Individual loss reserving using activation patterns

Marie Michaelides

Avril 2022

Contents

Contents

Introduction

- Increased complexity of the insurers' databases: granular data, multi-level data,...
- Insurance world = competitive world: diversification of the insurance products and coverages offered to the policyholders
- Goal: build a granular reserving model that takes into account the dependence between multiple insurance coverages.
- Use of copulas for dependency modelling between coverages: Shi et al. (2016) [8], Yang et al. (2001) [9].
- Extension of the hierarchical models of Frees et Valdez (2008) [5], Frees et al. (2009) [6] and Côté et al. (2021) [3].

Contents

General framework

- Let c denote an insurance coverage with $c = 1, ..., C$.
- Goal: model the activation of the C coverages in development year j.
- Let $\boldsymbol{v}_{i,j}^{v} \in \mathcal{V}$ be the v^{th} possible vector for the activation of the coverages for claim i in year j, such that:

$$\boldsymbol{v}_{i,j}^{v} = \begin{bmatrix} v_{i,j,1}^{v} & v_{i,j,2}^{v} & \dots & v_{i,j,C}^{v} \end{bmatrix}.$$

  with

$$v_{i,j,c}^{v} = \begin{cases} 1, & \text{if coverage } c \text{ is activated for claim } i \text{ in year } j \\ 0, & \text{otherwise.} \end{cases}$$

- $\mathcal{V}$ is the set of the V possible coverage activation vectors with $V = 2^C - 1$.

General framework - Example

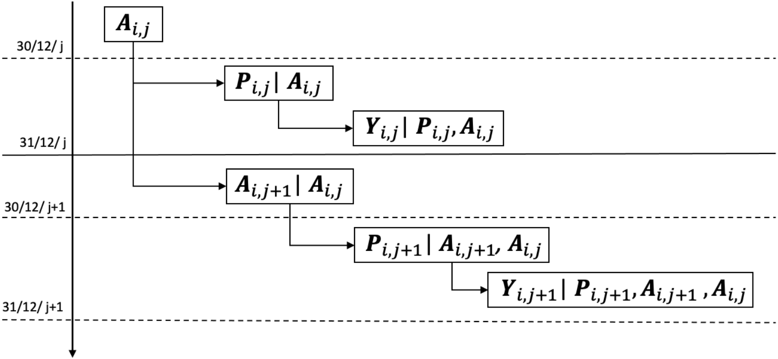
- $C = 2 \rightarrow$V=$2^2 - 1 = 3$
- $\boldsymbol{v}^v_{i,j}$ with $v = 1, 2, 3$
- $\boldsymbol{v}^v_{i,j} = \begin{bmatrix} v^v_{i,j,1} & v^v_{i,j,2} \end{bmatrix}$
- $\mathcal{V} = \{\boldsymbol{v}^1_{i,j} \;\; \boldsymbol{v}^2_{i,j} \;\; \boldsymbol{v}^3_{i,j}\} = \{(0\;1)\;\;(1\;0)\;\;(1\;1)\}$

General framework - Hypothesis

Hypothesis:

1. Once activated, a coverage remains active until the settlement of the claim.
2. - Activation of the coverages at any time until the 30^{th} of December.
   - Payments are made on the 31^{st} of December.
3. The activation patterns remain constant for $j > 2$.

Model - Overview



Model - Development year j

a) **Activation patterns: $\boldsymbol{A}_{i,j}$**

- $\boldsymbol{A}_{i,j} = \left[A_{i,j,1} \ \ A_{i,j,2} \ \ \ldots \ \ A_{i,j,C} \right]$: activation pattern for claim i in year j.

$$A_{i,j,c} = \begin{cases} 1 & \text{if coverage } c \text{ is activated for claim } i \text{ in } j \\ 0 & \text{otherwise.} \end{cases}$$

- Multinomial logit regression:

$$\mathrm{P}[\boldsymbol{A}_{i,j} = \boldsymbol{v}^{v}_{i,j} | \boldsymbol{x}_i, \beta] = \frac{\exp\left(\boldsymbol{x}_i' \beta_{j,v}\right)}{\sum_{k=1}^{V} \exp\left(\boldsymbol{x}_i' \beta_{j,k}\right)}.$$

Model - Development year j

b) **Payment patterns:** $\boldsymbol{P}_{i,j}|\boldsymbol{A}_{i,j}$

- $\boldsymbol{P}_{i,j}|\boldsymbol{A}_{i,j}$ payment pattern for claim i:

$$P_{i,j,c}|A_{i,j,c} = 1 = \begin{cases} 1 & \text{if a payment is made in year } j \\ & \text{for claim } i \text{ and coverage } c \\ 0 & \text{otherwise.} \end{cases}$$

- Bernoulli regression:

$$P_{i,j,c}|A_{i,j,c} = 1 \sim \text{Bernoulli}\big(\pi_{j,c}(\boldsymbol{x}_i'\gamma_{j,c})\big)$$

$$\pi_{j,c}(\boldsymbol{x}_i'\gamma_{j,c}) = \frac{\exp\left(\boldsymbol{x}_i'\gamma_{j,c}\right)}{1 + \exp\left(\boldsymbol{x}_i'\gamma_{j,c}\right)}$$

c) **Claim severity:** $\boldsymbol{Y}_{i,j}|\boldsymbol{P}_{i,j}, \boldsymbol{A}_{i,j}$

- GAMLSS with link function g:

$$\text{E}[Y_{i,j,c}|P_{i,j,c} = 1, A_{i,j,c} = 1] = g^{-1}(\boldsymbol{x}_i'\boldsymbol{\alpha}_{j,c} + \alpha^*_{j,c}j),$$

Model - Development year $j+1$

a) **Activation patterns:** $\boldsymbol{A}_{i,j+1}|\boldsymbol{A}_{i,j}$

$\boldsymbol{A}_{i,j+1}|\boldsymbol{A}_{i,j} \in \mathcal{V}^* \subset \mathcal{V}$ avec:

$$\mathrm{P}[\boldsymbol{A}_{i,j+1} = \boldsymbol{v}_{i,j}^{v*}|\boldsymbol{A}_{i,j}, \boldsymbol{x_i}, \beta] = \begin{cases} \frac{\mathrm{P}[\boldsymbol{A}_{i,j+1}=\boldsymbol{v}_{i,j}^{v*}|\boldsymbol{x_i},\boldsymbol{\beta}_{j+1}]}{\sum_{\boldsymbol{v}_{i,j}^{v*}\in\mathcal{V}^*} \mathrm{P}[\boldsymbol{A}_{i,j+1}=\boldsymbol{v}_{i,j}^{v*}|\boldsymbol{x}_i,\boldsymbol{\beta}_{j+1}]}, & \text{if } \boldsymbol{v}_{i,j}^{v*} \in \mathcal{V}^* \\ 0, & \text{otherwise} \end{cases}$$

b) **Payment patterns:** $\boldsymbol{P}_{i,j+1}|\boldsymbol{A}_{i,j+1}, \boldsymbol{A}_{i,j}$

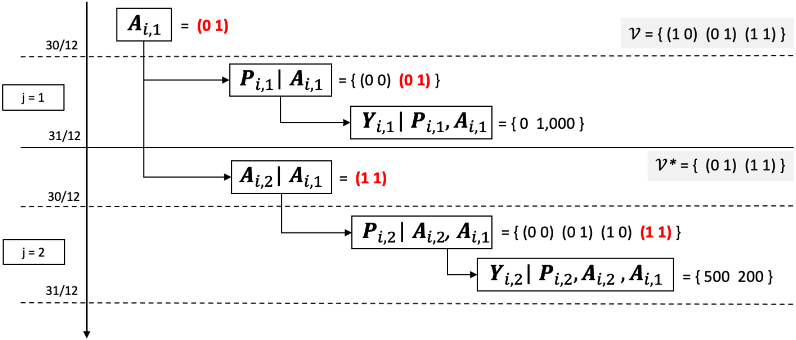
$P_{i,j+1,c}|A_{i,j+1} = 1, \boldsymbol{A}_{i,j} \sim \mathrm{Bernoulli}\left(\pi_{j+1,c}(\boldsymbol{x_i'}, \boldsymbol{\gamma_{j+1,c}})\right)$, with:

$$\pi_{j+1,c}(\boldsymbol{x}_i'\gamma_{j+1,c}) = \frac{\exp\left(\boldsymbol{x}_i'\gamma_{j+1,c}\right)}{1+\exp\left(\boldsymbol{x}_i'\gamma_{j+1,c}\right)}$$

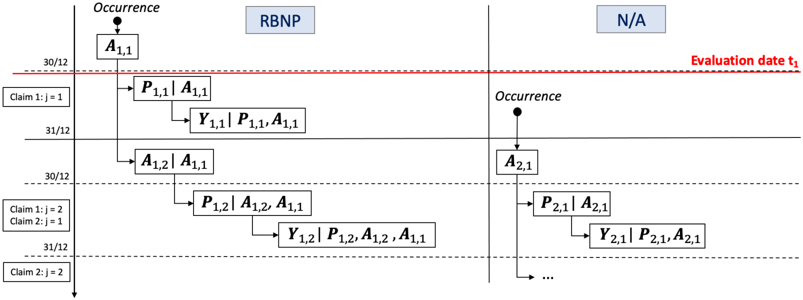
c) **Claim severity:** $\boldsymbol{Y}_{i,j+1}|\boldsymbol{P}_{i,j+1}, \boldsymbol{A}_{i,j+1}, \boldsymbol{A}_{i,j}$

$$\mathrm{E}[Y_{i,j+1,c}|P_{i,j+1,c} = 1, A_{i,j+1,c} = 1, \boldsymbol{A}_{i,j}] = g^{-1}(\boldsymbol{x}_i'\boldsymbol{\alpha}_{j+1,c} + \alpha_{j+1,c}^*(j+1))$$

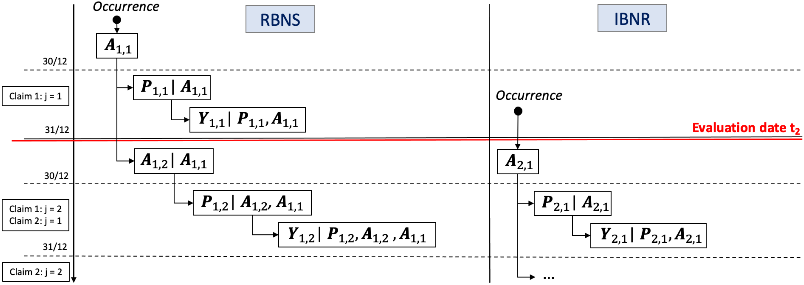
Model - Example with $C = 2$ insurance coverages



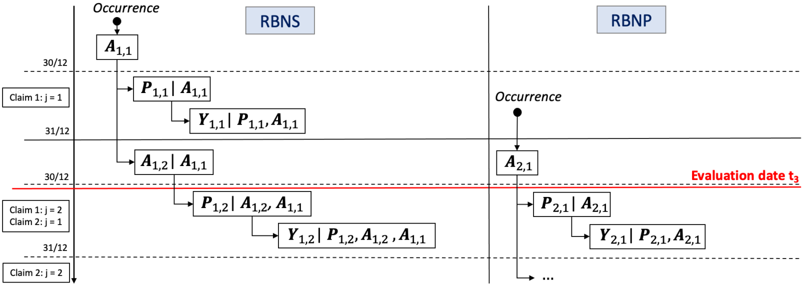
Treatment of different types of claim



Treatment of different types of claim



Treatment of different types of claim



Contents

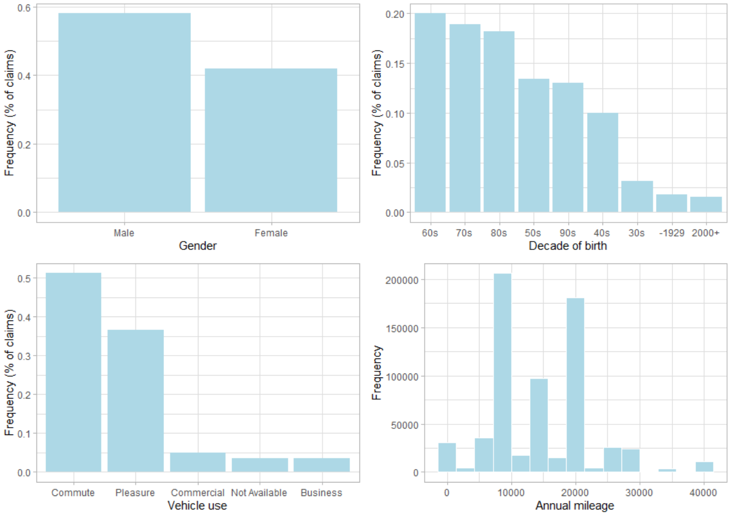
Data Exploration - Risk Factors

656,153 claims occurred between the 1st of January 2015 and the 30th of June 2021.

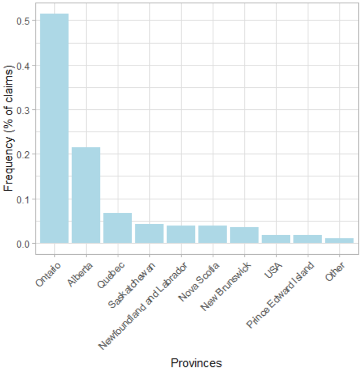
Risk factors	
GENDER	Gender of the insured.
YOB	Year of birth: decade during which the insured was born.
VU	Use of the vehicle made by the insured.
AM	Annual mileage driven by the insured (in km).
PROV	Canadian province (or USA) where the claim occurred.
FR	Fault rating: evaluation of the insured's level of responsability in the accident.

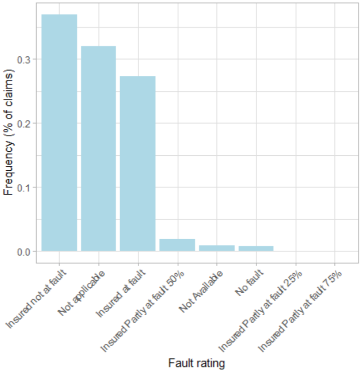
Table 1: Description of the risk factors provided for each claim in the dataset.

Data Exploration - Risk Factors



Data Exploration - Risk Factors





Data Exploration - Insurance coverages

Insurance coverages	
Accident Benefits	Compensation for loss of revenue, funeral expenses, medical expenses, death,...
Bodily Injury	Compensation for medical expenses or loss of revenue of a third party.
Vehicle Damage	Compensation for the damages incurred to the insured's or another party's vehicle.
Loss of Use	Compensation of costs in case of the temporary replacement of a vehicle or any other alternative transportation means used during vehicle repairs.

Table 2: Description of the four insurance coverages provided by the company.

Data Exploration - Insurance coverages

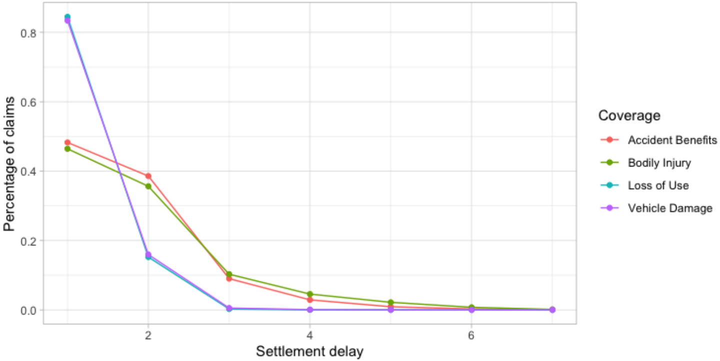
Coverage	% of claims	% of the total cost
Accident Benefits	9.42	12.82
Bodily Injury	5.70	13.13
Vehicle Damage	96.39	70.44
Loss of Use	51.89	3.61

Table 3: Weight of each coverage in the portfolio.

Coverage	μ	σ	Quantiles					
			0.5	0.75	0.9	0.95	0.99	1
Accident Benefits	12,386	53,561	3,215	6,909	27,472	47,757	127,896	2,435,334
Bodily Injury	23,271	76,027	4,000	15,150	53,314	98,449	322,612	2,039,570
Vehicle Damage	5,040	8,121	2,605	5,830	11,502	17,984	40,611	149,399
Loss of Use	545	620	419	714	1,000	1,260	2,336	52,777

Table 4: Descriptive statistics for the four insurance coverages.

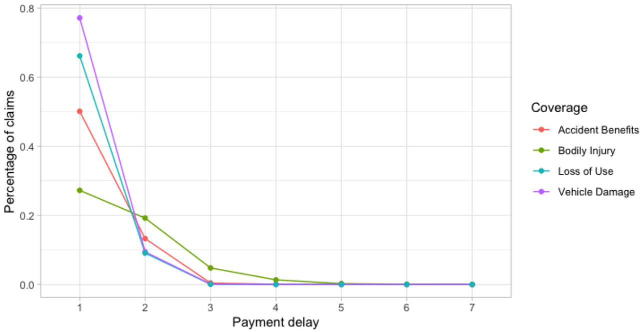
Data Exploration - Insurance coverages



Data Exploration - Insurance coverages

Coverage	Activation delays		
	No delay	1 year	$\geq$2 years
Accident Benefits	96.65	3.18	0.11
Bodily Injury	89.89	7.70	2.11
Vehicle Damage	98.20	1.73	0.07
Loss of Use	97.80	2.17	0.03

Table 5: Percentage of claims with different activation delays for the four coverages.



Estimation - Models

- Activation patterns: Multinomial logit model (1)
- Payment patterns: Bernoulli GLMs (8)
- Severities: GAMLSS (8)
- Maximum likelihood estimation

Estimation - Activation patterns

v	AB	BI	VD	LoU	Observed	$\overline{\mathrm{P}}[\boldsymbol{A}_{i,1} = \boldsymbol{v}_{i,1}^{v} \mid \boldsymbol{x}_i, \beta]$
1	0	0	0	1	0.0104	0.0177
2	0	0	1	0	0.4295	0.4541
3	0	0	1	1	0.4272	0.4144
4	0	1	0	0	0.0061	0.0052
5	0	1	0	1	0.0004	0.0005
6	0	1	1	0	0.0124	0.0110
7	0	1	1	1	0.0199	0.0171
8	1	0	0	0	0.0144	0.0128
9	1	0	0	1	0.0010	0.0014
10	1	0	1	0	0.0107	0.0106
11	1	0	1	1	0.0498	0.0408
12	1	1	0	0	0.0038	0.0029
13	1	1	0	1	0.0001	0.0001
14	1	1	1	0	0.0043	0.0035
15	1	1	1	1	0.0101	0.0078

Table 6: Average probabilities of observing each activation pattern obtained with the multinomial logit model.

Estimation - Payment patterns

Probability	$j = 1$	$j = 2+$
$\pi_{j,\text{AB}}$	0.5155	0.3578
$\pi_{j,\text{BI}}$	0.3484	0.2812
$\pi_{j,\text{VD}}$	0.8203	0.1138
$\pi_{j,\text{LoU}}$	0.7070	0.0720

Table 7: Average probabilities of having a payment in each time period for the different coverages.

Estimation - Severities

Coverage	Model	$j = 1$ AIC	BIC	$j = 2+$ AIC	BIC
Accident Benefits	Log-Normal	266,408	266,651	181,577	181,827
	Gamma	264,411	264,654	183,670	183,921
	Pareto	262,903	263,146	181,281	181,532
	Generalized Beta II	**261,110**	**261,368**	**181,222**	**181,487**
	Weibull	264,101	264,344	182,545	182,796
Bodily Injury	Log-Normal	100,523	100,735	106,985	107,214
	Gamma	99,675	99,887	107,096	107,325
	Pareto	**98,971**	**99,183**	106,691	106,919
	Generalized Beta II	99,966	100,185	106,884	107,119
	Weibull	99,187	99,400	**106,673**	**106,902**
Vehicle Damage	Log-Normal	5,375,907	5,376,246	620,136	620,431
	Gamma	5,385,180	5,385,519	619,456	619,751
	Pareto	5,342,022	5,342,361	617,247	**617,542**
	Generalized Beta II	**5,330,989**	**5,331,349**	**617,245**	617,557
	Weibull	5,371,787	5,372,126	618,531	618,827
Loss of Use	Log-Normal	1,874,185	1,874,498	193,705	193,967
	Gamma	1,868,867	1,869,180	192,929	193,191
	Pareto	1,894,927	1,895,241	194,376	194,638
	Generalized Beta II	**1,861,585**	**1,861,918**	**192,212**	**192,489**
	Weibull	1,877,508	1,877,821	193,459	193,721

Table 8: Choice of the distributions for the severity of the payments.

Results - Predictive distributions (RBNS)

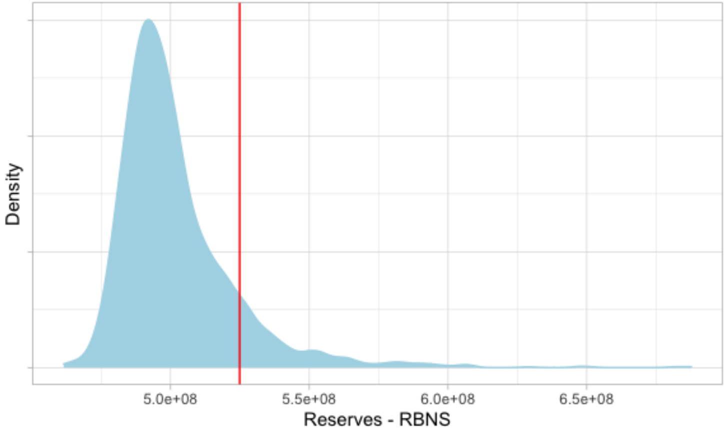


Figure 1: Evaluation date: 01/01/2019.

Results - Predictive distributions (RBNS)

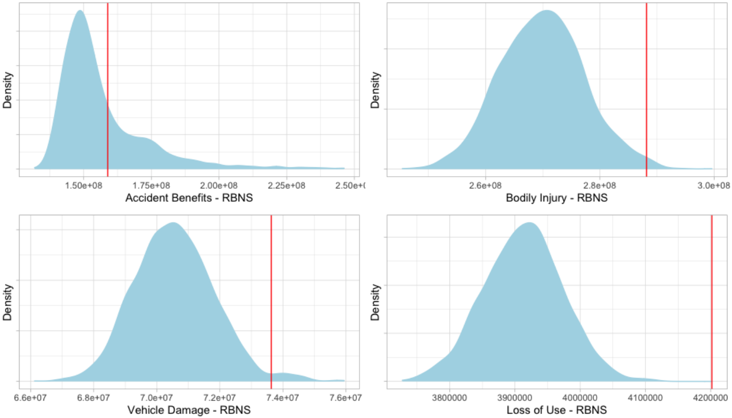


Figure 2: Evaluation date: 01/01/2019.

Results - RBNS Reserves

		Reserve	$VaR_{0.95}$
True reserve		524,909,605	/
Aggregated models	ODP Chain-Ladder	505,546,620	547,545,023
	Log-Normal Chain-Ladder	504,965,940	546,238,648
Independence model		514,189,692	554,286,625
Activation patterns model		504,363,024	542,753,541

Table 9: Comparison between the different models (evaluation date 01/01/2019).

Coverage	True reserves	Activation Patterns			Independence		
		Avg.	$VaR_{0.95}$	$VaR_{0.99}$	Avg	$VaR_{0.95}$	$VaR_{0.99}$
Accident Benefits	158.90	159.94	194.13	260.24	185.11	218.11	286.56
Bodily Injury	288.17	269.93	282.42	288.05	256.27	268.98	273.66
Vehicle Damage	73.63	70.58	72.65	74.19	68.78	70.81	72.08
Loss of Use	4.20	3.91	4.01	4.05	3.75	3.86	3.90
Total	524.91	504.36	542.75	605.27	513.91	551.60	612.29

Table 10: Quantiles of the simulated reserves (activation patterns model and independence model - in M$).

Contents

Conclusion

- Advantages of the Activation Patterns Model
  - Reduction of the higher quantiles used to estimate the reserves compared to the aggregate or the independence model
  - More accurate predictions reached for the four coverages
- Disadvantages of the Activation Patterns Model
  - Number of parameters
- Next steps?
  - Alternative ways to model the dependency
  - Switch to a non-parametric model

References I

Antonio, K. and Plat, R. (2014), Micro-level stochastic loss reserving for general insurance, *Scandinavian Actuarial Journal*, 2014:7, 649-669.

Bermüdez, L., Karlis, D. and Guillën, M. (2018), Allowing for time and cross dependence assumptions between claim counts in ratemaking models, *Insurance: Mathematics and Economics*, 83, 161-169.

Côté, M-P.,Genest, C. and Stephens, D.A. (2021), A Bayesian Approach to Modeling Multivariate Multilevel Insurance Claims in the Presence of Unsettled Claims, *Bayesian Analysis, Bayesian Anal. Advance Publication*, 1-27.

Denuit, M. and Trufin, J. (2017), Beyond the Tweedie Reserving Model : The Collective Approach to Loss Development, *North American Actuarial Journal*, 21:4, 611-619.

Frees, E. and Valdez, E. (2008), Hierarchical Insurance Claims Modeling, *Journal of the American Statistical Association*, 103:484, 1457-1469.

Frees, E., Shi, P., Valdez, E. (2009), Actuarial Applications of a Hierarchical Insurance Claims Model, *ASTIN Bulletin*, 39(1), 165-197.

References II

Pybus, M., Luisi, P.,Dall'Olio, G.M.,Uzkudun, M., Laayouni, H.,Bertranpetit, J.,Engelken, J. (2015), Hierarchical boosting: a machine-learning framework to detect and classify hard selective sweeps in human populations, *Bioinformatics*, 31:24, 3946–3952

Shi, P., Feng, X. and Boucher, J-P. (2016), Multilevel modeling of insurance claims using copulas, *The Annals of Applied Statistics*, 10, 834-863.

Yang, X., Frees, E.W. and Zhang, Z. (2011), A generalized beta copula with applications in modeling multivariate long-tailed data, *Insurance: Mathematics and Economics*, 49, 265:284.