I N S T I T U T D E S T A T I S T I Q U E
B I O S T A T I S T I Q U E E T
S C I E N C E S A C T U A R I E L L E S
( I S B A )

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

2014/29

# Multivariate Skew-Normal Individual Excess-of-Loss Reserving

PIGEON, M. and M. DENUIT

**Noname manuscript No.**
(will be inserted by the editor)

# Multivariate Skew-Normal Individual Excess-of-Loss Reserving

**Mathieu Pigeon · Michel Denuit**

the date of receipt and acceptance should be inserted later

**Abstract** In excess-of-loss reinsurance, only claims with incurred loss above a reporting threshold specified in the treaty are notified to the reinsurer. The data available for reinsurance pricing and reserving are, thus, incomplete and failing to account for this limitation may bias the actuarial analysis. The collective approach based on aggregate development triangles cannot deal with this kind of incompleteness. This is why we develop here an individual reserving method combining paid and incurred losses which can cope with any non-proportional reinsurance treaty. Extensive numerical illustrations performed on a motor third party liability reinsurance data set demonstrate the relevance of the approach proposed in this paper.

*Key words and phrases:* Stochastic loss reserving, reinsurance, Multivariate Skew Normal distribution, chain–ladder.

Mathieu Pigeon
Département de mathématiques, Université de Québec à Montréal, Montréal, Canada, E-mail: Mathieu.Pigeon@uclouvain.be

Michel Denuit
Institute of Statistics, Biostatistics and Actuarial Science (ISBA), Université catholique de Louvain, Louvain-la-Neuve, Belgium, E-mail: Michel.Denuit@uclouvain.be

## 1 Introduction and Motivation

Loss reserving in non-life insurance has traditionally been performed on the basis of collective data summarized in run–off triangles. See for instance Wüthrich and Merz [17] and England and Verrall [7] for an overview. However, this collective approach cannot accurately quantify the effect of non-proportional reinsurance treaties, such as excess-of-loss covers commonly found in motor insurance. The impact of a modification of the priority on solvency can only be assessed if data are available per claim and their incomplete nature is explicitly taken into account in the analysis. This is why modelling the development of each claim separately is needed to appropriately evaluate the effect of non-proportional reinsurance.

In order to project future cash flows of reinsurance treaties, we propose a Paid-Incurred Chain (PIC) loss reserving model, designed for individual claims developing in discrete time. Starting from the initial case estimate determined by the insurance analyst, the subsequent stream of payments and adjustments to the reserve is expressed in terms of chain-ladder development factors (or link ratios) and modeled on the log scale with the help of the Multivariate Skew Normal (MSN) distribution.

Individual loss reserving can be traced back to the late 1990s with the development of a mathematical framework in continuous time by Arjas [4] and Norberg [12]. Many extensions have been proposed, e.g. by Larsen [10], Zhao et al [20], Drieskens et al [6], Rosenlund [16], Antonio and Plat [3] and Pigeon et al [14]. Among these previous works, Drieskens et al [6] deal with reinsurance data. Specifically, Drieskens et al [6] present a discrete time model for the development of individual claims in reinsurance based on chain-ladder development factors. Similar to the discrete time approach in Drieskens et al [6] and in contrast to the continuous time approach from Norberg [12] and Antonio and Plat [3], we model in the present paper the reporting delay, the first evaluation delay, the number of payments, the number of adjustments to the reserve and the number of periods between two consecutive movements by discrete random variables. Individual development factors structure the development pattern of each claim. The Multivariate Skew Normal (MSN) distribution is used to model the resulting, dependent development factors at individual claim level. Although Drieskens et al [6] rely on the empirical distribution of each model component, our approach uses parametric or semi-parametric distributions.

Pigeon et al [15] propose a new model combining two sources of information, claim payments and reserve amounts, in a parametric and discrete time framework, extending the collective PIC approach pioneered by Wüthrich and Merz [18] to the case of individual claims data. In the present paper, we adapt this individual PIC reserving model to deal with incomplete reinsurance data. Compared to Pigeon et al [15], the present paper

- explicitly accounts for the incomplete nature of reinsurance data as claims are notified to the reinsurance company only when the incurred loss exceeds a reporting threshold specified in the treaty. Most excess-of-loss reinsurance treaties require that the direct insurer has to report to the reinsurer all claims with incurred loss exceeding a threshold equal to 50 to 75% of the excess-of-loss priority, depending on reinsurance conditions. In case of renewals, reinsurers also require information about claims with incurred loss above a threshold smaller than the priority of the treaty to be priced.

- models the first incurred loss (as in Happ and Wüthrich [19]) as opposed to the first paid amount in Pigeon et al [15]. As a realisation of this variable can be very large, some care is required to select an appropriate model and to combine it with the claim development pattern.

The approach proposed in the present paper is particularly well-suited for internal models under Solvency II. The joint distribution for the individual link factors can be used to simulate the run off of open claims and obtain the sequence of the associated future cash-flows as well as their ultimate value. The number of incurred but not reported (IBNR) claims can also be simulated as well as the development of these claims until final settlement. Thus, we have for each claim a specific development pattern, which allows the actuary to analyze specific risk transfer mechanisms such as the impact of an excess-of-loss coverage.

Notice that IBNR claims in this particular setting are those claims which have not yet reached the priority of the reinsurance treaty. Therefore, IBNR claims must here be understood as claims which are not yet large enough, adopting the point of view of an excess-of-loss reinsurer. Even if this paper adopts the reinsurer's point of view, cedants may also use this method in order to correctly assess their reserve risk and the appropriateness of the reinsurance program in force.

The remainder of this paper is structured as follows. The model is described in Section 2. Section 3 is devoted to a case study on a portfolio of motor reinsurance claims and Section 4 concludes.

## 2 Individual Loss Reserving Model

### 2.1 Individual paid-incurred model

In an excess-of-loss treaty, the claim is notified to the reinsurer when the incurred loss exceeds some reporting threshold specified by the reinsurance conditions. This threshold is typically in the range 50-75% of the priority of the treaty. Recall that the incurred loss comprises the payments made so far as well as the amount reserved for this claim (that is, the forecast of all future payments related to this claim by the claim adjuster). Since the condition to be notified to the reinsurer is based on incurred amounts rather than paid ones, we introduce here a slightly modified version of the individual Paid-Incurred Chain model proposed by Pigeon et al [15]. Specifically, we construct successive paid and incurred amounts from the initial case estimate rather than from the initial paid amount, as was the case in the original model.

Consider a reinsurance portfolio observed during $I$ calendar years. Time is here supposed to be measured in calendar years but other units can be used, such as semesters, quarters or months. Let $K_i$, $i = 1, \ldots, I$, be the number of claims reported by the cedants to the reinsurer in year $i$. These claims may have been reported much before to the direct insurer but their incurred amount reaches for the first time the reporting threshold during period $i$. Index $k$ corresponds to the $k$th claim in the data set.

Let the random variable $T_k$ represent the reporting delay for claim $k$, i.e. the number of years elapsed between claim occurrence and its notification to the reinsurer. Here, $T_k$ aggregates the reporting delay of the claim to the direct insurer as well as the time needed for the incurred loss to exceed for the first time the reporting threshold specified in the treaty. This claim then produces a bivariate random cluster of events: a sequence of claim payments and a sequence of adjustments to the incurred amount. Once a claim has been reported, it stays in the data set even if the associated incurred loss falls below the priority at some point in the future. Let us mention at this stage that the direct insurer sometimes provides the reinsurer with all the information at its disposal concerning the claims exceeding the reporting threshold. In the present paper, we adopt the reinsurer's point of view and we consider that the first incurred loss is the one established by the reinsurance claim analysts when the claim is notified by the ceding company. The analysts scrutinize the claims based on the information provided by the cedant and evaluate the reinsurer's liability with respect to this case.

Two sources of information are combined in our model: payment data and incurred loss data. The following random variables are needed to represent the individual claim trajectory until final settlement:

$Q_k^I =$ number of periods between the reporting at time $T_k$ and the initial case estimate;
$Q_k^P =$ number of periods between the reporting at time $T_k$ and the first payment;
$U_k^I =$ number of periods with non-zero adjustment to the reserve, after the period in which the initial case estimate is set.
$U_k^P =$ number of periods with non-zero payments.
$N_{kj}^I =$ number of periods between adjustments $j$ and $j+1$; here, $N_{k,U_{k+1}^I}^I$ is the number of periods between the last adjustment and the settlement of the claim;
$N_{kj}^P =$ number of periods between payments $j$ and $j+1$; here, $N_{k,U_{k+1}^P}^P$ is the number of periods between the last payment and the settlement of the claim.

We assume that all these random variables are independent and non-negative.

From the reinsurer point of view, the first payment is denoted as $P_{k,1}$ and consecutive cumulative payments are $P_{k,2},\ldots,P_{k,U_k^P+1}$. Notice that $P_{k,j}$ refers to calendar year $T_k+Q_k^P+\sum_{i=1}^{j-1} N_{ki}^P$, with the convention that the empty sum is zero as only periods with strictly positive payments are considered. By construction, $P_{k,j}$ differs from $P_{k,j-1}$. The first incurred loss is denoted with $I_{k,1}$ and consecutive incurred losses are $I_{k,2},\ldots,I_{k,U_k^I+1}$. We recall that the incurred loss amount is the sum of paid and outstanding reserve amounts. By construction, $I_{k,j+1}$ differs from $I_{k,j}$. As in the collective PIC approach proposed by Wüthrich and Merz [18], we reconcile both sources of information for each claim $k$ with the help of the condition

$$I_{k,U_k^I+1} = P_{k,U_k^P+1}. \tag{2.1}$$

We present an example of individual claim development trajectory in Figure 2.1. We removed the reference to the claim index $k$ for a simpler presentation. The upper part represents the viewpoint of the insurer (marked with stars) and the lower part represents the viewpoint of the excess-of-loss reinsurer. In this paper, we adopt the reinsurer's point of view. In this example, we suppose that the claim exceeds the priority of the treaty at time $t_4^*$. It will be considered for the first time in the portfolio at time $t_1 = t_4^*$ with an initial payment of $P_1 = P_1^*$ and an initial incurred amount of $I_1$. This initial estimate is established by the reinsurance company based on its own experts. The reinsurer may trust the analysis

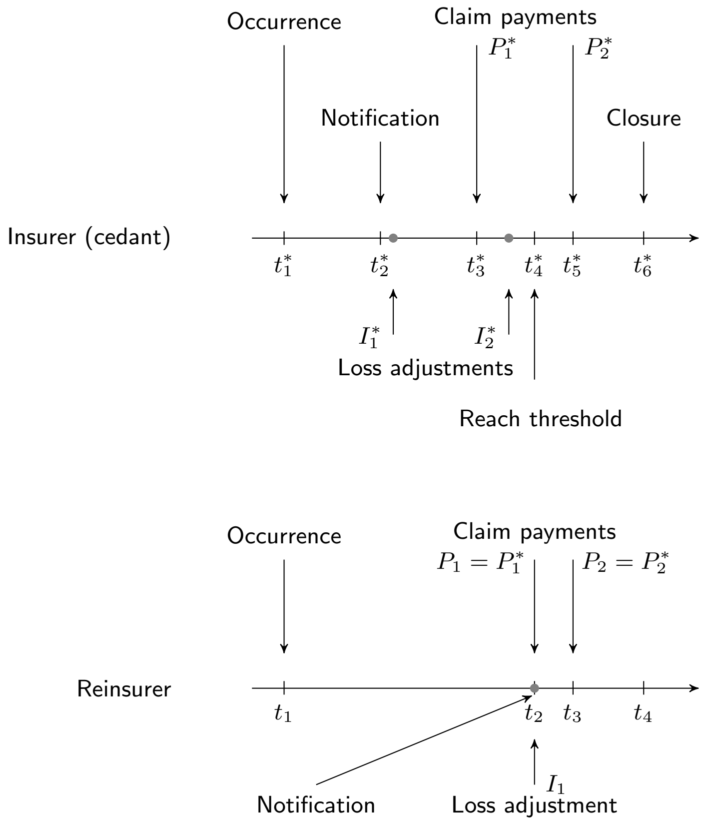


**Fig. 2.1** *Time line representing an individual claim development trajectory.*

conducted by the ceding company and simply set $I_j^* = I_j$ but it may also depart from this evaluation based on its expert knowledge. The reporting delay is the number of periods between $t_2$ and $t_1$ and the claim is closed at time $t_4 = t_6^*$ with 2 payments (at $t_2 = t_4^*$ and $t_3 = t_5^*$) and 1 claim adjustment (at time $t_2 = t_4^*$).

Given $U_k^P = u_k^P$ and $U_k^I = u_k^I$, we define adjustment-to-adjustment link ratios $\gamma_j^{(k)}$ for claim $k$ as

$$I_{k,j} = I_{k,j-1} \cdot \gamma_{j-1}^{(k)} \qquad j = 2, \dots, u_k^I + 1, \tag{2.2}$$

with initial value $I_{k,1}$. Similarly, we define payment-to-payment link ratios $\lambda_j^{(k)}$ for claim $k$ as

$$P_{k,j-1} = P_{k,j} \cdot \lambda_{j-1}^{(k)} \qquad j = 2, \dots, u_k^P + 1, \tag{2.3}$$

with initial value $P_{k,u_k^P+1}$. Notice that $j$ refers to an event (payment or adjustment) instead of development period in both (2.2) and (2.3). Formulas (2.2) and (2.3) are similar to the chain-ladder method. However, with chain-ladder, the index $j$ is for development period. Using a development-to-development model (as chain-ladder does) with individual claims can be problematic because the length of the development pattern differs among claims, and many development factors will have value one, which is not the case under the payment-to-payment or adjustment-to-adjustment approach (in discrete time) adopted in the present paper.

Define $u^P_{\max} = \max_k u^P_k$, $u^I_{\max} = \max_k u^I_k$, $M = u^P_{\max} + u^I_{\max} + 1$, and the random vector

$$\boldsymbol{\Omega} = \left(\log I_1, \log\gamma_1, \ldots, \log\gamma_{u^I_{\max}}, \log\lambda_{u^P_{\max}}, \ldots, \log\lambda_1\right)' = (\Omega_1, \ldots, \Omega_M)'. \quad (2.4)$$

Notice that $\boldsymbol{\Omega}$ comprises all individual claim information recorded on the log scale. The initial incurred loss $I_1$ is developed to the ultimate by the link ratios $\gamma_j$ according to (2.2) and the stream of payments is then obtained backward by means of the $\lambda_j$ according to (2.3). This explains the order of the elements appearing in $\boldsymbol{\Omega}$ defined in (2.4). We also define a location vector $\boldsymbol{\mu} = (\mu_1 \ \ldots \ \mu_M)$, a scale matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2_{11} & \sigma_{12} & \ldots & \sigma_{1M} \\ \sigma_{21} & \sigma^2_{22} & \ldots & \sigma_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M1} & \sigma_{M2} & \ldots & \sigma^2_{MM} \end{pmatrix}$$

and a shape vector $\boldsymbol{\Delta} = (\Delta_1 \ \ldots \ \Delta_M)$. We assume that the random vector $\boldsymbol{\Omega}$ in (2.4) obeys the Multivariate Skew Normal (MSN) distribution with location vector $\boldsymbol{\mu}$, scale matrix $\boldsymbol{\Sigma}$ and shape vector $\boldsymbol{\Delta}$, i.e. its joint probability density function $f_{\boldsymbol{\Omega}}$ is given by

$$f_{\boldsymbol{\Omega}}(\boldsymbol{\omega}) = \frac{2^M}{\det(\boldsymbol{\Sigma})^{1/2}} \cdot \left(\phi\left(\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\omega}-\boldsymbol{\mu})\right)\prod_{j=1}^{M}\Phi\left(\Delta_j \mathbf{e}'_j \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\omega}-\boldsymbol{\mu})\right)\right), \quad (2.5)$$

where $\mathbf{e}'_j$ are the basis vectors of $\mathbb{R}^M$ and where $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density function and cumulative distribution function of the standard Normal distribution, respectively. Notice that the scale matrix $\boldsymbol{\Sigma}$ is not the usual variance-covariance matrix, contrarily to the Multivariate Normal distribution. Formula (2.5) is generally valid whatever the square root considered. In the remainder of the paper, we consider that $\boldsymbol{\Sigma}^{1/2}$ is the square root of matrix $\boldsymbol{\Sigma} = \left(\boldsymbol{\Sigma}^{1/2}\right)\left(\boldsymbol{\Sigma}^{1/2}\right)'$ by Cholesky decomposition.

For an ordered subset $v = \{v_1, \ldots, v_{|v|}\}$, $|v| \leq M$ of $\{1, \ldots, M\}$, we construct the random vector $\boldsymbol{\Omega}_v$ by selecting elements from $\boldsymbol{\Omega}$ corresponding to the subscripts collected in $v$, i.e. $\boldsymbol{\Omega}_v := (\Omega_{v_1}, \Omega_{v_2}, \ldots, \Omega_{v_{|v|}})$. Similarly, we construct the corresponding location vector $\boldsymbol{\mu}_v = (\mu_{v_1}, \ldots, \mu_{v_{|v|}})$, the shape vector $\boldsymbol{\Delta}_v = (\Delta_{v_1}, \ldots, \Delta_{v_{|v|}})$ and the scale matrix $\boldsymbol{\Sigma}_v$

$$\boldsymbol{\Sigma}_v = \begin{pmatrix} \sigma^2_{v_1v_1} & \sigma_{v_1v_2} & \ldots & \sigma_{v_1v_{|v|}} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{v_{|v|}v_1} & \sigma_{v_{|v|}v_2} & \ldots & \sigma^2_{v_{|v|}v_{|v|}} \end{pmatrix}.$$

We assume that the random vector $\boldsymbol{\Omega}_v$ follows the Multivariate Skew Normal distribution with location vector $\boldsymbol{\mu}_v$, scale matrix $\boldsymbol{\Sigma}_v$ and shape vector $\boldsymbol{\Delta}_v$. The individual development of claim $k$ is then represented by

$$\boldsymbol{\Omega}_{v_k} = \left(\log I_{k,1}, \log\gamma^{(k)}_1, \ldots, \log\gamma^{(k)}_{u^I_k}, \log\lambda^{(k)}_{u^P_k}, \ldots, \log\lambda^{(k)}_1\right)', \quad (2.6)$$

where subscripts $v_k = \{1, \ldots, u_k^I + 1, u_{\max}^I + 1 + \left(u_{\max}^P - u_k^P\right), \ldots, M\}$ correspond to vector (2.4). The random vector $\boldsymbol{\Omega}_{v_k}$ in (2.6) contains the initial case estimate together with the link ratios to form the individual development pattern of claim $k$.

## 2.2 Incomplete data

Often, reinsurance data consist in claims for which information is available only after the incurred amount has exceeded a certain threshold specified in the treaty. Typically, the reinsurer receives information from its ceding company only if the incurred amount exceeds the reporting threshold equal to a percentage (in the range 50-75%) of the priority of the reinsurance treaty.

We consider a portfolio with claims from $c$ different companies. Each of them may have a different reporting threshold $\xi_i$, $i = 1, \ldots, c$ and we denote as $\xi^{(k)}$ the threshold of the company from which claim $k$ originates. As $\Omega_1$ is the first component of a MSN random vector, it obeys a univariate Skew-Normal (SN) distribution (see Pigeon et al [15] for more details) with parameters $\left(\mu_1, \sigma_{11}^2, \Delta_1\right)$. The associated probability density function is given by

$$f_{\Omega_1}(y) = \left(\frac{2}{\sigma_{11}}\right) \phi\left(\frac{y - \mu_1}{\sigma_{11}}\right) \Phi\left(\Delta_1 \left(\frac{y - \mu_1}{\sigma_{11}}\right)\right), \qquad -\infty < y < \infty. \tag{2.7}$$

Given $U_k^P = u_k^P$ and $U_k^I = u_k^I$, let us partition the $|v_k|$-dimensional random vector $\boldsymbol{\Omega}_{v_k}$ following the MSN distribution with location parameter $\boldsymbol{\mu}_{v_k}$, scale parameter $\boldsymbol{\Sigma}_{v_k}$ and shape parameter $\boldsymbol{\Delta}_{v_k}$ into

$$\boldsymbol{\Omega}_{v_k} = \begin{pmatrix} \Omega_1 \\ \boldsymbol{\Omega}_2 \end{pmatrix}, \quad \boldsymbol{\mu}_{v_k} = \begin{pmatrix} \mu_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma}_{v_k}^{1/2} = \begin{pmatrix} \sigma_{11} & \mathbf{0} \\ \boldsymbol{\Sigma}_{21}^{1/2} & \boldsymbol{\Sigma}_{22}^{1/2} \end{pmatrix}, \quad \boldsymbol{\Delta}_{v_k} = \begin{pmatrix} \Delta_1 \\ \boldsymbol{\Delta}_2 \end{pmatrix}.$$

The conditional probability density function of $\boldsymbol{\Omega}_{v_k}$ given $\Omega_1 \geq \log \xi^{(k)}$ is

$$f\left(\boldsymbol{\omega}_{v_k} | \Omega_1 \geq \log \xi^{(k)}\right) = \mathbb{I}[\Omega_1 \geq \log \xi^{(k)}] \cdot \left( \frac{\dfrac{2^{|v_k|}}{\det\left(\boldsymbol{\Sigma}_{v_k}\right)^{1/2}} \cdot \phi\left(\boldsymbol{\Sigma}_{v_k}^{-1/2}\left(\boldsymbol{\omega}_{v_k} - \boldsymbol{\mu}_{v_k}\right)\right) \displaystyle\prod_{j=1}^{|v_k|} \Phi\left(\Delta_{[v_k, j]} \boldsymbol{e}_j' \boldsymbol{\Sigma}_{v_k}^{-1/2}\left(\boldsymbol{\omega}_{v_k} - \boldsymbol{\mu}_{v_k}\right)\right)}{\displaystyle\int_{\log \xi^{(k)}}^{\infty} \left(\frac{2}{\sigma_{11}}\right) \phi\left(\frac{y - \mu_1}{\sigma_{11}}\right) \Phi\left(\Delta_1 \left(\frac{y - \mu_1}{\sigma_{11}}\right)\right) dy} \right). \tag{2.8}$$

Expected value $E\left[\boldsymbol{\Omega}_{v_k} | \Omega_1 \geq \log \xi^{(k)}\right]$, as well as higher moments, can be obtained with a numerical procedure.

## 3 Case Study

### 3.1 Presentation of the dataset

The individual loss reserving model described in the preceding section is now applied to a motor third-party liability reinsurance data set gathering information about 19 direct insurance companies operating in the EU. As the company-specific priority often evolves over time, we display the average reporting threshold per company in Table 3.1, together with the corresponding number of reinsurance claims comprised in the data set. The observation period starts in 1991 and ends in 2010. Our evaluation date is supposed to be the first of January 2011. All the amounts in the data set have been corrected to reflect costs in calendar year 2011. To do so, inflation has been taken into account, as well as super-inflation affecting large claims in motor third party liability.

| Company | Threshold (average) | Number of claims |
|---|---|---|
| C1 | 154,508 | 129 |
| C2 | 154,508 | 24 |
| C3 | 432,624 | 100 |
| C4 | 154,508 | 349 |
| C5 | 156,885 | 654 |
| C6 | 855,739 | 154 |
| C7 | 156,885 | 162 |
| C8 | 497,992 | 226 |
| C9 | 97,459 | 130 |
| C10 | 233,546 | 581 |
| C11 | 185,410 | 847 |
| C12 | 285,246 | 145 |
| C13 | 30,902 | 115 |
| C14 | 231,763 | 125 |
| C15 | 130,738 | 40 |
| C16 | 463,525 | 170 |
| C17 | 216,312 | 38 |
| C18 | 231,763 | 79 |
| C19 | 231,763 | 28 |

**Table 3.1** *Average reporting threshold and corresponding claim numbers for each insurance company included in the data set.*

We have at our disposal a total of 4,096 claims reported to the reinsurer, represented by 15,813 entries in the data set. Recall that a claim is included in the data base of the reinsurance company as soon as the incurred loss, i.e. the sum of the outstanding reserve and the total amount paid so far for the claim, exceeds the reporting threshold described in Table 3.1. At the evaluation date, there are 2,715 open claims. Table 3.2 displays descriptive statistics for closed claims, only.

As mentioned earlier, time is measured in calendar years and the data are transformed accordingly. Specifically, all the payments relating to the same development year are aggregated in a single data point, and the same procedure is applied to incurred loss data. As the data set sometimes contains some negligible movements in paid and/or incurred amounts, i.e. with very low amplitude in comparison with amounts under consideration, the data have been smoothed in order to concentrate the analysis on relevant movement in paid or incurred amounts. Figure 3.1 illustrates the process where events representing less than 10% of the reporting threshold are aggregated with the next ones until this limit is exceeded. The same analysis has been conducted by replacing 10% with 5% and 15%, without significantly affecting the results. Again, let us mention that calculations may be performed more frequently, if needed, such as every semester, quarter or month. The procedure is similar to the one described here except that the aggregation of the data now proceeds along this finer time grid.

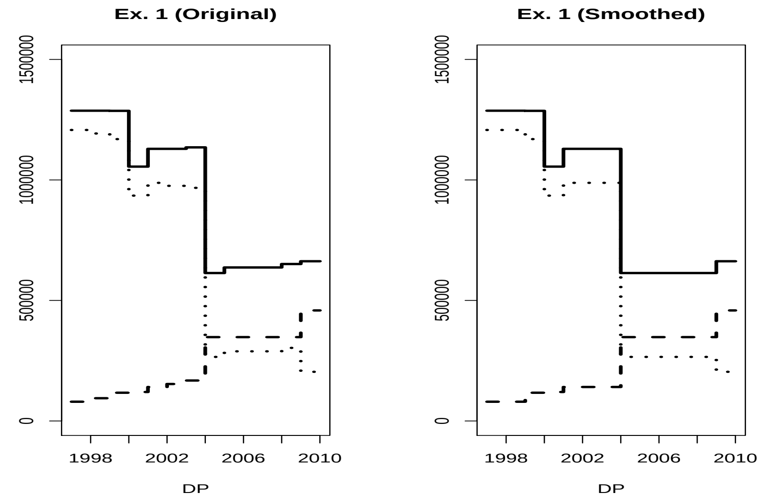


**Fig. 3.1** *Example of smoothed development with the incurred amount (solid line), the total paid amount (broken line) and the total outstanding claim reserve (dotted line).*

| Variables | Mean | Median | s.e. | Minimum | Maximum | Number of Observations |
|---|---|---|---|---|---|---|
| $P_1$ | 67,310 | 37,600 | 105,291 | 5 | 1,070,000 | 1,193 |
| $\lambda_1$ | 0.508 | 0.519 | 0.285 | 0.000 | 3.332 | 855 |
| $\lambda_2$ | 0.640 | 0.670 | 0.239 | 0.014 | 1.394 | 504 |
| $\lambda_3$ | 0.687 | 0.740 | 0.239 | 0.100 | 1.632 | 252 |
| $\lambda_4$ | 0.764 | 0.821 | 0.218 | 0.189 | 1.397 | 121 |
| $I_1$ | 394,900 | 310,900 | 345,469 | 31,280 | 4,265,000 | 1,380 |
| $\gamma_1$ | 1.031 | 0.994 | 0.870 | 0.001 | 20.46 | 1,152 |
| $\gamma_2$ | 1.017 | 0.991 | 0.974 | 0.005 | 24,33 | 754 |
| $\gamma_3$ | 0.954 | 0.963 | 0.380 | 0.006 | 3.28 | 407 |
| $\gamma_4$ | 0.973 | 0.972 | 0.343 | 0.244 | 2.84 | 180 |

**Table 3.2** *Descriptive statistics for closed claims: first payment $P_1$, first incurred amount $I_1$, payment development factors $\lambda_j$, $j = 1,2,3,4$, and incurred development factors $\gamma_j$, $j = 1,2,3,4$.*

## 3.2 Estimation results for $T$, $U^P$ and $U^I$

For the time components of the model (random variables $T$, $U^P$ and $U^I$), we investigate the use of a discrete Geometric distribution combined with some de-

| Reporting delay | # Claim payments | # Incurred loss adjustments |
|---|---|---|
| $p = 1$ | $p = 0$ | $p = 1$ |
| $(T; \boldsymbol{\alpha})$ | $(U^P; \boldsymbol{\beta^P})$ | $(U^I; \boldsymbol{\beta^I})$ |
| (s.e.) | (s.e.) | (s.e.) |
| 0.4400 | 0.1185 | 0.0712 |
| (0.008) | (0.006) | (0.005) |
| 0.3102 | 0.1889 | 0.1496 |
| (0.007) | (0.006) | (0.007) |
| 0.3147 | | 0.2258 |
| (0.062) | | (0.008) |

**Table 3.3** *Parameter estimates and standard errors for $\boldsymbol{\alpha}$, $\boldsymbol{\beta}^P$ and $\boldsymbol{\beta}^I$ entering* (3.1).

generate components. Specifically, we assume that the reporting delay of each claim has probability mass function

$$\Pr[T = t] = \begin{cases} \alpha_0, & \text{if } t = 0 \\ \alpha_1, & \text{if } t = 1 \\ \vdots & \\ \alpha_p, & \text{if } t = p \\ \left(1 - \sum_{s=0}^{p} \alpha_s\right) \alpha_{p+1}^{t-p} \left(1 - \alpha_{p+1}\right), & \text{if } t \geq p + 1. \end{cases} \tag{3.1}$$

In (3.1), empirical estimates are used as long as enough data are available to estimate the probability masses $\alpha_0, \ldots, \alpha_p$. The Geometric distribution then plays the role of a tail factor extrapolating the claim behavior after horizon $p$. We consider similar specifications for $U^P$ and $U^I$ with parameters $\boldsymbol{\beta}^P$ and $\boldsymbol{\beta}^I$ replacing $\boldsymbol{\alpha}$. Table 3.3 displays the estimated parameters for the distribution of $T$, $U^P$ and $U^I$. The optimal value of $p$ has been selected by means of BIC.

The number of claims $K_i$ reported to the reinsurer in calendar year $i$ is assumed to follow a Poisson distribution with mean value $\theta \cdot w(i)$, where $w(i)$ is the exposure registered for period $i$, based on premium income. Since the reinsurer only observes claims reaching the reporting threshold, this Poisson distribution is thinned in the following way

$$\theta \cdot w(i) \cdot \Pr\left(T \leq r_i^\star - 1\right) \tag{3.2}$$

where $r_i^\star$ is the number of periods separating evaluation date and occurrence period $i$. The estimated intensity $\theta$ entering (3.2) is $\hat{\theta} = 0.1467$ (with standard error 0.0025).

### 3.3 Covariance structure

We visualize the empirical correlation matrix corresponding to $\boldsymbol{\Omega}$ with a heat map in Figure 3.2. As the MSN covariance structure is given by

$$\mathrm{Cov}[\boldsymbol{\Omega}] = \boldsymbol{\Sigma}^{1/2} \left( \mathbb{I}_8 - \left(\frac{2}{\pi}\right) \mathrm{diag}\left( \frac{\Delta_1^2}{1 + \Delta_1^2}, \ldots, \frac{\Delta_8^2}{1 + \Delta_8^2} \right) \right) \left(\boldsymbol{\Sigma}^{1/2}\right)',$$

where $\mathbb{I}$ is the identity matrix. Figure 3.2 suggests to test for the following three covariance structures, henceforth denoted as $\boldsymbol{\Sigma}_A$, $\boldsymbol{\Sigma}_B$ and $\boldsymbol{\Sigma}_C$, respectively, for the scale matrix $\boldsymbol{\Sigma}$:

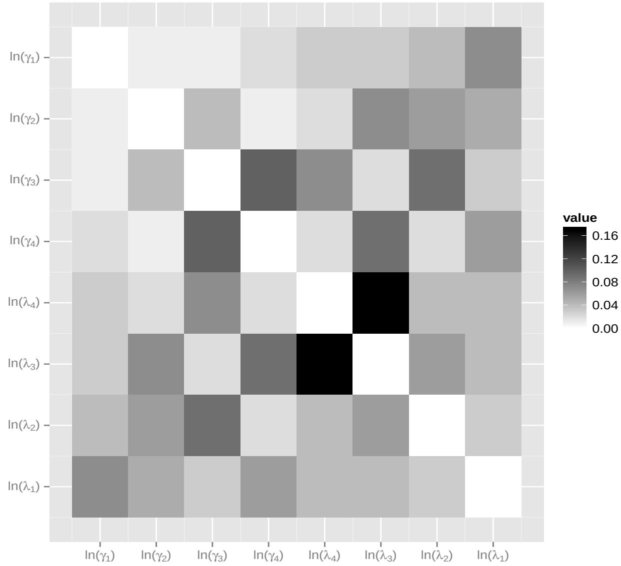


**Fig. 3.2** *Empirical correlation matrix heat map (in absolute value). For the sake of presentation, the unit values along the main diagonal were removed and the corresponding squares have been left blank.*

(i) $\boldsymbol{\Sigma}_A$ has diagonal structure with no correlation between link ratios. Thus, it corresponds to mutual independence between the first incurred loss and the subsequent development factors and is given by

$$\boldsymbol{\Sigma}_A = \begin{pmatrix} \sigma_{11}^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_{22}^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma_{33}^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_{88}^2 \end{pmatrix}.$$

This covariance structure serves as a benchmark to test whether some correlations are significant.

(ii) $\boldsymbol{\Sigma}_B$ is of Toeplitz form and allows for correlation between neighboring elements of the random vector $\boldsymbol{\Omega}$, i.e. between successive development factors as well as between the first incurred loss and its development factor

$$\boldsymbol{\Sigma}_B = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12} & 0 & \dots & 0 & 0 \\ \sigma_{12} & \sigma_{22}^2 & \sigma_{23} & \dots & 0 & 0 \\ 0 & \sigma_{23} & \sigma_{33}^2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_{78} & \sigma_{88}^2 \end{pmatrix}.$$

(iii) $\boldsymbol{\Sigma}_C$ authorizes correlation between corresponding elements of $\boldsymbol{\Omega}$, i.e. between $\log \lambda_i$ and $\log \gamma_i$, $i = 1, 2, 3, 4$. Thus, corresponding loss and payment development factors may be correlated, as well as the first incurred loss and the first payment link ratio, i.e.

$$\boldsymbol{\Sigma}_C = \begin{pmatrix} \sigma_{11}^2 & 0 & \dots & \dots & \dots & \dots & \sigma_{18} \\ 0 & \sigma_{22}^2 & \dots & \dots & \dots & \sigma_{27} & 0 \\ 0 & 0 & \sigma_{33}^2 & \dots & \sigma_{36} & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \sigma_{81} & 0 & \dots & \dots & \dots & \dots & \sigma_{88}^2 \end{pmatrix}.$$

For each structure for the scale matrix ($\boldsymbol{\Sigma}_A$, $\boldsymbol{\Sigma}_B$ and $\boldsymbol{\Sigma}_C$), we consider a model designed for complete data, i.e. based on the erroneous assumption $\xi = 0$, and with claim–specific reporting thresholds $\xi_k$ henceforth denoted as $\xi \neq 0$. We present in Table 3.4, the number of parameters, the maximum log–likelihood and the $p$-value of likelihood–ratio tests considering model $A$ and $B$ as the null model. For the preferred models (with $\boldsymbol{\Sigma}_B$ in both cases), we present estimation results in Table 3.5.

| Model | | # of par. | $-ll$ | $p$-value ($\boldsymbol{\Sigma}_A$) | $p$-value ($\boldsymbol{\Sigma}_B$) |
|---|---|---|---|---|---|
| | $\boldsymbol{\Sigma}_A$ | 18 | 11,529 | – | – |
| $\xi = 0$ | $\boldsymbol{\Sigma}_B$ | 26 | 11,495 | $< 0.001$ | – |
| | $\boldsymbol{\Sigma}_C$ | 22 | 11,514 | $< 0.001$ | $< 0.001$ |
| | $\boldsymbol{\Sigma}_A$ | 18 | 13,939 | – | – |
| $\xi \neq 0$ | $\boldsymbol{\Sigma}_B$ | 26 | 13,904 | $< 0.001$ | – |
| | $\boldsymbol{\Sigma}_C$ | 22 | 13,931 | 0.003 | $< 0.001$ |

**Table 3.4** *Comparison of covariance structures for different models in the cases $\xi \neq 0$ and $\xi = 0$.*

| $\xi = 0$ | | | $\xi \neq 0$ | | |
|---|---|---|---|---|---|
| Loc. par. ($\boldsymbol{\mu}$) (s.e.) | Sc. par. ($\boldsymbol{\Sigma}$) | Sh. par. ($\boldsymbol{\Delta}$) | Loc. par. ($\boldsymbol{\mu}$) (s.e.) | Sc. par. ($\boldsymbol{\Sigma}$) | Sh. par. ($\boldsymbol{\Delta}$) |
| $\mu_1 = 13.074$ | $\sigma_{11}^2 = 0.89$ | $\Delta_1 = -0.33$ | $\mu_1 = 14.046$ | $\sigma_{11}^2 = 0.95$ | $\Delta_1 = -0.33$ |
| (0.0003) | $\sigma_{22}^2 = 0.82$ | $\Delta_2 = -2.38$ | (0.0004) | $\sigma_{22}^2 = 0.82$ | $\Delta_2 = -2.38$ |
| $\mu_2 = 0.569$ | $\sigma_{33}^2 = 0.45$ | $\Delta_3 = -1.87$ | $\mu_2 = 0.497$ | $\sigma_{33}^2 = 0.44$ | $\Delta_3 = -1.87$ |
| (0.0002) | $\sigma_{44}^2 = 0.33$ | $\Delta_4 = -1.73$ | (0.0003) | $\sigma_{44}^2 = 0.33$ | $\Delta_4 = -1.73$ |
| $\mu_3 = 0.415$ | $\sigma_{55}^2 = 0.27$ | $\Delta_5 = 1.45$ | $\mu_3 = 0.388$ | $\sigma_{55}^2 = 0.28$ | $\Delta_5 = 1.45$ |
| (0.0003) | $\sigma_{66}^2 = 0.25$ | $\Delta_6 = -4.57$ | (0.0002) | $\sigma_{66}^2 = 0.25$ | $\Delta_6 = -4.57$ |
| $\mu_4 = 0.325$ | $\sigma_{77}^2 = 0.40$ | $\Delta_7 = -4.78$ | $\mu_4 = 0.358$ | $\sigma_{77}^2 = 0.39$ | $\Delta_7 = -4.78$ |
| (0.0004) | $\sigma_{88}^2 = 0.57$ | $\Delta_8 = -8.62$ | (0.0003) | $\sigma_{88}^2 = 0.59$ | $\Delta_8 = -8.62$ |
| $\mu_5 = -0.352$ | $\sigma_{99}^2 = 1.63$ | $\Delta_9 = -8.10$ | $\mu_5 = -0.382$ | $\sigma_{99}^2 = 1.63$ | $\Delta_9 = -8.10$ |
| (0.0008) | $\sigma_{21} = -0.06$ | | (0.0008) | $\sigma_{21} = -0.10$ | |
| $\mu_6 = 0.037$ | $\sigma_{32} = -0.02$ | | $\mu_6 = 0.032$ | $\sigma_{32} = -0.05$ | |
| (0.0001) | $\sigma_{43} = -0.01$ | | (0.0004) | $\sigma_{43} = 0.02$ | |
| $\mu_7 = 0.031$ | $\sigma_{54} = -0.04$ | | $\mu_7 = -0.005$ | $\sigma_{54} = -0.05$ | |
| (0.0001) | $\sigma_{65} = -0.02$ | | (0.0003) | $\sigma_{65} = -0.02$ | |
| $\mu_8 = -0.001$ | $\sigma_{76} = -0.03$ | | $\mu_8 = 0.007$ | $\sigma_{76} = -0.03$ | |
| (0.0003) | $\sigma_{87} = -0.01$ | | (0.0001) | $\sigma_{87} = -0.03$ | |
| $\mu_9 = 0.014$ | $\sigma_{98} = -0.02$ | | $\mu_9 = 0.004$ | $\sigma_{98} = -0.01$ | |
| (0.0004) | | | (0.0001) | | |

**Table 3.5** *Estimation results and standard errors for logarithms of development factors in the cases $\xi \neq 0$ and $\xi = 0$.*

## 3.4 Prediction results

For both models, Table 3.6 contains expected values, standard deviations as well as 95% and 99.5% Value-at-Risk (VaR) and Conditional Tail Expectation (CTE) measures for the ultimate loss experienced by the reinsurer. We clearly see that disregarding the limitation present in the data (i.e. putting $\xi = 0$) leads to a lower evaluation of the reserve amounts. When compared to the collective approach which also neglects the incomplete nature of the data, we see that the expected ultimate loss of the portfolio is underestimated (with the Chain-Ladder method applied only on paid amounts) or almost triples (with the Munich Chain-Ladder method). This shows the necessity of the approach developed in the present paper. It is also interesting to consider the sum of case estimates determined by claim adjusters as listed in the data set, which amounts to 1,335,980,984, a value quite close to the 99.5% VaR reported in Table 3.6 for the model with $\xi \neq 0$. We display in Figure 3.3 the probability density function of the ultimate loss obtained in the cases $\xi = 0$ and $\xi \neq 0$. We note that the allowance for $\xi \neq 0$ in the model leads to a horizontal shift in the distribution but does not affect its shape.

| Model | | Exp. Value | s.d. | Risk Measures | |
|---|---|---|---|---|---|
| | | | | (0.95) | (0.995) |
| MSN<br>($\xi = 0$) | $\boldsymbol{\Sigma}_B$ | 925,431,788 | 55,611,430 | $VaR$ = 1,017,999,987<br>$CTE$ = 1,048,118,367 | $VaR$ = 1,084,724,881<br>$CTE$ = 1,110,869,714 |
| MSN<br>($\xi \neq 0$) | $\boldsymbol{\Sigma}_B$ | 1,015,528,427 | 59,950,009 | $VaR$ = 1,118,273,629<br>$CTE$ = 1,148,184,977 | $VaR$ = 1,178,920,869<br>$CTE$ = 1,226,479,449 |
| C-L (paid) | | 877,357,349 | 92,521,517 | $VaR$ = 1,038,626,764<br>$CTE$ = 1,095,218,538 | $VaR$ = 1,168,329,326<br>$CTE$ = 1,232,096,313 |
| Munich C-L | | 2,304,960,343 | – | – | – |

**Table 3.6** *Prediction results based on 10,000 simulations*

## 3.5 Alternative modelling for the first incurred amount

### *3.5.1 Combining MSN development factors with initial incurred loss*

Until now, it has been assumed that $I_1$ obeyed the log-SN distribution, i.e. the probability density function of $\log I_1$ is (2.7). By using the Von Mises condition [5], it can be shown that the univariate log-SN distribution for $I_1$ belongs to the maximum domain of attraction of the Gumbel distribution. The associated tail is thus rather thin. If the initial case estimate data do not support this specification for $I_1$, we propose a strategy to allow for different distributional assumptions such as the Lognormal, Pareto, Generalized Pareto and some composite Lognormal–Pareto models [13] for the first incurred amount $I_1$. This approach may also be relevant for the individual PIC reserving model developed by Pigeon et al [15] but as this approach models the first paid loss amount $P_1$ instead of $I_1$, the log-SN specification is often appropriate. Table 3.2 clearly shows that the distribution of $I_1$ is more dispersed and exhibits a heavier tail than $P_1$.

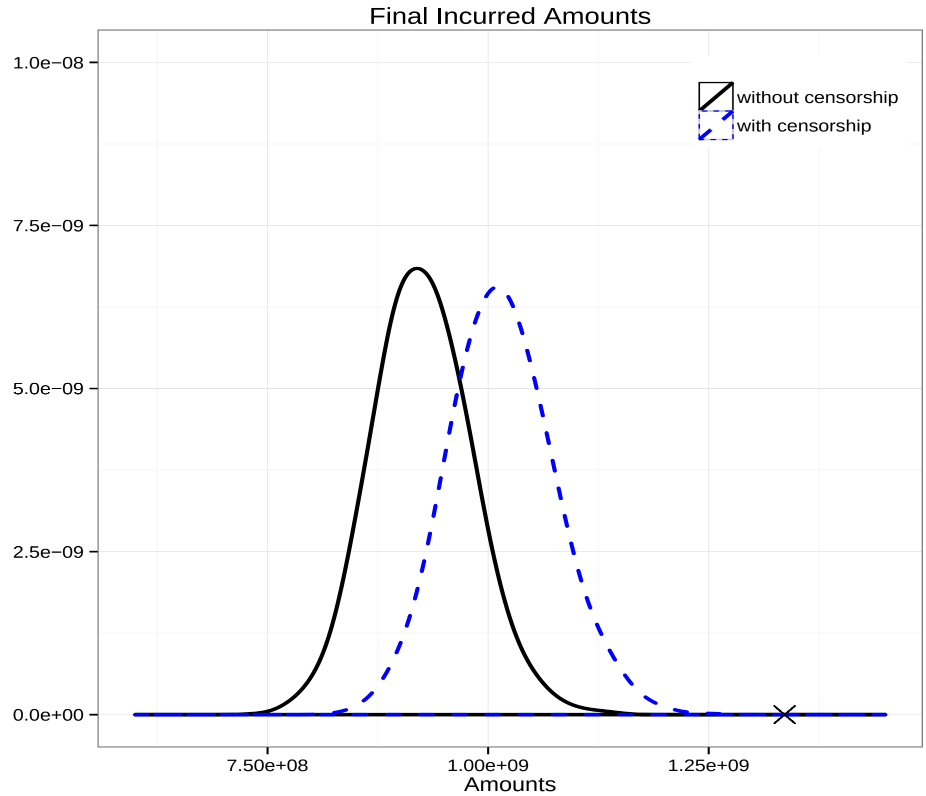


**Fig. 3.3** *Probability density function of the ultimate loss for $\xi \neq 0$ and $\xi = 0$. The cross along the horizontal axis represents the sum of case estimates listed in the data set (*1,335,980,984*).*

For a claim $k$, we remove the first incurred amount of the random vector (2.6) and the MSN distribution is used to describe the behavior of

$$\boldsymbol{\Omega}^{-}_{v_k} = \left(\log\gamma_1^{(k)}, \log\gamma_2^{(k)}, \ldots, \log\gamma_{u_k^I}^{(k)}, \log\lambda_{u_k^P}^{(k)}, \ldots, \log\lambda_1^{(k)}\right)'.$$

The scale parameter $\boldsymbol{\Sigma}^{-}_{v_k}$ and the shape parameter $\boldsymbol{\Delta}^{-}_{v_k}$ are defined in a similar way. Then, we select an appropriate probability density function $g_1(\cdot;\boldsymbol{\nu})$ for $I_1$ depending on a vector of parameters $\boldsymbol{\nu}$. As the first incurred loss amount may not be SN-distributed, there are at least two ways to combine it with the MSN random vector $\boldsymbol{\Omega}^{-}_{v_k}$ containing all the development factors for paid and incurred amounts. This first incurred loss amount could be mapped to the unit interval by transforming it by its marginal distribution function. It could then be transformed by the SN quantile function and the MSN distribution then applied to the resulting random vector. This is in essence a copula approach. We do not follow this route here but we adopt a regression approach instead, as described next.

The idea is to specify the conditional distribution of the development factors, given the first incurred loss amount by letting the location parameter vector $\boldsymbol{\mu}$ depend on the value of $I_1$. The MSN distribution can then be used in a regression approach, following a similar approach as in Antonio and Plat [3], where $I_1$ is used as covariate information in an individual model. Specifically, given $I_{k1} = i$, the random vector $\boldsymbol{\Omega}_{v_k}^{-}$ obeys the MSN distribution with location vector $g_2(i; \boldsymbol{\kappa})$, scale matrix $(\boldsymbol{\Sigma}_{v_k}^{-})^{1/2}$ and shape vector $\boldsymbol{\Delta}_{v_k}^{-}$, for some function $g_2(\cdot) : \mathbb{R} \to \mathbb{R}^{|v_k|-1}$. The development vector and the first incurred amount are linked with the function $g_2(\cdot)$.

### *3.5.2 Estimation results*

Estimation results for time and exposure components are identical to those presented earlier. The location vector is constructed by considering three link functions $g_2^{(I)}$, $g_2^{(II)}$ and $g_2^{(III)}$, respectively given by

(I) linear function with different intercepts ($\boldsymbol{\kappa_1}$) and different slopes ($\boldsymbol{\kappa_2}$), i.e.

$$g_2^{(I)}(i; \boldsymbol{\kappa}) = \boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_2 i.$$

(II) linear function with different intercepts ($\boldsymbol{\kappa_1}$) and identical slope ($\kappa_3$), i.e.

$$g_2^{(II)}(i; \boldsymbol{\kappa}) = \boldsymbol{\kappa}_1 + \kappa_3 i.$$

(III) constant function $\boldsymbol{\kappa_1}$, i.e.

$$g_2^{(III)}(i; \boldsymbol{\kappa}) = \boldsymbol{\kappa}_1.$$

For each link function, we consider the three structures $\boldsymbol{\Sigma}_A$, $\boldsymbol{\Sigma}_B$ and $\boldsymbol{\Sigma}_C$ presented above for the scale matrix restricted to the development factor components. In Table 3.7, we report the Akaike and Bayesian Information Criteria (AIC and BIC), the number of parameters and the maximum log–likelihood. For each structure, we also include $p$-values from likelihood–ratio tests against simpler models. Based on the BIC, the preferred model is constructed with the scale matrix $\boldsymbol{\Sigma}_A$ and the link function $g_2^{(II)}$. For this model, we present estimation results in Table 3.8. The negative estimated $\kappa_3$ suggests that location parameters decrease in the first incurred amount $i$.

| Model | | # of par. | $-ll$ | AIC | BIC | $p$ value |
|---|---|---|---|---|---|---|
| $g_2$ | $\boldsymbol{\Sigma}$ | | | | | |
| $(I)$ | A | 24 | 7,732 | 15,512 | 15,655 | – |
| $(II)$ | A | 17 | 7,733 | 15,500 | 15,602 | 0.99 vs $(I)$ |
| $(III)$ | A | 16 | 7,742 | 15,515 | 15,611 | 0.003 vs $(II)$ |
| | | | | | | 0.27 vs $(I)$ |
| $(I)$ | B | 31 | 7,719 | 15,501 | 15,686 | – |
| $(II)$ | B | 24 | 7,723 | 15,494 | 15,637 | 0.78 vs $(I)$ |
| $(III)$ | B | 23 | 7,731 | 15,508 | 15,645 | 0.005 vs $(II)$ |
| | | | | | | 0.15 vs $(I)$ |
| $(I)$ | C | 28 | 7,710 | 15,476 | 15,643 | – |
| $(II)$ | C | 21 | 7,728 | 15,498 | 15,623 | 0.01 vs $(I)$ |
| $(III)$ | C | 20 | 7,736 | 15,512 | 15,631 | 0.005 vs $(II)$ |
| | | | | | | 0.001 vs $(I)$ |

**Table 3.7** *Comparison of results for different models for development pattern.*

| Loc. par. ($\boldsymbol{\kappa}$) (s.e.) | Sc. par. ($\boldsymbol{\Sigma}$) | Sh. par. ($\boldsymbol{\Delta}$) |
|---|---|---|
| $\kappa_{1,1} = 0.84$ | $\sigma_1^2 = 0.89$ | $\Delta_1 = -2.66$ |
| (0.0036) | $\sigma_2^2 = 0.45$ | $\Delta_2 = -1.93$ |
| $\kappa_{1,2} = 0.65$ | $\sigma_3^2 = 0.32$ | $\Delta_3 = -1.73$ |
| (0.0037) | $\sigma_4^2 = 0.25$ | $\Delta_4 = 1.45$ |
| $\kappa_{1,3} = 0.57$ | $\sigma_5^2 = 1.61$ | $\Delta_5 = -8.18$ |
| (0.0038) | $\sigma_6^2 = 0.57$ | $\Delta_6 = -6.84$ |
| $\kappa_{1,4} = -0.08$ | $\sigma_7^2 = 0.37$ | $\Delta_7 = -4.74$ |
| (0.0040) | $\sigma_8^2 = 0.24$ | $\Delta_8 = -4.63$ |
| $\kappa_{1,5} = 0.23$ | | |
| (0.0036) | | |
| $\kappa_{1,6} = 0.24$ | | |
| (0.0037) | | |
| $\kappa_{1,7} = 0.25$ | | |
| (0.0038) | | |
| $\kappa_{1,8} = 0.28$ | | |
| (0.0040) | | |
| $\kappa_3 = -0.02$ | | |
| (0.00002) | | |

**Table 3.8** *Estimation results for logarithms of development factors.*

In our analysis, we account for data limitation by considering for $g_1$ the censored probability density function

$$g_1(x;\xi,\boldsymbol{\nu}) = \frac{g_1^*(x;\boldsymbol{\nu})}{1 - G_1^*(\xi;\boldsymbol{\nu})}$$

corresponding to the uncensored distribution function $G_1^*(\cdot)$ with associated probability density function $g_1^*(\cdot)$. We consider different choices for $g_1^*(\cdot)$: Lognormal, Pareto, Generalized Pareto as defined in Klugman et al [9] are considered, as well as the Skew Lognormal (SLN) distribution and the composite Lognormal–Pareto distribution with Gamma threshold defined in Pigeon and Denuit [13]. The corresponding probability density functions are listed in Appendix A. Notice that SLN case is in line with previous sections, except that $I_1$ is now used as a regressor to explain the behavior of the development factors. Table 3.9 displays the corresponding AIC and BIC values. We can see there that both criteria assume close values so that no clear optimal model seems to emerge and they all equally well fit the data. We can see there that the SLN model outperforms other ones with differences in AIC and BIC higher than 10. We report estimated parameters in Table 3.10.

| Distribution | # of param. | $-ll$ | AIC | BIC |
|---|---|---|---|---|
| Lognormal | 2 | 46,903 | 93,811 | 93,823 |
| Pareto | 2 | 46,908 | 93,820 | 93,832 |
| Gen. Pareto | 3 | 46,902 | 93,809 | 93,828 |
| SLN | 3 | 46,888 | 93,783 | 93,801 |
| Comp. Log.–Par. | 4 | 46,893 | 93,794 | 93,819 |

**Table 3.9** *Comparison of results for different distributional choices for the first incurred amount.*

| Distribution | # of param. | $\boldsymbol{\nu}$ | | | |
|---|---|---|---|---|---|
| | | $\nu_1$ | $\nu_2$ | $\nu_3$ | $\nu_4$ |
| Lognormal | 2 | 11.92 | 1.02 | – | – |
| Pareto | 2 | 3.18 | 460,943 | – | – |
| Gen. Pareto | 3 | 2.45 | 5,496 | 0.010 | – |
| $\text{MSL}_1$ | 3 | 11.87 | 1.03 | 5.00 | – |
| Comp. Log.–Par. | 4 | 3.39 | 1.00 | 0.0000054 | 24.66 |

**Table 3.10** *Estimation of the parameters involved in the probability density function $g_1(\cdot)$ of the first incurred amount.*

### *3.5.3 Prediction results*

Let us now again predict the ultimate loss for this reinsurance portfolio using this alternative model allowing for flexible distributional choice for the first incurred amount. The probability density function of this random variable is displayed in Figure 3.4 for the different choices of the distribution for $I_1$. Table 3.11 contains the corresponding expected value, standard deviation as well as 95% and 99.5% Value-at-Risk and Conditional Tail Expectation measures. We can see there that the different candidate distributions for $I_1$ produce similar results for the expected aggregate ultimate loss. Figure 3.4 nevertheless reveals different shapes for the corresponding probability density function and the standard deviations listed in Table 3.11 considerably vary according to the distribution selected for $I_1$. The VaRs at the two probability levels considered are relatively similar.

When compared to the collective approach neglecting the censoring present in the data, we see again that the expected ultimate loss of the portfolio almost doubles which clearly illustrates the necessity of the approach developed in the present paper. As before, the sum of case estimates determined by claim adjusters as listed in the data set, which amounts to 1,335,980,984 is quite close to the 99.5% risk measures reported in Table 3.11.

## 4 Conclusion

In this paper, we have adapted the individual PIC model proposed in Pigeon et al [15] to deal with incomplete reinsurance data. In a first specification, data limitation induced by the first incurred loss amount in excess of some reporting threshold has been taken into account when fitting the multivariate Skew Normal distribution. In a second specification, some flexibility has been added to capture the tail behavior of the first incurred loss.

The model proposed in this paper has been illustrated by means of a case study based on a motor third-party liability reinsurance data set. When tested against a simpler specification disregarding the limitation of the data, our model clearly shows its superiority based on AIC, BIC and likelihood-ratio tests. The expected ultimate loss, its standard errors, 95% and 99.5% VaR and CTE risk measures are obtained by simulation. Failing to account for censorship biases downward the total reserve amount. As in Pigeon et al [15], comparison with collective reserving methods (performed in Table 3.6) reveals that – for the case study developed here – the proposed model in the present paper reduces the outstanding loss reserve and lowers the associated standard errors and risk measures. Similar conclusions apply to the model allowing for several distributional choices for the initial incurred loss.

| Model | | Exp. Value | s.d. | Risk Measures (0.95) | Risk Measures (0.995) |
|---|---|---|---|---|---|
| LN | A–(II) | 945,649,174 | 73,411,023 | $VaR$ = 1,071,152,637 | $VaR$ = 1,193,143,687 |
| | | | | $CTE$ = 1,124,470,725 | $CTE$ = 1,279,951,153 |
| | B–(II) | 936,006,404 | 71,567,812 | $VaR$ = 1,061,856,485 | $VaR$ = 1,165,737,262 |
| | | | | $CTE$ = 1,108,468,522 | $CTE$ = 1,222,890,615 |
| | C–(I) | 1,004,341,393 | 78,286,803 | $VaR$ = 1,139,535,397 | $VaR$ = 1,258,118,803 |
| | | | | $CTE$ = 1,194,852,648 | $CTE$ = 1,331,951,140 |
| Gen. Pareto | A–(II) | 952,125,461 | 84,321,950 | $VaR$ = 1,091,582,238 | $VaR$ = 1,273,826,685 |
| | | | | $CTE$ = 1,169,434,518 | $CTE$ = 1,434,980,469 |
| | B–(II) | 943,588,026 | 89,397,885 | $VaR$ = 1,082,447,363 | $VaR$ = 1,259,054,284 |
| | | | | $CTE$ = 1,171,316,660 | $CTE$ = 1,476,750,295 |
| | C–(I) | 1,011,783,937 | 92,696,436 | $VaR$ = 1,161,091,902 | $VaR$ = 1,371,695,613 |
| | | | | $CTE$ = 1,255,607,380 | $CTE$ = 1,555,331,874 |
| Pareto | A–(II) | 946,391,004 | 79,125,631 | $VaR$ = 1,077,070,298 | $VaR$ = 1,226,416,594 |
| | | | | $CTE$ = 1,145,619,929 | $CTE$ = 1,369,233,306 |
| | B–(II) | 937,680,146 | 79,487,110 | $VaR$ = 1,069,762,304 | $VaR$ = 1,212,891,917 |
| | | | | $CTE$ = 1,136,031,110 | $CTE$ = 1,354,972,098 |
| | C–(I) | 1,005,109,350 | 80,434,915 | $VaR$ = 1,141,155,465 | $VaR$ = 1,279,620,616 |
| | | | | $CTE$ = 1,203,950,011 | $CTE$ = 1,378,034,069 |
| USN | A–(II) | 949,089,824 | 72,988,478 | $VaR$ = 1,074,663,130 | $VaR$ = 1,203,967,996 |
| | | | | $CTE$ = 1,127,350,751 | $CTE$ = 1,257,195,969 |
| | B–(II) | 940,347,073 | 72,992,555 | $VaR$ = 1,065,169,073 | $VaR$ = 1,185,779,880 |
| | | | | $CTE$ = 1,119,455,705 | $CTE$ = 1,259,068,493 |
| | C–(I) | 1,007,745,464 | 77,827,274 | $VaR$ = 1,141,492,393 | $VaR$ = 1,263,274,237 |
| | | | | $CTE$ = 1,198,310,201 | $CTE$ = 1,340,908,840 |
| LN-Par. | A–(II) | 949,954,346 | 56,069,694 | $VaR$ = 1,045,187,869 | $VaR$ = 1,202,851,329 |
| | | | | $CTE$ = 1,071,589,043 | $CTE$ = 1,229,276,488 |
| | B–(II) | 938,624,577 | 54,425,352 | $VaR$ = 1,028,965,282 | $VaR$ = 1,188,970,075 |
| | | | | $CTE$ = 1,055,978,447 | $CTE$ = 1,213,145,576 |
| | C–(I) | 1,005,794,986 | 59,666,162 | $VaR$ = 1,105,618,555 | $VaR$ = 1,275,232,747 |
| | | | | $CTE$ = 1,183,278,266 | $CTE$ = 1,333,197,221 |
| C-L (paid) | | 877,357,349 | 92,521,517 | $VaR$ = 1,038,626,764 | $VaR$ = 1,168,329,326 |
| | | | | $CTE$ = 1,095,218,538 | $CTE$ = 1,232,096,313 |
| Munich C-L | | 2,304,960,343 | – | – | – |

**Table 3.11** *Prediction results based on 10,000 simulations for models with with (i) scale matrix* $\boldsymbol{\Sigma}_A$ *and link function* $g_2^{(II)}$ *("A-(II)"), (ii) scale matrix* $\boldsymbol{\Sigma}_B$ *and link function* $g_2^{(II)}$ *("B-(II)") and (iii) scale matrix* $\boldsymbol{\Sigma}_C$ *and link function* $g_2^{(I)}$ *("C-(I)").*

Numerical results obtained in this case study suggest that expected ultimate losses are quite robust with respect to the choice of distribution for the initial incurred amount but that the standard errors and high quantiles are impacted. Comparisons with collective approaches (performed in Table 3.11) lead to the same conclusion as before and demonstrates the interest of individual reserving methods.

**Acknowledgements**

We would like to express our deepest gratitude to Jean-François Walhin for extensive discussions about reinsurance practice and for deep comments on earlier versions of this manuscript, which helped us to considerably improve our modelling approach and to properly interpret the results.

**References**

1. Akdemir D (2009) A class of multivariate skew distributions: properties and inferential issues. PhD thesis, Bowling Green State University, Ohio

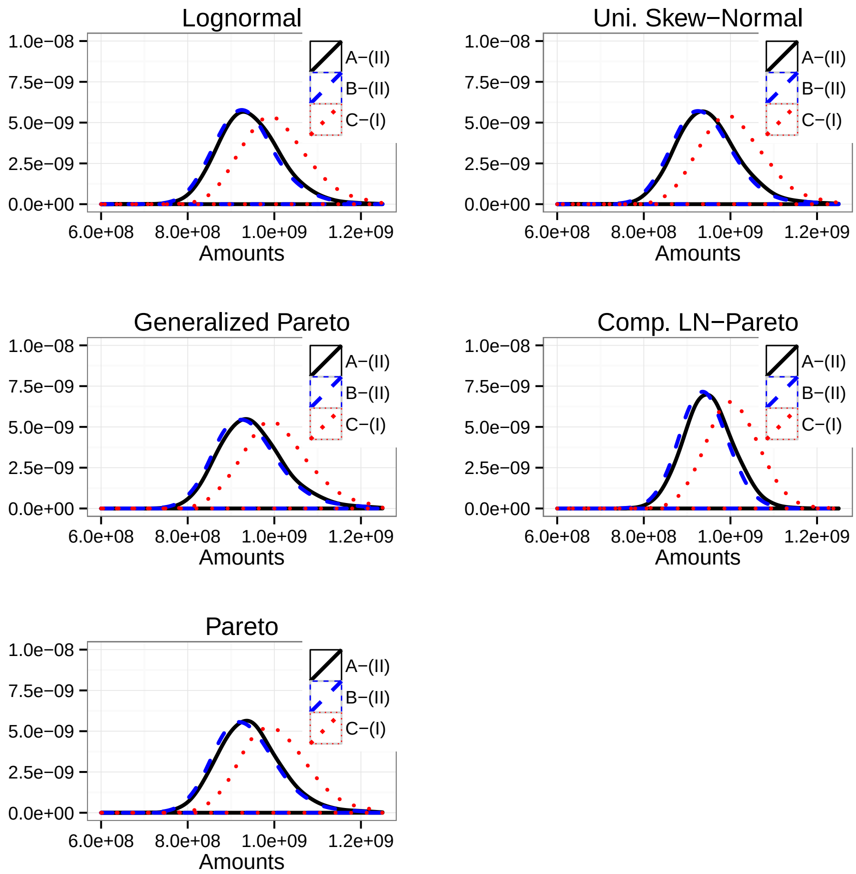


**Fig. 3.4** *Prediction results based on 10,000 simulations for models with (i) scale matrix* $\boldsymbol{\Sigma}_A$ *and link function* $g_2^{(II)}$ *("A-(II)"), (ii) scale matrix B and link function* $g_2^{(II)}$ *("B-(II)") and (iii) scale matrix C and link function* $g_2^{(I)}$ *("C-(I)").*


2. Akdemir D, Gupta A (2010) A Matrix Variate Skew Distribution. European Journal of Pure and Applied Mathematics 3(2):128–140
3. Antonio K, Plat R (2013) Micro–level stochastic loss reserving for general insurance. Scandinavian Actuarial Journal, *in press*
4. Arjas E (1989) The claims reserving problem in non–life insurance: some structural ideas. ASTIN Bulletin 19(2):140–152
5. de Haan L, Ferreira A (2006) Extreme Value Theory. Springer, New York
6. Drieskens D, Henry M, Walhin J-F, Wielandts J (2012) Stochastic projection for large individual losses. Scandinavian Actuarial Journal 2012(1):1–39
7. England P, Verrall R (2002) Stochastic claims reserving in general insurance. British Actuarial Journal 8(3):443–544
8. Gupta A, Chen J (2004) A class of multivariate skew normal models. The Annals of the Institute of Statistical Mathematics 56(2):305–315

9. Klugman S, Panjer H, Willmot G (2004) Loss Models: from Data to Decisions, Second Edition. Wiley, New Jersey
10. Larsen C (2007) An individual claims reserving model. ASTIN Bulletin 37(1):113–132
11. Mack T (1993) Distribution–free calculation of the standard error of chain–ladder reserve estimates. ASTIN Bulletin 23(2):213–225
12. Norberg R (1993) Prediction of outstanding liabilities in non-life insurance. ASTIN Bulletin 23(1):95–115
13. Pigeon M, Denuit M (2011) Composite Lognormal-Pareto model with random threshold. Scandinavian Actuarial Journal 2011(3):177–192
14. Pigeon M, Antonio K, Denuit M (2013) Individual loss reserving with the multivariate skew normal framework. ASTIN Bulletin 43(3):399–428
15. Pigeon M, Antonio K, Denuit M (2014) Individual loss reserving using paid–incurred data, *Accepted for publication in Insurance: Mathematics and Economics*
16. Rosenlund S (2012) Bootstrapping individual claim histories. ASTIN Bulletin 42(1):291–324
17. Wüthrich M, Merz M (2008) Stochastic claims reserving methods in insurance. Wiley Finance
18. Wüthrich M, Merz M (2010) Paid–incurred chain claims reserving method. Insurance: Mathematics and Economics 46(3):568–579
19. Happ S, Wüthrich M (2013) Paid–incurred chain reserving method with dependence modelling. ASTIN Bulletin 43(1):1–20
20. Zhao X, Zhou X, Wang J (2009) Semiparametric model for prediction of individual claim loss reserving. Insurance: Mathematics and Economics 45(1):1–8


## A Probability Density Functions

Lognormal: $g_1^*(x;\boldsymbol{\nu}) = \left(\frac{1}{x\nu_2\sqrt{2\pi}}\right)\exp\left(-0.5\left(\frac{\log(x)-\nu_1}{\nu_2}\right)^2\right)$.

Pareto: $g_1^*(x;\boldsymbol{\nu}) = \frac{\nu_1\nu_2^{\nu_1}}{(x+\nu_2)^{\nu_1+1}}$.

Generalized Pareto: $g_1^*(x;\boldsymbol{\nu}) = \left(\frac{\Gamma(\nu_1+\nu_3)}{\Gamma(\nu_1)\,\Gamma(\nu_3)}\right)\left(\frac{\nu_2^{\nu_1}x^{\nu_3-1}}{(x+\nu_2)^{\nu_1+\nu_3}}\right)$.

Skew Lognormal: $g_1^*(x;\boldsymbol{\nu}) = \left(\frac{1}{x}\right)\phi\left(\frac{\log(x)-\nu_1}{\nu_2}\right)\Phi\left(\nu_3\left(\frac{\log(x)-\nu_1}{\nu_3}\right)\right)$.

Composite Lognormal–Pareto with Gamma threshold:

$$\begin{aligned} g_1^*(x;\boldsymbol{\nu}) = {} & \frac{r}{\Phi(\nu_1\nu_2)x\nu_2}\int_x^\infty \phi\left(\frac{\ln(x)-\ln(y)+\nu_1\nu_2^2}{\nu_2}\right) g(y;\nu_3,\nu_4))\,dy \\ & + (r+1)g(x;\nu_3,\nu_4) \\ & - (1-r)\left(\frac{\Gamma(\nu_1+\nu_3)}{\Gamma(\nu_3)}\right)\left(\frac{-\nu_1}{\nu_4^{\nu_1}x^{\nu_1+1}}G(x;\nu_1+\nu_3,\nu_4) + \frac{1}{(\nu_4 x)^{\nu_1}}g(x;\nu_1+\nu_3,\nu_4)\right), \end{aligned}$$

where $G(x;a,b)$ denotes the Gamma cumulative distribution function with shape parameter $a$ and scale parameter $1/b$ evaluated at $x$ and

$$r = \frac{\sqrt{2\pi}\nu_1\nu_2\Phi(\nu_1\nu_2)\exp\left(\frac{1}{2}(\nu_1\nu_2)^2\right)}{\sqrt{2\pi}\nu_1\nu_2\Phi(\nu_1\nu_2)\exp\left(\frac{1}{2}(\nu_1\nu_2)^2\right)+1}.$$