longevity model, it is possible to assess the systematic and non-systematic risk parts more precisely. The present paper therefore proposes a partial internal model for longevity risk within the Solvency 2 framework, where the systematic and the non-systematic risk parts are modeled and assessed separately. As mentioned above, the proposed method is closely linked to the mechanisms of the Danish longevity benchmark and takes the following risks into account:

- the risk associated with the yearly updating procedure for the benchmark index of the current mortality intensity (the sector mortality),
- the risk associated with the yearly updating procedure for the benchmark index of the trend in the mortality intensity (estimated from data for the total population), and
- the company specific risk associated with the yearly analysis of the deviation from the benchmark index of the current mortality intensity.

If we view the mortality intensity as random, the loss in basic own funds from an instantaneous permanent decrease in the intensity (for each age and each gender) also becomes random with a certain, but unknown, distribution. The longevity risk is then defined as 99.5% quantile of this distribution calculated with a one year time horizon.

Thus, the multi-dimensionally distributed mortality intensities are transformed into a one dimensional loss distribution, which depends on the exact composition of the company's insurance portfolio. In order to avoid this complexity, we derive the 99.5% quantile for the longevity risk finding the 99.5% quantile in the distribution of the remaining lifetime.

3.3 The systematic risk

Randomness of the longevity benchmark (interpreted as the systematic risk parts) can be divided into randomness associated with the benchmark for the current mortality intensity and randomness associated with the benchmark for the trend.

The benchmark for the current mortality intensity is based on mortality data provided by several large pension companies in Denmark (sector data), while the benchmark for the trend is based on national mortality data. Although the data sources differ the data are clearly correlated since sector data are a subset of national data.

For a model of the stochastic nature of the national mortality intensity, we apply the Poisson Lee-Carter model of Brouhns et al. (2002), see Appendix A. For simplicity, and in order not to underestimate the effect of simultaneous improvements in national and sector mortality, we will assume that improvements in national and sector mortality are fully correlated. This assumption is more fully discussed in Section 4.

We will measure longevity risk in terms of changes in (expected) remaining life time under the benchmark. In order to calibrate the longevity stress we calculate the 99.5% quantile for the distribution of the remaining life time on a one year time horizon for each gender and each age x. This is done by Monte Carlo simulation in the following way:

1. Simulate underlying national and sector mortality for the next year.

2. Simulate observed deaths for sector and national data for the next year using the simulated mortality intensities above.

3. For each simulation, a new current mortality and trend benchmark is derived, and from this benchmark the remaining life time for each age and each gender is calculated.

4. Calculate the 99.5% quantile of the distribution for the remaining life time for each gender and each age.

5. Calibration: Stress the benchmark for current mortality and trend such that the life expectancy under the stressed benchmark matches the calculated 99.5% quantiles.

The simulations in steps 1–3 are performed 10,000 times. In principle it is possible to stop at step 4 above and use a full age and gender dependent stress. However, in order to obtain a simple model we include the last step to derive an age- and gender-independent stress for the current mortality and the trend, respectively.

3.4 The non-systematic risk

Under the Danish regulatory regime each company has to perform a statistical analysis each year to determine whether the mortality experience of the company differs significantly from the benchmark mortality. If the difference is found to be insignificant the company should use the benchmark for current mortality, otherwise it should use an adjusted intensity estimated from the data. The analysis is performed as a Poisson regression treating the benchmark for observed mortality as fixed.

For companies with sparse data, either because they hold a small portfolio or new portfolio, a few extra death counts or the absence of these, can have a strong impact on the regression. On the other hand in large and mature portfolios a few extra death counts have only a small impact on the regression. In other words, small/new portfolios will have a high volatility in their estimation of the mortality, whereas large/mature portfolios will have a low volatility in their estimation of the mortality.

In this paper we will equate non-systematic risk with estimation uncertainty, i.e. the uncertainty associated with determining the underlying company specific mortality from (finite) data. We approach this in a simplified one-parameter theoretical setting where we show that the standard deviation of the overall level of (excess) mortality is approximately inversely proportional to the square root of the expected number of deaths in the company in the estimation period. Using this result, we can quantify the effect of the annual analysis of company mortality and derive an approximate 99.5% stress for non-systematic risk.

4 Technical description: Longevity stress analysis

The following sections contain the motivation and details of the calibration analysis carried out by the working group. The aim of the analysis is twofold:

1. to assess the uncertainty associated with the annual update of the Danish FSA longevity benchmark (systematic longevity risk);

2. to assess the uncertainty associated with the annual company specific calibration of mortality assumptions (non-systematic longevity risk).

Both the systematic and non-systematic longevity risks are calibrated to a 99.5% level on a one year horizon.

The analysis relies on the current Danish regulatory regime which is characterized by: the inclusion of expected, future improvements in life expectancy, calibration to current observed mortality, and annual updates of all assumptions based on the most recent mortality experience. The longevity stress is calibrated on the basis of a well-established stochastic model for describing mortality, taking the details of the Danish regulatory regime into account.

The working group has taken it as a premise for the analysis that the resulting longevity stress should have a simple form which is easy to implement in practice.

4.1 Longevity stress within the Danish regulatory regime

As described in the previous sections, the Danish FSA benchmark consists of annual rates of improvements of age- and gender-specific mortality intensities and a current observed level of mortality of insured lives. In the following we refer to the former as the (benchmark) *trend* and the latter as the (benchmark) *level*.

The gender-specific, benchmark mortality intensity for a person of age x in year t of gender k (F for females, M for males) takes the form

$$\mu_k^{FSA}(x,t) = \mu_k^{FSA}(x,T)\,(1 - R_k(x))^{t-T}\,, \tag{1}$$

where T is the reference year for the benchmark, i.e. the year of the current observed level of mortality. The Danish FSA provides the level, $\mu_k^{FSA}(x,T)$, and the trend, $R_k(x)$, from which the users of the benchmark can construct the full benchmark intensity surface by use of formula (1).

Given the benchmark each company estimates its own company-specific mortality relative to the benchmark. The company-specific mortality is termed the *model* mortality and has the form[2]

$$\mu_k^{model}(x,t) = \exp\left(\beta_1^k r_1(x) + \beta_2^k r_2(x) + \beta_3^k r_3(x)\right)\mu_k^{FSA}(x,t), \tag{2}$$

where the regressors are given by

$$r_m(x) = \begin{cases} 1 & \text{for } x \leq x_{m-1}, \\ (x_m - x)/(x_m - x_{m-1}) & \text{for } x_{m-1} < x < x_m, \\ 0 & \text{for } x \geq x_m, \end{cases} \tag{3}$$

[2]For the analysis performed by the companies a centralized version of the benchmark intensity is used. But this is a technical detail not relevant to the present analysis and is therefore omitted.

for $m = 1, 2, 3$ and $(x_0, x_1, x_2, x_3) = (40, 60, 80, 100)$. The β-parameters are estimated by the company and subject to a significance test in which non-significant parameters are set to zero. This estimation is based on the mortality experience of the company's portfolio over the last 5 years.

For the longevity stress to be easy to use in practice we need it to conform with the benchmark parametrization. As the analysis will show the systematic longevity risk can be adequately captured by stressing the two components of the FSA benchmark mortality in the following simple way

$$\tilde{\mu}_k^{FSA}(x,t) = (1 - S_{level})\mu_k^{FSA}(x,T)\left(1 - (1 + S_{trend})R_k(x)\right)^{t-T}. \tag{4}$$

That is, by reducing the current observed level by the factor S_{level} and increasing the future rates of improvement by the factor S_{trend}. The two factors depend neither on gender nor age. However the impact of S_{trend} in terms of both mortality rates and life expectancy is larger for young ages than for old ages due to the longer time horizon. The stress thereby agrees with the general conception that the level of uncertainty increases with horizon.

The non-systematic longevity risk is interpreted as the variation in the estimated current level of company-specific mortality from year to year due to the randomness of deaths. It can be described by a reduction of the model mortality on top of the systematic risk above. The total longevity stress taking account of both the systematic and non-systematic longevity thus takes the form

$$\tilde{\mu}_k^{model}(x,t) = (1 - S_{rl})\exp\left(\beta_1^k r_1(x) + \beta_2^k r_2(x) + \beta_3^k r_3(x)\right)\tilde{\mu}_k^{FSA}(x,t), \tag{5}$$

where $S_{rl} = 2.6/\sqrt{5H}$ and H is the total expected number of deaths over the last 5 years in the company's portfolio assuming the FSA benchmark mortality (1). The subscript rl refers to the Danish term "realisationsrisiko" coined by the Danish FSA. The reduction factor S_{rl} does not depend on gender nor age, and H is calculated for females and males combined.

4.2 Terminology and model framework

The benchmark trend and benchmark level are estimated each year by the FSA; the trend is estimated from the mortality experience of the Danish population over a 30 year period, while the level is estimated from a large pool of Danish insured lives over a 5 year period. The model mortality used by the company is also estimated each year. The estimation is based on the mortality experience over a 5 year period of the company's portfolio. The actual longevity assumptions used by a given company thus depends on the mortality experience of both the Danish population, the sector (pool of insured lives), and its own portfolio.[3]

In order to assess the longevity risk faced by a company we need to model jointly the mortality experience of these three populations. For a given population we denote by $\mu(x,t)$ the true underlying, but unobservable, mortality intensity for people of age x in year t, and we denote by $D(x,t)$ and $E(x,t)$ respectively the number of deaths and

[3]The procedures for calculating the benchmark trend and benchmark level, as well as the company-specific analysis are described in detail on the following homepage of the Danish FSA: http://www.finanstilsynet.dk/da/Tal-og-fakta/Oplysninger-for-virksomheder/Oplysningstal-om-forsikring-og-pension/Levetidsmodel.aspx [in Danish].

the exposure ("the number of people at risk of dying") of age x in year t. To distinguish between the different populations we superscript these quantities with N for national data (Danish), S for sector data (insured lives), and C for company data. Further, we make the standard assumption that death counts are Poisson distributed with mean μE. In summary, we have

National data:	$D_k^N(x,t) \sim \text{Poisson}\left(\mu_k^N(x,t)E_k^N(x,t)\right)$
Sector data:	$D_k^S(x,t) \sim \text{Poisson}\left(\mu_k^S(x,t)E_k^S(x,t)\right)$
Company data:	$D_k^C(x,t) \sim \text{Poisson}\left(\mu_k^C(x,t)E_k^C(x,t)\right)$

The three underlying intensities, μ_k^N, μ_k^S, and μ_k^C, are stochastic and dependent, while death counts are assumed to be independent conditioned on the underlying intensities.

We think of the benchmark intensity, μ_k^{FSA}, as an estimate of μ_k^S, and the model intensity μ_k^{model} as an estimate of μ_k^C. Within this framework we can study the effect of the uncertainty in the underlying intensities and death counts on the benchmark and model mortality.

4.3 Systematic longevity risk

The analysis of the systematic longevity risk proceeds as follows. First we fit a Poisson Lee-Carter stochastic mortality model for Danish national mortality. We then use this model to generate the joint distribution of national and sector data one year ahead. We assume that the two populations evolve in parallel to guard us from underestimating the double impact of simultaneous changes in both benchmark trend and level. Second we calculate the (remaining) life expectancy distribution for each age and gender one year ahead. Third we calibrate a longevity stress of the form (4) to reproduce the 99.5% quantiles of these life expectancy distributions. The calibration is initially done for each gender separately, and afterwards averaged to arrive at a unisex stress.

4.3.1 Simulating the longevity benchmark

The current benchmark published by the Danish FSA in August 2012 has reference year $T = 2011$. The trend is estimated from Danish data available on the Human Mortality Database (HMD) at www.mortality.org, while the level for the sector is estimated from data for a pool of insured lives gathered by The Danish Centre of Health and Insurance (dk. Helbred og Forsikring).

The most recent Danish data on HMD is for 2009 and the trend is therefore estimated on the basis of data for the 30 year period 1980–2009, while the level is estimated from sector data for the 5 year period 2007–2011. We let $T_{trend} = 2009$ and $T_{level} = 2011$ denote the last data year used in estimation of respectively the trend and the level.

To describe the evolution in Danish mortality we will use the Poisson variant of the classical Lee-Carter model, see Brouhns et al. (2002) and Lee and Carter (1992). We fit this model to Danish data from 1980-2009 for ages 0-105. A description of the model and plots of fit can be found in Appendix A.

With this model at hand we are able to simulate from the distribution of the Danish mortality one year ahead, i.e. we can simulate from the distribution of $\mu_k^N(x, T_{trend}+1)$

for ages $x = 0, \ldots, 105$. In order to calculate the benchmark level we also need to simulate the sector mortality one year ahead. We will assume that the sector experiences the same rates of improvement as in the national data. Specifically we set

$$\mu_k^S(x, T_{level} + 1) = \mu_k^{FSA}(x, T_{level}) \frac{\mu_k^N(x, T_{trend} + 1)}{\mu_k^N(x, T_{trend})}, \tag{6}$$

where the numerator in the last term is the simulated value from the Poison Lee-Carter model, and the denominator is the fitted value of the model to the Danish data at the last data year.

Assumption (6) represents in some sense the worst case since improvements affect both benchmark trend and level simultaneously. On the other hand it can be argued that rates of improvements for insured lives might be higher that for the population at large. Given the short time series of sector data we cannot verify this hypothesis. Also, if that were indeed to be the case it would be reasonable to assume a less than perfect correlation between national and sector data which would offset this effect to some degree. Overall, we find that the chosen model captures the main effect of the longevity risk in a simple way.

Having the joint distribution of national and sector intensities one year ahead we then simulate the actual number of deaths as independent Poisson variates. As we do not yet know the exposures we use the exposures for the last data year as proxy. That is, we simulate national and sector deaths for ages $x = 0, \ldots, 105$ as[4]

$$D_k^N(x, T_{trend} + 1) \sim \text{Poisson}\left(\mu_k^N(x, T_{trend} + 1) E_k^N(x, T_{trend})\right), \tag{7}$$

$$D_k^S(x, T_{level} + 1) \sim \text{Poisson}\left(\mu_k^S(x, T_{level} + 1) E_k^S(x, T_{level})\right). \tag{8}$$

Finally, we use the FSA algorithm for estimating the benchmark trend and level. The trend is estimated from Danish data for the 30 year period from 1981 to $T_{trend}+1 = 2010$, where we use the historic data for the first 29 years and the simulated data for the last year. Similarly, we estimate the level from sector data for the 5 year period from 2008 to $T_{level} + 1 = 2012$, where we use historic data for the first 4 years and simulated data for the last year.

To sum up, we generate a new benchmark by:

- Simulate $\mu_k^N(x, T_{trend} + 1)$ for ages $x = 0, \ldots, 105$ from the Poisson Lee-Carter model
- Calculate $\mu_k^S(x, T_{level} + 1)$ for ages $x = 0, \ldots, 105$ using relation (6) with the simulated $\mu_k^N(x, T_{trend} + 1)$ ....
- Simulate national and sector data using the simulated intensities and the last known exposures by (7)–(8)
- Compute a new benchmark trend from 29 years of historic data and the simulated national data for year $T_{trend} + 1$
- Compute a new benchmark level from 4 years of historic data and the simulated sector data for year $T_{level} + 1$

We repeat this procedure 10000 times which gives us the benchmark distribution with reference year $T_{level} + 1 = 2012$.

[4] Technical note. For ages 0-25 the benchmark level is based on Danish data. We therefore use the exposure for Denmark from year T_{trend} for these ages when simulating sector data for year $T_{level}+1$.

4.3.2 Calibration of systematic longevity stress

The next step in the analysis is to use the benchmark distribution to find a longevity stress corresponding to a 99.5% confidence level. We are here faced with the usual problem when dealing with multidimensional distributions that it is not obvious which of the 10000 simulated benchmarks belong to the "worst" 0.5%. One way to proceed would be to calculate the marginal 99.5% quantiles separately for the trend and for the level for each age (and each gender). This however clearly exaggerates the stress since the probability that all of these events happen at once is less than 0.5%.

Since we are looking at longevity risk it is safe to assume that the worst benchmarks (in terms of economically most expensive for the companies) are those with the highest (remaining) life expectancy. We will therefore calibrate the longevity stress of the form (4) such that the life expectancy of $\tilde{\mu}_k^{FSA}$ equals the 99.5% quantile in the life expectancy distribution of the simulated benchmarks for a range of (economically) relevant age groups.

First we calculate for each of the simulated benchmarks (pair of trend and level) the cohort life expectancy for ages $x = 0, \ldots, 100$ in 2012

$$e_k^i(x, 2012) = e(x, 2012; \mu_k^i, R_k^i) \text{ for } i = 1, \ldots, 10000, \tag{9}$$

where μ_k^i and R_k^i denote respectively the simulated benchmark level and trend in 2012 for gender k. We now have the life expectancy distribution in 2012 for each age and gender. In each of these distributions we calculate the 99.5% quantile, $q_k(x, 2012)$.

We also calculate the life expectancy in 2012 under the current FSA benchmark for ages $x = 0, \ldots, 100$

$$e_k^{FSA}(x, 2012) = e(x, 2012; \mu_k^{FSA}(\cdot, 2011), R_k), \tag{10}$$

where $\mu_k^{FSA}(\cdot, 2011)$ and R_k denote respectively the current benchmark level and trend for gender k. The difference,

$$\Delta_k(x) = q_k(x, 2012) - e_k^{FSA}(x, 2012), \tag{11}$$

can be interpreted as the increase in the benchmark life expectancy for age x which can happen with probability 0.5% by the next annual update of the benchmark. By calculating both life expectancies in 2012 we only look at the increase in longevity in excess of what is already anticipated by the current benchmark.

Figure 1 shows the life expectancy increase Δ_k (solid black line) together with the life expectancy increase resulting from a uniform reduction in the current benchmark of 5%, 10%, 15% and 20% (dotted lines). The figure also features two dashed green and blue lines which we will return to shortly. We can see from the figure that Δ_k corresponds to a uniform stress in the range from 5% to 10%. We also note that compared to a uniform stress the profile of Δ_k is different. The life expectancy increase under Δ_k is more rapidly decreasing with age which seems to conform better with intuition than a uniform stress. (The peculiar bump for old women is presumable due to parametric extrapolation used after age 80 when constructing the benchmark level. The bump is also visible for men but less pronounced.)

The last step in the analysis is to calibrate a longevity stress of the form (4) such that the "stressed" life expectancy corresponds to the 99.5% quantiles $q_k(x, 2012)$. To formalize the calibration procedure we define the difference

$$\tilde{\Delta}_k(x; S_{trend}, S_{level}) = \tilde{e}_k^{FSA}(x, 2012; S_{trend}, S_{level}) - e_k^{FSA}(x, 2012), \tag{12}$$

where $\tilde{e}_k^{FSA}(x, 2012; S_{trend}, S_{level})$ denotes the life expectancy in year 2012 for age x and gender k calculated from the stressed benchmark (4). Using this terminology the dotted lines in Figure 1 correspond to $\tilde{\Delta}_k(x; 0, S_{level})$ for S_{level}=5%, 10%, 15% and 20%.

To find the combination of S_{trend} and S_{level} we minimize the squared distance

$$\sum_{x=x_{min}}^{x_{max}} \left(\tilde{\Delta}_k(x; S_{trend}, S_{level}) - \Delta_k(x) \right)^2 \tag{13}$$

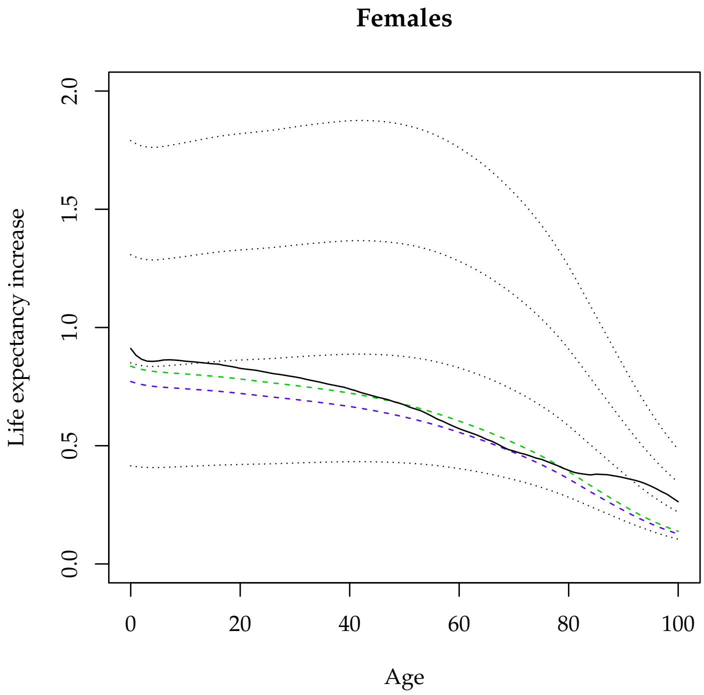
for the economically relevant age span $(x_{min}, x_{max}) = (30, 90)$. We minimize (13) over S_{trend} and S_{level} in steps of 0.5%. The gender-specific calibrated values are shown in Table 11. It is seen that the stress are of the same magnitude for females and males, with the stress for females being 1%-point higher for both parameters. The table also contains a column labeled unisex of the average values.

Parameter	Females	Males	Unisex
S_{trend}	6.5%	5.5%	6.0%
S_{level}	6.5%	5.5%	6.0%

Table 11: Calibrated values of gender-specific and unisex parameters S_{trend} and S_{level}.

The green and blue dashed lines in Figure 1 show the life expectancy increase for respectively the calibrated gender-specific and unisex values. That is, the green line is $\tilde{\Delta}_k(x; 6.5\%, 6.5\%)$ for females and $\tilde{\Delta}_k(x; 5.5\%, 5.5\%)$ for males, while the blue line is $\tilde{\Delta}_k(x; 6.0\%, 6.0\%)$ for both females and males. The fit of the green line is remarkable good from age 40 to 80 for both genders, while deviating somewhat above age 80 in particular for females. This demonstrates that a stress of the form (4) is able to capture in a simple way the systematic longevity risk for the economically most important age groups.

The higher stress for females than for males is driven by a more volatile development of female mortality than of male mortality in the historic period used to estimate the underlying Lee-Carter model. It might be argued that we have no reason to believe that this is an intrinsic feature of female mortality, and that looking forward it would be more reasonable to assume the same level of uncertainty for both genders. This can be obtained by using the unisex stress in Table 11. The unisex stress is shown as the blue dashed line in Figure 1. By construction the unisex stress is slightly lower for females and slightly higher for males than the gender-specific stress but it still provides a very good description of the systematic longevity risk overall.



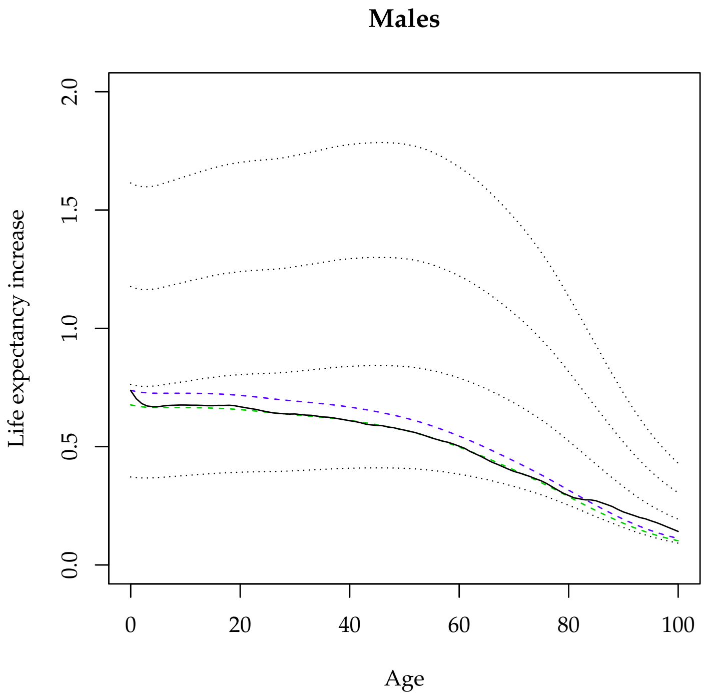


Figure 1: Increase in life expectancy for ages 0–100 in 2012 relative to current FSA benchmark for females (top) and males (bottom): 99.5% quantile in simulated life expectancy distribution (solid black), gender-specific stress (dashed green) and unisex stress (dashed blue), see Table 11 for values of S_{trend} and S_{level}. Also shown, life expectancy increase for stress with S_{trend}=0 and S_{level}=5%, 10%, 15% and 20% (dotted).

4.4 Non-systematic longevity risk

Non-systematic longevity risk is interpreted as the risk associated with the variation in the estimated current level of company-specific mortality from year to year due to the randomness of deaths (Poisson variation). Like in the analysis of systematic longevity risk we will analyze the *change* in assumptions from year to year taking account of the overlap in data used in consecutive years. We will derive the non-systematic longevity stress by theoretical means based on the general framework of Section 4.2.

Under the current Danish regulatory regime companies can only use a model mortality differing from the benchmark if the deviations are statistically significant. This implies that small companies will tend to use the benchmark every year and therefore — in the strict sense of the definition above — faces *no* non-systematic longevity risk. This is clearly not the intention of the regulators. In order to arrive at a stress which is meaningful for all companies regardless of size we will therefore disregard this aspect of the regulation, and perform the analysis as if all companies used an estimated model mortality. In other words, we will quantify the change in the annual estimate of the level of company-specific mortality whether or not this is actually used. This gives the true picture of uncertainty in a given population, and gives rise to a stress that scales (inversely) with the size of the population.

4.4.1 Company data

We will use the following model for a company where we use a single parameter α, independent of age and gender, to measure the overall excess mortality relative to sector mortality

$$D_k^C(x,t) \sim \text{Poisson}\left(\mu_k^C(x,t)E_k^C(x,t)\right), \text{where} \tag{14}$$

$$\mu_k^C(x,t) = \alpha\mu_k^S(x,t). \tag{15}$$

This model allows us to derive a simple formula for non-systematic longevity risk which depends only on a single summary statistic of the member base. The model although simple captures the effect of the Poisson variation in a succinct way.

In the actual regulation the companies are required to perform the Poisson regression (2), where μ_k^{FSA} can be viewed as an estimate of μ_k^S. It would be possible to take this parametric form as our starting point and study the distribution of the β-estimates under this model. The results, however, will depend on the entire age-profile of the company which leaves us with little hope of arriving at a general formula. On the other hand, companies may well want to perform this analysis on their own.

One could also consider to let the level of excess mortality depend on sex, i.e. to have α_k instead of α in (15). At first sight this would perhaps appear more natural considering that the company-specific mortality is estimated for each sex separately. However, if we assume that the economic impact is, at least approximately, proportional to the estimated level multiplied by the size of the population for which that level applies it follows that the combined effect is proportional to the estimated level of *average* excess mortality multiplied by the size of the total population. Therefore we can assess the combined effect in the simpler one-parameter model above.[5]

[5] The argument can be made more mathematically rigorous as follows. In the two-parameter model, $\hat{\alpha}_F$ and $\hat{\alpha}_M$ are estimated by (16) where the summations extend only over t and x. Denoting by

The rationale behind linking the company specific mortality to the sector mortality is the same as in Section 4.3.1. We assume that the mortality of the member base of the company evolves in parallel to the sector mortality such that the non-systematic risk is indeed an add-on risk on top of the systematic risk. As in Section 4.3.1 it could well be argued that the evolution of company specific mortality may differ from that of the sector, but then it would also be reasonable to assume a less than perfect dependency and thereby a diversification discount. Like before we find that the chosen model captures the main effect of non-systematic longevity risk in a simple way.

4.4.2 Estimation uncertainty

Before turning to the variation in the estimate of α for two successive years we first look at the estimate of α itself. Assuming μ_k^S is know, the maximum likelihood estimator for α is given by

$$\hat{\alpha} = \frac{\sum_{x,t,k} D_k^C(x,t)}{H}, \quad \text{with } H = \sum_{x,t,k} \mu_k^S(x,t) E_k^C(x,t), \tag{16}$$

where the summation extends over the same ages and years used to estimate the model mortality. Under Danish regulation the model mortality is estimated on the basis of data from a 5 year period, and hence we will assume an estimation period of 5 years. The quantity H is the number of deaths expected in the data period had the company mortality been equal to the sector mortality.

Under the model we have that

$$\sum_{x,t,k} D_k^C(x,t) \sim \text{Poisson}(\alpha H), \tag{17}$$

and thereby

$$\hat{\alpha} \sim \frac{\text{Poisson}(\alpha H)}{H}, \quad \text{E}(\hat{\alpha}) = \alpha, \quad \text{Var}(\hat{\alpha}) = \frac{\alpha}{H}, \quad \text{Std}(\hat{\alpha}) = \frac{\sqrt{\alpha}}{\sqrt{H}}. \tag{18}$$

We see that the standard deviation of the estimator depends on the true value of α (and H). Since the true value of α is unknown a possible solution would be to use the estimated value of α instead. However, for small companies the estimation uncertainty is substantial and this approach will therefore lead to variation in the perceived uncertainty from year to year.

A more stable estimate of the standard deviation which is applicable to all companies regardless of size can be obtained by noting that α and thereby $\sqrt{\alpha}$ will (on average) be close to 1 since α measures the level of mortality for a specific group of insured lives relative to other insured lives. We can therefore approximate the standard deviation by

$$\text{Std}(\hat{\alpha}) \sim \frac{1}{\sqrt{H}}. \tag{19}$$

Note that the approximation does not imply that we assume α to equal 1. The approximation can be used for all α but of course it is more precise the closer α is to 1.

H_F and H_M the denominator in the fraction defining $\hat{\alpha}_F$ and $\hat{\alpha}_M$ respectively, we have the following relation between the estimators: $\hat{\alpha}H = \hat{\alpha}_F H_F + \hat{\alpha}_M H_M$. Hence, if the combined effect is a function of $\hat{\alpha}_F H_F + \hat{\alpha}_M H_M$ it is also a function of $\hat{\alpha}H$, and we need only consider the one-parameter model to assess the risk.

4.4.3 Uncertainty of successive estimates

We now extend the analysis of the preceding section to study the variability in the change of the α estimate from one year to the next. To present the analysis we define the partial sums

$$D_t = \sum_{x,k} D_k^C(x,t) \quad \text{and} \quad H_t = \sum_{x,k} \mu_k^S(x,t)_k E_k^C(x,t). \tag{20}$$

We will also need to distinguish between different estimation periods and for this purpose we will subscript the estimator with the last data year of the estimation period, i.e. we will write $\hat{\alpha}_T$ for the estimate of α based on data from the 5 year period from $T-4$ to T. In the notation introduced above we have

$$\hat{\alpha}_T = \frac{D_{T-4} + \ldots + D_T}{H_{T-4} + \ldots + H_T}, \tag{21}$$

and for the ratio of two successive estimators we have

$$\frac{\hat{\alpha}_{T+1}}{\hat{\alpha}_T} = \frac{H_{T-4} + \ldots + H_T}{H_{T-3} + \ldots + H_{T+1}} \left(\frac{D_{T-3} + \ldots + D_T}{D_{T-4} + \ldots + D_T} + \frac{D_{T+1}}{D_{T-4} + \ldots + D_T} \right). \tag{22}$$

We are interested in the variability of (22) caused by inclusion of the new data D_{T+1}. In statistical terms we want to find the standard deviation of (22) conditioned on data up to and including time T

$$\begin{aligned}
\operatorname{Std}\left(\frac{\hat{\alpha}_{T+1}}{\hat{\alpha}_T} | \{D_t\}_{t \le T} \right) &= \frac{H_{T-4} + \ldots + H_T}{H_{T-3} + \ldots + H_{T+1}} \operatorname{Std}\left(\frac{D_{T+1}}{D_{T-4} + \ldots + D_T} \right) \\
&\approx \frac{\operatorname{Std}(D_{T+1})}{\alpha (H_{T-4} + \ldots + H_T)} \\
&= \frac{\sqrt{\alpha H_{T+1}}}{\alpha (H_{T-4} + \ldots + H_T)} \\
&\approx \frac{1}{\sqrt{\alpha}\sqrt{5H}} \\
&\approx \frac{1}{\sqrt{5H}},
\end{aligned} \tag{23}$$

where $H = H_{T-4} + \ldots + H_T$ (as in Section 4.4.2). In the first approximation above $D_{T-4} + \ldots + D_T$ is replaced by its expectation $\alpha (H_{T-4} + \ldots + H_T)$, and the H-ratio is replaced by 1. The latter rests on the assumption that the exposure and thereby H_t is almost constant over time. This assumption is also used in the second approximation to replace H_{T+1} by $H/5$. Finally, we approximate $\sqrt{\alpha}$ by 1, by the same arguments as in Section 4.4.2, to arrive at a stable, easily implementable formula.

Comparing formulas (19) and (23) we see that the effect of reusing 4 years of data in successive estimations is to reduce the standard deviation of the ratio of α estimates by a factor of $\sqrt{5}$ relative to the standard deviation of α itself.

4.4.4 Calibration of non-systematic longevity stress

The theoretical developments above show that the standard deviation of the change in the estimate of the level of company mortality from one year to the next attributable to non-systematic (Poisson) variation can be approximated by $1/\sqrt{5H}$. From this we conclude that a 99.5% non-systematic longevity stress can be obtained by setting S_{rl} of (5) equal to

$$S_{rl} = \frac{2.6}{\sqrt{5H}}. \tag{24}$$

The stress is obtained by utilizing that the distribution of the ratio follows a (scaled and translated) Poisson distribution which is well approximated by a normal distribution. The stress S_{rl} corresponds to the 99.5% quantile in the approximating normal distribution.

As seen in the derivations in Sections 4.4.2–4.4.3 leading to (24) we approximate α by 1 to achieve a simple and stable formula. We justified the approximation by the observation that on average company mortality ought to be close to sector mortality. On the other hand, for a specific company the value of α will deviate to a smaller or larger extent from 1 and the approximation will therefore (from a theoretical perspective) introduce a bias in the estimate of the non-systematic risk of the given company. For $\alpha = 0.8$ and $\alpha = 0.6$ the stress in (24) should be multiplied by respectively 1.12 and 1.29, while for $\alpha = 1.2$ and $\alpha = 1.4$ the stress should be multiplied by 0.91 and 0.85.

If instead we approximate α by its estimator $\hat{\alpha}$ we can reduce the bias at the expense of introducing variability in the estimate of the risk. This is the usual variance-bias tradeoff. For large populations the use of $\hat{\alpha}$ will tend to improve the estimate of the risk, while for small populations the use of $\hat{\alpha}$ will tend to degrade the estimate of the risk. Overall we judge that the benefits of simplicity and stability outweigh the potential bias introduced by the approximation.

In the theoretical derivations H is expressed in terms of the underlying sector mortality μ_k^S. In practice we will use the FSA benchmark instead and calculate H by the formula

$$H = \sum_{x,t,k} \mu_k^{FSA}(x,t) E_k^C(x,t), \tag{25}$$

where the summation extends over the same 5 year period used to estimate the model mortality of the company.

The magnitude of the stress depends on the size of the population through H. Table 12 shows the stress for various values of H ranging from a very small population with only 1 expected death per year ($H = 5$) to a very large population with 10000 expected deaths per year ($H = 50000$). The expected number of deaths are calculated under the FSA benchmark. We note that for very small populations the size of the stress might be excessive, and one might consider to introduce a cap.

Expected number of deaths per year	1	10	100	1000	10000
Expected number of deaths in 5 years (H)	5	50	500	5000	50000
Non-systematic longevity stress (S_{rl})	52.0%	16.4%	5.2%	1.6%	0.5%

Table 12: Non-systematic longevity stress, S_{rl} for various population sizes. The size of the population is measured by the expected number of deaths under the FSA benchmark mortality.

Let us finally illustrate the combined effect of the systematic and non-systematic longevity stress. We will illustrate the effect by the increase in life expectancy relative to the current benchmark for different population sizes. To make the results comparable to those of Section 4.3 we use a company with all β's equal to zero, and we use the unisex systematic stress derived in Section 4.3.2 with $S_{trend} = S_{level} = 6.0\%$. In Figure 2 the life expectancy increase in 2012 for the systematic stress relative to the current benchmark is shown as the solid blue line (as in Figure 1).

The additional effect of including non-systematic risk for populations with respectively 10, 100 and 1000 expected number of deaths per year are shown as dashed blue lines. For comparison the figure also shows the life expectancy increase for a uniform reduction in the benchmark of 5%, 10%, 15% and 20% as dotted black lines (as in Figure 1). For a population with only 10 expected deaths per year (uppermost dashed blue line) the combined effect is more severe than a uniform reduction of 20%, while for a population with 1000 expected deaths per year the combined effect corresponds to a uniform reduction of approximately 10%.

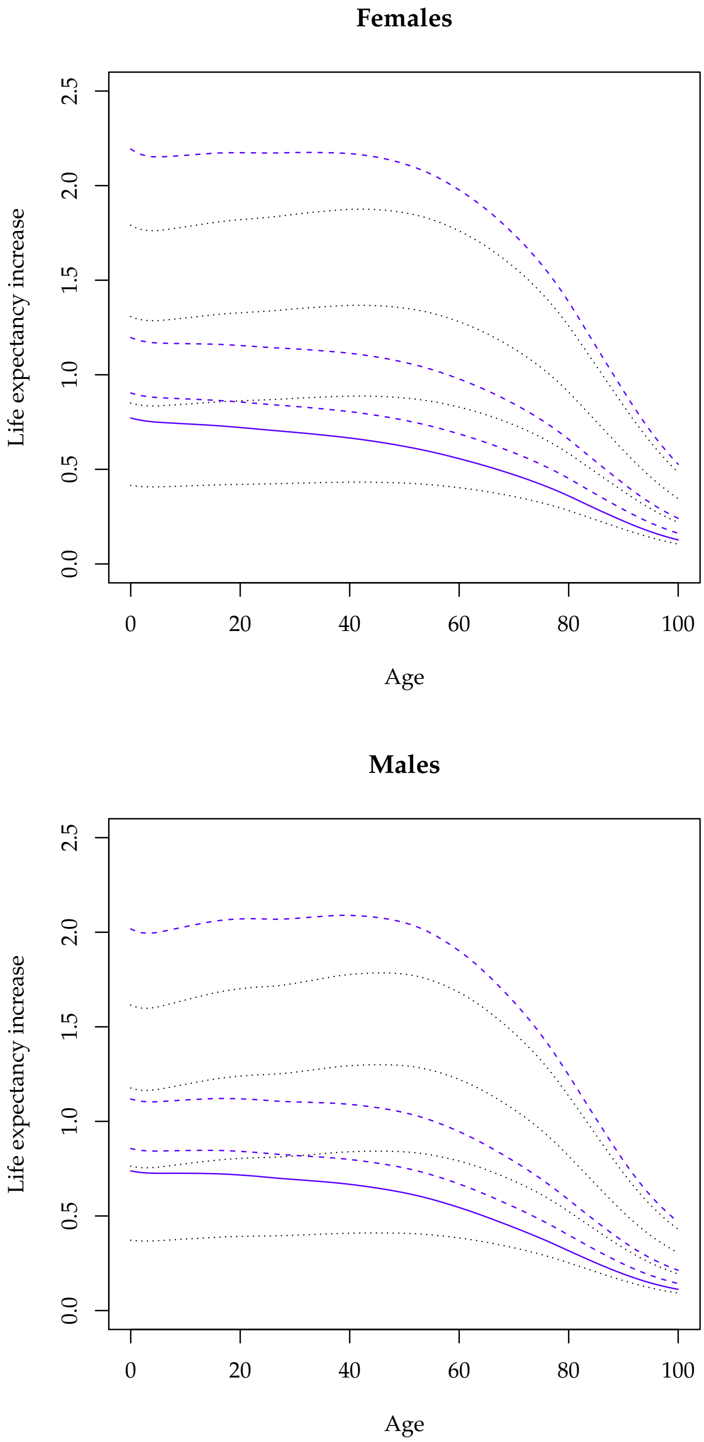


Figure 2: Increase in life expectancy for ages 0–100 in 2012 relative to current FSA benchmark for females (top) and males (bottom): unisex stress without non-systematic risk (solid blue) and with non-systematic risk (dashed blue) for populations with respectively 10, 100 and 1000 expected number of deaths per year. Also shown, life expectancy increase for stress with $S_{rl} = S_{trend}$=0 and S_{level}=5%, 10%, 15% and 20% (dotted).

A Poisson Lee-Carter model

We will use the Poisson variant of the classical Lee-Carter model to describe the evolution in Danish mortality, see Brouhns et al. (2002) and Lee and Carter (1992). The model assumes that

$$D_k^N(x,t) \sim \text{Poisson}\left(\mu_k^N(x,t)E_k^N(x,t)\right), \text{with} \tag{26}$$

$$\mu_k^N(x,t) = \exp\left(a_k(x) + b_k(x)h_k(t)\right), \tag{27}$$

where $a(x)$ and $b(x)$ are respectively an age-dependent level and "rate of improvement", while $h(t)$ is an index common to all ages of accumulated "improvement" evolving in time. Note that it is the product of b and h that determines the actual improvements, and that b and h themselves are only measures of improvement (hence the quotation marks in the previous sentence). In order to identify the parameters constraints must be imposed. We use the usual parameter constraints,

$$\sum_t h_k(t) = 0 \quad \text{and} \quad \sum_x b_k(x) = 1, \tag{28}$$

and the maximum likelihood fitting procedure described in Brouhns et al. (2002).

The Lee-Carter model is a standard model used in numerous mortality studies. It is a one factor model in which all improvements are fully correlated. From a longevity risk perspective this is in a sense the worst that can happen, and therefore this model although simple seems well-suited to study longevity risk. The actual number of deaths are assumed independent (conditioned on μ) and will therefore make only a small contribution to the total risk since random excess mortality is unlikely to occur in many age groups at once. The Poisson version is chosen because it complies with our general framework and because it handles cells with no deaths or no exposure better than the original Lee-Carter model. Also, with the original Lee-Carter model one typically adjusts the fit in the jump-off year to reproduce the actual number of deaths in order to avoid bias in the forecast. This is not needed with the Poisson version since it is fitted to the number of deaths, ref. Brouhns et al. (2002).

We have fitted the model to Danish data for the period 1980-2009 for ages 0-105 for each sex separately. The data are available at the Human Mortality Database (HMD) at www.mortality.org. These are the same data used by the Danish FSA to estimate the trend in the current benchmark. The parameter estimates for a and b are shown in Tables 14–16.

Figure 3 shows the raw death rates (solid lines), the Lee-Carter fit (dashed lines) and 95% confidence intervals obtained from the model (dotted lines) for selected ages. The model provides a decent fit to data and, importantly, the fluctuations in observed death rates are within the confidence intervals predicted by the model. Figure 4 shows the fit for all ages in the first and last data year. This confirms that data are well described by the model.

Having estimated the model we need to specify the dynamics of the index of mortality, h_k, in order to simulate from the model. Again we will follow the standard route and treat the estimated index of mortality as an observed time series and assume that it follows a random walk with drift

$$h_k(t+1) = h_k(t) + \epsilon_k(t+1), \quad \text{where } \epsilon_k \text{ are independent } N(\xi_k, \sigma_k^2). \tag{29}$$

The mean and standard deviation of the innovations, ϵ, are estimated from the "observed" part of h_k by the usual estimators

$$\hat{\xi}_k = \frac{1}{n}\sum_{t=1981}^{2009} \Delta h_k(t), \quad \hat{\sigma}_k = \left\{ \frac{1}{n-1} \sum_{t=1981}^{2009} \left(\Delta h_k(t) - \hat{\xi}_k \right)^2 \right\}^{1/2}, \tag{30}$$

where $\Delta h_k(t) = h_k(t) - h_k(t-1)$ and $n = 29$ equals the number of differences of the series. The estimated drift and standard deviation of the innovations are shown in Table 13. The table also shows the value of the index in the last data year, $h_k(2009)$, used as jump-off value when forecasting.

The estimated standard deviation is higher for females than for males reflecting a more volatile historic evolution in female mortality. The higher standard deviation for females gives rise to a higher stress for females than for males when we perform the gender-specific analysis in Section 4.3.2. Although we have no reason to believe that female mortality is intrinsically more volatile than male mortality we have chosen to preserve the estimated quantities in the analysis. Instead we will propose a unisex stress on the basis of the results of the gender-specific analysis to reflect the point of view that the level of uncertainty should be the same for the two genders going forward.

With the estimates in place it is now straightforward to simulate from the model one or more years into the future

- Draw independent innovations, $\epsilon_k(2010), \ldots$, from a normal distribution with mean and standard deviation given by (30)
- Forecast the mortality index, h_k, by repeated use of (29) starting from the value of the index in the last data year, $h_k(2009)$
- Calculate the future mortality intensity, $\mu_k^N(x,t)$, by the relation in (27) using the forecasted value of the mortality index and $a_k(x)$ and $b_k(x)$ replaced by their estimated values

The assumption that the mortality index evolves like a random walk rules out the possibility of a structural break. For long term projections this might be a questionable assumption. However, since we will be using the model only to quantify the uncertainty one year ahead and since this uncertainty is dominated by the standard deviation of the innovations this is deemed not to pose a problem.

Gender	$\hat{\xi}_k$	$\hat{\sigma}_k$	$h_k(2009)$
Females	-1.8953	4.0983	-35.1604
Males	-1.9511	2.4744	-35.4906

Table 13: Estimated drift and standard deviation for innovations together with the value of the index in the last data year for females ($k = F$) and males ($k = M$).

Age (x)	a_F	a_M	b_F	b_M
0	-5.2852	-5.0516	0.014977	0.016817
1	-7.6082	-7.5357	0.018646	0.018602
2	-8.2271	-8.0282	0.016182	0.021751
3	-8.5601	-8.2628	0.018628	0.021908
4	-8.6725	-8.4399	0.025759	0.017943
5	-9.0276	-8.7128	0.027802	0.022109
6	-8.9507	-8.6677	0.022831	0.023158
7	-9.0287	-8.6197	0.018354	0.016113
8	-9.1070	-8.5860	0.012625	0.026694
9	-9.0957	-8.6309	0.029581	0.025271
10	-9.0320	-8.7879	0.018463	0.021003
11	-8.9457	-8.6930	0.015746	0.013293
12	-8.9864	-8.7546	0.020783	0.011154
13	-8.9570	-8.3318	0.019682	0.014655
14	-8.6902	-8.2291	0.012106	0.012718
15	-8.5894	-8.1033	0.007132	0.017914
16	-8.3135	-7.6249	0.013635	0.010525
17	-8.2710	-7.4874	0.008275	0.007875
18	-8.0740	-7.0391	0.009997	0.005805
19	-8.0483	-7.0570	0.012894	0.013158
20	-8.1060	-7.0291	0.014640	0.010795
21	-8.1621	-7.0110	0.012403	0.010692
22	-8.1284	-7.0816	0.019284	0.008493
23	-8.0492	-7.0516	0.010351	0.011294
24	-8.0187	-6.9606	0.013171	0.009419
25	-7.9687	-6.9996	0.009467	0.012524
26	-8.0036	-6.9440	0.013336	0.013062
27	-7.8116	-6.9354	0.017240	0.012062
28	-7.8876	-6.9270	0.012597	0.010654
29	-7.5962	-6.8774	0.010159	0.014895
30	-7.6814	-6.8197	0.013495	0.014925
31	-7.5437	-6.7338	0.014562	0.012697
32	-7.4469	-6.6945	0.013075	0.014604
33	-7.4176	-6.6261	0.018115	0.010598
34	-7.3102	-6.5775	0.014210	0.012364
35	-7.1450	-6.5470	0.008642	0.013219

Table 14: Estimated Lee-Carter parameters for females and males for ages 0–35.

Age (x)	a_F	a_M	b_F	b_M
36	-7.0603	-6.4882	0.011779	0.009046
37	-7.0139	-6.3828	0.015954	0.012594
38	-6.8675	-6.3150	0.013743	0.011802
39	-6.7152	-6.1867	0.013199	0.008049
40	-6.6931	-6.1119	0.010152	0.009443
41	-6.5041	-6.0345	0.014909	0.007922
42	-6.4448	-5.9505	0.007704	0.008370
43	-6.3396	-5.8666	0.013490	0.008288
44	-6.1933	-5.7661	0.012097	0.007205
45	-6.1129	-5.7035	0.010188	0.006668
46	-6.0254	-5.5974	0.011186	0.006662
47	-5.9129	-5.5049	0.007958	0.009189
48	-5.7907	-5.3763	0.010137	0.006426
49	-5.7025	-5.2915	0.008440	0.005481
50	-5.6324	-5.1750	0.009528	0.006041
51	-5.5232	-5.1146	0.009787	0.006467
52	-5.3994	-4.9992	0.008565	0.005912
53	-5.3231	-4.9284	0.008397	0.006790
54	-5.2343	-4.8131	0.008440	0.008089
55	-5.1443	-4.7144	0.007768	0.009337
56	-5.0569	-4.6338	0.008655	0.008022
57	-4.9912	-4.5414	0.009411	0.010773
58	-4.8898	-4.4257	0.009529	0.009484
59	-4.8147	-4.3234	0.008972	0.010508
60	-4.7133	-4.2269	0.008183	0.010783
61	-4.6291	-4.1353	0.008200	0.010079
62	-4.5236	-4.0492	0.007761	0.009887
63	-4.4459	-3.9491	0.007205	0.010870
64	-4.3640	-3.8491	0.008106	0.010363
65	-4.2651	-3.7661	0.007017	0.010563
66	-4.1807	-3.6807	0.005731	0.010438
67	-4.0937	-3.5736	0.005777	0.010481
68	-4.0156	-3.4896	0.005642	0.009615
69	-3.9052	-3.3947	0.004471	0.009674
70	-3.8164	-3.3044	0.004594	0.009742

Table 15: Estimated Lee-Carter parameters for females and males for ages 36–70.

Age (x)	a_F	a_M	b_F	b_M
71	-3.7207	-3.2069	0.004298	0.009263
72	-3.6293	-3.1151	0.003523	0.008919
73	-3.5335	-3.0084	0.003493	0.008318
74	-3.4513	-2.9236	0.003531	0.008405
75	-3.3460	-2.8166	0.002773	0.007815
76	-3.2380	-2.7322	0.003591	0.008080
77	-3.1467	-2.6268	0.003420	0.007429
78	-3.0455	-2.5321	0.003242	0.006713
79	-2.9377	-2.4457	0.004314	0.006722
80	-2.8290	-2.3582	0.004203	0.006268
81	-2.7240	-2.2678	0.004205	0.006049
82	-2.6136	-2.1744	0.004361	0.004914
83	-2.4934	-2.0878	0.004779	0.004293
84	-2.3902	-2.0046	0.005149	0.003718
85	-2.2715	-1.9174	0.004754	0.003746
86	-2.1722	-1.8127	0.004660	0.003045
87	-2.0458	-1.7324	0.004021	0.003151
88	-1.9370	-1.6355	0.004393	0.003031
89	-1.8267	-1.5275	0.003375	0.002428
90	-1.7111	-1.4474	0.003886	0.001610
91	-1.6074	-1.3625	0.003514	0.001124
92	-1.4861	-1.2501	0.003326	0.001903
93	-1.3999	-1.1862	0.003322	0.001292
94	-1.2920	-1.1174	0.002384	0.000291
95	-1.2104	-1.0065	0.002535	0.002192
96	-1.0976	-0.9331	0.002413	0.000690
97	-1.0362	-0.8483	0.002406	0.002432
98	-0.9481	-0.7941	0.000965	0.000090
99	-0.8824	-0.7518	0.000744	-0.000934
100	-0.7458	-0.6718	0.003339	0.005341
101	-0.6991	-0.6403	0.001434	0.002872
102	-0.6428	-0.4206	0.002888	0.005089
103	-0.5863	-0.3604	0.002164	0.001786
104	-0.5090	-0.3570	0.001976	0.002522
105	-0.3851	-0.1120	0.005097	0.017565

Table 16: Estimated Lee-Carter parameters for females and males for ages 71–105.

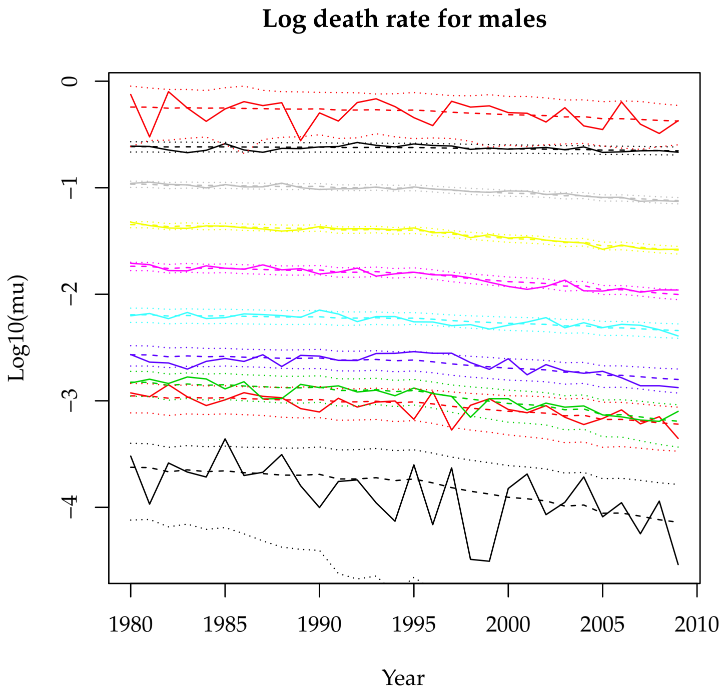


Figure 3: Data and model fit for Danish females (top) and males (bottom) for the period 1980-2009. The plots show observed death rates (solid), Lee-Carter fit (dashed) and 95% confidence intervals (dotted) for ages 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 years.

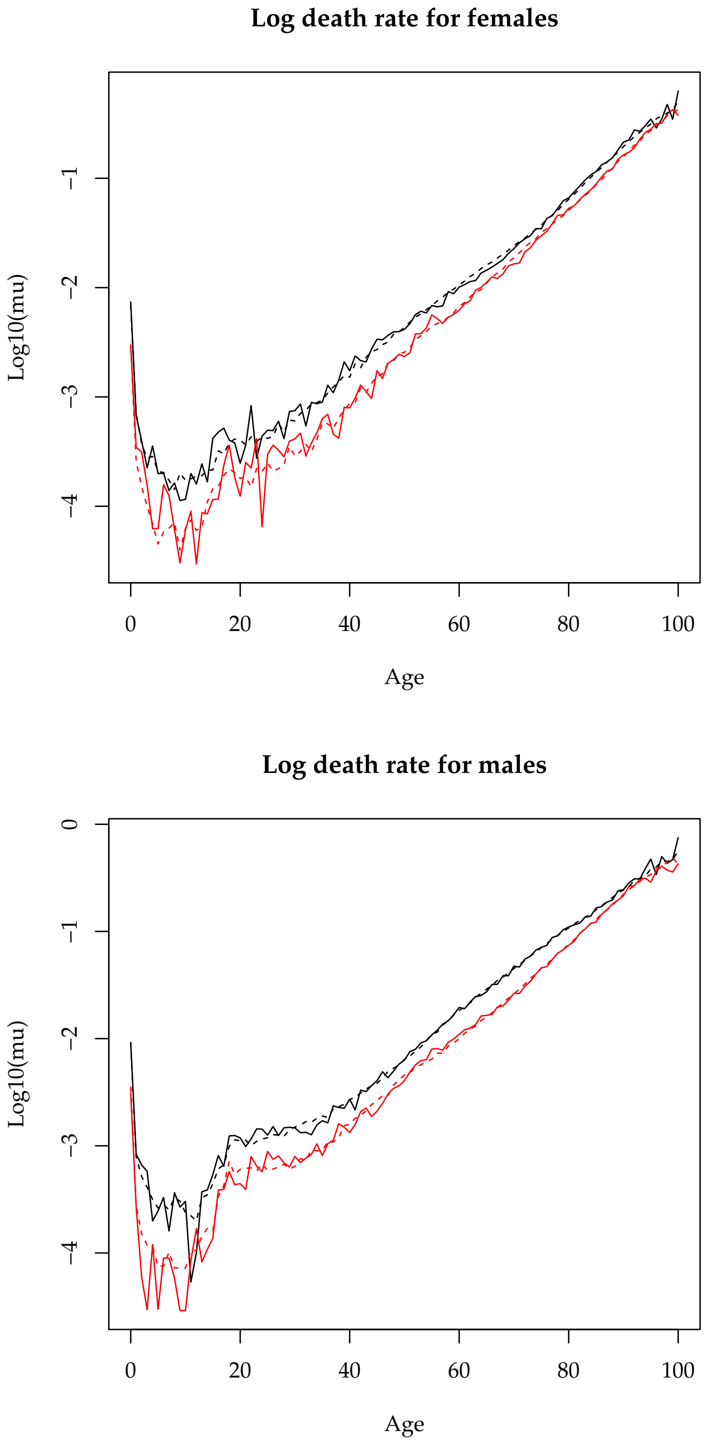


Figure 4: Data and model fit for Danish females (top) and males (bottom) for ages 0–100 years. The plots show observed death rates (solid) and Lee-Carter fit (dashed) in 1980 (black) and 2009 (red).

References

Brouhns, N., Denuit, M. and Vermunt, J.K. (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables, *Insurance: Mathematics and Economics* **31**, 373–393.

Lee, R.D. and Carter, L.R. (1992). Modelling and Forecasting of U.S. Mortality, *Journal of the American Statistical Association* **87**, 659–675.