Working with a Parametric Copula-Based model for Individual Non-Life Loss Reserving

ROXANE TURCOTTE[a,b,†], HÉLÈNE COSSETTE[a,*] and MATHIEU PIGEON[b,‡]

[a]École d'actuariat, Université Laval, Québec, Canada

[b] Département de mathématiques, Université du Québec à Montréal, Montréal, Canada

Emails: turcotte.roxane@courrier.uqam.ca[†], helene.cossette@act.ulaval.ca[*], pigeon.mathieu.2@uqam.ca[‡].

June 2020

Abstract

In this paper, we propose a generalization of the individual loss reserving model introduced by Pigeon et al. (2013) considering a discrete time framework for claims development. We use a copula to model the potential dependence within the development structure of a claim, which allows a wide variety of marginal distributions. We also add a specific component to consider claims closed without payment. We provide a case study based on a detailed personal auto insurance data set from a North American insurance company.

Key words: individual loss reserving, non-life insurance, micro-level model, copula.

1 Introduction

Insurance companies are subject to numerous regulatory standards to which they must comply to guarantee their financial engagements arising from insurance policies. In particular, they must follow the IFRS 17 *Insurance Contracts* issued by the International Accounting Standards Board (IASB) for the recognition, measurement and disclosure of insurance contracts. For non-life insurance companies, this results in the establishment of a loss reserve covering the liability for incurred claims (LIC), i.e. the unpaid part of insured claims that have occurred. A significant part of this non-financial risk is the frequency risk, that is the uncertainty related to the number of claims, and the severity risk, i.e. the uncertainty related to the total amount of each of the claims.

Most of the loss reserving models proposed in the literature can be classified into two categories: collective and individual approaches. This is mainly due to the granularity of the underlying data set. Collective loss reserving models are better known, more used in practice and have been investigated more extensively by researchers. These models rest on aggregated data, usually collected by accident year and development year, whose dynamic we try to capture. The information is usually gathered in a table referred to as a run-off triangle (see Figure 4). The actuarial literature on such models is vast: we refer readers to the overviews by Wüthrich and Merz (2008) and England and Verrall (2002).

Individual loss models aim to explain the dynamic of the development on a claim-level basis, before the claims are aggregated (see Figure 1). These approaches advantageously use detailed information on each payment and each contract to model the reserve of the portfolio. The first individual models were developed in the 1980s by Bühlmann et al. (1980), Hachemeister (1980) and Norberg (1986) but it is really at the turn of the century that such a loss reserving approach became popular. Many modeling strategies were pursued in parallel, in particular, the parametric or semi-parametric approaches presented by Antonio and Plat (2013), Pigeon et al. (2013), Larsen (2007), Zhao et al. (2009) or Zhao and Zhou (2010), and the approaches based on machine learning techniques, such as those of Wüthrich (2018), Baudry and Robert (2019) and Duval and Pigeon (2019). The idea to replace the Poisson distribution by a Cox process has been investigated by Avanzi, Wong and Yang (2016), Badescu, Lin and Tang (2019) and Badescu, Lin and Tang (2019). The goal is to account for dependence in the claim arrival process to refine the estimate of the IBNR reserve. Finally, a few research papers have compared collective and individual approaches (see Huang et al. (2015), Hiabu et al. (2016) or Charpentier and Pigeon (2016) for some examples).

In this paper, we introduce a model that has many features patterned on the basic structure of the models proposed by Antonio and Plat (2013) and Pigeon et al. (2013). We thoroughly analyze a portfolio from a Canadian casualty and property insurer, focusing on three popular car insurance coverages: automobile physical damage (APD), bodily injury (BI) and accident benefit (AB). The main objective of this paper is threefold:

- propose a generalization of an existing parametric model that allows more flexibility in the modeling of the dependence structure;
- discuss model implementation strategies;
- analyze the differences and similarities of the models obtained for the three available coverages.

Note that there is dependence between these three coverages, but in the data set considered, there are too few relevant observations for a model to be statistically valid at the individual level. Collective approaches, see for example Shi and Free (2011), De Jong (2012), Zhang and Dukic (2013), Abdallah et al. (2015) and Côté et al. (2016), have been proposed to model such dependence and, with an adequate data set, these approaches could be adapted to an individual setting.

In Section 2, we present the structure of the individual model proposed here, highlighting the generalizations of Antonio and Plat (2013) and Pigeon et al. (2013). In Section 3, we implement the model on a data set and we analyze the results. We conclude the paper in Section 4.

2 Individual Loss Reserving

In this section, we summarize the structure of the model, mainly inspired by Pigeon et al.

(2013). In Figure 1, we illustrate a typical development pattern of a claim and show the different reserves that must be modeled: incurred but not reported (IBNR), reported but not paid (RBNP), and reported but not settled (RBNS). Identically to Pigeon et al. (2013), time

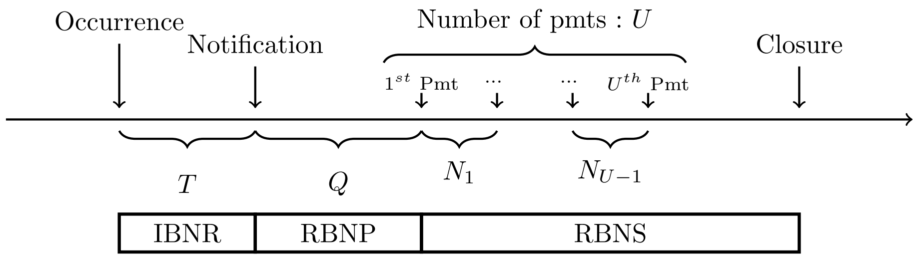

Figure 1: Typical development pattern of a claim in non-life insurance

is discretized in periods, which will be years in our case. Each claim is uniquely identified by (i, k) where $i \in \{1, \ldots, I\}$ corresponds to the i^{th} occurrence period and $k \in \{1, \ldots, K_i\}$ is the k^{th} claim from occurrence period i. To simplify the notation throughout, note that we do not make reference to (i, k) if unnecessary. To model such a claim, three basic components are needed: a model for the occurrence of a claim, which is the starting point of its development (see subsection 2.2), a series of discrete random variables (rvs) representing the time structure (see subsection 2.1), meaning the delays between the different events, and a random vector for the development of the severity of a claim (see subsection 2.3).

2.1 Time Components

Let us characterize the evolution of a claim (i, k) from its occurrence to its settlement through the following three random variables:

- $T_{i,k}$: rv representing the reporting delay (in years or, more generally, in periods) for the k^{th} claim of the i^{th} occurrence period, i.e., the number of periods between the occurrence period of a claim and the reporting time to the insurance company;
- $Q_{i,k}$: rv which represents the first payment delay (in periods) for the k^{th} claim of the i^{th} occurrence period, i.e., the number of periods between the reporting time to the insurance company and the first payment;
- $U_{i,k}$: rv which represents the total number of payments made for the k^{th} claim of the i^{th} occurrence period.

In practice, we often observe claims for which the file is closed without payment ($U_{i,k} = 0$). In order to have more flexibility to model this situation, we introduce a dichotomous rv $I_{i,k}$ to take into account the occurrence of a first payment and we define

$$U_{i,k} = \begin{cases} 0, & \text{if } I_{i,k} = 0 \\ U^*_{i,k} + 1, & \text{if } I_{i,k} = 1, \end{cases}$$

where $U^*_{i,k}$ corresponds to the number of payments made after the first one. At this point, we remind the reader that (i,k) is a unique identifier for each independent claim. Therefore, each claim is associated with a unique value for each time component $T_{i,k}$, $Q_{i,k}$ and $U_{i,k}$.

Because we are solely interested in the total amount of the reserve and not in the payment calendar, it is not necessary to construct a model for the delays between each payment. Each time component $T_{i,k}$, $Q_{i,k}$ and $U_{i,k}$ is a discrete rv with a cumulative distribution function (cdf) denoted by $F_1(\cdot;\boldsymbol{\nu})$, $F_2(\cdot;\boldsymbol{\psi})$ and $F_3(\cdot;\boldsymbol{\beta})$, $\mathbb{N}\to[0,1]$, with respective probability mass function (pmf) given by $f_1(\cdot;\boldsymbol{\nu})$, $f_2(\cdot;\boldsymbol{\psi})$ and $f_3(\cdot;\boldsymbol{\beta})$. Finally, we assume that $I_{i,k}\sim$ Bernoulli(p).

Depending on the stage of development, each observation of the data set provides more or less information on the distribution of each of these components. Indeed, given the category (IBNR, RBNS, RBNP or closed, see Figure 1) to which a claim at the evaluation date belongs, we present, in Table 1, the information available for the estimation of the parameters of the models. This information will be used to derive the likelihood functions maximized in subsection 2.4.

Table 1: Information available in the data set

Component	Category	Information
T	IBNR	NA
	RBNS	$(T=t\|T\leq t^*-1)$
	RBNP	$(T=t\|T\leq t^*-1)$
	Closed	$(T=t\|T\leq t^*-1)$
Q	IBNR	NA
	RBNS	$(Q=q\|Q\leq t^*-t-1)$
	RBNP	$(Q>t^*-t-1)$
	Closed	$(Q=q\|Q\leq t^*-t-1)$
U	IBNR	NA
	RBNS	$(U>u-1)$
	RBNP	NA
	Closed	$(U=u\|U\leq t^*-t-q-1)$

Note: The evaluation occurs t^* periods after the occurrence date.

2.2 Occurrence of a Claim

Let the rv K_i represent the number of claims for the occurrence period i with cdf denoted by $G_i:\mathbb{N}\to[0,1]$ such that $\mathrm{E}[K_i]=\int x dG_i(x)=\theta\omega_i$, where ω_i corresponds to the exposure for the occurrence period i. In order to choose and estimate a model for K_i, it is important to note that only the claims reported to the insurer are observed. This means that we must adjust the distribution of K_i, $i=1,\ldots,I$, as follows

$$\Pr(K_i=k\,|T_{i,1}\leq t_i^*-1,\ldots,T_{i,k}\leq t_i^*-1),$$

where the evaluation takes place t_i^* periods after the occurrence and the evaluation is done at the beginning of the period, before payments are made by the insurer. In the simplest case, $K_i \sim \text{Poisson}(\theta\omega_i)$ and

$$\left(K_i \,|T_{i,1} \le t_i^* - 1, \ldots, T_{i,k} \le t_i^* - 1\right) \sim \text{Poisson}\left(\theta\omega_i F_1\left(t_i^* - 1; \boldsymbol{\nu}\right)\right),$$

where F_1 is the cdf of the reporting delay rv $T_{i,k}$.

2.3 Severity Model

We have chosen to model the severity of the claims based on development factors like those of the Chain-Ladder method but adequately adapted in an individual loss reserving framework. In the well-known Chain-Ladder or Mack's model for aggregated data, the multiplicative development factors link cumulative payments from one development period to the next. In the model here, the development factors rather link the cumulative payments for each claim individually.

Let $Y_{i,k,j},\ j \in \{1, 2, \ldots, U_{i,k}\}$ denote the severity of the j^{th} incremental payment of claim (i,k) and $\boldsymbol{\Lambda}^{(i,k)}$ be a random vector of dimension $U_{i,k}$, assuming that a vector of dimension 0 is simply equal to 0.

More precisely, let $\boldsymbol{\Lambda}^{(i,k)}$ be given by

$$\boldsymbol{\Lambda}^{(i,k)} = \begin{bmatrix} Y_{i,k,1} & \lambda_1^{(i,k)} & \cdots & \lambda_{U_{i,k}^*}^{(i,k)} \end{bmatrix}^T,$$

where $\lambda_j^{(i,k)} = \sum_{r=1}^{j+1} Y_{i,k,r} / \sum_{r=1}^{j} Y_{i,k,r}$. Thus, the ultimate amount of claim (i,k), denoted by $C_{i,k}$, is given by

$$C_{i,k} = \prod_{j=1}^{U_{i,k}} \Lambda_j^{(i,k)} = Y_{i,k,1} \times \prod_{j=1}^{U_{i,k}^*} \lambda_j^{(i,k)},$$

where $\boldsymbol{\Lambda}_j^{(i,k)}$ corresponds to the j^{th} component of $\boldsymbol{\Lambda}^{(i,k)}$.

In this paper, we choose a multivariate distribution for $\boldsymbol{\Lambda}^{(i,k)}$ defined through a copula, i.e.

$$F_{\boldsymbol{\Lambda}}(\boldsymbol{\lambda}) = \mathcal{C}\left(F_{Y_1}(y_1), \ldots, F_{\lambda_{u^*}}(\lambda_{u^*})\right),$$

where $\mathcal{C}$ is the copula and $F_{Y_1}, F_{\lambda_1}, \ldots, F_{\lambda_{u^*}}$ are respectively the marginal cdf of each component of the development vector $\boldsymbol{\Lambda}$. For the sake of simplicity, we assume that $U_{i,k}$ only affects the dimension of the copula, but not the joint distribution of $\boldsymbol{\Lambda}$ in the same way as the multivariate skew normal distribution in Pigeon et al. (2013). This assumption should be challenged in future research. We assume these distributions to be continuous on $\mathbb{R}^+$. For the first payment Y_1, we consider various classical distributions such as the gamma and lognormal distributions while for the development factors, a wide range of distributions are investigated. For the choice of the marginal distributions as well as for the dependence structure, this approach is more flexible than multivariate distributions used in other papers.

For the choice of the copula, we consider first and foremost elliptical copulas. The normal and Student copulas have the advantage of being defined through a correlation matrix $\boldsymbol{\alpha}$ that allows us to consider diversified correlation structures. The Student copula may be preferable given its ability to capture dependence in the extreme values. The degree of freedom obtained will be an indicator of the more suitable model.

In practice, to establish the model, we have to choose the maximum dimension M of the random vector $\boldsymbol{\Lambda}^{(i,k)}$ beforehand. This will be done considering the data set used and a tail factor, e.g. a geometric mean could be defined for cases where the number of payments exceeds M.

2.4 Estimation

Let us now define the likelihood functions that will be maximized to estimate the parameters of the models. Of course, these functions will consider the use of the density c of the corresponding copula $\mathcal{C}$ for the dependence structure linking the components of $\boldsymbol{\Lambda}^{(i,k)}$. For closed claims, we denote the likelihood function by L^{Cl} which is given by

$$\begin{aligned} L^{\text{Cl}} \propto & \prod_{(i,k)_{\text{Cl}}} c\left(F_{Y_1}(y_1), \ldots, F_{\lambda_{u^*_{i,k}}}(\lambda_{u^*_{i,k}}); \boldsymbol{\alpha} | u_{i,k}\right) \times f_{Y_1}(y_1) \times \ldots \times f_{\lambda_{u^*_{i,k}}}(\lambda_{u^*_{i,k}}) \\ & \times \prod_{(i,k)_{\text{Cl}}} f_1(t_{i,k}; \boldsymbol{\nu} | T_{i,k} \leq t^*_{i,k} - 1) \times f_2(q_{i,k}; \boldsymbol{\psi} | Q_{i,k} \leq t^*_{i,k} - t_{i,k} - 1) \\ & \times \prod_{(i,k)_{\text{Cl}}} f_3(u_{i,k}; \boldsymbol{\beta} | U_{i,k} \leq t^*_{i,k} - q_{i,k} - t_{i,k} - 1). \end{aligned}$$

For the RBNS claims, the likelihood function, denoted by L_c^{RBNS}, is

$$\begin{aligned} L^{\text{RBNS}} \propto & \prod_{(i,k)_{\text{RBNS}}} c\left(F_{Y_1}(y_1), \ldots, F_{\lambda_{u^*_{i,k}}}(\lambda_{u^*_{i,k}}); \boldsymbol{\alpha} | u^{ed}_{i,k}\right) \times f_{Y_1}(y_1) \times \ldots \times f_{\lambda_{u^*_{i,k}}}(\lambda_{u^*_{i,k}}) \\ & \times \prod_{(i,k)_{\text{RBNS}}} f_1(t_{i,k}; \boldsymbol{\nu} | T_{i,k} \leq t^*_{i,k} - 1) \times f_2(q_{i,k}; \boldsymbol{\psi} | Q_{i,k} \leq t^*_{i,k} - t_{i,k} - 1) \\ & \times \prod_{(i,k)_{\text{RBNS}}} (1 - F_3(u^{ed}_{i,k} - 1; \boldsymbol{\beta})), \end{aligned}$$

where $u^{ed}_{i,k}$ represents the observed number of payments at the evaluation date. Finally, for the RBNP claims, the likelihood function is given by

$$L^{\text{RBNP}} \propto \prod_{(i,k)_{\text{RBNP}}} f_1(t_{i,k}; \boldsymbol{\nu} | T_{i,k} \leq t^*_{i,k} - 1)(1 - F_2(t^*_{i,k} - t_{i,k} - 1; \boldsymbol{\psi})).$$

The likelihood functions are evaluated considering all observations classified in the respective categories, $(i,k)_{\text{Cl}}$, $(i,k)_{\text{RBNS}}$ and $(i,k)_{\text{RBNP}}$, at the evaluation date. The marginal cdfs $F_{Y_1}, \ldots, F_{\lambda_{u^*_{i,k}}}$ and their associated probability density functions (pdfs) $f_{Y_1}, \ldots, f_{\lambda_{u^*_{i,k}}}$ can be estimated by the fully parametric inference function for margins (IFM) method. This parametric two-step procedure is more flexible than the maximum likelihood method and more amenable to computations. We can also use a pseudo maximum likelihood method by first estimating each marginal nonparametrically by the empirical distribution function not requiring specifying functional forms for the marginals. Then, the dependence parameters of the parametric copula family are estimated by the values that maximize the log pseudo likelihood function. See Joe, H. (2014) and Genest and Favre (2007) for more details on these estimation methods.

2.5 Predictions

Once all components of the model are estimated, we can proceed with the prediction of the total reserve amount. An approach by Monte Carlo simulation allows us to find the predictive distribution of the reserve and for the insurer to use different risk measures to determine its economic capital. For each reserve category, we simulate the missing part of the development, based on observed information. We will obtain a reserve for each claim, for each category and hence for the entire portfolio. Here is a summary of the components to be simulated for each reserve category:

1. For the IBNR reserve, we will need to simulate, for each occurrence period, a realization of the conditional rv

$$(K|T_1 \leq t^* - 1, \ldots, T_{k^*} \leq t^* - 1, T_{k^*+1} > t^* - 1, \ldots, T_K > t^* - 1),$$

   where k^* represents the observed number of claims in the considered period. For the case where $K_i \sim$ Poisson $(\theta\omega_i)$, we obtain a conditional rv with a Poisson $(\theta\omega_i F_1 (t_i^* - 1; \nu))$ distribution. Then, for each IBNR claim, we simulate a realization of the total number of payments U and of the development vector $\mathbf{\Lambda}$. Finally, the IBNR reserve amount is given by $Y_1 \times \prod_{j=2}^{U} \Lambda_j$.

2. For each observed RBNP claim, we simulate a realization of the number of payments U and a realization of $\mathbf{\Lambda}$ of dimension U. We obtain a realization of the RBNP reserve with $Y_1 \times \prod_{j=2}^{U} \Lambda_j$. The only difference between the IBNR reserve and the RBNP reserve is that in the former case, we do not know the number of claims, whereas this number is known in the latter case.

3. For each observed RBNS claim, we simulate a realization of the conditional rv $\left(U \,\middle|\, U \geq u^{ed}\right)$ to obtain the number of missing payment(s). Note that the development vector $\mathbf{\Lambda}$ can be written as

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}^* & \mathbf{\Lambda}^- \end{bmatrix},$$

   where $\mathbf{\Lambda}^*$ corresponds to the observed part of $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^-$ to the unobserved development factors. We will then have to simulate a realization of the complete vector conditionally on the observed values $(\mathbf{\Lambda} \,|\mathbf{\Lambda}^* = \boldsymbol{\ell})$ where $\boldsymbol{\ell} = \left(Y_1, \lambda_1, \ldots, \lambda_{(u^{ed}-1)}\right)$. We sum all the claims to obtain an estimation of the RBNS reserve given by $Y_1 \times \left(\prod_{j=2}^{U} \Lambda_j - \prod_{j=2}^{u^{ed}} \Lambda_j^* \right)$.

3 Numerical Analysis

3.1 Data set

In this section, we provide a detailed analysis of a personal auto insurance data set from a Canadian property and casualty insurance company. The initial data set contains information regarding more than 100,000 car insurance claims for a period running from 2004 to 2016. We consider claims for three well-known types of coverage in the data set: automobile physical damage (APD), bodily injury (BI), and accident benefit (AB). APD coverage includes all guarantees regarding the loss or the partial loss of the car, such as collision or theft. BI and AB coverages are related to medical expenses and other related fees for insureds and a not-at-fault third party. All calculations have been done in R, using the **copula** package. This package is mainly used to manipulate elliptical and Archimedean copula families. More specifically, the function cCopula based on the Rosenblatt transformation (see Rosenblatt (1952) and Hofert et al. (2012)) has been used to generate realizations of rvs conditionally on what has been observed, i.e., to compute $\mathcal{C}(u_d|u_1, \ldots, u_{d-1})$.

We will distinguish two types of statuses for a claim: "open" or "closed." Because the data set did not explicitly include information regarding the status of a claim, a claim will be considered "open" if the amount of the reserve of the current period is nonzero or if the annual paid amount in the current year is different from zero. Also, for BI coverage, we consider as being "open" any claim with the following two additional characteristics:

- the development year is below or equal to 3, and
- no payment or reserve has ever been allocated to this claim.

Otherwise, the claim is considered "closed". Adding these conditions for the BI coverage is justified given the slower development of these claims, as observed in Table 5, in which the most important payments occur in the third or even fourth development year. The absence of a payment or a reserve in the first years does not guarantee that there will be no future development. Finally, a claim whose status goes from "closed" to "open" is considered reopened. In Table 2, we present other characteristics of the data set such as the number of claims with a refund, with multiple refunds, with reopenings, etc.

Table 2: Additional features of the data set

Class	Status	Refund (mult.)	Reopening (mult.)	No payment	One payment or more
AB	Open	281 (83)	364 (12)	1,251	8,501
	Closed	463 (64)	474 (4)	2,410	12,883
BI	Open	26 (1)	642 (2)	1,447	3,220
	Closed	46 (1)	628 (1)	3,370	5,529
APD	Open	242 (8)	310 (8)	754	8,490
	Closed	803 (5)	481 (4)	961	26,108

Many reopenings tend to indicate that the file is complicated to settle because its development differs from the typical development pattern of a claim presented in Figure 1. Indeed, there is something uncommon about the claims that required further developments from the original settlement. The coverage with the most reopenings is BI. Table 3 presents descriptive statistics for total payments of closed claims. The total payments for the BI coverage vary greatly, yet it is the class for which we have the fewest observations. The median total payment for BI is the lowest, while the average is the highest. Those statistics show that under the BI coverage, a large number of small payments are made as well as a non-negligible number of large payments, leading to the very large standard error observed. In comparison with the other two classes, these statistics illustrate how much riskier AB and BI coverages are: standard deviations and maximums are much higher than for APD. We also notice low minimums for all three classes, which are most likely administration fees that cannot be separated from the actual indemnity, adding to the complexity of the data set.

Table 3: Descriptive statistics for total payments of closed claims

Class	Mean	Median	Std. err.	Min	Max	n
AB	18,260	4,050	60,539	3.40	2,312,000	15,565
BI	21,300	1,206	80,767	0.10	2,147,000	5,529
APD	6,152	4,596	5,717	0.18	102,300	26,108

To validate the results, we consider January 1, 2012 as the evaluation date. Hence, only the information known at that date will be used for the model adjustment and for the analysis. The incremental loss development triangles for the three types of coverages, AB, BI and APD, are respectively given in Tables 4, 5 and 6. The amounts in gray correspond to

Table 4: Incremental run-off triangle for AB (in \$100,000)

	1	2	3	4	5	6	7	8
2004	113	142	87	61	70	61	43	19
2005	120	186	117	83	95	61	53	24
2006	132	201	100	7	51	40	23	21
2007	157	220	135	95	98	77	26	17
2008	146	259	136	116	68	33	66	63
2009	191	358	189	129	98	85	27	23
2010	221	324	176	111	72	86	28	?
2011	115	180	112	134	125	120	?	?

those observed, and are used only to validate the results. We note that the development of claims in AB and BI is slower than for APD claims. The amounts paid for AB claims are more important in the first two/three years and then decrease slowly thereafter. For BI claims, we observe the opposite: the development of the claims is rather slow in the first two/three years and then accelerates, before slowing down again around the seventh year. Unlike BI claims, APD claims take essentially two years for their development, leaving only minor payments

Table 5: Incremental run-off triangle for BI (in $100,000)

	1	2	3	4	5	6	7	8
2004	9	17	26	64	85	124	81	55
2005	7	21	29	86	138	113	68	36
2006	9	24	57	84	97	114	47	35
2007	7	31	100	108	133	87	43	27
2008	8	33	91	96	108	61	59	34
2009	8	28	87	128	157	94	82	50
2010	8	41	86	150	124	142	71	?
2011	15	41	85	128	141	224	?	?

Table 6: Incremental run-off triangle for APD (in $100,000)

	1	2	3	4	5	6	7	8
2004	256	36	0.33	0.19	0.06	−0.12	−0.02	−0.02
2005	258	43	0.86	0.24	−0.04	0.03	0.00	−0.09
2006	264	36	0.40	0.01	0.03	−0.01	−0.10	−0.03
2007	230	34	0.33	0.02	−0.18	−0.11	−0.01	0.00
2008	199	28	−0.36	−0.03	0.02	−0.04	0.00	0.00
2009	209	31	0.10	0.17	0.08	0.05	0.06	0.00
2010	222	41	1.29	0.30	0.20	0.27	0.45	?
2011	278	46	0.38	0.58	0.40	0.55	?	?

thereafter. From the information available between 2012 and 2016 (in gray in Tables 4, 5 and 6), it is possible to obtain an approximation of the total paid amount for each coverage: $206,253,310 for the AB reserve, $246,404,599 for the BI reserve and $5,015,952 for the APD reserve. It is important to note that the "true" reserve amounts should be slightly higher than the ones previously given, in particular for AB and BI coverages given that the total amount of claims that occurred between 2004 and 2011 are not completely settled at the end of 2016.

For comparison purposes, we present the results obtained with classical collective approaches in subsection 3.2 before presenting the results obtained with the individual copula-based approach in subsection 3.3.

3.2 Collective Approaches

We compare our individual loss reserving approach to classical collective models such as Mack's model and parametric models based on the Poisson, gamma and Tweedie distribution (see Wüthrich and Merz (2008) for a detailed presentation). These models are benchmarks that insurers use to evaluate solvency. Results obtained are given in Table 7. We have also included in Table 7 the 95th and 99th percentiles of the distribution of the approximated reserve obtained by a bootstrap approach. For AB claims, we note that the results are rather interesting as the observed total amount is within one standard deviation of the

Table 7: Prediction results for collective approaches

		Mack	GLM (Poisson)	GLM (gamma)	GLM (Tweedie)
AB	Exp. value	226,785,847	226,785,847	222,354,920	226,749,037
	std.-er.	14,639,535	16,518,916	16,207,136	15,180,596
	95^{th} quantile	252,589,598	252,584,138	248,999,637	252,208,184
	99^{th} quantile	262,982,300	264,338,648	259,956,817	261,149,898
BI	Exp. value	395,259,746	395,259,746	364,125,084	376,913,901
	std.-er.	70,725,189	70,870,255	58,043,349	60,340,681
	95^{th} quantile	565,297,170	532,238,561	474,503,383	486,860,190
	99^{th} quantile	701,490,917	595,634,011	530,830,765	523,073,345
APD	Exp. value	4,267,915	-	-	-
	std.-er.	493,837			
	95^{th} quantile	5,016,298	-	-	-
	99^{th} quantile	5,379,680	-	-	-

Note: Observed total amounts are $206,253,310 for the AB reserve, $246,404,599 for the BI reserve and $5,015,952 for the APD reserve.

expected value, and the variance represents less than 10% of the expected amount predicted. Otherwise, the estimated parameter $\widehat{p}$ of the Tweedie model is approximately 1.01, which explains the similarity of the results of the Poisson and Tweedie models. However, the quantiles are considerably above the true value of the reserve. In fact, the values are in the same range as the observed value.

For BI coverage, the estimations are much higher than the true value of the reserve, and the variance is high compared with the estimated value (more than 15%). The estimated parameter $\widehat{p}$ of the Tweedie model is 1.72. The parametric models perform better than Mack's model in that the 99^{th} quantiles are not unreasonably higher than the 95^{th} quantiles. As can be seen in the incremental loss development triangle (see Table 5), the highest amounts are not in the first years of development. This is due to claims with a large first payment, for which the delay for the first payment was longer than claims with a smaller first payment. For BI coverage, there are numerous dynamics for the settlement of claims, which complicates the treatment of aggregated data. Indeed, if the proportion of claims between these different dynamics change, adjustments must be made to pursue with a common settlement dynamic for all claims. However, incorporating this type of information is not consistent with the ideas behind such models. Also, the quantiles are considerably higher than the realizations for this type of coverage. In such a case, an insurer using this model would set aside a reserve amount much higher than that truly needed.

Results for GLM models are not given for APD claims. Such models are not appropriate due to numerous negative entries, which leads to conclude that these models do not generally capture the dynamic behind the data. As for Mack's model, it underestimates the necessary reserve. We could explain this underestimation by the fact that the negative elements of the development triangle (see Table 6) lead to development factors below 1, starting from the fourth factor. Refunds are specific to each file, and evolve over time. For instance, the insurance company can sell the iron from a car destroyed in an accident to recover some money from the total loss. More generally, an insurer can use its right of subrogation to regain money from insured loss payments. Collective approaches cannot easily handle this kind of portfolio characteristics because it is not consistent over time. Given the number of refunds received for APD claims (see Table 2) and the short settlement time (see Table 9), it is not surprising to have negative entries in Table 6.

Considering an individual approach would be more appropriate in this context given that a collective approach is based on the idea of stability through time. However, this portfolio contains many settlement patterns and a significant variation in the number of insured exposure units. In Tables 4, 5 and 6, unsteady variation can be observed down the first column (especially for year 2011), when more stability could have been expected for the first development year.

3.3 Individual Approaches

3.3.1 Time structure

Various parametric models have been tested to model rvs $T_{i,k}$, $Q_{i,k}$ and $U^*_{i,k}$. Besides the classical counting distributions, we have considered mixtures with degenerate components such as

$$f\left(x;\boldsymbol{\xi}\right)=\sum_{s=0}^{p}\xi_s\mathbb{I}_{\{x=s\}}+\left(1-\sum_{s=0}^{p}\xi_s\right)g\left(x\right), \tag{3.1}$$

where $g()$ can be the pmf of the Poisson, negative binomial or binomial distribution, $\boldsymbol{\xi}=\begin{bmatrix}\xi_0 & \cdots & \xi_p\end{bmatrix}$ is a vector of parameters and $p=0,1,2$ or 3. Final choices were made using the AIC and BIC criteria. We present the results in Table 8.

As suggested in subsection 2.1, we include a Bernoulli distribution to model the possibility of closure without any payment ($I=0$). Estimated values and standard deviations of the Bernoulli parameter for each class are $\widehat{p}_{AB}=0.356\,(0.004)$, $\widehat{p}_{BI}=0.140\,(0.003)$ and $\widehat{p}_{APD}=0.658\,(0.004)$ for IBNR claims and $\widehat{p}_{AB}=0.111\,(0.003)$, $\widehat{p}_{BI}=0.066\,(0.002)$ and $\widehat{p}_{APD}=0.170\,(0.005)$ for RBNP claims. The probability that an IBNR claim closes with a payment is significantly higher than for an RBNP claim. Therefore, we conclude that the additional information that no payment has been made during the first year is relevant to predict the probability that payments will be made on that claim.

In the data set used, we observed that a non-negligible number of claims were closed without payment. When we modeled the occurrence of claims, all claims in the data set were included since we are interested in a claim being open and not being settled. However, we use this distribution to predict the number of IBNR claims, whereas the number of predicted IBNR claims and RBNP claims does not reflect the number of claims for which the insurer will have to make a payment to settle the claim. We will determine the number of predicted IBNR claims and RBNP claims for which we will have to simulate a claim severity by making an adjustment for the number of predicted IBNR claims and observed RBNP claims.

We use a binomial distribution to model the number of claims for which a development will occur. To estimate parameter p of the binomial distribution, we consider only claims that have been in the data set for at least five years. This allows us to have a reasonable

Table 8: Final models for the time structure

Class	$(T, \widehat{\boldsymbol{\nu}})$	$(Q, \widehat{\boldsymbol{\psi}})$	$(U^*, \widehat{\boldsymbol{\beta}})$
	Negative binomial	2-Geometric	Negative Binomial
AB	0.136 (0.027)	0.785 (0.028)	2.166 (0.075)
	0.853 (0.026)	0.652 (0.041)	0.682 (0.008)
	-	0.179 (0.006)	-
	1-Poisson	2-Poisson	2-Poisson
BI	0.560 (0.049)	0.935 (0.039)	1.859 (0.049)
	0.900 (0.007)	0.274 (0.028)	0.645 (0.006)
	-	0.055 (0.016)	0.078 (0.006)
	Negative binomial	2-Geometric	2-Geometric
APD	0.097 (0.027)	0.842 (0.058)	0.755 (0.029)
	0.906 (0.024)	0.878 (0.025)	0.810 (0.017)
	-	0.085 (0.003)	0.113 (0.003)

Note: A x-distribution means a mixture of the distribution with x degenerate component(s) as defined in Equation (3.1). We give the standard errors next to each of the estimated values.

protection against future claim developments. We estimate this parameter by its empirical counterpart.

The probability that an RBNP claim is not settled varies depending on whether we possess the additional information that no payment was made in the first year. We therefore estimate the probability that an RBNP claim will eventually lead to a payment by calculating the ratio of the number of closed claims for which no payment was made in the first year but the total amount paid is not zero to the total number of closed claims with no payment in the first year.

It could be possible to refine this method if we had the underwriting rules or information concerning the construction of the data set. Individual models allow to handle such targeted adjustments. This kind of analysis is not possible in a collective approach. For example, Table 2 shows for BI coverage that more than 35% of closed claims have been settled without payment. In a collective approach, such a particularity in the data set would certainly have been missed because only the aggregated information is used in the modeling process.

3.3.2 Development structure

Modeling the development vector $\boldsymbol{\Lambda}$ constitutes one of the key questions in our analysis given the potential strong associations between the components of this vector. We present in Table 9 descriptive statistics of $\boldsymbol{\Lambda}$ for closed claims. First, we observe, at an individual level this time, a situation similar to that observed in Tables 4, 5 and 6 with respect to the development pattern. Indeed, we see faster development for APD coverage than for the other coverages. We also note the same slower development of the claims at the beginning for BI coverage with a higher development factor λ_1 on average. Moreover, high values for standard errors of the development vector $(1269, 91.03, 12.63, 11.87)$ seem to support the idea that it could be risky to model the development of a BI claim from an aggregated point of view.

Table 9: Descriptive statistics for closed claims

Class	Variable	Mean	Median	Std. err.	Min	Max	n
AB	Y_1	6,213	2,286	10,280	3.36	273,200	15,565
	λ_1	8.00	1.90	46.97	0.00	2,769	8,929
	λ_2	1.53	1.29	1.24	0.01	44.10	4,142
	λ_3	1.37	1.17	1.08	0.01	35.11	1,715
	λ_4	1.41	1.15	1.13	0.01	18.35	628
	λ_5	1.38	1.11	0.84	0.16	7.82	232
	λ_6	1.28	1.20	0.52	0.15	2.73	68
BI	Y_1	3,782	720.60	17,280	0.10	789,400	5,529
	λ_1	150.40	6.25	1,269	0.11	34,370	1,374
	λ_2	20.55	3.63	91.03	0.16	1,499	635
	λ_3	7.73	2.70	12.63	0.71	94.19	280
	λ_4	7.45	2.61	11.87	0.98	90.57	114
APD	Y_1	5,963	4,378	5,798	7,743	102,300	26,108
	λ_1	17.58	1.10	119.70	0.00	5,385	3,444
	λ_2	1.25	1.05	1.48	0.02	14.20	87

Marginal Distributions. To adjust and select a marginal distribution for each component of the development vector $\boldsymbol{\Lambda}$, we include all available information, independently of the claim status (open or closed). We consider various distributions such as the gamma, normal, lognormal, Weibull, Gumbel, Pareto Type I, Pareto Type 2 (Lomax), generalized Pareto, loggamma, etc. We estimate all parameters using maximum likelihood techniques and appropriate R functions. We made our final choice based on the AIC and BIC. For each rv, Table 10 presents the selected distribution with the estimated values of the parameters (see Appendix A for the parametrization of the distributions). For the third component of the development vector (λ_2) for AB coverage, we find that the loglogistic distribution is slightly better than the Burr distribution (AIC of 8,123 versus 8,147). We have selected the latter in order to avoid complicating the model unnecessarily. Further, we observe that the same distribution (lognormal) is selected to model the first payment for the two coverages for medical expenses and other related fees while a different distribution (Pareto II) is chosen for material damage coverage.

Dependence. To characterize the strength of the dependence relation between the components of the development vector $\boldsymbol{\Lambda}$, we estimate Spearman's rho and Kendall's tau empirically, and we present the results in Table 11. We have estimated only the components for which we had at least 150 observations in the data set. These results will guide us in the choice of the dependence structure within the individual models.

To estimate the dependence structure, we use a semi-parametric estimation method based on ranks (see e.g., Genest and Rivest (1993) and Genest and Favre (2007)) since the marginal

Table 10: Information criteria and estimated values for the selected marginal distributions

Class	Components	Distributions	AIC	BIC	$\widehat{\theta}_1$	$\widehat{\theta}_2$
AB	Y_1	Lognormal	417,984	418,000	7.855	1.662
	λ_1	Burr	54,753	54,768	0.145	6.624
	λ_2	Burr	8,147	8,161	0.326	8.592
	λ_3	Burr	2,574	2,585	0.435	8.132
	λ_4	Burr	1,265	1,275	0.518	6.632
BI	Y_1	Lognormal	147,700	147,714	6.417	1.944
	λ_1	Burr	21,316	21,328	0.048	9.921
	λ_2	Burr	7,549	7,560	0.043	16.716
	λ_3	Burr	3,012	3,021	0.035	0.258
APD	Y_1	Pareto II	671,681	671,698	144,532	877,025,786
	λ_1	Burr	18,832	18,845	0.230	5.373

distributions are unknown. For indication only, we also consider a classical parametric estimation method (IFM) proposed by Joe and Xu (1996). The values obtained for the different association measures indicate that we must choose a copula allowing negative dependence within the model. For AB and BI claims, we must also consider a copula flexible enough to permit a level of association that differs between the components of the development vector $\mathbf{\Lambda}$. For that reason, we will try to adjust only elliptical copulas for these two types of coverages. For APD claims, we will also test the Frank copula. Finally, the Gumbel copula is not appropriate to model negative dependence and the atypical behavior of the Clayton copula when the dependence parameter is negative excludes it from being considered in dependence modeling for a claim.

To estimate the dependence structure, we will only use closed claims for which we have the complete information. The dependence structure could be different for claims needing only a few payments and claims with a long settlement period. However, we must mention that by taking out the censored data, claims that are paid and closed rapidly necessarily represent a larger proportion of the data used in the estimation, which could impact the results.

In Table 12, we present the AIC criteria for each copula, and each estimation method and Table 11 presents the estimation results obtained for each coverage with the semi-parametric method for a normal copula with parameter Σ^{Normal} and a Student copula with parameters Σ^{Student} and υ. Note that for the latter copula, we obtain υ = 9,800, 47.7 and 11.2 for the AB, BI and APD claims respectively. The results for the Frank copula for APD claims have not been included given that it was the copula with the worst adjustment.

Predictive distribution. Figure 2 presents the predictive distribution for AB and BI coverages using the normal copula, and in Figure 3 the predictive distributions for BI and APD coverages using the Student copula. As expected ($\upsilon = 47.7$), we observe that both graphs for BI coverage are quite similar. This is confirmed with the values given in Table 13,

Table 11: Results for elliptical copulas

Class	Components	Spearman's rho	Kendall's tau	$\widehat{\Sigma}^{\text{Normal}}$	$\widehat{\Sigma}^{\text{Student}}$
AB	$(2,1)$	-0.450	-0.310	-0.45	-0.45
	$(3,1)$	-0.147	-0.097	-0.14	-0.14
	$(4,1)$	-0.078	-0.051	-0.10	-0.10
	$(5,1)$	0.024	0.018	0.00	0.00
	$(3,2)$	0.202	0.136	0.18	0.18
	$(4,2)$	0.247	0.167	0.19	0.19
	$(5,2)$	0.163	0.108	0.09	0.09
	$(4,3)$	0.141	0.096	0.06	0.06
	$(5,3)$	0.231	0.158	0.27	0.27
	$(5,4)$	0.223	0.156	0.11	0.11
BI	$(2,1)$	-0.559	-0.392	-0.56	-0.56
	$(3,1)$	-0.191	-0.129	-0.20	-0.20
	$(4,1)$	-0.135	-0.092	-0.09	-0.09
	$(3,2)$	-0.089	-0.059	-0.24	-0.24
	$(4,2)$	-0.014	-0.010	-0.04	-0.03
	$(4,3)$	-0.121	-0.081	-0.36	-0.37
APD	$(2,1)$	-0.776	-0.597	-0.69	-0.70

Note: All the values in columns Spearman's rho and Kendall's tau are significantly different from zero ($\alpha = 1\%$) excepted for the values in gray.

which shows that the two models are not significantly different than the expected value, the standard deviation and the 95^{th} and 99^{th} quantiles.

We note that the model for AB claims gives good results. The estimation of the expected value is slightly lower than its realization, but only by \$2 million, which is less than the standard deviation. Furthermore, the variability represents more than an acceptable proportion of the estimated value of the reserve. The empirical coefficient of variation is around 4%. Consequently, the upper quantiles of the predictive distribution represent a considerable amount to be put aside as a reserve to cover possible future developments or reopening of claims, without being excessively high. For BI claims, the estimated value of the reserve is around \$50 million over the realization, which represents a little more than 20% of the observed reserve. Finally, for APD coverage, the predicted amount is again slightly higher than the realization, but the performance of the model is satisfactory given that it predicts an amount above the observed reserve, in a similar range of values.

Out of the three coverages, AB provided the best overall results for both individual and collective approaches. The prediction for the individual model is closer to the observed value. On data containing numerous reopenings, like BI coverage, the advantages of an individual model are more perceptible. Even though neither approach could easily get the average close to the observed value, the results from the individual approach are a lot more sensible, especially in the upper quantiles. As for APD coverage, refunds affect this coverage considerably, and no special adjustment has been made in the individual model to take into

Table 12: Information criteria for the dependence structure. "Ranks" stands for the semi-parametric inference method based on ranks

	AB		BI		APD		
	Normal	Student	Normal	Student	Normal	Student	Frank
Ranks	$\mathbf{-1{,}868}$	$-1{,}866$	$\mathbf{-593.23}$	$\mathbf{-593.25}$	$-3{,}360$	$\mathbf{-3{,}456}$	$-3{,}324$
IFM	$-2{,}154$	$-2{,}243$	-487	-501	$-3{,}498$	$-3{,}752$	$-3{,}026$
# parameters	10	11	6	7	1	2	1

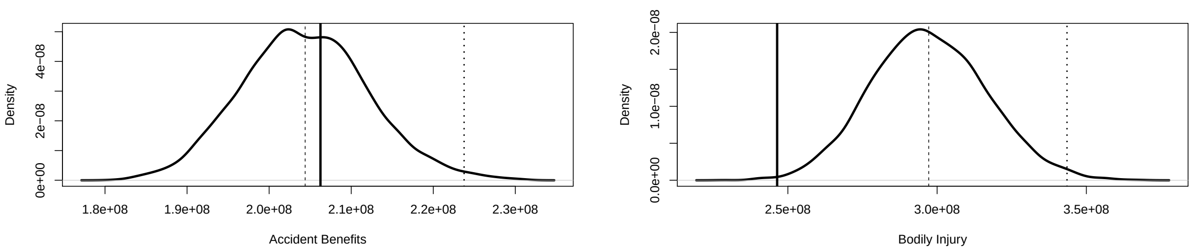

Figure 2: Predictive distribution based on a normal copula for AB (left-hand side) and BI (right-hand side) claims. The solid vertical line represents the observed reserve and broken lines represent the simulated average value, and the 99^{th} quantile

account the numerous refunds, even though it was possible to do so. We did not pay special attention to this characteristic because it is negligible for bodily injury coverage, leading to a slightly conservative loss reserve for APD coverage. The model uses the total number of payments without necessarily considering if the payment has been made after a reopening. In our situation, the information about the claim status (open or closed) is an approximation. With a data set including reliable information on the claim status, this should be part of the model.

4 Conclusion

In this paper, we propose a generalization of the loss reserving model introduced in Pigeon et al. (2013). Compared with the existing model, estimating the marginals and the dependence structure separately increases flexibility and facilitates the estimation procedure given that fewer parameters must be optimized simultaneously. However, strategy has a price: the total number of parameters, and therefore the complexity of the model, increase. Moreover, an individual parametric approach makes it possible to better understand and model the complexity of the underlying data set.

We also perform a detailed case study based on a micro-level data set from the industry. Our main conclusions are

- the presented individual parametric model allows us to capture the complexity of claims development for AB and BI coverages, yet it is less relevant for APD coverage;

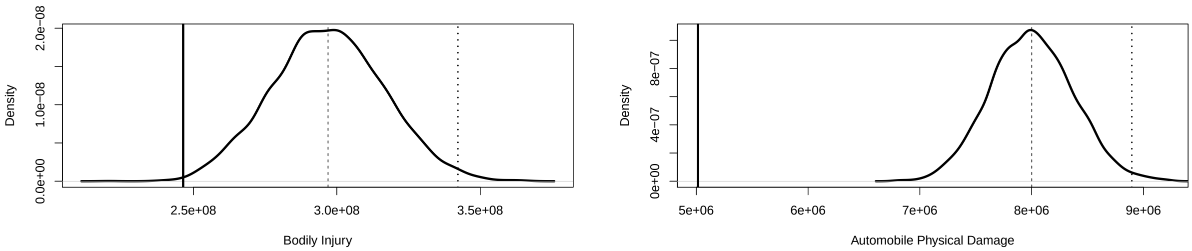

Figure 3: Predictive distribution based on a Student copula for BI (left-hand side) and APD (right-hand side) claims. The solid vertical line represents the observed reserve and broken lines represent the simulated average value and the 99^{th} quantile

Table 13: Prediction results for the copula-based individual approach

Class	Copula	Exp. value	std. err.	95^{th} quantile	99^{th}
AB	Normal	204,387,329	7,715,968	217,330,973	223,757,558
BI	Normal	297,202,295	19,268,393	329,387,448	343,521,171
	Student	296,935,058	19,466,274	329,213,722	342,230,083
APD	Student	8,001,200	369,584	8,608,782	8,896,145

Note: Observed total amounts are \$206,253,310 for the AB reserve, \$246,404,599 for the BI reserve and \$5,015,952 for the APD reserve.

- closed claims without payment must be modeled separately given their divergent behavior with respect to the coverage (AB/BI/APD) and the reserve stage reached by the development (IBNR/RBNP).

In future research, it would be interesting to add covariates in the three components of the model and to slightly expand the model to allow the prediction of the insurer's payment schedule. Finally, it seems essential to perform a meta-analysis, like the one carried out by Huang et al. (2015), which would compare different individual parametric approaches as well as models based on statistical learning techniques.

Acknowledgements

The authors thank an anonymous referee and the associate editor for useful comments which helped to improve the paper. Roxane Turcotte would like to thank the financial support from the Natural Sciences and Engineering Research Council of Canada and the Fonds de recherche du Québec – Nature et technologies. Hélène Cossette would like to thank the financial support from the Natural Sciences and Engineering Research Council of Canada, under grants 04273. Mathieu Pigeon would like to thank the financial support from the Natural Sciences and Engineering Research Council of Canada, under grants 07094. The authors would also like to thank the industrial partner for supplying the database.

References

Abdallah, A., J.-P. Boucher and H. Cossette (2015). Modeling dependence between loss triangles with hierarchical Archimedeaan copulas. *ASTIN Bulletin*, 45(3): 577–599.

Antonio, K. and R. Plat (2013). Micro-level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*, 2014(7): 649–669.

Avanzi, B., B. Wong and X. Yang (2016). A micro-level claim count model with overdispersion and reporting delays. *Insurance: Mathematics and Economics*, 71: 1–14.

Badescu, A. L., X. S. Lin and D. Tang (2019). A marked Cox model for the number of IBNR claims: Theory. *Insurance: Mathematics and Economics*, 69: 29–37.

Badescu, A. L., X. S. Lin and D. Tang (2019). A marked Cox model for the number of IBNR claims: Estimation and application. *ASTIN Bulletin*, 49(3): 709–739.

Baudry, M. and C. Y. Robert (2019). A machine learning approach for individual claims reserving in insurance. *Applied Stochastic Models in Business and Industry*, 35(5): 1127–1155.

Bühlmann, H., R. Schnieper and E. Straub (1980). Claims reserves in casualty insurance based on a probabilistic model. *Bulletin of the Association of Swiss Actuaries*, 21–45.

Charpentier, A. and M. Pigeon (2016). Macro vs. micro methods in non-life claims reserving (an econometric perspective). *Risks* 4(2): 12.

Côté, M.-P., C. Genest and A. Abdallah (2016). Rank-based methods for modeling dependence between loss triangles. *European Actuarial Journal*, 6(2): 377–408.

De Jong, P. (2012). Modeling dependence between loss triangles. *North American Actuarial Journal*, 16(1): 74–86.

Duval, F. and M. Pigeon (2019). Individual Loss Reserving using a Gradient Boosting-Based Approach. *Risks*, 7(3): 79.

England, P. and R. Verrall (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal*, 8(3): 443–544.

Genest, C. and L. P. Rivest (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, 88(423): 1034–1043.

Genest, C. and A. C. Favre (2007). Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12(4): 347–368.

Hachemeister, C. A. (1980). A stochastic model for loss reserving. *Transactions of the 21st International Congress of Actuaries*, 1: 185–194.

Hiabu, M., C. Margraf, M. D. Martínez-Miranda and J.P. Nielsen (2016). The link between classical reserving and granular reserving through double chain ladder and its extensions, *British Actuarial Journal*, 21(1): 97–116.

Hofert, M., M. Mächler and A. J. McNeil (2012). Likelihood inference for Archimedean copulas in high dimensions under known margins. *Journal of Multivariate Analysis*, 110: 133–150.

Huang, J., C. Qiu and X. Wu (2015). Stochastic loss reserving in discrete time: individual vs. aggregate data models. *Communications in Statistics – Theory and Methods*, 44: 2180–2206.

Joe, H. and J. J. Xu (1996). The estimation method of inference functions for margins for multivariate models. *Technical report No. 166. Department of Statistics, University of British Columbia.*

Joe, H. (2014). *Dependence modeling with copulas.* Chapman and Hall/CRC.

Larsen, C. (2007). An individual claims reserving model. *ASTIN Bulletin*, 37(1): 113–132.

Norberg, R. (1986). A contribution to modelling of IBNR claims. *Scandinavian Actuarial Journal*, 1986: 155–203.

Pigeon, M., K. Antonio and M. Denuit (2013). Individual loss reserving with the multivariate skew normal framework. *ASTIN Bulletin*, 43(3): 399–428.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The Annals of Mathematical Statistics*, 23(3): 470–472.

Shi, P. and E.W. Free (2011). Dependent loss reserving using copulas. *ASTIN Bulletin*, 41(2): 449–486.

Wüthrich, M. V. and M. Merz (2008). *Stochastic claims reserving methods in insurance.* Wiley Finance.

Wüthrich, M. V. (2018). Machine learning in individual claims reserving. *Scandinavian Actuarial Journal*, (6): 465–480.

Zhang, Y. and V. Dukic (2013). Predicting multivariate insurance loss payments under the bayesian copula framework. *Journal of Risk and Insurance*, 80(4): 891–919.

Zhao, X., X. Zhou and J. Wang (2009). Semiparametric model for prediction of individual claim loss reserving. *Insurance: Mathematics and Economics*, 45(1): 1–8.

Zhao, X. and X. Zhou (2010). Applying copula models to individual claim loss reserving methods. *Insurance: Mathematics and Economics*, 46(2): 290–299.

Appendix A. Distributions

The probability density function (pdf) of a lognormal rv is

$$f(x) = \frac{1}{x\theta_2\sqrt{2\pi}} \exp\left(\frac{-\left(\ln(x) - \theta_1\right)^2}{2\theta_2^2}\right) \mathbb{I}_{(0,\infty)}(x),$$

where $\theta_1 \in \mathbb{R}$ and $\theta_2 \in \mathbb{R}_+^*$.

The pdf of a Pareto Type 2 rv is

$$f(x) = k\frac{\theta_1^{\theta_2}}{(\theta_1 + x)^{\theta_2+1}} \mathbb{I}_{(0,\infty)}(x),$$

where $\theta_1 \in \mathbb{R}_+^*$ and $\theta_2 \in \mathbb{R}_+^*$.

Finally, the pdf of a Burr rv is

$$f(x) = \theta_1\theta_2 \times \frac{x^{\theta_2-1}}{\left(1 + x^{\theta_2}\right)^{\theta_1+1}} \mathbb{I}_{(0,\infty)}(x),$$

where $\theta_1 \in \mathbb{R}_+^*$ and $\theta_2 \in \mathbb{R}_+^*$.