Jeong, Himchan, and Dipak Kumar Dey. 2024. "Multi-Peril Frequency Credibility Premium via Shared Random Effects." *Variance* 17 (1).

Financial and Statistical Methods

Multi-Peril Frequency Credibility Premium via Shared Random Effects

Himchan Jeong[1a], Dipak Kumar Dey[2b]

[1] Department of Statistics and Actuarial Science, Simon Fraser University, [2] Department of Statistics, University of Connecticu

Keywords: Bonus-malus system, credibility, insurance ratemaking, multi-peril insurance, overdispersion, shared random effects

Variance

Vol. 17, Issue 1, 2024

It is not uncommon to observe that an insurance policy consists of multiple coverages that cover different perils over multiple years. In this article, we propose a shared random effects model to not only capture the unobserved heterogeneity in risks but also induce a natural dependence structure among the claims from multiple perils. The proposed model has a nice interpretation and comes with closed forms of joint likelihood and credibility premiums that are desirable for ease of implementation in practice. Both the proposed method and benchmark methods are calibrated using a public insurance dataset provided by the Wisconsin Local Government Property Insurance Fund, where one can observe both policy characteristics and historical claims information. It turns out that the proposed method shows acceptable level of performance compared to many pre-existing benchmarks in terms of out-of-sample validation.

Address for Correspondence: himchan_jeong@sfu.ca

1. INTRODUCTION

Usually, an auto insurance policy consists of various types of coverage against claims, such as collision, bodily injury liability, or property damage liability. Upon the presence of available covariates, their impacts for each type of claim could be different so that one needs to apply different regression coefficients for each claim type. Besides, it is natural to expect that various types of claims have different correlations with the unobserved heterogeneity in risk, such as driving habits. For example, an at-fault liability claim is usually at the control of a driver, so that the occurrence of such a claim is strongly related with driving habits. However, glass damage claims (for example, due to a heavy hail storm) are almost out of the driver's control, so that association with driving habits is quite low. In the end, a company needs to develop a predictive model that utilizes the covariate information to calculate a presumptive premium as well as the unobserved heterogeneity in risks, which can be indirectly captured by past claims history of each policyholder.

The idea of capturing the unobserved heterogeneity via past claims history has been discussed under the name of credibility theory. Bailey (1950) and Bühlmann (1967) suggested the credibility procedure, which usually arrives at the following formula:

$$\begin{aligned}\text{Posterior premium} = Z\times&\text{Claim experience}\\ &+(1-Z)\times\text{Prior premium},\end{aligned}$$

where Z is the so-called credibility weight. Since then, credibility theory has been developed and explored in the actuarial science literature. For example, Mayerson (1964) and Jewell (1974) analyzed credibility theory from a Bayesian perspective. While these researchers working on credibility theory provided a way to incorporate the unobservable heterogeneity, both the observed and unobserved heterogeneities need to be addressed in actuarial ratemaking practice, as shown in Norberg (1986). In this regard, Frees, Young, and Luo (1999) provided a general framework that integrated well-known credibility theory and regression analysis based on the use of linear mixed models. The linear mixed model is an extension of the linear model, whereby the response variable is affected by both the observed covariates via associated regression coefficients (fixed effects) and unobserved quantities (random effects). Upon the use of both fixed and random effects with a mixed model, it is noteworthy that credibility premiums have a

a Himchan is a Fellow of the Society of Actuaries (SOA) and holds a Ph.D. from the University of Connecticut. He has been actively involved in teaching and conducting research in actuarial science for several years. In recognition for his academic achievements and excellence, he has been awarded the James C. Hickman Scholarship from SOA recently in 2018-2020. His current research interest is predictive modeling for ratemaking and reserving of property and casualty insurance.

b Dipak Kumar Dey is an Indian-American statistician best known for his work on Bayesian methodologies. He is currently the Board of Trustees Distinguished Professor in the Department of Statistics at the University of Connecticut. Dey has an international reputation as a statistician as well as a data scientist.

close connection with bonus-malus systems, which reward or penalize a policyholder based on past claims history. Frangos and Vrontos (2001) showed that, upon the presence of fixed effects, credibility premium (or posterior premium) for claim takes the following form:

$$\text{Posterior premium} = \frac{r+\text{Actual claim experience}}{r+\text{Expected claim experience}} \times \text{Prior premium},$$

where r denotes a smoothing factor in the bonus-malus system so that larger values of r put less weight on past claims history. Similar results have been shown in the recent literature such as Jeong (2020) and Jeong and Valdez (2020a). Boucher and Denuit (2008) and Boucher, Denuit, and Guillen (2009) considered the use of a zero-inflated Poisson distribution in panel data to derive credibility premiums.

However, one should also consider possible dependence among the claims from either multiple lines of business or multiple coverages, in order to apply to ratemaking practices. Frees, Meyers, and Cummings (2010) proposed dependence modeling for multi-peril insurance under the individual risk model, which decomposes the compound loss into occurrence (a binary response whether there is a claim or not) and total amounts given a claim. To model dependence among the occurrence of claims in multiple coverages, they utilized multivariate binary regression via Gaussian copulas. On the other hand, Frees, Lee, and Yang (2016) proposed dependence modeling for multi-peril insurance under the collective risk model, which decomposes the compound loss into claim frequency and severity components. They applied copulas to describe possible dependence among the multivariate frequencies and severities, respectively. Andrade e Silva and Centeno (2017) suggested using various generalized bivariate frequency distributions to capture multi-peril dependence, which also allow capturing possible overdispersion in the observed claim counts. Quan and Valdez (2018) proposed the use of a multivariate decision tree in multi-peril claim amounts regression.

Interestingly, there are only a few research papers that incorporated such multi-peril dependence and longitudinal property simultaneously. For example, Frees (2003) proposed a way to calculate credibility premium of compound loss for multi-peril policy by specifying only the mean and covariance structures of claim frequencies and severities. While Frees took a non-parametric approach, Bermúdez and Karlis (2015) applied bivariate Poisson distributions for a posteriori ratemaking of two types of claims. However, these papers are limited to non-parametric or parametric setting without covariates, so that it is hard to incorporate both the observed and unobserved heterogeneities of a policyholder simultaneously with their models. Abdallah, Boucher, and Cossette (2016) provided a sophisticated approach to utilizing a bivariate Sarmanov distribution to account for the dependence between two lines of business with time weight for past claims, which is appropriate to deal with two types of claims but difficult to extend to the case where there are more than two types of claims. Recently, Pechon et al. (2018; 2019, 2020) proposed credibility premium formulas with multiple types of claims using correlated lognormal random effects. Although they provided a comprehensive framework that considers possible dependence among multi-peril claims in a longitudinal setting, their method requires high-dimensional integration (six or seven) in order to obtain credibility factors and marginal frequency distributions for every single policyholder, which may be burdensome in practice where the number of policyholders is usually large.

A novel method is proposed in this article that enables us to consider both the observed and unobserved heterogeneities of a policyholder in a longitudinal setting for multi-peril insurance via shared random effects. The proposed method can handle more than two types of claims and leads to a readily available closed-form formula for multi-peril credibility premiums. In short, the shared random effect for a policyholder is a latent variable that cannot be observed but affects the claims from multiple perils over time and induces a natural dependence structure among the claims. This article is organized as follows. In Section 2, the proposed method—a shared random effects model for multi-peril frequency—is specified in detail, and important characteristics of the model are provided including the multi-peril credibility premium formula. In Section 3, an empirical data analysis is conducted using a data set provided by the Wisconsin Local Government Property Insurance Fund (LGPIF). We conclude this article in Section 4 with some remarks and possible directions for future work.

2. METHODOLOGY

Let us consider a usual data structure for multi-peril frequency. For an insurance policy of i^{th} policyholder where $i = 1, 2, \ldots, I$, we may observe the multi-peril claim frequencies over time $t = 1, 2, \ldots, T$ for $j = 1, \ldots, J$ types of coverage as follows:

$$\mathcal{D} = \left\{ \left(N_{it}^{(1)}, \ldots, N_{it}^{(j)}, \ldots, N_{it}^{(J)}, \mathbf{x}_{it} \right) \,\middle|\, i = 1, \ldots, I,\, t = 1, \ldots, T \right\}, \quad (2.1)$$

where $\mathbf{x}_{it}$ is a p-dimensional vector that captures the observable characteristics of the policy and $N_{it}^{(j)}$ is defined as the number of accident(s) from claim type $j \in \{1, 2, \ldots, J\}$, for the i^{th} policyholder in year t, respectively.

Based on the data structure, one can consider the following issues for multi-peril frequency modeling. Although the Poisson distribution is widely used for modeling frequency, it is often questionable whether it is valid due to possible overdispersion. In order to handle this issue, according to Wedderburn (1974), one can use a quasi-Poisson distribution to capture overdispersion in the observed data as follows:

$$N \sim \mathcal{QP}(\nu, w) \iff wN \sim \mathcal{P}(w\nu),$$

for which $\mathbb{E}[N] = \nu$ and $\text{Var}[N] = \nu/w$. Here $X \sim \mathcal{P}(\lambda)$ means X follows a Poisson distribution with mean λ. In that regard, the Poisson distribution is a special case of the quasi-Poisson distribution where $w = 1$. Due to this nested property, some research results enable to test

$$H_0 : w = 1 \text{ versus } H_1 : w \neq 1.$$

For details, see Cameron and Trivedi (1990).

Table 1. A hypothetical dataset on claims experience of two identical policyholders

Year	Gender	Age	Vehicle Year	Policyholder A		Policyholder B	
				# of Claim 1	# of Claim 2	# of Claim 1	# of Claim 2
2018	F	25	8	0	0	1	0
2019	F	26	9	0	0	2	0
2020	F	27	10	1	0	1	1

Secondly, recall that the data structure described in (2.1) has longitudinality, or repeated measurements of the same policyholder over time. Thus, it is natural to incorporate random effects, which may account for unobserved heterogeneity in risks of each policyholder. The following hypothetical example in Table 1 shows that one can capture the unobserved heterogeneity in risks by observing the residuals, after controlling for the effect of the observed covariates. The latter can be explained in terms of random effects for policyholders A and B. In Table 1, it is possible to observe that both policyholders have the same observable characteristics, whereas they exhibit quite different past claims experiences. In such a case, one may suspect that the observed covariates might not be sufficient to fully explain the heterogeneity in risks. Random effects can help capture such (unobserved) heterogeneity for both types of claims.

As we can see, the random effects model has connections with credibility theory and bonus-malus systems as well. Some papers have incorporated random effects in a ratemaking perspective, including but not limited to Gómez-Déniz and Vázquez-Polo (2005) and Jeong and Valdez (2020b).

Based on the observations from Table 1, let us assume that the number of claims is affected by both observable covariates via associated regression coefficients and the (common) unobserved heterogeneity factor θ_i. It is shared for all types of coverages and every year of claims of policyholder i as follows:

$$\begin{aligned} &N_{it}^{(j)}|\mathbf{x}_{it}, e_{it}, \theta_i \overset{indep}{\sim} \mathcal{QP}(\theta_i \nu_{it}^{(j)}, w^{(j)}) \\ &\quad \iff w^{(j)} N_{it}^{(j)}|\mathbf{x}_{it}, e_{it}, \theta_i \overset{indep}{\sim} \mathcal{P}(\theta_i w^{(j)} \nu_{it}^{(j)}) \text{ and } \mathbb{E}[\theta_i] = 1, \end{aligned} \tag{2.2}$$

where $\nu_{it}^{(j)} = e_{it} \exp(\mathbf{x}_{it}\alpha^{(j)})$, $e_{it} \in (0, 1]$ is the exposure, and $w^{(j)}$ is the (unknown) weight for each j^{th} line of business or type of coverage. Here $\nu_{it}^{(j)}$ accounts for the observed heterogeneity in risks of policyholder i at time t for coverage j, while multiplicative random effect θ_i accounts for the unobserved heterogeneity in risks of policyholder i. One can also find a very comprehensive setup of the Poisson model with a shared random effect in Shi and Valdez (2014a), which is helpful to understand the proposed setup of the quasi-Poisson model with a shared random effect. Though N does not necessarily have integer-valued quantities with quasi-Poisson distribution, the proposed framework allows for obtaining the posterior distribution of wN. Subsequently, distributional quantities of N are easy to derive with the distribution, which are our main interest.

With the specification above, it is straightforward to derive that

$$\mathbb{E}\left[w^{(j)} N_{it}^{(j)}|\theta_i\right] = w^{(j)} \nu_{it}^{(j)} \theta_i \text{ and } \mathbb{E}\left[N_{it}^{(j)}\right] = \nu_{it}^{(j)},$$

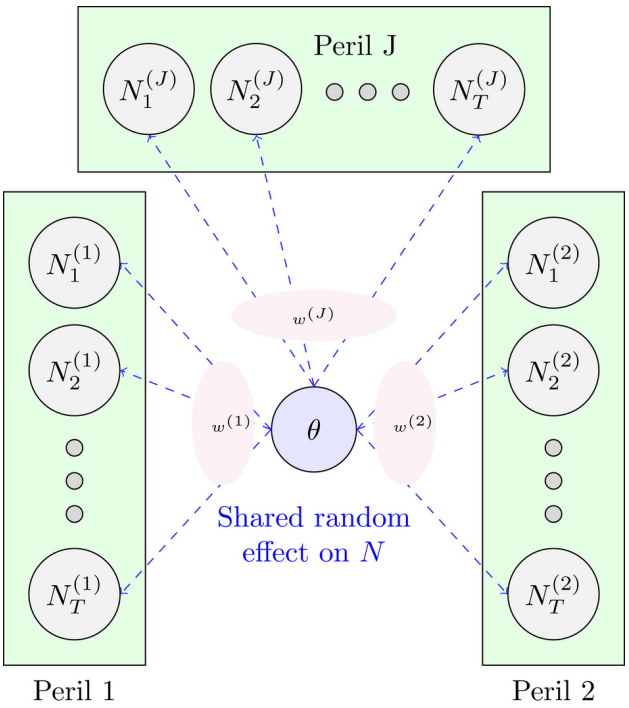

Figure 1. Graphical description of the shared random effects model for multi-peril frequency

since $\mathbb{E}[\theta_i] = 1$. Therefore, $\nu_{it}^{(j)}$ can be interpreted as a prior mean of the number of claims at time t, for policyholder i, without any information on her/his unobserved heterogeneity. For notational convenience, we suppress the subscript i from now on when there is no confusion. Figure 1 provides a graphical description of the proposed model specified in (2.2). Although we assume that the common random effect affects multiple types of peril simultaneously, we also allow $w^{(j)}$ to vary so that the impacts of θ to $(N^{(1)}, \ldots, N^{(J)})$ are not necessarily identical.

Note that the use of shared random effects is not the only way to capture possible dependence among multi-peril claims in a longitudinal setting. For example, one can utilize copulas to capture such dependence. For details, see Yang and Shi (2018).

The impact of the unobserved heterogeneity on the claim frequency is inherently unknown. Therefore, it needs to be described with a specified prior distribution. In a Bayesian analysis, it is usually recommended to use a non-informative prior or less informative prior on θ unless we have enough knowledge on the dynamics of the random parameter. In this sense, one can suggest to use Jeffreys' prior, perhaps the most widely used noninformative prior

in Bayesian analyses. According to Jeffreys (1946), Jeffreys' prior is defined as the square root of the determinant of the Fisher information matrix. One can see that under the model specification in (2.2), Jeffreys' prior of θ is given as $\theta^{-1/2}$ and the corresponding posterior distribution is proportional to $\theta^{N-1/2}\exp(-\nu\theta)$ so that $\theta|N \sim \mathcal{G}(N+0.5,\, \nu^{-1})$. Here $X \sim \mathcal{G}(\alpha, \lambda)$ means X follows a gamma distribution with mean $\alpha\lambda$ and variance $\alpha\lambda^2$.

The corresponding posterior is proper although Jeffreys' prior itself is improper. Nevertheless, there is an identifiability issue, unless we impose a condition on the mean of the random effects. For example, it is customary to set $\mathbb{E}[\theta] = 0$ when θ is an additive random effect. Likewise, since θ is a multiplicative random effect in our specified model, we need to impose that $\mathbb{E}[\theta] = 1$. Therefore, in order to satisfy the identifiability condition as well as have a similar posterior as in the case of Jeffreys' prior, one can propose the following prior on θ:

$$\begin{aligned} &\pi(\theta) \propto \theta^{r-1}e^{-\theta r} \\ &\quad \text{so that } \theta \sim \mathcal{G}(r, 1/r) \\ &\quad \text{and } \mathbb{E}[\theta] = 1, \mathrm{Var}[\theta] = \frac{1}{r}. \end{aligned} \tag{2.3}$$

Here $\pi(\theta)$ is the so-called conjugate prior. This choice is not ad hoc but from the argument of prior elicitation, considering both informativeness of the prior and identifiability of the fixed effects. Based on the model specification in (2.2) and (2.3), it follows that

$$\mathbb{E}\left[N_t^{(j)}\right] = \nu_t^{(j)} \text{ and } \mathrm{Var}\left[N_t^{(j)}\right] = \nu_t^{(j)}/w^{(j)} + \left(\nu_t^{(j)}\right)^2/r, \tag{2.4}$$

since $w^{(j)}N_t^{(j)} \sim \mathcal{NB}\left(r, \frac{w_t^{(j)}\nu_t^{(j)}}{r+w_t^{(j)}\nu_t^{(j)}}\right)$ from Theorem 1. Here $X \sim \mathcal{NB}(r,p)$ means X follows a negative binomial distribution with mean $\frac{pr}{1-p}$ and variance $\frac{pr}{(1-p)^2}$.

It is also shown that

$$\begin{aligned} \mathrm{Cov}\left(N_t^{(j)}, N_{t'}^{(k)}\right) &= \mathbb{E}\left[\mathrm{Cov}\left(N_t^{(j)}, N_{t'}^{(k)}\middle|\theta\right)\right] \\ &\quad + \mathrm{Cov}\left(\mathbb{E}\left[N_t^{(j)}\middle|\theta\right], \mathbb{E}\left[N_{t'}^{(k)}\middle|\theta\right]\right) \\ &= \mathrm{Cov}\left(\nu_t^{(j)}\theta, \nu_{t'}^{(k)}\theta\right) = \nu_t^{(j)}\nu_{t'}^{(k)}/r, \\ \mathrm{Corr}\left(N_t^{(j)}, N_{t'}^{(k)}\right) &= \frac{\mathrm{Cov}\left(N_t^{(j)}, N_{t'}^{(k)}\right)}{\mathrm{Var}\left[N_t^{(j)}\right]\mathrm{Var}\left[N_{t'}^{(k)}\right]} \\ &= \frac{1}{\sqrt{1+r/w^{(j)}\nu_t^{(j)}}\sqrt{1+r/w^{(k)}\nu_{t'}^{(k)}}}, \end{aligned} \tag{2.5}$$

for $t, t' = 1, \ldots, T$ and $j, k = 1, \ldots, J$ due to conditional independence of $N_t^{(j)}$ and $N_{t'}^{(k)}$ given θ. Therefore, the proposed model induces a natural dependence structure among the frequencies of the multiple insurance coverages over time. For example, if $r \to \infty$, then the proposed model is reduced to an independent quasi-Poisson model. Further, if $r \to \infty$ and $w^{(j)} = 1$ for all $j = 1, \ldots, J$, then the proposed model is reduced to an independent Poisson model.

Since $\mathbb{E}\left[N_t^{(j)}|\theta\right] = \theta\nu_t^{(j)}$, the multiplicative random effect θ can be understood as a bonus-malus factor on top of $\nu_t^{(j)}$ for determination of insurance premium. The value of hyperparameter r can be determined by using either prior knowledge or the method of moments. For example, from 54% to 200% was once a common range of bonus-malus factors for the frequency premium according to Lemaire (1998). Therefore, one can incorporate this idea on choosing the hyperparameter r for our proposed prior so that the 95% highest posterior density (HPD) interval of θ can include $(0.54, 2.00)$. One can also consistently estimate r since the expectation of $\sum_{t\neq t', j\neq k}(N_{it}^{(j)} - \nu_{it}^{(j)})(N_{it'}^{(k)} - \nu_{it'}^{(k)})$ is $\frac{1}{r}\sum_{t\neq t', j\neq k}\nu_{it}^{(j)}\nu_{it'}^{(k)}$ for $i = 1, \ldots, I$ (Sutradhar and Jowaheer 2003) so that

$$\hat{r} = \frac{\sum_{i=1}^{I}\sum_{t\neq t', j\neq k}\hat{\nu}_{it}^{(j)}\hat{\nu}_{it'}^{(k)}}{\sum_{i=1}^{I}\sum_{t\neq t', j\neq k}(N_{it}^{(j)} - \hat{\nu}_{it}^{(j)})(N_{it'}^{(k)} - \hat{\nu}_{it'}^{(k)})}, \tag{2.6}$$

where $\hat{\nu}$ is consistently estimated by solving generalized estimating equations (GEE) with a pre-specified mean structure (Liang and Zeger 1986; Purcaru, Guillén, and Denuit 2004; Denuit et al. 2007). According to this specification, one can easily observe that posterior density for θ as follows:

Lemma 1. *Based on the model specification described in (2.2) and (2.3), we have*

$$\theta|\mathbf{n}_T^{(1)}, \ldots, \mathbf{n}_T^{(J)} \sim \mathcal{G}\left(\sum_{j=1}^{J}\sum_{t=1}^{T} w^{(j)}n_t^{(j)} + r,\, \frac{1}{\sum_{j=1}^{J}\sum_{t=1}^{T} w^{(j)}\nu_t^{(j)} + r}\right),$$

where $\mathbf{n}_T^{(j)} = (n_1^{(j)}, \ldots, n_T^{(j)})$.

Proof. It is easy to show that

$$\begin{aligned} \pi(\theta|\mathbf{n}_T^{(1)}, \ldots, \mathbf{n}_T^{(J)}) &\propto \pi(\theta)\prod_{j=1}^{J}\prod_{t=1}^{T} p(w^{(j)}n_t^{(j)}|\theta) \\ &\propto \theta^{r-1}e^{-\theta r}\prod_{j=1}^{J}\prod_{t=1}^{T}\theta^{w^{(j)}n_t}e^{-\nu_t^{(j)}\theta} \\ &= \theta^{\sum_{j=1}^{J}\sum_{t=1}^{T}w^{(j)}n_t^{(j)}+r-1}\exp\left(-\theta\left(\sum_{j=1}^{J}\sum_{t=1}^{T}w^{(j)}\nu_t^{(j)} + r\right)\right), \end{aligned}$$

which leads to

$$\theta|\mathbf{n}_T^{(1)}, \ldots, \mathbf{n}_T^{(J)} \sim \mathcal{G}\left(\sum_{j=1}^{J}\sum_{t=1}^{T} w^{(j)}n_t^{(j)} + r,\, \frac{1}{\sum_{j=1}^{J}\sum_{t=1}^{T} w^{(j)}\nu_t^{(j)} + r}\right).$$

□

Further, it is not difficult to show that the predictive distribution of N_{T+1} given $\mathbf{n}_T^{(1)}, \ldots, \mathbf{n}_T^{(J)}$ follows a negative binomial distribution as follows:

Theorem 1. *Suppose that* $(w^{(1)}\mathbf{n}_T^{(1)}, \ldots, w^{(J)}\mathbf{n}_T^{(J)})$ *and* $(w^{(j)}N_{T+1}^{(j)})$ *follow the model specified in (2.2) and (2.3). Then we have*

$$w^{(j)}N_{T+1}^{(j)}|\mathbf{n}_T^{(1)}, \ldots, \mathbf{n}_T^{(J)} \sim \mathcal{NB}\left(r_T,\, \frac{w^{(j)}\nu_{T+1}^{(j)}}{\tilde{r}_T + w^{(j)}\nu_{T+1}^{(j)}}\right),$$

where $r_T = r + \sum_{j=1}^{J}\sum_{t=1}^{T} w^{(j)}n_t^{(j)}$, $\hat{r}_T = r + \sum_{j=1}^{J}\sum_{t=1}^{T} w^{(j)}\nu_t^{(j)}$, *and*

$$\mathbb{E}\left[N_{T+1}^{(j)}|\mathbf{n}_T^{(1)}, \ldots, \mathbf{n}_T^{(J)}\right] = \frac{r + \sum_{j=1}^{J}(w^{(j)}\sum_{t=1}^{T}n_t^{(j)})}{r + \sum_{j=1}^{J}(w^{(j)}\sum_{t=1}^{T}\nu_t^{(j)})}\nu_{T+1}^{(j)}. \tag{2.7}$$

Proof. It is sufficient to show that if $N|\theta \sim \mathcal{P}(\nu\theta)$ and $\theta \sim \mathcal{G}(r, 1/\beta)$, then

$$\begin{aligned} p(n) &= \int_0^\infty p(n|\theta)\pi(\theta)d\theta \\ &= \int_0^\infty \lambda^n\frac{\theta^n}{n!}e^{-\theta\lambda}\frac{\beta^r\theta^{r-1}}{\Gamma(r)}e^{-\theta\beta} \\ &= \frac{\nu^n\beta^r}{n!\Gamma(r)}\int_0^\infty \theta^{n+r-1}e^{-\theta(\beta+\nu)} \\ &= \frac{\nu^n\beta^r}{n!\Gamma(r)}\frac{\Gamma(n+r)}{(\beta+\nu)^{n+r}} \\ &= \frac{\Gamma(n+r)}{n!\Gamma(r)}\left(\frac{\nu}{\beta+\nu}\right)^n\left(\frac{\beta}{\beta+\nu}\right)^r, \end{aligned}$$

so that $N \sim \mathcal{NB}\left(r, \frac{\nu}{\beta+\nu}\right)$ and $\mathbb{E}[N] = \frac{r\nu}{\beta}$.

□

Therefore, the predictive premium of $N_{T+1}^{(j)}$, with the previous T years' information, is given in the form of a product of a prior premium that depends on regression coefficients associated with observable covariates, and with a pooled estimate of the credibility factor, which accounts for the unobserved (shared) heterogeneity on the claim frequency of a specific policyholder. Note that (2.7) is a natural extension of Frangos and Vrontos (2001), whereby $J = 1$ and $w^{(1)} = 1$. For example, if $J = 2$, and $N^{(1)}$, $N^{(2)}$ denote the claim frequencies due to at-fault liability and glass damage, respectively, then the pooled estimate of unobserved heterogeneity is given as

$$\begin{aligned}
&\mathbb{E}\left[\theta \mid \mathbf{n}_T^{(1)}, \mathbf{n}_T^{(2)}\right] \\
&= \frac{r + w^{(1)}\sum_{t=1}^{T} n_t^{(1)} + w^{(2)}\sum_{t=1}^{T} n_t^{(2)}}{r + w^{(1)}\sum_{t=1}^{T} \nu_t^{(1)} + w^{(2)}\sum_{t=1}^{T} \nu_t^{(2)}} \\
&= \frac{\text{smoothing factor } + \text{ weighted sum of actual frequencies}}{\text{smoothing factor } + \text{ weighted sum of expected frequencies}}.
\end{aligned} \tag{2.8}$$

Therefore, the proposed model specification enables us to consider different levels of contribution to the credibility factor, $\mathbb{E}\left[\theta|\mathbf{n}_T^{(1)}, \ldots, \mathbf{n}_T^{(J)}\right]$, by allowing varying dispersions $w^{(j)}$ for multiple perils.

In (2.8), r, the hyperparameter of the prior distribution of θ, is working as a smoothing factor. For example, if $r \to \infty$, then $\pi(\theta|\mathbf{n})$ converges to a Dirac delta function at 1, which implies that $\mathbb{P}(\theta = 1) = 1$ and ends up with a very informative point-mass prior. On the other hand, any choice of a smaller r leads to use of a less informative prior.

In light of the empirical Bayes method, the estimation of parameters $\alpha^{(j)}$ for $j = 1, 2, \ldots, J$ can be done by maximizing the following joint loglikelihood:

$$\begin{aligned}
&\ell\left(\alpha^{(1)}, \ldots, \alpha^{(J)} \mid \mathbf{n}_T^{(1)}, \ldots, \mathbf{n}_T^{(J)}\right) \\
&\quad = \sum_{i=1}^{M} \ln p\left(w^{(1)}\mathbf{n}_{iT}^{(1)}, \ldots, w^{(J)}\mathbf{n}_{iT}^{(J)}\right),
\end{aligned} \tag{2.9}$$

where

$$\begin{aligned}
&p(w^{(1)}\mathbf{n}_{iT}^{(1)}, \ldots, w^{(J)}\mathbf{n}_{iT}^{(J)}) \\
&= \int \prod_{j=1}^{J}\prod_{t=1}^{T} p(w^{(j)}n_{it}^{(j)}|\theta_i)\pi(\theta_i)d\theta_i \\
&\propto \prod_{j=1}^{J}\prod_{t=1}^{T}\left[\left(\frac{w^{(j)}\nu_{it}^{(j)}}{\sum_{j=1}^{J}\sum_{t=1}^{T} w^{(j)}\nu_{it}^{(j)} + r}\right)^{w^{(j)}n_{it}^{(j)}}\right. \\
&\qquad \left.\times\left(\frac{r}{\sum_{j=1}^{J}\sum_{t=1}^{T} w^{(j)}\nu_{it}^{(j)} + r}\right)^{r}\right],
\end{aligned} \tag{2.10}$$

which follows a multivariate negative binomial (MVNB) distribution indeed.

3. DATA ANALYSIS

In this article, we use a public data set of insurance claims that has been provided by the Wisconsin Local Government Property Insurance Fund (LGPIF). The data set consists of six years of observations, from 2006 to 2011. We used a training set of 5,677 observations from 2006 to 2010 and a test set of 1,098 observations from 2011. Since there is a unique identifier that allows tracking of the claims of a policyholder over time, it is indeed longitudinal data. Furthermore, the data set contains the claims information for multiple lines of business so that one can try multivariate longitudinal frequency modeling. Here, information on inland marine claims (IM), collision claims from new vehicles (CN), and comprehensive claims from new vehicles (PN) is used with six categorical and four continuous explanatory variables. Table 2 provides a detailed summary of the observable policy characteristics of all lines of insurance business. Note that all of the categorical variables, which are dummy variables for location types, are commonly used for all types of coverage, but their impact is not necessarily identical.

Before the fixed effects are incorporated in the marginal models of the coverages, a variable selection was performed by using penalized Poisson likelihood with elastic net, which considers both L_1 and L_2 penalties. Since the location parameters are correlated with each other, it is desirable to consider L_2 penalty for performing variable selection group-wise on top of L_1 penalty. The variable selection was implemented using `glmnet` (Friedman et al. 2021) by setting $\alpha = 0.5$ so that weights on L_1 and L_2 penalties are equal. As shown in Figure 2, it turns out that the validation errors measured by Poisson deviances are minimized with no variable selection in all three coverages. Therefore, all the available covariates were used for the calibration of marginal models eventually.

Since the proposed model specified in (2.2) and (2.3) assumes the presence of both overdispersion and association among the numbers of claims from multiple coverages, one needs to investigate their effect in order to validate the application of more complicated models than the naive independent Poisson model. Table 3 shows the frequency tables for the number of IM, CN, and PN claims. By applying the usual measures to investigate the association between each pair of categorical responses, summarized in Table 4, one can observe that there are significant positive associations among the numbers of IM, CN, and PN claims, at least without controlling the possible impacts of common covariates. Note that these measures need to be used with precautions since they may not be appropriate for insurance claim counts with excessive zeros, due to possible ties in ranking. For discussions on effects of ties on the rank-based correlation, see Kendall (1945).

There has been some research on ways to estimate the dispersion parameter in a quasi-Poisson distribution. For example, McCullagh and Nelder (1989) proposed $\hat{\phi} = \frac{1}{m-p}\sum_{k=1}^{m}\frac{(n_k - \hat{\nu}_k)^2}{\hat{\nu}_k}$, while Cameron and Trivedi (1990) proposed $\tilde{\phi} = \frac{1}{m}\sum_{k=1}^{m}\tilde{R}_k$ where $\tilde{R}_k = \frac{(n_k - \hat{\nu}_k)^2 - n_k}{\hat{\nu}_k}$, $\hat{\nu}_k$ is expected value of k^{th} frequency response depending on the estimated parameter(s) from the calibrated model, p is the number of estimated parameter(s), and m is the number of total observations. According to Table 5, for all types of coverages, we can observe that dispersion parameters are estimated to be quite above one. (Remember that the dispersion parameter should equal one under the Poisson assumption.)

Table 2. Observable policy characteristics used as covariates

Categorical variables	Description		Proportions	
TypeCity	Indicator for city entity:	Y=1	14%	
TypeCounty	Indicator for county entity:	Y=1	5.78%	
TypeMisc	Indicator for miscellaneous entity:	Y=1	11.04%	
TypeSchool	Indicator for school entity:	Y=1	28.17%	
TypeTown	Indicator for town entity:	Y=1	17.28%	
TypeVillage	Indicator for village entity:	Y=1	23.73%	
Continuous variables		Minimum	Mean	Maximum
CoverageIM	Log coverage amount of IM claim in mm	0	0.85	46.75
lnDeductIM	Log deductible amount for IM claim	0	5.34	9.21
CoverageCN	Log coverage amount of CN claim in mm	0	0.1	4.14
CoveragePN	Log coverage amount of PN claim in mm	0	0.16	25.67

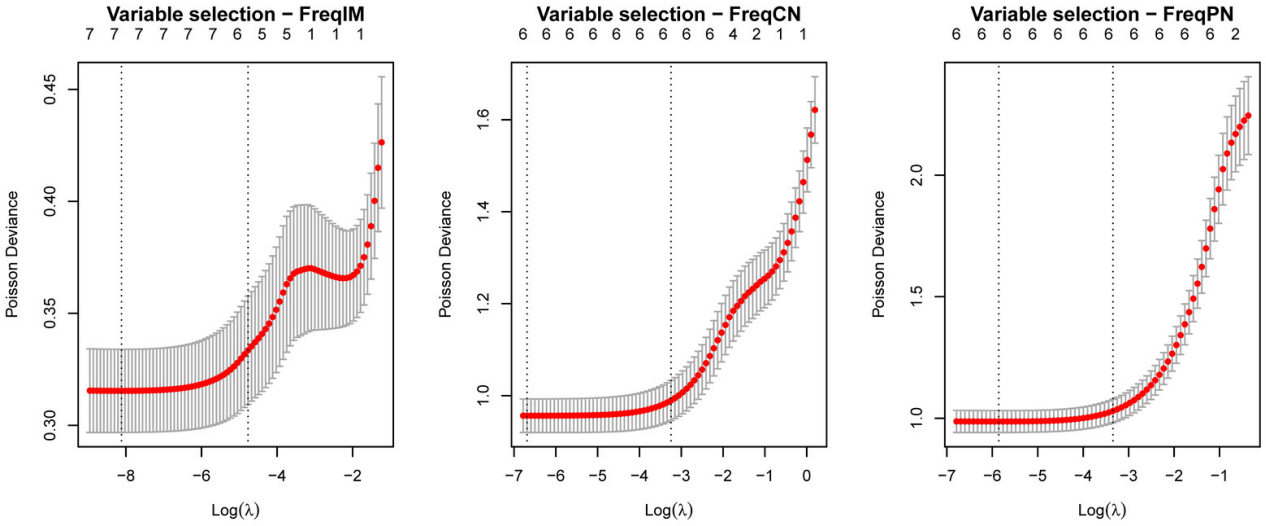

Figure 2. Variable selection via elastic net

Cameron and Trivedi (1990) also showed that

$$\tilde{Z} = \frac{\tilde{\phi}}{\sqrt{\frac{1}{n}\sum_{k=1}^{m}\left(\tilde{R}_k - \tilde{\phi}\right)^2}} \simeq \mathcal{N}(0,1) \text{ under } H_0 : \phi = 1,$$

when m is large enough and n_k, for $k = 1, \ldots, m$, are independent. For the given data, the values of $\hat{Z}$ for IM, CN, and PN claims are 2.7844, 4.1829, and 3.8956, respectively. Therefore, one can find significant evidence to reject the hypothesis that there is no overdispersion in all lines of business.

Since we have observed possible overdispersion from the data, it is natural to consider the following models that can capture overdispersion or possible dependence, except for the independent Poisson model:

- **Poisson**: Independent Poisson model as an industry benchmark. Note that it is a special case of the proposed model where $w^{(1)} = w^{(2)} = w^{(3)} = 1$ and $r = \infty$.
- **Poisson/gamma**: Poisson random effects model where each coverage has unique unobserved heterogeneity as follows:
$$N_{it}^{(j)} | \mathbf{x}_{it}, e_{it}, \theta_i \overset{indep}{\sim} \mathcal{P}(\theta_i^{(j)} \nu_{it}^{(j)}),\ \theta^{(j)} \sim \mathcal{G}(r_j, 1/r_j).$$
Note that the marginal mean of $\mathbb{E}\left[N_{it}^{(j)}\right]$ under Poisson/gamma model is $\nu_{it}^{(j)}$, as in the independent Poisson model, so that we use the same regression coefficients of $\alpha^{(j)}$ as the independent Poisson model.
- **ZI-Poisson**: Zero-inflated Poisson model with the following density:
$$p(n_t^{(j)}; \nu_t^{(j)}, \eta_t^{(j)}) = \begin{cases} \eta_t^{(j)} + (1-\eta_t^{(j)}) \exp(-\nu_t^{(j)}), & n_t^{(j)} = 0 \\ (1-\eta_t^{(j)}) \dfrac{(\nu_t^{(j)})^{n_t^{(j)}} \exp(-\nu_t^{(j)})}{n_t^{(j)}!}, & n_t^{(j)} \neq 0, \end{cases}$$
which also assumes $n_t^{(j)} \perp n_{t'}^{(j')}$ for all $t \neq t'$ or $j \neq j'$. Since $\eta_t^{(j)}$ captures the zero-inflated probability, it is usual to model $\eta_t^{(j)}$ as $\eta_t^{(j)} = \dfrac{\exp(\mathbf{x}_t \gamma^{(j)})}{1 + \exp(\mathbf{x}_t \gamma^{(j)})}$.

Table 3. Frequency tables for IM, CN, and PN claims

# of CN claims	# of IM claims					
	0	1	2	3	4	5 +
0	5142	134	19	4	3	3
1	206	18	4	2	0	0
2	57	12	5	0	0	0
3	13	9	2	0	1	1
4	12	2	2	0	0	0
5 +	11	7	8	0	0	0
# of PN claims	# of IM claims					
	0	1	2	3	4	5 +
0	5210	126	15	4	3	2
1	130	16	7	1	0	1
2	39	8	3	1	0	0
3	21	9	3	0	0	0
4	11	5	2	0	0	1
5 +	30	18	10	0	1	0
# of CN claims	# of PN claims					
	0	1	2	3	4	5 +
0	5153	92	27	11	6	16
1	167	32	13	7	5	6
2	34	15	5	6	2	12
3	3	4	2	3	4	10
4	2	4	1	2	2	5
5 +	1	8	3	4	0	10

Table 4. Measures of association among the frequencies of IM, CN, and PN claims

	Pearson correlation	Kendall's τ	Spearman's ρ
IM and CN	0.2093	0.2120	0.2576
PN and IM	0.2820	0.2860	0.2705
PN and CN	0.4520	0.4597	0.4555

Table 5. Estimates of overdispersion parameters for IM, CN, and PN claims

	# of IM claims	# of CN claims	# of PN claims
$\hat{\phi}$	1.1420	1.1072	1.5729
$\tilde{\phi}$	1.2165	1.3045	1.4751

- **Copula**: While each marginal component is modeled with a Poisson distribution, possible dependence among the lines of business is captured by a copula, as proposed in Shi and Valdez (2014b), so that the joint probability of $n_t^{(1)}$ and $n_t^{(2)}$ is given as follows:

$$\begin{aligned} p(n_t^{(1)}, n_t^{(2)}, n_t^{(3)}; \nu_t^{(1)}, \nu_t^{(2)}, \nu_t^{(3)}, \phi) = {} & C_\phi(u_{t+}^{(1)}, u_{t+}^{(2)}, u_{t+}^{(3)}) \\ & - C_\phi(u_{t-}^{(1)}, u_{t+}^{(2)}, u_{t+}^{(3)}) \\ & - C_\phi(u_{t+}^{(1)}, u_{t-}^{(2)}, u_{t+}^{(3)}) \\ & + C_\phi(u_{t-}^{(1)}, u_{t-}^{(2)}, u_{t+}^{(3)}) \\ & - C_\phi(u_{t+}^{(1)}, u_{t+}^{(2)}, u_{t-}^{(3)}) \\ & + C_\phi(u_{t-}^{(1)}, u_{t+}^{(2)}, u_{t-}^{(3)}) \\ & + C_\phi(u_{t+}^{(1)}, u_{t-}^{(2)}, u_{t-}^{(3)}) \\ & - C_\phi(u_{t-}^{(1)}, u_{t-}^{(2)}, u_{t-}^{(3)}), \end{aligned}$$

  where C_ϕ means a copula function parametrized by ϕ, $u_{t+}^{(j)} = \mathbb{P}(N_t^{(j)} \le n_t^{(j)})$, and $u_{t-}^{(j)} = \mathbb{P}(N_t^{(j)} < n_t^{(j)})$.
- **Proposed model**: The model specified in (2.2) and (2.3).

In the calibration of the **Proposed** model, values of $w^{(j)}$ should be determined either by prior knowledge of the characteristics of the multiple perils or by some statistics determined by observations. As mentioned in (2.8), a pooled estimate of the credibility factor is given as the ratio of weighted sum of actual frequencies to weighted sum of expected frequencies. Therfore, $w^{(j)}$ can be interpreted as a peril-specific factor that provides information about the association between the peril and the unobserved heterogeneity, such as driving habits. For instance, suppose there are two types of coverage: property damage (PD) liability and glass damage to the driver's own car. In that case, claims from PD liability are more positively associated with

the unobserved heterogeneity in risks than claims from glass damage, since claims from PD liability are at-fault of the policyholder, whereas claims from the driver's own glass damage may be due to external factors such as a heavy hail storm. In this regard, it is natural to put more weight on the past claims experiences from PD liability than those from drivers' own glass damage in calculating the pooled estimate of the credibility factor, or, equivalently, set $w^{(j)}$ for PD liability higher than $w^{(j)}$ for drivers' own glass damage. Therefore, the proposed method enables practitioners to incorporate their prior knowledge about the characteristics of multiple perils in the modeling process with flexibility.

If there is no strong evidence to pre-specify the values of $w^{(j)}$ with prior knowledge, then one can directly use the estimated overdispersion parameters to determine $w^{(j)}$ by matching the moments. Recall that from (2.4) we have

$$\begin{aligned}\sum_{t=1}^{T_i}\mathbb{E}\left[\left(N_{it}^{(j)}-\nu_{it}^{(j)}\right)^2\right] &= \sum_{t=1}^{T_i}\mathrm{Var}\left[N_{it}^{(j)}\right]\\ &= \sum_{t=1}^{T_i}\frac{\nu_{it}}{w^{(j)}}+\frac{\left(\nu_{it}^{(j)}\right)^2}{r},\end{aligned}$$

for $j=1,\ldots,J$ and $i=1,\ldots,I$ under the proposed model so that one can consistently estimate $w^{(j)}$ in a similar manner to (2.6):

$$\hat{w}^{(j)}=\frac{\sum_{i=1}^{I}\sum_{t=1}^{T_i}\hat{\nu}_{it}^{(j)}}{\sum_{i=1}^{I}\sum_{t=1}^{T_i}\left[\left(N_{it}^{(j)}-\hat{\nu}_{it}^{(j)}\right)^2-\left(\hat{\nu}_{it}^{(j)}\right)^2/\hat{r}\right]}.$$

Once the overdispersion parameters $w^{(j)}$ for $j=1,2,3$ are specified, either from the prior knowledge or moment matching, one can estimate $\alpha^{(1)}$, $\alpha^{(2)}$, and $\alpha^{(3)}$ from the joint likelihood after integrating out the shared random effect, as given in (2.10). In our analysis, we used the moment matching approach so that r, $w^{(1)}$, $w^{(2)}$, and $w^{(3)}$ are estimated to be 2.256, 0.7226, 0.4828, and 0.3732, respectively. Since $w^{(1)}>w^{(2)}>w^{(3)}$ in our case, the relative contribution of IM frequencies to the proposed posterior ratemaking scheme is greater than those of CN and PN frequencies.

Tables 6, 7, and 8 provide the estimated regression coefficients for IM, CN, and PN frequencies, respectively. One can observe that the regression coefficients from the **Poisson** and **Proposed** models are similar while the coefficients from the other models are different. Further, as mentioned above, all types of frequencies share some dummy variables indicating location types, but their estimated coefficients are quite different from each other. This supports our model specification in that one can capture the fixed effects separately for each line of business while retaining a natural dependence structure by the shared random effects.

Figure 3 shows the distributions of the credibility factors under the proposed model, which are defined as the posterior expectations of θ given observations of past claims:

$$\begin{aligned}&\mathbb{E}\left[\theta|\mathbf{n}_T^{(1)},\mathbf{n}_T^{(2)},\mathbf{n}_T^{(3)}\right]\\ &\quad=\frac{r+w^{(1)}\sum_{t=1}^{T}n_t^{(1)}+w^{(2)}\sum_{t=1}^{T}n_t^{(2)}+w^{(3)}\sum_{t=1}^{T}n_t^{(3)}}{r+w^{(1)}\sum_{t=1}^{T}\nu_t^{(1)}+w^{(2)}\sum_{t=1}^{T}\nu_t^{(2)}+w^{(3)}\sum_{t=1}^{T}\nu_t^{(3)}}.\end{aligned}$$

Blue squares mean the average of credibility factors for each group. From the figure, one can see that, if there are no past claims, then the insured can get a discount since $n_t^{(j)}=0$ while $\nu_t^{(j)}>0$ for all j and t. It is also observed that, as a policyholder has more past claims, the credibility factor tends to rise on average. However, regardless of such general positive association between credibility factors and number of past claims, the latter varies because the proposed credibility premium formula utilizes both the actual claims and expected claims, which is based on the observable heterogeneity of risks. By doing so, an insurance company may avoid double discounts or surcharges due to segregating the impacts of the observable policy characteristics and unobservable heterogeneity in risks.

After all five models are calibrated, one can test the validity of each model by comparing the predicted performance via out-of-sample validation. In this case, it is natural to predict $N_{T+1}^{(j)}$ as $\hat{N}_{T+1}^{(j)}:=\mathbb{E}\left[N_{T+1}^{(j)}|\mathbf{n}_T^{(1)},\mathbf{n}_T^{(2)},\mathbf{n}_T^{(3)}\right]$ under each model. Recall that

$$\begin{aligned}&\mathbb{E}\left[N_{T+1}^{(j)}|\mathbf{n}_T^{(1)},\mathbf{n}_T^{(2)},\mathbf{n}_T^{(3)}\right]\\ &\quad=\frac{r+w^{(1)}\sum_{t=1}^{T}n_t^{(1)}+w^{(2)}\sum_{t=1}^{T}n_t^{(2)}+w^{(3)}\sum_{t=1}^{T}n_t^{(3)}}{r+w^{(1)}\sum_{t=1}^{T}\nu_t^{(1)}+w^{(2)}\sum_{t=1}^{T}\nu_t^{(2)}+w^{(3)}\sum_{t=1}^{T}\nu_t^{(3)}}\\ &\qquad\cdot e^{\mathbf{x}_{T+1}\alpha^{(j)}}\end{aligned}$$

under the **Proposed** model from Theorem 1, whereas

$$\mathbb{E}\left[N_{T+1}^{(j)}|\mathbf{n}_T^{(1)},\mathbf{n}_T^{(2)},\mathbf{n}_T^{(3)}\right]=\mathbb{E}\left[N_{T+1}^{(j)}\right]=e^{\mathbf{x}_{T+1}\alpha^{(j)}}$$

under the **Poisson** model.

To compare the out-of-sample validation performance, we use root-mean-square error (RMSE), mean absolute error (MAE), and Poisson deviance (PDEV) for each coverage that are defined as follows:

$$\begin{aligned}\text{RMSE}&:\sqrt{\frac{1}{I}\sum_{i=1}^{I}\left(N_{i,T_i+1}^{(j)}-\hat{N}_{i,T_i+1}^{(j)}\right)^2},\\ \text{MAE}&:\frac{1}{I}\sum_{i=1}^{I}|N_{i,T_i+1}^{(j)}-\hat{N}_{i,T_i+1}^{(j)}|,\\ \text{PDEV}&:2\sum_{i=1}^{I}\left[N_{i,T_i+1}^{(j)}\log(N_{i,T_i+1}^{(j)}/\hat{N}_{i,T_i+1}^{(j)})-(N_{i,T_i+1}^{(j)}-\hat{N}_{i,T_i+1}^{(j)})\right],\end{aligned}$$

where I means the number of observed policyholders in the validation set. We prefer a model with lower RMSE, MAE, and/or PDEV. As shown in Tables 9, 10, and 11, the **Proposed** model outperforms Poisson, ZI-Poisson, and Copula models in all validation measures. It is also shown that the **Proposed** model is comparable to Poisson-gamma model in RMSE and MAE while prediction performance of Poisson-gamma in PDEV is outstanding. Such empirical results show that the interdependence among the unobserved heterogeneities of coverages may exist but not be substantial in the given data set.

4. CONCLUSION

It is quite natural that an insurance company observes the same policyholder over many years with multiple coverages. Inspired by such characteristics, a new method is explored in this article that allows us to incorporate distinct fixed effects capturing inherent characteristics of each type of coverage (or line of business), while retaining a natural dependence structure among the claims from multiple perils over time, via the shared random effects. The proposed

Table 6. Estimation results for IM frequency

	Poisson		ZI-Poisson		Copula		Proposed	
	Estimate	p-value	Estimate	p-value	Estimate	p-value	Estimate	p-value
(Intercept)	-0.9897	0.0438	-0.9065	0.2776	-0.7749	0.1235	-0.9216	0.1636
TypeCity	-0.9538	0.0000	0.4245	0.0964	-1.0287	0.0000	-0.9318	0.0000
TypeMisc	-4.3685	0.0000	-4.7670	0.2480	-4.4375	0.0000	-4.2862	0.0003
TypeSchool	-2.7157	0.0000	0.0831	0.8759	-2.7790	0.0000	-2.6120	0.0000
TypeTown	-2.2558	0.0000	-1.4776	0.1378	-2.3341	0.0000	-2.1530	0.0000
TypeVillage	-1.8980	0.0000	-1.7966	0.0073	-1.9737	0.0000	-1.8120	0.0000
CoverageIM	0.0765	0.0000	0.0465	0.0000	0.0780	0.0000	0.0734	0.0000
lnDeductIM	-0.0684	0.3368	0.0159	0.8898	-0.0896	0.2226	-0.0870	0.3520

Table 7. Estimation results for CN frequency

	Poisson		ZI-Poisson		Copula		Proposed	
	Estimate	p-value	Estimate	p-value	Estimate	p-value	Estimate	p-value
(Intercept)	-0.0907	0.2691	0.4176	0.0000	-0.1619	0.0532	0.1064	0.4873
TypeCity	-0.9465	0.0000	-0.7764	0.0000	-0.8852	0.0000	-0.9542	0.0000
TypeMisc	-2.0818	0.0000	-1.5408	0.1527	-1.9922	0.0000	-2.1298	0.0002
TypeSchool	-2.0033	0.0000	-1.9537	0.0000	-1.9337	0.0000	-2.1467	0.0000
TypeTown	-3.1464	0.0000	-3.6159	0.0000	-3.0687	0.0000	-3.1514	0.0000
TypeVillage	-1.5165	0.0000	-1.3240	0.0000	-1.4527	0.0000	-1.6062	0.0000
CoverageCN	0.5975	0.0000	0.3989	0.0000	0.6079	0.0001	0.3446	0.0001

Table 8. Estimation results for PN frequency

	Poisson		ZI-Poisson		Copula		Proposed	
	Estimate	p-value	Estimate	p-value	Estimate	p-value	Estimate	p-value
(Intercept)	0.8214	0.0000	1.1213	0.0000	0.8302	0.0000	0.9848	0.0000
TypeCity	-2.9648	0.0000	-3.2390	0.0000	-2.8984	0.0000	-2.7868	0.0000
TypeMisc	-3.9064	0.0000	-1.1185	0.1818	-3.8191	0.0000	-3.7059	0.0001
TypeSchool	-3.3801	0.0000	-2.1860	0.0000	-3.3910	0.0000	-3.4867	0.0000
TypeTown	-4.2487	0.0000	-2.1314	0.0281	-4.2479	0.0000	-4.3704	0.0000
TypeVillage	-3.6947	0.0000	-3.2285	0.0010	-3.6899	0.0000	-3.6953	0.0000
CoveragePN	0.0835	0.0000	0.0744	0.0011	0.0636	0.0031	-0.0468	0.3797

Table 9. Out-of-sample validation for IM frequency prediction

	Poisson	Poisson-gamma	ZI-Poisson	Copula	Proposed
RMSE	0.6647	0.4996	0.7093	0.4813	0.4936
MAE	0.1232	0.1069	0.1260	0.1114	0.1127
PDEV	374.9218	270.2795	377.2521	319.7245	307.8984

Table 10. Out-of-sample validation for CN frequency prediction

	Poisson	Poisson-gamma	ZI-Poisson	Copula	Proposed
RMSE	0.4699	0.3564	0.4731	0.4654	0.3510
MAE	0.1268	0.0959	0.1261	0.1249	0.1002
PDEV	231.8781	158.9138	232.5247	226.7693	173.5909

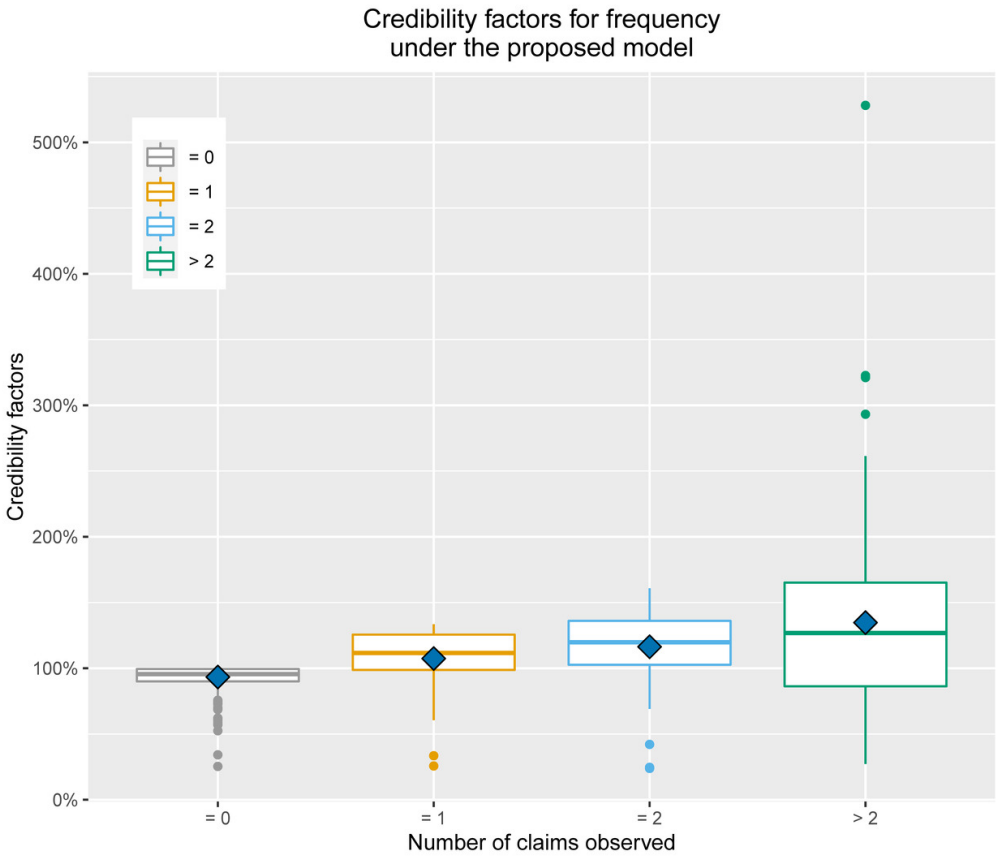

Figure 3. Credibility factors for multi-peril frequency

Table 11. Out-of-sample validation for PN frequency prediction

	Poisson	Poisson-gamma	ZI-Poisson	Copula	Proposed
RMSE	0.9775	0.6555	0.9768	0.9785	0.7613
MAE	0.1914	0.1366	0.1904	0.1918	0.1537
PDEV	379.5177	213.3153	378.3951	380.7286	243.8167

model has a good interpretation that can explain both overdispersion and serial/multi-peril dependence. The model is also easy to implement due to the presence of the closed form of joint likelihood and multi-peril premium formula. The proposed method is tested using a public property and casualty insurance data set provided by LGPIF. From the empirical study, it is observed that the proposed method outperforms many existing benchmarks but is less powerful than the model with unique heterogeneity for each coverage. Based on the results, one can consider a compromise between the complete shared random effects model (all coverages share the same random effect) and the unique random effects model (each coverage has its own unique random effect) by investigating underlying interdependence among the unobserved heterogeneities, as a future research topic.

ELECTRONIC SUPPLEMENTARY MATERIAL

The code for data analysis and out-of-sample validation is available at https://github.com/ssauljin/multi-peril_cred_premium/.

Submitted: November 18, 2020 EDT. Accepted: April 22, 2022 EDT. Published: January 08, 2024 EDT.

REFERENCES

Abdallah, Anas, Jean-Philippe Boucher, and Hélène Cossette. 2016. "Sarmanov Family of Multivariate Distributions for Bivariate Dynamic Claim Counts Model." *Insurance: Mathematics and Economics* 68 (May):120–33. https://doi.org/10.1016/j.insmatheco.2016.01.003.

Andrade e Silva, J. M., and M. de Lourdes Centeno. 2017. "Ratemaking of Dependent Risks." *ASTIN Bulletin* 47 (3): 875–94. https://doi.org/10.1017/asb.2017.16.

Bailey, A. L. 1950. "Credibility Procedures: Laplace's Generalization of Bayes' Rule and the Combination of Collateral Knowledge with Observed Data." In *Proceedings of the Casualty Actuarial Society*, 37:7–23.

Bermúdez, Lluís, and Dimitris Karlis. 2015. "A Posteriori Ratemaking Using Bivariate Poisson Models." *Scandinavian Actuarial Journal* 2017 (2): 148–58. https://doi.org/10.1080/03461238.2015.1094403.

Boucher, Jean-Philippe, and Michel Denuit. 2008. "Credibility Premiums for the Zero-Inflated Poisson Model and New Hunger for Bonus Interpretation." *Insurance: Mathematics and Economics* 42 (2): 727–35. https://doi.org/10.1016/j.insmatheco.2007.08.003.

Boucher, Jean-Philippe, Michel Denuit, and Montserrat Guillen. 2009. "Number of Accidents or Number of Claims? An Approach with Zero-Inflated Poisson Models for Panel Data." *Journal of Risk and Insurance* 76 (4): 821–46. https://doi.org/10.1111/j.1539-6975.2009.01321.x.

Bühlmann, Hans. 1967. "Experience Rating and Credibility." *ASTIN Bulletin* 4 (3): 199–207. https://doi.org/10.1017/s0515036100008989.

Cameron, A.Colin, and Pravin K. Trivedi. 1990. "Regression-Based Tests for Overdispersion in the Poisson Model." *Journal of Econometrics* 46 (3): 347–64. https://doi.org/10.1016/0304-4076(90)90014-k.

Denuit, Michel, Xavier Maréchal, Sandra Pitrebois, and Jean-François Walhin. 2007. *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems*. Wiley. https://doi.org/10.1002/9780470517420.

Frangos, Nicholas E., and Spyridon D. Vrontos. 2001. "Design of Optimal Bonus-Malus Systems with a Frequency and a Severity Component on an Individual Basis in Automobile Insurance." *ASTIN Bulletin* 31 (1): 1–22. https://doi.org/10.2143/ast.31.1.991.

Frees, Edward W. 2003. "Multivariate Credibility for Aggregate Loss Models." *North American Actuarial Journal* 7 (1): 13–37. https://doi.org/10.1080/10920277.2003.10596074.

Frees, Edward W., Gee Lee, and Lu Yang. 2016. "Multivariate Frequency-Severity Regression Models in Insurance." *Risks* 4 (1): 4. https://doi.org/10.3390/risks4010004.

Frees, Edward W., G. Meyers, and A. D. Cummings. 2010. "Dependent Multi-Peril Ratemaking Models." *ASTIN Bulletin* 40 (2): 699–726.

Frees, Edward W., Virginia R. Young, and Yu Luo. 1999. "A Longitudinal Data Analysis Interpretation of Credibility Models." *Insurance: Mathematics and Economics* 24 (3): 229–47. https://doi.org/10.1016/s0167-6687(98)00055-9.

Friedman, J., T. Hastie, R. Tibshirani, and B. Narasimhan. 2021. *Package "Glmnet."* CRAN R Repositary.

Gómez-Déniz, Emilio, and Francisco J. Vázquez-Polo. 2005. "Modelling Uncertainty in Insurance Bonus–Malus Premium Principles by Using a Bayesian Robustness Approach." *Journal of Applied Statistics* 32 (7): 771–84. https://doi.org/10.1080/02664760500079746.

Jeffreys, H. 1946. "An Invariant Form for the Prior Probability in Estimation Problems." *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 186 (1007): 453–61. https://doi.org/10.1098/rspa.1946.0056.

Jeong, Himchan. 2020. "Testing for Random Effects in Compound Risk Models via Bregman Divergence." *ASTIN Bulletin* 50 (3): 777–98. https://doi.org/10.1017/asb.2020.19.

Jeong, Himchan, and Emiliano A. Valdez. 2020a. "Bayesian Credibility Premium with GB2 Copulas." *Dependence Modeling* 8 (1): 157–71. https://doi.org/10.1515/demo-2020-0009.

———. 2020b. “Predictive Compound Risk Models with Dependence.” *Insurance: Mathematics and Economics* 94 (September):182–95. https://doi.org/10.1016/j.insmatheco.2020.07.011.

Jewell, William S. 1974. “Credible Means Are Exact Bayesian for Exponential Families.” *ASTIN Bulletin* 8 (1): 77–90. https://doi.org/10.1017/s0515036100009193.

Kendall, M. G. 1945. “The Treatment of Ties in Ranking Problems.” *Biometrika* 33 (3): 239–51. https://doi.org/10.1093/biomet/33.3.239.

Lemaire, Jean. 1998. “Bonus-Malus Systems: The European and Asian Approach to Merit-Rating.” *North American Actuarial Journal* 2 (1): 26–38. https://doi.org/10.1080/10920277.1998.10595668.

Liang, Kung-Yee, and Scott L. Zeger. 1986. “Longitudinal Data Analysis Using Generalized Linear Models.” *Biometrika* 73 (1): 13–22. https://doi.org/10.1093/biomet/73.1.13.

Mayerson, A. L. 1964. “A Bayesian View of Credibility.” In *Proceedings of the Casualty Actuarial Society*, 51:7–23.

McCullagh, P., and J. Nelder. 1989. *Generalized Linear Models*. 2nd ed. Chapman & Hall/CRC Monographs on Statistics and Applied Probability.

Norberg, Ragnar. 1986. “Hierarchical Credibility: Analysis of a Random Effect Linear Model with Nested Classification.” *Scandinavian Actuarial Journal* 1986 (3–4): 204–22. https://doi.org/10.1080/03461238.1986.10413807.

Pechon, Florian, Michel Denuit, and Julien Trufin. 2019. “Multivariate Modelling of Multiple Guarantees in Motor Insurance of a Household.” *European Actuarial Journal* 9 (2): 575–602. https://doi.org/10.1007/s13385-019-00201-5.

———. 2020. “Home and Motor Insurance Joined at a Household Level Using Multivariate Credibility.” *Annals of Actuarial Science* 15 (1): 82–114. https://doi.org/10.1017/s1748499520000160.

Pechon, Florian, Julien Trufin, and Michel Denuit. 2018. “Multivariate Modelling of Household Claim Frequencies in Motor Third-Party Liability Insurance.” *ASTIN Bulletin* 48 (3): 969–93. https://doi.org/10.1017/asb.2018.21.

Purcaru, O., M. Guillén, and M. Denuit. 2004. “Linear Credibility Models Based on Time Series for Claim Counts.” STAT Discussion Paper 0422, University of Louvain, Belgium. https://dial.uclouvain.be/pr/boreal/object/boreal:114667.

Quan, Zhiyu, and Emiliano A. Valdez. 2018. “Predictive Analytics of Insurance Claims Using Multivariate Decision Trees.” *Dependence Modeling* 6 (1): 377–407. https://doi.org/10.1515/demo-2018-0022.

Shi, Peng, and Emiliano A. Valdez. 2014a. “Longitudinal Modeling of Insurance Claim Counts Using Jitters.” *Scandinavian Actuarial Journal* 2014 (2): 159–79. https://doi.org/10.1080/03461238.2012.670611.

———. 2014b. “Multivariate Negative Binomial Models for Insurance Claim Counts.” *Insurance: Mathematics and Economics* 55 (March):18–29. https://doi.org/10.1016/j.insmatheco.2013.11.011.

Sutradhar, Brajendra C., and Vandna Jowaheer. 2003. “On Familial Longitudinal Poisson Mixed Models with Gamma Random Effects.” *Journal of Multivariate Analysis* 87 (2): 398–412. https://doi.org/10.1016/s0047-259x(03)00062-9.

Wedderburn, R. W. M. 1974. “Quasi-Likelihood Functions, Generalized Linear Models, and the Gauss—Newton Method.” *Biometrika* 61 (3): 439–47. https://doi.org/10.1093/biomet/61.3.439.

Yang, Lu, and Peng Shi. 2018. “Multiperil Rate Making for Property Insurance Using Longitudinal Data.” *Journal of the Royal Statistical Society Series A: Statistics in Society* 182 (2): 647–68. https://doi.org/10.1111/rssa.12419.