The efficiency of the Swiss bonus-malus system

B. Wissenschaftliche Mitteilungen

FRANÇOIS DUFRESNE, University of Lausanne, Switzerland

The Efficiency of the Swiss Bonus-malus System

1 Introduction

Most empirical studies found in the actuarial literature on bonus-malus systems in automobile insurance compare the systems of different countries (e.g. Vepsäläinen (1972), Lemaire (1988), Venter (1991)). They consider the rules of the different systems, including transition rules from class to class, premium scales, possible deductibles, etc. And, since the question of the performance of such systems will be eventually raised, they use one or more measures of efficiency to rate the systems among them. In these comparative studies, the Swiss bonus-malus system always gets some of the highest marks.

In 1990, the bonus-malus system of Switzerland has been modified. It was decided to change the transition rule from one class to another but to keep unchanged all the other parameters of the system. In short, in the new system, a claim is now penalized by increasing the premium class of the driver by four instead of by three. Of course, the driver cannot pay more than the premium of the 22th class which is the highest class of the system. The rules of the Swiss bonus-malus system will be summarized in the next section.

The aim of this paper is to study the effect of such a change in the transition rules on the efficiency of the Swiss bonus-malus system. Three measures of efficiency are used: the asymptotic efficiency defined by Loimaranta (1972), the efficiency of the 2nd kind according to Lemaire (1985), and the predictive accuracy suggested by Venter (1991). These are mathematical concepts of efficiency and address only one aspect of the performance of bonus-malus systems, namely their aptitude at discriminating among the good and the bad risks.

In addition to the step sizes, 3 and 4, that have really been used in the Swiss bonus-malus system, the computations are carried for smaller and bigger step sizes. This shows how the different efficiency measures vary as functions of the step size and what would be the "improvement" (if any) of a still bigger step size in the real system.

2 The Swiss Bonus-malus System

For a particular type of car (for example, a passenger car with an engine having a capacity between 1393 and 2963 cubic centimeters), there are 22 premium classes or "states", to use the terminology of Markov chains. Table 1 shows the premium scale in percentage of the basic premium, that is the premium of class 9 (the 10th class). Let b_x be the percentage of the basic premium corresponding to state x.

Table 1: Premium Scale in %

State	Premium	State	Premium
0	45	10	110
1	50	11	120
2	55	12	130
3	60	13	140
4	65	14	155
5	70	15	170
6	75	16	185
7	80	17	200
8	90	18	215
9	100	19	230
		20	250
		21	270

A new car owner enters the system at state 9 and, thus, pays for his first year the basic premium. The following rule applies to each of the subsequent years: if the previous state was x, the new state will be

$x - 1$ if no claims were reported

$x + n \cdot s$ if n claims were reported,

subject to the condition that the state cannot be lower than zero or higher than 21. Hence after a claim free year the state is reduced by one and for each reported claim the state is increased by s. The number s used to be equal to 3 but since 1990 it is now set to 4. The condition means that a policyholder who is in state 0 (maximal bonus) stays there after a claim free year. And, in the case of very bad or unlucky drivers, if the state was x and the number of claims n is such that $x + n \cdot s \geq 21$, the new state will be 21.

3 The Stationary Distribution

For each policyholder, the sequence of states forms a Markov chain. The transition probabilities may differ from one policyholder to the other but it is assumed (for simplicity) that they do not vary in time. If the actual (total) step in the premium scale for year $t+1$ is denoted by Y_{t+1}, then it is defined by

$$Y_{t+1} = \begin{cases} -1 & \text{if no claims} \\ s \cdot n & \text{if } n \text{ claims}. \end{cases} \tag{1}$$

It is assumed that $Y_1, Y_2, \ldots$ are mutually independent and identically distributed according to the probability function

$$q(y) = \Pr[Y_t = y], \qquad y = -1, s, 2s, 3s, \ldots \tag{2}$$

If X_{t+1} is the state of the policyholder at time $t+1$, then according to the previous section, X_{t+1} can be defined as follows:

$$X_{t+1} = \begin{cases} X_t + Y_{t+1} & \text{if } 0 \leq X_t + Y_{t+1} \leq 21, \\ 0 & \text{if } X_t + Y_{t+1} = -1, \\ 21 & \text{if } X_t + Y_{t+1} > 21. \end{cases} \tag{3}$$

It is easily shown (see Dufresne (1988)) that the distribution function of X_{t+1} satisfies the relation

$$F(x, t+1) = \sum_{y=-1}^{x} F(x-y, t) \cdot q(y), \qquad x = 0, 1, \ldots, 21, \tag{4}$$

with $F(21, t+1) = 1$. The stationary distribution function $F(x)$ is given by

$$\begin{aligned} F(x) &= \lim_{t \to \infty} F(x, t) \\ &= \sum_{y=-1}^{x} F(x-y) \cdot q(y), \qquad x = 0, 1, \ldots, 21, \end{aligned} \tag{5}$$

with $F(21) = 1$, and can be computed recursively with the following algorithm

(Dufresne (1988)):

1. Set $A(0) = 1$.
2. Compute for $x = 0, 1, 2, \ldots, 20$:

$$A(x+1) = \frac{1}{q(-1)}\left\{A(x) - \sum_{y=0}^{x} A(x-y)\cdot q(y)\right\}. \tag{6}$$

3. Set $F(x) = \frac{A(x)}{A(21)}$ for $x = 0, 1, 2, \ldots, 21$.

One of the efficiency measures considered requires the calculation of stationary distributions. The recursive formula above is computationally efficient but has also the virtue of giving some insight about the analytical form of the stationary distribution and, consequently, providing some information on the efficiency measure itself.

4 The Portfolio Model

It will be assumed that, for a policyholder with expected claim frequency of λ (per year), the probability that he or she has n claims follows a Poisson distribution of parameter λ. That is,

$$\Pr[N_t = n] = \frac{e^{-\lambda}\lambda^n}{n!}, \qquad n = 0, 1, 2, \ldots, \tag{7}$$

where N_t is the claim number random variable.
For a policyholder taken at random from a portfolio of automobile insurance, the Poisson parameter is unknown and considered as a random variable Λ. It will be assumed as in Lemaire (1988) and Venter (1991) that the distribution of Λ is a Gamma distribution with parameters $\alpha = \frac{10}{7}$ and $\beta = \frac{100}{7}$. Thus N_t has a Negative binomial distribution with

$$E[N_t] = \frac{\alpha}{\beta} = 0.1 \tag{8}$$

and

$$\mathrm{Var}[N_t] = \frac{\alpha}{\beta} + \frac{\alpha}{\beta^2} = 0.107, \tag{9}$$

since the conditional distribution of N_t given $\Lambda = \lambda$ is Poisson of parameter λ. If the stationary and the transient distribution functions, given $\Lambda = \lambda$, are denoted $F^{(\lambda)}(x)$ and $F^{(\lambda)}(x,t)$ respectively, the unconditional corresponding distributions are given by

$$F(x) = \int_0^\infty F^{(\lambda)}(x)\,dU(\lambda) \tag{10}$$

and

$$F(x,t) = \int_0^\infty F^{(\lambda)}(x,t)\,dU(\lambda) \tag{11}$$

where $U(\lambda)$ is the Gamma distribution function.
A bayesian point of view has been adopted here, but an "empirical Bayes" approach would lead to the very same mathematics. Perhaps in this latter case it should be mentioned that the portfolio is closed: no new entrants and no exits.
Under these assumptions, the probability function $q(y)$ is given, when $s = 3$, by

$$\begin{aligned} &q(-1) = e^{-\lambda}, \qquad q(0) = q(1) = q(2) = 0\,, \\ &q(3) = \lambda e^{-\lambda}, \qquad q(4) = q(5) = 0\,, \qquad q(6) = \lambda^2 e^{-\lambda}/2, \text{ etc.} \end{aligned} \tag{12}$$

The application of the recursive procedure of section 3 to this (conditional) Poisson case produces the following analytical form for the auxiliary function

$$A^{(\lambda)}(x) = \sum_{j=0}^{x}\sum_{i=0}^{x} C_{ij}(x)\lambda^i e^{j\lambda}\,, \qquad x = 0, 1, \ldots, 21\,, \tag{13}$$

where the $C_{ij}(x)$ are constant with respect to λ (they depend on x, i, j and s). These constants are most easily obtained with the use of a symbolic algebra computer software. The analytical form of the auxiliary function $A^{(\lambda)}(x)$ is interesting in itself and may have some computational advantages. The conditional asymptotic distribution function $F^{(\lambda)}(x)$ is given by the ratio $A^{(\lambda)}(x)/A^{(\lambda)}(21)$, $x = 0, 1, \ldots, 21$. The backward differences $F^{(\lambda)}(x) - F^{(\lambda)}(x-1)$ give the conditional asymptotic probability function $f^{(\lambda)}(x)$.

The asymptotic distribution and probability functions for different values of the step size are presented in Table 2 and 3 respectively. A numerical integration technique was used to evaluate (10). It is easily seen that the distribution tends to spread out as s increases. The percentage of drivers enjoying the maximal bonus decreases steadily. This percentage is 66.8% for the old system and 58.9% for the new one (according to our model).

Table 2: Stationary Distribution Functions (in percent)

state	step size						
x	$s = 1$	$s = 2$	$s = 3$	$s = 4$	$s = 5$	$s = 6$	$s = 7$
0	88.0	76.6	66.8	58.9	52.6	47.6	43.7
1	96.5	83.0	71.7	62.7	55.7	50.2	45.9
2	98.6	90.3	77.1	66.9	59.1	53.1	48.4
3	99.2	92.6	83.2	71.6	62.8	56.2	51.0
4	99.5	94.5	85.4	76.7	66.9	59.5	53.9
5	99.7	95.5	87.4	78.8	71.4	63.2	57.1
6	99.7	96.2	89.2	80.8	73.4	67.3	60.5
7	99.8	96.8	90.3	82.7	75.3	69.2	64.3
8	99.8	97.2	91.3	84.4	77.3	71.2	66.2
9	99.8	97.5	92.2	85.6	79.1	73.2	68.2
10	99.9	97.8	92.9	86.9	80.9	75.1	70.3
11	99.9	98.0	93.6	88.0	82.3	77.1	72.4
12	99.9	98.2	94.2	89.1	83.8	79.0	74.5
13	99.9	98.4	94.8	90.1	85.3	80.8	76.8
14	99.9	98.6	95.3	91.1	86.7	82.6	79.0
15	99.9	98.7	95.9	92.1	88.2	84.5	81.2
16	99.9	98.9	96.4	93.2	89.8	86.5	83.6
17	99.9	99.1	97.0	94.3	91.4	88.7	86.2
18	100.0	99.3	97.7	95.5	93.2	91.0	89.0
19	100.0	99.5	98.3	96.8	95.2	93.6	92.2
20	100.0	99.7	99.1	98.3	97.4	96.6	95.8
21	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 3: Stationary Distribution Functions (in percent)

state	step size						
x	$s = 1$	$s = 2$	$s = 3$	$s = 4$	$s = 5$	$s = 6$	$s = 7$
0	88.0	76.6	66.8	58.9	52.6	47.6	43.7
1	8.5	6.4	4.9	3.8	3.1	2.6	2.3
2	2.0	7.3	5.4	4.2	3.4	2.8	2.5
3	0.7	2.3	6.1	4.6	3.7	3.1	2.7
4	0.3	1.9	2.2	5.1	4.1	3.4	2.9
5	0.1	1.0	2.0	2.1	4.5	3.7	3.1
6	0.1	0.8	1.8	2.0	2.0	4.0	3.4
7	0.0	0.5	1.1	1.9	2.0	1.9	3.8
8	0.0	0.4	1.0	1.7	1.9	2.0	1.9
9	0.0	0.3	0.9	1.3	1.9	2.0	2.0
10	0.0	0.3	0.7	1.2	1.7	2.0	2.1
11	0.0	0.2	0.7	1.1	1.5	2.0	2.1
12	0.0	0.2	0.6	1.1	1.5	1.9	2.2
13	0.0	0.2	0.6	1.0	1.5	1.7	2.2
14	0.0	0.2	0.6	1.0	1.5	1.8	2.3
15	0.0	0.2	0.6	1.0	1.5	1.9	2.2
16	0.0	0.2	0.6	1.1	1.5	2.0	2.4
17	0.0	0.2	0.6	1.1	1.6	2.2	2.6
18	0.0	0.2	0.6	1.2	1.8	2.3	2.8
19	0.0	0.2	0.7	1.3	2.0	2.6	3.2
20	0.0	0.2	0.8	1.5	2.2	3.0	3.6
21	0.0	0.3	0.9	1.7	2.6	3.4	4.2

5 Efficiency

In the next sections, three measures of efficiency are considered: the asymptotic efficiency defined by Loimaranta (1972), the efficiency of the 2nd kind according to Lemaire (1985), and the predictive accuracy suggested by Venter (1991).
If we know the Poisson parameter λ of the policy holder, then his asymptotic efficiency $\eta(\lambda)$, defined in the next section, can be computed. But this parameter is usually unknown and the measure of interest is then the expected asymptotic

efficiency η obtained by conditioning:

$$\eta = \int_0^\infty \eta(\lambda)\, dU(\lambda)\,. \tag{14}$$

The same argument applies to the efficiency of the 2nd kind $\mu_i(\lambda)$ which depends also on the starting class. The expected efficiency of the 2nd kind will be denoted by μ_i. The predictive accuracy $\nu(\lambda)$ will have an expected value (over the portfolio) of ν.

6 Asymptotic Efficiency

In the long run, a policyholder should pay an average premium proportional to his expected frequency. This assertion relies on the fact that the severity of the claims is not taken into account (or would be uncorrelated with the frequency). It relies also on the assumption that there is no solidarity expected between the insureds.
The expected long run premium as a function of the Poisson parameter, denoted by $b(\lambda)$, can be expressed by

$$b(\lambda) = \sum_{x=0}^{21} f^{(\lambda)}(x) b_x\,. \tag{15}$$

It is then required that the expected long run premium should be proportional to the frequency. In other words, $db(\lambda)/b(\lambda)$ should be equal to $d\lambda/\lambda$. This motivated the definition of the asymptotic efficiency $\eta(\lambda)$ of Loimaranta (1972):

$$\eta(\lambda) = \frac{\lambda}{b(\lambda)} \frac{db(\lambda)}{d\lambda}\,. \tag{16}$$

A perfectly efficient bonus-malus system would have an asymptotic efficiency of 1 for all values of λ, but that is impossible (at least for discrete state bonus-malus systems).
Figure 1 shows the asymptotic efficiency (as a function of the Poisson parameter λ) for different values of the step size s ($s = 1,2, \ldots, 7$). For $s = 1$ or

$s = 2$, the hypothetical bonus-malus system would have a very high efficiency (even greater than one) but for very unlikely (high) values of the Poisson parameter. As s increases, the maximal efficiency (for a given s) decreases. The mean asymptotic efficiency, $E[\eta(\Lambda)]$, which is also a function of s, attains a maximum for $s = 6$. This is shown in Table 4. It should be observed that the mean asymptotic efficiency of the new system ($s = 4$) is noticeably higher than the one of the old system ($s = 3$).

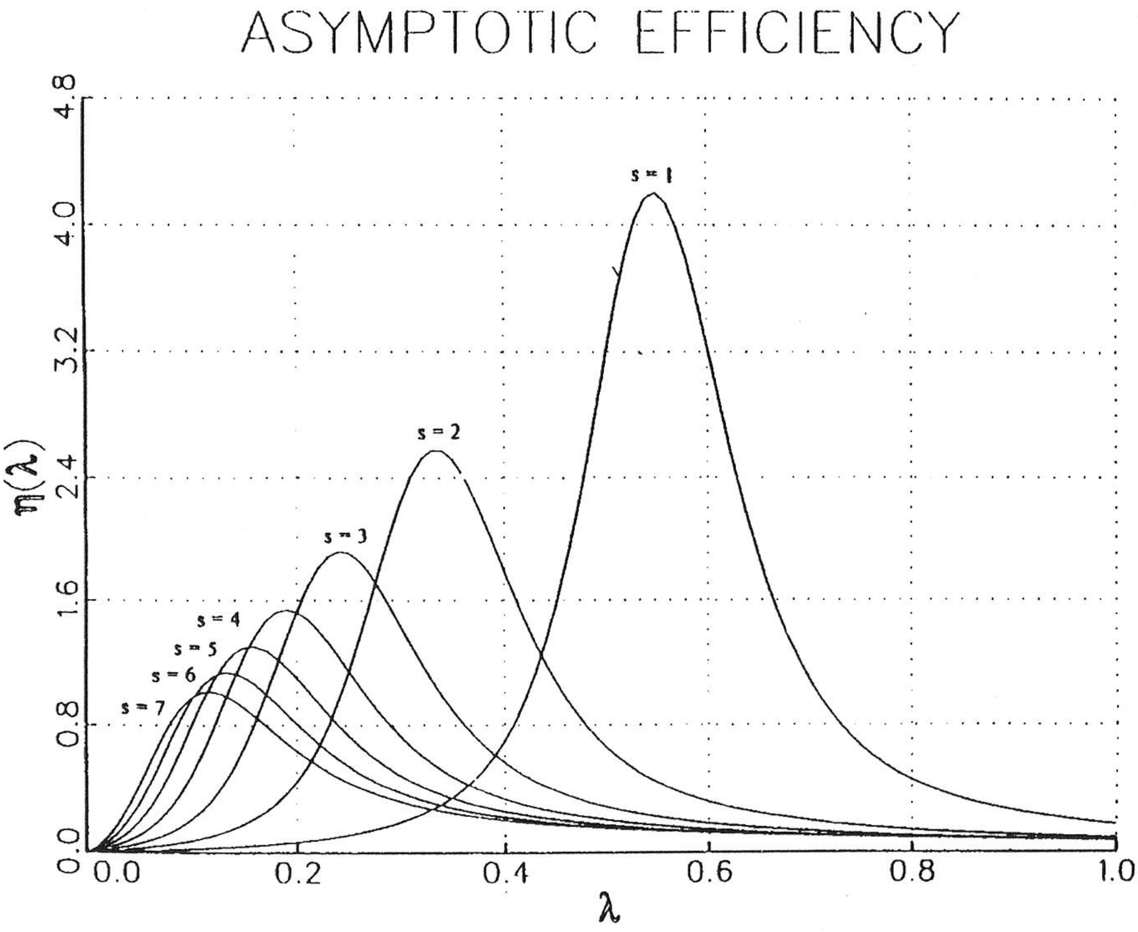

Figure 1

Table 4: Mean asymptotic efficiency

s	$E[\eta(\Lambda)]$
1	0.0462
2	0.2130
3	0.3807
4	0.4861
5	0.5382
6	0.5567
7	0.5565

7 Efficiency of the 2nd Kind

The efficiency of the 2nd kind is based on the expected total discounted premium paid by a policyholder who is at a given date in the state i, $i = 0, 1, \ldots, 21$. If we denote this quantity by $s_i(\lambda)$ and by $\underline{s(\lambda)}$ the vector $(s_0(\lambda), s_1(\lambda), \ldots, s_{21}(\lambda))^T$, it easily shown (see Lemaire (1985)) that $\underline{s(\lambda)}$ satisfies equation

$$\underline{s(\lambda)} = \underline{b} + vM^{(\lambda)}\underline{s(\lambda)} \tag{17}$$

where $M^{(\lambda)}$ is the matrix of transition probabilities of the Markov chain associated with the system and v is a discount factor corresponding to an appropriate rate of interest. Equation (17) represents a system of linear equation in $s_i(\lambda)$, $i = 0, 1, \ldots, 21$, whose solution, in vector notation, is

$$\underline{s(\lambda)} = (I - vM^{(\lambda)})^{-1}\underline{b} \tag{18}$$

Again, for a given initial (or current) state i, $s_i(\lambda)$ should be proportional to λ or, equivalently, $ds_i(\lambda)/s_i(\lambda)$ should be equal to $d\lambda/\lambda$. This motivates the definition of the efficiency of the 2nd kind, a function of λ and i (a given initial or current state):

$$\mu_i(\lambda) = \frac{\lambda}{s_i(\lambda)} \frac{ds_i(\lambda)}{d\lambda}, \qquad i = 0, 1, \ldots, 21\,. \tag{19}$$

Table 5 presents the values of this efficiency measure for $i = 9$, the actual starting class of the swiss bonus-malus system, and a discount factor $v = 1/1.06$.

In the new system, the mean asymptotic efficiency is 0.3235 compared to 0.2610 in the old system, a substantial increase if one notes that the maximal value seems to be about 0.4.

Table 5: Mean asymptotic efficiency of the 2nd kind for class 9

s	$E[\mu_9(\Lambda)]$
1	0.0745
2	0.1709
3	0.2610
4	0.3235
5	0.3610
6	0.3813
7	0.3903

8 Predictive Accuracy

Another measure of the efficiency of a bonus-malus system has been suggested by Venter (1991). He called it *predictive accuracy*. The idea is the following: a bonus-malus system is good at discriminating among the good and the bad risks if the premium they pay is close to their "*true*" premium. This suggests that, for a given period of time, the following expression is a measure of the discriminating power of a bonus malus-system:

$$\nu(\lambda) = \frac{1}{n} \sum_{t=0}^{n-1} \sum_{x=0}^{21} f^{(\lambda)}(x,t)(\lambda - \tilde{b}_x)^2 \tag{20}$$

where $\tilde{b}_x$ is the rescaled premium for class x. The premiums b_x must be rescaled in order that, over a given period of time, the expected average premium paid by an insured is equal to the expected number of claims per period. In other words, we define $\tilde{b}_x = c \cdot b_x$ where c is a constant such that

$$\frac{1}{n} \sum_{t=0}^{n-1} \sum_{x=0}^{21} f(x,t) \cdot \tilde{b}_x = E[N_1], \tag{21}$$

which is 0.1 in our numerical example. If such a rescaling is not done (as it seems to be the case in Venter (1991)), a built-in bias is introduced: it is well

known that the basic premium (class 9) is not the average premium. Over any period of time (of more than one year), the average premium is less than the basic premium and that means that there is a penalty for newcomers.
It should be noted that the lower the measure given by (20) is, the better the system is from the point of view of the discriminating power.
Table 6 shows the predictive accuracy of the Swiss-bonus malus system under the assumption that over each of these periods of time (10 years, 20 years, etc.), the premiums are rescaled according to (21) and the preceding considerations. The figures show that the Swiss bonus-malus system has had an improved discriminating power since the introduction of the new rule in 1990 (step size of 4 vs step size of 3).

Table 6: Predictive accuracy ($\times$ 10 000)

Number of Years	Step size s								
	1	2	3	4	5	6	7	8	9
10	65.2	60.1	57.0	56.0	56.2	57.1	58.1	59.2	60.1
20	65.9	55.9	50.7	49.2	49.4	50.2	51.3	52.2	53.1
30	66.4	54.0	47.5	45.6	45.6	46.3	47.3	48.4	49.4
40	66.4	52.7	45.4	43.2	43.1	43.9	45.0	46.1	47.2
50	66.3	51.7	44.0	41.6	41.5	42.3	43.4	44.6	45.8
60	66.1	51.0	42.9	40.5	40.3	41.2	42.3	43.6	44.8

Table 7: Average premium per period (unrescaled)
(basic premium is one unit)

Number of Years	Step size s								
	1	2	3	4	5	6	7	8	9
10	0.741	0.777	0.818	0.862	0.904	0.946	0.984	1.016	1.045
20	0.613	0.663	0.727	0.794	0.859	0.919	0.972	1.018	1.059
30	0.564	0.616	0.688	0.766	0.841	0.910	0.971	1.024	1.070
40	0.538	0.591	0.667	0.751	0.833	0.907	0.972	1.028	1.077
50	0.523	0.575	0.654	0.742	0.828	0.905	0.973	1.032	1.082
60	0.513	0.565	0.646	0.736	0.824	0.904	0.974	1.034	1.086

Table 7 shows the average premium for different periods of time if the basic premium were of 1 unit. It is seen that for almost all step sizes considered except the unreasonably big step sizes of 8 and 9, the average premium is noticeably lower than the basic premium.

9 Conclusion

The Swiss bonus-malus system has improved its efficiency significantly with the modification introduced in 1990. The three measures considered revealed this. Of course, the goal of a bonus-malus system is not to be optimally efficient in the sense that was used in this paper. Other elements of the problem must be considered and analyzed, in particular the bonus hunger.

Acknowledgment

The author is grateful to Professor Rob Kaas for his valuable comments.

References

Dufresne, F. (1988) Distributions stationnaires d'un système bonus-malus et probabilité de ruine, ASTIN Bulletin **18**, p. 31-46.

Gerber, H.U. (1990) The New Bonus-Malus System of Switzerland, Paper presented at the XXII ASTIN Colloquium in Montreux, p. 2.27-2.29.

Lemaire, J. (1976) Driver versus Company, Scand. Actuarial J., p. 209-219.

Lemaire, J. (1982) L'assurance automobile: problèmes, Labor, Bruxelles, 178 pp.

Lemaire, J. (1985) Automobile Insurance, Kluwer-Nijhoff, Boston, 248 pp.

Lemaire, J. (1988) A comparative Analysis of Most European And Japanese Bonus-malus Systems, Journal of Risk and Insurance **LV**, p 660-681.

Loimaranta, K. (1972) Some asymptotic properties of bonus systems, ASTIN Bulletin **6**, p. 233-245.

Venter, G.G. (1991) A comparative Analysis of Most European And Japanese Bonus-malus Systems: Extension, Journal of Risk and Insurance **LVIII**, p. 542-547.

Vepsäläinen, S. (1972) Applications to a theory of bonus systems, ASTIN Bulletin **6**, p. 212-221.

François Dufresne
Ecole des HEC
University of Lausanne
CH-1 015 Lausanne, Switzerland

Zusammenfassung

Die Effizienz des (neuen) Schweizer Bonus Malus Systems wird anhand von drei Kriterien analysiert. Bei zwei Effizienzmassen wird die stationäre Verteilung über den Prämienstufen benötigt. Schliesslich wird deren analytische Form hergeleitet.

Résumé

L'auteur analyse l'efficacité du (nouveau) système bonus-malus suisse selon trois critères. Il déduit aussi la forme analytique la distribution stationnaire du système à partir de la formule récursive qui permet de la calculer. Cette distribution stationnaire intervient dans deux des mesures d'efficacité considérées.

Summary

The efficiency of the (new) Swiss bonus-malus system is analyzed according to three efficiency measures. In addition, the analytical form of the stationary distribution of the system, which is involved in two of the efficiency measures, is obtained as a byproduct of its recursive calculation scheme.