Prospective modelling of biometric laws in a Long-Term Care Insurance portfolio

Guillaume Biessy

SCOR

12th of September 2019

SCOR
The Art & Science of Risk

Stakes in LTC prospective modelling

- Long-Term Care is a complex biometric risk that arises alongside longevity due to the ageing population
- Most insurance products offer lifetime guarantee in exchange for a lifetime premium.
- For France average age at subscribing/claim is 60/85 meaning a **25 years duration on liabilities**
- In case of deviation in the risk, premium increases, when allowed, must be applied early to
  - Have bigger impact
  - Avoid reputation risk
  - Not trigger mass lapse events among younger policyholders

Under those conditions, it makes sense to look for a **prospective modelling approach**.

First obstacle comes from lack of data:

- Short history (first French LTC product launched in 1985)
- Low volumes compared to national population
- Data aggregation usually not possible due to different definitions

SCOR
The Art & Science of Risk

LTC modelling

- LTC may be modelled using the so-called *illness-death model*.
- In the French definition, LTC must be permanent in order to claim benefit
- We rely on portfolio data with $\sim$ 200,000 insured and $\sim$ 20,000 claims

Notations used:

- x: age
- y: calendar year
- z: time spent in LTC
- g: gender
- $\lambda^{01}(x, y, g)$: incidence in LTC
- $\lambda^{02}(x, y, g)$: autonomous mortality
- $\lambda^{12}(x, y, z, g)$: disabled mortality

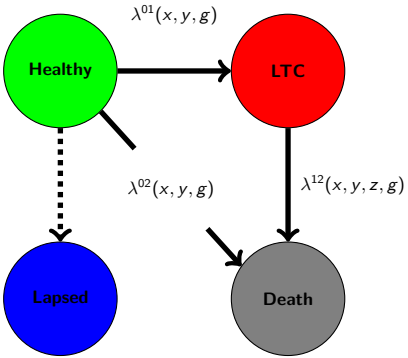


Figure 1: The Long-Term Care model with one level of care and no recoveries from LTC

Poisson GLM framework

We work on the random vector $Y = D$, the number of observed deaths (or entries in LTC when working on the incidence rate) and assume it follows a Poisson distribution of parameter λe^c where:

- λ is the transition intensity
- e^c the central exposure

The canonical link for the Poisson distribution is the log link.
Therefore our model specification is given by

$$\log \mathbb{E}(D) = \log \mu = \eta = \log \lambda + \log e^c$$

If we further note $\log \lambda = X\theta + \xi$ where ξ does not depend on θ then we can write

$$\log \mathbb{E}(D) = X\theta + \text{offset}$$

where $\text{offset} = \log e^c + \xi$.

This framework allows us to specify and fit a wide range of models from parametric models to smoothing in one or several dimensions, as well as prospective models such as the Lee-Carter model.

SCOR The Art & Science of Risk

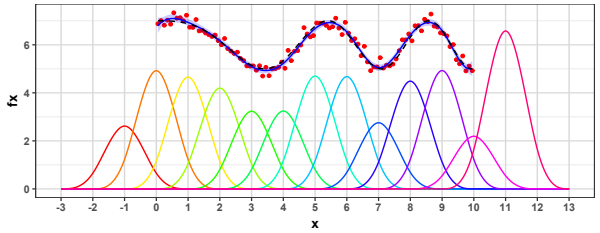
P-splines smoothing with cubic splines (Eilers & Marx, 1996)

- To build the fitted curve we determine the multiplier to apply to each spline
- Inference relies on maximum likelihood
- The more spline, the more accurate with the risk of overfitting the data
- We add to the log-likelihood a term to penalize huge variations in the curve:

$$\ell_p(\theta) = \ell(\theta) + \frac{1}{2}\theta' P \theta$$

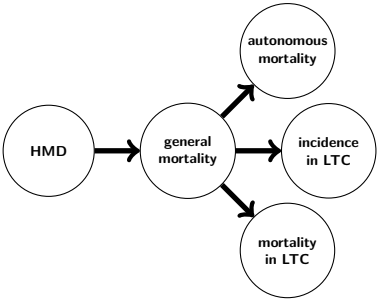
Metaparameters for the method:

- Degree of the splines: 3
- Number of splines: high
- Order of penalty: we may penalize non-zero coefficients (order 0), difference between consecutive values (order 1) or non-alignment between 3 consecutives values (2)
- Scale of the penalty: determined by minimizing a criterion such as the BIC



Inference process and constraints

- To ensure consistency between biometric laws we set a number of constraints listed in the table below.
- This sets an order in which the biometric laws must be derived, as on the right figure.



biometric law	reference	constraint
HMD	x	x
general mortality	HMD	$<$ HMD
autonomous mortality	general mortality	$<$ general mortality
incidence in LTC	general mortality	x
mortality in LTC	general mortality	$>$ general mortality

1D smoothing

- In the case of 1D smoothing the model is summarized by $\log \lambda_x = \overline{\alpha}_x$
- It may be expressed as a Poisson GLM: $\log \mathbb{E}(D) = X\theta + \text{offset}$ where $D = D_x$, $X = B_x$, $\theta = \theta_x$ and $\text{offset} = \log e_x^c$ with the penalty $P = P_x$.
- A mortality reference may be included in the model as an additional offset.

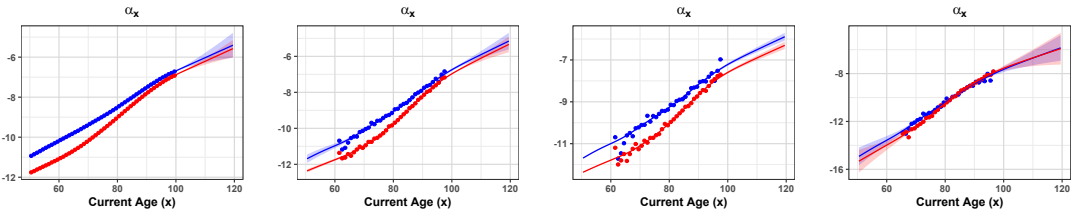


Figure 2: From left to right: HMD, general mortality, autonomous mortality, incidence in LTC

Age-Period (AP) model

The Age-Period model adds period coefficients $\log \lambda_{x,y} = \overline{\alpha}_x + \kappa_y$ with $\sum \kappa_y = 0$ corresponding to the Poisson GLM: $\log \mathbb{E}(D) = X\theta + \text{offset}$ where $D = D_{x,y}$, $X = (\mathbf{1}_y \otimes B_x : \mathbf{I}_y \otimes \mathbf{1}_x)$, $\theta = (\theta_x^{'} : \kappa_y^{'})^{'}$ and $\text{offset} = \log e_{x,y}^c$

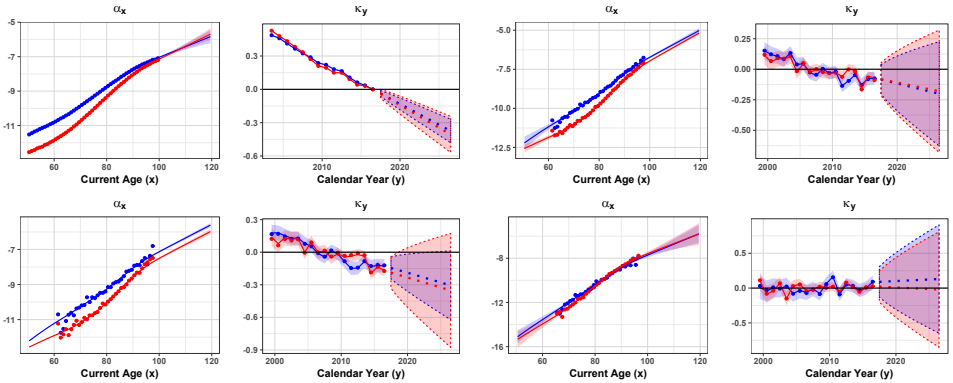


Figure 3: From left to right, top to bottom: HMD, general mortality, autonomous mortality, incidence in LTC

Drift inference

To keep things simple we assume κ_y to follow a random walk with drift.

$$\kappa_y = \mu + \kappa_{y-1} + \epsilon_y \quad \text{or} \quad \Delta\kappa_y = \mu + \epsilon_y \quad \text{where} \quad \epsilon_y \sim \mathcal{N}(0, \sigma^2)$$

The parameters μ and σ^2 called drift parameter and innovation variance may be estimated from the data

$$\hat{\mu} = \frac{\sum \Delta\kappa_y}{n_y - 1} = \frac{\kappa_{n_y} - \kappa_1}{n_y - 1} \quad \text{and} \quad \hat{\sigma}^2 = \frac{\sum (\Delta\kappa_y - \hat{\mu})^2}{n_y - 2}$$

with variance

$$\text{var}(\hat{\mu}) \simeq \frac{\hat{\sigma}^2}{n_y - 1} + \underbrace{\hat{\Psi}^{\kappa}_{1,1} + \hat{\Psi}^{\kappa}_{n_y,n_y} - 2\hat{\Psi}^{\kappa}_{1,n_y}}_{\text{estimation error on the } \kappa_y}$$

where $\hat{\Psi}^{\kappa}$ is the diagonal block of $\hat{\Psi}$ which contains the covariance of the κ_y coefficients.
In practice the estimation error may be neglected even on portfolio data.

Forecasting

Forecasting future values of κ_{n_y+m} for $m \geq 1$ may be performed easily :

- The expectancy of forecasted value is $\mathbb{E}(\kappa_{n_y+m}) = \kappa_{n_y} + m\hat{\mu}$.
- The associated variance is $\text{var}(\kappa_{n_y+m}) \simeq \text{var}(\kappa_{n_y}) + m^2\text{var}(\hat{\mu}) + 2m\ \text{covar}(\hat{\mu}, \kappa_{n_y}) + m\hat{\sigma}^2$ with $\text{var}(\kappa_{n_y}) = \hat{\Psi}^{\kappa}_{n_y,n_y}$ and $\text{covar}(\hat{\mu}, \kappa_{n_y}) = \frac{1}{n_y-1}\left[\hat{\Psi}^{\kappa}_{n_y,n_y} - \hat{\Psi}^{\kappa}_{1,n_y}\right]$. It is the sum of :
  - A parameter error which reflect the uncertainty on the value of the drift due to the presence of the random noise ϵ_y and uncertainty on the value of the κ_y estimates
  - A stochastic error $m\hat{\sigma}^2$ which reflect the presence of random noise ϵ_y in future observations

  The relative weight of the parameter error increases when looking at observations that are farther away.

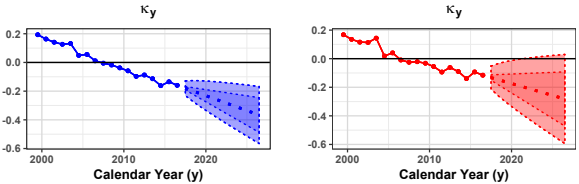


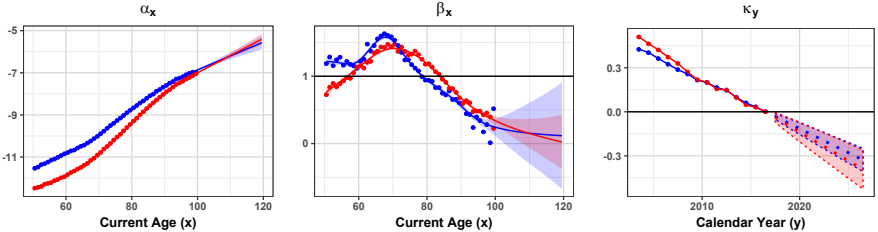
Figure 4: Forecasted κ_y for HMD mortality. Prediction intervals for parameter error, parameter & stochastic error

Lee-Carter model: HMD data for France

The Lee-Carter model improves the Age-Period model, modulating the period effect by age through an additional set of parameters $\overline{\beta}_x$ such that $\log \lambda_{x,y} = \overline{\alpha}_x + \overline{\beta}_x \kappa_y$ where $\sum \kappa_y = 0$ and $\sum \overline{\beta}_x = n_x$.
As the $\overline{\beta}_x \kappa_y$ term is non-linear, the model can no longer be expressed as a single GLM. However, it is possible to alternatively iterate the scoring algorithms for two GLMs.

$$\text{GLM 1} : \log \mathbb{E}(D_{x,y}) = \underbrace{[\kappa_y \otimes B_x]}_{X_1} \theta_x^1 \quad + \quad \underbrace{\mathbf{1}_y \otimes \overline{\alpha}_x + \log e_{x,y}^c}_{\text{offset for GLM 1}}$$

$$\text{GLM 2} : \log \mathbb{E}(D_{x,y}) = \underbrace{\left[\mathbf{1}_y \otimes B_x : \mathbf{I}_y \otimes \overline{\beta}_x\right]}_{X_2} (\theta_x^{2'} : \kappa_y')' \quad + \quad \underbrace{\log e_{x,y}^c}_{\text{offset for GLM 2}}$$



The case of mortality in LTC

- We could use the Age-Period model for mortality in LTC as well: $\log \lambda_{x,y} = \alpha_x^{\text{gen}} + \overline{\alpha}_x + \kappa_y$
- However, analysis of residuals on a grid with Age and Duration in LTC shows a strong impact of duration in LTC, not captured by the model...
- Higher mortality during the first few months in LTC may indeed be linked to a mixture effect caused by the heterogeneity of pathologies that lead to LTC (see for example Biessy, 2017)

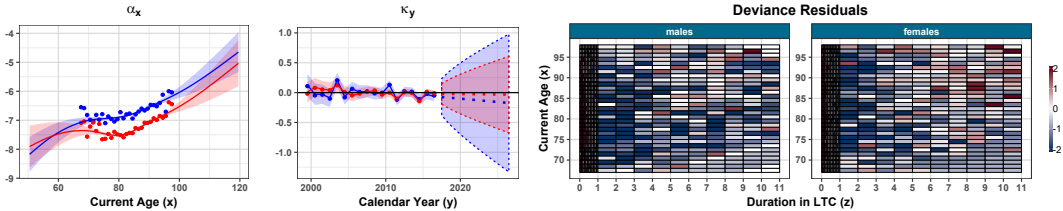


Figure 5: Left: fitted coefficients; Right: residuals on an age x duration in LTC grid

2D smoothing

- To account for the role of duration in LTC we use a 2D smoothing model: $\log \lambda_{x,z} = \overline{A}_{x,z}$.
- It may be expressed as a Poisson GLM: $\log \mathbb{E}(D) = X\theta + \text{offset}$ where $D = D_{x,z}$, $X = B_z \otimes B_x$, $\theta = \theta_{x,z}$ and $\text{offset} = \log e^c_{x,z}$
- The penalty in this case is the sum of a row penalty and a column penalty: $P = P_x \otimes \mathbf{I}_{\text{nb}_z} + \mathbf{I}_{\text{nb}_x} \otimes P_z$

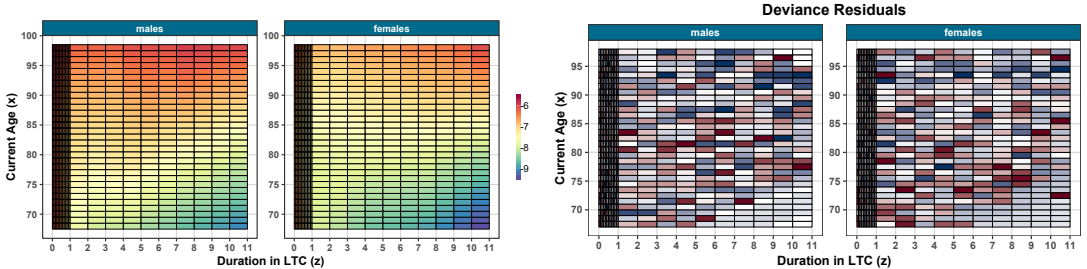


Figure 6: Left: fitted values; Right: residuals

Age x Duration smoothing with Period effect (AxDP) model

- We combine smoothing over the age and duration in LTC dimensions with a period effect.
- The resulting model is $\log \lambda_{x,y,z} = \overline{A}_{x,z} + \kappa_y$ with $\sum \kappa_y = 0$.
- It may be expressed as a Poisson GLM: $\log \mathbb{E}(D) = X\theta + \text{offset}$ where $D = D_{x,y,z}$, $X = (B_z \otimes \mathbf{1}_y \otimes B_x : \mathbf{1}_z \otimes \mathbf{l}_y \otimes \mathbf{1}_x)$, $\theta = (\theta'_{x,z} : \kappa'_y)'$ and $\text{offset} = \log e^c_{x,y,z}$

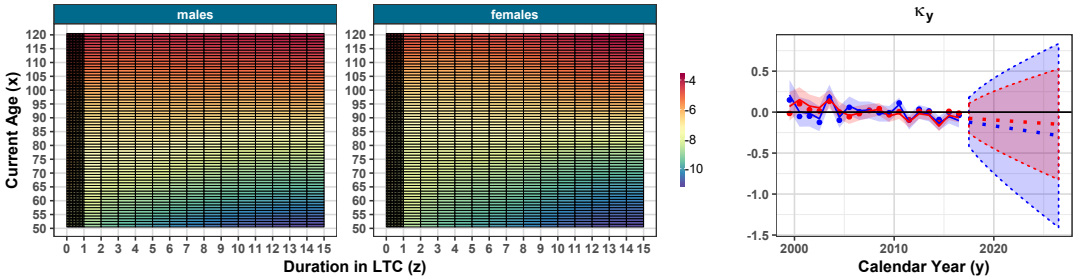


Figure 7: Left: fitted age x duration in LTC coefficients; Right: fitted period coefficients

Age Period Duration (APD) model

- We introduce the Age-Period-Duration model: $\log \lambda_{x,y,z} = \overline{\alpha}_x + \kappa_y + \overline{\gamma}_z$ with $\sum \kappa_y = 0$ and $\sum \overline{\gamma}_z = 0$
- It may be expressed as a Poisson GLM: $\log \mathbb{E}(D) = X\theta + \text{offset}$ where $D = D_{x,y,z}$, $X = (\mathbf{1}_z \otimes \mathbf{1}_y \otimes B_x : \mathbf{1}_z \otimes \mathbf{I}_y \otimes \mathbf{1}_x : B_z \otimes \mathbf{1}_y \otimes \mathbf{1}_x)$, $\theta = (\theta_x^{'} : \kappa_y^{'}\theta_z^{'})^{'}$ and $\text{offset} = \log e_{x,y,z}^{c}$

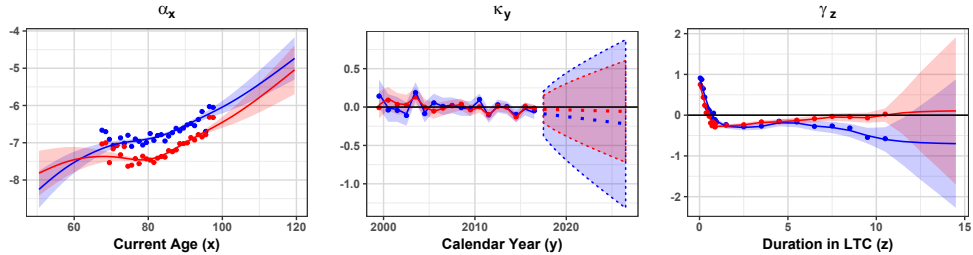


Figure 8: Mortality in LTC: age (left), period (middle) and duration in LTC (right) coefficients

Augmented Age-Period-Duration (AAPD) model

It may be further refined into an Augmented Age-Period-Duration model: $\log \lambda_{x,y,z} = \overline{\alpha}_x + \kappa_y + \overline{\beta}_x \overline{\gamma}_z$.
The $\overline{\beta}_x \overline{\gamma}_z$ term is non-linear as in the Lee-Carter model and must be fitted using 2 GLMs.

$$\text{GLM 1: } \log \mathbb{E}(D_{x,y,z}) = \underbrace{[\overline{\gamma}_z \otimes \mathbf{1}_y \otimes B_x]\theta_x^2}_{X_1} + \underbrace{\mathbf{1}_z \otimes \mathbf{I}_y \otimes \overline{\alpha}_x + \log e^c_{x,y,z}}_{\text{offset for GLM 1}}$$

$$\text{GLM 2: } \log \mathbb{E}(D_{x,y,z}) = \underbrace{\left[\mathbf{1}_z \otimes \mathbf{1}_y \otimes B_x : \mathbf{1}_z \otimes \mathbf{I}_y \otimes \mathbf{1}_x : B_z \otimes \mathbf{1}_y \otimes \overline{\beta}_x\right](\theta_x^{1'} : \kappa_y' : \theta_z')'}_{X_2} + \underbrace{\log e^c_{x,y,z}}_{\text{offset for GLM 2}}$$

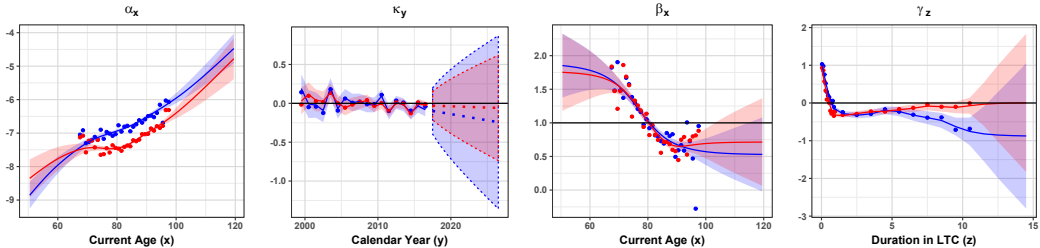


Figure 9: From left to right age , period, age-modulating and duration in LTC coefficients

Comparison of models for mortality in LTC

Model	Gender	deviance	df	BIC
AP	male	8581.37	20.66	8768.35
AxDP	male	7361.55	34.79	7676.43
APD	male	7437.78	27.46	7686.36
AAPD	male	7335.34	29.69	7604.15
AP	female	9330.57	21.59	9528.61
AxDP	female	8245.59	42.25	8633.17
APD	female	8326.49	30.54	8606.58
AAPD	female	8225.03	32.46	8522.76

Table 1: Comparison of models for mortality in LTC

- AP and APD models are doing poorly
- AxDP model is purely non-parametric while AAPD is semi-parametric
- AAPD is slightly more accurate than AxDP and also more parsimonious
- AAPD coefficients may be easily interpreted
- AAPD allows for comparison with other biometric laws

Comparison of model drifts

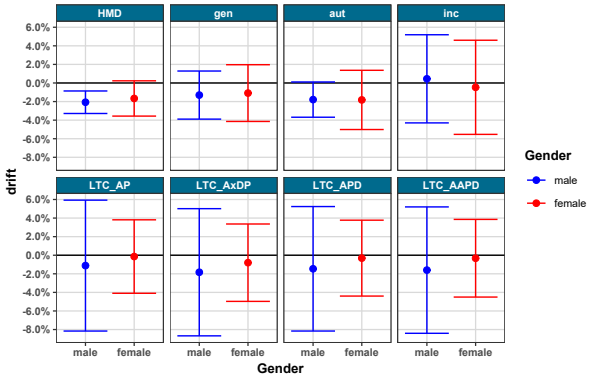


Figure 10: Model drifts with 95 % confidence intervals

law	male	female
HMD	-2.1 %	-1.7 %
gen	-1.3 %	-1.1 %
aut	-1.8 %	-1.8 %
inc	0.4 %	-0.5 %
LTC_AP	-1.1 %	-0.1 %
LTC_AxDP	-1.8 %	-0.8 %
LTC_APD	-1.5 %	-0.3 %
LTC_AAPD	-1.6 %	-0.3 %

Table 2: Best estimate model drifts

- The estimated trends are very volatile. . .
- Extreme scenarios must be considered in addition to best estimate scenario

Conclusion

Summary:

- Prospective modelling must be done on each of the 3 biometric laws involved in LTC modelling
- Simple model inspired from longevity modelling may be used
- For mortality in LTC, the duration in LTC dimension must be included as well
- Two approaches have been proposed a non-parametric approach with 2D smoothing and a robust and interpretable semi-parametric approach inspired from the Lee-Carter model

Future development:

- Impact of trends on pricing and reserving to be quantified
- More complex time series models could be used (ARIMA)
- Alternatively a mixed model approach could replace the 2-step inference / forecast procedure
- Correlation between the different trends should be considered to create more realistic scenarios

References

- Lee, R. D., & Carter, L. R. (1992). Modelling and forecasting US mortality. Journal of the American statistical association, 87(419), 659-671.
- Eilers, P. H., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. Statistical science, 89-102.
- Currie, I. D. & all (2006). Generalized linear array models with applications to multidimensional smoothing. J. R. Statist. Soc. B 68(2), 259-280.
- Currie, I. D. (2013). Smoothing constrained generalized linear models with an application to the Lee-Carter model. Statistical Modelling, 13(1), 69-93.
- Biessy, G. (2017). Continuous-time semi-Markov inference of biometric laws associated with a long-term care insurance portfolio. ASTIN Bulletin: The Journal of the IAA, 47(2), 527-561.
- Biessy, G. (2019a). Smoothing of multidimensional biometric laws in a Long-Term Care Insurance portfolio. Working Paper.
- Biessy, G. (2019b). Prospective modelling of biometric laws in a Long-Term Care Insurance portfolio. Working Paper.

SCOR The Art & Science of Risk