

Bonus–Malus systems with Weibull distributed claim severities

Weihong Ni

Department of Mathematical Sciences, Institute for Actuarial and Financial Mathematics, University of Liverpool, UK

Corina Constantinescu

Department of Mathematical Sciences, Institute for Actuarial and Financial Mathematics, University of Liverpool, UK

Athanasios A. Pantelous*

Department of Mathematical Sciences, Institute for Actuarial and Financial Mathematics, University of Liverpool, UK; Institute for Risk and Uncertainty, University of Liverpool, UK

Abstract

One of the pricing strategies for Bonus–Malus (BM) systems relies on the decomposition of the claims' randomness into one part accounting for claims' frequency and the other part for claims' severity. By mixing an exponential with a Lévy distribution, we focus on modelling the claim severity component as a Weibull distribution. For a Negative Binomial number of claims, we employ the Bayesian approach to derive the BM premiums for Weibull severities. We then conclude by comparing our explicit formulas and numerical results with those for Pareto severities that were introduced by Frangos & Vrontos.

Keywords

Bonus–Malus Systems; Weibull Distribution; Bayesian Estimator

1. Introduction

A Bonus–Malus (BM) system is referring to a merit rating system where policyholders get discounts when having no claims and they are punished for making claims. That is to say, the premiums for a policyholder to pay depends very much on his past claim records. This system has now been widely used in many European insurance companies for automobile insurance. However, due to the particular feature of a BM system, one of the biggest issues related to introducing the system is the bonus hunger problem. This means that in order to achieve a premium discount in the following year, the insured may not report some small claims and pay the costs themselves. As a consequence the insurance company is missing some useful information for calculating individual premiums. To our knowledge, this phenomenon is still present in many models used so far in the literature (e.g. Frangos & Vrontos, 2001). In this paper, we introduce a new BM premium model, which we

*Correspondence to: Athanasios A. Pantelous, Department of Mathematical Sciences, Institute for Actuarial and Financial Mathematics, University of Liverpool, UK. E-mail: A.Pantelous@liverpool.ac.uk

numerically analyse using the same structure as in Frangos & Vrontos (2001), namely, keeping the total claim size fixed. Under this restriction, our model generates a premium function that reflects a discouragement on the bonus hunger reaction.

The designing of BM systems can date back to the middle of the 20th century, since when only the number of claims reported in the past was considered in the calculation of the future premium a policyholder is to pay. In this case, Picard (1976) claimed that problems might arise if one had a claim worth much more money while the other had lots of small claims. The original objective to develop a fair system is then violated and the sharing of the premiums is not fair among these two policyholders. Thus, introducing the claim severity component is a natural consideration.

Following this thought, Lemaire (1995) first applied Picard's (1976) method to Belgian data that distinguishes between small and large claims. Lemaire (1995) found that this classification would lead to serious practical problems, since it is time consuming to assess the exact amount and many policyholders who have claims just over the limit argue a lot. Thus, instead of categorising the claim types, Frangos & Vrontos (2001) proposed a specific modelling of claim severity. As the claim sizes in motor insurance seem to exhibit long-tails, distributions with this property are thought-out. For instances Valdez & Frees (2005) used a distribution called the Burr XII long-tailed distribution to fit the claim sizes data from Singapore Motor Insurance. Another obvious choice to obtain, e.g., Pareto (Frangos and Vrontos, 2001) is mixing the exponential parameter with an Inverse Gamma. In this paper, we will combine the result of mixing distribution presented in Albrecher *et al.* (2011) and the idea of applying the Bayes' theorem as proposed by Frangos & Vrontos (2001) for Weibull severities instead of Pareto. We will discuss in detail the motivation and the consequences of choosing a Weibull distribution for the severities.

Our motivation could be interpreted from both academic and practical perspectives. Technically speaking, since Pareto in Frangos & Vrontos (2001) turned out to fit well on the data, especially in the tails, it is preferable to have some similar-shaped distributions, and Weibull distribution is a reasonable candidate. Even though Weibull distribution does not have tails as long as Pareto, in reality, it can rely on reinsurance that usually alleviates the burden of extremely large claims. Thus, practically speaking, if Weibull distribution fits well the moderate claims of an insurance company, this combined with reinsurance would be the best choice. The additional value of choosing a Weibull fit in some instances is that this could address the bonus hunger problem, as it will be illustrated later in the paper. This is an advantage of such choice, since by carrying out this model, the tendency of policyholders for not reporting small claims could be discouraged.

We conclude this paper by fitting an exponential, a Weibull and a Pareto on a given data set. All the upper tails of the three distributions are written as follows with one common parameter θ :

$$\text{Exponential : } P(X > x) = \exp(-\theta x)$$

$$\text{Weibull : } P(X > x) = \exp(-\theta x^\gamma)$$

$$\text{Pareto : } P(X > x) = \left(\frac{\theta}{\theta + x} \right)^s$$

It can be seen that if $\gamma < 1$, the tail of a Weibull distribution is fatter than that of the Exponential but thinner than that of the Pareto distribution. However, when $\gamma = 1$, the Weibull tail is lighter than that of the Exponential and a special case appears when $\gamma = 1$ where the Weibull distribution becomes an Exponential distribution. Thus, we aim to find a Weibull distribution with its shape parameter less

than 1 so that a longer-tail property can be retained. Fortunately enough, we have found that the heterogeneity of the claim severity could be described by a Lévy (1/2 stable) distribution. Consequently, when the mixing of this distribution is carried out on an Exponential distribution, a Weibull distribution is obtained in the end (Albrecher *et al.*, 2011). What is more motivating is that the shape parameter γ is known and equal to 1/2, which is less than 1, fitting our aim.

This paper will be presented via the structure described below. Section 2 discusses the procedure of deriving the premiums for our model. The core of this article lies in section 2.2. It involves explanation of how the mixing distribution is achieved as well as the Bayesian approach. The premium formula is obtained in the end with several analysis described in the subsequent subsection. Section 3 is dedicated to applying both our Weibull model and the one using Pareto claim severities on some given data. Discussions on the findings are concluded in section 4.

2. Premium Calculation Based on an Individual Policyholder

2.1. Modelling of the claim frequency

The modelling of the claim counts is as in Frangos & Vrontos (2001), so that a comparison of the results can be clearer. We give a very brief explanation here, as details are available in Frangos & Vrontos (2001).

As in Frangos & Vrontos (2001), mixing the Poisson intensity Λ with a Gamma (α, τ) is a Negative Binomial:

$$P(N = n) = \int_0^\infty \frac{e^{-\lambda} \lambda^n}{n!} \times \frac{\lambda^{\alpha-1} \tau^\alpha e^{-\tau\lambda}}{\Gamma(\alpha)} d\lambda = \binom{n+\alpha-1}{n} \tau^\alpha \left(\frac{1}{1+\tau}\right)^{\alpha+n}$$

Furthermore, by applying the Bayesian approach, the posterior distribution is given by

$$\mu(\lambda \mid n_1, n_2, \dots, n_t) = \frac{(\tau+t)^{K+\alpha} \lambda^{K+\alpha-1} e^{-(t+\tau)\lambda}}{\Gamma(\alpha+K)}$$

where $K = \sum_{i=1}^t n_i$ represents the total claim frequency over t years with n_i denoting the claim numbers in each year, respectively. It is easily seen that this posterior distribution is still a Gamma but with new parameters $K + \alpha$ and $t + \tau$. According to the Bayesian analysis, the posterior mean is the best estimate, which is given as follows:

$$\lambda_{t+1}(n_1, n_2, \dots, n_t) = \frac{\alpha+K}{t+\tau} \tag{1}$$

This also represents the expected claim frequency for the coming period, as the mean of a Poisson is λ itself.

2.2. Modelling of the claim severity

In this paper, our focus lies in the claim severity distribution. We use the Weibull distribution to model the claim severities whose applications appear not only in reliability engineering and failure analysis, but also in insurance survival analysis and sometimes in reinsurance (Boland, 2007), its probability density function (PDF) is

$$f(x) = c\gamma x^{\gamma-1} \exp(-cx^\gamma), x \geq 0, c > 0, \gamma > 0$$

and its cumulative density function (CDF) is

$$F(x) = 1 - \exp(-cx^\gamma), x \geq 0$$

It was found by Albrecher *et al.* (2011) that a mixing of a Lévy distribution on the exponential distribution would result in a Weibull distribution with its shape parameter known as 1/2. Suppose the exponential distribution is denoted as

$$f(X = x | \theta) = \theta e^{-\theta x}$$

with a distribution function (CDF)

$$F(X \leq x | \theta) = 1 - e^{-\theta x}$$

And the parameter θ is assumed to be Lévy distributed that is also referred to as the stable (1/2) distribution. Then we have a prior distribution described as follows:

$$\pi(\Theta = \theta) = \frac{c}{2\sqrt{\pi\theta^3}} \exp\left(-\frac{c^2}{4\theta}\right), \theta > 0$$

Hence, we obtained the distribution function as follows (proof in Appendix A):

$$F(x) = 1 - \exp(-c\sqrt{x}), x \geq 0$$

It is the Weibull distribution with shape parameter equal to 1/2.

Furthermore, with similar approach to Frangos & Vrontos (2001), we need to find the posterior distribution using the Bayes' theorem. Suppose the insurance company receives a sequence of claim costs $\{x_1, x_2, \dots, x_K\}$ from a policyholder with total of K claims over the time horizon considered. If we let $M = \sum_{i=1}^K x_i \geq 0$ to describe the total amount of all these claims, the posterior structure function is written in the following form according to Bayes' theorem.

$$\pi(\theta | x_1, x_2, \dots, x_K) = \frac{\left[\prod_{i=1}^K f(x_i | \theta)\right] \pi(\theta)}{\int_0^\infty \left[\prod_{i=1}^K f(x_i | \theta)\right] \pi(\theta) d\theta} = \frac{\theta^{K-\frac{2}{3}} \exp\left(-\left(M\theta + \frac{c^2}{4\theta}\right)\right)}{\int_0^\infty \theta^{K-\frac{2}{3}} \exp\left(-\left(M\theta + \frac{c^2}{4\theta}\right)\right) d\theta} \tag{2}$$

We know that after the integration, the denominator will become independent from θ . By omitting all terms that are not related to θ , we can obtain the kernel of this distribution.

$$\pi(\theta | x_1, x_2, \dots, x_K) \propto \theta^{K-\frac{2}{3}} \exp\left(-\left(M\theta + \frac{c^2}{4\theta}\right)\right)$$

$$\pi(\theta | x_1, x_2, \dots, x_K) \propto \theta^p \exp\left(-\left(\frac{\theta}{q} + \frac{r}{\theta}\right)\right)$$

where $p = K - \frac{3}{2}, q = \frac{1}{M}, r = \frac{c^2}{4}$. This is a form of a Generalised Inverse Gaussian distribution (Tremblay, 1992). Going back to equation (2), by slightly modifying the variables, it can be rewritten as

$$\pi(\theta | x_1, x_2, \dots, x_K) = \frac{\left(\frac{c}{2\sqrt{M}}\right)^{-(K-\frac{1}{2})} \theta^{K-\frac{2}{3}} \exp\left(-\left(M\theta + \frac{c^2}{4\theta}\right)\right)}{\int_0^\infty \left(\frac{2\sqrt{M}\theta}{c}\right)^{K-\frac{3}{2}} \exp\left(-\frac{c\sqrt{M}}{2} \left(\frac{2\sqrt{M}\theta}{c} + \frac{c}{2\sqrt{M}\theta}\right)\right) d\left(\frac{2\sqrt{M}\theta}{c}\right)} \tag{3}$$

The integral on the denominator can be transformed to a modified Bessel function, whose integral representation is normally given as follows (Abramowitz & Stegun, 1964):

$$B_\nu(x) = \int_0^\infty e^{-x \cosh t} \cosh(\nu t) dt$$

However, we cannot make a direct connection with what we have simply from this expression.

Proposition. An alternative integral representation of the modified Bessel function is given as follows

$$B_\nu(x) = \frac{1}{2} \int_0^\infty \exp\left(-\frac{1}{2}x\left(y + \frac{1}{y}\right)\right) y^{\nu-1} dy, x > 0$$

Proof of Appendix B. □

As compared to the integral in equation (3), it is not difficult to rewrite the posterior distribution in the form below:

$$\pi(\theta) = \frac{\left(\frac{c}{2\sqrt{M}}\right)^{-(K-\frac{1}{2})} \theta^{K-\frac{3}{2}} \exp\left(-\left(M\theta + \frac{c^2}{4\theta}\right)\right)}{2B_{K-\frac{1}{2}}(c\sqrt{M})}$$

Or alternatively as

$$\pi(\theta) = \frac{\left(\frac{\alpha'}{\beta'}\right)^{\frac{\nu}{2}} \theta^{\nu-1} \exp\left(-\frac{1}{2}\left(\alpha'\theta + \frac{\beta'}{\theta}\right)\right)}{2B_\nu\left(\sqrt{\alpha'\beta'}\right)}$$

where $\alpha' = 2M, \beta' = \frac{c^2}{2}, \nu = K - \frac{1}{2}$. From the properties of a Generalised Inverse Gaussian distribution, the expectation is shown below (Ebrechts, 1983):

$$E[\text{GIG}] = \sqrt{\frac{\beta'}{\alpha'}} \frac{B_{\nu+1}\left(\sqrt{\alpha'\beta'}\right)}{B_\nu\left(\sqrt{\alpha'\beta'}\right)}$$

Since our model distribution was assumed to be exponential whose conditional mean is given by $E(X|\theta) = \frac{1}{\theta}$. By integrating $\frac{1}{\theta}$ with respect to the posterior distribution $\pi(\theta)$, one gets

$$E[\text{Claim severity}] = \frac{2\sqrt{M} B_{K-\frac{3}{2}}(c\sqrt{M})}{c B_{K-\frac{1}{2}}(c\sqrt{M})}$$

With the claim frequency (equation (1)) considered, this expression contributes to the explicit formula of the estimated net premium for the following period.

$$\text{Premium} = \frac{\alpha + K}{t + \tau} \times \left(\frac{2\sqrt{M} B_{K-\frac{3}{2}}(c\sqrt{M})}{c B_{K-\frac{1}{2}}(c\sqrt{M})}\right)$$

Now the problem has reduced to calculate the ratio of the above two Bessel functions. As described by Lemaire (1995), two properties of the modified Bessel function could be considered here, i.e., for any $x > 0, B_\nu(x)$ satisfies the following two conditions:

$$B_{-\nu}(x) = B_\nu(x)$$

$$B_{\nu+1}(x) = \frac{2\nu}{x} B_\nu(x) + B_{\nu-1}(x)$$

If we let:

$$\frac{B_{\nu-1}(c\sqrt{M})}{B_\nu(c\sqrt{M})} = \frac{B_{K-\frac{3}{2}}(c\sqrt{M})}{B_{K-\frac{1}{2}}(c\sqrt{M})} = Q_K(c\sqrt{M})$$

then it can be easily seen that $Q_1 = 1$ from the first condition. In addition, we can write a recursive function for Q_k based on the second condition:

$$\frac{1}{Q_{K+1}(c\sqrt{M})} = \frac{2K-1}{c\sqrt{M}} + Q_K(c\sqrt{M}), K \geq 1$$

This will finally contribute to the calculation of the premium.

On the other hand, it is not difficult to see that our premium model is not defined when $M = 0$. This denotes the scenario where there are no claims. So at the same time $K = 0$. Hence, we will redefine the premium rates for this case. As we assumed the claim severity is Weibull distributed in a single year for each policyholder, it would be reasonable to assume that our initial premium is equal to the product of the mean claim frequency (Negative Binomial) and severity (Weibull). Therefore, the base premium for any new entrant is set as follows:

$$P_0 = \frac{\alpha}{\tau} \times \frac{2}{c^2}$$

Then after, the premium for the following years if no claims are filed will be given by

$$\text{Premium} |_{M=0} = \left(\frac{\alpha}{t+\tau}\right) \left(\frac{2}{c^2}\right) \tag{4}$$

This means that when there are no claims reported the premium would be discounted with the time a policyholder is within the company.

2.3. Analysis on the premiums

In this subsection, we would like to analyse further on the premium functions from Frangos & Vrontos (2001) as well as ours. Their premium function is concave. Nevertheless, our model presents a more complex shape, as one can see by analysing the second-order derivatives of the premium functions.

Initially, let us see the premium expression given by Frangos & Vrontos (2001).

$$P_{FV} = \frac{\alpha+K}{t+\tau} \times \frac{m+M}{s+K-1}$$

where $m > 0, s > 1$ are the parameters in the Pareto distribution originally coming from the Inverse Gamma distribution, and other notations are the same as above. The difference equation with respect to K is obtained as follows:

$$P_{FV}(K+1) - P_{FV}(K) = \frac{M+m}{t+\tau} \times \frac{s-\alpha-1}{(s+K-1)(s+K)} \tag{5}$$

Since $K \geq 0, s > 1$, if $s-\alpha-1 > 0$, which is normally the case, we can conclude that the premium is strictly increasing with K . This will be further illustrated with our numerical example in section 3.

Subsequently, we look at the monotonicity of our Weibull model regarding the variable K when we keep M fixed. By analysing the ratio of two successive K values for the premium function we obtained above, we have

$$\frac{\text{Premium}(K+1)}{\text{Premium}(K)} = \frac{\alpha+K+1}{\alpha+K} \times \frac{Q_{K+1}(c\sqrt{M})}{Q_K(c\sqrt{M})}$$

Clearly, the left half of the above formula is larger than 1. However, the right half is less than 1, which is explained as follows:

$$\frac{Q_{K+1}(c\sqrt{M})}{Q_K(c\sqrt{M})} = \frac{B_{K-\frac{1}{2}}^2(c\sqrt{M})}{B_{K-\frac{3}{2}}(c\sqrt{M})B_{K+\frac{1}{2}}(c\sqrt{M})} < 1$$

The “<” comes from the Turán-type inequalities whose proof can be found in Ismail & Muldoon (1978) and Lorch (1994):

$$B_\nu^2(x) < B_{\nu-1}(x)B_{\nu+1}(x)$$

Thus, one sees that the monotonicity depends on chosen parameters. This gives evidence that our premium function distinguishes from the Pareto model. In our numerical examples we have identified the specific pattern of our premium.

3. Numerical illustration

3.1. Data description

In this section, application of the model proposed in section 2 will be illustrated in more details. Initially, a brief description of the data will be given. The data structure is originally obtained from table 2.27 of Klugman *et al.* (1998). But we modified the scale of the data by dividing 5 so that it makes sense in the British environment. By keeping the group structure unchanged, the sample size was also shrunken to 250. The grouped data could be found in Appendix C. However, for some computational reasons, we have randomly generated a data set based on this grouped data. A summary of the data is shown below with its histogram underneath.

As can be seen from this histogram, most of the claims lie below £20,000. The situation where claim severities exceed £45,000 is very rare. In our sample, there are a total of three policyholders claiming more than this amount. We are not concerned much about the tail here, since practically there is settlement for severe claims, through reinsurance, for instance (Table 1).

3.2. Fitting the distributions

In order to compare with Frangos & Vrontos (2001), both of the two distributions used to model claim severity will be fitted using our data set. Initially, the Pareto distribution is applied (Figure 1). Our estimates are obtained using *R*. The results of our maximum likelihood estimation for Pareto distribution is

$$\begin{cases} m = 1,999.985031 \\ s = 1.343437 \end{cases}$$

And the dot–dashed curve in Figure 2 is the fitted Pareto distribution for the data without the three outliers. On the other hand, the parameter in our Weibull distribution was also estimated through maximum likelihood estimation. (We have a known shape parameter equal to 1/2.) The estimated

Table 1. Description of the data.

Minimum	1 st Qu.	Median	Mean	3 rd Qu.	Maximum
10	240	1,395	4,538	4,103	102,722

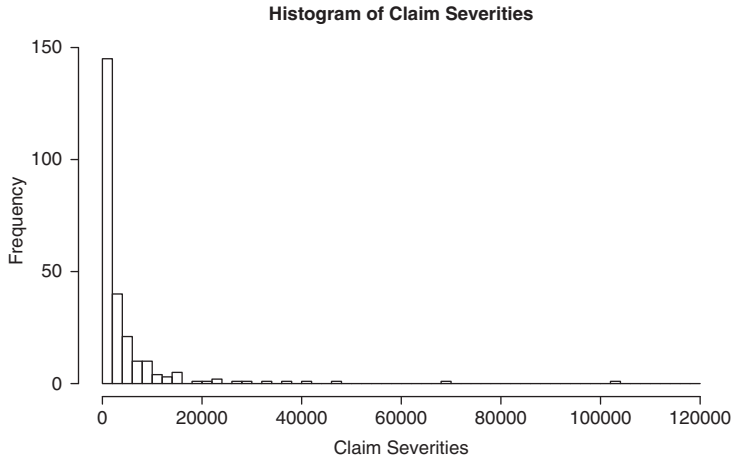


Figure 1. Histogram of the data.

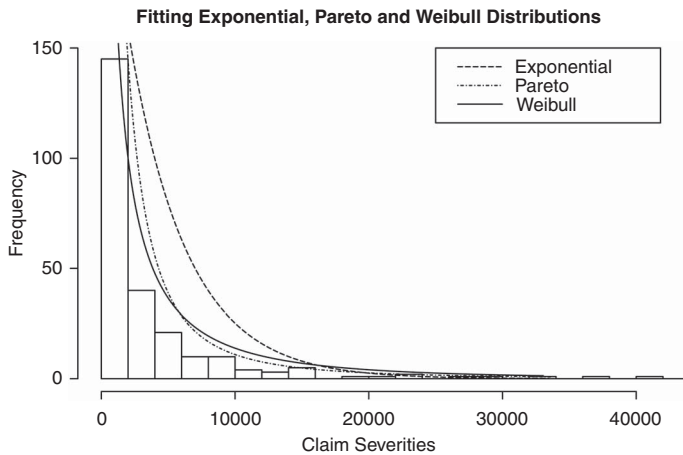


Figure 2. The fitted curves on the data without the outliers.

scale parameter is equal to 2,227.752. However, in our case, the PDF of the Weibull distribution was written in the form

$$f(x) = \frac{c}{2} x^{-\frac{1}{2}} e^{-c\sqrt{x}}, x > 0$$

Hence, the estimate for our parameter c is obtained by modifying the scale parameter, where $c = 2,227.752^{-\frac{1}{2}} = 0.02118686$. The fitted curve is shown by the solid line in the Figure 2.

There is another fitted curve in this graph, which is drawn by the dashed line, and it is an exponential distribution fitting to the data. The two mixing distributions appear to fit much better than the exponential distribution. On the other hand, in order to see clearly how the curves are fitted to the

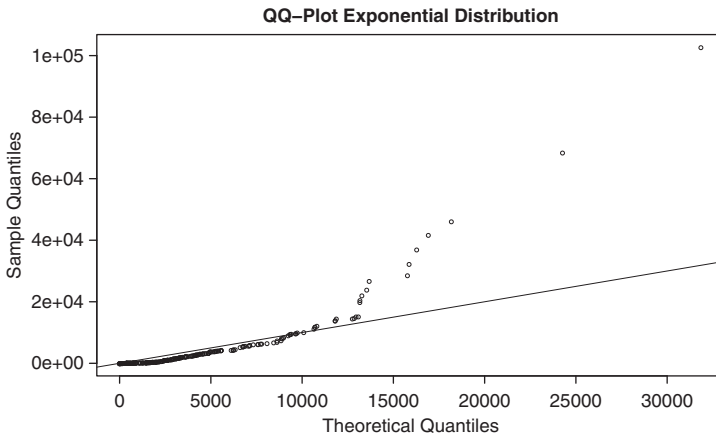


Figure 3. QQ-plot of the exponential distribution versus the sample data.

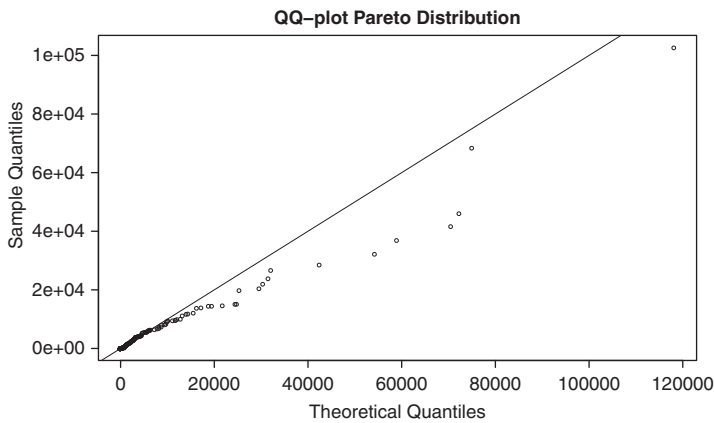


Figure 4. QQ-plot of the Pareto distribution versus the sample data.

data, corresponding QQ plots are presented as follows. The tail behaviours of the three distributions can be seen clearly and are consistent with what is discussed section 1.

It is shown on this sketch that the exponential distribution fits the data relatively better up to around £12,500. At the tail where there are extremely large claim sizes, it fails to estimate the probability accurately (Figure 3).

Figure 4 demonstrates the goodness-of-fit of the Pareto distribution to our data. Several over-estimations for some small- and medium-sized claims are present. Nevertheless, it fits well especially for very expensive claims. This has emphasised its heavy-tail feature.

It seems that the Weibull distribution fits very well up to £40,000, although there is slight perturbation. In the tail, it fits better than the Exponential distribution but worse than the Pareto distribution, as what is expected (Figure 5).

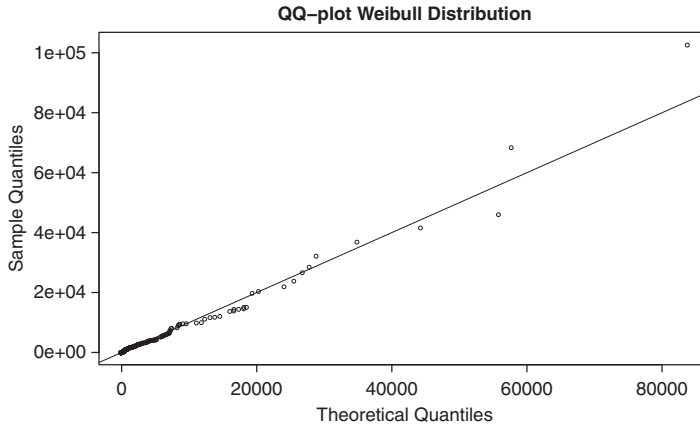


Figure 5. QQ-plot of the Weibull distribution versus the sample data.

Overall, the exponential distribution does not perform well compared to the other two. While Weibull fits better for smaller claims, Pareto yields the best performance for very large claim sizes. From these plots, it is likely to suggest a mixture of strategies. When the claim sizes are moderate, the insurer is advised to adopt the Weibull claim severities. Particularly, when reinsurance is in place, Weibull distribution can be the better choice. On the contrary, for very large claim sizes, Pareto distribution plays the key role due to its long-tail property.

3.3. Calculations for net premiums

As mentioned before, the net premiums are calculated via the product of the expected claim frequency and the expected claim severity with independence between the two components assumed. Regarding the claim frequency component, we proceed as Frangos & Vrontos (2001). Their estimates for the parameters α and τ are

$$\begin{cases} \alpha = 0.228 \\ \tau = 2.825 \end{cases}$$

In terms of the claim severity component, our first task is to estimate Q_k , as mentioned in the final part of last section. Since we have a recursive function for Q_k and its initial value is known as 1, this can be easily calculated. We used MATLAB to generate the solutions for $K = 1, 2, 3, 4, 5$ and $t = 1, 2, 3, 4, 5$. Tables 2 and 3 underneath demonstrate the resulting premium rates for our Weibull model, and the other two show those for the Pareto model. The first column in each table denotes the scenario where no claims are reported to the company ($M = 0$). While we derive these values from equation (4), Frangos & Vrontos (2001) used the following formula to calculate the first column premium rates. Notice that tables with different total claim severities have the same first column.

$$P_{FV} | (M = 0) = \frac{\alpha}{\tau + t} \times \frac{m}{s - 1}$$

First, let us see our premiums. Table 2 describes the premiums for the situation where the accumulative claim costs are £7,500. Table 3 gives the ones where the total claim amounts are £10,000. Overall, it follows the pattern that the premium decreases over time if claim frequencies are kept constant. How the BM system works will be illustrated in the next example. For instance, if a policyholder has a claim that costs £7,500 in his 1st policy year, his premium will rise to

Table 2. Optimal net premiums with Weibull severities and total claim cost ($M = 7,500$).

Year (t)	Number of claims					
	0	1	2	3	4	5
0	359.6			NA		
1	265.6	2,624.6	3,082.1	3,022.9	2,856.7	2,704.7
2	210.5	2,080.6	2,443.3	2,396.4	2,264.7	2,144.2
3	174.4	1,723.4	2,023.9	1,985.0	1,875.9	1,776.1
4	148.8	1,470.9	1,727.3	1,694.2	1,601.0	1,515.8
5	129.8	1,282.9	1,506.6	1,477.7	1,396.4	1,322.1

Table 3. Optimal net premiums with Weibull severities and total claim cost ($M = 10,000$).

Year (t)	Number of claims					
	0	1	2	3	4	5
0	359.6			NA		
1	265.6	3,030.6	3,735.4	3,802.0	3,677.7	3,528.7
2	210.5	2,402.5	2,961.3	3,014.0	2,915.5	2,797.4
3	174.4	1,990.1	2,452.9	2,496.6	2,415.0	2,317.1
4	148.8	1,698.5	2,093.5	2,130.8	2,061.1	1,977.6
5	129.8	1,481.4	1,826.0	1,858.5	1,797.7	1,724.9

£2,624.6 (Table 2). If in the subsequent year, he has another claim whose severity is £2,500. The total accumulated number of claims is now two and the total size amounts to £10,000 (Table 3). He is then subject to pay £2,961.3 in the next year. And if no more claims are filed in the following 3 years, his payment will reduce to £1,826.0 from the beginning of year 6. Now it is essential to see how the BM system using the Pareto model works. Again Tables 4 and 5 represent the total claim cost of £7,500 and £10,000, respectively.

For the same insured person described before, this system will in general punish less severely than the previous one. Specifically, this customer will pay £470 at the start. Owing to one claim reported in the 1st year, his premium is raised to £2,270.2 (Table 4). Then after a second claim in the subsequent year, an increase to £2,364.5 occurs (Table 5). If he has no more claims till the end of the 5th year, his premium payment could go down to £1,458.0. Up to now, the flow of the two systems appears to be similar except that the punishment is less severe in the latter one.

However, for this data set, there is an unexpected finding of our results. Unlike the Pareto fitting (Frangos & Vrontos, 2001), when the total claim size is fixed, our premium is not strictly increasing with the number of claims, but starts to drop slightly for more than two or three claims in our example. The last few columns in Tables 2 and 3 have demonstrated this phenomenon. In order to see how our calculated premiums behave when compared to the Pareto model we have plotted the premiums for large quantities of claims ($K = 100$) when the total claim severity is kept unchanged (Figure 6). Again, two cases where the total cost of claims is £7,500 and £10,000, respectively, are analysed. This irregular behaviour of the severity component of this premium formula may be different depending on the nature of the data set to which the Weibull model is

Table 4. Optimal net premiums with pareto severities and total claim cost ($M = 7,500$).

Year (t)	Number of claims					
	0	1	2	3	4	5
0	470.0			NA		
1	347.1	2,270.2	2,361.3	2,397.9	2,417.6	2,430.0
2	275.2	1,799.7	1,871.9	1,900.9	1,916.6	1,926.4
3	227.9	1,490.8	1,550.6	1,574.6	1,587.6	1,595.7
4	194.5	1,272.3	1,323.4	1,343.9	1,354.9	1,361.9
5	169.7	1,109.7	1,154.3	1,172.1	1,181.8	1,187.8

Table 5. Optimal net premiums with Pareto severities and total claim cost ($M = 10,000$).

Year (t)	Number of claims					
	0	1	2	3	4	5
0	470.0			NA		
1	347.1	2,867.7	2,982.7	3,028.9	3,053.9	3,069.5
2	275.2	2,273.3	2,364.5	2,401.2	2,420.9	2,433.3
3	227.9	1,883.1	1,958.6	1,989.0	2,005.3	2,015.6
4	194.5	1,607.2	1,671.6	1,697.5	1,711.5	1,720.3
5	169.7	1,401.8	1,458.0	1,480.6	1,492.8	1,500.4

applied because the monotonicity of our premium function is affected by the parameters, as mentioned in section 2.

As presented in Figure 6, under a fixed total claim cost, our premium function is not strictly increasing with the number of claims. It reaches a peak at some point and then decreases and finally asymptotically converges. To our knowledge so far, this behaviour has not appeared in classical models. However, with this effect, the bonus hunger problems could in some way be alleviated. As seen in our previous tables, if one client has already claimed twice with a total cost slightly less than £7,500 in one particular year. In the system using Pareto claim severities, it is very likely that he will bear the cost himself if he has one more small accident and does not disclose this information to the insurer because of growth in premium payment. However, in our system, there will be a reward if he reveals the true information to the company for a further small claim. It is where the premiums start to decrease. Notice that our premiums only drop a little when the additional claim size is not very large. If one has one more severe claim, he will be paying more than the current amount. Hence, our model is helpful to encourage the insured to be truthful to insurance companies. In this way, it is a flexible decision for the insurer to either fix the premium rates after a certain level of claim frequency or to reward the policyholders for disclosing true information about cheap claims.

On the other hand, this model actually distinguishes the premium payments more according to various claim costs. It punishes more those who make a few large claims than the ones who report many small claims. An extreme example could be seen in Figure 6. When comparing an insured person who makes one claim worth £7,500 and the other who has ten claims summing up to £7,500,

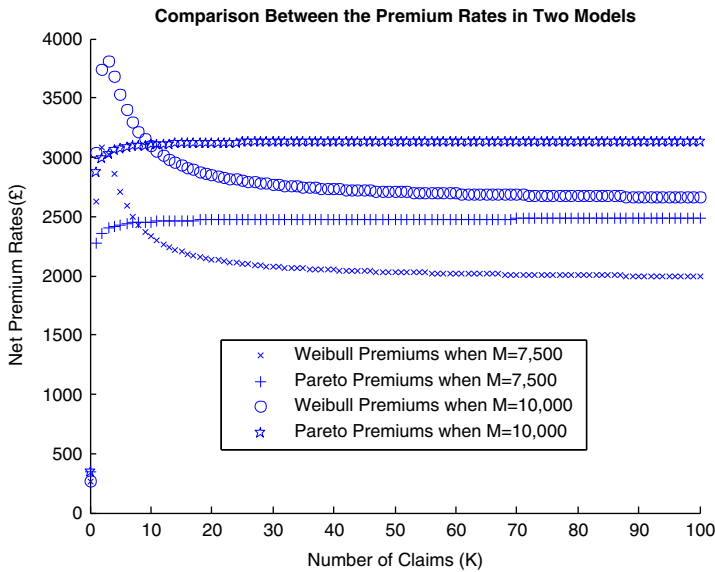


Figure 6. Behaviours of premium rates with respect to the number of claims when the total claim sizes is £7,500 and £10,000.

which means he has an average of £750 for each claim, we could see that the former one is paying a similar amount to the latter one or even slightly more. Obviously, this is not the case in the Pareto model. This observation implies that our model emphasises more on the claim severity component while the Pareto model addresses more on the frequency level.

It is also noticeable that our initial premium payments are lower than the Pareto model that might be more preferable to starting policyholders thus creating more competitiveness to the insurer. Hence, we proposed a slightly different pricing strategy for insurance companies here. This additional option for insurers leads to a diversity of premium strategies where a suitable one could be chosen from to adapt to certain needs.

4. Discussions and Conclusions

As we have seen from our results, the BM system adopting the Weibull severities provides a lower entrance premium level and punishes more for fewer large-sized claims and less for numerous small claims. This could be considered as one of the options when insurers are deciding on pricing strategies. We would suggest our model to be adopted by those who prefer to offer a mild treatment to policyholders with many small claims, discouraging thus the hunger for bonus phenomenon. We recommend that the insurance company would consider all these factors when choosing among models. Sometimes a mixture of the models would be the most rational choice.

To conclude, this paper extends (Frangos & Vrontos, 2001) by choosing a different severity distribution. We analysed the BM system only considering the a posteriori information. By comparing the two models when applied to the same data set, although we provide worse estimation in the tail, we offer cheaper initial premiums and more reasonable, especially to those who claim many times

but with small severities. Furthermore, on this data set our model seems to discourage the hunger for bonus phenomenon, which is empirically significant.

Further extensions are possible if the a priori information is considered. A potentially fruitful direction of research might be the consideration of a BMS that integrates the a priori and a posteriori information on an individual basis based on the framework developed by Dionne & Vanasse (1989) and Frangos & Vrontos (2001).

Acknowledgements

The authors would like to acknowledge Nicholas Frangos and Spyros Vrontos for the introduction to the topic, Dominik Kortschak and Bo Li for helpful remarks during the revision of this paper. Moreover, we would like to thank the anonymous reviewer for his/her suggestions, and the idea to consider the generalised BMS that integrates the a priori and a posteriori information on an individual basis. The authors also kindly acknowledge partial support from the RARE-318984 project, a Marie Curie IRSES Fellowship within the 7th European Community Framework Programme. The paper has received the Best Paper Award at the 2013 Perspectives on Actuarial Risks in Talks of Young researchers (PARTY) conference, Ascona, Switzerland, January 27th–February 1st.

References

- Abramowitz, M. & Stegun, I.A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Dover, New York, NY.
- Albrecher, H., Constantinescu, C. & Loisel, S. (2011). Explicit ruin formulas for models with dependence among risks. *Insurance: Mathematics and Economics*, **48**(2), 265–270.
- Boland, P.J. (2007). *Statistical and Probabilistic Methods in Actuarial Science*. Chapman & Hall/CRC, Boca Raton, FL.
- Dionne, G. & Vanasse, C. (1989). A generalization of automobile insurance rating models: the negative binomial distribution with a regression component. *ASTIN Bulletin*, **19**(2), 199–212.
- Embrechts, P. (1983). A property of the generalized inverse Gaussian distribution with some applications. *Journal of Applied Probability*, **3**, 537–544.
- Frangos, N.E. & Vrontos, S.D. (2001). Design of optimal Bonus-Malus systems with a frequency and a severity component on an individual basis in automobile insurance. *ASTIN Bulletin*, **31**, 1–22.
- Ismail, M.E.H. & Muldoon, M.E. (1978). Monotonicity of the zeros of a cross-product of Bessel functions. *SIAM Journal on Mathematical Analysis*, **9**(4), 759–767.
- Klugman, S.A., Panjer, H.H. & Willmot, G.E. (1998). *Loss Models: From Data to Decisions*. John Wiley & Sons, New Jersey.
- Lemaire, J. (1995). *Bonus-Malus Systems in Automobile Insurance*. Kluwer Academic Publishers, Boston, MA.
- Lorch, L. (1994). Monotonicity of the zeros of a cross product of Bessel functions. *Methods and Applications of Analysis*, **1**(1), 75–80.
- Picard, P. (1976). Généralisation de l'étude sur la survenance des sinistres en assurance automobile. *Bulletin Trimestriel de l'Institut des Actuaire Français*, **296**, 204–268.
- Tremblay, L. (1992). Using the poisson inverse Gaussian in Bonus-Malus systems. *ASTIN Bulletin*, **22**(1), 97–106.
- Valdez, E.A. & Frees, E.W. (2005). Longitudinal modeling of Singapore Motor Insurance, working paper.

Appendix A: Derivation of Weibull distribution

Some techniques were used to calculate the distribution function of the mixing distribution:

$$F(x) = \int_0^\infty (1 - e^{-\theta x}) \frac{c}{2\sqrt{\pi\theta^3}} \exp\left(-\frac{c^2}{4\theta}\right) d\theta$$

By the change of variables, let

$$\delta = \frac{c}{2\sqrt{\theta}}, \quad \theta = \frac{c^2}{2\delta^2}$$

with $\theta \in (0, \infty)$, ν will decrease from ∞ to 0. Then $d\delta = -\frac{1}{2} \frac{c}{2\sqrt{\theta^3}} d\theta$ and $-2d\delta = \frac{c}{2} \theta^{-\frac{3}{2}} d\theta$. The integral is then modified to the following form:

$$\begin{aligned} F(x) &= \int_0^\infty \left(1 - e^{-\frac{c^2 x}{4\delta^2}}\right) \frac{2}{\sqrt{\pi}} e^{-\delta^2} d\delta \\ &= 1 - \frac{2}{\sqrt{\pi}} \int_0^\infty \exp\left\{-\left(\frac{c^2}{4\delta^2} x + \delta^2\right)\right\} d\delta \end{aligned} \tag{A.1}$$

The equating of $\int_0^\infty \frac{2}{\sqrt{\pi}} e^{-\delta^2} d\delta = 1$ comes from the well-known fact that $\int_0^\infty e^{-\delta^2} d\delta = \frac{\sqrt{\pi}}{2}$. Before continuing the above equation, we let

$$\frac{c^2 x}{4} = a \geq 0$$

and use the letter I to denote the integral in equation (A.1), then it becomes

$$I = \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-\left(\frac{a}{\delta^2} + \delta^2\right)} d\delta$$

On the other hand, if we look at the integral:

$$I_0 = \int_0^\infty e^{-\left(\delta - \frac{\sqrt{a}}{\delta}\right)^2} d\left(\delta - \frac{\sqrt{a}}{\delta}\right) = \int_0^\infty e^{-\left(\delta - \frac{\sqrt{a}}{\delta}\right)^2} d\delta + \int_0^\infty e^{-\left(\delta - \frac{\sqrt{a}}{\delta}\right)^2} \frac{\sqrt{a}}{\delta^2} d\delta$$

Now let $\epsilon = \frac{\sqrt{a}}{\delta^2}$ with δ increasing from 0 to infinity. ϵ is decreasing from infinity to 0. And again we have $d\epsilon = -\frac{\sqrt{a}}{\delta^2} d\delta$. Hence, the latter term in the above integral is altered. We have:

$$I_0 = 2 \int_0^\infty e^{-\left(\delta - \frac{\sqrt{a}}{\delta}\right)^2} d\delta \tag{A.2}$$

On the other hand, if we let

$$g = \delta - \frac{\sqrt{a}}{\delta} \in (-\infty, \infty)$$

then

$$\frac{dg}{d\delta} = 1 + \frac{\sqrt{a}}{\delta^2} > 0$$

This indicates that g is monotonically increasing from minus infinity to infinity. Thus, the integral I_0 is alternatively written as

$$I_0 = \int_0^\infty e^{-(\delta-\frac{\sqrt{a}}{\delta})^2} \left(1 + \frac{\sqrt{a}}{\delta^2}\right) d\delta = \int_{-\infty}^\infty e^{-g^2} dg = \sqrt{\pi} \tag{A.3}$$

Combining the results of equations (A.1) and (A.2), we know that

$$\int_0^\infty e^{-(\delta-\frac{\sqrt{a}}{\delta})^2} d\delta = \frac{\sqrt{\pi}}{2}$$

Recall that the integral we would like to solve is actually

$$I = \frac{2}{\sqrt{\pi}} e^{-2\sqrt{a}} \int_0^\infty e^{-(\delta-\frac{\sqrt{a}}{\delta})^2} d\delta = e^{-2\sqrt{a}} = \exp(-c\sqrt{x})$$

Therefore, we have proved that the resulting distribution function is

$$F(x) = 1 - \exp(-c\sqrt{x}) \quad \square$$

Appendix B: Proof of the Transformed Integral Representation of the Modified Bessel Function

The form we often see on various mathematical handbooks regarding the modified Bessel function is presented below:

$$B_\nu(x) = \int_0^\infty e^{-x \cosh t} \cosh(\nu t) dt$$

Expanding the cosh terms, we can write,

$$B_\nu(x) = \int_0^\infty \exp\left(-\frac{x}{2}(e^t + e^{-t})\right) \frac{e^{\nu t} + e^{-\nu t}}{2} dt$$

It can be transformed initially by changing of the variable. We let $y = e^t$ with $t \in (0, \infty)$, $y \in (1, \infty)$ and $t = \ln y$, $dt = \frac{1}{y} dy$. Substituting into the above formula yields

$$B_\nu(x) = \frac{1}{2} \int_1^\infty \exp\left(-\frac{x}{2}\left(y + \frac{1}{y}\right)\right) y^{\nu-1} dy + \frac{1}{2} \int_1^\infty \exp\left(-\frac{x}{2}\left(y + \frac{1}{y}\right)\right) y^{-(\nu+1)} dy$$

Before continuing, we set $z = \frac{1}{y}$, $y = \frac{1}{z}$ with $y \in (1, \infty)$, $z \in (0, 1)$ and $dy = -\frac{1}{z^2} dz$. The second integral is then modified. Notice here the sign of the integral will change due to the alteration of the domain of the variable. It follows that

$$\begin{aligned} B_\nu(x) &= \frac{1}{2} \int_1^\infty \exp\left(-\frac{x}{2}\left(y + \frac{1}{y}\right)\right) y^{\nu-1} dy - \frac{1}{2} \int_0^1 \exp\left(-\frac{x}{2}\left(\frac{1}{z} + z\right)\right) z^{\nu+1} \left(-\frac{1}{z^2}\right) dz \\ &= \frac{1}{2} \int_0^\infty \exp\left(-\frac{x}{2}\left(y + \frac{1}{y}\right)\right) y^{\nu-1} dy \end{aligned}$$

This has completed the proof. □

Appendix C: Grouped Data for Claim Severities

Groups (range of claims severities in GBP)	Number of claims
0–250	52
250–500	43
500–1,500	33
1,500–2,500	28
2,500–3,500	22
3,500–4,500	19
4,500–6,500	14
6,500–9,500	11
9,500–13,500	9
13,500–17,500	7
17,500–25,000	4
25,000–35,000	3
35,000–45,000	2
45,000–65,000	1
65,000–95,000	1
95,000–135,000	1