

Institute and Faculty of Actuaries

A VaR approach to mis-estimation risk

Stephen J. Richards

10th May 2021, 08:30hrs



Institute
and Faculty
of Actuaries

Copyright © Longevitas Ltd. All rights reserved. This presentation may be freely distributed, provided it is unaltered and has this copyright notice intact.

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



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

Woods [2016]

“[...] 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]

“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]

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

2 Portfolio features

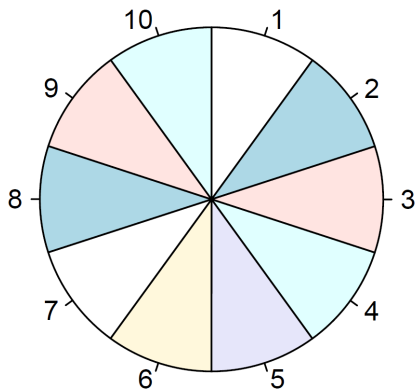


- Medium-sized pension scheme: 15,698 lives.
- Pensioners only.
- 5,956 deaths in 2007–2012.

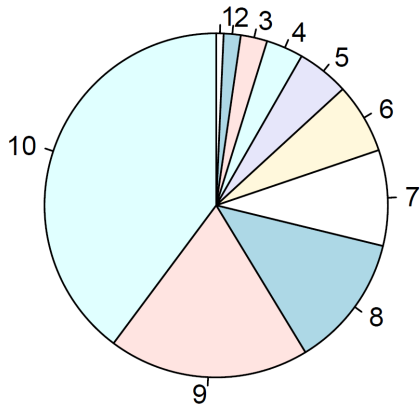
2 Liability concentration



Lives



Pensions



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

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

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.

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.

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].

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

Measures relative uncertainty over parameter estimate.

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].

- Liabilities highly concentrated.
- Sub-groups with most liability have lowest mortality...
... *and* high relative uncertainty.

3 Parameter estimation



In a statistical model with m parameters:

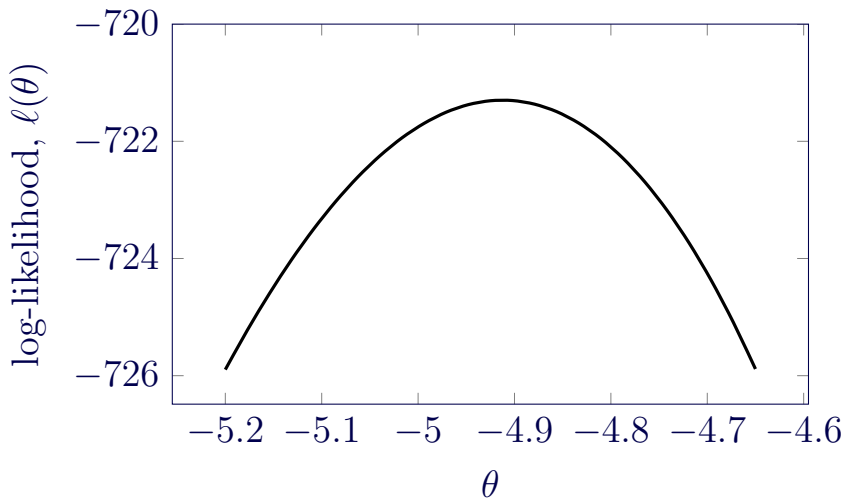
- Consider a parameter vector, $\underline{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_m \end{pmatrix}$.
- We have an estimate, $\hat{\underline{\theta}}$, which is uncertain.
- Uncertainty over $\hat{\underline{\theta}}$ is estimation risk.

3 Assumption 1



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

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



Source: Richards [2016, Figure 1].

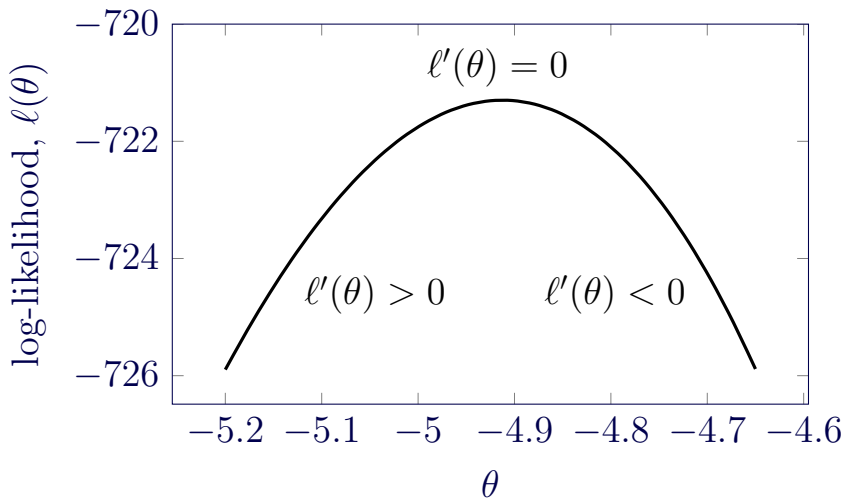
3 Assumption 2



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

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

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



Source: Richards [2016, Figure 1].

3 Assumption 3



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

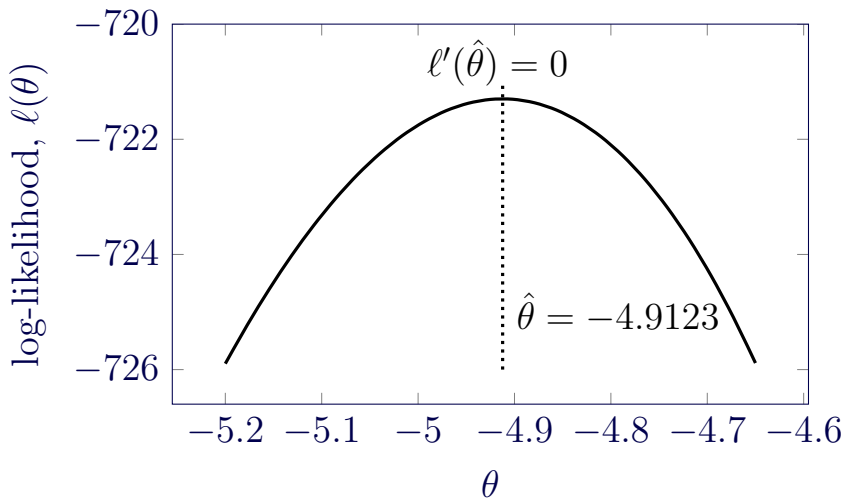
$$\mathbf{H}(\underline{\boldsymbol{\theta}}) = \begin{pmatrix} \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_1^2} & \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 \ell(\underline{\boldsymbol{\theta}})}{\partial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

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

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

[†] $\underline{\mathbf{x}}^T \mathbf{H}(\hat{\underline{\theta}}) \underline{\mathbf{x}} \leq 0, \forall \underline{\mathbf{x}} \in \mathbb{R}^m$, where $\underline{\mathbf{x}}^T$ denotes the transpose of $\underline{\mathbf{x}}$.

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



Source: Richards [2016, Figure 1].

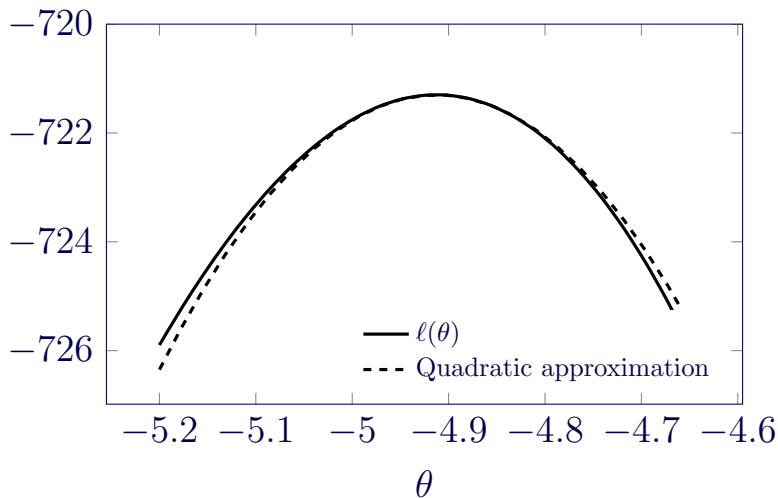


$\hat{\boldsymbol{\theta}}$ has an approximate multivariate normal (MVN) distribution[†]:

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

[†] Cox and Hinkley [1996, Chapter 9].

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



Source: Richards [2016, Figure 1].

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

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

3 (Mis-)Estimation risk



Estimation risk

Statistical parameter uncertainty.

Mis-estimation risk

Financial impact of parameter uncertainty.

Estimation risk

Uncertainty over $\hat{\theta}$.

Mis-estimation risk

Uncertainty over $V(\hat{\theta})$, where V is the liability function depending on $\hat{\theta}$.

4 Preconditions



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

- 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].

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



What is the uncertainty over $V(\hat{\underline{\theta}})$ caused by the uncertainty over $\underline{\hat{\theta}}$?



Two approaches to uncertainty over $V(\hat{\underline{\theta}})$:

- Delta method.
- Richards [2016].



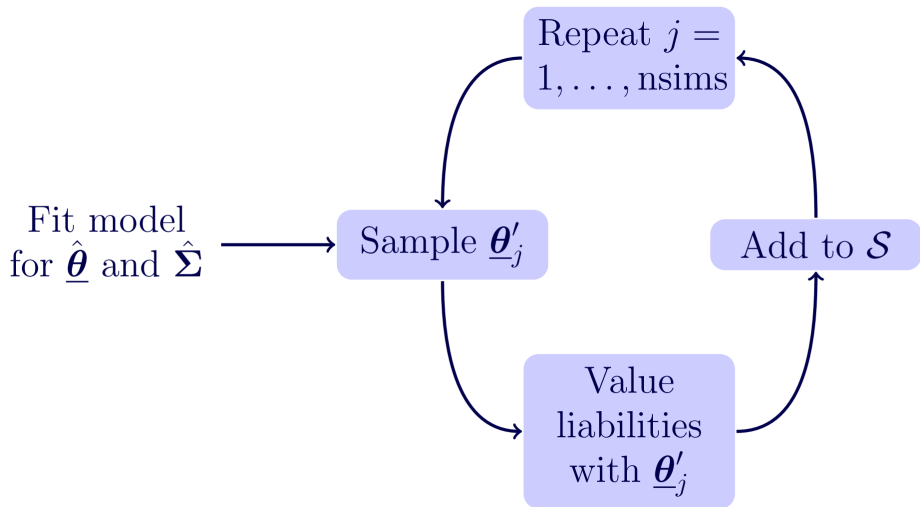
- $\hat{\theta}$ is a random variable.
- $V(\hat{\theta})$ is a function of a random variable.
- $\text{Var}[V(\hat{\theta})] \approx \underline{\mathbf{a}}^T \hat{\Sigma} \underline{\mathbf{a}}$, where $\underline{\mathbf{a}} = \frac{\partial V(\hat{\theta})}{\partial \hat{\theta}}$.



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

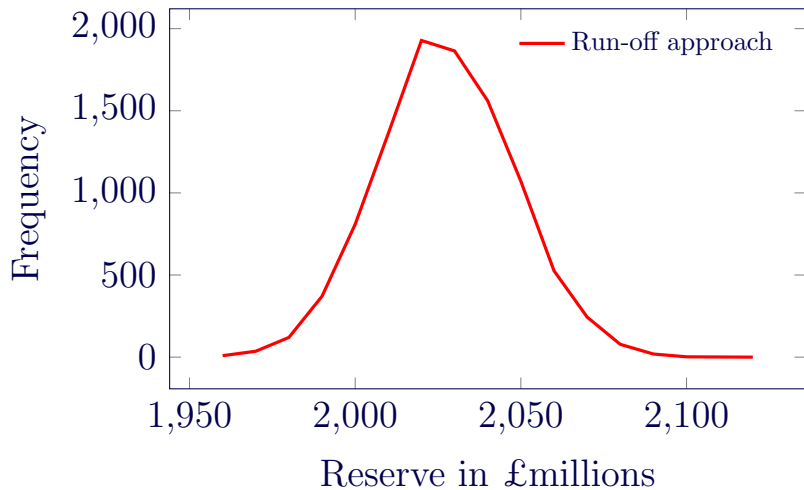


- Best-estimate parameter vector $\underline{\hat{\theta}}$.
- $\hat{\Sigma}$ is estimated covariance matrix for $\underline{\hat{\theta}}$.
- Alternative parameter vector, $\underline{\theta}'$, can be sampled from $MVN(\underline{\hat{\theta}}, \hat{\Sigma})$ using Monte Carlo simulation.
- Liability value $V(\underline{\theta}')$ added to set \mathcal{S} .



- Larger pension scheme: 44,616 lives.
- Pensioners only.
- 10,663 deaths in 2001–2009.

5 Distribution of \mathcal{S}



Source: Richards [2020c, Figure 4].



What is the relative capital percentage, RCP , required to cover a proportion p of mis-estimation risk?

$$RCP = \Phi^{-1}(p) \frac{\sqrt{\underline{\mathbf{a}}^T \hat{\underline{\Sigma}} \underline{\mathbf{a}}}}{V(\hat{\underline{\theta}})} \times 100\%$$

where $\Phi^{-1}(p)$ is the inverse of the normal distribution function.

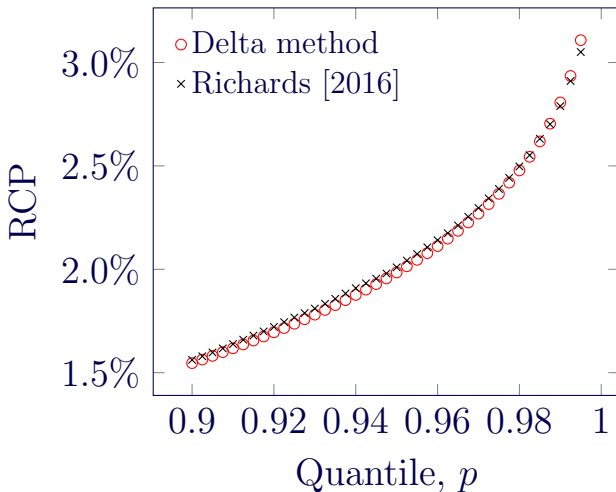
Note that $\frac{\sqrt{\underline{\mathbf{a}}^T \hat{\underline{\Sigma}} \underline{\mathbf{a}}}}{V(\hat{\underline{\theta}})}$ is the coefficient of variation.



$$RCP = \left[\frac{Q_p(\mathcal{S})}{V(\hat{\theta})} - 1 \right] \times 100\%$$

where $Q_p(\mathcal{S})$ is the p -quantile of \mathcal{S} .

5 Run-off mis-estimation capital



Source: Own calculations using model in Richards [2020c, Table 4].



- Both methods agree closely for this liability, even far into the upper tail.
- Delta method is quicker.
- Richards [2016] is better for skewed sets \mathcal{S} .

6 Value-at-risk mis-estimation





Run-off mis-estimation

What is the uncertainty over $V(\hat{\underline{\theta}})$ caused by the uncertainty over $\underline{\theta}$?

VaR mis-estimation

What is the uncertainty over $V(\hat{\underline{\theta}})$ caused by the re-estimation of $\underline{\theta}$ with n years additional data?

The VaR question is about recalibration risk [Cairns, 2013].

The run-off approach:

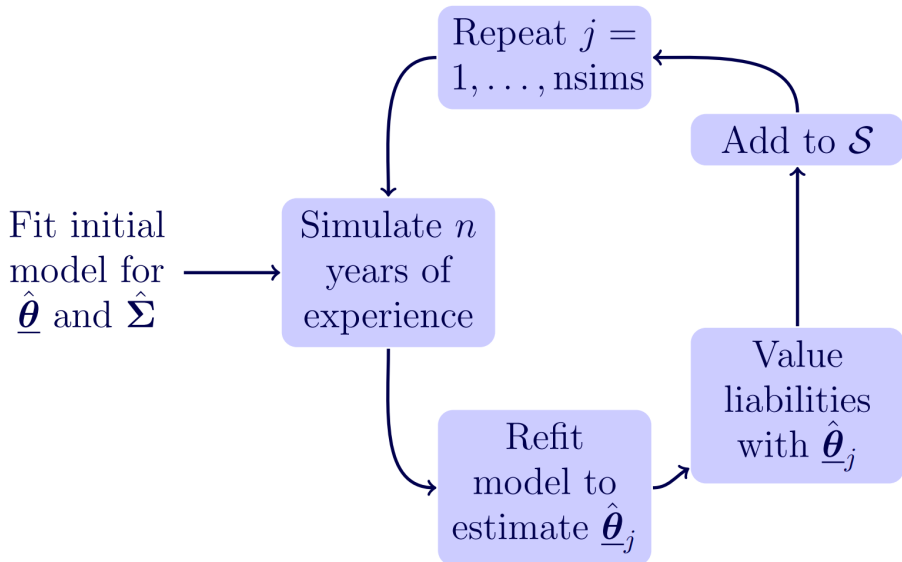
- Has no “new experience” element...
- Has no recalibration element...
- Is not a plug-in fit for VaR applications like Solvency II.



We make two changes to the algorithm of Richards [2016]:

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

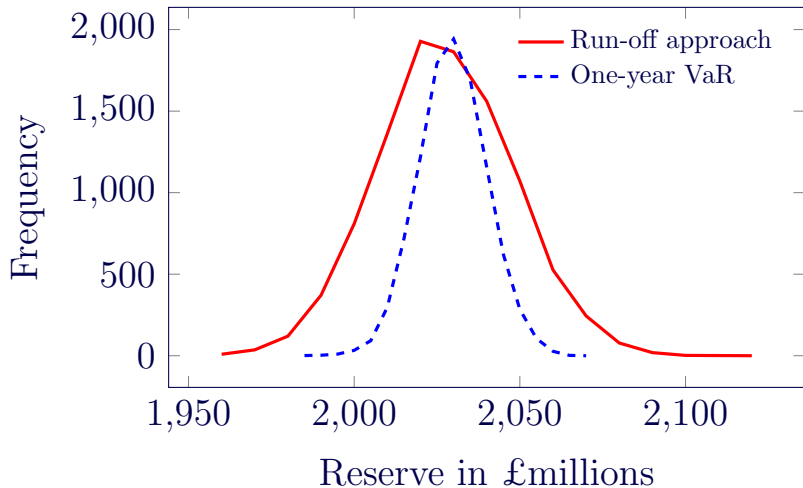
and proceed to generate set \mathcal{S} as before.





- $\hat{\theta}_j$ 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



- 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:

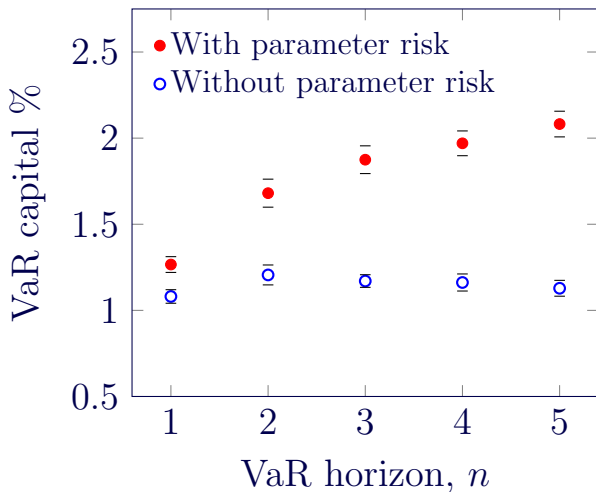
Without parameter risk

Use $\hat{\theta}$ each time.

With parameter risk

Sample from $MVN(\hat{\theta}, \hat{\Sigma})$.

7 VaR capital by horizon



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

For mis-estimation VaR capital:

- Including parameter risk 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 ($n = 3, 4, 5$).
- How correlated is VaR mis-estimation capital with idiosyncratic risk?

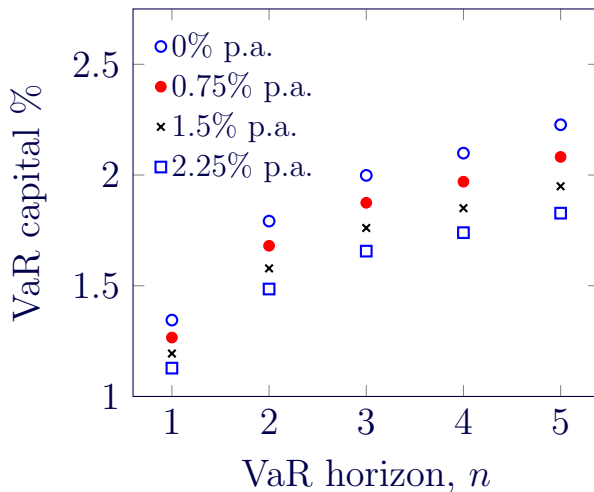
Correlations between the mis-estimation VaR reserves at $y_1 = 2010$ and three measures of the simulated pseudo-experience underlying the recalibration.

VaR horizon	Time lived	New deaths	Payments made
1	80.2%	-90.1%	76.5%
2	80.8%	-88.9%	80.0%
3	81.1%	-87.7%	82.5%
4	81.7%	-86.9%	84.8%
5	81.7%	-86.2%	86.4%



- VaR mis-estimation capital strongly correlated with idiosyncratic risk.
- Must be reflected in aggregation matrix for Solvency II.

7 Role of discount rate



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



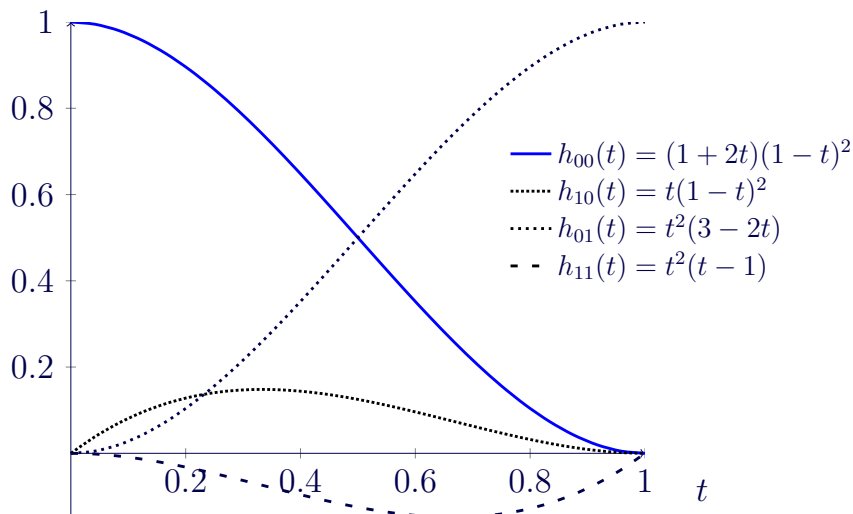
Lower discount rates mean higher mis-estimation VaR capital requirements.

Various options for post-retirement mortality:

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

- 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].

- 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].

- Mortality differentials in α reduce with age due to path of h_{00} .
 - don't need age interactions as in other models.
 - requires fewer parameters.

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

By lives

Use information criterion like AIC [Akaike, 1987].

By amounts

Use pension-weighted resampling approach like bootstrapping [Macdonald et al., 2018, Section 6.7].

7 What is the best fit?



By lives

Lowest AIC.

By amounts

Bootstrap percentage closest to 100%.

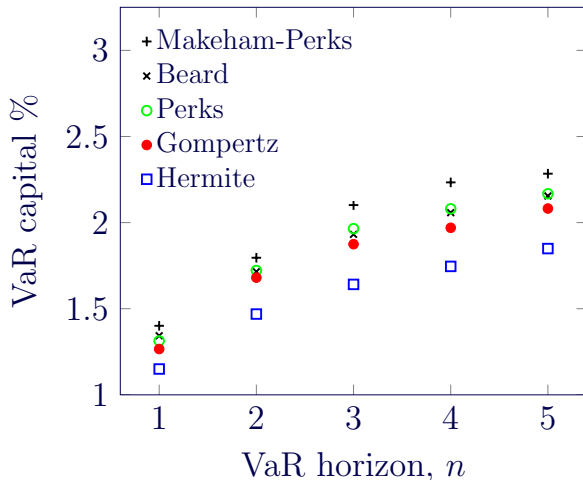
Sorted by descending AIC:

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].

- 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.

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

8 Conclusions





- 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 alternative views of mis-estimation:

Run-off

For pricing block transactions like reinsurance treaties, bulk annuities or longevity swaps.

Value-at-risk

For Solvency II reporting and ORSA.



For one-year VaR mis-estimation capital:

- Parameter risk is a surprisingly minor driver.
- Instead, recalibration risk dominates.
- Strong correlation with idiosyncratic risk.

Thanks to:

- Gavin Ritchie of Longevitas,
- Patrick Kelliher of Crystal Risk Consulting and
- Prof. Andrew Cairns of Heriot-Watt University

for comments on earlier drafts of this presentation.

Any errors or omissions remain the responsibility of the author.

- H. Akaike. Factor analysis and AIC. *Psychometrika*, 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.

- 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.
- A. J. G. Cairns. Robust hedging of longevity risk. *Journal of Risk and Insurance*, 80:621–648, 2013.
- D. R. Cox and D. V. Hinkley. *Theoretical Statistics*. Chapman and Hall, 1996. ISBN 0-412-16160-5.
- B. Gompertz. On the nature of the function expressive of the law of human mortality. *Philosophical Transactions of the Royal Society*, 115:513–585, 1825.

- 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.

- S. J. Richards. Modelling mortality by continuous benefit amount. *Longevity 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. *Longevity Ltd*, 2020c.

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/pru/Pages/publications/ps/2016/ps3316.aspx>.

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

Longevity is a registered trademark of Longevity Ltd:

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