

## Incorporating Information on Insured Amounts to Improve Survival Rate Estimates from a Liability Perspective.

Andrey Ugarte Montero, Frank van Berkum, Michel Vellekoop & Torsten Kleinow

University of Amsterdam

L19: The Nineteenth International Longevity Risk and Capital Markets Solutions Conference, 16-17 September 2024

# Outline

Introduction

Model framework

Implementation of experiments

- Data

- Computation of liabilities

- Thoughts on weight determination

- Results

Food for thought and pending work

References

## Hypothetical case

Think, for instance, of an insurer with just 2 policyholders for a product whose pay-out depends on death/survival:

- ▶ First person has a benefit of €1 000 euro
- ▶ Second person has a benefit of €1 000 000 euro

From a liability perspective,

- ▶ the financial relevance of the survival of these two individuals is very different
- ▶ The outcome from the second life will generate differences that are much more financially material to the insurer.

**Key question from the industry:** Should mortality assumptions be estimated giving these two individuals the same level of importance (as a classical survival model would), or should models account for the difference in benefits when setting assumptions?

## Practical solution

- ▶ Some practitioners think that **mortality assumptions should try to predict as well as possible the survival of the second individual** even if this means increasing the chances of getting wrong the survival of the first one.
- ▶ They advocate for the use of **insured amount-weighted mortality models** (see, for instance, [Gordon \[2011\]](#)).
- ▶ It is usually said that these models will result in **lower mortality** (see, for example, [Richards \[2008\]](#)).
- ▶ Practitioners expect sum insured to be an indicator of affluence [[Haberman et al., 2014](#)].

# Maximum likelihood estimation (MLE) refresher

Under the MLE approach...

- ▶ We are interested in estimating the unknown parameters  $P$  of a function  $g$ . We do this in a way that the parameters can be considered the “most likely to be true” given data observed
- ▶ Consider a random variable  $T$ , and a subscript  $i$  denoting an individual, out of a group of  $N$  members.
- ▶ Assume the existence of a related joint probability (density) function that can be written as  $g(t; P)$

MLE will find the value of  $P$  that maximizes the probability of observing the data collected, which can be written as

$$L(P) = \prod_{i=1}^N g(t_i; P) \quad (1)$$

## “Issues” related to the classical MLE approach

Note how we have assumed that the experience of every observation is equally important. **This is what some insurers/practitioners do not like.**

Because of this, they resort to a methodology where they weight the observations by the sum insured/benefit. Consider for this a variation of the previous model,

$$L_w(P) = \prod_{i=1}^N g(t_i; P)^{w_i} \quad (2)$$

where  $w_i$  refers to a weight associated to each individual. In this particular case,  $w_i$  is a function of the size of the sum insured.

## Weighted MLE estimation in survival modeling

Consider a data set on survival of  $N$  individuals, where individual  $i$  is observed for  $t_i$  units of time. Moreover, assume that individual survival for individual  $i$  is weighted by  $w_i$ , which is some function of the sum insured. Then equation 2 can be written as

$$\begin{aligned} L &= \prod_{i=1}^N \{Pr(T_{x_i} = t_i)^{\delta_i} Pr(T_{x_i} > t_i)^{1-\delta_i}\}^{w_i} = \prod_{i=1}^N \{f_{x_i}(t_i)^{\delta_i} S_{x_i}(t_i)^{1-\delta_i}\}^{w_i} = \\ &= \prod_{i=1}^N \{\{S_{x_i}(t_i)\mu_{x_i+t_i}\}^{\delta_i} S_{x_i}(t_i)^{1-\delta_i}\}^{w_i} = \prod_{i=1}^N \{\mu_{x_i+t_i}^{\delta_i} S_{x_i}(t_i)\}^{w_i} = \prod_{i=1}^N \{\mu_{x_i+t_i}^{\delta_i} e^{-H_{x_i}(t_i)}\}^{w_i} \end{aligned} \quad (3)$$

where  $T_{x_i}$  is the future lifetime of individual  $i$ ,  $F_{x_i}(t) = Pr(T_{x_i} \leq t)$  is the probability of the individual surviving at most  $t$  years,  $S_{x_i}(t) = Pr(T_{x_i} > t)$  is the probability of surviving at least  $t$  years, and  $\delta_i$  is an indicator of whether the individual survived ( $\delta_i = 0$ ) or died ( $\delta_i = 1$ ) during the period of observation.

Notice that we have assumed that individuals aged  $x$  are subject to a force of mortality  $\mu_x$ , and a cumulative hazard function  $H_x(t)$ .

## Weighted MLE estimation in survival modeling

We will be assuming classical mortality laws from literature as shown in the table below.

Law	$\mu_x$	$H_x(t)$
Gompertz	$e^{\alpha+\beta x}$	$\frac{(e^{\beta t}-1)}{\beta} e^{\alpha+\beta x}$
Makeham	$e^\epsilon + e^{\alpha+\beta x}$	$te^\epsilon + \frac{(e^{\beta t}-1)}{\beta} e^{\alpha+\beta x}$
Perks	$\frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$	$\frac{1}{\beta} \log\left\{\frac{1+e^{\alpha+\beta(x+t)}}{1+e^{\alpha+\beta x}}\right\}$
Makeham-Perks	$\frac{e^\epsilon + e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$	$te^\epsilon + \frac{1-e^\epsilon}{\beta} \log\left\{\frac{1+e^{\alpha+\beta(x+t)}}{1+e^{\alpha+\beta x}}\right\}$
Makeham-Beard	$\frac{e^\epsilon + e^{\alpha+\beta x}}{1+e^{\alpha+\rho+\beta x}}$	$te^\epsilon + \frac{e^{-\rho} - e^\epsilon}{\beta} \log\left\{\frac{1+e^{\alpha+\rho+\beta(x+t)}}{1+e^{\alpha+\rho+\beta x}}\right\}$

Table: Key functions per mortality law

## Merits of this methodology

### **Advantage:**

We can incorporate the effect of covariates in our estimations when computing the parameters associated to an individual. For a set of risk factors  $\mathcal{V} = \{C_1, C_2, \dots\}$ , where  $C_{j,i}$  as the value of  $C_j$  for individual  $i$ , and a parameter  $\zeta$ , we get

$$\zeta_i = \zeta_0 + \sum_{C_j \in \mathcal{V}} \Delta_{C_j} C_{j,i} \quad (4)$$

where  $\zeta_0$  denotes the baseline effect for the parameter,  $C_{j,i}$  denotes the value of risk factor  $C_j$  for individual  $i$ , and  $\Delta_{C_j}$  denotes the marginal change in the parameter per unit of change in  $C_j$ .

### **Disadvantage:**

Model does not consider time trends.

# Implementation of experiments

In our experiments, we will...

- ▶ estimate mortality rates using both a classical and a weighted approach in maximum likelihood estimation.
- ▶ estimate mortality rates using pension information as a covariate instead of a weight.
- ▶ compute the liabilities associated to an annuity portfolio

For our purposes, we will use a dataset provided by one of the largest insurance companies in the Netherlands.

This data set offers valuable information on the survival of individuals buying annuity products from the company.

# Outline

Introduction

Model framework

**Implementation of experiments**

**Data**

Computation of liabilities

Thoughts on weight determination

Results

Food for thought and pending work

References

# Data available: the basics

## Relevant elements:

- ▶ Survival information given in the form of yearly snapshots.
- ▶ The time period covers from 2010 to 2023.
- ▶ Estimations made with information on 986 369 individuals (705 123 men and 281 246 women).
- ▶ Total exposure used amounts to 9 470 576 years lived (6 765 289 years for males and 2 705 287 for females).
- ▶ Ages go from 35 to 100 .
- ▶ The data includes policyholder characteristics:
  - ▶ Age
  - ▶ Gender
  - ▶ Estimated salary
  - ▶ Sum insured
  - ▶ postcode based characteristics

# Characteristics available

Variable	Label	Description (categories)
<i>AG</i>	Age	Age
<i>BY</i>	Building year of houses	Before 1945, 1945 to 1995, around 2000 (after 1995)
<i>CH</i>	Children	Households in postcode mostly with or without children
<i>GE</i>	Gender	Male or female
<i>HT</i>	Housing terms	Owners, Renters
<i>OR</i>	Origin of individuals	Dutch (born in the Netherlands with and without Dutch background), other (born outside of the Netherlands)
<i>PC</i>	Pension class	We created pension classes: C1 (annuity amount in the lowest 20% of individuals of the same age and gender), C2, C3, and C4, and C5 (pension in the highest 20% of pensions).
<i>SA</i>	Salary	Annual salary of less than 25k, 25k to less than 50k, 50k to 250k, Over 250k, Unknown
<i>SB</i>	Level of social benefits received	Level 1, 2, 3, 4 (higher level, higher social benefits received)
<i>VH</i>	Value of house	Less than 250k, 250k to 400k, over 400k

**Table:** Summary of the variables used in the study.

The variables related to postcode correspond to the dominant category in the zipcode of the person. For instance, if most houses were built before 1945 in the postcode of a policyholder, the policyholder counts in the category "Before 1945" for variable BY.

# Outline

Introduction

Model framework

**Implementation of experiments**

Data

**Computation of liabilities**

Thoughts on weight determination

Results

Food for thought and pending work

References

# Computation of liabilities

Based on mortality estimations, we compute the value of a portfolio.

The benefit of the product is an annuity payable yearly from age 65 until death.

## Characteristics of the portfolio

- ▶ 30 000 policyholders
- ▶ Younger than 55 years old

Let  $S$  as the random variable denoting the total value of payments coming from this annuity portfolio. We adopt two approaches:

1. The results presented as “deterministic” are obtained as  
 $E(S) = \sum_{i=1}^{np} B_i \cdot {}_{65-x_i}E_{x_i} \ddot{a}_{65}$ , with  $B_i$  the annual pension received by individual  $i$  and  ${}_{65-x_i}E_{x_i} = v^{65-x_i} {}_{65-x_i}p_{x_i}$ .
2. A more stochastic approach involving the simulation of 1 000 possible scenarios. We disclose  $E(S)$ , the  $VaR_\alpha(S)$ , and the  $TVaR_\alpha = E(S | VaR_\alpha(S))$ .

We define  $VaR_\alpha(S)$  and  $TVaR_\alpha(S)$  as

$$VaR_\alpha(S) = \inf\{x \mid \mathcal{P}(S > x) \leq (1 - \alpha)\} \quad (5)$$

and

$$TVaR_\alpha(S) = E(S > VaR_\alpha(S)) \quad (6)$$

# Outline

Introduction

Model framework

**Implementation of experiments**

Data

Computation of liabilities

**Thoughts on weight determination**

Results

Food for thought and pending work

References

## General thoughts on weight determination

Notice that the loglikelihood derived from Equation 3 turns out to be

$$l = \log(L) = \sum_{i=1}^N w_i \{ \delta_i \log(\mu_{x_i+t_i}) - H_{x_i}(t_i) \} = \sum_{i=1}^N w_i \delta_i \log(\mu_{x_i+t_i}) - \sum_{i=1}^N w_i H_{x_i}(t_i) \quad (7)$$

- ▶ Equation 7 is the equation to be maximized using the experience of a population.
- ▶ Weights can take multiple forms!
- ▶ The industry tends to **use directly the sum insured** of the individual.
- ▶ As suggested by Richards [2008], however, it makes sense if  $\sum_{i=1}^N w_i = N$ . We adopt this suggestion.

## Alternative weight 1

For an individual  $i$ ,

$$w_{1,i} = \frac{B_{i,2015}}{\frac{\sum_{j=1}^N B_{j,2015}}{N}} = \frac{B_{i,2015}}{\bar{B}_{2015}} \quad (8)$$

where  $\bar{B}_{2015}$  denotes the average benefit in 2015-euro values.

Notice how this weight gives an indication of benefit amount of an individual.

Clearly,

$$\sum_{i=1}^N w_{1,i} = \sum_{i=1}^N \frac{B_{i,2015}}{\bar{B}_{2015}} = \frac{1}{\bar{B}_{2015}} \sum_{i=1}^N B_{i,2015} = \frac{1}{\bar{B}_{2015}} N \cdot \bar{B}_{2015} = N \quad (9)$$

which satisfies our imposed constraint.

## Alternative weight 2( $w_2$ )

In a more granular determination of weight, we explore the results of introducing a weight that does not only depend on sum insured but considers as well the age and gender structure of the pool. Consider as  $N_g$  the respective number of individuals in our pool per gender  $g = \{m, f\}$ , for males and females respectively. Similarly, consider  $N_{g,x}$  as the number of individuals of a given gender aged  $x$ . Clearly,

$$N = N_m + N_f = \sum_{x=35}^{100} N_{m,x} + \sum_{x=35}^{100} N_{f,x} \quad (10)$$

Refining the notation previously used, consider  $B_i^{g,x}$  as the benefit corresponding to individual  $i$  in the group with gender  $g$  and age  $x$ . Then we will consider weights by age and gender calculated as

$$w_i^{(g,x)} = \frac{B_{i,2015}^{g,x}}{\frac{\sum_{j=1}^{N_{g,x}} B_{j,2015}^{g,x}}{N_{g,x}}} = \frac{B_{i,2015}^{g,x}}{\bar{B}_{2015}^{g,x}} \quad (11)$$

## Alternative weight 2 ( $w_2$ )

Notice that under this approach,

$$\sum_{i=1}^{N_{g,x}} w_i^{(g,x)} = \sum_{i=1}^{N_{g,x}} \frac{B_{i,2015}^{g,x}}{\bar{B}_{2015}^{g,x}} = \frac{\sum_{i=1}^{N_{g,x}} B_{i,2015}^{g,x}}{\bar{B}_{2015}^{g,x}} = \frac{N_{g,x} \bar{B}_{2015}^{g,x}}{\bar{B}_{2015}^{g,x}} = N_{g,x} \quad (12)$$

and so,

$$\begin{aligned} \sum_{i=1}^N w_i^{(g,x)} &= \sum_{x=35}^{100} \sum_{k=1}^{N_{m,x}} w_k^{(m,x)} + \sum_{x=35}^{100} \sum_{k=1}^{N_{f,x}} w_k^{(f,x)} = \\ &= \sum_{x=35}^{100} N_{m,x} + \sum_{x=35}^{100} N_{f,x} = N_m + N_f = N \end{aligned} \quad (13)$$

which is again comparable to the number of terms in the unweighted version.

# Outline

Introduction

Model framework

**Implementation of experiments**

Data

Computation of liabilities

Thoughts on weight determination

**Results**

Food for thought and pending work

References

# Computation of liabilities

Portfolio: 30 000 lives younger than 55 Simulations: 1 000 Gini index model: 0.7168 Gini index portfolio: 0.7362 Age annuity: 65 Discount rate: 1%				
	Unweighted (in MM EUR)	Weighted $w_1$ (in MM EUR)	Weighted $w_2$ (in MM EUR)	Pension as covariate (in MM EUR)
<b>Gompertz</b>				
<i>Deterministic</i>				
Best estimate	14 250.8	14 019.1	13 770.31	14 068.21
<i>Simulation</i>				
Best estimate	14 246.72	14 022.85	13 772.35	14 073.32
95-th quantile	14 394.39	14 180.63	13 928.03	14 227.91
TVaR	14 430.92	14 225.11	13 962.51	14 258.52
<b>Makeham</b>				
<i>Deterministic</i>				
Best estimate	14 199.96	13 711.95	13 597.23	13 911.4
<i>Simulation</i>				
Best estimate	14 198.2	13 712.99	13 595.61	13 911.12
95-th quantile	14 343.04	13 859.45	13 748.39	14 056.34
TVaR	14 376.09	13 892.3	13 788.71	14 086.26
<b>Perks</b>				
<i>Deterministic</i>				
Best estimate	14 427.54	14 189.63	13 926.69	14 250.2
<i>Simulation</i>				
Best estimate	14 422.56	14 188.33	13 926.98	14 252.78
95-th quantile	14 565.41	14 361.68	14 080.15	14 400.81
TVaR	14 608.61	14 400.91	14 115.75	14 438.29

Table: Summary of the liability values.

# Computation of liabilities

Portfolio: 30 000 lives younger than 55				
Simulations: 1 000				
Gini index model: 0.7168	Unweighted	Weighted $w_1$	Weighted $w_2$	Pension as
Gini index portfolio: 0.7362	(in MM EUR)	(in MM EUR)	(in MM EUR)	covariate
Age annuity: 65				(in MM EUR)
Discount rate: 1%				
<b>Makeham-Perks</b>				
<i>Deterministic</i>				
Best estimate	14 369.9	14 369.9	14 369.9	14 203.34
<i>Simulation</i>				
Best estimate	14 368.48	14 370.51	14 367.51	14 200.62
95-th quantile	14 514.68	14 518.35	14 511.2	14 348.41
TVaR	14 553.17	14 546.43	14 540.77	14 390.37
<b>Makeham-Beard</b>				
<i>Deterministic</i>				
Best estimate	14 215.43	13 713.85	13 600.39	13 906.06
<i>Simulation</i>				
Best estimate	14 215.52	13 714.65	13 599.81	13 907.02
95-th quantile	14 361.54	13 863.01	13 751.61	14 051.42
TVaR	14 390.04	13 900.95	13 782.97	14 090.27

Table: Summary of the liability values.

# Weighted mortality using $w_1$ is higher than the unweighted equivalent?

Let's take a look at some basic descriptive statistics to understand what the weighting under  $w_1$  is doing.

Pension band	Pension value	Headcount	Average age	Total weight	Total exposure	Headcount females	Total weight females	Average age females	Proportion females (%)	Proportion weight females (%)
B1	Less than 1 000	541 062	57.94	72 250.55	5 315 202.00	184 138	22 541.22	54.57	34.03	31.20
B2	[1 000, 2 000[	154 841	59.05	82 499.95	1 553 965.00	41 253	21 827.28	55.21	26.64	26.46
B3	[2 000, 3 000[	77 139	59.71	70 490.66	763 995.75	18 238	16 553.71	55.77	23.64	23.48
B4	[3 000, 4 000[	46 571	60.49	60 062.15	451 128.25	9 914	12 744.72	56.05	21.29	21.22
B5	[4 000, 5 000[	31 931	61.05	53 151.20	304 803.50	6 372	10 605.72	56.16	19.96	19.95
B6	[5 000, 6 000[	22 663	61.48	46 159.09	211 120.75	4 267	8 692.48	55.83	18.83	18.83
B7	[6 000, 7 000[	17 121	61.43	41 289.03	151 786.75	3 037	7 321.57	55.17	17.74	17.73
B8	7 000 or more	95 041	60.28	560 466.37	7 185 74.25	14 027	72 110.22	52.24	14.76	12.87

Table: Descriptive statistics by pension band

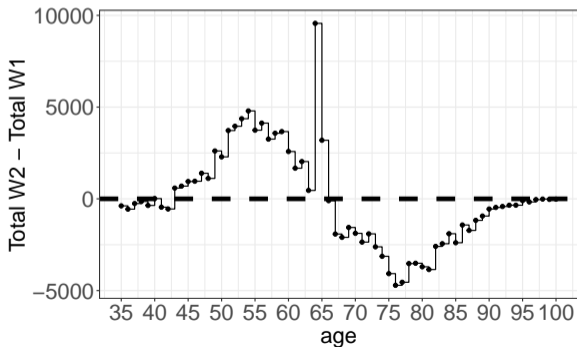
This suggests that the model is paying disproportionate attention to the experience of (fewer) older males, which results in higher mortality. Using directly the sum insured as weights yields the same type of result.

This effect can be an example of the situation described by [Madrigal et al. \[2011\]](#): *It is highly likely that these members (members representing a bigger proportion of liabilities) are also towards the top end of the life spectrum.*

## Alternative weight 2 ( $w_2$ )

Given the previous results, it makes sense to adjust the weights in a way that limits the influence of a particular group of older men in parameter estimation. The definition of weights that are gender and age-dependent such as the one given for  $w_2$  is helpful in this sense.

Figure: Difference in total representation by age



## Weighted mortality continues to be higher under $w_2$ ?

Weighting mortality using  $w_2$  continues to result in higher mortality than in the unweighted model!

Let's take a look at more granular descriptive statistics. In the case of males,

Variable Category	Unweighted				Weight 1 ( $w_1$ )				Weight 2 ( $w_2$ )			
	N	Deaths	Exposure	Rate	N	Deaths	Exposure	Rate	N	Deaths	Exposure	Rate
<b>Males</b>												
<i>Salary</i>												
Less 25k	145667	10512.00	1289843.75	8.15	114913.00	13383.23	945874.33	14.15	96043.80	10276.48	779043.53	13.20
25k to less 50k	196333	6770.00	1958570.50	3.457	193315.91	11143.10	1690471.36	6.592	171996.88	9186.36	1494324.91	6.147
50k to less 250k	128419	3376.00	1155961.25	2.921	318980.96	12507.35	2477895.55	5.048	287540.49	10583.21	2210382.66	4.788
Over 250k	611	18.00	4683.25	3.843	11749.39	418.08	81219.25	5.148	10370.28	356.95	68283.57	5.23
Unknown	234093	42782.00	2356230.25	18.157	175012.82	47936.56	1625733.77	29.49	139171.54	35537.92	1284500.70	27.67
<i>Value of House</i>												
Less 250K	163890	20875	1559675.75	13.384	130861.94	19904.96	1098552.39	18.119	112045.57	15321.3	932924.21	16.423
250k to 400k	333775	27998.00	3221486.75	8.691	360351.68	36733.06	3034616.82	12.105	312638.34	28427.02	2604220.67	10.916
Over 400k	207458	14585	1984126.50	7.351	322758.45	28750.30	2688025.06	10.696	280439.08	22192.552	2299390.50	9.651
<i>Pension class</i>												
C1	141052	12984	1360816.75	9.541	5473.99	555.46	54315.40	10.227	4717.01	424.34	46507.73	9.124
C2	141013	11995	1401875.25	8.556	22789.27	2118.53	230815.09	9.178	19752.76	1624.506	198913.95	8.167
C3	141012	12673	1418097.75	8.937	53489.38	5098.889	544551.72	9.363	46549.50	3936.78	472059.50	8.340
C4	141010	12598	1400702.50	8.994	129634.32	12966.00	1304339.52	9.941	112753.64	9999.70	1128568.09	8.861
C5	141036	13208	1183796.75	11.157	602585.12	64649.45	4687172.52	13.793	521350.08	49955.60	3990486.10	12.519

## Weighted mortality continues to be higher under $w_2$ ?

We observe a similar behaviour in the case of females.

Variable Category	Unweighted				Weight 1 ( $w_1$ )				Weight 2 ( $w_2$ )			
	N	Deaths	Exposure	Rate	N	Deaths	Exposure	Rate	N	Deaths	Exposure	Exposure
<b>Females</b>												
<i>Salary</i>												
Less 25k	73 568	1 725.00	673 519.25	2.56	24 840.29	973.88	206 521.60	4.72	41 798.43	1 597.18	351 478.10	4.54
25k to less 50k	107 569	1 353.00	1 050 191.75	1.29	71 308.12	1 278.44	577 132.52	2.22	115 204.95	2 201.81	970 113.89	2.27
50k to less 250k	25 261	209.00	212 743.50	0.98	53 693.56	534.62	348 319.22	1.53	85 516.84	894.56	590 856.17	1.51
Over 250k	40		306.00		333.02		2 285.36		10 370.29		4 360.96	
Unknown	74 808	5 052.00	768 526.75	6.57	22 221.94	2 747.50	221 287.04	12.42	38 124.23	4 265.52	381 397.34	11.18
<i>Value of house</i>												
Less 250K	61 239	3 115.00	575 821.75	5.41	28 253.88	1 831.85	222 565.00	8.23	45 836.69	2 919.75	373 269.41	7.82
250k to 400k	133 948	3 540.00	1 295 392.00	2.73	76 096.35	2 288.73	605 058.29	3.78	124 109.72	3 733.45	1 023 833.54	3.65
Over 400k	86 059	1 684.00	834 073.50	2.02	68 046.70	1 413.84	527 922.45	2.68	111 299.59	2 305.86	901 103.50	2.56
<i>Pension class</i>												
C1	56 282	1 623.00	543 403.50	2.99	1 132.54	47.66	11 159.71	4.27	1 854.05	75.34	18 354.51	4.10
C2	56 237	1 556.00	558 439.00	2.79	4 960.70	174.04	49 999.73	3.48	8 186.73	276.50	83 085.10	3.33
C3	56 231	1 624.00	568 097.50	2.86	11 667.00	388.28	119 039.23	3.26	19 379.35	626.94	199 070.46	3.15
C4	56 234	1 724.00	563 077.75	3.06	26 646.06	891.87	267 656.01	3.33	44 368.17	1 447.88	450 266.62	3.22
C5	56 262	1 812.00	471 639.50	3.84	127 990.62	4 032.58	907 691.06	4.44	207 457.70	6 532.40	1 547 429.76	4.22

- ▶ Mortality shows a clear decreasing trend when considering groups likely to be richer (higher estimated salaries and more expensive housing are linked with lower rates).
- ▶ The pattern is just not there when it comes to the pension categories!

## Weighted mortality is actually higher?

Take a look at the following table:

Variable/Pension class	C1	C2	C3	C4	C5	Total Variable
Salary						
Less 25k	61 156	52 812	42 531	35 866	26 870	219 235
25k to less 50k	28 592	56 181	70 920	79 886	68 323	303 902
50k to less 250k	9 546	14 487	24 364	35 899	69 384	153 680
Over 250k	15	17	62	61	496	651
Unknown	98 025	73 753	59 366	45 532	32 225	308 901
Value of house						
Less 250K	57 668	49 082	44 221	40 341	33 817	225 129
250k to 400k	93 100	94 858	94 331	94 578	90 856	467 723
Over 400k	46 566	53 310	58 691	62 325	72 625	293 517
Total pension class	197 334	197 250	197 243	197 244	197 298	986 369

Table: Distribution of salary by pension class

- ▶ Results suggest that the highest pension group has experienced higher mortality.
- ▶ Pension classes present a diverse composition regarding the salary categories. This will of course result in a mortality effect that is not fully aligned with income.
- ▶ Similarly, we can see that the highest pension classes are also diverse by the potential value of houses based on post-code.

# Results

- ▶ This indeed suggest that pension amounts are not necessarily a good predictor to understand how wealthy individuals are in our dataset, which may result in more atypical mortality outcomes like the ones we have observed.
- ▶ *Why would we expect them to have exceptionally lower mortality?*

Work by [Madrigal et al. \[2011\]](#) indicate that salary can be a better “affluence-based longevity predictor” than pension in certain populations.

## Food for thought

- ▶ Annuity amounts are not always the best indicator of financial wealth, and so mortality patterns may not be as clear.
- ▶ Weighting these models by sum insured can introduce significant bias in parameter estimation. It's important to understand the implications of this in your models!
- ▶ It is also important to be transparent of what this may mean in liability estimation. The difference is not necessarily towards a more conservative result.
- ▶ Can we really say that these models are better tuned with liabilities when dealing with annuities/pensions?
- ▶ We will analyze the results using generalized additive models including time trends at a later stage.

*Thanks!*  
*Dank u wel!*  
*¡ Gracias!*  
*Merci Beaucoup!*  
*Obrigado!*

Tim Gordon. Weighted mortality experience analysis (Presentation available on-line from Aon Hewitt), 2011.

Steven Haberman, Vladimir Kaishev, Pietro Millosovich, Andrés M. Villegas, Steven Baxter, Andrew Gaches, Sveinn Gunnlaugsson, and Mario Sison. Longevity basis risk. A methodology for assessing basis risk. *Institute and Faculty of Actuaries*, (December), 2014. URL <http://www.actuaries.org.uk/documents/longevity-basis-risk-methodology-assessing-basis-risk>.

A. M. Madrigal, F. E. Matthews, D. D. Patel, A. T. Gaches, and S. D. Baxter. What Longevity Predictors Should Be Allowed for When Valuing Pension Scheme Liabilities? *British Actuarial Journal*, 16(1):1–38, 2011. ISSN 1357-3217. doi: 10.1017/s1357321711000018.

S. J. Richards. Applying Survival Models to Pensioner Mortality Data. *British Actuarial Journal*, 14(2):257–303, jul 2008. ISSN 1357-3217. doi: 10.1017/S1357321700001720. URL [https://www.cambridge.org/core/product/identifier/S1357321700001720/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S1357321700001720/type/journal_article).