COMPOUND POISSON CLAIMS RESERVING MODELS: EXTENSIONS AND INFERENCE

BY

SHENGWANG MENG AND GUANGYUAN GAO

ABSTRACT

We consider compound Poisson claims reserving models applied to the paid claims and to the number of payments run-off triangles. We extend the standard Poisson-gamma assumption to account for over-dispersion in the payment counts and to account for various mean and variance structures in the individual payments. Two generalized linear models are applied consecutively to predict the unpaid claims. A bootstrap is used to estimate the mean squared error of prediction and to simulate the predictive distribution of the unpaid claims. We show that the extended compound Poisson models make reasonable predictions of the unpaid claims.

KEYWORDS

Stochastic claims reserving, compound Poisson model, generalized linear model, quasi-likelihood estimation, bootstrap, Tweedie's distribution.

1. INTRODUCTION

Stochastic claims reserving models have been studied extensively. A majority of the models are univariate fitting to the historical claim amounts recorded in a run-off triangle. Three monographs of claims reserving models are Taylor (2000), England and Verrall (2002) and Wüthrich and Merz (2008). For the more recent work, we refer to Wüthrich and Merz (2015). Adequate reserves are required by the insurance authority, the same related to the policyholders' interests. It is equivalently crucial for insurance companies not to hold redundant reserves, which would affect the premium adversely and put the insurance companies at disadvantage in a competitive market. Every insurance company has the motivation to setup a reserve as accurately as possible.

This paper focuses on the situation when paid claims and number of payments are available in two run-off triangles, the same setting as Wüthrich (2003). Wüthrich (2003) applied a Tweedie's compound Poisson model (compound

Astin Bulletin 48(3), 1137-1156. doi: 10.1017/asb.2018.12 © 2018 by Astin Bulletin. All rights reserved.

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

Poisson-gamma model). The parameter estimation procedure had three iterative steps. First, the power parameter was assumed known, and the mean parameters were estimated in a Tweedie generalized linear model (GLM). Second, the maximum likelihood estimate (MLE) of the dispersion parameter was derived as a function of the mean parameters and the power parameter. Third, the profile likelihood of the power parameter was derived by replacing the mean and dispersion parameters in the likelihood function with their MLEs, and the power parameter was estimated by maximizing the profile likelihood. Following Wüthrich (2003), Peters *et al.* (2009) fitted the paid claims by a Tweedie's compound Poisson model without considering the number of payments. They presented three equivalent representations of the Tweedie's compound Poisson model. The model selection was discussed, using the posterior model probabilities calculated on the Markov chain Monte Carlo samples. Boucher and Davidov (2011) studied the same data set as in Wüthrich (2003). In contrast to Wüthrich (2003), they assumed the dispersion parameter as a regression function. A double generalized linear model (DGLM) (Smyth and Jørgensen, 2002) was applied to estimate the mean and dispersion parameters. The power parameter was estimated by a trial-and-error method. Gigante *et al.* (2013, 2016) also assumed the dispersion parameter as a regression function. Gigante *et al.* (2016) proposed a mixture model framework, and applied the hierarchical log-likelihood method (Lee and Nelder, 2001) to estimate parameters via three iterative GLMs.

There is another series of claims reserving literature using claim amounts and counts. The fundamental idea originated from Wright (1990), where the proposed model is known as Hoerl curve model, fitting to claim amounts. Following Wright (1990), Verrall *et al.* (2010) and Martínez-Miranda *et al.* (2011, 2012, 2013, 2015) built a double chain ladder (DCL) model fitting to the paid claims and the number of reported claims. The paid claims run-off triangle did not "match" with the number of reported claims run-off triangle cell-wisely, since the payment delay was considered. Instead, the paid claims run-off triangle matched with the number of payments run-off triangle cell-wisely. In the DCL model, the numbers of payments were related with the numbers of reported claims via a multinomial distribution. Another work using claim amounts and counts is Taylor and Xu (2016), where they proposed stochastic versions for the payments per claim incurred model and the payments per claim finalization (PPCF) model. In the following, we will compare our model with the Tweedie's compound Poisson model (Wüthrich, 2003), the DCL model (Verrall *et al.*, 2010) and the PPCF model (Taylor and Xu, 2016). Note that the Tweedie's compound Poisson model uses the paid claims and the numbers of payments, the same data as our model. The DCL model uses the paid claims and the numbers of reported claims, while the PPCF model uses the paid claims, the numbers of reported claims and the numbers of open cases.

The paper is structured as follows. Section 2 specifies the extended compound Poisson models and states several associated conditional and marginal distributions. Section 3 discusses model inference, including the parameter

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

estimation and the prediction of unpaid claims. Section 4 demonstrates the proposed models and the inference methods using a real data set from Wüthrich (2003). Section 5 concludes with findings.

2. Extended compound Poisson claims reserving models

Denote the accident years by $i = 1, \ldots, I$, the development years by $j = 1, \ldots, I$, and the experienced years by $i + j - 1 = 1, \ldots, 2I - 1$. The incremental paid claims and the incremental number of payments of the accident year i during the development year j are denoted by Y_{ij} and N_{ij}, respectively. Denote the index in the upper triangle by $\Omega_u := \{(i, j) : i, j \in \mathbb{N}, i + j \leq I + 1\}$ and the index in the lower triangle by $\Omega_l := \{(i, j) : i, j \in \mathbb{N}, i+j \geq I+2, i \leq I, j \leq I\}$. Denote the two upper triangles by $Y^u := \{Y_{ij} : (i, j) \in \Omega_u\}$ and $N^u := \{N_{ij} : (i, j) \in \Omega_u\}$. Assume all the claims are settled by the development year I. The claims reserving problem is to predict the lower triangle $Y^l := \{Y_{ij} : (i, j) \in \Omega_l\}$ and $N^l := \{N_{ij} : (i, j) \in \Omega_l\}$. The unpaid claims of accident years $i = 2, \ldots, I$ are denoted by $R_i := \sum_{j=I-i+2}^{I} Y_{ij}$ and the total unpaid claims are denoted by $R := \sum_{i=2}^{I} R_i$.

2.1. Model specification

Following Verrall *et al.* (2010), we restrict our analysis to a line of business whose claims are usually settled with one payment. Under this assumption, different payments are corresponding to different claims, and they are assumed to be independent from each other. Denote the kth individual payment of accident year i during development year j as X_{ij}^k, $k = 1, \ldots, N_{ij}$. The total paid claims during this period are given by

$$Y_{ij} = \sum_{k=1}^{N_{ij}} X_{ij}^k,$$

with three standard assumptions:

1. N_{ij} is a discrete random variable taking values from non-negative integers;
2. $(X_{ij}^k)_{k=1:N_{ij}}$ are independent and identically distributed (i.i.d.);
3. N_{ij} and $(X_{ij}^k)_{k=1:N_{ij}}$ are independent.

Under the no partial payments assumption, the number of payments is equal to the number of finalized claims, so our model has a connection with the PPCF model (discussed later). The compound distribution of Y_{ij} is fully determined by the distributions of N_{ij} and X_{ij}^k. In the following, we make assumptions for the number of payments and the individual payments.

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

2.1.1. *Models for number of payments.* Claim counts are usually modeled by a Cox Poisson process; see Norberg (1993), Antonio and Plat (2014), Avanzi *et al.* (2016) and Badescu *et al.* (2016). Here, we also use a Poisson process to model the number of payments. We note that the payment counts may be subject to an insurer's administrative change. So, before using the following results, one should make sure that the insurer's operation does not change dramatically to unduly affect the payment counts (Verrall and Wüthrich, 2016). Denote by M_{it}, the number of payments up to the development time t of the accident year i. M_{it} is assumed to follow a Poisson process with non-constant intensity $\Lambda_i(t)$:

$$M_{it} \sim \text{Poi}\left(e_i \int_0^t \Lambda_i(s)ds\right),$$

where Poi(a) denotes a Poisson distribution with mean a, and e_i is an exposure measure of accident year i. The non-constant intensity accounts for the fact that many small payments tend to be made during the first few development years and few large payments tend to be made in the later development periods. We have the following relationship between M_{it} and N_{ij}:

$$N_{ij} = M_{it(j)} - M_{it(j-1)},$$

where $t(\cdot)$ is a mapping from discrete development years to continuous development time. Hence, for $j \neq j'$, N_{ij} and $N_{ij'}$ cover non-overlapping time intervals, and they are independent. N_{ij} follows a Poisson distribution with mean:

$$e_i \int_{t(j-1)}^{t(j)} \Lambda_i(s)ds,$$

which is determined by two factors: the accident year i and the development year j. Without other information available, we parameterize the mean as a multiplicative chain ladder mean:

$$e_i \int_{t(j-1)}^{t(j)} \Lambda_i(s)ds := e_i\lambda_{ij} := e_i\gamma_i^{\text{Acc}}\gamma_j^{\text{Dev}},$$

where λ_{ij} is the expected payment frequency, γ_i^{Acc} is the accident year i effect, and γ_j^{Dev} is the development year j effect. For compactness, we define the following notations:

$$\lambda = \{\lambda_{ij} : (i, j) \in \Omega_u \cup \Omega_l\} \text{ and } \gamma = \{\gamma_1^{\text{Acc}}, \ldots, \gamma_I^{\text{Acc}}, \gamma_1^{\text{Dev}}, \ldots, \gamma_I^{\text{Dev}}\}.$$

We work on the number of payments due to the data availability. One may find that the number of reported claims is more robust and unlikely subject to administrative changes. In the DCL model, the numbers of reported claims are observed, from which the numbers of payments are estimated. In the PPCF model, the numbers of reported claims and open cases are observed, from which

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

the finalization rates (operational time) are predicted and used in the PPCF submodel.

2.1.2. *Models for individual payments.* An individual payment is assumed to follow one of the three distributions: Poisson, gamma or inverse-Gaussian. If a Poisson distribution is assumed, we apply the quasi-likelihood approach. So, the individual payment can take values from real numbers and over-dispersion is considered. Please refer to McCullagh and Nelder (1989) and Appendix B for the quasi-likelihood approach. The regression function for the mean parameter is assumed as

$$\mathbb{E}(X_{ij}^k) = \tau_{ij} = \eta(i+j-1)\delta(i)\upsilon(j), \quad k = 1, \ldots, N_{ij}, \tag{1}$$

where $\eta(\cdot)$ is the claim cost escalation function, $\delta(\cdot)$ is the accident year effect function and $\upsilon(\cdot)$ is the development year effect function. Note that Tweedie's compound Poisson model assumed a chain ladder mean; the DCL model assumed η, δ and υ as constant; the PPCF model assumed δ as constant and υ as a quadratic function of the finalization rate by the development year j. Whether including the calendar year η and/or the accident year covariates δ depends on the data and actuarial judgment. We suggest to first fit a model with only the development year effects υ, and then plot the residuals against the accident years and the calendar years, respectively. An obvious pattern typically indicates the need of the corresponding covariate. The variance of individual payment is given by

$$\mathrm{Var}(X_{ij}^k) = \psi V(\tau_{ij}) = \psi \tau_{ij}^m, \quad k = 1, \ldots, N_{ij},$$

where ψ is the dispersion parameter and $V(\cdot)$ is the variance function with $m = 1, 2, 3$ corresponding to Poisson, gamma and inverse-Gaussian distribution. For compactness, we denote $\tau = \{\tau_{ij} : (i,j) \in \Omega_u \cup \Omega_l\}$. Tweedie's compound Poisson model assumed a gamma distribution ($m = 2$); the DCL model assumed a constant variance ($m = 0$); the PPCF model assumed an over-dispersed Poisson (ODP) distribution ($m = 1$). We suggest to select a suitable distribution according to the residual plots, if no particular actuarial judgment is available. Sometimes, listing and comparing the results for different distributions is also helpful. Note that most of the analysis below can be applied to the Tweedie distributed individual payment with power parameter in [1, 3]. The Poisson, gamma and inverse-Gaussian distributions are three special cases of the Tweedie distribution family.

2.2. Conditional and marginal distributions

In a compound model, the conditional distribution and the marginal distribution of Y_{ij} are often used, depending on the available data and the parameters of interest. If we only have paid claims, the marginal distribution of Y_{ij} would be used. On the other hand, if we have both paid claims and number of payments,

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

we would consider their joint distribution or the conditional distribution of paid claims given the number of payments.

2.2.1. *The conditional distribution of* Y_{ij}. It is known that the sum of Poisson/gamma/inverse-Gaussian random variables still follows Poisson/gamma/inverse-Gaussian distributions. Hence, the conditional distribution of Y_{ij} given N_{ij} has the same form as the distribution of the individual payments X_{ij}^k. The mean and variance of Y_{ij} given N_{ij} are

$$\mathbb{E}(Y_{ij}|N_{ij}) = N_{ij}\tau_{ij} \text{ and } \mathrm{Var}(Y_{ij}|N_{ij}) = N_{ij}\psi\tau_{ij}^m = N_{ij}^{1-m}\psi(N_{ij}\tau_{ij})^m, \quad (2)$$

with $m = 1, 2, 3$ corresponding to Poisson, gamma and inverse-Gaussian distributions. Note that the conditional distribution has mean $N_{ij}\tau_{ij}$ and dispersion $N_{ij}^{1-m}\psi$ rather than ψ. One benefit of incorporating the numbers of payments is to simplify the mean function of paid claims. For example, according to equation (1) τ_{ij} might be constant or only involve the development year covariate. If the numbers of payments were not given, a chain ladder mean is assumed for paid claims.

2.2.2. *The marginal distribution of* Y_{ij}. As a special case when N_{ij} follows a Poisson distribution and X_{ij}^k follows a gamma distribution, the marginal distribution of Y_{ij} is a Tweedie's compound Poisson distribution with mean and variance as follows:

$$\mathbb{E}(Y_{ij}) = e_i\lambda_{ij}\tau_{ij} \text{ and } \mathrm{Var}(Y_{ij}) = e_i^{1-p}\frac{\lambda_{ij}^{1-p}\tau_{ij}^{2-p}}{2-p}(e_i\lambda_{ij}\tau_{ij})^p,$$

where the power parameter

$$p = \frac{1+2\psi}{1+\psi}. \quad (3)$$

Wüthrich (2003) approximated $(\lambda_{ij}^{1-p}\tau_{ij}^{2-p})/(2-p)$ as a constant. Here, we do not make such an approximation. When N_{ij} displays an over-dispersion ϕ, i.e., $\mathrm{Var}(N_{ij}) = \phi\mathbb{E}(N_{ij})$, the variance of Y_{ij} is inflated to

$$\mathrm{Var}(Y_{ij}) = \frac{\phi+\psi}{1+\psi}e_i^{1-p}\frac{\lambda_{ij}^{1-p}\tau_{ij}^{2-p}}{2-p}(e_i\lambda_{ij}\tau_{ij})^p,$$

where $(\phi+\psi)/(1+\psi)$ is the inflation factor due to the over-dispersion in the number of payments. In a Tweedie's compound Poisson model only using the paid claims, we might choose a suitable power parameter by comparing the residual plots for different p's. By incorporating the number of payments, we estimate p from equation (3) rather than from subjective judgment. This is another benefit of using the numbers of payments. Finally, we point out that if X_{ij}^k

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

follows a Poisson or inverse-Gaussian distribution, Y_{ij} does not follow a standard marginal distribution. Nevertheless, we have the following general result for the variance:

$$\mathrm{Var}(Y_{ij}) = \phi e_i \lambda_{ij} \tau_{ij}^2 + \psi e_i \lambda_{ij} \tau_{ij}^m,$$

where $\phi e_i \lambda_{ij} \tau_{ij}^2$ is the contribution of variation in numbers of payments, and $\psi e_i \lambda_{ij} \tau_{ij}^m$ is the contribution of variation in individual payments.

2.2.3. *The marginal distribution of R_i.* Rather than looking at each cell in the lower triangle, we pay more attention to the unpaid claims of each accident year. Appendix 5 shows that $(R_i)_{i=2:I}$ is a compound Poisson variable

$$R_i = \sum_{l=1}^{\tilde{N}_i} \tilde{X}_i^l, \tag{4}$$

where $\tilde{N}_i$ follows a Poisson distribution with mean $e_i \gamma_i^{\mathrm{Acc}} \sum_{j=I+2-i}^{I} \gamma_j^{\mathrm{Dev}}$, and $(\tilde{X}_i^l)_{l=1:\tilde{N}_i}$ follows a mixture distribution with probability density function

$$p(\tilde{X}_i^l) = \sum_{j=I+2-i}^{I} w_{ij}\, p_X\left(\tilde{X}_i^l \middle| \tau_{ij}, \psi\right), \tag{5}$$

where $w_{ij} = \gamma_j^{\mathrm{Dev}} / \sum_{t=I+2-i}^{I} \gamma_t^{\mathrm{Dev}}$ is the weight of component distribution, and $p_X(x|a, b)$ denotes the assumed probability density function of individual payments with mean a and dispersion b evaluated at x. Note that the above results are valid when N_{ij} follows a Poisson distribution. If an over-dispersion was assumed in N_{ij}, we would only have the following general mean and variance formulas:

$$\mathbb{E}(R_i) = e_i \sum_{j=I+2-i}^{I} \lambda_{ij} \tau_{ij} \quad \text{and} \quad \mathrm{Var}(R_i) = \phi e_i \sum_{j=I+2-i}^{I} \lambda_{ij} \tau_{ij}^2 + \psi e_i \sum_{j=I+2-i}^{I} \lambda_{ij} \tau_{ij}^m. \tag{6}$$

3. Model inference

We have seen that over-dispersion in payment counts would affect the (marginal) variance of paid claims, so the quasi-likelihood approach is applied when modeling the numbers of payments. The quasi-likelihood approach is also applied when the individual payments are assumed to follow a Poisson distribution. Please refer to Appendix B for the quasi-likelihood approach.

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

3.1. Parameter estimation

The joint likelihood of the data N^u, Y^u is

$$\begin{aligned} & L(\lambda, \phi, \tau, \psi | N^u, Y^u) \\ &= \prod_{i,j \in \Omega_u} p(Y_{ij}, N_{ij}; \lambda, \phi, \tau, \psi) \\ &= \prod_{i,j \in \Omega_u} p(Y_{ij} | N_{ij}; \tau, \psi) p(N_{ij}; \lambda, \phi) \\ &= \prod_{i,j \in \Omega_u} p(Y_{ij} | N_{ij}; \tau, \psi) \prod_{i,j \in \Omega_u} p(N_{ij}; \lambda, \phi). \end{aligned}$$

From the above, we can see that $\{\lambda, \phi\}$ and $\{\tau, \psi\}$ are allocated to two likelihood functions. So, we can estimate the parameters by fitting two GLMs consecutively. First, a Poisson GLM is fitted to N^u using the quasi-likelihood approach. We treat accident years and development years as factor covariates, using the logarithm link function, and adding the offset term $\log e_i$. In R, the argument family=quasipoisson is used in the glm function to implement the quasi-likelihood approach. Denote the estimated parameters by $\hat{\lambda}^{\text{MLE}}$ and $\hat{\phi}$. Second, a Poisson/gamma/inverse-Gaussian GLM is fitted to Y^u. We apply equations (1) and (2), use the logarithm link function, add the offset term $\log N_{ij}$ and specify the weights N_{ij}^{m-1} with $m = 1, 2, 3$ corresponding to Poisson, gamma and inverse-Gaussian distributions. Denote the estimated parameters by $\hat{\tau}^{\text{MLE}}$ and $\hat{\psi}$.

3.2. Prediction of unpaid claims and its MSEP

The predicted unpaid claims in Y^l are

$$\hat{Y}_{ij} = e_i \hat{\lambda}_{ij}^{\text{MLE}} \hat{\tau}_{ij}^{\text{MLE}} = \hat{N}_{ij} \hat{\tau}_{ij}^{\text{MLE}}.$$

Hence, we first predict the number of payments in N^l as $(\hat{N}_{ij})_{(i,j)\in\Omega_l}$ using the first fitted GLM, and then predict the unpaid claims in Y^l given $(\hat{N}_{ij})_{(i,j)\in\Omega_l}$, i.e., specifying the offset term $\log \hat{N}_{ij}$ and the weights $\hat{N}_{ij}^{m-1}$ in the second fitted GLM. We denote the predicted unpaid claims of accident year $i = 2, \ldots, I$ and the predicted total unpaid claims by

$$\hat{R}_i = \sum_{j=I+2-i}^{I} \hat{Y}_{ij} \quad \text{and} \quad \hat{R} = \sum_{i=2}^{I} \hat{R}_i.$$

The mean squared error of prediction (MSEP) of $(\hat{R}_i)_{i=2:I}$ is

$$\text{MSEP}(\hat{R}_i) = \mathbb{E}(R_i - \hat{R}_i)^2 = \text{Var}(R_i) + \text{Var}(\hat{R}_i) + (\mathbb{E}(R_i) - \mathbb{E}(\hat{R}_i))^2.$$

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

The evaluation of the process variance $\mathrm{Var}(R_i)$ is straightforward. We can substitute the estimated parameters into the process variance (6). The evaluation of the estimation variance $\mathrm{Var}(\hat{R}_i)$ is more complicated. A routine method is to apply Fisher's scoring method. The biased term $(\mathbb{E}(R_i) - \mathbb{E}(\hat{R}_i))^2$ is difficult to evaluate. Wüthrich (2003) simply assumed $\hat{R}_i$ is an unbiased estimation of $\mathbb{E}(R_i)$.

Here, we propose a parametric bootstrap method to approximate the process variance and the estimation variance. Based on the bootstrap sample of unpaid claims, we can also approximate tail-based risk measures such as VaR, TVaR, etc. This is an advantage of the bootstrap method over Fisher's scoring method. The tth bootstrap iteration includes the following steps:

1. Simulate a bootstrap sample of payment counts $N_{ij}^{u_t}$ in the upper triangle from Tweedie's distribution with mean $e_i \hat{\lambda}_{ij}^{\mathrm{MLE}}$, dispersion $\hat{\phi}$ and power parameter 1.
2. Fit a new Poisson GLM to the bootstrap sample N^{u_t} using the quasi-likelihood approach, estimate the parameters as $\hat{\lambda}^t$, $\hat{\phi}^t$ and predict the lower triangle as $\hat{N}_{ij}^{l_t} = e_i \hat{\lambda}_{ij}^t$.
3. If $N_{ij}^{u_t} = 0$, set the bootstrap sample of paid claims $Y_{ij}^{u_t} = 0$. If $N_{ij}^{u_t} \neq 0$, simulate a bootstrap sample of $Y_{ij}^{u_t}$ from Tweedie's distribution with mean $N_{ij}^{u_t} \hat{\tau}_{ij}^{\mathrm{MLE}}$, dispersion $\hat{\psi}/(N_{ij}^{u_t})^{m-1}$ and power parameter m, where $m = 1, 2, 3$ corresponds to Poisson, gamma and inverse-Gaussian distributions.
4. Fit a new GLM to the positive bootstrap sample Y^{+u_t} given N^{+u_t}, estimate the parameters as $\hat{\tau}^t$, $\hat{\psi}^t$, and predict the lower triangle as $\hat{Y}_{ij}^{l_t} = \hat{N}_{ij}^{l_t} \hat{\tau}_{ij}^t$. Note that if the column sum of the paid claims in a development year was zero, the coefficient of that development year could not be estimated. In such cases, we would use the estimated coefficient of that development year from the original GLM as its bootstrap estimate.
5. Calculate the predicted unpaid claims of accident years $i = 2, \ldots, I$ and the predicted total unpaid claims as

$$\hat{R}_i^t = \sum_{j=I+2-i}^{I} \hat{Y}_{ij}^{l_t} \quad \text{and} \quad \hat{R}^t = \sum_{i=2}^{I} \hat{R}_i^t.$$

6. Approximate the process variances of the accident year i unpaid claims and the total unpaid claims as

$$\widehat{\mathrm{Var}}(R_i^t) = \hat{\phi}^t e_i \sum_{j=I+2-i}^{I} \hat{\lambda}_{ij}^t (\hat{\tau}_{ij}^t)^2 + \hat{\psi}^t e_i \sum_{j=I+2-i}^{I} \hat{\lambda}_{ij}^t (\hat{\tau}_{ij}^t)^m$$

$$\text{and} \quad \widehat{\mathrm{Var}}(R^t) = \sum_{i=2}^{I} \widehat{\mathrm{Var}}(R_i^t).$$

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

Suppose the number of total bootstrap iterations is T. The bootstrap predictions of accident years i unpaid claims and the total unpaid claims are

$$\hat{R}_i^B = \frac{1}{T}\sum_{t=1}^{T}\hat{R}_i^t \text{ and } \hat{R}^B = \frac{1}{T}\sum_{t=1}^{T}\hat{R}^t.$$

The process variances of accident years i unpaid claims and the total unpaid claims are approximated as

$$\widehat{\text{Var}}(R_i) = \frac{1}{T}\sum_{t=1}^{T}\widehat{\text{Var}}(R_i^t) \text{ and } \widehat{\text{Var}}(R) = \frac{1}{T}\sum_{t=1}^{T}\widehat{\text{Var}}(R^t).$$

The estimation variances of accident years i unpaid claims and the total unpaid claims are approximated by the sample variance of $(\hat{R}_i^t)_{t=1:T}$ and $(\hat{R}^t)_{t=1:T}$ as follows:

$$\widehat{\text{Var}}(\hat{R}_i) = \frac{1}{T-1}\sum_{t=1}^{T}(\hat{R}_i^t - \hat{R}_i^B)^2 \text{ and } \widehat{\text{Var}}(\hat{R}) = \frac{1}{T-1}\sum_{t=1}^{T}(\hat{R}^t - \hat{R}^B)^2.$$

Note that the estimation variance does not have the additivity property as the process variance, since $(\hat{R}_i)_{i=2:I}$ are not independent, i.e.,

$$\text{Var}(\hat{R}) = \text{Var}\left(\sum_{i=2}^{I}\hat{R}_i\right) \neq \sum_{i=2}^{I}\text{Var}(\hat{R}_i).$$

Finally, the MSEP of unpaid claims can be approximated by summing up the process variance and the estimation variance.

3.3. The predictive distribution of unpaid claims

We can generate the unpaid claims $Y_{ij}^{I_t}$ for each bootstrap sampled parameter as follows:

1. Simulate the number of payments $N_{ij}^{I_t}$ from Tweedie's distribution with mean $e_i\hat{\lambda}_{ij}^t$, dispersion $\hat{\phi}$ and power parameter 1.
2. Simulate the unpaid claims $Y_{ij}^{I_t}$ from Tweedie's distribution with mean $N_{ij}^{I_t}\hat{\tau}_{ij}^t$, dispersion $\hat{\psi}^t/(N_{ij}^{I_t})^{m-1}$ and power parameter m, where $m = 1, 2, 3$ corresponds to Poisson, gamma and inverse-Gaussian distributions.
3. Predict accident years i unpaid claims and the total unpaid claims as

$$R_i^t = \sum_{j=I+2-i}^{I} Y_{ij}^{I_t} \text{ and } R^t = \sum_{i=2}^{I} R_i^t.$$

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

If no over-dispersion in payment counts was observed, we could apply the marginal distribution of R_i; see equations (4) and (5). For each bootstrap sampled parameter, we simulate R_i^t according to the following steps:

1. Simulate $\tilde{N}_i^t$ from a Poisson distribution with mean $e_i \sum_{j=I+2-i}^{I} \hat{\lambda}_{ij}^t$.
2. Simulate an i.i.d. $\tilde{N}_i^t$ sample from a mixture distribution with the following probability density function:

$$p(\tilde{X}_i) = \sum_{j=I+2-i}^{I} w_{ij}^t p_X\left(\tilde{X}_i \middle| \hat{\tau}_{ij}^t, \hat{\psi}^t\right),$$

   where $w_{ij}^t = (\hat{\gamma}_j^{\text{Dev}})^t / \sum_{s=I+2-i}^{I} (\hat{\gamma}_s^{\text{Dev}})^t$ is the weight of component distribution, and $p_X(x|a, b)$ denotes the assumed probability density function of individual payment with mean a and dispersion b evaluated at x.
3. Sum up these $\tilde{N}_i^t$ values as R_i^t.

Note that the variation in $(R_i^t)_{t=1:T}$ and $(R^t)_{t=1:T}$ contains both process variance and estimation variance, while the variation in $(\hat{R}_i^t)_{t=1:T}$ and $(\hat{R}^t)_{t=1:T}$ only contains estimation variance. The empirical distributions of $(R_i^t)_{t=1:T}$ and $(R^t)_{t=1:T}$ approximate the predictive distributions of R_i and R. The sample variance of $(R_i^t)_{t=1:T}$ and $(R^t)_{t=1:T}$ approximate the MSEP of $\hat{R}_i^B$ and $\hat{R}^B$. The tail-based risk measures such as VaR, TVaR, etc can also be approximated from the empirical distribution of $(R_i^t)_{t=1:T}$ and $(R^t)_{t=1:T}$.

4. Example: Swiss motor insurance

We analyze the Swiss Motor Insurance data which was studied by Wüthrich (2003). It contains a paid claims run-off triangle and a payment counts run-off triangle, both of which have 9 accident years and 11 development years. The exposure measure for each accident year is the number of reported claims. We assume $I = 11$ and treat the paid claims and the number of payments in the accident years 10 and 11 as missing values. We do not predict for the accident years 10 and 11.

First, we fit a Poisson GLM to the numbers of payments. The quasi-likelihood approach is applied to account for possible over-dispersion. The dispersion parameter is estimated as $\hat{\phi} = 4.91$, much larger than 1. The ignorance of over-dispersion would underestimate the variance in paid claims. We draw the studentized residual plots in Figure 1. The first plot indicates a heterosedasticity. However, this is not necessarily due to an inappropriate distribution assumed. It might be caused by the triangle structured data, i.e., large values are more than small values. A further investigation using a Tweedie GLM indicates that the power parameter of 1.2 leads to a satisfying residual plot. Here, we continue with the Poisson distribution. The last three plots in Figure 1 indicate that the chain ladder mean is a suitable assumption for this particular data set.

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

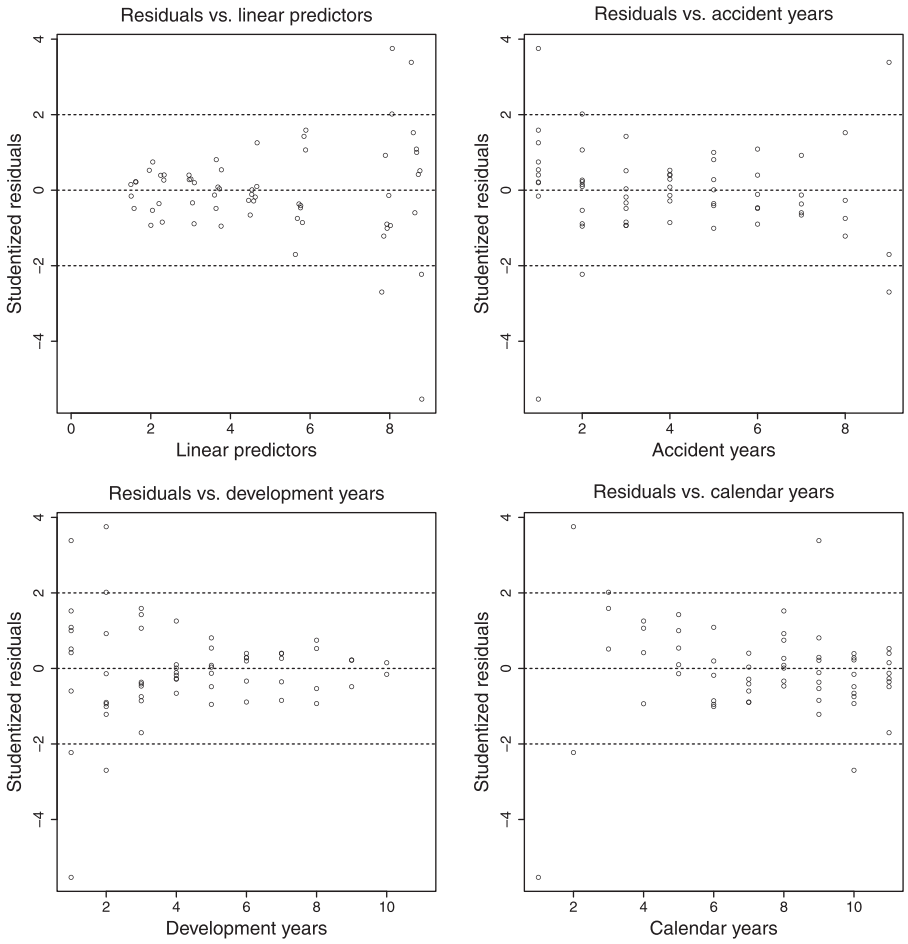


FIGURE 1: The studentized residuals from the Poisson GLM fitting to the numbers of payments.

Next, we fit a gamma GLM to the paid claims given the number of payments. The regression function for the mean parameter is assumed as

$$\mathbb{E}(Y_{ij}|N_{ij}) = N_{ij}\mathbb{E}(X_{ij}^{\kappa}) = N_{ij}\upsilon(j) = N_{ij}\beta_j^{\text{Dev}}.$$

The dispersion parameter is estimated as $\hat{\psi} = 6.00$. Hence, the process variance is inflated by $(4.91 + 6.00)/(1 + 6.00) = 1.56$ due to the over-dispersion in the number of payments. We back-test the assumption of constant accident year effects and constant calendar year effects. The top-row plots in Figure 2 indicate that the average individual payment changes with the development year. The bottom-row plots indicate that both the accident year effects and the calendar effects are roughly constant. We apply the bootstrap method with total $T = 5,000$ iterations. The predicted unpaid claims and the associated uncertainty

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

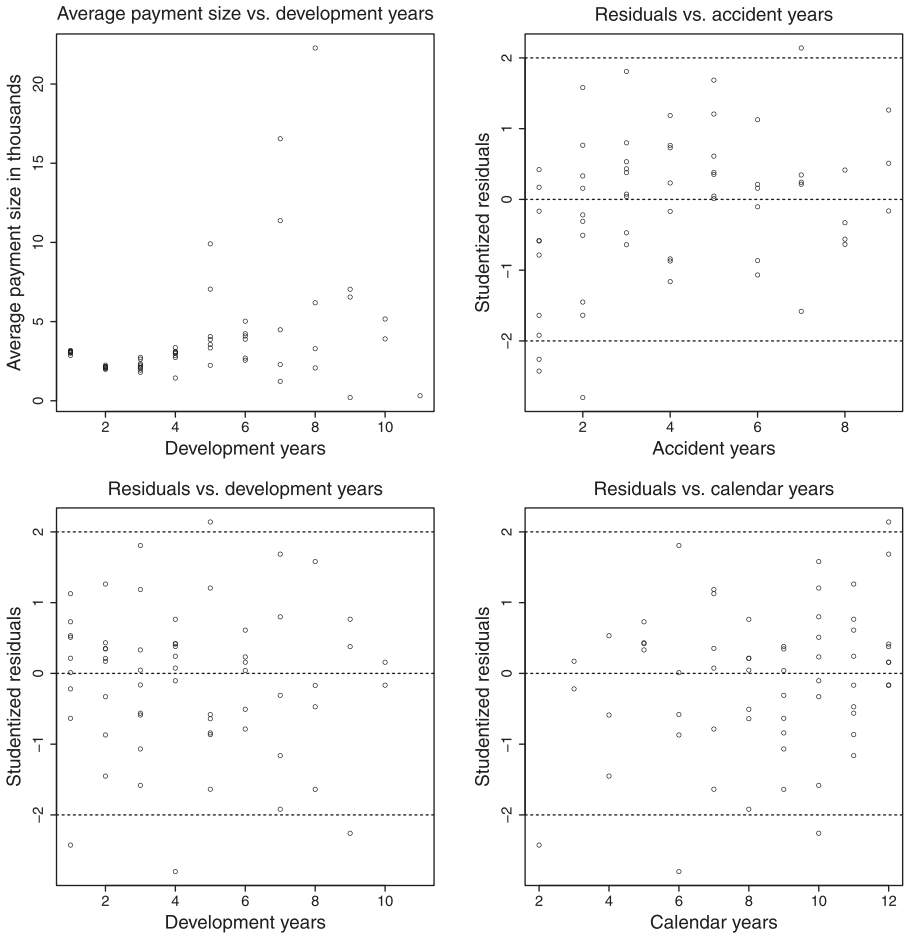


FIGURE 2: Top-left: the relationship between the average individual payment and the development year. The rest: the studentized residuals from the gamma GLM fitting to the paid claims given the numbers of payments.

measures are listed in Table 1. For the last two columns, we generate a sample of $(R_i^t)_{t=1:T}$ using the first method in Section 3.3. The sample standard deviation of $(R_i^t)_{t=1:T}$ is shown in the "Sample Sd". column. The 95% VaR is approximated from the empirical distribution of $(R_i^t)_{t=1:T}$. We denote this model by Model 1.

For comparison, we fit another five models: Model 2 considers (a) the over-dispersion in the payment counts, (b) the development year effects and (c) the accident year effects in the individual payments, i.e., assuming a chain ladder mean for the individual payments; Model 3 only considers (a), i.e., assuming a constant mean for the individual payments; Model 4 only considers (b); Model A is exactly the proposed model by Wüthrich (2003), considering both (b) and (c); Model B is an ODP model fitting to paid claims (England and Verrall, 2002). Remind that Model 1 considers both (a) and (b). Note that the inference

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

TABLE 1

PREDICTIONS FROM THE COMPOUND POISSON-GAMMA MODEL.

AY	Prediction	Estimation Err.	Process Err.	$\text{MSEP}^{1/2}$	Sample Sd.	95% VaR
2	319	1,040	1,486	1,813	2,424	1,366
3	20,694	21,933	39,486	45,168	45,337	103,599
4	39,730	26,785	52,151	58,627	63,334	166,379
5	85,120	38,210	82,119	90,573	93,254	272,832
6	134,572	44,624	100,154	109,646	109,698	347,961
7	200,410	47,798	112,638	122,360	121,773	427,504
8	365,251	57,715	146,539	157,495	160,087	649,394
9	577,817	62,961	165,454	177,029	176,052	892,337
Total	1,423,914	249,778	287,387	380,763	385,076	2,141,646

The "$\text{MSEP}^{1/2}$" column is calculated as the square root of the sum of estimation variance and process variance. The "Sample Sd". column is the sample standard deviation of $(R_i^t)_{t=1:T}$.

TABLE 2

PREDICTIONS OF TOTAL UNPAID CLAIMS FROM SIX MODELS.

Models	Attributes	Prediction	Estimation Err.	Process Err.	$\text{MSEP}^{1/2}$
Model 1	(a) (b)	1,423,914	249,778	287,387	380,763
Model 2	(a) (b) (c)	1,443,156	255,072	292,699	388,245
Model 3	(a)	911,691	107,568	381,249	396,134
Model 4	(b)	1,415,869	192,537	228,907	299,113
Model A	(b) (c)	1,451,299	179,890	203,355	271,503
Model B	using paid claims	1,461,360	216,965	231,403	371,208

Model 1 considers (a) the over-dispersion in payment counts and (b) the development year effects in individual payment. Model 2 considers (a), (b) and (c) the accident year effects in individual payment. Model 3 only considers (a); Model 4 only considers (b); Model A is the model proposed by Wüthrich (2003); Model B is the ODP model fitting to paid claims.

methods for Models A and B follow the corresponding references rather than our proposed methods. We list the predicted total unpaid claims and the associated uncertainty measures from the six models in Table 2. Comparing Models 1, 4, A and B, we conclude that the ignorance of over-dispersion in payment counts would underestimate the MSEP of unpaid claims. Comparing Models 1 and B, we conclude that using counts data and multiple sub-models might slightly degrade the MSEP (around 3% for this data set). Comparing Models 1 and 2, we conclude that a suitable simplified mean function for individual payment would decrease both the estimation error and the prediction error. Comparing Models 1 and 3, we conclude that an over-simplified mean function for individual payment would decrease the estimation error, but increase the process error and the bias.

Finally, we test the assumption of gamma distributed individual payments. We draw the residual plots from the Poisson/gamma/inverse-Gaussian GLMs

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

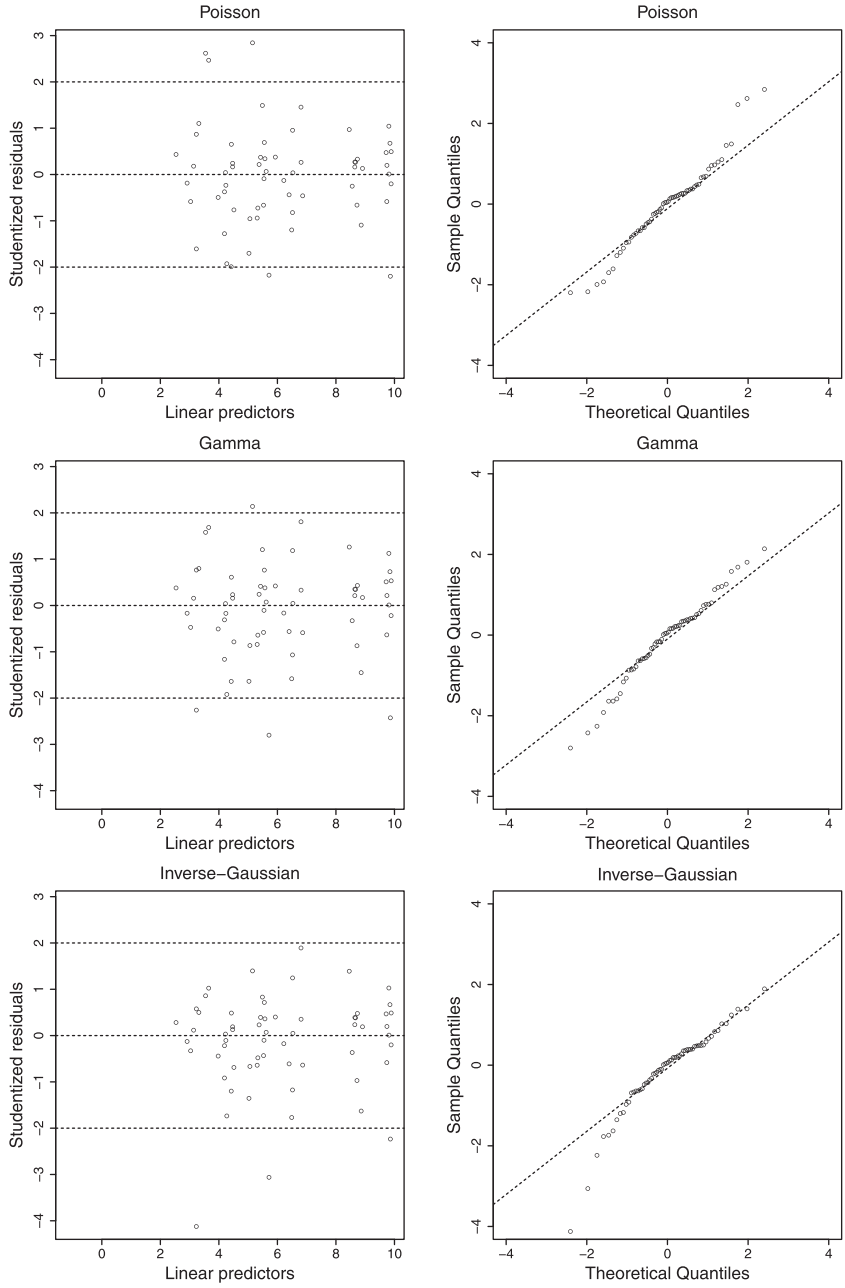


FIGURE 3: The scatter plots and the qq plots of the studentized residuals from the Poisson GLM (top row), the gamma GLM (middle row) and the inverse-Gaussian GLM (bottom row), fitting to the paid claims given the numbers of payments.

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

TABLE 3

PREDICTIONS FROM THE COMPOUND POISSON–POISSON MODEL.

AY	Prediction	Estimation Err.	Process Err.	$\text{MSEP}^{1/2}$	Sample Sd.	95% Var
2	243	2,488	3,375	4,193	3,485	69
3	20,373	21,297	34,115	40,216	41,104	105,550
4	39,483	26,332	46,441	53,387	53,659	149,979
5	85,013	34,921	70,588	78,753	80,481	244,687
6	134,055	40,854	87,149	96,249	95,382	313,804
7	199,593	44,970	101,893	111,375	111,584	408,342
8	364,550	54,525	135,537	146,094	145,437	621,450
9	578,032	61,610	162,017	173,336	172,068	879,745
Total	1,421,341	235,352	266,292	355,390	355,612	2,042,128

The "$\text{MSEP}^{1/2}$" column is calculated as the square root of the sum of estimation variance and process variance. The "Sample Sd". column is the sample standard deviation of $(R_i^t)_{t=1:T}$.

fitting to the paid claims given the number of payments in Figure 3. The Poisson distribution also seems reasonable. The prediction results are listed in Table 3. Whether setting reserves according to Tables 1 or 3 might depend on actuaries' judgment. The Poisson distribution is light tailed compared to the gamma distribution and the inverse-Gaussian distribution. We must be careful with the Poisson assumption even when the residual plot supports it. A better decision would be made if we could obtain the individual payments data set.

5. CONCLUSIONS

We extend the assumptions in compound Poisson claims reserving model and propose a new inference method. Our proposed models are more general, facilitating the over-dispersion in number of payments, various mean and variance structures in individual payments. The parameter estimation method is straightforward and natural, applying two GLMs consecutively. A bootstrap method is proposed based on the two GLMs, involving bootstrapping the number of payments and the paid claims consecutively.

There are three important findings. First, the over-dispersion in the number of payments should be addressed. It inflates the process variance of paid claims by a factor $(\phi + \psi)/(1 + \psi)$. Second, the individual payments distribution affects the uncertainty of predicted unpaid claims. We might rely on the residual plots from the GLM fitting to the paid claims given the number of payments to select a suitable individual payment distribution. Ideally, we hope get access to the individual payments data to make a better judgment. Third, for individual payments, we can always assume a simpler mean function than the chain ladder mean, and incorporate the extracted covariates from the payment counts into the mean function. In the example, we only consider the development year covariate and use the logarithm of payment counts as the intercept. The first

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

two points affect the uncertainty measures of unpaid claims much more than their means, while the third point affects both the uncertainty measures and the means. When setting reserves according to some risk measure of unpaid claims, one must be aware of these three points.

Acknowledgements

We appreciate the constructive comments made by three anonymous reviewers and the editor Mario Wüthrich. Their comments helped improve the quality of this paper significantly. We acknowledge the financial support from National Social Science Fund of China (Grant no. 16ZDA052) and MOE National Key Research Bases for Humanities and Social Sciences (Grant no. 16JJD910001).

References

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

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

Badescu, A.L., Lin, S. and Tang, D. (2016) A marked Cox model for the number of IBNR claims: Theory. *Insurance: Mathematics and Economics,* **69**, 29–37.

Boucher, J.-P. and Davidov, D. (2011) On the importance of dispersion modeling for claims reserving: An application with the Tweedie distribution. *Variance,* **5**, 158–172.

England, P.D. and Verrall, R.J. (2002) Stochastic claims reserving in general insurance. *British Actuarial Journal,* **8**, 443–518.

Gigante, P., Picech, L. and Sigalotti, L. (2013) Claims reserving in the hierarchical generalized linear model framework. *Insurance: Mathematics and Economics,* **52**, 381–390.

Gigante, P., Picech, L. and Sigalotti, L. (2016) A mixture model for payments and payment numbers in claims reserving. *ASTIN Bulletin,* **47**, 1–29.

Klugman, S.A., Panjer, H.H. and Willmot, G.E. (2012) *Loss Models: From Data to Decisions*, 4th ed. New York: John Wiley & Sons.

Lee, Y. and Nelder, J.A. (2001) Hierarchical generalised linear models: A synthesis of generalised linear models, random-effect models and structured dispersions. *Biometrika,* **88**, 987–1006.

Martínez-Miranda, M.D., Nielsen, B., Nielsen, J.P. and Verrall, R. (2011) Cash flow simulation for a model of outstanding liabilities based on claim amounts and claim numbers. *ASTIN Bulletin,* **41**, 107–129.

Martínez-Miranda, M.D., Nielsen, J.P. and Verrall, R. (2012) Double chain ladder. *ASTIN Bulletin,* **42**, 59–76.

Martínez-Miranda, M.D., Nielsen, J.P. and Verrall, R. (2013) Double chain ladder and Bornhuetter–Ferguson. *North American Actuarial Journal,* **17**, 101–113.

Martínez-Miranda, M.D., Nielsen, J.P., Verrall, R. and Wüthrich, M.V. (2015) Double chain ladder, claims development inflation and zero-claims. *Scandinavian Actuarial Journal,* **2015**, 383–405.

McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models*, 2nd ed. New York: Chapman & Hall.

Norberg, R. (1993) Prediction of outstanding liabilities in non-life insurance. *ASTIN Bulletin,* **23**, 95–115.

Peters, G.W., Shevchenko, P.V. and Wüthrich, M.V. (2009) Model uncertainty in claims reserving within Tweedie's compound Poisson models. *ASTIN Bulletin,* **39**, 1–33.

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

RENSHAW, A.E. and VERRALL, R.J. (1998) A stochastic model underlying the chain-ladder technique. *British Actuarial Journal,* **4**, 903–923.
SMYTH, G.K. and JØRGENSEN, B. (2002) Fitting Tweedie's compound Poisson model to insurance claims data: Dispersion modelling. *ASTIN Bulletin,* **32**, 143–157.
TAYLOR, G. (2000) *Loss Reserving: An Actuarial Perspective*. Huebner International Series on Risk, Insurance and Economic Security. Boston: Kluwer Academic Publishers.
TAYLOR, G. and XU, J. (2016) An empirical investigation of the value of finalisation count information to loss reserving. *Variance,* **10**, 75–120.
VERRALL, R., NIELSEN, J.P. and JESSEN, A.H. (2010) Prediction of RBNS and IBNR claims using claim amounts and claim counts. *ASTIN Bulletin,* **40**, 871–887.
VERRALL, R.J. and WÜTHRICH, M.V. (2016) Understanding reporting delay in general insurance. *Risks,* **4**, 25.
WRIGHT, T.S. (1990) A stochastic method for claims reserving in general insurance. *Journal of the Institute of Actuaries,* **117**, 677–731.
WÜTHRICH, M.V. (2003) Claims reserving using Tweedie's compound Poisson model. *ASTIN Bulletin,* **33**, 331–346.
WÜTHRICH, M.V. and MERZ, M. (2008) *Stochastic Claims Reserving Methods in Insurance*. Chichester: John Wiley & Sons.
WÜTHRICH, M.V. and MERZ, M. (2015) *Stochastic claims reserving manual: Advances in dynamic modeling*. Swiss Finance Institute Research Paper no. 15–34.

SHENGWANG MENG
Center for Applied Statistics and School of Statistics
Renmin University of China
Beijing 100872, China
E-Mail: mengshw@ruc.edu.cn

GUANGYUAN GAO (Corresponding author)
Center for Applied Statistics and School of Statistics
Renmin University of China
Beijing 100872, China
E-Mail: guangyuan.gao@ruc.edu.cn

APPENDIX A. THE DISTRIBUTION OF R_i

Denote by $P_{N_{ij}}(t)$ the probability generating function (pgf) of N_{ij} and $M_{X_{ij}}(t)$ the moment generating function (mgf) of $(X_{ij}^k)_{k=1:N_{ij}}$. It is straightforward to get

$$P_{N_{ij}}(t) = e^{e_i \lambda_{ij}(t-1)}.$$

From the introductory statistics course (Klugman *et al.*, 2012), it can be shown that the mgf of Y_{ij} is

$$M_{Y_{ij}}(t) = P_{N_{ij}}(M_{X_{ij}}(t)) = \exp\left[e_i \lambda_{ij}\left(M_{X_{ij}}(t) - 1\right)\right].$$

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

The mgf of R_i is

$$\begin{aligned}
M_{R_i}(t) &= \prod_{j=I+2-i}^{I} M_{Y_{ij}(t)} \\
&= \prod_{j=I+2-i}^{I} \exp\left[e_i \lambda_{ij} \left(M_{X_{ij}}(t) - 1\right)\right] \\
&= \exp\left[\sum_{j=I+2-i}^{I} e_i \lambda_{ij} M_{X_{ij}}(t) - \sum_{j=I+2-i}^{I} e_i \lambda_{ij}\right] \\
&= \exp\left[e_i \gamma_i^{\text{Acc}} \sum_{j=I+2-i}^{I} \gamma_j^{\text{Dev}} \times \left(\sum_{j=I+2-i}^{I} \frac{\gamma_j^{\text{Dev}}}{\sum_{s=I+2-i}^{I} \gamma_s^{\text{Dev}}} M_{X_{ij}}(t) - 1\right)\right] \\
&= \exp\left[e_i \gamma_i^{\text{Acc}} \sum_{j=I+2-i}^{I} \gamma_j^{\text{Dev}} \times \left(\sum_{j=I+2-i}^{I} w_{ij} M_{X_{ij}}(t) - 1\right)\right],
\end{aligned}$$

where $w_{ij} = \gamma_j^{\text{Dev}} / \sum_{s=I+2-i}^{I} \gamma_s^{\text{Dev}}$ in the last line. The mgf of R_i is the same as the following random sum:

$$\sum_{l=1}^{\tilde{N}_i} \tilde{X}_i^l,$$

where $\tilde{N}_i$ follows a Poisson distribution with the mean of $e_i \gamma_i^{\text{Acc}} \sum_{j=I+2-i}^{I} \gamma_j^{\text{Dev}}$ and $(\tilde{X}_i^l)_{1 \le l \le \tilde{N}_i}$ follows a mixture distribution with the pdf

$$p(\tilde{X}_i^l) = \sum_{j=I+2-i}^{I} w_{ij} \, p_X\left(\tilde{X}_i^l | \tau_{ij}, \psi\right),$$

where w_{ij} is the weight of component distribution, and $p_X(x|a, b)$ denotes the assumed pdf of individual payment with mean a and dispersion b evaluated at x.

APPENDIX B. QUASI-LIKELIHOOD APPROACH, REVISITED

We apply the quasi-likelihood approach twice: first, it provides means for modeling payment counts, which are more variable than the Poisson distribution predicts; second, it provides means for modeling paid claims for which the variance is expected to be proportional to the mean. We demonstrate the quasi-likelihood approach in the first scenario. A similar proce-

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press

dure applies for the second scenario. The log quasi-likelihood function for N_{ij} is defined as

$$l_{ij}(\lambda_{ij}) = \int_{N_{ij}}^{e_i\lambda_{ij}} \frac{N_{ij}-z}{\phi V(z)}dz = \int_{N_{ij}}^{e_i\lambda_{ij}} \frac{N_{ij}-z}{\phi z}dz,$$

where the variance function $V(\mu) = \mu$ and $\lambda_{ij} = \gamma_i^{\text{Acc}}\gamma_j^{\text{Dev}}$. Then, the log quasi-likelihood function for N^u is defined as

$$l(\lambda) = \sum_{(i,j)\in\Omega^u} l_{ij}(\lambda_{ij}).$$

Consider obtaining the MLEs of γ. Differentiating l w.r.t. γ_i^{Acc} or γ_j^{Dev} yields

$$\frac{\partial l}{\partial \gamma_i^{\text{Acc}}} = \sum_{(s,t)\in\Omega^u} \frac{N_{st}-e_s\lambda_{st}}{\phi e_s\lambda_{st}} \frac{\partial e_s\lambda_{st}}{\partial \gamma_i^{\text{Acc}}} = \sum_{(s,t)\in\Omega^u} 1_{\{s=i\}} \frac{N_{st}-e_s\gamma_s^{\text{Acc}}\gamma_t^{\text{Dev}}}{\phi\gamma_s^{\text{Acc}}},$$

$$\frac{\partial l}{\partial \gamma_j^{\text{Dev}}} = \sum_{(s,t)\in\Omega^u} \frac{N_{st}-e_s\lambda_{st}}{\phi e_s\lambda_{st}} \frac{\partial e_s\lambda_{st}}{\partial \gamma_j^{\text{Dev}}} = \sum_{(s,t)\in\Omega^u} 1_{\{t=j\}} \frac{N_{st}-e_s\gamma_s^{\text{Acc}}\gamma_t^{\text{Dev}}}{\phi\gamma_t^{\text{Dev}}},$$

so that the parameter estimates are solutions to the equations

$$\sum_{(s,t)\in\Omega^u} 1_{\{s=i\}} \frac{N_{st}-e_s\gamma_s^{\text{Acc}}\gamma_t^{\text{Dev}}}{\phi\gamma_s^{\text{Acc}}} = 0 \quad \text{for } i = 1,\ldots,I,$$

$$\sum_{(s,t)\in\Omega^u} 1_{\{t=j\}} \frac{N_{st}-e_s\gamma_s^{\text{Acc}}\gamma_t^{\text{Dev}}}{\phi\gamma_t^{\text{Dev}}} = 0 \quad \text{for } j = 1,\ldots,I.$$

The above equations are equivalent to

$$\gamma_i^{\text{Acc}} = \frac{\sum_{t=1}^{I-i+1} N_{it}}{e_i \sum_{t=1}^{I-i+1} \gamma_t^{\text{Dev}}} \quad \text{for } i = 1,\ldots,I,$$

$$\gamma_j^{\text{Dev}} = \frac{\sum_{s=1}^{I-j+1} N_{sj}}{\sum_{s=1}^{I-j+1} e_s\gamma_s^{\text{Acc}}} \quad \text{for } j = 1,\ldots,I.$$

Iterating these equations will give the MLEs of γ, and lead to the same prediction as the chain ladder prediction for N^l (Renshaw and Verrall, 1998). Note that the MLEs of γ can be obtained without knowing the dispersion parameter ϕ. Finally, the dispersion parameter is estimated as

$$\hat{\phi} = \frac{1}{n-q} \sum_{(i,j)\in\Omega^u} \frac{(N_{ij}-e_i\hat{\lambda}_{ij}^{\text{MLE}})^2}{e_i\hat{\lambda}_{ij}^{\text{MLE}}},$$

where the number of observations $n = I(I+1)/2$ and the number of parameters $q = 2I-1$.

https://doi.org/10.1017/asb.2018.12 Published online by Cambridge University Press