Insurance: Mathematics and Economics 46 (2010) 568–579

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Paid–incurred chain claims reserving method

Michael Merz [a], Mario V. Wüthrich [b,*]

[a] University of Hamburg, Department of Business Administration, 20146 Hamburg, Germany

[b] ETH Zurich, Department of Mathematics, 8092 Zurich, Switzerland

ARTICLE INFO

Article history:
Received July 2009
Received in revised form
February 2010
Accepted 16 February 2010

Keywords:
Claims reserving
Outstanding loss liabilities
Ultimate loss
Claims payments
Claims incurred
Incurred losses
Prediction uncertainty

ABSTRACT

We present a novel stochastic model for claims reserving that allows us to combine claims payments and incurred losses information. The main idea is to combine two claims reserving models (Hertig's (1985) model and Gogol's (1993) model) leading to a log-normal paid–incurred chain (PIC) model. Using a Bayesian point of view for the parameter modelling we derive in this Bayesian PIC model the full predictive distribution of the outstanding loss liabilities. On the one hand, this allows for an analytical calculation of the claims reserves and the corresponding conditional mean square error of prediction. On the other hand, simulation algorithms provide any other statistics and risk measure on these claims reserves.

1. Paid–incurred chain model

1.1. Introduction

The main task of reserving actuaries is to predict ultimate loss ratios and outstanding loss liabilities. In general, this prediction is based on past information that comes from different sources of information. In order to get a unified prediction of the outstanding loss liabilities one needs to rank these information channels by assigning credibility weights to the available information. Often this is a difficult task. Therefore, most classical claims reserving methods are based on one information channel only (for instance, claims payments *or* incurred losses data).

Halliwell (1997, 2009) was probably one of the first who investigated the problem of combining claims payments and incurred losses data for claims reserving from a statistical point of view. The analysis of Halliwell (1997, 2009) as well as of Venter (2008) is done in a regression framework.

A second approach to unify claims prediction based on claims payments and incurred losses is the Munich chain ladder (MCL) method. The MCL method was introduced by Quarg and Mack (2004) and their aim is to reduce the gap between the two chain ladder (CL) predictions that are based on claims payments and incurred losses data, respectively. The idea is to adjust the CL factors with incurred–paid ratios to reduce the gap between the two predictions (see Quarg and Mack, 2004; Verdier and Klinger, 2005; Merz and Wüthrich, 2006 and Liu and Verrall, 2008). The difficulty with the MCL method is that it involves several parameter estimations whose precision is difficult to quantify within a stochastic model framework.

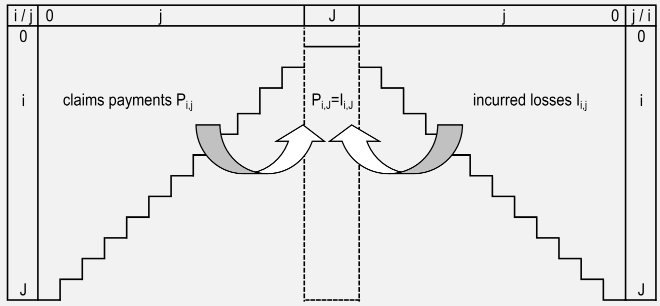
A third approach was presented in Dahms (2008). Dahms considers the complementary loss ratio method (CLRM) where the underlying volume measures are the case reserves which is the basis for the regression and CL analysis. Dahms' CLRM can also be applied to incomplete data and he derives an estimator for the prediction uncertainty.

In this paper we present a novel claims reserving method which is based on the combination of Hertig's log-normal claims reserving model (Hertig, 1985) for claims payments and of Gogol's Bayesian claims reserving model (Gogol, 1993) for incurred losses data. The idea is to use Hertig's model for the prior ultimate loss distribution needed in Gogol's model which leads to a paid–incurred chain (PIC) claims reserving method. Using basic properties of multivariate Gaussian distributions we obtain a mathematically rigorous and consistent model for the combination of the two information channels claims payments and incurred losses data. The analysis will attach credibility weights to these sources of information and it will also involve incurred–paid ratios (similar to the MCL method, see Quarg and Mack (2004)). Our PIC model will provide *one single* estimate for the claims reserves (based on both information channels) and, moreover, it has the advantage that we can quantify the prediction uncertainty and that it allows for complete model simulations. This means that this PIC model allows for the derivation of the full predictive distribution

* Corresponding author.
E-mail address: mario.wuethrich@math.ethz.ch (M.V. Wüthrich).

0167-6687/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.insmatheco.2010.02.004

Table 1
Left-hand side: cumulative claims payments $P_{i,j}$ development triangle; Right-hand side: incurred losses $I_{i,j}$ development triangle; both leading to the same ultimate loss $P_{i,J} = I_{i,J}$.



of the outstanding loss liabilities. Endowed with the simulated predictive distribution one is not only able to calculate estimators for the first two moments but one can also calculate any other risk measure, like Value-at-Risk or expected shortfall.

Posthuma et al. (2008) were probably the first who studied a PIC model. Under the assumption of multivariate normality they formed a claims development chain for the increments of claims payments and incurred losses. Their model was then treated in the spirit of generalized linear models similar to Venter (2008). Our model will be analyzed in the spirit of the Bayesian chain ladder link ratio models (see Bühlmann et al., 2009).

1.2. Notation and model assumptions

For the PIC model we consider two channels of information: (i) claims payments, which refer to the payments done for reported claims; (ii) incurred losses, which correspond to the reported claim amounts. Often, the difference between incurred losses and claims payments is called case reserves for reported claims. Ultimately, claims payments and incurred losses must reach the same value (when all the claims are settled).

In many cases, statistical analysis of claims payments and incurred losses data is done by accident years and development years, which leads to the so-called claims development triangles (see Table 1, and Chapter 1 in Wüthrich and Merz (2008)). In the following, we denote accident years by $i \in \{0, \ldots, J\}$ and development years by $j \in \{0, \ldots, J\}$. We assume that all claims are settled after the Jth development year. Cumulative claims payments in accident year i after j development periods are denoted by $P_{i,j}$ and the corresponding incurred losses by $I_{i,j}$. Moreover, for the ultimate loss we assume $P_{i,J} = I_{i,J}$ with probability 1, which means that ultimately (at time J) they reach the same value. For an illustration we refer to Table 1.

We define for $j \in \{0, \ldots, J\}$ the sets

$$\mathcal{B}_j^P = \left\{P_{i,l} : 0 \le i \le J,\ 0 \le l \le j\right\},$$

$$\mathcal{B}_j^I = \left\{I_{i,l} : 0 \le i \le J,\ 0 \le l \le j\right\},$$

$$\mathcal{B}_j = \mathcal{B}_j^P \cup \mathcal{B}_j^I,$$

the paid, incurred and joint paid and incurred data, respectively, up to development year j.

Model Assumption 1.1 (*Log-Normal PIC Model*).

- Conditionally, given $\boldsymbol{\Theta} = (\Phi_0, \ldots, \Phi_J, \Psi_0, \ldots, \Psi_{J-1}, \sigma_0, \ldots, \sigma_J, \tau_0, \ldots, \tau_{J-1})$, we have:
  - the random vector $(\xi_{0,0}, \ldots, \xi_{J,J}, \zeta_{0,0}, \ldots, \zeta_{J,J-1})$ has a multivariate Gaussian distribution with uncorrelated components given by
    $\xi_{i,j} \sim \mathcal{N}\left(\Phi_j, \sigma_j^2\right)$ for $i \in \{0, \ldots, J\}$ and $j \in \{0, \ldots, J\}$,
    $\zeta_{k,l} \sim \mathcal{N}\left(\Psi_l, \tau_l^2\right)$ for $k \in \{0, \ldots, J\}$ and $l \in \{0, \ldots, J-1\}$;
  - cumulative payments $P_{i,j}$ are given by the recursion
    $P_{i,j} = P_{i,j-1} \exp\left\{\xi_{i,j}\right\}$, with initial value $P_{i,0} = \exp\left\{\xi_{i,0}\right\}$;
  - incurred losses $I_{i,j}$ are given by the (backwards) recursion
    $I_{i,j-1} = I_{i,j} \exp\left\{-\zeta_{i,j-1}\right\}$, with initial value $I_{i,J} = P_{i,J}$.
- The components of $\boldsymbol{\Theta}$ are independent and $\sigma_j, \tau_j > 0$ for all j. □

Remarks.

- This PIC model combines both cumulative payments and incurred losses data to get a unified predictor for the ultimate loss that is based on both sources of information. Thereby, the model assumption $I_{i,J} = P_{i,J}$ guarantees that the ultimate loss coincides for claims payments and incurred losses data. This means that in this PIC model there is no gap between the two predictors based on cumulative payments and incurred losses, respectively. This is similar to Section 4 in Posthuma et al. (2008) and to the CLRM (see Dahms, 2008), but this is different to the MCL method (see Quarg and Mack, 2004).
- The cumulative payments $P_{i,j}$ satisfy Hertig's (1985) model, conditional on the parameters $\boldsymbol{\Theta}$. The model assumption $I_{i,J} = P_{i,J}$ also implies that we assume

$$E\left[P_{i,J}\middle|\boldsymbol{\Theta}\right] = E\left[I_{i,J}\middle|\boldsymbol{\Theta}\right] = \exp\left\{\sum_{m=0}^{J} \Phi_m + \sigma_m^2/2\right\}, \tag{1.1}$$

  see also (2.2). Henceforth, incurred losses $I_{i,j}$ satisfy Gogol's (1993) model with prior ultimate loss mean $E\left[P_{i,J}|\boldsymbol{\Theta}\right]$.
- The assumption $I_{i,J} = P_{i,J}$ means that all claims are settled after J development years and there is no so-called tail development factor. If there is a claims development beyond development year J, then one can extend the PIC model for the estimation of a tail development factor. Because this inclusion of a tail development factor requires rather extended derivations and discussions we provide the details in Merz and Wüthrich (submitted for publication).
- We assume conditional independence between all $\xi_{i,j}$'s and $\zeta_{k,l}$'s. One may question this assumption, especially, because Quarg and Mack (2004) found high correlations between incurred–paid ratios. In the Example section (see Section 5) we calculate the implied posterior correlation between the Φ_j's and the Ψ_l's (see Table 12). Our findings are that these correlations for our data set are fairly small (except in regions where we have only a few observations). Therefore, we refrain from introducing dependence between the components. However, this dependence could be implemented but then the solutions can only be found numerically and, moreover, the estimation of the correlation matrix is not obvious.
- We choose a log-normal PIC model. This has the advantage that the conditional distributions of $P_{i,J}$, given $\boldsymbol{\Theta}$ and $\mathcal{B}_j$, $\mathcal{B}_j^P$ or $\mathcal{B}_j^I$, respectively, can be calculated explicitly. Other distributional assumptions only allow for numerical solutions using simulations with missing data (see, for instance van Dyk and Meng, 2001).

Organisation. In the next section we are going to give the model properties and first model interpretations conditional on the knowledge of $\mathbf{\Theta}$. In Section 3 we discuss the estimation of the underlying model parameters $\mathbf{\Theta}$. In Section 4 we discuss prediction uncertainty and in Section 5 we provide an example. All proofs of the statements are given in the Appendix.

2. Simultaneous payments and incurred losses consideration

2.1. Cumulative payments

Our first observation is that, given $\mathbf{\Theta}$, cumulative payments $P_{i,j}$ satisfy the assumptions of Hertig's (1985) log-normal CL model (see also Section 5.1 in Wüthrich and Merz, 2008). That is, conditional on $\mathbf{\Theta}$, we have for $j \geq 0$

$$\log \frac{P_{i,j}}{P_{i,j-1}}\Bigg|_{\{\mathcal{B}^P_{j-1},\mathbf{\Theta}\}} \sim \mathcal{N}\left(\Phi_j, \sigma_j^2\right),$$

where we have set $P_{i,-1} = 1$. This gives the CL property (see also Lemma 5.2 in Wüthrich and Merz, 2008)

$$E\left[P_{i,j}|\mathcal{B}^P_{j-1}, \mathbf{\Theta}\right] = P_{i,j-1}\exp\left\{\Phi_j + \sigma_j^2/2\right\}. \tag{2.1}$$

The tower property for conditional expectations (see, for example Williams, 1991, 9.7 (i)) then implies for the expected ultimate loss, given $\{\mathcal{B}^P_j, \mathbf{\Theta}\}$,

$$E\left[P_{i,J}|\mathcal{B}^P_j, \mathbf{\Theta}\right] = P_{i,j}\exp\left\{\sum_{l=j+1}^{J}\Phi_l + \sigma_l^2/2\right\}. \tag{2.2}$$

2.2. Incurred losses

The model properties of incurred losses $I_{i,j}$ are in the spirit of Gogol's (1993) model. Namely, given $\mathbf{\Theta}$, the ultimate loss $I_{i,J} = P_{i,J}$ has a log-normal distribution and, conditional on $I_{i,J}$ and $\mathbf{\Theta}$, the incurred losses $I_{i,j}$ have also a log-normal distribution. This then allows for Bayesian inference on $I_{i,J}$, given $\mathcal{B}^I_j$, similar to Lemma 4.21 in Wüthrich and Merz (2008). The key lemma is the following well-known property for multivariate Gaussian distributions (see e.g. Appendix A in Posthuma et al., 2008):

Lemma 2.1. *Assume $(X_1, \ldots, X_n)'$ is multivariate Gaussian distributed with mean $(m_1, \ldots, m_n)'$ and positive definite covariance matrix Σ. Then we have for the conditional distribution*

$$X_1|_{\{X_2,\ldots,X_n\}} \sim \mathcal{N}\left(m_1 + \Sigma_{1,2}\Sigma_{2,2}^{-1}\left(X^{(2)} - m^{(2)}\right),\right.$$
$$\left.\Sigma_{1,1} - \Sigma_{1,2}\Sigma_{2,2}^{-1}\Sigma_{2,1}\right),$$

where $X^{(2)} = (X_2, \ldots, X_n)'$ is multivariate Gaussian with mean $m^{(2)} = (m_2, \ldots, m_n)'$ and positive definite covariance matrix $\Sigma_{2,2}$, $\Sigma_{1,1}$ is the variance of X_1 and $\Sigma_{1,2} = \Sigma'_{2,1}$ is the covariance vector between X_1 and $X^{(2)}$.

Lemma 2.1 gives the following proposition whose proof is provided in the Appendix.

Proposition 2.2. *Under Model Assumption 1.1 we obtain for $0 \leq j < j + l \leq J$*

$$\log I_{i,j+l}|_{\{\mathcal{B}^I_j,\mathbf{\Theta}\}}$$
$$\sim \mathcal{N}\left(\mu_{j+l} + \frac{v_{j+l}^2}{v_j^2}\left(\log I_{i,j} - \mu_j\right), v_{j+l}^2(1 - v_{j+l}^2/v_j^2)\right),$$

where the parameters are given by (an empty sum is set equal to 0)

$$\mu_j = \sum_{m=0}^{J}\Phi_m - \sum_{n=j}^{J-1}\Psi_n \quad \text{and} \quad v_j^2 = \sum_{m=0}^{J}\sigma_m^2 + \sum_{n=j}^{J-1}\tau_n^2.$$

Note that $\mu_J = \sum_{m=0}^{J}\Phi_m$ and $v_J^2 = \sum_{m=0}^{J}\sigma_m^2$.

Henceforth, we have the Markov property and we obtain the following corollary:

Corollary 2.3. *Under Model Assumption 1.1 we obtain for the expected ultimate loss $I_{i,J}$, given $\{\mathcal{B}^I_j, \mathbf{\Theta}\}$,*

$$E\left[I_{i,J}|\mathcal{B}^I_j, \mathbf{\Theta}\right] = I_{i,j}^{1-\alpha_j}\exp\left\{(1-\alpha_j)\sum_{l=j}^{J-1}\Psi_l + \alpha_j\left(\mu_J + v_J^2/2\right)\right\}$$
$$= I_{i,j}\exp\left\{\sum_{l=j}^{J-1}\Psi_l + \tau_l^2/2\right\}$$
$$\times \exp\left\{\alpha_j\left(\mu_j - \log I_{i,j} - \sum_{l=j}^{J-1}\tau_l^2/2\right)\right\},$$

with credibility weight

$$\alpha_j = 1 - \frac{v_J^2}{v_j^2} = \frac{1}{v_j^2}\sum_{l=j}^{J-1}\tau_l^2.$$

Remark. Compare the statement of Corollary 2.3 with formula (2.2). We see that under Model Assumption 1.1 cumulative payments $P_{i,j}$ fulfill the classical CL assumption (2.1) whereas incurred losses $I_{i,j}$ do in general not satisfy the CL assumption, given $\mathbf{\Theta}$. This is different from the MCL method where one assumes that both cumulative payments and incurred losses satisfy the CL assumption (see Quarg and Mack, 2004). At this stage one may even raise the question about interesting stochastic models such that cumulative payments $P_{i,j}$ and incurred losses $I_{i,j}$ simultaneously fulfill the CL assumption. Our Model 1.1 does not fall into that class. In our model, the classical CL factor gets a correction term

$$\exp\left\{\alpha_j\left(\mu_j - \log I_{i,j} - \sum_{l=j}^{J-1}\tau_l^2/2\right)\right\},$$

which adjusts the CL factor $\exp\left\{\sum_{l=j}^{J-1}\Psi_l + \tau_l^2/2\right\}$ to the actual claims experience with credibility weight α_j. The smaller the development year j the bigger is the credibility weight α_j. On the other hand, we could also rewrite the right-hand side of Corollary 2.3 as

$$E\left[I_{i,J}|\mathcal{B}^I_j, \mathbf{\Theta}\right] = \exp\left\{(1-\alpha_j)\left(\log I_{i,j} + \sum_{l=j}^{J-1}\Psi_l\right) + \alpha_j\sum_{m=0}^{J}\Phi_m\right\}$$
$$\times \exp\left\{\alpha_j v_J^2/2\right\},$$

the first factor on the right-hand side shows that we consider a credibility weighted average between incurred losses $\log I_{i,j} + \sum_{l=j}^{J-1}\Psi_l$ and cumulative payments $\mu_J = \sum_{m=0}^{J}\Phi_m$.

2.3. Cumulative payments and incurred losses

Finally, we would like to predict the ultimate loss $P_{i,J} = I_{i,J}$ when we jointly consider payments and incurred losses information $\mathcal{B}_j$. We therefore apply the full PIC model, given the model parameters $\mathbf{\Theta}$.

Theorem 2.4. *Under Model Assumption 1.1 we obtain for the ultimate loss $P_{i,J} = I_{i,J}$, given $\{\mathcal{B}_j, \mathbf{\Theta}\}$, $0 \leq j < J$,*

$$\log P_{i,J}|_{\{\mathcal{B}_j,\mathbf{\Theta}\}} \sim \mathcal{N}\left(\mu_J + (1-\beta_j)(\log P_{i,j} - \eta_j) + \beta_j(\log I_{i,j} - \mu_j),\right.$$
$$\left.(1-\beta_j)(v_J^2 - w_j^2)\right),$$

where the parameters are given by

$$\eta_j = \sum_{m=0}^{j}\Phi_m \quad \text{and} \quad w_j^2 = \sum_{m=0}^{j}\sigma_m^2,$$

and the credibility weight is given by

$$\beta_j = \frac{v_J^2 - w_j^2}{v_j^2 - w_j^2} > 0.$$

Remarks. • The conditional distribution of $\log P_{i,J}$, given $\{\mathcal{B}_j, \mathbf{\Theta}\}$, changes accordingly to the observations $\log P_{i,j}$ and $\log I_{i,j}$. Thereby is the prior expectation μ_J for the ultimate loss $P_{i,J} = I_{i,J}$ updated by a credibility weighted average between the paid residual $\log P_{i,j} - \eta_j$ and the incurred residual $\log I_{i,j} - \mu_j$, where the credibility weight is given by

$$\beta_j = \frac{\sum_{m=j+1}^{J} \sigma_m^2}{\sum_{m=j+1}^{J} \sigma_m^2 + \sum_{n=j}^{J-1} \tau_n^2}.$$

Analogously, the prior variance $v_J^2 - w_j^2$ is reduced by the credibility weight $1 - \beta_j$, this is typical in credibility theory, see for example Theorem 4.3 in Bühlmann and Gisler (2005).

• Theorem 2.4 shows that in our log-normal PIC model we can calculate analytically the posterior distribution of the ultimate loss, given $\mathcal{B}_j$ and conditional on $\mathbf{\Theta}$. Henceforth, we can calculate the conditionally expected ultimate loss, see Corollary 2.5.

Corollary 2.5 (*PIC Ultimate Loss Prediction*)**.** *Under Model Assumption* 1.1 *we obtain for the expected ultimate loss $I_{i,J} = P_{i,J}$, given $\{\mathcal{B}_j, \mathbf{\Theta}\}$,*

$$\begin{aligned}
E\left[P_{i,J}|\mathcal{B}_j, \mathbf{\Theta}\right] &= P_{i,j} \exp\left\{\sum_{l=j+1}^{J} \Phi_l + \frac{\sigma_l^2}{2}\right\} \\
&\quad \times \exp\left\{\beta_j \left(\log \frac{I_{i,j}}{P_{i,j}} - (\mu_j - \eta_j) - \sum_{l=j+1}^{J} \frac{\sigma_l^2}{2}\right)\right\} \\
&= I_{i,j} \exp\left\{\sum_{l=j}^{J-1} \Psi_l\right\} \\
&\quad \times \exp\left\{(1-\beta_j)\left(\log \frac{P_{i,j}}{I_{i,j}} - (\eta_j - \mu_j) + \sum_{l=j+1}^{J} \frac{\sigma_l^2}{2}\right)\right\} \\
&= \exp\left\{(1-\beta_j)\left(\log P_{i,j} + \sum_{l=j+1}^{J} \Phi_l\right) + \beta_j \left(\log I_{i,j} + \sum_{l=j}^{J-1} \Psi_l\right)\right\} \\
&\quad \times \exp\left\{(1-\beta_j)(v_J^2 - w_j^2)/2\right\}.
\end{aligned}$$

Remark. Henceforth, if we consider simultaneously claims payments and incurred losses information, we obtain a correction term

$$\exp\left\{\beta_j \left(\log \frac{I_{i,j}}{P_{i,j}} - (\mu_j - \eta_j) - \sum_{l=j+1}^{J} \frac{\sigma_l^2}{2}\right)\right\} \tag{2.3}$$

to the classical CL predictor $E\left[P_{i,J}|\mathcal{B}_j^P, \mathbf{\Theta}\right]$. This adjustment factor compares incurred–paid ratios and corresponds to the observed residuals $\log \frac{I_{i,j}}{P_{i,j}} - (\mu_j - \eta_j)$. For example, a large incurred–paid ratio $I_{i,j}/P_{i,j}$ gives a large correction term (2.3) to the classical CL predictor $E\left[P_{i,J}|\mathcal{B}_j^P, \mathbf{\Theta}\right]$. This is a similar mechanism as in the MCL method that also adjusts the predictors according to incurred–paid ratios (see Quarg and Mack, 2004). The last formula in the statement of Corollary 2.5 shows that we can also understand the PIC ultimate loss predictor as a credibility weighted average between claims payments and incurred losses information.

3. Parameter estimation

So far, all consideration were done for known parameters $\mathbf{\Theta}$. However, in general, they are not known and need to be estimated from the observations. Assume that we are at time J and that we have observations (see also Table 1)

$$\mathcal{D}_J^P = \left\{P_{i,j} : i+j \le J\right\}, \qquad \mathcal{D}_J^I = \left\{I_{i,j} : i+j \le J\right\} \quad \text{and}$$
$$\mathcal{D}_J = \mathcal{D}_J^P \cup \mathcal{D}_J^I.$$

We estimate the parameters in a Bayesian framework. Therefore we define the following model:

Model Assumption 3.1 (*Bayesian PIC Model*)**.** Assume Model Assumption 1.1 hold true with deterministic $\sigma_0, \ldots, \sigma_J$ and $\tau_0, \ldots, \tau_{J-1}$ and

$$\Phi_m \sim \mathcal{N}\left(\phi_m, s_m^2\right) \quad \text{for } \in \{0, \ldots, J\},$$
$$\Psi_n \sim \mathcal{N}\left(\psi_n, t_n^2\right) \quad \text{for } n \in \{0, \ldots, J-1\}. \quad \square$$

In a full Bayesian approach one chooses an appropriate prior distribution for the whole parameter vector $\mathbf{\Theta}$. We will only use a prior distribution for Φ_m and Ψ_n and assume that σ_m and τ_n are known. This has the advantage that we can analytically calculate the posterior distributions that will allow for explicit model calculations and interpretations.

3.1. Cumulative payments

For claims payments we only need the parameters $\mathbf{\Phi} = (\Phi_0, \ldots, \Phi_J)$. The posterior density of $\mathbf{\Phi}$, given $\mathcal{D}_J^P$, is given by (we set $P_{i,-1} = 1$)

$$\begin{aligned}
u\left(\mathbf{\Phi}|\mathcal{D}_J^P\right) \propto \prod_{j=0}^{J}\prod_{i=0}^{J-j} \exp\left\{-\frac{1}{2\sigma_j^2}\left(\Phi_j - \log\frac{P_{i,j}}{P_{i,j-1}}\right)^2\right\} \\
\times \prod_{j=0}^{J} \exp\left\{-\frac{1}{2s_j^2}\left(\Phi_j - \phi_j\right)^2\right\}.
\end{aligned}$$

This immediately provides the next theorem:

Theorem 3.2. *Under Model Assumption 3.1 the posterior distribution of $\mathbf{\Phi}$, given $\mathcal{D}_J^P$, has independent components with*

$$\Phi_j|_{\{\mathcal{D}_J^P\}} \sim \mathcal{N}\left(\phi_j^{P,\mathrm{post}} = \gamma_j^P \frac{1}{\sharp(j)} \sum_{i=0}^{J-j} \log\frac{P_{i,j}}{P_{i,j-1}} + (1-\gamma_j^P)\phi_j, \right.$$
$$\left. (s_j^{P,\mathrm{post}})^2 = \left(\frac{1}{s_j^2} + \frac{\sharp(j)}{\sigma_j^2}\right)^{-1}\right),$$

with $\sharp(j) = J - j + 1$ and credibility weight

$$\gamma_j^P = \frac{\sharp(j)}{\sharp(j) + \sigma_j^2/s_j^2}.$$

Henceforth, the posterior mean is a credibility weighted average between the prior mean ϕ_j and the empirical mean

$$\overline{\phi}_j = \frac{1}{\sharp(j)} \sum_{i=0}^{J-j} \log\frac{P_{i,j}}{P_{i,j-1}},$$

see also formula (5.2) in Wüthrich and Merz (2008). The posterior distribution of $P_{i,J}$, given $\mathcal{D}_J^P$, is now completely determined. Moments can be calculated in closed form and Monte Carlo simulation provides the empirical posterior distribution of the ultimate losses vector $(P_{1,J}, \ldots, P_{J,J})|_{\{\mathcal{D}_J^P\}} = (I_{1,J}, \ldots, I_{J,J})|_{\{\mathcal{D}_J^P\}}$. In view of (2.2) and Theorem 3.2 the ultimate loss predictor, given

$\mathcal{D}_J^P$, is given by

$$E\left[P_{i,J}|\mathcal{D}_J^P\right] = P_{i,J-i}\prod_{l=J-i+1}^{J}\exp\left\{\phi_l^{P,\text{post}}+\sigma_l^2/2+(s_l^{P,\text{post}})^2/2\right\}. \tag{3.1}$$

3.2. Incurred losses

In this subsection we concentrate on the parameter estimation given the data $\mathcal{D}_J^I$ of incurred losses. We define the underlying parameter for $I_{i,J}$ by (see also (1.1))

$$\Psi_J = -\mu_J = -\sum_{j=0}^{J}\Phi_j \sim \mathcal{N}\left(\psi_J = -\sum_{j=0}^{J}\phi_j, t_J^2 = \sum_{j=0}^{J}s_j^2\right),$$

which is independent from $\Psi_0,\ldots,\Psi_{J-1}$. For incurred losses we then only need the parameters $\boldsymbol{\Psi}=(\Psi_0,\ldots,\Psi_J)$. The posterior density of $\boldsymbol{\Psi}$, given $\mathcal{D}_J^I$, is given by

$$u\left(\boldsymbol{\Psi}|\mathcal{D}_J^I\right) \propto \prod_{i=0}^{J}\exp\left\{-\frac{1}{2v_{J-i}^2}\left(\sum_{n=J-i}^{J}\Psi_n+\log I_{i,J-i}\right)^2\right\} \times \prod_{j=0}^{J-1}\prod_{i=0}^{J-j-1}\exp\left\{-\frac{1}{2\tau_j^2}\left(\Psi_j+\log\frac{I_{i,j}}{I_{i,j+1}}\right)^2\right\} \times \prod_{j=0}^{J}\exp\left\{-\frac{1}{2t_j^2}\left(\Psi_j-\psi_j\right)^2\right\}. \tag{3.2}$$

This immediately provides the next theorem:

Theorem 3.3. *Under Model Assumption 3.1 the posterior distribution of $\boldsymbol{\Psi}$, given $\mathcal{D}_J^I$, is a multivariate Gaussian distribution with posterior mean $\boldsymbol{\psi}^{\text{post}}(\mathcal{D}_J^I)$ and posterior covariance matrix $\Sigma(\mathcal{D}_J^I)$. The inverse covariance matrix $\Sigma(\mathcal{D}_J^I)^{-1}=(a_{n,m}^I)_{0\le n,m\le J}$ is given by*

$$a_{n,m}^I = \left(t_n^{-2}+(J-n)\tau_n^{-2}\right)1_{\{n=m\}}+\sum_{i=0}^{n\wedge m}v_i^{-2} \quad \text{for } 0\le n,m\le J.$$

The posterior mean $\boldsymbol{\psi}^{\text{post}}(\mathcal{D}_J^I)=(\psi_0^{I,\text{post}},\ldots,\psi_J^{I,\text{post}})'$ is obtained by

$$\boldsymbol{\psi}^{\text{post}}(\mathcal{D}_J^I)=\Sigma(\mathcal{D}_J^I)(b_0^I,\ldots,b_J^I)',$$

with vector $(b_0^I,\ldots,b_J^I)$ given by

$$b_j^I = t_j^{-2}\psi_j-\tau_j^{-2}\sum_{i=0}^{J-j-1}\log\frac{I_{i,j}}{I_{i,j+1}}-\sum_{i=0}^{j}v_i^{-2}\log I_{J-i,i}.$$

Observe that b_j^I can be rewritten so that it involves a credibility weighted average between prior mean Ψ_j and the incurred losses observations, namely

$$b_j^I = \left(t_j^{-2}+(\sharp(j)-1)\tau_j^{-2}\right)\left[\gamma_j^I\overline{\psi}_j+(1-\gamma_j^I)\psi_j\right] - \sum_{i=0}^{j}v_i^{-2}\log I_{J-i,i}, \tag{3.3}$$

with credibility weight

$$\gamma_j^I = \frac{\sharp(j)-1}{\sharp(j)-1+\tau_j^2/t_j^2},$$

and empirical mean

$$\overline{\psi}_j = -\frac{1}{\sharp(j)-1}\sum_{i=0}^{J-j-1}\log\frac{I_{i,j}}{I_{i,j+1}}.$$

The posterior distribution of $P_{i,J}=I_{i,J}$, given $\mathcal{D}_J^I$, is now completely determined. We obtain for the ultimate loss predictor, given $\mathcal{D}_J^I$,

$$E\left[P_{i,J}|\mathcal{D}_J^I\right] = I_{i,J-i}^{1-\alpha_{J-i}}\exp\left\{(1-\alpha_{J-i})\sum_{l=J-i}^{J-1}\psi_l^{I,\text{post}} + \alpha_{J-i}\left(-\psi_J^{I,\text{post}}+\frac{v_J^2}{2}\right)+\left(s_i^{I,\text{post}}\right)^2/2\right\}, \tag{3.4}$$

where

$$\left(s_i^{I,\text{post}}\right)^2 = (\mathbf{e}_i^I)'\Sigma(\mathcal{D}_J^I)\mathbf{e}_i^I,$$

with $\mathbf{e}_i^I=(0,\ldots,0,1-\alpha_{J-i},\ldots,1-\alpha_{J-i},-\alpha_{J-i})'\in\mathbb{R}^{J+1}$.

3.3. Cumulative payments and incurred losses

Similar to the last section we determine the posterior distribution of $\boldsymbol{\Theta}$, given $\mathcal{D}_J$. Observe that

$$\log I_{i,j}|_{\{\mathcal{B}_j^P,\boldsymbol{\Theta}\}} \sim \mathcal{N}\left(\mu_j-\eta_j, v_j^2-w_j^2\right) = \mathcal{N}\left(\sum_{m=j+1}^{J}\Phi_m-\sum_{n=j}^{J-1}\Psi_n, \sum_{m=j+1}^{J}\sigma_m^2+\sum_{n=j}^{J-1}\tau_n^2\right).$$

This implies that the joint likelihood function of the data $\mathcal{D}_J$ is given by (we set $P_{i,-1}=1$)

$$l_{\mathcal{D}_J}(\boldsymbol{\Theta}) = \prod_{j=0}^{J}\prod_{i=0}^{J-j}\frac{1}{\sqrt{2\pi}\sigma_jP_{i,j}}\exp\left\{-\frac{1}{2\sigma_j^2}\left(\Phi_j-\log\frac{P_{i,j}}{P_{i,j-1}}\right)^2\right\} \times \prod_{i=1}^{J}\frac{1}{\sqrt{2\pi(v_{J-i}^2-w_{J-i}^2)}I_{i,J-i}} \times \exp\left\{-\frac{1}{2(v_{J-i}^2-w_{J-i}^2)}\left(\mu_{J-i}-\eta_{J-i}-\log\frac{I_{i,J-i}}{P_{i,J-i}}\right)^2\right\} \times \prod_{j=0}^{J-1}\prod_{i=0}^{J-j-1}\frac{1}{\sqrt{2\pi}\tau_jI_{i,j}}\exp\left\{-\frac{1}{2\tau_j^2}\left(\Psi_j+\log\frac{I_{i,j}}{I_{i,j+1}}\right)^2\right\}. \tag{3.5}$$

Under Model Assumption 3.1 the posterior distribution $u\left(\boldsymbol{\Theta}|\mathcal{D}_J\right)$ of $\boldsymbol{\Theta}$, given $\mathcal{D}_J$, is given by

$$u\left(\boldsymbol{\Theta}|\mathcal{D}_J\right) \propto l_{\mathcal{D}_J}(\boldsymbol{\Theta})\prod_{m=0}^{J}\exp\left\{-\frac{1}{2s_m^2}\left(\Phi_m-\phi_m\right)^2\right\} \times \prod_{n=0}^{J-1}\exp\left\{-\frac{1}{2t_n^2}\left(\psi_n-\Psi_n\right)^2\right\}. \tag{3.6}$$

This immediately implies the following theorem:

Theorem 3.4. *Under Model Assumption 3.1, the posterior distribution $u\left(\boldsymbol{\Theta}|\mathcal{D}_J\right)$ is a multivariate Gaussian distribution with posterior mean $\theta^{\text{post}}(\mathcal{D}_J)$ and posterior covariance matrix $\Sigma(\mathcal{D}_J)$. The inverse covariance matrix $\Sigma(\mathcal{D}_J)^{-1}=(a_{n,m})_{0\le n,m\le 2J}$ is given by*

$$a_{n,m} = \left(s_n^{-2}+(J-n+1)\sigma_n^{-2}\right)1_{\{n=m\}}+\sum_{i=0}^{(n-1)\wedge(m-1)}\left(v_i^2-w_i^2\right)^{-1} \quad \text{for } 0\le n,m\le J,$$

$$a_{J+1+n,J+1+m} = \left(t_n^{-2}+(J-n)\tau_n^{-2}\right)1_{\{n=m\}}+\sum_{i=0}^{n\wedge m}\left(v_i^2-w_i^2\right)^{-1} \quad \text{for } 0\le n,m\le J-1,$$

$$a_{n,J+1+m}=a_{J+1+m,n} = -\sum_{i=0}^{(n-1)\wedge m}\left(v_i^2-w_i^2\right)^{-1} \quad \text{for } 0\le n\le J,\ 0\le m\le J-1.$$

The posterior mean $\theta^{\text{post}}(\mathcal{D}_J) = \left(\phi_0^{\text{post}}, \ldots, \phi_J^{\text{post}}, \psi_0^{\text{post}}, \ldots, \psi_{J-1}^{\text{post}}\right)'$ *is obtained by*

$$\theta^{\text{post}}(\mathcal{D}_J) = \Sigma(\mathcal{D}_J)(c_0, \ldots, c_J, b_0, \ldots, b_{J-1})',$$

with vector $(c_0, \ldots, c_J, b_0, \ldots, b_{J-1})$ *given by*

$$c_j = s_j^{-2}\phi_j + \sigma_j^{-2}\sum_{i=0}^{J-j}\log\frac{P_{i,j}}{P_{i,j-1}} + \sum_{i=J-j+1}^{J}\left(v_{J-i}^2 - w_{J-i}^2\right)^{-1}\log\frac{I_{i,J-i}}{P_{i,J-i}},$$

$$b_j = t_j^{-2}\psi_j - \tau_j^{-2}\sum_{i=0}^{J-j-1}\log\frac{I_{i,j}}{I_{i,j+1}} - \sum_{i=J-j}^{J}\left(v_{J-i}^2 - w_{J-i}^2\right)^{-1}\log\frac{I_{i,J-i}}{P_{i,J-i}}.$$

Henceforth, this implies for the expected ultimate loss in the Bayesian PIC model, given $\mathcal{D}_J$, (see also Corollary 2.5)

$$\begin{aligned} &E\left[P_{i,J}|\mathcal{D}_J\right] \\ &= P_{i,J-i}^{1-\beta_{J-i}} I_{i,J-i}^{\beta_{J-i}} \exp\left\{(1-\beta_{J-i})\sum_{l=J-i+1}^{J}\phi_l^{\text{post}} + \beta_{J-i}\sum_{l=J-i}^{J-1}\psi_l^{\text{post}}\right\} \\ &\quad\times \exp\left\{(1-\beta_{J-i})\frac{v_J^2 - w_{J-i}^2}{2} + \left(s_i^{\text{post}}\right)^2/2\right\}, \end{aligned} \tag{3.7}$$

where

$$\left(s_i^{\text{post}}\right)^2 = (\mathbf{e}_i)'\Sigma(\mathcal{D}_J)\mathbf{e}_i,$$

with $\mathbf{e}_i = (0, \ldots, 0, 1-\beta_{J-i}, \ldots, 1-\beta_{J-i}, 0, \ldots, 0, \beta_{J-i}, \ldots, \beta_{J-i})' \in \mathbb{R}^{2J+1}$.

4. Prediction uncertainty

The ultimate loss $P_{i,J} = I_{i,J}$ is now predicted by its conditional expectations

$$E\left[P_{i,J}|\mathcal{D}_J^P\right], \qquad E\left[P_{i,J}|\mathcal{D}_J^I\right] \quad\text{or}\quad E\left[P_{i,J}|\mathcal{D}_J\right],$$

depending on the available information $\mathcal{D}_J^P$, $\mathcal{D}_J^I$ or $\mathcal{D}_J$ (see (3.1), (3.4) and (3.7)). With Theorems 3.2–3.4 all posterior distributions in the Bayesian PIC Model 3.1 are given analytically. Therefore any risk measure for the prediction uncertainty can be calculated with a simple Monte Carlo simulation approach. Here, we consider the conditional mean square error of prediction (MSEP) as risk measure. The conditional MSEP is probably the most popular risk measure in claims reserving practice and has the advantage that it is analytically tractable in our context. We derive it for the posterior distribution, given $\mathcal{D}_J$. The cases $\mathcal{D}_J^P$ and $\mathcal{D}_J^I$ are completely analogous. The conditional MSEP is given by

$$\begin{aligned} &\operatorname{msep}_{\sum_{i=1}^J P_{i,J}|\mathcal{D}_J}\left(E\left[\left.\sum_{i=1}^J P_{i,J}\right|\mathcal{D}_J\right]\right) \\ &= E\left[\left.\left(\sum_{i=1}^J P_{i,J} - E\left[\left.\sum_{i=1}^J P_{i,J}\right|\mathcal{D}_J\right]\right)^2\right|\mathcal{D}_J\right] \\ &= \operatorname{Var}\left(\left.\sum_{i=1}^J P_{i,J}\right|\mathcal{D}_J\right), \end{aligned}$$

see Wüthrich and Merz (2008), Section 3.1. For the conditional MSEP, given the observations $\mathcal{D}_J$, we obtain:

Theorem 4.1. *Under Model Assumption 3.1 we have, using information* $\mathcal{D}_J$,

$$\begin{aligned} &\operatorname{msep}_{\sum_{i=1}^J P_{i,J}|\mathcal{D}_J}\left(E\left[\left.\sum_{i=1}^J P_{i,J}\right|\mathcal{D}_J\right]\right) \\ &= \sum_{1\le i,k\le J}\left(e^{(1-\beta_{J-i})(v_J^2 - w_{J-i}^2)1_{\{i=k\}} + \mathbf{e}_i'\Sigma(\mathcal{D}_J)\,\mathbf{e}_k} - 1\right) \\ &\quad\times E\left[P_{i,J}|\mathcal{D}_J\right]E\left[P_{k,J}|\mathcal{D}_J\right] \end{aligned}$$

with $E\left[P_{i,J}|\mathcal{D}_J\right]$ *given by (3.7).*

Similarly we obtain for the conditional MSEP w.r.t. $\mathcal{D}_J^I$ and $\mathcal{D}_J^P$, respectively:

Theorem 4.2. *Under Model Assumption 3.1 we have, using information* $\mathcal{D}_J^I$,

$$\begin{aligned} &\operatorname{msep}_{\sum_{i=1}^J P_{i,J}|\mathcal{D}_J^I}\left(E\left[\left.\sum_{i=1}^J P_{i,J}\right|\mathcal{D}_J^I\right]\right) \\ &= \sum_{1\le i,k\le J}\left(e^{\alpha_{J-i}v_J^2 1_{\{i=k\}} + (\mathbf{e}_i^I)'\Sigma(\mathcal{D}_J^I)\mathbf{e}_k^I} - 1\right) \\ &\quad\times E\left[P_{i,J}|\mathcal{D}_J^I\right]E\left[P_{k,J}|\mathcal{D}_J^I\right] \end{aligned}$$

with $E\left[P_{i,J}|\mathcal{D}_J^I\right]$ *given by (3.4). Using information* $\mathcal{D}_J^P$ *we obtain*

$$\begin{aligned} &\operatorname{msep}_{\sum_{i=1}^J P_{i,J}|\mathcal{D}_J^P}\left(E\left[\left.\sum_{i=1}^J P_{i,J}\right|\mathcal{D}_J^P\right]\right) \\ &= \sum_{1\le i,k\le J}\left(e^{(v_J^2 - w_{J-i}^2)1_{\{i=k\}} + (\mathbf{e}_i^P)'\Sigma(\mathcal{D}_J^P)\mathbf{e}_k^P} - 1\right) \\ &\quad\times E\left[P_{i,J}|\mathcal{D}_J^P\right]E\left[P_{k,J}|\mathcal{D}_J^P\right] \end{aligned}$$

with $E\left[P_{i,J}|\mathcal{D}_J^P\right]$ *given by (3.1),* $\mathbf{e}_i^P = (0, \ldots, 0, 1, \ldots, 1)' \in \mathbb{R}^{J+1}$ *and* $\Sigma(\mathcal{D}_J^P) = \operatorname{diag}((s_j^{P,\text{post}})^2)$.

5. Example

We revisit the first example given in Dahms (2008) and Dahms et al. (2009) (see Tables 10 and 11). We do a first analysis of the data under Model Assumption 3.1 where we assume that σ_j and τ_j are deterministic parameters (using plug-in estimates). In a second analysis we also model these parameters in a Bayesian framework.

5.1. Data analysis in the Bayesian PIC Model 3.1

Because we do not have prior parameter information and because we would like to compare our results to Dahms' (2008) results, we choose non-informative priors for Φ_m and Ψ_n. This means that we let $s_m^2 \to \infty$ and $t_n^2 \to \infty$. This then implies for the credibility weights $\gamma_m^P = \gamma_m^I = 1$ which means that our claims payments prediction is based on $\mathcal{D}_J^P$ only (see Theorem 3.2) and our incurred losses prediction is based on $\mathcal{D}_J^I$ only (see (3.3)), i.e. no prior knowledge ϕ_m and ψ_n is used. Similarly for the joint PIC prediction no prior knowledge is needed for non-informative priors, because then the prior values ϕ_m and ψ_n disappear in c_m and b_n for $s_m^2 \to \infty$ and $t_n^2 \to \infty$, see Theorem 3.4.

Henceforth, there remains to provide the values for σ_j and τ_j. For the choice of these parameters we choose in this subsection an empirical Bayesian point of view and use the standard plug-in estimators (sample standard deviation estimator, see e.g. (5.3) in Wüthrich and Merz, 2008). Since the last variance parameters cannot be estimated from the data (due to the lack of observations)

Table 2
Standard deviation parameter choices for σ_j and τ_j.

j	0	1	2	3	4	5	6	7	8	9
σ_j	0.1393	0.0650	0.0731	0.0640	0.0264	0.0271	0.0405	0.0227	0.0494	0.0227
τ_j	0.0633	0.0459	0.0415	0.0122	0.0083	0.0017	0.0019	0.0011	0.0006	

Table 3
Claims reserves in the Bayesian PIC Model 3.1.

	$\widehat{R}(\mathcal{D}_J^P) = E\left[P_{i,J}\|\mathcal{D}_J^P\right] - P_{i,J-i}$	$\widehat{R}(\mathcal{D}_J^I) = E\left[P_{i,J}\|\mathcal{D}_J^I\right] - P_{i,J-i}$	$\widehat{R}(\mathcal{D}_J) = E\left[P_{i,J}\|\mathcal{D}_J\right] - P_{i,J-i}$
1	115,470	337,994	337,799
2	428,272	31,526	31,686
3	642,664	331,526	331,890
4	729,344	1,018,924	1,018,308
5	1,284,545	1,102,580	1,104,816
6	1,183,781	1,869,284	1,842,669
7	1,692,632	1,990,260	1,953,767
8	2,407,438	1,465,661	1,602,229
9	2,027,245	2,548,242	2,402,946
Total	10,511,390	10,695,996	10,626,108

Table 4
Claims reserves from the CL method for claims payments and incurred losses (see Mack, 1993), from the CLRM (see Dahms, 2008), and from the MCL method for claims payments and incurred losses (see Quarg and Mack, 2004).

	CL paid Mack (1993)	CL incurred Mack (1993)	CLRM Dahms (2008)	MCL paid Quarg and Mack (2004)	MCL incurred Quarg and Mack (2004)
1	114,086	337,984	314,902	104,606	338,200
2	394,121	31,884	66,994	457,484	30,850
3	608,749	331,436	359,384	664,871	330,205
4	697,742	1,018,350	981,883	615,436	1,021,361
5	1,234,157	1,103,928	1,115,768	1,271,110	1,102,396
6	1,138,623	1,868,664	1,786,947	919,102	1,894,861
7	1,638,793	1,997,651	1,942,518	1,498,163	2,020,310
8	2,359,939	1,418,779	1,569,657	3,181,319	1,320,492
9	1,979,401	2,556,612	2,590,718	1,602,089	2,703,242
Total	10,165,612	10,665,287	10,728,771	10,314,181	10,761,918

we use the extrapolation used in Mack (1993). This gives the parameter choices provided in Table 2.

The expected outstanding loss liabilities then provide the PIC claims reserves:

$$\widehat{R}(\mathcal{D}_J) = E\left[P_{i,J}|\mathcal{D}_J\right] - P_{i,J-i},$$

if the ultimate loss prediction is based on the whole information $\mathcal{D}_J$ (and similarly for $\mathcal{D}_J^P$ and $\mathcal{D}_J^I$, respectively). This gives the claims reserves provided in Table 3.

We observe that in all accident years the PIC reserves $\widehat{R}(\mathcal{D}_J)$ based on the whole information $\mathcal{D}_J$ are between the estimates based on $\mathcal{D}_J^P$ and $\mathcal{D}_J^I$. The deviations from $\widehat{R}(\mathcal{D}_J^I)$ are comparably small which comes from the fact that $\tau_j \ll \sigma_j$.

In Table 4 we provide the claims reserves estimates for other popular methods. We observe that the reserves $\widehat{R}(\mathcal{D}_J^P)$ are close to the ones from CL paid (the differences can partly be explained by the variance terms $\sigma_l^2/2$ in the expected value of log-normal distributions, see (2.2)). The reserves $\widehat{R}(\mathcal{D}_J^I)$ are very close to the ones from CL incurred. The PIC reserves $\widehat{R}(\mathcal{D}_J)$ from the combined information look similar to the ones from the CLRM and to MCL incurred method. We also mention that for our data set the MCL does not really reduce the gap between the two predictions. This was also already previously observed for other data sets.

In Table 5 we compare the corresponding prediction uncertainties measured by the square root of the conditional MSEP. Under Model Assumption 3.1 these are calculated analytically with Theorems 3.2–3.4, 4.1 and 4.2. First of all, we observe that in our model the PIC predictor $\widehat{R}(\mathcal{D}_J)$ has a smaller prediction uncertainty compared to $\widehat{R}(\mathcal{D}_J^P)$ and $\widehat{R}(\mathcal{D}_J^I)$. This is clear because increasing the set of information reduces the uncertainty. Therefore, the PIC predictor $\widehat{R}(\mathcal{D}_J)$ should be preferred within Model Assumption 3.1.

Table 5
Total claims reserves and prediction uncertainty.

	Claims reserves	$\text{msep}^{1/2}$
$\widehat{R}(\mathcal{D}_J^P)$	10,511,390	1,559,228
$\widehat{R}(\mathcal{D}_J^I)$	10,695,996	421,298
$\widehat{R}(\mathcal{D}_J)$	10,626,108	389,496
CL paid	10,165,612	1,517,480
CL incurred	10,665,287	455,794
CLRM paid	10,728,771	467,814
CLRM incurred	10,728,771	471,873
MCL paid	10,314,181	Not available
MCL incurred	10,761,918	Not available

Furthermore, our prediction uncertainties are comparable to the ones from the other models. We would also like to mention that in the CLRM there are two values for the prediction uncertainty due to the fact that one can use different parameter estimators in the CLRM, see Corollary 4.4 in Dahms (2008) (the claims reserves coincide).

As mentioned in Section 4 we cannot only calculate the conditional MSEP, but Theorem 3.4 allows for a Monte Carlo simulation approach to the full predictive distribution of the outstanding loss liabilities in the Bayesian PIC Model 3.1. This is done as follows: Firstly, we generate multivariate Gaussian samples $\mathbf{\Theta}^{(t)}$ with mean $\theta^{\text{post}}(\mathcal{D}_J)$ and covariance matrix $\Sigma(\mathcal{D}_J)$ according to Theorem 3.4. Secondly, we generate samples $(I_{1,J}^{(t)}, \ldots, I_{J,J}^{(t)})|_{\{\mathbf{\Theta}^{(t)}\}}$ according to Theorem 2.4. In Table 6 we provide the empirical density for the outstanding loss liabilities from 20,000 simulations. Moreover, we compare it to the Gaussian density with the same mean and standard deviation (see also Table 5, line $\widehat{R}(\mathcal{D}_J)$). We observe

Table 6
Empirical density from 20,000 simulations and a comparison to the Gaussian density.

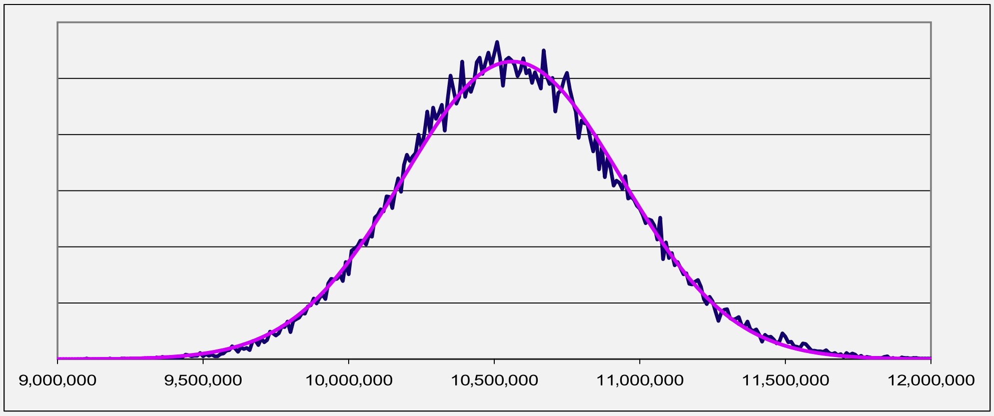
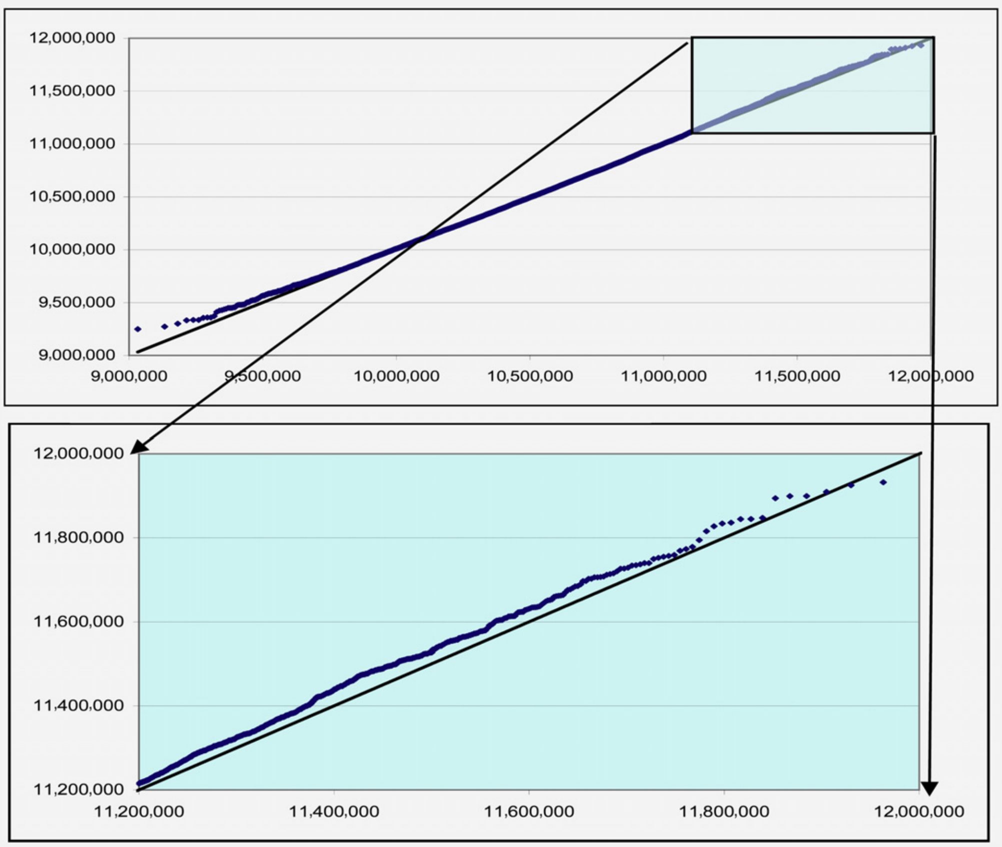


Table 7
Q–Q-plot from 20,000 simulations of the outstanding loss liabilities with the Gaussian distribution. Upper panel: full Q–Q-Plot; Lower panel: Q–Q-Plot of the upper 90% quantile.



that these densities are very similar, the Gaussian density having slightly more probability mass in the lower tail and slightly less probability mass in the upper tail (less heavy tailed). To further investigate this issue we provide the Q–Q-Plot in Table 7. We only observe differences in the tails (as described above). The lower panel in Table 7 gives the upper tail for values above the 90% quantile. There we see that a Gaussian approximation underestimates the true risks. However, the differences are comparably small.

In Table 3 we have observed that the resulting PIC reserves $\widehat{R}(\mathcal{D}_J)$ are close to the claims reserves $\widehat{R}(\mathcal{D}_J^I)$ from incurred losses information only. The reason therefore is that the standard deviation parameters τ_j for incurred losses are rather small. We now calculate a second example where we double these standard deviation parameters, i.e. $\tau_j^* = 2\tau_j$. The results are presented in Table 8. Firstly, we observe that the conditional MSEP using information $\mathcal{D}_J^I$ and $\mathcal{D}_J$ strongly increases. This is clear, because doubling the standard deviation parameters increases the uncertainty. More interestingly, we observe that the PIC reserves for each accident year $i = 2, \dots, J$ are now closer to the claims reserves from cumulative payments (especially for young accident years). The reason for this is that increasing the τ_j parameters means that we give less credibility weight to the incurred losses observations $\mathcal{D}_J^I$ and more credibility weight to the claims payments observations $\mathcal{D}_J^P$.

Table 8
Total claims reserves and prediction uncertainty for τ_j and $\tau_j^* = 2\tau_j$.

	res. $\widehat{R}(\mathcal{D}_J^P)$ τ_j	res. $\widehat{R}(\mathcal{D}_J^I)$ τ_j	PIC res. $\widehat{R}(\mathcal{D}_J)$ τ_j	res. $\widehat{R}(\mathcal{D}_J^P)$ $\tau_j^* = 2\tau_j$	res. $\widehat{R}(\mathcal{D}_J^I)$ $\tau_j^* = 2\tau_j$	PIC res. $\widehat{R}(\mathcal{D}_J)$ $\tau_j^* = 2\tau_j$
1	115,470	337,994	337,799	115,470	338,025	337,246
2	428,272	31,526	31,686	428,272	31,574	32,212
3	642,664	331,526	331,890	642,664	331,580	333,028
4	729,344	1,018,924	1,018,308	729,344	1,019,091	1,016,637
5	1,284,545	1,102,580	1,104,816	1,284,545	1,101,948	1,110,585
6	1,183,781	1,869,284	1,842,669	1,183,781	1,869,904	1,774,059
7	1,692,632	1,990,260	1,953,767	1,692,632	1,981,419	1,882,341
8	2,407,438	1,465,661	1,602,229	2,407,438	1,581,122	1,903,155
9	2,027,245	2,548,242	2,402,946	2,027,245	2,549,115	2,242,048
Total	10,511,390	10,695,996	10,626,108	10,511,390	10,803,778	10,631,310
$\text{msep}^{1/2}$	1,559,228	421,298	389,496	1,559,228	741,829	614,453

Table 9
Total claims reserves and prediction uncertainty in the Full Bayesian PIC model for different coefficients of variations for σ_j and τ_j.

$\text{Vco}(\sigma_j) = \gamma_{\sigma_j}^{-1/2} = \text{Vco}(\tau_j) = \gamma_{\tau_j}^{-1/2} =$	0%	10%	100%
Claims reserves $\widehat{R}(\mathcal{D}_J)$	10,626,108	10,589,180	10,701,455
Prediction uncertainty $\text{msep}^{1/2}$	389,496	392,832	472,449

Finally, in Table 12 we provide the posterior correlation matrix of $\boldsymbol{\Theta} = (\Phi_0, \ldots, \Phi_9, \Psi_0, \ldots, \Psi_8)$, given $\mathcal{D}_J$, which can directly be calculated from the posterior covariance matrix $\Sigma(\mathcal{D}_J)$ (see Theorem 3.4). We observe only small posterior correlations which may justify the assumption of prior uncorrelatedness.

5.2. Full Bayesian PIC model

A full Bayesian approach suggests that one also models the standard deviation parameters σ_j and τ_j stochastically. We therefore modify Model Assumption 3.1 as follows: Assume that σ_j and τ_j have independent gamma distributions with

$$\sigma_j \sim \Gamma\left(\gamma_{\sigma_j}, c_{\sigma_j}\right) \quad \text{and} \quad \tau_j \sim \Gamma\left(\gamma_{\tau_j}, c_{\tau_j}\right).$$

Of course, the choice of gamma distributions for the standard deviation parameters is rather arbitrary and any other positive distribution would also fit. Then the posterior distribution of the parameters $\boldsymbol{\Theta} = \left(\Phi_0, \ldots, \Phi_J, \Psi_0, \ldots, \Psi_{J-1}, \sigma_0, \ldots, \sigma_J, \tau_0, \ldots, \tau_{J-1}\right)$ is given by

$$u\left(\boldsymbol{\Theta}|\mathcal{D}_J\right) \propto l_{\mathcal{D}_J}(\boldsymbol{\Theta}) \prod_{m=0}^{J} \exp\left\{-\frac{1}{2s_m^2}\left(\Phi_m - \phi_m\right)^2\right\} \times \prod_{n=0}^{J-1} \exp\left\{-\frac{1}{2t_n^2}\left(\psi_n - \Psi_n\right)^2\right\} \times \prod_{m=0}^{J} \sigma_j^{\gamma_{\sigma_j}-1} \exp\left\{-c_{\sigma_j}\sigma_j\right\} \prod_{n=0}^{J-1} \tau_j^{\gamma_{\tau_j}-1} \exp\left\{-c_{\tau_j}\tau_j\right\}, \tag{5.1}$$

with $l_{\mathcal{D}_J}(\boldsymbol{\Theta})$ given in (3.5). This distribution can no longer be handled analytically because the normalizing constant takes a non-trivial form. But because we can write down its likelihood function up to the normalizing constant, we can still apply the Markov chain Monte Carlo (MCMC) simulation methodology. MCMC methods are very popular for this kind of problems. For an introduction and overview to MCMC methods we refer to Gilks et al. (1996), Asmussen and Glynn (2007) and Scollnik (2001). Because MCMC methods are widely used we refrain from describing them in detail. We will use the Metropolis–Hastings algorithm as described in Section 4.4 in Wüthrich and Merz (2008). The aim is to construct a Markov chain $(\boldsymbol{\Theta}_t)_{t\geq 0}$ whose stationary distribution is $u\left(\boldsymbol{\Theta}|\,\mathcal{D}_J\right)$. Then, we run this Markov chain for sufficiently long, so that we obtain approximate samples $\boldsymbol{\Theta}_t$'s from that stationary distribution. This is achieved by defining an acceptance probability

$$\alpha\left(\boldsymbol{\Theta}_t, \boldsymbol{\Theta}^*\right) = \min\left\{1, \frac{u\left(\boldsymbol{\Theta}^*|\mathcal{D}_J\right) q\left(\boldsymbol{\Theta}_t|\boldsymbol{\Theta}^*\right)}{u\left(\boldsymbol{\Theta}_t|\mathcal{D}_J\right) q\left(\boldsymbol{\Theta}^*|\boldsymbol{\Theta}_t\right)}\right\},$$

for the next step in the Markov chain, i.e. the move from $\boldsymbol{\Theta}_t$ to $\boldsymbol{\Theta}_{t+1}$. Thereby, the proposal distribution $q(\cdot|\cdot)$ is chosen in such a way that we obtain an average acceptance rate of roughly 24% because this satisfies certain optimal mixing properties for Markov chains (see Roberts et al., 1997, Corollary 1.2).

We apply this algorithm to different coefficients of variations $\gamma_{\sigma_j}^{-1/2}$ and $\gamma_{\tau_j}^{-1/2}$ of σ_j and τ_j, respectively. Moreover, we keep the means $E[\sigma_j] = \gamma_{\sigma_j}/c_{\sigma_j}$ and $E[\tau_j] = \gamma_{\tau_j}/c_{\tau_j}$ fixed and choose them equal to the deterministic values provided in Table 2. The results are provided in Table 9. We observe, as expected, an increase of the prediction uncertainty. The increase from Model 3.1 with deterministic σ_j's and τ_j's to a coefficient of variation of 10% is moderate, but it starts to increase strongly for larger coefficients of variations.

6. Conclusions

We have defined a stochastic PIC model that simultaneously considers claims payments information and incurred losses information for the prediction of the outstanding loss liabilities by assigning appropriate credibility weights to these different channels of information. The benefits of our method are that

- it combines two different channels of information to get a unified ultimate loss prediction;
- for claims payments observation the CL structure is preserved using credibility weighted correction terms to the CL factors;
- for deterministic standard deviation parameters we can calculate both the claims reserves and the conditional MSEP analytically;
- the full predictive distribution of the outstanding loss liabilities is obtained from Monte Carlo simulations, this allows one to consider any risk measure;
- for stochastic standard deviation parameters all key figures and the full predictive distribution of the outstanding loss liabilities are obtained from the MCMC method.
- a model extension will allow the inclusion of tail development factors, for details see Merz and Wüthrich (submitted for publication).

Appendix. Proofs

In this appendix we prove all the statements. The proof of Theorem 3.2 easily follows from its likelihood function (and it is a special version of Theorem 6.4 in Gisler and Wüthrich, 2008), therefore we omit its proof.

Proof of Proposition 2.2. Note that we only consider conditional distributions, given the parameter $\mathbf{\Theta}$, and for this conditional distributions claims in different accident years are independent. Therefore we can restrict ourselves to one fixed accident year i. The vector

$$\left(\log I_{i,j+l}, \log I_{i,j}, \log I_{i,j-1}, \ldots, \log I_{i,0}\right)'\Big|_{\{\mathbf{\Theta}\}}$$

has a multivariate Gaussian distribution with mean $(\mu_{j+l}, \mu_j, \mu_{j-1}, \ldots, \mu_0)$ and covariance matrix Σ with elements given by: for $n \geq m \in \{j+l, j, j-1, \ldots, 0\}$

$$\mathrm{Cov}\left(\log I_{i,n}, \log I_{i,m}|\mathbf{\Theta}\right) = v_n^2.$$

Henceforth, we can apply Lemma 2.1 to the random variable $\log I_{i,j+l}|_{\{\mathcal{B}_j^I,\mathbf{\Theta}\}}$ with parameters $m_1 = \mu_{j+l}$, $m^{(2)} = (\mu_j, \ldots, \mu_0)$, $\Sigma_{1,1} = v_{j+l}^2$, $\Sigma_{2,2}$ is the covariance matrix of $X^{(2)} = \left(\log I_{i,j}, \ldots, \log I_{i,0}\right)'|_{\{\mathbf{\Theta}\}}$ and

$$\Sigma_{1,2} = \left(v_{j+l}^2, \ldots, v_{j+l}^2\right) \in \mathbb{R}^{j+1}.$$

We obtain from Lemma 2.1 a Gaussian distribution and there remains the calculation of the explicit parameters of the Gaussian distribution. Note that the covariance matrix $\Sigma_{2,2}$ has the following form

$$\Sigma_{2,2} = \left(v^2_{(j+1-n)\vee(j+1-m)}\right)_{1\leq n,m\leq j+1},$$

where $(j+1-n) \vee (j+1-m) = \max\{j+1-n, j+1-m\}$. Henceforth, $\Sigma_{2,2}$ has a fairly simple structure which gives a nice form for its inverse

$$\Sigma_{2,2}^{-1} = \left(b_{n,m}\right)_{1\leq n,m\leq j+1},$$

with diagonal elements

$$b_{1,1} = \frac{v_{j-1}^2}{v_j^2(v_{j-1}^2 - v_j^2)},$$

$$b_{n,n} = \frac{v_{j-n}^2 - v_{j+2-n}^2}{(v_{j+1-n}^2 - v_{j+2-n}^2)(v_{j-n}^2 - v_{j+1-n}^2)} \quad \text{for } n \in \{2, \ldots, j\},$$

$$b_{j+1,j+1} = \frac{1}{v_0^2 - v_1^2},$$

and off-diagonal elements 0 except for the side diagonals

$$b_{n,n+1} = \frac{-1}{v_{j-n}^2 - v_{j+1-n}^2} \quad \text{for } n \in \{1, \ldots, j\},$$

and its symmetric counterpart $b_{n,n-1}$ for $n \in \{2, \ldots, j+1\}$. This matrix has the following property

$$\Sigma_{1,2}\Sigma_{2,2}^{-1} = \left(v_{j+l}^2/v_j^2, 0, \ldots, 0\right) \in \mathbb{R}^{j+1},$$

from which the claim follows. □

Proof of Corollary 2.3. Proposition 2.2 implies for the conditional expectation

$$E\left[I_{i,J}|\mathcal{B}_j^I, \mathbf{\Theta}\right] = \exp\left\{\mu_J + \frac{v_J^2}{v_j^2}\left(\log I_{i,j} - \mu_j\right) + \frac{v_J^2}{2}\left(1 - \frac{v_J^2}{v_j^2}\right)\right\}$$

$$= \exp\left\{\mu_J + (1-\alpha_j)\left(\log I_{i,j} - \mu_j\right) + \alpha_j v_J^2/2\right\}$$

$$= I_{i,j}\exp\left\{\sum_{l=j}^{J-1}\Psi_l\right\}\exp\left\{-\alpha_j\left(\log I_{i,j} - \mu_j\right) + \alpha_j v_J^2/2\right\}.$$

Finally, observe that

$$\alpha_j v_J^2 - \sum_{l=j}^{J-1}\tau_l^2 = \sum_{l=j}^{J-1}\tau_l^2\left(\frac{v_J^2}{v_j^2} - 1\right) = -\alpha_j\sum_{l=j}^{J-1}\tau_l^2.$$

This completes the proof. □

Proof of Theorem 2.4. The proof is similar to the one of Proposition 2.2 and uses Lemma 2.1. Again we only consider conditional distributions, given the parameter $\mathbf{\Theta}$, therefore we can restrict ourselves to one fixed accident year i. Using the Markov property of cumulative payments, we see that it suffices to consider the vector

$$\left(\log I_{i,J}, \log P_{i,j}, \log I_{i,j}, \ldots, \log I_{i,0}\right)'\Big|_{\{\mathbf{\Theta}\}},$$

which has a multivariate Gaussian distribution with mean $(\mu_J, \eta_j, \mu_j, \ldots, \mu_0)$ and covariance matrix similar to Proposition 2.2 but with an additional column and row for

$$\mathrm{Cov}\left(\log P_{i,j}, \log I_{i,l}|\mathbf{\Theta}\right) = w_j^2 \quad \text{for } l \in \{J, j, \ldots, 0\}.$$

Henceforth, we can apply Lemma 2.1 to the random variable $\log I_{i,J}|_{\{\mathcal{B}_j,\mathbf{\Theta}\}}$ with parameters $m_1 = \mu_J$, $m^{(2)} = (\eta_j, \mu_j, \ldots, \mu_0)$, $\Sigma_{1,1} = v_J^2$, $\Sigma_{2,2}$ is the covariance matrix of $X^{(2)} = \left(\log P_{i,j}, \log I_{i,j}, \ldots, \log I_{i,0}\right)'\Big|_{\{\mathbf{\Theta}\}}$ and

$$\Sigma_{1,2} = \left(w_j^2, v_J^2, \ldots, v_J^2\right) \in \mathbb{R}^{j+2}.$$

Thus, there remains the calculation of the explicit parameters of the Gaussian distribution. Note that the covariance matrix $\Sigma_{2,2}$ has now the following form

w_j^2 for elements in the first column or first row,

$v^2_{(j+2-n)\vee(j+2-m)}$ for elements in the remaining right lower square $2 \leq n, m \leq j+2$.

Therefore, $\Sigma_{2,2}$ has again a simple form whose inverse can easily be calculated and has a similar structure to the one given in Proposition 2.2.

$$\Sigma_{1,2}\Sigma_{2,2}^{-1} = \left(\frac{v_j^2 - v_J^2}{v_j^2 - w_j^2}, \frac{v_J^2 - w_j^2}{v_j^2 - w_j^2}, 0, \ldots, 0\right)$$
$$= \left(1-\beta_j, \beta_j, 0, \ldots, 0\right) \in \mathbb{R}^{j+2}.$$

This implies (note that $v_j^2 > v_J^2 > w_j^2$)

$$m_1 + \Sigma_{1,2}\Sigma_{2,2}^{-1}\left(X^{(2)} - m^{(2)}\right) = \mu_J + (1-\beta_j)\left(\log P_{i,j} - \eta_j\right) + \beta_j\left(\log I_{i,j} - \mu_j\right).$$

Moreover, we have

$$\Sigma_{1,1} - \Sigma_{1,2}\Sigma_{2,2}^{-1}\Sigma_{2,1} = v_J^2 - (1-\beta_j)w_j^2 - \beta_j v_J^2$$
$$= (1-\beta_j)(v_J^2 - w_j^2),$$

from which the claim follows. □

Proof of Corollary 2.5. Theorem 2.4 implies

$$E\left[I_{i,J}|\mathcal{B}_j, \mathbf{\Theta}\right]$$
$$= \exp\left\{\mu_J + (1-\beta_j)(\log P_{i,j} - \eta_j) + \beta_j(\log I_{i,j} - \mu_j) + (1-\beta_j)(v_J^2 - w_j^2)/2\right\}$$
$$= P_{i,j}\exp\left\{\sum_{l=j+1}^{J}\Phi_l + \sigma_l^2/2\right\}\exp\left\{\beta_j(\eta_j - \log P_{i,j} + \log I_{i,j} - \mu_j - (v_J^2 - w_j^2)/2)\right\}. \quad \square$$

Proof of Theorem 3.3. From the likelihood (3.2) it immediately follows that the posterior distribution $\mathbf{\Psi}$, given $\mathcal{D}_j^I$, is Gaussian.

Table 10
Observed cumulative payments $P_{i,j}$.

i/j	0	1	2	3	4	5	6	7	8	9
0	1,216,632	1,347,072	1,786,877	2,281,606	2,656,224	2,909,307	3,283,388	3,587,549	3,754,403	3,921,258
1	798,924	1,051,912	1,215,785	1,349,939	1,655,312	1,926,210	2,132,833	2,287,311	2,567,056	
2	1,115,636	1,387,387	1,930,867	2,177,002	2,513,171	2,931,930	3,047,368	3,182,511		
3	1,052,161	1,321,206	1,700,132	1,971,303	2,298,349	2,645,113	3,003,425			
4	808,864	1,029,523	1,229,626	1,590,338	1,842,662	2,150,351				
5	1,016,862	1,251,420	1,698,052	2,105,143	2,385,339					
6	948,312	1,108,791	1,315,524	1,487,577						
7	917,530	1,082,426	1,484,405							
8	1,001,238	1,376,124								
9	841,930									

Hence, there remains the calculation of the posterior parameters. First we square out all the terms for obtaining the Ψ_j^2 and $\Psi_j\Psi_n$ terms

$$\sum_{i=0}^{J}\frac{1}{v_{J-i}^2}\left(\sum_{n=J-i}^{J}\Psi_n\right)^2+\sum_{j=0}^{J-1}\sum_{i=0}^{J-j-1}\frac{1}{\tau_j^2}\Psi_j^2+\sum_{j=0}^{J}\frac{1}{t_j^2}\Psi_j^2$$
$$=\sum_{j=0}^{J}\Psi_j^2\left(\frac{J-j}{\tau_j^2}+\frac{1}{t_j^2}+\sum_{i=0}^{j}\frac{1}{v_i^2}\right)+2\sum_{j=0}^{J}\Psi_j\sum_{n=j+1}^{J}\Psi_n\sum_{i=0}^{j}\frac{1}{v_i^2}.$$

This gives the covariance matrix $\Sigma(\mathcal{D}_J^I)$. The posterior mean $\boldsymbol{\psi}^{\text{post}}(\mathcal{D}_J^I)$ is obtained by solving the posterior maximum likelihood functions for Ψ_j. They are given by

$$\frac{\partial \log u\left(\boldsymbol{\Psi}|\mathcal{D}_J^I\right)}{\partial\Psi_j}=\left[t_j^{-2}\psi_j-\sum_{i=0}^{J-j-1}\tau_j^{-2}\log\frac{I_{i,j}}{I_{i,j+1}}\right.$$
$$\left.-\sum_{i=J-j}^{J}v_{J-i}^{-2}\log I_{i,J-i}\right]-\sum_{n=0}^{J}a_{j,n}^{I}\Psi_n\overset{!}{=}0,$$

from which the claim follows. □

Proof of Theorem 3.4. The proof is similar to the proof of Theorem 3.3. From (3.6) it is straightforward that the posterior distribution is again a multivariate Gaussian distribution. Henceforth, there remains the calculation of the first two moments.

The inverse covariance matrix Σ^{-1} is obtained by considering all the terms Φ_m^2, Ψ_n^2 and $\Phi_m\Psi_n$ in the posterior distribution $u\left(\boldsymbol{\Theta}|\mathcal{D}_J\right)$ similar to the proof of Theorem 3.3. The posterior mean is obtained by solving the posterior maximum likelihood functions for Φ_j and Ψ_j. They are given by

$$\frac{\partial\log u\left(\boldsymbol{\Theta}|\,\mathcal{D}_J\right)}{\partial\Phi_j}=\left[s_j^{-2}\phi_j+\sum_{i=0}^{J-j}\sigma_j^{-2}\log\frac{P_{i,j}}{P_{i,j-1}}\right.$$
$$\left.+\sum_{i=J-j+1}^{J}\left(v_{J-i}^2-w_{J-i}^2\right)^{-1}\log\frac{I_{i,J-i}}{P_{i,J-i}}\right]$$
$$-\sum_{m=0}^{J}\Phi_m a_{j,m}-\sum_{m=0}^{J-1}\Psi_m a_{j,J+1+m}\overset{!}{=}0,$$

and

$$\frac{\partial\log u\left(\boldsymbol{\Theta}|\,\mathcal{D}_J\right)}{\partial\Psi_j}=\left[t_j^{-2}\psi_j-\sum_{i=0}^{J-j-1}\tau_j^{-2}\log\frac{I_{i,j}}{I_{i,j+1}}\right.$$
$$\left.-\sum_{i=J-j}^{J}\left(v_{J-i}^2-w_{J-i}^2\right)^{-1}\log\frac{I_{i,J-i}}{P_{i,J-i}}\right]$$
$$-\sum_{m=0}^{J}\Phi_m a_{J+1+j,m}-\sum_{m=0}^{J-1}\Psi_j a_{J+1+j,J+1+m}\overset{!}{=}0.$$

Henceforth, this implies

$$(c_0,\ldots,c_J,b_0,\ldots,b_{J-1})=\Sigma(\mathcal{D}_J)^{-1}\left(\Phi_0,\ldots,\Phi_J,\Psi_0,\ldots,\Psi_{J-1}\right)'$$

from which the claim follows. □

Proof of Theorem 4.1. We use the usual decomposition for the variance given by

$$\operatorname{Var}\left(\sum_{i=1}^{J}P_{i,J}\middle|\,\mathcal{D}_J\right)=\operatorname{Var}\left(E\left[\sum_{i=1}^{J}P_{i,J}\middle|\,\mathcal{D}_J,\boldsymbol{\Theta}\right]\middle|\,\mathcal{D}_J\right)$$
$$+E\left[\operatorname{Var}\left(\sum_{i=1}^{J}P_{i,J}\middle|\,\mathcal{D}_J,\boldsymbol{\Theta}\right)\middle|\,\mathcal{D}_J\right].$$

For the second term we use that the accident years are independent, conditionally given $\boldsymbol{\Theta}$. Using Theorem 2.4 this leads to

$$E\left[\operatorname{Var}\left(\sum_{i=1}^{J}P_{i,J}\middle|\,\mathcal{D}_J,\boldsymbol{\Theta}\right)\middle|\,\mathcal{D}_J\right]$$
$$=E\left[\sum_{i=1}^{J}\operatorname{Var}\left(P_{i,J}|\mathcal{D}_J,\boldsymbol{\Theta}\right)\middle|\,\mathcal{D}_J\right]$$
$$=E\left[\sum_{i=1}^{J}E\left[P_{i,J}|\mathcal{D}_J,\boldsymbol{\Theta}\right]^2\left(\mathrm{e}^{(1-\beta_{J-i})(v_J^2-w_{J-i}^2)}-1\right)\middle|\,\mathcal{D}_J\right]$$
$$=\sum_{i=1}^{J}\left(\mathrm{e}^{(1-\beta_{J-i})(v_J^2-w_{J-i}^2)}-1\right)E\left[E\left[P_{i,J}|\mathcal{D}_J,\boldsymbol{\Theta}\right]^2|\mathcal{D}_J\right].$$

For the first term we obtain

$$\operatorname{Var}\left(E\left[\sum_{i=1}^{J}P_{i,J}\middle|\,\mathcal{D}_J,\boldsymbol{\Theta}\right]\middle|\,\mathcal{D}_J\right)$$
$$=\operatorname{Var}\left(\sum_{i=1}^{J}E\left[P_{i,J}|\mathcal{D}_J,\boldsymbol{\Theta}\right]|\mathcal{D}_J\right)$$
$$=\sum_{1\le i,k\le J}E\left[E\left[P_{i,J}|\mathcal{D}_J,\boldsymbol{\Theta}\right]E\left[P_{k,J}|\mathcal{D}_J,\boldsymbol{\Theta}\right]|\mathcal{D}_J\right]$$
$$-\sum_{1\le i,k\le J}E\left[P_{i,J}|\mathcal{D}_J\right]E\left[P_{k,J}|\mathcal{D}_J\right].$$

Henceforth, this provides

$$\operatorname{Var}\left(\sum_{i=1}^{J}P_{i,J}\middle|\,\mathcal{D}_J\right)=\sum_{i=1}^{J}\mathrm{e}^{(1-\beta_{J-i})(v_J^2-w_{J-i}^2)}E\left[E\left[P_{i,J}|\mathcal{D}_J,\boldsymbol{\Theta}\right]^2|\mathcal{D}_J\right]$$
$$+2\sum_{1\le i<k\le J}E\left[E\left[P_{i,J}|\mathcal{D}_J,\boldsymbol{\Theta}\right]E\left[P_{k,J}|\mathcal{D}_J,\boldsymbol{\Theta}\right]|\mathcal{D}_J\right]$$
$$-\sum_{1\le i,k\le J}E\left[P_{i,J}|\mathcal{D}_J\right]E\left[P_{k,J}|\mathcal{D}_J\right].$$

Table 11
Observed incurred losses $I_{i,j}$.

i/j	0	1	2	3	4	5	6	7	8	9
0	3,362,115	5,217,243	4,754,900	4,381,677	4,136,883	4,094,140	4,018,736	3,971,591	3,941,391	3,921,258
1	2,640,443	4,643,860	3,869,954	3,248,558	3,102,002	3,019,980	2,976,064	2,946,941	2,919,955	
2	2,879,697	4,785,531	4,045,448	3,467,822	3,377,540	3,341,934	3,283,928	3,257,827		
3	2,933,345	5,299,146	4,451,963	3,700,809	3,553,391	3,469,505	3,413,921			
4	2,768,181	4,658,933	3,936,455	3,512,735	3,385,129	3,298,998				
5	3,228,439	5,271,304	4,484,946	3,798,384	3,702,427					
6	2,927,033	5,067,768	4,066,526	3,704,113						
7	3,083,429	4,790,944	4,408,097							
8	2,761,163	4,132,757								
9	3,045,376									

Table 12
Posterior correlation matrix corresponding to $\Sigma(\mathcal{D}_J)$.

	Φ_0 (%)	Φ_1 (%)	Φ_2 (%)	Φ_3 (%)	Φ_4 (%)	Φ_5 (%)	Φ_6 (%)	Φ_7 (%)	Φ_8 (%)	Φ_9 (%)	Ψ_0 (%)	Ψ_1 (%)	Ψ_2 (%)	Ψ_3 (%)	Ψ_4 (%)	Ψ_5 (%)	Ψ_6 (%)	Ψ_7 (%)	Ψ_8 (%)
Φ_0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Φ_1	0	100	−2	−1	−1	0	−1	0	−1	0	1	1	1	0	0	0	0	0	0
Φ_2	0	−2	100	−4	−2	−1	−2	−1	−2	−1	2	3	3	1	0	0	0	0	0
Φ_3	0	−1	−4	100	−3	−3	−4	−2	−4	−1	1	3	5	1	1	0	0	0	0
Φ_4	0	−1	−2	−3	100	−2	−3	−2	−4	−1	0	1	2	1	1	0	0	0	0
Φ_5	0	0	−1	−3	−2	100	−6	−3	−7	−2	0	1	2	1	1	0	0	0	0
Φ_6	0	−1	−2	−4	−3	−6	100	−8	−18	−6	1	1	2	2	2	1	1	0	0
Φ_7	0	0	−1	−2	−2	−3	−8	100	−18	−6	0	1	1	1	1	0	1	0	0
Φ_8	0	−1	−2	−4	−4	−7	−18	−18	100	−34	1	1	3	2	2	1	1	2	1
Φ_9	0	0	−1	−1	−1	−2	−6	−6	−34	100	0	0	1	1	1	0	1	1	2
Ψ_0	0	1	2	1	0	0	1	0	1	0	100	−1	−1	0	0	0	0	0	0
Ψ_1	0	1	3	3	1	1	1	1	1	0	−1	100	−2	0	0	0	0	0	0
Ψ_2	0	1	3	5	2	2	2	1	3	1	−1	−2	100	−1	−1	0	0	0	0
Ψ_3	0	0	1	1	1	1	2	1	2	1	0	0	−1	100	0	0	0	0	0
Ψ_4	0	0	0	1	1	1	2	1	2	1	0	0	−1	0	100	0	0	0	0
Ψ_5	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	100	0	0	0
Ψ_6	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	100	0	0
Ψ_7	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	100	0
Ψ_8	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	0	100

Using Theorem 3.4 provides the claim:

$$E\left[E\left[P_{i,J}|\mathcal{D}_J,\mathbf{\Theta}\right]E\left[P_{k,J}|\mathcal{D}_J,\mathbf{\Theta}\right]|\mathcal{D}_J\right]$$
$$= E\left[P_{i,J}|\mathcal{D}_J\right]E\left[P_{k,J}|\mathcal{D}_J\right]\exp\left\{\mathbf{e}_i'\Sigma(\mathcal{D}_J)\mathbf{e}_k\right\}. \quad \square$$

The proof of Theorem 4.2 is similar to the proof of Theorem 4.1.

References

Asmussen, S., Glynn, P.W., 2007. Stochastic Simulation. Springer.

Bühlmann, H., De Felice, M., Gisler, A., Moriconi, F., Wüthrich, M.V., 2009. Recursive credibility formula for chain ladder factors and the claims development result. Astin Bulletin 39 (1), 275–306.

Bühlmann, H., Gisler, A., 2005. A Course in Credibility Theory and its Applications. In: Springer Universitext.

Dahms, R., 2008. A loss reserving method for incomplete claim data. Bulletin Swiss Association of Actuaries 127–148.

Dahms, R., Merz, M., Wüthrich, M.V., 2009. Claims development result for combined claims incurred and claims paid data. Bulletin Francais d'Actuariat 9 (18), 5–39.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov Chain Monte Carlo in Practice. Chapman & Hall.

Gisler, A., Wüthrich, M.V., 2008. Credibility for the chain ladder reserving method. Astin Bulletin 38 (2), 565–600.

Gogol, D., 1993. Using expected loss ratios in reserving. Insurance: Mathematics and Economics 12 (3), 297–299.

Halliwell, L.J., 1997. Cojoint prediction of paid and incurred losses. CAS Forum Summer. pp. 241–379.

Halliwell, L.J., 2009. Modeling paid and incurred losses together. CAS E-Forum Spring.

Hertig, J., 1985. A statistical approach to the IBNR-reserves in marine insurance. Astin Bulletin 15 (2), 171–183.

Liu, H., Verrall, R.J., 2008. Bootstrap estimation of the predictive distributions of reserves using paid and incurred claims. In: Conference Paper, 38th Astin Colloquium 2008. Manchester, UK.

Mack, T., 1993. Distribution-free calculation of the standard error of chain ladder reserve estimates. Astin Bulletin 23 (2), 213–225.

Merz, M., Wüthrich, M.V., 2006. A credibility approach to the Munich chain-ladder method. Blätter DGVFM Band XXVII, 619–628.

Merz, M., Wüthrich, M.V., 2010. Estimation of tail factors in the paid–incurred chain reserving method, Preprint (submitted for publication).

Posthuma, B., Cator, E.A., Veerkamp, W., Zwet, van E.W., 2008. Combined analysis of paid and incurred losses. CAS E-Forum Fall. pp. 272–293.

Quarg, G., Mack, T., 2004. Munich chain ladder. Blätter DGVFM Band XXVI, 597–630.

Roberts, G.O., Gelman, A., Gilks, W.R., 1997. Weak convergence and optimal scaling of random walks Metropolis algorithm. Annals of Applied Probability 7, 110–120.

Scollnik, D.P.M., 2001. Actuarial modeling with MCMC and BUGS. North American Actuarial Journal 5 (2), 96–125.

Venter, G.G., 2008. Distribution and value of reserves using paid and incurred triangles. CAS E-Forum Fall. pp. 348–375.

van Dyk, D.A., Meng, X.-L., 2001. The art of data augmentation. Journal of Computational and Graphical Statistics 10 (1), 1–50.

Verdier, B., Klinger, A., 2005. JAB chain: a model-based calculation of paid and incurred loss development factors. In: Conference Paper, 36th Astin Colloquium 2005. Zürich, Switzerland.

Williams, D., 1991. Probability with Martingales. In: Cambridge Mathematical Textbooks.

Wüthrich, M.V., Merz, M., 2008. Stochastic Claims Reserving Methods in Insurance. Wiley.