# SETTING A BONUS-MALUS SCALE IN THE PRESENCE OF OTHER RATING FACTORS: TAYLOR'S WORK REVISITED

BY

SANDRA PITREBOIS, MICHEL DENUIT AND JEAN-FRANÇOIS WALHIN

ABSTRACT

In this paper, we propose an analytic analogue to the simulation procedure described in Taylor (1997). We apply the formulas to a Belgian data set and discuss the interaction between a priori and a posteriori ratemakings.

KEYWORDS AND PHRASES

Bonus-Malus system, Markov chains, *a priori* ratemaking, experience rating

## 1. INTRODUCTION AND MOTIVATION

One of the main tasks of the actuary is to design a tariff structure that will fairly distribute the burden of claims among policyholders. To this end, he often has to partition all policies into homogeneous classes with all policyholders belonging to the same class paying the same premium. The classification variables introduced to partition risks into cells are called *a priori* variables (as their values can be determined before the policyholder starts to drive). In motor third-party liability (MTPL, in short) insurance, they include age, gender and occupation of the policyholders, type and use of their car, place where they live and sometimes even number of cars in the household or marital status. It is convenient to achieve *a priori* classification by resorting to generalized linear models (e.g. Poisson regression).

However, many important factors cannot be taken into account at this stage; think for instance of swiftness of reflexes, aggressiveness behind the wheel or knowledge of the highway code. Consequently, risk classes are still quite heterogeneous despite the use of many *a priori* variables. But it is reasonable to believe that these hidden factors are revealed by the number of claims reported by the policyholders over the successive insurance periods. Hence the amount of premium is adjusted each year on the basis of the individual claims experience in order to restore fairness among policyholders.

Rating systems penalizing insureds responsible for one or more accidents by premium surcharges (or *maluses*), and rewarding claim-free policyholders by awarding them discounts (or *bonuses*) are now in force in many developed countries. This *a posteriori* ratemaking is a very efficient way of classifying

ASTIN BULLETIN, Vol. 33, No. 2, 2003, pp. 419-436

policyholders according to their risk. Besides encouraging policyholders to drive carefully (i.e. counteracting moral hazard), they aim to better assess individual risks. Such systems are called no-claim discounts, experience rating, merit rating, or Bonus-Malus systems (BMS, in short). We will adopt here the latter terminology. For a thorough presentation of the techniques relating to BMS, see Lemaire (1995).

When a BMS is in force, the amount of premium paid by the policyholder depends on the rating factors of the current period but also on claim history. In practice, a BMS consists of a finite number of levels, each with its own relative premium. New policyholders have access to a specified class. After each year, the policy moves up or down according to transition rules and to the number of claims at fault. The premium charged to a policyholder is obtained by applying the relative premium associated to his current level in the BMS to a base premium depending on his observable characteristics incorporated into the price list.

The problem addressed in this paper is the determination of the relative premiums attached to each of the levels of the BM scale when *a priori* classification is used by the company. The severity of the *a posteriori* corrections must depend on the extent to which amounts of premiums vary according to observable characteristics of policyholders. The key idea is that both *a priori* classification and *a posteriori* corrections aim to create tariff cells as homogeneous as possible. The residual heterogeneity inside each of these cells being smaller for insurers incorporating more variables in their *a priori* ratemaking, the *a posteriori* corrections must be softer for those insurers.

This paper is not conceptually innovating. All the ideas are contained in the seminal work by Taylor (1997). Our only contribution is to show how it is possible to avoid simulations by providing analytical formulas for the relative premiums attached to each level of the BM scale.

Our work is organized as follows. In Section 2, we briefly present the modelling used to compute pure premiums. Section 3 describes BM scales and their representation as Markov chains. Section 4 explains how to determine the relative premiums when *a priori* classification is in force or not. Section 5 describes several numerical illustrations. In Section 6, we show that it is possible to apply different *a posteriori* corrections according to *a priori* characteristics. The final Section 7 discusses some possible improvements and concludes.

## 2. Credibility updating formulas

Let $N_{it}, t = 1, 2, \ldots$, represent the number of claims incurred by policyholder $i$ in period $t$. The annual expected claim frequency for policy $i$ in year $t$ is $\lambda_{it} = \mathbb{E}[N_{it}]$. It is expressed as the exponential transform of some predictor involving the characteristics of policyholder $i$ in period $t$. Of course, all the risk factors cannot be taken into account at this stage.

Risk classes remain heterogeneous despite the use of many *a priori* risk characteristics. This residual heterogeneity can be represented by a random effect $\Theta_i$ superposed to the annual expected claim frequency. Specifically, given $\Theta_i = \theta$ the annual numbers of claims $N_{it}$ are assumed to be independent and to conform to a Poisson distribution with mean $\lambda_{it}\theta$, i.e.

$$\Pr\left[N_{it} = k \mid \Theta_i = \theta\right] = \exp\left(-\theta\lambda_{it}\right)\frac{\left(\lambda_{it}\theta\right)^k}{k!}, \quad k \in \mathbb{N}.$$

Moreover, all the $\Theta_i$'s are assumed to be independent and to follow a standard Gamma distribution with probability density function

$$u(\theta) = \frac{1}{\Gamma(a)} a^a \theta^{a-1} \exp(-a\theta), \quad \theta \in \mathbb{R}^+. \tag{2.1}$$

The latter is often referred to as the structure function of the portfolio. Since $\mathbb{E}[\Theta_i] = 1$ we have that $\mathbb{E}[N_{it}] = \lambda_{it}; \lambda_{it}$ is the expected claim number for a policyholder for which no information about past claims is available.

The premium is then adjusted over time with the help of credibility techniques. We assume that each policyholder has an unknown expected claim frequency $\vartheta_i$, constant over time. Following the seminal work of Dionne and Vanasse (1989), the company approaches this unknown value with annual predictions of the form $\hat{\vartheta}_{i1} = \lambda_{i1}$ and for $t \geq 2$,

$$\hat{\vartheta}_{it} = \mathbb{E}\left[N_{it} \mid N_{i1}, \ldots, N_{it-1}\right] = \lambda_{it} \frac{a + \sum_{\tau=1}^{t-1} N_{i\tau}}{a + \sum_{\tau=1}^{t-1} \lambda_{i\tau}}. \tag{2.2}$$

The latter Bayesian credibility estimator cannot be enforced in practice for MTPL, essentially due to commercial reasons and legal constraints. Instead, companies resort to BM scales, that may be considered as simplified versions of credibility theory formulas. Those are presented in the next section.

## 3. Markov Models for Practical BMS

### 3.1. BMS as Markov chains

In practice, insurance companies often resort to BM scales similar to those in Tables 5.4-5.6-5.8 and not on credibility coefficients like those of (2.2). Such scales possess a number of levels, $s + 1$ say, numbered from 0 to $s$. A specified level is assigned to a new driver (often according to the use of the vehicle). Each claim free year is rewarded by a bonus point (i.e. the driver goes one level down). Claims are penalized by malus points (i.e. the driver goes up a certain number of levels each time he files a claim). We assume that the penalty is a given number of classes per claim. This facilitates the mathematical treatment of the problem but more general systems could also be considered. After sufficiently many claim-free years, the driver enters level 0 where he enjoys the maximal bonus.

In commercial BMS, the knowledge of the present level and of the number of claims of the present year suffice to determine the next level. This ensures that the BMS may be represented by a Markov chain: the future (the class for year $t + 1$) depends on the present (the class for year $t$ and the number of accidents reported during year $t$) and not on the past (the complete claim history and the levels occupied during years $1, 2, \ldots, t-1$). Sometimes, fictitious classes have to be introduced in order to meet this memoryless property. Indeed, in some BMS, policyholders occupying high levels are sent to the starting class after a few claimless years.

The relativity associated to level $\ell$ is denoted as $r_\ell$; the meaning is that an insured occupying that level pays an amount of premium equals to $r_\ell$% of the *a priori* premium determined on the basis of his observable characteristics.

### 3.2. Transient distributions

Let $p_{\ell_1\ell_2}(\vartheta)$ be the probability of moving from level $\ell_1$ to level $\ell_2$ for a policyholder with mean frequency $\vartheta$. Further, $M(\vartheta)$ is the one-step transition matrix, i.e. $M(\vartheta) = \{p_{\ell_1\ell_2}(\vartheta)\}$, $\ell_1, \ell_2 = 0, 1, \ldots, s$. Taking the $\nu$th power of $M(\vartheta)$ yields the $\nu$-step transition matrix whose element $(\ell_1\ell_2)$, denoted as $p^{(\nu)}_{\ell_1\ell_2}(\vartheta)$, is the probability of moving from level $\ell_1$ to level $\ell_2$ in $\nu$ transitions.

### 3.3. Stationary distribution

All BMS in practical use have a "best" level, with the property that a policy in that level remains in the same level after a claim-free period. In the following, we restrict attention to such non-periodic bonus rules. The transition matrix $\boldsymbol{M}(\vartheta)$ associated to such a BMS is regular, i.e. there exists some integer $\xi_0 \geq 1$ such that all entries of $\{\boldsymbol{M}(\vartheta)\}^{\xi_0}$ are strictly positive. Consequently, the Markov chain describing the trajectory of a policyholder with expected claim frequency $\vartheta$ accross the levels is ergodic and thus possesses a stationary distribution $\boldsymbol{\pi}(\vartheta) = (\pi_0(\vartheta), \pi_1(\vartheta), \ldots, \pi_s(\vartheta))^t$; $\pi_\ell(\vartheta)$ is the stationary probability for a policyholder with mean frequency $\vartheta$ to be in level $\ell$ i.e.

$$\pi_{\ell_2}(\vartheta) = \lim_{\nu \to +\infty} p^{(\nu)}_{\ell_1\ell_2}(\vartheta).$$

Note that $\boldsymbol{\pi}(\vartheta)$ does not depend on the starting class.

Let us now recall how to compute the $\pi_\ell(\vartheta)$'s. The vector $\boldsymbol{\pi}(\vartheta)$ is the solution of the system

$$\begin{cases} \boldsymbol{\pi}^t(\vartheta) = \boldsymbol{\pi}^t(\vartheta)\boldsymbol{M}(\vartheta), \\ \boldsymbol{\pi}^t(\vartheta)\boldsymbol{e} = 1 \end{cases}$$

where $\boldsymbol{e}$ is a column vector of 1's. Let $\boldsymbol{E}$ be the $(s+1)\times(s+1)$ matrix all of whose entries are 1, i.e. consisting of $s+1$ column vectors $\boldsymbol{e}$. Then, it can be shown that

$$\boldsymbol{\pi}^t(\vartheta) = \boldsymbol{e}^t(\mathbf{I} - \boldsymbol{M}(\vartheta) + \boldsymbol{E})^{-1},$$

which provides a direct method to get $\boldsymbol{\pi}(\vartheta)$. For a derivation of the latter result, see e.g. Rolski et al. (1999).

## 4. Determination of the relativities

### 4.1. Interaction between the BM scale and *a priori* ratemaking

Since the relativities attached to the different levels are the same whatever the risk class to which the policyholders belong, those scales overpenalize *a priori* bad risks. Let us explain this phenomenon, put in evidence by Taylor (1997). Over time, policyholders will be distributed over the levels of the bonus-malus scale. Since their trajectory is a function of past claims history, policyholders with low *a priori* expected claim frequencies will tend to gravitate in the lowest levels of the scale. Conversely for individuals with high *a priori* expected claim frequencies. Consider for instance a policyholder with a high *a priori* expected claim frequency, a young male driver living in a urban area, say. This driver is expected to report many claims (this is precisely why he has been penalized *a priori*) and so to be transferred to the highest levels of the BM scale. On the contrary, a policyholder with a low *a priori* expected claim frequency, a middle-aged lady living in a rural area, say, is expected to report few claims and so to gravitate in the lowest levels of the scale. The level occupied by the policyholders in the BM scale can thus be partly explained by their observable characteristics included in the price list. It is thus fair to isolate that part of the information contained in the level occupied by the policyholder that does not reflect observables characteristics. *A posteriori* corrections should be only driven by this part of the BM information.

Let us try to quantify these findings. To this end, we introduce the random variable $L_\vartheta$ valued in $\{0, 1, \ldots, s\}$ such that $L_\vartheta$ conforms to the distribution $\boldsymbol{\pi}(\vartheta)$ i.e.

$$\Pr[L_\vartheta = \ell] = \pi_\ell(\vartheta), \quad \ell = 0, 1, \ldots, s.$$

The variable $L_\vartheta$ thus represents the level occupied by a policyholder with annual expected claim frequency $\vartheta$ once the steady state has been reached.

Let us now pick at random a policyholder from the portfolio. Let us denote as $\Lambda$ his (unknown) *a priori* expected claim frequency and as $\Theta$ the residual effect of the risk factors not included in the ratemaking. The actual (unknown) annual expected claim frequency of this policyholder is then $\Lambda\Theta$. Since the random effect $\Theta$ represents residual effects of hidden covariates, the random variables $\Lambda$ and $\Theta$ may reasonably be assumed to be mutually independent. Let $w_k$ be the weight of the $k$th risk class whose annual expected claim frequency is $\lambda_k$. Clearly, $\Pr[\Lambda = \lambda_k] = w_k$.

Now, let $L$ be the BM level occupied by this randomly selected policyholder once the steady state has been reached. The distribution of $L$ can be written as

$$\Pr[L = \ell] = \sum_k w_k \int_{\theta > 0} \pi_\ell(\lambda_k \theta) u(\theta) d\theta; \tag{4.1}$$

$\Pr[L = \ell]$ represents the proportion of the policyholders in level $\ell$.

https://doi.org/10.2143/AST.33.2.503701 Published online by Cambridge University Press

### 4.2. Norberg's predictive accuracy in segmented tariffs

Predictive accuracy is a useful measure of the efficiency of a BMS. The idea behind this notion is as follows. A BMS is good at discriminating among the good and the bad risks if the premium they pay is close to their "true" premium. According to Norberg (1976), once the number of classes and the transition rules have been fixed, the optimal relativity $r_\ell$ associated to level $\ell$ is determined by maximizing the asymptotic predictive accuracy.

As above, let $\Lambda\Theta$ be the true (unknown) expected claim frequency of a policyholder picked at random from the portfolio, where $\Theta$ admits the pdf (2.1) and $\Pr[\Lambda = \lambda_k] = w_k$, with $\mathbb{E}[\Lambda] = \overline{\lambda}$. Our aim is to minimize the expected squared difference between the "true" relative premium $\Theta$ and the relative premium $r_L$ applicable to this policyholder (after the stationary state has been reached), i.e. the goal is to minimize

$$
\begin{aligned}
\mathbb{E}[(\Theta - r_L)^2] &= \sum_{\ell=0}^{s} \mathbb{E}\left[(\Theta - r_\ell)^2 \mid L=\ell\right] \Pr[L = \ell] \\
&= \sum_{\ell=0}^{s} \int_{\theta>0} (\theta - r_\ell)^2 \Pr[L = \ell \mid \Theta = \theta]\, u(\theta)\, d\theta \\
&= \sum_k w_k \int_{\theta>0} \sum_{\ell=0}^{s} (\theta - r_\ell)^2 \pi_\ell(\lambda_k \theta)\, u(\theta)\, d\theta.
\end{aligned}
$$

The solution is given by

$$
\begin{aligned}
r_\ell &= \mathbb{E}[\Theta \mid L = \ell] \\
&= \mathbb{E}\left[\mathbb{E}[\Theta \mid L = \ell, \Lambda \mid L=\ell]\right] \\
&= \sum_k \mathbb{E}\left[\Theta \mid L = \ell, \Lambda = \lambda_k\right] \Pr\left[\Lambda = \lambda_k \mid L = \ell\right] \\
&= \sum_k \int_{\theta>0} \theta \frac{\Pr[L = \ell \mid \Theta = \theta, \Lambda = \lambda_k]\, w_k}{\Pr[L = \ell, \Lambda = \lambda_k]} u(\theta)\, d\theta \frac{\Pr[\Lambda = \lambda_k, L = \ell]}{\Pr[L = \ell]} \\
&= \frac{\sum_k w_k \int_{\theta>0} \theta \pi_\ell(\lambda_k \theta)\, u(\theta)\, d\theta}{\sum_k w_k \int_{\theta>0} \pi_\ell(\lambda_k \theta)\, u(\theta)\, d\theta}.
\end{aligned}
\tag{4.2}
$$

It is easily seen that $\mathbb{E}[r_L] = 1$, resulting in financial equilibrium once steady state is reached.

To end with, let us mention that if the insurance company does not enforce any *a priori* ratemaking system, all the $\lambda_k$'s are equal to $\overline{\lambda}$ and reduces to the formula

$$
r_\ell = \frac{\int_{\theta>0} \theta \pi_\ell(\overline{\lambda}\theta)\, u(\theta)\, d\theta}{\int_{\theta>0} \pi_\ell(\overline{\lambda}\theta)\, u(\theta)\, d\theta}
$$

that has been derived in Norberg (1976).

## 5. Numerical illustrations

### 5.1. A priori ratemaking

The data used to illustrate this paper relate to a Belgian MTPL portfolio observed during the year 1997. The data set comprises 158,061 policies. The claim number distribution in the portfolio is described in Table 5.1. The overall mean claim frequency is 11.25%.

TABLE 5.1

Observed claims distribution in the Belgian MTPL portfolio.

| *Number k of claims reported* | *Observed number of policies having reported k claims* |
|---|---|
| 0 | 140 276 |
| 1 | 16 085 |
| 2 | 1 522 |
| 3 | 159 |
| 4 | 17 |
| 5 | 2 |
| ≥ 6 | 0 |

The following information is available on an individual basis: in addition to the number of claims filed by each policyholder and the exposure-to-risk from which these claims originate (i.e. the number of days the policy has been in force during 1997), we know the age of the policyholder in 1997 (18-21 years, 22-30, 31-55 or above 56), his/her gender (male-female), the kind of district where he/she lives (rural or urban), the fuel oils of the vehicle (gasoline or diesel), the power of the vehicle in kilowatts (less than 40 Kw, between 40 and 70 Kw or more than 70Kw), the use of the vehicle (leisure and commuting only, or also professionnal use), whether the vehicle has been classified as a sportscar by the company, whether the policyholder splits the payment of the premium (premium paid once a year versus premium splitted up), whether the policyholder subscribed other guarantees than MTPL (for instance material damage, theft, or comprehensive coverage in addition to MTPL).

A segmented tariff has been built on the basis of a Poisson regression model. Afterwards, geographical ratemaking has been performed following the method proposed by Boskov and Verrall (1994); see also Brouhns, Denuit, Masuy and Verrall (2002). This resulted in the definition of four zones. The final model was fitted by Poisson regression with the four zones that can be seen in Figure 5.1. A backward-type selection procedure eliminated some risk factors: use and sport were considered as non significant and were excluded from the Poisson model. This resulted in 1536 risk classes, each with its own *a priori* annual expected claim frequency. Table 5.2 displays the point estimates

https://doi.org/10.2143/AST.33.2.503701 Published online by Cambridge University Press

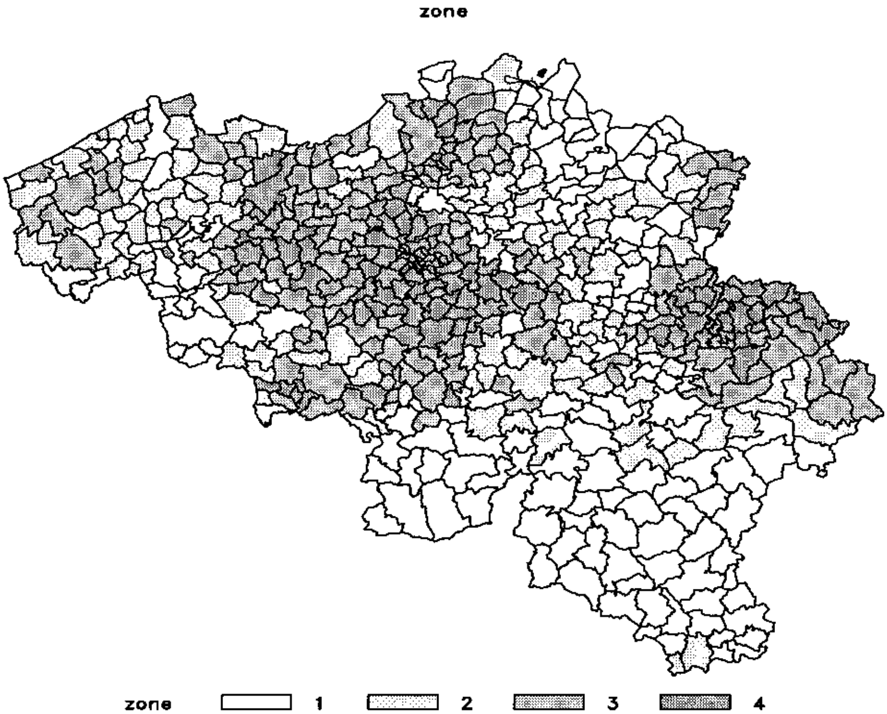


Figure 5.1: The four zones obtained with the Boskov-Verrall method.

of the regression coefficients $\beta_0$, $\beta_1$, … together with confidence intervals and $p$-values of test for the null hypothesis $\beta_j = 0$. Table 5.2 has been obtained with the SAS/STAT procedure GENMOD. Table 5.3 gives a part of the resulting price list. A "1" indicates the presence of the characteristic corresponding to the column. For a thorough description of the tariff construction, we refer the interested reader to Brouhns and Denuit (2003).

TABLE 5.2

SUMMARY OF THE POISSON FIT TO THE BELGIAN MTPL PORTFOLIO

| *Parameter* | *DF* | *Estimate* | *Standard Error* | *Wald 95% Confidence Limits* | | *Chi-Square* | *Pr > ChiSq* |
|---|---|---|---|---|---|---|---|
| Intercept $\beta_0$ | 1 | –1.7326 | 0.0197 | –1.7713 | –01.6939 | 7701.76 | <.0001 |
| AGE 18-21 | 1 | 0.8219 | 0.0578 | 0.7086 | 0.9352 | 202.26 | <.0001 |
| AGE 22-30 | 1 | 0.3996 | 0.0184 | 0.3636 | 0.4357 | 472.45 | <.0001 |
| AGE >56 | 1 | –0.2254 | 0.0185 | –0.2618 | –0.1891 | 147.92 | <.0001 |
| AGE 31-55 | 0 | 0 | 0 | 0 | 0 | . | . |
| GENDER woman | 1 | 0.066 | 0.0165 | 0.0338 | 0.0983 | 16.1 | <.0001 |
| GENDER man | 0 | 0 | 0 | 0 | 0 | . | . |

| *Parameter* | *DF* | *Estimate* | *Standard Error* | *Wald 95% Confidence Limits* | | *Chi-Square* | *Pr > ChiSq* |
|---|---|---|---|---|---|---|---|
| DISTRICT urban | 1 | 0.2439 | 0.0153 | 0.214 | 0.2738 | 255.06 | <.0001 |
| DISTRICT rural | 0 | 0 | 0 | 0 | 0 | . | . |
| FUEL diesel | 1 | 0.2074 | 0.0158 | 0.1764 | 0.2383 | 172.21 | <.0001 |
| FUEL gasoline | 0 | 0 | 0 | 0 | 0 | . | . |
| PAYMENT yearly | 1 | –0.2487 | 0.0147 | –0.2776 | –0.2198 | 284.53 | <.0001 |
| PAYMENT splitted | 0 | 0 | 0 | 0 | 0 | . | . |
| GARACCESS MTPL+ | 1 | –0.1701 | 0.015 | –0.1994 | –0.1407 | 128.97 | <.0001 |
| GARACCESS MTPL only | 0 | 0 | 0 | 0 | 0 | . | . |
| POWER $>$70 | 1 | 0.1243 | 0.0198 | 0.0855 | 0.1631 | 39.38 | <.0001 |
| POWER $<$40 | 1 | –0.0925 | 0.0185 | –0.1288 | –0.0562 | 24.95 | <.0001 |
| POWER 40-70 | 0 | 0 | 0 | 0 | 0 | . | . |
| ZONE 1 | 1 | –0.5492 | 0.0225 | –0.5933 | –0.5051 | 594.8 | <.0001 |
| ZONE 2 | 1 | –0.3525 | 0.0199 | –0.3916 | –0.3135 | 313.2 | <.0001 |
| ZONE 3 | 1 | –0.2301 | 0.0178 | –0.2649 | –0.1952 | 167.63 | <.0001 |
| ZONE 4 | 0 | 0 | 0 | 0 | 0 | . | . |

## 5.2. Scale –1/top

In this BM scale, the policyholders are classified according to the number of claim-free years since their last claim (0, 1, 2, 3, 4 or at least 5). After a claim all premiums reductions are lost. The transition rules are described in Table 5.4. Specifically, the starting class is the highest level 5. Each claim-free year is rewarded by one bonus class. In case an accident is reported, all the discounts are lost and the policyholder is transferred to level 5.

TABLE 5.4

TRANSITION RULES FOR THE BMS –1/TOP.

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

TABLE 5.3

A PRIORI RISK CLASSIFICATION

| *Fuel oil* | | *Premium splitting* | | *Coverage* | | *District* | | *Gender* | | *Power* | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Diesel* | *Gas.* | *Annual* | *Splitted* | *MTPL* | *MTPL +* | *Rural* | *Urban* | *Woman* | *Man* | *40-70Kw* | *$>$70Kw* | *<40Kw* |
| 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |
| 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 |
| 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 |

| *Age* | | | | *Zone* | | | | *Annual* |
|---|---|---|---|---|---|---|---|---|
| ***18-21*** | ***31-55*** | ***22-30*** | ***>55*** | ***Zone 1*** | ***Zone 2*** | ***Zone 3*** | ***Zone 4*** | ***Claim Freq.*** |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0.09858 |
| 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0.09864 |
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0.09873 |
| 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0.09904 |
| 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0.09909 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0.09914 |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0.09940 |
| 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0.09948 |
| 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0.09979 |
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0.09980 |
| 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0.09981 |
| 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0.09992 |
| 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0.10011 |
| 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0.10013 |
| 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0.10013 |
| 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0.10034 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0.10047 |
| 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0.10054 |
| 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0.10062 |
| 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0.10089 |
| 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0.10096 |
| 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0.10106 |
| 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0.10118 |
| 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0.10121 |
| 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0.10139 |

https://doi.org/10.2143/AST.33.2.503701 Published online by Cambridge University Press

Note that the philosophy behind such a BMS is different from credibility theory. Indeed, this BMS only aims to counteract moral hazard: it is in fact more or less equivalent to a deductible which is not paid at once but smoothed over the time needed to go back to the lowest class. Note however that this smoothed deductible only applies to the first claim.

The transition matrix $\boldsymbol{M}(\vartheta)$ associated to this BMS is given by

$$\boldsymbol{M}(\vartheta) = \begin{pmatrix} \exp(-\vartheta) & 0 & 0 & 0 & 0 & 1-\exp(-\vartheta) \\ \exp(-\vartheta) & 0 & 0 & 0 & 0 & 1-\exp(-\vartheta) \\ 0 & \exp(-\vartheta) & 0 & 0 & 0 & 1-\exp(-\vartheta) \\ 0 & 0 & \exp(-\vartheta) & 0 & 0 & 1-\exp(-\vartheta) \\ 0 & 0 & 0 & \exp(-\vartheta) & 0 & 1-\exp(-\vartheta) \\ 0 & 0 & 0 & 0 & \exp(-\vartheta) & 1-\exp(-\vartheta) \end{pmatrix}$$

It is easily checked that $p_{5\ell}^{(5)}(\vartheta) = \pi_\ell(\vartheta)$ for $\ell = 0, 1, \ldots, 5$, so that the system needs 5 years to reach stationarity (i.e. the time needed by the best policyholders starting from level 5 to arrive in level 0).

TABLE 5.5

NUMERICAL CHARACTERISTICS FOR THE SYSTEM –1/TOP

| *Level* $\ell$ | **Pr** $[L = \ell]$ | *Relativity* $r_\ell = \mathbb{E}[\Theta \mid L = \ell]$ ***without*** **a priori** *ratemaking* | *Relativity* $r_\ell = \mathbb{E}[\Theta \mid L = \ell]$ ***with*** **a priori** *ratemaking* | *Average a priori expected claim frequency in level* $\ell$ $\mathbb{E}[\Lambda \mid L = \ell]$ ***with*** **a priori** *ratemaking* |
|---|---|---|---|---|
| 5 | 10.2% | 166.6% | 142.7% | 12.8% |
| 4 | 8.6% | 154.4% | 135.3% | 12.5% |
| 3 | 7.2% | 143.8% | 128.9% | 12.2% |
| 2 | 6.2% | 134.6% | 123.3% | 12.0% |
| 1 | 5.3% | 126.5% | 118.3% | 11.8% |
| 0 | 62.4% | 70.8% | 80.5% | 10.6% |

The results for the BM scale –1/top are displayed in Table 5.5. Specifically, the values in the third column are computed with the help of (4.3) with $\hat{a} = 1.3671$ and $\hat{\overline{\lambda}} = 0.1125$ Those values were obtained by fitting a Negative Binomial distribution to the portfolio observed claim frequencies given in Table 5.1. Integrations have been performed numerically with the QUAD procedure of SAS/IML. The fourth column is based on (4.2) with $\hat{a} = 2.1368$ and the $\hat{\lambda}_k$'s obtained from *a priori* risk classification (i.e. from the $\hat{\beta}_j$'s displayed in Table 5.2). Once the steady state has been reached, the majority of the policies (62.4%) occupy level 0 and enjoy the maximum discount. The remaining 47.6% of the portfolio are distributed over levels 1-5, with about 10% in level 5 (those policyholders who just claimed). Concerning the relativities, the minimum percentage of 70.8% when the *a priori* ratemaking is not recognized becomes 80.5% where the relativities are adapted to the *a priori* risk classification. Similarly, the relativity attached to the highest level of 166.6% gets reduced to 142.7%. The severity of the *a posteriori* corrections is thus weaker once the *a priori* ratemaking is taken into account in the determination of the $r_\ell$'s. The last column of Table 5.5 indicates the extent to which *a priori* and *a posteriori* ratemakings interact. The numbers in this column are computed as

$$\begin{aligned} \mathbb{E}[\Lambda \mid L = \ell] &= \sum_k \lambda_k \Pr[\Lambda = \lambda_k \mid L = \ell] \\ &= \sum_k \lambda_k \frac{\Pr[L = \ell \mid \Lambda = \lambda_k] w_k}{\Pr[L = \ell]} \\ &= \frac{\sum_k \lambda_k w_k \int_{\theta > 0} \pi_\ell(\lambda_k \theta) u(\theta) d\theta}{\sum_k w_k \int_{\theta > 0} \pi_\ell(\lambda_k \theta) u(\theta) d\theta}. \end{aligned} \tag{5.1}$$

If $\mathbb{E}[\Lambda \mid L = \ell]$ is indeed increasing in the level $\ell$, those policyholders who have been granted premium discounts at policy issuance (on the basis of their observable characteristics) will be also rewarded *a posteriori* (because they occupy the lowest levels of the BM scale). Conversely, the policyholders who have been penalized at policy issuance (because of their observable characteristics) will cluster in the highest BM levels and will consequently be penalized again. The average *a priori* expected claim frequency clearly increases with the level $\ell$ occupied by the policyholder.

### 5.3. Soft Taylor's scale (–1/+2)

Let us now consider the soft experience rating system defined in Taylor (1997). There are 9 BM levels. Level 6 is the starting level. A higher level number indicates a higher premium. If no claims have been reported by the policyholder then he moves one level down. If a number of claims, $n_t > 0$, has been reported during year $t$ then the policyholder moves $2n_t$ levels up. The transition rules are described in Table 5.6.

Results are displayed in Table 5.7 which is the analogue of Table 5.5 for the BMS –1/+2. The BMS is perhaps too soft since the vast majority of the portfolio (about 75%) clusters in the super bonus level 0. The higher levels are occupied by a very small minority of drivers. Such a system does not really discriminate between good and bad drivers. Consequently, only those policyholders in level 0 get some discount whereas occupancy of any level 1-8 implies some penalty. Again, the a posteriori corrections are softened when a priori risk classification is taken into account in the determination of the $r_\ell$'s. The comments made for the scale –1/top still apply to this BMS.

TABLE 5.6

TRANSITION RULES FOR THE BMS –1/+2

| ***Starting level*** | ***Level occupied if*** **0** | **1** | **2** | **3** | **≥ 4** |
|---|---|---|---|---|---|
| ***claim(s) 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 5.7

NUMERICAL CHARACTERISTICS FOR THE SYSTEM –1/+2

| ***Level*** $\ell$ | **Pr** $[L = \ell]$ | ***Relativity*** $r_\ell = \mathbb{E}[\Theta \mid L = \ell]$ ***without*** **a priori** ***ratemaking*** | ***Relativity*** $r_\ell = \mathbb{E}[\Theta \mid L = \ell]$ ***with*** **a priori** ***ratemaking*** | ***Average a priori expected claim frequency in level*** $\ell$ $\mathbb{E}[\Lambda \mid L = \ell]$ ***with*** **a priori** ***ratemaking*** |
|---|---|---|---|---|
| 8 | 1.1% | 325.3% | 238.1% | 17.2% |
| 7 | 1.1% | 294.0% | 220.9% | 16.2% |
| 6 | 1.4% | 258.0% | 200.6% | 15.2% |
| 5 | 1.6% | 234.0% | 187.0% | 14.5% |
| 4 | 2.6% | 194.5% | 163.0% | 13.5% |
| 3 | 2.9% | 179.2% | 153.9% | 13.1% |
| 2 | 7.9% | 133.9% | 124.1% | 12.0% |
| 1 | 6.8% | 127.2% | 119.9% | 11.8% |
| 0 | 74.7% | 75.6% | 84.4% | 10.7% |

## 5.4. Severe Taylor's scale (-1/+4)

Let us finally consider the severe experience rating system defined in Taylor (1997). Again, there are 9 BM levels. Level 6 is the starting level. A higher level number indicates a higher premium. If no claims have been reported by the policyholder then he moves down one level. Each claim is now penalized by 4 levels (instead of 2 in the soft Taylor's scale). The transition rules are described in Table 5.8.

TABLE 5.8

TRANSITION RULES FOR THE BMS –1/+4.

| *Starting level* | *Level occupied if* 0 *claim is reported* | *Level occupied if* 1 *claim is reported* | *Level occupied if* ≥ 2 *claim is reported* |
|---|---|---|---|
| 8 | 7 | 8 | 8 |
| 7 | 6 | 8 | 8 |
| 6 | 5 | 8 | 8 |
| 5 | 4 | 8 | 8 |
| 4 | 3 | 8 | 8 |
| 3 | 2 | 7 | 8 |
| 2 | 1 | 6 | 8 |
| 1 | 0 | 5 | 8 |
| 0 | 0 | 4 | 8 |

Results are displayed in Table 5.9, the analogue of Tables 5.5 and 5.7. The interesting point is to compare results for the scale –1/+2 to those obtained for the scale –1/+4. The higher severity of the –1/+4 system results in more important premium discounts in the lowest part of the scale, and in reduced penalties for those occupying the highest levels. Similarly, the average *a priori* expected claim frequency for each level diminishes when the claims are more heavily penalized.

TABLE 5.9

NUMERICAL CHARACTERISTICS FOR THE SYSTEM –1/+4

| *Level* $\ell$ | **Pr** $[L = \ell]$ | *Relativity* $r_\ell = \mathbb{E}[\Theta \mid L = \ell]$ *without* **a priori** *ratemaking* | *Relativity* $r_\ell = \mathbb{E}[\Theta \mid L = \ell]$ *with* **a priori** *ratemaking* | *Average a priori expected claim frequency in level* $\ell$ $\mathbb{E}[\Lambda \mid L = \ell]$ *with* **a priori** *ratemaking* |
|---|---|---|---|---|
| 8 | 4.6% | 225.1% | 180.7% | 14.3% |
| 7 | 4.3% | 203.0% | 167.3% | 13.7% |
| 6 | 4.0% | 185.7% | 156.9% | 13.2% |
| 5 | 3.8% | 171.7% | 148.6% | 12.9% |
| 4 | 7.0% | 130.0% | 121.1% | 11.9% |
| 3 | 6.1% | 123.0% | 116.8% | 11.7% |
| 2 | 5.3% | 116.7% | 112.8% | 11.6% |
| 1 | 4.7% | 111.1% | 109.2% | 11.5% |
| 0 | 60.3% | 64.9% | 76.5% | 10.5% |

## 6. *A POSTERIORI* CORRECTIONS DEPENDING ON *A PRIORI* CHARACTERISTICS

We know from credibility theory that the *a posteriori* corrections are functions of the *a priori* characteristics; see (2.2). On the contrary, when a BMS is in force, the same *a posteriori* corrections apply to all policyholders, whatever their *a priori* expected claim frequency. This of course induces unfairness in the portfolio.

In order to reduce the unfairness of the tariff, we could propose several BM scales, according to the *a priori* characteristics. Table 6.1 describes such a system where the company differentiates policyholders according to the type of district where they live (urban or rural). People living in urban areas have higher *a priori* expected claim frequencies. Thus, they should be more rewarded in case they do not file any claim and less penalized when they report accidents compared to people living in rural zones. This is indeed what we observe when we compare the relative premiums obtained for the system –1/+4: the maximal discount is 73.1% for urban policyholders, compared to 77.7% for rural ones. Similarly, the highest penalty is 176.6% for urbans against 183.0% for rurals.

TABLE 6.1

NUMERICAL CHARACTERISTICS FOR THE SYSTEM –1/+4 WITH THE DICHOTOMY URBAN/RURAL.

| | *Urban* | | *Rural* | |
|---|---|---|---|---|
| *Level* $\ell$ | *Relativity* $r_\ell = \mathbb{E}[\Theta \mid L = \ell]$ *with* a priori *ratemaking* | *Average a priori expected claim frequency level* $\ell$ $\mathbb{E}[\Lambda \mid L = \ell]$ *with* a priori *ratemaking* | *Relativity* $r_\ell = \mathbb{E}[\Theta \mid L = \ell]$ *with* a priori *ratemaking* | *Average a priori expected claim frequency in in level* $\ell$ $\mathbb{E}[\Lambda \mid L = \ell]$ *with* a priori *ratemaking* |
| 8 | 176.6% | 16.5% | 183.0% | 13.0% |
| 7 | 162.5% | 15.8% | 169.8% | 12.5% |
| 6 | 151.6% | 15.3% | 159.6% | 12.2% |
| 5 | 142.9% | 14.9% | 151.4% | 11.9% |
| 4 | 116.8% | 13.8% | 122.9% | 11.1% |
| 3 | 112.2% | 13.6% | 118.7% | 10.9% |
| 2 | 108.1% | 13.4% | 114.8% | 10.8% |
| 1 | 104.3% | 13.3% | 111.2% | 10.7% |
| 0 | 73.1% | 12.2% | 77.7% | 9.8% |

## 7. DISCUSSION

All the techniques used in this paper resort to the stationary distribution of the scale. Therefore they can only be recommended if the steady state is reached after a relatively short period, as it is the case for the BM scale –1/top. It is worth mentioning that for the scale –1/top, the use of the stationary distribution for the computation yields higher premiums than those obtained using transient distributions, with the method of Børgan, Hoem and Norberg (1981).

The method described in the present paper can be extended to transient distributions, in the spirit of Børgan, Hoem and Norberg (1981). This may be interesting when a new scale is introduced or for BMS needing many years to reach their stationay regime.

If on a given market companies start to compete on the basis of BMS many policyholders could leave the portfolio after the occurrence of an accident, in order to avoid the resulting penalties. Those attritions can be incorporated in the model by adding an additional level to the Markov chain (in the spirit of Centeno and Silva (2001)). Transitions from a level of the BMS to this state represents a policyholder leaving the portfolio whereas transitions from this state to any level of the BMS means that a new policy enters the portfolio.

It has been assumed throughout this paper that the unknown expected claim frequencies were constant and that the random effects representing hidden characteristics were time-invariant. Dropping these assumptions makes the determination of the relativities much harder. We refer the interested reader to Brouhns, Guillén, Denuit and Pinquet (2003) for a thorough study of this general situation.

## Acknowledgements

The authors wish to express their gratitude to the anonymous referees who have corrected a mistake in the original manuscript and whose numerous remarks and suggestions have considerably improved the text.

## References

Borgan, Ø, Hoem, J.M. and Norberg, R. (1981) A nonasymptotic criterion for the evaluation of automobile bonus systems. *Scandinavian Actuarial Journal*, 265-178.

Boskov, M. and Verrall, R.J. (1994) Premium rating by geographical area using spatial models. *ASTIN Bulletin* **24**, 131-143.

Brouhns, N. and Denuit, M. (2003) Applications des modèles généralisés additifs à l'assurance automobile. Manuscript.

Brouhns, N., Denuit, M., Masuy, B. and Verrall, R. (2002) Ratemaking by geographical area in the Boskov and Verrall model: a case study using Belgian car insurance data. *actu-L* **2**, 3-28.

Brouhns, N., Guillén, M., Denuit, M. and Pinquet, J. (2003) Optimal bonus-malus scales in segmented tariffs. *Journal of Risk and Insurance*, in press.

Centeno, M. and Silva, J.M.A. (2001) Bonus systems in an open portfolio. *Insurance: Mathematics & Economics* **28**, 341-350.

Dionne, G., and Vanasse, C. (1989) A generalization of actuarial automobile insurance rating models: the Negative Binomial distribution with a regression component. *ASTIN Bulletin* **19**, 199-212.

Lemaire, J. (1995) *Bonus-Malus Systems in Automobile Insurance*. Kluwer Academic Publisher, Boston.

NORBERG, R. (1976) A credibility theory for automobile bonus system. *Scandinavian Actuarial Journal*, 92-107.

ROLSKI, T., SCHMIDLI, H., SCHMIDT, V. and TEUGELS, J. (1999) *Stochastic Processes for Insurance and Finance*. John Wiley & Sons, New York.

TAYLOR, G. (1997) Setting a Bonus-Malus scale in the presence of other rating factors. *ASTIN Bulletin* **27**, 319-327.

SANDRA PITREBOIS & JEAN-FRANÇOIS WALHIN
*Secura Belgian Re*
*Avenue des Nerviens, 9-13 boîte 6*
*B-1040 Bruxelles, Belgium*

MICHEL DENUIT
*Institut de Statistique*
*Université Catholique de Louvain*
*Voie du Roman Pays, 20*
*B-1348 Louvain-la-Neuve, Belgium*

JEAN-FRANÇOIS WALHIN & MICHEL DENUIT
*Institut des Sciences Actuarielles*
*Université Catholique de Louvain*
*Grand-Rue, 54*
*B-1348 Louvain-la-Neuve, Belgium*