

The Design of an Optimal Bonus-Malus System Based on the Sichel Distribution

George Tzougas and Nicholas Frangos

Abstract This paper presents the design of an optimal Bonus-Malus System (BMS) using the Sichel distribution to model the claim frequency distribution. This system is proposed as an alternative to the optimal BMS obtained by the traditional Negative Binomial model (Lemaire, 1995). The Sichel distribution has a thicker tail than the Negative Binomial distribution and it is considered as a plausible model for highly dispersed count data. We also consider the optimal BMS provided by the Poisson-Inverse Gaussian distribution (PIG), which is a special case of the Sichel distribution. Furthermore, we develop a generalized BMS that takes into account both the a priori and a posteriori characteristics of each policyholder. For this purpose we consider the generalized additive models for location, scale and shape (GAMLSS) in order to use all available information in the estimation of the claim frequency distribution. Within the framework of the GAMLSS we propose the Sichel GAMLSS for assessing claim frequency as an alternative to the Negative Binomial Type I (NBI) regression model used by Dionne and Vanasse (1989, 1992). We also consider the NBI and PIG GAMLSS for assessing claim frequency.

Keywords: Optimal BMS; Claim frequency; Negative Binomial; Poisson-Inverse Gaussian; Sichel; Generalized Additive Models For Location, Scale And Shape (GAMLSS); A priori classification criteria; A posteriori classification criteria.

George Tzougas

Department of Statistics, Athens University of Economics and Business, 76 Patission Str., 104 34, Athens, Greece, e-mail: gtzougas@otenet.gr

Nicholas Frangos

Department of Statistics, Athens University of Economics and Business, 76 Patission Str., 104 34, Athens, Greece, e-mail: nef@aueb.gr

1 Introduction

A Bonus-Malus System (BMS) penalizes policyholders responsible for one or more claims by a premium surcharge (malus) and rewards the policyholders who had a claim-free year by awarding discount of the premium (bonus). An optimal BMS is financially balanced for the insurer and fair for the policyholder. Optimal BMSs can be broadly derived in two ways; based only on the a posteriori classification criteria and based on both the a priori and the a posteriori classification criteria. Typically, classification criteria such as the number of accidents of the policyholder and the severity of each accident are considered as a posteriori, while variables such as the characteristics of the driver and the automobile are considered as a priori classification criteria.

Lemaire (1995) developed the design of an optimal BMS based on the number of claims of each policyholder, following the game-theoretic framework introduced by Bichsel (1964) and Buhlmann (1964). Given that the premium is proportional to the unknown claim frequency and an estimate has to be employed instead, the insurer faces a loss. Minimizing this loss gives the optimal estimate of the policyholder's claim frequency. Lemaire (1995) considered, among other BMS, the optimal BMS obtained using the quadratic error loss function and the expected value premium calculation principle approximating the claim frequency distribution by the Negative Binomial. Walhin and Paris (1997) obtained an optimal BMS using as the claim frequency distribution the Hofmann's distribution, which encompasses the Negative Binomial and the Poisson-Inverse Gaussian, and also using as a claim frequency distribution a finite Poisson mixture. For more on BMS one can see Frangos and Vrontos (2001), Coene and Doray (1996), Walhin and Paris (1997), Lemaire (1995), and Denuit et al. (2007), Mahmoudvand and Hassani (2009), Mahmoudvand and Aziznasiri (2013), Frangos et al. (2010) and Tzougas and Frangos (2013) and the references therein.

Our first contribution is the development of an optimal BMS using the Sichel distribution for assessing claim frequency. This system is proposed as an alternative to the optimal BMS provided by the Negative Binomial distribution (Lemaire, 1995). In fact the Sichel distribution (Sichel, 1985) differs from the standard Negative Binomial one by using an Generalized Inverse Gaussian (GIG) mixing distribution for the parameter of the Poisson density, i.e. the expected claim frequency, instead of the Gamma one, which the derivation of the Negative Binomial distribution is based on. It is important to note that different parameterizations of the Generalized Inverse Gaussian distribution may lead to other models. An additional advantage of the Sichel model is that it can be considered as a candidate model for highly dispersed count data. We also consider the optimal BMS obtained by the Poisson-Inverse Gaussian distribution (PIG), which is a special case of the Sichel distribution.

Our second contribution is the development of a generalized BMS that integrates the a priori and the a posteriori information on an individual basis, extending the framework developed by Dionne and Vanasse (1989, 1992). This is achieved by using the generalized additive models for location, scale and shape (GAMLSS).

The (GAMLSS) were introduced by Rigby and Stasinopoulos (2001, 2005) and Akantziliotou, Rigby and Stasinopoulos (2002) as a way of overcoming some of the limitations associated with the popular generalized linear models, GLM, and generalized additive models, GAM. In the GAMLSS, the exponential family distribution assumption for the response variable is relaxed and replaced by a general distribution family, including highly skewed continuous and discrete distributions. Thus, the GAMLSS are suited to model highly dispersed count data. The GAMLSS are a general framework for univariate regression analysis that allows modelling not only of the mean (or location) but other parameters of the distribution of the response variable as, linear and/or non-linear, parametric and/or additive non-parametric functions of explanatory variables and/or random effects. Within the framework of the GAMLSS we present the Sichel GAMLSS for assessing claim frequency as an alternative to the Negative Binomial regression model of Dionne and Vanasse (1989, 1992). Furthermore, we consider the Negative Binomial Type I (NBI) and the PIG GAMLSS for assessing claim frequency. With the aim of constructing an optimal BMS by updating the posterior mean claim frequency, we adopt the parametric linear formulation of these models and we allow only their mean parameter to be modelled as a function of the significant a priori rating variables for the number of claims. In the resulting generalized system, the premium is a function of the years that the policyholder is in the portfolio, the number of accidents and the significant a priori rating variables for the number of accidents.

The rest of the paper proceeds as follows. In Section 2 we consider the design of an optimal BMS based on the a posteriori classification criteria. The design presented in Section 3 is based on both the a posteriori and the a priori classification criteria. Section 4 contains an application to a data set concerning car-insurance claims while Section 5 summarizes the main findings of the paper.

2 The Design of an Optimal BMS Based on the a Posteriori Criteria

This Section presents the development of an optimal BMS using the Sichel distribution for assessing claim frequency. This system is proposed as an alternative to the optimal BMS provided by the Negative Binomial distribution (see Lemaire, 1995). In fact the Sichel distribution works very well when the data is highly dispersed. In other situations, it works similar to the Negative Binomial distribution. Furthermore, we consider the optimal BMS obtained by the Poisson-Inverse Gaussian (PIG) distribution, which is a special case of the Sichel distribution.

2.1 The Negative Binomial Model

We consider first the design of an optimal BMS using the Negative Binomial distribution for assessing claim frequency. The portfolio is considered to be heterogeneous and all policyholders have constant but unequal underlying risks of having an accident. We assume that the number of claims k given the parameter λ is distributed as a Poisson(λ),

$$P(k|\lambda) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad (1)$$

for $k = 0, 1, 2, 3, \dots$ and $\lambda > 0$, where the parameter λ is the mean claim frequency which varies from individual to individual, denoting the different underlying risk of each policyholder having an accident. Following the setup of Lemaire (1995) we consider that the structure function follows a Gamma distribution which has a probability density function of the form

$$u(\lambda) = \frac{\lambda^{\alpha-1} \tau^\alpha \exp(-\tau\lambda)}{\Gamma(\alpha)}, \quad (2)$$

for $\lambda > 0, \alpha > 0, \tau > 0$, with mean $E(\lambda) = \frac{\alpha}{\tau}$ and variance $Var(\lambda) = \frac{\alpha}{\tau^2}$. Then it can be proved that the unconditional distribution of the number of claims k is a Negative Binomial (α, τ) distribution with probability density function

$$P(k) = \binom{k+\alpha-1}{k} p^\alpha q^k, \quad p = \left(\frac{\tau}{1+\tau} \right), q = \left(\frac{1}{1+\tau} \right), \quad (3)$$

for $k = 0, 1, 2, 3, \dots$, where $\lambda > 0, \alpha > 0, \tau > 0$. The mean and the variance of k are given by $E(k) = \mu = \frac{\alpha}{\tau}$ and $Var(k) = \frac{\alpha}{\tau} \left(1 + \frac{1}{\tau} \right)$ respectively.

Consider a policyholder with claim history k_1, \dots, k_t where k_i is the number of claims that the policyholder had in year $i, i = 1, \dots, t$. Let us denote with $K = \sum_{i=1}^t k_i$ the total number of claims that the policyholder had in t years. Applying Bayes theorem we obtain the posterior structure function of λ for a policyholder or a group of policyholders with claim history k_1, \dots, k_t , denoted as $u(\lambda|k_1, \dots, k_t)$ and given by

$$u(\lambda|k_1, \dots, k_t) = \frac{(\tau+t)^{K+\alpha} \lambda^{K+\alpha-1} e^{-(\tau+t)\lambda}}{\Gamma(\alpha+K)}, \quad (4)$$

which is the probability density function of a gamma $(\alpha+K, \tau+t)$.

Consequently, by using the quadratic error loss function the optimal choice of λ at time $t+1$ for a policyholder with claim history k_1, \dots, k_t , denoted as $\hat{\lambda}_{t+1}$, is the mean of the posterior structure function given by Eq. (4), that is

$$\hat{\lambda}_{t+1}(k_1, \dots, k_t) = \frac{K+\alpha}{\tau+t}. \quad (5)$$

2.2 The Sichel Model

Let us consider now the construction of an optimal BMS using the Sichel distribution to model the claim frequency distribution. The Sichel is a compound Poisson distribution and it can be derived by assuming that the mixing distribution of the Poisson rate λ is a Generalized Inverse Gaussian distribution.

As previously, the portfolio is considered to be heterogeneous and all policyholders have constant but unequal underlying risks of having an accident and $k|\lambda$ is distributed according to a $\text{Poisson}(\lambda)$. Let us assume that the mean claim frequency λ follows a Generalized Inverse Gaussian distribution, denoted as $\text{GIG}(\mu, \sigma, \nu)$, with probability density function given by

$$u(\lambda) = \frac{\left(\frac{c}{\mu}\right)^\nu \lambda^{\nu-1} \exp\left[-\frac{1}{2\sigma}\left(\frac{c}{\mu}\lambda + \frac{\mu}{c}\frac{1}{\lambda}\right)\right]}{2B_\nu\left[\frac{1}{\sigma}\right]}, \quad (6)$$

for $\lambda > 0$, where $\mu > 0, \sigma > 0$ and $-\infty < \nu < \infty$ and where $c = \frac{B_{\nu+1}\left(\frac{1}{\sigma}\right)}{B_\nu\left(\frac{1}{\sigma}\right)}$, where

$$2B_\nu(z) = \int_0^\infty x^{\nu-1} \exp\left[-\frac{1}{2}z\left(x + \frac{1}{x}\right)\right] dx,$$

is the modified Bessel function of the third kind of order ν with argument z . Eq. (6) is obtained from a reparameterization of equation (2.2) of the Generalized Inverse Gaussian distribution of Jørgensen (1982) or equation (15.74) from Johnson et al. (1994) p 284. The mean and the variance of λ are given by $E(\lambda) = \mu$ and $\text{Var}(\lambda) = \mu^2 \left[\frac{2\sigma(\nu+1)}{c} + \frac{1}{c^2} - 1 \right]$ respectively.

Considering the assumptions of the model, it can be proved that the unconditional distribution of the number of claims k is given by a Sichel (μ, σ, ν) distribution, which has a probability density function of the form

$$P(k) = \frac{\left(\frac{\mu}{c}\right)^k B_{k+\nu}(a)}{k! (a\sigma)^{k+\nu} B_\nu\left(\frac{1}{\sigma}\right)}, \quad (7)$$

for $k = 0, 1, 2, 3, \dots$, where $a^2 = \sigma^{-2} + 2\mu(c\sigma)^{-1}$.

The mean of k is equal to $E(k) = \mu$ and the variance of k is equal to $\text{Var}(k) = \mu + \mu^2 \left[\frac{2\sigma(\nu+1)}{c} + \frac{1}{c^2} - 1 \right]$. Like the Negative Binomial, the variance of the Sichel exceeds its mean, a desirable property which is common for all mixtures of Poisson distributions and allows us to deal with data that present overdispersion.

Let us consider now the special case in which $\nu = -0.5$. If we let $\nu = -0.5$ in (6) then the Generalized Inverse Gaussian distribution reduces to an Inverse Gaussian distribution with pdf given by

$$u(\lambda) = \frac{\sqrt{\mu}}{\sqrt{2\pi\sigma\lambda^3}} \exp\left[-\frac{1}{2\sigma\mu\lambda}(\lambda - \mu)^2\right], \quad (8)$$

for $\lambda > 0$ and $\sigma > 0$, where $E(\lambda) = \mu$ and where $Var(\lambda) = \mu^2 \sigma$. Another interesting connection is the following. The special case $\nu = -0.5$ in (7) gives the parameterization of the Poisson-Inverse Gaussian (PIG) distribution used by Dean et al. (1989). The probability density function of the PIG given by

$$P(k) = \left(\frac{2a}{\pi}\right)^{\frac{1}{2}} \frac{\mu^k e^{\frac{1}{\sigma}} B_{k-\frac{1}{2}}(a)}{(a\sigma)^k k!}, \quad (9)$$

for $k = 0, 1, 2, 3, \dots$, where $a^2 = \sigma^{-2} + 2\frac{\mu}{\sigma}$. The mean and the variance of the PIG distribution are given by $E(k) = \mu$ and $Var(k) = \mu + \mu^2 \sigma$ respectively. Thus the $PIG(\mu, \sigma)$ can arise if we assume that the mixing distribution of the Poisson rate λ is an Inverse Gaussian distribution with probability density function given by Eq. (8). Note also that the Poisson-gamma, i.e. Negative Binomial Type I is a limiting case of (7) obtained by letting $\sigma \rightarrow \infty$ for $\nu > 0$.

In what follows we present the design of an optimal Bonus-Malus System (BMS) using the Sichel distribution for assessing claim frequency. We also present the optimal BMS provided by the PIG distribution, which is a special case of the Sichel distribution. Consider again a policyholder observed during t years and denote by k_i the number of accidents in which they were at fault in year $i = 1, \dots, t$, so their claim frequency history will be in a form of a vector (k_1, \dots, k_t) . Let us denote by

$K = \sum_{i=1}^t k_i$ the total number of claims that this insured had in t years. Also, let $k_i | \lambda$,

for $i = 1, \dots, t$, be distributed according to a Poisson (λ) and let the prior structure function of the parameter λ be the $GIG(\mu, \sigma, \nu)$. The posterior structure function of λ for a policyholder or a group of policyholders with claim history k_1, \dots, k_t , denoted as $u(\lambda | k_1, \dots, k_t)$, is a $GIG(w_1, w_2, K + \nu)$ distribution with probability density function of the form

$$u(\lambda | k_1, \dots, k_t) = \frac{\left(\frac{w_1}{w_2}\right)^{\frac{K+\nu}{2}} \lambda^{K+\nu-1}}{2B_{K+\nu}(\sqrt{w_1 w_2})} \exp\left[-\frac{1}{2}\left[w_1 \lambda + w_2 \frac{1}{\lambda}\right]\right], \quad (10)$$

for $\lambda > 0$, where $w_1 = \frac{c}{\sigma\mu} + 2t$ and $w_2 = \frac{\mu}{\sigma c}$, with $\sigma > 0, -\infty < \nu < \infty$ and $c = \frac{B_{\nu+1}\left[\frac{1}{\sigma}\right]}{B_{\nu}\left[\frac{1}{\sigma}\right]}$.

Subsequently, by using the quadratic error loss function, the optimal choice of λ at time $t+1$ for a policyholder with claim history k_1, \dots, k_t is the mean of the $GIG(w_1, w_2, K + \nu)$, i.e. the posterior structure function given by Eq. (10), that is

$$\begin{aligned} \hat{\lambda}_{t+1}(k_1, \dots, k_t) &= \int_0^{\infty} \lambda u(\lambda | k_1, \dots, k_t) d\lambda \\ &= \left(\sqrt{\frac{w_2}{w_1}}\right) \frac{B_{K+\nu+1}(w_1 w_2)}{B_{K+\nu}(w_1 w_2)}. \end{aligned} \quad (11)$$

In the special case when $\nu = -0.5$, i.e. when the simple Inverse Gaussian (given by Eq. (8)) is the structure function of λ , the optimal choice of λ at time $t+1$ for a policyholder with claim history k_1, \dots, k_t is

$$\begin{aligned}\hat{\lambda}_{t+1}(k_1, \dots, k_t) &= \int_0^\infty \lambda u(\lambda | k_1, \dots, k_t) d\lambda \\ &= \left(\sqrt{\frac{h_2}{h_1}} \right) \frac{B_{K+\frac{1}{2}}(h_1 h_2)}{B_{K-\frac{1}{2}}(h_1 h_2)},\end{aligned}\quad (12)$$

where $h_1 = \frac{1}{\sigma\mu} + 2t$ and $h_2 = \frac{\mu}{\sigma}$ and where $\sigma > 0$.

3 The Design of an Optimal BMS Based on Both the a Priori and the a Posteriori Criteria

In this section we develop a generalized BMS that integrates the a priori and the a posteriori information on an individual basis. For this purpose we consider the generalized additive models for location, scale and shape, GAMLSS, (see Rigby and Stasinopoulos, 2001, 2005 and Akantziliotou et al., 2002) in order to use all available information in the estimation of the claim frequency distribution. Within the framework of the GAMLSS we propose the Sichel GAMLSS for assessing claim frequency as an alternative to the Negative Binomial regression model of Dionne and Vanasse (1989, 1992). Furthermore, we consider the NBI and the PIG GAMLSS for approximating the number of claims. With the aim of constructing an optimal BMS by updating the posterior mean claim frequency, we adopt the parametric linear formulation of these models and we allow only their mean parameter to be modelled as a function of the significant a priori rating variables for the number of claims. In this generalized BMS, the premium is a function of the years that the policyholder is in the portfolio, the number of accidents and the explanatory variables for the number of accidents.

3.1 The Negative Binomial Model

This generalized optimal BMS is developed according to the design of Dionne and Vanasse (1989, 1992), Frangos and Vrontos (2001)¹ and Mahmoudvand and Hassani (2009). Consider a policyholder i with an experience of t periods whose number of claims for period j , denoted as K_i^j are independent. If we assume that K_i^j follows the Poisson distribution with parameter λ^j , the expected number of claims for period j , then the probability of having k accidents is

¹ We use the same notation as in Frangos and Vrontos (2001).

$$P(K_i^j = k) = \frac{e^{-\lambda^j} (\lambda^j)^k}{k!},$$

for $k = 0, 1, 2, 3, \dots$ and $\lambda^j > 0$, where $E(K_i^j) = \lambda^j$ and $Var(K_i^j) = \lambda^j$.

We can allow the λ^j parameter to vary from one individual to another. Let $\lambda_i^j = \exp(c_i^j \beta^j)$, where $c_i^j = (c_{i,1}^j, \dots, c_{i,h}^j)$ is the $1 \times h$ vector of h individual characteristics, which represent different a priori rating variables and β^j is the vector of the coefficients. The exponential form ensures the non-negativity of λ_i^j . The conditional to c_i^j probability that policyholder i will be involved in k accidents during the period j will become

$$P(K_i^j = k | c_i^j) = \frac{e^{-\exp(c_i^j \beta^j)} [\exp(c_i^j \beta^j)]^k}{k!}, \quad (13)$$

for $k = 0, 1, 2, 3, \dots$ and $\lambda_i^j > 0$, where $E(K_i^j | c_i^j) = Var(K_i^j | c_i^j) = \lambda_i^j = \exp(c_i^j \beta^j)$.

For the determination of the expected number of claims in this model we assume that the h individual characteristics provide enough information. However, if one assumes that the a priori rating variables do not contain all the significant information for the expected number of claims then a random variable ε_i has to be introduced into the regression component. According to Gouriéroux, Montfort and Trognon (1984 a), (1984 b) we can write

$$\lambda_i^j = \exp(c_i^j \beta^j + \varepsilon_i) = \exp(c_i^j \beta^j) u_i,$$

where $u_i = \exp(\varepsilon_i)$, yielding a random λ_i^j .

Assume that u_i follows a Gamma distribution with probability density function

$$v(u_i) = \frac{u_i^{\frac{1}{\alpha}-1} \frac{1}{\alpha} \exp(-\frac{1}{\alpha} u_i)}{\Gamma(\frac{1}{\alpha})}, \quad (14)$$

$u_i > 0, \alpha > 0$, with mean $E(u_i) = 1$ and variance $Var(u_i) = \alpha$. Under this assumption the conditional distribution of $K_i^j | c_i^j$ becomes

$$P(K_i^j = k | c_i^j) = \binom{k + \frac{1}{\alpha} - 1}{k} \frac{[\alpha \exp(c_i^j \beta^j)]^k}{[1 + \alpha \exp(c_i^j \beta^j)]^{k + \frac{1}{\alpha}}}, \quad (15)$$

which is a Negative Binomial Type I (NBI) distribution with parameters α and $\exp(c_i^j \beta^j)$. It can be shown that the above parameterization does not affect the results if there is a constant term in the regression. We choose $E(u_i) = 1$ in order to have $E(\varepsilon_i) = 0$. The mean and the variance of the NBI distribution are given

by $E(K_i^j | c_i^j) = \mu_i^j = \exp(c_i^j \beta^j)$ and $Var(K_i^j | c_i^j) = \exp(c_i^j \beta^j) [\alpha \exp(c_i^j \beta^j)]$ respectively. More details about the Negative Binomial regression can be found in Lawless (1987) and Hilbe (2011). Note also that Eq. (15) gives the parametric linear GAMLSS where only the mean parameter of the distribution of the response variable, i.e. the NBI distribution, is modelled as a function of the explanatory variables.

We are going to build an optimal BMS based on the number of past claims and on an individual's characteristics in order to adjust that individual's premiums over time. The problem is to determine, at the renewal of the policy, the expected claim frequency of the policyholder i for the period $t + 1$ given the observation of the reported accidents in the preceding t periods and observable characteristics in the preceding $t + 1$ periods and the current period.

Consider a policyholder i with K_i^1, \dots, K_i^t claim history and c_i^1, \dots, c_i^{t+1} characteristics and denote as $K = \sum_{j=1}^t K_i^j$ the total number of claims that they had. The mean claim frequency of the individual i for period $t + 1$ is $\lambda_i^{t+1}(c_i^{t+1}, u_i)$, a function of both the vector of the individual's characteristics and a random factor u_i with probability density function given by Eq. (14). The posterior distribution of λ_i^{t+1} is obtained using Bayes theorem and is given by a Gamma with updated parameters $\frac{1}{\alpha} + K$ and S_i^j , with pdf

$$f(\lambda_i^{t+1} | K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^t) = \frac{(S_i^j)^{K + \frac{1}{\alpha}} (\lambda_i^{t+1})^{K + \frac{1}{\alpha} - 1} \exp[-S_i^j \lambda_i^{t+1}]}{\Gamma(\frac{1}{\alpha} + K)}, \quad (16)$$

where $S_i^j = \frac{\frac{1}{\alpha} + \sum_{j=1}^t \exp(c_i^j \beta^j)}{\exp(c_i^{t+1} \beta^{t+1})}$, with $\lambda_i^{t+1} > 0$ and $\alpha > 0$.

Using the quadratic loss function one can find that the optimal estimator of λ_i^{t+1} is the mean of the posterior structure function, given by

$$\begin{aligned} \hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^{t+1}) &= \int_0^\infty \lambda_i^{t+1}(c_i^{t+1}, u_i) f(\lambda_i^{t+1} | K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^t) d\lambda_i^{t+1} \\ &= \exp(c_i^{t+1} \beta^{t+1}) \left[\frac{\frac{1}{\alpha} + \sum_{j=1}^t K_i^j}{\frac{1}{\alpha} + \sum_{j=1}^t \exp(c_i^j \beta^j)} \right]. \end{aligned} \quad (17)$$

This estimator defines the premium and corresponds to the multiplicative tariff formula where the base premium is the a priori frequency $\exp(c_i^{t+1} \beta^{t+1})$ and where the Bonus-Malus factor is represented by the expression in brackets.

Let us consider, as a special case, the situation in which the vector of the individual characteristics remains the same from one year to the next, i.e. $c_i^1 = c_i^2 = \dots = c_i^{t+1} = c_i$ and $\beta_z^1 = \beta_z^2 = \dots = \beta_z^t = \beta$. Then $\hat{\lambda}_i^{t+1}$ is simplified to

$$\hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^{t+1}) = \exp(c_i \beta) \left[\frac{\frac{1}{\alpha} + \sum_{j=1}^t K_i^j}{\frac{1}{\alpha} + t \exp(c_i \beta)} \right].$$

When $t = 0$, $\hat{\lambda}_i^1(c_i^1) = \exp(c_i^1 \beta)$, which implies that only a priori rating is used in the first period. Moreover, when the regression component is limited to a constant β_0 one obtains

$$\hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t) = \exp(\beta_0) \left[\frac{\frac{1}{\alpha} + \sum_{j=1}^t K_i^j}{\frac{1}{\alpha} + t \exp(\beta_0)} \right],$$

which corresponds to the ‘univariate’, without regression component, model.

3.2 The Sichel Model

Let us now consider the generalized BMS obtained by using the Sichel parametric linear GAMLSS for assessing claim frequency. The Sichel distribution (Sichel, 1985) can be considered as a candidate model for highly dispersed claim count data when the observed high dispersion cannot be efficiently handled by the Negative Binomial regression model.

Consider a policyholder i with an experience of t periods whose number of claims for period j , denoted as K_i^j are independent. We assume again that K_i^j follows the Poisson distribution with parameter $\lambda_i^j = \exp(c_i^j \beta^j)$, where $c_i^j = (c_{i,1}^j, \dots, c_{i,h}^j)$ is the vector of h individual characteristics and β^j is the vector of the coefficients. The conditional to c_i^j probability that policyholder i will be involved in k accidents during the period j is given by Eq. (13).

For the determination of the expected number of claims in this model we assume that the h individual characteristics provide enough information. Nevertheless, if one assumes that the a priori rating variables do not contain all the significant information for the expected number of claims then a random variable ε_i has to be introduced into the regression component, and for $u_i = \exp(\varepsilon_i)$ we have

$$\lambda_i^j = \exp(c_i^j \beta^j + \varepsilon_i) = \exp(c_i^j \beta^j) u_i,$$

yielding a random λ_i^j . Let u_i have a Generalized Inverse Gaussian distribution $\text{GIG}(1, \sigma, \nu)$. This parameterization ensures that $E(u_i) = 1$. Note also that $\text{Var}(u_i) = \frac{2\sigma(\nu+1)}{c} + \frac{1}{c^2} - 1$.

Considering the assumptions of the model, the conditional distribution of $K_i^j | c_i^j$ will be a Sichel $(\exp(c_i^j \beta^j), \sigma, \nu)$ distribution. The above parametrization of the Sichel distribution ensures that the location parameter is the mean of $K_i^j | c_i^j$, given by

$$E(K_i^j | c_i^j) = \mu_i^j = \exp(c_i^j \beta^j).$$

Note that the variance of $K_i^j | c_i^j$ is given by

$$\text{Var}(K_i^j | c_i^j) = \exp(c_i^j \beta^j) + \left(\exp(c_i^j \beta^j)\right)^2 \left[\frac{2\sigma(\nu+1)}{c} + \frac{1}{c^2} - 1 \right]$$

, where $c = \frac{B_{\nu+1}(\frac{1}{\sigma})}{B_{\nu}(\frac{1}{\sigma})}$ and $B_{\nu}(z)$ is the modified Bessel function of the third kind of order ν with argument z . This gives the parametric linear GAMLSS where only the mean parameter of the distribution of the response, i.e. the Sichel distribution, is modelled as a function of the significant a priori rating variables for the number of claims.

Our goal is to build an optimal BMS which integrates a priori and a posteriori information on an individual basis, using the Sichel GAMLSS for assessing claim frequency. We will also consider the optimal BMS provided by the PIG GAMLSS, which is a special case of the Sichel model for $\nu = -0.5$. Similarly to the case of the Negative Binomial model, the problem is to determine at the renewal of the policy the expected claim frequency of the policyholder i for the period $t+1$ given the observation of the reported accidents in the preceding t periods and observable characteristics in the preceding $t+1$ periods and the current period.

Consider again a policyholder i with claim history K_i^1, \dots, K_i^t and c_i^1, \dots, c_i^{t+1} characteristics and denote by $K = \sum_{j=1}^t K_i^j$ the total number of claims that they had. The

mean claim frequency of the individual i for period $t+1$ is $\lambda_i^{t+1}(c_i^{t+1}, u_i)$, a function of both the vector of individual characteristics and a random factor u_i with pdf $\text{GIG}(1, \sigma, \nu)$. The posterior distribution of the expected claim frequency λ_i^{t+1} for an individual i observed over $t+1$ periods with periods with K_i^1, \dots, K_i^t claim history and c_i^1, \dots, c_i^{t+1} characteristics is obtained by using Bayes theorem and is given by a

$$\text{GIG}(w_1, w_2, K + \nu), \text{ where } w_1 = \frac{c + 2\sigma \sum_{j=1}^t \exp(c_i^j \beta^j)}{\sigma \exp(c_i^{t+1} \beta^{t+1})} \text{ and } w_2 = \frac{\exp(c_i^{t+1} \beta^{t+1})}{\sigma c}.$$

Using the quadratic loss function one can find that the optimal estimator of λ_i^{t+1} is the mean of the $\text{GIG}(w_1, w_2, K + \nu)$, i.e. the posterior structure function, that is

$$\begin{aligned}
& \hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^{t+1}) \\
&= \int_0^\infty \lambda_i^{t+1}(c_i^{t+1}, u_i) f(\lambda_i^{t+1} | K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^t) d\lambda_i^{t+1} \\
&= \left(\sqrt{\frac{w_2}{w_1}} \right) \frac{K_{K+v+1}(w_1 w_2)}{K_{K+v}(w_1 w_2)}. \tag{18}
\end{aligned}$$

When the vector of the individual characteristics remains the same from one year to the next, $\hat{\lambda}_i^{t+1}$, given by Eq. (18), is simplified to

$$\hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^{t+1}) = \left(\sqrt{\frac{\omega_2}{\omega_1}} \right) \frac{K_{K+v+1}(\omega_1 \omega_2)}{K_{K+v}(\omega_1 \omega_2)},$$

for $\omega_1 = \frac{c}{\sigma \exp(c_i \beta)} + 2t$ and $\omega_2 = \frac{\exp(c_i \beta)}{\sigma c}$.

When $t = 0$, $\hat{\lambda}_i^1(c_i^1) = \exp(c_i^1 \beta)$, which implies that only a priori rating is used in the first period. Moreover, when the regression component is limited to a constant β_0 one obtains

$$\hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t) = \left(\sqrt{\frac{\omega_2}{\omega_1}} \right) \frac{K_{K+v+1}(\omega_1 \omega_2)}{K_{K+v}(\omega_1 \omega_2)},$$

for $\omega_1 = \frac{c}{\sigma \exp(\beta_0)} + 2t$ and $\omega_2 = \frac{\exp(\beta_0)}{\sigma c}$, which corresponds to the ‘univariate’, without regression component, model.

Let us consider now the special case when $v = -0.5$. In this case, the posterior structure function of λ_i^{t+1} for a policyholder with K_i^1, \dots, K_i^t claim history and

c_i^1, \dots, c_i^{t+1} characteristics is a GIG($h_1, h_2, K - \frac{1}{2}$), where $h_1 = \frac{1+2\sigma \sum_{j=1}^t \exp(c_i^j \beta^j)}{\sigma \exp(c_i^{t+1} \beta^{t+1})}$ and $h_2 = \frac{\exp(c_i^{t+1} \beta^{t+1})}{\sigma}$.

Using again the quadratic error loss function, the optimal choice of $\hat{\lambda}_{t+1}$ for a policyholder with claim history k_1, \dots, k_t is the mean of the GIG($h_1, h_2, K - \frac{1}{2}$), that is

$$\begin{aligned}
& \hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^{t+1}) \\
&= \int_0^\infty \lambda_i^{t+1}(c_i^{t+1}, u_i) f(\lambda_i^{t+1} | K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^t) d\lambda_i^{t+1} \\
&= \left(\sqrt{\frac{h_2}{h_1}} \right) \frac{K_{K+v+1}(h_1 h_2)}{K_{K+v}(h_1 h_2)}. \tag{19}
\end{aligned}$$

When the vector of the individual characteristics remains the same from one year to the next, $\hat{\lambda}_i^{t+1}$ is simplified to

$$\hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^{t+1}) = \left(\sqrt{\frac{\eta_2}{\eta_1}} \right) \frac{K_{K+v+1}(\eta_1 \eta_2)}{K_{K+v}(\eta_1 \eta_2)},$$

for $\eta_1 = \frac{1}{\sigma \exp(c_i \beta)} + 2t$ and $\eta_2 = \frac{\exp(c_i \beta)}{\sigma}$.

When $t = 0$, $\hat{\lambda}_i^1(c_i^1) = \exp(c_i^1 \beta)$, which implies that only a priori rating is used in the first period. Moreover, when the regression component is limited to a constant β_0 one obtains

$$\hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t) = \left(\sqrt{\frac{\eta_2}{\eta_1}} \right) \frac{K_{K+v+1}(\eta_1 \eta_2)}{K_{K+v}(\eta_1 \eta_2)},$$

for $\eta_1 = \frac{1}{\sigma \exp(\beta_0)} + 2t$ and $\eta_2 = \frac{\exp(\beta_0)}{\sigma}$, which corresponds to the ‘univariate’, without regression component, model.

4 Application

The data were kindly provided by a Greek insurance company and concern a motor third party liability insurance portfolio. The data refer to the policyholders at the end of the year 2011 and specifically they describe the number of claims at fault at that year. The data set consists of 4469 policyholders. The mean of claims at fault is 0.138 and the variance is 0.73085. The a priori rating variables we employ are the sex of the driver, Bonus-Malus (BM) class and the horsepower of the car. The drivers were divided into four categories according to the horsepower of their car. Those who had a car with a horsepower between 0-33, between 34-66, between 67-99 and between 100-132. This Bonus-Malus System has 20 classes and the transition rules are described as follows: Each claim free year is rewarded by one class discount and each accident in given year is penalized by one class. The drivers were divided into five categories according to their BM class. Those who belong to BM classes 1 and 2, those who belong to BM classes 3-5, those who belong to BM classes 6-9, those who belong to BM class 10 and those who belong to BM classes 11-20. Firstly, the Negative Binomial, Poisson-Inverse Gaussian (PIG) and Sichel distributions were fitted on the number of claims. Secondly, the NBI, PIG and Sichel GAMLSS were applied to model claim frequency. For the GAMLSS models we selected the parametric linear formulation considering a linear model in the explanatory variables only for the log of their mean parameter in order to derive optimal an optimal BMS by updating the posterior mean. The distributions and the GAMLSS models were estimated using the GAMLSS package in the software R. The ratio of Bessel functions of the third kind whose orders are different was calculated using the HyperbolicDist package in software R. Subsequently, we are able to compute the premiums determined by the optimal BMS based on the a posteri-

ori criteria and the premiums determined by the optimal BMS based both on the a priori and the a posteriori criteria according to the net premium principle and the current methodology as presented in Sections 2 and 3 respectively. For the calculation of the premiums based on both criteria we assume that the sex of the driver and the horsepower of the car remain the same from one year to the next and BM class changes because the policy moves up or down according to the transition rules described previously.

4.1 Claim Frequency Models Comparison

In this subsection we compare the fit of the models for the observed claim frequencies in the portfolio of 4469 policyholders analyzed in the preceding section. These models are all non-nested. In order to accept or reject some models, classical hypothesis/specification tests for non-nested models can be used (see, Boucher et al., 2007, 2008).

Firstly, we compare the non-nested distributions presented in Section 2. In this case, information criteria like AIC or SBC are useful as well as the Vuong test (Vuong, 1989). Table 1 (Panels A and B) reports our results with respect to the aforementioned non-nested comparisons. Specifically, from Panel A and Panel B we observe the superiority of the Poisson-Inverse Gaussian distribution vs the Negative Binomial distribution. Overall, the best fit is given by the Sichel distribution.

Table 1 Comparison of Distributions for the Greek Data Set

Panel A: Based on AIC, SBC				
Model	df	AIC	SBC	
Negative Binomial	2	29338.6	29353.9	
PIG	2	29313.2	29328.5	
Sichel	3	29311.9	29334.9	
Panel B: Based on Vuong Test Statistic				
Model 1	Model 2	Vuong Test	p-value	Decision
Negative Binomial	PIG	-2.38	0.00	PIG
PIG	Sichel	-0.71	0.00	Sichel

Secondly, we compare the non-nested GAMLSS models presented in Section 3 employing Global Deviance, AIC, SBC (see, Rigby and Stasinopoulos, 2009) and the Vuong test. The results are displayed in Table 2. Specifically, when the Global Deviance, AIC and SBC are used (Table 2, Panel A)) our findings suggest that the PIG GAMLSS is superior to the NBI GAMLSS. However, when the Vuong test is used, (Table 2, Panel B) we observe the superiority of the NBI GAMLSS vs the PIG GAMLSS. Finally, with respect to the Global Deviance, AIC, SBC and the Vuong test results, the Sichel GAMLSS provided the best fitting performances.

Table 2 Comparison of GAMLSS Models for the Greek Data Set

Panel A: Based on Global Deviance, AIC, SBC				
Model	df	Global Deviance	AIC	SBC
NBI	10	28417.9	28437.9	28514.5
PIG	10	28380.1	28380.1	28476.6
Sichel	11	28347.28	28369.3	28453.5
Panel B: Based on Vuong Test Statistic				
Model 1	Model 2	Vuong Test	p-value	Decision
NBI	PIG	26.45	0.00	NBI
NBI	Sichel	-2.46	0.00	Sichel

4.2 Optimal BMS Based on the a Posteriori Criteria

In this subsection we consider the premiums determined by the optimal BMS based on the a posteriori classification criteria. In the following examples, the premiums will be divided by the premium when $t = 0$, since we are not so much interested in the absolute premium values as in the differences between various classes. We will present the results so that the premium for a new policyholder is 100.

Let us consider a policyholder observed for 7 years whose number of claims range from 1 to 6. In the following tables we compute this individual's scaled premiums for the case of the Negative Binomial, PIG and Sichel models respectively. We consider first the Negative Binomial model, following Lemaire (1995). The maximum likelihood estimators of the parameters are $\hat{\tau} = 7.868$ and $\hat{\alpha} = 1.089$.

Table 3 Optimal BMS Based on the a Posteriori Classification Criteria, Negative Binomial Model

Number of Claims							
Year	k						
t	0	1	2	3	4	5	6
0	100.00	0.00	0.00	0.00	0.00	0.00	0.00
1	88.72	170.14	251.55	332.95	414.37	495.77	577.19
2	79.73	152.89	226.05	299.21	372.40	445.54	518.70
3	72.40	138.82	205.25	271.68	338.11	404.55	471.00
4	66.29	127.13	187.96	248.79	309.63	370.46	431.30
5	61.14	117.25	173.35	229.46	285.56	341.67	397.80
6	56.73	108.79	160.85	212.91	265.00	317.04	369.09
7	52.92	101.48	150.03	198.60	247.15	295.71	344.27

Let us consider next the Poisson-Inverse Gaussian (PIG) distribution. The maximum likelihood estimators of the parameters are $\hat{\mu} = 0.138$ and $\hat{\sigma} = 0.989$. The BMS derived by the PIG distribution will be defined by Eq. (12) and is presented in Table 4.

Table 4 Optimal BMS Based on the a Posteriori Classification Criteria, PIG Model

Year t	Number of Claims						
	0	1	2	3	4	5	6
0	100.00	0.00	0.00	0.00	0.00	0.00	0.00
1	88.60	156.62	254.20	371.02	497.35	628.04	760.81
2	80.37	131.16	201.63	286.00	378.14	474.22	572.34
3	74.08	113.86	167.54	231.65	302.14	376.17	452.15
4	69.08	101.31	143.81	194.37	250.22	309.22	370.06
5	64.95	91.77	126.42	167.44	212.90	261.14	311.11
6	61.50	84.25	113.16	147.20	185.00	225.25	267.10
7	58.54	78.17	102.72	131.50	163.45	197.61	233.24

Finally, we consider the Sichel distribution. The maximum likelihood estimators of the parameters are $\hat{\mu} = 0.138$, $\hat{\sigma} = 0.990$ and $\hat{\nu} = -1.244$. This system provided by this model will be defined by Eq. (11) and is presented in Table 5.

Table 5 Optimal BMS Based on the a Posteriori Classification Criteria, Sichel Model

Year t	Number of Claims						
	0	1	2	3	4	5	6
0	100.00	0.00	0.00	0.00	0.00	0.00	0.00
1	94.32	158.79	262.10	400.55	561.44	733.96	912.19
2	88.83	134.68	201.93	289.09	390.60	500.82	615.94
3	83.96	118.78	166.88	227.55	298.02	375.08	456.26
4	79.71	107.34	143.89	188.97	241.06	298.24	358.87
5	76.00	98.62	127.58	162.66	202.94	247.21	294.36
6	72.72	91.71	115.37	143.60	175.81	211.21	249.02
7	69.82	86.05	105.86	129.17	155.60	184.63	215.69

It is interesting to compare the optimal BMS provided by the Sichel distribution with the systems obtained from the Poisson-Inverse Gaussian and Negative Binomial distributions respectively. From Table 3, Table 4 and Table 5 we observe that these three systems are fair since if the policyholder has a claim free year the premium is reduced, while if the policyholder has one or more claims the premium is increased. Furthermore, we notice that they can be considered generous with good risks and strict with bad risks. For example, the bonuses given for the first claim free year are 11.28%, 11.4% and 5.68% of the basic premium in the case of the Negative Binomial (Table 3), Poisson-Inverse Gaussian (Table 4) and Sichel (Table 5) models respectively. On the contrary, policyholders who had one claim over the first year of observation will have to pay a malus of 70.14%, 56.62% and 58.79% of the basic premium in the case of the Negative Binomial, Poisson-Inverse Gaussian and Sichel models respectively. Also, policyholders who had one claim over the second year of

observation will have to pay a malus of 51.55%, 54.20% and 62.10% in the case of the Negative Binomial, Poisson-Inverse Gaussian model and Sichel models respectively. Overall, we observe that the bonuses given for the first claim free year in the case of the Negative Binomial model are the highest among three models. However, the high rate of malus for large number of claims in the case of the Sichel model is a superiority. The features of these models must be taken under consideration for determining bonus and malus in practice.

4.3 Optimal BMS Based on Both the a Priori and the a Posteriori Criteria

In this subsection we consider the premiums determined by the generalized optimal BMS that integrates the a priori and the a posteriori information on an individual basis. In what follows the premiums will be divided again by the premium when $t = 0$, as it is interesting to see the percentage change in the premiums after one or more claims.

Let us see an example in order to understand better how this BMS works. Consider a group of policyholders who share the following common characteristics. The policyholder i is a woman who has a car with horsepower between 0-33 and her Bonus-Malus (BM) class varies over time, starting from BM class 1. Implementing the NBI GAMLSS we found that $\hat{\alpha} = 0.655$, implementing the PIG GAMLSS we found that $\hat{\sigma} = 0.725$, and implementing the Sichel GAMLSS we found that $\hat{\sigma} = 0.889$ and $\hat{\nu} = -3.023$. As we have already mentioned, the mean (or location) parameter of these models is given by $E(K_i^j | c_i^j) = \mu_i^j = \exp(c_i^j \beta^j)$, where $c_i^j(c_{i,1}^j, \dots, c_{i,h}^j)$ is the $1 \times h$ vector of h individual characteristics, which represent different a priori rating variables and β^j is the vector of the coefficients. Note also that all the explanatory variables of these models were statistically significant at a 5% threshold. The estimation of the vector β^j and therefore of the mean parameter, $\hat{\mu}_i^j$, for the NBI, PIG and Sichel distributions respectively led to the following results presented in Table 6.

Table 6 Estimation of the Mean Parameter, Women, Horse Power 0-33

BM Category	NBI	PIG	Sichel
1	0.1339	0.1323	0.1314
2	0.2459	0.2483	0.2514
3	0.3123	0.3088	0.3073
4	0.0523	0.0515	0.0490
5	0.9571	1.0610	1.0642

Based on the above estimates for this group of individuals we are now able to derive the generalized optimal BMSs resulting from the Eqs (17, 19 and 18) for the case of the NBI, PIG and Sichel models respectively. These BMSs are presented in Table 7. Note that the explanatory variable Bonus-Malus class varies substantially depending on the number of claims of policyholder i for period j . For this reason in Table 7 we specify the exact order of the claims history in order to derive the scaled premiums that must be paid by this group of policyholders, assuming that the age of the policy is up to 2 years. For example, consider a policyholder who at $t = 2$ has a total number of claims $K = 2$. From Table 7 we observe that if she has claim frequency history $k_1 = 0, k_2 = 2$ then her premium increases from 100 to 361.15, 297.74 and 371.50, in the case of the NBI, PIG and Sichel models respectively. On the contrary, if she has claim frequency history $k_1 = 1, k_2 = 1$ then her premium increases from 100 to 339.90, 265.83 and 318.49 in the case of the NBI, PIG and Sichel models respectively.

Table 7 Women, Horse Power 0-33, Varying Bonus-Malus Class

Year	Number of Claims k_t	Optimal BMS NBI	Optimal BMS PIG	Optimal BMS Sichel
t=0	$k_0=0$	100	100	100
	$k_1=0$	91.93	91.60	98.31
t=1	$k_1=1$	279.57	247.92	297.21
	$k_1=2$	390.29	347.12	463.42
	$k_1=0, k_2=0$	85.06	85.01	92.77
t=2	$k_1=0, k_2=1$	258.69	220.31	257.80
	$k_1=0, k_2=2$	361.15	297.74	371.50
	$k_1=1, k_2=0$	132.56	107.47	121.53
t=2	$k_1=1, k_2=1$	339.90	265.83	318.49
	$k_1=1, k_2=2$	554.13	423.95	520.57
	$k_1=2, k_2=0$	339.90	265.83	318.49
t=2	$k_1=2, k_2=1$	554.13	423.95	520.57
	$k_1=2, k_2=2$	676.59	527.64	672.74

Consider now another group of policyholders who share the following common characteristics. The policyholder i is now a man who has a car with horsepower between 0-33 and his Bonus-Malus class varies over time, starting from BM class 1. The estimation of the vector β^j and thus of the mean parameter, $\hat{\mu}_i^j$, of the NBI, PIG and Sichel distributions respectively led to the following results displayed in Table 8.

Table 8 Estimation of the Mean Parameter, Men, Horse Power 0-33

BM Category	NBI	PIG	Sichel
1	0.1237	0.1215	0.1194
2	0.2272	0.2282	0.2286
3	0.2886	0.2837	0.2795
4	0.0483	0.0472	0.0446
5	0.8844	0.9745	0.9678

Based on the above estimates for this new group of policyholders we can derive the generalized optimal BMSs provided by the Eqs (17, 19 and 18) for the case of the NBI, PIG and Sichel models respectively. In Table 9 we specify again the exact order of the claims history in order to compute the scaled premiums that must be paid by this new group of policyholders assuming again that the age of the policy is up to 2 years. For instance, consider again a policyholder who at $t = 2$ has a total number of claims $K = 2$. From Table 9 we can see that if he has claim frequency history $k_1 = 0, k_2 = 2$ then his premium increases from 100 to 365.29, 304.69 and 384.82, in the case of the NBI, PIG and Sichel models respectively. On the contrary, if he has claim frequency history $k_1 = 1, k_2 = 1$ then his premium increases from 100 to 345.13, 273.55 and 331.80 in the case of the NBI, PIG and Sichel models respectively.

Table 9 Men, Horse Power 0-33, Varying Bonus-Malus Class

Year	Number of Claims k_t	Optimal BMS NBI	Optimal BMS PIG	Optimal BMS Sichel
t=0	$k_0=0$	100	100	100
	$k_1=0$	92..49	92.21	98.85
t=1	$k_1=1$	281.29	250.58	301.71
	$k_1=2$	392.70	351.99	474.61
	$k_1=0, k_2=0$	86.04	85.98	93.72
t=2	$k_1=0, k_2=1$	261.66	224.27	263.82
	$k_1=0, k_2=2$	365.29	304.69	384.82
	$k_1=1, k_2=0$	134.60	109.89	124.95
t=2	$k_1=1, k_2=1$	345.13	273.55	331.80
	$k_1=1, k_2=2$	562.66	438.37	548.06
	$k_1=2, k_2=0$	345.13	273.55	331.80
t=2	$k_1=2, k_2=1$	562.66	438.37	548.06
	$k_1=2, k_2=2$	687.00	547.37	713.46

Overall, from Table 7 and Table 9 we observe that the premiums that should be paid by a woman who has a car with horsepower between 0-33 and her Bonus-Malus class varies over time do not differ much from those that should be paid by a man who shares common characteristics. Note that other combinations of a priori characteristics could be used and also different claim frequency histories. Note also

that the optimal BMSs resulting from Eqs (17, 19 and 18) can be computed using other characteristics that vary by times such as age of drivers, age of car and mileage.

It is interesting to compare these BMSs with those obtained when only the a posteriori classification criteria are used. Using these BMSs we saw from Table 3, Table 4 and Table 5 that a policyholder who at $t = 2$ has two claims faces a malus of 126.05%, 101.63% and 101.93% of the basic premium in the case of the Negative Binomial, Poisson-Inverse Gaussian and Sichel distributions respectively. Using the generalized optimal BMSs based both on the a priori and the a posteriori classification criteria we consider first a woman, who has a car with horsepower between 0-33 and her Bonus-Malus class varies over time, starting from BM class 1. From Table 7 we saw that if at $t = 2$ she has claim frequency history $k_1 = 0, k_2 = 2$, she faces a malus of 261.15%, 197.74% and 271.50% of the basic premium in the case of the NBI, PIG and Sichel GAMLSS respectively, while if she has $k_1 = 1, k_2 = 1$ claim frequency history then she faces a malus of 239.90%, 165.83% and 218.49% of the basic premium in the case of the NBI, PIG and Sichel GAMLSS respectively. Consider also a man, who has a car with horsepower between 0-33 and his Bonus-Malus class varies over time, starting from BM class 1. From Table 9 we saw that if at $t = 2$ he has claim frequency history $k_1 = 0, k_2 = 2$, he faces a malus 265.29%, 204.69% and 284.82% of the basic premium, in the case of the NBI, PIG and Sichel GAMLSS respectively, while if he has $k_1 = 1, k_2 = 1$ claim frequency history then he faces a malus of 245.13%, 173.55% and 231.80% of the basic premium in the case of the NBI, PIG and Sichel GAMLSS respectively. These systems are more fair since they consider all the important a priori and a posteriori information for the number of claims of each policyholder in order to estimate their risk of having an accident and thus they permit the differentiation of the premiums for various number of claims based on the expected claim frequency of each policyholder as this is estimated both from the a priori and the a posteriori classification criteria.

5 Conclusions

In this paper we developed the design of an optimal BMS assuming that the number of claims is distributed according to a Sichel distribution. This system was proposed as an alternative to the optimal BMS resulting from the traditional Negative Binomial distribution, which cannot handle data with a long tail efficiently. We also considered the optimal BMS provided by the Poisson-Inverse Gaussian distribution, which is a special case of the Sichel distribution. These systems were obtained by updating the posterior mean claim frequency, following the setup of Lemaire (1995). We have also considered a generalized BMS that integrates the a priori and the a posteriori information on an individual basis, following the framework developed by Dionne and Vanasse (1989, 1992). This was achieved by using the Sichel GAMLSS to approximate the number of claims as an alternative to the Negative Binomial regression model used by Dionne and Vanasse (1989, 1992). The new model offers the advantage of being able to model count data with high dispersion.

Furthermore, we considered the NBI GAMLSS and the PIG GAMLSS for assessing claim frequency. With the aim of constructing an optimal BMS by updating the posterior mean claim frequency, we adopted the parametric linear formulation and we allowed only the mean parameter to be modelled as a function of the significant a priori rating variables for the number of claims. The modeling results showed that the Sichel distribution and the Sichel GAMLSS provided the best fitting performances for the data set examined in this study. The optimal BMSs obtained have all the attractive properties of the BMSs developed by Lemaire (1995) and Dionne and Vannasse (1989, 1992).

The above design can be employed by insurance companies which are free to set up their own tariff structures and rating policies according to recent European directives. A possible line of future research is the integration of claim severity into the BMSs presented above (see for example Frangos and Vrontos, 2001).

References

1. Akantziliotou, C., Rigby, R.A., Stasinopoulos, D.M.: The R Implementation of Generalized Additive Models for Location, Scale and Shape. In: Stasinopoulos, M and Touloumi, G. (eds.) *Statistical Modelling in Society: Proceedings of the 17th International Workshop on Statistical Modelling*, pp. 75-83. Chania, Greece (2002)
2. Bichsel, F.: Erfahrung-Tarifierung in der Motorfahrzeug-haftpflichtversicherung. *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, 119-129 (1964)
3. Boucher, J. P., Denuit, M., Guillen, M.: Risk Classification for Claim Counts: A Comparative Analysis of Various Zero-Inflated Mixed Poisson and Hurdle Models. *North American Actuarial Journal* **11**(4), 110–131 (2007)
4. Boucher, J. P., Denuit, M., Guillen, M.: Models of Insurance Claim Counts with Time Dependence Based on Generalisation of Poisson and Negative Binomial Distributions. *Variance* **2**(1), 135–162 (2008)
5. Buhlmann, H.: Optimale Pramienstufensysteme. *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker* 193-213 (1964)
6. Coene, G., Doray, L. G.: A Financially Balanced Bonus-Malus System. *ASTIN Bulletin*, **26**, 107-115 (1996)
7. Dean, C., Lawless, J.F. Willmot, G.E.: A mixed Poisson-inverse-Gaussian regression model. *Canadian Journal of Statistics* **17** **2**, 171-181 (1989)
8. Denuit, M., Marechal, X., Pitrebois, S., Walhin, J. F.: *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems*. Wiley, Chichester (2007)
9. Dionne, G., Vannasse, C.: A generalization of actuarial automobile insurance rating models: the negative binomial distribution with a regression component. *ASTIN Bulletin* **19**, 199–212 (1989)
10. Dionne, G., Vannasse, C.: Automobile insurance ratemaking in the presence of asymmetrical information. *Journal of Applied Econometrics* **7**, 149–165 (1992)
11. Frangos, N., Vrontos, S.: 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 (2001)
12. Frangos, N., Tzougas, G., Vrontos, S.: On the Design of Some Optimal Bonus-Malus Systems Using Frequency and Severity Components. In: *Proceedings of the 6th Conference in Actuarial Science and Finance*. Samos, Greece (2010)
13. Gourieroux, C., Montfort, A., Trognon, A.: Pseudo maximum likelihood methods: theory. *Econometrica* **52**, 681–700 (1984 a)

14. Gouriéroux, C., Montfort, A., Trognon, A.: Pseudo maximum likelihood methods: applications to Poisson models. *Econometrica* **52**, 701–720 (1984 b)
15. Hilbe, J. M.: *Negative Binomial Regression Extensions*. Cambridge University Press (2011)
16. Johnson, N. L., Kotz, S., Balakrishnan, N.: *Continuous Univariate Distributions*. Wiley, Chichester (1994)
17. Jørgensen, B.: Statistical Properties of the Generalized Inverse Gaussian Distribution. In: *Lecture Notes in Statistics*: Springer-Verlag, 9, New York (1982)
18. Lawless, J.F.: Negative Binomial Distribution and Mixed Poisson Regression. *Canadian Journal of Statistics* **15**(3) 209–225 (1987)
19. Lemaire, J.: *Bonus-Malus Systems in Automobile Insurance*. Kluwer, Boston (1995)
20. Mahmoudvand, R., Hassani, H.: Generalized Bonus-Malus systems with a frequency and a severity component on an individual basis in automobile insurance. *ASTIN Bulletin* **39**, 307–315 (2009)
21. Mahmoudvand, R., Aziznasiri, S.: Bonus-Malus Systems in Open and Close Portfolios. In: *International Cramér Symposium on Insurance Mathematics*. Stockholm, Sweden (2013)
22. Pinquet, J. et al.: Long-range contagion in automobile insurance data: estimation and implications for experience rating. *ASTIN Bulletin* **31**(2), 337–348 (2001)
23. Rigby, R.A., Stasinopoulos, D.M.: The GAMLSS project— a flexible approach to statistical modelling. In: Klein, B., Korsholm, L. (eds.) *New Trends in Statistical Modelling: Proceedings of the 16th International Workshop on Statistical Modelling*, 249–256, Odense, Denmark (2001)
24. Rigby, R.A., Stasinopoulos, D.M.: Generalized additive models for location, scale and shape, (with discussion). *Applied Statistics* **54**, 507–554 (2005)
25. Rigby, R.A., Stasinopoulos, D.M., Akantziliotou, C.: A framework for modeling overdispersed count data, including the Poisson-shifted generalized inverse Gaussian distribution. *Computational Statistics and Data Analysis* **53**, 381–393 (2008)
26. Rigby, R.A., Stasinopoulos, D.M.: A flexible regression approach using GAMLSS in R. Lancaster (2009) Available via DIALOG.
<http://www.gamlss.org/wp-content/uploads/2013/01/Lancaster-booklet.pdf>. Cited 13 Nov 2009
27. Sichel, H. S.: A bibliometric distribution which really works. *Journal of the American society for information science* **36**(5), 314–321 (1985)
28. Tzougas, G., Frangos, N.: Design of an Optimal Bonus-Malus System Using the Sichel Distribution as a Model of Claim Counts. In: *International Cramér Symposium on Insurance Mathematics*. Stockholm, Sweden (2013)
29. Vuong, Q.: Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* **57**, 307–333 (1989)
30. Walhin, J. F., Paris, J.: Using Mixed Poisson distributions in connection with Bonus-Malus Systems. *ASTIN Bulletin*, **29**, 81–99 (1999)