Annals of Actuarial Science (2020), **14**, pp. 42–71
doi:10.1017/S1748499519000034

PAPER

Insurance ratemaking using the Exponential-Lognormal regression model

George Tzougas*, Woo Hee Yik and Muhammad Waqar Mustaqeem

Department of Statistics, London School of Economics and Political Science, Holborn, London WC2A 2AE, UK
*Corresponding author. E-mail: g.tzougas@lse.ac.uk

(Received 10 October 2018; revised 2 March 2019; accepted 4 March 2019; first published online 26 June 2019)

Abstract

This paper is concerned with presenting the Exponential-Lognormal (ELN) regression model as a competitive alternative to the Pareto, or Exponential-Inverse Gamma, regression model that has been used in a wide range of areas, including insurance ratemaking. This is the first time that the ELN regression model is used in a statistical or actuarial context. The main contribution of the study is that we illustrate how maximum likelihood estimation of the ELN regression model, which does not have a density in closed form, can be accomplished relatively easily via an Expectation-Maximisation type algorithm. A real data application based on motor insurance data is examined in order to emphasise the versatility of the proposed algorithm. Finally, assuming that the number of claims is distributed according to the classic Negative Binomial and Poisson-Inverse Gaussian regression models, both the a priori and a posteriori, or Bonus–Malus, premium rates resulting from the ELN regression model are calculated via the net premium principle and compared to those determined by the Pareto regression model that has been traditionally used for modelling claim sizes.

Keywords: Exponential-Lognormal regression model; EM Algorithm; Motor Third Party Liability Insurance; ratemaking

1. Introduction

The Pareto distribution, which was discovered by Vilfredo Pareto (1848–1923), while he was studying distributions for modelling income in Switzerland, has been widely used for modelling heavy-tailed phenomena which appear frequently in a plethora of different scientific fields such as sociology, economics, physics and seismology among others. In actuarial science, where quantifying the risk posed by the more risky types of insurance has often been an imperative task for actuaries, the Pareto distribution and its generalisation, namely the Generalised Pareto (GP) distribution, which has been used in the context of Extreme Value Theory, see, for example, Embrechts, Klüppelberg & Mikosch (1997), are the most popular heavy-tailed models which have been employed by actuaries for effectively modelling extreme losses, which may have low frequencies but usually represent the biggest part of the indemnities paid by insurance companies. For instance, the economic losses from natural catastrophes in 2017 hit the second-highest level ever recorded, see Munich Re (2017). However, the Pareto distribution, similar to any other claim size distribution, has both merits and demerits. In what follows we provide a thorough discussion about the advantages and limitations of a special case of the GP distribution, the two parameter Pareto, or Exponential-Inverse Gamma, density with a regression structure, henceforth called the Pareto regression model, when it is used for premium determination in Motor Third Party Liability (MTPL) insurance which is the main focus of the present study.

As far as MTPL insurance is concerned, traditionally, a dual approach to ratemaking has been adopted by actuaries who developed both a priori ratemaking schemes and a posteriori ratemaking mechanisms or Bonus–Malus Systems (BMSs). The former process relies on the use of claim frequency and severity generalised linear models (GLMs) for computing the a priori premiums. References for a priori ratemaking include, for example, Haberman & Renshaw (1996), Denuit & Lang (2004), Boucher, Denuit & Guillen (2007), De Jong & Heller (2008), Kaas *et al.* (2008), Frees (2010) and Tzougas, Vrontos & Frangos (2015). The latter process uses information about the claim frequency and severity history of the policyholders to calculate a posteriori, or Bonus–Malus (BM), premium rates in a way which takes into account the a posterior criteria, i.e. all the factors that could not be identified, measured and introduced in the previous a priori tariff. An excellent account of BMSs can be found in Lemaire (1995). Further references for BMSs include, among many others, Trembley (1992), Picech (1994), Pinquet (1997, 1998), Brouhns *et al.* (2003), Mert & Saykan (2005), Gómez-Déniz & Calderín-Ojeda (2018), Gómez-Déniz, Hernández-Bastida & Fernández-Sánchez (2014), Ni, Constantinescu & Pantelous (2014), Ni *et al.* (2014), Santi, Purnaba & Mangku (2016), Karlis, Tzougas & Frangos (2018) and Gómez-Déniz & Calderín-Ojeda (2018). Furthermore, a basic interest of actuarial literature research is the design of BMSs for the number and costs of claims based on both the a priori and a posterior criteria, making the price discrimination even more fair and reasonable. Nevertheless, since the seminal work of Dionne & Vanasse (1989, 1992), who employed the Negative Binomial Type I (NBI), or Poisson-Gamma, regression model for constructing a BMSs by updating the posterior mean claim frequency based on explanatory variables for claim counts, the steady march of methodological innovation has mainly focused on deriving BMSs with a frequency component based on alternative count regression models to the NBI, such as the Poisson-Inverse Gaussian (PIG) regression model, which has also been a traditional choice, see Denuit *et al.* (2007), Boucher, Denuit & Guillen (2008), Tzougas & Frangos (2014) and Tzougas, Hoon & Lim (2018) among many others, while, unlike the case without covariates, the severity component has been largely ignored even if it is critical in the ratemaking process. Specifically, to the best of our knowledge, only the Pareto regression model[1] has been used so far for deriving BMS by updating the posterior mean claim severity based on covariate information for claim sizes, see, for instance, Frangos & Vrontos (2001), Mahmoudvand & Hassani (2009) and Tzougas, Vrontos & Frangos (2014, 2018). The main advantage of the Pareto model lies in the conjugacy, in a Bayesian sense, between the Inverse Gamma prior, or mixing, distribution and the Exponential distribution, which facilitates maximum likelihood (ML) estimation and a Bayesian approach towards calculating a posteriori, or BM, premiums. However, regardless of the statistical and mathematical convenience of conjugancy, there is no guarantee that variation in claim sizes has precisely the distributional forms implied by the Pareto model. In particular, a serious drawback of the Pareto model is that it is among the most heavy-tailed claim severity models and hence not flexible enough to adequately cover the behaviour of claims with moderate sizes. Thus, if very few observations are available in the tail area, meaning that claims with large amounts are so rare that their numerical impact is low, then an inappropriate imposition of the Pareto model may lead to biased estimates for moderate claim costs. Moreover, even if large claims have a significant contribution to the overall portfolio risk, as empirical evidence has shown, claims with moderate severities usually constitute the largest proportion of MTPL data and hence may also lead to substantial losses. More importantly, unless the assumption that the actual claim size distribution is a Pareto is valid, then due to its very heavy-tailed character, the Pareto model can be ill-suited for pricing risks since it will result in a severe penalisation of policyholders with moderate claim costs. Nevertheless, unlike large claims that will always be reported to the company, only moderate claims are subject to the Bonus–Hunger phenomenon. Consequently, this situation can lead to huge financial impacts for the insurance company since in those BMSs resulting from the Pareto model for claim severities it is very likely that moderate claim costs will be defrayed by the policyholders themselves and hence the insurer will have a false appreciation of

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

the real risks they are taking. Furthermore, taking into consideration that according to the latest report of Insurance Europe, in 2014 MTPL remained the most widely purchased non-life product in the European Union, accounting for 27.3% of non-life business, see Insurance Europe (2016), it becomes clear that the Bonus–Hunger phenomenon requires an accurate modelling for moderate claim amounts based on representative distributions for the claim sizes data that have the potential to capture more efficiently their stylised characteristics and thus determine the appropriate level of premiums.

The aim of the present work is to propose the Exponential-Lognormal (ELN) regression model as a competitive alternative to the Pareto regression model. The ELN model can be considered as a prominent candidate for MTLP claim severity data due to the following academic and practical reasons. Firstly, it is desirable to construct some distributions similar in nature to the Pareto distribution which can adequately capture the tail of claim size data, since large size claims generally affect liability coverages. The ELN distribution has the heavy-tailed property since, as is well known, mixing tends to produce heavy-tailed distributions, see, for example, Halliwell (2013). For instance, similar to the Pareto distribution, the ELN distribution can more adequately model the tail of the claim size data when compared to the Gamma distribution which has been very often used in the literature for modelling moderate claims, see Klugman, Panjer & Willmot (2012) and Denuit *et al.* (2007). Secondly, the advantage that the ELN distribution enjoys over the more heavy-tailed Pareto distribution is that it has a more promising shape for moderate claims. This effect is most notable when the a posteriori correction for claim amounts is calculated since the ELN model can enable the actuary to calculate a fair increment in the a posteriori, or BM, premiums that must be paid to cover all the expenses caused by a large number of medium size claims hitting the portfolio, alleviating thus the Bonus–Hunger phenomenon. We will investigate the right tail behaviour of the ELN distribution and compare it to that of the Gamma and Pareto distributions based on the set up of Wang (1998), who proposed the use of the right tail index for classifying claim severity distributions by their right tail behaviour without referring to higher moments.

At this point, we would also like to call attention to the fact that unlike the vast pricing literature on mixed Poisson models for claims counts, the study of mixed Exponential models stemming from continuous mixing distributions, which can be used for describing claim size heterogeneity and for deriving ratemaking mechanisms for claim costs, still remains a largely uncharted territory. In particular, regarding the case without covariates, except for the Pareto distribution the only other mixed Exponential distribution that has been used in an a posteriori ratemaking context so far is the Weibull, or Exponential-Lévy (Stable 1/2) distribution, which was studied in depth by Ni, Constantinescu & Pantelous (2014) who gave an excellent account of its statistical properties, putting special emphasis on its ability to fit moderate size claims well, and demonstrated how BM premiums can be derived based on the Bayesian approach. Furthermore, Ni *et al.* (2014) calculated the a posteriori correction for claim sizes by using a hybrid structure which was based on the Weibull distribution for modelling medium sized claims and the Pareto distribution for modelling larger ones. Additionally, the Exponential-Inverse Gaussian (EIG) distribution, which is also less heavy-tailed than the Pareto model and can apt for moderate claim costs, was presented by Bhattacharya & Kumar (1986), who used it for reliability purposes, while Hesselager, Wang & Willmot (1998) proposed a different parameterisation of the distribution and Frangos & Karlis (2004) considered the case with covariates by allowing a regression specification in the function for the mean parameter of the EIG distribution. However, this is the first time that the ELN model is used in a statistical or actuarial setting for the cases with and without covariate information because, due to the complexity of its log-likelihood, direct maximisation is difficult and has not been addressed in the literature so far. In particular there is no analytical form for the distribution of the cost of claims if the random effect variable, which follows the Lognormal distribution, is marginalised out. As a result, ML estimation of the ELN regression model is not straightforward to calculate and requires a special effort.

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

The main achievement of this study is that we propose a relatively simple Expectation-Maximisation (EM) type algorithm for ML estimation of the ELN regression model. The ML estimation framework we consider is based on the inherent latent structure of mixed Exponential models and is particularly useful for situations where the mixing distribution, such as the Lognormal, is not conjugate to the Exponential distribution. Furthermore, using the Negative Binomial Type I (NBI) and the PIG regression models for claim counts, both the a priori and a posteriori, or BM, premium rates resulting from the new model are calculated via the law of total expectation and the use of numerical approximation, and are compared to those determined by the Pareto model that has been widely used for modelling claim severity.

The rest of this paper proceeds as follows: Section 2 presents the derivation of the ELN regression model. Section 3 fully describes the ML estimation through the EM algorithm. Section 4 contains an application to a data set concerning car insurance claims at fault. Finally, concluding remarks can be found in section 5.

2. The ELN Regression Model

The ELN regression model which is considered in this study can be described as follows. Assume that the individual claim costs, y_i, arising from a policyholder i, $i = 1, ..., n$ are independent and identically distributed random variables according to an Exponential distribution with probability density function (pdf) given by

$$f\left(y_i|\mathbf{x}_i, z_i\right) = \frac{e^{-\frac{y_i}{\mu_i z_i}}}{\mu_i z_i} \tag{1}$$

where $y_i > 0$ and $z_i > 0$, with $\mu_i = \exp\left(\mathbf{x}_i^T \boldsymbol{\beta}\right)$, where $\mathbf{x}_i$ is the vector of covariate information regarding individual characteristics and characteristics of the vehicle related to the i^{th} insured person and where β is the vector of the regression coefficients.

The mean and the variance of $y_i|\mathbf{x}_i, z_i$ are given by

$$E(y_i|\mathbf{x}_i, z) = \exp\left(\mathbf{x}_i^T \boldsymbol{\beta} + \log\left(z_i\right)\right) \text{ and} \tag{2}$$

$$Var(y_i|\mathbf{x}_i, z) = \left[\exp\left(\mathbf{x}_i^T \boldsymbol{\beta} + \log\left(z_i\right)\right)\right]^2 \tag{3}$$

Let us now assume that z_i follows a Lognormal distribution with pdf given by

$$g\left(z_i\right) = \frac{1}{\sqrt{2\pi}\phi z_i} \exp\left[-\frac{\left(\log\left(z_i\right) + \frac{\phi^2}{2}\right)^2}{2\phi^2}\right] \tag{4}$$

with $\phi > 0$, where $E(z_i) = 1$ ensures the identifiability of the model and where $Var(z_i) = \exp\left(\phi^2\right) - 1$, for $i = 1, ..., n$.

Considering the assumptions of the model, i.e. equations (1 and 4), it is easy to see that the resulting distribution of $y_i|\mathbf{x}_i$ is the ELN distribution with pdf

$$f(y_i|\mathbf{x}_i) = \int_0^\infty \frac{e^{-\frac{y_i}{\mu_i z_i}}}{\mu_i z_i} \frac{\exp\left[-\frac{\left(\log(z_i) + \frac{\phi^2}{2}\right)^2}{2\phi^2}\right]}{\sqrt{2\pi}\phi z_i} dz_i \tag{5}$$

Unfortunately, the above integral cannot be simplified but it can be computed via numerical integration.

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Finally, based on the laws of total expectation and total variance and the moments of the Exponential distribution, one can easily see that the mean and the variance of $y_i|\mathbf{x}_i$ are given by

$$E(y_i|\mathbf{x}_i) = E_{z_i}\left[E(y_i|z_i)\right] = \exp\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)E_{z_i}\left[z_i\right] = \mu_i \tag{6}$$

and

$$\begin{aligned} Var(y_i|\mathbf{x}_i) &= E_{z_i}\left[Var(y_i|z_i)\right] + Var_{z_i}\left[E(y_i|z_i)\right] \\ &= \mu_i^2\left[2\exp\left(\phi^2\right) - 1\right] \end{aligned} \tag{7}$$

3. The EM Algorithm for ML Estimation of the ELN Regression Model

In this section we describe how an EM type algorithm can be used to facilitate the ML estimation of the ELN regression model. Let $(y_i, \mathbf{x_i})$, $i = 1, ..., n$, be a sample of independent observations, where y_i is the claim severity and where $\mathbf{x}_i$ is the vector of covariate information. Also, consider that the data are produced according to the ELN model. Then, the log-likelihood can be written as

$$l\left(\boldsymbol{\theta}\right) = \sum_{i=1}^{n}\log\left(f(y_i|\mathbf{x}_i)\right) \tag{8}$$

where $\boldsymbol{\theta} = (\phi, \boldsymbol{\beta})$ is the vector of the parameters and where $f(y_i|\mathbf{x}_i)$ is the pdf of the ELN model, which is given by equation (5). Maximisation of the above function with respect to the vector of parameters $\boldsymbol{\theta}$ is not easy because the pdf of the model does not exist in closed form and hence $\boldsymbol{\theta}$ cannot be estimated via traditional numerical maximisation methods, such as, for instance, the Newton–Raphson algorithm. Fortunately, ML estimation of the model can be achieved in easy manner via an EM type algorithm. The EM algorithm (see Dempster, Laird & Rubin 1997 and McLachlan & Krishnan 2007) is suitable for mixed Exponential models, since their stochastic mixture representation involving a non-observable random variable, denoted by z_i herein, can be considered to produce missing data. In particular, if one augments the unobserved data z_i to the observed data $(y_i, \mathbf{x_i})$, for $i = 1, ..., n$, then the complete data log-likelihood factorises into two parts

$$\begin{aligned} l_c\left(\boldsymbol{\theta}\right) &= \sum_{i=1}^{n}\left[-\frac{y_i}{\mu_i z_i} - \log\left(\mu_i\right) - \log\left(z_i\right)\right] + \\ &\quad \sum_{i=1}^{n}\left[-\frac{1}{2}\log\left(2\pi\right) - \log\left(\phi\right) - \log\left(z_i\right) - \frac{\left(\log\left(z_i\right) + \frac{\phi^2}{2}\right)^2}{2\phi^2}\right] \end{aligned} \tag{9}$$

for $i = 1, ...n$. The regression coefficients $\boldsymbol{\beta}$ are involved in the first term and the parameter ϕ is involved in the second term of equation (9) which correspond to the log-likelihoods of the Exponential and Lognormal distributions respectively.

The conditional expectation of the complete data log-likelihood is proportional to

$$\begin{aligned} Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(r)}\right) &\equiv E_{z_i}\left(l_c\left(\boldsymbol{\theta}\right)|y_i, \mathbf{x}_i, \boldsymbol{\theta}^{(r)}\right) \propto \\ &\propto \sum_{i=1}^{n}\left[-\frac{y_i E_{z_i}\left[\frac{1}{z_i}|y_i, \mathbf{x}_i, \boldsymbol{\theta}^{(r)}\right]}{\mu_i^{(r)}} - \log\left(\mu_i^{(r)}\right)\right] + \\ &\quad \sum_{i=1}^{n}\left[-\frac{E_{z_i}\left[\left(\log\left(z_i\right)\right)^2|y_i, \mathbf{x}_i, \boldsymbol{\theta}^{(r)}\right]}{2\left(\phi^{(r)}\right)^2} - \frac{\left(\phi^{(r)}\right)^2}{8} - \log\left(\phi^{(r)}\right)\right] \end{aligned} \tag{10}$$

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

where $\boldsymbol{\theta}^{(r)}$ is the estimate of $\boldsymbol{\theta}$ at the r^{th} iteration in the E-step of our EM type algorithm. In what follows, the expectations $E_{z_i}\left[\frac{1}{z_i}|y_i, \mathbf{x}_i, \boldsymbol{\theta}^{(r)}\right]$ and $E_{z_i}\left[\left(\log(z_i)\right)^2|y_i, \mathbf{x}_i, \boldsymbol{\theta}^{(r)}\right]$ have to be calculated for implementing the E-step of the algorithm, while the M-step involves maximising the $Q-$function with respect to $\boldsymbol{\theta}$. The EM type algorithm for the ELN regression model can be formally described as follows.

- E-Step: Given the current estimates, say $\boldsymbol{\theta}^{(r)}$, taken from the r^{th} iteration, calculate for all $i = 1, ..., n$, the pseudo-values

$$w_{1,i} = E_{z_i}\left[\frac{1}{z_i}|y_i, \mathbf{x}_i, \boldsymbol{\theta}^{(r)}\right] = \frac{\int_0^\infty \frac{1}{z_i}\frac{e^{-\frac{y_i}{\mu_i^{(r)}z_i}}}{\mu_i^{(r)}z_i}\frac{\exp\left[-\frac{\left(\log(z_i)+\frac{\left(\phi^{(r)}\right)^2}{2}\right)^2}{2\left(\phi^{(r)}\right)^2}\right]}{\sqrt{2\pi}\phi^{(r)}z_i}dz_i}{\int_0^\infty \frac{e^{-\frac{y_i}{\mu_i^{(r)}z_i}}}{\mu_i^{(r)}z_i}\frac{\exp\left[-\frac{\left(\log(z_i)+\frac{\left(\phi^{(r)}\right)^2}{2}\right)^2}{2\left(\phi^{(r)}\right)^2}\right]}{\sqrt{2\pi}\phi^{(r)}z_i}dz_i} \tag{11}$$

and

$$w_{2,i} = E_{z_i}\left[\left(\log(z_i)\right)^2|y_i, \mathbf{x}_i, \boldsymbol{\theta}^{(r)}\right] = \frac{\int_0^\infty \left(\log(z_i)\right)^2\frac{e^{-\frac{y_i}{\mu_i^{(r)}z_i}}}{\mu_i^{(r)}z_i}\frac{\exp\left[-\frac{\left(\log(z_i)+\frac{\left(\phi^{(r)}\right)^2}{2}\right)^2}{2\left(\phi^{(r)}\right)^2}\right]}{\sqrt{2\pi}\phi^{(r)}z_i}dz_i}{\int_0^\infty \frac{e^{-\frac{y_i}{\mu_i^{(r)}z_i}}}{\mu_i^{(r)}z_i}\frac{\exp\left[-\frac{\left(\log(z_i)+\frac{\left(\phi^{(r)}\right)^2}{2}\right)^2}{2\left(\phi^{(r)}\right)^2}\right]}{\sqrt{2\pi}\phi^{(r)}z_i}dz_i} \tag{12}$$

Clearly the expectations involved in the E-step of the algorithm do not have closed form expressions and thus numerical approximations are needed. Specifically, equations (11 and 12) can be evaluated numerically. Alternatively, a Monte Carlo approach can also be used based on a rejection algorithm. The latter case leads to variants of the EM algorithm such as the Monte Carlo EM algorithm (see, for instance, Booth & Hobert 1999 and Booth, Hobert & Jank 2001) which do not require knowledge of the pdf $f(y_i|\mathbf{x}_i)$, but it is sufficient to be able to simulate from the posterior density $g\left(z_i|y_i, \mathbf{x}_i, \boldsymbol{\theta}\right)$, where $g(z_i)$ in our case is the pdf of the Lognormal mixing distribution, which is given by equation (4).

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

- M-Step: In the M-Step, the pseudo-values $w_{1,i}$ and $w_{2,i}$ from the E-Step can be used to maximise the $Q-$function.
  - Firstly, the Newton–Raphson algorithm is employed to obtain ML estimates of the elements of $\boldsymbol{\beta}$. Taking the necessary derivatives of the $Q-$function with respect to $\boldsymbol{\beta}$ we obtain the following results:

$$h\left(\boldsymbol{\beta}\right)=\frac{\partial Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(r)}\right)}{\partial \boldsymbol{\beta}}=\sum_{i=1}^{n}\left(\frac{y_i}{\mu_i^{(r)}}w_{1,i}-1\right)\mathbf{x}_i \tag{13}$$

  and

$$H\left(\boldsymbol{\beta}\right)=\frac{\partial^2 Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^{(r)}\right)}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}^T}=\sum_{i=1}^{n}\left(-\frac{y_i}{\mu_i^{(r)}}w_{1,i}\right)\mathbf{x}_i\mathbf{x}_i^T=\mathbf{X}^T\mathbf{W}\mathbf{X} \tag{14}$$

  for $i=1,...,n$, where $\mathbf{W}=diag\left\{-\frac{y_i}{\mu_i^{(r)}}w_{1,i}\right\}$ and where $w_{1,i}$ is given by equation (11).

  The Newton–Raphson iterative procedure for obtaining ML estimates of the elements of β goes as follows:

$$\boldsymbol{\beta}^{(r+1)}\equiv\boldsymbol{\beta}^{(r)}-\left[H\left(\boldsymbol{\beta}^{(r)}\right)\right]^{-1}h\left(\boldsymbol{\beta}^{(r)}\right) \tag{15}$$

  - Secondly, update ϕ with

$$\phi^{(r+1)}\equiv\sqrt{2\left(\sqrt{1+\bar{w}_2}-1\right)} \tag{16}$$

  where $\bar{w}_2=\frac{\sum_{i=1}^{n}w_{2,i}}{n}$ and where $w_{2,i}$ is given by equation (12).
- Note also that when the regression component of the model is limited to a constant β_0 one obtains $E(y_i|\mathbf{x}_i)=\exp\left(\mathbf{x}_i^T\beta_0\right)=\mu$, and thus this EM type algorithm can be employed for the ML estimation of the univariate, without a regression component, model.

4. Numerical Illustration

In this paper, we worked with a sample of data that were kindly provided by a major insurance company operating in Greece and concern a MTPL insurance portfolio observed over 3 years. In our sample we considered only policyholders with complete records, i.e. with availability of all a priori rating variables that correspond to their characteristics, including the characteristics of their car. Response variables are the number of claims at fault reported to the company and the loss corresponding to each claim made. The sample comprises 9,986 policies which met our criteria. The available a priori rating variables we employ are the BM class of the policyholders, the horsepower (HP) of their car and the size of the city (CS) where they live.

- This BMS 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 a given year is penalised by one class. The variable BM class divides the classes of the current Greek BMS into four categories of drivers, those who belong to BM classes: C1= "Class 1–Class 2", C2 = "Class 3–Class 5", C3 = "Class 6–Class 9" and C4 = "Class 10–Class 20".
- The variable HP consists of three categories of cars, those with a HP: C1 = "0–1,400 cc", C2 = "1,400–1,800 cc" and C3 = "greater than 1,800 cc".
- The variable CS consists of three categories of policyholders, those who live in a: C1 = "large city", C2 = "middle sized city" and C3 = "small city".

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Table 1. Descriptive statistics of claim severities – size of the different categories of the explanatory variables

Statistic	Claim severities	Bonus–Malus (BM) category		Horsepower (HP) of the car		Size of the city (CS)	
Minimum	55.4	C1:	6,709	C1:	4,252	C1:	4,942
Median	1,931	C2:	1,393	C2:	2,037	C2:	2,539
Mean	2,685.8	C3:	1,168	C3:	3,697	C3:	2,505
Maximum	141,432	C4:	716		-		-

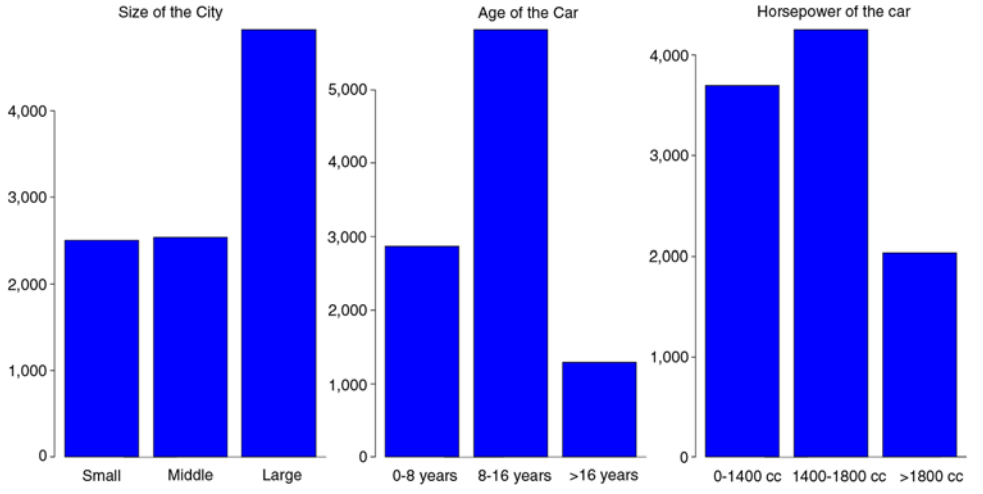

Figure 1. Descriptive histograms for the explanatory variables.

Furthermore, since in this study we focus on the claim severity component, in Table 1 we present some standard descriptive statistics for claim severities along with the number of observations in each category of the three explanatory variables.

Additionally, Figure 1 displays the corresponding to the explanatory variables descriptive histograms, giving us an indication of the range of their values.

Finally, Table 2 presents a summary of the effects of the covariates on claim severities for 36 different risk classes, which can be formed by dividing the portfolio into clusters defined by the combinations of the characteristics of the policyholders and their cars based on all 9,986 observations. In particular, Table 2 depicts the percentage of observations with claim sizes less than 2,500 euros, higher than or equal to 2,500 euros and less than 3,500 euros, higher than or equal to 3,500 euros and less than 4,500 euros, and higher than or equal to 4,500 euros for each of the 36 groups of policyholders.

In the following subsections, we fit the Negative Binomial Type I (NBI) and PIG distributions on the number of claims, and the ELN and Pareto distributions on claim sizes. Moreover, we will compare the two aforementioned claim severity distributions with the Generalised Beta of the second kind (GB2) distribution, which has been used in an abundance of actuarial settings for accommodating the long-tailed nature of claim sizes. Furthermore, regression components are introduced in their mean parameters, and all the risk classifying characteristics presented above are included so as to use all the available information in the estimation of the claim frequency and severity distributions. Additionally, the a priori and a posteriori, or BM, premium rates resulting from the combinations of the NBI and PIG claim frequency distributions/regression models with the Pareto and ELN distributions/regression models will be calculated via the net

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Table 2. Summary statistics of the claim severities of the different risk classes determined by the combination of the explanatory variables

Risk class	Explanatory variables			Claim severity y			
	BM	HP	CS	$y < 2,500$	$2,500 \leq y < 3,500$	$3,500 \leq y < 4,500$	$y \geq 4,500$
1	C1	C1	C1	75.80	8.52	8.68	7.00
2	C1	C1	C2	74.96	5.63	11.26	8.15
3	C1	C1	C3	76.70	7.52	8.19	7.59
4	C1	C2	C1	70.37	8.15	9.63	11.85
5	C1	C2	C2	68.39	7.69	11.19	12.73
6	C1	C2	C3	72.45	8.67	9.44	9.44
7	C1	C3	C1	74.52	7.81	10.31	7.36
8	C1	C3	C2	69.94	9.17	11.71	9.18
9	C1	C3	C3	73.33	7.97	10.11	8.59
10	C2	C1	C1	77.44	11.28	8.27	3.01
11	C2	C1	C2	77.85	4.70	10.74	6.71
12	C2	C1	C3	76.95	10.48	7.78	4.79
13	C2	C2	C1	67.86	3.57	21.43	7.14
14	C2	C2	C2	69.81	7.55	10.06	12.58
15	C2	C2	C3	72.45	8.16	10.20	9.19
16	C2	C3	C1	79.78	8.20	6.01	6.01
17	C2	C3	C2	70.59	11.76	8.82	8.83
18	C2	C3	C3	76.03	4.55	9.92	9.50
19	C3	C1	C1	78.81	6.78	5.93	8.47
20	C3	C1	C2	76.24	2.97	11.88	8.91
21	C3	C1	C3	72.89	9.89	8.79	8.42
22	C3	C2	C1	77.78	8.33	5.56	8.33
23	C3	C2	C2	70.73	9.76	9.75	9.76
24	C3	C2	C3	71.43	3.30	9.89	15.38
25	C3	C3	C1	76.16	6.62	11.92	5.30
26	C3	C3	C2	73.21	3.57	8.93	14.29
27	C3	C3	C3	79.28	6.31	4.05	10.36
28	C4	C1	C1	80.00	10.00	7.14	2.86
29	C4	C1	C2	80.88	1.48	11.76	5.88
30	C4	C1	C3	79.10	3.95	12.99	3.96
31	C4	C2	C1	66.67	8.33	16.67	8.33
32	C4	C2	C2	65.33	10.67	13.33	10.67
33	C4	C2	C3	65.96	6.38	12.77	14.89
34	C4	C3	C1	82.00	6.00	8.00	4.00
35	C4	C3	C2	78.38	5.41	5.41	10.80
36	C4	C3	C3	73.28	7.63	9.92	9.17

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

premium principle with independence between the claim frequency and severity components assumed[2].

- Because we will be comparing the ELN distribution/regression model with the Pareto and GB2 distributions/regression models, we give below some rudimentary facts concerning the latter two.
  - The Pareto model, taken for instance from Frangos & Vrontos (2001) and Mahmoudvand & Hassani (2009), can be constructed as follows. Following a similar approach, as in the case of the ELN model, consider that $y_i|\mathbf{x}_i, z_i$, for $i = 1, ..., n$, is distributed according to the Exponential distribution with mean $\mu_i z_i$, where $y_i > 0$, where $\mu_i = \exp\left(\mathbf{x}_i^T \boldsymbol{\beta}\right)$, where $\mathbf{x}_i$ is the vector of covariate information regarding individual characteristics and characteristics of the car related to the i^{th} insured person and β is the vector of the regression coefficients, and where $z_i > 0$ is a continuous random variable which follows an Inverse Gamma distribution with pdf

$$g(z_i) = \frac{\frac{1}{(\phi-1)} \exp\left(-\frac{(\phi-1)}{z_i}\right)}{\left(\frac{z_i}{\phi-1}\right)^{\phi+1} \Gamma(\phi)} \tag{17}$$

    with $\phi > 2$ and mean $E(z_i) = 1$. Then, the resulting distribution of $y_i|\mathbf{x}_i$ is a Pareto distribution with pdf given by

$$f(y_i|\mathbf{x}_i) = \phi \frac{\left[(\phi-1)\mu_i\right]^{\phi}}{\left[y_i + (\phi-1)\mu_i\right]^{\phi+1}} \tag{18}$$

    The mean and the variance of the Pareto distribution are given by

$$E(y_i|\mathbf{x}_i) = \mu_i \text{ and } Var(y_i|\mathbf{x}_i) = \frac{\left[(\phi-1)\mu_i\right]^2}{\phi-1}\left(\frac{2}{\phi-2} - \frac{1}{\phi-1}\right) \tag{19}$$

  - Regarding the GB2 distribution, it should be noted that many heavy-tailed distributions can be written as special or limiting cases of the GB2, see, for instance, McDonald & Xu (1995). However, despite the prominence of the GB2 distribution in fitting heavy-tailed data, relatively few applications use the GB2 in a regression context. Further details of the GB2 regression model can be found, for instance, in Frees & Valdez (2008), Frees, Derrig & Meyers (2014a) and Calderín-Ojeda, Fergusson & Wu (2017). Herein we use a simple specification of the GB2, see Rigby & Stasinopoulos (2009), which has as a special case the Pareto model given by equation (18) and allows us to parameterise the location parameter, μ_i, in terms of covariates, i.e. $\mu_i = \exp\left(\mathbf{x}_i^T \boldsymbol{\beta}\right)$, for $i = 1, .., n$. The pdf of the GB2 distribution is defined by

$$\begin{aligned} f(y_i|\mathbf{x}_i) &= |\phi| \, y_i^{\phi\nu-1} \left\{ \mu_i^{\phi\nu} B(\nu,\tau) \left[1 + \left(\frac{y_i}{\mu_i}\right)^{\phi}\right]^{\nu+\tau} \right\}^{-1} \\ &= \frac{\Gamma(\nu+\tau)}{\Gamma(\nu)\Gamma(\tau)} \frac{\phi \left(\frac{y_i}{\mu_i}\right)^{\phi\nu}}{y_i \left[1 + \left(\frac{y_i}{\mu_i}\right)^{\phi}\right]^{\nu+\tau}} \end{aligned} \tag{20}$$

    for $y_i > 0$, where $-\infty < \phi < \infty$, $\nu > 0$ and $\tau > 0$, see also McDonald & Xu (1995), equation (2.7), and where

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

$$B(\nu,\tau)=\int_0^1 x^{\nu-1}(1-x)^{\tau-1}\,dx=\frac{\Gamma(\nu)\,\Gamma(\tau)}{\Gamma(\nu+\tau)} \tag{21}$$

is the Beta function.

The first and the second moment of the GB2 distribution are given by

$$E(y_i|\mathbf{x}_i)=\frac{\mu_i B\left(\nu+\frac{1}{\phi},\tau-\frac{1}{\phi}\right)}{B(\nu,\tau)},\ \text{for}-\nu<\frac{1}{\phi}<\tau \tag{22}$$

and

$$E(y_i^2|\mathbf{x}_i)=\frac{\mu_i^2 B\left(\nu+\frac{2}{\phi},\tau-\frac{2}{\phi}\right)}{B(\nu,\tau)},\ \text{for}-\nu<\frac{2}{\phi}<\tau \tag{23}$$

see McDonald (1996). Hence, the variance of the GB2 distribution can be easily calculated using equations (22 and 23).

- Furthermore, the NBI and PIG claim frequency models can be derived as follows. Consider a policyholder i whose number of claims, denoted as k_i, with $k_i=0,1,2,3,...$, are independent and suppose that $\boldsymbol{\chi}_i$ is the vector of individual characteristics and characteristics of the car related to the i^{th} insured person, $i=1,...,n$, which represent different a priori rating variables. We assume that given a continuous random variable $u_i>0$ with pdf $\upsilon(u_i)$ defined on $\mathcal{R}^+$, $k_i|\boldsymbol{\chi}_i, u_i$ follows the Poisson distribution with mean $\lambda_i u_i$, where $\lambda_i=\exp\left(\boldsymbol{\chi}_i^T\mathbf{b}\right)$ and where $\mathbf{b}$ is the vector of the regression coefficients. Then, the unconditional distribution of k_i is a mixed Poisson distribution. The following are the well-known results applied to the above situation, see, for example, Dionne & Vanasse (1989, 1992), Boucher, Denuit & Guillen (2007, 2008) and Tzougas, Vrontos & Frangos (2018). We consider that $E(u_i)=1$ as this ensures the identifiability of the model.
  - Let u_i follow a Gamma distribution with pdf given by

$$\upsilon(u_i)=\frac{u_i^{\frac{1}{\sigma}-1}\frac{1}{\sigma}^{\frac{1}{\sigma}}\exp\left(-\frac{u_i}{\sigma}\right)}{\Gamma\left(\frac{1}{\sigma}\right)} \tag{24}$$

    where $\sigma>0$. Parameterisation (24) ensures that $E(u_i)=1$.

    Under this assumption the distribution of $k_i|\boldsymbol{\chi}_i$ becomes a NBI distribution, with probability mass function (pmf) given by

$$P(k_i|\boldsymbol{\chi}_i)=\frac{\Gamma\left(k_i+\frac{1}{\sigma}\right)}{k_i!\Gamma\left(\frac{1}{\sigma}\right)}\left(\frac{\sigma\lambda_i}{1+\sigma\lambda_i}\right)^{k_i}\left(\frac{1}{1+\sigma\lambda_i}\right)^{\frac{1}{\sigma}} \tag{25}$$

    The mean and the variance of the NBI distribution are given by

$$E(k_i|\boldsymbol{\chi}_i)=\lambda_i \text{ and } Var(k_i|\boldsymbol{\chi}_i)=\lambda_i+\lambda_i^2\sigma \tag{26}$$

  - Let u_i follow an Inverse Gaussian distribution with pdf given by

$$\upsilon(u_i)=\frac{1}{\sqrt{2\pi\sigma u_i^3}}\exp\left[-\frac{1}{2\sigma u_i}(u_i-1)^2\right] \tag{27}$$

    where $\sigma>0$. Parameterisation (27) also ensures that $E(u_i)=1$. Then, the distribution of $k_i|\boldsymbol{\chi}_i$ becomes a PIG distribution, with pmf given by

$$P(k_i|\boldsymbol{\chi}_i)=\left(\frac{2a}{\pi}\right)^{\frac{1}{2}}\frac{\lambda_i^{k_i}e^{\frac{1}{\sigma}}K_{k_i-\frac{1}{2}}(a)}{(a\sigma)^{k_i}k_i!} \tag{28}$$

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

where $a^2 = \sigma^{-2} + \frac{2\lambda_i}{\sigma}$ and where

$$K_\nu(\omega) = \int_0^\infty x^{\nu-1} \exp\left[-\frac{1}{2}\omega\left(x + \frac{1}{x}\right)\right] dx \tag{29}$$

is the modified Bessel function of the third kind of order ν with argument ω.
The mean and the variance of the PIG distribution are given by

$$E(k_i|\boldsymbol{\chi}_i) = \lambda_i \text{ and } Var(k_i|\boldsymbol{\chi}_i) = \lambda_i + \lambda_i^2\sigma \tag{30}$$

- Finally, note that when the regression component in each of the aforementioned claim severity and frequency regression models is limited to a constant, one obtains the univariate, without a regression component, models.

4.1 Modelling results

This subsection presents the modelling results of the ELN distribution/regression model and the traditional NBI, PIG, GB2 and Pareto distributions/regression models. The EM algorithm described in section 3 was used to estimate both the ELN distribution and regression models. The model, for both the cases with and without covariate information, converged after a few iterations using a rather strict stopping criterion. In particular, we iterated between the E-step and the M-step until the relative change in log-likelihood, which is given by equation (8), between two successive iterations was smaller than 10^{-12}. We also emphasise that for this model the choice of initial values for both the vector of the regression coefficients β and the parameter ϕ of the Lognormal mixing distribution needed special attention because one may obtain inadmissible values if the starting values are bad. Good starting values for β were obtained by fitting the exponential regression model. Also, a good initial value for ϕ was feasible by calculating $Var(y_i|\mathbf{x}_i)$ based on all observations, $i = 1, ..., n$, in our data set and solving equation (7) with respect to $\phi > 0$. Additionally, to ensure that the global maximum had been obtained and the algorithm had not been trapped in a local maximum, we checked with many initial values for ϕ, but for all cases we converged on the same solution. Furthermore, standard errors were obtained using the standard approach of Louis (1982). All computing was done using the statistical computing environment language R. Additionally, ML estimation of the NBI, PIG, GB2 and Pareto distributions/regression models, for which the definition of a log-likelihood function in closed form is feasible, was straightforward by using standard statistical packages in R, such as the GAMLSS package. For more details on the GAMLSS package, see Stasinopoulos, Rigby & Akantziliotou (2008). Finally, the computational time requirements of the ELN distribution/regression model were compared to those of the over-simplistic Exponential distribution/regression model. As anticipated, the Exponential distribution/regression model compared favourably to the ELN distribution/regression model in terms of computing times required for ML estimation since both the cases took fewer than 10 seconds of CPU time. Nevertheless, taking into account that there were 9,986 policies in the sample of MTPL data that was examined in this study, that we used a rather strict stopping criterion for EM iterations and that the expectations involved at the E-step of the algorithm do not have closed form expressions, the CPU times of the EM algorithm used for ML estimation of both the ELN distribution and the ELN regression models can be characterised as modest. In particular, the ML estimation of the ELN regression model was more chronologically demanding than that of the ELN distribution because the numerical evaluation of the integrals at the E-step for the case with covariates is more computationally time consuming than for the case without covariates. However, both cases took less than 2 minutes of CPU time. Finally, it should be mentioned that the trade-off between CPU time requirements and the efficiency of the ELN regression model for approximating claim costs in our sample and for deriving ratemaking mechanisms is sifted in favour of the latter two. In particular, regardless of the very low CPU time required for ML estimation of the Exponential model, the assumption of an

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Exponential response distribution is inadequate both for modelling the tail of the claim severity and for ratemaking purposes since firstly this model cannot effectively capture the tail of the claim size distribution and secondly all the possible important factors, such as reaction times and aggressive driving behaviour, etc., whichto a great extent reveal the riskiness of the insureds, but are either unmeasurable or unobservable, cannot be integrated into the model as it can only take into account covariate information of the policyholder and/or their car. As a result, heterogeneity may still be observed in tariff cells despite the use of many classification variables. On the contrary, as far as the ELN regression model is concerned, firstly, as illustrated later in this subsection, it has an upper tail which can sufficiently fit large claims in our sample and secondly, as was previously shown, it allows us to correct for this unobserved heterogeneity since it was constructed by assuming that $y_i|\mathbf{x}_i, z_i$, for $i = 1, ..., n$, is distributed according to the Exponential distribution with mean $\mu_i z_i$, where $y_i > 0$, where $\mu_i = \exp\left(\mathbf{x}_i^T \boldsymbol{\beta}\right)$, and where the risk parameter $z_i > 0$, which represents the risk proneness of policyholder i, i.e. unknown risk characteristics of the policyholder having a significant impact on the occurrence of claims, was regarded as a random variable that is distributed according to the Lognormal distribution, see equations (1 and 4).

The ML estimates of the parameters and the corresponding standard errors in parentheses for the NBI, PIG, GB2, Pareto and ELN distributions are presented in Panel A of Table 3, while Panel B of Table 3 reports our findings with respect to the NBI, PIG, GB2, Pareto and ELN regression models.

Furthermore, we rely on normalised quantile residuals[3], see Dunn & Smyth (1996), as an exploratory graphical device for investigating the adequacy of the fit of the competing NBI and PIG models for claim frequencies and GB2, Pareto and ELN models for the claim severities. Also, for comparison purposes, we fitted the simple Exponential regression model, which obviously has a thinner tail than the GB2, Pareto and ELN models, and included the corresponding plot. For continuous response distributions, the normalised (randomised) quantile residuals are defined as $\hat{r}_i = \Phi^{-1}(w_i)$, where Φ^{-1} is the inverse cumulative distribution function of a standard Normal distribution and where $w_i = F_i(y_i|\hat{\boldsymbol{\theta}})$, where F_i is the cumulative distribution function estimated for the ith individual, y_i is the corresponding observation and $\hat{\boldsymbol{\theta}}$ contains all estimated model parameters. For discrete response distributions, the aforementioned definition is extended and w_i is defined as a random value from the uniform distribution on the interval $[F_i(y_i - 1|\hat{\boldsymbol{\theta}}), F_i(y_i|\hat{\boldsymbol{\theta}})]$. In both cases, the model fit can be evaluated by means of usual quantile–quantile plots. Specifically, if the data indeed follow the assumed distribution, then the residual on the quantile–quantile plot will fall approximately on a straight line. Figure 2 shows the normalised (random) quantiles for the NBI, PIG, Exponential, GB2, Pareto and ELN regression models. From Figure 2, we see that the residuals of the NBI and PIG models are very close to the diagonal and indicate a very good fit to the distribution of the claim frequencies. Also, regarding claim severities, the residuals indicate that the GB2, Pareto and ELN models are better assumptions than the Exponential model since the residuals of the former three are very close to the diagonal and indicate a very good fit to the distribution of the claim sizes, while the sample quantiles of the Exponential model greater than 2 are significantly higher than the theoretical quantiles and thus, as expected, the Exponential model does not capture the tail of the claim size distribution which corresponds to significantly large claim sizes.

Additionally, in what follows, we will investigate the behaviour of the ELN distribution at large claim sizes and compare it to that of the Gamma and Pareto distributions. In particular, we will present a ranking of these claim severity distributions by the right tail index[4] which is a risk measure for right tail deviation that was suggested by Wang (1998). The right tail index is defined as

$$d(Y) = \frac{\int_0^\infty \sqrt{S_Y(t)}dt}{E(Y)} - 1 \tag{31}$$

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Table 3. Results of the fitted NBI, PIG, GB2, Pareto and ELN models

Panel A: Distributions									
NBI		PIG		GB2		PAR		ELN	
Variable		Variable		Variable		Variable		Variable	
λ	0.41070 (0.01971)	λ	0.41070(0.01951)	μ	1,174.41 (0.03551)	μ	2,725.50 (0.03536)	μ	2,688.20 (0.03521)
Variable		Variable		Variable		Variable		Variable	
σ	1.00570 (0.00041)	σ	1.08170(0.00038)	ϕ	3.81521(0.04071)	ϕ	5.16000 (0.03719)	ϕ	0.33040 (0.02272)
-	-	-	-	Variable		-	-	-	-
-	-	-	-	ν	3.29721(0.06894)	-	-	-	-
-	-	-	-	Variable		-	-	-	-
-	-	-	-	τ	0.60273 (0.06206)	-	-	-	-
Panel B: Regression models									
NBI		PIG		GB2		PAR		ELN	
Variable	Coeff. **b**	Variable	Coeff. **b**	Variable	Coeff. $\boldsymbol{\beta}$	Variable	Coeff. $\boldsymbol{\beta}$	Variable	Coeff. $\boldsymbol{\beta}$
Intercept	−0.68410 (0.01161)	Intercept	−0.68440 (0.01158)	Intercept	7.01870 (0.03859)	Intercept	7.84670 (0.03767)	Intercept	7.82290 (0.03697)
BM		BM		BM		BM		BM	
C2	−0.07752 (0.03672)	C2	−0.07622 (0.03543)	C2	−0.02332 (0.01131)	C2	−0.03933 (0.01057)	C2	−0.04183 (0.01007)
C3	−0.14731 (0.06998)	C3	−0.14751 (0.06978)	C3	0.02811 (0.01355)	C3	0.04953 (0.01257)	C3	0.05442 (0.01232)
C4	−0.03912 (0.01922)	C4	−0.04101 (0.01918)	C4	−0.00682 (0.00326)	C4	−0.02092 (0.00261)	C4	−0.01872 (0.00241)
HP		HP		HP		HP		HP	
C2	0.00613 (0.00281)	C2	0.00682 (0.00279)	C2	0.04733 (0.02017)	C2	0.09281 (0.01999)	C2	0.08801 (0.01958)
C3	0.06832 (0.02791)	C3	0.06893 (0.02788)	C3	0.01692 (0.00704)	C3	0.04222 (0.00607)	C3	0.06443 (0.00546)
CS		CS		CS		CS		CS	
C2	−0.16933 (0.07912)	C2	−0.17273 (0.07909)	C2	0.00493 (0.00222)	C2	0.07911 (0.00189)	C2	0.09313 (0.00170)
C3	−0.46392 (0.21731)	C3	−0.49493 (0.21725)	C3	−0.00313 (0.00139)	C3	0.04243 (0.00122)	C3	0.05152 (0.00105)
Variable		Variable		Variable		Variable		Variable	
σ	0.96700 (0.00794)	σ	1.03820 (0.00791)	ϕ	3.81210(0.04062)	ϕ	5.18670 (0.03710)	ϕ	0.32600 (0.02262)
-	-	-	-	Variable		-	-	-	-
-	-	-	-	ν	3.28361(0.06844)	-	-	-	-
-	-	-	-	Variable		-	-	-	-
-	-	-	-	τ	0.60491(0.06263)	-	-	-	-

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

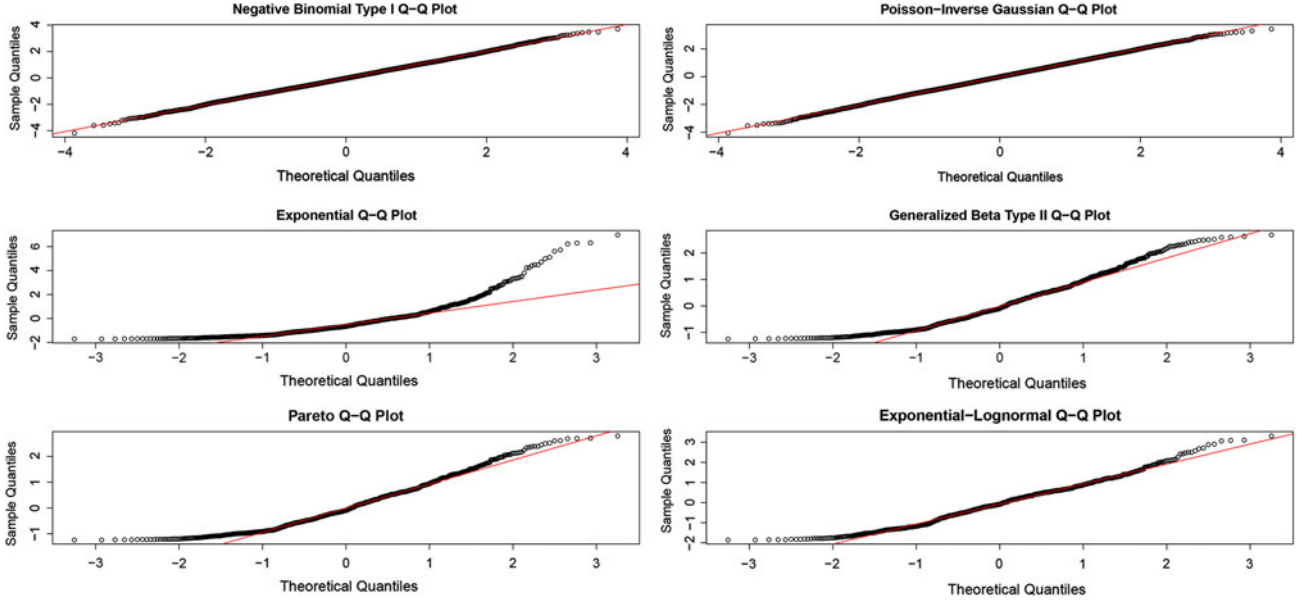

Figure 2. Normalised quantiles for the NBI, PIG, Exponential, GB2, Pareto and ELN regression models.

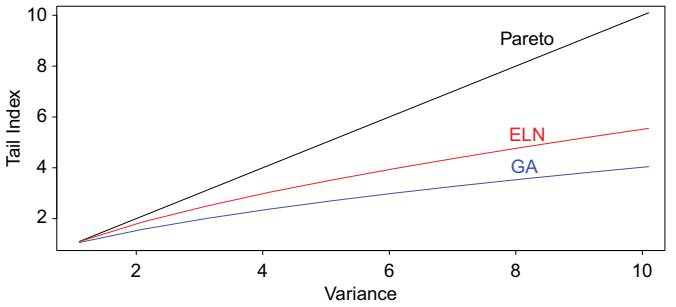

Figure 3. Plot of the right tail index as a function of the variance for the Gamma, Pareto and ELN distributions with unit mean.

where $S_Y(t) = P(Y \geq t)$ is the survival function, or the decumulative distribution function, of Y.

Figure 3 displays three plots of the right tail index, $d(Y)$, as a function of the variance for the Gamma, Pareto and ELN distributions respectively. The parameters were chosen so that the Gamma, Pareto and ELN distributions have a unit mean, i.e. $E(Y) = 1$, and varying variance, $Var(Y)$, taking on the same values for all densities. From Figure 3, we observe that the ranking of all models in terms of their right tail index values is consistent with what was discussed in section 1. Specifically, the right tail index ranks the Pareto distribution as having a fatter tail than ELN distribution, which in turn has a fatter tail than the Gamma distribution.

Finally, the empirical estimator of the right tail index, $\hat{d}(Y)$, which was considered by Jones & Zitikis (2003) can be calculated as

$$\hat{d}(Y) = \sum_{i=1}^{n} c_i \frac{Y_{(i)}}{\bar{Y}} \tag{32}$$

where $Y_{(i)}$ is the $i-th$ ordered observation of the sample $Y_1, ..., Y_n$ and where the coefficients c_i are given by

$$c_i = \sqrt{\left(\frac{n-i}{n}\right)} + \sqrt{\left(\frac{n-i+1}{n}\right)} - \frac{1}{n} \tag{33}$$

for $i = 1, ..., n$.

Regarding our data set, the value which we obtained for the empirical estimator of the right tail index is $\hat{d}(Y) = 1.362$, while the ELN distribution has a right tail index $d(Y) = 1.217$ which is close to the empirical result. Nevertheless, since for smaller data sets the empirical approach can lead to an underestimation of $d(Y)$, it makes sense to build a parametric bootstrap two-sided confidence interval (CI) for $\hat{d}(Y)$. Given our data, we generated $B = 100{,}000$ bootstrap samples of size 9,986 and we calculated the 95% bootstrap-based CI for $\hat{d}(Y)$ to be (1.113, 1.560). The value 1.217 is included in this CI and hence this is an additional indicator that the ELN distribution is able to effectively model the right tail of the data.

4.2 Models comparison

Thus far, we have the NBI, PIG and GB2, Pareto and ELN competing distributions/regression models for the claim frequency and severity component respectively. Consequently, to differentiate between these models, this subsection compares them so as to select the best for each case, employing the Global Deviance (DEV), Akaike information criterion (AIC) and the Schwarz Bayesian Criterion (SBC) which are classic hypothesis/specification tests. The (fitted) Global

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Table 4. Global deviance, AIC and SBC values for the NBI, PIG, GB2, Pareto and ELN models

Panel A: Distributions				Panel B: Regression models				
Model	df	AIC	SBC	Model	df	DEV	AIC	SBC
NBI	2	17,021.61	17,036.03	NBI	9	16,949.97	16,967.97	17,032.86
PIG	2	16,963.01	16,977.43	PIG	9	16,878.27	16,896.27	16,961.15
GB2	4	176,167.20	176,196	GB2	11	176,108	176,130	176,209.30
Pareto	2	176,518	176,532.40	Pareto	9	176,485.40	176,503.40	176,568.30
ELN	2	176,173	176,187.40	ELN	9	176,120.80	176,138.80	176,203.70

Deviance is defined as

$$DEV = -2\hat{l}(\hat{\boldsymbol{\theta}}) \tag{34}$$

where $\hat{l}$ is the maximum of the log-likelihood and $\hat{\theta}$ is the estimated parameter vector of the model. Furthermore, the AIC is given by

$$AIC = DEV + 2 \times df \tag{35}$$

and the SBC is given by

$$SBC = DEV + \log(n) \times df \tag{36}$$

where df are the degrees of freedom, that is, the number of fitted parameters in the model and n is the number of observations in the sample.

It should be noted that special emphasis should be placed on the comparison of the NBI distribution/regression model with the PIG distribution/regression model and the comparison of the Pareto distribution/regression model with the ELN distribution/regression model since their stochastic mixture representation, see sections 2 and 4, enables the use of a Bayesian approach towards deriving a posteriori, or BM, ratemaking mechanisms for the number and the cost of claims. Moreover, in a posteriori ratemaking it is of crucial importance to the actuary to have interpretable results in order to refine their a priori risk classification and restore fairness by using a premium structure that can be sufficiently explained to policyholders and regulators. Therefore, while the GB2 model is also included in this subsection for the sake of comparison with the two mixed Exponential models, in the following subsections we will limit our analyses only to the NBI, PIG, Pareto and ELN distributions/regression models since the differences between those models will produce different a posteriori, or BM, premiums. Note also that since the NBI and PIG models for the frequency component and the Pareto and ELN models for the severity component have the same number of parameters, it is sufficient to examine the respective log-likelihoods. The resulting Global Deviance, AIC and SBC values for the NBI, PIG, Pareto and ELN models are given in Table 4 (Panels A and B). As is well known, a commonly used rule-of-thumb states that a model significantly outperforms a competitor if the difference in their log-likelihoods exceeds five, corresponding to a difference in their AIC values of more than ten and to a difference in their SBC values of more than 5, see Burnham & Anderson (2003) and Raftery (1995) respectively. This means here that, as can be seen from Panels A and B, as far as claim frequencies are concerned, the best fit is given by the PIG distribution/regression model, while regarding the claim severities, the ELN distribution/regression model is superior to the Pareto distribution/regression model.

Finally, Table 4 (Panels A and B) also includes the Deviance, AIC and SBC values for the GB2 model. Our findings suggest that with respect to the AIC (see Panels A and B) and the Global Deviance (see Panel B) test results, the fit provided by the GB2 distribution/regression model is only marginally better than the fit given by the ELN distribution/regression model, which has fewer parameters. However, when the SBC test is used, we observe a slight superiority of the ELN distribution/regression model versus the GB2 distribution/regression model. The difference in the

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

outcome of the AIC and SBC tests for the GB2 and ELN distributions/regression models is understandable because, as has been mentioned on several occasions in the applied statistical literature, the AIC may fail to choose the most parsimonious model because the AIC function is largely based on the deviance function, whereas the BIC penalises model complexity, i.e. the number of parameters, more heavily. In particular, as it can be seen from equations (35 and 36), only the penalty term differs between each formula, while the goodness of fit parts remain the same. Therefore, in this case the BIC favours the more parsimonious ELN distribution/regression model since in the BIC the penalty for additional parameters is stronger than that of the AIC. Of course, for other data sets the GB2 distribution/regression model, which has more parameters, may perform better than both the Pareto and ELN distributions/regression models. In such cases, if we knew that the actual distribution is a GB2, we would prefer to fit a GB2 instead of a mixed exponential model. Nevertheless, before using any distribution for carrying out different tasks, such as setting the appropriate level of premiums, reserves and reinsurance, its appropriateness for modelling claims data should always be investigated. In the case of the GB2 distribution, if the data set is large, even if the GB2 provides parameter estimates with small standard errors, these estimates may be significantly biased if the assumption that the distribution is a GB2 is not valid. Moreover, if the data set is small, the parameter estimates of the GB2 distribution can be unstable, as for instance was reported in Calderín-Ojeda, Fergusson & Wu (2017). In particular, Calderín-Ojeda, Fergusson & Wu (2017) developed a novel EM algorithm for ML estimation of the parameters of the heavy-tailed Double Pareto-Lognormal (DPLN) regression model. As was discussed in Reed & Jorgensen (2004), the DPLN model exhibits Paretian behaviour in both tails, among other theoretical properties, and hence can be considered as a valid alternative to other parametric heavy-tailed models such as the GB2 model. With this in mind, Calderín-Ojeda, Fergusson & Wu (2017) compared the performance of the DPLN distribution with the GB2 distribution and reported that regarding small sample sizes the standard errors of the parameter estimates of the DPLN distribution are noticeably smaller, highlighting the consistency of the parameter estimation for the DPLN distribution and its advantage over the GB2 distribution in this case.

4.3 Calculation of the a priori premiums

In this subsection, based on the use of the net premium calculation principle, we analyse the differences between the Pareto and ELN claim severity regression models through the mean of the cost of claims of the policyholders who belong to the 36 different risk classes, which are determined by the relevant a priori characteristics. In particular, $E(y_i|\mathbf{x}_i)$, for $i = 1, ..., n$, forms the basis of the premium for each risk class. Note that in the case of the ELN model, which does not have a pdf in closed form, the estimated expected annual claim severity for each risk class was calculated in equation (6) using the law of total expectation. Additionally, we calculate the a priori premiums that must be paid by all the different groups of policyholders based on the combinations of the NBI and PIG regression models for approximating claim frequency and the two competing claim severity regression models. Specifically, assuming independence between the claim frequency and severity components, the premium rates calculated according to net premium principle are given by $E(k_i|\boldsymbol{\chi}_i) \times E(y_i|\mathbf{x}_i)$, for $i = 1, ..., n$, in the case of the ELN, Pareto, NBI and PIG regression models, see equations (6, 19, 26 and 30) respectively.

The results are summarised in Table 5. At this point we should emphasise that if the ratemaking exercise is only based on the number of claims at fault, regardless of their severity, which is usually the case encountered in the literature, this over-simplification can be regarded as problematic since it can limit any insights one can get into the extent to which certain explanatory variables can predict insurance outcomes as these can significantly differ between the number and the costs of claims. The key idea is that both the a priori and a posteriori corrections should aim at creating tariff cells as homogeneous as possible by integrating the severity of the claims since this can improve ratemaking by providing a more complete picture to the actuary about the extent to

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Table 5. A priori premium rates, NBI, Pareto, PIG and ELN regression models

Risk class	Explanatory variables			A priori premiums					
	BM	HP	CS	Pareto	ELN	NBI-Pareto	NBI-ELN	PIG-Pareto	PIG-ELN
1	C1	C1	C1	2,557.30	2,497.14	1,290.27	1,259.92	1,289.88	1,259.54
2	C1	C1	C2	2,767.52	2,740.79	1,178.87	1,167.48	1,174.51	1,163.16
3	C1	C1	C3	2,668.06	2,629.11	846.50	834.14	820.41	808.43
4	C1	C2	C1	2,805.98	2,726.84	1,424.40	1,384.23	1,424.97	1,384.78
5	C1	C2	C2	3,036.64	2,992.91	1,301.42	1,282.67	1,297.52	1,278.83
6	C1	C2	C3	2,927.51	2,870.96	934.50	916.44	906.33	888.82
7	C1	C3	C1	2,667.53	2,663.24	1,441.02	1,438.71	1,441.45	1,439.14
8	C1	C3	C2	2,886.81	2,923.10	1,316.60	1,333.15	1,312.53	1,329.02
9	C1	C3	C3	2,783.07	2,803.99	945.40	952.51	916.82	923.71
10	C2	C1	C1	2,458.75	2,394.91	1,148.04	1,118.23	1,149.18	1,119.35
11	C2	C1	C2	2,660.87	2,628.58	1,048.91	1,036.18	1,046.40	1,033.70
12	C2	C1	C3	2,565.24	2,521.48	753.18	740.33	730.92	718.45
13	C2	C2	C1	2,697.84	2,615.21	1,267.38	1,228.56	1,269.54	1,230.65
14	C2	C2	C2	2,919.62	2,870.38	1,157.95	1,138.42	1,155.99	1,136.49
15	C2	C2	C3	2,814.69	2,753.42	831.48	813.38	807.47	789.89
16	C2	C3	C1	2,564.73	2,554.21	1,282.17	1,276.91	1,284.22	1,278.96
17	C2	C3	C2	2,775.56	2,803.43	1,171.46	1,183.23	1,169.36	1,181.10
18	C2	C3	C3	2,675.81	2,689.20	841.18	845.39	816.81	820.90
19	C3	C1	C1	2,687.08	2,636.74	1,170.06	1,148.14	1,169.47	1,147.57
20	C3	C1	C2	2,907.96	2,894.02	1,069.03	1,063.90	1,064.87	1,059.76
21	C3	C1	C3	2,803.45	2,776.09	767.63	760.14	743.83	736.57
22	C3	C2	C1	2,948.37	2,879.29	1,291.69	1,261.43	1,291.95	1,261.68
23	C3	C2	C2	3,190.74	3,160.23	1,180.16	1,168.88	1,176.39	1,165.14
24	C3	C2	C3	3,076.07	3,031.46	847.43	835.14	821.73	809.81
25	C3	C3	C1	2,802.90	2,812.14	1,306.76	1,311.07	1,306.89	1,311.20
26	C3	C3	C2	3,033.31	3,086.52	1,193.93	1,214.88	1,190.00	1,210.88
27	C3	C3	C3	2,924.3	2,960.76	857.32	868.01	831.23	841.60
28	C4	C1	C1	2,504.41	2,450.87	1,215.13	1,189.16	1,212.46	1,186.54
29	C4	C1	C2	2,710.28	2,690.01	1,110.21	1,101.91	1,104.01	1,095.76
30	C4	C1	C3	2,612.88	2,580.40	797.20	787.29	771.17	761.58
31	C4	C2	C1	2,747.94	2,676.33	1,341.45	1,306.49	1,339.44	1,304.53
32	C4	C2	C2	2,973.84	2,937.46	1,225.63	1,210.63	1,219.64	1,204.72
33	C4	C2	C3	2,866.96	2,817.77	880.08	864.97	851.93	837.31
34	C4	C3	C1	2,612.36	2,613.90	1,357.10	1,357.91	1,354.93	1,355.74
35	C4	C3	C2	2,827.10	2,868.95	1,239.93	1,258.28	1,233.74	1,252.00
36	C4	C3	C3	2,725.50	2,752.05	890.34	899.02	861.79	870.18

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

which the amounts of premiums vary according to observable characteristics of policyholders and their cars. For instance, from Table 5 we observe that the premium payment for a policyholder who belongs to the first BM category lives in a large city and has a car with HP between 0 and 1,400 cc, i.e. for the reference class, is equal to 2,557.30 and 2,497.14 euros, while another insured who belongs to the same BM category has a car with similar characteristics and lives in a small city, i.e. risk class 3, has to pay a higher premium equal to 2,668.06 and 2,629.11 euros in the case of the Pareto and ELN models respectively. However, when the number of claims is also taken into consideration, the premium payment for the first individual we described before, i.e. for the reference class, reduces to 1,290.27 and 1,259.92 euros in the case of the NBI-Pareto and NBI-ELN models respectively, and to 1,289.88 and 1,259.54 euros in the case of the PIG-Pareto and PIG-ELN models respectively, and hence is now higher than the premium required to be paid by the second individual we described before, i.e. for risk class 3, which goes down to 846.50 and 834.14 euros in the case of the NBI-Pareto and NBI-ELN models respectively, and to 820.41 and to 808.43 euros in the case of the PIG-Pareto and PIG-ELN models respectively. Moreover, we observe that regarding all combinations of both claim frequency models with the two corresponding claim severity models, the group of policyholders with the lowest expected claim severity are those who belong to the second BM category, live in a large city and have a car with HP between 0 and 1,400 cc, i.e. risk class 10, whereas the group of insureds with the highest expected claim severity are those who belong to the third BM category, live in a middle sized city and have a car with HP between 1,400 and 1,800 cc, i.e. risk class 23. On the other hand, when claim frequencies are also taken into account, the lowest premium payment is required by those insureds who belong to the second BM category, live in a small city and have a car with HP between 0 and 1,400 cc, i.e. risk class 12, whereas the highest premium payment is required by those policyholders who belong to the first BM category, live in a large city and have a car with HP greater than 1,800 cc, i.e. risk class 7.

Overall, as expected, Table 5 shows that small differences lie in the a priori premiums resulting from the Pareto and ELN claim severity models and also in those determined by their combinations with the NBI and PIG claim frequency models because, as is well known, in this case only the mean parameters of the Pareto and ELN models, which are modelled using the same covariate information, affect the estimation of the premium rates. However, on the path towards actuarial relevance where the Bayesian view will be taken to calculate the severity of the a posteriori corrections, it is the value of the dispersion parameter of the Lognormal mixing distribution that will consequently affect the calculation of the a posteriori, or BM, premium rates. In particular, as was previously mentioned, the ELN model which is less heavy-tailed than the Pareto model will show much less extreme relative a posteriori, or BM, premiums for policyholders with some claim experience.

4.4 Calculation of the a posteriori premiums

In this subsection, we examine how the ELN model responds to claim experience. Consider a policyholder i who was observed for t years of their presence in the portfolio with claim frequency history $k_i^1, ..., k_i^t$ and denote by $y_{i,k}^j$ the loss incurred from their claim k for the period j, for $i = 1, ..., n$ and $j = 1, ..., t$. Then, the information we have for their claim size history will be in the form of a vector $y_{i,1}^1, ..., y_{i,k_i^t}^t$ and the total amount of their claims will be equal to $\sum_{j=1}^{t} \sum_{k=1}^{k_i^j} y_{i,k}^j$. The problem is to determine at the renewal of the policy the expected claim severity 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. Applying Bayes theorem, one can find that the pdf of the posterior distribution of z_i^{t+1}, given claim size records $y_{i,1}^1, ..., y_{i,k_i^t}^t$ and $x_i^1, ..., x_i^{t+1}$ individual characteristics records, is given by

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

$$g\left(z_i^{t+1}|y_{i,1}^1,...,y_{i,k_i^t}^t;x_i^1,...,x_i^{t+1}\right)=\frac{f\left(y_{i,1}^1,...,y_{i,k_i^t}^t|z_i^{t+1};x_i^1,...,x_i^{t+1}\right)g\left(z_i^{t+1}\right)}{\int_0^\infty f\left(y_{i,1}^1,...,y_{i,k_i^t}^t|z_i^{t+1};x_i^1,...,x_i^{t+1}\right)g\left(z_i^{t+1}\right)dz_i^{t+1}} \tag{37}$$

where $f\left(y_{i,k_i^t}^j|\mathbf{x}_i,z_i\right)$ is the pdf of the Exponential distribution, which is given by equation (1), and where $g\left(z_i^{t+1}\right)$ is the pdf of the Lognormal prior distribution[5] which is given by equation (4). Using the quadratic loss function and the net premium principle, one can find that the mean of the posterior structure function given by[6]

$$E\left(z_i^{t+1}|y_{i,1}^1,...,y_{i,k_i^t}^t;x_i^1,...,x_i^{t+1}\right)=\int_0^\infty z_i^{t+1}g\left(z_i^{t+1}|y_{i,1}^1,...,y_{i,k_i^t}^t;x_i^1,...,x_i^{t+1}\right)dz_i^{t+1} \tag{38}$$

The expectation in equation (38) does not have a closed form expression. However, it can be easily computed based on either numerical integration or a Monte Carlo approach since neither scheme requires knowledge of the pdf of the posterior distribution of z_i^{t+1}.

Based on the aforementioned methodology, we compute the a posteriori, or BM, premium rates resulting from the ELN model based only on different claim costs, i.e. the a posteriori criteria, and based both on different claim costs and the characteristics of the policyholder and the automobile, i.e. the a priori criteria. When both criteria are considered, we examine a group of policyholders who share the following common characteristics: We consider that the policyholder i belongs to the first BM category, lives in a large city and has a car with HP between 0 and 1,400 cc. The premium rates will be divided by the premium when $t=0$, i.e. we calculate the relative premiums, since we are interested in the differences between various classes and the results are presented so that the premium for a new policyholder is 100. Table 6 (Panels A and B) shows comparable relative premiums for the Pareto and ELN distributions/regression models respectively, assuming that the cost of one claim in the first year of observation ranges from 1,000 euros to 20,000 euros. Table 6 (Panels A and B) shows that when the claim size increases the premium rates also increase. Furthermore, we observe that while for very small claim sizes up to 2,000 euros the bonuses awarded by the ELN and Pareto distributions/regression models are almost indistinguishable, in all other cases, as expected, the less heavy-tailed ELN distribution/regression model in general penalises policyholders who reported claims with moderate amounts significantly less severely than the Pareto distribution/regression model. For example, from Panel A, when only the a priori criteria are considered, we see that policyholders who had one claim size of 12,000 euros in the first year will have to pay a malus of 65.94% and 31.75% of the basic premium, while those who had one claim size of 15,000 euros in the first year will have to pay a malus of 87.27% and 40.62% of the basic premium in the case of the Pareto and ELN distributions respectively. Additionally, from Panel B when both the a priori and the a posteriori criteria are considered, we see, for instance, that policyholders who had one claim size of 14,000 euros in the first year will have to pay a malus of 86.28% and 39.88% of the basic premium, while those who had one claim size of 17,000 euros in the first year will have to pay a malus of 108.90% and 48.79% of the basic premium in the case of the Pareto and ELN regression models respectively.

Let us now compute the BMSs with a frequency and a severity component using the net premium calculation principle. As far as the claim frequency component is concerned, similar to the severity component, employing a Bayesian approach and using the quadratic error loss function, one can easily see that the BM premium rates are given by the posterior[7] mean $E\left(u_i^{t+1}|k_i^1,...,k_i^t;\chi_i^1,...,\chi_i^{t+1}\right)$, where $\chi_i^1,...,\chi_i^{t+1}$ is the vector of individual characteristics. In what follows, based on the NBI and PIG distributions/regression models and the Pareto and ELN distributions/regression models for approximating the number and the cost of claims respectively, the relative premiums resulting from those systems are calculated via the product of the posterior

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Table 6. A posteriori, or BM, premium rates, Pareto and ELN distributions/regression models

Panel A: Distributions			Panel B:Regression models		
Claim severity	Pareto	ELN	Claim severity	Pareto	ELN
1,000	87.73	93.84	1,000	88.26	94.29
2,000	94.84	97.84	2,000	95.80	98.46
3,000	101.95	101.68	3,000	103.34	102.46
4,000	109.06	105.39	4,000	110.88	106.32
5,000	116.17	108.98	5,000	118.42	110.06
6,000	123.28	112.46	6,000	125.96	113.68
7,000	130.39	115.86	7,000	133.50	117.21
8,000	137.50	119.17	8,000	141.04	120.65
9,000	144.61	122.41	9,000	148.58	124.01
10,000	151.72	125.59	10,000	156.12	127.31
11,000	158.83	128.70	11,000	163.66	130.54
12,000	165.94	131.75	12,000	171.20	133.70
13,000	173.05	134.75	13,000	178.74	136.82
14,000	180.16	137.71	14,000	186.28	139.88
15,000	187.27	140.62	15,000	193.82	142.89
16,000	194.38	143.48	16,000	201.36	145.86
17,000	201.49	146.31	17,000	208.90	148.79
18,000	208.60	149.10	18,000	216.44	151.68
19,000	215.71	151.85	19,000	223.98	154.53
20,000	222.82	154.57	20,000	231.52	157.35

mean claim frequency and the posterior mean claim severity, i.e. $E\left(u_i^{t+1}|k_i^1,...,k_i^t;\chi_i^1,...,\chi_i^{t+1}\right)\times E\left(z_i^{t+1}|y_{i,1}^1,...,y_{i,k_i^t}^t;x_i^1,...,x_i^{t+1}\right)$, assuming that the accumulative claim severities range from 1,000 euros to 20,000 euros when the age of the policy is up to $j=2$ years. Tables 7 and 8 summarise our findings with respect to the a posteriori criteria in the case of the NBI-Pareto and NBI-ELN and PIG-Pareto and PIG-ELN distributions respectively. Also, Tables 9 and 10 present our results with respect to both the a priori and a posteriori criteria in the case of the NBI-Pareto and NBI-ELN and PIG-Pareto and PIG-ELN regression models respectively. Note that for the BMSs derived based on both criteria, since the explanatory variable BM category varies substantially depending on the number of claims of policyholder *i* for year *j*, the explicit claim frequency history determines the calculation of the premium rates resulting from the NBI, PIG, Pareto and ELN regression models and not just the total number of claims *K*, as in the case when only the a posteriori criteria are considered or when the policyholder is observed for a single year and both criteria are taken into account. Due to the aforementioned reason, in the examples presented in Tables 9 and 10, we consider two cases in which we specify the exact order of the claim frequency history in order to derive the relative premiums that must be paid by the same insured we described before when only the severity component was examined. Specifically, we assume that the specific policyholder has either reported one claim in the first year and another claim in the subsequent year, i.e. $k_i^1=1,k_i^2=1$, thus $K=2$ at $j=2$, or that they have made one claim in the first year and two claims in the subsequent year, i.e. $k_i^1=1,k_i^2=2$, hence $K=3$ at $j=2$. In what follows we discuss our findings.

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Table 7. A posteriori, or BM, premium rates, NBI, Pareto and ELN distributions

Year $j=1$	NBI-Pareto		NBI-ELN		Year $j=2$	NBI-Pareto		NBI-ELN	
Claim Fr.\ Claim Sev.	$K=1$	$K=2$	$K=1$	$K=2$	Claim Fr.\ Claim Sev.	$K=1$	$K=2$	$K=1$	$K=2$
1,000	156.55	196.88	167.45	226.51	1,000	140. 98	177. 32	150. 80	204.00
2,000	169.23	212.85	174.59	237.15	2,000	152. 41	191. 70	157. 23	213. 58
3,000	181.92	228.79	181.44	247.33	3,000	163. 83	206. 05	163. 40	222. 75
4,000	194.61	244.75	188.06	257.13	4,000	175. 26	220. 43	169. 36	231. 58
5,000	207.29	260.69	194.46	266.59	5,000	186. 69	234. 79	175. 13	240. 10
6,000	219.98	276.66	200.67	275.78	6,000	198. 11	249. 17	180. 72	248. 37
7,000	232.67	292.63	206.74	284.72	7,000	209. 54	263. 55	186. 19	256. 43
8,000	245.36	308.57	212.65	293.43	8,000	220. 96	277. 91	191. 51	264. 27
9,000	258.04	324.53	218.43	301.92	9,000	232. 39	292. 29	196. 71	271. 92
10,000	270.73	340.47	224.10	310.25	10,000	243. 81	306. 64	201. 82	279. 43
11,000	283.42	356.44	229.65	318.40	11,000	255. 24	321. 20	206. 82	286. 76
12,000	296.10	372.41	235.09	326.41	12,000	266. 67	335. 40	211. 72	293. 98
13,000	308.79	388.35	240.45	334.26	13,000	278. 09	349. 76	216. 54	301. 05
14,000	321.48	404.31	245.73	341.97	14,000	289. 52	364. 14	221. 30	307. 99
15,000	334.16	420.25	250.93	349.58	15,000	300. 94	378. 50	225. 98	314. 85
16,000	346.85	436.22	256.03	357.08	16,000	312. 37	392. 88	230. 57	321. 60
17,000	359.54	452.19	261.08	364.45	17,000	323. 79	407. 26	235. 12	328. 24
18,000	372.23	468.13	266.05	371.74	18,000	335. 22	421. 61	239. 60	334. 80
19,000	384.91	484.10	270.96	378.94	19,000	346. 65	435. 99	244. 02	341. 29
20,000	397.60	500.04	275.81	386.04	20,000	358. 07	450. 35	248. 39	347. 68

Firstly, from all Tables7–10, we observe that the systems resulting from the NBI-Pareto, NBI-ELN, PIG-Pareto and PIG-ELN distributions/regression models are fair since if the total accumulated number of claims K is kept constant the premium reduces over time while it increases proportionally to the total claim severity, whereas if time and the total claim size are fixed the premium increases when the total claim frequency increases. Secondly, for very small accumulative claim costs, specifically from 1,000 up to 5,000 euros in both the first and the second year of observation, when only the a posteriori criteria are considered, see Tables 7 and 8, and from 1,000 up to 5,000 euros and from 1,000 up to 7,000 euros in the first and the second year of observation respectively, when both criteria are taken into account, see Tables 9 and 10, we observe that the BMSs resulting from the NBI-ELN and PIG-ELN models punish slightly more those policyholders who had more than one claim in a given year than the systems determined by the NBI-Pareto and PIG-Pareto models respectively. On the contrary, those individuals who had only a single small claim in a given year of a cost which is equal to the total cost of those with more than one claim are, in the majority of cases, penalised slightly more under the BMSs provided by the NBI-Pareto and PIG-Pareto models. Those two observations imply that, regarding those claims with very small amounts, the ELN model generally puts more emphasis on the frequency component and hence distributes their burden among insureds in a more fair and equitable manner than the Pareto model. For instance, from Table 7, when $j=1$, we see that policyholders who had one claim size of 4,000 euros will have to pay a malus of 94.61% and 88.06% of the basic premium, while those who had two claims with total size amounting to 4,000 euros will have to pay a malus of 144.75% and

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Table 8. A posteriori, or BM, premium rates, PIG, Pareto and ELN distributions

Year $j=1$	PIG-Pareto		PIG-ELN		Year $j=2$	PIG-Pareto		PIG-ELN	
Claim Fr.\ Claim Sev.	$K=1$	$K=2$	$K=1$	$K=2$	Claim Fr.\ Claim Sev.	$K=1$	$K=2$	$K=1$	$K=2$
1,000	152.85	221.55	163.50	254.89	1,000	132.72	187.18	141.96	215.35
2,000	165.24	239.52	170.47	266.86	2,000	143.47	202.36	148.01	225.46
3,000	177.63	257.46	177.16	278.32	3,000	154.23	217.51	153.82	235.14
4,000	190.02	275.42	183.62	289.35	4,000	165.00	232.69	159.43	244.46
5,000	202.40	293.36	189.88	300.00	5,000	175.74	247.85	164.86	253.45
6,000	214.79	311.33	195.94	310.33	6,000	186.50	263.03	170.13	262.19
7,000	227.18	329.30	201.86	320.40	7,000	197.25	278.21	175.27	270.70
8,000	239.57	347.23	207.63	330.20	8,000	208.01	293.36	180.28	278.97
9,000	251.95	365.20	213.27	339.76	9,000	218.77	308.54	185.18	287.05
10,000	264.34	383.14	218.82	349.13	10,000	229.52	323.70	190.00	294.97
11,000	276.73	401.11	224.23	358.30	11,000	240.28	338.88	194.70	302.71
12,000	289.12	419.07	229.55	367.31	12,000	251.03	354.06	199.31	310.33
13,000	301.51	437.01	234.77	376.14	13,000	261.79	369.21	203.85	317.79
14,000	313.89	454.98	239.93	384.83	14,000	272.55	384.39	208.33	325.12
15,000	326.28	472.92	245.00	393.39	15,000	283.30	399.55	212.73	332.36
16,000	338.67	490.88	250.00	401.83	16,000	294.06	414.73	217.06	339.49
17,000	351.06	508.85	254.92	410.12	17,000	304.81	429.91	221.34	346.49
18,000	363.44	526.79	259.78	418.32	18,000	315.57	445.06	225.56	353.42
19,000	375.83	544.76	264.57	426.43	19,000	326.33	460.24	229.72	360.27
20,000	388.22	562.69	269.31	434.42	20,000	337.08	475.40	233.83	367.02

157.13% of the basic premium in the case of the NBI-Pareto and NBI-ELN distributions respectively. From Table 7, when $j=2$, we observe that policyholders who had one claim size of 4,000 euros will have to pay a malus of 75. 26% and 69.36% of the basic premium, while those who had two claims with total size equal to 4,000 euros will have to pay a malus of 120.43% and 131. 58% of the basic premium in the case of the NBI-Pareto and NBI-ELN distributions respectively. From Table 8, when $j=1$, we see that policyholders who had one claim size of 4,000 euros will have to pay a malus of 90.02% and 83.62% of the basic premium, while those who had two claims with total size amounting to 4,000 euros will have to pay a malus of 175.42% and 189.35% of the basic premium in the case of the PIG-Pareto and PIG-ELN distributions respectively. From Table 8, when $j=2$, we observe that policyholders who had one claim size of 4,000 euros will have to pay a malus of 65% and 59.43% of the basic premium, while those who had two claims with total size equal to 4, 000 euros will have to pay a malus of 132.69% and 144.46% of the basic premium in the case of the PIG-Pareto and PIG-ELN distributions respectively. From Table 9, when $j=1$, we see that policyholders who had one claim size of 5,000 euros will have to pay a malus of 103.16% and 88.82% of the basic premium, while those who had two claims with total size equal to 5,000 euros in the first year will have to pay a malus of 128.93% and 130.69% of the basic premium in the case of the NBI-Pareto and NBI-ELN regression models respectively. From Table 9, when $j=2$, we observe that policyholders who had claim frequency history $k_i^1=1, k_i^2=1$ (i.e. total number of claims $K=2$ at $j=2$) and the total size of their claims amounts to 7,000 euros will have to pay

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Table 9. A posteriori, or BM, premium rates, NBI, Pareto and ELN regression models

Year $j=1$	NBI-Pareto		NBI-ELN		Year $j=2$	NBI-Pareto		NBI-ELN	
Claim Fr.\ Claim Sev.	$K=1$	$K=2$	$K=1$	$K=2$	Claim Fr.\ Claim Sev.	$k_i^1=1$, $k_i^2=1$	$k_i^1=1$, $k_i^2=2$	$k_i^1=1$, $k_i^2=1$	$k_i^1=1$, $k_i^2=2$
1,000	151.42	169.04	161.76	193.96	1,000	151.16	173. 02	173. 44	208. 56
2,000	164.35	184.01	168.92	203.74	2,000	164.55	188. 34	182. 19	220. 10
3,000	177.29	199.00	175.78	213.07	3,000	177.95	203. 66	190. 53	231. 06
4,000	190.23	213.97	182.40	222.04	4,000	191.33	218. 98	198. 56	241. 56
5,000	203.16	228.93	188.82	230.69	5,000	204.72	234. 32	206. 29	251. 70
6,000	216.10	243.90	195.03	239.07	6,000	218.10	249. 64	213. 78	261. 49
7,000	229.03	258.87	201.09	247.22	7,000	231.49	264. 96	221. 07	270. 98
8,000	241.97	273.83	206.99	255.15	8,000	244. 87	280. 27	228. 16	280. 22
9,000	254.90	288.80	212.75	262.89	9,000	258. 25	295. 59	235. 09	289. 23
10,000	267.84	303.77	218.41	270.47	10,000	271. 64	310. 91	241. 86	298. 04
11,000	280.78	318.74	223.95	277.88	11,000	285. 02	326. 23	248. 49	306. 66
12,000	293.71	333.73	229.38	285.15	12,000	298. 43	341. 57	254. 99	315. 10
13,000	306.65	348.69	234.73	292.28	13,000	311. 81	356. 89	261. 37	323. 38
14,000	319.58	363.66	239.98	299.32	14,000	325. 19	372. 21	267. 66	331. 52
15,000	332.52	378.63	245.14	306.21	15,000	338. 58	112. 26	273. 82	339. 55
16,000	345.45	393.59	250.24	313.00	16,000	351. 96	402. 84	279. 90	347. 43
17,000	358.39	408.56	255.26	319.71	17,000	365. 35	418. 16	285. 89	355. 20
18,000	371.32	423.53	260.22	326.34	18,000	378. 73	433. 51	291. 82	362. 86
19,000	384.26	438.50	265.11	332.85	19,000	392. 11	448. 82	297. 64	370. 43
20,000	397.20	453.46	269.95	339.29	20,000	405. 50	464. 14	303. 40	377. 90

a malus of 131.49% and 121. 07% of the basic premium, while those who had claim frequency history $k_i^1=1, k_i^2=2$ (i.e. total number of claims $K=3$ at $j=2$) and the total size of their claims amounts to 7,000 euros will have to pay a malus of 164.96% and 170.98% of the basic premium, in the case of the NBI-Pareto and NBI-ELN regression models respectively. From Table 10, when $j=1$, we see that policyholders who had one claim size of 5, 000 euros will have to pay a malus of 87.98% and 74.71% of the basic premium, while those who had two claims with total size equal to 5,000 euros in the first year will have to pay a malus of 135.12% and 136.92% of the basic premium in the case of the PIG-Pareto and PIG-ELN regression models respectively. From Table 10, when $j=2$, we observe that policyholders who had claim frequency history $k_i^1=1, k_i^2=1$ (i.e. total number of claims $K=2$ at $j=2$) and the total size of their claims amounts to 7, 000 euros will have to pay a malus of 108.76% and 99.36% of the basic premium, while those who had claim frequency history $k_i^1=1, k_i^2=2$ (i.e. total number of claims $K=3$ at $j=2$) and the total size of their claims amounts to 7,000 euros will have to pay a malus of 157.25% and 163.10% of the basic premium, in the case of the PIG-Pareto and PIG-ELN regression models respectively. Finally, similar to the results presented in the Table 6, in every other case, Tables 7–10 show that the BMSs resulting from the NBI-ELN and PIG-ELN distributions/regression models penalise policyholders with moderate costs considerably less severely than the systems determined by the NBI-Pareto and PIG-Pareto distributions/regression models. For example, from Table 7, when $j=1$, we see that policyholders who had $K=1$ claim size of 14,000 euros will have to pay a malus of 221.48%

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Table 10. A posteriori, or BM, premium rates, PIG, Pareto and ELN regression models

Year $j=1$	PIG-Pareto		PIG-ELN		Year $j=2$	PIG-Pareto		PIG-ELN	
Claim Fr.\ Claim Sev.	$K=1$	$K=2$	$K=1$	$K=2$	Claim Fr.\ Claim Sev.	$k_i^1=1$, $k_i^2=1$	$k_i^1=1$, $k_i^2=2$	$k_i^1=1$, $k_i^2=1$	$k_i^1=1$, $k_i^2=2$
1,000	140.10	173.61	149.68	199.20	1,000	136.32	168.00	156.41	202.49
2,000	152.07	188.98	156.30	209.24	2,000	148.39	182.86	164.30	213.70
3,000	164.04	204.38	162.65	218.83	3,000	160.48	197.74	171.82	224.34
4,000	176.01	219.75	168.77	228.04	4,000	172.53	212.61	179.06	234.53
5,000	187.98	235.12	174.71	236.92	5,000	184.62	227.51	186.03	244.37
6,000	199.95	250.49	180.46	245.53	6,000	196.67	242.38	192.78	253.89
7,000	211.92	265.86	186.06	253.90	7,000	208.76	257.25	199.36	263.10
8,000	223.89	281.24	191.52	262.05	8,000	220.81	272.10	205.74	272.07
9,000	235.86	296.61	196.85	270.00	9,000	232.88	287.00	212.01	280.82
10,000	247.82	311.98	202.09	277.78	10,000	244.97	301.87	218.10	289.37
11,000	259.79	327.35	207.22	285.39	11,000	257.02	316.74	224.08	297.74
12,000	271.76	342.75	212.24	292.86	12,000	269.13	331.61	229.94	305.91
13,000	283.73	358.12	217.19	300.18	13,000	281.18	346.51	235.71	313.95
14,000	295.70	373.49	222.05	307.41	14,000	293.25	361.38	241.38	321.88
15,000	307.67	388.86	226.82	314.48	15,000	305.32	109.00	246.92	329.67
16,000	319.64	404.23	231.54	321.46	16,000	317.39	391.10	252.42	337.32
17,000	331.61	419.60	236.19	328.35	17,000	329.46	405.97	257.80	344.87
18,000	343.58	434.97	240.78	335.16	18,000	341.53	420.90	263.15	352.30
19,000	355.55	450.35	245.30	341.85	19,000	353.60	435.74	268.40	359.63
20,000	367.51	465.72	249.78	348.46	20,000	365.67	450.61	273.60	366.90

and 145.73% of the basic premium, while those who had $K=2$ claims with total size amounting to 14,000 euros will have to pay a malus of 304.31% and 241.97% of the basic premium in the case of the NBI-Pareto and NBI-ELN distributions respectively. From Table 7, when $j=2$, we observe that policyholders who had $K=1$ claim size of 17,000 euros will have to pay a malus of 223.79% and 135.12% of the basic premium, while those who had $K=2$ claims with total size equal to 17,000 euros will have to pay a malus of 307.26% and 228.24% of the basic premium in the case of the NBI-Pareto and NBI-ELN distributions respectively. From Table 8, when $j=1$, we see that policyholders who had $K=1$ claim size of 14,000 euros will have to pay a malus of 213.89% and 139.93% of the basic premium, while those who had $K=2$ claims with total size amounting to 14,000 euros will have to pay a malus of 354.98% and 284.83% of the basic premium in the case of the PIG-Pareto and PIG-ELN distributions respectively. From Table 8, when $j=2$, we observe that policyholders who had $K=1$ claim size of 17,000 euros will have to pay a malus of 204.81% and 121.34% of the basic premium, while those who had $K=2$ claims with total size equal to 17,000 euros will have to pay a malus of 329.91% and 246.49% of the basic premium in the case of the PIG-Pareto and PIG-ELN distributions respectively. From Table 9, when $j=1$, we see that policyholders who had $K=1$ claim size of 15,000 euros will have to pay a malus of 232.52% and 145.14% of the basic premium, while those who had $K=2$ claims with total size equal to 15,000 euros in the first year will have to pay a malus of 278.63% and 206.21% of the basic premium in the case of the NBI-Pareto and NBI-ELN regression models respectively. From Table 9, when $j=2$,

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

we observe that policyholders who had claim frequency history $k_i^1 = 1, k_i^2 = 1$ (i.e. total number of claims $K = 2$ at $j = 2$) and the total size of their claims amounts to 18,000 euros will have to pay a malus of 278.73% and 191.82% of the basic premium, while those who had claim frequency history $k_i^1 = 1, k_i^2 = 2$ (i.e. total number of claims $K = 3$ at $j = 2$) and the total size of their claims amounts to 18, 000 euros will have to pay a malus of 333.51% and 262.86% of the basic premium, in the case of the NBI-Pareto and NBI-ELN respectively. From Table 10, when $j = 1$, we see that policyholders who had $K = 1$ claim size of 15,000 euros will have to pay a malus of 207.67% and 126.82% of the basic premium, while those who had $K = 2$ claims with total size equal to 15, 000 euros in the first year will have to pay a malus of 288.86% and 214.48% of the basic premium in the case of the PIG-Pareto and PIG-ELN regression models respectively. From Table 10, when $j = 2$, we observe that policyholders who had claim frequency history $k_i^1 = 1, k_i^2 = 1$ (i.e. total number of claims $K = 2$ at $j = 2$) and the total size of their claims amounts to 18,000 euros will have to pay a malus of 241.53% and 163.15% of the basic premium, while those who had claim frequency history $k_i^1 = 1, k_i^2 = 2$ (i.e. total number of claims $K = 3$ at $j = 2$) and the total size of their claims amounts to 18,000 euros will have to pay a malus of 320.90% and 252.30% of the basic premium, in the case of the PIG-Pareto and PIG-ELN respectively.

Overall, as was illustrated in Tables 6–10, it is reasonable to agree that in MTPL data sets like the one used in this study, where moderate observations constitute the largest proportion of our sample, whereas large observations have very low frequencies, under the Pareto model it is very likely that policyholders will bear the cost of claims themselves because of the growth in premium payments, while the employment of the new model is beneficial for the insurance company as it can enable them to adopt a generally more mild pricing strategy for policyholders who reported a large number of claims with moderate severities, leading to a discouragement of the Bonus–Hunger phenomenon.

5. Conclusions

The main purpose of this paper was to propose an EM type algorithm that reduces the computational burden for ML estimation in the ELN regression model. The ELN regression model extends the commonly used specification that assumes that claims costs are distributed according to the Pareto regression model which was widely accepted for designing merit rating plans in accordance with the a priori ratemaking structure of the insurance company. The ELN model has just the appropriate level of generality for deriving both a priori and a posteriori ratemaking mechanisms, since while its upper tail can sufficiently fit large size claims, the ELN model can be considered as a candidate model for approximating moderate claim severities with high frequencies. Furthermore, the ELN regression model is suitable for application not only in insurance ratemaking but also in survival analysis, since, as is well known, all the distributions with decreasing failure rate can be retrieved as mixtures of the exponential distribution, see, for instance, Proschan (1963) and Barlow & Proschan (1975). Additionally, it should be noted that the novel EM type algorithm we developed in this study was based on the mixture representation of the ELN model, and did not require knowledge of its pdf, which could not be written in closed form, while it is computationally parsimonious and can avoid overflow problems which may occur via other numerical maximisation schemes. Therefore, it is obvious that the ML estimation scheme we presented has the considerable mathematical flexibility for fitting an abundance of mixed Exponential regression models stemming from several other mixing distributions which are not conjugate to the Exponential.

An interesting possible line of further research would be to go through the ratemaking exercise based on generalisations of the proposed model such as a finite mixture of the ELN model and two component mixture models, where the first component distribution is the ELN and the second component model can be a different more or less heavy-tailed claim severity distribution, thus providing alternative options to the insurer when they are deciding on their pricing strategies,

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

see Tzougas, Vrontos & Frangos (2014), and Ni *et al.* (2014) and Tzougas, Vrontos & Frangos (2018) respectively. The log-likelihood function of the general models can be maximised without special effort using standard techniques for finite mixtures, see Bohning (1999). Finally, the data augmentation which was used in the paper to derive the EM algorithm can be the basis for constructing Bayesian estimation methods, including functional forms other than the linear proceeding along similar lines as Klein *et al.* (2014) in which Bayesian generalised additive models for location, scale and shape claim frequency models were employed for non-life ratemaking and risk management.

Acknowledgements. The research reported here was supported by LSE LIFE and LSE Teaching and Learning Centre. We would like to thank the anonymous referee for their constructive comments and suggestions that have greatly improved the article. We also would like to thank the participants at the 11th International Conference of the ERCIM WG on Computational and Methodological Statistics. The usual disclaimer applies.

Notes

1 Note that the Pareto response distribution can also be derived as a mixture of the Exponential distribution with Gamma mixing weights. However, the Exponential-Inverse Gamma mixture representation of the Pareto is preferred for ratemaking purposes. Specifically, the latter parameterisation assumes that the mean is an explicit parameter of the Pareto distribution. This allows easier interpretation when the mean is modelled in terms of explanatory variables and provides a framework for analysing the extent to which a priori and a posteriori ratings interact. The Exponential-Lognormal model we consider in this study is derived in a similar way by using a Lognormal mixing distribution instead of the Inverse Gamma one.

2 Note that the BM premium rates calculated according to the net premium principle based only on the a posteriori criteria are obtained if the regression components of the NBI, PIG, Pareto and ELN regression models are limited to constants.

3 Note that the deviance residuals have also been traditionally utilised to examine goodness of fit of GLMs. Moreover, Pierce & Schafer (1986) indicated that the deviance residuals should be more nearly normal than the Pearson residuals. However, the deviance residuals cannot be guaranteed to be closely normal when the data are highly dispersed relative to the mean and their distribution is mostly skewed. As was mentioned in Dunn & Smyth (1996), normalised quantile residuals remedy the above-mentioned problem of the deviance residuals. Furthermore, the only information needed for computing normalised quantile residuals is knowing the cumulative distribution function of the response variable, which makes calculation much easier than deviance residuals, which might be challenging to define in more complex models. Finally, the randomised quantile residuals can also be applied for model diagnosis when the response variable does not belong to the GLM family of models.

4 Note also that the residuals and the right tail index for the ELN model which does not have a pdf in closed form can be accurately computed based on numerical integration methods.

5 Note that in the case of the Pareto model $g\left(z_i^{t+1}\right)$ is the pdf of the Inverse Gamma prior distribution which is given by equation (17).

6 Note also that for the sake of brevity, we have not included the BM premium functions resulting from the traditional NBI, PIG and Pareto models. Those functions can be easily computed in closed form based on the methodology presented, for instance, in Dionne & Vanasse (1989, 1992), Frangos & Vrontos (2001), Mahmoudvand & Hassani (2009) and Tzougas, Vrontos & Frangos (2014, 2018) respectively. Note also that the Bonus–Malus premium rates calculated according to the net premium principle based only on the a posteriori criteria are obtained if the regression components of the ELN, Pareto, NBI and PIG regression models, see equations (5, 18, 25 and 28), are limited to constants.

7 Note that u_i^{t+1} follows the Gamma prior distribution which is given by equation (24) in the case of the NBI model.

Note also that u_i^{t+1} follows the Inverse Gaussian prior distribution which is given by equation (27) in the case of the PIG model.

References

Barlow, R.E. & Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing: Probability Models*, Holt, Rinehart and Winston, New York, USA.

Bhattacharya, S.K. & Kumar, S. (1986). E-IG model in life testing. *Calcutta Statistical Association Bulletin,* **35**(1–2), 85–90.

Bohning, D. (1999). *Computer Assisted Analysis of Mixtures and Applications in Meta-Analysis, Disease Mapping and Others*. CRC Press, New York, USA.

Booth, J.G. & Hobert, J.P. (1999). Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **61**(1), 265–285.

Booth, J.G., Hobert, J.P. & Jank, W. (2001). A survey of Monte Carlo algorithms for maximizing the likelihood of a two-stage hierarchical model. *Statistical Modelling,* **1**(4), 333–349.

https://doi.org/10.1017/S1748499519000034 Published online by Cambridge University Press

Boucher, J.P., Denuit, M. & Guillen, M. (2007). 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.

Boucher, J.P., Denuit, M. & Guillen, M. (2008). Models of insurance claim counts with time dependence based on generalisation of Poisson and negative binomial distributions. *Variance,* **2**(1), 135–162.

Brouhns, N., Guillen, M., Denuit, M. & Pinquet, J. (2003). Bonus–Malus scales in segmented tariffs with stochastic migration between segments. *Journal of Risk and Insurance,* **70**(4), 577–599.

Burnham, K.P. & Anderson, D.R. (2003). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer, New York, USA.

Calderín-Ojeda, E., Fergusson, K. & Wu, X. (2017). An EM algorithm for Double-Pareto-Lognormal Generalized Linear Model applied to heavy-tailed insurance claims. *Risks*, **5**(4), 60.

De Jong, P. & Heller, G.Z. (2008). *Generalized Linear Models for Insurance Data.* Cambridge University Press, Cambridge, UK.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **39**(1), 1–38.

Denuit, M. & Lang, S. (2004). Nonlife ratemaking with Bayesian GAM's. *Insurance: Mathematics and Economics,* **35**(3), 627–647.

Denuit, M., Marechal, X., Pitrebois, S. & Walhin, J. F. (2007). *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus–Malus Systems*. John Wiley and Sons, Chichester, West Sussex, UK.

Dionne, G. & Vanasse, C. (1989). A generalization of actuarial automobile insurance rating models: the negative binomial distribution with a regression component. *ASTIN Bulletin,* **19**(2), 199–212.

Dionne, G. & Vanasse, C. (1992). Automobile insurance ratemaking in the presence of asymmetrical information. *Journal of Applied Econometrics,* **7**(2), 149–165.

Dunn, P.K. & Smyth, G.K. (1996). Randomized quantile residuals. *Computational and Graphical Statistics,* **5**(3), 236–245.

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997). Modelling extremal events. In *Applications of Mathematics*, vol. 33. Springer-Verlag, Berlin, Heidelberg, DE.

Frangos, N. & Vrontos, S. (2001). Design of optimal Bonus–Malus systems with a frequency and a severity component on an individual basis in automobile insurance. *ASTIN Bulletin*, **31**(1), 1–22.

Frangos, N. & Karlis, D. (2004). Modelling losses using an exponential-inverse Gaussian distribution. *Insurance: Mathematics and Economics,* **35**(1), 53–67.

Frees, E.W. & Valdez, E.A. (2008). Hierarchical insurance claims modeling. *Journal of the American Statistical Association*, **103**(484), 1457–1469.

Frees, E.W. (2010). *Regression Modelling with Actuarial and Financial Applications*. Cambridge University Press, Cambridge, UK.

Frees, E.W., Derrig, R.A. & Meyers, G. (2014a). *Predictive Modeling Applications in Actuarial Science,* vol. 1. Cambridge University Press, New York, USA.

Gómez-Déniz, E., Sarabia, J.M. & Calderín-Ojeda, E. (2008). Univariate and multivariate versions of the negative Binomial-Inverse Gaussian distributions with applications. *Insurance, Mathematics and Economics,* **42**(1), 39–49.

Gómez-Déniz, E., Hernández-Bastida, A. & Fernández-Sánchez, M.P. (2014). Computing credibility Bonus–Malus premiums using the aggregate claims distribution. *Hacettepe Journal of Mathematics and Statistics,* **43**(6), 1047–1061.

Gómez-Déniz, E. & Calderín-Ojeda, E. (2018). Multivariate credibility in Bonus–Malus systems distinguishing between different types of claims. *Risks*, **6**(2), 34.

Haberman, S. & Renshaw, A. E. (1996). Generalized linear models and actuarial science. *The Statistician,* **45**(4), 407–436.

Halliwell, Leigh J. (2013). Classifying the tails of loss distributions. *Casualty Actuarial Society E-Forum*, **2** (Spring 2013), 1.

Hesselager, O., Wang, S. & Willmot, G. (1998). Exponential and scale mixtures and equilibrium distributions. *Scandinavian Actuarial Journal,* **1998**(2), 125–142.

Insurance Europe. (2016). European motor insurance markets addendum. https://www.insuranceeurope.eu/sites/default/files/attachments/European%20motor%20insurance%20markets%20addendum.pdf.

Jones, B.L. & Zitikis, R. (2003). Empirical estimation of risk measures and related quantities. *North American Actuarial Journal,* **7**(4), 44–54.

Kaas, R., Goovaerts, M., Dhaene, J. & Denuit, M. (2008). *Modern Actuarial Risk Theory: Using R*. Springer-Verlag, Berlin, Heidelberg, DE.

Karlis, D., Tzougas, G. & Frangos, N. (2018). Confidence intervals of the premiums of optimal Bonus Malus systems. *Scandinavian Actuarial Journal,* **2018**(2), 129–144.

Klein, N., Denuit, M., Lang, S. & Kneib, T. (2014). Nonlife ratemaking and risk management with Bayesian generalized additive models for location, scale, and shape. *Insurance: Mathematics and Economics,* **55**, 225–249.

Klugman, S., Panjer, H. & Willmot, G. (2012). *Loss Models: From Data to Decisions*, vol. 715. John Wiley and Sons, New York, USA.

Lemaire, J. (1995). *Bonus–Malus Systems in Automobile Insurance*. Kluwer Academic Publishers, Dordrecht, NE.

Louis, T.A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology),* **44** (2), 226–233.

Mahmoudvand, R. & Hassani, H. (2009). Generalized Bonus–Malus systems with a frequency and a severity component on an individual basis in automobile insurance. *ASTIN Bulletin,* **39**(1), 307–315.

McDonald, J.B. & Xu, Y.J. (1995). A generalization of the beta distribution with applications. *Journal of Econometrics* **66**(1–2), 133–152.

McDonald, J.B. (1996). *Probability distributions for financial models*, vol. 14. In: G. S. Maddala & C. R. Rao (Eds.), *Handbook of Statistics*. Elsevier Science. Amsterdam, NL.

McLachlan, G. & Krishnan, T. (2007). *The EM Algorithm and Extensions,* vol. 382. John Wiley and Sons, Hoboken, NJ, USA.

Mert, M. & Saykan, Y. (2005). On a Bonus–Malus system where the claim frequency distribution is geometric and the claim severity distribution is Pareto. *Hacettepe Journal of Mathematics and Statistics,* **34**, 75–81.

Munich Re (2017). Group Annual Report 2017. https://www.munichre.com/site/corporate/get/documents_E-1569124979/mr/assetpool.shared/Documents/0_Corporate_Website/Financial_Reports/2018/annual-report-2017/302-09072_en.pdf.

Ni, W., Constantinescu, C. & Pantelous, A.A. (2014). Bonus–Malus systems with Weibull distributed claim severities. *Annals of Actuarial Science,* **8**(2), 217–233.

Ni, W., Li, B., Constantinescu, C. & Pantelous, A.A. (2014). Bonus-Malus systems with hybrid claim severity distributions. In *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management* - Proceedings of the 2nd International Conference on Vulnerability and Risk Analysis and Management, ICVRAM 2014 and the 6th International Symposium on Uncertainty Modeling and Analysis, ISUMA 2014 (pp. 1234–1244). https://doi.org/10.1061/9780784413609.124.

Pierce, D.A. & Schafer, D.W. (1986). Residuals in generalized linear models. *Journal of the American Statistical Association*, **81**(396), 977–986.

Picech, L. (1994). *The Merit-Rating Factor in a Multiplicating Rate-Making Model. ASTIN Colloquium*, Cannes, France.

Pinquet, J. (1997). Allowance for cost of claims in Bonus-Malus systems. *ASTIN Bulletin,* **27**(1), 33–57.

Pinquet, J. (1998). Designing optimal Bonus-Malus systems from different types of claims. *ASTIN Bulletin,* **28**(2), 205–220.

Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics,* **5**(3), 375–383.

Raftery, A.E. (1995). Bayesian model selection in social research. *Sociological Methodology,* **25**, 111–163.

Reed, W.J. & Jorgensen, M. (2004). The Double Pareto-Lognormal distribution – a new parametric model for size distributions. *Communications in Statistics-Theory and Methods* **33**(8), 1733–1753.

Rigby, R.A. & Stasinopoulos, D.M. (2009). *A flexible regression approach using GAMLSS in R*. http://www.gamlss.org.

Santi, D.N., Purnaba, I.G.P & Mangku, I.W. (2016). Bonus–Malus system with the claim frequency distribution is geometric and the severity distribution is truncated Weibull. In *Conference Series: Earth and Environmental Science,* vol. 31(1), pp. 1755-1315, IOP Publishing.

Stasinopoulos, D.M., Rigby, B. & Akantziliotou, C. (2008). *Instructions on how to use the gamlss package in R*. Second Edition. http://www.gamlss.org.

Trembley, L. (1992). Using the Poisson Inverse Gaussian in Bonus–Malus systems. *ASTIN Bulletin,* **22**(1), 97–106.

Tzougas, G. & Frangos, N. (2014). The design of an optimal Bonus–Malus system based on the Sichel distribution. In *Modern Problems in Insurance Mathematics* (pp. 239–260). Cham, Switzerland, Springer.

Tzougas, G., Vrontos, S. & Frangos, N. (2014). Optimal Bonus–Malus systems using finite mixture models. *ASTIN Bulletin,* **44**(2), 417–444.

Tzougas, G., Vrontos, S. & Frangos, N. (2015). Risk classification for claim counts and losses using regression models for location, scale and shape. *Variance,* **9**(1), 140–157.

Tzougas, G., Vrontos, S. & Frangos, N. (2018). Bonus–Malus systems with two component mixture models arising from different parametric families. *North American Actuarial Journal,* **22**(1), 55–91.

Tzougas, G., Hoon, W.L. & Lim, J.M. (2018). The Negative Binomial-Inverse Gaussian Regression Model with an application to insurance ratemaking. *European Actuarial Journal,* (2018), 1–22. https://doi.org/10.1007/s13385-018-0186-2.

Wang, S. (1998). An actuarial index of the right-tail risk. *North American Actuarial Journal,* **2**(2), 88–101.

Cite this article: Tzougas G, Yik WH and Mustaqeem MW (2020). Insurance ratemaking using the Exponential-Lognormal regression model. *Annals of Actuarial Science* **14**, 42–71. https://doi.org10.1017/S1748499519000034