01 Jun 2022

Analysis of IBNR Liabilities with Interevent Times Depending on Claim Counts

Daniel J. Geiger

Akim Adekpedjou
Missouri University of Science and Technology, akima@mst.edu

Recommended Citation

D. J. Geiger and A. Adekpedjou, "Analysis of IBNR Liabilities with Interevent Times Depending on Claim Counts," *Methodology and Computing in Applied Probability*, vol. 24, no. 2, pp. 815 - 829, Springer, Jun 2022.
The definitive version is available at https://doi.org/10.1007/s11009-022-09950-5

Methodology and Computing in Applied Probability (2022) 24:815–829
https://doi.org/10.1007/s11009-022-09950-5

Analysis of IBNR Liabilities with Interevent Times Depending on Claim Counts

Daniel J. Geiger[1] · **Akim Adekpedjou**[1]

Received: 28 February 2021 / Revised: 22 February 2022 / Accepted: 27 February 2022 /
Published online: 22 June 2022

Abstract
We extend a recently proposed stochastic loss reserving model for liabilities from incurred but not reported (IBNR) micro-level claims. We propose viewing the number of claims from an event as a measure of catastrophic severity. This view covers catastrophes with arbitrarily many classes of magnitude. Our Markovian model allows the time between disasters to depend on the previous event's level of severity. Simultaneously, we let the discount rate vary in the same manner. First, we find the moments of IBNR liabilities in our model. Then, we permit a later time horizon for IBNR claims when considered jointly with incurred and reported claims.

Keywords IBNR · Markov renewal models · Markovian discount rates · Random thresholds · Catastrophes

1 Introduction

Tarbell (1934) introduced the idea of insurance claims being "incurred but not reported" (IBNR) and created the "chain ladder" to model reserves. Tarbell multiplied the previous year's IBNR claims amount by the ratio of change in the incurred losses amount. Continuing to reserve aggregated claims, Bornhuetter and Ferguson (1972) advocated calling "IBNR" future developments on reported claims. Of course, modeling individual (micro-level) claims could be more detailed than aggregated claims. Micro-level claims used to be studied separately from IBNR reserving, like via compound random sums. For example, Léveillé and Garrido (2001a, b) found the moments of a "compound renewal present value process" with a non-zero force of discount.

Micro-level modeling of IBNR claims has recently received increasing attention. See Landriault et al. (2017) for a review of the literature on such modeling. They considered a compound renewal claims process, incorporating reporting lags and the time value of money at a constant discount rate. Importantly, they introduced events with claim

✉ Daniel J. Geiger
djgpnb@mst.edu

[1] Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla MO 65409, USA

batches to address the intuitive fact that many claims may result from catastrophes like natural disasters or a pandemic. They primarily studied the discounted IBNR claim amounts. They also modeled these jointly with the discounted incurred and reported (IR) claim amounts.

In this paper, micro-level claims with IBNR dynamics interest us rather than earlier, classical aggregated approaches. We build upon the main results of Landriault et al. (2017) for IBNR and IR liabilities. For motivation, consider clusters of record-breaking hurricanes. In 2017, the USA experienced hurricanes Harvey, Irma, and Maria. In 2020, Central America faced hurricanes Eta and Iota. Seasonal patterns aside, such storms challenge the assumption that interevent times are independent and identically distributed (iid). Storms such as hurricanes and typhoons come in several categories of severity. Yet not even all category 5 hurricanes are equal.

In light of the foregoing, we propose to let the interevent times for bulk claim arrivals depend on the severity of the previous event. The claim batches are the main source of catastrophic insurer losses in this model. So, we measure the severity of a catastrophic event by the number of claims the event caused. But we follow the actuarial modeling convention of keeping separate the frequency and severity of the individual claims themselves. Within our approach, we further reflect the fact that natural disasters may occur within more than two orders of magnitude. While we specifically mentioned hurricanes, extreme weather events are not the only such natural disasters. For another example, consider earthquakes and their associated after-shocks, which may not be iid events. Our model allows any number of categories, or orders of catastrophic magnitude.

We articulate now precisely how we extend the claim-batch feature of Landriault et al. (2017). For every k in the nonnegative integers $\mathbb{N}$, they let C_k be the number of claims from event k. Extending this assumption, we partition $\mathbb{N}$ into (possibly random) intervals. To do so, we let v from the positive integers $\mathbb{N}^+$ represent a number of classes of severity. For each event k, we introduce a row vector $\boldsymbol{Q}_k = \left(Q_{1,k}, \dots, Q_{v,k}\right)$ of integers. For all k in $\mathbb{N}$, we assume that

$$0 = Q_{1,k} < Q_{2,k} < \cdots < Q_{v-1,k} < Q_{v,k} < \infty \tag{1}$$

almost surely. Hence, the vector $\boldsymbol{Q}_k$ associates v increasing thresholds with event k. For all k in $\mathbb{N}$, we assume C_k and $\boldsymbol{Q}_k$ are independent. Then the random variable J_k defined by

$$J_k = \sup\{l : Q_{l,k} \leq C_k\} \tag{2}$$

represents the level of catastrophic magnitude of event k, with severity increasing from levels 1 to v. Notice that defining J_k does not require explicitly specifying the joint distribution of the components in $\boldsymbol{Q}_k$.

We let the distribution of the time τ_k between events $k-1$ and k depend on the value of J_{k-1} for k in $\mathbb{N}^+$. This idea previously appeared in a ruin-theoretic context, without batches of claims. Albrecher and Boxma (2004) and Li and Sendova (2015) adjusted the claim arrival rate by comparing the size of each individual claim to a random threshold. They only modeled two risk classes, albeit hinting at the possibility of more classes. Our innovation, then, uses claim counts as an analogue and indeed allows many risk classes. Our approach turns the random threshold into random intervals. This Markovian relaxing of the usual renewal assumption in modeling IBNR liabilities from catastrophes is our primary contribution to the reserving literature.

Usually, the literature assumes money grows compounded continuously at a constant rate δ. In other words, a unit payment at time t has the present value $\exp(-\delta t)$. Stochastic

discount rates have received some attention, such as Léveillé and Adékambi (2011); Léveillé and Hamel (2019). However, they still ignored multi-claim events. We assume a particular stochastic discount rate structure. Consider central banks cutting interest rates early in the current COVID-19 pandemic; catastrophic scenarios can change discount rates. On the occurrence of event k at time T_k for k in $\mathbb{N}$, we assume a constant discount rate δ_{k+1} takes effect until event $k+1$. The rate δ_k is possibly random, and its distribution depends on the value of J_{k-1} in the same manner as the distribution of the interevent time τ_k. This is our secondary contribution.

Some other approaches to micro-level reserving have recently appeared in the literature. Wahl et al. (2019) assumed the number of individual claims followed an overdispersed Poisson distribution with each claim resulting in multiple payments by a Poisson process in discrete time. Lindholm and Zakrisson (2022) extended the model of Wahl et al. (2019) to allow claims to close and reopen. Like Wahl et al. (2019), Bischofberger et al. (2020) extended the chain ladder to individual claims. However, they used hazard rates associated with payment-producing claims to develop a continuous-time version of the chain ladder. In a completely different approach, Wüthrich (2018) applied regression trees to predict how many payments are made. For a parametric model, see Yanez and Pigeon (2021); for variations on the expectation-maximization algorithm, see Verbelen et al. (2018) and Fung et al. (2021). For other examples of the burgeoning interest in individual-level reserving, see Avanzi et al. (2021), Boucher (2021), Chaoubi et al. (2022), Duval and Pigeon (2019), and Delong et al. (2021). For surveys we refer readers to Taylor (2019) and Blier-Wong et al. (2021).

We organize the paper as follows: Sect. 2 articulates how we make our model Markovian, and we derive a generalized form of the renewal function. In Sect. 3, we give recursions for the finite-time moments of claims in our proposed model. We look at IBNR moments alone in Sect. 3.1, and the joint IBNR/IR moments in Sect. 3.2. Finally, we discuss some distributional assumptions allowing tractable computation and fitting our model to data in Sect. 4.

2 A Markov Renewal Model

2.1 Dependence Structure

Now we specify the Markovian structure we impose on the quantities defined in the introduction. We will use ν in $\mathbb{N}^+$ to mean the number of severity levels throughout the rest of this paper. Likewise, l will indicate the severity class, and k will refer to the event number. Recall the quantities $\{C_k\}_{k\geq 0}$ and $\{\boldsymbol{Q}_k\}_{k\geq 0}$ from the introduction. Landriault et al. (2017) assumed $\{C_k\}_{k\geq 0}$ are iid. Including their assumption, we suppose that the random vectors $\{\left(C_k, Q_{1,k}, \dots, Q_{\nu,k}\right)\}_{k\geq 0}$ are iid, which makes $\{J_k\}_{k\geq 0}$ iid.

Let T_k be the time of event k, and assume $T_0 = 0$ almost surely. Then the interevent times are $\tau_k = T_k - T_{k-1}$ (almost surely) for $k \geq 1$. We assume that T_0 and the $\{\tau_k\}_{k\geq 1}$ are mutually independent. For any given k, we assume that $\boldsymbol{Q}_k$ as a vector-valued random variable, C_k, and τ_k are mutually independent. Similarly to Albrecher and Boxma (2004), for $k \geq 0$ and $1 \leq l \leq \nu$ we let τ_{k+1} depend on C_k and $\boldsymbol{Q}_k$ according to

$$\Pr\left(\tau_{k+1} \leq x \mid J_k = l\right) = \Pr\left(\tau_1 \leq x \mid J_0 = l\right) := F_l(x). \tag{3}$$

Each F_l is a proper distribution function (df) with $F_l(0) = 0$.

As mentioned previously, we assume discount rates follow the same structure. Let $\{\delta_k\}_{k\geq 1}$ be a sequence of random variables. For $k \geq 0$ and $1 \leq l \leq \nu$, we assume δ_{k+1} depends upon C_k and $\boldsymbol{Q}_k$ according to

$$\Pr\left(\delta_{k+1} \leq z \mid J_k = l\right) = \Pr\left(\delta_1 \leq z \mid J_0 = l\right) := D_l(z). \tag{4}$$

Each $D_l(z)$ is a proper df with $D_l(z) = 0$ for $z < 0$. Note that we do not say $D_l(0) = 0$, to permit discount rates of 0. Apart from J_k, δ_{k+1} is independent of all other model quantities. Permitting each δ_k to be random leads to expressions containing terms of the form $\tilde{D}_l(s) = \int_{0-}^{\infty} e^{-sz} D_l(\mathrm{d}z)$. Here, $0-$ means to include any jump in $D_l(z)$ at $z = 0$.

2.2 A Generalized Markov Renewal Function

We define the standard counting process as $N(t) = \sup\{k : T_k \leq t\}$, i.e. the number of claim-causing events by time t. We denote the "Markov renewal function" given $J_0 = l$ by $H_l(t) = \mathrm{E}\left(N(t) \mid J_0 = l\right)$ for $1 \leq l \leq \nu$. Meanwhile, we will write the "delayed" Markov renewal function as $H_0(t)$ in Corollaries 2 and 3. More generally, we denote $\mathrm{E}_l(X) = \mathrm{E}\left(X \mid J_0 = l\right)$ for a given random variable (rv) X.

For our purposes, we prefer to define the function $H_{n|l}(t)$ for $n \in \mathbb{N}$, given by

$$H_{n|l}(t) = \mathrm{E}\left(\sum_{k=1}^{N(t)} \exp\left(-\sum_{j=1}^{k} n\delta_j\tau_j\right) \mid J_0 = l\right).$$

The reason for the n is because terms of the form $\exp\left(-n\delta_1\right)$ appear in the recursion relations given in Theorems 1 and 2. We interpret $H_{1|l}(t)$ (where $n = 1$) as the expected present value of the sum of unit payments made at each event time until t. More generally, $H_{n|l}(t)$ can be interpreted the same as $H_{1|l}(t)$ except with discount rates $n\delta_k$ instead of δ_k. Let $\mathrm{I}(A)$ be the indicator function of a set A. Notice that when $n = 0$ or $D_l(z) = \mathrm{I}(z \geq 0)$ for all $1 \leq l \leq \nu$, $H_{n|l}(t)$ is simply the Markov renewal function $H_l(t) = \mathrm{E}\left(N(t) \mid J_0 = l\right)$.

We need the following proposition in order to express our main results in Sect. 3 without requiring the conventional constant force of discount. To express $H_{n|l}(t)$ in Proposition 1, we introduce some more notation first. We write $\chi_{C|l}(z) = \mathrm{E}\left(z^{C_k}\,\mathrm{I}\left(J_k = l\right)\right)$ for all k in $\mathbb{N}$. The $\chi_{C|l}(z)$ form a discrete analog to the functions introduced by Albrecher and Boxma (2004). We use the subscript C in the $\chi_{C|l}(z)$ to emphasize the importance of the claim batch sizes. For $1 \leq l \leq \nu$, we denote $G_{n|l}(t) = \int_0^t \tilde{D}_l(nx)F_l(\mathrm{d}x)$. Then we write $G_n(t) = \chi_C(1)\boldsymbol{G}_n(t)$, in which the ν-dimensional row vector $\chi_C(1)$ has $\chi_{C|l}(1)$ in entry l, and the ν-dimensional column vector $\boldsymbol{G}_n(t)$ has $G_{n|l}(t)$ in entry l. We adopt the convention of suppressing the subscript l when $\nu = 1$ (as in $\tau_k \sim F$). So our definition of $G_n(t)$ makes sense. We write the convolution of two df's F and G as $F \star G$. For $q \geq 1$, $F^{\star q}$ is the q-fold convolution of F with itself. When $q = 0$, $F^{\star q}(t)$ becomes $\mathrm{I}(t \geq 0)$.

Proposition 1 *For all $n \in \mathbb{N}^+$ and each $1 \leq l \leq \nu$, the function $H_{n|l}$ may be expressed as $H_{n|l}(t) = G_{n|l}(t) + \sum_{q=1}^{\infty} G_{n|l} \star G_n^{\star q}(t)$.*

Proof (Proof of Proposition 1) To obtain $H_{n|l}(t)$ for $1 \leq l \leq \nu$, we use iterated expectation conditioning on τ_1, δ_1, and J_1. The key observation is that $N\left(t + \tau_1\right) - 1 \mid J_1 = l$ and $N(t) \mid J_0 = l$ have the same distribution. This follows from the regenerative form of F_l (c.f. Janssen and Manca (2006)). First we find

$$H_{n|l}(t) = G_{n|l}(t) + \sum_{m=1}^{\nu} \chi_{C|m}(1) \int_0^t H_{n|m}(t-x) G_{n|l}(\mathrm{d}x). \tag{5}$$

Let $\boldsymbol{I}$ be the identity matrix with ν rows and columns. Now, take the Laplace-Stieltjes transform $\tilde{H}_{n|l}(s)$ of $H_{n|l}(t)$ in Eq. (5). Writing the result in matrix form and rearranging, we get $\left(\boldsymbol{I} - \tilde{\boldsymbol{G}}_n(s)\chi_C(1)\right)\tilde{\boldsymbol{H}}_n(s) = \tilde{\boldsymbol{G}}_n(s)$. Here, $\tilde{H}_{n|l}(s)$ is in row l of the ν-dimensional column vector $\tilde{\boldsymbol{H}}_n(s)$.

We solve for $\tilde{\boldsymbol{H}}_n(s)$ next. By the Sylvester determinant theorem, the matrix $\boldsymbol{I} - \tilde{\boldsymbol{G}}_n(s)\chi_C(1)$ has the determinant $1 - \chi_C(1)\tilde{\boldsymbol{G}}_n(s) = 1 - \tilde{G}_n(s)$. For all s with positive real part, this determinant is non-zero, verified as follows. For each $1 \le m \le \nu$, $|\tilde{G}_{n|m}(s)| < \tilde{F}_m(0) = 1$ because $|\tilde{D}_m(nx)| \in (0,1)$ for all x with positive real part. Hence, $|\chi_C(1)\tilde{\boldsymbol{F}}_n(s)| < \sum_{m=1}^{\nu} \chi_{C|m}(1) = 1$, implying that $\boldsymbol{I} - \tilde{\boldsymbol{G}}_n(s)\chi_C(1)$ is invertible in the positive half-plane. Finally,

$$\begin{aligned}\tilde{\boldsymbol{H}}_n(s) &= \left(\boldsymbol{I} - \tilde{\boldsymbol{G}}_n(s)\chi_C(1)\right)^{-1}\tilde{\boldsymbol{G}}_n(s) = \sum_{q=0}^{\infty}\left(\tilde{\boldsymbol{G}}_n(s)\chi_C(1)\right)^q\tilde{\boldsymbol{G}}_n(s) \\ &= \tilde{\boldsymbol{G}}_n(s)\sum_{q=0}^{\infty}\left(\chi_C(1)\tilde{\boldsymbol{G}}_n(s)\right)^q = \tilde{\boldsymbol{G}}_n(s)\left(1 - \chi_C(1)\tilde{\boldsymbol{G}}_n(s)\right)^{-1}.\end{aligned} \tag{6}$$

From Eq. (6), we can write

$$\chi_C(1)\tilde{\boldsymbol{G}}_n(s) = \tilde{G}_n(s) \implies \tilde{H}_{n|l}(s) = \tilde{G}_{n|l}(s)\frac{1}{1-\tilde{G}_n(s)},$$

where $\tilde{H}_{n|l}(s)$ is the entry in row l of $\tilde{\boldsymbol{H}}_n(s)$. Inverting the Laplace-Stieltjes transforms leads to

$$H_{n|l}(t) = G_{n|l}(t) + G_{n|l} \star \sum_{q=1}^{\infty} G_n^{\star q}(t) = G_{n|l}(t) + \sum_{q=1}^{\infty} G_{n|l} \star G_n^{\star q}(t).$$

□

To address the "delayed" Markov renewal case, we suppose that $\tau_1 \sim F_0$ and $\delta_1 \sim D_0$. Without assuming the proper df F_0 to be one of the F_l for $1 \le l \le \nu$, we will still impose $F_0(0) = 0$. Let $G_{n|0}(t) = \int_0^t \tilde{D}_0(nx)F_0(\mathrm{d}x)$ to express Corollary 1.

Corollary 1 *Let F_0 and D_0 be arbitrary proper df's with support on $(0,\infty)$ and $[0,\infty)$, respectively. The function $H_{n|0}(t) = \mathrm{E}\left(\sum_{k=1}^{N(t)} \exp\left(-\sum_{j=1}^{k} n\delta_j\tau_j\right) \mid J_0 = 0\right)$ may be expressed $H_{n|0}(t) = G_{n|0}(t) + \sum_{q=1}^{\infty} G_{n|0} \star G_n^{\star q}(t)$.*

Proof (Proof of Corollary 1) We condition on τ_1, δ_1, and J_1 again. This time, we have 0 instead of l in Eq. (5). Taking Laplace-Stieltjes transforms, we notice

$$\begin{aligned}\tilde{H}_{n|0}(s) &= \tilde{G}_{n|0}(s) + \chi_C(1)\tilde{\boldsymbol{H}}_n(s)\tilde{G}_{n|0}(s) \\ &= \tilde{G}_{n|0}(s) + \chi_C(1)\tilde{\boldsymbol{G}}_n(s)\Big(1 - \chi_C(1)\tilde{\boldsymbol{G}}_n(s)\Big)^{-1}\tilde{G}_{n|0}(s) \\ &= \tilde{G}_{n|0}(s) + \tilde{G}_{n|0}(s)\big(1 - \tilde{G}_n(s)\big)^{-1}\tilde{G}_n(s).\end{aligned} \tag{7}$$

In the second equation of Eq. (7), we substituted the last equality of Eq. (6). Recall that $\chi_C(1)\tilde{\boldsymbol{G}}_n(s) = \tilde{G}_n(s)$ is scalar. Laplace-Stieltjes inversion completes the proof. □

3 Modeling IBNR Liabilities for Several Severity Classes

This section contains the main results for our model. Section 3.1 addresses IBNR claims alone, while Sect. 3.2 covers IBNR and IR claims jointly.

Now we recall some quantities from Landriault et al. (2017). First, we introduce the claim severities and reporting lags. Consider the event k at time T_k. For claim i, denote the corresponding deflated severity and reporting lag by $X_{i,k}$ and $W_{i,k}$, respectively. We assume $\{X_{i,k}\}$ and $\{W_{i,k}\}$ are both iid, with respective df's P and K. Further, let $X_{i,m}$ and $W_{j,n}$ be independent whenever either $i \neq j$ or $m \neq n$.

We denote the conditional df of $W_{i,k}|\tau_k = t$ by $K_{W|\tau}(w|t)$, and of $X_{i,k}|\tau_k = t, W_{i,k} = w$ by $P_{X|\tau,W}(x|t,w)$. Thus, the joint distribution of $\left(\tau_k, W_{i,k}, X_{i,k}\right)$, conditioned upon $J_{k-1} = l$, is $F_l(t)K_{W|\tau}(w|t)P_{X|\tau,W}(x|t,w)$, extending Landriault et al. (2017). We will later use the functions which they called $\xi_n(x,t)$ and $\eta_n(x,t)$. As they did, we say $\ell(\cdot)$ is some nonnegative function of the lag $W_{i,k}$. This $\ell(\cdot)$ is meant to incorporate scenarios around the lags, such as inflation at rate ϵ: that is, $\ell(w) = e^{-\epsilon w}$. Note that when we speak of claim liabilities in this paper, we mean $\ell\left(W_{i,k}\right)X_{i,k}$. Then, with $\mu_n(t,w) = \mathrm{E}\left(X_{i,k}^n|\tau_k = t, W_{i,k} = w\right)$ for $n \in \mathbb{N}^+$, the functions ξ_n and η_n are

$$\xi_n(x,t) = \int_{t-x}^{\infty} (\ell(w))^n \mu_n(x,w)K_{W|\tau}(\mathrm{d}w|x),\ 0 \le x \le t, \tag{8}$$

$$\eta_n(x,t) = \int_{0-}^{t-x} (\ell(w))^n \mu_n(x,w)K_{W|\tau}(\mathrm{d}w|x),\ 0 \le x \le t. \tag{9}$$

By assumption, $\left(\tau_k, W_{i,k}, X_{i,k}\right)|J_{k-1} = l$ and $\left(\tau_1, W_{1,1}, X_{1,1}\right)|J_0 = l$ have the same joint distribution for all $k \ge 1$. So, it is easy to see from Eqs. (8) and (9) that

$$\begin{aligned}\xi_n(x,t) &= \mathrm{E}\left([\ell\left(W_{1,1}\right)X_{1,1}]^n\,\mathrm{I}\left(W_{1,1} > t-x\right)|\tau_1 = x\right), \\ \eta_n(x,t) &= \mathrm{E}\left([\ell\left(W_{1,1}\right)X_{1,1}]^n\,\mathrm{I}\left(W_{1,1} \le t-x\right)|\tau_1 = x\right).\end{aligned}$$

Therefore, we may interpret $\xi_n(x,t)$ (respectively $\eta_n(x,t)$) as the nth moment of the non-discounted IBNR (respectively IR) liability from one claim at the first event time, given $\tau_1 = x \le t$.

3.1 Moments of IBNR Claims

We want to study the total IBNR liabilities valued at time 0. Incorporating the Markovian discount rate structure, we write these IBNR liabilities $Z(t)$ as

$$Z(t) = \sum_{k=1}^{N(t)} \exp\left(-\sum_{j=1}^{k} \delta_j \tau_j\right) \sum_{i=1}^{C_k} \ell\left(W_{i,k}\right) \mathrm{I}\left(W_{i,k} + T_k > t\right) X_{i,k}.$$

With $\gamma \geq 0$, we write the Laplace transform of $Z(t)$ given $J_0 = l$, for $1 \leq l \leq \nu$, as $\tilde{L}_{\gamma|l}(t) = \mathrm{E}_l(\exp(-\gamma Z(t)))$. Further, we signify the Laplace transform of the rv X (conditioned upon τ and W) by $\tilde{P}_{X|\tau,W}(\gamma|x,w) = \mathrm{E}\left(e^{-\gamma X} \mid \tau = x, W = w\right)$. Before analyzing $Z(t)$, first let the function $\zeta_\gamma(x,t;z)$ be given by

$$\zeta_\gamma(x,t;z) = 1 + \int_{t-x}^{\infty} \left(\tilde{P}_{X|\tau,W}(\gamma e^{-xz}\ell(w)|x,w) - 1\right) K_{W|\tau}(\mathrm{d}w|x),\ 0 \leq x \leq t.$$

This ζ_γ is the Laplace transform of the (discounted) IBNR liability from one claim at the first event time, given $\delta_1 = z$ and $\tau_1 = x \leq t$. Extending (Landriault et al. (2017), Eq. (3.1)) we find, for each $1 \leq l \leq \nu$, that

$$\tilde{L}_{\gamma|l}(t) = \bar{F}_l(t) + \int_0^t \int_{0-}^{\infty} \chi_C\left(\zeta_\gamma(x,t;z)\right) \tilde{\boldsymbol{L}}_{\gamma e^{-xz}}(t-x) D_l(\mathrm{d}z) F_l(\mathrm{d}x). \tag{10}$$

Now, Eq. (10) is “Markov renewal-like”, but not truly of Markov renewal form. On the right-hand side, x appears in the Laplace transform parameter value γe^{-xz}. Similar comments apply to $\zeta_\gamma(x,t;z)$. But if we differentiate Eq. (10) n times with respect to (wrt) γ and set $\gamma = 0$, we do get a Markov renewal-form equation for the nth moment of $Z(t)$. So, we find a Markovian extension of the renewal-form Eqs. (3.2) and (3.3) of Landriault et al. (2017). According to our notational conventions, $\mathrm{E}_l(Z^n(t))$ signifies the nth moment of $Z(t)$, given $J_0 = l$. For matrix usage, we also denote $M_{n|l}(t) = \mathrm{E}_l(Z^n(t))$. We continue writing $l = 0$ to mean the delayed Markov case ($\tau_1 \sim F_0$ and $\delta_1 \sim D_0$).

For $1 \leq k \leq r \leq n$, let $\boldsymbol{\xi}(x,t)$ be a $(r-k+1)$-dimensional row vector with the ith entry given by Eq. (8). Denote the Bell polynomial in $\boldsymbol{\xi}(x,t)$ by $B_{r,k}(\boldsymbol{\xi}(x,t))$. This function $B_{r,k}$ arises when applying “Faà di Bruno’s formula” (e.g. (Johnson (2002), Eq. (2.2))) to $\chi_C\left(\zeta_\gamma(x,t)\right)$. Define $B^*_{r|l}(x,t) = \sum_{k=1}^{r} \chi^{(k)}_{C|l}(1) B_{r,k}(\boldsymbol{\xi}(x,t))$ for $1 \leq l \leq \nu$ and $0 \leq x \leq t$. Then let $\boldsymbol{B}^*_r(x,t)$ be the ν-dimensional row vector with $B^*_{r|l}(x,t)$ in entry l. Equations (3.2) and (3.3) of Landriault et al. (2017) extend to our delayed Markov renewal model as follows:

$$\begin{aligned} M_{n|0}(t) &= \int_0^t \chi_C(1) \boldsymbol{M}_n(t-x) G_{n|0}(\mathrm{d}x) + v_{n|0}(t); \\ v_{n|0}(t) &= \sum_{r=1}^{n} \binom{n}{r} \int_0^t \boldsymbol{B}^*_r(x,t) \boldsymbol{M}_{n-r}(t-x) G_{n|0}(\mathrm{d}x). \end{aligned} \tag{11}$$

In Eq. (11), $M_{n|l}(t)$ is the entry in row l of the ν-dimensional column vector $\boldsymbol{M}_n(t)$. For all $0 \leq l \leq \nu$, we set $\mathrm{E}_l\left(Z^0(t)\right) \equiv 1$ for $t > 0$. Now we may solve for $\mathrm{E}_0(Z^n(t))$ as a certain integral with respect to the generalization $H_{n|0}$ of the Markov renewal function introduced in Corollary 1. The following theorem generalizes (Landriault et al. (2017), Theorem 1), their main result for the IBNR moments. The ν-dimensional column vector $\boldsymbol{v}_n(t)$ has the entry $v_{n|l}(t)$ in row l.

Theorem 1 *Let $n \in \mathbb{N}^+$, $\tau_1 \sim F_0$, and $\delta_1 \sim D_0$. Then, we have*:

$$\mathrm{E}_0(Z^n(t)) = v_{n|0}(t) + \int_0^t \chi_C(1)\boldsymbol{v}_n(t-x)H_{n|0}(\mathrm{d}x). \tag{12}$$

Proof (Proof of Theorem 1) Setting $F_0 = F_l$ for each $1 \le l \le v$, we express Eq. (11) in matrix form as

$$\boldsymbol{M}_n(t) = \int_0^t \chi_C(1)\boldsymbol{M}_n(t-x)\boldsymbol{G}_n(\mathrm{d}x) + \boldsymbol{v}_n(t).$$

We follow the arguments regarding Eq. (6) to solve for the Laplace transform $\hat{\boldsymbol{M}}_n(s)$ of $\boldsymbol{M}_n(t)$. We get

$$\hat{\boldsymbol{M}}_n(s) = \hat{\boldsymbol{v}}_n(s) + \frac{\chi_C(1)\hat{\boldsymbol{v}}_n(s)\tilde{\boldsymbol{G}}_n(s)}{1-\chi_C(1)\tilde{\boldsymbol{G}}_n(s)}. \tag{13}$$

Notice in Eq. (13) that the matrix product $\chi_C(1)\hat{\boldsymbol{v}}_n(s)$ is scalar. Therefore, we can premultiply Eq. (13) by $\chi_C(1)$ to get

$$\begin{aligned}\chi_C(1)\hat{\boldsymbol{M}}_n(s) &= \chi_C(1)\hat{\boldsymbol{v}}_n(s) + \chi_C(1)\hat{\boldsymbol{v}}_n(s)\frac{\chi_C(1)\tilde{\boldsymbol{G}}_n(s)}{1-\chi_C(1)\tilde{\boldsymbol{G}}_n(s)}\\ &= \frac{\chi_C(1)\hat{\boldsymbol{v}}_n(s)}{1-\tilde{G}_n(s)},\end{aligned} \tag{14}$$

since $\chi_C(1)\tilde{\boldsymbol{G}}_n(s) = \tilde{G}_n(s)$. Now we substitute Eq. (14) into the Laplace transform of Eq. (11):

$$\begin{aligned}\hat{M}_{n|0}(s) &= \hat{v}_{n|0}(s) + \chi_C(1)\hat{\boldsymbol{M}}_n(s)\tilde{G}_{n|0}(s)\\ &= \hat{v}_{n|0}(s) + \chi_C(1)\hat{\boldsymbol{v}}_n(s)\tilde{H}_{n|0}(s).\end{aligned} \tag{15}$$

The second step of Eq. (15) uses Eq. (7). Laplace transform inversion gives Eq. (12).

□

Note that $\mathrm{E}_0(Z^n(t))$ depends on $\mathrm{E}_l(Z^r(t))$ for all $1 \le l \le v$ and $1 \le r < n$. This shows that the distributions F_l for all v risk classes collectively affect the IBNR moments. We can simplify Theorem 1 in two ways. First, let the discount rate be constant. Second, assume there is only one risk level, meaning the renewal assumption of Landriault et al. (2017).

Corollary 2 *Let $n \in \mathbb{N}^+$. Let τ_1 and δ_1 have arbitrary df's F_0 and D_0, respectively.*

1. *Assume the discount rate is constant, namely $D_l(\cdot) \equiv \mathrm{I}(\cdot \ge \delta)$ for $0 \le l \le v$. Then,*

$$\mathrm{E}_0(Z^n(t)) = v_{n|0}(t) + \int_0^t e^{-n\delta x}\chi_C(1)\boldsymbol{v}_n(t-x)H_0(\mathrm{d}x). \tag{16}$$

2. *Assume for $k \ge 2$ that τ_k are iid with df F and δ_k are iid with df D, equivalent to setting $v = 1$. Then,*

$$\mathrm{E}_0(Z^n(t)) = v_{n|0}(t) + \int_0^t v_n(t-x)H_{n|0}(\mathrm{d}x). \tag{17}$$

Both cases of Corollary 2 subsume (Landriault et al. (2017), Theorem 1). In Eq. (16), first set $F_0 = F_l$ for each $1 \le l \le \nu$ to get the ordinary Markov renewal case. The resulting matrix equation is $\boldsymbol{M}_n(t) = \boldsymbol{v}_n(t) + \int_0^t e^{-n\delta x}\boldsymbol{\chi}_C(1)\boldsymbol{v}_n(t-x)\boldsymbol{H}(\mathrm{d}x)$. Now let there be only one risk level. When $\nu = 1$, recall $\boldsymbol{\chi}_C(1)$ becomes $\chi_C(1) = \mathrm{E}\left(1^C\right) = 1$. Each of $\boldsymbol{M}_n$, $\boldsymbol{v}_n$, and $\boldsymbol{H}$ becomes scalar, retrieving (Landriault et al. (2017), Theorem 1). In Eq. (17), imposing the additional assumption $D_0(\cdot) = D(\cdot) = \mathrm{I}(\cdot \ge \delta)$ means that $\int_0^\infty e^{-sx}G_n(\mathrm{d}x) = \tilde{F}(s+n\delta)$. Setting $F_0 = F$, Corollary 1 implies that

$$\tilde{H}_{n|0}(s) = \frac{\tilde{F}(s+n\delta)}{1-\tilde{F}(s+n\delta)} = \tilde{H}(s+n\delta) = \int_0^\infty e^{-sx}e^{-n\delta x}H(\mathrm{d}x).$$

We have again found (Landriault et al. (2017), Theorem 1).

Proof (Proof of Corollary 2) For both cases, the proof involves Laplace transformation of Eq. (11), rearrangement, and Laplace inversion.

1. Under a constant discount rate δ, we see that $G_{n|0}(\mathrm{d}x) = e^{-n\delta x}F_0(\mathrm{d}x)$. This $G_{n|0}$ has Laplace-Stieltjes transform $\tilde{F}_0(s+n\delta)$. The Laplace transform of Eq. (11) becomes

$$\hat{M}_{n|0}(s) = \tilde{F}_0(s+n\delta)\boldsymbol{\chi}_C(1)\hat{\boldsymbol{M}}_n(s) + \hat{v}_{n|0}(s). \tag{18}$$

   Setting $F_0 = F_l$ for each $1 \le l \le \nu$, we may write Eq. (18) as a matrix equation for $\hat{\boldsymbol{M}}_n(s)$. So, follow the derivation of $\tilde{H}_{n|l}(s)$ in Sect. 2 to solve $\hat{\boldsymbol{M}}_n(s) = \tilde{\boldsymbol{F}}(s+n\delta)\boldsymbol{\chi}_C(1)\hat{\boldsymbol{M}}_n(s) + \hat{\boldsymbol{v}}_n(s)$ for $\hat{\boldsymbol{M}}_n(s)$. As a result, $\hat{\boldsymbol{M}}_n(s)$ satisfies

$$\hat{\boldsymbol{M}}_n(s) = \hat{\boldsymbol{v}}_n(s) + \boldsymbol{\chi}_C(1)\hat{\boldsymbol{v}}_n(s)\tilde{\boldsymbol{H}}(s+n\delta).$$

   Substitute this solution for $\hat{\boldsymbol{M}}_n(s)$ into Eq. (18) and repeat the steps used to derive Eq. (15) in the proof of Theorem 1. Then Eq. (16) follows upon Laplace transform inversion.
2. When $\nu = 1$, the vectors in Eq. (11) revert to scalars. Now Eq. (11) possesses the Laplace transform $\hat{M}_{n|0}(s) = \hat{M}_n(s)\tilde{G}_{n|0}(s) + \hat{v}_{n|0}(s)$. Letting $F_0 = F$, we get $\hat{M}_n(s) = \hat{M}_n(s)\tilde{G}_n(s) + \hat{v}_n(s)$. This "ordinary renewal" case has the solution $\hat{M}_n(s) = \hat{v}_n(s)\sum_{q=0}^\infty \left(\tilde{G}_n(s)\right)^q$. We insert $\hat{M}_n(s)$ into $\hat{M}_{n|0}(s)$ and recall by Corollary 1 that $\tilde{H}_{n|0}(s) = \tilde{G}_{n|0}(s)\sum_{q=0}^\infty \left(\tilde{G}_n(s)\right)^q$. Then Eq. (17) corresponds to Laplace inversion of $\hat{M}_{n|0}(s) = \hat{v}_{n|0}(s) + \hat{v}_n(s)\tilde{H}_{n|0}(s)$. □

3.2 Joint Moments of IBNR and IR Claims

We shall similarly generalize the main result of Landriault et al. (2017) for the IBNR and IR joint moments. As a counterpart to $Z(t)$ in Sect. 3.1, we define

$$Z_{ir}(t) = \sum_{k=1}^{N(t)} \exp\left(-\sum_{j=1}^{k}\delta_j\tau_j\right)\sum_{i=1}^{C_k}\ell\left(W_{i,k}\right)\mathrm{I}\left(W_{i,k}+T_k \le t\right)X_{i,k}.$$

Then $Z_{ir}(t)$ represents the total liabilities incurred and reported by time t, discounted to time 0. Letting $\Delta \geq 0$, for $\beta, \gamma \geq 0$ we define a joint Laplace transform, given $J_0 = l$, namely $\tilde{L}_{\beta,\gamma|l}(t;\Delta) = \mathrm{E}_l\left(\exp\left(-\beta Z_{ir}(t) - \gamma Z(t+\Delta)\right)\right)$. Now, to begin analyzing $\tilde{L}_{\beta,\gamma|l}(t;\Delta)$, we need not only $\zeta_\gamma(x,t;z)$ but also a function $\zeta_{\beta,\gamma}\left(x;t_1,t_2;z\right)$. For all $0 \leq x \leq t_1 \leq t_2$, we will find the function $\zeta_{\beta,\gamma}\left(x;t_1,t_2;z\right)$ is given by

$$\zeta_{\beta,\gamma}\left(x;t_1,t_2;z\right) = \mathrm{E}\left(\exp\left(-e^{-xz}\ell\left(W_{1,1}\right)X_{1,1}\{\beta\,\mathrm{I}\left(W_{1,1} \leq t_1 - x\right) + \gamma\,\mathrm{I}\left(W_{1,1} > t_2 - x\right)\}\right) \mid \tau_1 = x\right).$$

In contrast to ζ_β, $\zeta_{\beta,\gamma}$ is the joint Laplace transform of the discounted IR liability and the discounted IBNR liability from one claim at the first event time, given $\delta_1 = z$ and $\tau_1 = x \leq t_1 \leq t_2$. By conditioning on τ_1, δ_1, and J_1, we express the delayed case $\tilde{L}_{\beta,\gamma|0}(t;\Delta)$ in "Markov renewal-like" form:

$$\begin{aligned}\tilde{L}_{\beta,\gamma|0}(t;\Delta) &= \bar{F}_0(t+\Delta) + \int_t^{t+\Delta}\int_{0-}^{\infty} \chi_C\left(\zeta_\gamma(x,t+\Delta;z)\right)\tilde{\boldsymbol{L}}_{\gamma e^{-xz}}(t+\Delta-x)D_0(\mathrm{d}z)F_0(\mathrm{d}x)\\ &\quad + \int_0^t\int_{0-}^{\infty} \chi_C\left(\zeta_{\beta,\gamma}(x;t,t+\Delta;z)\right)\tilde{\boldsymbol{L}}_{\beta e^{-xz},\gamma e^{-xz}}(t-x;\Delta)D_0(\mathrm{d}z)F_0(\mathrm{d}x).\end{aligned} \tag{19}$$

We denote the moments of $Z_{ir}(t)$ and $Z(t+\Delta)$, given the initial risk class $J_0 = l$, by $M_{m,n|l}(t;\Delta) = \mathrm{E}_l\left(Z_{ir}^m(t)Z^n(t+\Delta)\right)$. As before, $M_{m,n|0}$ indicates $\tau_1 \sim F_0$ and $\delta_1 \sim D_0$. Recall the function $B^*_{r|l}(x,t)$ for $1 \leq l \leq v$, $1 \leq k \leq r \leq n$, and $0 \leq x \leq t$ from Sect. 3.1. For two numbers m and n, we denote $m \wedge n = \min(m,n)$. In the present bivariate situation, now consider the indices $0 \leq 1 \wedge q \leq j \leq q \leq m$ and $0 \leq 1 \wedge r \leq k \leq r \leq n$. Let the $(q-j+1)$-dimensional row vector $\boldsymbol{\eta}\left(x,t_1\right)$ have the ith entry according to Eq. (9). Once again, let the ith entry of the $(r-k+1)$-dimensional row vector $\boldsymbol{\xi}\left(x,t_2\right)$ be Eq. (8). Instead of $B^*_{r|l}(x,t)$, we define $B^*_{q,r|l}\left(x;t_1,t_2\right)$ for $1 \leq l \leq v$ and $0 \leq x \leq t_1 \leq t_2$. This $B^*_{q,r|l}\left(x;t_1,t_2\right)$, which we write in column l of the v-dimensional row vector $\boldsymbol{B}^*_{q,r}\left(x;t_1,t_2\right)$, satisfies

$$B^*_{q,r|l}\left(x;t_1,t_2\right) = \sum_{j=1\wedge q}^{q}\sum_{k=1\wedge r}^{r}\chi_{C|l}^{(j+k)}(1)B_{q,j}\left(\boldsymbol{\eta}\left(x,t_1\right)\right)B_{r,k}\left(\boldsymbol{\xi}\left(x,t_2\right)\right).$$

Now we are equipped to analyze Eq. (19) via Markov renewal equations. We take the partial derivative of Eq. (19) m times wrt β and n times wrt γ. Then we set $(\beta,\gamma) = (0,0)$. If $m = 0$ here, setting $\beta = 0$ simplifies Eq. (19) to Eq. (10) evaluated at $t+\Delta$. So, we assume $m \geq 1$ in the following. Analogous to $v_{n|0}(t)$ in Sect. 3.1, we have the function $v_{m,n|0}(t;\Delta)$, given by

$$v_{m,n|0}(t;\Delta) = \sum_{\substack{r=0\\ q+r>0}}^{n}\sum_{q=0}^{m}\binom{m}{q}\binom{n}{r}\int_0^t \boldsymbol{B}^*_{q,r}(x;t,t+\Delta)\boldsymbol{M}_{m-q,n-r}(t-x;\Delta)G_{(m+n)|0}(\mathrm{d}x).$$

Finally, we have the following extended version of the renewal-form (Landriault et al. (2017), Eq. (3.17)):

$$M_{m,n|0}(t;\Delta) = \int_0^t \chi_C(1)\boldsymbol{M}_{m,n}(t-x;\Delta)G_{(m+n)|0}(\mathrm{d}x) + v_{m,n|0}(t;\Delta). \tag{20}$$

Using Eq. (20), Theorem 2 generalizes (Landriault et al. (2017), Theorem 3).

Theorem 2 *Let $n \in \mathbb{N}$, $m \in \mathbb{N}^+$, and $\tau_1 \sim F_0$ and $\delta_1 \sim D_0$. Then, we have:*

$$\mathrm{E}_0\big(Z^m_{ir}(t)Z^n(t+\Delta)\big) = v_{m,n|0}(t;\Delta) + \int_0^t \chi_C(1)\boldsymbol{v}_{m,n}(t-x;\Delta)H_{(m+n)|0}(\mathrm{d}x). \tag{21}$$

Establishing Eq. (21) in Theorem 2 from Eq. (20) is analogous to the proof of Theorem 1. We may specialize Theorem 2 in the same fashion that we simplified Theorem 1. Both cases of Corollary 3 continue to generalize (Landriault et al. (2017), Theorem 3). Our comments in Sect. 3.2 about how Corollary 2 contains (Landriault et al. (2017), Theorem 1) equally apply here.

Corollary 3 *Let $n \in \mathbb{N}$ and $m \in \mathbb{N}^+$. Let τ_1 and δ_1 have arbitrary df's F_0 and D_0, respectively.*

1. *Assume the discount rate is constant, namely $D_l(\cdot) \equiv \mathrm{I}(\cdot \geq \delta)$ for $0 \leq l \leq v$. Then,*

$$\mathrm{E}_0\big(Z^m_{ir}(t)Z^n(t+\Delta)\big) = v_{m,n|0}(t;\Delta) + \int_0^t e^{-(m+n)\delta x}\chi_C(1)\boldsymbol{v}_{m,n}(t-x;\Delta)H_0(\mathrm{d}x).$$

2. *Assume for $k \geq 2$ that τ_k are iid with df F and δ_k are iid with df D, equivalent to setting $v = 1$. Then,*

$$\mathrm{E}_0\big(Z^m_{ir}(t)Z^n(t+\Delta)\big) = v_{m,n|0}(t;\Delta) + \int_0^t v_{m,n}(t-x;\Delta)H_{(m+n)|0}(\mathrm{d}x).$$

The proof of Corollary 3 is essentially the same as that of Corollary 2, so we omit it. In Theorem 2 and Corollary 3 (1), the joint moments $\mathrm{E}_0\big(Z^m_{ir}(t)Z^n(t+\Delta)\big)$ depend on the lower-order conditional joint moments for all risk classes $1 \leq l \leq v$.

4 Computational Aspects

Landriault et al. (2017) suggested using Erlang mixtures for the interevent time and reporting lag distributions. They observed that this allowed tractable computation of the moments $\mathrm{E}(Z^n(t))$ and $\mathrm{E}\big(Z^m_{ir}(t)Z^n(t+\Delta)\big)$. Now, one might expect less tractable computations with multiple orders of severity and our Markovian discount rates. But certain distributional assumptions still provide some tractability.

We discuss computation under multiple risk classes first. Now, recall that we never fully specified the joint distribution of $\boldsymbol{Q}_k$ to define J_k in Eq. (2). We only assumed the condition Eq. (1) and that $\{\boldsymbol{Q}_k\}_{k\geq 0}$ are iid. One natural choice for the distribution of $Q_{l,k}$ is deterministic thresholds. For random thresholds, one could let the $\{Q_{l,k} - Q_{l-1,k}\}_{l=2}^{v}$ be independent. (Recall $Q_{1,k} = 0$ almost surely.) Another choice would be to define iid random variables $\{\Gamma_k\}_{k\geq 0}$ on the positive integers. Then let $Q_{l,k} = \alpha_l \Gamma_k$ for constants α_l in $\mathbb{N}$ with $l \in \{1, \ldots, v\}$ satisfying $0 = \alpha_1 < \alpha_2 < \cdots < \alpha_v < \infty$. Both of these random choices generalize the deterministic thresholds choice. For the distributions of the $\{Q_{l,k} - Q_{l-1,k}\}_{l=2}^{v}$, we suggest using discrete phase-type distributions, which are closed under convolution (see Bladt and Nielsen (2017)).

Now let us consider general distributions D_l for the discount rates δ_k. One obvious choice would be a constant rate $\kappa_l \geq 0$ when in severity class l, namely $D_l(z) = \mathrm{I}\big(z \geq \kappa_l\big)$. Then the distribution functions G_n would maintain the Erlang-mixture form suggested for

the F_l. The same would be true for any discount-rate distributions D_l taking a finite number of values each. One interesting such choice is to let the discount rates δ_k be a common constant $\kappa_* \geq 0$ with probability $1-\rho$ or the value $\kappa_l \geq 0$ with probability ρ. In other words, the distribution function D_l is $D_l(z) = (1-\rho)\,\mathrm{I}\left(z \geq \kappa_*\right) + \rho\,\mathrm{I}\left(z \geq \kappa_l\right)$.

We look at independent discrete phase-type distributions for $\{Q_{l,k} - Q_{l-1,k}\}_{l=2}^{\nu}$ next. For all $k \in \mathbb{N}$, let $R_{l,k} = Q_{l+1,k} - Q_{l,k}$ almost surely for $1 \leq l \leq \nu - 1$. In the following, $\boldsymbol{e}$ is a column vector of 1's, $\mathbf{0}$ is a matrix of 0's, and $\boldsymbol{I}$ is an identity matrix. The dimensions of $\boldsymbol{e}$, $\mathbf{0}$, and $\boldsymbol{I}$ may be inferred from the context. Let $\{R_{l,k}\}_{l=1}^{\nu-1}$ be independent and assume for $1 \leq l \leq \nu - 1$ that $R_{l,k}$ has the discrete phase-type distribution $DPH\left(\boldsymbol{\pi}_l, \boldsymbol{\Theta}_l\right)$. The row vector $\boldsymbol{\pi}_l$ satisfies $\boldsymbol{\pi}_l \boldsymbol{e} = 1$ and has the same number of columns as the matrix $\boldsymbol{\Theta}_l$. This $\boldsymbol{\Theta}_l$ is a subtransition matrix with corresponding vector $\boldsymbol{\theta}_l = \left(\boldsymbol{I} - \boldsymbol{\Theta}_l\right)\boldsymbol{e}$. Then (Bladt and Nielsen (2017), Theorem 1.2.65) tells us that for $1 \leq l \leq \nu - 1$, $Q_{l+1,k}$ has the distribution $DPH\left(\boldsymbol{\alpha}_{l+1}, \boldsymbol{S}_{l+1}\right)$. For $2 \leq l \leq \nu - 1$, the row vectors $\boldsymbol{\alpha}_{l+1}$ and matrices $\boldsymbol{S}_{l+1}$ are given by

$$\boldsymbol{\alpha}_{l+1} = \left(\boldsymbol{\alpha}_l, \mathbf{0}\right) \text{ and } \boldsymbol{S}_{l+1} = \begin{pmatrix} \boldsymbol{S}_l & \boldsymbol{s}_l \boldsymbol{\pi}_{l+1} \\ \mathbf{0} & \boldsymbol{\Theta}_{l+1} \end{pmatrix},$$

where $\boldsymbol{\alpha}_2 = \boldsymbol{\pi}_1$ and $\boldsymbol{S}_2 = \boldsymbol{\Theta}_1$. The column vector $\boldsymbol{s}_l$ corresponding to $\boldsymbol{S}_l$ is $\boldsymbol{s}_l = \left(\boldsymbol{I} - \boldsymbol{S}_l\right)\boldsymbol{e}$, and $\boldsymbol{s}_2 = \boldsymbol{\theta}_1$. The following lemma is straightforward.

Lemma 1 *Let C and Q be independent random variables on $\mathbb{N}$. If Q has the distribution $DPH(\boldsymbol{\alpha}, \boldsymbol{S})$, then* $\mathrm{E}\left(z^C\,\mathrm{I}(C < Q)\right) = \mathrm{E}\left(\boldsymbol{\alpha}(z\boldsymbol{S})^C \boldsymbol{e}\right)$.

Because of assumption Eq. (1), it is clear that $\sum_{i=1}^{l} \chi_{C|i}(z) = \mathrm{E}\left(z^{C_k}\,\mathrm{I}\left(C_k < Q_{l+1,k}\right)\right)$ for $1 \leq l \leq \nu - 1$. Recall that $\sum_{i=1}^{\nu} \chi_{C|i}(z) = \mathrm{E}\left(z^{C_k}\right)$. Then by Lemma 1 we have the following proposition.

Proposition 2 *For all $k \in \mathbb{N}$, let $R_{l,k} = Q_{l+1,k} - Q_{l,k}$ almost surely for $1 \leq l \leq \nu - 1$. Let $\{R_{l,k}\}_{l=1}^{\nu-1}$ be independent. If $R_{l,k}$ has the discrete phase-type distribution $DPH\left(\boldsymbol{\pi}_l, \boldsymbol{\Theta}_l\right)$ for $1 \leq l \leq \nu - 1$, then the $\{\chi_{C|l}(z)\}_{l=1}^{\nu}$ may be expressed as*

$$\begin{aligned}
\chi_{C|1}(z) &= \mathrm{E}\left(\boldsymbol{\alpha}_2\left(z\boldsymbol{S}_2\right)^{C_k}\boldsymbol{e}\right), \\
\chi_{C|l}(z) &= \mathrm{E}\left(\boldsymbol{\alpha}_{l+1}\left(z\boldsymbol{S}_{l+1}\right)^{C_k}\boldsymbol{e}\right) - \mathrm{E}\left(\boldsymbol{\alpha}_l\left(z\boldsymbol{S}_l\right)^{C_k}\boldsymbol{e}\right), \text{ for } 2 \leq l \leq \nu - 1, \\
\chi_{C|\nu}(z) &= \mathrm{E}\left(z^{C_k}\right) - \mathrm{E}\left(\boldsymbol{\alpha}_\nu\left(z\boldsymbol{S}_\nu\right)^{C_k}\boldsymbol{e}\right).
\end{aligned}$$

The importance of Proposition 2 is that the $\chi_{C|l}(z)$ are expectations wrt C_k only. This is just like the case of deterministic thresholds $Q_{l,k}$. For some choices of the distribution of the C_k, we can therefore make the $\chi_{C|l}(z)$ entirely explicit. One example would be the so-called "Panjer class". Since (Sundt and Vernic (2009), Theorem 2.6) reminds us the Panjer class is the Poisson, Negative Binomial, and Binomial distributions, we can specialize Lemma 1. We ignore the distribution degenerate at zero.

Lemma 2 *Let C and Q be independent random variables on $\mathbb{N}$. If Q has the distribution $DPH(\boldsymbol{\alpha}, \boldsymbol{S})$ and the nondegenerate C is from the Panjer class, then* $\mathrm{E}\left(z^C\,\mathrm{I}(C < Q)\right)$ *is given as follows*:

1. $\mathrm{E}\left(z^C \,\mathrm{I}(C < Q)\right) = \boldsymbol{\alpha} \exp\left(\lambda(z\boldsymbol{S} - \boldsymbol{I})\right)\boldsymbol{e}$, *when* C *is* $\mathrm{Poi}(\lambda)$.
2. $\mathrm{E}\left(z^C \,\mathrm{I}(C < Q)\right) = (1-p)^n \boldsymbol{\alpha}(\boldsymbol{I} - pz\boldsymbol{S})^{-n}\boldsymbol{e}$, *when* C *is* $\mathrm{NB}(n,p)$.
3. $\mathrm{E}\left(z^C \,\mathrm{I}(C < Q)\right) = \boldsymbol{\alpha}((1-p)\boldsymbol{I} + pz\boldsymbol{S})^n\boldsymbol{e}$, *when* C *is* $\mathrm{Bin}(n,p)$.

Proof (Proof of Lemma 2) We can apply Lemma 1 to the expectation $\mathrm{E}\left(z^C \,\mathrm{I}(C < Q)\right)$. Then the proofs of (1) and (3) are straightforward. For (1), we just point out the series expansions of matrix exponentials are known to be convergent (Asmussen and Albrecher (2010), A3). The proof of (2) is slightly more involved, but still simple. By (Bladt and Nielsen (2017), Corollary 1.2.63), all the eigenvalues of $\boldsymbol{S}$ are strictly inside the unit circle. Since p is a probability, the expansion $(\boldsymbol{I} - pz\boldsymbol{S})^{-n} = \sum_{c=0}^{\infty} \binom{n+c-1}{c} (pz\boldsymbol{S})^c$ converges for $|z| \le 1$. □

We have the following corollary to Proposition 2.

Corollary 4 *For all* $k \in \mathbb{N}$, *let* $R_{l,k} = Q_{l+1,k} - Q_{l,k}$ *almost surely for* $1 \le l \le \nu - 1$. *Let* $\{R_{l,k}\}_{l=1}^{\nu-1}$ *be independent. If* $R_{l,k}$ *has the discrete phase-type distribution* $DPH\left(\boldsymbol{\pi}_l, \boldsymbol{\Theta}_l\right)$ *for* $1 \le l \le \nu - 1$, *then the* $\{\chi_{C|l}(z)\}_{l=1}^{\nu}$ *may be expressed as*

$$\chi_{C|1}(z) = \begin{cases} \boldsymbol{\alpha}_2 \exp\left(\lambda\left(z\boldsymbol{S}_2 - \boldsymbol{I}\right)\right)\boldsymbol{e} & \text{if } C_k \sim \mathrm{Poi}(\lambda) \\ (1-p)^n \boldsymbol{\alpha}_2\left(\boldsymbol{I} - pz\boldsymbol{S}_2\right)^{-n}\boldsymbol{e} & \text{if } C_k \sim \mathrm{NB}(n,p) \\ \boldsymbol{\alpha}_2\left((1-p)\boldsymbol{I} + pz\boldsymbol{S}_2\right)^n\boldsymbol{e} & \text{if } C_k \sim \mathrm{Bin}(n,p) \end{cases}$$

$$\chi_{C|l}(z) = \begin{cases} \boldsymbol{\alpha}_{l+1} \exp\left(\lambda\left(z\boldsymbol{S}_{l+1} - \boldsymbol{I}\right)\right)\boldsymbol{e} & \\ \qquad -\boldsymbol{\alpha}_l \exp\left(\lambda\left(z\boldsymbol{S}_l - \boldsymbol{I}\right)\right)\boldsymbol{e} & \text{if } C_k \sim \mathrm{Poi}(\lambda) \\ (1-p)^n \boldsymbol{\alpha}_{l+1}\left(\boldsymbol{I} - pz\boldsymbol{S}_{l+1}\right)^{-n}\boldsymbol{e} & \\ \qquad -(1-p)^n \boldsymbol{\alpha}_l\left(\boldsymbol{I} - pz\boldsymbol{S}_l\right)^{-n}\boldsymbol{e} & \text{if } C_k \sim \mathrm{NB}(n,p), \qquad \text{for } 2 \le l \le \nu - 1 \\ \boldsymbol{\alpha}_{l+1}\left((1-p)\boldsymbol{I} + pz\boldsymbol{S}_{l+1}\right)^n\boldsymbol{e} & \\ \qquad -\boldsymbol{\alpha}_l\left((1-p)\boldsymbol{I} + pz\boldsymbol{S}_l\right)^n\boldsymbol{e} & \text{if } C_k \sim \mathrm{Bin}(n,p) \end{cases}$$

$$\chi_{C|\nu}(z) = \begin{cases} \exp\left(\lambda(z-1)\right) - \boldsymbol{\alpha}_\nu \exp\left(\lambda\left(z\boldsymbol{S}_\nu - \boldsymbol{I}\right)\right)\boldsymbol{e} & \text{if } C_k \sim \mathrm{Poi}(\lambda) \\ (1-p)^n(1-pz)^{-n} - (1-p)^n \boldsymbol{\alpha}_\nu\left(\boldsymbol{I} - pz\boldsymbol{S}_\nu\right)^{-n}\boldsymbol{e} & \text{if } C_k \sim \mathrm{NB}(n,p) \\ (1-p+pz)^n - \boldsymbol{\alpha}_\nu\left((1-p)\boldsymbol{I} + pz\boldsymbol{S}_\nu\right)^n\boldsymbol{e} & \text{if } C_k \sim \mathrm{Bin}(n,p) \end{cases}$$

To apply our model to claim reserve data, note the particular features which we use beyond the model of Landriault et al. (2017). We just have to concern ourselves with fitting the functions $G_{n|l}$ and $\chi_{C|l}$. Due to our model assumptions, the rest is simply a matter of fitting the Landriault et al. (2017) model, which is beyond the scope of our paper. The Laplace-Stieltjes transforms $\tilde{D}_l$ are quite general. But if we limit the distributions D_l to take a finite number of values each as we suggested earlier, then one could use the usual expectation-maximization algorithm to fit the $G_{n|l}$ provided the F_l were assumed all phase-type. See (Bladt and Nielsen (2017), Theorem 12.7.1) for a formulation of the expectation-maximization algorithm.

Next we address fitting the $\chi_{C|l}$. One could say our model's definition of categories does not correspond to things like the categories of a hurricane, where these levels are not determined by insurance claim counts. So, whether one first fits the distribution of C_k or of $\boldsymbol{Q}_k$, we recommend fitting these so that $\chi_{C|l}(1) = \Pr\left(J_k = l\right)$ equals the probability

of a hurricane being category l (for example). Now, Lemma 1 and Proposition 2 serve to emphasize the centrality of the distribution of C_k in the functions $\chi_{C|l}$, since this C_k is the only rv to specify. Thus, we recommend fitting the distribution of C_k first, then finding the distributions of $\boldsymbol{Q}_k$ second.

Estimating the distribution of the C_k, whether through a bootstrap or otherwise, would require counts of claims from individual events. This would not be possible from traditional aggregated claim reserving triangles. Those do not attribute how many claims from each origin/development cell correspond to a particular event. With sufficient data granularity, one could use bootstrapping for nonparametric estimation of the batch-size distribution. But greater care would be required if we allow the C_k to be heavy-tailed, which is of interest in catastrophic situations. We also comment that data on claim counts per event would by necessity be for IR claims only. This would require further care, which we do not discuss here. We observe that the insufficient granularity of aggregated claims triangles prevents application of the traditional chain ladder method and its associated bootstrapping techniques to achieve the objectives of our model.

5 Conclusion

We have extended the methodology of Landriault et al. (2017). We allow insurers to change the distribution of the interevent time based upon the number of claims arising from the current event. We also permit modeling the IBNR and IR claims with a severity-class dependent discount rate.

Both Landriault et al. (2017) and our generalized form of their model implicitly assume claim count moments are finite. This implies a certain degree of light-tailedness in the claim count distribution. However, batch arrivals of claims are meant for catastrophic scenarios, so heavy-tailed properties may be more appropriate. Other methods may become necessary for modeling such IBNR liabilities. Furthermore, our model could be improved by incorporating context-specific modeling elements, like from meteorology or epidemiology.

Other useful directions to take our extension of Landriault et al. (2017) would include examining the performance of our model when fitting to data or estimating parameters. While we found the moments of total liabilities in this paper, finding their distributions should also be possible. Namely, choose $D_l(z)$ and $F_l(t)$ to make $G_{n|l}(t)$ phase-type, and let $X_{i,k}$ and $W_{i,k}$ have phase-type distributions.

Acknowledgements Daniel J. Geiger was supported by the Missouri University of Science and Technology [Chancellor's Fellowship; Graduate Assistantship].

The authors would like to thank the editors and the anonymous referee for helpful suggestions which greatly improved the manuscript. In particular, we thank the referee for their suggestion which led to the Panjer class results of Sect. 4 (Lemma 2 and Corollary 4).

Declarations

Conflicts of Interest The authors declare that they have no conflicts of interest.

References

Albrecher H, Boxma OJ (2004) A ruin model with dependence between claim sizes and claim intervals. Insur: Math Econ 35(2):245–254
Asmussen S, Albrecher H (2010) Ruin Probabilities, vol 14. World Scientific
Avanzi B, Taylor G, Wong B, Yang X (2021) On the modelling of multivariate counts with Cox processes and dependent shot noise intensities. Insur: Math Econ 99:9–24
Bischofberger SM, Hiabu M, Isakson A (2020) Continuous chain-ladder with paid data. Scand Actuar J 2020(6), 477–502
Bladt M, Nielsen BF (2017) Matrix-exponential distributions in applied probability, probability theory and stochastic modelling, vol 81, 1st edn. Springer Nature
Blier-Wong C, Cossette H, Lamontagne L, Marceau E (2021) Machine learning in P&C insurance: A review for pricing and reserving. Risks 9(1):4
Bornhuetter RL, Ferguson RE (1972) The actuary and IBNR. In: Proceedings of the Casualty Actuarial Society, vol 59, pp 181–195
Boucher JP (2021) Two incompatible objectives with individual reserve models: an approach with multivariate adaptative regression spline models
Chaoubi I, Besse C, Cossette H, Côté MP (2022) Micro-level reserving for general insurance claims using a long short-term memory network. arXiv preprint arXiv:220113267
Delong Ł, Lindholm M, Wüthrich MV (2021) Collective reserving using individual claims data. Scand Actuar J pp 1–28
Duval F, Pigeon M (2019) Individual loss reserving using a gradient boosting-based approach. Risks 7(3):79
Fung TC, Badescu AL, Lin XS (2021) Fitting censored and truncated regression data using the mixture of experts models. North Am Actuar J pp 1–25
Janssen J, Manca R (2006) Applied semi-Markov processes. Springer Science & Business Media
Johnson WP (2002) The curious history of Faà di Bruno's formula. Amer Math Monthly 109(3), 217–234
Landriault D, Willmot GE, Xu D (2017) Analysis of IBNR claims in renewal insurance models. Scand Actuar J 2017(7), 628–650
Léveillé G, Adékambi F (2011) Covariance of discounted compound renewal sums with a stochastic interest rate. Scand Actuar J 2011(2), 138–153
Léveillé G, Garrido J (2001a) Moments of compound renewal sums with discounted claims. Insur: Math Econ 28(2):217–231
Léveillé G, Garrido J (2001b) Recursive moments of compound renewal sums with discounted claims. Scand Actuar J 2001(2), 98–110
Léveillé G, Hamel E (2019) Compound trend renewal process with discounted claims: a unified approach. Scand Actuar J 2019(3), 228–246
Li Z, Sendova KP (2015) On a ruin model with both interclaim times and premiums depending on claim sizes. Scand Actuar J 2015(3), 245–265
Lindholm M, Zakrisson H (2022) A collective reserving model with claim openness. ASTIN Bulletin: The Journal of the IAA pp 1–27
Sundt B, Vernic R (2009) Recursions for convolutions and compound distributions with insurance applications. Springer Science & Business Media
Tarbell TF (1934) Incurred but not reported claim reserves. In: Proceedings of the Casualty Actuarial Society, vol 20, pp 275–280
Taylor G (2019) Loss reserving models: Granular and machine learning forms. Risks 7(3):82
Verbelen R, Antonio K, Claeskens G, Crevecoeur J (2018) An EM algorithm to model the occurrence of events subject to a reporting delay
Wahl F, Lindholm M, Verrall R (2019) The collective reserving model. Insur: Math Econ 87:34–50
Wüthrich MV (2018) Machine learning in individual claims reserving. Scand Actuar J 2018(6), 465–480
Yanez JS, Pigeon M (2021) Micro-level parametric duration-frequency-severity modeling for outstanding claim payments. Insur: Math Econ 98:106–119