2001

Exponential Bonus-Malus Systems Integrating A Priori Risk Classification

Llufs Bermudez
University of Barcelona, lbermu@eco.ub.es

Michel Denuit
Universite Catholique de Louvain, Belgium, denuit@stat.ucl.ac.be

Jan Dhaene
Catholic University of Leuven, Belgium, jan.dhaene@econ.kuleuven.ac.be

Exponential Bonus-Malus Systems Integrating A Priori Risk Classification

Lluís Bermúdez,* Michel Denuit,† and Jan Dhaene‡

Abstract§

This paper examines an integrated ratemaking scheme including a priori risk classification and a posteriori experience rating. In order to avoid the high penalties implied by the quadratic loss function, the symmetry between the overcharges and the undercharges is broken by introducing parametric loss functions of exponential type.

Key words and phrases: *quadratic loss function, exponential loss function, credibility estimation, explanatory variables, experience rating, risk classification*

*Lluís Bermúdez, Ph.D., is a professor of economics at the University of Barcelona.

Dr. Bremúdez's address is: Departament de Matemàtica Econòmica i Empreserial, Facultat de Econòmiques, Universitat de Barcelona, Avda. Diagonal 690, S-08034 Barcelona, SPAIN. Internet address: *lbermu@eco.ub.es*

†Michel Denuit, Ph.D., F.A.R.A.B., is professor of actuarial mathematics at the Université Catholique de Louvain, Belgium. He received his B.Sc. in mathematics and his Ph.D. in statistics from the Université Libre de Bruxelles, Belgium, where he worked five years as a lecturer. Dr. Denuit is a fellow of the Belgian Society of Actuaries (KVBA-ARAB), a former member of the Directorial Board of the Society, a member of the Education Committee, and has been a Belgian representative to the IAA and Groupe Consultatif Education Committees.

Dr. Denuit's address is: Institut de Statistique, Université Catholique de Louvain, Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, BELGIUM. Internet address: *denuit@stat.ucl.ac.be*

‡Jan Dhaene, Ph.D., is professor of actuarial mathematics at the Catholic University of Leuven, Belgium. He obtained his B.Sc. in mathematics from University of Ghent in 1985 and his Ph.D. in actuarial science from University of Leuven in 1991. His research interests include topics related to modeling dependencies, with applications in finance and insurance.

Dr. Dhaene's address is: Departement Toegepaste Economische Wetenschappen, Faculteit Economische And Toegepaste Economische Wetenschappen, Katholieke Universiteit Leuven, Minderbroedersstraat 5, B-3000 Leuven, BELGIUM. Internet address: *Jan.Dhaene@econ.kuleuven.ac.be*

§The authors would like to thank the referee for carefully reading the manuscript, as well as for several useful comments and suggestions.

1 Introduction and Motivation

1.1 A Priori Risk Classification Variables

One of the main tasks of an actuary is to design a tariff structure that fairly distributes the burden of current claims among its portfolio of current policyholders. If the insurance portfolio consists of heterogeneous risks (policyholders), then it is fair to partition the portfolio into homogeneous classes of policies with policyholders belonging to the same class paying the same premium.

The classification variables used to partition the portfolio into homogeneous cells are called a priori variables (as their values can be determined at the start of the policy). In automobile third-party liability insurance, for example, the commonly used classification variables include the age, gender, marital status, occupation, type and use of car, and residential address. Generalized linear models can be used to select the a priori classification variables.[1]

In most practical situations many important factors (such as driving style, reflexes, or knowledge of rules of the road) cannot be taken into account when selecting the a priori classification variables. Consequently, even after the a priori classification variables have been chosen, tariff cells may still be heterogeneous. It is reasonable to believe, however, that these characteristics are revealed by the number and sizes of claims reported by the policyholders over the successive insurance periods. Hence, at the end of each insurance period the next period's premium is adjusted on the basis of the individual's claims experience in order to ensure fair premiums among policyholders.

It is interesting to mention that in North America, emphasis has traditionally been laid on a priori ratings using many classifying variables, while in continental Europe just a few a priori classifying variables were used and much importance was placed on the a posteriori evaluation of drivers. Since July 1994, however, European Union (EU) directives have introduced complete rating freedom. Insurance companies operating in EU countries are now (theoretically) free to set up their own rates, select their own classification variables, and design their own bonus-malus system.[2] Companies in most EU countries have taken advantage of this freedom by introducing more rating variables.

---

[1]For more on generalized linear models, see, for example, Renshaw (1994) or Pinquet (1997,1999) for applications in actuarial science; or Mc Cullagh and Nelder (1989), Dobson (1990), or Fahrmeir and Tutz (1994).

[2]For a thorough presentation of the techniques relating to bonus-malus systems, we refer the interested reader to Lemaire (1995).

In a competitive insurance market the trend is toward portfolio segmentation because insurers tend to use all available relevant information to match the rating structure used by competitors. As the only item of interest is the unknown distribution function of the claim amounts produced by the driver during a period, it seems fair to correct the inadequacies of the *a priori* system by using an adequate bonus-malus system. Such an experience rating system should be better accepted by policyholders than arbitrary *a priori* classifications.

A bonus-malus system is a rating system based on the following mechanism:

1. Claim-free policyholders, i.e., those with zero claims within a single period, are rewarded by premium discounts called *bonuses*; and

2. Policyholders reporting one or more accidents at fault during a period are penalized by premium surcharges called *maluses*.

This a posteriori ratemaking system is an efficient way of classifying policyholders into cells according to their risk. As pointed out by Lemaire (1995), if insurers are allowed to use only one rating variable, it should be a merit rating variable because merit rating variables are the best predictor of the number of claims incurred by a driver. Besides encouraging policyholders to drive carefully (i.e., to counteract moral hazard), merit rating systems aim to better assess individual risks, so that everyone will pay in the long run a premium corresponding to her or his own claim frequency. Such systems are called no-claim discounts, experience rating, merit rating, or bonus-malus systems.

1.2 The Nature of Risk Transfers[3]

Consider a portfolio of automobile third-party liability insurance policies. Let Y denote a quantity of actuarial interest for a policy taken at random from the portfolio. For example, Y can be the amount of a claim, the aggregate claims in one period, or the number of accidents at fault reported by the policyholder during one period. The actuary has a set of observable risk classification variables, $\mathbf{X}$, pertaining to the selected policyholder, which may include such items as age, gender, marital status, occupation, home address, type and use of her or his car. In addition, Y also depends on a set of unknown characteristics $\mathbf{Z}$, which may include such items as annual mileage (i.e., risk exposure),

[3]The ideas presented in this section are inspired by De Wit and Van Eeghen (1984).

accuracy of judgment, aggressiveness behind the wheel, drinking behavior, etc. Some of the elements of $\mathbf{Z}$ are unobservable; others cannot be measured in a cost efficient way. If Ω denotes the entire set of risk factors for this policyholder then

$$\Omega = \mathbf{X} \cup \mathbf{Z}.$$

The true net premium for this policyholder is $\mathbb{E}[Y|\Omega]$; it is worth mentioning that this premium is a random variable but much less dispersed than Y itself, making insurance policies worth buying. The situation can be summarized as described in Table 1. In this case, the policyholder keeps the variations of the premiums due to the modifications in her or his personal characteristics Ω and transfers to the company the purely random fluctuations of Y (that is, the variance of the outcomes of Y once the personal characteristics $\mathbf{X}$ and $\mathbf{Z}$ have been taken into account). As the elements of $\mathbf{Z}$ are unknown to the insurer, the situation described in Table 1 is purely theoretical. Because the company only knows $\mathbf{X}$, the actual reality of the insurance business is rather as depicted in Table 2.

Table 1
Risk Transfer Between Insurance Company And Policyholder in Case of Full Information

	Amount Carried By	
	Policyholder	Insurer
Risk:	$\mathbb{E}[Y|\Omega]$	$Y - \mathbb{E}[Y|\Omega]$
Expectation:	$\mathbb{E}[Y]$	0
Variance:	$\mathbb{V}\mathrm{ar}\left[\mathbb{E}[Y|\Omega]\right]$	$\mathbb{E}\left[\mathbb{V}\mathrm{ar}[Y|\Omega]\right]$

It is well known to statisticians and actuaries that for a random variable A and a random vector B (possibly of dimension 1),

$$\mathbb{E}[A] = \mathbb{E}[\mathbb{E}[A|B]] \quad \text{and} \quad \mathbb{V}\mathrm{ar}[A] = \mathbb{E}[\mathbb{V}\mathrm{ar}[A|B]] + \mathbb{V}\mathrm{ar}[\mathbb{E}[A|B]].$$

If we let $A = Y|\mathbf{X}$ and $B = \Omega$, then

$$\mathbb{E}[\mathbb{V}\mathrm{ar}[Y|\mathbf{X}]] = \mathbb{E}[\mathbb{V}\mathrm{ar}[Y|\Omega]] + \mathbb{E}[\mathbb{V}\mathrm{ar}[\mathbb{E}[Y|\Omega]\,|\mathbf{X}]].$$

The first term on the right, i.e., $\mathbb{E}\left[\mathbb{V}\mathrm{ar}[Y|\Omega]\right]$, represents the purely random fluctuations of the risk and is supported by the insurance company. The second term on the right represents the variation in the expected claims due to the unknown risk characteristics **Z**. This quantity should be corrected by an experience rating mechanism.

Table 2
Risk Transfer Between Insurance Company And Policyholder in Case of Partial Information

	Amount Carried By	
	Policyholder	Insurer
Risk:	$\mathbb{E}[Y\|\mathbf{X}]$	$Y - \mathbb{E}[Y\|\mathbf{X}]$
Expectation:	$\mathbb{E}[Y]$	0
Variance:	$\mathbb{V}\mathrm{ar}\left[\mathbb{E}[Y\|\mathbf{X}]\right]$	$\mathbb{E}\left[\mathbb{V}\mathrm{ar}[Y\|\mathbf{X}]\right]$

Next, assume the insurance company incorporates more a priori variables in its pricing structure; that is, $\tilde{\mathbf{X}}$ (with $\mathbf{X} \subset \tilde{\mathbf{X}}$) is substituted for **X**.

$$\mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[\mathbb{E}[Y|\Omega]\,\middle|\tilde{\mathbf{X}}\right]\right] \leq \mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[\mathbb{E}[Y|\Omega]\,\middle|\mathbf{X}\right]\right],$$

that is, the residual heterogeneity in the portfolio is reduced. Consequently, the variance of the insurer's experience is also reduced, i.e.,

$$\mathbb{E}\left[\mathbb{V}\mathrm{ar}[Y|\tilde{\mathbf{X}}]\right] \leq \mathbb{E}\left[\mathbb{V}\mathrm{ar}[Y|\mathbf{X}]\right].$$

The severity of the a posteriori corrections thus decreases as the information used by the insurer increases.

1.3 Objectives

Let $\mathcal{F}_t$ denotes the entire past claims experience available about Y at time t. The central idea behind experience rating is that $\mathcal{F}_t$ reveals its hidden features **Z** as $t \to \infty$, i.e., the information contained in $(\mathbf{X}, \mathcal{F}_t)$ becomes comparable to $\mathbf{\Omega}$ as time goes on. Therefore, the a posteriori premium is $\mathbb{E}[Y|\mathbf{X}, \mathcal{F}_t]$.

The aim of this paper is to examine the interaction between a priori ratemaking (i.e., identification of the best predictors **X** and of the risk premium $\mathbb{E}[Y|\mathbf{X}]$) and a posteriori ratemaking (i.e., premium corrections according to past claims history $\mathcal{F}_t$ in order to reflect the unavailable information contained in **Z**).

The paper is organized as follows: Section 2 contains a brief review of the current methodology of automobile ratemaking in EU countries. It considers risk classification and credibility as two separate problems. This approach has flaws because the aim of experience rating is to reduce the residual heterogeneity of the portfolio, which obviously depends on the degree of a priori segmentation. Therefore a priori and a posteriori ratemaking have to be integrated in a continuous risk evaluation mechanism. In Section 3, we present the results of Dionne and Vanasse (1989, 1992) and Gisler (1996), as well as an alternative approach based on an exponential loss function. Such loss functions have been considered by Ferreira (1977), Lemaire (1979), Young (1996), and Denuit and Dhaene (2001), among others.

Our methods are illustrated by an example using a Spanish insurance portfolio. This example considers only two risk factors and allows for a deeper understanding of the technical concepts introduced. Adaptation of the methodology to real-life portfolio is then straightforward. Several optimization programs are used extensively throughout this paper (some of them are standard in actuarial science, others are less common). The appendix contains a description of all results, together with proofs for the sake of completeness.

2 Current Methodology

2.1 The Model

Consider an automobile portfolio consisting of N independent policies. These policies are split into M homogeneous risk classes. The premium paid by each policyholder depends on the policyholder's rating factors for the current period and also on her or his claim history. The premium charged is the product of a risk classification base premium and of a bonus-malus coefficient. The base premium for a risk class is a function of the current rating factors, whereas the bonus-malus coefficient only depends on the policyholder's history of reported claims at fault.

We assume the insurance company determines its risk classification factor using generalized linear models; see, e.g., in Renshaw (1994). We suppose the N risks are partitioned into M distinct (disjoint) risk classes. In each risk class, the policies are identical from the company point of view, whereas policies in different risk classes have distinct risk profiles.

For $m = 1, 2, \ldots, M$, the base premium for the m^{th} risk class is denoted by BP_m, which is the amount charged to a new policyholder entering the m^{th} risk class. Of course, inside each risk class, the policies are not strictly identical. Therefore, the premium is adjusted over time using a bonus-malus factor $\text{BMF}(k, t)$ where t is the number of years the policy is in force and k is the number of claims reported while the policy is in force.

Notice that while the base premium depends on the risk class, the *same* bonus-malus factor is applied to all drivers, i.e., it is independent of the risk class. This is erroneous because a bonus-malus system is supposed to correct the actual premium for the residual heterogeneity existing in the different risk classes, which implies that the severity of a bonus-malus system must depend on the policyholder's risk class. In fact, the more a priori risk factors used in the risk classification system, the less severe bonus-malus coefficients should be. Uniform bonus-malus systems imposed by regulatory authorities in some EU countries (e.g., Belgium and France) create cross-subsidization of insurance portfolios.

Let

K_{ij} = Number of claims incurred by the i^{th} policyholder during period $(j-1, j)$;

n_{ik} = Number of policies from class i reporting k claims;

Θ_i = Risk proneness parameter of policyholder i. It captures the propensity of policyholder i to produce claims and is regarded as a random variable; and

Z_{ijk} = Size (severity) of the k^{th} claim produced by the i^{th} policyholder during year $(j-1, j)$.

At the portfolio level, the vectors $(\Theta_i, K_{i1}, K_{i2}, K_{i3}, \ldots)$ are assumed to be independent and identically distributed for $i = 1, 2, \ldots, N$. Also, given $\Theta_i = \theta$, the random variables $K_{i1}, K_{i2}, K_{i3}, \ldots$ are assumed to be independent and identically distributed for fixed i. Unconditionally, these random variables are dependent. For fixed i, the Z_{ijk}s are assumed to be independent and identically distributed and independent of the claim frequencies K_{ij}. This assumption has been questioned by several authors because it implies that the cost of an accident is, for the most part, beyond the control of a policyholder. Though the degree of care exercised by a driver may mostly influence the number of accidents, it has less influence on the cost of these accidents. Nevertheless, this assumption seems acceptable in third-party liability insurance. The Z_{ijk} are also independent of Θ_i for any given i.

The total claim amount for policyholder i in year j is

$$S_{ij} = \sum_{k=1}^{K_{ij}} Z_{ijk}.$$

We put $\mathbb{E}[Z_{ijk}] = 1$, which means that the expected claim amount is chosen as monetary unit. The pure premium for policy i in year j is then given by

$$\mathbb{E}[S_{ij}|\Theta_i = \theta] = \mathbb{E}[K_{ij}|\Theta_i = \theta] = \theta.$$

A priori (i.e., without information about claims history), an identical amount of premium $\mathbb{E}[\Theta_i]$ is charged to new policyholders.

Given $\Theta_i = \theta$, the numbers of claims generated in $(j-1, j)$ by policyholder i are assumed to be independent and identically distributed (i.i.d.) Poisson random variables with mean θ, i.e.,

$$\mathbb{P}\mathrm{r}[K_{ij} = k|\Theta_i = \theta] = \exp(-\theta)\frac{\theta^k}{k!}, \tag{1}$$

where θ is the claim frequency of this policyholder. The cumulative distribution function (cdf) of Θ_i, $F_\Theta(\cdot)$, (often called the *structure function*), belongs to the two-parameter gamma family, i.e.,

$$F_\Theta(\theta) = \Gamma(\theta|\alpha,\tau) \tag{2}$$

where

$$\Gamma(\theta|\alpha,\tau) = \int_0^\theta \frac{\tau^\alpha v^{\alpha-1} e^{-\tau v}}{\Gamma(\alpha)}\, dv, \qquad \alpha,\tau,\theta > 0. \tag{3}$$

Combining equations (1) and (2) yields the well-known result that the number of claims for a policyholder randomly drawn from the portfolio follows a negative binomial distribution, i.e.,

$$\mathbb{P}\mathrm{r}[K_{ij} = k] = \frac{k+\alpha-1}{k}\left(\frac{\tau}{1+\tau}\right)^\alpha \left(\frac{1}{1+\tau}\right)^k. \tag{4}$$

Though $K_{i1}, K_{i2}, \ldots$ are identically distributed, they are not independent, because they are generated by the same policyholder and thus contingent on the same risk parameter Θ_i.

2.2 A Posteriori Premiums

Suppose policyholder i has been observed for t years and the number of claims reported during this period is $k_{i1}, k_{i2}, \dots, k_{it}$. The premium for year $t+1$ is defined as a function Ψ_t of the claims reported during the previous years, which is determined by minimizing $\mathbb{E}[L(\Theta_i - \Psi_t(k_{i1}, k_{i2}, \dots, k_{it}))]$ for some loss function L, taken to be non-negative, convex, and such that $L(0) = 0$. The loss functions considered in this paper are the quadratic loss where $L(x) = x^2$ and the exponential loss with positive parameter c where $L(x) = \exp(-cx)$.

From the results recalled in the appendix, we easily get the following proposition.

Proposition 1. *The best estimator of the pure premium Θ_i at time $t+1$ is given by*

$$W_{t+1}^{(q)} = \frac{\alpha}{\tau}(1 - \rho_q) + \frac{k_{i\bullet}(t)}{t}\rho_q \tag{5}$$

for the quadratic loss function where

$$\rho_q = \frac{t}{\tau + t} \quad \text{and} \quad k_{i\bullet}(t) = \sum_{j=1}^{t} k_{ij}(t) \quad \text{while}$$

$$W_{t+1}^{(e)} = \frac{\alpha}{\tau}(1 - \rho_e(c)) + \frac{k_{i\bullet}(t)}{t}\rho_e(c) \tag{6}$$

for the exponential loss function with $c > 0$, and

$$\rho_e(c) = \frac{t}{c}\ln\left(1 + \frac{c}{\tau + t}\right). \tag{7}$$

Notice that in Proposition 1, both expressions for W_{t+1} are convex combinations of the portfolio mean a/τ and the observed average number of claims $k_{i\bullet}(t)/t$ over the period $[0, t]$. In both cases the weight given to the past claims tends to 1 as t goes to ∞. The weight given to the claim history with the exponential loss function is smaller than the weight given to the claim history a quadratic loss function. i.e.,

$$\frac{t}{c}\ln\left(1 + \frac{c}{\tau + t}\right) \le \frac{t}{\tau + t}.$$

Note that in the Poisson-gamma model, the Bayesian approach coincides with the linear credibility estimator. In other words, Proposition

1 can be interpreted in a semi-parametric framework, as in the classical Bühlmann-Straub approach. Notice that

$$\lim_{c\to 0} W_{t+1}^{(e)} = W_{t+1}^{(q)},$$

i.e., the a posteriori premium associated with the exponential loss function converges to that associated with the quadratic loss function.

Also, as $c \to +\infty$ we have that

$$\lim_{c\to +\infty} \rho_e(c) = 0 \text{ so that } W_{t+1}^{(e)} \to \alpha/\tau.$$

This provides an intuitive meaning of the parameter c: if c increases, then the a posteriori merit-rating scheme becomes less severe, and at the limit, the premium no longer depends on the incurred claims. Moreover, routine calculations show that

$$\frac{d}{dc}\rho_e(c) < 0,$$

so that the weight given to the observed average claim number decreases as c increases.

Let $I_i(t) \in (1, 2, \ldots, M)$ denote the index of the risk class occupied by policyholder i during year t. Now, the a posteriori premium for year $t+1$ (i.e., for the time period $(t, t+1)$) charged to policyholder i having reported $k_{i\bullet}(t)$ claims during the first t years is given by

$$P_{t+1}^{(q)}(k_{i\bullet}(t), t) = \mathrm{BP}_{I_i(t+1)}\mathrm{BMF}^{(q)}(k_{i\bullet}(t), t) \tag{8}$$

with

$$\begin{aligned}\mathrm{BMF}^{(q)}(k_{i\bullet}(t), t) &= \frac{W_{t+1}^{(q)}}{\mathbb{E}[\Theta_i]} \\ &= \frac{\alpha + k_{i\bullet}(t)}{\tau + t} \times \frac{\tau}{\alpha}\end{aligned} \tag{9}$$

under a quadratic loss. Under an exponential loss, we get

$$P_{t+1}^{(e)}(k_{i\bullet}(t), t) = \mathrm{BP}_{I_i(t+1)}\mathrm{BMF}^{(e)}(k_{i\bullet}(t), t) \tag{10}$$

with

$$\begin{aligned}\mathrm{BMF}^{(e)}(k_{i\bullet}(t),t) &= \frac{W_{t+1}^{(e)}}{\mathbb{E}[\Theta_i]} \\ &= 1 - \frac{t}{c}\ln\left(1+\frac{c}{\tau+t}\right) \\ &\quad + \ln\left(1+\frac{c}{\tau+t}\right)\frac{k_{i\bullet}(t)}{c}\frac{\tau}{\alpha}. \end{aligned} \tag{11}$$

The model used to determine the bonus-malus coefficients assumes that all the risks of the portfolio have the same a priori claim frequency and that the differences in the claim frequency between the risks are only due to differences in the individual risk characteristics Θ_i. Hence, the model implicitly assumes that the tariff takes into account differences in claim frequencies only through the bonus-malus payments and that such differences are not reflected in the base premiums.

This approach is erroneous because the aim of the bonus-malus system is to adjust the amount of premium according to past claim experience. The effect of this premium adjustment is to reduce the residual heterogeneity within the different risk classes of the portfolio. As the bonus-malus coefficients of Proposition 1 do not take into account explanatory variables, they are functions of the total heterogeneity of the portfolio, before tariff segmentation. In other words, the bonus-malus factors penalize bad risks and reward good risks.

2.3 A Numerical Illustration

Consider Table 3, which displays data from a Spanish insurance company. As can be seen from Table 3, policies have been categorized into 12 classes according to the age of the driver (three categories) and and the power of the car (four categories). The three age categories are "Age $\leq$ 35," "36 $\leq$ Age $\leq$ 49," and Age $\geq$ 50." The four power categories are "Power $\leq$ 53," "54 $\leq$ Power $\leq$ 75," "76 $\leq$ Power $\leq$ 118," and "Power $\geq$ 119."

Let n_{ik} represent the number of policies from class i reporting k claims, $i = 1, 2, \ldots, 12$, and

$$n_{i\bullet} = \sum_{k=0}^{\infty} n_{ik}$$

is the number of policies in the i^{th} class, $i = 1, 2, \ldots, 12$.

Again, we assume that the number of claims reported by a policyholder in class i during a year follows a Poisson distribution with mean

Table 3
The Twelve Risk Classes
For Classification Factors Age and Power

Power of Car (In Horsepower)	Age of Driver (in Years)		
	Age ≤ 35	36 ≤ Age ≤ 49	Age ≥ 50
Power ≤ 53	1	2	3
54 ≤ Power ≤ 75	4	5	6
76 ≤ Power ≤ 118	7	8	9
Power ≥ 119	10	11	12

Table 4
Observed Mean Claim Frequencies
For Classification Factors Age and Power

Power of Car (In Horsepower)	Age of Driver (in Years)		
	Age ≤ 35	36 ≤ Age ≤ 49	Age ≥ 50
Power ≤ 53	0.1866	0.1572	0.1283
54 ≤ Power ≤ 75	0.2685	0.2279	0.1986
76 ≤ Power ≤ 118	0.2992	0.2526	0.2386
Power ≥ 119	0.3217	0.2846	0.2483

λ_i. Moreover, the random variables $K_{i1}, K_{i2}, \ldots$ are assumed to be independent. Therefore, the total number of claims $K_{i\bullet} = \sum_{j=1}^{n_{i\bullet}} K_{ij}$ reported by the $n_{i\bullet}$ policyholders in class i has a Poisson distribution with mean $n_{i\bullet}\lambda_i$. The realization of $K_{i\bullet}$ is $k_{i\bullet} = \sum_{k=1}^{\infty} k n_{ik}$.

Next we introduce the indicator variable J_{ik} such that

$$J_{ik} = \begin{cases} 1 & \text{if policyholder } i \text{ is in age category } k \text{ for } k = 2, 3; \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, define L_{ik} as

$$L_{ik} = \begin{cases} 1 & \text{if policyholder } i \text{ drives a car in category } k \text{ for } k = 2, 3, 4; \\ 0 & \text{otherwise} \end{cases}$$

The i^{th} policyholder is represented by a vector of classification information:

Table 5
Observed Claims Distribution of Number of Policyholders Submitting k Claims in the i^{th} Risk Class, n_{ik}, $i = 1, 2, \ldots, 12$

The First Six Risk Classes, $i = 1, 2, \ldots, 6$						
k	n_{1k}	n_{2k}	n_{3k}	n_{4k}	n_{5k}	n_{6k}
0	3,316	7,797	10,437	9,470	21,031	22,788
1	548	1,063	1,159	1,916	3,775	3,766
2	61	140	143	445	720	591
3	15	17	15	84	143	109
4	4	6	2	21	36	24
5	1	0	1	7	11	5
6	0	0	1	0	2	4
7	0	0	0	1	1	0
8	0	0	0	3	0	0
≥ 9	0	0	0	0	0	0
$n_{i\bullet} =$	3,945	9,023	11,758	11,947	25,719	27,287
$k_{i\bullet} =$	736	1,418	1,509	3,208	5,862	5,420
$\bar{x}_i =$	0.1866	0.1751	0.1283	0.2685	0.2279	0.1986
$s_i^2 =$	0.227	0.1828	0.1501	0.3635	0.2946	0.2451

$$\mathbf{X}_i = (1, J_{i2}, J_{i3}, L_{i2}, L_{i3}, L_{i4}),$$

and a corresponding vector of unknown parameters is

$$\boldsymbol{\eta}^T = (\epsilon, \gamma_2, \gamma_3, \delta_2, \delta_3, \delta_4)$$

where ^{T} denotes the transposed matrix.

When the claim numbers are small, which is typically the case in automobile insurance, the normal approximation is poor and fails to account for the discreteness of the data. Normal regression should be avoided in this case. Generalized linear models provide an appropriate framework for the analysis of discrete data. A linear model for $\ln(\lambda_i)$ is often used in actuarial science. [See, e.g., Pinquet (1997)]. This provides a regression model for count data analogous to the usual normal regression for continuous data. In addition, the standard methodology of generalized linear models uses the logarithmic function as the natural link function for the Poisson distribution. [See, e.g., Dobson (1990).] Thus, we specify a linear model for $\ln(\lambda_i) + \ln(n_{i\bullet})$ as

Table 5 (Continued)
Observed Claims Distribution of Number of Policyholders Submitting k Claims in the i^{th} Risk Class, n_{ik}, $i = 1, 2, \ldots, 12$

The Second Six Risk Classes, $i = 7, 8, \ldots, 12$						
k	n_{7k}	n_{8k}	n_{9k}	n_{10k}	n_{11k}	n_{12k}
0	6,570	15,702	15,158	1,125	4,554	4,680
1	1,423	3,112	2,848	274	902	900
2	321	603	510	69	224	187
3	89	148	123	9	55	25
4	33	31	33	7	15	12
5	6	11	11	1	9	5
6	3	2	1	1	2	1
7	1	0	3	0	0	1
8	1	0	1	0	1	1
≥ 9	0	0	0	0	0	0
$n_{i\bullet} =$	8,447	19,609	18,688	1,486	5,762	5,812
$t_{i\bullet} =$	2,527	4,953	4,459	478	1,640	1,443
$\bar{x}_i =$	0.2992	0.2526	0.2386	0.3217	0.2846	0.2483
$s_i^2 =$	0.4322	0.3288	0.3200	0.4376	0.4214	0.3408

$$\ln(\lambda_i) + \ln(n_{i\bullet}) = \mathbf{X}_i\boldsymbol{\eta} = \epsilon + \sum_{k=2}^{3} \gamma_k J_{ik} + \sum_{k=2}^{4} \delta_k L_{ik}. \tag{12}$$

In order to determine the maximum likelihood estimator of the parameter $\boldsymbol{\eta}$, we have to maximize $L(\boldsymbol{\eta})$ where

$$L(\boldsymbol{\eta}) = \prod_{i=1}^{12} \exp(-\lambda_i n_{i\bullet}) \frac{(\lambda_i n_{i\bullet})^{k_{i\bullet}}}{k_{i\bullet}!}.$$

The regularity conditions satisfied by the Poisson distribution ensure there is a unique solution to the system of equations $\partial \ln L / \partial \boldsymbol{\eta} = 0$. It is easy to check that the maximum likelihood estimator $\hat{\boldsymbol{\eta}}$ of the parameter $\boldsymbol{\eta}$ is the solution of the equations

$$\sum_{i=1}^{12} (k_{i\bullet} - n_{i\bullet}\lambda_i) X_{ij} = 0 \tag{13}$$

for $j = 1, 2, \ldots, 6$, where X_{ij} is the j^{th} element of $\mathbf{X}_i$. As pointed out by Pinquet (1997), equation (13) can be interpreted as an orthogonality relation between the residuals and the covariates. The estimates and the standard deviation and 95% confidence interval for the estimates are displayed in Table 6.

As the rating factors have a finite number of levels and the explanatory variables are indicators of these levels, equation (13) implies that, for every sub-portfolio corresponding to a given level, the sum of the fitted claim numbers is equal to the total number of claims incurred in that sub-portfolio for the observation period. As an example, equation (13) with $j = 2$ ensures that as far as policyholders in age category 2 are concerned the sum of the fitted claim frequencies equals the total number of claims. Consequently, such a system is expected not to create cross-subsidization in the portfolio

Table 6

Parameters Estimates of η in Equation (12)

η	$\hat{\eta}$	Standard Deviation	95% Confidence Interval
ϵ	-1.7219	0.0198	$[-1.7607, -1.6831]$
γ_2	-0.1634	0.0147	$[-0.1922, -0.1345]$
γ_3	-0.2800	0.0149	$[-0.3093, -0.2508]$
δ_2	0.3987	0.0185	$[0.3625, 0.4350]$
δ_3	0.5324	0.0189	$[0.4953, 0.5694]$
δ_4	0.6150	0.0236	$[0.5688, 0.6611]$

It is well known that the vector $\hat{\eta}$ is approximately normal for large sample sizes, with mean η and variance-covariance matrix $\hat{V}$, which is the inverse of the Fisher information matrix. The element (j, k) of $\hat{V}$ is

$$V_{jk} = \sum_{i=1}^{12} X_{ij} X_{ik} n_{i\bullet} \lambda_i.$$

Computing the variance-covariance matrix yields

$$\hat{V}^{-1} = 10^{-3}\begin{pmatrix} 0.392 & -0.144 & -0.151 & -0.277 & -0.277 & -0.265 \\ -0.144 & 0.217 & 0.145 & 0.002 & 0.000 & -0.014 \\ -0.151 & 0.145 & 0.223 & 0.008 & 0.009 & -0.006 \\ -0.277 & 0.002 & 0.008 & 0.342 & 0.274 & 0.273 \\ -0.277 & 0.000 & 0.009 & 0.274 & 0.357 & 0.273 \\ -0.265 & -0.014 & -0.006 & 0.273 & 0.273 & 0.555 \end{pmatrix}.$$

Considering Table 6, all the parameters are significantly different from 0 (because no confidence interval overlaps 0), so that all the covariates are statistically significant. The expected claim numbers for each of the 12 cells are given in Table 7. (It is interesting to compare the fitted results to their empirical counterparts given in Table 3.) Table 7 thus gives the base premiums attached to each of the 12 risk classes.

Table 7
Estimated Mean Claim Frequencies
Based on Classification Factors Age and Power

Power of Car	Age of Driver (in Years)		
(In Horsepower)	Age ≤ 35	36 ≤ Age ≤ 49	Age ≥ 50
Power ≤ 53	0.1787	0.1518	0.1351
54 ≤ Power ≤ 75	0.2663	0.2262	0.2013
76 ≤ Power ≤ 118	0.3044	0.2585	0.2300
Power ≥ 119	0.3306	0.2808	0.2498

In order to calculate the bonus-malus factors, let us consider the claim distribution for the whole portfolio, which is given in Table 8. The negative binomial is fitted using the maximum likelihood approach and is displayed in the third column. The a posteriori premiums are then given by equations (8) and (10) with the estimated values of α and τ given by $\hat{\alpha} = 0.8665$ and $\hat{\tau} = 3.9097$.

Consider for instance a 30-year-old female driver whose car is in the power category "≤ 53." Her a priori expected number of accidents is 0.1787 for the first five years; upon reaching age 35 her expected number of accidents becomes 0.1518. In the first half of Table 9, one can see the bonus-malus coefficients and premiums for that individual. The second column (entitled "BP_t") represents the expected number of accidents (i.e., the base premium) for each period. The BMF_{t+1} column represents the bonus-malus factor in case the policyholder does not cause

Table 8
Observed and Fitted Claim Distribution Using Data in Table 5

k	n_k	$\hat{n}_k$
0	122,628	122,713
1	21,686	21,656
2	4,014	4,116
3	832	801
4	224	158
5	68	31
6	17	6
7	7	1
8	7	0
≥ 9	0	0

Notes: The fit is a negative binomial distribution with parameters $\hat{\alpha} = 0.8665$ and $\hat{\tau} = 3.9097$.

any claims during $(0, t)$ computed on the basis of equation (8). Column $P_{t+1}^{(q)}$ gives the total corresponding premium ($P_{t+1}^{(q)} = \text{BP}_{t+1} \times \text{BMF}_{t+1}$). For power category "$\geq 119$," her expected claim frequency for the first five periods is 0.3306 and 0.2808 after. The second half of Table 9 shows the evolution of the premium amounts for this policyholder.

Table 10 is similar to Table 9 except an exponential loss function is used. The bonus-malus factors are computed from equation (10) with $c = 12.93$. This parameter has been set in such a way that the variance of the a posteriori premiums paid by a policyholder during the first 10 years represents 50% of the variance if the premiums were computed under a quadratic loss; for more details.[See Denuit and Dhaene (2001).] It is interesting to compare the bonus-malus factors in Tables 9 and 10. Notice that her bonus-malus factors are identical whatever the power of the car but the premiums differ substantially. When an exponential loss is used, the size of the maluses is reduced. Because the system is financially balanced, this implies that the size of the bonuses is also reduced.

Table 9
Bonus-Malus Coefficients and A Posteriori Premiums
Quadratic Loss Function for Policyholder Age 30

Car in Power Category "Power ≤ 53"							
		0 Claim in $(0,t)$		1 Claim in $(0,t)$		2 Claims in $(0,t)$	
t	BP_{t+1}	$\mathrm{BMF}_{t+1}^{(q)}$	$P_{t+1}^{(q)}$	$\mathrm{BMF}_{t+1}^{(q)}$	$P_{t+1}^{(q)}$	$\mathrm{BMF}_{t+1}^{(q)}$	$P_{t+1}^{(q)}$
1	0.1787	0.7963	0.1423	1.7154	0.3065	2.6344	0.4708
2	0.1787	0.6616	0.1182	1.4251	0.2547	2.1887	0.3911
3	0.1787	0.5658	0.1011	1.2189	0.2178	1.8719	0.3345
4	0.1787	0.4943	0.0883	1.0648	0.1903	1.6352	0.2922
5	0.1787	0.4388	0.0784	0.9453	0.1689	1.4517	0.2594
6	0.1518	0.3945	0.0599	0.8499	0.1290	1.3052	0.1981
7	0.1518	0.3584	0.0544	0.7720	0.1172	1.1856	0.1800
8	0.1518	0.3283	0.0498	0.7072	0.1073	1.0860	0.1649
9	0.1518	0.3028	0.0460	0.6524	0.0990	1.0019	0.1521
10	0.1518	0.2811	0.0427	0.6055	0.0919	0.9299	0.1412
Car in Power Category "Power ≥ 119"							
1	0.3306	0.7963	0.2633	1.7154	0.5671	2.6344	0.8709
2	0.3306	0.6616	0.2187	1.4251	0.4711	2.1887	0.7236
3	0.3306	0.5658	0.1871	1.2189	0.4030	1.8719	0.6189
4	0.3306	0.4943	0.1634	1.0648	0.3520	1.6352	0.5406
5	0.3306	0.4388	0.1451	0.9453	0.3125	1.4517	0.4799
6	0.2808	0.3945	0.1108	0.8499	0.2386	1.3052	0.3665
7	0.2808	0.3584	0.1006	0.7720	0.2168	1.1856	0.3329
8	0.2808	0.3283	0.0922	0.7072	0.1986	1.0860	0.3050
9	0.2808	0.3028	0.0850	0.6524	0.1832	1.0019	0.2813
10	0.2808	0.2811	0.0789	0.6055	0.1700	0.9299	0.2611

Table 10
Bonus-Malus Coefficients and A Posteriori Premiums
Exponential Loss Function ($c = 12.93$) for Policyholder Age 30

t	$BP^{(e)}_{t+1}$	0 Claim in $(0,t)$ $BMF^{(e)}_{t+1}$	0 Claim in $(0,t)$ $P^{(e)}_{t+1}$	1 Claim in $(0,t)$ $BMF^{(e)}_{t+1}$	1 Claim in $(0,t)$ $P^{(e)}_{t+1}$	2 Claims in $(0,t)$ $BMF^{(e)}_{t+1}$	2 Claims in $(0,t)$ $P^{(e)}_{t+1}$
Car in Power Category "Power ≤ 53"							
1	0.1787	0.9002	0.1609	1.3505	0.2413	1.8007	0.3218
2	0.1787	0.8207	0.1467	1.2253	0.2190	1.6299	0.2913
3	0.1787	0.7553	0.1350	1.1234	0.2007	1.4915	0.2665
4	0.1787	0.7003	0.1251	1.0384	0.1856	1.3765	0.2460
5	0.1787	0.6533	0.1167	0.9662	0.1727	1.2791	0.2286
6	0.1518	0.6125	0.0930	0.9039	0.1372	1.1953	0.1815
7	0.1518	0.5768	0.0876	0.8496	0.1290	1.1224	0.1704
8	0.1518	0.5452	0.0828	0.8017	0.1217	1.0583	0.1606
9	0.1518	0.5170	0.0785	0.7591	0.1152	1.0013	0.1520
10	0.1518	0.4916	0.0746	0.7210	0.1095	0.9504	0.1443
Car in Power Category "Power ≥ 119"							
1	0.3306	0.9002	0.2976	1.3505	0.4465	1.8007	0.5953
2	0.3306	0.8207	0.2713	1.2253	0.4051	1.6299	0.5388
3	0.3306	0.7553	0.2497	1.1234	0.3714	1.4915	0.4931
4	0.3306	0.7003	0.2315	1.0384	0.3433	1.3765	0.4551
5	0.3306	0.6533	0.2160	0.9662	0.3194	1.2791	0.4229
6	0.2808	0.6125	0.1720	0.9039	0.2538	1.1953	0.3356
7	0.2808	0.5768	0.1620	0.8496	0.2386	1.1224	0.3152
8	0.2808	0.5452	0.1531	0.8017	0.2251	1.0583	0.2972
9	0.2808	0.5170	0.1452	0.7591	0.2132	1.0013	0.2812
10	0.2808	0.4916	0.1381	0.7210	0.2025	0.9504	0.2669

3 Integrated Ratemaking

3.1 Claim Frequency Model

In seminal papers, Dionne and Vanasse (1989, 1992) proposed a bonus-malus system that integrates a priori and a posteriori information on an individual basis. Their system introduces a regression component in the Poisson counting model in order to use all available information in the estimation of accident frequency.

Let us assume that the number of claims K_{it} for the i^{th} policyholder of the portfolio during the year t conforms to a Poisson distribution with mean $\lambda_{I_i(t)}$, where $I_i(t)$ is the index of the risk class occupied by policyholder i in year t. A common problem for count data is that the fits obtained are poor even after allowing for important explanatory variables using the Poisson regression model. This indicates that, conditional upon the explanatory variables included in the final model, the variance of an observation is greater than its mean, implying that the Poisson assumption is incorrect. Most often, this is due to the fact that important explanatory variables may not have been measured and are consequently incorrectly excluded from the regression relationship.

A convenient way to avoid this problem is to introduce a random effect in this model; see, e.g., Pinquet (1999). We assume that K_{it} follows a Poisson distribution with mean $\lambda_{I_i(t)}\Theta_i$, where Θ_i has a gamma distribution but with unit mean, i.e., with parameters (α, α). Then, K_{it} follows a negative binomial law, i.e.,

$$\mathbb{P}\mathrm{r}[K_{it} = k | I_i(t)] = \frac{\alpha + k - 1}{k} \left(\frac{\lambda_{I_i(t)}}{\alpha + \lambda_{I_i(t)}} \right)^k \left(\frac{\alpha}{\alpha + \lambda_{I_i(t)}} \right)^\alpha .$$

We can view Θ_i as representing the impact on the mean claim frequency of all the policyholders' characteristics not taken into account a priori. Let us now derive the a posteriori distribution of Θ_i.

Lemma 1. *If the cdf of* Θ_i *is* $\Gamma(.|\alpha, \alpha)$ *then the cdf of* $[\Theta_i | K_{ij} = k_{ij}, j = 1, 2, \ldots, t]$ *is* $\Gamma(.|\alpha + k_{i\bullet}(t), \alpha + \lambda_{i\bullet}(t))$ *where*

$$\lambda_{i\bullet}(t) = \sum_{j=1}^{t} \lambda_{I_i(j)}.$$

Proof: Bayes Theorem yields

$$
\begin{aligned}
d\mathbb{P}\mathrm{r}[\Theta_i \le \theta | K_{ij} = k_{ij},\ j = 1,2,\dots,t]
&= \frac{\mathbb{P}\mathrm{r}[K_{ij} = k_{ij},\ j = 1,2,\dots,t | \Theta_i = \theta] d\mathbb{P}\mathrm{r}[\Theta_i \le \theta]}{\mathbb{P}\mathrm{r}[K_{ij} = k_{ij},\ j = 1,2,\dots,t]} \\
&= \frac{\theta^{k_{i\bullet}(t)} \exp(-\theta\lambda_{i\bullet}(t))\alpha^\alpha \theta^{\alpha-1}\exp(-\alpha\theta)d\theta}{\alpha^\alpha \int_{\xi\in\mathbb{R}^+} \xi^{k_{i\bullet}(t)+\alpha-1}\exp(-\alpha\lambda_{i\bullet}(t)\xi)d\xi},
\end{aligned}
$$

and the result follows.

In order to estimate the parameter α describing the residual heterogeneity of the portfolio, we use the maximum likelihood method. We maximize

$$
L(\alpha) = \prod_{i=1}^{12}\prod_{k=0}^{\infty}\left\{\frac{\alpha+k-1}{k}\left(\frac{\lambda_i}{\alpha+\lambda_i}\right)^k\left(\frac{\alpha}{\alpha+\lambda_i}\right)^\alpha\right\}^{n_{ik}},
$$

which yields $\hat{\alpha} = 0.8157$.

3.2 A Posteriori Premium Using a Quadratic Loss Function

In the model described in the preceding section, Dionne and Vanasse (1989, 1992) and Gisler (1996) have obtained the following result; it can be seen as a direct consequence of Proposition 4 and its proof is thus omitted.

Proposition 2. *Assuming the cdf of Θ_i is $\Gamma(\alpha,\alpha)$, then under a quadratic loss function, the a posteriori premium for policyholder i is given by*

$$
P_{t+1}^{(q)} = \lambda_{I_i(t+1)} BMF^{(q)}(k_{i\bullet}(t), \lambda_{i\bullet}(t)),
$$

where the bonus-malus coefficient is given by

$$
BMF^{(q)}(k_{i\bullet}(t), \lambda_{i\bullet}(t)) = \frac{\alpha + k_{i\bullet}(t)}{\alpha + \lambda_{i\bullet}(t)} = (1-\rho_q)\times 1 + \rho_q \frac{k_{i\bullet}(t)}{\lambda_{i\bullet}(t)}
$$

with

$$
\rho_q = \frac{\lambda_{i\bullet}(t)}{\alpha + \lambda_{i\bullet}(t)}.
$$

Note that the greater the variance of Θ_i (i.e., the smaller α) the greater ρ_q (i.e., the greater the weight given to the claim history of the policyholder). Moreover, ρ_q is clearly increasing in $\lambda_{i\bullet}$. If $\lambda_{i\bullet}$ is small, as is the case for policies with a high deductibles, then ρ_q is also small. The no-claim discount for such policies is thus also small and, as pointed out by Gisler (1996), the bonus-malus systems are of questionable utility.

3.3 A Posteriori Premium Using Exponential Loss

The use of a quadratic loss function leads to high maluses because of the symmetry of the loss function: overcharges and undercharges are equally penalized. Although theoretically correct, such a system is not accepted by policyholders. It is better to have a model with a parameter controlling the severity of the system. One approach is to incorporate a priori variables in the exponential loss function.

Proposition 3. *Assuming that the cdf of Θ_i is $\Gamma(.|\alpha,\alpha)$, then under an exponential loss with parameter $c > 0$ the a posteriori premium for policyholder i is given by*

$$P_{t+1}^{(e)} = \lambda_{l_i(t+1)} BMF^{(e)}(k_{i\bullet}(t), \lambda_{i\bullet}(t)) \tag{14}$$

where the bonus-malus coefficient is given by

$$BMF^{(e)}(k_{i\bullet}(t), \lambda_{i\bullet}(t)) = (1-\rho_e) \times 1 + \rho_e \times \frac{k_{i\bullet}(t)}{\lambda_{i\bullet}(t)} \tag{15}$$

with

$$\rho_e = \frac{\lambda_{i\bullet}(t)}{c} \ln\left(1 + \frac{c}{\alpha + \lambda_{i\bullet}(t)}\right). \tag{16}$$

Proof: From Lemma 1, we get

$$\mathbb{E}\left[e^{-c\Theta_i} | K_{ij} = k_{ij},\ j = 1,2,\ldots,t\right] = \left(\frac{\alpha + \lambda_{i\bullet}(t)}{\alpha + \lambda_{i\bullet}(t) + c}\right)^{\alpha + k_{i\bullet}(t)}.$$

It follows that

$$\begin{aligned}
&\ln \mathbb{E}\Big[e^{-c\Theta_i}|K_{ij} = k_{ij},\, j = 1,2,\dots,t\Big] \\
&\qquad = -(\alpha + k_{i\bullet}(t)) \ln\left(1 + \frac{c}{\alpha + \lambda_{i\bullet}(t)}\right) \\
&\mathbb{E}\Big[\ln \mathbb{E}\Big[e^{-c\Theta_i}|K_{ij},\, j = 1,2,\dots,t\Big]\Big] \\
&\qquad = -(\alpha + \lambda_{i\bullet}(t)) \ln\left(1 + \frac{c}{\alpha + \lambda_{i\bullet}(t)}\right).
\end{aligned}$$

The result then follows from Proposition 4.

Comparing the bonus-malus coefficients obtained with a quadratic and exponential loss functions we have, for any $c \geq 0$,

$$\ln\left(1 + \frac{c}{\alpha + \lambda_{i\bullet}}\right) \leq \frac{c}{\alpha + \lambda_{i\bullet}},$$

so that $\rho_e(c) \leq \rho_q$; the weight given to past claims is thus smaller under an exponential loss.

It can be shown that $\rho_e(c) \to 0$ as $c \to +\infty$. If the asymmetry factor c tends to $+\infty$ then all the risks within the same tariff class pay the same premium: there is no more experience rating. Conversely, $\rho_e(c) \to \rho_q$ as $c \to 0$. The results obtained by Dionne and Vanasse (1989, 1992) also appear as limit cases of those obtained with an exponential loss function.

3.4 Numerical Illustration

Computing the premiums for a 30-year old female policyholder using Dionne-Vanasse's methodology yields the results in Table 11. Unlike Table 9, the bonus-malus factors are not the same for both categories of car. The differences are explained by the presence of personal characteristics in the calculation of the factors in Table 11. Once the a priori variables are introduced the sizes of the bonuses and the maluses are reduced. Technically, this means that part of the heterogeneity has been taken into account in the a priori differentiation of the premiums, so that the residual heterogeneity is smaller and the magnitude of the a posteriori corrections is reduced.

It is interesting to note that even if a policyholder whose car is in category "Power $\leq$ 53" always pays a smaller premium that the corresponding premium for the driver in category "Power $\geq$ 119," her bonus-malus factors are always greater (i.e., she has less bonuses and more maluses). This is because good risks are rewarded in their base premiums (through the a priori variables incorporated in the tariff). Consequently, the size of bonus they require for equity is reduced. In other words, the premium discount awarded to risks judged as good a priori has to be smaller than the bonus awarded to those judged as bad a priori. Conversely, the penalties assessed to risks judged as good are larger than the penalties assessed to those judged as bad.

The same remarks hold for the bonus-malus coefficients obtained with an exponential loss function presented in Table 12. The severity of the a posteriori corrections is weaker than with a quadratic loss function, as expected.

4 Summary and Conclusions

As was pointed out earlier, the aim of this paper is to examine the interaction between a priori ratemaking (i.e., identification of the best predictors $\mathbf{X}$ and of the risk premium $\mathbb{E}[Y|\mathbf{X}]$) and a posteriori ratemaking (i.e., premium corrections according to the claims history up to time t). To this end, we propose an extension of the exponential bonus-malus systems introduced in Denuit and Dhaene (2001) in the presence of a priori risk classification. The main advantage of this extension is that it provides the actuary with a parameter for controlling the severity of the a posteriori corrections. The actuary is allowed to vary this parameter from one extreme where there is no a posteriori correction to the other extreme where the severity corresponds to the classical quadratic loss function. At the limit, previous results based on a quadratic loss function are thus obtained. The a posteriori corrections also depend on the a priori amount of premium, yielding an integrated ratemaking mechanism recognizing the continuous nature of risk evaluation.

To illustrate our methodology, an example is provided using data from a Spanish insurance portfolio. We show that good risks are rewarded in their base premiums and, consequently, they require a smaller bonus than the bonus awarded to those judged as bad a priori, as expected.

In the future, we purpose to study bonus-malus scales accounting for a priori risk classification in the spirit of Taylor (1997), substituting the exponential loss function for its classical quadratic counterpart.

Table 11
Bonus-Malus Coefficients and A Posteriori Premiums
Quadratic Loss Function for Policyholder Age 30

t	BP_{t+1}	0 Claim in $(0,t)$ $\mathrm{BMF}_{t+1}^{(q)}$	0 Claim in $(0,t)$ $P_{t+1}^{(q)}$	1 Claim in $(0,t)$ $\mathrm{BMF}_{t+1}^{(q)}$	1 Claim in $(0,t)$ $P_{t+1}^{(q)}$	2 Claims in $(0,t)$ $\mathrm{BMF}_{t+1}^{(q)}$	2 Claims in $(0,t)$ $P_{t+1}^{(q)}$
Car in Power Category "Power ≤ 53"							
1	0.1787	0.8203	0.1466	1.8259	0.3263	2.8316	0.5060
2	0.1787	0.6953	0.1243	1.5478	0.2766	2.4002	0.4289
3	0.1787	0.6034	0.1078	1.3432	0.2400	2.0829	0.3722
4	0.1787	0.5330	0.0952	1.1863	0.2120	1.8397	0.3288
5	0.1787	0.4772	0.0853	1.0623	0.1898	1.6474	0.2944
6	0.1518	0.4383	0.0665	0.9757	0.1481	1.5130	0.2297
7	0.1518	0.4053	0.0615	0.9021	0.1369	1.3989	0.2124
8	0.1518	0.3768	0.0572	0.8388	0.1273	1.3008	0.1975
9	0.1518	0.3521	0.0535	0.7838	0.1190	1.2155	0.1845
10	0.1518	0.3305	0.0502	0.7356	0.1117	1.1408	0.1732
Car in Power Category "Power ≥ 119"							
1	0.3306	0.7945	0.2626	1.4162	0.4682	2.0379	0.6737
2	0.3306	0.6590	0.2179	1.1747	0.3884	1.6905	0.5589
3	0.3306	0.5630	0.1861	1.0036	0.3318	1.4442	0.4775
4	0.3306	0.4914	0.1625	0.8760	0.2896	1.2606	0.4168
5	0.3306	0.4360	0.1441	0.7772	0.2569	1.1184	0.3697
6	0.2808	0.3979	0.1117	0.7092	0.1992	1.0206	0.2866
7	0.2808	0.3659	0.1027	0.6522	0.1831	0.9386	0.2635
8	0.2808	0.3387	0.0951	0.6037	0.1695	0.8687	0.2439
9	0.2808	0.3152	0.0885	0.5619	0.1578	0.8085	0.2270
10	0.2808	0.2948	0.0828	0.5255	0.1476	0.7562	0.2123

Table 12
Bonus-Malus Coefficients and A Posteriori Premiums
Exponential Loss Function (c = 12.93) for Policyholder Age 30

Car in Power Category "Power ≤ 53"							
		0 Claim in $(0,t)$		1 Claim in $(0,t)$		2 Claims in $(0,t)$	
t	$\text{BP}_{t+1}^{(e)}$	$\text{BMF}_{t+1}^{(e)}$	$P_{t+1}^{(e)}$	$\text{BMF}_{t+1}^{(e)}$	$P_{t+1}^{(e)}$	$\text{BMF}_{t+1}^{(e)}$	$P_{t+1}^{(e)}$
1	0.1787	0.9635	0.1722	1.1676	0.2087	1.3718	0.2451
2	0.1787	0.9313	0.1664	1.1236	0.2008	1.3159	0.2352
3	0.1787	0.9022	0.1612	1.0846	0.1938	1.2669	0.2264
4	0.1787	0.8758	0.1565	1.0495	0.1876	1.2232	0.2186
5	0.1787	0.8516	0.1522	1.0177	0.1819	1.1838	0.2115
6	0.1518	0.8324	0.1264	0.9927	0.1507	1.1531	0.1750
7	0.1518	0.8144	0.1236	0.9694	0.1472	1.1245	0.1707
8	0.1518	0.7974	0.1210	0.9476	0.1438	1.0978	0.1666
9	0.1518	0.7813	0.1186	0.9270	0.1407	1.0728	0.1628
10	0.1518	0.7660	0.1163	0.9076	0.1378	1.0492	0.1593
Car in Power Category "Power ≥ 119"							
1	0.3306	0.9359	0.3094	1.1298	0.3735	1.3238	0.4377
2	0.3306	0.8835	0.2921	1.0597	0.3503	1.2359	0.4086
3	0.3306	0.8390	0.2774	1.0013	0.3310	1.1636	0.3847
4	0.3306	0.8003	0.2646	0.9513	0.3145	1.1023	0.3644
5	0.3306	0.7660	0.2532	0.9075	0.3000	1.0491	0.3468
6	0.2808	0.7396	0.2077	0.8743	0.2455	1.0089	0.2833
7	0.2808	0.7154	0.2009	0.8439	0.2370	0.9724	0.2731
8	0.2808	0.6931	0.1946	0.8161	0.2292	0.9391	0.2637
9	0.2808	0.6723	0.1888	0.7904	0.2219	0.9084	0.2551
10	0.2808	0.6530	0.1834	0.7665	0.2152	0.8800	0.2471

References

Denuit, M., and Dhaene, J. "Bonus-Malus Scales Using Exponential Loss Functions." *Bulletin of the German Society of Actuaries* 25 (2001): 13–27.

De Wit, G.W., and Van Eeghen, J. "Rate Making and Society's Sense of Fairness." *ASTIN Bulletin* 14 (1984): 151–163.

Dionne, G., and Vanasse, C. "A Generalization of Actuarial Automobile Insurance Rating Models: The Negative Binomial Distribution with a Regression Component." *ASTIN Bulletin* 19 (1989): 199–212.

Dionne, G., and Vanasse, C. "Automobile Insurance Ratemaking in the Presence of Asymmetrical Information." *Journal of Applied Econometrics* 7 (1992): 149–165.

Dobson, A.J. *An Introduction to Generalized Linear Models.* London, England: Chapman & Hall/CRC, 1990.

Fahrmeir, L., and Tutz, G. *Multivariate Statistical Modelling Based on Generalized Linear Models.* New York, N.Y.: Springer Verlag, 1994.

Ferreira, J. "Identifying Equitable Insurance Premiums for Risk Classes: an Alternative to the Classical Approach." At 23rd International Meeting, Institute of Management Sciences, Athens, Greece, 1977.

Gisler, A. "Bonus-Malus and Tariff Segmentation." At 27th ASTIN Colloquium, Copenhagen, Denmark, September 1-5, 1996.

Lemaire, J. "How to Define a Bonus-Malus System with an Exponential Utility Function." *ASTIN Bulletin* 10, (1979): 274–282.

Lemaire, J. *Bonus-Malus Systems in Automobile Insurance.* Boston, Mass.: Kluwer, 1995.

Mc Cullagh, P., and Nelder, J.A. *Generalized Linear Models.* New York, N.Y.: Chapman & Hall, 1989.

Pinquet, J. "Allowance for Cost of Claims in Bonus-Malus Systems." *ASTIN Bulletin* 27 (1997): 33–57.

Pinquet, J. "Experience Rating Through Heterogeneous Models." In *Handbook of Insurance* G. Dionne, Editor. Boston, Mass.: Kluwer, 1999.

Renshaw, A.E. "Modelling the Claim Process in the Presence of Covariates." *ASTIN Bulletin* 24 (1994): 265–285.

Taylor, G. "Setting a Bonus-Malus Scale in the Presence of Other Rating Factors." *ASTIN Bulletin* 27 (1997): 319–327.

Young, V. "Credibility and Persistency." *ASTIN Bulletin* 26 (1996): 53–69.

Appendix: Credibility Models with Quadratic and Exponential Loss Functions

Let us consider a sequence of random variables $\{X_1, X_2, X_3, \dots\}$ and a risk parameter Θ where Θ is a random variable or possibly a sequence of random variables. We assume the sequence of random variables $\{X_1, X_2, X_3, \dots | \Theta\}$ are independent. The first two moments of the X_is are assumed to be finite. Moreover, the conditional mean of the X_is is given by

$$\begin{aligned} \mu_i(\Theta) &= \mathbb{E}[X_i|\Theta] \\ \mathbb{E}[\mu_i(\Theta)] &= \mu_i \end{aligned}$$

for $i = 1, 2, 3, \dots$.

Proposition 4.

(i) *The minimum of* $\mathbb{E}\Big[\mu_{n+1}(\Theta) - \Psi_n(X_1, X_2, \dots, X_n)\Big]^2$ *on all the measurable functions* $\Psi_n : \mathbb{R}^n \to \mathbb{R}$ *is obtained for*

$$\Psi_n^*(X_1, X_2, \dots, X_n) = \mathbb{E}[\mu_{n+1}(\Theta)|X_1, X_2, \dots, X_n].$$

(ii) *The minimum of* $\mathbb{E}\Big[\exp\Big[-c\Big(\mu_{n+1}(\Theta) - \Psi_n(X_1, X_2, \dots, X_n)\Big)\Big]\Big]$ *on all the measurable functions* $\Psi_n : \mathbb{R}^n \to \mathbb{R}$ *satisfying the constraint* $\mathbb{E}[\Psi_n(X_1, X_2, \dots, X_n)] = \mu_{n+1}$ *is obtained for*

$$\begin{aligned} \Psi_n^*(X_1, \dots, X_n) = \mu_{n+1} & \\ &+ \frac{1}{c}\mathbb{E}[\ln \mathbb{E}[\exp(-c\mu_{n+1}(\Theta))|X_1, \dots, X_n]] \\ &- \frac{1}{c}\ln \mathbb{E}[\exp(-c\mu_{n+1}(\Theta))|X_1, \dots, X_n]\Big]. \end{aligned}$$

(This constraint is made in order to guarantee financial equilibrium.)

Proof: (i) This is a classic result, and its proof can be found in many statistical textbooks. An easy way to see it consists in noting that

$$
\begin{aligned}
&\mathbb{E}\Big[\Big(\mu_{n+1}(\Theta) - \Psi_n(X_1,\dots,X_n)\Big)^2 \\
&\quad = \mathbb{E}\Big[\Big(\mu_{n+1}(\Theta) - \Psi_n^*(X_1,\dots,X_n) \\
&\qquad + \Psi_n^*(X_1,\dots,X_n) - \Psi_n(X_1,\dots,X_n)\Big)^2\Big] \\
&\quad = \mathbb{E}\Big[\Big(\mu_{n+1}(\Theta) - \Psi_n^*(X_1,\dots,X_n)\Big)^2\Big] \\
&\qquad + \mathbb{E}\Big[\Big(\Psi_n^*(X_1,\dots,X_n) - \Psi_n(X_1,\dots,X_n)\Big)^2\Big],
\end{aligned}
$$

which is clearly minimal for $\Psi_n \equiv \Psi_n^*$.
(ii) Starting from

$$
\begin{aligned}
&\mathbb{E}\Big[\exp\Big[-c\Big(\mu_{n+1}(\Theta) - \Psi_n(X_1,\dots,X_n)\Big)\Big] \\
&= \mathbb{E}\Big[\Big[\exp\Big[c\Psi_n(X_1,\dots,X_n)\}\mathbb{E}\Big[\exp\Big[-c\mu_{n+1}(\Theta)\Big]|X_1,\dots,X_n\Big]\Big] \\
&= \mathbb{E}\Big[\exp\Big[c(\Psi_n(X_1,\dots,X_n) - \Psi_n^*(X_1,\dots,X_n))\Big] \\
&\qquad \exp\Big[c\mu_{n+1}\Big]\exp\Big[\mathbb{E}\ln\mathbb{E}[\exp\Big[-c\mu_{n+1}(\Theta)\Big]|X_1,\dots,X_n]\Big].
\end{aligned}
$$

Now, let us apply Jensen's inequality to get

$$
\begin{aligned}
&\mathbb{E}\exp\Big[-c\Big(\mu_{n+1}(\Theta) - \Psi_n(X_1,\dots,X_n)\Big)\Big] \\
&\geq \exp\Big[c\mathbb{E}\left[\Psi_n(X_1,\dots,X_n) - \Psi_n^*(X_1,\dots,X_n)\right]\Big] \\
&\qquad \exp\Big[c\mu_{n+1}\Big]\exp\Big[\mathbb{E}\ln\mathbb{E}[\exp\Big[-c\mu_{n+1}(\Theta)\Big]|X_1,\dots,X_n]\Big].
\end{aligned}
$$

Because of the constraint on the expectation of the Ψ_ns, the first exponential is 1, thus completing the proof. □