How to Transfer Policyholders from one Bonus-malus Scale to the Other ?

S. Pitrebois [•], J.F. Walhin[*•] and M. Denuit[*†]

[*] *Institut des Sciences Actuarielles, Université Catholique de Louvain*

[•] *Secura*

[†] *Institut de Statistique, Université Catholique de Louvain*

Corresponding author :

Jean-François Walhin
Secura Belgian Re
Avenue des Nerviens, 9-31
B-1040 Bruxelles
Belgique
phone : + 32 2 504 82 22
fax : + 32 2 504 82 00
e-mail : jfw@secura-re.com

How to Transfer Policyholders from one Bonus-malus Scale to the Other ?

Abstract

With the deregulation of bonus-malus systems in the EU, it is important to obtain rules in order to transfer policyholders from one bonus-malus scale to another. The present paper proposes a solution to this problem.

Keywords

Bonus-malus scale, a posteriori random effect, Kolmogorov distance, total variation distance.

1 Introduction

Since the 90's insurance markets in the EU have been deregulated. More competition is allowed resulting a.o. with the use of segmented tariffs for motor third party liability (MTPL) covers. Until recently, most member states[1] of the EU did not allow insurance companies to use their own bonus-malus scales. Actually a posteriori rating through bonus-malus scales was still regulated. No competition was allowed on that tariff aspect.

Pitrebois et al. (2003) have argued that it is a nonsense to oblige insurance companies to use the same bonus-malus scale when they use different a priori segments. It is indeed clear that an insurer having no a priori segmentation has a portfolio which is more heterogeneous than an insurer using a very detailed segmentation. Both insurers should therefore use different bonus-malus scales.

Furthermore we may argue that an insurer having different a priori tariff classes should use specific bonus-malus scales for their a priori classes. Insurers should be allowed to compete on the basis of their bonus-malus scales as well. With the notable exception of France, insurers within the EU are now free to apply the bonus-malus scales of their choice. The Belgian experience shows they do. However it seems that they do so more for marketing reasons than for the technical reasons described hereabove.

Two related problems arise with the deregulation of the bonus-malus systems in the EU. The first one consists in transfering the policyholders in the new scales. The second one is more difficult : it consists in transfering a new policyholder in the scale of the company knowing his level in the scale of his previous insurer.

The aim of the present paper is to show how to develop rules allowing to transfer a policyholder in a bonus-malus scale knowing his level in his previous bonus-malus scale. The rest of the paper is organised as follows. Section 2 recalls the necessary background about bonus-malus systems and suggests a solution to our problem. Numerical illustrations are given in section 3. Section 4 concludes.

---

[1] This is however not the case for Germany where different scales are in use for several years.

2 Bonus-malus Systems

Let us assume a bonus-malus scale with $s+1$ levels numbered from 0 to s. To each level ℓ corresponds a premium level r_ℓ. When the driver causes a claim, he suffers an increase of his position in the scale, resulting in an increase of his premium: this is the malus. When the driver is claim free, he gets a bonus by going down in the scale, resulting in a lower premium (unless the lowest levels have the same premium level). Each bonus-malus system has well-defined transition rules showing the behaviour of the policyholders within the scale. The driver never goes below level $\ell = 0$ or above level $\ell = s$.

In practice, these bonus-malus scales may be represented by a Markov chain. Indeed, the knowledge of the current level and of the number of claims of the current year suffices to determine the next level in the scale. Moreover, the Markov chains are generally[2] irreducible, meaning that all states are always accessible in a finite number of steps from all other states (having a bonus-malus scale whose associated Markov chain is not irreducible would be difficult to justify to the policyholders). Moreover, bonus-malus scales have a bonus level with maximal reward : policyholders being in that level will remain in that level after a claim free year. Both these properties insure that the Markov chain associated to the bonus-malus scale has a limiting distribution, which is also the stationary distribution.

Let $\boldsymbol{M}(\vartheta)$ be the transition matrix of the Markov chain associated to the bonus-malus scale for a policyholder with mean frequency ϑ. Then the limiting distribution $\boldsymbol{\pi}(\vartheta) = (\pi_0(\vartheta), \ldots, \pi_s(\vartheta))^t$ is easily given by

$$\begin{aligned}\boldsymbol{\pi}(\vartheta) &= \lim_{\nu \to \infty} \boldsymbol{M}^\nu \boldsymbol{b} \\ &= \boldsymbol{e}^t (\boldsymbol{I} - \boldsymbol{M}(\vartheta) + \boldsymbol{E})^{-1}\end{aligned}$$

where $\boldsymbol{b}$ is any initial distribution of the policyholders within the bonus-malus scale, $\boldsymbol{e}$ is a column vector of 1's, $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{E}$ is a $(s+1) \times (s+1)$ matrix consisting of $s+1$ column vectors $\boldsymbol{e}$. See Rolski et al. (1999) for a derivation.

As from now, we will assume that our portfolio has reached stationarity following Norberg (1976). Extending the results of this paper to transient states is immediate in the spirit of Borgan, Hoem & Norberg (1981).

We will use the random variable L_ϑ, valued in $\{0, \ldots, s\}$ that conforms to the distribution $\boldsymbol{\pi}(\vartheta)$ i.e.

$$\mathbb{P}[L_\vartheta = \ell] = \pi_\ell(\vartheta) \quad , \quad \ell = 0, \ldots, s.$$

The variable L_ϑ represents the level occupied by a policyholder with annual expected frequency ϑ once the steady state has been reached.

Let us assume that m a priori tariff classes have been determined by the actuary. For a randomly picked policyholder, the probability of having a priori frequency λ_i is ω_i. Denoting as Λ the (random) a priori expected claim frequency for this randomly picked up policyholder, we thus have

$$\mathbb{P}[\Lambda = \lambda_i] = \omega_i$$

with $\sum_{i=1}^m \omega_i = 1$.

---

[2] This is not the case for the entrance class 0 for many bonus-malus systems in Germany.

Because it is not possible to observe all significant covariates (some features are hidden, e.g. speed of reflexes, aggressiveness in traffic, drug abuse,...), there remains some heterogeneity in the subportfolios of policyholders having a priori expected frequency λ_i which we model with a random effect Θ_i.

Let us assume that the number of claims, N_i caused by a policyholder a priori classified in segment class i in a period of length 1 is given by a Poisson distribution with parameter $\lambda_i\theta_i$. The portfolio being heterogeneous, we assume that the frequency may vary along the policyholders according to our multiplicative random effect Θ_i with $\mathbb{E}\Theta_i = 1$. So we have

$$\mathbb{P}[N_i = k|\Theta_i = \theta_i] = \exp(-\lambda_i\theta_i)\frac{(\lambda_i\theta_i)^k}{k!}\,,\; k \in \mathbb{N}, i = 1, \dots, m.$$

For mathematical convenience, we assume that the Θ_i are independent and all following the same Gamma distribution with probability density function

$$f_\Theta(\theta) = \frac{1}{\Gamma(a)}a^a\theta^{a-1}\exp(-a\theta)\,,\; \theta \in \mathbb{R}^+.$$

From this we conclude that the number of claims of a randomly picked up policyholder from the portfolio is Negative Binomial distributed :

$$\mathbb{P}[N_i = k] = \left(\begin{array}{c} a+k-1 \\ k \end{array}\right)\left(\frac{\lambda_i}{a+\lambda_i}\right)^k\left(\frac{a}{a+\lambda_i}\right)^a.$$

Within this framework, the probability of being at level ℓ at stationary state is given by

$$\mathbb{P}[L = \ell] = \sum_{i=1}^m \omega_i \int_0^\infty \pi_\ell(\lambda_i\theta)f_\theta(\theta)d\theta.$$

The a posteriori distribution of Θ given the level of the bonus-malus scale is given by

$$\begin{aligned} f_{\Theta|L=\ell}(\theta) &= \frac{\mathbb{P}[L = \ell|\Theta = \theta]f_\theta(\theta)}{\mathbb{P}[L = \ell]} \\ &= \frac{\sum_{i=1}^m \omega_i\pi_\ell(\lambda_i\theta)f_\Theta(\theta)}{\sum_{i=1}^m \omega_i\int_0^\infty \pi_\ell(\lambda_i\theta)f_\Theta(\theta)d\theta}. \end{aligned}$$

Because we want to move a policyholder from one scale to the other, we should try to put the policyholder at a level which is as close as possible from his level in his original bonus-malus scale. Close means here having the a posteriori random effect as close as possible.

A first measure may be to compare the expected a posteriori random effect, which actually is the relative premium at level ℓ in the bonus-malus scale :

$$r_{L=\ell} = \mathbb{E}[\Theta|L = \ell],$$

see Pitrebois et al. (2003) for more details.

So the closest level ℓ_j in scale 2 from the level ℓ_i in scale 1 is given by

$$\mathrm{argmin}_{\ell_j}(r_{L_1=\ell_i} - r_{L_2=\ell_j})^2.$$

This rule simply amounts to place the policyholder in the new scale at the level with the closest relativity to the one applicable in the previous scale. Because most commercial scales are normalized (to associate a unit relativity to the entry level), this means that the insurer has to compute the relativities for both scales. The implicit assumption is that the new entrant has the same characteristics than the policyholders in the portfolio (no adverse selection being allowed).

Another measure of discrepancy consists in comparing the distribution functions of the a posteriori random effects. This can be done by using the Kolmogorov distance or the total variation distance :

$$\begin{aligned} d_K(\Theta|L_1=\ell_i, \Theta|L=\ell_j) &= \max_{\theta} |\mathbb{P}[\Theta \le \theta|L_1=\ell_i] - \mathbb{P}[\Theta \le \theta|L_2=\ell_j]|, \\ d_{TV}(\Theta|L_1=\ell_i, \Theta|L=\ell_j) &= \int_0^{\infty} |f_{\Theta|L_1=\ell_i}(\theta) - f_{\Theta|L_2=\ell_j}(\theta)| d\theta. \end{aligned}$$

Summarizing, a policyholder being at level ℓ_i in scale 1 and moving to scale 2 will be put in the level ℓ_j of scale 2 that minimizes one of the following distances :

1. $(r_{L_1=\ell_i} - r_{L_2=\ell_j})^2$
2. $d_K(\Theta|L_1=\ell_i, \Theta|L=\ell_j)$
3. $d_{TV}(\Theta|L_1=\ell_i, \Theta|L=\ell_j)$

3 Numerical Application

In this paper, we use the same segmentation as in Pitrebois et al (2004). We apply Poisson regression modelling : we assume that the number of claims N_i for a driver having characteristics x_i in a period of length d_i is Poisson distributed with mean

$$\lambda_i = d_i \exp(\boldsymbol{\beta}^t \boldsymbol{x}_i).$$

The statistical analysis of the reference portfolio provides the following significant covariates and point estimates $\hat{\boldsymbol{\beta}}$ by maximum likelihood :

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 1: *A priori* segmentation

Gender-Age			Use of the car		Premium splitting		District		Annual	Weights
Fe 18-30 Ma 25-3	Ma 18-24	Rest	Private	Professional	Annual Claim	Splitted	Rural	Urban	Freq. (%) λ_i	(%) ω_i
1	0	0	1	0	1	0	1	0	11,73	10,49
1	0	0	1	0	1	0	0	1	14,05	13,96
1	0	0	1	0	0	1	1	0	18,72	3,98
1	0	0	1	0	0	1	0	1	22,43	7,05
1	0	0	0	1	1	0	1	0	14,54	0,76
1	0	0	0	1	1	0	0	1	17,42	1,22
1	0	0	0	1	0	1	1	0	23,21	0,13
1	0	0	0	1	0	1	0	1	27,81	0,14
0	1	0	1	0	1	0	1	0	17,29	2,93
0	1	0	1	0	1	0	0	1	20,72	2,99
0	1	0	1	0	0	1	1	0	27,60	1,52
0	1	0	1	0	0	1	0	1	33,08	2,42
0	1	0	0	1	1	0	1	0	21,44	0,07
0	1	0	0	1	1	0	0	1	25,69	0,09
0	1	0	0	1	0	1	1	0	34,22	0,02
0	0	1	1	0	1	0	1	0	9,27	13,38
0	0	1	1	0	1	0	0	1	11,11	19,73
0	0	1	1	0	0	1	1	0	14,80	2,94
0	0	1	1	0	0	1	0	1	17,73	6,61
0	0	1	0	1	1	0	1	0	11,49	3,72
0	0	1	0	1	1	0	0	1	13,77	5,17
0	0	1	0	1	0	1	1	0	18,35	0,25
0	0	1	0	1	0	1	0	1	21,98	0,44

Table 2: *A priori* risk classification (Poisson regression).

Because of the residual heterogeneity, a Gamma random effect has been added into the model. We then work within a Poisson mixture regression model, actually a Negative Binomial regression model. As in Pitrebois et al. (2004), the a parameter of the Gamma residual effect has been estimated at $\hat{a} = 1.2401$ according to a moment method. The estimator $\hat{\boldsymbol{\beta}}$ obtained by maximum likelihood being consistent in the Poisson mixture regression model, we adhere to its estimate in the Negative Binomial regression model.

Now let us assume that we want to move policyholders between the following two bonus-malus scales :

- The -1/Top scale with 6 levels, having transition rules

Starting level	Level occupied if 0	≥ 1
	claim is reported	
5	4	5
4	3	5
3	2	5
2	1	5
1	0	5
0	0	5

Table 3: -1/Top scale with 6 levels

- The -1/+2 scale of Taylor (1997) with 9 levels, having transition rules

Starting	Level occupied if				
level	0	1	2	3	≥ 4
	claim is/are reported				
8	7	8	8	8	8
7	6	8	8	8	8
6	5	8	8	8	8
5	4	7	8	8	8
4	3	6	8	8	8
3	2	5	7	8	8
2	1	4	6	8	8
1	0	3	5	7	8
0	0	2	4	6	8

Table 4: -1/+2 scale of Taylor

The following tables provide the distances between the conditional random effect for our three metrics :

$\ell_1 \backslash \ell_2$	0	1	2	3	4	5	6	7	8
0	0.04	0.35	0.38	0.55	0.59	0.68	0.72	0.77	0.81
1	0.30	0.01	0.04	0.23	0.29	0.41	0.47	0.55	0.61
2	0.33	0.05	0.03	0.19	0.25	0.37	0.43	0.51	0.57
3	0.37	0.10	0.07	0.15	0.20	0.32	0.38	0.46	0.53
4	0.41	0.15	0.12	0.10	0.15	0.27	0.33	0.41	0.48
5	0.46	0.21	0.18	0.06	0.10	0.21	0.27	0.35	0.42

Table 5: Kolmogorov distance between $\Theta | L_1 = \ell_1$ and $\Theta | L_2 = \ell_2$ (d_K)

$\ell_1 \backslash \ell_2$	0	1	2	3	4	5	6	7	8
0	0.08	0.71	0.76	1.10	1.18	1.36	1.44	1.54	1.62
1	0.60	0.07	0.10	0.47	0.58	0.82	0.95	1.11	1.23
2	0.67	0.11	0.08	0.39	0.50	0.74	0.86	1.03	1.15
3	0.74	0.20	0.14	0.31	0.40	0.64	0.77	0.93	1.07
4	0.82	0.30	0.24	0.26	0.32	0.54	0.66	0.83	0.96
5	0.92	0.43	0.37	0.25	0.27	0.43	0.54	0.70	0.84

Table 6: Total variation distance between $\Theta | L_1 = \ell_1$ and $\Theta | L_2 = \ell_2$ (d_{TV})

$\ell_1 \backslash \ell_2$	0	1	2	3	4	5	6	7	8
0	0.038	0.304	0.352	0.794	0.964	1.493	1.844	2.437	3.003
1	0.120	0.000	0.003	0.122	0.194	0.462	0.666	1.038	1.418
2	0.192	0.007	0.002	0.066	0.121	0.346	0.524	0.859	1.208
3	0.300	0.037	0.023	0.022	0.057	0.228	0.377	0.667	0.978
4	0.468	0.108	0.082	0.000	0.010	0.117	0.228	0.464	0.728
5	0.740	0.255	0.214	0.027	0.005	0.028	0.091	0.255	0.458

Table 7: $(\mathbb{E}[\Theta|L_1 = \ell_1] - \mathbb{E}[\Theta|L_2 = \ell_2])^2$ ($\mathbb{E}$)

The minima are attained for the following transition rules :

	d_K	d_{TV}	$\mathbb{E}$
ℓ_1	ℓ_2	ℓ_2	ℓ_2
0	0	0	0
1	1	1	1
2	2	2	2
3	2	2	3
4	3	2	3
5	3	3	4

Table 8: Transition rules from scale 1 to scale 2

	d_K	d_{TV}	$\mathbb{E}$
ℓ_2	ℓ_1	ℓ_1	ℓ_1
0	0	0	0
1	1	1	1
2	2	2	2
3	5	5	3
4	5	5	5
5	5	5	5
6	5	5	5
7	5	5	5
8	5	5	5

Table 9: Transition rules from scale 2 to scale 1

In the following tables, we give the transition rules for the case without segmentation :

	d_K	d_{TV}	$\mathbb{E}$
ℓ_1	ℓ_2	ℓ_2	ℓ_2
0	0	0	0
1	1	1	1
2	2	2	2
3	2	2	2
4	3	2	3
5	3	3	4

Table 10: Transition rules from scale 1 to scale 2 without a priori segmentation

	d_K	d_{TV}	$\mathbb{E}$
ℓ_2	ℓ_1	ℓ_1	ℓ_1
0	0	0	0
1	1	1	1
2	1	2	2
3	5	4	4
4	5	5	5
5	5	5	5
6	5	5	5
7	5	5	5
8	5	5	5

Table 11: Transition rules from scale 2 to scale 1 without a priori segmentation

In the following tables, we give the transition rules for a bad policyholder, e.g. a male driver aged more than 30 years, with premium splitted, having private use of the car and living in an urban environment. His a priori expected frequency is 0.1773.

We here compare the distance between $\Theta|L_1 = l_i, \Lambda = 0.1773$ and $\Theta|L_2 = l_j, \Lambda = 0.1773$. We have

$$f_{\Theta|L=l,\Lambda=\lambda}(\theta) = \frac{\pi_l(\lambda\theta) f_\Theta(\theta)}{\int_0^\infty \pi_l(\lambda\theta) f_\Theta(\theta) d\theta}.$$

	d_K	d_{TV}	$\mathbb{E}$
ℓ_1	ℓ_2	ℓ_2	ℓ_2
0	0	0	0
1	2	1	2
2	2	2	2
3	2	2	3
4	3	3	3
5	3	4	4

Table 12: Transition rules from scale 1 to scale 2 for a policyholder with expected frequency 0.1773

	d_K	d_{TV}	$\mathbb{E}$
ℓ_2	ℓ_1	ℓ_1	ℓ_1
0	0	0	0
1	1	1	1
2	1	1	1
3	5	4	4
4	5	5	5
5	5	5	5
6	5	5	5
7	5	5	5
8	5	5	5

Table 13: Transition rules from scale 2 to scale 1 for a policyholder with expected frequency 0.1773

Let us now take into account the a priori characteristics of the driver. It will be seen that good and bad drivers are not placed in the same way when they are transferred from one scale to the other.

In the following tables, we give the transition rules for a good policyholder, e.g. a male driver aged more than 30 years, with upfront premium, having private use of the car and living in an rural environment. His a priori expected frequency is 0.0927.

We here compare the distance between $\Theta|L_1 = l_i, \Lambda = 0.1773$ and $\Theta|L_2 = l_j, \Lambda = 0.0927$. We have

$$f_{\Theta|L=l,\Lambda=\lambda}(\theta) = \frac{\pi_l(\lambda\theta) f_\Theta(\theta)}{\int_0^\infty \pi_l(\lambda\theta) f_\Theta(\theta) d\theta}.$$

	d_K	d_{TV}	$\mathbb{E}$
ℓ_1	ℓ_2	ℓ_2	ℓ_2
0	0	0	0
1	1	1	1
2	1	1	2
3	2	2	2
4	2	2	2
5	2	2	3

Table 14: Transition rules from scale 1 to scale 2 for a policyholder with expected frequency 0.0927

	d_K	d_{TV}	$\mathbb{E}$
ℓ_2	ℓ_1	ℓ_1	ℓ_1
0	0	0	0
1	2	2	1
2	3	3	2
3	5	5	5
4	5	5	5
5	5	5	5
6	5	5	5
7	5	5	5
8	5	5	5

Table 15: Transition rules from scale 2 to scale 1 for a policyholder with expected frequency 0.0927

4 Conclusion

We observe that the different metrics we have chosen do not provide extremely different results. Because the expected posterior random effect has a financial meaning (i.e. it is the multiplier of the average cost to get the a posteriori premium), we may be tempted to choose its corresponding metric to transfer a policyholder from a scale to the other. Note also that the a priori characteristics of the driver influence the way the policy is transferred from one scale to the other. This results in a number of rules according to the risk classification scheme applied by the insurance company.

Transferring a policyholder from the bonus-malus scale of a given insurer to the bonus-malus scale of another insurer remains a more complicated task. Indeed, the actuary needs to know the a posteriori random effect in both situations. However the a priori random effect may be different due to another type of a priori tariff or due to antiselection. The distance to minimize then extends to $d(\Theta_1|L_1 = \ell_1, \Theta_2|L_2 = \ell_2)$.

Acknowledgments

We would like to thank the referee for his comments, in particular regarding the particularities of the German market.

References

1. Borgan, O., Hoem, J.M. and Norberg, R. (1981) A nonasymptotic criterion for the evaluation of automobile bonus systems. Scandinavian Actuarial Journal, 165-178.

2. Norberg, R. (1976) A credibility theory for automobile bonus systems. Scandinavian Actuarial Journal, 92-107.

3. Pitrebois, S., Denuit, M. and Walhin, J.F. (2003) Setting A Bonus-Malus Scale In The Presence of Other Rating Factors: Taylor's Work Revisited. ASTIN Bulletin 33, 419-436.

4. Pitrebois, S., Denuit, M. and Walhin, J.F. (2004) Bonus-malus Scales in Segmented Tariffs : Gilde & Sundt's Work Revisited. Australian Actuarial Journal, 10, 107-125.

5. Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J. (1999) Stochastic Processes for Insurance and Finance. John Wiley & Sons, New York.