CREATING PORTFOLIO-SPECIFIC MORTALITY TABLES: A CASE STUDY

By S. J. Richards, K. Kaufhold & S. Rosenbusch

LONGEVITAS

CREATING PORTFOLIO-SPECIFIC MORTALITY TABLES: A CASE STUDY

By S. J. Richards, K. Kaufhold and S. Rosenbusch

25th May 2013

Abstract

Effective risk management of a portfolio demands accurate and succinct models which explain the main risk factors. Since portfolios have detailed individual records, an ideal approach is to use survival models. We look at a case study of how the administrator of a large multi-employer pension scheme created its own mortality tables. In addition to looking at statistical tests of fit, we consider a process for checking the suitability of a model for financial purposes. We also illustrate how a given scheme can test whether its experience is significantly different from other schemes, even after allowing for various known risk factors.

1 Introduction

1.1 A mortality table is a simple kind of deterministic model. The purpose of the mortality table is to replace volatile mortality-experience data with a set of smooth rates for use in pricing, reserving and other forms of risk management. Actuaries refer to the process of creating mortality tables as graduation. In this paper we will be concerned with the creation of mortality tables for pensioners and annuitants, i.e. for longevity risk, but the principles apply to any demographic risk.

1.2 Historically, there were several barriers to pension schemes and insurance companies creating their own mortality tables. The first barrier was statistical: there were often not enough pensioners to form a statistically credible data set, especially for pension schemes. The second barrier was resource-related: often there was not the time or expertise within an organization to perform such a specialized task. For these and other reasons, the creation of mortality tables was typically outsourced to a third party, who would not only have the necessary expertise, but who would also aggregate the data from several organizations to get statistical credibility. In the UK this role has been historically played by the CMI (see Continuous Mortality Investigation, 2004).

1.3 However, today the statistical barrier often no longer applies, as many portfolios now have enough data to create their own mortality tables. Furthermore, modern computing power allows actuaries to quickly fit models which would have been computationally very difficult to fit in the past (and practically impossible in the pre-computer era). This paper describes a case study of how a multi-employer pension-scheme administrator created its own mortality tables with a rich set of risk factors for longevity. The pension scheme is in the Federal Republic of Germany, but the methods illustrated here apply to any portfolio of demographic risk in any country.

1.4   The importance of modern risk-modelling techniques to insurers has increased following the ban in the European Union on the use of a person's gender for setting insurance prices. In the case of longevity risk, this creates an unusual situation: as a highly significant risk factor, gender must be used in reserving and internal risk-management, but it cannot be used in pricing. We also demonstrate a testing procedure for ensuring that a given model is suitable for financial purposes, even if it excludes a significant risk factor like gender.

2   Data

2.1   The data used in this paper cover a multi-employer pension arrangement in the Federal Republic of Germany. The data are 253,444 individual records of pensions in payment, administered centrally on the same computer system. In total there were 1,042,437.6 life-years lived over 2007-2011. A summary of the numbers of deaths and time lived is given in Table 1 for key age bands.

Table 1: Deaths and time lived for ages 30+ over the period 2007–2011. Source: Own calculations.

Age range	Deaths: Males	Females	Time lived (years): Males	Females
30–34	3	4	71.5	428.3
35–39	5	17	424.7	1,483.1
40–44	21	67	1,411.1	4,209.7
45–49	68	101	3,352.4	9,135.1
50–54	165	219	5,994.7	16,558.3
55–59	293	382	11,130.9	26,778.8
60–64	659	660	39,207.2	94,942.4
65–69	1,649	1,380	103,458.2	172,171.7
70–74	2,339	1,896	88,512.5	140,643.4
75–79	2,563	2,350	51,893.1	86,282.2
80–84	2,499	3,888	29,180.8	69,687.9
85–89	1,837	4,914	12,521.2	43,456.8
90–94	646	2,158	2,799.2	10,587.4
95–99	195	751	528.3	2,222.2
100–104	18	78	27.5	149.6
105+	1	6	0.2	7.0
Ages 30+	12,961	18,871	350,513.5	678,744.0

Note: If a single pensioner is observed from 1st January 2007 to the end of 2011 without dying, then that life contributes 5.0 years to the time lived shown above.

2.2   The data are for mandatory occupational pensions, so there is no element of individual choice in joining the scheme. The employers include municipalities, councils, church organizations and some private employers. The scheme provides a top-up pension to the German national pension, which means that small pensions will materially under-state the total retirement income received by a pensioner. The final pension is related to salary, and so the pension amount will be an imperfect proxy for overall income and wealth due to the top-up nature of the pensions here. In particular, one cannot distinguish between (a) someone receiving a modest pension because they earned a modest full-time salary, and (b) someone who earned the same pension by working part-time for a larger full-time-equivalent salary.

2.3   The scheme as a whole has less concentration risk than, say, pension schemes and annuity portfolios in the UK. For this data set the Gini coefficient is 43.4% (see Gini, 1921) whereas in Richards (2008) the same statistic was 60.9% for a data set of private-sector UK pension schemes and 66.0% for a UK life-office annuity portfolio. Another way of looking at concentration risk is shown in Table 2 — 50% of the total pensions paid are received by just 23.5% of pensioners.

2.4   Pension increases over the period of observation were exactly 1% per annum. This introduces a source of potential bias, because deaths and early exits will receive fewer increases than survivors to the end of the period of

observation. We therefore revalue the pension amounts for deaths and other exits from the date of exit to the end of 2011 at a rate of 1% per annum. This keeps the pension values of exits on a common footing with survivors.

Table 2: Revalued pension amounts by decile.

Pension decile	Total pension (€)	Percentage of total
1	6,754,727	0.7%
2	16,551,597	1.7%
3	31,779,949	3.3%
4	51,947,989	5.4%
5	73,724,219	7.6%
6	95,056,868	9.9%
7	116,929,120	12.1%
8	141,321,770	14.7%
9	172,876,147	17.9%
10	257,390,571	26.7%
All	964,332,957	100.0%

Note: For pensions stopping before end-2011, either due to death or early termination, the pension amount is revalued at a constant 1% per annum to that date.

2.5 Pension rights are transferable between employers in this scheme, and, when an employee retires, they are deemed to have retired from their last employer. Names and addresses were not available, so deduplication could not be performed along the lines described in Richards (2008). However, this was not felt to be a major problem, as the nature of the administration of this pension system — the transfer of pension rights between employers — made it unlikely that a given person would have multiple benefit records. This would not necessarily be the case for other portfolios, and the importance of deduplication for annuities is illustrated in Richards and Currie (2009), where the tendency to have multiple benefit records was strongly correlated with income and socio-economic group. Note that the absence of address data makes it impossible to model mortality by geodemographic profile, as was done for the UK data in Richards (2008).

2.6 However, we note that the data as supplied did contain one scenario where the same individual might be represented by two separate records. Where a pensioner has a pension in their own right as a former employee, and where they are also the surviving spouse of someone who was also an employee, then that person will have two records: one for their own pension, and a second for their widow(er)'s pension. However, this scenario was felt to be uncommon enough to not materially distort the results.

2.7 The data as supplied comprised a series of in-force extracts on the 1st of January of each of the years 2007–2011. Information was provided on new entrants and exits during each calendar year, but without details of the precise date of movement. New entrants who died before the end of their year of entry were not listed in the movements. This required the exclusion of the fractional part of the exposure time until the life reached the following 1$^{\text{st}}$ of January, as deaths before this time would be missing from the movement records.

2.8 There were 31,842 deaths observed over 2007–2011, excluding the deaths of new entrants during the year of entry. Not all recorded movements were deaths, however, as pensions could cease for a variety of other reasons: a child's pension would cease upon reaching a given age or finishing full-time education, while widow(er)'s pensions terminated upon remarriage. Also, some employers could move the administration of the pension to another administrator. Table 3 summarizes the various movements and their treatment for the construction of the survival models in Section 5. Note that in the administration system an ill-health early retirement pension changes status after the normal retirement age. For modelling mortality, such cases were kept marked as ill-health retirements throughout. However, due to an earlier system migration it is possible that some ill-health cases remain incorrectly categorized as normal-age retirements (this is suggested by the rapid reduction in exposures in Table 6 and the sparseness of data at older ages shown in Figure 4).

Table 3: Summary of movements and their treatment. The column d_i indicates how the movement is treated: 0 is for a censoring event and 1 is for a death, while a blank entry means that the movement does not constitute an event of interest (death) or a censoring event. Censoring events and deaths are presumed to occur in the middle of each calendar year on average.

Code	d_i	Count	Description
ZF		0	Deduplication of multiple insured numbers for a single person.
ABF		3	Annulled new entry from deferred status.
AEM		5	Annulled entry from ill-health status.
AB	0	13	Commutation.
UB	0	14	Transfer to another administrator.
ABZ		48	Annulled new entry from active status.
AG		72	Change of gender code.
KLW	0	77	End of temporary widow(er)'s pension.
WDR	0	273	Widow(er)'s pension ends with remarriage.
AGZ	0	287	Orphan's pension ends upon reaching adulthood.
ZBZ	0	324	Disability pension ends with return to contributing status.
ZBF	0	413	Disability pension ends without return to contributing status.
ESB	0	1,804	Orphan's pension ends upon completing education.
VAR		3,200	Disability pension changed to early-retirement pension.
AR		6,337	Ill-health disability pension converted to normal pension.
TOD	1	7,468	Death of widow(er).
TMH	1	9,061	Death of main pensioner leaving a surviving spouse.
TOH	1	15,314	Death of main pensioner without leaving a surviving spouse.

2.9   Age and calendar time are used directly as real-valued variables. The remaining potential risk factors are categorical or ordinal factors, and they are described in Table 4.

Table 4: Risk factors for the multi-employer pension scheme considered in this case study.

Risk factor	Levels	Description
Gender	Male	The gender of the pensioner.
	Female	
Largest scheme	Yes	Whether the pensioner belongs to the
	No	scheme with the largest number of lives.
Region	B	The region of the last employer.
	P	
Scheme type	1	Municipalities, councils and public-sector.
	2	Private businesses, churches.
Size	1	Those with the smallest 85% of revalued pensions.
	2	Those with the second-largest 10% of revalued pensions.
	3	Those with the largest 5% of revalued pensions.
Status	Normal-age retiral	Whether the pension recipient retired at
	Ill-health retiral	normal age, retired early due to ill health
	Widow(er)	or is a surviving spouse.

Note: The derivation of the pension-size risk factor is detailed in ¶6.3.

3   Initial exploratory data analysis

3.1   Figure 1 is a plot by age of crude mortality hazards on a logarithmic scale. Mortality hazards are approximately log-linear above age 65 or so. Between ages 60 and 65 there appears to be an essentially constant rate of mortality, suggesting a Makeham-type mortality law in place of a purely Gompertz law — see Gompertz (1825) and Makeham (1859). Mortality hazards under age 60 (not shown) are particularly high, due to the dominance of pensioners who have retired early due to ill-health or disability (see Section 4, especially Figure 4 and Table 6).

Figure 1: $\log_e$(crude mortality hazard) from age 60, males and females combined.
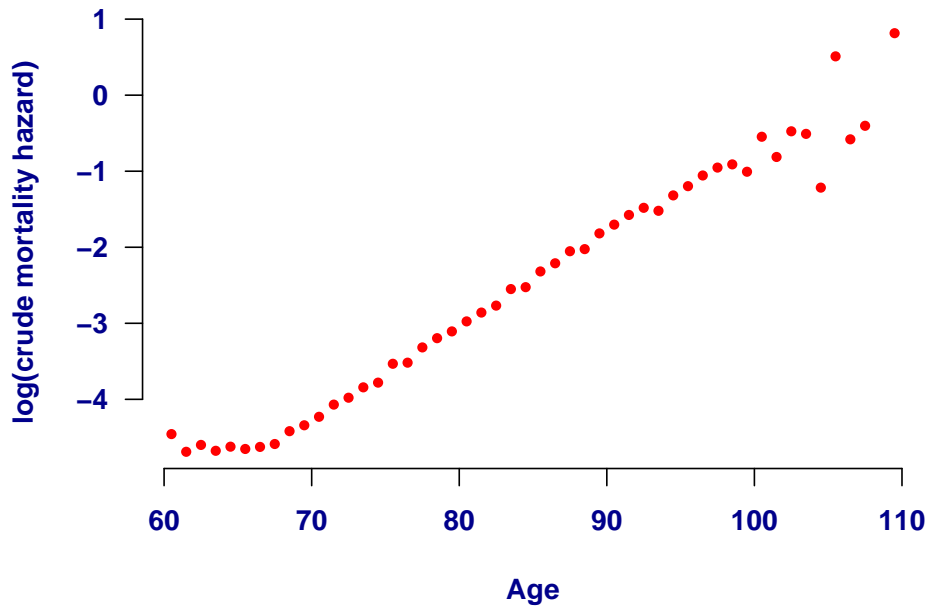
Note: The crude mortality hazard is the actual number of deaths in the age interval $[x,x+1)$ divided by the time lived in that interval.

3.2   A useful check on the data is to plot the survival curves proposed by Kaplan and Meier (1958) for males and females, as in Figure 2. This shows the shapes and patterns we expect for a modern population, with females having a higher survival probability at all ages. With some other data sets we have seen, the Kaplan-Meier curves show little distinction between males and females; this is usually a data error, often related to the processing of deaths and the setting up of pensions for surviving spouses. However, the clear separation between male and female survival curves in Figure 2 suggests that this particular corruption is not present in this data set.

3.3   Another useful check on the data is to compute the mortality ratios of actual deaths against the expected numbers according to a published table. Table 5 shows the ratios using a contemporaneous national mortality table, with the ratios weighted alternatively by lives and revalued annual pension. Mortality is slightly heavier than the national table, as the lives covered in this data set are municipal workers. The male ratios are proportionately heavier due to the occupations covered: policemen, firemen, foresters and the like. The female ratios are only slightly heavier due to the preponderance of office-based staff. As expected, mortality ratios weighted by annual pension size are lower than the lives-based measures, with the difference being more pronounced for males. Table 5 also shows why the administrator needs portfolio-specific mortality tables: the mortality hazards of the population table are not a good fit across the entire age range. While the excess mortality below age 65 is expected, the very light mortality experience at ages 65–69 also stands out.

4   The need for a parametric statistical model

4.1   One reason for wanting a mortality model is the strong differentials in survival between different sub-groups, and thus in the annuity factors which should be used in reserving or pricing. Figure 2 shows the

Figure 2: Kaplan-Meier product-limit estimators for males and females from age 60.
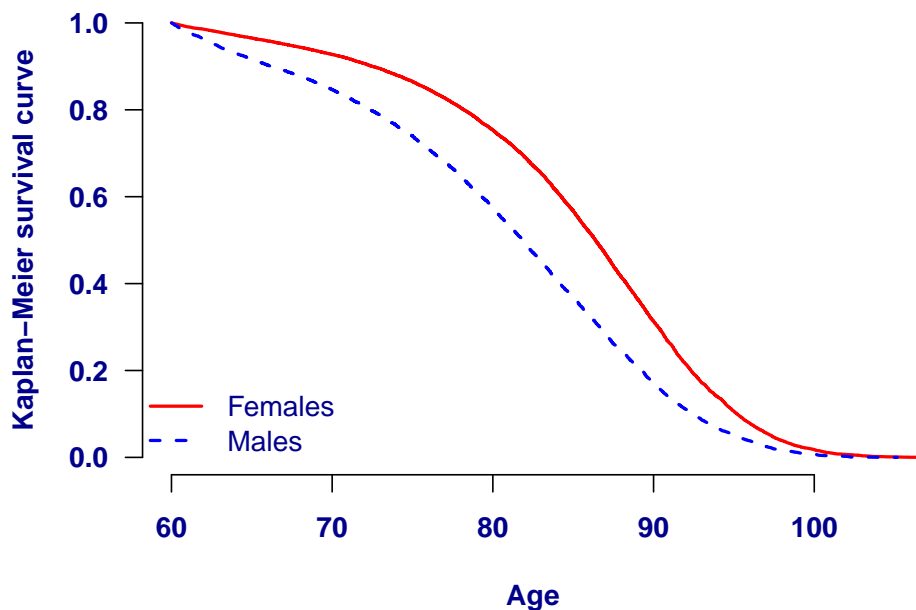
Note: The definition of the Kaplan-Meier estimator used here is that given in Richards (2012). The lines appear very smooth because of the large numbers of deaths — 12,406 for males and 18,081 for females — and the correspondingly small interval of time between events.

Table 5: Mortality ratios of actual deaths compared to expected deaths using German population tables for 2009–2011, Statistisches Bundesamt (2013). Source: Own calculations.

| Age range	Lives-weighted: Males	Females	Pension-weighted: Males	Females
50–54	528%	477%	605%	453%
55–59	317%	344%	307%	328%
60–64	134%	107%	121%	111%
65–69	87%	87%	78%	87%
70–74	96%	93%	86%	94%
75–79	104%	96%	96%	96%
80–84	105%	98%	102%	99%
85–89	105%	102%	104%	100%
90–94	101%	102%	96%	95%
95–99	105%	107%	106%	101%
50–99	103.9%	101.0%	95.9%	99.2%

Note: The denominator of the mortality ratio is calculated by summing $\mu_{x+\frac{1}{2}} E^c_{x+\frac{1}{2}}$, where $E^c_{x+\frac{1}{2}}$ is the central exposed-to-risk or time lived at average age $x+\frac{1}{2}$. $\mu_{x+\frac{1}{2}}$ is approximated from the published table by $-\log(1-q_x)$. Pensions to early terminations, both deaths and censored cases, are revalued at 1% per annum to the end of the investigation period.

significant differences between males and females, but the Kaplan-Meier curves in Figure 3 also show differences in survival rates by pension size-band.

4.2   Figures 2 and 3 suggest that both gender and pension size are important explanatory factors for mortality. However, this raises the problem of correlation: males account for 34.5% of lives overall, but 59.7% of lives in the largest pension-size category. How do we separately identify the impact of gender and pension size?

Figure 3: Kaplan-Meier product-limit estimators for the highest- and lowest-income males from age 60.
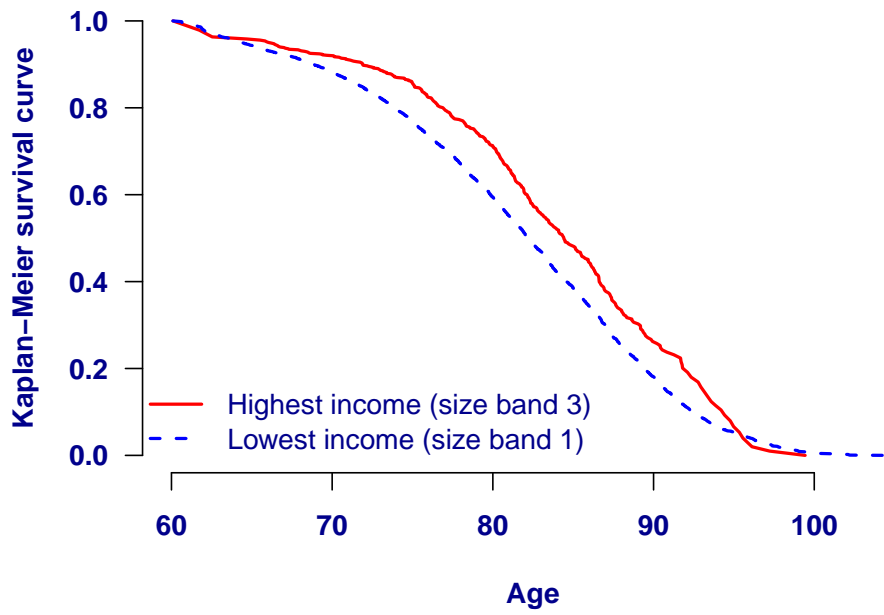
Note: The subgroups are size band 1 (smallest 85%) and 3 (largest 5%) of normal retirees in scheme type 1 in Region B. The definition of the Kaplan-Meier estimator used here is that given in Richards (2012). The lines appear less smooth than Figure 2 because of the smaller data volumes: 5,313 deaths for size-band 1 and 413 for size band 3 (see Table 8).

This is not merely a question of these two risk factors — Figure 4 shows that the health status at retirement is also an important explanatory variable for mortality. We therefore need a methodology which allows the simultaneous estimation of the impact of multiple risk factors.

Figure 4: $\log_e$(crude mortality hazard) from age 60 for ill-health and normal-age retirements.
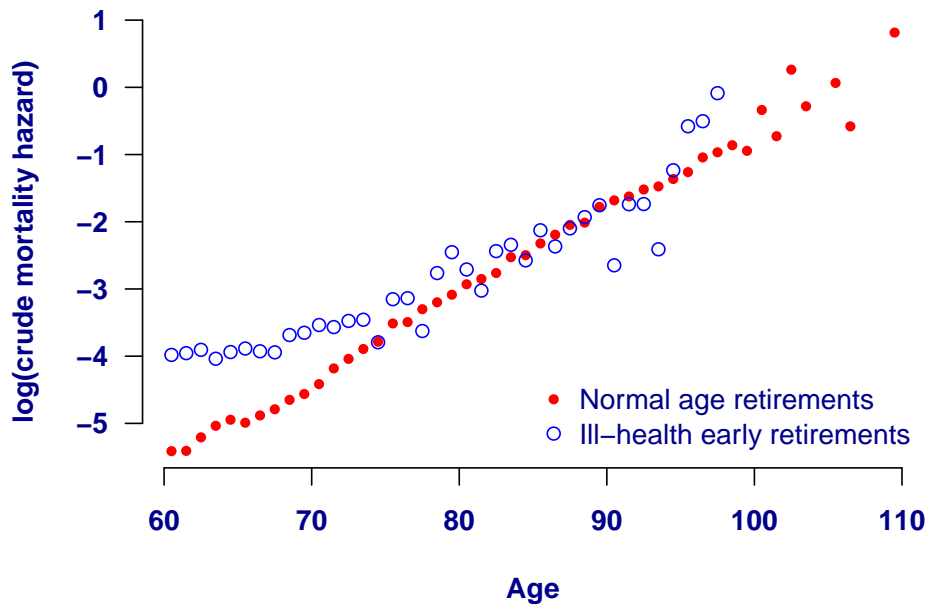
Note: The crude mortality hazard is the actual number of deaths in the age interval $[x, x+1)$ divided by the time lived in that interval.

4.3 A common-sense goal is that the greatest weight should be assigned to areas where we have the most data. However, the volumes of data are very different at different ages: Table 6 shows that ill-health retirals account for half of the time lived at age 60, but by age 75 they account for under 1%. We therefore need a methodology which automatically gives the greatest weight where the data are most numerous.

4.4 A related requirement is the need to ignore random variation, i.e. to smooth the fitted rates. For example, in Figure 1 we do not actually believe that the underlying mortality hazard at age $93\frac{1}{2}$ is truly lower than at age $92\frac{1}{2}$. Instead, it is far likelier that the observed reduction is a result of random volatility, and that the true underlying mortality hazard progresses relatively smoothly from one age to the next. By using a suitably chosen parametric formula, we can get automatically smoothed mortality hazards from the model which will iron out random variation.

4.5 Our final desire is to be able to sensibly extrapolate mortality hazards to older ages. Table 1 shows that there is very little information above age 100, while Figure 1 shows that there is no information at all above age 110. We therefore need a methodology which will give sensible fitted mortality hazards at higher ages. Fortunately, the same mechanism which provides the smoothing — a parametric formula — can also yield sensible extrapolated mortality hazards.

Table 6: Time lived between ages 60 and 76. The proportion of time exposed which comes from ill-health retirals varies dramatically over the age range. Source: Own calculations.

Age last birthday	Time lived (years): (a) All lives	(b) Ill-health retirals	Proportion of exposure from ill-health lives
60	14,659.8	7,337.7	50.1%
61	23,505.0	7,253.3	30.9%
62	26,246.9	7,149.9	27.2%
63	31,469.2	7,315.6	23.2%
64	38,256.8	7,546.8	19.7%
65	45,924.8	7,796.5	17.0%
66	55,315.5	8,172.0	14.8%
67	58,065.6	8,611.7	14.8%
68	58,418.6	8,728.4	14.9%
69	57,905.4	8,517.5	14.7%
70	55,225.0	7,106.8	12.9%
71	51,192.5	5,346.3	10.4%
72	45,914.3	3,584.9	7.8%
73	40,710.4	1,967.7	4.8%
74	36,113.7	621.4	1.7%
75	32,416.2	280.6	0.9%

4.6 A parametric statistical model will therefore fulfill all of our requirements: (i) the ability to simultaneously separate the impact of multiple risk factors, (ii) assigning the greatest weight to areas with the most data, (iii) smoothing out random variation and (iv) extrapolating rates to ages with little or no data. The only remaining questions are (a) whether the statistical model should be for grouped counts or individual lifetimes, and (b) whether to model the rate of mortality, q_x, or the force of mortality, μ_x. These last two questions are decided by the nature of the data, and are considered in the next section.

5 Model structure

5.1 Historical methods of graduation would often involve something like the application of a moving average to the crude mortality rates. However, as we saw in Section 4, a much better approach is to use a statistical model. Furthermore, historical methods would often separate the calculation of the mortality rates from the smoothing step. In contrast, a statistical model can use a smooth functional form for the mortality hazard, thus combining estimation and smoothing in a single stage. We therefore seek to build a statistical model for the mortality experience observed.

5.2   As we have individual records, it makes sense to model mortality at the level of the individual. It makes little sense to aggregate individual-level data for the purposes of modelling, as this will inevitably lose information. In ¶6.7 we will demonstrate why modelling grouped data is less effective than modelling individual data. The description of the data in ¶2.7 would initially seem to suggest a model for q-type annual mortality rates: we have in-force data at the start of each year, but data on the precise dates of death are not supplied. However, there are non-mortality decrements described in ¶2.8, which are not easily handled by q-type models without additional assumptions. These exits — both deaths and other terminations — are best handled by assuming that they contribute a half-year of exposure on average. This reality of so-called central exposures, not initial exposures, pushes us towards modelling the continuous-time force of mortality, μ_x, defined as follows:

$$\mu_x = \lim_{h \to 0^+} \frac{1}{h} \Pr(\text{death before age } x+h | \text{alive at age } x) \tag{1}$$

5.3   One further advantage of modelling μ_x is that it allows us to model on a simple logarithmic scale. If we were to use q-type rates, we would have to work on a scale which ensured the fitted rates lay within $(0,1)$, such as the logit transform. We can, of course, derive exact values for q_x if required, so long as we have full knowledge of μ_x.

5.4   To build our model, we first note that the probability of surviving from age x to age $x+t$, $_tp_x$, is given by:

$$_tp_x = e^{-H_x(t)} \tag{2}$$

where $H_x(t)$ is the integrated hazard function:

$$H_x(t) = \int_0^t \mu_{x+s} ds \tag{3}$$

5.5   Some implementations of survival models at the individual level deal with age-varying mortality through a variable transformation. This demands that the lives be observed from outset, i.e. from birth if chronological age is to be used directly. In contrast, people start life-insurance contracts or pension benefits when they are well into adult life. Such lifetimes are called *left-truncated*, since observation starts at age x_i and we have no data on deaths or exposure prior this age. Equally, when an extract of mortality data is taken not all lives will be dead at the extract date. Such data is called *right-censored*, since all that can be said of the mortality process is that it will occur some time after age x_i+t_i. This truncation/censorship model is illustrated in Figure 5.

Figure 5: Diagram of survival-model setup. The time observed, t_i, is shown in grey, while deaths are marked with a cross, ×. Since people do not usually enter into life-insurance contracts at birth, observations are left-truncated, i.e. lives start being observed at age $x_i > 0$. The upper case is an example of right-censored data, as death happens after the end of the observation period. Source: Richards (2008).
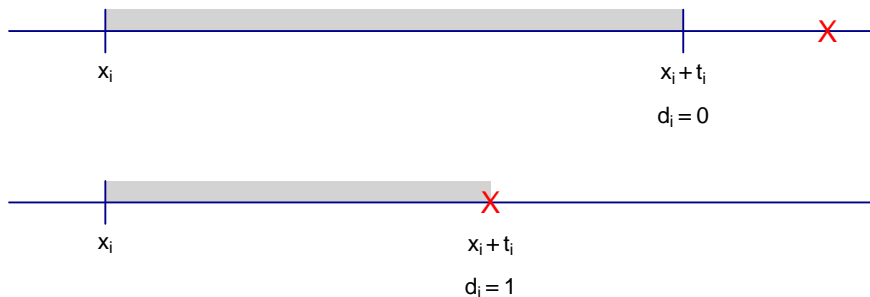
5.6   Since we have left-truncated lifetimes, we cannot fit our survival models by means of a variable transformation. We will therefore have to work directly with the likelihood function to fit our models. For each life i of n lives we have (i) an entry age, x_i, (ii) a time observed, t_i, and (iii) an indicator variable, d_i, for the state of the life at age x_i+t_i. We also have numerous other items of information on each life, as listed in Table 4. The variable d_i takes the value 0 on survival and 1 on the event of interest. This event can be death

(as in this case study) or any other decrement of interest, such as critical-illness claim, lapse or surrender. We are modelling the lifetime of an individual aged x, where the probability density function of the future lifetime, T_x, is $_tp_x\mu_{x+t}$. The likelihood function, L, is therefore given by:

$$L \propto \prod_{i=1}^{n} {}_{t_i}p_{x_i}\mu_{x_i+t_i}^{d_i} \tag{4}$$

and taking the natural logarithm of L gives us the log-likelihood function, ℓ:

$$\ell = \sum_{i=1}^{n} -H_{x_i}(t_i) + \sum_{i=1}^{n} d_i \log \mu_{x_i+t_i} \tag{5}$$

5.7   Thus, when applying survival models to individual data, it simply suffices to specify the structure of the hazard rate, μ_x, and subsequently derive $H_x(t)$. When fitting any model, we choose the set of parameter values which jointly maximizes the log-likelihood function in Equation 5.

5.8   Richards (2008) and Richards (2012) detail the functional forms of μ_x and $H_x(t)$ for a wide variety of models. The one we will focus on in this paper is a time-varying version of the Makeham-Beard law:

$$\mu_{x,y}^{i} = \frac{e^{\epsilon_i} + e^{\alpha_i + \beta_i x + \delta(y-2000)}}{1 + e^{\alpha_i + \rho_i + \beta_i x + \delta(y-2000)}} \tag{6}$$

which proved to be the best-fitting of the sixteen survival models in Richards (2012), as measured by the AIC (see Akaike, 1987).

5.9   The structure of the formula in Equation 6 is such that the force of mortality for life i at age x and year y, $\mu_{x,y}^{i}$, is always positive. The parameters are free to vary along the real line, however, giving us a regression model. The term e^{ϵ_i} is a constant rate of mortality independent of age or time, and is referred to as Makeham's constant after Makeham (1859). α_i is the overall mortality level for individual i, while β_i is the rate of increase in mortality with age for life i. ρ_i is a heterogeneity parameter, whose derivation and role is discussed in more detail in Richards (2008). δ models the constant rate of change in overall mortality levels with time, and is normalized by using $(y-2000)$ instead of y to keep the other parameters well scaled; the choice of 2000 is somewhat arbitrary, and a case could be made for using the mid-point of the period spanned by the data. The logistic nature of the function, $\frac{e^{\bullet}}{1+e^{\bullet}}$, was suggested by Perks (1932) and ensures that the mortality hazard increases approximately linearly on a logarithmic scale until advanced ages, whereupon the rate of increase slows. This logistic form was modified by Beard (1959) to include the heterogeneity parameter, ρ_i. Richards (2008) shows how the time-stationary version of the Makeham-Beard law in Equation 6 arises for individuals themselves following a Makeham law with heterogeneity in e^{α} following a gamma distribution (known as a frailty model).

5.10   For the models fitted in this paper, the ith individual has their own values of α_i, β_i, ρ_i and ϵ_i. δ will be constant across all sub-groups. The value of α_i is built up as follows:

$$\begin{aligned}
\alpha_i = \alpha_0 &+ \alpha_{male} z_{i,male} \\
&+ \alpha_{largest\ scheme} z_{i,largest\ scheme} \\
&+ \alpha_{region\ P} z_{i,region\ P} \\
&+ \alpha_{scheme\ type\ 2} z_{i,scheme\ type\ 2} \\
&+ \alpha_{size\ 2} z_{i,size\ 2} \\
&+ \alpha_{size\ 3} z_{i,size\ 3} \\
&+ \alpha_{ill-health} z_{i,illhealth} \\
&+ \alpha_{widow(er)} z_{i,widow(er)}
\end{aligned} \tag{7}$$

where α_0 is the baseline, which will be for a female who is (i) not in the largest scheme, (ii) in Region B, (iii) of scheme type 1, (iv) in the lowest pension size-band, and (v) a retiree at normal age and not a widow. The

values of the other α_j represent the main effects of being male, being a member of the largest scheme, etc. The $z_{i,j}$ are indicator variables taking the value 0 when the life i does not belong to group j, and the value 1 when life does belong to the group. By construction, two sets of indicators are mutually exclusive: one cannot be simultaneously in size band 2 and in size band 3, for example, and one cannot be both an ill-health retiral and a widow(er). Note that the latter is not strictly true in reality, but will at least be true for the records here bearing in mind the comments in ¶2.6. Where age interactions are present, β_i is built up analogously to Equation 7 and using the same indicator variables, $z_{i,j}$. ϵ_i is treated similarly, and represents different constant Makeham parameters for some sub-groups. ρ_i is handled analogously. As we will see in Table 9, not all risk factors provide strong enough evidence to justify varying β, ρ or ϵ by sub-group, even where there is evidence for varying α.

5.11 In presenting model results, we will denote α_0 as the intercept and β_0 as the coefficient of age. Similarly, for the α_j we will use the names and levels in Table 4. Further background reading on survival models can be found in Collett (2003), with additional actuarial considerations covered in Richards (2008) and Richards (2012).

6 A model with main effects only

6.1 We start with a simple model with the main effects only, i.e. β, ρ, δ and ϵ are constant across all lives. The only difference between the various lives is their categorization, and therefore their level of mortality (as represented by α_i defined in Equation 7). The force of mortality for life i, $\mu_{x,y}^i$, is therefore given by the following formula:

$$\mu_{x,y}^i = \frac{e^\epsilon + e^{\alpha_i + \beta x + \delta(y-2000)}}{1 + e^{\alpha_i + \rho + \beta x + \delta(y-2000)}} \tag{8}$$

which shows that only α is varying between lives.

6.2 The formula in Equation 8 is plugged into the log-likelihood function in Equation 5, which is then optimized to yield the maximum-likelihood estimates of the various parameters. The variance-covariance matrix of the parameters is estimated from inverting the matrix of all negative second partial derivatives, and from the leading diagonal we can get estimated variances of the parameter estimates. The results are shown in Table 7.

Table 7: Model with main effects only, as per Equation 8. AIC is 238,366.

Parameter	Estimate	Standard error	Z-value	p-value	Significance level	Lives	Deaths
Age (β_0)	0.150817	0.0033	45.54	0	***	231,641	30,487
Beard (ρ_0)	0.427666	0.0928	4.61	0	***	231,641	30,487
Gender.M (α_{male})	0.629587	0.0211	29.81	0	***	80,789	12,406
Intercept (α_0)	-15.1662	0.2704	-56.09	0	***	231,641	30,487
Largest scheme.Y	-0.119731	0.0331	-3.62	0.0003	***	12,392	1,480
Makeham (ϵ)	-6.30107	0.1904	-33.09	0	***	231,641	30,487
Region.P	0.12165	0.0215	5.67	0	***	24,257	3,832
Scheme type.2	-0.0938358	0.0171	-5.48	0	***	73,827	6,766
Size.2	-0.0859703	0.0234	-3.67	0.0002	***	23,167	2,962
Size.3	-0.220652	0.0324	-6.82	0	***	11,562	1,477
Status.Ill-health	0.94188	0.0345	27.27	0	***	25,354	2,393
Status.Widow(er)	0.155541	0.0183	8.50	0	***	36,687	7,374
Time (δ)	-0.0132796	0.0049	-2.72	0.0064	**	231,641	30,487

Note: Parameter significance is labelled according to the same scheme used in R [R Development Core Team (2011)], i.e. . for 10%, * for 5%, ** for 1% and *** for 0.1%.

6.3 One aspect to note about the size-band factor is that it is an optimization of the vigintiles. The revalued pensions are sorted and grouped into twenty ordinal categories of approximately equal numbers of lives in each. To use this category directly as a risk factor would require nineteen parameters: the twenty categories less one for the baseline. A more practical size-band factor would only involve, say, three bands and thus two parameters. To select the grouping of vigintiles into the three-level size-band factor, we try out each pair of break-points in turn and compare an information criterion such as the AIC (see Akaike, 1987). The assignment which produces the lowest AIC is selected as the best-fitting definition of the size bands. In Table 7 the baseline is the smallest 85% of pensions, while the second size-band contains the next largest 10% and the third size-band contains the largest 5%. This is the same size-band optimization as used in Richards (2008). Note that it is an ordinal optimization, i.e. if vigintiles 1 and 5 are in the same size-band, then vigintiles 2, 3 and 4 must also be.

6.4 All the parameters in Table 7 are significant, albeit the time trend is less strong in both magnitude and significance than all the other risk factors. As a general rule, the time-trend parameter in this kind of modelling has more uncertainty over its value than most other parameters. Note that the parameters in Table 7 — and elsewhere in this paper — are partial regression coefficients. They are dependent parameters in the sense that each parameter value in Table 7 is dependent on the presence, action and value of all the other parameters. They are also dependent on the structure of the model, i.e. the choice of mortality law in Equation 6. As will be seen in Table 9, parameter values and their interpretation can change dramatically, even where they bear the same name. A particular point to note is that parameter estimates are often strongly correlated with each other, which can be investigated by inspecting the variance-covariance matrix described in ¶6.2. These are general features of regression-type models, and they also apply to other models such as Generalized Linear Models (see McCullagh and Nelder, 1989).

6.5 It is tempting to read off the parameter values in Table 7 as being approximate percentage differences. For example, the age parameter of 0.150817 could be interpreted as meaning that the mortality hazard is increasing by around 15% for each additional year of age. Furthermore, males have a mortality hazard around 63% higher than females (females form the baseline category, so a single parameter is required for gender as a factor). The departure from the baseline, i.e. being male, causes an extra 0.629587 to be added to α. Similarly, pensioners of employers in Region P have around 12% higher mortality, whereas widow(er)s have around 15–16% higher mortality. The precise percentages vary by age, and can be calculated exactly using Equation 6 if need be. However, even in a simple model the isolated interpretation of a parameter value can be misleading. Many of the parameters in even this simple model are strongly correlated, which complicates the interpretation of even the most straightforward-seeming values. This is discussed in more detail in Appendix 2, but a consequence is that one should be cautious about interpreting a parameter in isolation of the others on whose values it depends.

6.6 Table 7 also shows a considerable benefit of modelling mortality at the level of the individual instead of modelling grouped counts. All the risk factors in Table 7 are significant, so we would want to use them all in a mortality model. However, an alternative model based on grouped counts would require 90 different cells for each combination of risk factors ($90 = 72 + 18 =$ two genders $\times$ three status categories $\times$ three pension-size bands $\times$ two regions $\times$ two scheme types, plus for the largest city scheme a further two genders $\times$ three status categories $\times$ three pension-size bands). Even for a large data set such as this, this would involve a lot of cells with very few or zero deaths, as demonstrated in Table 8.

6.7 A contingency-table model for Poisson-distributed grouped counts would require a minimum number of, say, five expected deaths per cell. Table 8 shows that 21 out of 90 cells would be below this minimum for a model for grouped counts. The usual response to this would be to collapse the categories to achieve the minimum cell count, but this would be at the expense of several of the risk factors one would want to use. However, this constraint does not apply to survival models at the level of the individual, which is why they are preferable wherever there is rich detail to be considered.

7 A model with second-order interactions

7.1 The model in Table 7 assumes that mortality differentials are constant, e.g. that males have an extra 0.629587 units of mortality on a logarithmic scale whatever their age. However, mortality differentials are seldom constant — Richards (2012) showed that the gender differential varied strongly by age for a portfolio of life-office annuitants, while Figure 4 shows that the additional mortality for ill-health retirals also reduces with increasing age. We therefore want to capture this non-constant feature in our model.

Table 8: Deaths categorized by risk-factor combination for model in Table 7.

Member of largest scheme	Region	Scheme type	Pension size-band	Normal retirees: Females	Males	Ill-health retirees: Females	Males	Widow(er)s: Females	Males
No	B	1	1	5,142	5,313	525	738	4,434	618
			2	824	725	39	98	36	0
			3	282	413	14	33	24	1
		2	1	2,200	1,323	308	183	628	222
			2	305	275	20	39	18	0
			3	140	206	15	18	15	1
	P	1	1	695	811	51	99	798	89
			2	138	122	7	22	9	0
			3	59	72	1	5	3	1
		2	1	174	274	26	33	166	23
			2	26	56	3	4	4	0
			3	8	41	5	2	5	0
Yes	B	1	1	480	338	41	45	224	47
			2	108	65	12	3	4	0
			3	60	45	1	3	4	0
Totals				10,641	10,079	1,068	1,325	6,372	1,002

7.2 To improve the model, therefore, we can consider interacting the main effects in Table 7 with age. We can also consider interactions with the Beard (ρ) and Makeham (ϵ) parameters. Not all of the resulting new parameters are significant, however, so we drop the interactions which do not improve the model. We use the information criterion from Akaike (1987), the AIC, as our decision-making metric as to whether parameters sufficiently improve the log-likelihood to justify being kept in the model.

7.3 A process of iteration using the AIC as a target yields the final model as shown in Table 9. The first point to note is that the model with the various interactions is a better fit than the model in Table 7: the AIC has reduced by 601 units, suggesting that the new model is a material improvement. This is also reflected in the fact that the new age-, Beard- and Makeham-interaction parameters are highly significant, telling us that mortality differentials vary strongly enough by age and subgroup that this needs to be reflected in the model.

7.4 Table 9 demonstrates some common features of interactions with age. The first point of note is that the sign of an age interaction is usually the opposite of the sign of the main effect. This is the phenomenon of *mortality convergence*, i.e. the tendency for mortality differentials to narrow with increasing age. The second, related aspect is that the larger the absolute size of the main effect, the larger the absolute size of the interaction with age. This means that the rate of mortality convergence is faster for larger initial differentials. This is demonstrated in Table 9, where the excess mortality of Region P reaches zero around age 92–93 ($|\frac{0.943612}{-0.0102169}| = 92.36$ years). This phenomenon of mortality convergence around age 90–95 is common to mortality differentials in many other pensioner and annuitant data sets we have modelled. However, there can be considerable uncertainty over the true age of convergence: for example, the ill-health mortality excess reaches zero around age 83–84, as confirmed by Figure 4. As discussed in ¶4.3, the data available at older ages is very sparse for ill-health retirements, so there would be particular uncertainty over whether this was the true convergent age for ill-health differentials. The phenomenon of mortality convergence is discussed in Gavrilov and Gavrilova (2001).

Table 9: Final model with main effects and selected interactions with age, the Beard parameter and the Makeham parameter. AIC is 237,765.

Parameter	Estimate	Standard error	Z-value	p-value	Significance level	Lives	Deaths
Age (β_0)	0.174283	0.004	43.94	0	***	231,641	30,487
Beard (ρ_0)	0.247414	0.1086	2.28	0.0227	*	231,641	30,487
Gender.M (α_{male})	3.53738	0.2179	16.23	0	***	80,789	12,406
Gender.M:Age (β_{male})	-0.0361882	0.0026	-13.77	0	***	80,789	12,406
Intercept (α_0)	-17.1472	0.3256	-52.66	0	***	231,641	30,487
Largest scheme.Y	-0.136226	0.0355	-3.84	0.0001	***	12,392	1,480
Makeham (ϵ_0)	-5.96492	0.1135	-52.53	0	***	231,641	30,487
Region.P	0.943612	0.2375	3.97	0.0001	***	24,257	3,832
Region.P:Age	-0.0102169	0.0029	-3.47	0.0005	***	24,257	3,832
Scheme type.2	-0.104973	0.0182	-5.76	0	***	73,827	6,766
Size.2	-3.05163	0.4981	-6.13	0	***	23,167	2,962
Size.2:Age	0.0395201	0.0067	5.91	0	***	23,167	2,962
Size.2:Beard	0.944189	0.1588	5.94	0	***	23,167	2,962
Size.3	-3.17912	0.761	-4.18	0	***	11,562	1,477
Size.3:Age	0.0396146	0.0102	3.89	0.0001	***	11,562	1,477
Size.3:Beard	1.09962	0.207	5.31	0	***	11,562	1,477
Status.Ill-health	4.4203	0.4824	9.16	0	***	25,354	2,393
Status.Ill-health:Age	-0.0531685	0.0064	-8.32	0	***	25,354	2,393
Status.Ill-health:Makeham	1.15514	0.135	8.56	0	***	25,354	2,393
Status.Widow(er)	1.41502	0.3528	4.01	0.0001	***	36,687	7,374
Status.Widow(er):Age	-0.0159419	0.0041	-3.91	0.0001	***	36,687	7,374
Status.Widow(er):Makeham	0.373138	0.2192	1.70	0.0887	.	36,687	7,374
Time (δ)	-0.0110258	0.0051	-2.16	0.0309	*	231,641	30,487

Note: Parameter significance is labelled according to the same scheme used in R [R Development Core Team (2011)], i.e. . for 10%, * for 5%, ** for 1% and *** for 0.1%.

8   Testing the quality of fit

8.1   We have used the AIC to select the best-fitting model in Table 9, and now we must check the quality of fit. For actuarial work, quality of fit needs to be tested both in statistical and financial terms. We first use some statistical tests, starting with the χ^2 goodness of fit test. As it happens, the model in Table 9 manages to fail this test, and the reason can be seen in the plot of deviance residuals — see McCullagh and Nelder (1989) — in Figure 6 (the χ^2 test statistic is the sum of squared deviance residuals). Figure 6 shows that there are seven residuals which are larger than one would ordinarily expect if they were truly drawn from a N(0,1) distribution. One reason for this might be an unexplained source of variation, or else over-dispersion arising from not being able to deduplicate the data (see ¶2.6). However, a modest amount of over-dispersion will not affect the parameter estimates too badly, and the residuals in Figure 6 pass (i) a runs test, (ii) a lag-1 autocorrelation test, (iii) a standard-deviations test, and (iv) a bias test. Visually, we are not too unhappy with the quality of fit for the model in Table 9, despite some over-large residuals.

8.2   We can also plot the fitted survival curves against the Kaplan-Meier empirical survival curves, as in the example in Figure 7. This shows that the model is producing an acceptable smoothed equivalent of the actual survival curve.

Figure 6: Deviance residuals by age for model in Table 9, together with 95% lines for N(0,1) variates
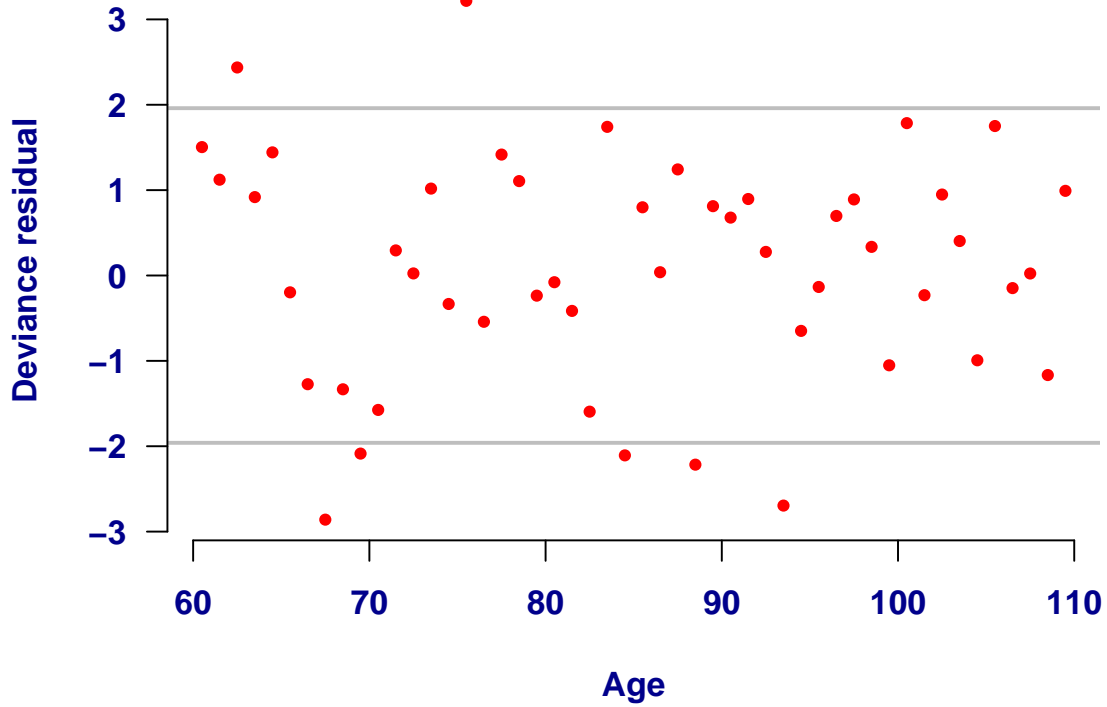
Note: Poisson deviance residuals are calculated using the observed deaths in the interval $[x, x+1)$ and a Poisson parameter of $\sum_i H_{x_i}(t_i)$, according to the result from Cox and Miller (1987, page 153). Each individual contributes at most one year of exposure to each age interval.

8.3 We have seen in the preceding paragraphs and figures that the model is broadly acceptable from a statistical perspective, i.e. where each life has equal weight. However, not all lives are equal in their financial impact: all other things being equal, a pensioner receiving €10,000 per annum has ten times the financial importance of a pensioner receiving €1,000 per annum. Thus, one area where actuaries' needs differ from those of statisticians is testing the *financial* suitability of a model, in addition to its statistical acceptability. This can be done through a process of repeated sampling to test whether the model not only predicts variation in the number of deaths, but also in the amounts-weighted number of deaths. We therefore use a process of bootstrapping, as documented in Richards (2008), where we (i) randomly sample 10,000 records with replacement, (ii) use the model in Table 9 to predict the number of deaths and the pension amounts ceasing due to death, and (iii) calculate mortality ratios of the actual number of deaths and pension amounts ceasing compared to the model's predicted numbers. We repeat this process 1,000 times and get the results given in Table 10. This shows that the model accurately predicts the number of lives dying, although this is what we would expect for a rich model whose parameters were maximum-likelihood estimates. We can also see that the model tends to slightly over-state mortality levels weighted by pension amount, i.e. by up to 1%. Depending on the application, this slight over-statement could be ignored, or it could be dealt with by applying a factor of 0.991 to the final mortality hazards. Alternatively, one could do further work on the model to get the bootstrapped A/E ratios closer to 100% for both lives- and amounts-weighted measures. A further option, discussed by Richards (2008), is to weight the log-likelihood in Equation 5 by pension size, but this should be viewed very much as a last resort when all other approaches fail.

9 Scheme-specific effects

9.1 The advantage of pooling data is the greater statistical power it gives over a stand-alone scheme model. One drawback of this is that there are often relevant scheme-specific mortality features which can be lost through pooling. However, fitting a statistical model with the scheme itself as a risk factor allows the best of both worlds: the pooled data set brings greater power in measuring things like age- and gender-related mortality, while the

Figure 7: Kaplan-Meier survival curve for high-income males from Figure 3, together with fitted survival curve from the model in Table 9 at $y = 2009.5$, i.e. the approximate mid-point of the investigation period.
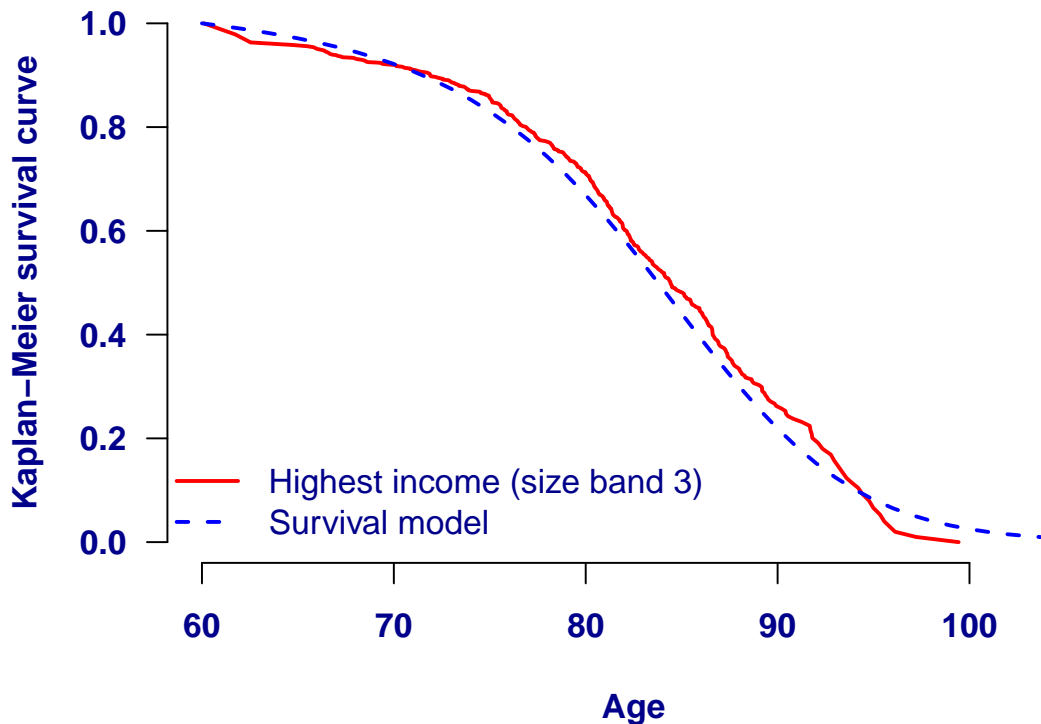

Table 10: Bootstrapping results for model in Table 9. Ratios of actual versus expected mortality from 1,000 samples with replacement of 10,000 lives from the data set.

Mortality measure	Median	Mean
Lives-weighted	100.0%	99.9%
Amounts-weighted	99.1%	99.2%

use of scheme identifier as a factor allows the formal testing of whether there are mortality dynamics in the scheme which are not accounted for by the standard risk factors.

9.2   In Tables 7 and 9 we see that the largest scheme has lighter mortality than can be explained by the other seven risk factors. This is perhaps to be expected: the largest scheme in question is for the employees of a large and wealthy city with a notably high standard of living. The inclusion of a scheme-specific identifier allows this scheme's mortality characteristics to be identified as statistically significant, and managed accordingly.

10   Conclusions

10.1   This case study has shown that it is straightforward for large portfolios to graduate their own mortality tables for pricing and reserving. This reduces the basis risk from using tables or models calibrated using other data, and ensures that the shape of the tables matches the experience of the portfolio's own risks. The methodology described here enables the use of more risk factors than typically found in published mortality tables, thus improving the potential to manage risk. The use of a parametric statistical model allows the separate identification of the effect of various risk factors, while automatically smoothing and extrapolating rates to higher ages. Modelling mortality at the level of the individual permits a wider range of risk factors to be considered than would be the case with a model for groups.

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrica 52*, 317–333.

Beard, R. E. (1959). Note on some mathematical mortality models. In G. E. W. Wolstenholme and M. O'Connor (Eds.), *The Lifespan of Animals*, pp. 302–311. Little, Brown, Boston.

Collett, D. (2003). *Modelling Survival Data in Medical Research*. Chapman and Hall.

Continuous Mortality Investigation (2004). *Continuous Mortality Investigation Report No. 21*. Institute of Actuaries and Faculty of Actuaries.

Cox, D. R. and H. D. Miller (1987). *The Theory of Stochastic Processes*, Volume 134 of *Science Paperbacks*. Chapman and Hall.

Currie, I. D. and A. Korabinksi (1984). Some comments on bivariate regression. *The Statistician 33*, 283–293.

Gavrilov, L. A. and N. S. Gavrilova (2001). The reliability theory of aging and longevity. *Journal of Theoretical Biology 213*, 527–545.

Gini, C. (1921). Measurement of inequality of incomes. *The Economic Journal 31(121)*, 124–126.

Gompertz, B. (1825). The nature of the function expressive of the law of human mortality. *Philosophical Transactions of the Royal Society 115*, 513–585.

Kaplan, E. L. and P. Meier (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association 53*, 457–481.

Longevitas Development Team (2013). *Longevitas v2.7 User Guide*. Longevitas Ltd, Edinburgh, United Kingdom.

Makeham, W. M. (1859). On the law of mortality and the construction of annuity tables. *Journal of the Institute of Actuaries 8*, 301–310.

McCullagh, P. and J. A. Nelder (1989). *Generalized Linear Models*, Volume 37 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, London (2nd edition).

Perks, W. (1932). On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries 63*, 12–40.

R Development Core Team (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, http://www.r-project.org.

Richards, S. J. (2008). Applying survival models to pensioner mortality data. *British Actuarial Journal 14(II)*, 257–326 (with discussion).

Richards, S. J. (2012). A handbook of parametric survival models for actuarial use. *Scandinavian Actuarial Journal 2012 (4)*, 233–257.

Richards, S. J. and I. D. Currie (2009). Longevity risk and annuity pricing with the Lee-Carter model. *British Actuarial Journal 15(II) No. 65*, 317–365 (with discussion).

Richards, S. J. and G. L. Jones (2004). *Financial aspects of longevity risk*. Staple Inn Actuarial Society (SIAS), London.

Statistisches Bundesamt (2013). *Sterbetafel Deutschland 2009–2011*. Statistisches Bundesamt.

11   Appendix 1: Life expectancies and annuity factors

11.1   The model in Table 9 is relatively rich with rating factors. In all there are 144 combinations of risk factors: 2 genders, 3 statuses, 2 regions, 2 scheme types, 3 size-bands and a binary flag for membership of the largest scheme. However, since the largest scheme is from a specific region and has a given scheme type, in practice there are only 90 different combinations of risk factors. Table 11 shows the approximate complete life expectancies according to the model in Table 9 for each of the various risk-factor combinations. The range of life expectancies is wide: from 13.83 years to 23.00 years. The equivalent life expectancies using a contemporaneous national mortality table — Sterbetafel Deutschland 2009–2011, Statistisches Bundesamt (2013) — are 17.47 years for males and 20.66 years for females.

Table 11: Complete life expectancies at age 65 in 2012 according to model in Table 9.

Member of largest scheme	Region	Scheme type	Pension size-band	Normal retirees: Females	Normal retirees: Males	Ill-health retirees: Females	Ill-health retirees: Males	Widow(er)s: Females	Widow(er)s: Males
No	B	1	1	21.24	17.71	19.31	14.92	20.33	16.47
			2	21.58	18.61	19.71	16.18	20.76	17.56
			3	22.42	19.52	20.58	17.14	21.61	18.49
		2	1	21.77	18.32	19.93	15.62	20.90	17.11
			1	22.02	19.12	20.21	16.74	21.22	18.09
			3	22.86	20.03	21.07	17.70	22.08	19.01
	P	1	1	20.65	16.83	18.55	13.83	19.67	15.51
			2	21.07	17.89	19.09	15.31	20.20	16.78
			3	21.93	18.82	19.98	16.29	21.07	17.72
		2	1	21.21	17.47	19.20	14.54	20.26	16.17
			2	21.53	18.42	19.61	15.88	20.68	17.32
			3	22.39	19.35	20.50	16.87	21.56	18.27
Yes	B	1	1	21.93	18.51	20.12	15.83	21.07	17.30
			2	22.16	19.27	20.36	16.91	21.36	18.24
			3	23.00	20.19	21.22	17.87	22.21	19.17

Note: Approximation of complete life expectancy using $0.5+\sum_{t=1}^{\infty} {}_t p_{65}$. Period life expectancies for 2012, i.e. with no allowance for future improvements.

11.2   Table 12 shows the corresponding annuity factors for each of the 90 risk-factor combinations. Once again, the range is very wide: from 10.599 to 16.181. The equivalent annuity factors using the mortality rates to age 100 in Sterbetafel Deutschland 2009–2011, Statistisches Bundesamt (2013), are 13.04 for males and 14.93 for females.

11.3   In Richards and Jones (2004) an equivalent-annuity calculation was performed to express bespoke mortality bases in terms of a published table. We can also do this for the model in Table 9, and we can calculate the equivalent percentages of the population table which produce the same annuity factors in Table 12. These equivalences are given in Table 13.

Table 12: Annuity factors at age 65 in 2012 according to model in Table 9.

Member of largest scheme	Region	Scheme type	Pension size-band	Normal retirees: Females	Normal retirees: Males	Ill-health retirees: Females	Ill-health retirees: Males	Widow(er)s: Females	Widow(er)s: Males
No	B	1	1	15.300	13.229	13.964	11.308	14.744	12.435
			2	15.482	13.771	14.238	12.131	14.987	13.106
			3	15.886	14.248	14.664	12.655	15.404	13.603
		2	1	15.581	13.580	14.291	11.717	15.043	12.808
			1	15.715	14.057	14.501	12.457	15.233	13.408
			3	16.114	14.529	14.920	12.974	15.644	13.900
	P	1	1	14.955	12.681	13.504	10.599	14.350	11.821
			2	15.190	13.328	13.869	11.576	14.660	12.615
			3	15.610	13.826	14.317	12.125	15.096	13.134
		2	1	15.253	13.051	13.854	11.025	14.668	12.213
			2	15.436	13.629	14.148	11.916	14.920	12.932
			3	15.851	14.121	14.590	12.459	15.350	13.445
Yes	B	1	1	15.664	13.684	14.387	11.839	15.131	12.919
			2	15.784	14.142	14.579	12.554	15.306	13.498
			3	16.181	14.612	14.996	13.069	15.715	13.988

Note: Cashflows paid continuously and discounted at 3% per annum. Approximation of continuously paid annuity factor using $0.5 + \sum_{t=1}^{\infty} v^t \, _tp_{65}$, where v^t is the discount function for t years. Period mortality rates for 2012, i.e. with no allowance for future improvements.

12 Appendix 2: Covariance and correlation

12.1 In ¶6.4 we discussed how the parameter values in a statistical model are dependent on each other. Using the full variance-covariance matrix we can calculate the extent to which parameter estimates are correlated with each other by dividing the covariance of two parameters by their respective standard errors. The resulting correlation lies between -100% (for perfect negative correlation) and +100% (for perfect positive correlation), with 0% signifying no correlation. Table 14 shows that quite a number of parameter values in the model in Table 7 are highly correlated. For example, the correlation between the Age and Intercept parameters is -98%, showing that a reduction in the Age parameter would cause an increase in the Intercept (and vice versa). These strong correlations have important consequences for the assessment of mis-estimation risk, i.e. the uncertainty over the best-estimate mortality rates, but this is beyond the scope of this case study.

12.2 Currie and Korabinksi (1984) discuss features of multivariate regression, including concealment (where the significance of a risk factor is hidden until another risk factor is included in the model) and enhancement (where the inclusion of a new risk factor improves the explanatory power of an existing risk factor). We are aware of instances where the inclusion of a new risk factor can even change the sign of an existing risk factor, which demonstrates the danger of drawing strong conclusions from a dependent parameter in isolation.

Table 13: Equivalent percentages of Sterbetafel Deutschland 2009–2011 to model in Table 9 at age 65 in 2012.

Member of largest scheme	Region	Scheme type	Pension size-band	Normal retirees: Females	Males	Ill-health retirees: Females	Males	Widow(er)s: Females	Males
No	B	1	1	90.9	95.7	129.0	148.9	105.2	114.9
			2	86.6	84.5	120.1	123.2	98.7	98.5
			3	77.9	75.8	107.4	109.2	88.4	87.9
		2	1	84.4	88.3	118.5	135.5	97.2	105.4
			2	81.5	79.2	112.1	114.3	92.5	91.9
			3	73.4	71.1	100.4	101.5	83.0	82.1
	P	1	1	99.5	108.6	145.4	175.6	116.6	132.3
			2	93.6	93.6	132.3	140.0	107.5	110.2
			3	83.8	83.5	117.6	123.4	95.9	97.8
		2	1	92.0	99.7	132.8	159.0	107.3	120.9
			2	87.7	87.3	123.0	129.4	100.4	102.5
			3	78.6	78.0	109.5	114.2	89.7	91.1
Yes	B	1	1	82.6	86.2	115.5	131.7	95.0	102.8
			2	80.0	77.6	109.8	111.8	90.8	90.0
			3	72.1	69.7	98.4	99.3	81.5	80.4

Note: Annuity factors are equated using cashflows discounted at 3% per annum to yield the same figures in Table 12. Sterbetafel Deutschland stops at age 100, so the mortality rate at age 100 was used for all ages higher than this.

Table 14: Percentage correlations between coefficients in the model summarized in Table 7.

	Age	Beard	Gender.M	Intercept	Makeham	Largest scheme.Y	Scheme type.2	Size.2	Size.3	Region.P	Time	Status.Ill-health	Status.Widow(er)
Age	100	85	70	-98	86	-9	-9	-6	-13	13	-11	65	6
Beard	85	100	63	-82	63	-7	-10	-9	-14	14	-9	56	11
Gender.M	70	63	100	-70	61	-4	-5	-6	-15	8	-8	44	27
Intercept	-98	-82	-70	100	-87	8	8	4	12	-14	-6	-64	-9
Makeham	86	63	61	-87	100	-10	-14	-2	-10	12	-8	51	9
Largest.Y	-9	-7	-4	8	-10	100	14	-2	-3	7	1	-5	2
Scheme type.2	-9	-10	-5	8	-14	14	100	-1	-2	-2	-2	-5	6
Size.2	-6	-9	-6	4	-2	-2	-1	100	11	-2	0	-5	14
Size.3	-13	-14	-15	12	-10	-3	-2	11	100	-2	1	-6	7
Region.P	13	14	8	-14	12	7	-2	-2	-2	100	-1	10	-2
Time	-11	-9	-8	-6	-8	1	-2	0	1	-1	100	-9	0
Status.Ill-health	65	56	44	-64	51	-5	-5	-5	-6	10	-9	100	13
Status.Widow(er)	6	11	27	-9	9	2	6	14	7	-2	0	13	100

Note: As with the variance-covariance matrix, the correlation matrix is symmetric about the leading diagonal. The leading diagonal is 100% because a parameter value is perfectly correlated with itself.

Contact

More information including case studies, latest features, technical documentation and demonstration videos can be found on our website at www.longevitas.co.uk

Conference House, 152 Morrison Street,
The Exchange, Edinburgh, EH3 8EB
Telephone 0131 315 4470
Email info@longevitas.co.uk

LONGEVITAS