

Article

Multivariate Credibility in Bonus-Malus Systems Distinguishing between Different Types of Claims

Emilio Gómez-Déniz ^{1,*} and Enrique Calderín-Ojeda ² 

¹ Department of Department of Quantitative Methods, Faculty of Economics and Business Sciences, TiDES Institute, University of Las Palmas de Gran Canaria, Canary Islands, E-35017 Las Palmas de Gran Canaria, Spain

² Centre for Actuarial Studies, Department of Economics, University of Melbourne, Melbourne VIC 3010, Australia; enrique.calderin@unimelb.edu.au

* Correspondence: emilio.gomez-deniz@ulpgc.es

Received: 5 March 2018; Accepted: 8 April 2018; Published: 11 April 2018



Abstract: In the classical bonus-malus system the premium assigned to each policyholder is based only on the number of claims made without having into account the claims size. Thus, a policyholder who has declared a claim that results in a relatively small loss is penalised to the same extent as one who has declared a more expensive claim. Of course, this is seen unfair by many policyholders. In this paper, we study the factors that affect the number of claims in car insurance by using a trivariate discrete distribution. This approach allows us to discern between three types of claims depending whether the claims are above, between or below certain thresholds. Therefore, this model implements the two fundamental random variables in this scenario, the number of claims as well as the amount associated with them. In addition, we introduce a trivariate prior distribution conjugated with this discrete distribution that produce credibility bonus-malus premiums that satisfy appropriate traditional transition rules. A practical example based on real data is shown to examine the differences with respect to the premiums obtained under the traditional system of tarification.

Keywords: Bayesian; bonus-malus system; claim number; claim size; conjugate distribution

MSC: 62-07; 62P05; 62E99

1. Introduction

In an attempt of reducing the economic and casualty losses, the bonus-malus systems (BMS) have been introduced in the actuarial community. BMS is a pricing system mainly used in Europe in vehicle insurance. In this systems, the insured may have his/her premium discounted or penalized based on his/her own experience of claims. Actuarial literature about this topic is extensive see for example [Lemaire; \(1985, 1995\)](#); [Boucher et al. \(2007\)](#); [Mert and Saykan \(2005\)](#); [Sarabia et al. \(2004\)](#); [Denuit et al. \(2009\)](#), among other papers. Different methodologies have been used to determine the fair premium that policyholders must pay for the different classes in which the system is configured. Among these methods, the most popular ones are Bayesian methods and discrete Markov chains.

In Bonus-Malus systems, the premium is usually computed by using only the random variable number of claims. However, as not all the events produce the same individual claim amount Then, as different claims produce different claim sizes, it would be sensible to develop BMS based on both the number of claims and the corresponding severity (see [Gómez-Déniz \(2016\)](#)). In addition to this, if the severity is not included in the bounus-malus premium (BMP), the independence between the claims number and severity is implicitly assumed (see [Lemaire \(2004\)](#)). In this regard, several papers have discussed the

question of implementing both variables in the BMS. See, for example, [Frangos and Vrontos \(2001\)](#); [Gómez-Déniz et al. \(2014\)](#) and the recent paper [Gómez-Déniz \(2016\)](#).

In this article we develop a trivariate model that discriminates between three types of claims. Since this distinction is made with respect to claims that are above, between or below certain thresholds, the resulting model takes into account the two fundamental random variables in this scenario, the number of claims as well as the corresponding claim size. Furthermore, we incorporate a trivariate prior distribution which is conjugated with the latter discrete trivariate model. As a result, we obtain credibility BMP's which satisfy desirable transition rules. We present an example consisting of real data corresponding to an Australian portfolio of automobile insurance claims. Our findings reveal that the BMP's computed by using the methodology proposed in this work (unlike those ones derived under the traditional Poisson-Gamma model) does not modify the discounts made in the absence of claims. However, the methodology used in this paper is different to the recent developments in multivariate credibility models that can be found in the recent actuarial literature (see for example [Frees \(2003\)](#); [Bühlmann and Gisler \(2005\)](#)). Multidimensional credibility models was considered by [Englund et al. \(1999\)](#). In that paper the authors assumed that each dimension of the risk parameter represented one cover from the business. However, they only used frequency information in the credibility approach. Similarly, [Thuring \(2011\)](#) investigated the effect of assuming that one out of two insurance products is inactive when estimating the latent risk profile. Moreover [Thuring et al. \(2012\)](#) used a multivariate credibility model that allows the practitioner to consider the positive correlation in customer behaviour between different financial products and estimate the customer specific risk profiles for a specific product not owned by the customer. Again, this approach uses only two quantities, the a priori expected number of events and the observed number of events.

The Bayesian methodology has been used in actuarial science since the mid-twentieth century and it has proved to be a useful tool for the calculation of insurance premiums. It generally consists of accepting that each policy or insured is represented by a risk parameter that is unknown but random with a certain probability distribution (in the insurance portfolio), called a priori distribution or structure function. This way of proceeding is even more useful in the BMS scenario since the premium obeys certain transition rules that classify the policyholders as a bonus or malus. In other words, it lower the premiums to be paid for the bonuses and it increase the premium for the malus. The fundamental Bayesian tool here is simply the Bayes's Theorem, so by dividing the a posteriori mean of the parameter by the a priori mean, when the net premium principle or the quadratic loss function are used, see [Denuit et al. \(2009\)](#); [Lemaire \(1985, 1995, 2004\)](#); [Sarabia et al. \(2004\)](#), among others, we obtain an estimator of the risk/s parameter/s that will indirectly divide the insured into good risks and bad risks. In addition, the empirical results illustrated in [Frangos and Vrontos \(2001\)](#); [Mert and Saykan \(2005\)](#); [Gómez-Déniz et al. \(2014\)](#) show that many auto insurance portfolios present a positive correlation between these two random variables, and therefore, the assumption of some kind of dependence between them should be considered in the calculation of BMP's. The pioneering research in this field was that of [Picard \(1976\)](#) (see also [Lemaire \(1995\)](#), chp. 13), who divided the claims into two types: small, those ones that are below a limit value, say ψ , and large, above ψ . Then, as this assumption did not produce a good fit, the author proposed to distinguish between accidents that caused property damage and those ones that caused personal injury.

The rest of this paper is organised as follows. Section 2 describes the basic distributional assumptions formulated for the numbers of either type of claims. Section 3 discusses a trivariate conjugate, with respect to the discrete model, prior distribution. Next, Bayesian BMP's are derived and written as credibility formula. Numerical applications are displayed in Section 4. Finally, Section 5 concludes the article.

2. Basic Model

There is a lot of criticism about assuming that the number of claims in an auto insurance portfolio can follow a Poisson distribution due to the fact that for this distribution the dispersion index (ratio between the variance and the mean) is one, when in the auto insurance portfolios have been empirically proven to be a value slightly higher than the unit. However, as an initial starting point and facilitating the methodology that we will develop, we will assume that, in effect, the number of claims has a Poisson distribution with parameter $\theta > 0$ and probability function given by

$$f(x|\theta) = \frac{1}{x!} \theta^x \exp(-\theta), \quad x = 0, 1, \dots \tag{1}$$

When the i th policyholder makes a claim x_i , it has associated a certain size, say $y_i \geq 0$. It is our interest now to distinguish between different types of claims (three in our case). For that reason, we include two new random variables that give rise to the consideration of three separate sub-events as follows. Let us to consider Z_i^0, Z_i^1 and Z_i^2 , the following random variables

$$Z_i^0 = \begin{cases} 1, & y_i \leq \phi_1, \\ 0, & \text{otherwise,} \end{cases} \quad Z_i^1 = \begin{cases} 1, & \phi_1 < y_i \leq \phi_2, \\ 0, & \text{otherwise,} \end{cases} \quad Z_i^2 = \begin{cases} 1, & \phi_2 < y_i, \\ 0, & \text{otherwise,} \end{cases}$$

where ϕ_1 and $\phi_2, \in \mathbb{R}^+$ with $\phi_2 > \phi_1$.

The Z_i^j 's, $j = 0, 1, 2$ are modelled as independent and identically distributed random variables with the following Bernoulli distributions:

$$f(z_i^j|p_i) = \begin{cases} p_i, & \text{if } z_i^j = 1, \\ 1 - p_i, & \text{if } z_i^j = 0, \end{cases}$$

where $0 < p_i < 1$. Observe that these assumptions imply that $E(Z_i) = p_i, (i = 0, 1, 2)$. Since in practise majority of policyholders in the portfolio does not make a claim and those ones that declare claims with large claim size are sparse, we will also assume also $p_0 \geq p_1 \geq p_2$ with $\sum_{j=0}^2 p_j = 1$.¹

We now assume that $Z_1 = \sum_{i=1}^x Z_i^1$ is the total number of claims with a claim size between $\phi_1 > 0$ and $\phi_2 > 0$ and $Z_2 = \sum_{i=1}^{x-z_1} Z_i^2$ is the total number of claims with a claims size larger than ϕ_2 . Thus, if the $Z_i^j, i = 1, \dots, x, j = 1, 2$, are assumed to be mutually independent, then the conditional probability function of Z_1 , given that $X = x$, is binomial with parameters x and p_1 and the conditional probability function of Z_2 , given that $X = x$ and $Z_1 = z_1$ is also binomial with parameters $x - z_1$ and p_2 . That is,

$$f(z_1|x, p_1) = \binom{x}{z_1} p_1^{z_1} (1 - p_1)^{x-z_1}, \quad z_1 = 0, 1, \dots, x, \tag{2}$$

$$f(z_2|x, z_1, p_2) = \binom{x - z_1}{z_2} p_2^{z_2} (1 - p_2)^{x-z_1-z_2}, \quad z_2 = 0, 1, \dots, x - z_1. \tag{3}$$

Thus, the conditional mean and variance are linear. They are given by $E(Z_i|x, (i - 1)z_{i-1}, p_i) = (x - (i - 1)z_{i-1})p_i$ and $var(Z_i|x, (i - 1)z_{i-1}, p_i) = (x - (i - 1)z_{i-1})p_i(1 - p_i), i = 1, 2$.

Now, by conditioning it is easy to get the joint probability function of the random variable (X, Z_1, Z_2) which results,

$$f(x, z_1, z_2|\theta, p_1, p_2) = \frac{(\theta q_1 q_2)^x \exp(-\theta)}{z_1! z_2! (x - z_1 - z_2)!} \left(\frac{p_1}{q_1 q_2} \right)^{z_1} \left(\frac{p_2}{q_2} \right)^{z_2}, \tag{4}$$

¹ As a reviewer has pointed out if this inequality is not sustained, then the likelihood function and posterior distribution that will be defined later are not correct.

for $x = 0, 1, \dots, z_1 = 0, 1, \dots, x, z_2 = 0, 1, \dots, x - z_1$, being $q_i = 1 - p_i, i = 1, 2$. Observe that the distribution depends on three parameters every one related with the three types of claims.

Straightforward algebra provides moments and the cross moment,

$$E(X|\theta) = \theta, \tag{5}$$

$$E(Z_1|\theta, p_1) = \theta p_1, \tag{6}$$

$$E(Z_2|\theta, p_1, p_2) = \theta p_2 q_1, \tag{7}$$

$$E(XZ_1Z_2|\theta, p_1, p_2) = (2 + \theta)\theta^2 p_1 p_2 q_1. \tag{8}$$

Estimation

Given a sample $(\tilde{x}, \tilde{z}_1, \tilde{z}_2)$, where t is the sample size, estimation of the parameters θ, p_1 and p_2 via maximum likelihood method are easily obtained. They result $\hat{\theta} = \bar{x}, \hat{p}_1 = \tilde{z}_1/\bar{x}$ and $\hat{p}_2 = \tilde{z}_2/(\bar{x} - \tilde{z}_1)$, where \bar{x}, \tilde{z}_1 and \tilde{z}_2 are the sample mean of the three random variables, respectively. These estimators coincide with the moment estimators obtained by using (5)–(7). The Fisher information matrix is given by

$$\mathcal{J}(\hat{\theta}, \hat{p}_1, \hat{p}_2) = \text{diag} \left[\frac{t}{\hat{\theta}}, \frac{t\hat{\theta}}{\hat{p}_1(1 - \hat{p}_1)}, \frac{t\hat{\theta}(1 - \hat{p}_1)}{\hat{p}_2(1 - \hat{p}_2)} \right],$$

from which the asymptotic variance-covariance matrix of $(\hat{\theta}, \hat{p}_1, \hat{p}_2)$ is obtained by inverting this information matrix. The score equations used to estimate the parameters and the Fisher’s information matrix are provided in the Appendix A.

3. Contemplating Heterogeneity

Let us assume now that the model includes a certain level of heterogeneity and it allows parameters θ, p_1 and p_2 to vary among insureds in the portfolio. In this regard, we suppose that the parameter θ follows a gamma prior distribution (structure function) with a shape parameter $\alpha > 0$, a scale parameter $\beta > 0$ and a probability density function given by

$$\pi_1(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} \exp(-\beta\theta), \quad \theta > 0,$$

with mean and variance are given by $E(\theta) = \alpha/\beta$ and $var(\theta) = \alpha/\beta^2$, respectively.

The p_i parameters are assumed to follow a beta prior distribution with parameters $\alpha_i > 0$ and $\beta_i > 0, i = 1, 2$. That is, the probability density function of p_i are given by

$$\pi_i(p_i) = \frac{1}{B(\alpha_i, \beta_i)} p_i^{\alpha_i-1} q_i^{\beta_i-1}, \quad 0 < p_i < 1, \tag{9}$$

respectively where $B(a, b)$ represents the beta function given by $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$ and $\Gamma(\cdot)$ is the Euler gamma function. The mean and variance of these prior distributions, given by (9) for $i = 1, 2$, are provided by $E(p_i) = \alpha_i/(\alpha_i + \beta_i)$ and $var(p_i) = \alpha_i\beta_i/((\alpha_i + \beta_i)^2(\alpha_i + \beta_i + 1)), i = 1, 2$, respectively.

The flexibility of the beta distribution allows it to take up different shapes depending on the values of its two parameters. The choice of these prior distributions obeys to the fact that they are conjugate with the likelihoods (see Heilmann (1989); Denuit et al. (2009); Klugman et al. (2008); among others). For that reason, we will assume independence between the random variables θ, p_1 and p_2 by taking $\pi(\theta, p_1, p_2) = \pi_1(\theta)\pi_2(p_1)\pi_3(p_2)$.

Given a sample $(\tilde{x}, \tilde{z}_1, \tilde{z}_2)$, where t is the sample size, the posterior distribution of $\vartheta = (\theta, p_1, p_2)$ given the sample information is computed according to Bayes’s theorem and is proportional to the product of the likelihood and the prior distribution. Thus we find that the likelihood function is proportional to

$$\mathcal{L}((\bar{x}, \bar{z}_1, \bar{z}_2) | \vartheta) \propto (\theta q_1 q_2)^{t\bar{x}} \left(\frac{p_1}{q_1 q_2}\right)^{t\bar{z}_1} \left(\frac{p_2}{q_2}\right)^{t\bar{z}_2} \exp(-t\theta) \tag{10}$$

and the prior distribution is proportional to

$$\pi(\vartheta) \propto \theta^{\alpha-1} \exp(-\beta\theta) \prod_{i=1}^2 p_i^{\alpha_i-1} q_i^{\beta_i-1}.$$

Thus, the posterior distribution is conjugated with respect to the likelihood (10) and it is described by

$$\pi^*(\vartheta | (\bar{x}, \bar{z}_1, \bar{z}_2)) \propto \theta^{\alpha+t\bar{x}-1} \exp(-(\beta+t)\theta) \prod_{i=1}^2 p_i^{\alpha_i+t\bar{z}_i-1} q_i^{\beta_i+t(\bar{x}-\bar{z}_1-(i-1)\bar{z}_2)-1},$$

where the constant of proportionality does not depend on θ , p_1 and p_2 . Here, $\bar{x} = (1/t) \sum_{i=1}^t \tilde{x}_i$, $\bar{z}_i = (1/t) \sum_{i=1}^t \tilde{z}_i$, $i = 1, 2$, are the sample means of X, Z_1 and Z_2 , respectively.

Therefore, the posterior distribution is the product of a gamma and two beta distributions, with the updated parameters given in Table 1.

Table 1. Updated parameters of the posterior distribution.

Parameter	Updated Parameter
α	$\alpha^* = \alpha + t\bar{x}$
β	$\beta^* = \beta + t$
α_1	$\alpha_1^* = \alpha_1 + t\bar{z}_1$
β_1	$\beta_1^* = \beta_1 + t(\bar{x} - \bar{z}_1)$
α_2	$\alpha_2^* = \alpha_2 + t\bar{z}_2$
β_2	$\beta_2^* = \beta_2 + t(\bar{x} - \bar{z}_1 - \bar{z}_2)$

In the numerical applications Section later we will adopt an empirical Bayes approach where the parameters of the prior distributions can be estimated from the data (see [Robbins \(1964\)](#); [Casella \(1985\)](#)). In order to do this, we need the marginal (unconditional) distribution of (X, Z_1, Z_2) , that can be easily obtained by compounding. Due that the variables are separated the integration process is simple. Thus, the unconditional distribution results,

$$\begin{aligned} f(x, z_1, z_2) &= \int_0^\infty \int_0^1 \int_0^1 f(x, z_1, z_2 | \vartheta) \pi(\vartheta) d\vartheta \\ &= \frac{1}{z_1! z_2! (x - z_1 - z_2)!} \mathcal{NB}\left(\alpha, \frac{\beta}{1 + \beta}\right) \mathcal{BB}(\alpha_1 + z_1, \beta_1 + x - z_1) \\ &\times \mathcal{BB}(\alpha_2 + z_2, \beta_2 + x - z_1 - z_2), \end{aligned} \tag{11}$$

where \mathcal{NB} represents the negative binomial distribution and \mathcal{BB} the compound binomial-beta distribution. Some algebra provides that the probability function (12) of this trivariate unconditional model can be rewritten in a compact form as,

$$f(x, z_1, z_2) = \frac{\beta^\alpha \binom{x+\alpha-1}{x} \binom{z_1+\alpha_1-1}{z_1} \binom{z_2+\alpha_2-1}{z_2} \binom{x-z_1+\beta_1-1}{x-z_1} \binom{x-z_1-z_2+\beta_2-1}{x-z_1-z_2}}{(1 + \beta)^{\alpha+x} \binom{x+\alpha_1+\beta_1-1}{x} \binom{x-z_1+\alpha_2+\beta_2-1}{x-z_1}}. \tag{12}$$

The unconditional means are obtained using (5)–(8) by compounding and given by

$$\begin{aligned}
 E(X) &= \frac{\alpha}{\beta}, \\
 E(Z_1) &= \frac{\alpha}{\beta} \frac{\alpha_1}{\alpha_1 + \beta_1}, \\
 E(Z_2) &= \frac{\alpha}{\beta} \frac{\alpha_2}{\alpha_2 + \beta_2} \frac{\beta_1}{\alpha_1 + \beta_1}, \\
 E(XZ_1Z_2) &= \frac{\alpha\alpha_1\alpha_2\beta_1(\alpha + 1)(\alpha + 2\beta + 2)}{\beta^3(\alpha_1 + \beta_1)(\alpha_2 + \beta_2)(\alpha_1 + \beta_1 + 1)}.
 \end{aligned}$$

3.1. The Premiums

Premiums can be derived by following the ideas displayed in Gómez-Déniz (2016). Let

$$g(x, z_1, z_2) = p_{z_2}z_2 + p_{z_1}z_1 + p_x(x - z_1 - z_2), \tag{13}$$

be an appropriate function of the number of claims with claim size below $\phi_1 > 0$, between ϕ_1 and $\phi_2 > 0$ and above ϕ_2 , where p_x, p_{z_1} and p_{z_2} are appropriate weights assigned to the number of claims with different types of size. It is also reasonable to assume that $p_{z_2} > p_{z_1} > p_x$.

Now, by using the net premium principle, i.e., the squared-error loss function, and simple algebra, we obtain the risk premium,

$$P(\theta) = (q_1p_{z_2}p_2 + p_1p_{z_1} + q_1q_2p_x) \theta. \tag{14}$$

Observe that if $p_{z_1} = p_{z_2} = p_x = 1$, then the risk premium in (14) is simply $P(\theta) = \theta$, that is, the risk premium obtained under the traditional model (net premium principle). The premium depends only on the number of claims, irrespective of their size.

This is the fair premium to be charged to a policyholder if θ, p_1 and p_2 were known. However, these quantities are unobservable in practice and then, the risk premium is a theoretical one which cannot be determined exactly and it must be estimated from the data. On the other hand, the a priori premium is obtained for a policyholder about whom nothing is known, i.e., the average premium for all possible risk premiums.

We now obtain the a priori (collective) premium, as follows:

$$P = \int_0^\infty \int_0^1 \int_0^1 P(\theta)\pi(\theta) d\theta = \frac{\alpha p_{z_1}\alpha_1(\alpha_2 + \beta_2) + \beta_1(p_{z_2}\alpha_2 + p_x\beta_2)}{\beta(\alpha_1 + \beta_1)(\alpha_2 + \beta_2)}. \tag{15}$$

Again, by inserting $p_{z_1} = p_{z_2} = p_x = 1$ in (15) we obtain the collective premium computed under the traditional model. That is, $P = \alpha/\beta$.

The Bayesian premium $P^*(t, x, z_1, z_2)$, which is not reproduced here, is derived from (15) by interchanging the parameters α, β, α_i and β_i ($i = 1, 2$) with the updated parameters by using the expressions displayed in Table 1.

Note that $P^*(0, 0, 0, 0) = P$. That is, the Bayesian premium coincides with the a priori premium when no information is available. Furthermore, the expression of the Bayesian premium can be written as

$$P^*(x, z_1, z_2, t) = \gamma(x, z_1, t)P + [1 - \gamma(x, z_1, t)]h(x, z_1, z_2, t),$$

where

$$\begin{aligned}
 \gamma(x, z_1, t) &= \frac{\alpha^*(\alpha_1 + \beta_1)(\alpha_2 + \beta_2)}{\beta^*(\alpha_1^* + \beta_1^*)(\alpha_2^* + \beta_2^*)}, \\
 h(x, z_1, z_2, t) &= \frac{\alpha^* [\mathcal{H}_1 + \mathcal{H}_2 + (x - z_1)(p_{z_2}\alpha_2 + p_x\beta_2)]}{\beta^*(\alpha_1^* + \beta_1^*)(\alpha_2^* + \beta_2^*) - \alpha^*(\alpha_1 + \beta_1)(\alpha_2 + \beta_2)},
 \end{aligned} \tag{16}$$

where

$$\begin{aligned} \mathcal{H}_1 &= p_{z_1} \alpha_1^* (x - z_1) + p_{z_1 z_1} (\alpha_2 + \beta_2), \\ \mathcal{H}_2 &= \beta_1^* [p_{z_2} z_2 + p_x (x - z_1 - z_2)]. \end{aligned}$$

Additionally, $\gamma(x, z_1, z_2, t)$ can also be written as

$$\gamma(x, z_1, z_2, t) = Z(t) \frac{\alpha (\alpha_1 + \beta_1)(\alpha_2 + \beta_2)}{\beta (\alpha_1^* + \beta_1^*)(\alpha_2^* + \beta_2^*)} + [1 - Z(t)] \frac{(\alpha_1 + \beta_1)(\alpha_2 + \beta_2)\bar{x}}{(\alpha_1^* + \beta_1^*)(\alpha_2^* + \beta_2^*)},$$

where

$$Z(t) = \frac{\beta}{\beta + t} = \frac{\kappa}{t + \kappa} \in [0, 1],$$

with $\kappa = E[\text{var}(X|\theta)] / \text{var}[E(X|\theta)]$, coincides with the classical credibility factor usually appearing in this setting in actuarial science (see Bühlmann (1967); Jewell (1974); Gómez-Déniz (2008), among others for details). Now it is simple to see that:

- When $t \rightarrow 0$, $Z(0) \rightarrow 1$, $\gamma(x, z_1, z_2, 0) \rightarrow \frac{\alpha}{\beta}$ and therefore $P^*(x, z_1, z_2, 0) \rightarrow \frac{\alpha}{\beta}P$. Then, the premium is based only in the prior information about the risk. Therefore, the case is the one in which experience is ignored and external information is used as the sole basis for the process of ratemaking.
- When $t \rightarrow \infty$, $Z(\infty) \rightarrow 0$, $\gamma(x, z_1, z_2, \infty) \rightarrow 0$ and therefore $P^*(x, z_1, z_2, \infty) \rightarrow h(x, z_1, z_2, \infty)$. Then, the premium is based only in the sample information.

The Bayesian BMP can now be obtained by considering the rate (see Lemaire (1995); Gómez-Déniz et al. (2002); among others).

$$P^{**}(x, z_1, z_2, t) = \frac{P^*(x, z_1, z_2, t)}{P^*(0, 0, 0, 0)} = \frac{P^*(x, z_1, z_2, t)}{P}, \tag{17}$$

which ensures that the initial premium, i.e., when $(x, z_1, z_2, t) = (0, 0, 0, 0)$, the rate in (17) is precisely P and the rates achieved for the first year is given by $P^*(x, z_1, z_2, 1)$, for the second year $P^*(x, z_1, z_2, 2)$, etc.

4. Numerical Applications

Now, in order to compute the premiums based on the models introduced in this paper, we will examine a dataset that include information based on one-year vehicle insurance policies taken out in 2004 or 2005. This dataset is available on the website of the Faculty of Business and Economics, Macquarie University (Sydney, Australia) (see also De Jong and Heller (2008)). The total portfolio contains 67,856 policies of which 4624 have at least one claim. With respect to the number of claims, the minimum and maximum are 0 and 4 respectively. The mean is 0.072 and standard deviation is 0.278. On the other hand, regarding the claim size, the minimum and maximum are 0 and 55,922.10 respectively. The mean is 137.27 and the standard deviation is 1056.30. This latter measure is very large for the size of the claims, therefore it means that a premium based only on the mean claim size is not adequate for calculating the bonus-malus premiums. Due to this portfolio only includes the aggregate value of the claims severity, a simulation analysis was completed to randomly determine the exact value that corresponds to all claims. Then, we proceed to allocate the claims that correspond to each interval, i.e., \$0–\$500, \$500–\$1000 and >\$1000. Thus, we are assuming that $\phi_1 = 500$ and $\phi_2 = 1000$. This simulation analysis was carried by using Mathematica software package. We have taken the integer part of the individual claim amount, this does not seem very relevant in our analysis. It is important to mention that due to RandomChoice function, the partition of the aggregate claim amount is different every time the program is run.

Empirical values and fitted values by using the discrete trivariate distribution, Fitted (1), and the mixing model (12), Fitted (2), are illustrated in Table 2. These sample values were taken from the results

obtained after dividing the claims by using the simulation scheme mentioned above. The estimated parameter values (the standard errors appear in brackets) are shown in Table 3. We also show the value obtained for two measures of model selection: Akaike’s information criterion (AIC) and the consistent Akaike information criterion (CAIC). See Akaike (1974); Bozdogan (1987) for details. The goodness of fit was determined by standard Pearson’s chi squared test statistics with the following grouping procedure: the outermost classes were consolidated to produce theoretical class sizes of 5 or larger. It is observable that the fit to data is reasonably good for the mixture model and not very promising for the basic model. For the mixture model, the maximum likelihood estimates were obtained by directly maximizing the log-likelihood surface.

Table 2. Empirical and fitted data for the basic model (1) and mixture model (2).

(x, z_1, z_2)	Empirical	Fitted (1)	Fitted (2)	(x, z_1, z_2)	Empirical	Fitted (1)	Fitted (2)
(0, 0, 0)	63,232	63,094.30	63,233.20	(3, 0, 0)	0	0.29	1.49
(1, 0, 0)	1840	1921.02	1812.24	(3, 1, 0)	5	1.05	4.75
(1, 1, 0)	2084	2257.62	2128.08	(3, 0, 1)	0	0.19	0.44
(1, 0, 1)	409	411.91	387.96	(3, 0, 2)	0	0.04	0.22
(2, 0, 0)	31	29.24	51.83	(3, 1, 1)	3	0.45	1.28
(2, 1, 0)	134	68.73	113.61	(3, 2, 1)	0	0.26	1.11
(2, 0, 1)	7	12.54	13.98	(3, 1, 2)	0	0.05	0.51
(2, 1, 1)	16	14.74	24.32	(3, 3, 0)	3	0.48	2.04
(2, 2, 0)	79	40.40	66.82	(3, 0, 3)	0	0.00	0.10
(2, 0, 2)	4	1.34	5.60	(4, ·, ·)	4	0.35	0.89

Table 3. Parameters estimated and standard errors (SE) for the basic and mixture model without including covariates.

Basic Model			Mixture Model		
Parameter	Estimate	SE	Parameter	Estimate	SE
$\hat{\theta}$	0.072	0.001	$\hat{\alpha}$	1.157	0.121
\hat{p}_1	0.492	0.007	$\hat{\beta}$	15.903	1.681
\hat{p}_2	0.176	0.007	$\hat{\alpha}_1$	575.261	2.779
			$\hat{\beta}_1$	594.757	2.961
			$\hat{\alpha}_2$	0.365	0.268
			$\hat{\beta}_2$	1.705	1.257
χ^2	137.06			7.18	
<i>p</i> -value	0.00			0.007	
df	4			1	
AIC	45,129.80			45,027.70	
CAIC	45,160.20			45,088.50	

4.1. The Proposed Premiums

Table 4 illustrates the relativities (Bayesian BMP’s) obtained by applying (17) and the parameter estimates displayed in Table 3. It is noticeable that the structure of this table is built in a similar way the one derived in traditional BMS. Namely, at the beginning of the system the relativity is set equal to 1.000; then this relativity decreases within the year in the absence of claims, and it increases when claims are declared. Nevertheless, for $x > 1$ the system now discerns whether the number of claims corresponds to those below the size ϕ_1 , between ϕ_1 and ϕ_2 and above ϕ_2 . For the sake of comparison, the reader is referred to Gómez-Déniz (2016) where the Bayesian bonus-malus premiums calculated for the Poisson-Gamma model under the net premium principle (see also Dionne and Vanasse (1989)) and those ones computed by using expression (20) in Gómez-Déniz (2016). It is observable that the bonus-malus premiums are the same for the bonus class ($x = 0$) and different for the rest of the malus ($x \geq 1$) classes with respect to the first of the models mentioned above. In this regard, it can be distinguished

now between claims with a severity below ϕ_1 , between ϕ_1 and ϕ_2 and above ϕ_2 . In this sense if we consider Table 3 in Gómez-Déniz (2016), for example for the case $(t, x, z) = (1, 0, 1)$ (i.e., at the end of the first year the policyholder has declared one claim with size higher than \$500) the Bayesian bonus-malus premium is 1.800. However, under the scheme introduced in Table 4, the Bayesian BMP to be paid is 1.754 if the claim size belong to the interval \$500–\$1000, i.e., $(t, x, z_1, z_2) = (1, 1, 1, 0)$. On the other hand, when the claim amount is $> \$1000$, i.e., $(t, x, z_1, z_2) = (1, 1, 0, 1)$ the Bayesian BMP is 2.040. Therefore, the premiums as shown in Table 4, may be larger or smaller than those ones shown in Table 3 in Gómez-Déniz (2016). This methodology would ensure the financial viability of the company.

Table 4. BMP’s for claims when there are x claims, z_1 with a claim size between ψ_1 and ψ_2 , z_2 claims with a size larger than ψ_2 and $x - z_1 - z_2$ claims with a claim size smaller than ψ_1 with $p_x = 0.25$, $p_y = 0.50$ and $p_z = 0.75$.

(x, z_1, z_2)	t					
	0	1	2	3	4	5
(0, 0, 0)	1.000	0.940	0.888	0.841	0.799	0.760
(1, 0, 0)	1.798	1.692	1.597	1.513	1.437	1.368
(1, 1, 0)	1.864	1.754	1.656	1.568	1.489	1.418
(1, 0, 1)	2.168	2.040	1.926	1.824	1.732	1.649
(2, 0, 0)	2.583	2.430	2.295	2.173	2.064	1.965
(2, 1, 0)	2.633	2.477	2.339	2.215	2.104	2.003
(2, 1, 1)	3.174	2.986	2.819	2.670	2.536	2.414
(2, 2, 0)	2.729	2.568	2.424	2.296	2.180	2.076
(2, 2, 1)	3.530	3.321	3.135	2.969	2.820	2.685

5. Final Comments and Future Research

In this paper, a simple model that distinguishes, among three types of claims in bonus-malus settings has been introduced. This distinction is based on discriminating between those claims with associated amount below a threshold, between two values of thresholds or greater than a certain threshold. This methodology presented is based on the use of a trivariate distribution (not common in any statistical scenario) that depends on parameters that in turn are considered as random variables that follow certain a priori probability distributions. As a consequence, it is possible to express the bonus-malus premium based on the net premium principle (quadratic error loss function) as a credibility formula that writes the premium as a convex combination of sample information and a priori information. The bonuses of the premium obtained are undoubtedly fairer than those ones computed by using the classical methodology that does not discern between different types of claims. We shall conclude with an interesting comment made by a referee with respect to the range of values that parameter p_2 can take on. In this work we have assumed that $p_2 \in (0, 1)$, however, in the third section when the probabilities are randomized it could be sensible to suppose that p_2 is dependent on the value of p_1 . In this sense, we believe it should be more realistic to consider $\pi(\vartheta) = \pi(\theta)\pi(p_1, p_2)$, where the latter factor is a bivariate distribution that assumes dependency between p_1 and p_2 . This could be subject of future research. In addition, it would be interesting to examine how the premiums behave when both the distribution based on the classical model and the marginal model derived after including heterogeneity are normalized to implement a generalized linear model. This is likely to refine the premiums according to the individual factors of each insured.

Appendix A

In this Appendix we provide the score equations which provide the maximum likelihood estimators of the parameters and the elements of the Fisher’s information matrix.

For that, let us now consider a sample $(\tilde{x}, \tilde{z}_1, \tilde{z}_2)$ of size t . The likelihood function is given by

$$\mathcal{L}(\vartheta; \tilde{x}, \tilde{z}_1, \tilde{z}_2) \propto \theta^{t\tilde{x}} q_1^{t\tilde{x}} q_2^{t\tilde{x}} \exp(-t\theta) p_1^{t\tilde{z}_1} q_1^{-t\tilde{z}_1} q_2^{-t\tilde{z}_1} p_2^{t\tilde{z}_2} q_2^{-t\tilde{z}_2},$$

from which the log-likelihood results

$$\begin{aligned}\ell(\vartheta; \bar{x}, \bar{z}_1, \bar{z}_2) &\propto t\bar{x} \log \theta + t\bar{x} \log q_1 + t\bar{x} \log q_2 - t\theta \\ &+ t\bar{z}_1 \log p_1 - t\bar{z}_1 \log q_1 - t\bar{z}_1 \log q_2 \\ &+ t\bar{z}_2 \log p_2 - t\bar{z}_2 \log q_2.\end{aligned}$$

The score equations are given by

$$\begin{aligned}\frac{\partial \ell(\vartheta; \bar{x}, \bar{z}_1, \bar{z}_2)}{\partial \theta} &= \frac{t\bar{x}}{\theta} - t = 0, \\ \frac{\partial \ell(\vartheta; \bar{x}, \bar{z}_1, \bar{z}_2)}{\partial p_1} &= -\frac{t\bar{x}}{q_1} + \frac{t\bar{z}_1}{p_1} + \frac{t\bar{z}_1}{q_1} = 0, \\ \frac{\partial \ell(\vartheta; \bar{x}, \bar{z}_1, \bar{z}_2)}{\partial p_2} &= -\frac{t\bar{x}}{q_2} + \frac{t\bar{z}_1}{q_2} + \frac{t\bar{z}_2}{p_2} + \frac{t\bar{z}_2}{q_2} = 0.\end{aligned}$$

The second partial derivatives of the log-likelihood function with respect to the parameters are given by

$$\begin{aligned}\frac{\partial^2 \ell(\vartheta)}{\partial \theta^2} &= -\frac{t\bar{x}}{\theta^2}, \quad \frac{\partial^2 \ell(\vartheta, p_1, p_2)}{\partial \theta \partial p_1} = 0, \quad \frac{\partial^2 \ell(\vartheta, p_1, p_2)}{\partial \theta \partial p_2} = 0, \\ \frac{\partial^2 \ell(\vartheta)}{\partial p_1^2} &= \frac{t(\bar{z}_1 - \bar{x})}{q_1^2} - \frac{t\bar{z}_1}{p_1^2}, \quad \frac{\partial^2 \ell(\vartheta)}{\partial p_1 \partial p_2} = 0, \\ \frac{\partial^2 \ell(\vartheta)}{\partial p_2^2} &= \frac{t(\bar{z}_2 + \bar{z}_1 - \bar{x})}{q_2^2} - \frac{t\bar{z}_2}{p_2^2}.\end{aligned}$$

Now by taking expectations, we have

$$\begin{aligned}E\left(-\frac{\partial^2 \ell(\vartheta)}{\partial \theta^2}\right) &= E\left(\frac{t\bar{x}}{\theta^2}\right) = \frac{t}{\theta}, \\ E\left(-\frac{\partial^2 \ell(\vartheta)}{\partial p_1^2}\right) &= E\left(-\frac{t(\bar{z}_1 - \bar{x})}{q_1^2} + \frac{t\bar{z}_1}{p_1^2}\right) = \frac{t\theta q_1}{q_1^2} + \frac{t\theta}{p_1} = \frac{t\theta}{p_1 q_1}, \\ E\left(-\frac{\partial^2 \ell(\vartheta)}{\partial p_2^2}\right) &= E\left(-\frac{t(\bar{z}_2 + \bar{z}_1 - \bar{x})}{q_2^2} + \frac{t\bar{z}_2}{p_2^2}\right) = \frac{t\theta}{q_2^2} - \frac{p_1 t\theta}{q_2^2} \\ &\quad - \frac{p_2 q_1 t\theta}{q_2^2} + \frac{q_1 t\theta}{p_2} = \frac{q_1 t\theta}{p_2 q_2},\end{aligned}$$

from which the Fisher's information matrix are obtained in the conventional way.

Acknowledgments: This work was partially funded by grant ECO2013–47092 (Ministerio de Economía y Competitividad, Spain and ECO2017–85577–P (Ministerio de Economía, Industria y Competitividad. Agencia Estatal de Investigación)).

Author Contributions: Both authors contributed equally to this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Akaike, Hirotugu. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19: 716–23.
- Boucher, Jean-Philippe, Michel Denuit, and Montserrat Guillén. 2007. Risk classification for claim counts: A comparative analysis of various zero-inflated mixed Poisson and hurdle models. *North American Actuarial Journal* 11: 110–31.
- Bozdogan, Hamparsum. 1987. The general theory and its analytical extension. *Psychometrika* 52: 345–70.

- Bühlmann, Hans. 1967. Experience rating and credibility. *ASTIN Bulletin* 4: 199–7.
- Bühlmann, Hans, and Alois Gisler. 2005. *A Course in Credibility Theory and its Applications*. Berlin: Springer.
- Casella, George. 1985. An introduction to empirical Bayes data analysis. *American Statistician* 39: 83–87.
- Denuit, Michel, Xavier Maréchal, Sandra Pitrebois, and Jean-François Walhin. 2009. *Actuarial Modelling of Claim Counts Risk Classification, Credibility and Bonus-Malus Systems*. Hoboken: John Wiley & Sons.
- De Jong, Piet, and Gillian Z. Heller. 2008. *Generalized Linear Models for Insurance Data*. Cambridge: Cambridge University Press.
- Dionne, Georges, and Charles Vanasse. 1989. A generalization of actuarial automobile insurance rating models: The negative binomial distribution with a regression component. *ASTIN Bulletin* 19: 199–12.
- Englund, Martin, Jim Gustafsson, Jens Perch Nielsen, and Fredrik Thuring. 1999. Multidimensional credibility with time effects: An application to commercial business lines. *The Journal of Risk and Insurance* 76: 443–53.
- 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: 1–22.
- Frees, Edward W. 2003. Multivariate credibility for aggregate loss models. *North American Actuarial Journal* 7: 13–37.
- Gómez-Déniz, E. 2008. A generalization of the credibility theory obtained by using the weighted balanced loss function. *Insurance: Mathematics and Economics* 42: 850–54.
- Gómez-Déniz, E. 2016. Bivariate credibility bonus-malus premiums distinguishing between two types of claims. *Insurance: Mathematics and Economics* 70: 117–24.
- Gómez-Déniz, Emilio, Agustín Hernández-Bastida, and M. P. Fernández-Sánchez. 2014. Computing credibility bonus-malus premiums using the total claim amount distribution. *Hacettepe Journal of Mathematics and Statistics* 43: 1047–61.
- Gómez, E., A. Hernández, J. M. Pérez, and F. J. Vázquez-Polo. 2002. Measuring sensitivity in a bonus-malus system. *Insurance: Mathematics and Economics* 31: 105–13.
- Heilmann, Wolf-Rüdiger. 1989. Decision theoretic foundations of credibility theory. *Insurance: Mathematics and Economics* 8: 75–95.
- Jewell, William S. 1974. Credible means are exact Bayesian for exponential families. *ASTIN Bulletin* 8: 77–90.
- Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot. 2008. *Loss Models: From Data to Decisions*, 3rd ed. Hoboken: Wiley.
- Lemaire, Jean. 1985. *Automobile Insurance. Actuarial Models*. Dordrecht: Kluwer-Nijhoff Publishing.
- Lemaire, Jean. 1995. *Bonus-Malus Systems in Automobile Insurance*. London: Kluwer Academic Publishers.
- Lemaire, Jean. 2004. Bonus-Malus systems. In *Encyclopedia of Actuarial Science*. Hoboken: Wiley.
- Mert, Mehmet, and Yasemin Saykan. 2005. On a bonus malus system where the claim frequency distribution is geometric and the claim severity distribution is Pareto. *Hacettepe Journal of Mathematics and Statistics* 34: 75–81.
- Picard, Philippe. 1976. Généralisation de l'étude sur la survenance des sinistres en assurance automobile. *Bulletin Trimestriel de l'Institut des Actuaries Français* 296: 204–67.
- Robbins, Herbert. 1964. The empirical Bayes approach to statistical decision problems. *Annals of Mathematical Statistics* 35: 1–20.
- Sarabia, José María, Emilio Gómez-Déniz, and Francisco J. Vázquez-Polo. 2004. On the use of conditional specification models in claim count distributions: An application to bonus-malus systems. *ASTIN Bulletin* 34: 85–89.
- Thuring, Fredrik. 2011. A credibility method for profitable cross-selling of insurance products. *Annals of Actuarial Science* 6: 65–75.
- Thuring, Fredrik, Jens Perch Nielsen, Montserrat Guillén, and Catalina Bolancé. 2012. Selecting projects for cross-sellings financial products using multivariate credibility. *Expert Systems with Applications* 39: 8809–16.

