Risk Classification for Claim Counts and Losses Using Regression Models for Location, Scale and Shape

by George Tzougas, Spyridon Vrontos, and Nicholas Frangos

ABSTRACT

This paper presents and compares different risk classification models for the frequency and severity of claims employing regression models for location, scale and shape. The differences between these models are analyzed through the mean and the variance of the annual number of claims and the costs of claims of the insureds who belong to different risk classes, and interesting results about claiming behavior are obtained. Furthermore, the resulting a priori premiums rates are calculated via the expected value and standard deviation principles with independence between the claim frequency and severity components assumed.

KEYWORDS

Claim frequency, claim severity, regression models for location, scale and shape, a priori risk classification, expected value premium calculation principle, standard deviation premium calculation principle

1. Introduction

The idea behind a priori risk classification is to split an insurance portfolio into classes that consist of risks with all policyholders belonging to the same class paying the same premium. In view of the economic importance of motor third party liability (MTPL) insurance in developed countries, actuaries have made many attempts to find a probabilistic model for the distribution of the number and costs of claims reported by policyholders.

Recent actuarial literature research assumes that the risks can be rated a priori using generalized linear models, GLM (Nelder and Wedderburn 1972) and generalized additive models, GAM (Hastie and Tibshirani 1990). For motor insurance, typical response variables in these regression models are the number of claims (or claim frequency) and its corresponding severity. References for a priori risk classification include, for example, Dionne and Vanasse (1989, 1992), Dean, Lawless, and Willmot (1989), Denuit and Lang (2004), Yip and Yau (2005), and Boucher, Denuit, and Guillen (2007). Dionne and Vanasse used a negative binomial type I regression model. Dean, Lawless, and Willmot used a Poisson-inverse Gaussian regression model. Denuit and Lang used generalized additive models. Yip and Yau presented several parametric zero-inflated count distributions and Boucher, Denuit, and Guillen presented a comparison of various zero-inflated Mixed Poisson and Hurdle Models. Also, a review of actuarial models for risk classification and insurance ratemaking can be found in Denuit et al. (2007).

The models briefly described above assume that only the mean is modeled as a function of risk factors. However, any model for the mean in terms of a priori rating variables indirectly yields a model for scale and/or shape. Also, even if the mean is the most commonly used measure of the expected claim frequency and of the expected claim severity it does not provide a good description of a distribution's scale and shape. The scale and shape parameters are not adequately described due to the unobserved heterogeneity changes with explanatory variables. In this study, we extend this setup by assuming that all the parameters of the claim frequency/severity distributions can be modeled as functions of explanatory variables with parametric linear functional forms. Joint modeling of all the parameters in terms of covariates improves rate making and estimation of the scale and shape of the claim frequency/severity distributions. In light of a priori ratemaking there is a substantial benefit in this approach, since by modeling all the parameters jointly, both mean and variance may be assessed by choosing a marginal distribution and building a predictive model using all the available ratemaking factors as independent variables. In this respect, risk heterogeneity is modeled as the distribution of frequency and/or severity of claims changes between classes of policyholders by a function of the level of ratemaking factors underlying the analyzed classes. We model the claim frequency using the Poisson, negative binomial type II, Delaporte, Sichel and zero-inflated Poisson models and the claim severity using the gamma, Weibull, Weibull type III, generalized gamma and generalized Pareto models. Our contribution puts focus on the comparison of these models through their variance values and not only the mean values as usually considered in risk classification literature. To the best of our knowledge, it is the first time that the variance of the claim frequency and severity is modeled in the context of ratemaking. Furthermore, the variance of the claim frequency and severity is an important risk measure of the specific class of policyholders, as it can provide a measure of the uncertainty regarding the mean claim frequency and the mean claim severity of the specific class, and the difference in the premium that it implies can act as a cushion against adverse experience.

The difference between the premium and the mean loss is the premium loading. Estimates of variance values are produced by employing a parametric regression for the scale and/or the shape parameters in addition to the mean parameter. However, the commonly used specification that only the mean claim frequency/severity is modeled in terms of risk factors was widely accepted for ratemaking. In this respect, a priori ratemaking is refined by taking in to account the

variance values yielded by modeling jointly all the parameters in terms of risk factors. Furthermore, the differences in the variance values alter significantly the premiums calculated through the standard deviation principle since it is understood that in this case the loading is related to the variability of the loss. Thus, joint modeling of location, scale and shape parameters is justified because it enables us to use all the available information in the estimation of these values through the use of the important explanatory variables for the claim frequency and severity, respectively.

The rest of this paper proceeds as follows. Section 2 introduces the alternative distributions we employ for modeling claim frequency and severity. Section 3 contains an application to a data set concerning car-insurance claims at fault. These classification models are compared on the basis of a sample of the automobile portfolio of a major company operating in Greece employing the generalized Akaike information criterion (GAIC) which is valid for both nested or non-nested model comparisons (as suggested by Rigby and Stasinopoulos, 2005 and 2009). The differences between these models are analyzed through the mean and the variance of the annual number of claims and the costs of claims of the policyholders who belong to different risk classes, which are formed by dividing the portfolio into clusters defined by the relevant ratemaking factors. Finally, the resulting premium rates are calculated via the expected value and standard deviation principles with independence between the claim frequency and severity components assumed.

2. Regression models for location, scale and shape

This section summarizes the characteristics of the various count and loss models used in this study. As we have mentioned, in the setup we extend the recent a priori risk classification research by assuming that every parameter of the conditional response frequency/severity distribution is modeled in terms of covariates through the use of known monotonic link functions chosen to ensure a valid range for the distribution parameters.[1]

2.1. Frequency component

Consider a policyholder i whose number of claims, denoted as K_i, are independent, for $i = 1, \ldots, n$. The probability that the policyholder i has reported k claims to the insurer, $k = 0, 1, 2, \ldots$, is denoted by $P(K_i = k)$. In this study, besides the traditional Poisson regression model, we model the claim frequency using a negative binomial type II, Delaporte, Sichel and zero-inflated mixed Poisson regression model for location scale and shape.

- The probability density function (pdf) of the Poisson distribution is given by[2]

$$P(K_i = k) = \frac{e^{-\mu}\mu^k}{k!}. \tag{1}$$

We allow the μ parameter to vary from one individual to another. Let $\mu_i = e_i \exp(c_{1i}\beta_1)$, where e_i denotes the exposure of policy i and where $\beta_1^T(\beta_{1,1}, \ldots, \beta_{1,J_1'})$ is the $1 \times J_1'$ vector of the coefficients. The mean and the variance of K_i are given by[3]

$$E(K_i) = Var(k_i) = \mu_i = e_i \exp(c_{1i}\beta_1). \tag{2}$$

- The pdf of negative binomial type II (NBII) distribution is given by[4]

$$P(K_i = k) = \frac{\Gamma\left(k + \frac{\mu}{\sigma}\right)\sigma^k}{\Gamma\left(\frac{\mu}{\sigma}\right)\Gamma(k+1)[1+\sigma]^{k+\frac{\mu}{\sigma}}}, \tag{3}$$

[1]For more details about the claim frequency/severity models and the associated link functions used in this paper, we refer the reader to Rigby and Stasinopoulos (2005 and 2009).

[2]The Poisson regression model has been widely used by insurance practitioners for modeling claim count data. See, for example, Renshaw (1994).

[3]Equidispersion implied by the Poisson distribution is usually corrected by the introduction of a random variable into the regression component. Then the marginal distribution of the number of claims is a mixed Poisson distribution. For well-known results applied to the above situation, we refer the interested reader to Gourieroux, Montfort and Trognon (1984a, 1984b), Boyer, Dionne, and Vanesse (1992), Lemaire (1995), and Boucher, Denuit, and Guillen (2007, 2008).

[4]This parameterization was used by Evans (1953) as pointed out by Johnson, Kotz, and Balakrishnan (1994). Note also that a negative binomial type I distribution arises if σ is reparameterized to $\sigma_1\mu$. A priori ratemaking using the NBI where regression is not only performed on the mean parameter has been recommended by, for example, Boucher, Denuit, and Guillen (2007, 2008).

for $\mu > 0$ and $\sigma > 0$. Following Rigby and Stasinopoulos (2005; 2009), we assume that $\mu_i = e_i \exp(c_{1i}\beta_1)$ and $\sigma_i = \exp(c_{2i}\beta_2)$, where c_{ji} $(c_{ji,1}, \ldots, c_{ji,J_1'})$ and β_j^T $(\beta_{j,1}, \ldots, \beta_{j,J_1'})$ are the $1 \times J_1'$ vectors of the a priori rating variables and the coefficients respectively, for $j = 1, 2$. The mean and the variance of K_i are given by

$$E(K_i) = e_i \exp(c_{1i}\beta_1) \tag{4}$$

and

$$Var(K_i) = e_i \exp(c_{1i}\beta_1)[1 + \exp(c_{2i}\beta_2)]. \tag{5}$$

- The pdf of the Delaporte distribution is given by[5]

$$P(K_i = k) = \frac{e^{-\mu\nu}}{\Gamma\left(\frac{1}{\sigma}\right)}[1 + \mu\sigma(1-\nu)]^{-\frac{1}{\sigma}} S, \tag{6}$$

where $\sigma_i > 0$ and $0 \le \nu < 1$ and where

$$S = \sum_{m=0}^{k} \binom{k}{m} \frac{\mu^k \nu^{k-m}}{k!} \left[\mu + \frac{1}{\sigma(1-k)}\right]^{-m} \Gamma\left(\frac{1}{\sigma} + m\right). \tag{7}$$

Following Rigby and Stasinopoulos (2008; 2009), we assume that $\mu_i = e_i \exp(c_{1i}\beta_1)$, $\sigma_i = \exp(c_{2i}\beta_2)$ and $\nu_i = \frac{\exp(c_{3i}\beta_3)}{1 + \exp(c_{3i}\beta_3)}$, where c_{ji} $(c_{ji,1}, \ldots, c_{ji,J_j'})$ and $\beta_j^T(\beta_{j,1}, \ldots, \beta_{j,J_j'})$ are the $1 \times J_j'$ vectors of the a priori rating variables and the coefficients respectively, for $j = 1, 2, 3$. The mean and variance of K_i are given by

$$E(K_i) = e_i \exp(c_{1i}\beta_1) \tag{8}$$

and

$$Var(K_i) = e_i \exp(c_{1i}\beta_1) + [e_i \exp(c_{1i}\beta_1)]^2 \exp(c_{2i}\beta_2)\left[1 - \frac{\exp(c_{3i}\beta_3)}{1 + \exp(c_{3i}\beta_3)}\right]^2. \tag{9}$$

- The pdf of the Sichel distribution is given by[6]

$$P(K_i = k) = \frac{\left(\frac{\mu}{c}\right)^k K_{k+\nu}(a)}{k!(a\sigma)^{k+\nu} K_\nu\left(\frac{1}{\sigma}\right)}, \tag{10}$$

where $\sigma > 0$ and $-\infty < \nu < \infty$ and where $c = \frac{K_{\nu+1}\left(\frac{1}{\sigma}\right)}{K_\nu\left(\frac{1}{\sigma}\right)}$ where

$$K_\nu(z) = \frac{1}{2}\int_0^\infty x^{\nu-1} \exp\left[-\frac{1}{2}z\left(x + \frac{1}{x}\right)\right] dx, \tag{11}$$

is the modified Bessel function of the third kind of order ν with argument z and where $a^2 = \sigma^{-2} + 2\mu(c\sigma)^{-1}$. Following Rigby, Stasinopoulos and Akantziliotou (2008) and Rigby and Stasinopoulos (2009), we assume that $\mu_i = e_i \exp(c_{1i}\beta_1)$, $\sigma_i = \exp(c_{2i}\beta_2)$ and $\nu_i = c_{3i}\beta_3$, where c_{ji} $(c_{ji,1}, \ldots, c_{ji,J_j'})$ and $\beta_j^T(\beta_{j,1}, \ldots, \beta_{j,J_j'})$ are the $1 \times J_j'$ vectors of the a priori rating variables and the coefficients respectively, for $j = 1, 2, 3$. The mean and variance of K_i are given by

$$E(K_i) = e_i \exp(c_{1i}\beta_1) \tag{12}$$

and

$$Var(K_i) = e_i \exp(c_{1i}\beta_1) + [e_i \exp(c_{1i}\beta_1)]^2 \left\{\frac{2\exp(c_{2i}\beta_2)[c_{3i}\beta_3 + 1]}{c_i} + \frac{1}{c_i^2} - 1\right\}, \tag{13}$$

where $c_i = \frac{K_{c_{3i}\beta_3+1}\left(\frac{1}{\exp(c_{2i}\beta_2)}\right)}{K_{c_{3i}\beta_3}\left(\frac{1}{\exp(c_{2i}\beta_2)}\right)}$.

[5]This parameterization of Delaporte was given by Rigby, Stasinopoulos and Akantziliotou (2008).

[6]Parameterization (10) was given by Rigby, Stasinopoulos, and Akantziliotou (2008). The use of the Sichel distribution for modeling claim frequency where regression is only performed on the mean parameter has been recommended by Tzougas and Frangos (2014).

- The pdf of the zero-inflated Poisson (ZIP) distribution is given by[7]

$$P(K_i=k)=\begin{cases}\pi+(1-\pi)e^{-\mu}, \text{ if } k=0\\ (1-\pi)\dfrac{e^{-\mu}\mu^k}{k!}, \text{ if } k=1,2,3,\ldots\end{cases} \quad (14)$$

Following Rigby and Stasinopoulos (2005 and 2009), we assume that $\mu_i = e_i \exp(c_{1i}\beta_1)$ and $\pi=\dfrac{\exp(c_{2i}\beta_2)}{1+\exp(c_{2i}\beta_2)}$, where c_{ji} $(c_{ji,1}, \ldots, c_{ji,J_j'})$ and β_j^T $(\beta_{j,1}, \ldots, \beta_{j,J_j'})$ are the $1\times J_j'$ vectors of the a priori rating variables and the coefficients respectively, for $j = 1, 2$. The mean and the variance of K_i are given by

$$E(K_i)=e_i\exp(c_{1i}\beta_1)[1-\exp(c_{2i}\beta_2)] \quad (15)$$

and

$$Var(K_i)=e_i\exp(c_{1i}\beta_1)[1-\exp(c_{2i}\beta_2)][1+e_i\exp(c_{1i}\beta_1)\exp(c_{2i}\beta_2)]. \quad (16)$$

2.2. Severity component

In this section, we need to consider the claim severities. Let $X_{i,k}$ be the cost of the kth claim reported by policyholder i, $i = 1, \ldots, n$ and assume that the individual claim costs $X_{i,1}, X_{i,2}, \ldots$ are independent and identically distributed (i.i.d). Different models are used to describe the behavior of the costs of claims as a function of the explanatory variables including gamma, Weibull, Weibull type III, generalized gamma, and generalized Pareto regression models for location, scale and shape.

- The pdf of the gamma distribution is given by[8]

$$f(x)=\frac{1}{(s^2m)^{\frac{1}{s^2}}}\frac{x^{\frac{1}{s^2}-1}\exp\left(-\dfrac{x}{s^2m}\right)}{\Gamma\left(\dfrac{1}{s^2}\right)}, \quad (17)$$

for $X_{i,k} > 0$, where $m > 0$ and $s > 0$. Following Rigby and Stasinopoulos (2009), we assume that $m_i = \exp(d_{1i}\gamma_1)$ and $s_i=\exp(d_{2i}\gamma_2)$, where d_{ji} $(d_{ji,1}, \ldots, d_{ji,J_j'})$ and $\gamma_j^T(\gamma_{j,1}, \ldots, \gamma_{j,J_j'})$ are the $1\times J_j'$ vectors of the exogenous variables and the coefficients respectively for $j = 1, 2$. The mean and variance of $X_{i,k}$ are given by

$$E(X_{i,k})=\exp(d_{1i}\gamma_1) \quad (18)$$

and

$$Var(X_{i,k})=[\exp(d_{2i}\gamma_2)]^2[\exp(d_{1i}\gamma_1)]^2. \quad (19)$$

- The pdf of the Weibull distribution is given by[9]

$$f(x)=\frac{sx^{s-1}}{m^s}\exp\left[-\left(\frac{x}{m}\right)^s\right], \quad (20)$$

where $m > 0$ and $s > 0$. Following Rigby and Stasinopoulos (2009), we assume that $m_i = \exp(d_{1i}\gamma_1)$ and $s_i = \exp(d_{2i}\gamma_2)$, where d_{ji} $(d_{ji,1}, \ldots, d_{ji,J_j'})$ and $\gamma_j^T(\gamma_{j,1}, \ldots, \gamma_{j,J_j'})$ are the $1\times J_j'$ vectors of the exogenous variables and coefficients respectively, for $j = 1, 2$. The mean and the variance of $X_{i,k}$ are given by

$$E(X_{i,k})=\exp(d_{1i}\gamma_1)\Gamma\left(\frac{1}{\exp(d_{2i}\gamma_2)}+1\right) \quad (21)$$

and

[7]This parameterization was used by Johnson, Kotz, and Balakrishnan (1994) and Lambert (1992). The ZIP model is a special case of a mixed Poisson distribution. However, if overdispersion in the Poisson part is still present then all the distributions seen before can be used since a heterogeneity term may be incorporated in the model. For instance, see Yip and Yau (2005) for an application to insurance claim count data. For more details about zero-inflated count models see Lambert (1992) and Green and Silverman (1994).

[8]We use the parameterization of the two parameter gamma distribution given by Rigby and Stasinopoulos (2009). Note also that a priori ratemaking using the gamma distribution where regression is not only performed on the mean parameter can be found in, for example, Denuit et al. (2007).

[9]The specific parameterization of the two parameter Weibull distribution used here was that used by Johnson, Kotz, and Balakrishnan (1994).

$$Var(X_{i,k}) = [\exp(d_{1i}\gamma_1)]^2 \left\{\Gamma\left(\frac{2}{\exp(d_{2i}\gamma_2)}+1\right) - \left[\Gamma\left(\frac{1}{\exp(d_{2i}\gamma_2)}+1\right)\right]^2\right\}. \quad (22)$$

- The pdf of the Weibull type III (WEI3) distribution is given by[10]

$$f(x) = \frac{s}{m}\Gamma\left(\frac{1}{s}+1\right)\left[\frac{x}{m}\Gamma\left(\frac{1}{s}+1\right)\right]^{s-1} \exp\left\{-\left[\frac{x}{m}\Gamma\left(\frac{1}{s}+1\right)\right]^s\right\}, \quad (23)$$

where $m > 0$ and $s > 0$. Following Rigby and Stasinopolous (2009), we assume that $m_i = \exp(d_{1i}\gamma_1)$ and $s_i = \exp(d_{2i}\gamma_2)$, where d_{ji} $(d_{ji,1}, \ldots, d_{ji,J'_j})$ and $\gamma_j^T(\gamma_{j,1}, \ldots, \gamma_{j,J'_j})$ are the $1 \times J'_j$ vectors of the exogenous variables and the coefficients respectively, for $j = 1, 2$. The mean and the variance of $X_{i,k}$ are given by

$$E(X_{i,k}) = \exp(d_{1i}\gamma_1) \quad (24)$$

and

$$Var(X_{i,k}) = [\exp(d_{1i}\gamma_1)]^2 \left\{\Gamma\left(\frac{2}{\exp(d_{2i}\gamma_2)}+1\right)\left[\Gamma\left(\frac{1}{\exp(d_{2i}\gamma_2)}+1\right)\right]^{-2} - 1\right\}. \quad (25)$$

- The pdf of the generalized gamma (GG) distribution is given by[11]

$$f(x) = \frac{|n|\theta^\theta\left(\frac{x}{m}\right)^{n\theta}\exp\left[-\theta\left(\frac{x}{m}\right)^n\right]}{\Gamma(\theta)x}, \quad (26)$$

where $m > 0, s > 0, -\infty < n < \infty$ and $\theta = \frac{1}{s^2 n^2}$. Following Rigby, Stasinopoulos, and Akantziliotou (2008), we assume that $m_i = \exp(d_{1i}\gamma_1)$, $s_i = \exp(d_{2i}\gamma_2)$ and $n_i = d_{3i}\gamma_3$, where d_{ji} $(d_{ji,1}, \ldots, d_{ji,J'_j})$ and $\gamma_j^T(\gamma_{j,1}, \ldots, \gamma_{j,J'_j})$ are the $1 \times J'_j$ vectors of the exogenous variable and the coefficients respectively, for $j = 1, 2, 3$. The mean and the variance of $X_{i,k}$ are given by

$$E(X_{i,k}) = \frac{\exp(d_{1i}\gamma_1)\Gamma\left(\theta_i + \frac{1}{d_{3i}\gamma_3}\right)}{\theta_i^{\frac{1}{d_{3i}\gamma_3}}\Gamma(\theta_i)} \quad (27)$$

and

$$Var(X_{i,k}) = [\exp(d_{1i}\gamma_1)]^2 \frac{\left\{\Gamma(\theta_i)\Gamma\left(\theta_i + \frac{2}{d_{3i}\gamma_3}\right) - \left[\Gamma\left(\theta_i + \frac{1}{d_{3i}\gamma_3}\right)\right]^2\right\}}{\theta_i^{\frac{2}{d_{3i}\gamma_3}}[\Gamma(\theta_i)]^2}, \quad (28)$$

where $\theta_i = \frac{1}{s_i^2 n_i^2} = \frac{1}{(\exp(d_{2i}\gamma_2))^2 (d_{3i}\gamma_3)^2}$.

- The pdf of the generalized Pareto distribution is given by[12]

$$f(x) = \frac{\Gamma(n+t)}{\Gamma(n)\Gamma(t)}\frac{m^t x^{n-1}}{(x+m)^{n+t}}, \quad (29)$$

where $m > 0$, $n > 0$ and $t > 0$. Following Rigby, Stasinopoulos, and Akantziliotou (2008), we assume that $m_i = \exp(d_{1i}\gamma_1)$, $n_i = \exp(d_{2i}\gamma_2)$ and $t_i = \exp(d_{3i}\gamma_3)$, where d_{ji} $(d_{ji,1}, \ldots, d_{ji,J'_j})$ and $\gamma_j^T(\gamma_{j,1}, \ldots, \gamma_{j,J'_j})$ are the $1 \times J'_j$ vectors of the exogenous variables and the

[10]This is a parameterization of the Weibull distribution where m is the mean of the distribution.

[11]The parameterization of the generalized gamma distribution we use was that used by Lopatatzidis and Green (2000).

[12]The above parameterization of the generalized Pareto distribution can be found, for example, in Klugman, Panjer, and Willmot (2004). Note that if we let $n = 1$ in Eq. (29), the generalized Pareto distribution reduced to the Pareto distribution. The use of the Pareto distribution for modeling claim severity where regression is not only performed on the mean parameter can be found in Frangos and Vrontos (2001).

coefficients respectively for $j = 1, 2, 3$. The mean and variance of $X_{i,k}$ are given by

$$E(X_{i,k}) = \frac{\exp(d_{1i}\gamma_1)\exp(d_{2i}\gamma_2)}{\exp(d_{3i}\gamma_3) - 1} \tag{30}$$

and

$$Var(X_{i,k}) = \frac{[\exp(d_{1i}\gamma_1)]^2 \exp(d_{2i}\gamma_2)}{\exp(d_{3i}\gamma_3) - 1} \left\{ \frac{\exp(d_{2i}\gamma_2) + \exp(d_{3i}\gamma_3) - 1}{[\exp(d_{3i}\gamma_3) - 1][\exp(d_{3i}\gamma_3) - 2]} \right\}. \tag{31}$$

3. Application

The data were kindly provided by a Greek insurance company and concern a motor third party liability insurance portfolio observed during 3.5 years. The data set comprises 15641 policies. Both private cars and fleet vehicles have been considered in this sample.[13] The available a priori rating variables we employ are the Bonus Malus (BM) class,[14] the horsepower (HP) of the car and gender of the driver. Only policyholders with complete records, i.e., with availability of all the variables under consideration were considered. Records for fleet data were not available for the case of the claim frequency. Furthermore, in light of the heterogeneity which exists within the portfolio, consideration was given to grouping the levels of each explanatory variable with respect to risk profiles with similar number and costs of claims at fault reported to the company over the 3.5 years of observation. This was done in order to achieve ratemaking accuracy and homogeneity within rating cells, for the claim frequency and severity component respectively. Also, by balancing homogeneity and sufficiency of the volume of data in each cell credible patterns were provided. As a result of the aforementioned methodology, Bonus-Malus and horsepower variables were segmented into different categories for claim frequency and claim severity component. This will affect the a priori ratemaking, since the claim frequency and severity component will contain a different number of homogeneous classes, generating a ratemaking structure that is fair to the policyholders. Claim counts are modeled for all 15641 policies. The Bonus-Malus class consists of four categories: A, B, C and D, where: A = "drivers who belong to BM classes 1 and 2," B = "drivers who belong to BM classes 3–5," C = "drivers who belong to BM classes 6–9 & 11–20" and D = "drivers who belong to BM class 10." The horsepower of the car consists of three categories: A, B and C, where: A = "drivers who had a car with a HP between 0–33 & 100–132," B = "drivers who had a car with a HP between 34–66" and C = "drivers who had a car with a HP between 67–99." The gender consists of two categories: M = "male" and F = "female" drivers. Regarding the amount paid for each claim, there were 5590 observations that met our criteria. The Bonus-Malus class consists of three categories: A, B and C, where: A = "drivers who belong to BM classes 1 and 2," B = "drivers who belong to BM classes 3–5 & 6–9 & 11–20" and C = "drivers who belong to BM class 10." The horsepower of the car consists of four categories A, B, C and D, where: A = "drivers who had a car with a HP between 100–110 & 111–121 & 122–132," B = "drivers who had a car with a HP between 0–33 & 34–44 & 45–55 & 56–66," C = "drivers who had a car with a HP between 67–74" and D = "drivers who had a car with a HP between 75–82 & 83–90 & 91–99." Finally, the gender consists of three categories: M = "male," F = "female" and B = "both," since in this case, data for fleet vehicles used by either male or female drivers were also available, i.e., shared use.

The claim frequency and severity models presented in Sections 2 and 3 were estimated using the GAMLSS package in software R.[15] The ratio of Bessel functions

[13] All the characteristics we consider are observable.

[14] 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).

[15] Note that the same models can be fitted to larger data sets in order to study the effect of other rating factors such as age of driver, driving experience or driving zone, which have been traditionally used in MTPL insurance.

of the third kind whose orders are different was calculated using the HyperbolicDist package in software R.

3.1. Modeling results

This subsection describes the modeling results of the Poisson, negative binomial type II (NBII), Delaporte (DEL), Sichel and zero-inflated Poisson (ZIP), and gamma (GA), Weibull (WEI), Weibull type III (WEI3), generalized gamma (GG) and generalized Pareto (GP) regression models for location scale and shape that have been applied to model claim frequency and claim severity respectively.

Claim frequency and severity models have been calibrated with respect to GAIC goodness of fit index as suggested by Rigby and Stasinopoulos (2005, 2009). We followed a model selection technique similar to the one presented in Heller et al. (2007).[16] Specifically, our variable selection started with the examination of the mean parameter of each frequency and severity model. This was achieved by adding all available explanatory variables and testing whether the exclusion of each one lowered the Global Deviance, AIC and SBC values. After having selected the best predictor for the mean parameter, we continued in determining the remaining predictors by testing which rating variable between those used in the mean parameter would lead to a further decrease of the GAIC when inserted in the scale and shape parameters of the claim frequency and severity models respectively. Furthermore, if between the same frequency/severity distributions with different parameter specifications several models have similar AIC and SBC values, we preferred the simpler model in order to avoid overfitting. Therefore, the scale and shape parameters of the models have fewer predictors than the mean parameter (see Tables 1 and 2). In the above respect, the final claim frequency and severity models we selected are those that yield the lowest Global deviance (DEV), Akaike information criterion (AIC), and Bayesian information criterion (BIC) values. Also, every explanatory variable they contain is statistically significant at a 5% threshold.

Tables 1 and 2 summarize our findings with respect to the aforementioned claim frequency and severity models respectively.[17]

From Table 1 we observe, for all frequency models, that BM category A, HP category A and male drivers are the reference categories of μ. HP category A and male drivers are the reference categories for σ in the case of the NBII model. HP category A is the reference category for σ in the case of the Delaporte and Sichel models. BM category A and male drivers are the reference categories for σ in the case of the ZIP model. Furthermore, we see that HP category appears in model equations for both μ and σ in the case of the NBII, Delaporte and Sichel models. Gender appears in model equations for both μ and σ in the case of the NBII and ZIP models. BM category appears in the models equation for both μ and σ in the case of the ZIP model. These a priori rating variables do not always have a similar effect (positive and/or negative) on μ and σ.

The results summarized in Table 2 show that BM category A, HP category A and fleet vehicles used by both male or female drivers are the reference categories for m and s in the case of gamma, Weibull, Weibull type III and generalized gamma models. BM category A, HP category A and fleet vehicles are the reference categories for m and n, and BM category A and HP category A are the reference categories for t in the case of the generalized Pareto model. Note also that BM category, HP category, and gender appear in the model equations for both m and s in the case of the gamma, Weibull and Weibull type III and generalized gamma models. Furthermore, in the case of the generalized gamma model, BM category and gender are also in the model equations for n. Finally, in the case of the generalized Pareto model we observe that

[16]Heller et al. (2007) used generalized additive models for location scale and shape (GAMLSS) for the statistical analysis of the total amount of insurance paid out on a policy.

[17]Note that in Tables 1 and 2 the significant at a probability level of 5% p-values are included in parentheses.

Table 1. Results of the fitted claim frequency models

Poisson		NBII		DEL		Sichel		ZIP	
Variable μ	Estimate	Variable μ	Estimate	Variable μ	Estimate	Variable μ	Estimate	Variable μ	Estimate
Intercept	–0.8150 (0.0000)	Intercept	–0.8131 (0.0000)	Intercept	–0.8221 (0.0000)	Intercept	–0.8201 (0.0000)	Intercept	–0.2210 (0.0000)
BM Cat.		BM Cat.		BM Cat.		BM Cat.		BM Cat.	
B	0.6078 (0.0000)	B	0.6328 (0.0000)	B	0.6429 (0.0000)	B	0.6387 (0.0000)	B	0.1571 (0.0000)
C	0.8834 (0.0000)	C	0.8388 (0.0000)	C	0.8679 (0.0000)	C	0.8694 (0.0000)	C	0.7160 (0.0000)
D	–0.9423 (0.0000)	D	–0.9736 (0.0000)	D	–0.9561 (0.0000)	D	–0.9804 (0.0000)	D	–0.2085 (0.0021)
HP Cat.		HP Cat.		HP Cat.		HP Cat.		HP Cat.	
B	–02371 (0.0000)	B	–0.2351 (0.0000)	B	–0.2434 (0.0000)	B	–0.2458 (0.0000)	B	–0.2492 (0.0000)
C	–0.0725 (0.0120)	C	–0.0730 (0.0318)	C	–0.0742 (0.0403)	C	–0.0759 (0.0357)	C	–0.0939 (0.0005)
Gender		Gender		Gender		Gender		Gender	
F	0.0683 (0.0044)	F	0.0687 (0.0107)	F	0.0880 (0.0010)	F	0.0908 (0.0013)	F	–0.1010 (0.0000)
—	—	Variable σ	Estimate	Variable σ	Estimate	Variable σ	Estimate	Variable σ	Estimate
—	—	Intercept	–0.3728 (0.0000)	Intercept	1.5821 (0.0000)	Intercept	1.2100 (0.0158)	Intercept	–0.2036 (0.0000)
—	—	HP Cat.		HP Cat.		HP Cat.		BM Cat.	
—	—	B	–0.7777 (0.0000)	B	–0.9700 (0.0000)	B	–1.664 (0.0024)	B	–2.8671 (0.0000)
—	—	C	–0.6716 (0.0000)	C	–0.8971 (0.0000)	C	–1.598 (0.0018)	C	–0.4926 (0.0000)
—	—	Gender		Parameter ν	Estimate	Parameter ν	Estimate	D	1.2694 (0.0000)
—	—	F	–0.4313 (0.0005)	Intercept	–0.2013 (0.0021)	Intercept	–2.1040 (0.0000)	Gender	–0.5648 (0.0000)
—	—	—	—	—	—	—	—	F	—

Table 2. Results of the fitted claim severity models

GA		WEI		WEI3		GG		GP	
Variable *m*	Estimate	Variable *m*	Estimate	Variable *m*	Estimate	Variable *m*	Estimate	Variable *m*	Estimate
Intercept	6.3699 (0.0000)	Intercept	6.4939 (0.0000)	Intercept	6.3880 (0.0000)	Intercept	6.3277 (0.0000)	Intercept	7.2849 (0.0000)
BM Cat.		BM Cat.		BM Cat.		BM Cat.		BM Cat.	
B	–0.6786 (0.0000)	B	–0.7118 (0.0000)	B	–0.6649 (0.0000)	B	–1.2020 (0.0000)	B	–1.8305 (0.0000)
C	0.0294 (0.0103)	C	0.0307 (0.0203)	C	0.0312 (0.0192)	C	0.0548 (0.0000)	C	0.0734 (0.0000)
HP Cat.		HP Cat.		HP Cat.		HP Cat.		HP Cat.	
B	–0.6833 (0.0000)	B	–0.6838 (0.0000)	B	–0.6968 (0.0000)	B	–0.6223 (0.0000)	B	–0.3370 (0.0000)
C	–0.5807 (0.0000)	C	–0.5851 (0.0000)	C	–0.5978 (0.0000)	C	–0.5142 (0.0000)	C	–0.2263 (0.0000)
D	–0.4082 (0.0000)	D	–0.4066 (0.0000)	D	–0.4208 (0.0000)	D	–0.3608 (0.0000)	D	–0.1463 (0.0000)
Gender		Gender		Gender		Gender		Gender	
M	–0.1127 (0.0002)	M	–0.1166 (0.0005)	M	–0.1184 (0.0003)	M	–0.1839 (0.0000)	M	–0.4307 (0.0000)
F	–0.0711 (0.0206)	F	–0.0790 (0.0202)	F	–0.0798 (0.0174)	F	–0.1602 (0.0006)	F	–0.4227 (0.0006)
Variable *s*	Estimate	Variable *s*	Estimate	Variable *s*	Estimate	Variable *s*	Estimate	Variable *n*	Estimate
Intercept	–0.4621 (0.0000)	Intercept	0.3899 (0.0000)	Intercept	0.3883 (0.0000)	Intercept	–0.4366 (0.0000)	Intercept	1.3215 (0.0000)
BM Cat.		BM Cat.		BM Cat.		BM Cat.		BM Cat.	
B	0.5946 (0.0000)	B	–0.5492 (0.0000)	B	–0.5498 (0.0000)	B	0.5872 (0.0000)	B	–0.7347 (0.0000)
C	–0.0443 (0.0308)	C	0.0455 (0.0216)	C	0.0442 (0.0261)	C	–0.0520 (0.0224)	C	0.0445 (0.0024)
HP Cat.		HP Cat.		HP Cat.		HP Cat.		HP Cat.	
B	–0.3130 (0.0000)	B	0.4145 (0.0000)	B	0.4139 (0.0000)	B	–0.2622 (0.0000)	B	0.2362 (0.0000)
C	–0.3797 (0.0000)	C	0.4199 (0.0000)	C	0.4197 (0.0000)	C	–0.3410 (0.0000)	C	0.2984 (0.0000)
D	–0.2535 (0.0000)	D	0.2806 (0.0000)	D	0.2799 (0.0000)	D	–0.2311 (0.0000)	D	0.2250 (0.0000)
Gender		Gender		Gender		Gender		Gender	
M	–0.1589 (0.0000)	M	0.0962 (0.0135)	M	0.0975 (0.0123)	M	–0.2133 (0.0000)	M	0.3062 (0.0000)
F	–0.1788 (0.0000)	F	0.0967 (0.0153)	F	0.1016 (0.0109)	F	–0.2423 (0.0000)	F	0.3400 (0.0000)
—	—	—	—	—	—	Variable *n*	Estimate	Variable *t*	Estimate
—	—	—	—	—	—	Intercept	0.7189 (0.0001)	Intercept	2.3395 (0.0000)
—	—	—	—	—	—	BM Cat.		BM Cat.	
—	—	—	—	—	—	B	–0.9809 (0.0014)	B	–1.5622 (0.0000)
—	—	—	—	—	—	C	0.2763 (0.0056)	C	0.0537 (0.0000)
—	—	—	—	—	—	Gender		HP Cat.	
—	—	—	—	—	—	M	–0.3272 (0.0246)	B	0.5190 (0.0000)
—	—	—	—	—	—	F	–0.3516 (0.0321)	C	0.5859 (0.0000)
—	—	—	—	—	—	—	—	D	0.4332 (0.0000)

BM category, HP category and gender appear in the model equations for both m and n, and BM category and HP category are in the model equations for t. These explanatory variables do not always have the same effect (positive and/or negative) on the parameters m, s, n and t.

Most of the models presented in Tables 1 and 2, their reparameterizations and special cases have already been employed for modeling claim frequency/severity data. However, as we have already mentioned, the commonly used specification that only the mean claim frequency/severity is modeled in terms of risk factors was widely accepted for ratemaking. Also, the results for the location parameter of the claim frequency/severity models are in line with the existing results, based on the examination of the relative data sets, in recent actuarial literature research. Specifically, as expected, the values of the estimated regression coefficients of the explanatory variables for this parameter will lead to mean claim frequency/severity values which will not differ much under different distributional assumptions. Within the framework we adopted, the systematic part of these models was expanded to allow modeling of all the parameters of the claim frequency/severity distribution as functions of a priori rating variables. This approach is especially suited to modeling insurance response data which often exhibit heterogeneity, i.e., a situation where the scale or shape of the distribution of the response variable changes with explanatory variables. Furthermore, joint modeling of all the parameters in an a priori ratemaking scheme breaks the nexus between the mean and variance implied by the standard procedure using GLM models, leading to a more complete comparison of these models through their variance values. Finally, in this way we will be able to use all the available information in the estimation of the claim frequency/severity distribution in order to group risks with similar risk characteristics and to establish fair premium rates. Furthermore, our analysis shows that the employment of more advanced models that capture the stylized characteristics of the data is beneficial for the insurance company.

3.2. Models comparison

So far, we have several competing models for the claim frequency and severity components. The differences between models produce different premiums. Consequently, to distinguish between these models, this section compares them so as to select the best for each case. As suggested by Rigby and Stasinopoulos (2005; 2009) the models have been calibrated with respect to generalized Akaike information criterion (GAIC) which is valid for both nested or non-nested model comparisons. The generalized Akaike information criterion (GAIC) is defined as

$$GAIC = \hat{D} + \kappa \times df, \tag{32}$$

where $\hat{D} = -2\hat{l}$ is the fitted (global) deviance, $\hat{l}$ is the fitted log-likelihood, df is the degrees of freedom used in the model (i.e., the sum of the degrees of freedom used for the location, scale and shape parameters) and κ is a constant. The Akaike information criterion (AIC) and the Schwartz Bayesian criterion (SBC) are special cases of the GAIC. Specifically, if we let $\kappa = 2$ we have the AIC, while if we let $\kappa = \log(n)$ we have the SBC.

The resulting Global Deviance, AIC and SBC are given in Table 3 for the different claim frequency (Panel A) and claim severity (Panel B) fitted models.

Table 3. Models comparison

Model	df	Global Deviance	AIC	SBC
Panel A: Claim Frequency Models				
Poisson	7	29115.29	29129.29	29182.90
NBII	11	28323.32	28345.32	28429.55
Delaporte	11	28357.99	28379.99	28464.23
Sichel	11	28348.97	28370.97	28455.20
ZIP	12	28503.22	28527.22	28619.11
Panel B: Claim Severity Models				
Gamma	16	69665.05	69697.05	69803.11
WEI	16	70794.96	70826.96	70933.02
WEI3	16	70793.02	70825.02	70931.08
GG	21	69427.16	69469.16	69608.37
GP	22	69582.12	69526.12	69771.96

Overall, with respect to the Global Deviance, AIC and SBC indices, from Panel A we observe the best fitted claim frequency model is the negative binomial type II model, followed closely by the Sichel and Delaporte models. From the claim severity models in Panel B we see that the best fitting performances are provided by the generalized gamma model followed by the generalized Pareto and gamma models. Negative binomial type II and generalized gamma capture more efficiently the stylized characteristics of the data, such as overdispersion of the number of claims and the tail behavior of losses and performed better than the other distributions.

3.3. A priori risk classification

In this subsection differences between the claim frequency and severity models, presented in Sections 2 and 3 respectively, are analyzed through the mean and the variance of the number and costs of claims of the policyholders who belong to different risk classes, which are determined by the availability of the relevant a priori characteristics.

The final a priori ratemaking for the claim frequency models contains 24 classes. The estimated expected annual claim frequency and the variance for each risk class are obtained by Eqs (2, 4, 8, 12 and 15) and the Eqs (2, 5, 9, 13 and 16) for the case of the Poisson, negative binomial type II (NBII), Delaporte (DEL), Sichel and zero-inflated Poisson (ZIP) model respectively. The results are summarized in Table 4. As expected, the variance of the NBII, Delaporte, Sichel and ZIP model exceeds the mean and these models allow for overdispersion. Furthermore, we observe that the biggest differences lie in the variance values of these models. For example, the variance of the expected number of claims for a man who belongs to BM category A and has a car that belongs to HP category A, i.e., for the reference class, is equal to 0.1264, 0.2140, 0.1868, 0.1884 and 0.1391 while the variance of the expected number of claims for a woman who shares common characteristics is equal to 0.1354, 0.1964, 0.2100, 0.2128 and 0.1507 in the case of the Poisson, NBII, Delaporte, Sichel and ZIP model, respectively.

The final a priori ratemaking for the claim severity models contains 36 classes. Table 5 gives the estimated expected claim severity and the variance for each risk class obtained from the gamma (GA), Weibull (WEI), Weibull type III (WEI3), generalized gamma (GG) and generalized Pareto (GP) model according to the Eqs (18, 21, 24, 27 and 30) and the Eqs (19, 22, 25, 28 and 31) respectively. As expected, similarly to the case of the claim frequency models, we see that the biggest differences between the claim severity models lie in their variance values. For instance, the variance of the expected claim costs for a fleet vehicle that belongs to HP category A, used by both a man and a woman, and belongs to BM category A, i.e., for the reference class, is equal to 135347.30, 169637.36, 168267.90, 148196.45 and 142078.20, while the variance of the expected claim costs for a private car that belongs to HP category A and is used by a man who belongs to BM category A is equal to 78621.46, 110315.30, 111018.27, 72875.39 and 89891.64 in the case of the gamma, WEI, WEI3, generalized gamma and generalized Pareto model.

Overall, the results summarized in Tables 4 and 5 show the following trends by type of frequency/severity model as to which the lowest/highest variances are observed. First, from Table 4 we see that the NBII model has the highest variance values among all models in eleven risk classes. The Delaporte model has the highest variance values among all models in six risk classes, while it has the lowest variance value among all mixed Poisson models[18] in one risk class. The Sichel model has the highest variance values among all models in five risk classes, while it has the lowest variance values among all mixed Poisson models in eight risk classes. The ZIP model has the highest variance values among all models in two risk classes, while it has the lowest variance values among all mixed Poisson models in fifteen risk classes. Second, from Table 5 we observe that the gamma model has the highest variance value among

[18]The Poisson regression model has the lowest variance values among all models since they are equal to its mean values.

Table 4. A priori risk classification using claim frequency models

Risk Class		Poisson		NBII		DEL		Sichel		ZIP	
		Mean	Var	Mean	Var	Mean	Var	Mean	Var	Mean	Var
1	BMA, HP A, M	0.1264	0.1264	0.1267	0.2140	0.1255	0.1868	0.1258	0.1884	0.1261	0.1391
2	BMA, HP A, W	0.1354	0.1354	0.1357	0.1964	0.1371	0.2100	0.1377	0.2128	0.1414	0.1507
3	BMA, HP B, M	0.0997	0.0997	0.1001	0.1318	0.0984	0.1127	0.0984	0.1046	0.0983	0.1062
4	BMA, HP B, W	0.1068	0.1068	0.1072	0.1293	0.1075	0.1245	0.1078	0.1152	0.1102	0.1158
5	BMA, HP C, M	0.1176	0.1176	0.1178	0.1592	0.1165	0.1381	0.1166	0.1260	0.1148	0.1256
6	BMA, HP C, W	0.1259	0.1259	0.1261	0.1550	0.1273	0.1529	0.1277	0.1390	0.1288	0.1365
7	BMB, HP A, M	0.2323	0.2323	0.2385	0.4029	0.2388	0.4602	0.2383	0.4629	0.2742	0.2777
8	BMB, HP A, W	0.2486	0.2486	0.2555	0.3699	0.2608	0.5247	0.2610	0.5302	0.2527	0.2543
9	BMB, HP B, M	0.1832	0.1832	0.1885	0.2483	0.1872	0.2388	0.1863	0.2089	0.2136	0.2158
10	BMB, HP B, W	0.1961	0.1961	0.2020	0.2435	0.2044	0.2659	0.2040	0.2311	0.1969	0.1980
11	BMB, HP C, M	0.2160	0.2160	0.2217	0.2998	0.2217	0.2995	0.2208	0.2548	0.2496	0.2524
12	BMB, HP C, W	0.2312	0.2312	0.2375	0.2918	0.2422	0.3349	0.2418	0.2825	0.2300	0.2314
13	BMC, HP A, M	0.3059	0.3059	0.2931	0.4950	0.2991	0.6462	0.3001	0.6564	0.3127	0.3616
14	BMC, HP A, W	0.3276	0.3276	0.3140	0.4545	0.3266	0.7406	0.3286	0.7559	0.3301	0.3610
15	BMC, HP B, M	0.2413	0.2413	0.2317	0.3050	0.2344	0.3153	0.2347	0.2705	0.2438	0.2734
16	BMC, HP B, W	0.2584	0.2584	0.2482	0.2992	0.2560	0.3525	0.2571	0.2999	0.2573	0.2761
17	BMC, HP C, M	0.2845	0.2845	0.2725	0.3684	0.2777	0.3997	0.2782	0.3320	0.2847	0.3252
18	BMC, HP C, W	0.3047	0.3047	0.2919	0.3586	0.3032	0.4487	0.3047	0.3692	0.3005	0.3261
19	BMD, HP A, M	0.0493	0.0493	0.0478	0.0808	0.0482	0.0573	0.0486	0.0579	0.0476	0.0542
20	BMD, HP A, W	0.0527	0.0527	0.0512	0.0742	0.0527	0.0634	0.0532	0.0645	0.0634	0.0701
21	BMD, HP B, M	0.0388	0.0388	0.0378	0.0498	0.0378	0.0399	0.0380	0.0389	0.0371	0.0411
22	BMD, HP B, W	0.0416	0.0416	0.0405	0.0489	0.0413	0.0438	0.0417	0.0427	0.0494	0.0534
23	BMD, HP C, M	0.0458	0.0458	0.0444	0.0601	0.0448	0.0480	0.0450	0.0465	0.0433	0.0488
24	BMD, HP C, W	0.0490	0.0490	0.0476	0.0585	0.0489	0.0527	0.0493	0.0510	0.0577	0.0632

all models in one risk class, while it has the lowest variance values among all models in fourteen risk classes. The Weibull model has the highest variance values among all models in five risk classes. The Weibull type III model has the highest variance values among all models in ten risk classes. The generalized gamma model has the lowest variance values among all models in nineteen risk classes. The generalized Pareto model has the highest variance value among all models in twenty risk classes, while it has the lowest variance values among all models in three risk classes.

The claim frequency and severity models are better compared through their variance values, leading to a better classification of the policyholders and thus modeling jointly the location, scale and shape parameters in terms of a priori rating variables is justified because it enables us to use all the available information in the estimation of these values through the use of the important a priori rating variables for the number and the costs of claims respectively.

3.4. Calculation of the premiums according to the expected value and standard deviation principles

Consider a policyholder i who belongs to a group of policyholders, whose number of claims, denoted as K_i, are independent, for $i = 1, \ldots, n$. Let $X_{i,k}$ be the cost of the kth claim reported by the policyholder i and assume that the individual claim costs $X_{i,1}$,

Table 5. A priori risk classification using claim severity models

Risk Class		GA		WEI		WEI3		GG		GP	
		Mean	Var	Mean	Var	Mean	Var	Mean	Var	Mean	Var
1	BMA, HP A, B	584.00	135347.30	597.96	169637.36	594.66	168267.90	591.62	148196.45	583.03	142078.20
2	BMA, HP A, M	521.75	78621.46	526.73	110315.30	528.26	111018.27	504.93	72875.39	514.78	89891.64
3	BMA, HP A, W	543.92	82108.76	546.89	118812.19	549.06	119033.67	516.38	72022.76	536.75	95624.76
4	BMA, HP B, B	294.89	18453.33	295.51	19539.26	296.25	19714.32	310.72	24073.97	300.72	26138.91
5	BMA, HP B, M	263.46	10719.29	263.36	13061.64	263.17	13063.90	262.37	11431.24	265.51	16207.29
6	BMA, HP B, W	274.65	11194.75	273.45	14069.47	273.53	14009.16	268.44	11300.70	276.84	17199.88
7	BMA, HP C, B	326.75	19827.00	326.18	23575.68	327.07	23782.38	344.55	25257.37	333.03	29934.69
8	BMA, HP C, M	291.93	11517.24	290.72	15762.85	290.55	15759.88	290.30	11905.58	294.05	18551.62
9	BMA, HP C, W	304.32	12028.09	301.85	16979.11	301.99	16900.22	297.05	11770.71	306.59	19686.62
10	BMA, HP D, B	388.27	36033.34	390.33	43363.58	390.41	43561.39	404.41	43421.71	394.23	46566.35
11	BMA, HP D, M	346.88	20931.28	346.96	28820.08	346.82	28847.75	341.83	20685.10	348.08	29009.46
12	BMA, HP D, W	361.62	21859.70	360.26	31043.01	360.47	30934.51	349.72	20448.12	362.94	30803.37
13	BMB, HP A, B	296.28	114416.43	352.27	172055.65	305.85	130297.57	265.02	129671.66	250.44	178704.02
14	BMB, HP A, M	264.70	66462.96	297.20	100325.75	271.70	84002.18	164.63	25281.89	221.13	121573.35
15	BMB, HP A, W	275.95	69410.96	308.51	107997.62	282.39	89988.87	165.62	23924.98	230.56	130384.50
16	BMB, HP B, B	149.60	15599.59	151.45	13989.85	152.36	14234.31	119.36	13878.38	108.56	11957.62
17	BMB, HP B, M	133.66	9061.59	132.51	8946.20	135.36	9359.92	83.06	3737.71	95.85	7832.20
18	BMB, HP B, W	139.34	9463.52	137.58	9634.51	140.68	10034.46	84.12	3595.64	99.94	8364.86
19	BMB, HP C, B	165.77	16760.83	166.98	16833.12	168.23	17162.40	127.92	13265.26	118.52	12850.63
20	BMB, HP C, M	148.10	9736.14	146.14	10772.70	149.44	11287.22	91.28	3837.95	104.64	8402.63
21	BMB, HP C, W	154.39	10167.99	151.73	11601.59	155.32	12100.73	92.58	3705.93	109.11	8972.35
22	BMB, HP D, B	196.98	30460.93	206.75	33670.27	200.79	31936.98	157.66	26065.04	145.27	23671.24
23	BMB, HP D, M	175.98	17694.34	179.31	21059.67	178.37	20903.82	108.54	6804.49	128.26	15622.57
24	BMB, HP D, W	183.46	18479.18	186.15	22677.63	185.39	22406.46	109.84	6535.28	133.74	16699.22
25	BMC, HP A, B	601.42	131373.54	613.31	164111.60	613.51	165097.24	591.91	131860.30	618.24	151126.30
26	BMD, HP A, M	537.32	76313.11	541.27	107216.06	545.01	109018.66	511.81	65142.30	545.87	95523.00
27	BMD, HP A, W	560.14	79698.02	561.99	115476.70	566.45	116893.25	524.41	64612.06	569.17	101603.68
28	BMD, HP B, B	303.69	17911.53	304.87	19167.52	305.64	19385.37	317.57	22467.66	319.63	28068.18
29	BMD, HP B, M	271.32	10404.57	271.88	12831.92	271.51	12847.07	270.40	10712.80	282.22	17391.14
30	BMD, HP B, W	282.84	10866.07	282.31	13822.11	282.20	13776.66	276.98	10614.65	294.27	18454.66
31	BMD, HP C, B	336.50	19244.87	336.52	23129.37	337.44	23385.76	353.80	23820.14	354.06	32168.91
32	BMD, HP C, M	300.64	11179.09	300.14	15486.68	299.76	15498.33	300.25	11270.97	312.61	19922.38
33	BMD, HP C, W	313.40	11674.94	311.64	16681.73	311.56	16619.75	307.55	11165.35	325.96	21139.48
34	BMD, HP D, B	399.85	34975.39	402.16	42412.94	402.78	42819.65	412.50	40339.87	418.90	49941.43
35	BMD, HP D, M	357.23	20316.73	357.83	28251.56	357.81	28364.50	351.74	19299.08	369.87	31088.48
36	BMD, HP D, W	372.41	21217.89	371.54	30430.94	371.89	30416.57	360.31	19124.05	385.65	33007.98

$X_{i,2}, \ldots, X_{i,n}$ are independent. It is assumed that the number of claims of each policyholder that belongs to a certain group is independent of the severity of each claim in order to deal with the frequency and the severity components separately.

A premium principle is a rule for assigning a premium to an insurance risk. In this section the premiums rates will be calculated via two well-known premium principles, the expected value and the standard deviation premium principles. More details about the use of the expected value premium principle in MTPL insurance can be found in Lemaire (1995). Furthermore, regarding the use of the standard deviation premium principle one can refer to Bühlmann (1970) and Lemaire (1995) who used the variance principle in MTPL insurance, which is closely related to the standard deviation principle. The standard deviation principle can be used as an alternative and complementary of the expected value principle. It provides a more complete picture to the actuary since it takes into account an additional characteristic of the distribution, i.e., the standard deviation of the number of claims and of losses.

- The premium rates calculated according to the expected value principle are given by

$$P_1 = (1 + w_1)\, E(K_i)(1 + w_2)\, E(X_{i,k}), \qquad (33)$$

where $w_1 > 0$ and $w_2 > 0$ are risk loads.

- The premium rates calculated according to the standard deviation principle are given by

$$P_2 = \left[E(K_i) + \omega_1 \sqrt{Var(K_i)}\right] \left[E(X_{i,k}) + \omega_2 \sqrt{Var(X_{i,k})}\right], \qquad (34)$$

where $\omega_1 > 0$ and $\omega_2 > 0$ are risk loads.

In the following example (Table 6), six different groups of policyholders have been considered. In Table 6 a YES indicates the presence of the characteristic corresponding to the column.

We will calculate the premiums P_1 and P_2 that must be paid by a specific group of policyholders based on the alternative models for assessing claim frequency and the various claim severity models. We assume that

Table 6. The six different groups of policyholders to be compared

Group	BM Category A	HP 0-33	HP 34-66	HP 100-132	Male	Female
1	YES	YES	NO	NO	YES	NO
2	YES	YES	NO	NO	NO	YES
3	YES	NO	YES	NO	YES	NO
4	YES	NO	YES	NO	NO	YES
5	YES	NO	NO	YES	YES	NO
6	YES	NO	NO	YES	NO	YES

$w_1 = w_2 = \omega_1 = \omega_2 = \frac{1}{10}$. The premiums P_1 and P_2 are obtained in Table 7 by substituting into Eqs (33 and 34) the corresponding $E(K_i)$ and $Var(K_i)$, and $E(X_{i,k})$ and $Var(X_{i,k})$ values to these six different groups of policyholders, which were displayed in Tables 4 and 5 for the case of the Poisson, NBII, Delaporte, Sichel and ZIP, and the gamma, Weibull, Weibull type III, generalized gamma and generalized Pareto regression models for location scale and shape respectively.

From Table 7 consider, for instance, a man who belongs to BM category A and has a car with a HP between 34–66. In the case of the Poisson model and the corresponding claim severity models, P_1 is equal to 31.78, 31.77, 31.75, 31.65 and 32.03 euros, while P_2 equals 35.95, 36.07, 36.05, 35.85 and 36.5 euros. In the case of the NBII model and the corresponding claim severity models, P_1 is equal to 31.91, 31.90, 31.88, 31.78 and 32.16 euros, while P_2 equals 37.35, 37.48, 37.46, 37.25 and 37.95 euros. In the case of the Delaporte model and the corresponding claim severity models, P_1 is equal to 31.37, 31.36, 31.33, 31.24 and 31.61 euros, while P_2 equals 36.14, 36.26, 36.24, 36.04 and 36.72 euros. In the case of the Sichel model and the corresponding severity models, P_1 is equal to 31.37, 31.36, 31.33, 31.24 and 31.61 euros, while P_2 equals 35.80, 35.93, 35.90, 35.70 and 36.38 euros. In the case of the ZIP model and the corresponding claim severity models, P_1 is equal to 31.34, 31.33, 31.30, 31.20 and 31.58 euros, while P_2 equals 35.84, 35.97, 35.94, 35.74 and 36.42 euros. Overall, we observe that all the claim frequency models which were combined with the generalized gamma model for assessing claim severity have the lowest P_1 and P_2 values among their

Table 7. Premium rates calculated via the expected value and standard deviation principles

	PO-GA		PO-WEI		PO-WEI3		PO-GG		PO-GP	
Group	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
1	40.2946	44.3448	40.2793	44.5028	40.2503	44.4722	40.1279	44.2231	40.6082	45.0619
2	44.9970	49.1158	44.8004	49.1297	44.8135	49.1391	43.9796	48.0550	45.3558	49.9293
3	31.7830	35.9450	31.7710	36.0730	31.7480	36.0482	31.6515	35.8463	32.0303	36.5261
4	35.4925	39.7840	35.3374	39.7953	35.3477	39.8030	34.6900	38.9248	35.7755	40.4430
5	79.7985	89.0400	80.5602	90.6845	80.7942	90.9493	77.2260	86.1468	78.7325	88.2257
6	89.1126	98.5955	89.5992	100.1082	89.9547	100.4874	84.6006	93.5402	87.9379	97.7515
	NBII-GA		NBII-WEI		NBII-WEI3		NBII-GG		NBII-GP	
	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
1	40.3903	47.3588	40.3750	47.5275	40.3458	47.4948	40.2232	47.2290	40.7045	48.1246
2	45.0967	51.3464	44.9000	51.3610	44.9128	51.3708	44.0770	50.2374	45.4563	52.1968
3	31.9105	37.3493	31.8984	37.4824	31.8754	37.4566	31.7785	37.2468	32.1588	37.9532
4	35.6254	40.8331	35.4700	40.8447	35.4801	40.8525	34.8200	39.9513	35.9095	41.5094
5	79.9880	95.0917	80.7514	96.8480	80.9860	97.1309	77.4093	92.0020	78.9194	94.2221
6	89.3100	103.0732	89.7977	104.6550	90.1540	105.0510	84.7881	97.7883	88.1327	102.1909
	DEL-GA		DEL-WEI		DEL-WEI3		DEL-GG		DEL-GP	
	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
1	40.0077	46.1980	39.9925	46.3625	39.9637	46.3306	39.8422	46.0711	40.3190	46.9449
2	45.5620	52.1760	45.3630	52.1908	45.3762	52.2008	44.5318	51.0492	45.9253	53.0402
3	31.3686	36.1354	31.3567	36.2641	31.3341	36.2392	31.2388	36.0362	31.6127	36.7197
4	35.7251	40.7265	35.5690	40.7381	35.5794	40.7459	34.9173	39.8470	36.0100	41.4011
5	79.2304	92.7607	79.9866	94.4740	80.2190	94.7500	76.6762	89.7467	78.1720	91.9124
6	90.2314	104.7387	90.7241	106.3456	91.0841	106.7484	85.6628	99.3684	89.0420	103.8421
	SI-GA		SI-WEI		SI-WEI3		SI-GG		SI-GP	
	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
1	40.1034	46.3306	40.0881	46.4957	40.0592	46.4637	39.9374	46.2034	40.4154	47.0800
2	45.7614	52.4340	45.5614	52.4489	45.5748	52.4590	44.7267	51.3016	46.1263	53.3025
3	31.3686	35.7989	31.3567	35.9264	31.3341	35.9017	31.2388	35.7006	31.6127	36.3777
4	35.8248	40.4289	35.6683	40.4404	35.6787	40.4481	35.0148	39.5558	36.1105	41.0985
5	79.4197	93.0272	80.1778	94.7453	80.4107	95.0221	76.8594	90.0045	78.3588	92.1765
6	90.6263	105.2565	91.1212	106.8714	91.4827	107.2762	86.0377	99.8600	89.4317	104.3555
	ZIP-GA		ZIP-WEI		ZIP-WEI3		ZIP-GG		ZIP-GP	
	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2	P_1	P_2
1	40.1990	44.7401	40.1837	44.8994	40.1547	44.8685	40.0327	44.6172	40.5118	45.4635
2	46.9910	51.4043	46.7857	51.4189	46.7993	51.4287	45.9285	50.2941	47.3657	52.2557
3	31.3367	35.8390	31.3248	35.9666	31.3022	35.9420	31.2071	35.7406	31.5806	36.4185
4	36.6224	41.1386	36.4624	41.1503	36.4730	41.1582	35.7943	40.2502	36.9144	41.8200
5	79.6091	89.8335	80.3690	91.4926	80.6024	91.7600	77.0427	86.9146	78.5457	89.0120
6	93.0615	103.1895	93.5696	104.7726	93.9409	105.1700	88.3495	97.8986	91.8347	102.3061

combinations with the other claim severity models. Also, PO-GP, NBII-GP, DEL-GP, SI-GP and ZIP-GP have the highest P_1 and P_2 values in groups 1, 2, 3 and 4, while PO-WEI3, NBII-WEI3, DEL-WEI3, SI-WEI3 and ZIP-WEI3 have the highest P_1 and P_2 values in groups 5 and 6 among their combinations with the other claim severity models. Finally, with respect to the NBII and GG models which performed best, we see that NBII-GG has the lowest P_1 values in groups 2, 4 and 6 and the lowest P_2 values in groups 2 and 6 among all the combinations of the mixed Poisson models for approximating claim frequency and the claim severity models.

4. Conclusions

In this paper, we examined the use of regression models for location, scale and shape for pricing risks through ratemaking based on a priori risk classification. Specifically, we assumed that the number of claims was distributed according to a Poisson, negative binomial type II, the Delaporte, Sichel and zero-inflated Poisson and that the losses were distributed according to a gamma, Weibull, Weibull type III, generalized gamma and generalized Pareto regression model for location, scale and shape respectively. These classification models were calibrated employing a generalized Akaike information criterion (GAIC) which is valid for both nested or non-nested model comparisons (as suggested by Rigby and Stasinopoulos, 2005; 2009). The best fitted claim frequency model was the negative binomial type II model, followed closely by the Sichel and Delaporte models while regarding the claim severity models, the best fitting performances were provided by the generalized gamma model followed by the generalized Pareto and gamma models. Furthermore, the difference between these models was analyzed through the mean and the variance of the annual number of claims and the severity of claims of the policyholders, who belong to different risk classes. The resulting a priori premiums rates were calculated via the expected value and standard deviation principles with independence between the claim frequency and severity components assumed.

Extensions to other frequency/severity regression models for location scale and shape can be obtained in a similar straightforward way. Moreover, these models are parametric and a possible line of further research is to explore the semiparametric approach and go through the ratemaking exercise when functional forms other than the linear are included, based on the generalized additive models for location scale and shape (GAMLSS) approach of Rigby and Stasinopoulos (2001; 2005; 2009). Also see, for example, a recent paper by Klein et al. (2014) in which Bayesian GAMLSS models are employed for nonlife ratemaking and risk management.

Acknowledgments

The authors would like to thank the *Variance* Editor in Chief and the referees for their constructive comments and suggestions.

References

Boucher, J.P., M. Denuit, and M. Guillen, "Risk Classification for Claim Counts: A Comparative Analysis of Various Zero-Inflated Mixed Poisson and Hurdle Models," *North American Actuarial Journal* 11: 4, 2007, pp. 110–131.

Boucher, J.P., M. Denuit, and M. Guillen, "Models of Insurance Claim Counts with Time Dependence Based on Generalization of Poisson and Negative Binomial Distributions," *Variance* 2:1, 2008, pp. 135–162.

Boyer, M., G. Dionne, and C. Vanasse, "Econometric Models of Accident Distribution" in *Contributions to Insurance Economics,* ed. G. Dionne, pp. 169–213. Boston: Kluwer, 1992.

Bühlmann, H., *Mathematical Models in Risk Theory,* New York: Springer-Verlag, 1970.

Dean, C., J.F. Lawless, and G.E. Willmot, "A Mixed Poisson-Inverse-Gaussian Regression Model," *Canadian Journal of Statistics* 17 (2), 1989, pp. 171–181.

Denuit, M., and S. Lang, "Nonlife Ratemaking with Bayesian GAM's," *Insurance, Mathematics and Economics* 35, 2004, pp. 627–647.

Denuit, M., X. Marechal, S. Pitrebois, and J.F. Walhin, *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems,* Hoboken, NJ: Wiley, 2007.

Dionne, G., and C. Vanasse, "A Generalization of Actuarial Automobile Insurance Rating Models: The Negative Binomial Distribution with a Regression Component," *ASTIN Bulletin* 19, 1989, pp. 199–212.

Dionne, G., and C. Vanasse, "Automobile Insurance Ratemaking in the Presence of Asymmetrical Information," *Journal of Applied Econometrics* 7, 1992, pp. 149–165.

Evans, D.A., "Experimental Evidence Concerning Contagious Distributions in Ecology," *Biometrika* 40, 1953, pp. 186–211.

Frangos, N., and S. Vrontos, "Design of Optimal Bonus-Malus Systems with a Frequency and a Severity Component on an Individual Basis in Automobile Insurance," *ASTIN Bulletin* 31:1, 2001, pp. 1–22.

Gourieroux, C., A. Montfort, and A. Trognon, "Pseudo Maximum Likelihood Methods: Theory," *Econometrica* 52, 1984a, pp. 681–700.

Gourieroux, C., A. Montfort, and A. Trognon, "Pseudo Maximum Likelihood Methods: Applications to Poisson Models," *Econometrica* 52, 1984b, pp. 701–720.

Green, P.J., and B.W. Silverman, *Nonparametric Regression and Generalized Linear Models,* London: Chapman and Hall, 1994.

Hastie, T.J., and R.J. Tibshirani, *Generalized Additive Models,* London: Chapman and Hall, 1990.

Heller, G.Z., M.D. Stasinopoulos, R.A. Rigby, and P. de Jong, "Mean and Dispersion Modeling for Policy Claims Costs," *Scandinavian Actuarial Journal* 4, 2007, pp. 281–292.

Johnson, N.L., S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions,* New York: Wiley, 1994.

Klein, N., M. Denuit, S. Lang, and T. Kneib, "Nonlife Ratemaking and Risk Management with Bayesian Generalized Additive Models for Location, Scale, and Shape," *Insurance: Mathematics and Economics* 55, 2014, pp. 225–249.

Klugman, S., H. Panjer, and G. Willmot, *Loss Models: From Data to Decisions,* New York, Wiley, 2004.

Lambert, D., "Zero-Inflated Poisson Regression with an Application to Defects in Manufacturing," *Technometrics* 34, 1992, pp. 1–14.

Lemaire, J., *Bonus-Malus Systems in Automobile Insurance,* Norwell, MA: Kluwer Academic Publishers, 1995.

Lopatatzidis, A., and P.J. Green, "Nonparametric Quantile Regression using the Gamma Distribution," working paper, 2000.

Nelder, J.A., and R.W.M. Wedderburn, "Generalized Linear Models," *Journal of the Royal Statistical Society* A, 135, 1972, pp. 370–384.

Renshaw, A.E., "Modeling the Claims Process in the Presence of Covariates," *ASTIN Bulletin,* 24, 1994, pp. 265–285.

Rigby, R.A., and D.M. Stasinopoulos, "The GAMLSS Project: A Flexible Approach to Statistical Modelling," in B. Klein and L. Korsholm (eds.), *New Trends in Statistical Modelling, Proceedings of the 16th International Workshop on Statistical Modelling,* pp. 249–256, Odense, Denmark, 2001.

Rigby, R.A., and D.M. Stasinopoulos, "Generalized Additive Models for Location, Scale and Shape," (with discussion), *Applied Statistics* 54, 2005, pp. 507–554.

Rigby, R.A., D.M. Stasinopoulos, and C. Akantziliotou, "A Framework for Modeling Overdispersed Count Data, Including the Poisson-Shifted Generalized Inverse Gaussian Distribution," *Computational Statistics and Data Analysis* 53, 2008, pp. 381–393.

Rigby, R.A., and D.M. Stasinopoulos, *A Flexible Regression Approach Using GAMLSS in R,* University of Lancaster, 2009, http://www.gamlss.org/wp-content/uploads/2013/01/Lancaster-booklet.pdf.

Tzougas, G., and N. Frangos, "The Design of an Optimal Bonus-Malus System Based on the Sichel Distribution," pp. 239–260 in *Modern Problems in Insurance Mathematics,* New York: Springer, 2014.

Yip, K., and K. Yau, "On Modeling Claim Frequency Data in General Insurance with Extra Zeros," *Insurance, Mathematics and Economics* 36, 2005, pp. 153–63.