A Practitioner's Guide to Statistical Mortality Graduation

Society of Actuaries

A Practitioner's Guide to Statistical Mortality Graduation

AUTHORS   Stefan J. Ramonat, FSA, FCIA
            Kai F. Kaufhold, Aktuar (DAV)

REVIEWERS   Philip Adams, FSA, CERA, MAAA
            Tom Edwalds, FSA, ACAS, MAAA
            Jean-Marc Fix, FSA, MAAA
            Stuart Klugman, FSA, CERA
            Mervyn Kopinsky, FSA, EA
            Allen Pinkham, ASA, MAAA
            Marianne Purushotham, FSA, MAAA

Caveat and Disclaimer

This study is published by the Society of Actuaries (SOA) and contains information from a variety of sources. The study is for informational purposes only and should not be construed as professional or financial advice. The SOA does not recommend or endorse any particular use of the information provided in this study. The SOA makes no warranty, express or implied, or representation whatsoever and assumes no liability in connection with the use or misuse of this study.

TABLE OF CONTENTS

List of Figures

List of Tables

Abstract

Mathematicians and actuaries have been applying relatively simple mathematical formulae to model mortality since de Moivre in 1725. Some of these "mortality laws" have stood the test of time and even today allow mortality experience to be captured across a wide range of ages. In fact, in the United Kingdom standard mortality tables have long been graduated according to such laws. In this paper, a statistical modelling framework from first principles is outlined that readily permits mortality laws to be fitted to mortality experience data. The dataset underlying the 2014 Canadian Pensioners' Mortality study (CPM2014) is invoked as a case study relevant to North American actuaries. Parametric mortality laws readily allow risk factors to be incorporated that describe the mortality of different groups within the experience data in an intuitive and precise manner. The practitioner can thus graduate mortality tables from heterogeneous datasets and gain an understanding of the drivers of mortality. The incorporation of risk factors is explored in general, and in particular in the context of reflecting amounts such as pensions-in-pay, annuity payments, policy face amount, or other benefit amounts, which are financially significant. The concepts discussed in this paper permit any actuary concerned with life contingencies to implement a statistical graduation framework in a simple, accessible manner.

1   Introduction

Any mortality basis used in the valuation of employee pensions and other benefit plans or the pricing or reserving of life insurance products is a model. It is a model whether calibrated directly to the experience of the particular group of interest or to that of other relevant populations. A mortality model generally comprises two main components: (1) the current (baseline) mortality, and (2) the improvement trend. This paper is concerned with modelling the former, and its purpose is to provide practitioners access to simple but powerful tools to graduate mortality tables and in the process discover more about the drivers of mortality.

As in any other modelling exercise, the best model for baseline mortality is the simplest that can produce predictions that are sufficiently accurate for the intended application. We therefore need to be able to measure the quality and applicability of the model's predictions. In addition, the model should be built making as few assumptions about the outcome of the analysis as possible, with the impact of each decision made along the way measured and documented. To ensure that no bias is introduced, the experience data to which the model is fitted should not be adjusted, or any data discarded, without good cause.

In what follows, a framework is built up following these principles. Very few, if any, of the methods are new or original, as the methodology explored is largely the same as that used by the Continuous Mortality Investigation (CMI) in the United Kingdom in the graduation of their standard baseline tables for the past thirty years. The majority of the CMI's methodology is comprehensively documented in Forfar et al. (1988), with other considerations explored over a number of CMI working papers (CMI WP), some of which are referenced below.[1]

However, despite the prevalence of such statistical models in the United Kingdom, many practitioners in Canada and the United States still gravitate to more traditional, restrictive approaches. The authors believe that the application of statistical concepts to Canadian and American mortality experience will illustrate how statistical methods can improve understanding of mortality.

The main focus of this paper is a method known as *graduation by mathematical formula*, where rela-

---

[1]With few exceptions, CMI outputs since March 2013 are available only to CMI subscribers and may not be distributed freely.

tively simple and intuitive formulae with only 2 to 5 parameters are fitted to mortality experience through a statistical model. Figure 1 demonstrates how such a 3-parameter model can capture the shape and level of a standard mortality table that was constructed by "traditional" means, including the use of Whittaker-Henderson smoothing at core ages. The standard table invoked is the lives-weighted version of the RP-2014 table for males from the Society of Actuaries (RPH-2014).[2]

Figure 1: Fitting a 3-Parameter Formula to the RPH-2014 Table for Males
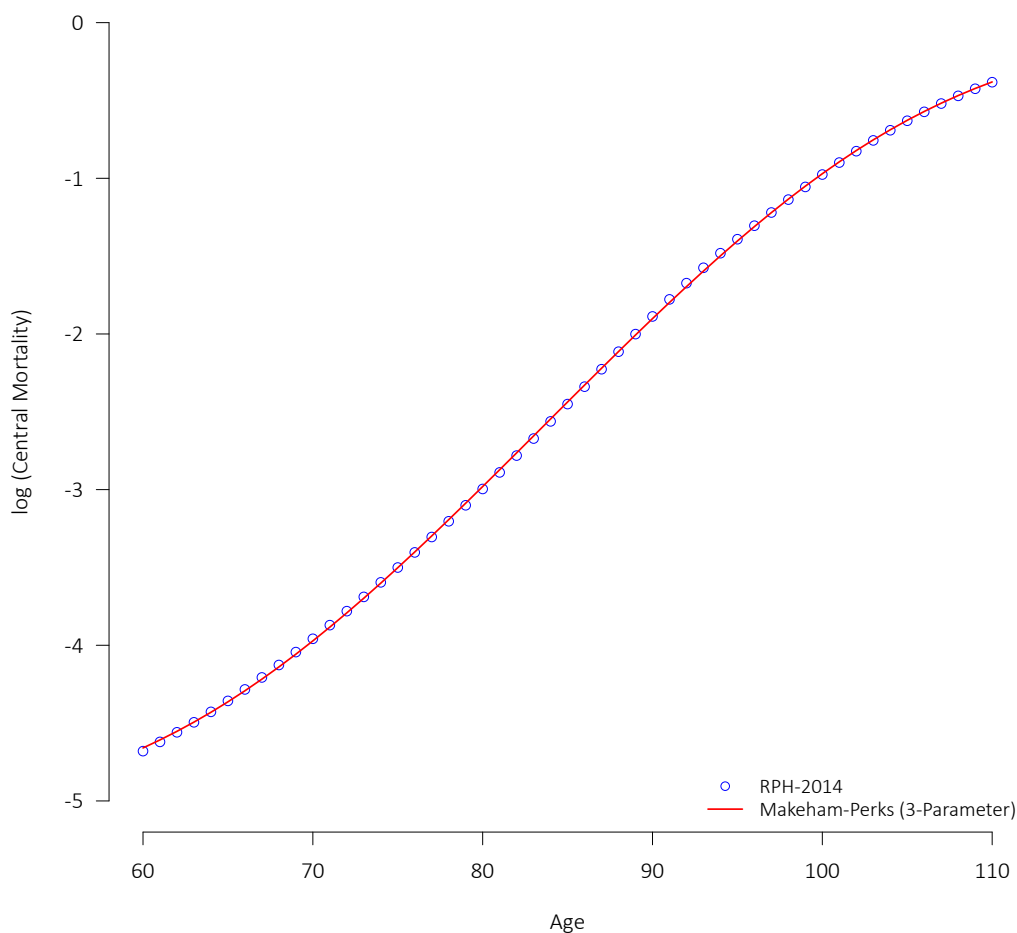
Clearly, the 3-parameter formula allows for a very simple model that still captures the same information in the 51 rates of RPH-2014 table.

The dataset used for the CPM2014 tables from the Canadian Pensioners Mortality study released in 2014

---

[2]See SOA (2014). For purposes of this illustration, the Makeham-Perks mortality law found in Table 3 is "fitted" to the logarithm of the graduated RPH-2014 probabilities of death (converted to approximate central rates via $m_x \approx -log\left(1 - q_x\right)$) using ordinary least-squares regression with no weightings over the entire 60 to 110 age range. The estimated parameter values, fitting Makeham-Perks directly to m_x, are -13.2042 for the "intercept" (α), 0.1269 for the "slope" (β), and -5.1565 for the "Makeham" parameter (ϵ). It should be noted that the RP-2014 family of tables were extended to the higher ages with a logistic function (Kannisto) consistent with the Makeham-Perks shape at those same ages, and thus the good fit at higher ages is not itself unexpected.

(see CIA (2014)) will be employed as a case study to illustrate the application of the concepts to be examined below. The grouped data is publicly available and the main subsets of the CPM2014 dataset used in this paper can be found in Appendix A.

All graphs below were created with R (R Core Team (2017)), as was the case for the majority of the statistical modelling. A version of the R code used for most of the modelling is made freely available as Appendix B. Typesetting was done in LaTeX. References to the logarithm log are to be taken to refer to the natural logarithm log_e unless otherwise stated.

2 Mortality Experience Data

The data typically available for mortality analysis at a grouped level includes the following:

- Deaths counts at each age x (denoted d_x)

- Population exposed-to-risk at each age x, or time lived, ideally the central (as opposed to initial) exposure (denoted E_x)[3]

- Benefit/policy amounts

- Additional attributes to allow analysis by risk factors, including gender, product, duration since commencement of benefit or policy, coverage (pensioner or surviving spouse, etc.), health status, smoking status where available, geographical and socio-economic information, etc.

For grouped data, the crude experience can be expressed in terms of the central death rate $m_x = \frac{d_x}{E_x}$ if the central exposure is used.[4] The central death rate m_x is closely related to the force of mortality μ_x, and thus models for m_x can be considered to be ones in continuous time. The relationships between the various rates, probabilities, and quantities, as well as other considerations regarding central exposure and the force of mortality, are provided in Appendix C.

It is crucial that the experience be as free of systematic bias as possible. In terms of the basic experience such as the deaths counts and exposure measures, that mainly means ensuring that certain individuals are not systematically excluded, particularly those who died during the observation period. The potential bias introduced into the risk factors can be especially problematic and subtle. For instance, a low pension amount may well indicate low service, such as may be the case for professionals changing employers frequently, as opposed to lower salary. Also, the place of residence (the postal code or ZIP code) may be more likely to be missing for deceased individuals than for survivors.

Addressing potential bias in the experience data is particularly important in relation to risk differentiators as the mortality differentials between various risk groupings are rarely proportional across all ages. The impact of nearly every risk differential lessens over age. For example, consider the central death rates m_x for males and females in Canada in 2011 that are provided in Figure 2.[5]

The difference between the two sexes lessens over age. As age is too vital a factor to ignore, we must account for the age shape of the impact of a risk factor. In this case, gender and age can either be modelled separately by splitting the experience data into separate parts for males and females (stratified) or modelled simultaneously using a multivariate approach. Otherwise, any conclusions drawn from the resulting model will be systematically flawed. The age structure across risk factors is one of the key motivations for developing a statistical modelling framework for mortality graduation.

---

[3] Central exposure is often denoted E_x^c to distinguish it from initial exposure. The superscript will generally be excluded in this paper, but references to E_x should be taken to mean the central exposure unless stated otherwise.
[4] Note that the crude experience would be modelled through the probability of death q_x if the initial exposure was instead used.
[5] Source: CHMD (2014)

Figure 2: Crude Male and Female Mortality for Canadians in 2011
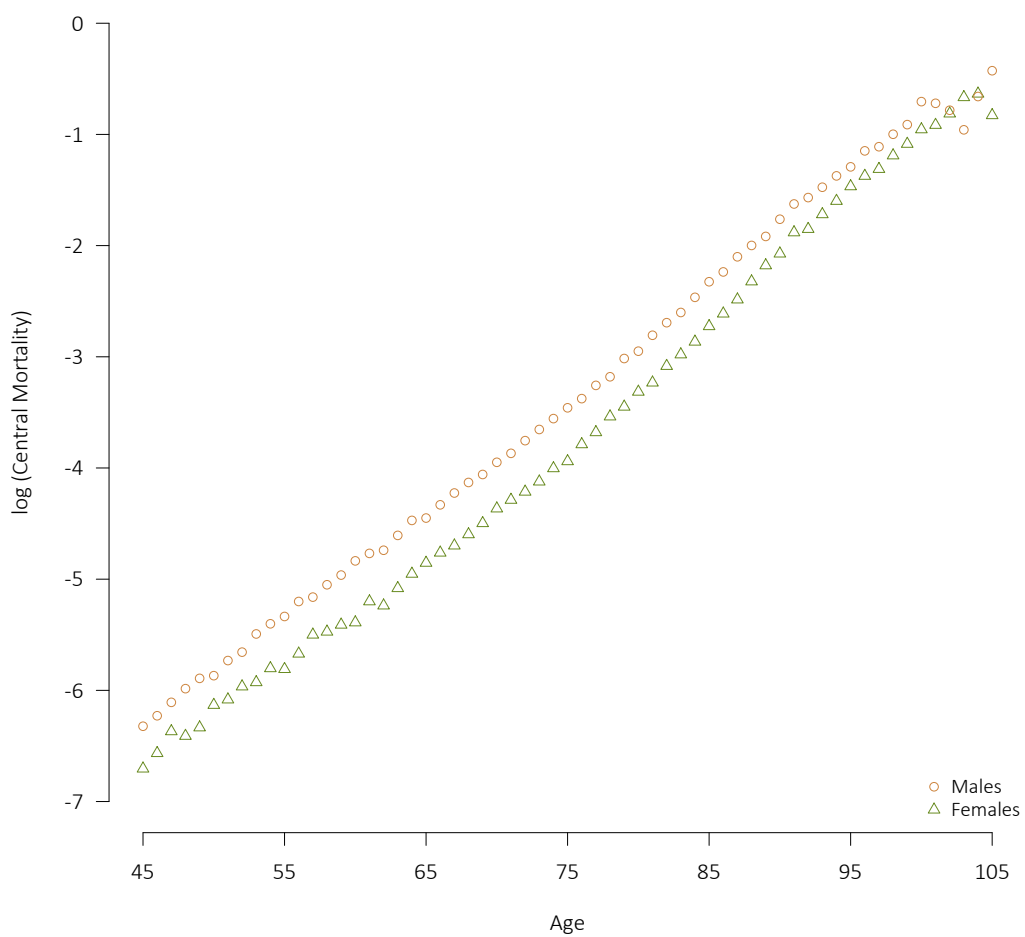
3   Introduction to Graduation

The main example in this and subsequent sections draws upon the male CPM2014 dataset, on a lives basis with no further stratification (i.e., across both public and private sectors and any of the other risk factors available).[6] The crude central deaths rates (m_x) for ages 50 through 100 are provided as Figure 3.

Figure 3: Crude Central Mortality Rates under the CPM2014 Dataset on a Lives Basis
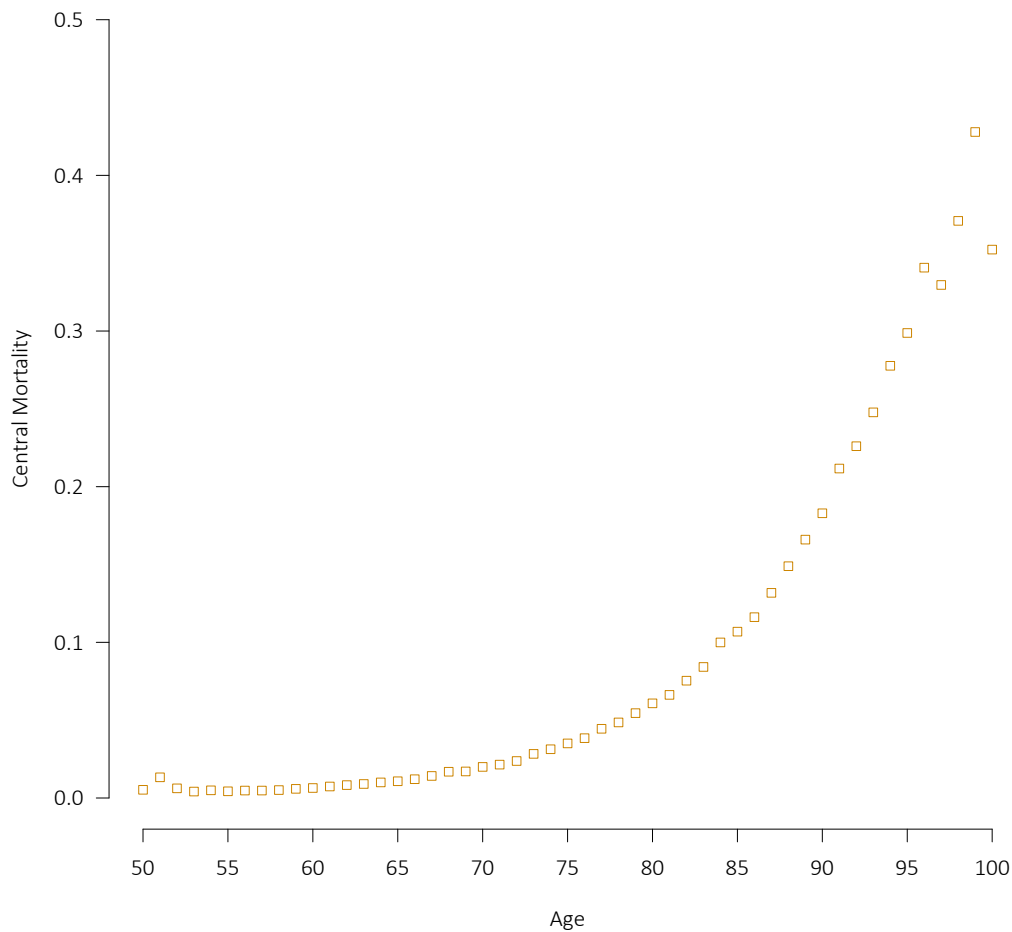
The mortality pattern demonstrates an exponential increase with age. Reverting to a logarithmic y-axis makes it easier to observe the inherent age structure, which is largely log-linear, as indicated in Figure 4.

The simplest model to capture the pattern in Figure 4 is a (log-)linear relationship, i.e. $m_x = e^{\alpha + \beta x} \implies log(m_x) = \alpha + \beta x$ with α the "intercept" and β the "slope" by age. The crude experience appears best behaved between ages 55 and 95, so analysis will be restricted to that range. To start, we fit the simple

---

[6]The data available with respect to the CPM2014 dataset was tabulated on an age nearest birthday (ANB) basis and only the initial exposure is available. For purposes of this paper, the age definition is effectively ignored, and the central exposure estimated from the initial exposure and deaths according to equation (C.4).

Figure 4: Crude Central Mortality Rates under the CPM2014 Dataset on Logarithmic Scale
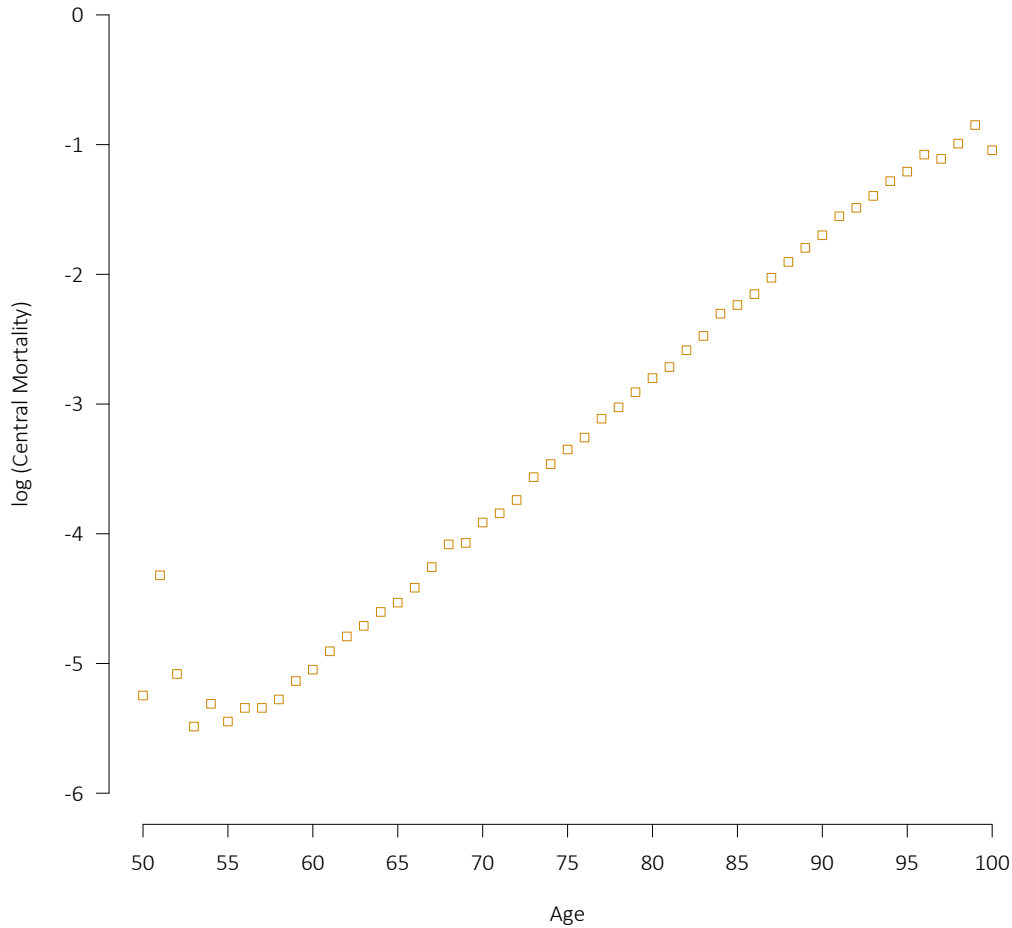
linear model by least-squares regression without any weightings,[7] which results in $log\left(m_x\right) = -11.6253 + 0.1102x$, as provided in Figure 5.

Even though it is a very simple model, the straight line appears to fit reasonably well. So, for this set of crude mortality experience, we appear to have a workable model. However, there is a major drawback to the least-squares approach; no allowance has been made for the fact that greater statistical credibility applies to those ages with a larger death count and more exposure.

As the aim is to push the limits of mortality analysis and find methods to measure the reliability of our results, the following sections will generalize the modelling beyond simple linear regression and introduce a fully developed statistical framework.

---

[7] In this case, minimizing the following function with respect to the parameters α and (β): $\sum_{x} \left[log\left(\frac{d_x}{E_x}\right) - (\alpha + \beta x) \right]^2$

Figure 5: Unweighted Linear Regression Applied to CPM2014 Dataset
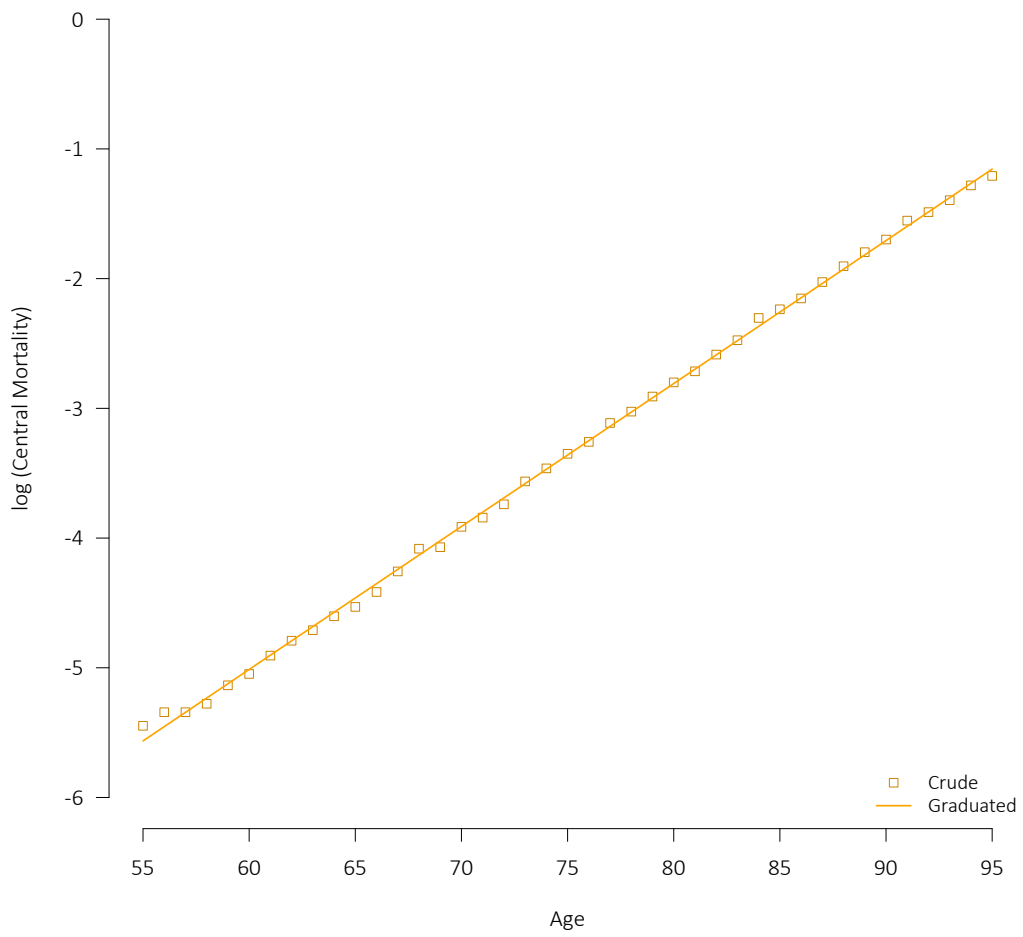
4   A Statistical Approach to Graduation

In the preceding section, the least-squares linear regression was unweighted. However, conceptually it would make sense to place more weight on ages where more experience is available. Critically, with unweighted least-squares the variance of the mortality process is assumed to be constant across all ages (*homoscedasticity*), which is not realistic as there will be different levels of deaths and exposures at each age, each suggesting a different variance.

To allow for differing variances by age and achieve a more appropriate fit, the least-squares regression could be modified to weight the squared differences at each age. In a least-squares context, such a weight would ideally be greatest where the variance of the quantity under measurement is smallest, which can be achieved by using a function of exposures or observed deaths. Another approach is to determine natural weightings following from an assumption regarding the distribution of deaths at a given age, which is developed in what follows.

The random variable of interest is that for the deaths count at each age, which we will denote D_x. So long as the lives are independent over an age x, we can assume that the deaths count at each age is Poisson-distributed,[8] where the expected value (and variance) in terms of the estimator of m_x, $\hat{m}_x$, is $\lambda = E_x \hat{m}_x$:[9]

$$D_x \sim Poisson\left(E_x \hat{m}_x\right) \tag{1}$$

Under the Poisson assumption, the probability mass function with expected value λ is as follows:

$$\Pr\left(K = k\right) = e^{-\lambda}\frac{\lambda^k}{k!} \tag{2}$$

The above holds for each age x. Therefore, if we can assume that the observed deaths counts are independent, multiplying the probability for observing the age-specific deaths count for each age yields the (conditional) likelihood function, where $\hat{m}_x(\theta)$ is a function of parameters θ:

$$L\left(\theta\right) \propto \prod_x e^{-E_x \hat{m}_x(\theta)}\frac{\left(E_x \hat{m}_x(\theta)\right)^{d_x}}{d_x!} \tag{3}$$

To obtain the optimal estimates of the parameters θ given the observations and model form, we determine the parameter set that maximizes the value of the above likelihood function, i.e., maximum likelihood estimation (MLE). Maximizing generally entails taking derivatives, and as the sum rule for derivatives is much simpler than the product rule, it is almost always more straightforward to maximize the log-likelihood function:[10]

$$l\left(\theta\right) = \sum_x \left[-E_x \hat{m}_x(\theta) + d_x log\left(E_x \hat{m}_x(\theta)\right) - log\left(d_x!\right)\right] \tag{4}$$

---

[8] Also see Forfar et al. (1988) for discussion of some additional considerations. There are a number of situations where the basic Poisson model may potentially be insufficient, particularly when the variance of D_x is greater than the mean, which violates the Poisson assumption. Refer to Appendix D for a brief discussion of how to handle such *overdispersion*.

[9] In this paper, m_x is effectively a convenient shorthand for $\mu_{x+\frac{1}{2}}$. Alternatively, the ratio $m_x = \frac{d_x}{E_x}$ could be assumed to refer to the integrated hazard: $\int_0^1 \mu_{x+t}\, dt$. Refer to Appendix C for additional details

[10] Since the logarithm is a strictly monotonic function, the set of parameters that maximizes the log-likelihood also maximizes the likelihood function itself.

The above can be simplified even further, as $d_x!$ does not depend on $\hat{m}_x$ and disappears when you take the derivative, and the $log\left(E_x\hat{m}_x(\theta)\right)$ can be replaced by $log\left(\hat{m}_x(\theta)\right)$ for the same reason.[11]

The Poisson assumption automatically weights based on how much experience is available at each age of observation and allows the variance to vary across the age range, which is a distinct improvement over the simple least squares approach.

We continue to use the simple log-linear model for m_x and estimate the parameters by maximizing equation (4). The resulting death rates are shown in Figure 6 and the corresponding *deviance residuals* in Figure 7. Deviance residuals represent the deviations between the observed and modelled deaths.[12] If the residuals are randomly distributed according to a standard normal distribution, the residuals will be randomly scattered about zero, with only 1-in-20 falling outside ± 1.96.

The deviance residuals in Figure 7 appear to be fairly random. Any clear patterns would imply that a more complex model is necessary to capture the information not explained by our model. However, for the case at hand this does not appear necessary, and so the simple log-linear model already provides a good fit.

Beyond visual inspection of the fit and residuals, operating under a statistical framework permits us to apply statistical tests, such as:

- **Pearson's χ^2 Test**: A goodness-of-fit test with the test statistic[13] assumed to follow a χ^2 distribution with degrees of freedom equal to the number of observations (number of ages in our case) less the number of parameters.

- **Sign Test**: Residuals should be equally likely to be above or below zero. The sign test involves calculating the probability that the observed split between positive and negative residuals follows a binomial distribution. If that probability is low, say below 0.05 or 0.10, then the residuals potentially would be considered to be insufficiently random.

- **Runs Test**: Involves calculating the probability of observing a number of runs of positive or negative residuals given the counts of positive and negative residuals. If the resulting probability of observing a particular number of runs or fewer in the residuals is quite low (suggesting the presence of too few runs) or high (suggesting too many runs), the randomness of the residuals might be called into question.

- **Information Criteria**: As mentioned in the introduction, our aim is to retain the simplest possible model that captures the optimal amount of information from the experience and provides us with reliable predictions. One measure of "simplicity" is the number of parameters within a model. For

[11]In this paper, the "full" log-likelihood function defined by equation (4) is employed as it is generally more conveniently scaled, and introduces no computational issues. However, the simplified log-likelihood function, which is given by $l\left(\theta\right) \propto \sum_x \left[-E_x\hat{m}_x(\theta) + d_x log\left(\hat{m}_x(\theta)\right)\right]$ is sufficient in most cases, and also is equivalent to assuming a constant force of mortality within individual ages x, as illustrated in Forfar et al. (1988).

[12] Under the Poisson assumption, with d_x the observed deaths and f_x the deaths fitted under the model at a given age, the deviance residual at that age is calculated as follows, with $0log\left(0\right)$ taken to be zero:

$$sign\left(d_x - f_x\right)\sqrt{2\left[d_x log\left(\frac{d_x}{f_x}\right) - (d_x - f_x)\right]}$$

[13]$\chi^2 = \sum_x \frac{(d_x - f_x)^2}{f_x}$, where d_x are the observed deaths (assumed to be independent of each other) and f_x the deaths expected (fitted) under the model.

Figure 6: Log-Linear Poisson Model Applied to CPM2014 Dataset
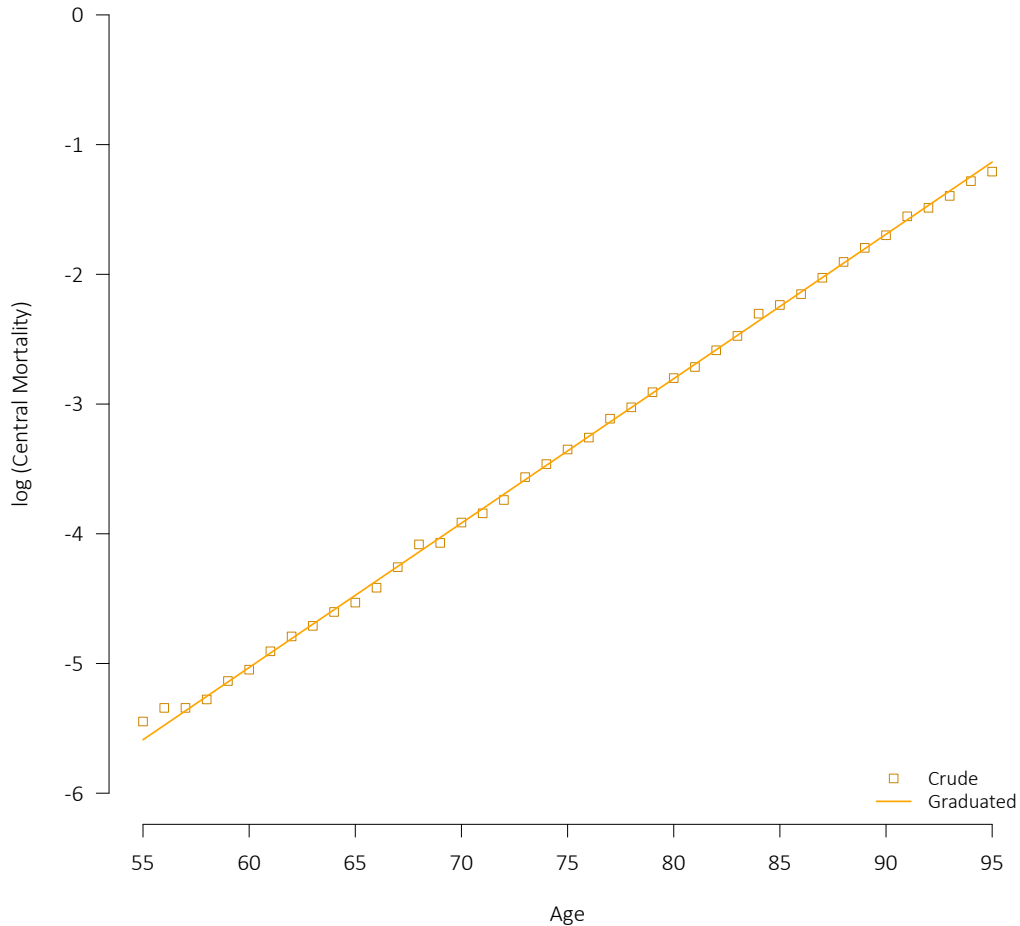
instance, a mortality table from ages 55 to 95 years based on the crude death rates themselves would consist of 41 parameters, one for each age. Our log-linear model, which appears to capture all the same information, requires only two parameters. Thus, the simple straight line "wins" easily.

A more general way of quantifying the preference for simple models with the fewest parameters (i.e., *parsimonious* models), is to balance the goodness-of-fit as expressed through the log-likelihood with the number of parameters in the model. That is precisely what various information criteria do:

– **Akaike's Information Criterion (AIC)**: $-2l + 2K$, where l is the maximum of the log-likelihood under the model and K is the number of parameters.

– **Bayesian Information Criterion (BIC)**: $-2l + log(N)K$, where l and K are as above and N is the number of observations (number of ages in our context).

With both the AIC and BIC, the lower the score, the better the relative fit of the model, balanced against the number of parameters. The BIC penalizes complex models to a greater degree than the AIC so long

Figure 7: Deviance Residuals – Log-Linear Poisson Model
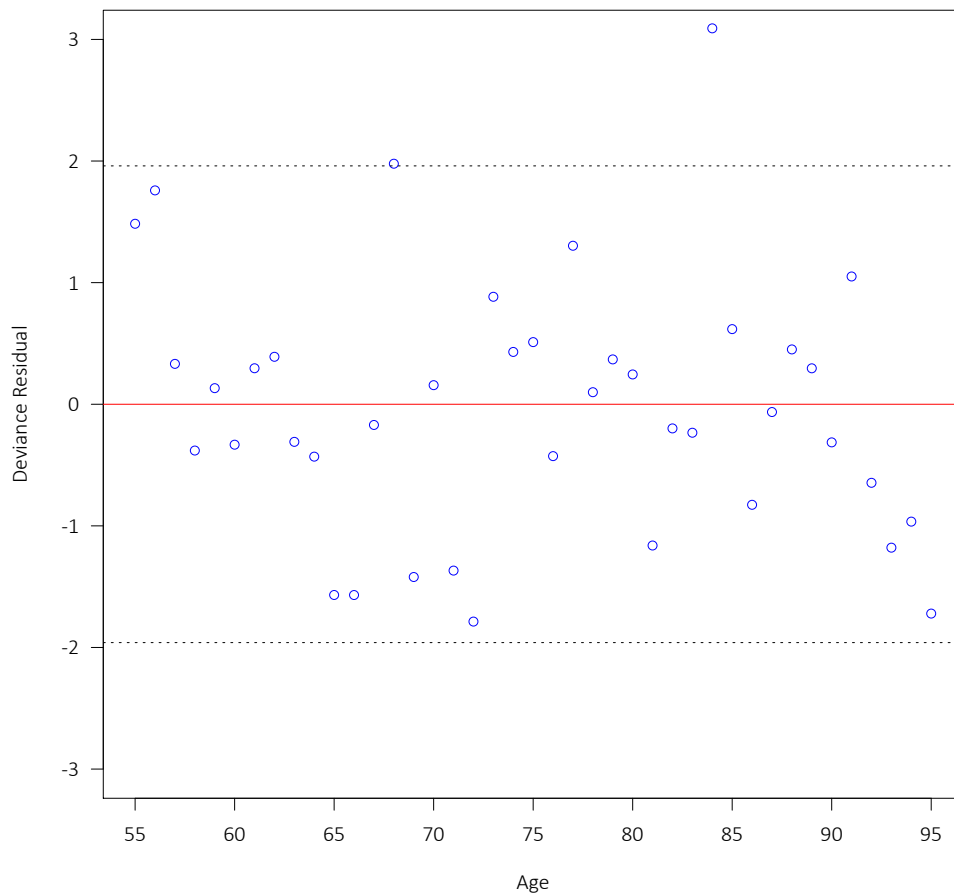
as there are more than seven observations.

Table 1 provides the values for the above measures and tests for the linear models. The Poisson-based model yields a much higher p-value under the χ^2 test and a non-significant one under the runs test, suggesting a better fit.

The significance legend herein is the same as that used by R Core Team (2017) ($***$ for 0.1%, $**$ for 1%, $*$ for 5%, and $\cdot$ for 10%).

It is important to ensure that extraneous parameters are not included in the models to avoid overfitting or misleading conclusions. The AIC, BIC, and similar criteria help greatly in that respect, but at the same time we have to ensure that each individual parameter estimate is statistically significant, which means we have not just fitted to random fluctuations. That can be determined by calculating the standard errors for the parameters, calculating the two-way Z statistic and determining the resulting p-values. The relevant quantities with respect to the log-linear model under the Poisson distribution assumption are provided in Table 2. Unsurprisingly, both parameters are found to be highly significant.

But what happens if the simple straight line does not capture all the information? The following section

Table 1: Model Fit under Log-Linear Modelling

	Log-Linear Least-Squares	Log-Linear Poisson Model
Parameters	2	2
Intercept (log scale)	-11.6253	-11.7109
Slope (log scale)	0.1102	0.1113
Log-Likelihood		-204.60
AIC		413.20
BIC		416.63
χ^2 **Statistic**	51.39	44.99
Degrees of Freedom	39	39
χ^2 **p-value**	0.0885	0.2355
χ^2 **Significance**	.	*not significant*
Positive Residuals	21	20
Negative Residuals	20	21
Sign Test p-value	0.5000	0.5000
Signs Test Significance	*not significant*	*not significant*
Runs	14	20
Runs Test p-value	0.0130	0.3789
Runs Test Significance	$*$	*not significant*

Table 2: Parameters under Log-Linear Poisson Model

Parameter	Estimate	Standard Error	Z	p-value	Significance
Intercept (α)	-11.7109	0.0377	-310.61	0	$* * *$
Slope (β)	0.1113	0.0005	235.51	0	$* * *$

will generalize the form of μ_x or m_x beyond a log-linear model.

5 Generalizing the Functional Form

As illustrated in the two preceding sections, a log-linear form already fits the experience of Canadian male pensioners. However, there are many situations where a more complex age shape is warranted, and we may even be able to do better than the log-linear form for the dataset at hand. It is important to get the general age shape right before including parameters to describe additional risk factors. As age is such an important risk factor, any systematic bias that might be present if the age shape is not captured accurately will confound the impact of other risk factors.

The log-linear form explored above assumes that mortality increases at an exponential rate with age. That assumption is equivalent to the mortality law put forth in Gompertz (1825):[14]

$$\mu_x = e^{\alpha+\beta x} \tag{5}$$

The mortality laws explored in this paper are properly expressed directly in terms of the force of mortality μ_x. However, for simplicity we are applying the relevant law to m_x instead. The results would not have been materially different had an approach applying the relevant law to μ_x had been employed. Refer to Appendix C for details on how μ_x and m_x relate.

Makeham (1860) and Makeham (1867) describe an extension to Gompertz' law that involves adding a positive constant to better capture mid-age mortality:

$$\mu_x = e^{\epsilon} + e^{\alpha+\beta x} \tag{6}$$

In Perks (1932), a logistic extension to Makeham's law is first presented, with a very insightful heterogeneity/frailty argument for the resulting form conceived subsequently in Beard (1959). This quite general mortality law is referred to as "Makeham-Beard" in Richards (2012) and in what follows, and is described by the intercept and slope parameters α and β, the "Makeham" parameter ϵ, and the "Beard" parameter ρ:

$$\mu_x = \frac{e^{\epsilon} + e^{\alpha+\beta x}}{1 + e^{\alpha+\rho+\beta x}} \tag{7}$$

In Richards (2012), the key mortality laws employed are Gompertz, Makeham, and the simplifications of the Makeham-Beard law where $\rho = 0$ (the "Makeham-Perks" law), $\epsilon \to -\infty$ (the "Beard" law), and where $\rho = 0$ and $\epsilon \to -\infty$ (the "Perks" law). A comparison of the force of mortality μ_x under certain laws from Richards (2012) is given in Figure 8, where in the relevant laws $\alpha = -12$, $\beta = 0.12$, $\epsilon = -6$, and the Beard parameter ρ set to 1 or -1 as indicated.

The laws shown in Figure 8 are particularly well-behaved and suitable for extrapolation to the higher and lower ages. However, if the practitioner suspects further structure in the mortality experience data, it is possible to generalize by replacing the additive Makeham constant and Gompertz linear form with higher-order polynomials, described in Forfar et al. (1988) as the generalized Gompertz-Makeham family of laws $GM(r, s)$:[15]

---

[14]Gompertz' original, equivalent expression was more similar to $\mu_x = Bc^x$, with B and c being positive constants. We adopt the convention in Richards (2012) where when a parameter θ must be positive, $e^{\theta'} = \theta$ is used instead. That aids in fitting the parameters as they are then free to vary over the entire real line and need not be explicitly constrained. It is also convenient, where possible, to have the parameters on a logarithmic scale.

[15]See Appendix E for details on how the $GM(r, s)$ family is generally fit in practice, which includes fitting to a polynomial function of a transformation of x as opposed to x^k directly.

Figure 8: Comparison of Main Mortality Laws under Richards (2012)
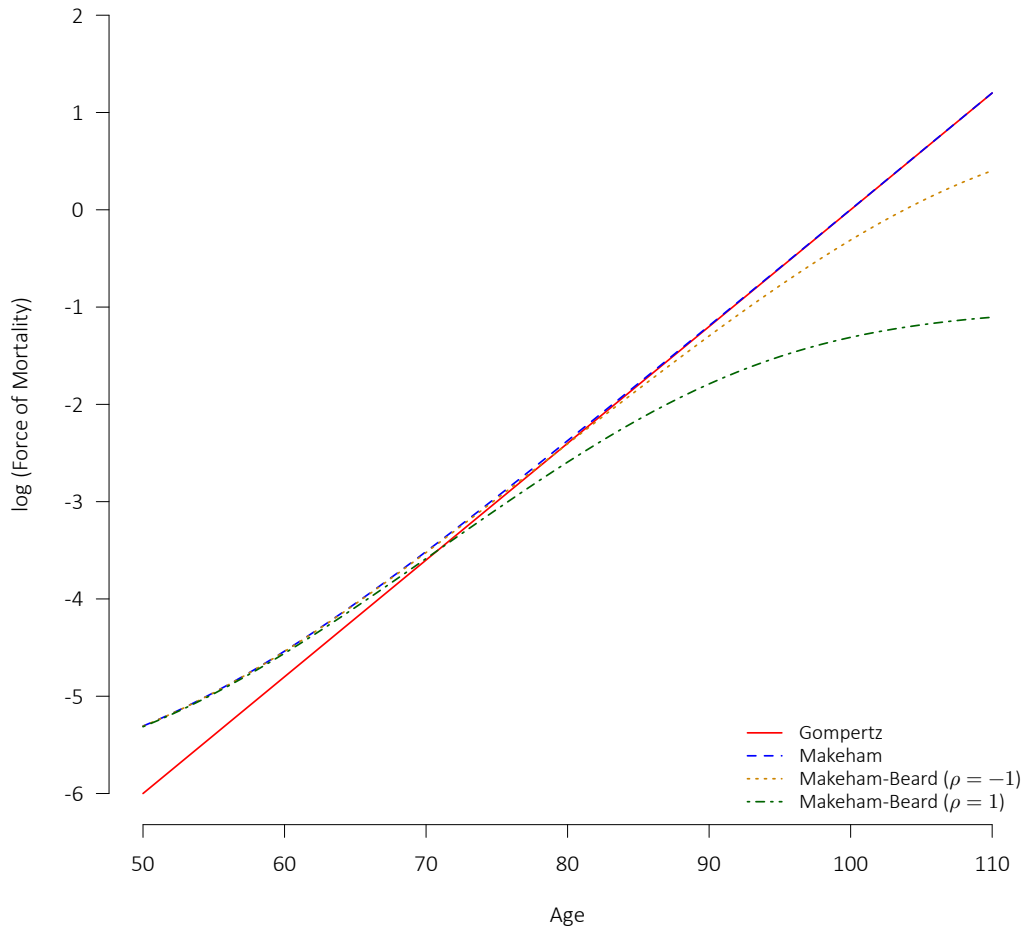
$$\mu_x = GM(r,s) = \sum_{i=0}^{r-1} a_i x^i + e^{\sum_{j=0}^{s-1} b_j x^j} \tag{8}$$

Note that the Gompertz law under Richards (2012) is equivalent to a $GM(0,2)$, and the Makeham law is equivalent to a $GM(1,2)$. Additionally, Makeham's second law described in Makeham (1890) is the same as a $GM(2,2)$. Table 3 summarizes the various laws.

The CMI has utilized some form of the $GM(r,s)$ family in the graduation of mortality tables over key ages for many years. The most recent standard tables constructed from the experience of employer-sponsored pension plans in the United Kingdom (the "S2" Series) mainly relied upon $GM(1,3)$ or $GM(0,4)$ forms (see CMI WP 071 (2014)). The most recent annuitant standard tables (the "08" Series) opted for simpler forms, effectively $GM(0,3)$, $GM(0,4)$, or $GM(0,5)$, as found in CMI WP 078 (2015). In all cases, relatively few parameters are involved.

Table 3: Selected Mortality Laws

Richards (2012)	$GM(r,s)$	Force of Mortality μ_x
Gompertz	$GM(0,2)$	$e^{\alpha+\beta x}$
Makeham	$GM(1,2)$	$e^{\epsilon} + e^{\alpha+\beta x}$
Perks		$\dfrac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$
Beard		$\dfrac{e^{\alpha+\beta x}}{1+e^{\alpha+\rho+\beta x}}$
Makeham-Perks		$\dfrac{e^{\epsilon}+e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$
Makeham-Beard		$\dfrac{e^{\epsilon}+e^{\alpha+\beta x}}{1+e^{\alpha+\rho+\beta x}}$
	$GM(r,s)$	$\displaystyle\sum_{i=0}^{r-1} a_i x^i + e^{\sum_{j=0}^{s-1} b_j x^j}$

In North America, while graduations based on Makeham's original laws were once common, graduation by mathematical formula has fallen out of favour. One relatively recent example is found in Panjer and Tan (1995) for Canadian insured mortality for the period from 1986 to 1992. Makeham's second law (i.e., a $GM(2,2)$) is invoked to graduate the experience over ages 40 to 99 for males and 44 to 99 for females. However, the fit is made using least-squares regression weighted with the exposure at the relevant age as opposed to maximum likelihood estimation.

Continuing the North American perspective, Chan and Panjer (1979) describe a method of employing maximum likelihood estimation to graduate mortality according to laws such as Gompertz and Makeham. The Poisson assumption is not invoked, as the paper focuses on building the likelihood function up from the level of the individual. As shown in Forfar et al. (1988), maximizing the likelihood function developed for grouped data in that paper results in equivalent parameterizations to maximizing the Poisson likelihood. Tenenbein and Vanderhoof (1980) suggest that there are many advantages of working with mathematical models of mortality and presents some extensions of Gompertz and Makeham.

To determine whether one of the model forms in Table 3 may perform better than the Gompertz model that we have already examined, we fit each law one-by-one, and find values of the AIC and BIC. The results are tabulated in Table 4.

There is not much to separate the "top" models, but the 4-parameter $GM(1,3)$ slightly edges out the others under both measures. The parameter estimates under the $GM(1,3)$ model (as provided through equation (8)) are provided in Table 5, along with the standard errors, the Z score, the p-value, and the significance.

The fit and residuals are not shown as there is little change from that of the Gompertz. However, given that the $GM(1,3)$ does appear to result in a slightly better-fitting model with parameter estimates all significant, we can consider it as the final model for fitting the CPM2014 male dataset on a lives basis.

Table 4: Measures of Model Fit for Males in the CPM2014 Dataset under Various Laws

Law	Parameters	AIC	BIC
Gompertz / $GM(0,2)$	2	413.20	416.63
Makeham / $GM(1,2)$	3	415.03	420.17
Perks	2	453.30	456.73
Beard	3	412.40	417.54
Makeham-Perks	3	411.64	416.78
Makeham-Beard	4	407.47	414.32
$GM(0,3)$	3	414.62	419.76
$GM(0,4)$	4	407.86	414.72
$GM(0,5)$	5	409.65	418.21
$GM(1,3)$	4	407.34	414.20
$GM(1,4)$	5	409.34	417.91
$GM(2,2)$	4	407.61	414.47
$GM(2,3)$	5	409.23	417.80

Table 5: Parameters under $GM(1,3)$ Model for CPM2014 Males

Parameter	Estimate	Standard Error	Z	p-value	Significance
a_0	0.001440	0.000412	3.4944	$4.7506e-04$	$***$
b_0	-4.474429	0.172141	-25.9928	0	$***$
b_1	6.129081	0.173795	35.2662	0	$***$
b_2	-0.466494	0.146658	-3.1808	$1.4686e-03$	$**$

It should be noted that logistic formulae tend to become the better fitting models at higher ages and this is especially so if the underlying population is particularly heterogeneous. To illustrate, we can invoke the frailty arguments explored in Beard (1959) and Richards (2008). Assume three risk groups are represented in the population, the mortality for each being captured through Makeham's law with $\beta = 0.13$ and $\epsilon = -5$, and with the "intercept" α being -12, -13, and -14 for the high, middle, and low mortality groups, respectively. Now, we simulate a process where 50,000 individuals commence a benefit or policy at exact age 50 at the

beginning of each year, where 15,000 (30%), 27,500 (55%), and 7,500 (15%) are assumed to fall within the high, middle, and low mortality groups, respectively. For each individual the remaining lifetime is simulated directly using the cumulative distribution function, as described in Richards (2012). The simulated experience in the 61st year of this process in then tabulated and treated as "crude experience" in Figure 9.

Figure 9: Illustration of the Frailty Argument
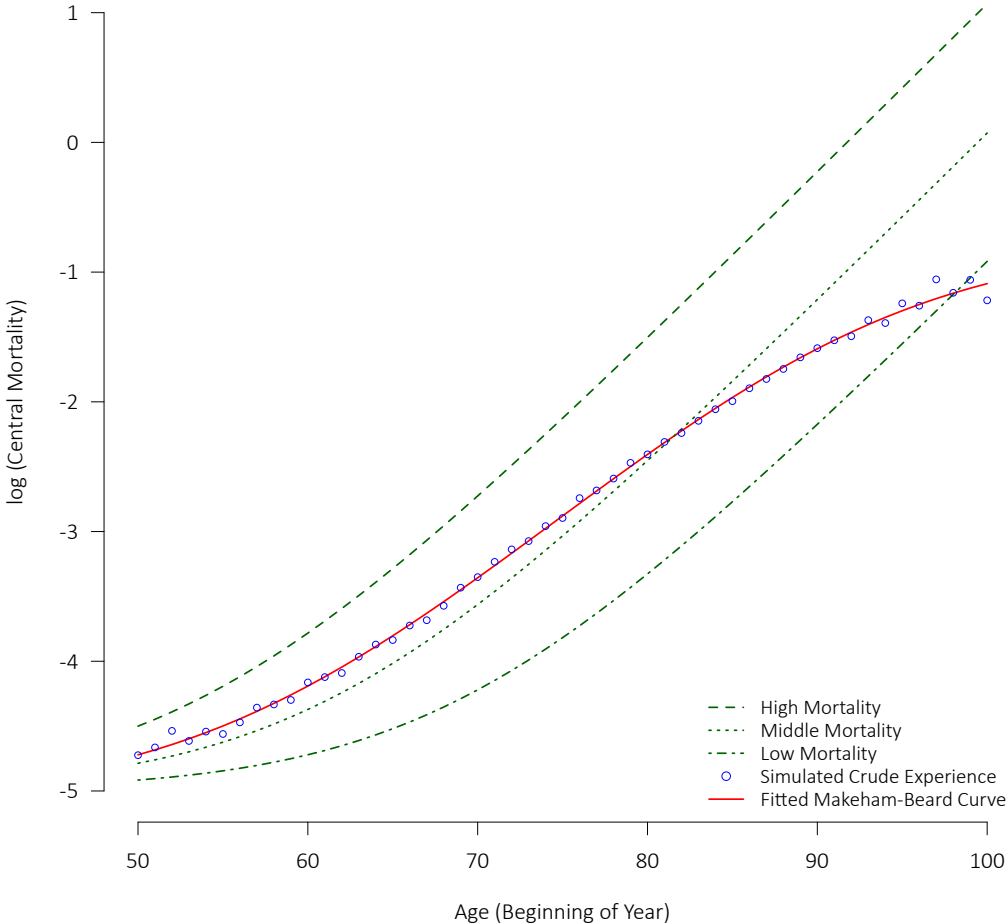
The simulated experience resulting from combining the three risk groups is clearly not well explained by any of the Makeham laws. In fact, the best fitting model is a Makeham-Beard (with parameters for the force of mortality μ_x of $\alpha = -12.0251$, $\beta = 0.1213$, $\epsilon = -5.0800$, and $\rho = 0.7588$ with the particular set of simulated data). This is a basic illustration of the frailty argument for the Beard parameter. Effectively, with advancing age the survivors tend more and more towards to lower mortality groups, creating the Beard shape.[16]

---

[16]The frailty result is developed mathematically in Richards (2008), where it is demonstrated that a Makeham-Beard form at the population level is obtained in a scenario where the underlying individuals' mortality patterns follow a Makeham form where e^α is drawn from a gamma distribution.

The illustrative simulation described above provides a useful intuition for the Beard parameter and shape. A positive value of the Beard parameter thus can be directly associated with heterogeneity within a pool of mortality experience data. In some applications, it will not be possible to identify the different risk groups that are contributing different mortality patterns to the overall experience, such as the study of population level mortality or aggregated within an industry experience study. However, in instances where the practitioner has access to additional information on the lives in question and where it is possible to model the different risk classes within the overall model, this will be preferable to the aggregate model with the accompanying heterogeneity. Using an aggregate model without accounting for the heterogeneity always introduces *distribution risk*, i.e. the risk that the distribution of the group being modelled is different from the group which supplied the experience data. In addition, capturing the shape of the heterogeneous experience may require a complex functional form, if a suitable one can be found at all. Incorporating differentials into the model or stratifying the dataset according to key risk classes will often allow a simpler underlying structure to be employed. Section 8 introduces risk factors in the context of our case study.

6 Extension to Higher & Lower Ages

It is often necessary to extend the graduated rates to older and younger ages where there is insufficient experience. Extending mortality rates to higher ages is especially important for annuitant and pensioner tables. Methods for extending the tables to higher ages include the logistic extension, akin to the Perks or Beard laws in Table 3, a polynomial extension (usually cubic), or to gradually converge towards a population table. Along a different vein, Li et al. (2008) applies extreme value theory to model mortality at the highest ages.

CMI WP 085 (2015) and CMI WP 100 (2017) provide an excellent overview of the various approaches available for extension to higher ages, including a summary of which approaches have been used in standard tables in the UK, the United States, and Canada. Those working papers also provide a summary of recent research into mortality patterns at high ages. Both are publicly and freely available.

Ages below those at which there is ample experience can actually pose even greater problems than the ages above. In some cases, it may not make sense to extrapolate to younger ages as the younger age mortality applies to a different population. For example, pensioner mortality rates in the 50s and early 60s only capture the portion of the pension plan's membership who have retired. The portion still in active service are likely to be in better health than their counterparts who have already retired. Therefore, extending the mortality rates to the younger ages based solely on the experience of retired members may not be appropriate. In such a situation, a separate mortality table would ideally be graduated for the active members, as was the case for the RP-2014 tables. The large differences between the active and healthy annuitant (approximate) central mortality rates under the lives-weighted versions of the RP-2014 tables for males (i.e., RPH-2014) are illustrated under Figure 10.[17]

If the rates are to be extended to the lower ages, some version of Makeham's law (or a $GM(r,s)$ law), or otherwise some form of polynomial, will likely prove satisfactory in many situations. For instance, extensions using $GM(r,s)$ forms were adopted for the "S2" series of pensioner tables in the UK (see CMI WP 071 (2014)).

One of the advantages of graduating mortality tables by mathematical formula is that certain laws such as Makeham-Perks or Makeham-Beard in Table 3 automatically extend to both younger and higher ages in a manner consistent with the underlying experience (assuming that the populations to which the extended rates will be applied are consistent with the population included in the graduation). For example, the Makeham-Perks law can be applied to the lives-based (unweighted) CPM2014 dataset for males explored above across the 55 to 95 age range and the resulting function allowed to naturally extend over the 50 to 110 range. The fit of the Makeham-Perks law over 55 to 95 is good, and provides for a particularly clean and smooth extension.

The fit and extension is provided in Figure 11, with the parameters estimated as $\alpha = -12.6846$, $\beta = 0.1242$, and $\epsilon = -6.6287$. The 97.5^{th} and 2.5^{th} percentiles (i.e., the higher and lower bounds of the 95% confidence intervals) for the central mortality rates at each age based on the raw data are also provided for context.[18] The fit appears more than adequate at all ages, including the extensions at the lower and higher ages, and thus the Makeham-Perks model could be considered as an alternative to the $GM(1,3)$ model

---

[17]See SOA (2014)

[18]The percentiles are calculated directly, by determining the corresponding percentile from the inverse of the Poisson cumulative distribution function for the deaths count. The mean of the Poisson is equal to the observed deaths (the d_x for that age), and the result is then divided by the exposure E_x at the age.

Figure 10: Comparison of the Healthy Annuitant and Employee (Active) RPH-2014 Rates
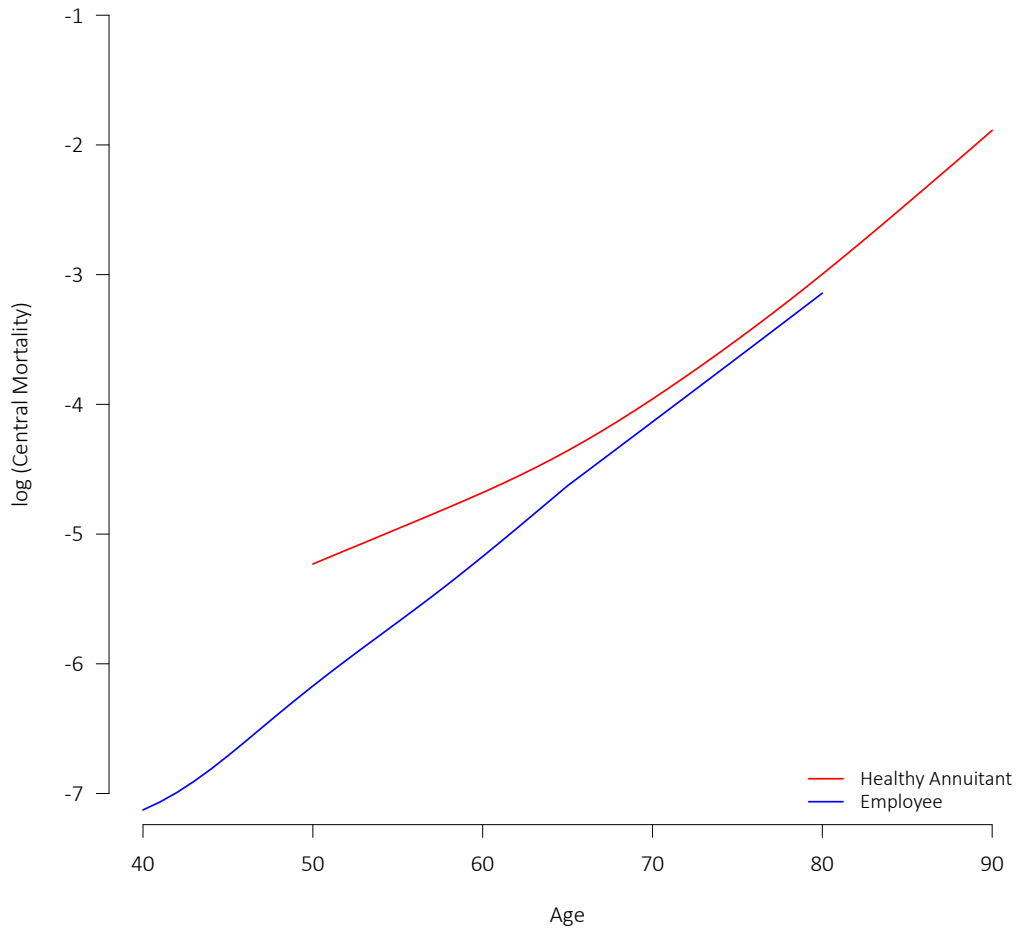
selected above where simple, reasonable extensions are desired.

A disadvantage of employing $GM(r, s)$ forms involving higher-order (i.e., greater than first degree) polynomials is that extrapolations outside the age range graduated are generally poorly behaved. Thus, when such higher-order polynomials are invoked, extensions must be made in subsequent steps following graduation.

Figure 11: Graduation and Extension of CPM2014 Male Experience Using Makeham-Perks
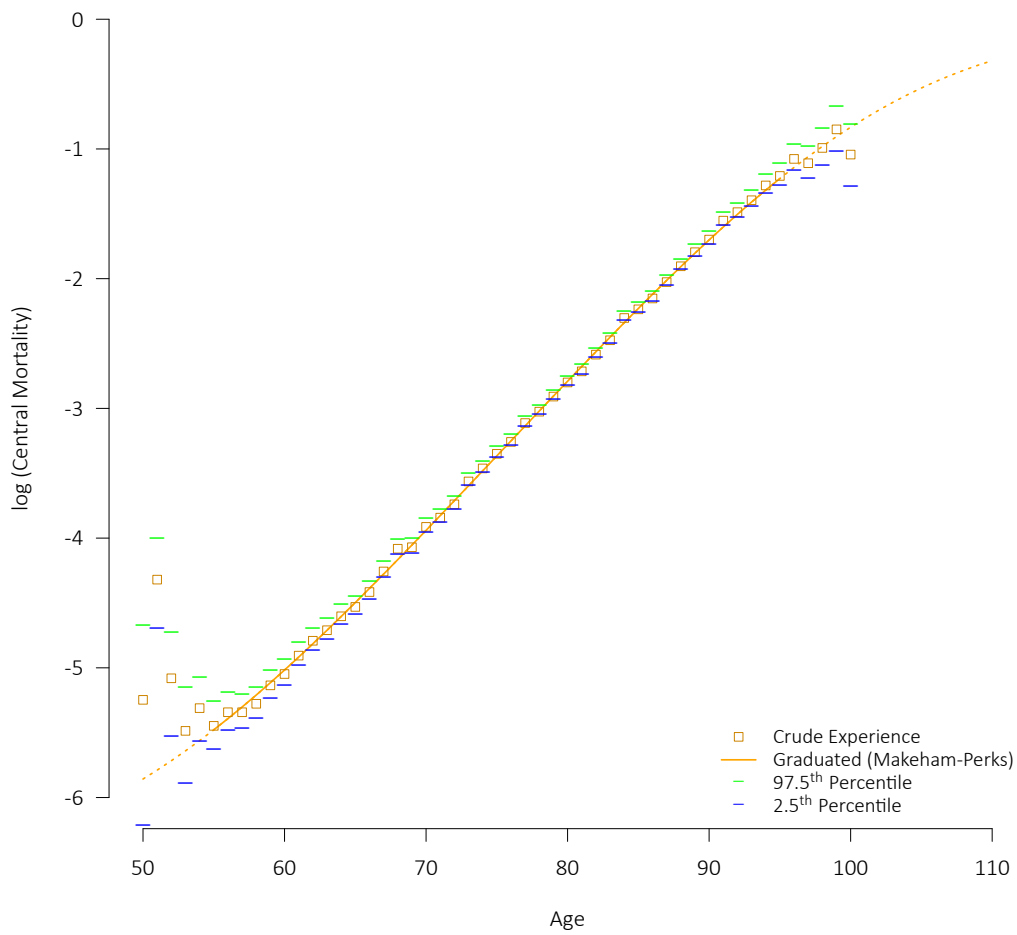
7 Smoothing through P-splines and Similar

The mortality laws discussed above are indeed very flexible and powerful, but there may be instances where the practitioner desires even greater flexibility. There are many smoothing techniques available. One which has become very popular in actuarial applications uses *B-splines* (basis splines) with difference penalties, referred to as *P-splines*. P-splines were introduced by Eilers and Marx (1996) and were extended to modelling two-dimensional mortality data by Currie et al. (2004). The basic concept is to carry out a regression to generate a smooth function by "splicing together" piecewise polynomials at certain intervals, with the meeting points called *knots*.[19]

We can find an expression for $log(\mu_x)$ using B-splines $B_i(x)$ with coefficients θ_i as follows:

$$log\left(\mu_x\right) = \sum_{i=1}^{s} \theta_i B_i(x) \tag{9}$$

The number of splines required (s) is determined by the knot-spacing chosen and the age range under consideration. The B-splines $B_i(x)$ are usually piecewise cubic polynomials.

By adding a penalty function to the log-likelihood we can force the splines into a chosen level of smoothness. That penalty function can be defined in terms of finite differences of order N with respect to the coefficients θ_i as:

$$P(\theta) = \sum_{i=1}^{s-N} \left(\Delta_{\theta_i}^N\right)^2$$
$$= \sum_{i=1}^{s-N} \left(\sum_{j=0}^{N} \binom{N}{j} (-1)^{N-j} \theta_{i+j}\right)^2 \tag{10}$$

For example, with the common choice of finite differences of order two ($N = 2$) the penalty function takes this form:

$$P(\theta) = (\theta_1 - 2\theta_2 + \theta_3)^2 + ... + (\theta_{s-2} - 2\theta_{s-1} + \theta_s)^2$$

The penalized log-likelihood function in terms of θ can then be expressed by a base log-likelihood function $l(\theta)$ (say as determined under equation (4)) and a *smoothing parameter* λ:

$$l_\lambda(\theta) = l(\theta) - \frac{1}{2}\lambda P(\theta) \tag{11}$$

The system of B-splines returned by maximizing equation (11) is referred to as a *penalized B-spline*, or a P-spline. The λ parameter can be selected arbitrarily or with reference to a measure that balances the absolute log-likelihood with the effective number of parameters[20] in the model such as the AIC or BIC.

With P-splines, the key modelling decisions are with respect to the knot-spacing, the finite difference order, and the smoothing parameter λ.

---

[19] In this paper, the knots are constrained to be equidistant apart, which simplifies the required calculations. However, the knot spacing need not be equidistant. Early work involving applying splines to mortality experience focused on attempts to determine the optimal number of splines and corresponding spacing, which the use of P-splines effectively circumvents at the cost of the introduction of the penalty term to the log-likelihood function.

[20] Note that by introducing the penalty function, we are reducing the degrees of freedom. This means that the greater the penalty, the lower the "effective" number of parameters.

The mortality (whether the q_x, m_x, or a logarithm of either) can also be fitted using a "mechanical" method that precedes modern statistics that is usually referred to as Whittaker-Henderson. This method of graduation was first presented in Whittaker (1923). Retaining much of the parameterization above, the Whittaker-Henderson smoothing of the crude values y_x (whether those be m_x, the q_x, or other) over k ages, subject to smoothing parameter λ, optional weights w_x, and finite difference order N, are those θ_x that *minimize* the following expression:

$$\sum_{x=1}^{k} w_x \left(y_x - \theta_x\right)^2 + \lambda \sum_{x=1}^{k-N} \left(\Delta_{\theta_x}^N\right)^2 \tag{12}$$

Equation (12) is reminiscent of the penalized log-likelihood function under equation (11). Indeed, as discussed in Currie (2015), if the knot-spacing in a P-spline is set equal to a single year-of-age, the order of the underlying B-splines is set to zero, and the Poisson log-likelihood function replaced with the (potentially weighted) least-squares difference, the Whittaker-Henderson result is equivalent to this special case of P-splines.

Let us compare the graduations obtained through the graduation by the mathematical formula process above to what we would get using other approaches. Figure 12 compares the resulting graduations under three methods for the CPM2014 male dataset with no further stratification on a lives basis:

- **Mathematical Formula**: The $GM(1,3)$ curve developed under Section 5.

- **P-splines**:[21] Modelling $log(m_x)$, using equidistant knot spacing of 10 years (requiring seven piecewise cubic B-splines over the 55 to 95 age range), a finite difference order (N) of 2 for the penalty function, and a smoothing parameter (λ) of 350. The value for λ was chosen as the approximate value of the smoothing parameter found to minimize the BIC under the implementation employed.

- **Whittaker-Henderson**:[22] This is the method that was used to graduate the published CPM2014 tables. For consistency, we will also model $log(m_x)$ over ages 55 to 95 with Whittaker-Henderson. The weights are based on the deaths,[23] with the sum of the weights set equal to 41 (the number of ages), the smoothing parameter (λ) to 500,[24] and the finite difference order (N) to 3.

As can be clearly seen, the three approaches result in nearly identical graduations. However, graduating by mathematical formula results in a model with only 4 parameters, as compared to 7 (physical) parameters for the P-splines and 41 "parameters" (being the 41 graduated m_x) under Whittaker-Henderson. The smoothing mechanisms under the P-splines and Whittaker-Henderson significantly decrease the *effective dimension* (or *effective degrees of freedom*) of the resulting model,[25] but a model with fewer physical pa-

---

[21]The P-splines are fitted using the R package *MortalitySmooth* (Camarda (2012)), using the function `Mort1Dsmooth` with the `ndx` parameter (the number of internal knots less one) set equal to 4 in order to achieve 10-year knot spacing over the 40 ages.

[22]The Whittaker-Henderson method is solved in R code in terms of crude observations `y`, smoothing parameter `L`, and finite difference order `N`. By setting `Id = diag(length(y))` as the identity matrix, `D = diff(Id, diff = N)` the finite differencing matrix, and `W` the vector of weightings, the smoothed results under Whittaker-Henderson are obtained directly through:

$$\texttt{solve}\left(\texttt{W} * \texttt{Id} + \texttt{L} * \texttt{crossprod}\left(\texttt{D}\right), \texttt{W} * \texttt{y}\right)$$

[23]Using deaths as the weight is a reasonable approximation to assuming Poisson-distributed deaths counts.

[24]Note that although the P-spline and Whittaker-Henderson approaches have a similar penalty term, under the implementations herein a smoothing parameter λ of 500 for Whittaker-Henderson represents a significantly higher degree of smoothing than a λ of 500 under the Poisson P-spline framework.

[25]See Eilers and Marx (1996) and Currie et al. (2004) for details. For example, the effective dimension of the P-splines fitted here is approximately 3.7, while under the Whittaker-Henderson approach it is about 6.1.

Figure 12: Graduated CPM2014 Central Mortality Rates for Males under Various Methods
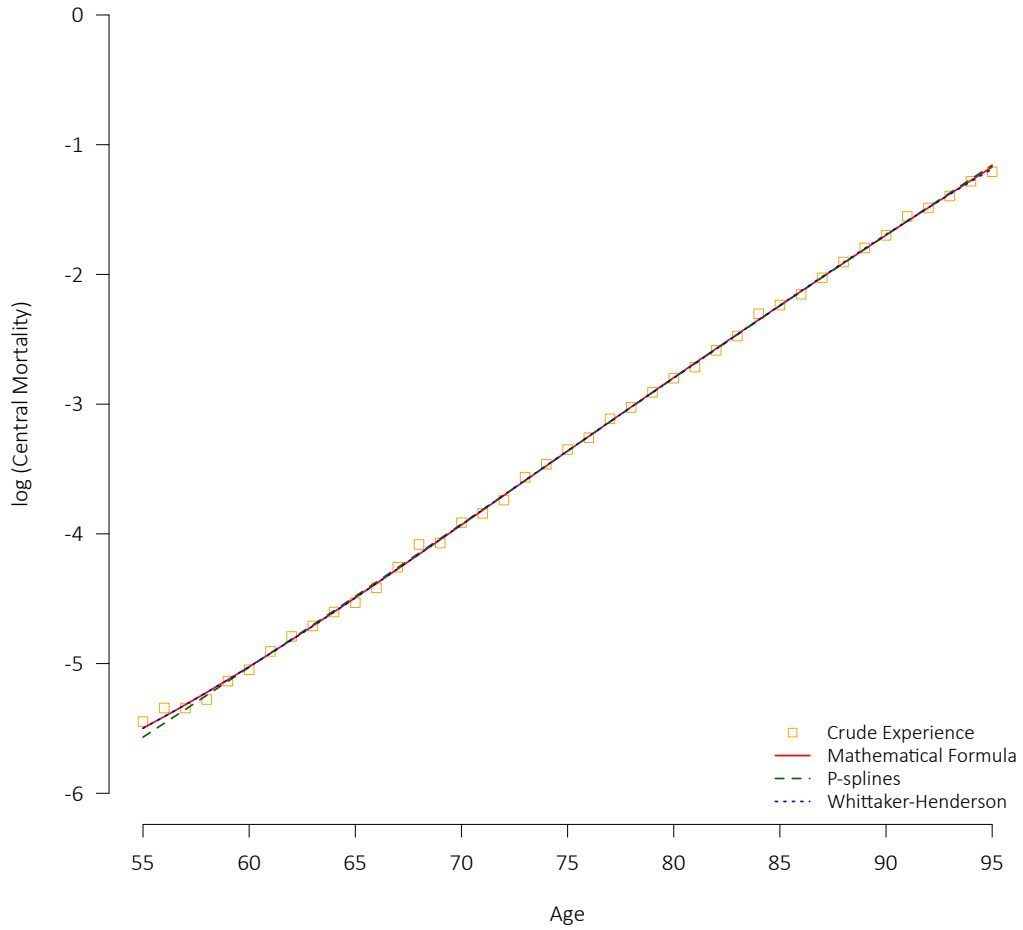
rameters with some intuitive interpretations provides considerable benefit when desiring to incorporate risk factors. Those advantages will be introduced in the following section.

8   Reflecting Risk Factors

To this point in the paper, we have limited our modelling to the CPM2014 male-only data on a lives-basis. However, age and gender are not the only risk factors that have an influence on mortality. Therefore, we need to be able to analyze the mortality differentials arising from other key risk factors. In general, it is critical to capture factors that are financially significant, which often will include the benefit amount, e.g. pension, annuity, sum insured, etc. As already discussed with respect to gender, the mortality differential will likely depend on age.

One of the potential risk factors available in the CPM2014 dataset is whether a plan covers the employees of corporations in the "private" or the "public" sector. At least in this dataset, the plans considered "private" are mainly blue-collar, while the "public" plans mainly cover public administration and education employees, meaning that the public sector mortality experience is expected to be lighter than the private sector experience. Figure 13 provides the raw central mortality rates for males in the two sectors on a lives basis, with separate Makeham-Perks (chosen for simplicity) curves fitted over ages 55 to 95 and then extended to the lower and higher ages.

The Makeham-Perks parameters for the private, public, and total graduations are provided in Table 6. In all cases, the three parameters are highly significant.

Table 6: Makeham-Perks Parameters for Male Dataset Split by Sector

Dataset	α	β	ϵ
Private	-12.621360	0.123481	-5.395322
Public	-13.207571	0.129613	-6.577481
Total	-12.684571	0.124192	-6.628663

The above is an example of stratification, where we split the experience across various classifications (in this case, we split males into "sectors") and then model the subgroups independently. However, we could instead attempt to model the experience of the risk groupings simultaneously, with certain parameters shared between the groupings but allowing the others to vary. To illustrate that concept, we turn to the female experience in the CPM2014 dataset, the corresponding raw experience split by sector found in Figure 14.

There is much less experience for females in the private sector than the public, particularly at lower ages. However, the experience across the two sectors does appear to converge well. Fundamentally, we have three choices: (a) we could model the total experience, without any distinction between the sectors, (b) we could stratify the female experience by sector and model the experience completely independently as we did above for the males, (c) or, finally, we could model the experience simultaneously, with certain parameters shared between the sectors and the remainder allowed to vary. In this case, if a Makeham-Perks model is again chosen, we could have the Makeham (ϵ) parameter shared between the female private and public experience with the intercept (α) and slope (β) parameters allowed to vary by sector, or have the intercept and slope shared but with the Makeham parameter varying by sector.

For purposes of this analysis, equation (4) must be modified to allow for the female public and private

Figure 13: CPM2014 Male Experience and Graduation across the Private & Public Sectors
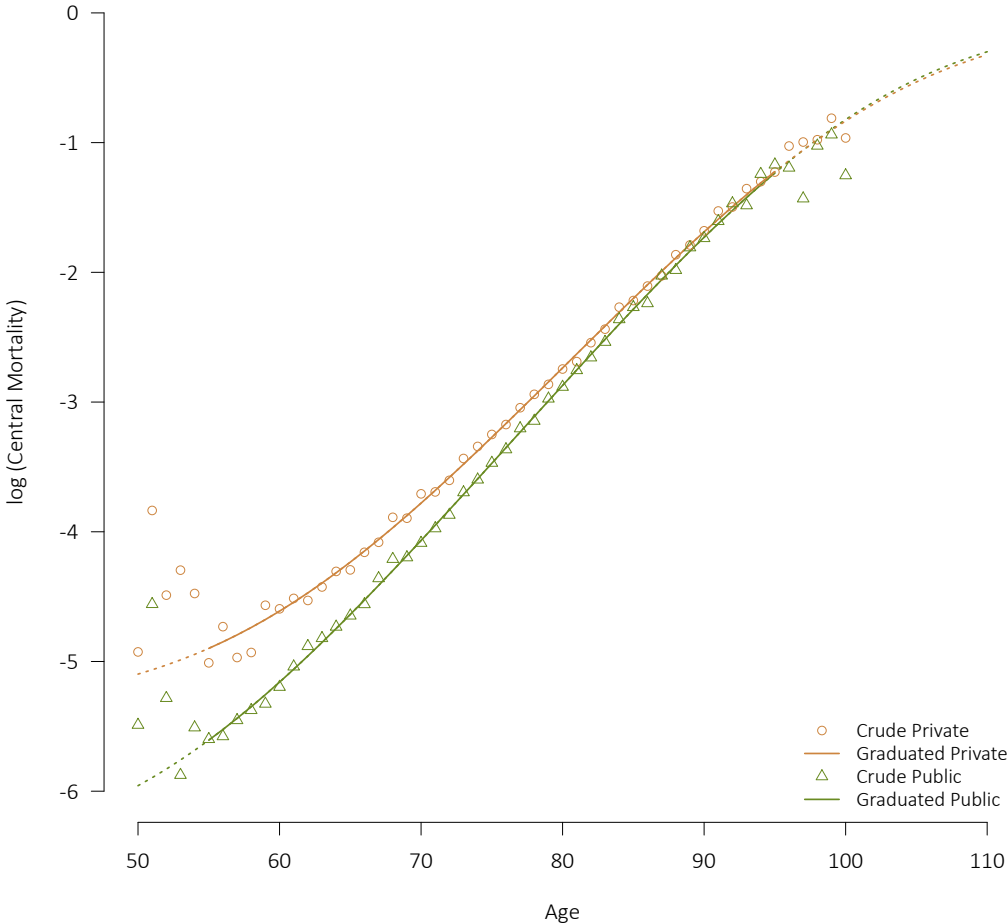
experience to be separately considered.[26] The resulting AIC and BIC scores under the four options considered are provided as Table 7.

The 4-parameter approach, where the α and β parameters are shared but the ϵ allowed to vary, results in the lowest AIC and BIC. The resulting parameters are shown in Table 8, where ϵ is the Makeham parameter for public sector females (the baseline) and $\epsilon + \epsilon{:}\mathrm{PRIVATE}$ that for those in the private sector. The fitted curves and extensions are provided as Figure 15. The interaction between the Makeham parameter and sector captures the apparent non-linear convergence between the public and private sector female mortality.

Similar logic as the above allows us to incorporate any risk factor available in a dataset. One of the most important factors is often the benefit amount applicable to an individual. Incorporating amounts in the analysis is discussed in the following section.

---

[26]Specifically, the log-likelihood function in equation (4) is replaced by $l^{pu}(\theta^{pu}) + l^{pr}(\theta^{pr})$, where the parameter sets, death counts, and exposures relate to the public sector for l^{pu} and θ^{pu} and the private sector for l^{pr} and θ^{pr}.

Figure 14: CPM2014 Female Crude Experience across the Private & Public Sectors
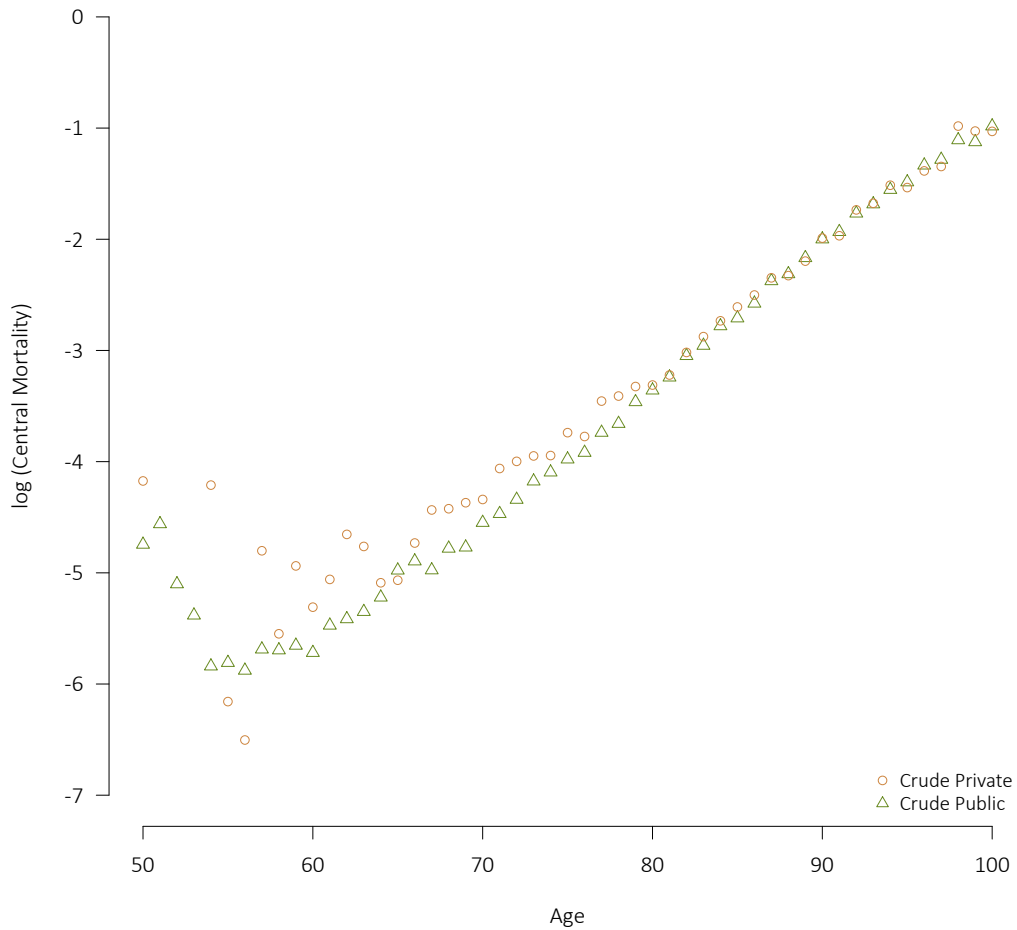
Table 7: Measures of Model Fit for Female CPM2014 Dataset under Four Options

Model	Parameters	AIC	BIC
No Distinction	3	794.34	801.56
All Parameters Varied	6	697.26	711.70
All but Makeham Varied	5	704.09	716.12
All but Makeham Shared	4	696.51	706.13

It should be noted that models applied to grouped data work well only if a small number of risk factors is considered. Incorporating a wider set of risk factors is possible at the level of the individual. For case studies

Table 8: Makeham-Perks Parameters for Female Dataset

Parameter	Estimate	Standard Error	Z	p-value	Significance
α	-14.763021	0.110190	-133.9776	0	$* * *$
β	0.142683	0.001286	110.9788	0	$* * *$
ϵ	-6.290117	0.063950	-98.3594	0	$* * *$
ϵ:**PRIVATE**	1.033823	0.078870	13.1079	$2.9658e-39$	$* * *$

Figure 15: Graduation and Extension of CPM2014 Female Rates Split by Sector
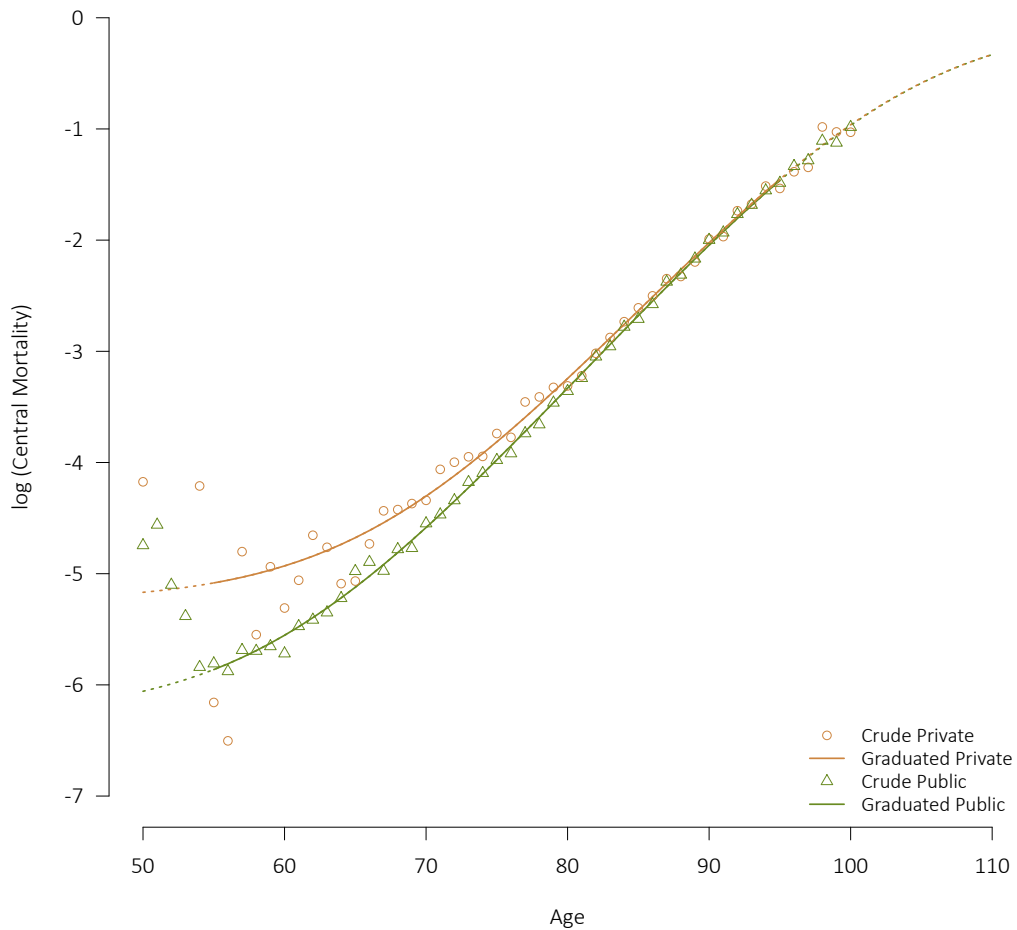
of applying similar concepts on an individual basis, the reader is referred to Richards (2008) and Richards et al. (2013).

9   Concerning Amounts

In actuarial practice, the aim of mortality analysis is to quantify the financial impact of death or survivorship. Therefore, the mortality rate for an insured life or the survival probability of an annuitant or pensioner will be multiplied by a sum at risk or benefit amount to calculate the resulting payout of a life insurance policy, annuity, or pension.

In addition to this, a correlation is observed between benefit amount or sum insured and mortality. Mortality often is lighter for those with higher of these amounts, and this skew in the mortality by amounts has direct financial implications. To illustrate the point, the experience across the thirteen amounts bands available within the CPM2014 dataset for males will be analyzed.[27]

We apply a lives-based Makeham-Perks model to the CPM2014 male dataset for ages 65 to 95. This constitutes the *expected* mortality basis for the actual-to-expected (A/E) ratios across all thirteen pension amount bands as provided under Figure 16, along with the corresponding deviance residuals. The deviance residuals are scaled to reflect the amount of experience available under each band, and can effectively be considered standardized or weighted versions of the A/E ratios.

Figure 16 implies that using the average lives-based mortality rates graduated without accounting for mortality differentials by amounts would lead to an underestimation of mortality for lives with lower than average amounts and an overestimation of mortality for lives with higher than average amounts. Considering a financial application such as life insurance or pensions, the lives with the greater amounts also have a greater financial impact. Therefore, on an amounts-weighted basis, the overall impact of mortality is lower than on a lives-weighted basis; in other words, weighting the mortality rates by sum assured or benefit amount leads to a lower average rate of mortality.[28]

As seen in Figure 17, the simple amounts-weighting shifts the average mortality lower, but does not overcome the bias that the mortality of lives with lower than average amounts will be underestimated and overestimated for lives with higher than average amounts. For short-term applications, such a crude approximation may be sufficient. It must always be top of mind that the best model is always the simplest that manages to do the job. However, there are instances in which it will be necessary to improve upon the model and increase its complexity to make it useful.

Consider pensions for example. As stated above, applying an amounts-weighted mortality assumption will overestimate mortality rates for lives with the greatest pensions. Mortality impacts upon pension liabilities, because these signify the amount released to pay ongoing benefits when a pensioner dies. Fewer deaths than expected will lead to a deficit, which means that applying an average mortality assumption leads to a cross-subsidy between pensioners with lower amounts and those with higher amounts. As more and more pensioners with lower pensions and higher than average mortality rates die, the total pension plan gradually accumulates a growing deficit if the basis is not updated frequently.

Such systematic bias can be avoided by grouping lives with similar mortality rates together and applying these different mortality rates in the liability calculation. As long as each group consists of persons with

---

[27] The experience data is not republished herein, but is available at the CPM2014 data tool at `http://www.cia-ica.ca/docs/default-source/2014/214013t5.zip`. As the bridging benefits payable to age 65 are unfortunately included while they are in pay, we restrict our analysis to ages 65 and above (see Appendix G for some related details).

[28] Reference can be made to Appendix F for details on how to extend the same statistical framework developed above for lives-based data to amounts-weighted experience, including a continuation of the case study with the CPM2014 dataset illustrating the graduation of the corresponding amounts-weighted data.

Figure 16: A/E Ratios and Residuals for the CPM2014 Dataset by Band, Lives Basis
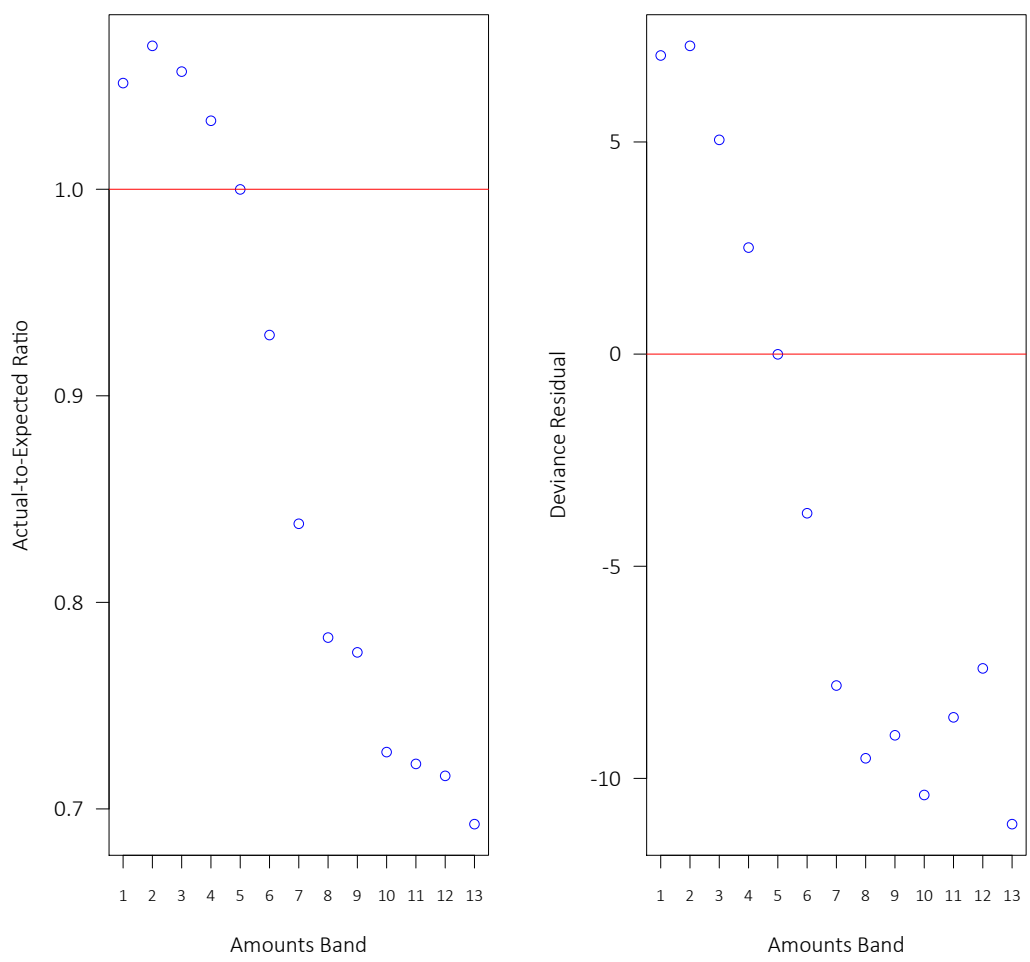
sufficiently similar mortality, which we call a *homogeneous* group with respect to mortality, the threat of a deficit due to inappropriate mortality assumptions is assuaged.

These considerations also give us a method of determining whether the right groupings have been chosen: so long as the residuals for a particular group appear to be randomly distributed, i.e. the observed rates fluctuate randomly around the estimated mortality rates, the group can be considered homogeneous.

From the residuals under Figure 16, there is little to distinguish the top seven bands (i.e., bands 7 through 13) from each other, and the bottom 4 bands (i.e., bands 1 through 4) also appear to belong together based on their residuals and A/E ratios. Three aggregate bandings will be considered: **Band A**: bands 1 through 4, **Band B**: bands 5 and 6, and **Band C**: bands 7 though 13.[29] The graduations (again using a Makeham-Perks model for males of ages 65 to 95) for the three aggregate bands are shown in Figure 18 on both a lives and an amounts basis. Note that (1) the three bandings each exhibit their own distinct age structure and (2) the

---

[29]These choices, particularly the decision to keep bands 5 and 6 in a separate group, are just for illustration, and different groupings may be just as or more appropriate.

Figure 17: A/E Ratios and Residuals for the CPM2014 Dataset by Band, Amounts Basis
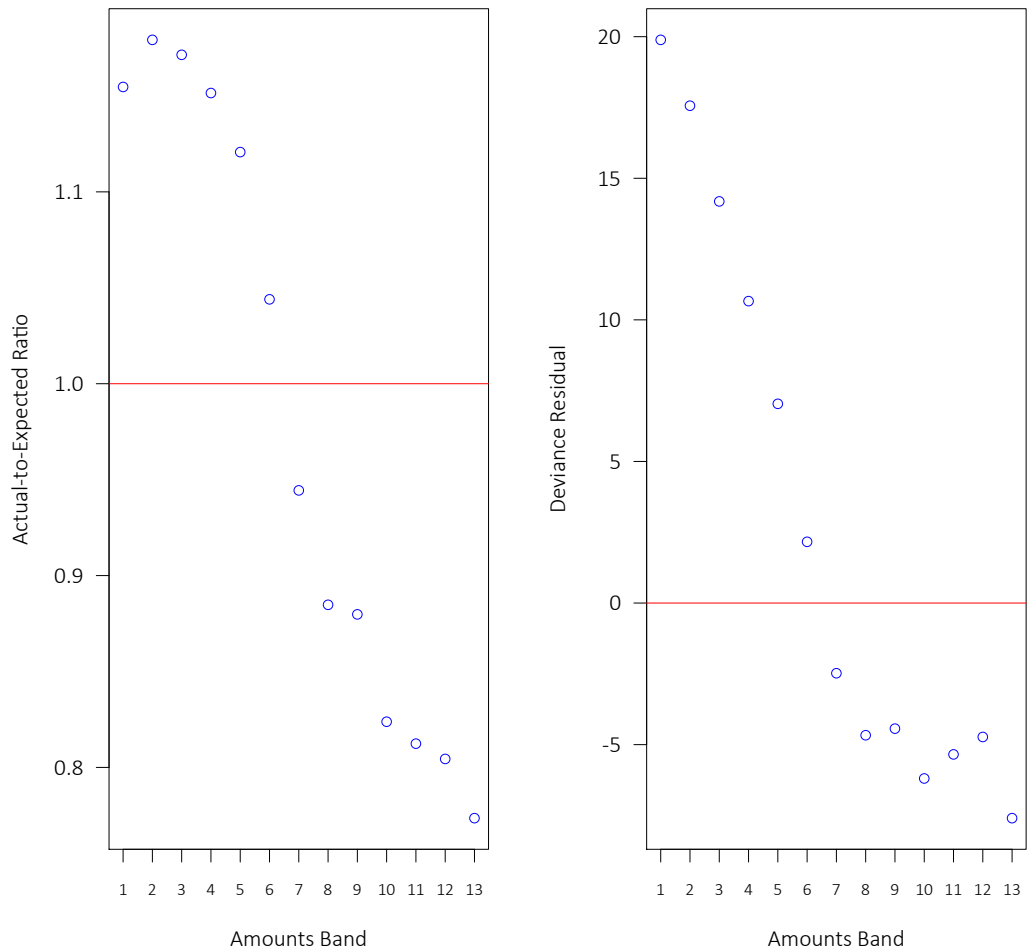
lives and amounts graduations are close, implying homogeneity within the bands.

Conversely, Figure 17 demonstrates that using only a single group and adjusting the mortality for all lives in the same direction, i.e. by applying an amounts-weighting, introduces bias which may undermine the purpose of the analysis, i.e. to predict the financial outcome of the group's survivorship. This is one of the most basic instances of model risk, which is unfortunately often ignored.

Figure 18: Central Mortality over Aggregate Amounts Bandings for the CPM2014 Dataset
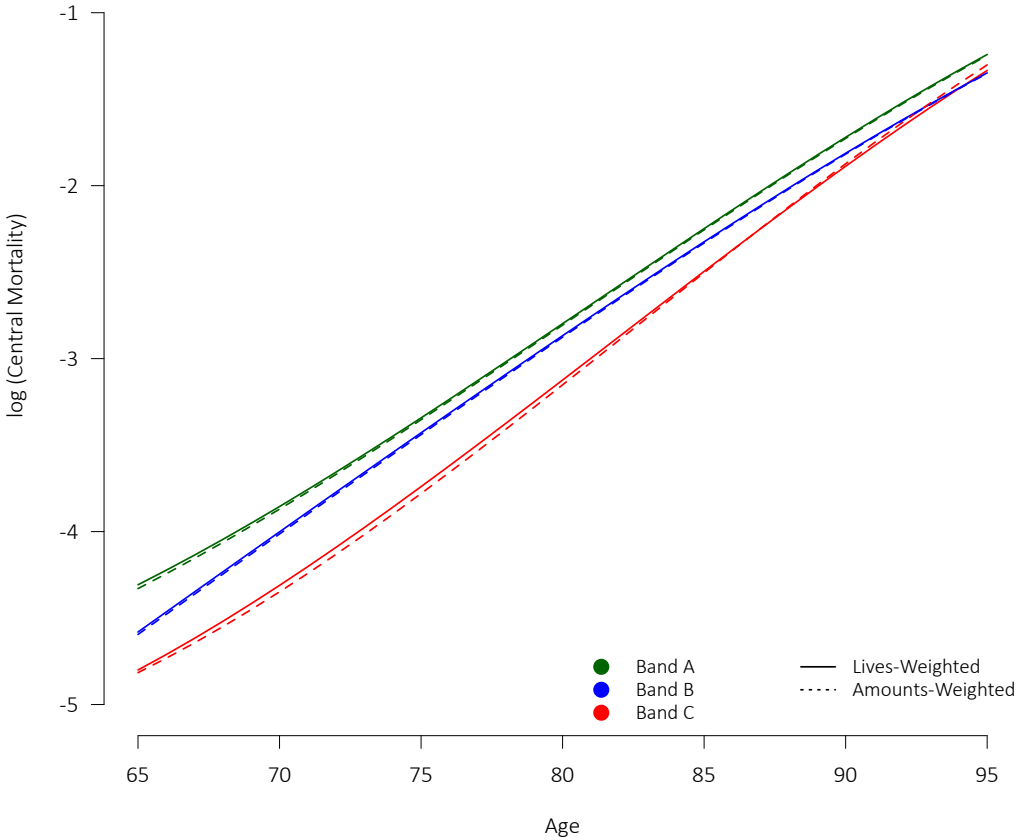
10   Conclusion

We have seen how mortality experience is well captured through simple parametric laws, and how those laws can be fitted to the experience data using maximum likelihood estimation. The resulting modelling framework allows for a variety of useful statistical tests to quantify how well the model fits the experience data, particularly the deviance residual analysis. These tests reveal whether the graduation was performed correctly and also whether the model can be deemed sufficiently complex or additional investigation is warranted. By their nature, parametric models allow for the incorporation of multiple risk factors, while simultaneously reflecting their age-dependence. An important advantage of parametric models is that their parameter estimates have explanatory value and allow an intuitive interpretation.

Graduation of mortality experience by mathematical formula is a valuable tool for any actuary whose work involves life contingencies. By employing the concepts explored in this paper, hopefully this tool is made more accessible to practitioners.

References

Beard, R. E. (1959), Note on some mathematical mortality models, *in* G. E. W. Wolstenholme and M. O'Connor, eds, 'The Lifespans of Animals', Little, Brown & Co., Boston, MA., pp. 302–311.

Caeiro, F. and Mateus, A. (2014), *randtests: Testing randomness in R*. R package version 1.0.
URL: *https://CRAN.R-project.org/package=randtests*

Camarda, C. G. (2012), 'MortalitySmooth: An R package for smoothing Poisson counts with P-splines', *Journal of Statistical Software* **50**(1), 1–24.
URL: *http://www.jstatsoft.org/v50/i01/*

Chan, L. K. and Panjer, H. H. (1979), 'A statistical approach to graduation by mathematical formula', *Actuarial Research Clearing House (ARCH)* **1979.2**, 27–60.

CHMD: Canadian Human Mortality Database (2014), Department of Demography, Université de Montréal (Canada). Available at `http://www.bdlc.umontreal.ca/chmd/` (data downloaded on 2017-10-22).

CIA Pension Experience Subcommittee – Research Committee (2014), *Final Report – Canadian Pensioners' Mortality*, Canadian Institute of Actuaries.

Continuous Mortality Investigation – Self-administered Pension Schemes Mortality Committee (2014), 'Graduations of the CMI SAPS 2004–2011 Mortality Experience Based on Data Collected by 30 June 2012', ***Working Paper 71*** .

Continuous Mortality Investigation – Graduation and Modelling Working Party (2015), 'Report of the Graduation and Modelling Working Party', ***Working Paper 77*** *(with accompanying CMI graduation software v0.2 and user guide)* .

Continuous Mortality Investigation – Annuities Committee (2015), 'Proposed "08" Series annuities tables', ***Working Paper 78*** .

Continuous Mortality Investigation – High Age Working Party (2015), 'Initial report on the features of high age mortality', ***Working Paper 85*** .

Continuous Mortality Investigation – High Age Working Party (2017), 'A second report on high age mortality', ***Working Paper 100*** .

Currie, I. D. (2015), 'Back to the future with Whittaker smoothing'.
URL: *https://www.longevitas.co.uk/site/informationmatrix/whittaker.html*

Currie, I. D., Durban, M., and Eilers, P. H. C. (2004), 'Smoothing and forecasting mortality rates', *Statistical Modelling* **4**, 279–298.

Eilers, P. H. C. and Marx, B. D. (1996), 'Flexible smoothing with B-splines and penalties', *Statistical Science* **11**(2), 89–102.

Forfar, D. O., McCutcheon, J. J., and Wilkie, A. D. (1988), 'On graduation by mathematical formula', *Journal of the Institute of Actuaries* **115**, 1–149.

Gilbert, P. and Varadhan, R. (2016), *numDeriv: Accurate Numerical Derivatives*. R package version 2016.8-1.
URL: *https://CRAN.R-project.org/package=numDeriv*

Gompertz, B. (1825), 'On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies', *Philosophical Transactions of the Royal Society of London* **115**, 513–585.

Li, J. S. H., Hardy, M. R., and Tan, K. S. (2008), 'Threshold life tables and their applications', *North American Actuarial Journal* **12**(2), 99–115.

Longevitas Development Team (2017), *Longevitas 2.7.7*, Longevitas Ltd, Edinburgh, United Kingdom.
URL: *https://www.longevitas.co.uk*

Makeham, W. M. (1860), 'On the law of mortality and the construction of annuity tables', *The Assurance Magazine, and Journal of the Institute of Actuaries* **8**, 301–310.

Makeham, W. M. (1867), 'On the law of mortality', *Journal of the Institute of Actuaries* **13**, 325–358.

Makeham, W. M. (1890), 'On the further development of Gompertz's law', *Journal of the Institute of Actuaries* **28**, 152–159, 185 − 192, 316 − 332.

Panjer, H. H. and Tan, K. S. (1995), *Graduation of Canadian Individual Insurance Mortaltiy Experience: 1986–1992*, Canadian Institute of Actuaries.

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

R Core Team (2017), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org*

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

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

Richards, S. J., Kaufhold, K., and Rosenbusch, S. (2013), 'Creating portfolio-specific mortality tables: a case study', *European Actuarial Journal* **3**(2), 295–319.

SOA Retirement Plans Experience Committee (2014), *RP-2014 Mortality Tables Report (with supplement)*, Society of Actuaries.

Tenenbein, A. and Vanderhoof, I. T. (1980), 'New mathematical laws of select and ultimate mortality', *Transactions of the Society of Actuaries* **32**, 119–183 (with discussion).

Whittaker, E. T. (1923), 'On a new method of graduation', *Proceedings of the Edinburgh Mathematical Society* **41**, 63–75.

Appendix A   Datasets Used as Main Case Study

The data used for purposes of the main case study are based on the original, unaltered dataset reflected in CIA (2014).

The datasets for males are provided in Table A.1, and for females in Table A.2. There are separate datasets with the "private sector", "public sector", and total experience, both on a lives (unweighted) and amounts (weighted by pensions-in-pay) basis. The data consists of the deaths counts and the *initial* exposure.

The datasets were extracted from the data summary tool provided publicly by the Canadian Institute of Actuaries in 2014.[30] The data reflected below is **not** equivalent to that used in the graduations of the tables published in CIA (2014). In particular, the graduations published reflect numerous adjustments to the original datasets. The data summary tool permits direct extraction of the unaltered dataset, with the exception of the application of an adjustment for incurred-but-not-reported (IBNR) deaths described in CIA (2014). For producing the datasets below, that adjustment was backed out, by decreasing the death counts (both on lives and amounts bases) in each individual year of experience by the appropriate factor. That process appears to have been successful, as the death counts on a lives basis decrease slightly and become integers at each age, as opposed to the fractional death counts provided in the direct extracts.

Some additional details and comments regarding the adjustments made to the datasets used in graduating the published CPM2014 tables are found in Appendix G.

The authors of this paper had no part in the collection, validation, or tabulation of the CPM2014 datasets, and therefore cannot speak to the completeness or appropriateness of the corresponding testing, adjustments, exposure calculation, etc. The datasets are merely meant for purposes of illustration only, and as such the information should not necessarily be relied upon for any specific purpose.

---

[30]Available online at `http://www.cia-ica.ca/docs/default-source/2014/214013t5.zip`.

Table A.1: Male Datasets Invoked as Case Study

Age	Total Dataset				Private Sector				Public Sector			
	Lives		Amounts		Lives		Amounts		Lives		Amounts	
	Deaths	Initial Exp.	Deaths	Initial Exp.	Deaths	Initial Exp.	Deaths	Initial Exp.	Deaths	Initial Exp.	Deaths	Initial Exp.
50	8.00	1,523.64	163.01	40,296.73	4.00	553.43	79.38	12,631.69	4.00	970.21	83.63	27,665.05
51	34.00	2,573.63	774.71	76,295.63	14.00	655.56	286.12	14,947.56	20.00	1,918.07	488.59	61,348.07
52	27.00	4,357.73	484.09	140,919.95	9.00	805.48	163.22	18,757.72	18.00	3,552.25	320.87	122,162.23
53	32.00	7,737.12	718.32	271,290.94	13.00	960.79	208.27	23,365.19	19.00	6,776.32	510.05	247,925.75
54	68.00	13,804.35	1,678.14	508,299.94	19.00	1,678.94	391.79	45,991.34	49.00	12,125.41	1,286.35	462,308.60
55	117.00	27,218.52	3,052.20	970,104.05	37.00	5,567.78	879.08	169,712.47	80.00	21,650.74	2,173.12	800,391.58
56	180.00	37,712.77	4,943.92	1,337,728.67	66.00	7,521.49	1,506.82	222,879.34	114.00	30,191.29	3,437.10	1,114,849.34
57	216.00	45,254.96	6,745.37	1,608,443.22	59.00	8,521.68	1,697.87	240,631.64	157.00	36,733.28	5,047.50	1,367,811.58
58	261.00	51,240.61	7,864.32	1,827,856.72	68.00	9,449.78	1,546.34	255,173.36	193.00	41,790.82	6,317.99	1,572,683.36
59	327.00	55,714.27	9,195.93	1,979,945.01	107.00	10,348.44	2,265.18	264,906.57	220.00	45,365.83	6,930.75	1,715,038.44
60	393.00	61,379.12	11,232.48	2,126,245.80	120.00	11,928.32	2,427.57	284,598.77	273.00	49,450.81	8,804.91	1,841,647.03
61	471.00	63,872.04	12,722.14	2,158,791.59	144.00	13,213.54	2,729.45	295,383.55	327.00	50,658.51	9,992.69	1,863,408.04
62	523.00	63,261.81	14,343.41	2,084,253.12	153.00	14,259.73	2,745.46	299,501.49	370.00	49,002.08	11,597.95	1,784,751.63
63	570.00	63,586.11	14,761.94	2,039,342.38	183.00	15,382.43	3,256.28	305,140.24	387.00	48,203.67	11,505.66	1,734,202.13
64	644.00	64,562.76	16,881.09	1,869,058.53	226.00	16,863.32	4,032.13	307,856.23	418.00	47,699.44	12,848.96	1,561,202.31
65	764.00	71,295.59	15,693.47	1,700,679.15	282.00	20,796.11	3,669.51	317,618.27	482.00	50,499.48	12,023.97	1,383,060.88
66	868.00	72,265.38	16,857.21	1,639,990.27	351.00	22,627.12	4,516.77	325,926.10	517.00	49,638.26	12,340.44	1,314,064.16
67	997.00	70,896.20	18,308.42	1,548,993.29	398.00	23,765.21	4,903.52	328,839.92	599.00	47,130.99	13,404.90	1,220,153.37
68	1,169.00	69,866.44	20,127.23	1,464,072.63	512.00	25,265.69	6,312.45	338,356.03	657.00	44,600.75	13,814.78	1,125,716.60
69	1,166.00	68,890.27	19,547.63	1,382,862.30	535.00	26,575.92	6,289.16	344,184.55	631.00	42,314.35	13,258.48	1,038,677.75
70	1,350.00	68,275.90	22,115.15	1,315,781.26	680.00	28,055.12	8,194.68	352,867.80	670.00	40,220.78	13,920.47	962,913.46
71	1,434.00	67,592.34	22,107.61	1,248,435.55	722.00	29,353.59	8,005.12	357,201.62	712.00	38,238.74	14,102.50	891,233.92
72	1,585.00	67,496.84	24,047.02	1,197,450.35	828.00	30,823.67	9,110.31	364,693.46	757.00	36,673.17	14,936.71	832,756.89
73	1,869.00	66,870.80	26,932.20	1,140,117.58	1,014.00	31,978.22	11,019.76	369,076.37	855.00	34,892.58	15,912.44	771,041.21
74	2,037.00	66,019.25	29,407.71	1,089,174.58	1,142.00	32,866.42	12,011.58	370,918.47	895.00	33,152.84	17,396.12	718,256.10
75	2,227.00	64,607.63	30,574.41	1,029,504.36	1,268.00	33,326.96	13,577.85	365,959.87	959.00	31,280.67	16,996.56	663,544.48
76	2,359.00	62,543.03	31,989.20	968,669.22	1,367.00	33,334.53	13,575.47	356,574.72	992.00	29,208.50	18,413.74	612,094.50
77	2,617.00	60,154.44	37,015.92	901,547.30	1,536.00	33,002.61	16,560.83	343,470.30	1,081.00	27,151.83	20,455.09	558,077.00
78	2,706.00	57,094.10	34,400.21	826,518.43	1,654.00	32,130.10	15,329.03	323,686.82	1,052.00	24,964.00	19,071.18	502,831.61
79	2,870.00	54,062.55	35,527.47	755,843.96	1,732.00	31,216.64	15,922.42	304,953.51	1,138.00	22,845.91	19,605.05	450,890.45
80	2,961.00	50,172.25	36,020.10	683,782.54	1,840.00	29,570.25	17,085.87	280,997.67	1,121.00	20,602.00	18,934.24	402,784.87
81	2,957.00	46,118.03	35,207.58	613,895.41	1,823.00	27,721.03	16,288.31	255,783.74	1,134.00	18,397.00	18,919.27	358,111.67
82	3,026.00	41,665.32	36,463.64	541,750.66	1,930.00	25,494.65	17,151.26	227,081.04	1,096.00	16,170.67	19,312.38	314,669.62
83	2,989.00	37,001.14	34,700.26	466,897.45	1,913.00	22,862.14	15,570.50	193,988.90	1,076.00	14,139.00	19,129.76	272,908.55
84	3,103.00	32,599.46	34,174.77	399,075.35	1,993.00	20,261.46	15,423.60	163,850.78	1,110.00	12,338.00	18,751.18	235,224.56
85	2,842.00	28,016.42	32,616.55	334,400.41	1,818.00	17,604.42	14,426.71	136,837.09	1,024.00	10,412.00	18,189.84	197,563.32
86	2,620.00	23,861.40	28,530.42	277,838.17	1,740.00	15,164.40	12,872.51	113,644.81	880.00	8,697.00	15,657.91	164,193.36
87	2,453.00	19,838.07	26,099.73	224,381.67	1,567.00	12,676.07	10,767.31	90,238.32	886.00	7,162.00	15,332.42	134,143.34
88	2,250.00	16,233.58	22,824.50	178,291.41	1,516.00	10,536.58	9,812.85	71,163.21	734.00	5,697.00	13,011.65	107,128.20
89	1,972.00	12,866.61	20,310.41	139,197.86	1,307.00	8,477.61	8,811.33	55,126.34	665.00	4,389.00	11,499.08	84,071.52
90	1,725.00	10,293.08	17,348.64	109,662.55	1,165.00	6,832.08	7,213.84	42,436.26	560.00	3,461.00	10,134.80	67,226.29
91	1,519.00	7,936.15	15,241.53	83,764.44	1,037.00	5,294.15	6,062.37	31,254.37	482.00	2,642.00	9,179.16	52,510.07
92	1,196.00	5,889.68	12,409.56	61,069.45	793.00	3,937.68	4,172.31	22,209.57	403.00	1,952.00	8,237.24	38,859.89
93	943.00	4,277.85	8,515.43	43,422.99	661.00	2,892.85	3,474.60	16,199.82	282.00	1,385.00	5,040.83	27,223.16
94	732.00	3,002.27	7,642.06	30,498.48	485.00	2,023.27	2,830.44	10,952.34	247.00	979.00	4,811.62	19,546.14
95	533.00	2,050.67	5,609.17	20,262.22	353.00	1,380.67	1,768.80	7,034.90	180.00	670.00	3,840.37	13,227.32
96	406.00	1,394.67	3,864.07	13,189.58	291.00	957.67	1,508.26	4,906.33	115.00	437.00	2,355.82	8,283.25
97	257.00	908.26	1,766.41	8,117.89	200.00	641.26	968.86	3,183.00	57.00	267.00	797.54	4,934.89
98	186.00	594.67	1,593.07	5,662.13	130.00	410.67	574.02	2,047.97	56.00	184.00	1,019.05	3,614.16
99	129.00	366.00	1,414.70	3,664.57	93.00	256.00	528.62	1,242.13	36.00	110.00	886.08	2,422.44
100	65.00	217.00	473.98	1,980.66	49.00	153.00	252.88	738.34	16.00	64.00	221.10	1,242.32

Table A.2: Female Datasets Invoked as Case Study

Age	Total Dataset Lives Deaths	Total Dataset Lives Initial Exp.	Total Dataset Amounts Deaths	Total Dataset Amounts Initial Exp.	Private Sector Lives Deaths	Private Sector Lives Initial Exp.	Private Sector Amounts Deaths	Private Sector Amounts Initial Exp.	Public Sector Lives Deaths	Public Sector Lives Initial Exp.	Public Sector Amounts Deaths	Public Sector Amounts Initial Exp.
50	14.00	1,567.08	225.79	28,678.72	1.00	65.46	14.16	631.56	13.00	1,501.62	211.63	28,047.16
51	28.00	2,783.76	496.27	60,747.87	0.00	92.10	0.00	805.19	28.00	2,691.66	496.27	59,942.68
52	35.00	5,896.01	610.94	161,897.69	0.00	132.83	0.00	1,274.27	35.00	5,763.18	610.94	160,623.42
53	50.00	11,054.98	1,090.91	342,225.47	0.00	156.25	0.00	1,398.11	50.00	10,898.72	1,090.91	340,827.36
54	55.00	18,106.45	1,435.30	577,760.76	3.00	203.76	22.47	1,981.61	52.00	17,902.68	1,412.83	575,779.15
55	100.00	33,488.06	2,859.17	958,761.14	1.00	472.99	9.65	5,654.60	99.00	33,015.07	2,849.52	953,106.54
56	127.00	45,727.87	3,146.28	1,262,957.09	1.00	667.96	5.50	7,948.29	126.00	45,059.91	3,140.78	1,255,008.79
57	187.00	54,072.89	4,553.46	1,469,633.01	7.00	855.45	55.24	9,640.28	180.00	53,217.44	4,498.22	1,459,992.74
58	203.00	60,311.47	5,323.64	1,611,643.66	4.00	1,029.17	4.72	11,133.52	199.00	59,282.29	5,318.92	1,600,510.14
59	230.00	64,416.59	5,239.72	1,680,603.87	9.00	1,259.60	36.79	12,669.54	221.00	63,156.99	5,202.93	1,667,934.34
60	239.00	72,052.64	5,420.04	1,787,823.91	8.00	1,620.93	78.26	13,979.28	231.00	70,431.71	5,341.79	1,773,844.62
61	317.00	74,672.17	6,199.35	1,787,072.21	12.00	1,895.40	94.89	15,226.86	305.00	72,776.77	6,104.45	1,771,845.35
62	332.00	72,397.86	6,810.03	1,683,193.58	20.00	2,110.52	114.68	15,652.39	312.00	70,287.34	6,695.34	1,667,541.19
63	346.00	71,154.79	6,882.82	1,606,896.62	20.00	2,351.19	139.71	16,610.53	326.00	68,803.60	6,743.11	1,590,286.09
64	384.00	70,848.39	7,846.25	1,383,390.72	17.00	2,767.59	123.97	17,955.34	367.00	68,080.80	7,722.28	1,365,435.38
65	517.00	75,598.32	9,769.03	1,185,565.99	24.00	3,818.76	170.11	20,228.80	493.00	71,779.55	9,598.92	1,165,337.19
66	552.00	73,281.95	7,548.87	1,101,589.87	38.00	4,332.43	159.96	21,239.32	514.00	68,949.52	7,388.91	1,080,350.54
67	494.00	68,453.18	7,054.03	999,650.01	55.00	4,665.56	188.69	22,291.26	439.00	63,787.62	6,865.34	977,358.74
68	553.00	64,039.74	7,422.12	903,047.32	60.00	5,033.66	205.62	23,438.85	493.00	59,006.07	7,216.50	879,608.48
69	532.00	60,374.54	6,919.59	822,984.08	68.00	5,406.51	344.21	25,019.83	464.00	54,968.02	6,575.38	797,964.25
70	613.00	56,900.43	7,262.46	746,746.20	75.00	5,791.81	347.51	25,890.57	538.00	51,108.62	6,914.95	720,855.63
71	647.00	53,645.64	7,230.18	680,966.39	107.00	6,267.22	483.32	27,597.89	540.00	47,378.42	6,746.87	653,368.50
72	697.00	51,162.56	8,748.00	629,455.96	121.00	6,650.49	589.99	28,768.57	576.00	44,512.08	8,158.01	600,687.39
73	767.00	48,555.32	8,190.07	573,936.32	134.00	7,017.32	483.22	30,250.40	633.00	41,538.00	7,706.84	543,685.93
74	785.00	46,356.24	7,982.06	532,135.14	141.00	7,357.41	611.49	31,512.00	644.00	38,998.83	7,370.57	500,623.14
75	858.00	44,213.45	8,438.89	493,830.98	181.00	7,704.29	697.19	32,637.87	677.00	36,509.16	7,741.69	461,193.11
76	850.00	41,976.82	8,578.92	455,037.29	181.00	7,972.07	745.65	33,681.12	669.00	34,004.75	7,833.26	421,356.17
77	999.00	39,853.32	9,184.35	418,389.53	254.00	8,171.41	932.34	33,732.36	745.00	31,681.91	8,252.02	384,657.17
78	1,024.00	37,889.53	9,488.32	384,865.62	275.00	8,461.29	1,182.68	34,893.71	749.00	29,428.25	8,305.65	349,971.91
79	1,145.00	35,856.37	10,527.73	352,048.04	303.00	8,570.37	1,152.26	34,124.46	842.00	27,286.00	9,375.47	317,923.57
80	1,170.00	33,795.20	10,400.27	324,592.22	309.00	8,624.87	1,208.65	33,474.37	861.00	25,170.33	9,191.62	291,117.85
81	1,223.00	31,706.37	10,591.41	297,476.88	335.00	8,559.37	1,247.48	32,254.23	888.00	23,147.00	9,343.93	265,222.64
82	1,380.00	29,522.29	11,862.38	274,303.17	397.00	8,315.29	1,323.44	31,000.56	983.00	21,207.00	10,538.94	243,302.60
83	1,420.00	27,356.09	12,331.94	253,358.47	435.00	7,932.84	1,378.57	29,350.20	985.00	19,423.25	10,953.37	224,008.27
84	1,526.00	25,024.71	13,199.25	232,769.11	462.00	7,340.05	1,733.25	26,635.92	1,064.00	17,684.67	11,466.00	206,133.20
85	1,498.00	22,565.73	13,483.01	210,895.20	477.00	6,717.73	1,727.57	23,835.37	1,021.00	15,848.00	11,755.44	187,059.83
86	1,507.00	20,122.64	13,003.22	190,232.20	474.00	6,018.64	1,548.63	21,112.45	1,033.00	14,104.00	11,454.59	169,119.74
87	1,565.00	17,463.33	14,038.06	168,064.62	471.00	5,157.33	1,589.97	17,824.12	1,094.00	12,306.00	12,448.09	150,240.49
88	1,404.00	14,937.61	12,897.92	145,711.69	409.00	4,395.61	1,202.65	14,488.85	995.00	10,542.00	11,695.27	131,222.84
89	1,357.00	12,630.48	13,487.45	126,826.82	388.00	3,683.48	1,371.68	11,963.26	969.00	8,947.00	12,115.76	114,863.57
90	1,371.00	10,778.74	13,322.37	109,880.21	398.00	3,109.74	1,131.80	9,639.71	973.00	7,669.00	12,190.57	100,240.50
91	1,204.00	9,002.05	12,125.32	94,392.58	331.00	2,535.05	1,010.55	7,553.76	873.00	6,467.00	11,114.77	86,838.82
92	1,177.00	7,415.78	11,942.50	79,509.40	333.00	2,055.78	781.20	5,662.62	844.00	5,360.00	11,161.30	73,846.78
93	995.00	5,845.48	10,544.89	64,028.22	273.00	1,596.06	731.22	4,345.70	722.00	4,249.42	9,813.67	59,682.52
94	872.00	4,515.64	9,238.05	50,089.08	239.00	1,205.64	547.30	3,051.40	633.00	3,310.00	8,690.75	47,037.68
95	674.00	3,352.86	7,239.26	37,949.34	169.00	868.86	412.97	2,209.05	505.00	2,484.00	6,826.30	35,740.30
96	567.00	2,465.36	6,672.40	28,202.60	142.00	638.36	319.91	1,576.88	425.00	1,827.00	6,352.49	26,625.72
97	418.00	1,741.01	4,694.82	20,228.76	103.00	447.01	279.02	1,134.63	315.00	1,294.00	4,415.80	19,094.14
98	351.00	1,201.34	3,825.55	13,768.77	102.00	323.34	198.92	800.42	249.00	878.00	3,626.63	12,968.35
99	222.00	776.01	2,456.14	8,952.19	65.00	214.01	215.47	558.28	157.00	562.00	2,240.66	8,393.90
100	156.00	500.00	1,945.67	5,935.72	40.00	132.00	107.56	327.96	116.00	368.00	1,838.11	5,607.76

Appendix B   Sample R Implementation

Much of the modelling and graphing in this paper was conducted with R code nearly identical to what is provided below. The code is provided in the interest of reproducibility, and for interested parties to further explore the concepts presented in this paper on their own.

What is provided below is simply a sample implementation of how one might approach various items. While care was taken to ensure that the code works as would be expected over a wide variety of scenarios, as it was meant to be a relatively basic implementation there are certain to be situations where the code fails or the result returned not correct. Thus, it should not be relied upon without independent verification.

Where feasible, the results herein were checked against those generated by the group counts module in Longevitas (Longevitas Development Team (2017)) and the CMI graduation software released alongside CMI WP 077 (2015). Any errors remain the responsibility of the authors.

The implementation relies mainly on the standard R packages, with the exceptions of the numerical derivatives package *numDeriv* (Gilbert and Varadhan (2016) and the package *randtests* (Caeiro and Mateus (2014) for direct access to a function providing the precise calculation of the p-value under the runs test.

THE CODE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT. IN NO EVENT SHALL THE AUTHORS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

```
 1  ###########################################################################
 2  #
 3  #   GRADUATE.POISSON
 4  #
 5  #       Sample implementation of procedure to graduate mortality experience
 6  #       grouped by age using the Poisson model and certain mortality laws
 7  #
 8  ###########################################################################
 9
10  # REQUIRED PACKAGES
11  #       numDeriv:  grad, hessian
12  #       randtests: runs.test
13  library(numDeriv)
14  library(randtests)
15
16  ###########################################################################
17  #
18  #       CONSTANTS and CONTROL
19  #
20  ###########################################################################
21
22  # If TRUE,  model central mortality as m {appox =} mu @ x + 0.5
23  # If FALSE, model central mortality as integrated hazard
24  MODEL.M <- TRUE
25
26  # Initial parameters for fitting laws from Richards (2012)
27  INIT.ALPHA   <- -10
28  INIT.BETA    <- 0.1
29  INIT.EPSILON <- -10
30  INIT.RHO     <- 0
```

```
31
32 # Initial parameters for fitting Gompertz-Makeham - GM(r,s) - family
33 INIT.B1 <- -5
34 INIT.B2 <- 5
35 INIT.BX <- 0
36 INIT.AX <- 0
37
38 # For controlling optimization and fitting procedures.
39 # Setting the relative tolerance (REL.TOLERANCE) so low may be overkill,
40 # but it assisted in testing and comparing results to other implementations.
41 # Forcing the integrated hazard or central mortality (m) to MINIMUM.MU
42 # is meant to avoid errors being raised when a negative value would be
43 # attempted in the fitting process (which can happen when fitting to a
44 # GM(r,s) formula), but should not have impact on the parameters fitted.
45 OPTIM.METHOD    <- "BFGS"
46 GRAD.METHOD     <- "Richardson"
47 REL.TOLERANCE   <- 1e-20
48 MAX.ITERATIONS  <- 1000
49 MININIMUM.MU    <- 1e-20
50
51 ##############################################################################
52 #
53 #      HELPER FUNCTIONS FOR GM(r,s) LAWS
54 #
55 ##############################################################################
56
57 ##############################################################################
58 # cheby(N, X) -> Return value of Chebyshev Polynomial of the first
59 #                kind of order N evaluated at X
60 ##############################################################################
61 cheby <- function(N, X) {
62     if (N == 0) {
63         return(1)
64     }
65     else if (N == 1) {
66         return(X)
67     }
68     else {
69         return(2 * X * cheby(N - 1, X) - cheby(N - 2, X))
70     }
71 }
72
73 ##############################################################################
74 # boole(f) -> Numerically approximate integral using integrand f
75 #             evaluated over 0 to 1 using Boole's rule
76 ##############################################################################
77 boole <- function(f)
78           { (7*f(0) + 32*f(0.25) + 12*f(0.5) + 32*f(0.75) + 7*f(1)) / 90 }
79
80 ##############################################################################
81 #
82 #      MORTALITY LAW FUNCTIONS
83 #
84 #      Defined by functions for:
85 #          - force of mortality (mu)
86 #          - integrated hazard function (int)
87 #
88 ##############################################################################
89
90 ##############################################################################
91 # Mortality laws from Richards (2012)
92 ##############################################################################
93 gompertz.mu  <- function(p, x) { exp(p[1] + p[2] * x) }
94 gompertz.int <- function(p, x)
95     { ((exp(p[2]) - 1) / p[2]) * exp(p[1] + p[2] * x) }
96 gompertz     <- list(mu = gompertz.mu, int = gompertz.int)
97
98 makeham.mu  <- function(p, x) { exp(p[3]) + exp(p[1] + p[2] * x) }
```

```
 99  makeham.int <- function(p, x)
100        { exp(p[3]) + ((exp(p[2]) - 1) / p[2]) * exp(p[1] + p[2] * x) }
101  makeham      <- list(mu = makeham.mu, int = makeham.int)
102
103  perks.mu  <- function(p, x)
104      { exp(p[1] + p[2] * x) / (1 + exp(p[1] + p[2] * x)) }
105  perks.int <- function(p, x)
106      { (1/p[2]) * log((1 + exp(p[1] + p[2] * (x + 1)))
107                       / (1 + exp( p[1] + p[2] * x ))) }
108  perks     <- list(mu = perks.mu, int = perks.int)
109
110  beard.mu  <- function(p, x) { exp(p[1] + p[2] * x) /
111                               (1 + exp(p[1] + p[3] + p[2] * x)) }
112  beard.int <- function(p, x)
113      { (exp(-p[3])/p[2]) * log((1 + exp(p[1] + p[3] + p[2] * (x + 1)))
114                               / (1 + exp(p[1] + p[3] + p[2] * x))) }
115  beard     <- list(mu = beard.mu, int = beard.int)
116
117  makeham.perks.mu  <- function(p, x) { (exp(p[3]) + exp(p[1] + p[2] * x)) /
118                                        (1 + exp(p[1] + p[2] * x)) }
119  makeham.perks.int <- function(p, x)
120      { exp(p[3]) + ((1 - exp(p[3]))/p[2]) * log((1 + exp(p[1] + p[2] * (x + 1)))
121                                                / (1 + exp(p[1] + p[2] * x))) }
122  makeham.perks     <- list(mu = makeham.perks.mu, int = makeham.perks.int)
123
124  makeham.beard.mu  <- function(p, x) { (exp(p[3]) + exp(p[1] + p[2] * x)) /
125                                        (1 + exp(p[1] + p[4] + p[2] * x)) }
126  makeham.beard.int <- function(p, x)
127      { exp(p[3])  + ((exp(-p[4]) - exp(p[3]))/p[2]) *
128                       log((1 + exp(p[1] + p[4] + p[2]*(x+1))) /
129                          (1 + exp(p[1] + p[4] + p[2] * x))) }
130  makeham.beard     <- list(mu = makeham.beard.mu, int = makeham.beard.int)
131
132  ########################################################################
133  # Generalized Gompertz-Makeham Family - GM(r,s)
134  # See Forfar et al. (1988)
135  ########################################################################
136  gmrs.mu <- function(p, r, s, x) {
137      TX <- (x - 70) / 50
138      sum1 <- 0
139      if (r != 0) {
140          for (i in 1:r) { sum1 <- sum1 + p[i] * cheby(i - 1, TX) }
141      }
142      sum2 <- 0
143      if (s != 0) {
144          for (j in 1:s) { sum2 <- sum2 + p[j + r] * cheby(j - 1, TX) }
145      }
146
147      return(sum1 + exp(sum2))
148  }
149
150  gmrs.int <- function(p, r, s, x) {
151      integrand <- function(tt) { gmrs.mu(p, r, s, x + tt) }
152      boole(integrand)
153  }
154
155  gmrs <- function(r, s) {
156      list (mu  = function(p, x) { gmrs.mu (p, r, s, x) },
157            int = function(p, x) { gmrs.int(p, r, s, x) } )
158  }
159
160  ########################################################################
161  # law.set(p, law) -> Create force of mortality (mu) and integrated
162  #                    hazard (int) functions for law using parameters p
163  ########################################################################
164  law.set <- function(p, law) {
165      off <- if(!MODEL.M) 0 else -0.5
166      list(mu  = function(x) { law$mu (p, x + off) },
```

```
167          int = function(x) { law$int(p, x + off) })
168 }
169
170 ############################################################################
171 #
172 #       MAIN FITTING FUNCTIONS
173 #
174 ############################################################################
175
176 ########################################################################
177 # ll(p, f, d) -> Exact log-likelihood function under the Poisson model
178 #     Poisson parameter determined by multiplying central exposure by
179 #     function f evaluated using the parameters in p and using the data
180 #     under d
181 ########################################################################
182 ll <- function(p, f, d) {
183     mu <- pmax(MININIMUM.MU, f(p, d$x))
184     sum(ifelse(d$e == 0, 0, -d$e * mu + d$d * log(d$e*mu) - lfactorial(d$d)))
185 }
186
187 ########################################################################
188 # fit(p.int, f, d) -> Fit model and return key results
189 #     Model fitted using mortality law defined by f, data under d, with
190 #     p.init the initial parameters for optimization
191 ########################################################################
192 fit <- function(p.init, f, d) {
193
194     ll.fit <- function(p) { ll(p, if (!MODEL.M) f$int else f$mu, d) }
195
196     grr <- function(p) { grad(ll.fit, p, method = GRAD.METHOD) }
197
198     model<-optim(par = p.init, gr = grr, fn = ll.fit, method=OPTIM.METHOD,
199                  control = list(fnscale=-1,
200                                 reltol=REL.TOLERANCE,
201                                 maxit=MAX.ITERATIONS))
202
203     hess.calc <- hessian(ll.fit, model$par, method = GRAD.METHOD)
204
205     return(list(ll=model$value, pars = model$par, npars = length(model$par),
206                 hessian = hess.calc, fun=law.set(model$par, f)))
207 }
208
209 ############################################################################
210 #
211 #       LAW-SPECIFIC FITTING FUNCTIONS
212 #
213 ############################################################################
214
215 ########################################################################
216 # For laws from Richards (2012)
217 # Order of fitting below chosen to provide optimal stability
218 ########################################################################
219
220 fit.gompertz <- function(d) { fit(c(INIT.ALPHA, INIT.BETA), gompertz, d) }
221
222 # For Makeham, fit Gompertz (ALPHA and BETA) first and then add Makeham
223 # (EPSILON) term
224 fit.makeham <- function(d) {
225     model.gompertz <- fit.gompertz(d)
226     fit(c(model.gompertz$pars, INIT.EPSILON), makeham, d)
227 }
228
229 fit.perks <- function(d) { fit(c(INIT.ALPHA, INIT.BETA), perks, d) }
230
231 # For Beard, fit Perks (ALPHA and BETA) first and then add Beard (RHO)
232 # parameter
233 fit.beard <- function(d) {
234     model.perks <- fit.perks(d)
```

```
235    fit(c(model.perks$pars, INIT.RHO), beard, d)
236 }
237
238 # For Makeham-Perks, fit Perks first and then add Makeham (EPSILON) term
239 fit.makeham.perks <- function(d) {
240    model.perks <- fit.perks(d)
241    fit(c(model.perks$pars, INIT.EPSILON), makeham.perks, d)
242 }
243
244 # For Makeham-Beard, fit Makeham-Perks first and then add Beard (RHO) term
245 fit.makeham.beard <- function(d) {
246    model.makeham.perks <- fit.makeham.perks(d)
247    fit(c(model.makeham.perks$pars, INIT.RHO), makeham.beard, d)
248 }
249
250 ##########################################################################
251 # For Generalized Gompertz-Makeham Family - GM(r,s)
252 # Fit Gompertz (s) terms first, and then add non-linear Makeham (r)
253 # terms one-by-one
254 ##########################################################################
255 fit.gmrs <- function(d, r, s) {
256    p.init <- if (s == 0) NULL
257              else if (s == 1) INIT.B1
258              else c(INIT.B1, INIT.B2, rep(INIT.BX, s - 2))
259    gmrs.fit <- if (s != 0) fit(p.init, gmrs(0, s), d) else NULL
260
261    if (r == 0) {
262        return(gmrs.fit)
263    }
264    else {
265        for (i in 1:r) {
266            r.pars <- if (i == 1) INIT.AX else c(gmrs.fit$pars[1:(i-1)],
267                                                 INIT.AX)
268            s.pars <- if(s != 0) gmrs.fit$pars[i:(i+s-1)] else NULL
269            gmrs.fit <- fit(c(r.pars, s.pars), gmrs(i, s), d)
270        }
271        return(gmrs.fit)
272    }
273 }
274
275 fit.gm <- function(r, s) { function(d) { fit.gmrs(d, r, s) } }
276
277 ##########################################################################
278 #
279 #     RUN GRADUATION and STATISTICAL TESTS
280 #
281 #
282 #     Fitting done by function f over data d, with the underlying model
283 #     assumed to be Poisson
284 #
285 #
286 #     The fitting function f must return a list that includes the following:
287 #        - ll:      the log-likelihood under the Poisson model
288 #        - pars:    the fitted parameters
289 #        - npars:   the number of parameters
290 #        - hessian: the Hessian matrix (used to estimate the standard errors)
291 #        - fun:     the functions defining the mortality law, as determined
292 #                   by the fitted parameters
293 #
294 #     The data to which the mortality law is fitted must include certain
295 #     columns or elements:
296 #        - x: the ages (x)
297 #        - d: the number of deaths at age (x)
298 #        - e: the central exposure at age (x)
299 #
300 ##########################################################################
301
302 graduate.poisson <- function(f, d) {
```

```
303     # Set results
304     result   <- f(d)
305
306     ll       <- result$ll
307     pars     <- result$pars
308     npars    <- result$npars
309     hessian <- result$hessian
310
311     m <- if(!MODEL.M) function(x) { result$fun$int(x) }
312         else function(x) { result$fun$mu(x + 0.5) }
313     fun <- list(mu  = result$fun$mu,
314                 int = result$fun$int,
315                 qx  = function(x) { 1 - exp(-result$fun$int(x)) },
316                 m   = m)
317     rates <- m(d$x)
318
319     # Calculate standard errors and p-values for parameter estimates
320     std.dev   <- sqrt(diag(solve(-hessian)))
321     z.stat    <- pars / std.dev
322     p.values <- 2*pnorm(abs(z.stat), 0, 1, lower.tail = FALSE)
323
324     # Calculate deviance and residuals
325     O <- d$d
326     E <- rates * d$e
327     dev <- sum(2 * (ifelse(O == 0, 0, O * log(O/E)) - (O - E)))
328     res <- sign(O - E) * sqrt(2 * (ifelse(O == 0, 0, O * log(O/E)) - (O - E)))
329
330     # Determine dispersion coefficient
331     dis <- dev / (length(d$x) - npars)
332
333     # Chi-squared calculations
334     chi   <- sum(((O - E)^2) / ifelse(E == 0, 1, E))
335     chi.p <- pchisq(chi, df = length(d$x) - npars, lower.tail = FALSE)
336
337     # Information criteria
338     aic <- -2 * ll + 2 * npars
339     bic <- -2 * ll + log(length(d$x)) * npars
340
341     # Signs test
342     signs.p    <- sum(res > 0)
343     signs.n    <- sum(res < 0)
344     signs.a    <- if(signs.p <= signs.n) "less" else "greater"
345     signs.test <- binom.test(signs.p, signs.p + signs.n,
346                             alternative = signs.a)$p.value
347
348     # Runs test
349     runs      <- length(rle(sign(res))$lengths)
350     runs.p    <- runs.test(res, "left.sided", 0, "exact", FALSE)$p.value
351
352     # Consolidate main results and statistics
353     details <- list(pars = pars, npar = npars, std.dev = std.dev,
354                     z.stat = z.stat, p.values = p.values, ll = ll, aic = aic,
355                     bic = bic, dis = dis, chi = chi, chi.p = chi.p,
356                     signs.p = signs.p, signs.n = signs.n,
357                     signs.test = signs.test, runs = runs, runs.p = runs.p)
358
359     # Calculate 95% confidence intervals
360     lower = log(qpois(0.025, O) / d$e)
361     upper = log(qpois(0.975, O) / d$e)
362     par(mfrow=c(2, 1))
363
364     # Plot crude versus fitted central mortality rates
365     obs <- log(d$d / d$e)
366     mod <- log(rates)
367
368     if(length(mod) == 1) mod <- rep(mod, length(obs))
369     y.min <- min(obs[is.finite(obs)],
370                  mod[is.finite(mod)],
```

```
371                   lower[is.finite(lower)])
372      y.max <- max(obs[is.finite(obs)],
373                   mod[is.finite(mod)],
374                   upper[is.finite(lower)])
375
376      plot (d$x, lower, pch = 150, cex=1.25, col = "blue",
377            xlab = "Age", ylab = "log (Central Mortality)",
378            ylim = c(y.min,y.max), axes = FALSE)
379      axis(1, at=c(seq(from=floor(min(d$x)/5)*5,to=ceiling(max(d$x)/5)*5,by=5)),
380           las = 1)
381      axis(2, las = 1)
382      box()
383      points(d$x, upper, pch = 150, cex = 1.25, col = "green")
384      points(d$x, obs, col = "black", cex = 1.2)
385      lines(d$x, mod, col = "red", lwd = 2)
386      legend("bottomright", lwd=c(NA, 2, NA, NA), lty=c(NA, 1, NA, NA),
387             pch=c(1, NA, 150, 150), col=c("black", "red", "green", "blue"),
388             legend=c("Crude","Graduated","97.5th Percentile",
389                      "2.5th Percentile"),
390             bty="n", cex=0.7)
391
392      # Plot deviance residuals
393      plot(d$x, res, col = "blue",
394           xlab = "Age", ylab = "Deviance Residual", axes = FALSE)
395      axis(1, at=c(seq(from=floor(min(d$x)/5)*5,to=ceiling(max(d$x)/5)*5,by=5)),
396           las = 1)
397      axis(2, las = 1)
398      box()
399      abline(h = 0,   col = "red")
400      abline(h = 1.96,  lty = 3)
401      abline(h = -1.96, lty = 3)
402
403      return(list(fun = fun, details = details))
404 }
405
406 ###############################################################################
407 ###############################################################################
408
409 # Sample Data
410 # Deaths and central exposures over ages 50 to 100 for females in British
411 # Columbia in 2011 from the Canadian Human Mortality Database (CHMD)
412 dat.sam <-
413     list (x = seq(50,100),
414           d = c(74, 71, 95, 87, 93, 100, 117, 123, 127, 125, 114, 143, 131,
415                 145, 164, 170, 167, 168, 182, 214, 220, 227, 216, 247, 251,
416                 254, 295, 316, 331, 360, 372, 422, 444, 489, 508, 531, 593,
417                 604, 671, 652, 647, 604, 467, 486, 383, 350, 331, 226, 222,
418                 148, 101),
419           e = c(36480.48, 36200.11, 35699.54, 35172.46, 34484.91, 33914.96,
420                 33190.50, 32112.22, 30779.34, 29843.52, 29366.44, 28930.82,
421                 28453.91, 28413.12, 27003.91, 24287.67, 22254.62, 21462.25,
422                 20502.33, 19431.93, 18418.72, 17283.89, 16324.33, 15669.11,
423                 15088.03, 14434.73, 13689.43, 13146.68, 12723.69, 12384.01,
424                 11983.83, 11432.55, 10718.27, 10083.34, 9463.17, 8855.87,
425                 8259.08, 7513.45, 6756.52, 5910.47, 5116.49, 4104.05,
426                 3119.46, 2537.77, 1984.37, 1569.65, 1254.94, 938.53, 634.19,
427                 403.07, 258.23))
428
429 # Examples of Graduations using Sample Data
430 ## graduate.poisson(fit.gompertz,      dat.sam)
431 ## graduate.poisson(fit.makeham.perks, dat.sam)
432 ## graduate.poisson(fit.gm(2,4),       dat.sam)
433
434 ###############################################################################
435 ###############################################################################
436 ###############################################################################
437 ###############################################################################
438 ###############################################################################
```

Appendix C   Identities Regarding the Force of Mortality

The following provides various key definitions, identities, and relationships regarding the key mortality measures, as well as explaining the main motivations for modelling the force of mortality μ_x, or at least the central death rate m_x, over the probability of death q_x.

Base Identities

The force of mortality μ_x, called the *hazard rate* by statisticians, is formally defined as:

$$\mu_x = \lim_{h \to 0+} \frac{1}{h} Pr\left(death\ before\ age\ x+h \mid alive\ at\ age\ x\right)$$
$$= \lim_{h \to 0+} \frac{{}_h q_x}{h}$$

The *integrated hazard function* is then defined as:

$$H_x(t) = \int_0^t \mu_{x+s}\,\mathrm{d}s \tag{C.1}$$

The probability of survival from x to $x+t$ is exactly given by:

$$_t p_x = e^{-H_x(t)} \tag{C.2}$$

And thus there is a precise relationship between q_x and μ_x:

$$q_x = 1 - p_x = 1 - e^{-H_x(1)} = 1 - e^{-\int_0^1 \mu_{x+t}\,\mathrm{d}t} \tag{C.3}$$

Definitions, Identities, and Approximations for Grouped Data

When dealing with mortality experience at the level of the individual, the above relationships are sufficient for statistical modelling purposes. With grouped data, such as the usual case in the construction of standard base tables where data is aggregated by age and other factors, we need to define some additional quantities:

- l_x: The number of lives alive at age (x)

- d_x: The number of deaths for those aged between (x) and $(x+1)$

- E_x^c: The total time lived between ages (x) and $(x+1)$, which can be termed the **central** exposed-to-risk/exposure

- E_x^i: The **initial** exposure at age (x). In the absence of entrants or withdrawals (for reasons other than death), $E_x^i = l_x$

To tabulate the central exposure E_x^c in practice, days can be counted between the earliest of (1) the appropriate birthday, (2) observation start date, (3) entry date, etc. to the latest of (1) the following birthday, (2) death date (or exit for reasons other than death), (3) observation end date, etc. The exact tabulation will depend on the age definition employed, but an age last birthday (ALB) basis is generally the most straightforward and effective, and will result in the implied μ_x and q_x applying at exact age (x).

If the central exposure E_x^c is not available, it can be roughly approximated through the following identity:

$$E_x^c \approx E_x^i - \frac{d_x}{2} \tag{C.4}$$

Now, we can define the central death rate m_x:

$$m_x = \frac{d_x}{E_x^c} = \frac{\int_0^1 l_{x+t}\mu_{x+t}\, \mathrm{d}t}{\int_0^1 l_{x+t}\, \mathrm{d}t}$$

Under the assumption where l_x does not change rapidly over the age:

$$m_x = \frac{d_x}{E_x^c}$$

$$\approx \int_0^1 \mu_{x+t}\, \mathrm{d}t = H_x(1) \tag{C.5}$$

Alternatively, under the assumption where μ_x is constant over the age, taken to be equal to $\mu_{x+\frac{1}{2}}$ (i.e., $\int_0^1 \mu_{x+t}\, \mathrm{d}t \approx \mu_{x+\frac{1}{2}}$):

$$m_x = \frac{d_x}{E_x^c}$$

$$\approx \mu_{x+\frac{1}{2}} \tag{C.6}$$

Note that if we instead work with initial exposure:

$$q_x = \frac{d_x}{E_x^i} \tag{C.7}$$

Advantages of Using the Force of Mortality

From equation (C.3), q_x can be directly returned from μ_x, but not the other way around. Indeed, graduating mortality directly according to μ_x, or to m_x, provides numerous advantages over dealing with q_x:

- As noted above, μ_x directly defines $_tp_x$ and thus q_x, but q_x does not necessarily return the exact μ_x. To see how this one-way relationship can impact estimates in practice, we use a similar example as provided in Richards (2008) and consider an extreme scenario where there are 100 lives alive at the beginning of the year and 50 die during the year. No matter when during the year those 50 lives died, q_x will equal $\frac{50}{100} = 0.5$. If on average the 50 lives die exactly halfway through the year, then $m_x = \frac{50}{50+\frac{50}{2}} = 0.6667$. However, if all 50 lives die at the end of the January, then m_x is estimated as $\frac{50}{50+\frac{50}{12}} = 0.9231$. Thus, using μ_x or m_x retains additional information compared to q_x.

- Since q_x is the probability of those aged (x) dying within the next year, it must tend to 1 as the higher ages are approached. There are no such restrictions on μ_x or m_x; they can theoretically go well above 1 without limit. As such, more complex models are required to deal with q_x than for μ_x or m_x for higher ages. For example, Figure C.1 illustrates a situation where μ_x is log-linear at all ages (i.e., Gompertz' law), the resulting q_x, as per equation (C.3), will follow a logistic shape. Clearly, the shape of the force of mortality is much simpler to capture.

Figure C.1: Force of Mortality (μ_x) and Mortality Rate (q_x) under Gompertz' Law
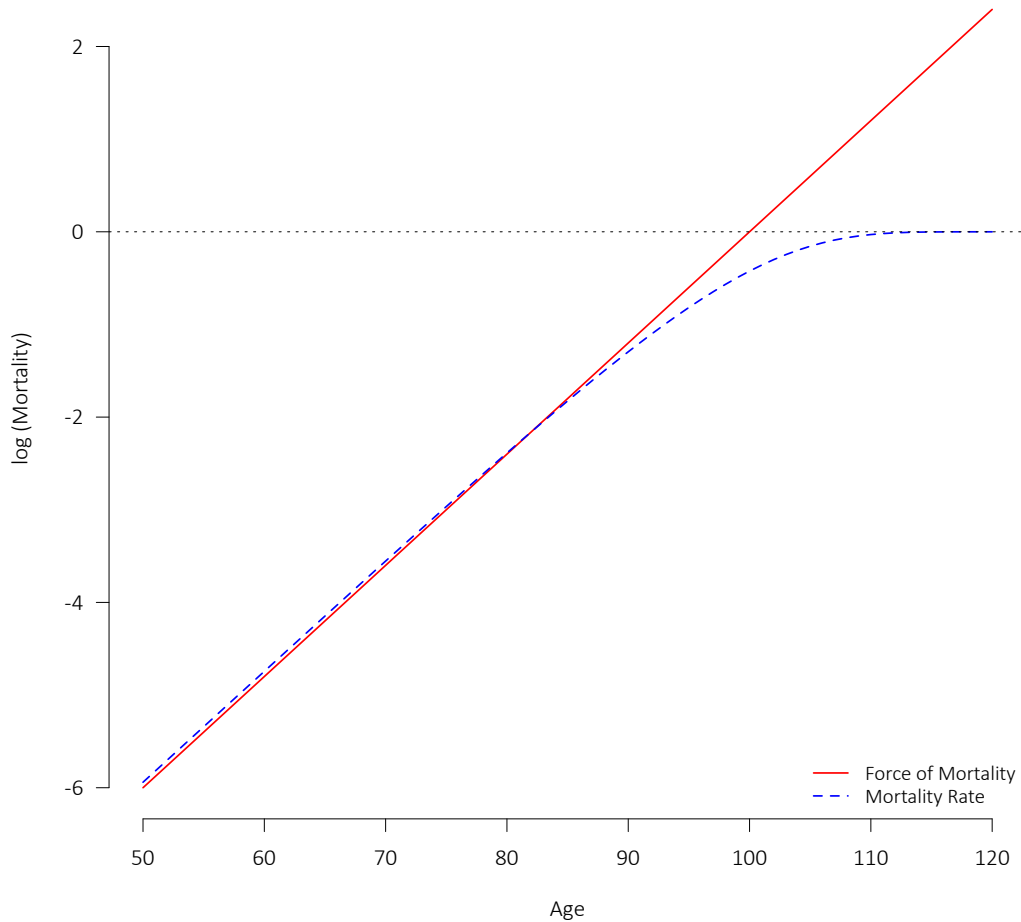
- Whenever q_x is used, even at the individual level, the initial exposure E_x^i must be invoked. Unlike with the central exposure E_x^c, it is not possible to deal with entrants during the year, as might be the case with retirements, the commencement of a new annuity or insurance coverage, or the succession of a spouse or other beneficiary to a pension or policy, in a completely satisfying manner. It is especially problematic to capture cases where a life dies within the year of entry. For illustration, say a life enters halfway through the year and dies before the end of the year. In such a situation, d_x is clearly incremented by 1. If the life had survived the year, E_x^i would presumably need to be increased by 0.5. However, as the life died, if 0.5 is added then mortality would be overstated (the implied q_x could even be brought above 1 in extreme cases); adding 1 to E_x^i instead is better and the common practice, but clearly understates mortality (this approach follows from the Balducci hypothesis, which implies that the force of mortality decreases over a year-of-age) . The use of initial exposure also leads to other issues, including in situations where there are withdrawals for reasons over than death (such as lapse). All those issues are avoided if μ_x is used at the individual level, or if the central exposure E_x^c and m_x is used for aggregated experience.

Many of the points listed above relate to one of the most fundamental advantages of central over initial exposure, which is the ability to properly reflect *left-truncation* and *right-censorship*. Experience data is generally left-truncated, as information on deaths falling before the observation start date is unavailable, as well as right-censored, as not all lives will be deceased by the end of the study period. These considerations are especially critical when analyzing mortality at the level of the individual.

As the mathematics are simpler and the limitations fewer, this paper focuses exclusively on working with μ_x or m_x. As such, all references to exposure generally refer to central exposure E_x^c unless explicitly stated. As described in detail in Forfar et al. (1988), a similar statistical framework (though invoking a binomial[31] as opposed to a Poisson assumption) can be applied to q_x instead if absolutely necessary. However, the simple identities and approximations provided through equations (C.3), (C.4), (C.5), and (C.6) should permit the use of central exposure in most situations.

Implementation Notes

The approximation under equation (C.5) should generally be considered slightly more statistically valid than that under equation (C.6). As explicit, closed-form expressions for $H_x(t)$ are available for many of the mortality laws explored in this paper (for instance, see Richards (2012) for integrated hazard functions corresponding to the laws presented therein), the former approximation does not necessarily pose any particular complications.

Nevertheless, for simplicity and consistency, the results in this paper are calculated using the approximation under equation (C.6), though m_x is parameterized directly as opposed to μ_x. The results and conclusions would not have been materially different had the equation (C.5) been employed instead.

The sample R implementation in Appendix B allows for either approximation to be applied, with the choice controlled through a single variable.

---

[31]Where the conditional likelihood function under the binomial assumption is $L\left(\theta\right) \propto \prod_{x} q_x(\theta)^{d_x} \left(1 - q_x(\theta)\right)^{E_x^i - d_x}$.

Appendix D The Poisson Model and Overdispersion

Under the assumption that the count of deaths at age x follows a Poisson distribution as described under equation (1), the mean and variance of the corresponding random variable D_x are assumed to be equal:

$$\mathbb{E}(D_x) = Var(D_x) = \lambda = E_x \hat{m}_x \tag{D.1}$$

This is a suitable assumption in many cases, but there are certain reasons why the Poisson model may not perform well in practice. One of the most common such situations is where a large amount of *duplicates* is present in the dataset. That is, where an individual has multiple policies, pensions, annuities, etc. and each is treated separately as opposed to being combined. Ideally, all such duplicates would be identified and the corresponding records merged. However, that may simply not be possible in some cases, such as when dealing with grouped data.

Note that the use of amounts-weighted exposure and deaths also creates a form of duplication, as each unit of exposure and/or death no longer will be independent.

In cases of duplication or other violations of the Poisson properties, the variance of D_x is likely higher than the mean, which is referred to as *overdispersion*. One way to quantify dispersion within the Poisson model framework is by dividing the Poisson deviance (or, alternatively and nearly equivalently, the Pearson's χ^2 statistic) by the degrees of freedom, i.e. the number of observations less the number of model parameters. This ratio can be referred to as the "dispersion coefficient" DCF:

$$Deviance = \sum_x 2 \left[d_x \times log\left(\frac{d_x}{f_x}\right) - (d_x - f_x) \right] \tag{D.2}$$

$$DCF = \frac{Deviance}{N - K} \tag{D.3}$$

where N is the number of observations (ages) and K is the number of parameters.

If the dispersion coefficient is close to 1, then there is no real evidence of dispersion. A coefficient under 1 would imply *underdispersion* (which is relatively rare in practice), while one over 1 implies overdispersion. If the amount of dispersion is deemed significant, then a different model may be warranted.

If the dispersion is significant but within reasonable bounds, then a *quasi-Poisson* model may make sense. Under such a model, the parameter estimates are the same as under the basic Poisson model (i.e., the mean is the same as under the Poisson), but the p-values and certain statistics are altered by adjusting the variance by a dispersion factor ψ:

$$\mathbb{E}(D_x) = E_x \hat{m}_x$$
$$Var(D_x) = \psi \mathbb{E}(D_x) = \psi \left(E_x \hat{m}_x \right) \tag{D.4}$$

If the amount of dispersion is more significant, then it could make more sense to switch from the Poisson model to the more general negative binomial model, where the overdispersion will impact both the parameter estimates and the variance. The negative binomial distribution is a generalization of the Poisson distribution, which includes an additional parameter so that the variance can vary from the mean. Under a negative binomial distribution with ψ the dispersion parameter, $\lambda_x(\theta) = E_x \hat{m}_x(\theta)$, $p_x(\theta) = \frac{\lambda_x(\theta)}{\lambda_x(\theta) + \psi}$, and $\Gamma(\cdot)$ the gamma function, the corresponding conditional likelihood function is:

$$L\left(\theta, \psi\right) \propto \prod_x \frac{\Gamma\left(d_x + \psi\right)}{d_x! \times \Gamma\left(\psi\right)} p(\theta)^{d_x} \left(1 - p(\theta)\right)^{\psi} \tag{D.5}$$

Now, in the case that there is access to the underlying individual data and there is the ability to merge any individual duplicate records, remaining overdispersion could be evidence that the assumption of a Poisson distribution is problematic for other reasons. Working with mortality at the level of the individual eliminates any necessity to make an assumption regarding the distribution of deaths at any particular age. Instead, we can work with the conditional likelihood function defined for each individual i who entered the observation period at age x_i and was observed for time t_i, where d_i is 1 if the individual died at age $\left(x_i + t_i\right)$ and 0 otherwise:

$$L\left(\theta\right) \propto \prod_i {}_{t_i}p_{x_i} \hat{\mu}_{x_i + t_i}^{\ d_i} \tag{D.6}$$

As ${}_tp_x$ is precisely determined by μ_x through equation (C.2), the above likelihood function is dependent solely on the function for the force of mortality μ_x (itself being a function of parameters θ).

Refer to Richards (2008) and Richards et al. (2013) for case studies where survival models are applied to the mortality experience at the level of the individual.

Appendix E   Technical Notes on Implementing $GM(r,s)$

The $GM(r,s)$ family described by equation (8) represents a general form that can capture a wide variety of mortality patterns, and it has been widely used by the Continuous Mortality Investigation (CMI) in the United Kingdom. There are a number of technical considerations that should be noted regarding how the $GM(r,s)$ curves are fitted in practice.

Most importantly, for fitting purposes and for better scaling and stability of the parameters, the $GM(r,s)$ structure relies on polynomials in age based on the transformation $T(x) = \frac{x-70}{50}$ and Chebyshev polynomials of the first kind defined by $C(N,X)$, where when r or s is equal to 0 the corresponding summation is completely excluded:

$$\mu_x = GM(r,s) = \sum_{i=0}^{r-1} a_i C(i, T(x)) + exp\left\{\sum_{j=0}^{s-1} b_j C(j, T(x))\right\} \tag{E.1}$$

Where the Chebyshev polynomials $C(N,X)$ are defined recursively as:

$$C(N,X) = \begin{cases} 1 & N = 0 \\ X & N = 1 \\ 2X \times C(N-1,X) - C(N-2,X) & N \geq 2 \end{cases}$$

By employing the Chebyshev polynomial structure, it is easier to fit the parameters, the parameter values are more consistently scaled, and the values are more stable when additional parameters are added. Other *orthogonal* polynomial bases, such as Legendre polynomials or the basis provided through the `poly` function in R Core Team (2017), could be used instead to achieve the same benefits. In this paper, the Chebyshev polynomials are invoked primarily to maintain consistency with the CMI's established approach when graduating mortality tables.

When the mortality probability rates q_x are required, it is most accurate to determine them through the precise relationship with μ_x provided through equation (C.3). Explicit formulae for the integrated hazard function $H_x(t)$ could be determined for the $GM(r,s)$ family where $s \leq 2$, though even then they could be unwieldy, especially in the presence of the Chebyshev polynomials and the transformation of x. Sufficiently precise results can instead be obtained through a numerical integration approach, and the CMI has employed Boole's rule for that purpose in recent graduations:

$$H_x(1) \approx \frac{1}{90}\left(7\mu_x + 32\mu_{x+\frac{1}{4}} + 12\mu_{x+\frac{1}{2}} + 32\mu_{x+\frac{3}{4}} + 7\mu_{x+1}\right) \tag{E.2}$$

The sample R implementation included in Appendix B is consistent with the above, as is that described in CMI WP 077 (2015).

Appendix F   On Amounts-Weighted Mortality Experience

As discussed under Section 9, an approximation to the financial impact of amounts (of pension, policy face value, other benefits, etc.) can be achieved by weighting the deaths and exposure by the pension, annuity, or policy amount applicable to the underlying lives. So, the amounts-weighted central mortality rate m_x^a is simply the ratio of the amounts-weighted deaths d_x^a to the amounts-weighted exposure E_x^a, that is:

$$m_x^a = \frac{d_x^a}{E_x^a} \tag{F.1}$$

The amounts-weighting is mathematically (and conceptually) equivalent to multiplying the lives-based m_x by a ratio of average pensions:

$$m_x^a = m_x \times \frac{Average\,Pension\,for\,Lives\,Dying\,at\,Age\,x}{Average\,Pension\,for\,all\,Lives\,at\,Age\,x} \tag{F.2}$$

As the average pension for members dying is expected to be lower than the overall average, the ratio will generally be ≤ 1 and so $m_x^a \leq m_x$.

We can choose to apply the same mathematical forms described in Section 5 with respect to lives-based mortality to that under amounts-based weightings. However, in order to produce more meaningfully scaled values for the residuals and other statistical tests, we need to scale the exposures and/or deaths in a manner that preserves the crude amounts-weighted mortality but that removes the level of the absolute amounts. One option is to scale at each individual age, multiplying both the amounts-weighted deaths and exposure by the ratio of the unweighted to amounts-weighted deaths count. Effectively, the deaths figure simply becomes the unweighted count (d_x) and the exposure is adjusted accordingly to retain the proper amounts-weighted crude rates. That is the approach that will be taken in what follows, and is also what was invoked in Section 9.

Following the same approach as under Section 5, we can try each model form in turn and determine which returns the optimal values of the AIC and BIC. As was the case for the lives analysis, the best-fitting model is found to be a $GM(1,3)$. Details on the corresponding parameter estimates are found in Table F.1.

Table F.1: Parameters under $GM(1,3)$ Model for CPM2014 Males on Amounts Basis

Parameter	Estimate	Standard Error	Z	p-value	Significance
a_0	0.001130	0.000331	3.4144	$6.3918e-04$	$***$
b_0	-4.541924	0.169649	-26.7725	0	$***$
b_1	6.329917	0.170193	37.1926	0	$***$
b_2	-0.379542	0.145501	-2.6085	$9.0934e-03$	$**$

It is an almost necessary feature of amounts analysis that the experience be more dispersed (see Appendix D) than it is on a lives basis. That is borne out in the χ^2 p-value, which under the amounts-weighted $GM(1,3)$ is a highly-significant 0.000207, compared to the very insignificant 0.5579 it was with the lives $GM(1,3)$. The deviance residuals provided in Figure F.1 are also more dispersed than they were on the lives basis.

Figure F.1: Deviance Residuals – Amounts-Weighted Graduation
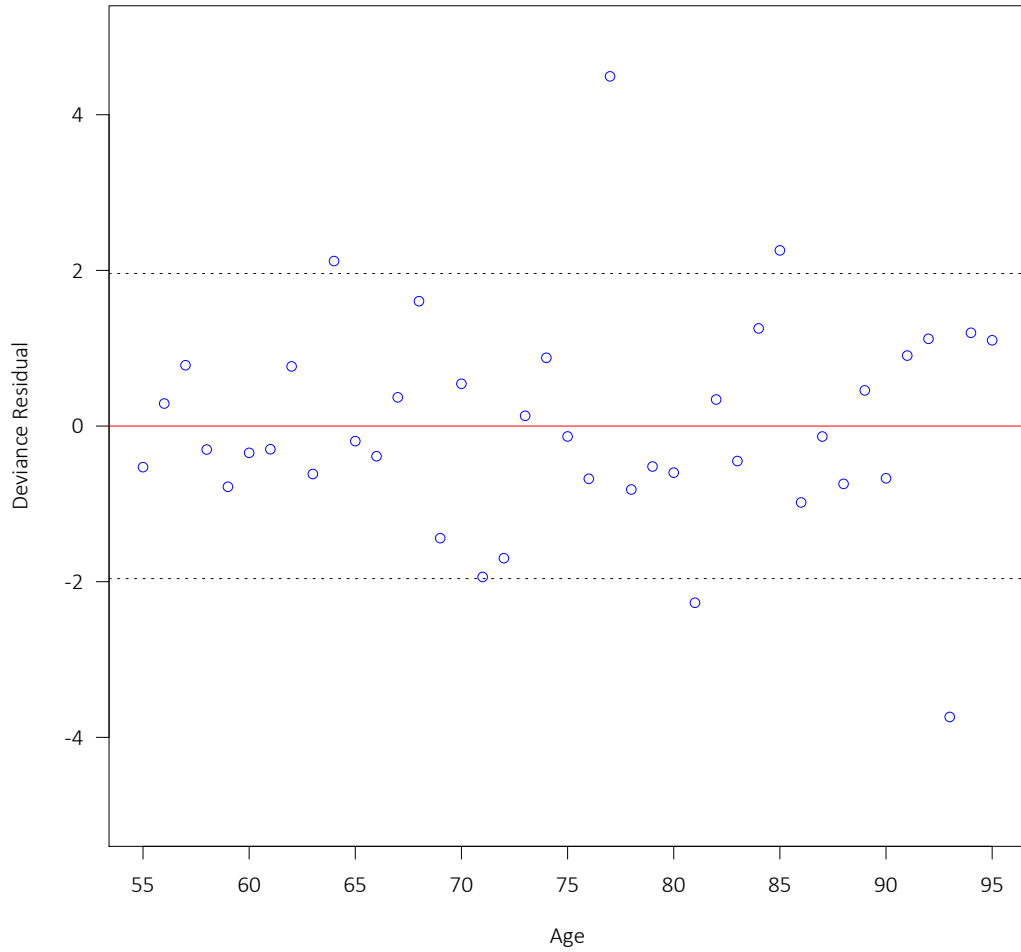
The approach explored under Section 9 allows for the most flexible model, though employing amounts-weighted mortality can be a useful tool. However, analysis should still first be performed on a lives basis, as doing so helps to gain better understanding of the experience.

Appendix G   Notes on the Datasets Used as Main Case Study

As mentioned in Appendix A, the data used as a case study throughout the majority of this paper represents the original, unaltered data underlying the CPM2014 set of Canadian pensioner tables published in CIA (2014), but differs significantly from what was actually used to graduate the published tables. That is a consequence of numerous adjustments made to the crude data prior to graduation. Few, if any, of these or comparable adjustments are found in the construction of standard tables produced by the Continuous Mortality Investigation in the United Kingdom, the Society of Actuaries in the United States, or the Canadian Institute of Actuaries in Canada for tables released prior to CIA (2014).

There were four main adjustments applied to the data before graduating the CPM2014 tables:

1. Adjustment to deaths for incurred-but-not-reported (IBNR) deaths

2. Adjustment to deaths for assumed mortality improvement

3. Adjustment to exposure and deaths to modify weightings by "industry"

4. Adjustment to deaths and lives-based exposure to "standardize" amounts distribution by age

We will reconcile the most significant of the above adjustments, that being the "standardization" of the amounts distribution by age. That modification appears to have had the greatest impact on the shape and the level of the resulting table. Exploring some of the observations that appear to have led to its introduction are quite instructive.

Before describing the adjustment itself, it is necessary to make some observations regarding the evolution of pension amount by age for the total CPM2014 dataset. The average pension by age is shown in Figure G.1 for males.

From the mid-50s, the average pension-in-pay decreases markedly. Some of the apparent reasons for that pattern include the following items:

(i) **Bridging Pensions**: The decrease in pensions between 63 and 65 is largely attributable to the cessation of bridging pensions. At lower ages in this dataset, much of the exposure is weighted towards "public sector" plans, where additional, automatic pensions are provided up to age 65 when government benefits such as those provided under the Canada Pension Plan or Québec Pension Plan typically commence. Only few "private sector" plans provide such additional pensions automatically, though "level income" options are relatively common, where the pension payable before age 65 is increased and that payable after age 65 reduced accordingly on a present value basis. A portion of the decrease may also be attributable to the commencement of deferred pensions, especially within the private sector.

(ii) **Wage Inflation**: A significant portion of the decrease past age 65 would be attributable to wages outpacing increases to the pensions (if any). Even in the case of plans fully indexed to increases in the Consumer Price Index (CPI), the indexation adjustments would still be expected to be less than the increase in wages that determine the pension amounts at retirement. Complicating matters further, not all pension plans included in the CPM2014 dataset are fully indexed; some would be partially indexed, and some not indexed at all. Pensions in the public sector are much more likely to be indexed than those in the private sector.

(iii) **Socio-Economic Shifts**: Even after accounting for the cessation of bridging pensions and the effects of wage inflation, part of the decrease may be attributable to socio-economic shifts, both within and

Figure G.1: Average Pensions for Males by Age in Total CPM2014 Dataset
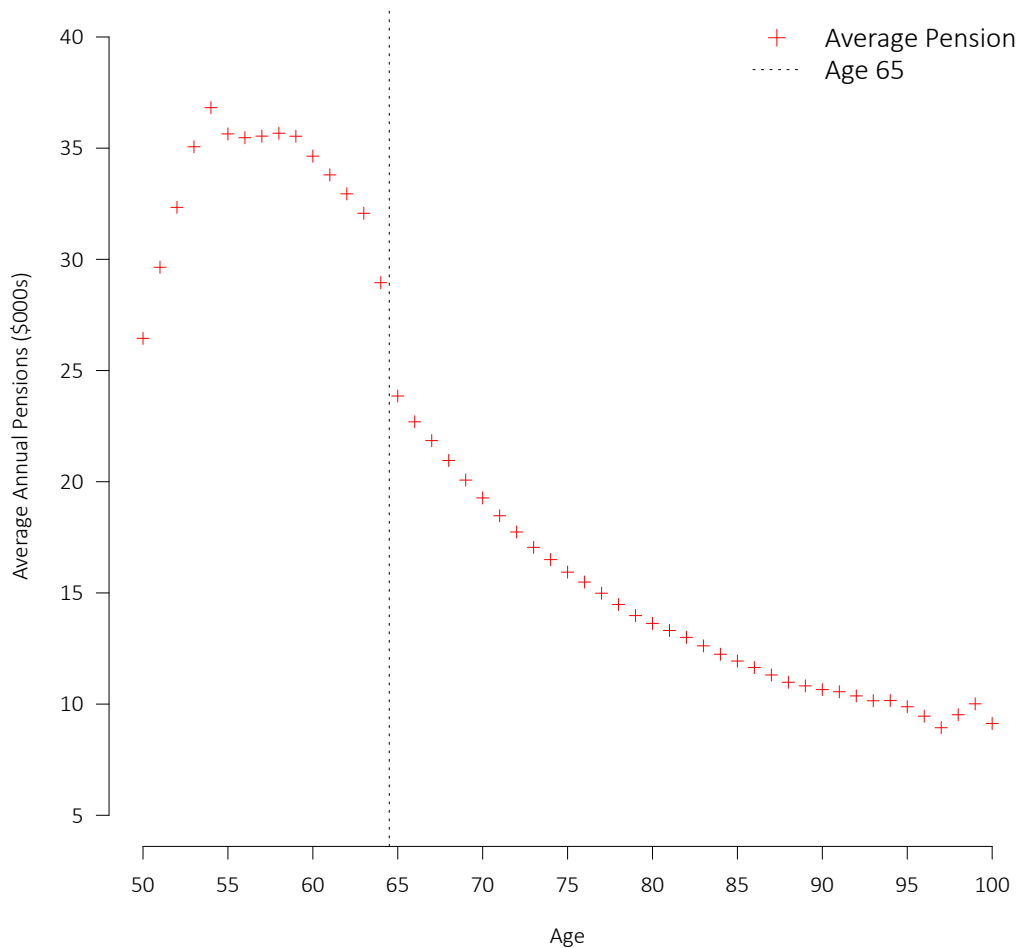
between plans. For example, part of the reason for the strength of the decrease in the average pension in the total dataset is a shift towards the private sector at the higher ages.

In terms of the graduation of the crude mortality experience:

- The effect of the bridging pensions under (i) will have absolutely no impact on a lives basis and very little on an amounts basis. The only impact on an amounts basis is that more weighting than warranted may be applied to ages or plans where bridge pensions are more common. The amounts effect across ages can be mitigated by adjusting the deaths and/or exposure within each individual age to account for the average pension at that particular age, as described in Appendix F. However, the inclusion of bridging pensions may well have implications for the weighting between plans that provide and do not provide such benefits before age 65. As such, automatic bridging pensions (such as those provided under "public sector" plans) should almost certainly be excluded for any analysis if possible, though those created under level-income options are trickier.

- The wage inflation considerations under (ii) also have little effect on the main graduations, especially

with the mitigation of the impact that can be provided by adjusting for the average pension in the weightings by age, again as described in Appendix F. However, once again the impact of wage inflation can also be expected to have the potentially undesired effect of shifting the weightings by amounts from less generous plans with lesser indexation adjustments to those with automatic adjustments.

The effects of socio-economic shifts are more subtle (especially across a pooled dataset) and much more difficult to attribute between differences in benefit formula, indexation provisions, service periods, and salary levels, as each has very different implications depending on the particular circumstance of the plan in question. However, in general, none of the specific concerns described above should have any bearing on the resulting best-estimate mortality rates.

However, in CIA (2014) the deaths and exposure were adjusted prior to graduation in such a manner that the exposure across the thirteen size amount bands at each individual age was adjusted ("standardized") to match the distribution across the entire age range. The dataset produced by the adjustment was not published, but it can be reconstructed by the data tool made available by the Canadian Institute of Actuaries.

The data for the total male dataset before the "standardization" adjustment (but with the adjustment for IBNR deaths) and with the adjustment incorporated is provided as Table G.1 at the end of the appendix. The chart provided as Figure G.2 shows the curves fitted with Makeham-Perks over ages 55 to 95 on an amounts-weighted basis with and without the adjustment.

As would be expected, the adjustment increases mortality before age 65 (i.e., while the bridging pensions are still in pay) and correspondingly decreases it after age 65. The distortion to the shape of the graduations is quite substantial.

While the presence of the bridging pension and wage inflation effects are not particularly problematic for the graduation, those considerations certainly have implications when attempting to analyze the experience by pension amount. Those implications are amplified by pension amounts not being revalued to take into account indexation adjustments (where applicable) to ensure that observed experience is consistent across all years. By way of example, Figure G.3 provides the crude and graduated rates for males in the second of thirteen CPM2014's size amount bands. The group captured before and after the cessation of the bridging pensions appears to be completely different. This observation was our main motivation to restrict the analysis in Section 9 to ages 65 and above.

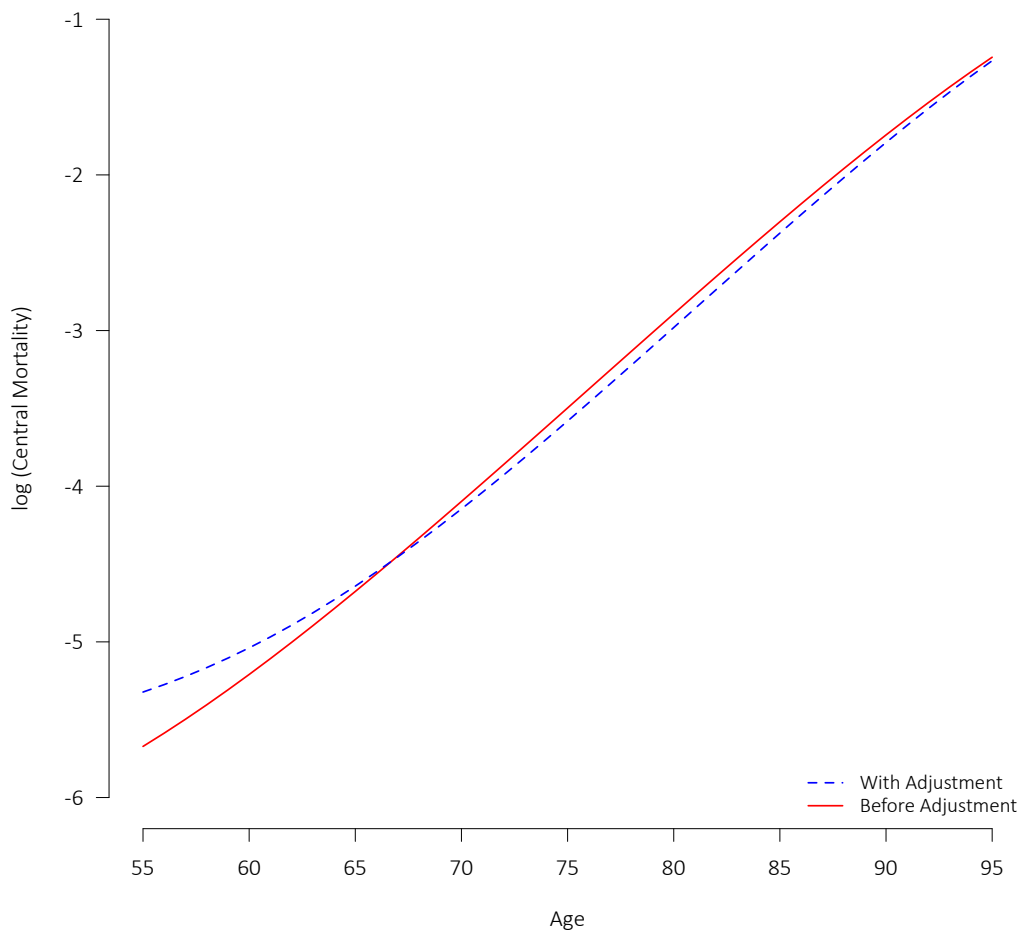Figure G.2: Graduations for Total Male Dataset with and without "Standardization"

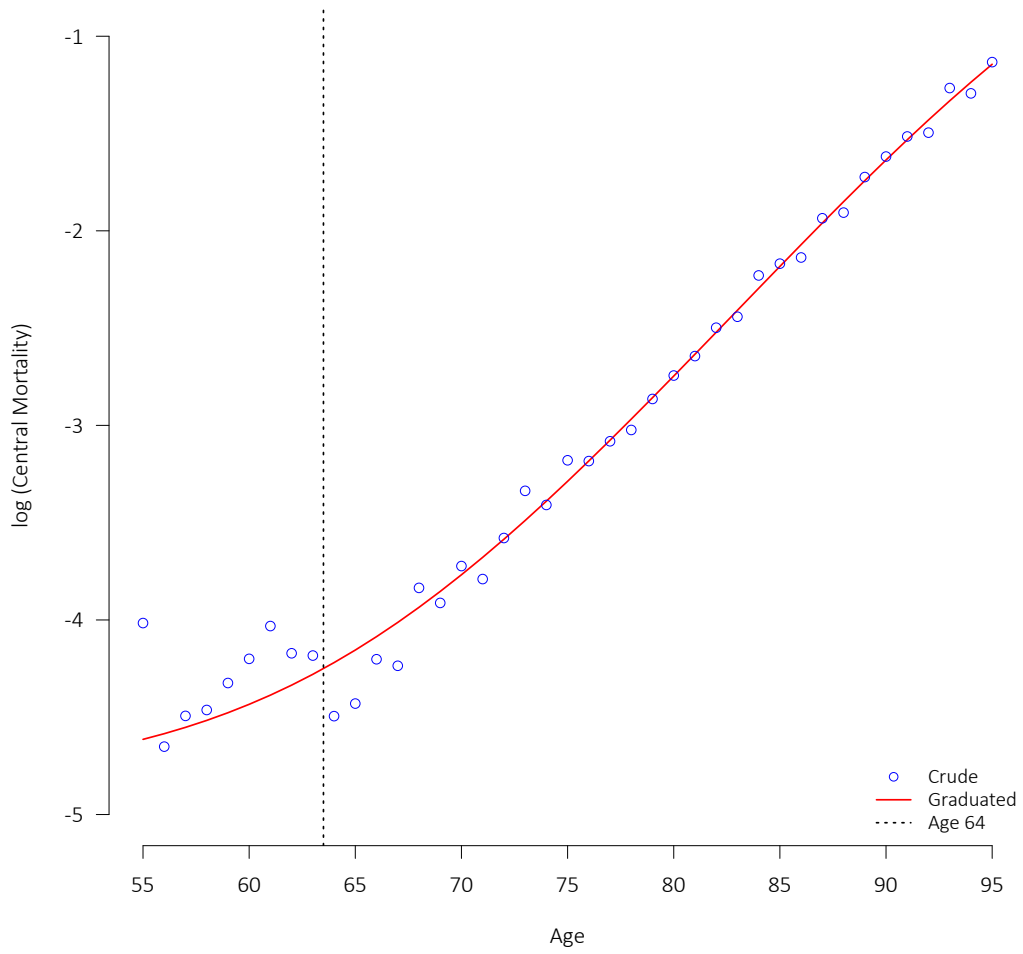Figure G.3: Crude and Graduated Rates for CPM2014's Size Band II for Males

Table G.1: Male Dataset with and without "Standardization" Adjustment

Age	Before Adjustment				After Adjustment			
	Lives		Amounts		Lives		Amounts	
	Deaths	Initial Exp.	Deaths	Initial Exp.	Deaths	Initial Exp.	Deaths	Initial Exp.
50	8.05	1,523.64	164.21	40,296.73	9.60	1,834.18	169.48	40,296.73
51	34.19	2,573.63	780.35	76,295.63	47.87	3,469.61	791.87	76,295.63
52	27.16	4,357.73	486.67	140,919.95	66.76	6,380.04	780.33	140,919.95
53	32.18	7,737.12	723.01	271,290.94	91.41	12,191.23	1,117.31	271,290.94
54	68.38	13,804.35	1,688.00	508,299.94	222.53	22,885.70	3,218.03	508,299.94
55	117.40	27,218.52	3,063.13	970,104.05	334.65	43,711.34	4,760.29	970,104.05
56	180.98	37,712.77	4,971.59	1,337,728.67	474.28	60,263.56	6,823.71	1,337,728.67
57	217.38	45,254.96	6,791.85	1,608,443.22	432.88	72,593.76	7,749.55	1,608,443.22
58	262.31	51,240.61	7,905.39	1,827,856.72	540.14	82,181.80	9,520.21	1,827,856.72
59	329.14	55,714.27	9,258.99	1,979,945.01	786.23	88,979.72	12,088.87	1,979,945.01
60	395.20	61,379.12	11,298.30	2,126,245.80	802.44	95,240.29	13,626.92	2,126,245.80
61	473.97	63,872.04	12,804.34	2,158,791.59	939.94	96,987.93	15,479.09	2,158,791.59
62	525.77	63,261.81	14,424.38	2,084,253.12	954.10	93,805.89	16,161.43	2,084,253.12
63	573.09	63,586.11	14,851.92	2,039,342.38	1,002.59	92,019.21	16,969.00	2,039,342.38
64	647.74	64,562.76	16,992.26	1,869,058.53	886.05	84,283.58	17,681.37	1,869,058.53
65	768.38	71,295.59	15,797.94	1,700,679.15	857.60	77,338.67	16,022.11	1,700,679.15
66	872.67	72,265.38	16,966.70	1,639,990.27	908.51	74,644.97	16,935.33	1,639,990.27
67	1,003.11	70,896.20	18,437.15	1,548,993.29	997.42	70,607.06	18,117.21	1,548,993.29
68	1,175.44	69,866.44	20,254.92	1,464,072.63	1,111.16	67,006.80	19,741.50	1,464,072.63
69	1,172.25	68,890.27	19,666.98	1,382,862.30	1,054.25	63,430.44	18,933.47	1,382,862.30
70	1,356.47	68,275.90	22,231.00	1,315,781.26	1,167.47	60,483.31	21,175.99	1,315,781.26
71	1,441.13	67,592.34	22,235.91	1,248,435.55	1,177.12	57,525.83	20,855.06	1,248,435.55
72	1,592.55	67,496.84	24,177.79	1,197,450.35	1,251.52	55,337.51	22,758.64	1,197,450.35
73	1,877.74	66,870.80	27,079.88	1,140,117.58	1,397.16	52,788.08	24,787.56	1,140,117.58
74	2,047.70	66,019.25	29,580.15	1,089,174.58	1,477.74	50,512.55	26,947.77	1,089,174.58
75	2,237.78	64,607.63	30,742.82	1,029,504.36	1,552.78	47,688.42	28,091.67	1,029,504.36
76	2,371.26	62,543.03	32,171.96	968,669.22	1,612.06	44,861.19	29,912.73	968,669.22
77	2,631.37	60,154.44	37,240.51	901,547.30	1,768.20	41,801.56	35,422.29	901,547.30
78	2,720.53	57,094.10	34,597.24	826,518.43	1,724.66	38,415.11	32,273.10	826,518.43
79	2,885.16	54,062.55	35,738.30	755,843.96	1,761.29	35,137.08	32,700.42	755,843.96
80	2,977.46	50,172.25	36,237.69	683,782.54	1,739.33	31,747.41	31,811.79	683,782.54
81	2,974.13	46,118.03	35,434.89	613,895.41	1,705.99	28,428.00	31,989.71	613,895.41
82	3,044.66	41,665.32	36,714.27	541,750.66	1,724.68	25,062.15	33,548.98	541,750.66
83	3,007.15	37,001.14	34,935.30	466,897.45	1,666.24	21,567.47	32,609.94	466,897.45
84	3,122.01	32,599.46	34,410.40	399,075.35	1,639.30	18,436.92	31,646.49	399,075.35
85	2,860.99	28,016.42	32,853.48	334,400.41	1,518.83	15,470.60	30,933.15	334,400.41
86	2,636.80	23,861.40	28,729.49	277,838.17	1,351.39	12,864.43	27,308.62	277,838.17
87	2,469.07	19,838.07	26,283.22	224,381.67	1,217.32	10,402.34	24,486.15	224,381.67
88	2,264.96	16,233.58	22,994.51	178,291.41	1,081.82	8,295.90	21,608.97	178,291.41
89	1,985.18	12,866.61	20,453.81	139,197.86	964.10	6,491.04	19,572.10	139,197.86
90	1,736.13	10,293.08	17,469.17	109,662.55	827.92	5,112.45	16,765.38	109,662.55
91	1,529.05	7,936.15	15,344.67	83,764.44	707.29	3,909.41	14,323.16	83,764.44
92	1,204.14	5,889.68	12,501.48	61,069.45	576.03	2,841.23	12,412.89	61,069.45
93	949.84	4,277.85	8,582.01	43,422.99	405.77	2,014.39	8,125.15	43,422.99
94	737.23	3,002.27	7,701.07	30,498.48	364.31	1,416.85	7,915.47	30,498.48
95	536.83	2,050.67	5,647.15	20,262.22	263.76	945.48	6,219.31	20,262.22
96	408.68	1,394.67	3,890.83	13,189.58	190.10	615.78	4,204.75	13,189.58
97	258.78	908.26	1,778.71	8,117.89	92.90	382.51	1,516.13	8,117.89
98	187.21	594.67	1,605.01	5,662.13	80.24	265.75	1,580.72	5,662.13
99	129.88	366.00	1,426.78	3,664.57	69.40	177.62	1,503.87	3,664.57
100	65.41	217.00	476.90	1,980.66	29.78	95.05	524.15	1,980.66

About The Society of Actuaries

The Society of Actuaries (SOA), formed in 1949, is one of the largest actuarial professional organizations in the world dedicated to serving 24,000 actuarial members and the public in the United States, Canada and worldwide. In line with the SOA Vision Statement, actuaries act as business leaders who develop and use mathematical models to measure and manage risk in support of financial security for individuals, organizations and the public.

The SOA supports actuaries and advances knowledge through research and education. As part of its work, the SOA seeks to inform public policy development and public understanding through research. The SOA aspires to be a trusted source of objective, data-driven research and analysis with an actuarial perspective for its members, industry, policymakers and the public. This distinct perspective comes from the SOA as an association of actuaries, who have a rigorous formal education and direct experience as practitioners as they perform applied research. The SOA also welcomes the opportunity to partner with other organizations in our work where appropriate.

The SOA has a history of working with public policymakers and regulators in developing historical experience studies and projection techniques as well as individual reports on health care, retirement, and other topics. The SOA's research is intended to aid the work of policymakers and regulators and follow certain core principles:

Objectivity: The SOA's research informs and provides analysis that can be relied upon by other individuals or organizations involved in public policy discussions. The SOA does not take advocacy positions or lobby specific policy proposals.

Quality: The SOA aspires to the highest ethical and quality standards in all of its research and analysis. Our research process is overseen by experienced actuaries and non-actuaries from a range of industry sectors and organizations. A rigorous peer-review process ensures the quality and integrity of our work.

Relevance: The SOA provides timely research on public policy issues. Our research advances actuarial knowledge while providing critical insights on key policy issues, and thereby provides value to stakeholders and decision makers.

Quantification: The SOA leverages the diverse skill sets of actuaries to provide research and findings that are driven by the best available data and methods. Actuaries use detailed modeling to analyze financial risk and provide distinct insight and quantification. Further, actuarial standards require transparency and the disclosure of the assumptions and analytic approach underlying the work.