INDIVIDUAL LOSS RESERVING WITH THE MULTIVARIATE SKEW NORMAL FRAMEWORK

BY

MATHIEU PIGEON, KATRIEN ANTONIO AND MICHEL DENUIT

ABSTRACT

The evaluation of future cash flows and solvency capital recently gained importance in general insurance. To assist in this process, our paper proposes a novel loss reserving model, designed for individual claims developing in discrete time. We model the occurrence of claims, as well as their reporting delay, the time to the first payment, and the cash flows in the development process. Our approach uses development factors similar to those of the well-known chain–ladder method. We suggest the Multivariate Skew Normal distribution as a multivariate distribution suitable for modeling these development factors. Empirical analysis using a real portfolio and out-of-sample prediction tests demonstrate the relevance of the model proposed.

KEYWORDS

Stochastic loss reserving, general insurance, Multivariate Skew Normal distribution, chain–ladder, individual claims.

1. INTRODUCTION

We develop a novel stochastic model for loss reserving in general insurance. The model uses detailed information on the development of individual claims. A vector of discrete random variables describes the claim's evolution over time, which evolves from occurrence of the accident until settlement or censoring of the claim. The corresponding stream of payments is expressed in terms of chain–ladder alike development factors (or: link ratios) and modeled with a multivariate, parametric distribution. The model leads to a theoretical expression for the expected value of the outstanding amount for each claim, and a corresponding predictive distribution follows by simulation.

We divide the time structure of a general insurance claim in three parts (see Figure 1). Between occurrence of the accident and notification to the insurance company, the insurer is liable for the claim amount but is unaware of the claim's

Astin Bulletin 43(3), 399-428. doi: 10.1017/asb.2013.20

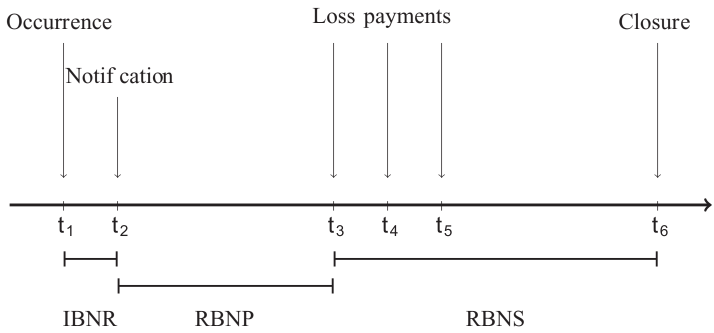

FIGURE 1: Evolution of a general insurance claim.

existence. The claim is said to be Incurred But Not Reported (IBNR). After notification, the company knows the claim and the first payment (if any) will follow. In this paper, we use the expression Reported But Not Paid (RBNP) to describe an incurred and reported claim for which no payments have been made yet. Then, the initial payment occurs and several partial payments (and refunds) follow. The claim finally closes at the closure or settlement date. From reporting until closure of the claim, the insurer is aware of its existence, but the final amount is still unknown: the claim is Reported But Not Settled (RBNS). This structure provides a flexible framework that can be simplified or extended if necessary.

At the evaluation date, the actuary should estimate technical provisions. Loosely speaking, the insurer must predict, with maximum accuracy, the total amount needed to pay claims that he has legally committed to cover. One part of the total amount comes from the completion of RBNS claims. Predictions for costs related to RBNP and IBNR claims form the second part of the total amount.

With the introduction of Solvency II and IFRS 4 Phase 2, the evaluation of future cash flows and regulatory-required solvency capital becomes more important and current techniques for loss reserving may have to be improved, adjusted or extended. In general, existing methods for claims reserving are designed for aggregated data, conveniently summarized in a run-off triangle with occurrence and development years. The chain–ladder approach (as studied in Mack, 1993, 1999) is the most popular member of this category. A rich literature exists about those techniques, see England and Verrall (2002) or Wüthrich and Merz (2008) for an overview.

Leaving the track of data aggregated in run-off triangles, Arjas (1989), Norberg (1993) and Norberg (1999) develop a mathematical framework for the development of individual claims in continuous time. More recent contributions in this direction are Zhao *et al.* (2009) and Antonio and Plat (2013). Verrall *et al.* (2010), Martinez *et al.* (2011) and Martinez *et al.* (2012) extend the traditional chain–ladder framework towards the use of extra data sources.

Inspired by Murphy and McLennan (2006), Drieskens *et al.* (2012) present a discrete time model for the development of individual large claims in reinsurance. Building blocks in Drieskens *et al.* (2012) are the large claim status, the open claim status and (chain–ladder alike) development factors. Rosenlund (2012) introduces the so-called "*Reserve by Detailed Conditioning*" (RDC) method (see Appendix *A* of his paper). This method is designed for the development of individual claims in discrete time. RDC leads to a point prediction of the outstanding loss amount by conditioning on specific characteristics of the observed development of an individual claim.

Similar to the discrete time approach in Drieskens *et al.* (2012) and Rosenlund (2012) (and in contrast to the continuous time approach from Norberg (1993) and Antonio and Plat (2013)), we use discrete random variables — at the level of an individual claim — for the reporting delay, the first payment delay, the number of payments and the number of periods between two consecutive payments. Individual development factors structure the development pattern. We propose the family of Multivariate Skew Symmetric (MSS) distributions (more specifically: the Multivariate Skew Normal (MSN)) to model the resulting, dependent development factors at individual claim level. Although Drieskens *et al.* (2012) and Rosenlund (2012) rely on the empirical distribution of each model component, our approach uses parametric distributions. The joint modeling of the development factors proposed in our paper is an alternative for Rosenlund's conditioning on the cumulative paid amount.

Our paper is organized as follows. We introduce the statistical model in Section 2. We present the data in Section 3 and develop this real example in Sections 4 and 5. Finally, we conclude in Section 6. Some technical developments are gathered in an Appendix, for the sake of completeness.

2. The model

Suppose we have a data set at our disposal with detailed information on the development of individual claims. More specifically, the model uses the occurrence date, the reporting date, the date(s) of payment(s) (and refund(s)) made for the claim, the amount(s) paid for the claim and the closure date. Figure 1 illustrates the development of an individual claim in continuous time.

2.1. Model specification

2.1.1. *Time components.* We leave the continuous time framework from Figure 1 and work in discrete time (e.g. with periods of one year). We denote the *kth* claim from occurrence period i with (ik). Hereby, $k = 1, \ldots, K_i$ and $i = 1, \ldots, I$ where I denotes the number of occurrence periods under consideration, and K_i is the number of claims originating in period i. In our discrete time framework, we identify:

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

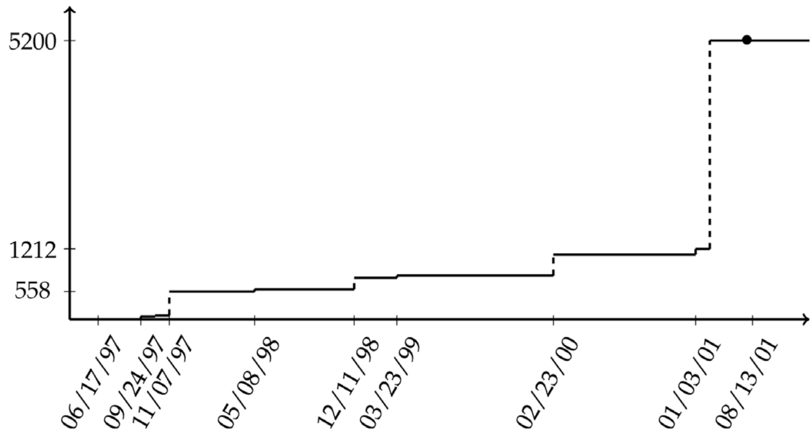

FIGURE 2: Development of a random claim in continuous time. The x-axis represents the date of each event and the y-axis represents the cumulative amount paid for the claim.

- the random variable T_{ik} is the *reporting delay* for claim (ik), i.e. the number of periods between the occurrence period of the claim and its notification to the insurance company;
- the random variable Q_{ik} is the *first payment delay*, representing the number of periods between notification and the first payment for claim (ik);
- the random variable U_{ik} denotes the *number of period(s)* with partial payment (> 0) after the first one; and
- the random variable N_{ikj} represents the *delay between two periods with payment* which is the number of periods between payments j and $j+1$. We use $N_{ik,U_{ik}+1}$ to denote the number of periods between the last payment and the settlement of the claim. Consequently, $N_{ik} = \sum_{j=1}^{U_{ik}+1} N_{ikj}$ is the number of periods between the first payment and the period of settlement of the claim.

Each component follows a discrete distribution $f : \mathbb{N} \to [0, 1]$, respectively $f_1(t; \boldsymbol{\nu})$, $f_2(q; \boldsymbol{\psi})$, $f_3(u; \boldsymbol{\beta})$ and $f_4(n; \boldsymbol{\phi})$ (with corresponding cdf $F_1(.; \boldsymbol{\nu}), \ldots, F_4(.; \boldsymbol{\phi})$). By definition, $\Pr(N_{ikj} = 0) = 0$, $j = 1, \ldots, U_{ik}$. In the sequel of the text, we will interpret "periods" as years. Per claim and per discrete time period, we aggregate all intermediate payments. Figure 2 represents the development of a random claim from the data set. Following the approach presented in this paper, Figure 3 transforms the data set to discrete time periods.

The accident occurs at 06/17/1997, thus the occurrence period corresponds to the year "1997". The claim is reported to the company on 07/22/1997, thus: $t_{(ik)} = 0$. A first payment is done on 09/24/1997, implying a first payment delay of 0 periods ($q_{(ik)} = 0$). Consequently, payments follow on 10/21/1997, 11/07/1997, 05/08/1998, 12/11/1998, 03/23/1999, 02/23/2000, 01/03/2001 and 02/24/2001. Therefore, $u_{(ik)} = 4$ and $n_{(ik),1} = n_{(ik),2} = n_{(ik),3} = n_{(ik),4} = 1$. Closure is at 08/13/2001, thus $n_{(ik),5} = 0$.

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

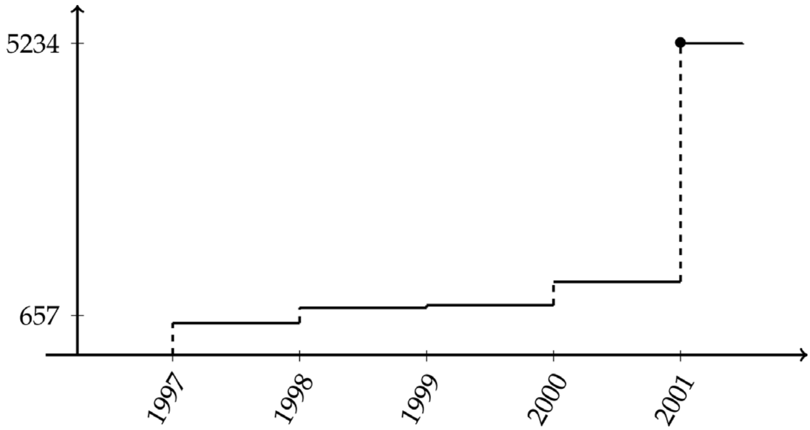

FIGURE 3: Development of the claim from Figure 2 in a discrete time framework (with yearly periods).

2.1.2. *Exposure and occurrence measures.* To distinguish explicitly between IBNR and RBNS/RBNP claims, we need a stochastic process driving the occurrence of claims, while accounting for the exposure in a specific occurrence period. The number of claims for occurrence period i, say K_i, follows a Poisson distribution with occurrence measure $\theta w(i)$. $w(i)$ is the exposure registered for occurrence period i ($i = 1, \ldots, I$). However, since we only observe reported claims, the Poisson distribution should be thinned in the following way:

$$\theta w(i) F_1(t_i^* - 1; \boldsymbol{\nu}). \tag{1}$$

Evaluation then takes place t_i^* periods after occurrence period i, and is done at the beginning of this period. As introduced in Section 2.1.1, $F_1(.; \boldsymbol{\nu})$ is the cdf assumed for reporting delay.

2.1.3. *Development pattern.*

Structuring the development pattern. Let the random variable Y_{ikj} (> 0) represent the jth incremental partial amount for the kth claim ($k = 1, \ldots, K_i$) from occurrence period i ($i = 1, \ldots, I$). We hereby aggregate intermediate payments from the same discrete time period in the development of a claim (as illustrated in Figure 3). The total cumulative amount paid for claim (ik) follows by multiplying the initial amount, Y_{ik1}, with one or more *link ratios*. The initial amount, together with the vector of link ratio(s), forms the *development pattern* of the claim. This approach is similar to the chain–ladder method (see Mack 1993,1999). However, with chain–ladder, the index j is for development period instead of partial payment. Using a *development-to-development* period model (as chain–ladder does) with individual claims can be problematic because the length of the development pattern differs among claims, and many development

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

factors will have value one. We avoid this in the *payment-to-payment* approach (in discrete time) used in our paper.

For a claim (ik) with a strict positive value of $U_{ik} = u_{ik}$, the vector $\mathbf{\Lambda}^{(ik)}_{u_{ik}+1}$ of length $u_{ik} + 1$ gives the development pattern

$$\mathbf{\Lambda}^{(ik)}_{u_{ik}+1} = \left[Y_{ik1}\ \lambda^{(ik)}_1\ \ldots\ \lambda^{(ik)}_{u_{ik}} \right]', \tag{2}$$

where

$$\lambda^{(ik)}_j = \frac{\sum_{r=1}^{j+1} Y_{ikr}}{\sum_{r=1}^{j} Y_{ikr}}, \tag{3}$$

for $j = 1, \ldots, u_{ik}$. In the stochastic version of the chain–ladder model, successive development factors are supposed to be non-correlated given past information. Moreover, independence is assumed between the initial payment and the vector of development factors. The Paid Incurred Chain (PIC) model from Merz and Wüthrich (2010) is an exception. The study by Happ and Wüthrich (2013) examines dependence structures for link ratios in the PIC model. In our individual framework, the assumption of independence is problematic and unrealistic (as demonstrated empirically in Section 4.2.3, Figures 6 (Bodily Injury) and 7 (Material Damage)). This motivates the use of a flexible multivariate distribution for $\mathbf{\Lambda}^{(ik)}_{u_{ik}+1}$ ($i = 1, \ldots, I$ and $k = 1, \ldots, K_i$). Such a distribution should be able to account for the dependence present in the development pattern vector, as well as the asymmetry in each of its components.

A flexible multivariate distribution for the development pattern. Our paper uses the family of MSS distributions (see Gupta and Chen, 2004; Akdemir and Gupta, 2010) to model the development pattern of a claim (ik) on log scale. More specifically, we will use the MSN distribution, a multivariate extension of the Univariate Skew Normal (USN), distribution (from Roberts and Geisser, 1966 and Azzalini, 1985).

Definition 2.1 (MSS and MSN distribution). *Let $\boldsymbol{\mu} = [\mu_1\ \ldots\ \mu_k]'$ be a vector of location parameters, $\mathbf{\Sigma}$ a $(k \times k)$ positive definite symmetric scale matrix and $\mathbf{\Delta} = [\Delta_1\ \ldots\ \Delta_k]'$ a vector of shape parameters. The $(k \times 1)$ random vector $\mathbf{X}$ follows a MSS distribution if its density function is of the form*

$$\begin{aligned} &MSS\left(\mathbf{x}; \boldsymbol{\mu}, \mathbf{\Sigma}^{1/2}, \mathbf{\Delta}\right) \\ &\quad = \frac{2^k}{det(\mathbf{\Sigma})^{1/2}} g^*\left(\mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\mu})\right) \prod_{j=1}^{k} H\left(\Delta_j \mathbf{e}'_j \mathbf{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\mu})\right), \end{aligned} \tag{4}$$

where $g^*(\mathbf{x}) = \prod_{j=1}^{k} g(x_j)$, $g(\cdot)$ is a density function symmetric around 0*,* $H(\cdot)$

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

is an absolutely continuous cumulative distribution function with $H'(\cdot)$ *symmetric around* 0 *and* $\mathbf{e}'_j$ *are the elementary vectors of the coordinate system* $\mathbb{R}^k$.[1]

The MSN distribution is obtained from (4) *by replacing* $g(\cdot)$ *and* $H(\cdot)$ *with the pdf and cdf of the standard Normal distribution, respectively.*

2.2. The likelihood

For the sake of clarity, the likelihood function will be divided into three parts: an expression for the likelihood of closed, RBNP and RBNS claims.

Closed claims. For closed claims (Cl), the likelihood function is given below. Hereby, t^*_{ik} refers to the evaluation date, expressed as number of periods after occurrence, and evaluation is performed at the beginning of this period. $(ik)_{\text{Cl}}$ refers to a closed claim.

$$\begin{aligned}
\mathcal{L}^{\text{Cl}} \propto & \prod_{(ik)_{\text{Cl}}} \text{MSN}(\ln\left(\mathbf{\Lambda}_{u_{ik}+1}\right); \boldsymbol{\mu}_{u_{ik}+1}, \mathbf{\Sigma}^{1/2}_{u_{ik}+1}, \mathbf{\Delta}_{u_{ik}+1}|u_{ik}) \\
& \cdot \prod_{(ik)_{\text{Cl}}} f_1(t_{ik}; \boldsymbol{\nu}|T_{ik} \leq t^*_{ik} - 1) \cdot f_2(q_{ik}; \boldsymbol{\psi}|Q_{ik} \leq t^*_{ik} - t_{ik} - 1) \\
& \cdot \prod_{(ik)_{\text{Cl}}} f_3(u_{ik}; \boldsymbol{\beta}|U_{ik} \leq t^*_{ik} - q_{ik} - t_{ik} - 1) \\
& \cdot \prod_{(ik)_{\text{Cl}}} I(u_{ik} = 0)(1) + I(u_{ik} = 1) f_4(n_{ik1}; \boldsymbol{\phi}|0 < N_{ik1} \leq t^*_{ik} - t_{ik} - q_{ik} - u_{ik}) \\
& + I(u_{ik} > 1) f_4(n_{ik1}; \boldsymbol{\phi}|0 < N_{ik1} \leq t^*_{ik} - t_{ik} - q_{ik} - u_{ik}) \\
& \cdot \prod_{j=2}^{u_{ik}} f_4\left(n_{ikj}; \boldsymbol{\phi}|0 < N_{ikj} \leq t^*_{ik} - t_{ik} - q_{ik} - (u_{ik} - j + 1) - \sum_{p=1}^{j-1} n_{ikp}\right). \quad (5)
\end{aligned}$$

The first component in this likelihood (i.e. “MSN(...)”) is the multivariate distribution of the development pattern, given the total number of link ratio(s). The other components, $f_1(.)$, $f_2(.)$, $f_3(.)$ and $f_4(.)$, refer to reporting delay, first payment delay, the number of periods with payment and the delay between two periods with payment.

RBNS claims. For RBNS claims, the likelihood is (with u^*_{ik} the observed number of periods with payment after the first one, and $(ik)_{\text{RBNS}}$ indicating an

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

RBNS claim)

$$\mathcal{L}^{\text{RBNS}} \propto \prod_{(ik)_{\text{RBNS}}} \text{MSN}(\ln\left(\mathbf{\Lambda}_{u^*_{ik}+1}\right); \boldsymbol{\mu}_{u^*_{ik}+1}\mathbf{\Sigma}^{1/2}_{u^*_{ik}+1}, \mathbf{\Delta}_{u^*_{ik}+1}|u^*_{ik})$$
$$\cdot \prod_{(ik)_{\text{RBNS}}} f_1(t_{ik}; \boldsymbol{\nu}|T_{ik} \leq t^*_{ik} - 1) \cdot f_2(q_{ik}; \boldsymbol{\psi}|Q_{ik} \leq t^*_{ik} - t_{ik} - 1)$$
$$\cdot \prod_{(ik)_{\text{RBNS}}} (1 - F_3(u^*_{ik} - 1; \boldsymbol{\beta}))$$
$$\cdot \prod_{(ik)_{\text{Cl}}} I(u^*_{ik} = 0)(1) + I(u^*_{ik} = 1) f_4(n_{ik1}; \boldsymbol{\phi}|0 < N_{ik1} \leq t^*_{ik} - t_{ik} - q_{ik} - u^*_{ik})$$
$$+ I(u^*_{ik} > 1) f_4(n_{ik1}; \boldsymbol{\phi}|0 < N_{ik1} \leq t^*_{ik} - t_{ik} - q_{ik} - u^*_{ik})$$
$$\cdot \prod_{j=2}^{u^*_{ik}} f_4(n_{ikj}; \boldsymbol{\phi}|0 < N_{ikj} \leq t^*_{ik} - t_{ik} - q_{ik} - (u^*_{ik} - j + 1) - \sum_{p=1}^{j-1} n_{ikp}). \quad (6)$$

RBNP claims. Finally, for RBNP claims, the likelihood function is (with $(ik)_{\text{RBNP}}$ indicating an RBNP claim)

$$\mathcal{L}^{\text{RBNP}} \propto \prod_{(ik)_{\text{RBNP}}} f_1(t_{ik}; \boldsymbol{\nu}|T_{ik} \leq t^*_{ik} - 1) \cdot (1 - F_2(t^*_{ik} - t_{ik} - 1; \boldsymbol{\psi})). \quad (7)$$

2.3. Analytical results for best estimates of outstanding reserves

The model specified in Sections 2.1 and 2.2 allows to derive analytical results for the nth moment of an IBNR, RBNP and RBNS claim, as well as for the expected value of the IBNR, RBNP and RBNS reserve. Proofs are deferred to Appendix A. We drop the (ik) subscript for reasons of simplicity.

Proposition 2.2 (nth moment of an IBNR or RBNP claim.). *Let C be the random variable representing the total claim amount of an IBNR (or RBNP) claim*

$$C = Y_1 \cdot \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_U. \quad (8)$$

Using the model assumptions from Sections 2.1 and 2.2 with location vector $\boldsymbol{\mu}$, scale matrix $\mathbf{\Sigma}$ and shape vector $\mathbf{\Delta}$, the nth moment of C is given by

$$E\left[2^{U+1} \exp\left(\mathbf{t}'_n \boldsymbol{\mu}_{U+1} + 0.5\mathbf{t}'_n \mathbf{\Sigma}^{1/2}_{U+1}\left(\mathbf{\Sigma}^{1/2}_{U+1}\right)' \mathbf{t}_n\right) \cdot \prod_{j=1}^{U+1} \Phi\left(\frac{\Delta_j \cdot \left(\left(\mathbf{\Sigma}^{1/2}_{U+1}\right)' \mathbf{t}_n\right)_j}{\sqrt{1 + \Delta^2_j}}\right)\right]. \quad (9)$$

$\mathbf{t}_n$ is an $((U+1) \times 1)$ vector, specified as $[n\ n\ \ldots\ n]'$. In formula (9) (as well as in formula (13) and elsewhere in the text), the expected value is taken with respect to the random variable U.

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

Proposition 2.3 gives the corresponding result for an RBNS claim. The distinguishing feature between Propositions 2.2 and 2.3 is the fact that for an RBNS claim part of the development pattern has already been observed.

Proposition 2.3 (nth moment of an RBNS claim.). *Define*

$$\boldsymbol{\Lambda}_{U+1} = \begin{bmatrix} \boldsymbol{\Lambda}_A \\ \boldsymbol{\Lambda}_B \end{bmatrix}, \quad \boldsymbol{\mu}_{U+1} = \begin{bmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{bmatrix},$$
$$\boldsymbol{\Sigma}_{U+1}^{1/2} = \begin{bmatrix} \boldsymbol{\Sigma}_{AA} & \mathbf{0} \\ \boldsymbol{\Sigma}_{BA} & \boldsymbol{\Sigma}_{BB} \end{bmatrix}, \quad \boldsymbol{\Delta}_{U+1} = \begin{bmatrix} \boldsymbol{\Delta}_A \\ \boldsymbol{\Delta}_B \end{bmatrix}, \tag{10}$$

where $\boldsymbol{\Lambda}_A$, $\boldsymbol{\mu}_A$ and $\boldsymbol{\Delta}_A$ are $U_A \times 1$ (with $U_A < U+1$). $\boldsymbol{\Sigma}_{AA}$ is a $U_A \times U_A$ lower triangular matrix with positive diagonal elements and $\boldsymbol{\Sigma}_{BB}$ is a $U_B \times U_B$ lower triangular matrix with positive diagonal elements. Hereby, $U_B + U_A = U+1$, the total number of periods with partial payment.

We define $[\boldsymbol{\mu}^*_{U+1}|\boldsymbol{\Lambda}_A = \boldsymbol{\ell}_A] := \boldsymbol{\mu}_B + \boldsymbol{\Sigma}_{BA}\boldsymbol{\Sigma}_{AA}^{-1}(\boldsymbol{\ell}_A - \boldsymbol{\mu}_A)$, $\boldsymbol{\Sigma}^*_{U+1} = \boldsymbol{\Sigma}_{BB}$ and $\boldsymbol{\Delta}^*_{U+1} = \boldsymbol{\Delta}_B$. The conditional final amount of a claim C, given past information, is defined as

$$[C|\boldsymbol{\Lambda}_A = \boldsymbol{\ell}_A] = y_1 \cdot \ell_1 \cdot \ldots \cdot \ell_{u_A - 1} \cdot \lambda_{u_A} \cdot \ldots \cdot \lambda_U. \tag{11}$$

Using the model assumptions from Sections 2.1 and 2.2, the nth conditional moment of C is given by

$$\begin{aligned} E[C^n|\boldsymbol{\Lambda}_A = \boldsymbol{\ell}_A] &= \left(y_1 \cdot \ell_1 \cdot \ldots \cdot \ell_{u_A-1}\right)^n \\ &\cdot E\left[2^{U_B} \exp\left(\boldsymbol{h}_n'\boldsymbol{\mu}^*_{U+1} + 0.5\boldsymbol{h}_n'\left(\boldsymbol{\Sigma}^*_{U+1}\right)^{1/2}\left(\left(\boldsymbol{\Sigma}^*_{U+1}\right)^{1/2}\right)'\boldsymbol{h}_n\right)\right. \\ &\left.\cdot \prod_{j=1}^{U_B} \Phi\left(\frac{\Delta^*_j \cdot \left(\left(\left(\boldsymbol{\Sigma}^*_{U+1}\right)^{1/2}\right)'\boldsymbol{h}_n\right)_j}{\sqrt{1 + \left(\Delta^*_j\right)^2}}\right)\right], \end{aligned} \tag{12}$$

with the $(U_B \times 1)$ vector $\boldsymbol{h}_n := [n\ n\ \ldots\ n]'$.

Analytical expressions for the expected values of the total outstanding IBNR, RBNP and RBNS reserves follow immediately from Propositions 2.2 and 2.3.

Proposition 2.4 (Best estimates for the IBNR, RBNP and RBNS reserves.). *Let $\mathcal{I}$ denote the observed information for all claims in the data set. We define $\mathbf{t}_n$, $\mathbf{h}_n$, $\boldsymbol{\mu}^*_{U+1}$, $\boldsymbol{\Sigma}^*_{U+1}$ and $\boldsymbol{\Delta}^*_{U+1}$ as in Propositions 2.2 and 2.3, respectively. Using the model assumptions from Sections 2.1 and 2.2, the best estimate of the outstanding IBNR, RBNP and RBNS reserves follow:*

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

(*a*) *The expected value of the total amount outstanding for IBNR and RBNP claims, respectively, is*

$$\begin{aligned} &E[IBNR|\mathcal{I}] \text{ versus } E[RBNP|\mathcal{I}] \\ &= (x) \cdot E\left[2^{U+1} \exp(\mathbf{t}_1' \boldsymbol{\mu}_{U+1} + 0.5\mathbf{t}_1' \boldsymbol{\Sigma}_{U+1}^{1/2} \left(\boldsymbol{\Sigma}_{U+1}^{1/2}\right)' \mathbf{t}_1) \right. \\ &\left. \cdot \prod_{j=1}^{U+1} \Phi\left(\frac{\Delta_j \cdot ((\boldsymbol{\Sigma}_{U+1}^{1/2})'\mathbf{t}_1)_j}{\sqrt{1+\Delta_j^2}}\right)\right], \end{aligned} \tag{13}$$

where (x) *should be replaced with* $E[K_{IBNR}]$ *in case of IBNR reserves, and with* k_{RBNP}, *the observed number of open claims without payment, in case of RBNP reserves. The expected number of IBNR claims follows from the Poisson distribution driving the occurrence of claims (appropriately thinned to represent IBNR claims).*

(*b*) *The expected value of the total amount outstanding for RBNS claims is*

$$\begin{aligned} &E[RBNS|\mathcal{I}] \\ &= \sum_{(ik)_{RBNS}} y_1 \cdot \ell_1 \cdot \ldots \cdot \ell_{u_1-1} \\ &\cdot E\left[2^{U_B} \exp(\mathbf{h}_1' \boldsymbol{\mu}_{U+1}^* + 0.5\mathbf{h}_1' \left(\boldsymbol{\Sigma}_{U+1}^*\right)^{1/2} \left(\left(\boldsymbol{\Sigma}_{U+1}^*\right)^{1/2}\right)' \mathbf{h}_1) \right. \\ &\left. \cdot \prod_{j=1}^{U_B} \Phi\left(\frac{\Delta_j^* \cdot \left(\left(\left(\boldsymbol{\Sigma}_{U+1}^*\right)^{1/2}\right)' \mathbf{h}_1\right)_j}{\sqrt{1+\left(\Delta_j^*\right)^2}}\right)\right], \end{aligned} \tag{14}$$

where the sum goes over all RBNS claims.

In Propositions 2.2, 2.3 and 2.4, except in the case of degenerate distribution, we need to empirically average over the outcome of U to obtain numerical results.

3. The data

3.1. Background

We use the data set from Antonio and Plat (2013) on a portfolio of general liability insurance policies for private individuals.[2] Available information is from January 1997 until December 2004. Originally, information is available until August 2009, but to enable out-of-sample prediction, we remove the observations from January 2005 to August 2009. Two types of payments are registered in the data set: Bodily Injury (BI) and Material Damage (MD).[3] Section 2 (see Figures 2 and 3) visualizes a claim from the data set.

TABLE 1

DESCRIPTIVE STATISTICS FOR CLOSED CLAIMS: FIRST PAYMENT, TOTAL CLAIM AMOUNT, AND DEVELOPMENT FACTORS λ_j WITH $j \leq 4$ FOR BI CLAIMS AND $j \leq 2$ FOR MD CLAIMS.

Class	Variables	Mean	Median	s.e.	Minimum	Maximum	Number of Observations
	Y_1	1, 008	351	3, 274	0.18	148, 900	2, 961
	λ_1	10.24	3.23	31.52	1.01	653.33	991
BI	λ_2	4.50	1.95	10.80	1.00	127.74	253
	λ_3	2.73	1.80	2.18	1.00	11.94	89
	λ_4	2.67	1.92	2.22	1.00	11.44	37
	Total Claim	2, 961	624	11, 825	6.3	410, 500	2, 961
	Y_1	298	151	528	0.35	68, 810	181, 828
MD	λ_1	5.44	2.18	11.71	1.00	371.40	1, 555
	λ_2	2.16	1.41	1.73	1.01	6.93	13
	Total Claim	305	153	679	0.35	108, 300	181, 828

3.2. Descriptive statistics

The data set consists of 279, 094 reported claims; 273, 977 of these claims are related to MD and 5, 117 to BI. A total of 268, 484 MD claims (181, 828 with at least one payment and 86, 656 with no payment) and 4, 098 BI claims (2, 961 with at least one payment and 1, 137 with no payment) are closed in the data set. We present descriptive statistics for closed claims with positive payments in Table 1. In Section 4.1, descriptive graphics follow representing reporting delay, first payment delay and the number of periods with payment (see Figure 4). We illustrate dependence between development factors in Figures 6 (BI) and 7 (MD).

4. DISTRIBUTIONAL ASSUMPTIONS AND ESTIMATION RESULTS

4.1. Distributional assumptions

Distributions for number of periods. For the random variables describing the time structure part of a claim's development (i.e. $\{T_{ik}\}$, $\{Q_{ik}\}$, $\{U_{ik}\}$ and $\{N_{ik}\}$ from Section 2.1), we consider mixtures of a discrete distribution with degenerate components (similar to Antonio and Plat, 2013). For reporting delay, for instance, we investigate distributions of the following type:

$$f_1(t; \boldsymbol{\nu}) = \sum_{s=0}^{p} \nu_s I_s(t) + \left(1 - \sum_{s=0}^{p} \nu_s\right) f_{T|T>p}(t), \tag{15}$$

where $I_s(t) = 1$ for reporting in the sth period after the period of occurrence and 0 otherwise. $f(t)$ is the pdf of a discrete distribution with parameter(s) $\nu_{p+1}, \ldots, \nu_{p+q}$. Further on, we investigate the use of a Geometric, Binomial, Poisson and Negative Binomial distribution for $f(.)$, combined with different values for p ($p = 0, 1, 2, 3$).

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

TABLE 2

MODEL SELECTION FOR $\{T_{ik}\}$, $\{Q_{ik}\}$, $\{U_{ik}\}$, USING THE STRUCTURE FROM (15) WITH A GEOMETRIC DISTRIBUTION FOR THE BASIC COUNT DISTRIBUTION.

	p	BI AIC	BI BIC	MD AIC	MD BIC
$(T; \nu)$	basic dist.	3, 120	3, 126	88, 730	88, 740
	0	2, 987	3, 000	85, 714	85, 735
	1	2, 966	2, 985	85, 484	85, 515
	2	2, 961	2, 987	85, 479	85, 521
	3	2, 963	2, 994	85, 485	85, 537
$(Q; \psi)$	basic dist.	4, 882	4, 888	116, 611	116, 621
	0	4, 605	4, 617	111, 045	111, 066
	1	4, 575	4, 594	110, 680	110, 711
	2	4, 577	4, 602	110, 676	110, 717
	3	4, 578	4, 609	110, 680	110, 732
$(U; \beta)$	basic dist.	6, 102	6, 108	18, 255	18, 265
	0	6, 096	6, 108	18, 250	18, 270
	1	6, 025	6, 043	18, 233	18, 264
	2	6, 026	6, 051	18, 233	18, 273
	3	6, 017	6, 048	18, 235	18, 285

TABLE 3

ESTIMATION RESULTS FOR THE SELECTED DISTRIBUTION FOR $\{T_{ik}\}$, $\{Q_{ik}\}$, $\{U_{ik}\}$, I.E. A GEOMETRIC DISTRIBUTION WITH DEGENERATE COMPONENTS. PARAMETERS ARE DENOTED AS IN (15).

Class	Parameter Index	Report delay $(T; \nu_s)$ (s.e.)	First pmt delay $(Q; \psi_s)$ (s.e.)	Number partial pmt $(U; \beta_s)$ (s.e.)
	0	0.8953	0.7127	0.5192
		(< 0.001)	(< 0.001)	(0.010)
	1	0.0819	0.2522	0.2470
		(0.003)	(0.003)	(0.008)
BI	2	0.5144	0.6431	0.3094
		(0.064)	(0.052)	(0.022)
	0	0.9565	0.9181	0.9896
		(< 0.001)	(< 0.001)	(< 0.001)
	1	0.0421	0.0794	0.0103
		(< 0.001)	(< 0.001)	(< 0.001)
MD	2	0.6820	0.6729	0.7184
		(0.031)	(0.026)	(0.125)

Development pattern. For the logarithm of the development pattern vector (as in (2)), we consider the MSN distribution on the one hand and the special case where $\mathbf{\Delta} = \mathbf{0}$, i.e. the Multivariate Normal (MN) distribution, on the other hand. The following structures are considered for the $z \times z$ matrix $\Sigma_c^{1/2}$ 4:

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

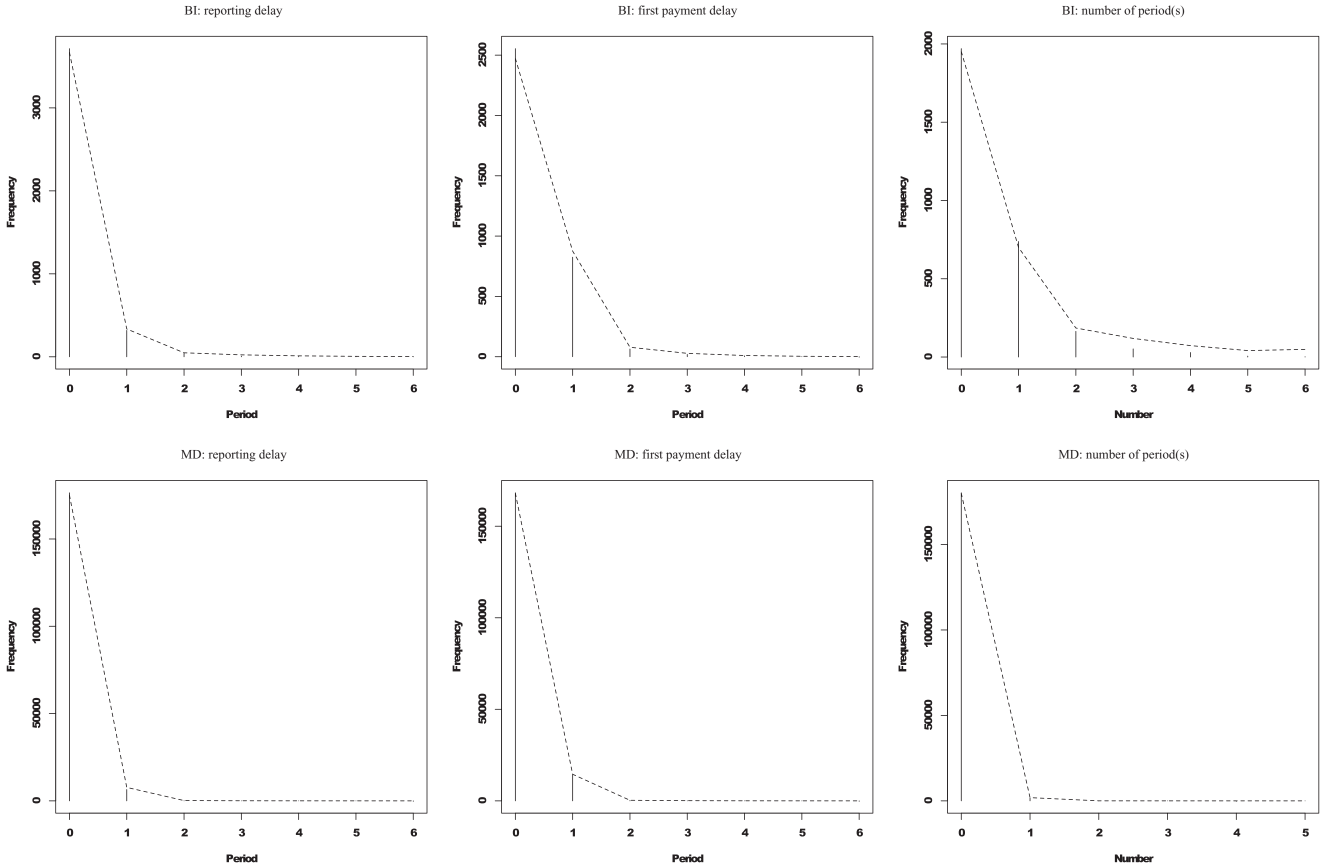

FIGURE 4: Observed (solid line) and estimated (broken line) frequency distributions for Bodily Injury (BI, top row) and Material Damage (MD, bottom row). From left to right: reporting delay, first payment delay and number of intermediate payments after the first one.

TABLE 4

LOGARITHM OF THE SEVERITY OF THE FIRST AND ONLY PAYMENT: ESTIMATION RESULTS FOR THE UNIVARIATE SKEW NORMAL DISTRIBUTION (USN) (WITH PARAMETERS μ, σ AND SCALE PARAMETER δ) AND THE NORMAL DISTRIBUTION (N) (WITH PARAMETERS μ AND σ).

	BI		MD	
	USN	N	USN	N
μ	5.9377	5.9226	4.9541	5.0428
(s.e.)	(1.04)	(0.03)	(0.06)	(< 0.01)
σ	1.3966	1.3968	1.1663	1.1637
(s.e.)	(0.02)	(0.02)	(0.01)	(< 0.01)
δ	−0.0139	–	0.0959	–
(s.e.)	(0.94)	–	(0.07)	–
AIC	4, 124	4, 122	284, 855	284, 853
BIC	4, 141	4, 133	284, 885	284, 873

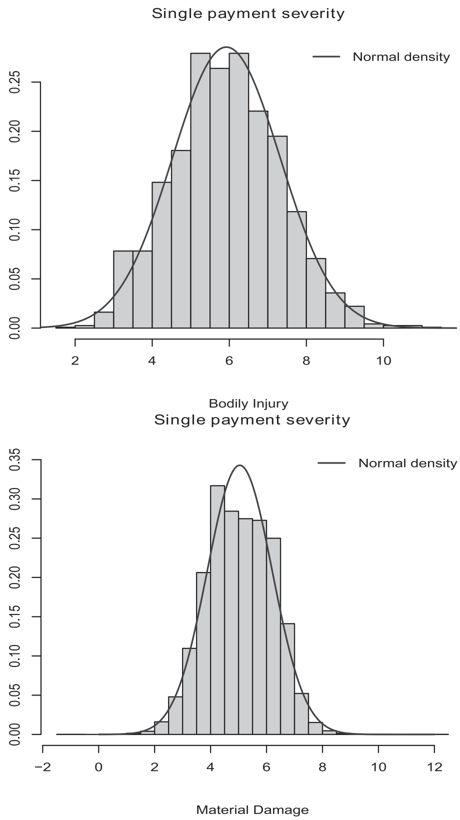

FIGURE 5: Logarithm of the severity of the first and only payment: empirical and fitted densities for Bodily Injury (top) and Material Damage (bottom).

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

unstructured (UN), Toeplitz (TOEP), Compound Symmetry (CS) and Diagonal (DIA) (see below).

$$\begin{pmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ \sigma_{21} & \sigma_2^2 & 0 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ \vdots & & & \ddots & \vdots \\ \sigma_{z1} & \sigma_{z2} & \sigma_{z3} & \dots & \sigma_z^2 \end{pmatrix} \quad \begin{pmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & 0 & \dots & 0 \\ \sigma_3\sigma_1\rho_2 & \sigma_3\sigma_2\rho_1 & \sigma_3^2 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ \sigma_z\sigma_1\rho_{z-1} & \sigma_z\sigma_2\rho_{z-2} & \sigma_z\sigma_3\rho_{z-3} & \dots & \sigma_z^2 \end{pmatrix}$$

(UN) (TOEP)

$$\begin{pmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & 0 & \dots & 0 \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ \sigma_z\sigma_1\rho & \sigma_z\sigma_2\rho & \sigma_z\sigma_3\rho & \dots & \sigma_z^2 \end{pmatrix} \quad \begin{pmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3^2 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma_z^2 \end{pmatrix}$$

(CS) (DIAG)

4.2. Estimation results

Following the discussion and approach in Antonio and Plat (2013), we fit the model separately for MD and BI payments. We perform data manipulations and likelihood optimization with R. We follow a procedure[5] inspired by Akdemir (2009) to estimate parameters in the MSN distribution. In a first step, we obtain shape parameters by the maximum product of spacings method (see Cheng and Amin, 1983 and Ranneby, 1984). Under general conditions, Cheng and Amin (1983) show that the maximum product of spacings estimator and the maximum likelihood estimator are asymptotically equivalent. In a second step, location and scale parameters follow from applying maximum likelihood method. In the likelihood optimization, numerical approximation of the Hessian matrix is used to estimate standard errors. For each component in the model, a model selection step is performed, comparing different models based on AIC and BIC. We highlight retained model specifications in gray in the tables that follow.

4.2.1. *Distributions for number of periods.* For the discrete random variables $\{T_{ik}\}$, $\{Q_{ik}\}$, $\{U_{ik}\}$ and $\{N_{ik}\}$, we investigate the use of a mixture of p degenerate distributions with a basic count distribution (see (15)). Consequently, $p+q+1$ parameters have to be estimated for each variable (with q the number of parameters used in the count distribution). Our model selection procedure (based on AIC and BIC) prefers a Geometric distribution, combined with degenerate components. Table 2 shows model selection steps assisting in the choice of the number of degenerate components. Table 3 displays parameter estimates and standard errors for the preferred specifications. Observed and estimated results

TABLE 5

BODILY INJURY: MODEL SELECTION STEPS EXAMINING MSN AND MN SPECIFICATIONS FOR THE DEVELOPMENT PATTERN VECTOR.

Model		No. of Parms.	$-ll$	AIC	BIC
MSN	UN	20	3,431	6,902	7,000
	TOEP	14	3,435	6,897	6,966
	CS	11	3,444	6,910	6,964
	DIA	10	3,605	7,230	7,279
MN	UN	20	3,496	7,032	7,128
	TOEP	14	3,499	7,025	7,094
	CS	11	3,531	7,083	7,137
	DIA	10	3,723	7,465	7,514

are compared in Figure 4, at least for the components necessary to project claims until settlement.

4.2.2. *Occurrence of claims.* Using the distributions selected for reporting delay, we estimate the thinned Poisson distribution from (1). Hereby, the exposure measure $w(.)$ is expressed in years. Results are: $\hat{\theta}_{BI} = 0.7445$ (s.e. 0.02) and $\hat{\theta}_{MD} = 38.96$ (s.e. 0.11).

4.2.3. *Development pattern.*

The development consists of a single payment. For the logarithm of the severity of the first and only payment, we explore the use of a Univariate Skew Normal (USN) as well as a Normal (N) distribution. The estimation results and a graphical goodness-of-fit check are in Table 4 and Figure 5. For the data at hand, the Normal distribution is to be preferred.

The development consists of more than one payment. We examine the use of the MSN, as well as the MN distribution for the logarithm of the development pattern vector $\boldsymbol{\Lambda}^{(ik)}_{U_{ik}+1}$ (see (2)).

For BI, we restrict the maximal dimension of the development vector, say m_p, to 5 and to $m_p = 3$ for MD.[6] Therefore, we fit a location vector of dimension $m_p \times 1$, a scale matrix of dimension $m_p \times m_p$ and a shape vector of dimension $m_p \times 1$. When observed claims use less development factors, appropriate sub-vectors and sub-matrices are used in the likelihood. If the simulated number of periods with payment is larger than m_p, we apply a tail factor.[7] Tables 5 and 6 (BI) and 7 and 8 (MD) present results of the model selection steps, as well as parameter estimates for the preferred MSN and the preferred MN distribution. For the MD case, we initially observed imprecision in dimension 3 estimates

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

TABLE 6

BODILY INJURY: PARAMETER ESTIMATES FOR PREFERRED MSN AND MN DISTRIBUTIONS.

MSN Model			MN Model	
Location $\boldsymbol{\mu}$ (s.e.)	Scale $\boldsymbol{\Sigma}_c^{1/2}$	Shape $\boldsymbol{\Delta}$	Location $\boldsymbol{\mu}$ (s.e.)	Scale $\boldsymbol{\Sigma}_c^{1/2}$
$\mu_1 = 5.44$	$\sigma_1 = 1.27$	$\Delta_1 = 0.51$	$\mu_1 = 6.04$	$\sigma_1 = 1.23$
(0.05)	$\sigma_2 = 1.18$	$\Delta_2 = 2.64$	(0.05)	$\sigma_2 = 0.97$
$\mu_2 = 0.53$	$\sigma_3 = 1.00$	$\Delta_3 = 2.29$	$\mu_2 = 1.43$	$\sigma_3 = 0.86$
(0.03)	$\sigma_4 = 0.83$	$\Delta_4 = -0.32$	(0.04)	$\sigma_4 = 0.82$
$\mu_3 = 0.63$	$\sigma_5 = 0.69$	$\Delta_5 = -0.002$	$\mu_3 = 0.95$	$\sigma_5 = 0.69$
(0.05)	$\rho = -0.28$		(0.05)	$\rho_1 = -0.49$
$\mu_4 = 1.49$			$\mu_4 = 0.64$	$\rho_2 = -0.23$
(0.09)			(0.08)	$\rho_3 = -0.003$
$\mu_5 = 1.12$			$\mu_5 = 0.66$	$\rho_4 = -0.26$
(0.10)			(0.11)	

TABLE 7

MATERIAL DAMAGE: MODEL SELECTION STEPS EXAMINING MSN AND MN SPECIFICATIONS FOR THE DEVELOPMENT PATTERN VECTOR.

Model		No. of Parms.	$-ll$	AIC	BIC
MSN	UN	9	4, 260	8, 538	8, 586
	TOEP	8	4, 282	8, 580	8, 622
	CS	7	4, 508	9, 031	9, 068
	DIA	6	4, 740	9, 492	9, 524
MN	UN	9	4, 260	8, 538	8, 586
	TOEP	8	4, 271	8, 557	8, 600
	CS	7	4, 510	9, 033	9, 071
	DIA	6	4, 743	9, 498	9, 530

TABLE 8

MATERIAL DAMAGE: PARAMETER ESTIMATES FOR PREFERRED MSN AND MN DISTRIBUTIONS.

MSN Model			MN Model	
Location $\boldsymbol{\mu}$ (s.e.)	Scale $\boldsymbol{\Sigma}_c^{1/2}$	Shape $\boldsymbol{\Delta}$	Location $\boldsymbol{\mu}$ (s.e)	Scale $\boldsymbol{\Sigma}_c^{1/2}$
$\mu_1 = 5.43$	$\sigma_{11} = 1.27$	$\Delta_1 = 0.01$	$\mu_1 = 5.43$	$\sigma_{11} = 1.27$
(0.03)	$\sigma_{22} = 1.07$	$\Delta_2 = 2.95$	(0.03)	$\sigma_{22} = 0.71$
$\mu_2 = 0.32$	$\sigma_{33} = 0.65$	$\Delta_3 = 2.95$	$\mu_2 = 1.12$	$\sigma_{33} = 0.42$
(0.02)	$\sigma_{12} = -0.55$		(0.02)	$\sigma_{12} = -0.66$
$\mu_3 = 0.32$	$\sigma_{13} = -0.27$		$\mu_3 = 1.12$	$\sigma_{13} = -0.41$
(0.02)	$\sigma_{23} = -0.04$		(0.02)	$\sigma_{23} = -0.12$

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

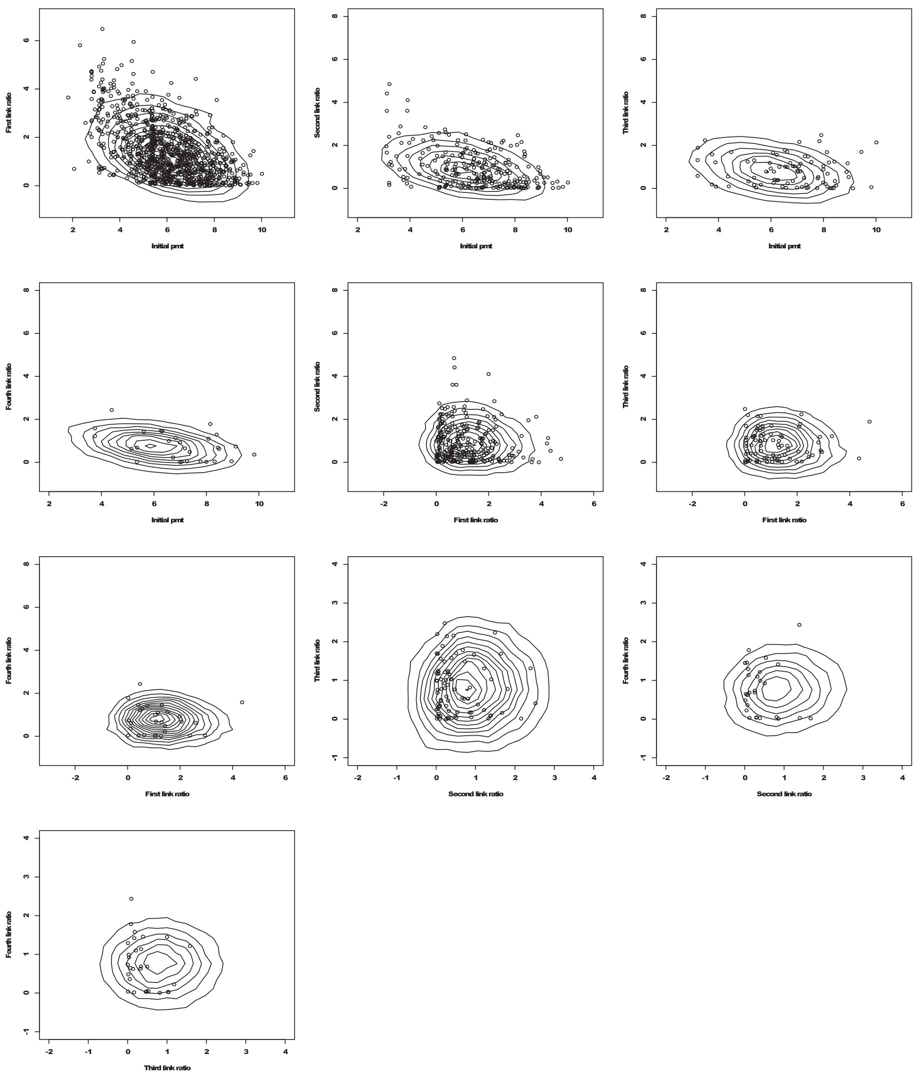

FIGURE 6: Bodily injury: empirical observations of the development vector (2) and contour plots obtained from selected MSN model. First row of plots (from left to right): first link ratio vs. initial payment, second link ratio vs. initial payment, third link ratio vs. initial payment. Second row (from left to right): fourth link ratio vs. initial payment, second vs. first link ratio, third vs. first link ratio. Third row (from left to right): fourth vs. first link ratio, third vs. second link ratio, fourth vs. second link ratio. Fourth row: fourth vs. third link ratio.

(namely $\mu_3 = 0.18$ with s.e. 0.20). Therefore, we reduced the M(S)N distribution by setting $\mu_2 = \mu_3$, $\Delta_2 = \Delta_3$ and by keeping the scale matrix unstructured. This modification improves the precision of our parameter estimates while having virtually no impact on the final results. Empirical data and contour plots for the chosen MSN multivariate densities are in Figures 6 (BI) and 7 (MD).

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

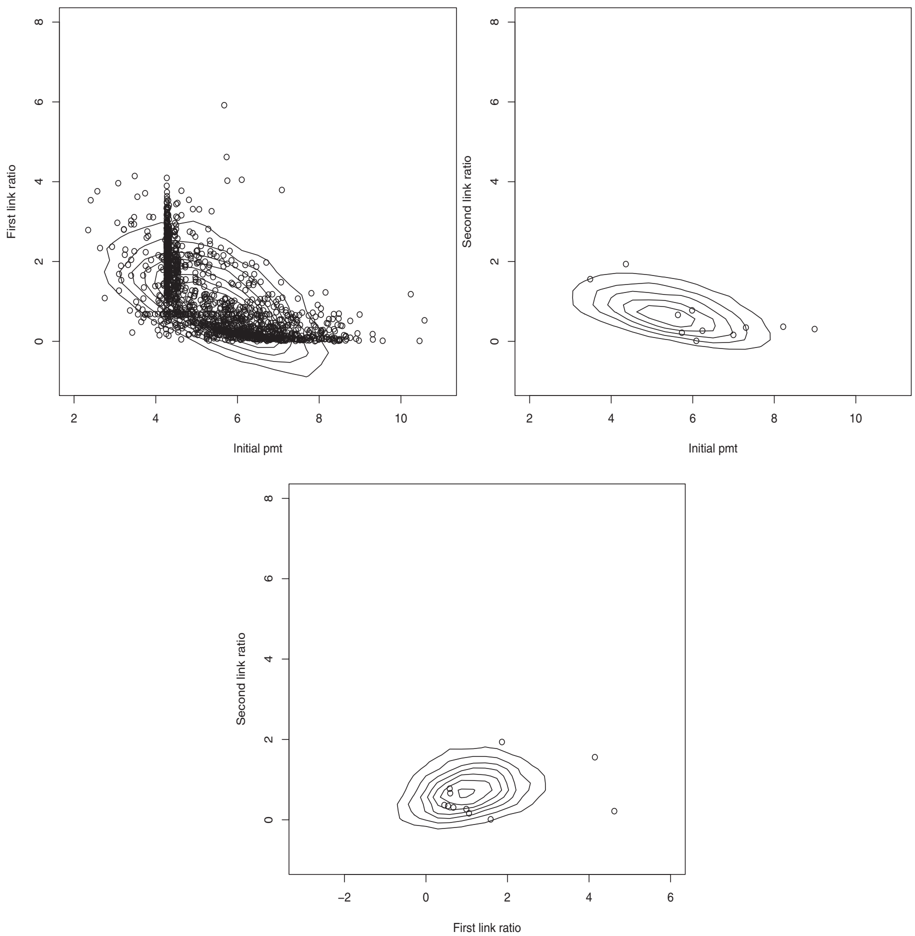

FIGURE 7: Material Damage: empirical observations of the development vector (2) and contour plots obtained from selected MSN model. First row of plots (from left to right): first link ratio vs. initial payment, second link ratio vs. initial payment. Second row: second vs. first link ratio.

5. PREDICTION RESULTS

We summarize the data set by occurrence and development year in run-off triangles, see Tables 9 and 10. Information with respect to occurrence years 2005 to 2009 (August) is available but not used in the analysis to enable out-of-sample prediction. This information is printed in bold in the run-off triangles.

5.1. Prediction of the IBNR and RBNP reserves

Analytical best estimate for outstanding IBNR and RBNP reserves. Analytical expressions for the expected value of the outstanding IBNR and RBNP reserve

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

TABLE 9

INCREMENTAL RUN-OFF TRIANGLE FOR BODILY INJURY (IN THOUSANDS).

	Development Year							
Arrival Year	1	2	3	4	5	6	7	8
1997	261	614	359	526	546	137	130	339
1998	202	473	307	336	269	56	179	**78**
1999	238	569	393	270	249	286	**132**	**97**
2000	237	557	429	496	406	**365**	**247**	**275**
2001	389	628	529	559	**446**	**375**	**147**	**239**
2002	260	570	533	**444**	**132**	**122**	**332**	**1, 082**
2003	236	743	**558**	**237**	**217**	**205**	**171**	
2004	248	**794**	**401**	**236**	**254**	**98**		

TABLE 10

INCREMENTAL RUN-OFF TRIANGLE FOR MATERIAL DAMAGE (IN THOUSANDS).

	Development Year							
Arrival Year	1	2	3	4	5	6	7	8
1997	4, 427	992	89	13	39	27	37	11
1998	4, 389	984	60	35	76	24	0.5	**16**
1999	5, 280	1, 239	76	110	113	12	**0.4**	**0**
2000	5, 445	1, 164	172	16	6	**10**	**0**	**10**
2001	5, 612	1, 838	156	127	**13**	**3**	**0.4**	**3**
2002	6, 593	1, 592	74	**71**	**17**	**15**	**9**	**9**
2003	6, 603	1, 660	**150**	**52**	**37**	**18**	**3**	
2004	7, 195	**1, 417**	**109**	**86**	**39**	**15**		

are available from Section 2.3, see Proposition 2.2, where unknown parameters should be replaced by estimates (as obtained in Section 4.2). Note that these expressions evaluate claims until settlement, even if this takes place beyond the boundary of the triangle. Moreover, these results do not incorporate parameter uncertainty (see Section 5.3 for further discussion). Table 12 displays these analytical results for BI and Table 13 for MD.

Simulation of outstanding IBNR and RBNP reserves. For each occurrence period, we simulate the number of IBNR claims (for BI and MD separately) from a Poisson distribution with occurrence measure

$$\hat{\theta} w(i)(1 - F_1(t_i^* - 1; \hat{\boldsymbol{\nu}})). \quad (16)$$

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

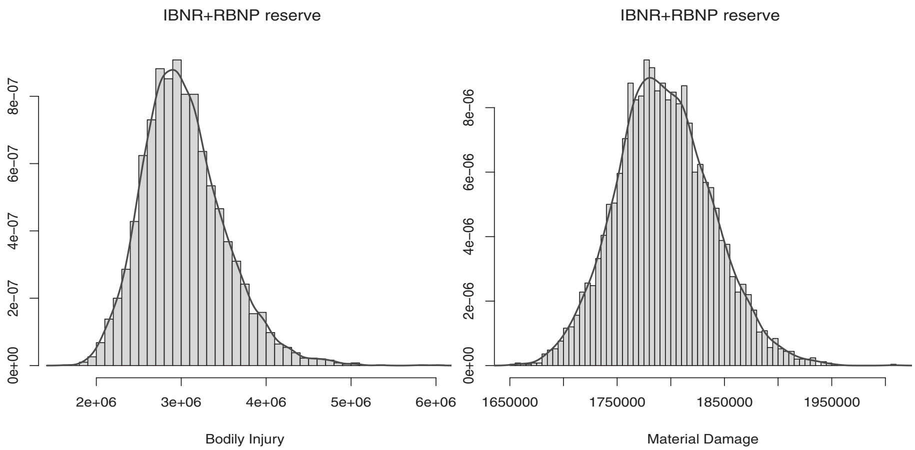

FIGURE 8: Histograms of the reserve obtained for IBNR and RBNP claims with the individual model for Bodily Injury (left) and Material Damage (right). Parameter uncertainty is not taken into account, claims are projected until settlement.

Consequently, for each IBNR claim (denoted with (ik)), we simulate the number of period(s) with partial payments U_{ik} and the corresponding development pattern vector $\mathbf{\Lambda}^{(ik)}_{U_{ik}}$. Note that — with this strategy — we develop a claim until settlement (which can be beyond the boundary of the triangle). Taking the timing of partial payments into account would require simulation of the random variables T_{ik}, Q_{ik} and N_{ikj} (see Tables 12 and 13 for results simulated until the boundary of the triangle).

The prediction routine for the RBNP reserve is similar to the routine for IBNR claims. However, the number of open RBNP claims is observed, and therefore does not require a simulation step. The variable Q_{ik} should be simulated from a truncated distribution, using the condition $Q_{ik} > t^*_{ik} - t_{ik} - 1$.

Graphical results based on 5,000 simulations are shown in Figure 8. Tables 12 (BI) and 13 (MD) display corresponding numerical results.

5.2. Prediction of the RBNS reserve

Analytical best estimate for outstanding RBNS reserve. Tables 12 (BI) and 13 (MD) display analytical results for BI and MD payments. Similar considerations apply as for IBNR and RBNP reserves.

Simulation of outstanding RBNS reserve. For each RBNS claim in the data set, we first simulate the number of period(s) with payment from the conditional pdf $f_3(u|u \geq u^*)$, where u^* is the observed number of periods with payment after the first one. Then, we simulate the missing part of the development pattern vector from the conditional MSN distribution (by conditioning on the observed part of the development pattern vector). Finally, we evaluate the RBNS reserve.

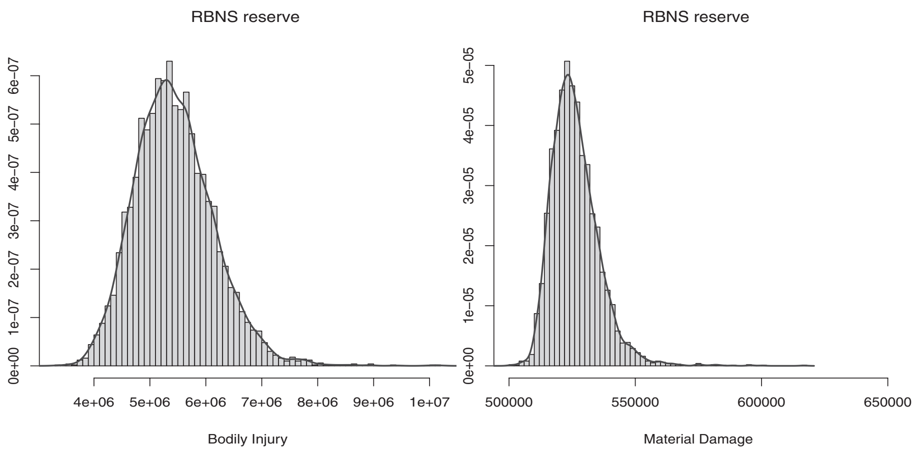

FIGURE 9: Histograms of the reserve obtained for RBNS claims with the individual model for Bodily Injury (left) and Material Damage (right). Parameter uncertainty is not taken into account, claims are projected until settlement.

Numerical results based on 5,000 simulations are in Tables 12 (BI) and 13 (MD) and corresponding graphical results are in Figure 9.

5.3. Discussion of results

Analytical versus simulation-based results, influence of parameter uncertainty, policy limit and design. Tables 12 (for BI) and 13 (for MD) show prediction results obtained with the individual claims reserving method. The *first two* "scenarios" in these tables display IBNR + RBNP (=IBNR$^+$), RBNS and Total reserves obtained with our preferred distributional assumptions (see Section 4.2.3) when claims are developed *until settlement*. Both analytical (first block of rows) and simulation-based results (second block of rows) are given. The analytical best estimate results are close to the mean of the corresponding predictive distribution based on simulation ("Sim."). This underpins the usefulness and appropriateness of the analytical formulas. The *third* block of rows shows simulation-based results, *including parameter uncertainty* ("Unc."). This is achieved by using the asymptotic normal distribution of the maximum likelihood estimators. At every iteration in the routine, we simulate for each building block in the model the parameter (vector) from its corresponding (uni/multi)variate asymptotic normal distribution, except for the (co)variance and the shape parameter(s) in the (MS)N distribution. This approach is inspired by Antonio and Plat (2013) and by Brouhns *et al.* (2002). Including parameter uncertainty with respect to (co)variance and shape parameters is a topic for future research. The *fourth* block of rows gives simulation-based results, including parameter uncertainty and developing claims until settlement, when the *policy limit* ("Pol. Limit") of 2.5 MEuro is taken into account (see Antonio and Plat (2013)). In a *fifth* block of rows, we present simulation-based results, accounting for parameter

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

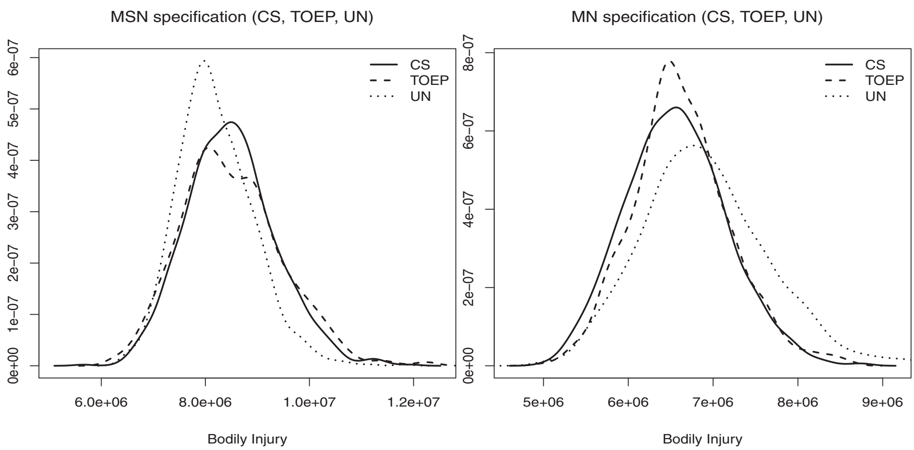

FIGURE 10: Bodily Injury: sensitivity of simulated predictive distributions with respect to the specification of the multivariate distribution for the development pattern vector. "MSN" refers to Multivariate Skew Normal and "MN" to Multivariate Normal. Parameter uncertainty is not taken into account, claims are projected until settlement and policy limit is not incorporated.

TABLE 11

SENSITIVITY OF ANALYTICAL BEST ESTIMATE RESULTS WITH RESPECT TO THE SPECIFICATION OF THE MULTIVARIATE DISTRIBUTION FOR THE DEVELOPMENT PATTERN VECTOR. "MSN" REFERS TO MULTIVARIATE SKEW NORMAL AND "MN" TO MULTIVARIATE NORMAL. PARAMETER UNCERTAINTY IS NOT TAKEN INTO ACCOUNT, CLAIMS ARE PROJECTED UNTIL SETTLEMENT AND POLICY LIMIT IS NOT INCORPORATED.

MSN	UN	BI	8,132,051	MD	2,320,735
	TOEP		8,476,498		2,331,575
	CS		8,404,192		2,339,406
MN	UN	BI	6,836,694	MD	2,327,497
	TOEP		6,578,931		2,327,733
	CS		6,547,580		2,345,500

uncertainty and policy limits, but restricting the development of claims to *the right boundary of the triangle* (i.e. development year 8), instead of developing claims until settlement.

Sensitivity analysis: distribution of development vector. According to Figure 10 (simulation based, for BI) and Table 11 (best estimate analytical results), the structure implied to $\boldsymbol{\Sigma}_c^{1/2}$ has minor impact on the expected value of the outstanding reserve, but it does influence its predictive distribution (obtained with MSN or MN assumption for (2)). The assumption of a MN versus MSN distribution for (2) also has a clear impact on the predictive distribution of the outstanding reserves, at least for BI payments. The impact is negligible for MD[8] (see Table 11). Recall from Tables 5 and 7 that all information criteria prefer the MSN distribution above the MN distribution. This sensitivity is a topic for future research.

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

TABLE 12

BODILY INJURY: COMPARISON OF ESTIMATION RESULTS. IBNR^+ DENOTES THE COMBINATION OF IBNR AND RBNP RESERVES. RESULTS ARE DISPLAYED FOR: ANALYTICAL BEST ESTIMATES (UNTIL SETTLEMENT OF EACH CLAIM), CORRESPONDING SIMULATION-BASED RESULTS, SIMULATION-BASED RESULTS INCORPORATING PARAMETER UNCERTAINTY ("UNC."), SIMULATION-BASED RESULTS ACCOUNTING FOR INDIVIDUAL POLICY LIMIT OF 2.5 M EURO ("POL. LIMIT"), SIMULATION-BASED RESULTS ACCOUNTING FOR POLICY LIMITS AND DEVELOPING UNTIL TRIANGLE BOUNDARY. CHAIN–LADDER RESULTS FOR TABLE 9 ARE DISPLAYED. OBSERVED AMOUNT (i.e. SUM OF BOLD NUMBERS IN TABLE 9) IS 7,684,000 EURO.

Model or Scenario	Item	Expected Value	S.E.	$\text{VaR}_{0.95}$	$\text{VaR}_{0.995}$
Individual MSN Analytical (Until Settlement)	IBNR^+	2, 970, 645			
	RBNS	5, 433, 548			
	Total	8, 404, 192			
Individual MSN Simulated (Until Settlement)	IBNR^+	3, 035, 519	494, 771	3, 912, 159	4, 673, 340
	RBNS	5, 439, 318	704, 701	6, 650, 958	7, 738, 003
	Total	8, 474, 837	853, 812	9, 927, 439	11, 105, 174
Individual MSN Sim. + Unc. (Until Settlement)	Total	8, 568, 506	922, 657	10, 134, 198	11, 406, 905
Individual MSN Sim. + Unc. + Pol. Limit (Until Settlement)	Total	8, 568, 355	902, 601	10, 141, 226	11, 320, 931
Individual MSN Sim. + Unc. + Pol. Limit (Until Triangle Bound)	Total	7, 251, 103	817, 878	8, 679, 618	9, 717, 771
Chain–Ladder (Bootstrap, ODP)	Total	9, 126, 639	1, 284, 793	11, 380, 743	13, 061, 937
Observed (Bold, Table 9)	Total	7, 684, 000			

Out-of-sample test and comparison with other reserving methods. The best estimate results reported in Tables 12 and 13 (simulation-based, until the boundary of the triangle and taking the policy limit into account) are close to the results obtained in Antonio and Plat (2013). Our out-of-sample test (see Figure 11) demonstrates the usefulness of the method developed in this paper. The realized outcomes are displayed in bold in the lower triangles in Tables 9 and 10. As discussed in Antonio and Plat (2013), the lower triangle for BI (see Table 9) shows an extreme payment (779,383 euro) in occurrence year 2002, development year 8. This is reflected in a realistic way by the individual loss reserving model. We also compare the results of the individual reserving method with the results of a reserving method for aggregated data. Figure 11 illustrates this comparison. Results for chain–ladder are obtained by bootstrapping an Overdispersed Poisson (ODP) model (with the `chainladder` package in `R`) for the incremental payments in a run-off triangle, i.e. X_{ij}, with i the arrival year and j the development year, (as

TABLE 13

MATERIAL DAMAGE: COMPARISON OF ESTIMATION RESULTS. IBNR$^+$ DENOTES THE COMBINATION OF IBNR AND RBNP RESERVES. RESULTS ARE DISPLAYED FOR: ANALYTICAL BEST ESTIMATES (UNTIL SETTLEMENT OF EACH CLAIM), CORRESPONDING SIMULATION-BASED RESULTS, SIMULATION-BASED RESULTS INCORPORATING PARAMETER UNCERTAINTY ("UNC."), SIMULATION-BASED RESULTS ACCOUNTING FOR INDIVIDUAL POLICY LIMIT OF 2.5 M EURO ("POL. LIMIT"), SIMULATION-BASED RESULTS ACCOUNTING FOR POLICY LIMITS AND DEVELOPING UNTIL TRIANGLE BOUNDARY. CHAIN–LADDER RESULTS FOR TABLE 10 ARE DISPLAYED. OBSERVED AMOUNT (i.e. SUM OF BOLD NUMBERS IN TABLE 9) IS 2,102,800 EURO.

Model or Scenario	Item	Expected Value	S.E.	$\text{VaR}_{0.95}$	$\text{VaR}_{0.995}$
Individual MSN	IBNR$^+$	1, 793, 545			
Analytical	RBNS	524, 945			
(Until Settlement)	Total	2, 318, 490			
Individual MSN	IBNR$^+$	1, 794, 044	44, 387	1, 869, 850	1, 917, 588
Simulated	RBNS	524, 984	15, 889	542, 146	562, 096
(Until Settlement)	Total	2, 318, 878	47, 109	2, 396, 281	2, 446, 362
Individual MSN Sim. + Unc. (Until Settlement)	Total	2, 443, 614	68, 364	2, 663, 801	2, 901, 795
Individual MSN Sim. + Unc. + Pol. Limit (Until Settlement)	Total	2, 442, 226	66, 472	2, 646, 649	2, 919, 822
Individual MSN Sim. + Unc. + Pol. Limit (Until Triangle Bound)	Total	2, 312, 270	70, 630	2, 429, 363	2, 497, 149
Chain–Ladder (Bootstrap, ODP)	Total	3, 053, 641	429, 266	3, 780, 437	4, 202, 478
Observed (Bold, Table 10)	Total	2, 102, 800			

in Table 9). This ODP model specifies: $X_{ij} = \phi Z_{ij}$ with $Z_{ij} \sim \text{POI}(\mu_{ij}/\phi)$ and $\mu_{ij} = \exp(\alpha_i + \beta_j)$.

6. CONCLUSIONS

This paper proposes a discrete time individual reserving model inspired by the chain–ladder model. The model is designed for a data set with the development of individual claims. Highlights of our contribution are twofold. Firstly, on a claim by claim as well as aggregate level, analytical expressions for the first moment of the outstanding reserve are available. Secondly, the predictive distribution of the outstanding reserve is available by simulation. The latter approach allows taking policy characteristics, such as a policy limit, into account. The case study performed on a real-life general liability insurance portfolio demonstrates the usefulness of the model.

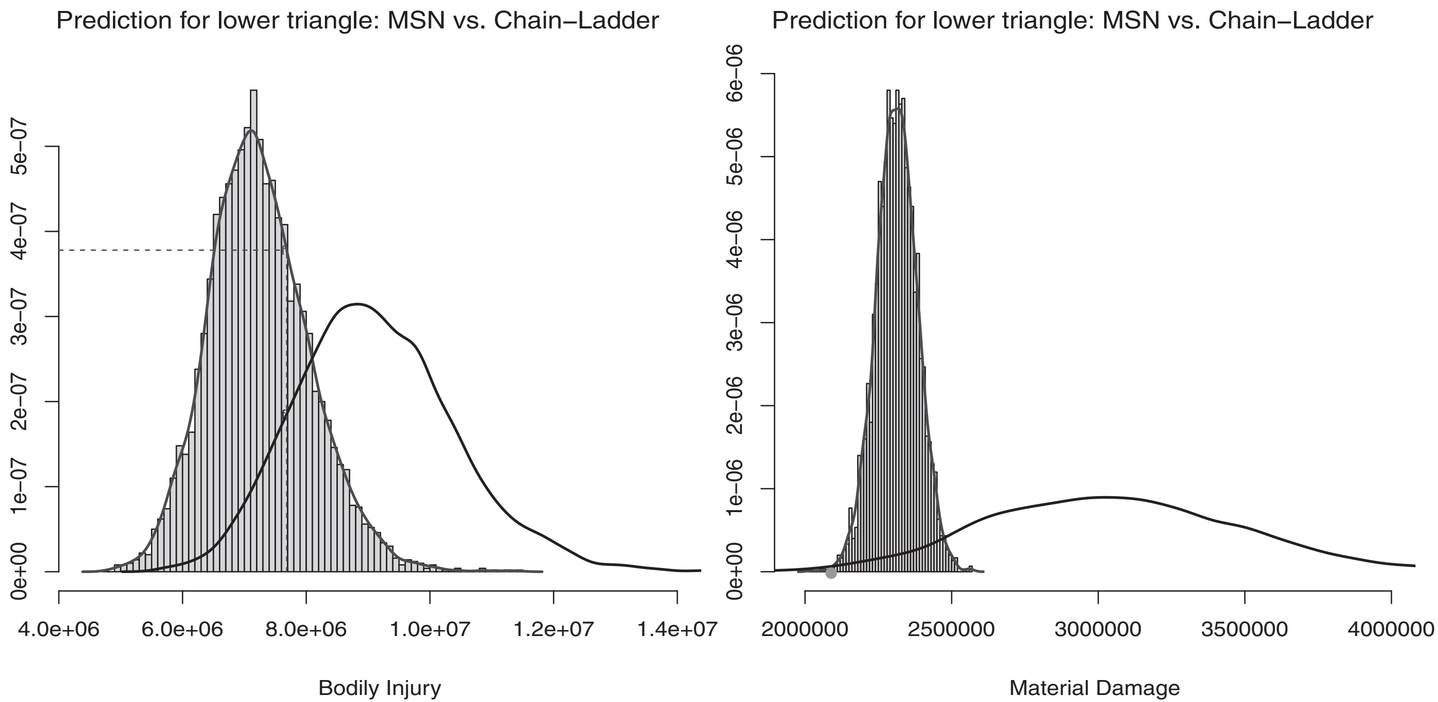

FIGURE 11: Histogram of the total reserve (light gray) obtained with the individual MSN model for Bodily Injury (left) and Material Damage (right). The histograms are based on 5,000 simulations (for BI) and 5,000 simulations (for MD) until the boundary of the triangle, taking parameter uncertainty and the policy limit into account. The black reference line is based on 5,000 bootstrap simulations of an Overdispersed Poisson model (with chain–ladder structure) for the Bodily Injury and Material Damage run-off triangle, respectively. Dotted lines (on the BI plot) and gray bullet (on the MD plot) represent the observed total payment for years 2005 to 2009 (August), i.e. the sum of the numbers in bold in Tables 9 and 10.

Several directions for future research can be envisaged. We plan further research with respect to the modeling of the first payment, using the Lognormal-Pareto distribution (see Pigeon and Denuit (2011)). Further investigation of the multivariate distribution for the development pattern vector is necessary, with an emphasis on research on inclusion of parameter uncertainty. Precise modeling of inflation effects and inclusion of the "time value of money" will be of importance in future work. Studying the approach in light of the new solvency guidelines, is another path to be explored, as well as extending the model to the reinsurance industry.

ACKNOWLEDGEMENTS

The authors thank two anonymous referees and the associate editor for useful comments which helped to improve the paper substantially. Katrien Antonio acknowledges financial support from the Casualty Actuarial Society, Actuarial Foundation and the Committee on Knowledge Extension Research of the Society of Actuaries, and from NWO through a Veni 2009 grant.

NOTES

1. The scale parameter $\mathbf{\Sigma}$ is not the usual variance–covariance matrix as in the Multivariate Normal distribution. An MSS random vector is defined by $\mathbf{\Sigma}^{1/2}$ in place of $\mathbf{\Sigma}$ because of the plurality of the square roots of $\mathbf{\Sigma}$. Without subscript, $\mathbf{\Sigma}^{1/2}$ designs any square root of the matrix $\mathbf{\Sigma}$.
2. As in Antonio and Plat (2013), we discount payments to 1/1/1997 with the appropriate consumer price index.
3. In contrast with Antonio and Plat (2013) a claim can have both BI payments, as well as MD payments. In Antonio and Plat (2013) a claim with at least one BI payment was considered as BI.
4. For MSN and MN, matrix $\Sigma_c^{1/2}$ refers to the square root of the covariance matrix $\mathbf{\Sigma}$, as obtained by the Cholesky decomposition.
5. A sample R program to verify this approach is available from the authors.
6. In the data set, we observe only eight BI claims with more than five periods with payment and two MD claims with more than three periods with payment.
7. This tail factor is the geometric average of empirically observed development factors.
8. Results for the Material Damage case are without constraint on $\boldsymbol{\mu}$ and $\mathbf{\Delta}$.

REFERENCES

AKDEMIR, D. (2009) *A Class of Multivariate Skew Distributions: Properties and Inferential Issues*. PhD thesis, Bowling Green State University, Ohio.

AKDEMIR, D. and GUPTA, A.K. (2010) A matrix variate skew distribution. *European Journal of Pure and Applied Mathematics*, **3**(2), 128–140.

ANTONIO, K. and PLAT, R. (2013) Micro–level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*, DOI:10.1080/03461238.2012.755938.

ARJAS, E. (1989) The claims reserving problem in non–life insurance: Some structural ideas. *ASTIN Bulletin*, **19**(2), 139–152.

AZZALINI, A. (1985) A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**, 171–178.

BROUHNS, N., DENUIT, M. and VERMUNT, J. K. (2002) A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, **31**, 373–393.

CHENG, J.T. and AMIN, N.A.K. (1983) Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society, Series* B *(Statistical Methodology)*, **45**(3), 394–403.
DRIESKENS, D., HENRY, M., WALHIN, J.-F. and WIELANDTS, J. (2012) Stochastic projection for large individual losses. *Scandinavian Actuarial Journal*, **1**, 1–39.
ENGLAND, P.D. and VERRALL, R.J. (2002) Stochastic claims reserving in general insurance. *British Actuarial Journal*, **8**, 443–544.
GUPTA, A.K. and CHEN, J.T. (2004) A class of multivariate skew normal models. *The Annals of the Institute of Statistical Mathematics*, **56**(2), 305–315.
HAPP, S. and WÜTHRICH, M. (2013) Paid–incurred chain reserving method with dependence modelling. *ASTIN Bulletin*, **43**(1), 1–20.
MACK, T. (1993) Distribution–free calculation of the standard error of chain–ladder reserve estimates. *ASTIN Bulletin*, **23**(2), 213–225.
MACK, T. (1999) The standard error of chain–ladder reserve estimates: Recursive calculation and inclusion of a tail-factor. *ASTIN Bulletin*, **29**, 361–366.
MARTINEZ, M.D., NIELSEN, B., NIELSEN, J.P. and VERRALL, R.J. (2011) Cash flow simulation for a model of outstanding liabilities based on claim amounts and claim numbers. *ASTIN Bulletin*, **41**(1), 107–129.
MARTINEZ, M.D., NIELSEN, J.P. and VERRALL, R.J. (2012) Double chain–ladder. *ASTIN Bulletin*, **42**(1), 59–76.
MERZ, M. and WÜTHRICH, M.V. (2010) PIC claims reserving method. *Insurance: Mathematics and Economics*, **46**(3), 568–579.
MURPHY, K. and MCLENNAN, A. (2006) A method for projecting individual large claims. *Casualty Actuarial Society Forum*, Fall Forum, 205–236.
NORBERG, R. (1993) Prediction of outstanding liabilities in non-life insurance. *ASTIN Bulletin*, **23**(1), 95–115.
NORBERG, R. (1999) Prediction of outstanding liabilities II. Model extensions variations and extensions. *ASTIN Bulletin*, **29**(1), 5–25.
PIGEON, M. and DENUIT, M. (2011) Composite lognormal-pareto model with random threshold. *Scandinavian Actuarial Journal*, **3**, 177–192.
RANNEBY, B. (1984) The maximum spacing method. An estimation method related to the maximum likelihood method. *Scandinavian Journal of Statistics*, **11**(2), 93–112.
ROBERTS, C. and GEISSER, S. (1966) A necessary and sufficient condition for the square of a random variable to be gamma. *Biometrika*, **53**(1/2), 275–278.
ROSENLUND, S. (2012) Bootstrapping individual claim histories. *ASTIN Bulletin*, **42**(1), 291–324.
VERRALL, R., NIELSEN, J.P. and JESSEN, A. (2010) Including count data in claims reserving. *ASTIN Bulletin*, **40**(2), 871–887.
WÜTHRICH, M.V. and MERZ, M. (2008) *Stochastic Claims Reserving Methods in Insurance*. Chichester, England: Wiley Finance.
ZHAO, X.B., ZHOU, X. and WANG, J.L. (2009) Semiparametric model for prediction of individual claim loss reserving. *Insurance: Mathematics and Economics*, **45**(1), 1–8.

MATHIEU PIGEON (Corresponding author)
Université Catholique de Louvain, UCL, Belgium
E-Mail: Mathieu.Pigeon@uclouvain.be

KATRIEN ANTONIO
KU Leuven, Belgium and University of Amsterdam, UvA, The Netherlands
E-Mail: K.Antonio@uva.nl

MICHEL DENUIT
Université Catholique de Louvain, UCL, Belgium
E-Mail: Michel.Denuit@uclouvain.be

APPENDIX

A. PROOF OF PROPOSITION 2.2

By definition,

$$\begin{aligned} M_{\ln(\mathbf{\Lambda})}(\mathbf{t}|u) &= E\left[\exp\left(\ln\left(\mathbf{\Lambda}_{U+1}\right)'\,\mathbf{t}\right)|U=u\right] \\ &= E[\exp\left(\ln\left(Y_1\right)t_1 + \ln\left(\lambda_1\right)t_2 + \ldots + \ln\left(\lambda_U\right)t_{U+1}\right)|U=u]\,. \end{aligned}$$

Taking $\mathbf{t} = \mathbf{t}_n = [n\ n\ \ldots\ n]'$, we obtain

$$\begin{aligned} M_{\ln(\mathbf{\Lambda})}(\mathbf{t}_n|u) &= E[\exp\left(n\left(\ln\left(Y_1\right) + \ln\left(\lambda_1\right) + \ldots + \ln\left(\lambda_U\right)\right)\right)|U=u] \\ &= E[(Y_1 \cdot \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_U)^n\,|U=u]\,. \end{aligned} \tag{A1}$$

For a conditional (on $U = u$) MSS random vector $((u+1)\times 1)$ and a $((u+1)\times 1)$ vector $\mathbf{t}$, the moment generating function is given by

$$\begin{aligned} M_{\ln(\mathbf{\Lambda})}(\mathbf{t}|u) &= \exp\left(\mathbf{t}_n'\boldsymbol{\mu}_{u+1}\right)\int\cdots\int_{\mathbb{R}^{u+1}} 2^{u+1}\cdot g^*(\mathbf{z}_{u+1}) \\ &\quad \cdot\exp\left(\left(\mathbf{\Sigma}_{u+1}^{1/2}\right)'\mathbf{t}_n'\mathbf{z}_{u+1}\right)\cdot\prod_{j=1}^{u+1}H\left(\Delta_j\mathbf{e}_j'\mathbf{z}_{u+1}\right)\,d\mathbf{z}_{u+1} \\ &= \exp\left(\mathbf{t}_n'\boldsymbol{\mu}_{u+1}\right)E_{g^*(\mathbf{z}_{u+1})}\left[\exp\left(\left(\mathbf{\Sigma}_{u+1}^{1/2}\right)'\mathbf{t}_n'\mathbf{z}_{u+1}\right)\cdot\prod_{j=1}^{u+1}H\left(\Delta_j\mathbf{e}_j'\mathbf{z}_{u+1}\right)\right] \end{aligned} \tag{A2}$$

with $\mathbf{z}_{u+1} = \mathbf{\Sigma}_{u+1}^{-1/2}(\ln\left(\mathbf{\Lambda}_{u+1}\right) - \boldsymbol{\mu}_{u+1})$.

The nth moment of an IBNR claim C is obtained by taking the expected value of equation (A1) with respect to random variable U and by including equation (A2)

$$\begin{aligned} E[C^n] &= E\left[E[(Y_1\cdot\lambda_1\cdot\lambda_2\cdot\ldots\cdot\lambda_U)^n\,|U=u]\right] \\ &= E\left[M_{\ln(\mathbf{\Lambda})}(\mathbf{t}_n|U)\right] \\ &= E\left[\exp\left(\mathbf{t}_n'\boldsymbol{\mu}_{U+1}\right)E_{g^*(\mathbf{z}_{U+1})}\left[\exp\left(\left(\mathbf{\Sigma}_{U+1}^{1/2}\right)'\mathbf{t}_n'\mathbf{z}_{U+1}\right)\right.\right. \\ &\quad \left.\left.\cdot\prod_{j=1}^{U+1}H\left(\Delta_j\mathbf{e}_j'\mathbf{z}_{U+1}\right)\right]\right]. \end{aligned} \tag{A3}$$

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press

For the specific case of a MSN distribution, the result becomes

$$E[C^n] = E\left[2^{U+1}\exp\left(\mathbf{t}_n'\boldsymbol{\mu}_{U+1} + 0.5\mathbf{t}_n'\boldsymbol{\Sigma}_{U+1}^{1/2}\left(\boldsymbol{\Sigma}_{U+1}^{1/2}\right)'\mathbf{t}_n\right)\right.$$
$$\left.\cdot\prod_{j=1}^{U+1}\Phi\left(\frac{\Delta_j\left(\left(\boldsymbol{\Sigma}_{U+1}^{1/2}\right)'\mathbf{t}_n\right)_j}{\sqrt{1+\Delta_j^2}}\right)\right]. \tag{A4}$$

B. PROOF OF PROPOSITION 2.3

Conditional on past information as well as the random variable U, the MSN random vector defined by

$$\ln\left(\boldsymbol{\Lambda}_B|\boldsymbol{\Lambda}_A = \boldsymbol{\ell}_A, U = u\right) = \left[\ln(y_1)\ \ln(\ell_1)\ \ldots\ \ln\left(\ell_{u_A-1}\right)\ \ln\left(\lambda_{u_A}\right)\ \ldots\ \ln(\lambda_u)\right] \tag{A5}$$

follows a MSN distribution with parameters $\boldsymbol{\mu}_{u+1}^*$, $\boldsymbol{\Sigma}_{u+1}^*$ and $\boldsymbol{\Delta}_{u+1}^*$ as defined in Proposition 2.3. This conditional result can be obtained in the same way as for a MN distribution. The rest of the proof is similar to the reasoning given in Section A.

C. PROOF OF PROPOSITION 2.4

(a) For IBNR claims, the expected value of the total claim amount is

$$E[\mathrm{IBNR}|\mathcal{I}] = E\left[\sum_{i=1}^{I}\sum_{k=1}^{K_{\mathrm{IBNR},i}} Y_1^{(ik)}\cdot\lambda_1^{(ik)}\cdot\ldots\cdot\lambda_{U_{ik}}^{(ik)}\right], \tag{A6}$$

where $K_{\mathrm{IBNR},i}$ is the random variable representing the number of IBNR claims from occurrence period i. Because $K_{\mathrm{IBNR},i}$ and $\boldsymbol{\Lambda}_{U+1}$ are independent, we obtain

$$\begin{aligned} E[\mathrm{IBNR}|\mathcal{I}] &= \sum_{i=1}^{I} E[K_{\mathrm{IBNR},i}]\,E\left[Y_1^{(ik)}\cdot\lambda_1^{(ik)}\cdot\ldots\cdot\lambda_{U_{ik}}^{(ik)}\right] \\ &= E[K_{\mathrm{IBNR}}]\cdot E\left[Y_1^{(ik)}\cdot\lambda_1^{(ik)}\cdot\ldots\cdot\lambda_{U_{ik}}^{(ik)}\right]. \end{aligned} \tag{A7}$$

The result then follows from Proposition 2.2. The proof is similar for RBNP claims.

(b) For RBNS claims, the expected value of the total claim amount is

$$\begin{aligned} E[\mathrm{RBNS}|\mathcal{I}] &= \sum_{(ik)_{\mathrm{RBNS}}} E\left[y_1^{(ik)}\cdot\ell_1^{(ik)}\cdot\ldots\cdot\ell_{u_A^{(ik)}-1}^{(ik)}\cdot\lambda_{u_A^{(ik)}}^{(ik)}\cdot\ldots\cdot\lambda_{U_{ik}}^{(ik)}\right] \\ &= \sum_{(ik)_{\mathrm{RBNS}}} y_1^{(ik)}\cdot\ell_1^{(ik)}\cdot\ldots\cdot\ell_{u_A^{(ik)}-1}^{(ik)}\cdot E\left[\lambda_{u_A}^{(ik)}\cdot\ldots\cdot\lambda_{U^{(ik)}}^{(ik)}\right]. \end{aligned} \tag{A8}$$

The proof then follows from Proposition 2.3.

https://doi.org/10.1017/asb.2013.20 Published online by Cambridge University Press