Fitting the Belgian Bonus-Malus System

S. Pitrebois,[1] M. Denuit[2] and J.F. Walhin[3]

Abstract. We show in this paper how to obtain the relativities of the Belgian Bonus-Malus System, including the special bonus rule sending the policyholders in the malus zone to initial level after four claim-free years. The model allows for *a priori* ratemaking. It is applied to a real-life portfolio.

Keywords: Bonus-Malus System, Markov Chain, Stationary distribution, *a priori* ratemaking.

1 Introduction

The Belgian Bonus-Malus System (BMS in short) is dying. As from 2004, Belgian companies will have complete freedom of using their own bonus-malus systems. The years 2002 and 2003 are transition periods. Companies are obliged to use the old scale but may set up the relativities they want. A company that wishes to abandon the bonus-malus system may do so by imposing the same relativities to each step of the scale.

Very few companies have used the possibilities offered by the new law. In this paper we will concentrate on the old Belgian BMS and show how to adapt the relativities to the experience of some automobile portfolio.

The Belgian BMS consists of a scale with 23 steps having relativities described in table 1. Business users enter the sytem in class 14 whereas commuters and pleasure users enter in class 11. Actually, we will not concentrate on this aspect because we will work within a framework where stationarity is assumed.

The transition rules are the following :

1. each year a one-class bonus is given.
2. each claim is penalized by five classes.
3. the maximal bonus is class 0.
4. the maximal malus is class 22.
5. a policyholder with four consecutive claim-free years may not be above class 14 (special bonus rule).

The latter point is usually ignored and we will concentrate on it in this paper. The motivation behind that special rule is that "bad" drivers who would suddenly improve should not be penalized on a too long period.

We will set up the relativities of the old Belgian BMS taking account of that rule and of a possible *a priori* ratemaking.

The paper is further organized as follows. Section 2 describes the *a priori* tariff that is used in the paper. Section 3 introduces the model for *a posteriori* ratemaking. Section 4 shows how to adapt Norberg's formulae in a setting where there are covariates and states that have to be constrained. Section 5 provides the numerical examples.

Class	Relativities	Class	Relativities
22	200%	10	81%
21	160%	9	77%
20	140%	8	73%
19	130%	7	69%
18	123%	6	66%
17	117%	5	63%
16	111%	4	60%
15	105%	3	57%
14	100%	2	54%
13	95%	1	54%
12	90%	0	54%
11	85%		

Table 1. Belgian BMS

2 The *a priori* tariff

We will not discuss in detail how the *a priori* tariff has been obtained. The interested reader will find details in Pitrebois et al. (2003a).

We work within a Poisson regression framework where it is assumed that the number of claims N_i, for a driver with characteristics $\boldsymbol{x}_i$ in a period of length d_i is Poisson distributed with mean

$$\lambda_i = d_i \exp(\boldsymbol{\beta}^t \boldsymbol{x}_i) \quad , \quad i = 1, 2, \ldots, n.$$

A statistical analysis of our reference portfolio provides the significant covariates. The point estimates of the regression coefficients are given in table 2.

Variable	Level	Coeff β_j
Intercept		-2.1975
Gender*Age	Female $18-30$ + Male $25-30$	0.2351
	Male $18-24$	0.6235
	Female > 30 + Male > 30	0
Kind of district	Rural	-0.1809
	Urban	0
Split of payment	Yes	0.4677
	No	0
Use of vehicle	Professionnal use	0.2150
	Leisure and commuting	0

Table 2. *A priori* segmentation

For example, the expected annual claims frequency for a male driver aged 35, living in suburbs, paying upfront premium and with professionnal use of the car is :

$$\exp(-2.1975 + 0 - 0.1809 + 0 + 0.2150) = 0.1149.$$

[1] Secura Belgian Re
[2] Institut des Sciences Actuarielles and Institut de Statistique, Université Catholique de Louvain
[3] Institut des Sciences Actuarielles, Université Catholique de Louvain and Secura Belgian Re, corresponding author, e-mail: jfw@secura-re.com

3 The *a posteriori* tariff

Even with the *a priori* segmentation proposed hereabove, there remains some heterogeneity within the defined classes. This is due to the unobservable variables, e.g. swiftness of reflexes, aggressiveness behind the wheel, consumption of drugs, ... This residual heterogeneity is taken into account with a random effect Θ_i. We have

$$\mathbb{P}[N_i = k|\Theta_i = \theta] = \exp(-\lambda_i\theta)\frac{(\lambda_i\theta)^k}{k!} \quad , \quad k = 0, 1, 2, \ldots$$

For the ease of mathematics, the Θ_i's are assumed to be iid and have common gamma density function (which is the natural conjugate of the Poisson distribution) :

$$u(\theta) = \frac{1}{\Gamma(a)}a^a\theta^{a-1}\exp(-a\theta) \quad , \quad \theta > 0.$$

It is possible to estimate the parameter a by maximum likelihood and we obtain, for our portfolio with the *a priori* tariff described in the previous section, $\hat{a} = 1.2401$.

For the sake of comparison, we also provide the estimate of a for the case where there is no *a priori* segmentation : $\hat{a} = 0.8888$. In this case, the average claims frequency is $\hat{\lambda} = 0.1474$.

4 Markov models

Let us denote by $s + 1$ the number of levels in our BMS. The index $j = 0, \ldots, s$ will denote a given level within the scale.

In most of the commercial BMS, the knowledge of the current level and of the number of claims during the current period suffice to determine the next level in the scale: these BMS are Markovian, which is a nice mathematical property ensuring easy calculations. Unfortunately the Belgian BMS is not Markovian due to the special bonus rule. Fortunately, it is possible to introduce fictitious classes in order to meet the memoryless property. Lemaire (1995) proposed to split the classes 16 to 21 into subclasses, depending on the number of consecutive years without accident. This authorizes to take account of the special bonus rule. A class $j.i$ is to be understood as level j and i consecutive years without accidents. Let n_j be the number of subclasses to be associated to bonus class j. The transition rules are completely defined in the table 3 and the different values for n_j are given in table 4. We take some liberty with the notations by using the value 0 for the subscript i and by not using a subscript when $n_j = 1$.

We now follow the notations of Pitrebois et al. (2003b). Let $\boldsymbol{\pi}(\nu)$ be the stationary distribution of the Markov chain for a policyholder with claims frequency ν. The stationary distribution exists because our Markov chain is ergodic. It is easily obtained by a matrix inversion as recalled in Pitrebois et al. (2003b). Let L be the BM level occupied by a randomly picked policyholder to be in class $j.i$ at stationarity. We have

$$\mathbb{P}[L = j.i] = \sum_k w_k \int_0^\infty \pi_{j.i}(\lambda_k\theta)u(\theta)d\theta, \quad \begin{cases} j = & 0, 1, \ldots, s \\ i = & 1, \ldots, n_j. \end{cases}$$

where w_k is the weight of the kth risk class whose annual expected claims frequency is λ_k. We have $w_k = \mathbb{P}[\Lambda = \lambda_k]$, where Λ represents the *a priori* claims frequency for a policyholder picked at random.

Class	Class after k accidents					
k	0	1	2	3	4	5
22	21.1	22	22	22	22	22
21.0	20.1	22	22	22	22	22
21.1	20.2	22	22	22	22	22
20.0	19.1	22	22	22	22	22
20.1	19.2	22	22	22	22	22
20.2	19.3	22	22	22	22	22
19.0	18.1	22	22	22	22	22
19.1	18.2	22	22	22	22	22
19.2	18.3	22	22	22	22	22
19.3	14	22	22	22	22	22
18.0	17	22	22	22	22	22
18.1	17.2	22	22	22	22	22
18.2	17.3	22	22	22	22	22
18.3	14	22	22	22	22	22
17	16	21.0	22	22	22	22
17.2	16.3	21.0	22	22	22	22
17.3	14	21.0	22	22	22	22
16	15	20.0	22	22	22	22
16.3	14	20.0	22	22	22	22
15	14	19.0	22	22	22	22
14	13	18.0	22	22	22	22
13	12	17	22	22	22	22
12	11	16	21.0	22	22	22
11	10	15	20.0	22	22	22
10	9	14	19.0	22	22	22
9	8	13	18.0	22	22	22
8	7	12	17	22	22	22
7	6	11	16	21.0	22	22
6	5	10	15	20.0	22	22
5	4	9	14	19.0	22	22
4	3	8	13	18.0	22	22
3	2	7	12	17	22	22
2	1	6	11	16	21.0	22
1	0	5	10	15	20.0	22
0	0	4	9	14	19.0	22

Table 3. Transition rules of the Belgian BMS

j	n_j	j	n_j
22	1	10	1
21	2	9	1
20	3	8	1
19	4	7	1
18	4	6	1
17	3	5	1
16	2	4	1
15	1	3	1
14	1	2	1
13	1	1	1
12	1	0	1
11	1		

Table 4. Subclasses of the Belgian BMS

By definition we have

$$\begin{aligned}\mathbb{P}[L = j] &= \sum_{i=1}^{n_j}\mathbb{P}[L = j.i] \\ &= \sum_k w_k \int_0^\infty \pi_j(\lambda_k\theta)u(\theta)d\theta,\end{aligned}$$

where

$$\pi_j = \sum_{i=1}^{n_j} \pi_{j.i} \quad , \quad j = 0, \ldots, s.$$

According to Norberg's criterion, the relativities, $r_{j.i}$ will be obtained by minimizing the squared difference between the true relative premium Θ and the relative premium r_L applicable to the policyholder when stationary state has been reached.

The current situation is more complicated because some states have to be constrained to have the same relativity. Indeed the artificial states $j.i$ have the property that

$$r_j = r_{j.1} = \cdots = r_{j.n_j} \quad , \quad j = 0, \ldots, s.$$

This point has been addressed by Centeno et al. (2002) in a framework without *a priori* segmentation. We extend it here in the general framework.

The Norberg's criterion becomes a minimization under constraints :

$$\min \mathbb{E}\left[(\Theta - r_L)^2\right]$$

such that

$$r_j = r_{j.1} = r_{j.2} = r_{j.n_j} \quad , \quad j = 0, 1, \ldots, s.$$

We want to minimize

$$\begin{aligned}
&\mathbb{E}\left[(\Theta - r_L)^2\right] = \\
&= \sum_{j=0}^{s} \sum_{i=1}^{n_j} \mathbb{E}\left[(\Theta - r_{j.i})^2 | L = j.i\right] \mathbb{P}[L = j.i] \\
&= \sum_{j=0}^{s} \sum_{i=1}^{n_j} \int_0^\infty (\theta - r_{j.i})^2 \mathbb{P}[L = j.i | \Theta = \theta] u(\theta) d\theta \\
&= \sum_k w_k \int_0^\infty \sum_{j=0}^{s} \sum_{i=1}^{n_j} (\theta - r_{j.i})^2 \pi_{j.i}(\lambda_k \theta) u(\theta) d\theta.
\end{aligned}$$

s.t.

$$r_j = r_{j.1} = r_{j.2} = r_{j.n_j} \quad , \quad j = 0, 1, \ldots, s.$$

Using the constraints we can rewrite the objective function as

$$\begin{aligned}
&\mathbb{E}\left[(\Theta - r_L)^2\right] = \\
&= \sum_k w_k \int_0^\infty \sum_{j=0}^{s} \sum_{i=1}^{n_j} (\theta - r_j)^2 \pi_{j.i}(\lambda_k \theta) u(\theta) d\theta \\
&= \sum_k w_k \int_0^\infty \sum_{j=0}^{s} (\theta - r_j)^2 \pi_j(\lambda_k \theta) u(\theta) d\theta.
\end{aligned}$$

Setting the derivatives wrt r_j equal to 0 :

$$\frac{\partial \mathbb{E}\left[(\Theta - r_L)^2\right]}{\partial r_j} = 0 \quad \forall j = 0, \ldots, s,$$

we obtain

$$r_j = \frac{\sum_k w_k \int_0^\infty \theta \pi_j(\lambda_k \theta) u(\theta) d\theta}{\sum_k w_k \int_0^\infty \pi_j(\lambda_k \theta) u(\theta) d\theta}.$$

This formula extends Norberg (1976), Centeno et al. (2002) and Pitrebois et al. (2003b). Indeed

1. Setting $n_j = 1 \quad \forall j$ gives the formula of Pitrebois et al. (2003b).
2. Setting $\lambda_k = \lambda \quad \forall k$ gives the formula of Centeno et al. (2002).
3. Combining the previous two particular cases gives the formula of Norberg (1976).

Note that it is easily seen that

$$r_j = \sum_{i=1}^{n_j} \frac{\mathbb{P}[L = j.i]}{\sum_{i=1}^{n_j} \mathbb{P}[L = j.i]} r_{j.i},$$

where the $r_{j.i}$'s represent the non constrained solution of $\min \mathbb{E}\left[(\Theta - r_L)^2\right]$.

We also immediately verify that

$$\sum_{j=0}^{s} r_j \mathbb{P}[L = j] = 1$$

which ensures that the BMS is financially balanced at steady state.

5 Numerical results

As expected, we observe in Table 5 that the relativities are higher with the special bonus rule than without that rule (with the exception of level 18). This is logical because the scale with the special bonus rule is less severe than the scale without the special bonus rule.

	Relativities	
Level j	Without Special Bonus Rule	With Special Bonus Rule
22	256.0%	267.5%
21	235.9%	246.1%
20	220.0%	228.8%
19	206.8%	214.4%
18	195.5%	193.6%
17	185.4%	188.8%
16	176.3%	182.1%
15	168.1%	175.0%
14	160.4%	186.7%
13	152.7%	175.2%
12	145.4%	164.7%
11	138.8%	155.4%
10	132.9%	147.2%
9	124.9%	137.0%
8	117.0%	127.2%
7	111.4%	120.0%
6	106.9%	114.3%
5	103.1%	109.6%
4	84.0%	88.2%
3	81.6%	85.3%
2	79.3%	82.7%
1	77.1%	80.1%
0	48.8%	49.9%

Table 5. Relativities $r_j = \mathbb{E}[\Theta | L = j]$ with *a priori* ratemaking

Table 6 displays the proportion at stationarity of policyholders in each level. We can see that if we use the special bonus rule, we have less policyholders in levels above 14 and more policyholders in levels below 14 than without the special rule. This is logical from the definition of this rule.

Table 7 shows us that the average *a priori* expected claims frequency in level j is always higher with the special bonus rule than without that rule. The effect is more pronounced in the highest levels of the scale and less pronounced in the lowest levels of the scale. This fact is obvious from the definition of the special bonus rule. The policyholders attaining the highest classes of the scale benefit from the rule. Those staying in these highest classes show therefore a higher expected frequency. Even below level 14 the effect remains true because the policyholders have benefitted of it before attaining the lowest levels. Obvioulsy the effect is less and less pronounced at the bottom of the scale.

Some insurance companies use the bonus-malus scale as an acceptation tool. They e.g. systematically refuse drivers with a BM level > 14. Our calculations show that this is not optimal because drivers at level 15 are less dangerous than drivers at level 14.

We observe that without the special bonus rule, the relativities are always increasing from level 0 to level 22. Although there is no certainty about that fact, it is logical and necessary from a commercial point of view. The same pattern is observed for $\mathbb{E}[\Lambda|L = j]$.

When looking at the results for the BMS with special bonus rule, we observe that the relativities at level 13, 14 and 15 are not ordered anymore. This fact may be explained as follows : there are many drivers at level 14 that have benefitted of the special bonus rule, i.e. they have made lot's of claims but are sent back to level 14, which is not very much representative of their claims frequency. Level 13 which is attained from level 14 after a claim free year is also polluted by this fact.

It is clear that such a situation is not acceptable from a commercial point of view. We may constrain the scale to be linear in the spirit of Gilde and Sundt (1989). However we propose a local adjustement to the scale in order to keep $\mathbb{E}[(\Theta - r_L)^2]$ as small as possible.

Let us constrain the scale to be linear between levels 13 and 16. We are looking for updated value for $r_j^{'}$, $j = 13, \ldots 16$. They are such that $r_j^{'} = r_{j-1}^{'} + a$, $j = 14, 15, 16$ where $a = \frac{r_{16}^{'} - r_{13}^{'}}{3}$. We also want to keep the financial equilibrium of the system. Therefore we constrain a local equilibrium :

$$\sum_{j=13}^{16} r_j \pi_j = \sum_{j=13}^{16} r_j^{'} \pi_j.$$

Choosing $r_{13}^{'} = 177.00\%$, we obtain $r_{16}^{'} = 185.69\%$.

Now let us compare the value of the expected error $Q = \mathbb{E}\left[(\Theta - r_L)^2\right]$ with the original model, Q_1 and with the constrained model, Q_2 :

$$Q_1 = 0.36857, \ Q_2 = 0.36876.$$

This shows that the error induced by the commercial constraint is really small. So we may adapt the scale without resorting to a full linear scale constraint.

$\Pr[L = j]$		
Level j	Without Special Bonus Rule	With Special Bonus Rule
22	5.0%	4.0%
21	3.7%	2.8%
20	2.8%	2.1%
19	2.3%	1.7%
18	1.9%	0.9%
17	1.7%	1.0%
16	1.5%	1.0%
15	1.4%	1.0%
14	1.3%	2.2%
13	1.3%	2.0%
12	1.3%	1.8%
11	1.3%	1.8%
10	1.3%	1.7%
9	1.5%	1.8%
8	1.7%	2.0%
7	1.8%	2.1%
6	1.9%	2.1%
5	1.9%	2.1%
4	4.5%	4.7%
3	4.0%	4.2%
2	3.6%	3.8%
1	3.3%	3.4%
0	49.1%	49.7%

Table 6. Distribution of L with *a priori* ratemaking

$\mathbb{E}[\Lambda\|L = j]$		
Level j	Without Special Bonus Rule	With Special Bonus Rule
22	18.3%	18.7%
21	17.6%	17.9%
20	17.1%	17.4%
19	16.7%	17.0%
18	16.4%	16.4%
17	16.2%	16.3%
16	15.9%	16.1%
15	15.7%	15.9%
14	15.6%	16.2%
13	15.4%	15.9%
12	15.3%	15.7%
11	15.1%	15.5%
10	15.0%	15.3%
9	14.9%	15.1%
8	14.7%	14.9%
7	14.6%	14.8%
6	14.5%	14.7%
5	14.5%	14.6%
4	14.2%	14.2%
3	14.1%	14.2%
2	14.1%	14.1%
1	14.1%	14.1%
0	13.7%	13.7%

Table 7. Average *a priori* claims frequency in level ℓ, $\mathbb{E}[\Lambda|L = j]$ with *a priori* ratemaking

We can perform numerically the local minimization without imposing a linear scale between levels 13 and 16. We use the following constraints :

$$
\begin{aligned}
r'_{13} &<= r'_{14}, \\
r'_{14} &<= r'_{15}, \\
r'_{15} &<= r'_{16}, \\
\sum_{j=13}^{16} r_j \pi_j &= \sum_{j=13}^{16} r'_j \pi_j, \\
r'_j &>= 165\% \quad j = 13, ..., 16, \\
r'_j &<= 188\% \quad j = 13, ..., 16.
\end{aligned}
$$

And we obtain $r'_{13} = 175.2\%$ and $r'_{14} = r'_{15} = r'_{16} = 182.7\%$. The value of Q is now 0.36867.

The next two tables show the proportion of the policyholders in each level (Table 8) and the relativities (Table 9) without *a priori* segmentation. We observe the same kind of behaviour along the levels $13, 14, 15$. We also observe that the scale without *a priori* segmentation is more elastic, which is logical because it has to take into account the full heterogeneity.

	$\Pr[L = j]$	
Level j	Without Special Bonus Rule	With Special Bonus Rule
22	5.6%	4.5%
21	4.0%	3.1%
20	3.0%	2.3%
19	2.4%	1.8%
18	2.0%	0.9%
17	1.7%	1.0%
16	1.5%	1.1%
15	1.4%	1.0%
14	1.3%	2.2%
13	1.2%	2.0%
12	1.2%	1.8%
11	1.2%	1.7%
10	1.2%	1.6%
9	1.3%	1.7%
8	1.5%	1.8%
7	1.6%	1.9%
6	1.7%	1.9%
5	1.7%	1.9%
4	4.0%	4.3%
3	3.6%	3.8%
2	3.3%	3.4%
1	2.9%	3.1%
0	50.5%	51.0%

Table 8. Distribution of L without *a priori* ratemaking

	Relativities	
Level j	Without Special Bonus Rule	With Special Bonus Rule
22	306.0%	324.2%
21	273.9%	290.2%
20	248.7%	262.6%
19	228.2%	240.0%
18	211.2%	209.0%
17	196.5%	202.2%
16	183.8%	192.8%
15	172.6%	183.0%
14	162.5%	201.0%
13	152.8%	184.9%
12	143.9%	170.7%
11	135.9%	158.4%
10	128.9%	147.9%
9	119.8%	135.5%
8	111.1%	124.0%
7	104.8%	115.6%
6	99.8%	109.0%
5	95.6%	103.5%
4	75.2%	80.1%
3	72.7%	77.0%
2	70.3%	74.1%
1	68.0%	71.4%
0	37.5%	38.6%

Table 9. Relativities $r_j = \mathbb{E}[\Theta | L = j]$ without *a priori* ratemaking

6 Conclusion

In this paper, we show how special bonus-malus rules can be taken into account when computing the relativities associated to each level of the scale. Specifically, fictitious classes are added to the physical scale in the spirit of Lemaire (1995) to cope with these rules maintaining the system Markovian. Our main contribution is to extend the results recently derived by Centeno and Silva (2002) to the situation where the company enforces *a priori* risk classification. The results contained in the present paper will certainly be useful to Belgian insurance companies during the deregulation phase 2002-2004.

REFERENCES

[1] Centeno, M. and Silva, J.M.A. (2002) Optimal Bonus Scales Under Path-Dependent Bonus Rules, Scandinavian Actuarial Journal, 24, 615-627.

[2] Gilde, V. and Sundt, B. (1989) On Bonus Systems with Credibility Scales, Scandinavian Actuarial Journal, 13-22.

[3] Lemaire, J. (1995) Bonus-Malus Systems in Automobile Insurance, Kluwer Academic Publishers, Boston/Dordrecht/London.

[4] Norberg, R. (1976) A Credibility Theory for Automobile Bonus Systems, Scandinavian Actuarial Journal, 2, 92-107.

[5] Pitrebois, S., Denuit, M. and Walhin, J.F. (2003a) Marketing et Systèmes Bonus-Malus, ACTU-L, 3, in press.

[6] Pitrebois, S., Denuit, M. and Walhin, J.F. (2003b) Setting A Bonus-Malus Scale In The Presence of Other Rating Factors: Taylor's Work Revisited, submitted.