Network of Consulting Actuaries Webinar

A VaR approach to mis-estimation risk

Stephen J. Richards

26th February 2021

Overview

1. Mis-estimation risk
2. Portfolio features
3. Parameter estimation
4. Preconditions
5. Run-off mis-estimation
6. Value-at-risk mis-estimation
7. Comparisons
8. Conclusions

1 Mis-estimation risk

1 Longevity risk

> "the PRA considers that longevity risk includes at least two sub-risks [. . .] namely, base mis-estimation risk and future improvement risk"
> Woods [2016]

1 Mis-estimation risk

"[. . .] the risk that the base mortality estimate is incorrect (i.e. the mortality estimate based on actual experience in the portfolio)" Burgess et al. [2010]

"How wrong could our base mortality assumptions be, or: what if our historical experience did not reflect the underlying mortality?" Armstrong [2013]

1 Measuring mis-estimation

"Mis-estimation risk lends itself to statistical analysis if there is sufficient accurate data" Armstrong [2013]

"The impact of uncertainty should always be quantified financially" Makin [2008]

1 Mis-estimation risk

LONGEVITAS

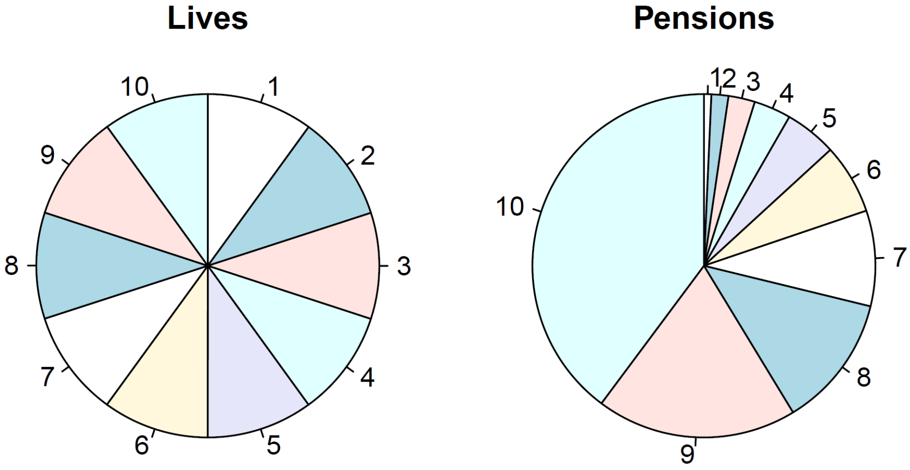
- Uncertainty over current mortality rates,
- Assessed using actual portfolio experience,
- Modelled statistically, and
- Measured financially.

2 Portfolio features

2 Portfolio features

- Medium-sized UK pension scheme.
- Pensioners only.
- Cashflows discounted at 0.75% p.a.

2 Liability concentration



Source: Data from Richards [2016, Appendix 1].

2 Estimation risk

- Top decile of pensioners receives 39.8% of pensions.
- Next two deciles receive further 31.4%.
- Liabilities highly concentrated.

2 Example model

Consider a time-varying model for the mortality hazard:

$$\mu_{x,y} = \frac{e^{\epsilon} + e^{\alpha+\beta x+\delta(y-2000)}}{1 + e^{\alpha+\beta x+\delta(y-2000)}}$$

where x is exact age and y is calendar time[†].

[†] -2000 is an offset to keep parameters well scaled.

2 Example model

Consider simple approach to gender and pension size for life i:

$$\begin{aligned}\alpha_i = \alpha_0 &+ \alpha_{\text{male}} z_{i,\text{male}} \\ &+ \alpha_{\text{decile 8 or 9}} z_{i,\text{decile 8 or 9}} \\ &+ \alpha_{\text{decile 10}} z_{i,\text{decile 10}}\end{aligned}$$

- α_j is the effect of risk factor j.
- $z_{i,j}$ is an indicator taking the value 1 if life i has risk factor j and zero otherwise.

2 Portfolio features

Parameter	Estimate	Std. Err	Lives
β	0.148	0.005	15,698
α_{male}	0.479	0.060	5,956
α_0	-14.731	0.491	15,698
ϵ	-5.420	0.154	15,698
$\alpha_{\text{decile 8 or 9}}$	**-0.180**	0.078	3,140
$\alpha_{\text{decile 10}}$	**-0.313**	0.108	1,567
δ	-0.046	0.016	15,698

Source: Parameter estimates from Richards [2016, Table 6].

2 Coefficient of variation

$$\text{Coefficient of variation} = \frac{\text{Standard error}}{|\text{Estimate}|}$$

Measures relative uncertainty over parameter estimate.

2 Coefficient of variation

Parameter	Coef. of variation
β	0.03
α_{male}	0.13
α_0	0.03
ϵ	0.03
$\alpha_{\text{decile 8 or 9}}$	**0.43**
$\alpha_{\text{decile 10}}$	**0.35**
δ	0.35

Source: own calculations from estimates in Richards [2016, Table 6].

2 Estimation risk

- Liabilities highly concentrated.
- Sub-groups with most liability have lowest mortality…

  … *and* high relative uncertainty.

A perfect storm of actuarial risk!

3 Parameter estimation

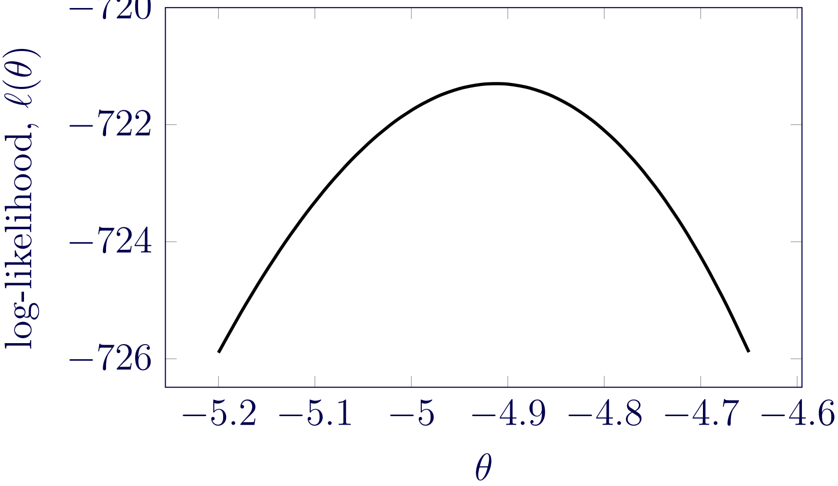
In a statistical model with m parameters:

- Consider a parameter vector, $\underline{\boldsymbol{\theta}} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \\ \vdots \\ \boldsymbol{\theta}_m \end{pmatrix}$.
- We have an estimate, $\hat{\underline{\boldsymbol{\theta}}}$, which is uncertain.
- Uncertainty over $\hat{\underline{\boldsymbol{\theta}}}$ is estimation risk.

3 Assumption 1

$\ell(\underline{\boldsymbol{\theta}})$ is the log-likelihood function for a model.

3 $\ell(\theta)$ in one dimension



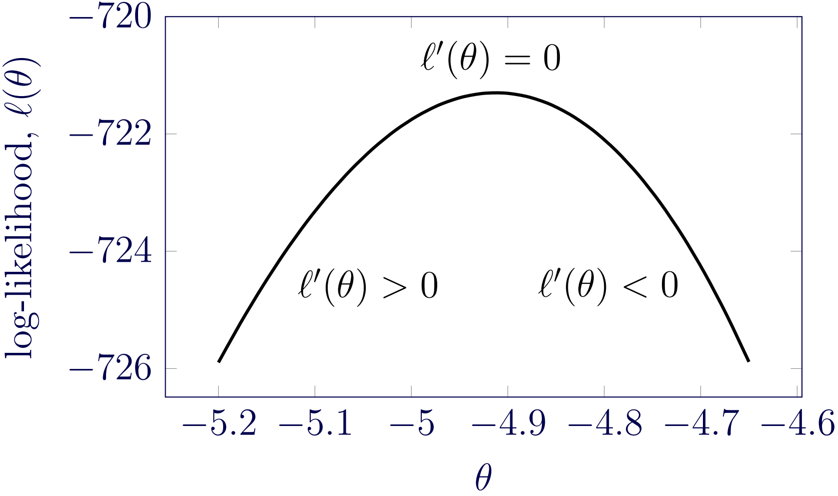
Source: Richards [2016, Figure 1].

3 Assumption 2

All first partial derivatives of $\ell(\underline{\boldsymbol{\theta}})$ exist, i.e.

$$\ell'(\underline{\boldsymbol{\theta}}) = \begin{pmatrix} \dfrac{\partial \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_1} \\ \dfrac{\partial \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_2} \\ \vdots \\ \dfrac{\partial \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_m} \end{pmatrix}$$

3 $\ell(\theta)$ in one dimension



Source: Richards [2016, Figure 1].

3 Assumption 3

The Hessian matrix, $\boldsymbol{H}(\underline{\boldsymbol{\theta}})$, of all second partial and cross-partial derivatives of $\ell(\underline{\boldsymbol{\theta}})$ exists:

$$\boldsymbol{H}(\underline{\boldsymbol{\theta}}) = \begin{pmatrix} \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_1^2} & \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_1 \partial \theta_m} \\ \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_2^2} & \cdots & \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_2 \partial \theta_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_m \partial \theta_1} & \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_m \partial \theta_2} & \cdots & \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_m^2} \end{pmatrix}$$

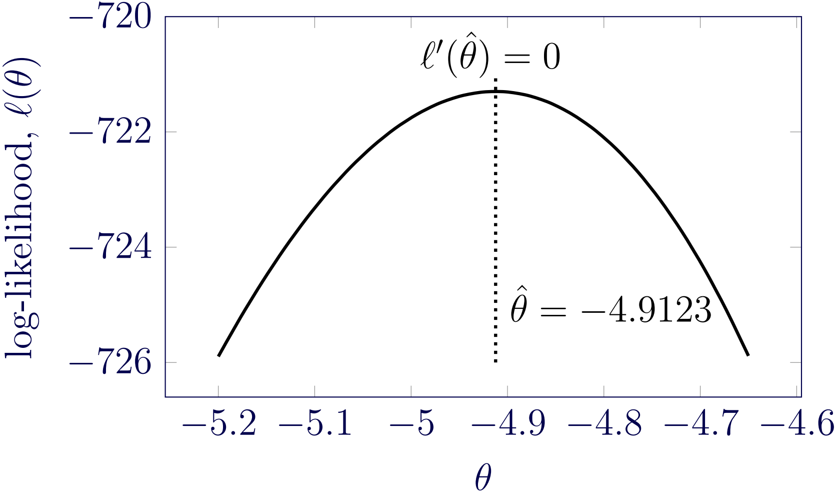
3 Maximum likelihood

$\underline{\hat{\boldsymbol{\theta}}}$ is the maximum-likelihood estimate of $\underline{\boldsymbol{\theta}}$ if:

- $\ell'(\underline{\hat{\boldsymbol{\theta}}}) = 0$, and
- $\boldsymbol{H}(\underline{\hat{\boldsymbol{\theta}}})$ is negative semi-definite†.

† $\underline{\hat{\boldsymbol{\theta}}}^{\mathrm{T}} \boldsymbol{H}(\underline{\hat{\boldsymbol{\theta}}}) \underline{\hat{\boldsymbol{\theta}}} \leq 0$, where $\underline{\hat{\boldsymbol{\theta}}}^{\mathrm{T}}$ is the transpose of $\underline{\hat{\boldsymbol{\theta}}}$.

3 $\ell(\theta)$ in one dimension



Source: Richards [2016, Figure 1].

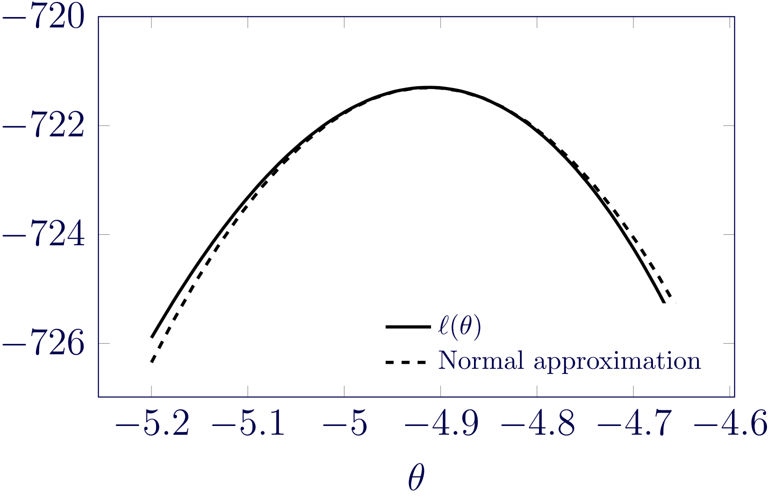
3 Maximum-likelihood theorem

$\underline{\hat{\boldsymbol{\theta}}}$ has a multivariate normal (MVN) distribution†:

- Mean vector $\underline{\hat{\boldsymbol{\theta}}}$, and
- Covariance matrix $\hat{\boldsymbol{\Sigma}} = -\boldsymbol{H}^{-1}(\underline{\hat{\boldsymbol{\theta}}})$.

† Cox and Hinkley [1996, Chapter 9].

3 $\ell(\theta)$ in one dimension



Source: Richards [2016, Figure 1].

3 Estimation risk

If $\hat{\underline{\boldsymbol{\theta}}}$ has a MVN distribution, all estimation risk is summarised in $\hat{\boldsymbol{\Sigma}}$:

- The leading diagonal has the variance of $\hat{\underline{\boldsymbol{\theta}}}$, and
- The off-diagonal entries have the covariances of $\hat{\underline{\boldsymbol{\theta}}}$.

3 (Mis-)Estimation risk

- Estimation risk is statistical parameter uncertainty.
- Mis-estimation risk is the financial impact of that uncertainty.

4 Preconditions

4 Preconditions

> "What assumptions are you making, e.g. independence? Duplicate policies? Amounts vs lives?"
> Armstrong [2013]

4 Preconditions

- Deduplicate records[†].
- Lives-based statistical model.
- Amounts effect on mortality handled as either:
  - Categorical factor, e.g. pension decile, or
  - Continuous covariate, e.g. using exact pension[‡].

† See Macdonald et al. [2018, Section 2.5].

‡ See Richards [2020a].

4 Preconditions

Mis-estimation capital underestimated if:

- Records not deduplicated, or
- Amounts effect on mortality ignored, or
- Time trend not included in model.

5 Run-off mis-estimation

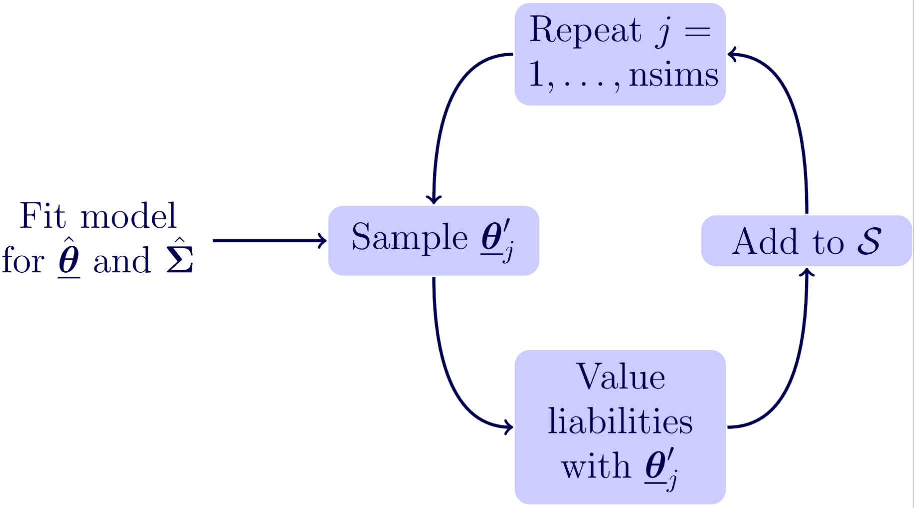
5 Run-off mis-estimation

- Need set, $\mathcal{S}$, of liability valuations subject to parameter risk.
- Can then calculate percentiles of $\mathcal{S}$.

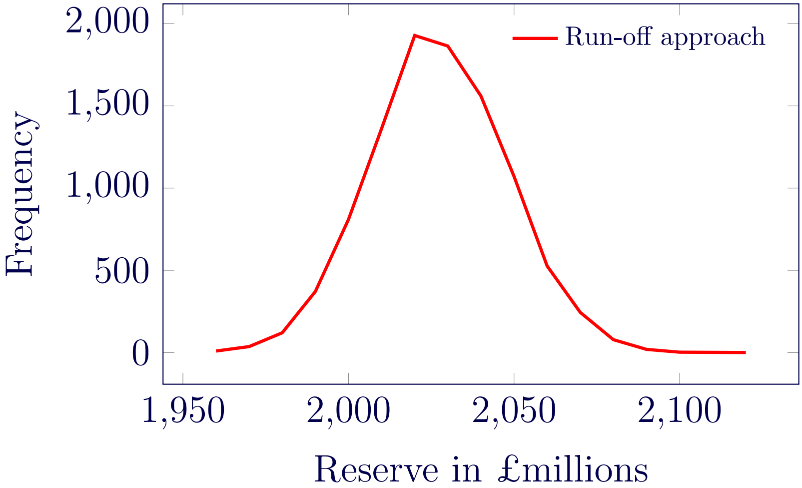
5 Run-off mis-estimation

- Best-estimate parameter vector $\hat{\underline{\boldsymbol{\theta}}}$.
- $\hat{\boldsymbol{\Sigma}}$ is estimated covariance matrix for $\hat{\underline{\boldsymbol{\theta}}}$.
- Alternative parameter vector, $\underline{\boldsymbol{\theta}}'$, can be sampled from $\mathrm{MVN}(\hat{\underline{\boldsymbol{\theta}}}, \hat{\boldsymbol{\Sigma}})$ using Monte Carlo simulation.

5 Richards [2016]



5 Distribution of $\mathcal{S}$



Source: Richards [2020c, Figure 4].

5 Richards [2016]

- $\underline{\boldsymbol{\theta}}'_j$ varies due to parameter risk only.
- Richards [2016] suitable for run-off valuation. . .

  . . . and pricing bulk annuities and reinsurance.

6 Value-at-risk mis-estimation

6 Issue with run-off approach

- No “one-year” element. . .
- Not an obvious fit for Solvency II.

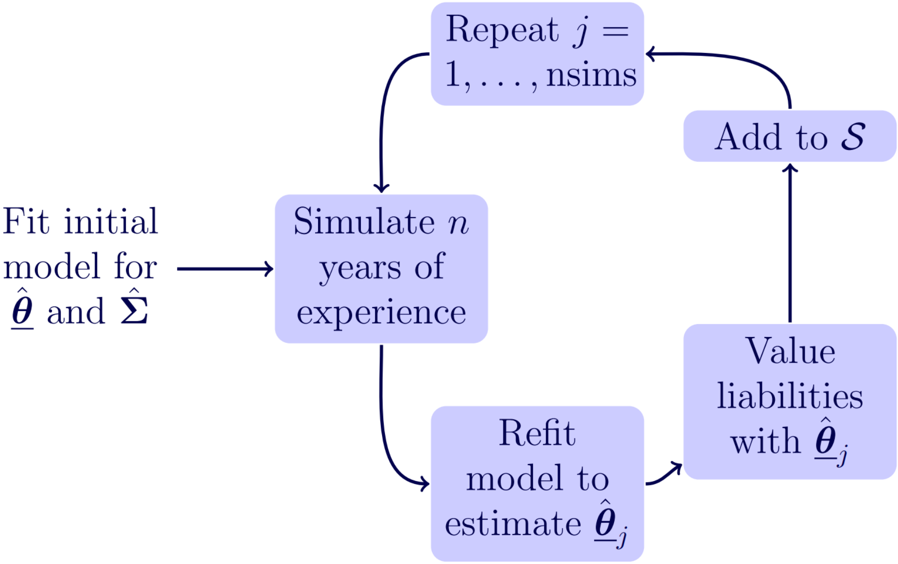
We make two changes to the algorithm:

1. Simulate n years of experience.
2. Refit model to re-estimate $\underline{\hat{\boldsymbol{\theta}}}_j$.

and proceed as before.

- $n = 1$ for Solvency II,
- $n = 3, 4, 5$ for ORSA.

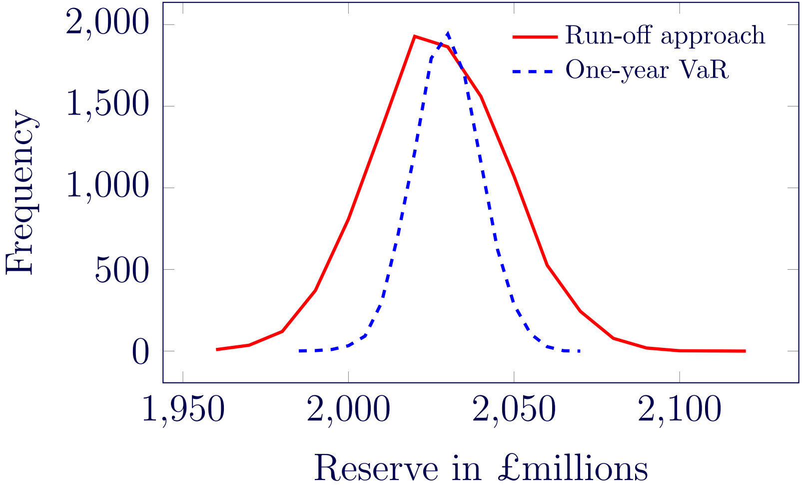
6 Richards [2020c]



6 Richards [2020c]

- $\underline{\hat{\boldsymbol{\theta}}}_j$ varies according to additional n years of experience data.
- True n-year VaR approach to mis-estimation.
- Calculate percentiles of $\mathcal{S}$ as before.

6 Distribution of $\mathcal{S}$



Source: Richards [2020c, Figures 4 and 6].

7 Comparisons

7 Model

- Survival model varying in age and time ($\mu_{x,y}$).
- Risk factors in model†:
  - Age
  - Gender
  - Normal v. early retirement
  - First life v. surviving spouse
  - Pension size
  - Time

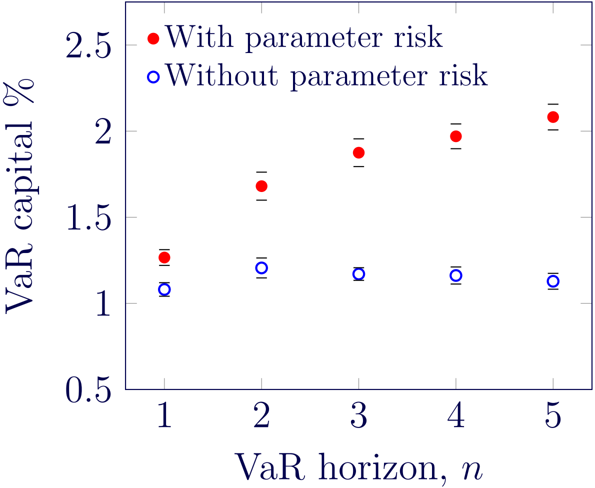
† Source: Richards [2020c, Table 3].

- Interested in 99.5% VaR capital requirement.
- VaR capital $= \left(\frac{99.5\% \text{ percentile of } \mathcal{S}}{\text{Mean of } \mathcal{S}} - 1 \right) \times 100\%.$
- 10,000 samples or simulations for each scenario.

Two options when simulating n years of experience:

1. Sample from $\text{MVN}(\underline{\hat{\boldsymbol{\theta}}}, \hat{\boldsymbol{\Sigma}})$, i.e. with parameter risk.
2. Use $\underline{\hat{\boldsymbol{\theta}}}$ each time, i.e. no parameter risk.

7 VaR capital by horizon



Source: Richards [2020c, Figure 5]. 99.5% VaR capital requirement from 10,000 simulations.

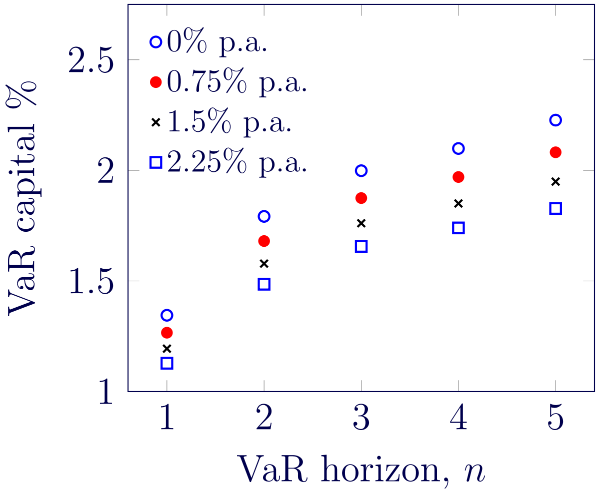
7 VaR capital by horizon

For mis-estimation VaR capital:

- Including parameter risk in simulations increases capital.
- However, most of 1-year capital not due to parameter risk.
- Only half of 5-year capital due to parameter risk.

Solvency ($n = 1$) different from ORSA ($n = 3, 4, 5$).

7 Role of discount rate



Source: Richards [2020c, Figure 7]. 99.5% VaR capital requirement from 10,000 simulations.

7 Role of discount rate

Lower discount rates mean higher mis-estimation VaR capital requirements.

7 Role of mortality law

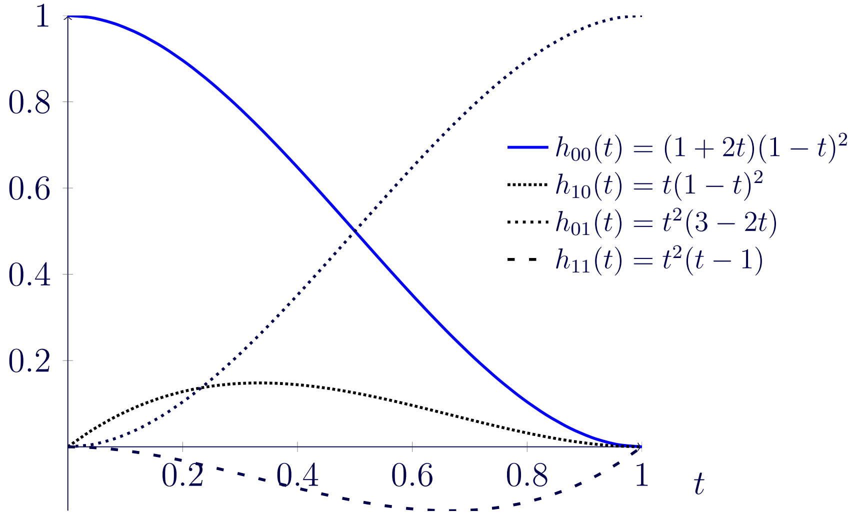
Various options for post-retirement mortality:

- Gompertz [1825], $\mu_x = e^{\alpha+\beta x}$.
- Perks [1932], $\mu_x = \dfrac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$.
- Beard [1959], $\mu_x = \dfrac{e^{\alpha+\beta x}}{1+e^{\alpha+\rho+\beta x}}$.
- Makeham-Perks, $\mu_x = \dfrac{e^{\epsilon}+e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$.
- Hermite-spline model of Richards [2020b].

7 A Hermite-spline model

- Designed for post-retirement mortality.
- Automatically allows for convergence with age.
- Uses Hermite splines for smoothness and flexibility.

7 A basis of Hermite splines



Source: Richards [2020b].

7 Hermite-spline model

- Let x_0 and x_1 be the minimum and maximum ages.
- Define $t = \frac{(x - x_0)}{(x_1 - x_0)}$, so $t \in [0, 1]$.
- $\log \mu_x = \alpha h_{00}(t) + \omega h_{01}(t) + m_0 h_{10}(t) + m_1 h_{11}(t)$

for parameters α, ω, m_0 and m_1 estimated from data.

Source: Richards [2020b].

7 Hermite-spline model

- Mortality differentials in α reduce with age due to path of h_{00}.

  $\rightarrow$ don't need age interactions as in other models.

  $\rightarrow$ requires fewer parameters.

7 Questions

1. Which models fit best by lives and amounts?
2. Do fewer parameters mean less mis-estimation risk?

7 Measures of fit

1. Measure fit by lives using AIC [Akaike, 1987].
2. Measure fit by amounts using bootstrapping [Richards, 2016, Section 8].

7 Measures of fit

Best fit:

- By lives: lowest AIC.
- By amounts: closest bootstrap percentage to 100%.

7 Model fits

Sorted by descending AIC:

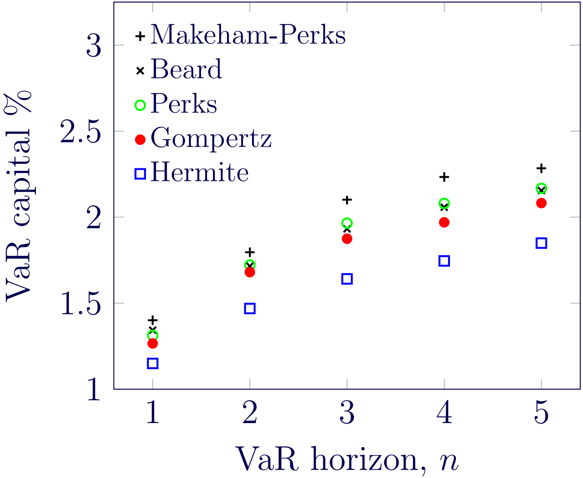
Mortality law	Parameters	AIC	Bootstrap
Gompertz	13	79,643	98.9%
Perks	13	79,638	99.3%
Beard	14	79,626	99.0%
Makeham-Perks	14	79,625	99.0%
Hermite	10	79,623	99.2%

Source: Data from Richards [2020c, Table 4].

7 Mortality laws

- Gompertz model has worst fit in terms of both AIC and explanation of amounts variation.
- Hermite model has:
  - Fewest parameters,
  - Best fit in terms of AIC, and
  - Second-best fit in terms of amounts.

7 Role of mortality law



Source: Richards [2020c, Figure 10]. 99.5% VaR capital requirement from 10,000 simulations. Discounting at 0.75% p.a.

7 Role of mortality law

- Models with fewer parameters have lower mis-estimation risk.
- Overly simple laws (Gompertz) don't explain enough amounts variation.

8 Conclusions

8 Conclusions — I

- Mis-estimation risk stems from having finite experience data.
- Quantification must be:
  - statistical to account for correlations, and
  - financial to account for concentration risk.

Two options for mis-estimation:

1. Richards [2016] — run-off approach for pricing block transactions.
2. Richards [2020c] — value-at-risk approach for one-year Solvency II reporting, ORSA etc.

8 Conclusions — III

For one-year VaR mis-estimation capital:

- Parameter risk accounts for a small proportion.
- Parsimonious models tend to have lower estimation risk.

8 Acknowledgements

Thanks to Gavin Ritchie of Longevitas and Patrick Kelliher of Crystal Risk Consulting for comments on earlier drafts.

Any errors or omissions remain the responsibility of the author.

References I

H. Akaike. Factor analysis and AIC. *Psychometrica*, 52:317–333, 1987. ISSN 0033–3123. doi: https://doi.org/10.1007/BF02294359.

K. Armstrong. Calibrating longevity stresses, 6 2013. URL `https://www.actuaries.org.uk/documents/rising-challenge-generating-longevity-stresses`

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

References II

S. Burgess, J. Kingdom, and J. Ayton. Longevity risk: Issues to be considered, 2 2010. URL `www.actuaries.org.uk/system/files/documents/pdf/burgesskingdomayton.pdf`.

D. R. Cox and D. V. Hinkley. *Theoretical Statistics*. Chapman and Hall, 1996. ISBN 0-412-16160-5.

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

References III

A. S. Macdonald, S. J. Richards, and I. D. Currie. *Modelling Mortality with Actuarial Applications*. Cambridge University Press, Cambridge, 2018. ISBN 978-1-107-04541-5.

S. Makin. Response to discussion paper on actuarial mortality assumptions, 2008.

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

S. J. Richards. Mis-estimation risk: measurement and impact. *British Actuarial Journal*, 21 (3):429–457, 2016. doi: 10.1017/S1357321716000040.

References IV

S. J. Richards. Modelling mortality by continuous benefit amount. *Longevitas Ltd*, 2020a.

S. J. Richards. A Hermite-spline model of post-retirement mortality. *Scandinavian Actuarial Journal*, 2020:2:110–127, 2020b. doi: 10.1080/03461238.2019.1642239.

S. J. Richards. A value-at-risk approach to mis-estimation risk. *Longevitas Ltd*, 2020c.

References V

S. Woods. Reflections on the 2015 Solvency II internal model approval process. *Bank of England Prudential Regulation Authority*, 2016. URL `http://www.bankofengland.co.uk/pra/Pages/publications/ps/2016/ps3316.aspx`.

More on longevity risk at www.longevitas.co.uk

9 Legal matters

Longevitas is a registered trademark of Longevitas Ltd:

- in the UK (No. 2434941),
- in the USA (No. 3707314), and
- in the European Union (No. 5854518).