BIVARIATE POISSON CREDIBILITY MODEL AND BONUS-MALUS SCALE FOR CLAIM AND NEAR-CLAIM EVENTS

Pierre-Alexandre Simon, Julien Trufin, Michel Denuit

LIDAM Discussion Paper ISBA
2023 / 14

ISBA

Voie du Roman Pays 20 - L1.04.01

B-1348 Louvain-la-Neuve

Email : lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/isba/publication.html

BIVARIATE POISSON CREDIBILITY MODEL AND BONUS-MALUS SCALE FOR CLAIM AND NEAR-CLAIM EVENTS

Pierre-Alexandre Simon
Department of Mathematics
Université Libre de Bruxelles (ULB)
Brussels, Belgium

Julien Trufin
Department of Mathematics
Université Libre de Bruxelles (ULB)
Brussels, Belgium

Michel Denuit
Institute of Statistics, Biostatistics and Actuarial Science - ISBA
Louvain Institute of Data Analysis and Modeling - LIDAM
UCLouvain
Louvain-la-Neuve, Belgium

May 3, 2023

Abstract

With the emergence of telematics, huge amounts of data have become available to insurers, raising the question of their integration into motor insurance pricing. Guillen (2021) and Sun (2022) introduced the notion of near-claim events, corresponding to dangerous events, which could have triggered a claim. These events typically correspond to harsh deceleration, acceleration or cornering. Based on the motor third-party liability insurance portfolio of an insurance company operating in Belgium, we highlight that deceleration events are the best candidates for characterizing near-claim events. Then, we propose to integrate these events in the pricing structure in addition to policyholder's claim history with the help of a bivariate Poisson-LogNormal credibility model. Finally, we design an original bonus-malus scale based on both claim and near-claim events.

Keywords: Telematics, Motor insurance pricing, Near-claim, Gradient Boosting Machine, Credibility theory, Bonus-Malus System.

1 Introduction

Premium calculation in motor insurance routinely uses explanatory variables including driver's profile (age, place of residence, etc.) or vehicle characteristics (power, make and model, etc.). In addition to these classical features, telematics data have now started to be collected by insurance companies, offering new opportunities for risk evaluation. These data are typically recorded by telematics boxes installed in each vehicle or policyholder's smartphone. They typically include distance traveled, speed or acceleration measured in real time.

As the telematics offer is nowadays not yet predominantly deployed, a standard pricing model is still missing. Several proposals have been made in the literature. A first research stream concentrates on the construction of new explanatory variables based on the processing of telematics signals, to supplement traditional risk factors. Gao et al. (2019) constructed driving scores from speed-acceleration heatmaps and included them as additional explanatory variables for claim frequency prediction. Gao et al. (2022b) applied a combination of feed-forward and convolutional neural networks to process these heatmaps. So et al. (2021) proposed a three-way simulation procedure to generate a telematics portfolio using SMOTE technique in conjunction with feed-forward neural networks. The analysis conducted by Gao et al. (2019) and Duval et al. (2022) suggests that only three months or 4,000 kilometers of driving experience are enough to characterize the driving style of a given policyholder. This illustrates the richness of telematics data. We refer the reader to Gao et al. (2022a) for further information about the pre-processing techniques of driving data.

Another approach consists in modeling telematics signals as supplementary response variable along claim counts. As part of this approach, Guillen et al. (2020) and Sun et al. (2021) defined different types of near-miss events (term borrowed to the aviation safety) such as excess speed, harsh acceleration, harsh deceleration or cornering events. They highlighted that significant risk factors (traditional and telematics) depend on the type of near-miss events under consideration. Both studies concluded that considering near-miss events all-together leads to misleading interpretation as all near-misses are not necessarily associated with a higher claim frequency.

A last stream of research assumes that the prediction of claim frequency based on traditional rating factors has to be adjusted based on policyholder's history including past claims and telematics signals. This echoes multiple-event experience rating systems initially proposed to combine several products or to integrate claim severities into premium updates, in a simplified way. See e.g. Dionne and Vanasse (1992) , Pinquet (1998), Gomez-Deniz (2016), Park et al. (2018), Cheung et al. (2021), Denuit and Lu (2021), Tzougas and Cerchiara (2021) and Verschuren (2022), as well as the references therein. To recognize that telematics data are recorded while the policyholders are driving, exactly as claim experience, Denuit et al. (2019) and Corradin et al. (2022) proposed to use multivariate credibility models (also called mixed models, in statistics) to capture the joint dynamics of telematics data and claim experience. Past telematics signals are not included in the score like ordinary risk factors but used for premium correction in experience rating. This is the approach followed in this paper.

Even if credibility models are powerful tools to integrate driving experience in future premiums, motor insurance generally favors their commercial simplification in the form of bonus-malus scales. These scales comprise a number of levels and policyholders' movements depend on the number of claims they reported to the insurer. The relativities associated to each level correspond to the relative corrections applied to the individual basis premium. Bonus-malus scales are effectively modeled as Markov Chains and relativities are generally computed from the corresponding stationary distribution. In this paper, we propose a novel scale, integrating bot claim and near-claim events in a simple and transparent way and we adapt the approach proposed by Pitrebois et al. (2003) to derive relativities.

The analysis conducted in this paper is based on the motor third-party liability (MTPL) insurance portfolio of an insurance company operating in Belgium, observed over the period 2017-2019. The remainder of the text is organized as follows. Sections 2 and 3 describe the MTPL data set. We first consider several candidates to define appropriate near-claim events. Then, we correct frequencies of claims and near-claim events for standard features used in motor insurance pricing, using Gradient Boosting Machine (GBM) in Section 3. In Section 4, we design a bivariate Poisson-LogNormal credibility model combining claim and near-claim events. The bivariate model is compared to the univariate credibility model only based on reported claims. Section 5 proposes a bonus-malus scale integrating claim and near-claim events. The final Section 6 briefly discusses the results obtained in this paper as well as possible extensions.

2 Near-claim events

2.1 Descriptive analysis

The telematic database considered in this paper gathers logs of vehicle events recorded during policyholders trips over the period 2017-2019. These events are typically harsh deceleration (DE), acceleration (AE), quick-lateral-movement (QLM), cornering (CE) and speed excess (SE) events. Together with event type, we retrieve the gps quality, distance traveled, instantaneous speed or acceleration. The information is aggregated on a daily basis by considering the count of each type of events and maximum signals values recorded for each type of event. The aggregated data set contains records from 2017-01-01 to 2019-07-15. It has been divided into two parts: one training part (called train set) containing data for 2017 and 2018 and a testing part (called test set) containing 2019 half-year data. Among all observations, the majority of records relate to CE events (130,665,548) followed by DE events (80,653,333), AE events (48,291,158), QLM events (40,065,254) and finally SE events (9,371,138).

2.2 Definition

Following Guillen et al. (2021), we propose to overcome the relatively low claim frequency in motor insurance by identifying relevant near-claim events. Since we consider MTPL insurance, we focus our study only on claims at fault.

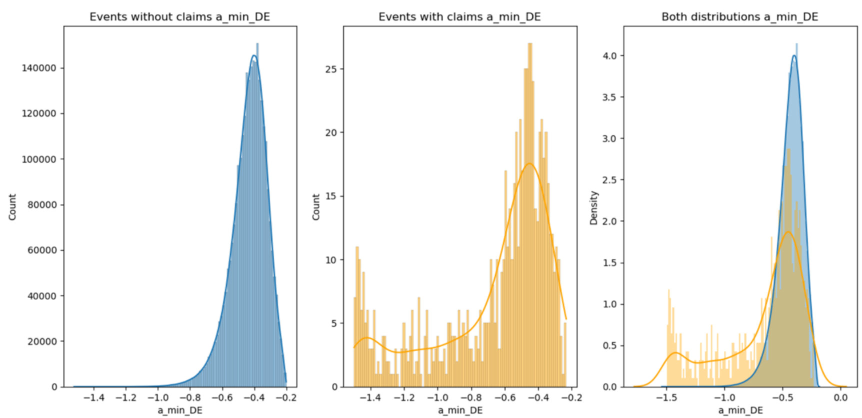


Figure 2.1: Minimum acceleration distribution for DE events (days without accident in blue and days with accident in yellow).

By inspecting marginal distributions of telematics signals separately for policyholders with and without claim at fault in Figures 2.1-2.2-2.3, we notice that a clear divergence is only observable for deceleration events. Thus, we decide to focus on policyholders the day of claim occurrence. Figures 2.4 and 2.5 displays the observed data available about two policyholders in the database. We observe that for each policyholder, we retrieve respectively 2 hours and 1 hour before the declared claim occurrence time a harsh deceleration, which suggests that the claim actually occurs at this particular time. This suggests that deceler-

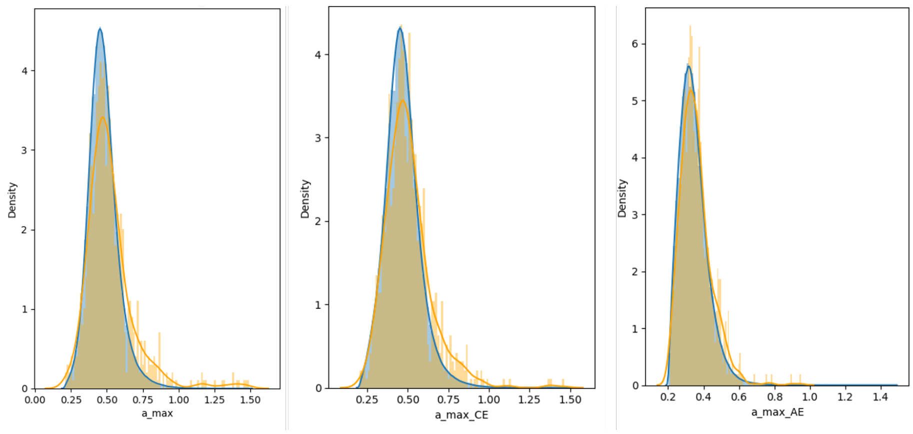


Figure 2.2: Maximum acceleration distribution for drivers without (blue) and with claim (orange) for all events (left), CE events (middle) and AE events (right).

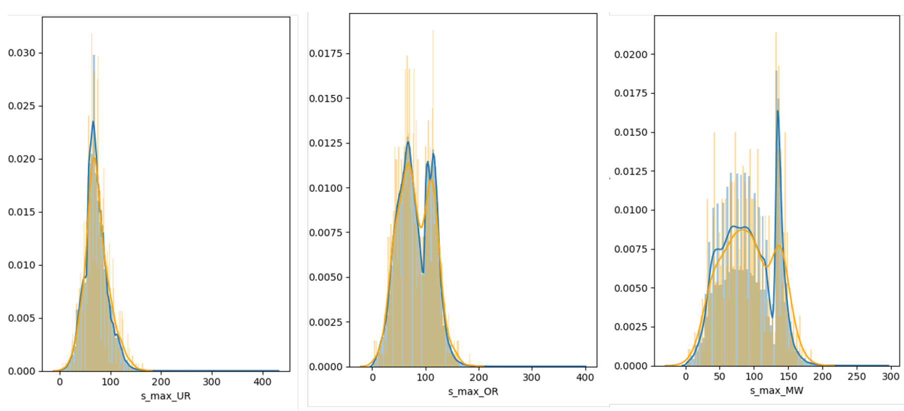


Figure 2.3: Maximum speed distribution for drivers without (blue) and with claim (orange) on urban roads (left), county roads (middle) and motorways (right).

ation events correspond to claim-at-fault occurrence because of late reaction in front of a danger for instance. It is also worth mentioning that more than half of the claim days are exhibiting a harsh deceleration.

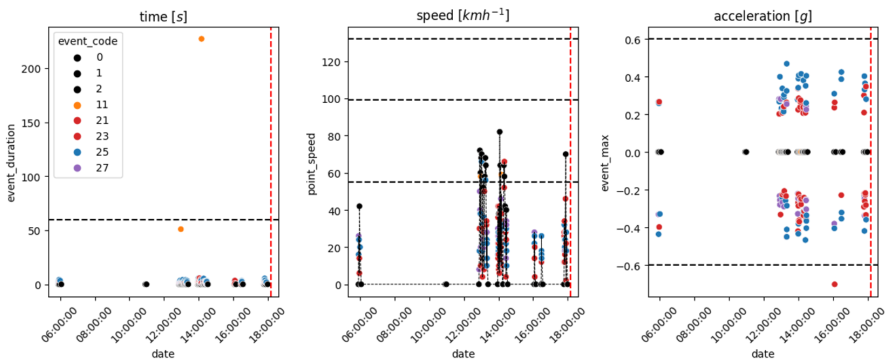


Figure 2.4: Telematics signals of a policyholder exhibiting harsh deceleration the day of claim occurrence.

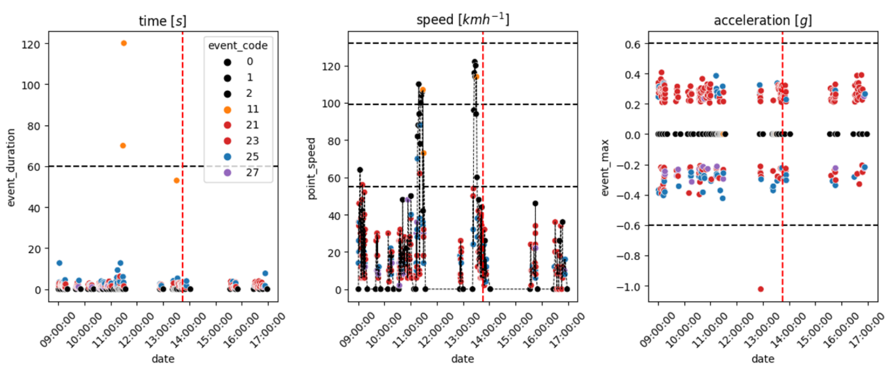


Figure 2.5: Telematics signals of a policyholder exhibiting harsh deceleration the day of claim occurrence.

Speed excess and acceleration events turn out to be more related to a specific driving attitude rather than directly to claim-at-fault occurrence. Moreover, this information is partially included in deceleration signals as harsh decelerations are more likely to happen if the policyholder drives at high speed and in an aggressive way. Therefore, we proposed to consider exclusively deceleration events in our definition of near-claim events.

The next step consists in determining the appropriate threshold to qualify as harsh deceleration. Guillen et al. (2021) suggest that some DE events (with intensity 1 and 3) are associated with higher claim frequency and some (with intensity 2) with lower claim frequency where DE events characterization is carried out by the telematics manufacturer. Guillen et al. (2020) consider the same absolute threshold of 0.6g for characterizing DE and AE events as near-miss events. However, based on our preliminary study, we see that empirical distributions exhibit some divergences suggesting that a proper threshold for DE events should be considered. From accident analysis and prevention literature, Chevalier et al. (2016) demonstrate the association between contrast sensitivity decline for old drivers and Rapid Deceleration Event (RDE). These events are characterized as such for a deceleration rate lower than -0.75g, corresponding to a 26.5km/h speed reduction in one second, and used as a surrogate safety measure. Fu and Sayed (2021) and Zheng et al. (2018) study the Deceleration Rate to Avoid a Crash (DRAC). They point out that the Maximum Available Deceleration Rate (MADR) above which the crash is unavoidable, is depending on vehicles weight, tire, braking systems and road pavement condition. They recommend to use the mean of the truncated Normal distribution (with mean 8.45 and standard deviation 1.402) as a threshold (in ms^{-2}) for crash occurrence. De Ceunynck (2017) advocate -0.8g as a conservative value for maximum deceleration rate that all automobiles can achieve in case of braking emergency. In our study, as the telematics portfolio contains heterogeneous vehicles, we decide to use this threshold for characterizing a particular DE event as a near-claim event, corresponding to a braking emergency situation. Table 2.1 describes the number of near-claim events and claims in the two respective data sets.

Dataset	Training	Testing
Nb of near-claim events	29,167	6,765
Nb of claims	2,336	435
Nb of claims at fault	1,111	234
Nb of claims not at fault	1,225	201
Exposure (km)	258,585,950	57,335,704

Table 2.1: Response variables statistics.

3 Risk classification

We denote by $N_{i,t}$ (resp. $Z_{i,t}$) the number of claims at fault (resp. the number of near-claim events) recorded for policyholder i, $i = 1,2,\dots,n$, during period t, $t = 1,2,\dots,T_i$, and by $d_{i,t}$ the corresponding exposure-to-risk, that can be the distance driven in kilometers or time spent behind the wheel in the context of telematics data. Boucher et al. (2017) and Ayuso et al. (2019) documented the relationship between distance driven and expected claim frequency. In this paper, we use the distance driven in kilometers as exposure-to-risk. At the beginning of each insurance period, the actuary has at his or her disposal some information about each policyholder summarized into p features $x_{i,t,j}$, $j = 1, \ldots, p$. The a priori information $\boldsymbol{x}_{i,t} = (x_{i,t,1}, \ldots, x_{i,t,p})$ is integrated into scores $\eta_{i,t} = \eta(\boldsymbol{x}_{i,t})$ for the number of claims at fault and $\nu_{i,t} = \nu(\boldsymbol{x}_{i,t})$ for the number of near-claim events for policyholder i in period t (taken to be a calendar year). Then, the expected number of claims at fault and near-claim events are

$$\mathrm{E}[N_{i,t}|X_{i,t} = x_{i,t}] = d_{i,t} \exp\big(\eta_{i,t}\big)$$

and

$$\mathrm{E}[Z_{i,t}|X_{i,t} = x_{i,t}] = d_{i,t} \exp\big(\nu_{i,t}\big),$$

respectively. Scores are estimated with the Gradient Boosting Machine (GBM) algorithm developed by Friedman (2008). To this end, the data set has been aggregated by policyholder on a yearly basis. The explanatory variables entering the score are described in Appendix A. Notice that exposure has been restrained in the range $[100km, 40,000km]$ to avoid to large volatility in the predicted expected frequencies.

3.1 Model calibration

In order to assess the robustness of each model, we perform 5-fold stratified cross validation. Each fold is designed so that the empirical probability mass function of the number of claims at fault (resp. near-claim events) remains similar. Table 3.1 displays the exposures (number of kilometers driven) on each fold as well as the claim (resp. near-claim) numbers per 1,000 km driven.

Fold	S1	S2	S3	S4	S5
Nb of rows	7707	7706	7708	7708	30829
Exposure (km)	51,413,880	52,122,710	52,345,880	52,580,810	50,867,180
Mean observed claim frequency (per 1000km)	0.0051	0.0050	0.0050	0.0050	0.0051
Mean observed near-claim frequency (per 1000km)	0.1099	0.1085	0.1079	0.1075	0.1096

Table 3.1: Stratified cross validation.

For tuning the GBM on each fold, we consider a bag fraction equals to 0.5. The number of trees is closely tied with the shrinkage parameter retained. By considering a standard shrinkage parameter $\nu = 0.01$, a preliminary study indicated that the minimum of the Poisson deviance on a out-of-sample set containing 20% of the training set observations was reached for a number of trees lower than $20,000$. Thus, we retain this value for the maximum number of trees M. We also observe that the minimum of the mean Poisson deviance was obtained for a lower number of trees for claim frequency model ($< 7,500$) than for near-claim frequency model ($< 15,000$). As a consequence, we consider $\nu = 0.01$ for near-claim model and $\nu = 0.001$ for claim model, letting the maximum number of trees sufficiently large to obtain the minimum of the mean Poisson deviance on the interval $[0; 20,000]$. As the minimum number of observations per leaf constrains the interaction depth of the trees, we decide to set it to one while letting the interaction depth vary in the range $[1, 10]$. Once the GBM tuning was performed on each fold of the stratified data set for claim and near-claim counts, the mean Poisson deviance for each interaction depth is computed over the different folds as depicted in Figure 3.1. This leads to the best model parameters selection for each model, which are summarized in Table 3.2.

Parameters	Claim GBM	Near-claim GBM
Shrinkage	0.001	0.01
Bag fraction	0.5	0.5
Minimum observations per leaf	1	1
Interaction depth	2	8
Number of trees	6,291	16,483

Table 3.2: Optimal GBM parameters.

3.2 Model interpretation

3.2.1 Importance of explanatory variables

For each best claim (resp. near-claim) frequency model trained on each fold, the respective importance of explanatory variables is represented in Figure 3.2. For both frequency models, the two most important variables responsible of more than 10% of the Poisson deviance reduction are the vehicle make followed by the main driver age. For the other variables responsible of more than 5% of the Poisson deviance reduction, we retrieve common explanatory variables for both models such as the geographical coordinates of the place of residence, vehicle value, geo mosaic code and the vehicle power and weight. The fact that both GBM models have been extensively using the same explanatory variables suggests similarities between each response variable. The main difference between both models are vehicle age variable and vehicle weight, whose relative importance order has been interchanged between the two models.

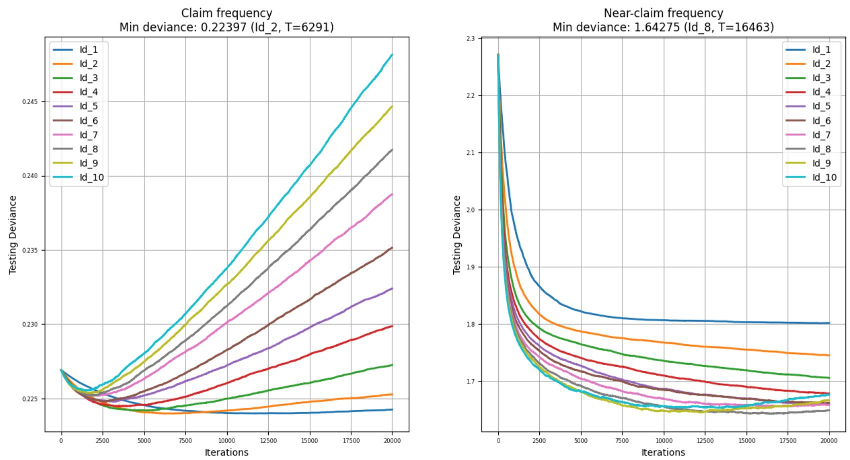


Figure 3.1: Mean Poisson deviance observed for different interaction depths for claim frequency models (left) and near-claim frequency models (right).

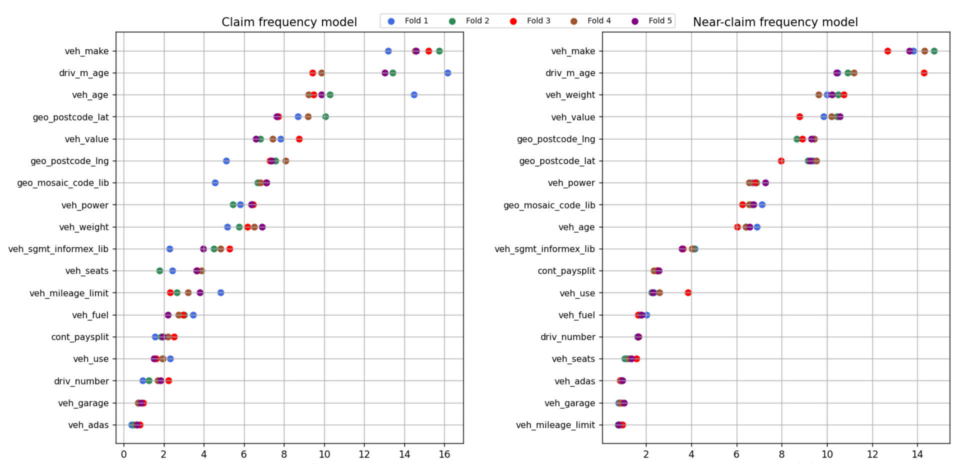


Figure 3.2: Explanatory variables importance for claim frequency model (left) and near-claim frequency model (right) on each of the five folds.

3.2.2 Partial dependencies

From the importance of the explanatory variables on each response variable identified in the previous section, we represent the effect of the most important variables on the predicted claim frequency (resp. near-claim frequency) in Figure 3.3. We choose to use the Accumulated Local Effect (ALE) plot of Apley and Zhu (2016) to isolate the interaction effect of the other explanatory variables on the dependence between the studied explanatory variable and the response. Notice that the ALE plot is centered at zero expressing the difference in expected claim (resp. near-claim) frequency with the respective mean expected frequency on the entire portfolio. For the vehicle make, partial dependence with expected claim frequency exhibits high volatility, mainly produced by outliers. This is due to the large number of vehicle makes related to the number of claims at fault present in the aggregated database. This prevents us to draw any robust conclusion about an eventual similarity in the dependence of some vehicle makes with claim frequency and those with expected near-claim frequency. However, we observe that dependence of the numerical explanatory variables on each response variable disclose the same behavior. Guillen et al. (2020) already emphasized that driver age was associated with lower near-miss frequency. Interestingly, we observe that vehicle age and weight are associated in both cases with higher (resp. lower) expected frequency. We propose in the following section to incorporate a posteriori individual claim and near-claim history.

4 Bivariate credibility model

Random effects Δ_i and Γ_i are added to the scores $\eta_{i,t}$ and $\nu_{i,t}$, respectively, to recognize the residual heterogeneity of the portfolio. Precisely, we assume that

$$\mathrm{E}[N_{i,t}|X_{i,t} = x_{i,t}, \Delta_i] = \frac{d_{i,t} \exp\left(\eta_{i,t} + \Delta_i\right)}{\mathrm{E}[\exp(\Delta_i)]}$$

and

$$\mathrm{E}[Z_{i,t}|X_{i,t} = x_{i,t}, \Gamma_i] = \frac{d_{i,t} \exp\left(\nu_{i,t} + \Gamma_i\right)}{\mathrm{E}[\exp(\Gamma_i)]}.$$

Let us denote

$$\lambda_{i,t} = \frac{d_{i,t} \exp\left(\eta_{i,t}\right)}{\mathrm{E}[\exp(\Delta_i)]}$$

and

$$\mu_{i,t} = \frac{d_{i,t} \exp\left(\nu_{i,t}\right)}{\mathrm{E}[\exp(\Gamma_i)]}.$$

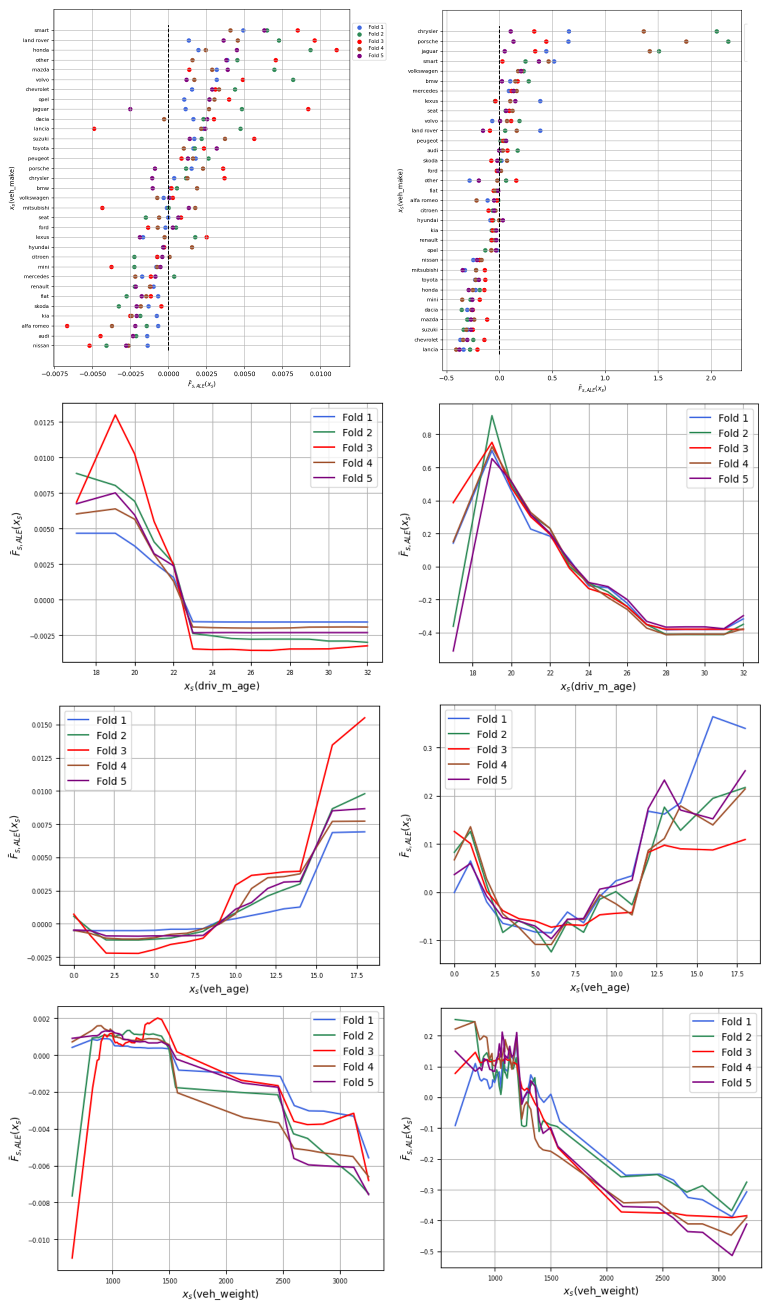


Figure 3.3: Most important explanatory variables partial dependence (ALE plot) for claim frequency model (left panel) and near-claim frequency model (right panel) on each of the five folds.

Given $\Delta_i = \delta$, the random variables $N_{i,t},\ t = 1,2,\ldots$, are assumed to be independent and Poisson distributed with means $\lambda_{i,t}\exp(\delta)$. At the portfolio level, the sequences $(\Delta_i, N_{i,1}, N_{i,2},\ldots)$, $i = 1,2,\ldots,n$, are assumed to be mutually independent. Similarly, given $\Gamma_i = \gamma$, the random variables $Z_{i,t},\ t = 1,2,\ldots$, are supposed to be independent and Poisson distributed with means $\mu_{i,t}\exp(\gamma)$, and the sequences $(\Gamma_i, Z_{i,1}, Z_{i,2},\ldots)$, $i = 1,2,\ldots,n$, are assumed to be mutually independent. Moreover, given Δ_i, claim counts $N_{i,1}, N_{i,2},\ldots$ are assumed to be independent of $\Gamma_i, Z_{i,1}, Z_{i,2},\ldots$, and given Γ_i, near-claim counts $Z_{i,1}, Z_{i,2},\ldots$ are assumed to be independent of $\Delta_i, N_{i,1}, N_{i,2},\ldots$. The random vectors (Δ_i,Γ_i), $i = 1,2,\ldots,n$, are supposed to be mutually independent and to obey the bivariate Normal distribution with zero mean, variances σ^2_Δ and σ^2_Γ and correlation $\rho_{\Delta,\Gamma}$, so that $\mathrm{E}[\exp(\Delta_i)] = \exp\big(\sigma^2_\Delta/2\big)$ and $\mathrm{E}[\exp(\Gamma_i)] = \exp\big(\sigma^2_\Gamma/2\big)$. This is henceforth denoted as $\mathscr{N}(\mathbf{0},\mathbf{\Sigma})$ with variance covariance matrix

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma^2_\Delta & \rho_{\Delta,\Gamma}\sigma_\Delta\sigma_\Gamma \\ \rho_{\Delta,\Gamma}\sigma_\Delta\sigma_\Gamma & \sigma^2_\Gamma \end{pmatrix},$$

and corresponding probability density function $f_{\mathbf{\Sigma}}(\cdot,\cdot)$. Unless $\rho_{\Delta,\Gamma} = 0$, the number of near-claim events will bring information about the number of claims.

4.1 Estimation of the covariance matrix

The likelihood associated to the observations (k_{it}, l_{it}) writes

$$\begin{aligned}
\mathscr{L} &= \prod_{i=1}^{n}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\prod_{t=1}^{T_i} e^{-\lambda_{it}e^{\delta}}\frac{(\lambda_{it}e^{\delta})^{k_{it}}}{k_{it}!}e^{-\mu_{it}e^{\gamma}}\frac{(\mu_{it}e^{\gamma})^{l_{it}}}{l_{it}!}\right) f_{\mathbf{\Sigma}}(\delta,\gamma)\,d\delta\,d\gamma \\
&= \prod_{i=1}^{n}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\prod_{t=1}^{T_i}\frac{\lambda_{it}^{k_{it}}\mu_{it}^{lit}}{k_{it}!l_{it}!}\right)\frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} e^{\delta\sum_{t=1}^{T_i}k_{it}+\gamma\sum_{t=1}^{T_i}l_{it}-\sum_{t=1}^{T_i}\lambda_{it}e^{\delta}-\sum_{t=1}^{T_i}\mu_{it}e^{\gamma}-\frac{1}{2}(\delta,\gamma)\Sigma^{-1}(\delta,\gamma)^\top}\,d\delta\,d\gamma \\
&= \prod_{i=1}^{n}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\left(\prod_{t=1}^{T_i}\frac{\lambda_{it}^{k_{it}}\mu_{it}^{lit}}{k_{it}!l_{it}!}\right)\frac{1}{2\pi|\Sigma|^{\frac{1}{2}}} e^{\delta k_{i\bullet}+\gamma l_{i\bullet}-\lambda_{i\bullet}e^{\delta}-\mu_{i\bullet}e^{\gamma}-\frac{1}{2}(\delta,\gamma)\Sigma^{-1}(\delta,\gamma)^\top}\,d\delta\,d\gamma,
\end{aligned}$$

where "$\bullet$" indicates a summation over time, like $k_{i\bullet} = \sum_{t=1}^{T_i}k_{it}$ for instance. Moreover, we have $|\mathbf{\Sigma}| = \sigma^2_\Delta\sigma^2_\Gamma - \rho^2_{\Delta,\Gamma}\sigma^2_\Delta\sigma^2_\Gamma$. The variance-covariance matrix $\mathbf{\Sigma}$ can be inverted only if $\rho_{\Delta,\Gamma} \neq \pm 1$. If the latter condition holds then

$$\mathbf{\Sigma}^{-1} = \frac{1}{|\mathbf{\Sigma}|}\begin{pmatrix} \sigma^2_\Gamma & -\rho_{\Delta,\Gamma}\sigma_\Delta\sigma_\Gamma \\ -\rho_{\Delta,\Gamma}\sigma_\Delta\sigma_\Gamma & \sigma^2_\Delta \end{pmatrix}$$

and the exponential term can be expressed as

$$\begin{aligned}
m &= \begin{pmatrix}\delta & \gamma\end{pmatrix} \boldsymbol{\Sigma}^{-1} \begin{pmatrix}\delta \\ \gamma\end{pmatrix} \\
&= \frac{1}{|\boldsymbol{\Sigma}|} \begin{pmatrix}\delta & \gamma\end{pmatrix} \begin{pmatrix}\sigma_\Gamma^2 & -\rho_{\Delta,\Gamma}\sigma_\Delta\sigma_\Gamma \\ -\rho_{\Delta,\Gamma}\sigma_\Delta\sigma_\Gamma & \sigma_\Delta^2\end{pmatrix} \begin{pmatrix}\delta \\ \gamma\end{pmatrix} \\
&= \frac{1}{|\boldsymbol{\Sigma}|} \left(\sigma_\Gamma^2\delta^2 - 2\rho_{\Delta,\Gamma}\sigma_\Delta\sigma_\Gamma\delta\gamma + \sigma_\Delta^2\gamma^2\right) \\
&= \frac{\sigma_\Gamma^2\delta^2 - 2\rho_{\Delta,\Gamma}\sigma_\Delta\sigma_\Gamma\delta\gamma + \sigma_\Delta^2\gamma^2}{\sigma_\Delta^2\sigma_\Gamma^2 - \rho_{\Delta,\Gamma}^2\sigma_\Delta^2\sigma_\Gamma^2}.
\end{aligned}$$

The log-likelihood finally writes

$$\log \mathscr{L} = -\frac{n}{2}\log(|\boldsymbol{\Sigma}|) + \sum_{i=1}^{n} \log\left(\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\delta k_{i\bullet} + \gamma l_{i\bullet} - \lambda_{i\bullet} e^{\delta} - \mu_{i\bullet} e^{\gamma} - \frac{m}{2}} \, d\delta \, d\gamma\right) + \text{constant}.$$

The maximization of the log-likelihood is performed using Nelder-Mead optimization method where the double integral is computed at each step using the function `dblquad` of the `scipy` library. Due to important computational time for the optimization process on the training set 2017-2018, we proceed in two steps. The first step consists in first estimating the random effect standard deviations σ_Δ and σ_Γ with $\rho_{\Delta,\Gamma} = 0$. We get

$$\begin{cases} \widehat{\sigma}_\Delta^{(0)} = 0.61889801 \\ \widehat{\sigma}_\Gamma^{(0)} = 0.73366228. \end{cases}$$

Then, we estimate the correlation coefficient $\rho_{\Delta,\Gamma}^{(0)}$ with the estimated values $(\widehat{\sigma}_\Gamma^{(0)}, \widehat{\sigma}_\Gamma^{(0)})$ for the random effect standard deviations σ_Δ and σ_Γ setting the initial value for $\rho_{\Delta,\Gamma}^{(0)}$ to 0.1. The objective function nearby the optimal solution is displayed in Figure 4.1, which leads to $\widehat{\rho}_{\Delta,\Gamma}^{(0)} = 0.57071$.

In a second step, we again perform the optimization estimating all parameters considering as initial values $(\widehat{\sigma}_\Gamma^{(0)}, \widehat{\sigma}_\Gamma^{(0)}, \widehat{\rho}_{\Delta,\Gamma}^{(0)})$. The joint estimation leads to the following estimators:

$$\begin{cases} \widehat{\sigma}_\Delta^{(1)} = 0.71943756 \\ \widehat{\sigma}_\Gamma^{(1)} = 0.73501123 \\ \widehat{\rho}_{\Delta,\Gamma}^{(1)} = 0.54464738. \end{cases}$$

The pretty large estimated correlation coefficient indicates that near-claim events bring relevant information about claims at fault.

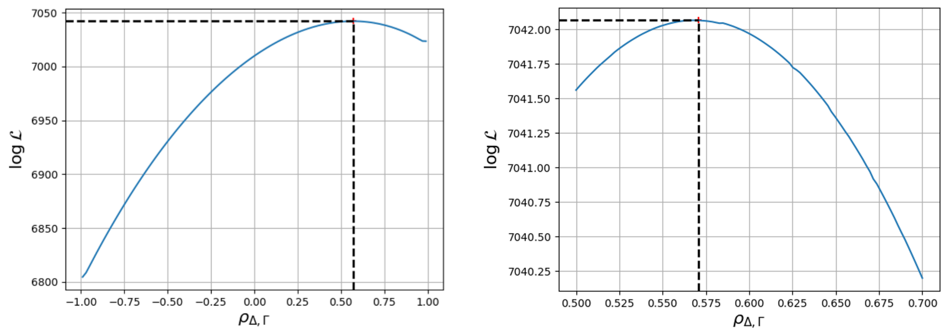


Figure 4.1: Maximum likelihood estimator $\widehat{\rho}^{(0)}_{\Delta,\Gamma}$ in the entire solution space (left) and nearby optimal solution (right).

4.2 Corrections based on experience

4.2.1 Univariate credibility model

In the univariate case, the expectation of the number of claims at fault for the year $T_i + 1$ given the claim history $\mathscr{H}_{T_i} = \{k_{i1}, \dots, k_{iT_i}\}$ is given by

$$\mathrm{E}[N_{iT_i+1}|\mathscr{H}_{T_i}] = \lambda_{iT_i+1}\mathrm{E}[\exp(\Delta_i)|\mathscr{H}_{T_i}]$$

where

$$\mathrm{E}[\exp(\Delta_i)|\mathscr{H}_{T_i}] = \frac{\int_{-\infty}^{+\infty} e^{-\lambda_{i\bullet}e^{\delta}+\delta(k_{i\bullet}+1)-\frac{\delta^2}{2\sigma_\Delta^2}}\,d\delta}{\int_{-\infty}^{+\infty} e^{-\lambda_{i\bullet}e^{\delta}+\delta k_{i\bullet}-\frac{\delta^2}{2\sigma_\Delta^2}}\,d\delta}.$$

4.2.2 Multivariate case

In the multivariate case, the expectation of the number of claim for the year T_i+1 given the claim and near-claim history $\mathscr{H}^*_{T_i} = \{k_{i1}, \dots, k_{iT_i}, l_{i1}, \dots, l_{iT_i}\}$ is given by

$$\mathrm{E}[N_{iT_i+1}|\mathscr{H}^*_{T_i}] = \lambda_{iT_i+1}E[\exp(\Delta_i)|\mathscr{H}^*_{T_i}]$$

where

$$\mathrm{E}[\exp(\Delta_i)|\mathscr{H}^*_{T_i}] = \frac{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\delta(k_{i\bullet}+1)+\gamma l_{i\bullet}-\lambda_{i\bullet}e^{\delta}-\mu_{i\bullet}e^{\gamma}-\frac{m}{2}}\,d\delta d\gamma}{\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{\delta k_{i\bullet}+\gamma l_{i\bullet}-\lambda_{i\bullet}e^{\delta}-\mu_{i\bullet}e^{\gamma}-\frac{m}{2}}\,d\delta d\gamma},$$

4.2.3 Comparison

Let us consider four different profiles corresponding to two different risk levels for claim and near-claim predicted frequencies, namely

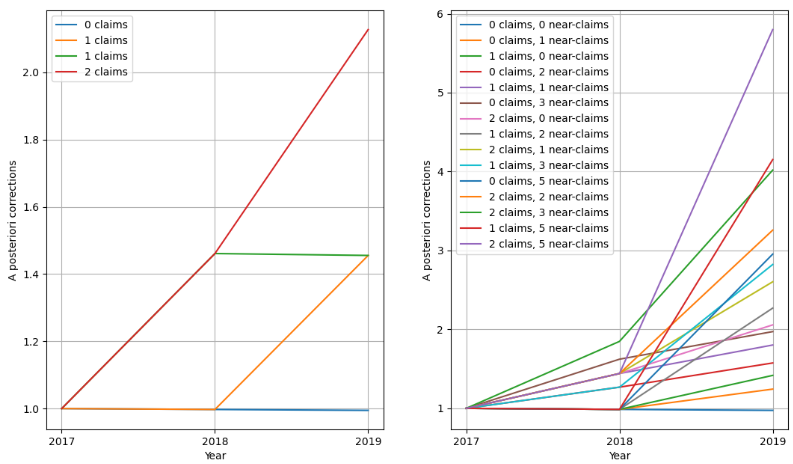
- Profile 1: Low claim and near-claim frequency;
- Profile 2: Low claim and high near-claim frequency;
- Profile 3: High claim and low near-claim frequency;
- Profile 4: High claim and near-claim frequency.

Low (resp. high) frequency profile corresponds to the 1/6-quantile (resp. 5/6-quantile) of the distribution of fitted frequencies. For each profile, we let the number of claims and near-claim events vary and compute the corresponding corrections based on experience according to univariate and bivariate credibility models. The latter are displayed in Figure 4.2 with claim numbers $\{0,1,2\}$ and near-claim numbers $\{0,1,2,3,5\}$. Corrections are applied for 2018 when 2017 history becomes available and for 2019 based on 2017 – 2018 history. For the univariate model, if the policyholder claim history is equal to 1, we represent the case where the claim occurred in 2017 and the case where the claim occurred in 2018, which leads mechanically to the same correction for 2019. For the near-claim history, we did not consider all possible combinations leading to the same near-claim history to not overload figures but corrections would be also identical as the random effects Δ_i and Γ_i do not depend on time. Figure 4.2 clearly shows that near-claim events considerably enrich corrections based on experience, resulting in a much finer grid of reevaluations at the end of the period.

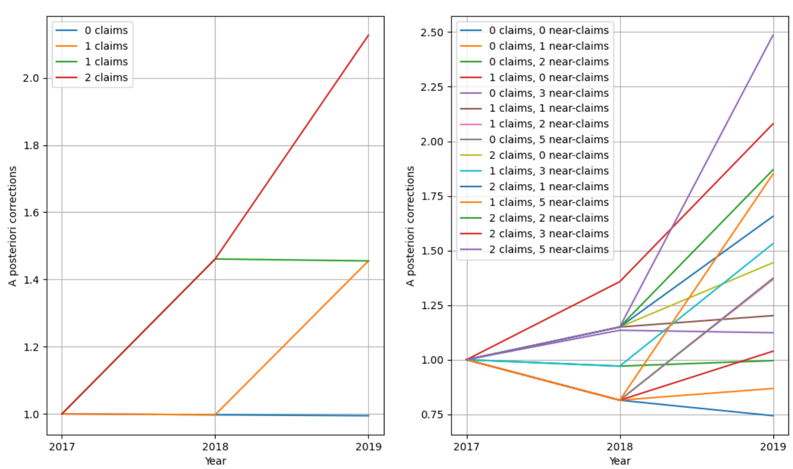
We observe that Profile 1 experiences the highest correction associated with an history of 2 claims at fault and 5 near-claim events observed whereas Profile 4 experiences the lowest correction associated with an empty history of claims at fault and near-claim events irrespective of the univariate or bivariate model. This reflects the fact that profile with low (resp. high) risk level and a contradicting claim and near-claim history will naturally experience the largest modification of their a priori frequency-premium. Note also that for Profiles 1 and 2, the univariate correction is slightly softer. Moreover, we can see that the bivariate model better rewards policyholder than the univariate model in case of empty near-claim history.

5 Bonus-malus scale with near-claim events

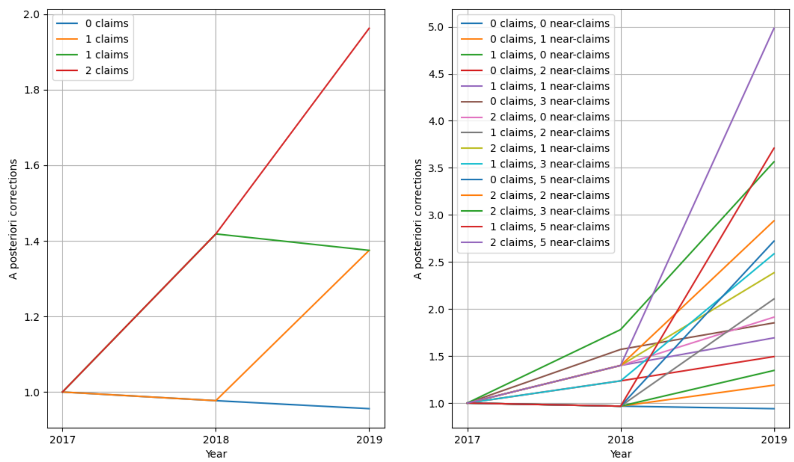
Bonus-malus scales are commercial simplifications to credibility corrections. A bonus-malus scale consists of a finite number of levels, each with its own relative premium called relativity. The relativity associated to level ℓ is denoted as r_ℓ. The premium charged is then obtained by multiplying by r_ℓ% a base premium depending on policyholder's characteristics.



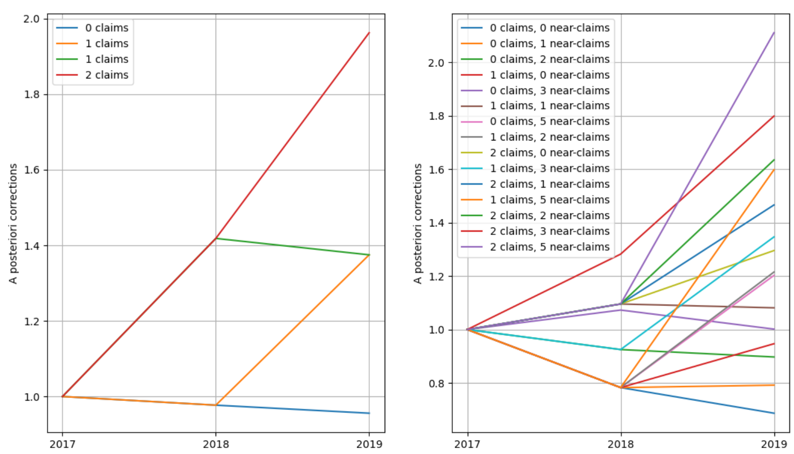
(a) Profile 1.



(b) Profile 2.



(c) Profile 3.



(d) Profile 4.

Figure 4.2: Experience corrections resulting from univariate (left) and bivariate (right) credibility models for each of the four risk profiles under consideration.

Whereas classical bonus-malus scales are based on a single type of event (occurrence of claims at fault), this section integrates near-claim events as a second type of event. Notice that systems based on several types of events are commonly encountered in North America where elements of the policyholders' driving record also impact on insurance premiums. For instance, safe driver insurance plans in the US encourage safe driving by rewarding drivers who do not cause an accident, nor incur a traffic law violation. Several types of events (major and minor at-fault accidents and traffic violations) each entail specific penalties: policyholders move up a certain number of levels in the bonus-malus scale according to the type of incident. Compared to Pitrebois et al. (2006) who developed multiple-event bonus-malus scales by assuming a Multinomial partitioning scheme, we work here with the bivariate credibility model described in Section 4. Since it seems to be difficult in practice to penalize drivers for near-claim events, we adapt the -1/top scale in Pitrebois et al. (2003) to account for both claim and near-claim counts. This scale is a pure no-claim discount scheme classifying drivers according to the number claim-free periods. It is commonly encountered in North America. Here, we adapt it by letting the annual discount depend on the near-claim events, not the penalty. This is precisely explained next.

The -1/top bonus-malus scale under consideration has $m+1$ levels (numbered 0 to m, where m denotes the maximum level). Policyholders are classified according to the number of claim-free periods since the last event (0, 1, ..., $m-1$ or at least m). New policyholders enter the system at the top level m where they pay the base premium (so that $r_m = 100\%$). Each event-free period (no claim reported nor near-claim event recorded) is rewarded by one bonus class. If no claim is reported but at least one near-claim event is recorded then the policyholder stays in his or her current level, thus losing the discount. In case a claim is reported, all the discounts are lost and the policyholder is transferred to level m.

Policyholder's trajectory in the scale is modeled by a sequence $\{L_t, \quad t = 0, 1, \ldots\}$ of random variables valued in $\{0, 1, \ldots, m\}$, with $L_0 = m$. Movements in the scale occur on a regular basis (yearly or monthly, say) and the policyholder occupies level L_t from time t until time $t+1$. Once the number N_t of reported claims at fault during $(t-1, t)$ and the number Z_t of near-claim events are both known, this information is used to reevaluate the position in the scale. Random variables L_t obey to the recursion

$$L_t = \begin{cases} \max\{L_{t-1}-1,0\} \text{ if } N_t = Z_t = 0 \\ \\ L_{t-1} \text{ if } N_t = 0 \text{ and } Z_t \geq 1 \\ \\ m \text{ if } N_t \geq 1. \end{cases}$$

This stochastic recursive equation shows that the future trajectory in the scale is independent of the levels occupied in the past, provided that the present level is given. This conditional independence property is at the heart of Markov models.

We follow Pitrebois et al. (2003) to determine relativities r_ℓ associated with each level $\ell \in \{0,1,\ldots,m\}$. Let us drop the policy index i and let us consider a generic policyholder with $\Delta = \delta$ and $\Gamma = \gamma$, taken at random from the portfolio. Given $\Delta = \delta$ and $\Gamma = \gamma$, all random variables N_t and Z_t are mutually independent, obeying Poisson distributions with respective means denoted as λ and μ. Let $p_{\ell_1\ell_2}(\lambda,\mu)$ be the probability of moving from level ℓ_1 to level ℓ_2 for that policyholder, that is, the probability that $L_{t+1} = \ell_2$ given that $L_t = \ell_1$, with $\ell_1, \ell_2 \in \{0,1,\ldots,m\}$. These probabilities are gathered in the transition matrix $\boldsymbol{P}(\lambda,\mu)$ assumed to be regular (that is, there exists some power of $\boldsymbol{P}(\lambda,\mu)$ whose entries are all strictly positive). For the bonus-malus scale considered in this paper, we get

$$\boldsymbol{P}(\lambda,\mu) = \begin{pmatrix} \exp(-\lambda) & 0 & 0 & \cdots & 0 & 0 & 1-\exp(-\lambda) \\ \exp(-\lambda-\mu) & \exp(-\lambda)(1-\exp(-\mu)) & 0 & \cdots & 0 & 0 & 1-\exp(-\lambda) \\ 0 & \exp(-\lambda-\mu) & \exp(-\lambda)(1-\exp(-\mu)) & \cdots & 0 & 0 & 1-\exp(-\lambda) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \exp(-\lambda-\mu) & \exp(-\lambda)(1-\exp(-\mu)) & 1-\exp(-\lambda) \\ 0 & 0 & 0 & \vdots & 0 & \exp(-\lambda-\mu) & 1-\exp(-\lambda-\mu) \end{pmatrix}.$$

The probability $p^{(k)}_{\ell_1\ell_2}(\lambda,\mu)$ that $L_{t+k} = \ell_2$ given $L_t = \ell_1$ of being transferred from level ℓ_1 to level ℓ_2 in k steps is then obtained from the kth power of the one-step transition matrix $\boldsymbol{P}(\lambda,\mu)$.

It is natural to expect that the system stabilizes in the long run, in that each policyholder ultimately oscillates around an equilibrium level corresponding to the expected number λ of reported claims at fault and expected number μ of near-claim events recorded by the telematics device. Formally, define the stationary distribution $\pi_\ell(\lambda,\mu)$ as

$$\pi_{\ell_2}(\lambda,\mu) = \lim_{k\to\infty} p^{(k)}_{\ell_1\ell_2}(\lambda,\mu).$$

Here, $\pi_\ell(\lambda,\mu)$ can be regarded as the fraction of the time a policyholder with expected claim frequency λ and expected near-claim frequency μ spends in level

ℓ, once the steady state has been reached. The stationary distribution can be obtained by solving a system of linear equations. It can also be obtained by matrix calculation. Specifically, let $\mathbf{1}$ be a column vector with all components equal to 1 and let $\boldsymbol{E}$ be the $(m+1) \times (m+1)$ matrix consisting of $m+1$ column vectors $\mathbf{1}$. Denoting as $\boldsymbol{I}$ the identity matrix, the vector $\boldsymbol{\pi}(\mu)$ of stationary probabilities is then given by

$$\boldsymbol{\pi}^\top(\lambda,\mu) = \mathbf{1}^\top(\boldsymbol{I} - \boldsymbol{P}(\lambda,\mu) + \boldsymbol{E})^{-1}. \tag{5.1}$$

The number of levels can be selected so that a policyholder drawn at random from the portfolio reaches level 0 in a given number of periods, on average, or derived from some other meaningful condition. Notice that basing calculation of r_ℓ on the stationary distribution is a conservative strategy with the -1/top scale.

Let w_k be the weight of the kth risk class with expected claim frequency λ_k per period and expected near-claim frequency μ_k per period. A very accurate risk classification (like the one produced by GBM in this paper) can result in a large number of risk classes k to consider. Model points, or reference risk profiles can then be created to ease calculations (applying CARTs, for instance). In this section, we consider all risk profiles so that we actually sum over all policyholders $(k = i)$, taking individual exposure to risk as weight w_k.

Let L be the level occupied by a randomly selected policyholder once the steady state has been reached. The distribution of L can be written as

$$\mathrm{P}[L=\ell] = \sum_k w_k \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \pi_\ell\big(\lambda_k \exp(\delta), \mu_k \exp(\gamma)\big) f_{\boldsymbol{\Sigma}}(\delta,\gamma)\, d\delta\, d\gamma. \tag{5.2}$$

Here, $\mathrm{P}[L=\ell]$ represents the proportion of the policyholders in level ℓ. Relativities are then obtained from

$$r_\ell = \frac{\mathrm{E}[\exp(\Delta)|L=\ell]}{\mathrm{E}[\exp(\Delta)]} \tag{5.3}$$

where

$$\mathrm{E}[\exp(\Delta)|L=\ell] = \frac{\sum_k w_k \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp(\delta)\pi_\ell\big(\lambda_k \exp(\delta), \mu_k \exp(\gamma)\big) f_{\boldsymbol{\Sigma}}(\delta,\gamma)\, d\delta\, d\gamma}{\sum_k w_k \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \pi_\ell\big(\lambda_k \exp(\delta), \mu_k \exp(\gamma)\big) f_{\boldsymbol{\Sigma}}(\delta,\gamma)\, d\delta\, d\gamma}, \tag{5.4}$$

as demonstrated by Pitrebois et al. (2003). It is easily seen that $\mathrm{E}[r_L] = 1$, resulting in financial equilibrium once steady state is reached.

When implemented in practice, it is very likely that the relativity r_m associated to the highest level in the scale differs from 100%. It then suffices to adapt the intercept in the frequency model to re-normalize it by dividing all relativities by r_m and change the intercept β_0 into $\beta_0 + \ln r_m$ in the Poisson regression model producing frequency-premiums. Relativities r_ℓ can be made linear in ℓ to entail the same discount when policyholders get one level down.

The relativities based on the MTPL insurance portfolio considered in this paper are computed based on the risk classes produced by the GBM. More precisely, we consider the a priori claim and near-claim frequencies for each policyholder of the portfolio entering the computation of stationary probabilities as well as integrals for each individual. As part of this approach, integrals involved in (5.4) have been approximated by the Simpson's rule using `python` package `scipy`. We parallelized the computation of the integrals for each individual reducing the computation time of the relativities for our MTPL insurance portfolio to 128 minutes on a 12 cores machine.

The relativities associated to each level of the scale are reported in Table 5.1. We observe that based on the transition rules considered, level 5 is occupied by 23.5% of the portfolio. Level 0 is occupied by 53.2% of the portfolio benefiting from the maximum discount, whereas other levels accommodate a lower proportion of policyholders. We observe that after reaching the lowest level of the scale, the discount is about 45%.

We also compare the relativities when near-claim events are taken into account to those of the standard -1/top scale based on claims, only. The levels L_t occupied in the scale then obey the recursion

$$L_t = \begin{cases} \max\{L_{t-1} - 1, 0\} \text{ if } N_t = 0 \\ \\ m \text{ if } N_t \geq 1. \end{cases}$$

Table 5.2 shows that policyholders are more concentrated in the lower levels of the bonus-malus scale when only claims at fault are considered. As expected, Table 5.2 shows that the univariate system grants less discount. The normalized relativities are larger for all levels 0-4 compared to Table 5.1. The maximum discount is reduced from 45% to about 33%.

6 Conclusion

Based on telematics signals recorded in a MTPL insurance portfolio, we have proposed in this paper a simple and effective definition of near-claims based on deceleration events. Similarly as for claim frequency, we use traditional insurance explanatory variables for explaining the near-claim frequency. Both responses disclose similar partial dependencies with numerical explanatory variables whereas exhibiting differences with high dimensional categorical explanatory variables. Then, we have integrated the relevant information about near-

Level	$\mathrm{P}[L = \ell]$	**Relativity** r_ℓ	**Norm. Relativity**
5	0.2353457	1.39674985	1.0
4	0.0872465	1.19670596	0.85677901
3	0.0601058	1.12348577	0.80435717
2	0.0466335	1.07353713	0.76859656
1	0.0382324	1.03497714	0.74098962
0	0.5324361	0.76950467	0.55092518

Table 5.1: Characteristics of the -1/top bonus-malus scale integrating both claim and near-claim events.

Level	$\mathrm{P}[L = \ell]$	**Relativity** r_ℓ	**Norm. Relativity**
5	0.06232492	1.42472536	1.0
4	0.05495674	1.35332398	0.949884109
3	0.04898448	1.29716995	0.91047018
2	0.04403181	1.25098174	0.87805115
1	0.03985431	1.21186835	0.85059786
0	0.74984774	0.89339133	0.62706214

Table 5.2: Characteristics of the -1/top bonus-malus scale integrating claims, only.

claim events related to driving experience with the help of a bivariate Poisson-LogNormal credibility model. We have demonstrated the added value of the bivariate model as the corrections resulting from the bivariate model adapt more dynamically to driving experience over the contract duration.

Then, we have designed a commercial simplification as a -1/top bonus-malus scale integrating both claim and near-claim events, but penalizing policyholders only for reporting claims at fault. More sophisticated transition rules could be considered. For instance, near-claim events could be used to transfer policyholders to upper levels in the scale.

Acknowledgements

The authors gratefully acknowledge the financial support of P&V Assurances through the research chair “Telematics and Prevention". We thank our colleagues from P&V Assurances, especially Nikias De Feyter and Carolina Aldana for interesting discussions that greatly contributed to the success of this research project.

A Explanatory variables

Table A.1: Explanatory variables description

Related to	Variable	Type	Unique values	Description
Driver	Postcode longitude	Float	-	Longitude of the policyholder's municipality
	Postcode latitude	Float	-	Latitude of the policyholder's municipality
	Main driver age	Integer	-	Age of the main driver insured
	Vehicle mileage limit	Binary	2	If the policyholder drives more than 10k km yearly
	Contract paysplit	Categorical	4	Premium payment frequency
	Vehicle use	Categorical	4	Use of the vehicle
	Number of drivers	Categorical	5	Number of drivers insured
	Mosaic code	Categorical	9	Policyholder's municipality socio-demographic code
Vehicle	Vehicle power	Float	-	Power of the vehicle power (in kWh)
	Vehicle weight	Float	-	Weight of the vehicle (in kg)
	Vehicle age	Float	-	Age of the vehicle
	Vehicle value	Float	-	Value of the vehicle (in EUR)
	Vehicle adas	Binary	2	If the vehicle is equiped with any driving assisting system
	Vehicle garage	Binary	2	If the vehicle is parked in a garage
	Vehicle fuel	Categorical	7	Vehicle fuel
	Vehicle segmentation	Categorical	7	Vehicle segmentation category
	Vehicle seats	Categorical	8	Number of seats in the vehicle
	Vehicle make	Categorical	32	Make of the vehicle

References

- Apley, D. W., Zhu, J. (2016). Visualizing the effects of predictor variables in black box supervised learning models. Journal of the Royal Statistical Society Series B 82, 1059-1086.

- Ayuso, M., Guillen, M., Nielsen, J.P. (2019). Improving automobile insurance ratemaking using telematics: incorporating mileage and driver behaviour data. Transportation 46, 735-752.

- Boucher, J-P., Côté, S., Guillen, M. (2017). Exposure as duration and distance in telematics motor insurance using Generalized Additive Models. Risks 5, 54.

- Chevalier, A., Coxon, K., Chevalier, A.J, Clarke, E., Rogers, K., Brown, J., Boufous, S., Ivers, R., Keay, L. (2016). Predictors of older drivers' involvement in rapid deceleration events. Accident Analysis & Prevention 98, 312-319.

- Cheung, E. C., Ni, W., Oh, R., Woo, J. K. (2021). Bayesian credibility under a bivariate prior on the frequency and the severity of claims. Insurance: Mathematics and Economics 100, 274-295.

- Corradin, A., Denuit, M., Detyniecki, M., Grari, V., Sammarco, M., Trufin, J. (2022). Joint modeling of claim frequencies and behavioral signals in motor insurance. Astin Bulletin 52, 33-54.

- De Ceunynck, T. (2017). Defining and Applying Surrogate Safety Measures and Behavioural Indicators Through Site-based Observations. Doctoral Dissertation. Department of Technology and Society, Lund University.

- Denuit, M., Guillen, M., Trufin, J. (2019). Multivariate credibility modeling for usage-based motor insurance pricing with behavioral data. Annals of Actuarial Science 13, 378-399.

- Denuit, M., Lu, Y. (2021). Wishart-Gamma random effects models with applications to nonlife insurance. Journal of Risk and Insurance 88, 443-481.

- Dionne, G., Vanasse, C. (1992). Automobile insurance ratemaking in the presence of asymmetrical information. Journal of Applied Econometrics 7, 149-165.

- Duval, F., Boucher, J-P., Pigeon, M. (2022). How much telematics information do insurers need for claim classification? North American Actuarial Journal 26, 570-590.

- Friedman, J. H. (2008). Greedy function approximation: a gradient boosting machine. Annals of Statistics 29, 1189-1232.

- Fu, C., Sayed, T. (2021). Comparison of threshold determination methods for the deceleration rate to avoid a crash (DRAC)-based crash estimation. Accident Analysis & Prevention 153, 106051.

- Gao, G., Meng, S., Wüthrich, M. (2019). Claims frequency modeling using telematics car driving data. Scandinavian Actuarial Journal 2019, 143-162.

- Gao, G., Meng, S., Wüthrich, M. (2022a). What we can learn from telematics car driving data: a survey. Mathematics and Economics 104, 185-199.

- Gao, G., Wang, H., Wüthrich, M. (2022b). Boosting Poisson regression models with telematics car driving data. Machine Learning 111, 243-272.

- Gomez-Deniz, E. (2016). Bivariate credibility bonus–malus premiums distinguishing between two types of claims. Insurance: Mathematics and Economics 70, 117-124.

- Guillen, M., Nielsen, J., Pérez-Marín, A. M. (2021). Near-miss telematics in motor insurance. Journal of Risk and Insurance 88, 569-589.

- Guillen, M., Nielsen, J.P., Pérez-Marín, A. M., Elpidorou, V. (2020). Can automobile insurance telematics predict the risk of near-miss events? North American Actuarial Journal 24, 141-152.

- Park, S. C., Kim, J. H., Ahn, J. Y. (2018). Does hunger for bonuses drive the dependence between claim frequency and severity? Insurance: Mathematics and Economics 83, 32-46.

- Pinquet, J. (1998). Designing optimal bonus-malus systems from different types of claims. Astin Bulletin 28, 205-220.

- Pitrebois, S., Denuit, M., Walhin, J.-F. (2003). Setting a bonus-malus scale in the presence of other rating factors: Taylor's work revisited. ASTIN Bulletin 33, 419-436.

- Pitrebois, S., Denuit, M., Walhin, J.-F. (2006). Multi-event bonus-malus scales. Journal of Risk and Insurance 73, 517-528.

- So, B., Boucher, J-P., Valdez, E. (2021). Synthetic dataset generation of driver telematics. Risks 9, 58.

- Sun, S., Bi, J., Guillen, M., Pérez-Marín, A. M. (2021). Driving risk assessment using near-miss events based on panel Poisson regression and panel Negative Binomial regression. Entropy 23, 829.

- Tzougas, G., di Cerchiara, A. P. (2021). The multivariate mixed Negative Binomial regression model with an application to insurance a posteriori ratemaking. Insurance: Mathematics and Economics 101, 602-625.

- Verschuren, R. M. (2022). Frequency-severity experience rating based on latent Markovian risk profiles. Insurance: Mathematics and Economics 107, 379-392.

- Zheng L., Sayed T., Essa, M. (2018). Validating the bivariate extreme value modeling approach for road safety estimation with different traffic conflict indicators. Accident Analysis & Prevention 123, 314-323.