AN APPLICATION OF LINEAR PROGRAMMING
TO BONUS MALUS SYSTEM DESIGN

BY

ANTONIO HERAS, JOSÉ A. GIL, PILAR GARCÍA-PINEDA AND JOSÉ L. VILAR[*]

ABSTRACT

The purpose of this paper is to show how linear programming methodology can help us to design Bonus-Malus premium scales with some interesting theoretical and practical attributes. Examples of these properties are the financial equilibrium of the system, the monotonicity and proper variability of the premium scale, and the improvement of some efficiency measures such as the RSAL and the elasticity of the system. We will conclude that the use of the linear programming methodology makes possible a high degree of interaction between the designer and the mathematical model.

KEYWORDS

Bonus-Malus systems, Linear Goal Programming, Premium variability, RSAL, Elasticity.

1. INTRODUCTION

It is well known that a priori rating techniques may not eliminate the risk heterogeneity within the policyholders' classes, because some of the most important risk factors may be unobservable. This fact forces many insurance companies to adopt bonus-malus rating systems, in order to adjust the premium to the policyholders claims experience.

Bonus-Malus systems (BMS) design requires to choose the number of classes, the transition rules between them, the initial class and the scale of premiums. It is clear that all these decisions should be optimal choices. In this paper we will refer only to the last problem, the calculation of the optimal set of premiums. The design of *optimal bonus-malus premium scales* for a given set of classes and transition rules has been addressed in the literature as early as in 1963 (see Pesonen (1963)), proposing that the premium for a given bonus-malus class should be defined as the expected claim amount per year of an infinitely old policy in that class. In Norberg (1976), it was proved that Pesonen's premiums minimize the *expected squared rating error* for a randomly chosen

---

[*] The authors are affiliated to Universidad Complutense de Madrid, Spain.

infinitely old policy. The premium scale obtained this way, called the ***Bayes Scale***, is the standard method for optimal BMS design and enjoys good properties, such as the ***financial equilibrium*** of the resulting BMS.

The importance of the financial equilibrium property is beyond all doubt, as it guarantees the equivalence between expected premiums and expected claims in the whole portfolio, and therefore the survival of the company in the long run. Nevertheless, financial equilibrium may be not the only BMS interesting property. In practice, it may be advisable that BMS enjoy other attributes. For example, BMS designers may wish to obtain a smoothed BMS if they want to prevent the bonus hunger phenomenon or even that policyholders may be tempted to quit the company. This can be done by means of the inclusion in the BMS of some desirable commercial features, such as the selection of the initial class and premium, upper (or lower) bounds for the differences between two consecutive premiums, or for the difference between the two extreme premiums, etc.

On the other hand, there exist in the actuarial literature several criteria that measure the degree of fulfilment of other theoretical reasonable properties for a given BMS. Examples of such criteria are the ***relative stationary average level*** (***RSAL***) and the ***elasticity*** or ***efficiency*** of a BMS (Loimaranta (1972), De Pril (1978), Lemaire (1985, 1995, 1998)). According to each one of these criteria it is possible to evaluate two different existing BMS and to conclude that one of them is "better" than the other. It could therefore be interesting to design a BMS with appropriate values of these criteria, or at least with improved valuations when compared with a pre-existing BMS.

In a recent paper (Heras, Vilar and Gil (2002)), we have proposed an alternative method for BMS design, based on the optimization of the ***global asymptotic fairness*** of the system. This objective function measures the global deviation between insured's mean premiums and mean claim frequencies, once the steady-state of the BMS is approached, and is mathematically represented by means of a linear function. Then the optimal premium scale is obtained as the optimal solution of some linear program.

The purpose of this paper is to show how this linear programming methodology becomes a useful tool when we have to design a BMS with desirable commercial and/or theoretical features, like for instance those exposed in the previous paragraphs (which can be easily represented by means of linear constraints to be added to the main linear program). Of course, the new linear program will often be unfeasible. In such cases the designers will know the impossibility of the existence of a BMS satisfying all the required properties, and will proceed to make some of them less severe. In fact, if the designers cannot satisfy all their aspirations about the good characteristics of the BMS, at least they can aim to get a reasonable degree of compromise between them.

2. Definition of a Bonus-Malus System

Consider a group of policies which is homogeneous with respect to some observable risk characteristics. Nevertheless, there remain risk differences within the group, due to unobservable factors. As it is usual in the literature, we assume

that the risk characteristics of each policy are summarized in the value of a parameter Λ, and that the claim numbers from different years are conditionally independent and identically distributed given the risk parameter of the policy. We also assume that the individual claim amounts are independent of the claim numbers and the risk parameter, and mutually independent and identically distributed.

Following several authors (see for instance Lemaire (1985, 1995)), we identify the value of the risk parameter of a policy with its mean claim frequency. This mean claim frequency is assumed to be stationary in time, i.e. not time dependent. In this case, taking the mean individual claim cost as one monetary unit, our objective is to calculate a pure premium for the insured as close as possible to the (unknown) true value of his parameter. We will try to perform this by means of a BMS. Of course, such system will be based on the number of claims and not on their amount. In fact, almost all the real BMS around the world are exclusively based on the number of claims (Lemaire (1985, p. 129)).

Finally, we assume that the risk parameter Λ is a random variable with known cumulative function $U(\lambda)$ (the **structure function**). This distribution is not a subjective distribution in the pure Bayesian sense, it has a frequency interpretation as different policies will have different values of their risk parameters.

Following Lemaire (1995, p. 6), we say that an insurance company uses a BMS when the following conditions hold:

- There exists a finite number of classes $(C_1, \ldots, C_n)$ such that each policy stays in one class throughout each insurance period (usually a year).
- The premium for each policy depends only on the class where it stays.
- The class for a given period is determined by the class in the preceding period and the number of claims reported in that period (**Markovian Condition**).

Every BMS is determined by three elements:

- The **initial class**, where new policies are assigned.
- The **premium scale** $\bar{P} = (P_1, \ldots, P_n)$, where P_i is the premium for policies in class C_i.
- The **transition rules**, that is, the rules establishing the conditions under which a policy in class C_i is transferred to class C_j in the next period.

These transition rules are usually defined by means of transformations T_k such that $T_k(i) = j$ when policyholders in class C_i reporting k claims are transferred to class C_j in the next period. Transformations T_k are usually defined by means of matrices,

$$T_k = (t_{ij}^k)$$

where

$$t_{ij}^k = 1 \text{ if } T_k(i) = j$$
$$t_{ij}^k = 0 \text{ if } T_k(i) \neq j$$

The **conditional transition probability** from C_i to C_j in one period, given that $\Lambda = \lambda$, can be calculated as

$$p_{ij}(\lambda) = \sum_{k=0}^{\infty} p_k(\lambda) \cdot t_{ij}^k$$

where $p_k(\lambda)$ is the conditional probability of reporting k claims in one period given that $\Lambda = \lambda$, that is,

$$p_k(\lambda) = \Pr[N = k \,/\, \Lambda = \lambda]$$

The **conditional transition matrix**, given that $\Lambda = \lambda$, is defined as

$$P(\lambda) = \big(p_{ij}(\lambda)\big)$$

These definitions allow us to look at the bonus-malus system as a **Markov chain**. This chain is **homogeneous**, since we have assumed that each claim frequency λ is stationary in time. The previously defined transition matrix $P(\lambda)$ will be the transition matrix of the corresponding Markov chain.

If we also assume that the Markov chain is ergodic and not cyclic then it is a regular chain, and it is well known (see for instance Kemeny and Snell (1976)) that there exists a **stationary (conditional) probability distribution** $\bar{\pi}(\lambda) = \big(\pi_1(\lambda),\ldots,\pi_n(\lambda)\big)$, where $\pi_i(\lambda)$ is defined as the limit value (when the number of periods tends to infinity) of the conditional probability that a policy belongs to the class C_i, given that $\Lambda = \lambda$. It can be shown that the stationary probability distribution coincides with the left eigenvector associated with the eigenvalue 1 of the corresponding transition matrix, and whose components add the unit.

It is also possible to define the stationary (unconditional) probability distribution $\bar{\pi} = (\pi_1,\ldots,\pi_n)$ for an arbitrary policy, as the mean value of the stationary conditional probability distributions $\big(\pi_1(\lambda),\ldots,\pi_n(\lambda)\big)$. That is,

$$\pi_i = \int \pi_i(\lambda) dU(\lambda)$$

It is clear that π_i and $\pi_i(\lambda)$ can be interpreted as the probabilities that an arbitrary policy and a policy conditioned to $\Lambda = \lambda$, respectively, belong to class C_i when **stationarity** is reached. The knowledge of these stationary distributions becomes a very useful tool when designing a BMS, because it informs us about the long term distribution of the policies. In fact, we will add to our previous set of hypotheses the assumption that the BMS has reached, or at least approached, its steady state.

3. Asymptotic Criteria for Optimal BMS Design

There are two main approaches for designing an optimal BMS based on the long term behaviour of the policies (when the system approaches the steady state). Both of them aim to minimize the expected rating error, that is, the

expected difference between premiums and claims for an arbitrary policy. But they differ in the way they calculate such a rating error.

The first approach considers an arbitrary policy in an arbitrary class when the steady state has been reached. If the mean claim frequency is λ and the class is C_i, then the rating error should be defined as a function of the difference $(P_i - \lambda)$. As mentioned in the Introduction, Norberg's approach (1976) follows this idea by considering a quadratic rating error function $(P_i - \lambda)^2$. Since the density function of the random variable (λ, C_i) can be calculated as $\pi_i(\lambda)u(\lambda)$, the expected squared rating error to be minimized will be

$$\int_0^\infty \sum_{i=1}^n (P_i - \lambda)^2 \pi_i(\lambda)u(\lambda)d\lambda$$

It is easy to show that the optimal premiums are, for $i = 1,\ldots,n$,

$$P_i = \frac{1}{\pi_i} \int_0^\infty \lambda\pi_i(\lambda)u(\lambda)d\lambda$$

that is, the expected claim amount of an arbitrary policy in each class (Pesonen (1963)). This premium scale is usually called the Bayes Scale. (See Baione, Levantesi and Menzietti (2002), for a similar (not identical) quadratic rating error function).

The second approach is based on the fact that when an arbitrary policy reaches the steady state, it does not necessarily stay forever in a fixed class: it is the probability $\pi_i(\lambda)$ of (temporarily) belonging to each class C_i that remains constant for a policy with risk parameter λ in the stationary state. The policy can change from one class to another along the class scale according to these probabilities, and therefore the mean value of the premiums paid by that policyholder will be (see Lemaire (1985), p. 166):

$$\sum_{i=1}^n P_i \pi_i(\lambda) = \overline{P}\overline{\pi}(\lambda)'$$

According to this approach, the rating error for an arbitrary policy should be defined as a function of the differences

$$\left(\sum_{i=1}^n P_i \pi_i(\lambda) - \lambda \right) = (\overline{P}\overline{\pi}(\lambda)' - \lambda)$$

Verico (2002) follows this idea by considering a quadratic rating error function. Consequently, Verico's optimal premiums are defined as those that minimize another kind of expected squared rating error function, namely

$$\int_0^\infty \left(\sum_{i=1}^n P_i \pi_i(\lambda) - \lambda \right)^2 u(\lambda)\,d\lambda$$

The differences between both approaches can be expressed in terms of **Bayesian Decision Theory**. In general decision problems it is well known (see, for instance,

DeGroot (1970)) that, if the decision maker's preferences over the possible consequences of his decisions are consistent with some axioms of rational behaviour, then it is possible to define a function over those consequences (called the ***utility*** of the consequences) such that one feasible decision will be preferred to another if, and only if, the expected utility of the possible consequences is larger for the first decision than it is for the second. In decision problems it is usual to specify the negative of the utility, and to call it the ***loss***. Then the decision maker should choose as optimal the decision that minimizes the expected loss of his consequences.

The decision maker must then proceed as follows: first, he must define a consequence for every feasible decision and every value of the random parameters; second, he must specify a numerical loss associated with every consequence; finally, he calculates the expected loss of every feasible decision, and chooses as optimal the decision with minimal expected loss.

In order to apply this general framework to our particular problem, let us consider an arbitrary policy in the stationary state. On the one hand, if we formulate a Bayesian decision problem in which the random states are the risk parameter λ of the policy and the class C_i to which the policy belongs, the feasible decisions are the possible scales $(P_1, \ldots, P_n)$, and the loss associated with every state (λ, C_i) and every feasible decision $(P_1, \ldots, P_n)$ is a quadratic loss $(P_i - \lambda)^2$, then the optimal decision minimizes the expected squared rating error as defined by Norberg (1976). In other words, the optimal decision is to adopt the Bayes Scale.

On the other hand, if we formulate a different Bayesian decision problem, where the feasible decisions are again the scales $(P_1, \ldots, P_n)$, the random states are the values of λ and the loss associated with every premium scale and every value of the random parameter λ is defined as

$$\left(\sum_{i=1}^{n} P_i \pi_i(\lambda) - \lambda \right)^2$$

then the optimal decision becomes Verico's premium scale.

It is clear that both approaches are rightful, and the decision to choose one or another could depend on the good properties of the resulting premium scales. In this sense, it is important to remember that Bayes Scales, which can be considered the standard method for optimal BMS design, enjoy good properties, such as the financial equilibrium of the resulting BMS.

But let us now pay attention to another component of the Bayesian decision problems, namely, the loss functions. Both Norberg's and Verico's approaches consider quadratic loss functions. This is a very usual choice when dealing with Bayesian estimation problems. We should not forget, however, that the Bayesian Methodology does not force the decision maker to always choose quadratic loss functions. For example, Box and Tiao (1973) say that "in a proper Bayesian decision analysis, the loss function is supposed to represent a realistic economic penalty associated with the available actions … it is easy to imagine practical situations where the squared error loss would be inappropriate" (Box

and Tiao (1973), pg. 309). Smith (1988) finds also difficult to justify the use of quadratic loss functions in most applications. Instead, "in many simple decision problems it is reasonable to assume that the appropriate loss function is piecewise linear … The simplest piecewise linear loss function is the absolute loss function" (Smith (1988), pg. 110). See also Raiffa and Schlaifer (1961), section 4.2, and Lee (1989), section 7.5.

These observations raise the question of the appropriateness of absolute loss functions as realistic penalties in our BMS design problems. In our opinion, it could be reasonable to use such absolute loss functions in this kind of decision problems, because it allows us to distinguish between policyholders paying more than their fair premiums, and those paying less. Instead, quadratic loss functions imply the equal valuation of overachievements and underachievements around the true value of the risk parameter λ, and do not allow to distinguish between policyholder's and insurer's extraordinary earnings.

In the next section we will use absolute instead of quadratic loss functions when the optimal BMS design follows the second approach previously mentioned. We will see that the resulting mathematical program becomes easy to solve, because it is equivalent to a linear program. We will see also that this fact allows the BMS designer to easily manage a great number of theoretical and/or commercial characteristics to be fulfilled by the premium scales.

4. THE LINEAR PROGRAMMING APPROACH TO OPTIMAL BMS DESIGN

Consider a policy with mean claim frequency $\Lambda = \lambda$. We assume that the BMS designer should try to calculate a pure premium for that policy as close as possible to its mean claim frequency. If we take absolute error functions in the second approach mentioned in Section 3, then the rating error should be defined as $\left| \sum_{i=1}^{n} P_i \pi_i(\lambda) - \lambda \right| = \left| \overline{P} \overline{\pi}(\lambda)' - \lambda \right|$. We want therefore to calculate the scale of premiums $(P_1, \ldots, P_n)$ that minimizes the following global rating error, which we have called the BMS ***Global Asymptotic Fairness*** (see Heras, Vilar and Gil (2002)):

$$\int \left| \sum_{i=1}^{n} P_i \pi_i(\lambda) - \lambda \right| dU(\lambda)$$

This mathematical program is, in general, rather difficult to solve. Nevertheless, if we assume, as it is usual in the literature, a discrete distribution for the parameter Λ (taking the values $\overline{\lambda} = (\lambda_1, \ldots, \lambda_m)$ with probabilities $\overline{q} = (q_1, \ldots, q_m)$), then it becomes equivalent to both

$$\text{Min} \sum_{j=1}^{m} \left| \sum_{i=1}^{n} P_i \pi_i(\lambda_j) - \lambda_j \right| q_j = \left| \overline{P} \Pi(\lambda) - \overline{\lambda} \right| \overline{q}'$$

and the following linear program (from hereafter, ***Program (P)***):

$$\text{Min} \sum_{j=1}^{m} \left(y_j^+ + y_j^- \right) q_j$$

$$\text{s.t.}$$
$$P_1\pi_1(\lambda_1) + \ldots + P_n\pi_n(\lambda_1) + y_1^- - y_1^+ = \lambda_1$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$P_1\pi_1(\lambda_m) + \ldots + P_n\pi_n(\lambda_m) + y_m^- - y_m^+ = \lambda_m$$
$$y_j^+, y_j^- \geq 0, \ \forall j = 1,\ldots,m; \ P_i \geq 0, \ \forall i = 1,\ldots,n$$

or, equivalently,

$$\text{Min} \ (\bar{y}^+ + \bar{y}^-)\bar{q}'$$
$$\text{s.t.}$$
$$\bar{P}.\Pi(\lambda) + \bar{y}^- - \bar{y}^+ = \bar{\lambda}$$
$$\bar{P}, \bar{y}^+, \bar{y}^- \geq \bar{0}$$

where $\bar{y}^+ = (y_1^+,\ldots,y_m^+)$, $\bar{y}^- = (y_1^-,\ldots,y_m^-)$, and $\Pi(\lambda)$ is the matrix $(\pi_i(\lambda_j))$. The optimal values of the new variables y_j^+, y_j^- (let us denote them as y_j^{*+}, y_j^{*-}) represent the positive and negative rating errors, respectively, for a policy with parameter λ_j. This fact, as well as the equivalence between the previous programs, are consequences of the fact that the optimal solution of the linear program verifies $y_j^{*+}. y_j^{*-} = 0, \forall j$ (see for instance Sawaragi, Nakayama and Tanino (1985, p. 253)), that is, only one of them can be non null, and therefore

$$\left| \bar{P}\Pi(\lambda) - \bar{\lambda} \right| = \bar{y}^{*+} + \bar{y}^{*-}$$

It is important to remark that the linear program has been obtained as a particular case of the multiobjective technique known as ***Goal Programming***. This fact has an easy interpretation: by means of Goal Programming the decision maker tries to find the values of the decision variables such that some objective functions take values as close as possible to a set of targets previously defined. In our problem these targets are $\lambda_1,\ldots,\lambda_m$, the possible values of the parameter Λ, and we try to find the values of the decision variables $P_1,\ldots,P_n$ in order to approximate, for every policy with parameter λ_j, the mean value of the premiums paid by that policyholder $\left(\sum_{i=1}^n P_i\pi_i(\lambda_j)\right)$ and his real mean claim frequency λ_j.

Goal Programming methodology has become one of the most important tools for solving multicriteria optimization problems, perhaps because it is an easy-to-handle technique with a high flexibility in fitting the characteristics of real-world problems (see Romero (1991)). We will see below that this flexibility allows to incorporate in our linear programming BMS all the reasonable features mentioned in the previous section.

But we can firstly investigate some of the properties of *Program (P)*. For example, it is evident that it is a feasible and lower-bounded program, and therefore it is solvable. The dual program can give us additional insight. The dual is ***Program (D)***:

$$\text{Max} \ \bar{\lambda}\bar{u}'$$
$$\text{s.t.}$$

$$\Pi(\lambda)\,\bar{u}' \le \bar{0}$$
$$-\bar{q} \le \bar{u} \le \bar{q}$$

where $\bar{u} = (u_1, \ldots, u_m)$ are the dual variables. Thus the dual optimal value is upper bounded by $\bar{\lambda}\bar{q}' = E[\Lambda]$, and by the duality theorems (see Ignizio (1982), chapter 8) this property holds also for the primal optimal value, that is, for the optimal value of *Program (P)*. The optimal global asymptotic fairness takes then values between zero and the expectation of the risk parameter.

5. Theoretical and Practical Requirements as Linear Constraints

We will show in this section how several theoretical and practical requirements can be easily included as new linear constraints of *Program (P)*. In the first set of requisites we will include the financial equilibrium property, as well as premium's monotonicity. As commercial requirements we will consider upper and/or lower bounds on premium variability, as well as the assignment of a pre-established value to some premium.

– First of all, let us begin by the ***financial equilibrium property***. It is not difficult to get a relationship between the optimal solution of our linear program and the degree of financial equilibrium: in fact, this is measured by the expression

$$\sum_{i=1}^{n} P_i \pi_i - \sum_{j=1}^{m} \lambda_j q_j = \overline{P}\overline{\pi}' - E[\Lambda]$$

where

$$\pi_i = \sum_{j=1}^{m} \pi_i(\lambda_j)\,q_j = \bar{\pi}(\lambda_j)\,\bar{q}'$$

Positive values of this expression denote a gain for the insurance company, while negative ones denote a loss (i.e. a gain for the policyholders). This fact allows us the introduction of the financial equilibrium condition in *Program (P)*, by means of the following linear constraint:

$$\sum_{i=1}^{n} P_i \pi_i = E[\Lambda]$$

– As in the case of Bayes scales, the premiums obtained as optimal solutions of *Program (P)* do not necessarily form a monotonous sequence. In order to get that ***monotonicity property*** we only need to include some additional linear constraints:

$$P_1 \le P_2 \le \ldots P_n$$

– On the other hand, the linear constraint

$$P_i = V_i$$

should be included if the BMS designer wants **to assign a pre-established value V_i to some premium P_i**. This could be useful, for instance, if the premium associated to the initial class must coincide with the pure premium for the whole portfolio.

– Let us consider now the **premium variability**. It is clear that linear constraints as

$$P_{i+1} - P_i \geq d$$

introduce a lower limit $d > 0$ for the difference between two consecutive premiums. Similarly, the constraints

$$0 \leq P_n - P_1 \leq D$$

introduce an upper limit $D > 0$ for the difference between the two extreme premiums. Obviously, changing the sign of the inequalities will transform upper into lower bounds and vice versa.

As an alternative, the distance between extreme or consecutive premiums could also be bounded on a percentage basis:

$$P_n \leq, \geq rP_1$$
$$P_{i+1} \leq, \geq rP_1$$

Finally, we must remember that the previous results allow to calculate a premium scale only for some given classes and transition rules. As in the case of Bayes scales, alternative sets of different classes and transition rules should be evaluated and compared according to the optimal value of the objective function.

The following easy example will help us to explain how linear programming methodology works out in practice. Let us assume three possible values of the random parameter Λ ($\lambda_1 = 0.05$ (with probability $\frac{1}{3}$), $\lambda_2 = 0.1$ (with probability $\frac{1}{3}$) and $\lambda_3 = 0.15$ (with probability $\frac{1}{3}$)), and three bonus-malus classes (C_1, C_2 and C_3). Assume also the following transition rules: claim-free years are rewarded by a one-class discount, the first claim leads to a one-class increase, and two or more claims during the year lead to a two-class increase. Assuming a portfolio with claims following a Poisson distribution mixed with the random parameter Λ, these rules imply the following conditional stationary probability distributions:

- If $\Lambda = \lambda_1$: $\pi_1(\lambda_1) = 0.9477$; $\pi_2(\lambda_1) = 0.0486$; $\pi_3(\lambda_1) = 0.0037$
- If $\Lambda = \lambda_2$: $\pi_1(\lambda_2) = 0.8917$; $\pi_2(\lambda_2) = 0.0938$; $\pi_3(\lambda_2) = 0.0145$
- If $\Lambda = \lambda_3$: $\pi_1(\lambda_3) = 0.8334$; $\pi_2(\lambda_3) = 0.1349$; $\pi_3(\lambda_3) = 0.0317$

The unconditional stationary probability distribution will be:

$$\pi_1 = 0.8910; \ \pi_2 = 0.0924; \ \pi_3 = 0.0166$$

The Bayes scale is defined as the set of premiums (P_1, P_2, P_3) that minimizes the value of the expected squared rating error. In this case, it takes the following values:

$$P_1 = 0.0979; \ P_2 = 0.1156; \ P_3 = 0.1281$$

We know that Bayes scales give rise to financially balanced bonus-malus systems. It is easy to check the financial equilibrium property in this example: the expected claims are $E(\Lambda) = 0.1$, while the expected premiums are $P_1\pi_1 + P_2\pi_2 + P_3\pi_3 = 0.1$.

But let us imagine that this system becomes inadequate for practical purposes, perhaps because the policyholders prefer larger differences between consecutive and extreme premiums. It is easy to check that our Bayes scale verifies that

$$\frac{P_2}{P_1} = 1.18; \ \frac{P_3}{P_2} = 1.108; \ \frac{P_3}{P_1} = 1.309$$

In this case marketing reasons could recommend the construction of a new bonus-malus system with increased values of these ratios, for example, such that

$$\frac{P_2}{P_1} \geq 1.20; \ \frac{P_3}{P_2} \geq 1.20; \ \frac{P_3}{P_1} = 1.50$$

If the BMS designer tries to calculate a new Bayes scale that minimizes the expected squared rating error constrained to these conditions, he will have to solve a constrained quadratic program whose optimal solution will not verify, in general, the financial equilibrium property. For this reason, this property should be included as a new constraint of the quadratic program. Note that, if the system has a great number of classes, this methodology may be computationally complex.

Of course, the BMS designer could try to define new classes and/or change the transition rules, in order to get a new Bayes scale with the desired properties. This methodology may be successful in many cases, but may also be inadequate when the system has a great number of classes and complex transition rules.

In that case, it may be reasonable to introduce the desired properties as linear constraints using the linear programming methodology. In this example, the linear program (from hereafter, ***Program (1)***) would be

$MIN(y_1^+ + y_1^- + y_2^+ + y_2^- + y_3^+ + y_3^-)\frac{1}{3}$
s.t.
$0.9477P_1 + 0.0486P_2 + 0.0037P_3 + y_1^- - y_1^+ = 0.05$
$0.8917P_1 + 0.0938P_2 + 0.0145P_3 + y_2^- - y_2^+ = 0.1$
$0.8334P_1 + 0.1349P_2 + 0.0317P_3 + y_3^- - y_3^+ = 0.15$

$$P_i \geq 0, \forall i$$
$$y_j^+, y_j^- \geq 0, \forall j$$
$$P_2 \geq 1.2P_1$$
$$P_3 \geq 1.2P_2$$
$$P_3 - 1.5P_1 = 0$$
$$0.8910P_1 + 0.0924P_2 + 0.0166P_3 = 0.1$$

The last constraint establishes the financial equilibrium condition. The optimal solution of this linear program is easy to find by means of the simplex method:

$$P_1 = 0.0970; \ P_2 = 0.1212; \ P_3 = 0.1454$$

It is easy to check that all the required properties hold for this new scale.

The inclusion of additional linear constraints is not the only way of taking into consideration new BMS properties. In fact, the objective function of *Program (P)* can be modified in order to account for some new properties. For example, this objective function assumes the equal valuation of overachievements and underachievements around the true value of the parameter λ. Nevertheless, it could be interesting to distinguish between the policyholders' and insurer's extraordinary earnings, respectively. Moreover, the objective function equally weights the errors in all the bonus-malus classes. But it could be interesting to make the weights dependent on the value of Λ: it seems clear that an error $\left| \sum_{i=1}^n P_i \pi_i(\lambda) - \lambda \right| = 0.1$ is less important when $\lambda = 0.15$ than when $\lambda = 0.05$, for instance.

It is easy to include these last properties in our mathematical program: if we redefine the objective function of *Program (P)* as

$$MIN \sum_{j=1}^m \left(\omega_j^+ y_j^+ + \omega_j^- y_j^- \right) q_j$$

then it would be possible to assign different weights (even depending on the different discrete values of Λ) to positive and negative rating errors.

In the example, all the weights in the objective function were equal, although we could easily make these weights dependent on the values of the parameter Λ. If, for instance, we take $\omega_i^+ = \omega_i^- = \frac{1}{\lambda_i}$, then the new objective function will be

$$\left(20y_1^+ + 20y_1^- + 10y_2^+ + 10y_2^- + \frac{100}{15} y_3^+ + \frac{100}{15} y_3^- \right) \frac{1}{3}$$

but the new optimal solution still remains the same: in fact, it is not necessary to solve a new linear program, since sensitivity analysis techniques in Linear Programming can easily be applied in order to check this property (see Ignizio (1982), chapter 10).

Let us remember the meaning of the linear programming premium scale: it optimizes the BMS global asymptotic fairness, by approximating as much as possible the mean premium paid by every policyholder to his real mean claim frequency, while preserving restrictions such as the financial equilibrium of the system and the commercial requirements. Moreover, the linear programming scale in this example is a robust scale, in the sense that the deviations between mean premiums and claim frequencies can be weighted according to the different values of these mean frequencies. All these different and desirable properties are joined together thanks to the linearity of their mathematical formulations.

6. Elasticity and RSAL as Linear Constraints

Let us turn now to the consideration of some a posteriori adequacy measures of BMS scales, that is, measures that allow the comparison between existing BMS in order to determine which one results to be "the best". As we will see in this section, two of these measures, namely Elasticity and RSAL, can be mathematically represented by means of linear constraints using the linear programming methodology, being therefore a priori controlled by the BMS designer.

Loimaranta (1972) defines the **_Elasticity_** of a BMS (in fact, he calls it **_Efficiency_**) for the parameter λ as the elasticity of the mean stationary premium with respect to λ. If we denote $P(\lambda) = \sum_{i=1}^{n} P_i \pi_i(\lambda)$ to such mean stationary premium associated with a claim frequency λ, the elasticity is defined as

$$\eta(\lambda) = \frac{\dfrac{dP(\lambda)}{P(\lambda)}}{\dfrac{d\lambda}{\lambda}} = \frac{dP(\lambda)}{d\lambda} \cdot \frac{\lambda}{P(\lambda)}$$

Ideally, the elasticity of a BMS should be close to one for the most common values of Λ, that is, a relative increment of the claim frequency should lead to a similar relative increment of the mean stationary premium. The general rule is, however, that the elasticity is usually less than one.

Actually, the value of the elasticity could be under the BMS designer control if he uses the linear programming methodology. In fact, the constraints $\eta(\lambda) \geq C_\lambda$ are linear functions of $(P_1, \ldots, P_n)$, so that it is possible to impose lower bounds C_λ for the elasticity values (of course, the same is valid for upper bounds).

In the case of the easy example that illustrates our reasoning along the previous section, the elasticity should be approximated by means of a discretization. The linear constraint

$$\frac{dP(\lambda)}{d\lambda} \cdot \frac{\lambda}{P(\lambda)} \geq C_\lambda$$

that is,

$$Lim_{\Delta\lambda\to 0} \frac{\Delta P(\lambda)}{\Delta\lambda} \cdot \frac{\lambda}{P(\lambda)}$$

can be discretized as

$$\frac{\sum_{i=1}^{n} P_i\left(\pi_i\left(\lambda_{j+1}\right) - \pi_i\left(\lambda_j\right)\right)}{\lambda_{j+1} - \lambda_j} \cdot \frac{\lambda_j}{\sum_{i=1}^{n} P_i \pi_i\left(\lambda_j\right)} \geq C_j$$

or, equivalently,

$$\sum_{i=1}^{n} P_i\left[\left(\pi_i\left(\lambda_{j+1}\right) - \pi_i\left(\lambda_j\right)\right).\lambda_j - \pi_i\left(\lambda_j\right).C_j.\left(\lambda_{j+1} - \lambda_j\right)\right] \geq 0$$

In fact, the last expressions correspond to the "right elasticity constraints". In a similar way one can define "left elasticity constraints" as

$$\frac{\sum_{i=1}^{n} P_i\left(\pi_i\left(\lambda_j\right) - \pi_i\left(\lambda_{j-1}\right)\right)}{\lambda_j - \lambda_{j-1}} \cdot \frac{\lambda_j}{\sum_{i=1}^{n} P_i \pi_i\left(\lambda_j\right)} \geq C_j$$

$$\sum_{i=1}^{n} P_i\left[\left(\pi_i\left(\lambda_j\right) - \pi_i\left(\lambda_{j-1}\right)\right).\lambda_j - \pi_i\left(\lambda_j\right).C_j.\left(\lambda_j - \lambda_{j-1}\right)\right] \geq 0$$

For example, if we want to control the right elasticity in $\lambda_2 = 0.1$ (the mean, median and mode of the random parameter Λ), we have to add to *Program (1)* the new linear constraint

$$P_1(0.8334 - 0.8917) + P_2(0.1349 - 0.0938) + P_3(0.0317 - 0.0145) \geq$$
$$\geq C.\frac{1}{2}.(0.8917 P_1 + 0.0938 P_2 + 0.0145 P_3)$$

that is,

$$P_1\left(-0.0583 - \frac{0.8917}{2}C\right) + P_2\left(0.0411 - \frac{0.0938}{2}C\right) + P_3\left(0.0172 - \frac{0.0145}{2}C\right) \geq 0$$

If we add these constraints to *Program (1)* for different values of the minimum right elasticity C, we obtain the same optimal solution

$$P_1 = 0.0970; \ P_2 = 0.1212; \ P_3 = 0.1454$$

for $C \leq 0.0364$. We also obtain that the linear program becomes unfeasible for the rest of the values. We can conclude that 0.0364 is the maximum value of the right elasticity, at least if we continue keeping the commercial requirements

that gave rise to *Program (1)*. Notice that there is a slight improvement when comparing with the Bayes scale, whose right elasticity value is 0.025. Of course, a similar analysis can be done with left elasticity: its value for the linear programming scale is 0.0323, while for the Bayes scale is 0.022.

Let us turn now to the second previously mentioned adequacy measure of a BMS scale, the so-called **Relative Stationary Average Level** (**RSAL**).

RSAL measures the severity of the BMS, expressing the relative premium paid by an arbitrary policyholder as a percentage, when the stationary state has been approached. This is defined as

$$RSAL = \frac{\sum_{i=1}^{n} P_i \pi_i - P_1}{P_n - P_1}$$

RSAL usually takes low values for most real BMS, indicating a high clustering of policies in the lowest classes, although its ideal value should be close to 50% (Lemaire (1995), p. 64). As in the case of elasticity, the value of RSAL could be under control if the BMS designer uses the linear programming methodology, because the constraints

$$RSAL_{\min} \leq RSAL \leq RSAL_{\max}$$

are linear constraints that could be added to *Program (P)*. For instance, the constraint $RSAL \geq RSAL_{\min}$ could be expressed as

$$\left(RSAL_{\min} + \pi_1 - 1\right) P_1 + \pi_2 P_2 + \left(\pi_3 - RSAL_{\min}\right) P_3 \geq 0$$

We know that, if the BMS designer is too optimistic taking bounds for elasticities and RSAL, the resulting linear program will often become unfeasible. This fact raises the question of the appropriate selection of those limits. That selection could be accomplished by means of a trial and error process, taking as starting point the observed values of elasticities and RSAL in an existing BMS. In our example these initial values could be those associated with the Bayes scale, namely 0.025 (right elasticity), 0.022 (left elasticity) and 0.07 (RSAL). Therefore, we should try to find a BMS scale with higher right and left elasticities and RSAL nearest to 0.5. Unfortunately, if we add to *Program (1)* three new linear constraints, two of them imposing that both right and left elasticities must be greater than 0.025 and 0.022, respectively, and the last one imposing that RSAL must be greater than 0.07, then the new program becomes unfeasible. The reason is that the commercial constraints

$$P_2 \geq 1.2 P_1; \quad P_3 \geq 1.2 P_2; \quad P_3 = 1.5 P_1$$

must also take the Bayes scale values, in order to guarantee feasibility:

$$P_2 \geq 1.18 P_1; \quad P_3 \geq 1.108 P_2; \quad P_3 \geq 1.309 P_1$$

Besides, we must include a constraint like $P_3 \leq 1.5P_1$, in order to reject useless solutions. If we solve the resulting linear program, the new optimal solution obtained is

$$P_1 = 0.0961; \; P_2 = 0.13; \; P_3 = 0.1441$$

The values of right elasticity, left elasticity and RSAL for this new premium scale are, respectively, 0.044 , 0.041 and 0.08. And the global asymptotic fairness is now 0.0319 (0.0326 for the Bayes scale). The new scale is financially balanced and achieves better values for right and left elasticity, RSAL and fairness. It seems clear that this scale could be preferred to the initial Bayes scale. Notice, however, the poor values obtained for both elasticities in this unreal example. In the next section we will study a more realistic example with a higher number of classes for the structure function, where the elasticities will take much better values.

7. A NUMERICAL EXAMPLE

We will illustrate the ideas presented in the previous sections by means of a more realistic numerical example. We will assume a portfolio with claims following a Poisson distribution mixed with the following discrete structure function:

λ_i	0.033	0.066	0.099	0.132	0.165	0.198	0.231	0.264	0.297	0.330
q_i	0.28770	0.21179	0.23174	0.06609	0.08872	0.02623	0.03636	0.01126	0.01592	0.00510
λ_i	0.363	0.396	0.429	0.462	0.495	0.528	0.561	0.594	0.627	0.660
q_i	0.00732	0.00240	0.00348	0.00116	0.00171	0.00058	0.00085	0.00029	0.00043	0.00078

This function is obtained by applying the discretization method explained in Vilar (2000) to the continuous structure function

$$u(\lambda) = \frac{g}{\lambda^{3/2}\sqrt{2\pi h}} \, \exp\left(-\frac{(\lambda - g)^2}{2h\lambda}\right)$$

with $g = 0.101081$ and $h = 0.062981$. This is the density function of an inverse Gaussian distribution, taken from Lemaire (1995, pp. 35-37).

We have chosen the following transition rules:

	CLASS AFTER CLAIMS				
	0	1	2	3	>3
Class					
10	9	10	10	10	10
9	8	10	10	10	10
8	7	10	10	10	10
7	6	9	10	10	10
6	5	8	10	10	10

	0	1	2	3	>3
Class					
5	**4**	**7**	**9**	**10**	**10**
4	**3**	**6**	**8**	**9**	**10**
3	**2**	**5**	**7**	**9**	**10**
2	**1**	**4**	**6**	**7**	**9**
1	**1**	**3**	**5**	**6**	**8**

The following expressions are the efficiency measures associated with the Bayes scale and with the linear programming scale, respectively, for a discrete structure function:

$$Q_c(P_1,...,P_n) = \sum_{j=1}^{m} \sum_{i=1}^{n} (P_i - \lambda_j)^2 \pi_i(\lambda_j) q_j$$

$$Q_m(P_1,...,P_n) = \sum_{j=1}^{m} \left| \sum_{i=1}^{n} P_i \pi_i(\lambda_j) - \lambda_j \right| . q_j$$

In the terminology of the previous sections, Q_c and Q_m correspond to the **expected squared rating error** and the **global asymptotic fairness**, respectively.

In the following tables we will summarize the most important properties of the Bayes scale and some linear programming scales for this example. In each case we will show the premium values and the ratio between the premiums of consecutive classes and the premiums of extreme classes. We will also calculate the values of Q_c and Q_m for each scale.

The first table shows these results for the Bayes scale (BS):

TABLE 1: THE BAYES SCALE

i	BS	$\frac{BS_{i+1}}{BS_i}$
1	0.0824	1.481
2	0.1222	1.046
3	0.1278	1.356
4	0.1734	1.088
5	0.1887	1.240
6	0.2324	1.118
7	0.2620	1.159
8	0.3039	1.115
9	0.3391	1.117
10	0.3789	$\frac{BS_{10}}{BS_1} = 4.6$
Q_c	0.00415	
Q_m	0.04036	

The properties of the Bayes scale seem reasonable, although we think that it is possible to improve some of them by means of the linear programming methodology. Let us see several examples:

1. Some BMS designers may judge unacceptable the variability of the penalties or ratios between some consecutive premiums (notice, for instance, that the ratio between the two first premiums (1.481) is much bigger than the next one (1.046)). In the same way, the ratio between extreme premiums (4.6) may be judged excessive. In Table 2 we obtain scales whose ratios have better properties. Both scales LP_1 and LP_2 enjoy the financial equilibrium property. Moreover, every penalty must vary between a minimum value of 1.05 and a maximum value of 1.30, so that the constraints $1.05 \times P_i \leq P_{i+1} \leq 1.3 \times P_i$ hold for both scales. Finally, the maximum premium for the scale LP_2 must be, at most, three times the value of the minimum premium, that is, the additional constraint $P_{10} \leq 3P_1$ holds for the scale LP_2. The ratio between extreme premiums for the scale LP_1 takes the same value as the Bayes scale (4.6).

TABLE 2: SMOOTHED BMS WITH LESS PREMIUM VARIABILITY

i	LP_1	$\frac{LP_{1,i+1}}{LP_{1,i}}$	LP_2	$\frac{LP_{2,i+1}}{LP_{2,i}}$
1	0.0802	1.3	0.0846	1.3
2	0.1043	1.3	0.1100	1.3
3	0.1356	1.3	0.1430	1.3
4	0.1764	1.3	0.1859	1.07
5	0.2293	1.3	0.1989	1.05
6	0.2981	1.07	0.2088	1.05
7	0.3190	1.05	0.2192	1.05
8	0.3350	1.05	0.2302	1.05
9	0.3517	1.05	0.2417	1.05
10	0.3696	$\frac{LP_{1,10}}{LP_{1,1}} = 4.6$	0.2538	$\frac{LP_{2,10}}{LP_{2,1}} = 3$
Q_c	0.00429		0.00440	
Q_m	0.03822		0.04251	

Notice that the global asymptotic fairness Q_m improves when we consider the scale LP_1 (0.03822) instead of the Bayes scale (0.04036), although this is not the case for the scale LP_2. Perhaps this fact suggests to make smoother the last LP_2-constraint relating to the values of the extreme premiums. Finally, notice that the Bayes scale gets the best score with respect to the expected squared rating error Q_c (as was to be expected).

2. It could seem reasonable that the class with an associated premium nearest to the expected loss of the portfolio should be chosen as the BMS initial class.

In our example, the mean value of the discrete structure function is 0.10100. In the particular case of the Bayes scale, this fact suggests to choose the first or second class as initial class. Assuming the last option, this gives one "bonus" class and eight "malus" classes for the Bayes scale. Because of commercial requirements, some BMS designers may wish to increase the number of "bonus" classes. If, for instance, they want to get two "bonus" classes, this could be accomplished by adding to the linear program the additional constraint $P_3 = 0.10100$.

Taking also into consideration the LP_1-constraints, this gives rise to the new scale LP_3:

TABLE 3: BMS WITH AN INCREASED NUMBER OF BONUS CLASSES

i	LP_3	$\dfrac{LP_{3,i+1}}{LP_{3,i}}$
1	0.08797	1.09
2	0.09619	1.05
3	0.10100	1.3
4	0.13130	1.3
5	0.17070	1.3
6	0.22191	1.3
7	0.28848	1.27
8	0.36705	1.05
9	0.38540	1.05
10	0.40467	$\dfrac{LP_{3,10}}{LP_{3,1}} = 4.6$
Q_c	0.00487	
Q_m	0.04031	

Notice that the global asymptotic fairness Q_m slightly improves when we take the scale LP_3 (0.04031) instead of the Bayes scale (0.04036).

In fact, real-world BMS usually choose starting premiums higher than the expected losses, since they assume that most of the new policies are inexperienced drivers. Of course, linear programming methodology allows to choose any premium value for the starting class, and not necessarily the expected loss as in the preceding example.

3. Next, we will try to improve the RSAL and Elasticity values of the Bayes scale by means of the linear programming methodology. The RSAL value and the elasticity (evaluated at the mean value 0.10100) for the Bayes scale are, respectively, 0.062 and 0.2006, so that we have to add two new linear constraints:

$$RSAL \geq 0.062$$
$$Elasticity \geq 0.2006$$

To make sure that the new linear program is feasible, we will add as new constraints the Bayes scale properties relating to the ratios between premiums (and, of course, the financial equilibrium property):

$$1.05 \times P_i \le P_{i+1} \le 1.5 \times P_i, \ i = 1,\ldots,9$$
$$P_{10} \le 4.6 \times P_1$$

Then we obtain the new LP_4 scale:

TABLE 4: BMS WITH IMPROVED RSAL AND ELASTICITY

i	LP_4	$\frac{LP_{4,i+1}}{LP_{4,i}}$
1	0.0743	1.5
2	0.1115	1.5
3	0.1672	1.5
4	0.2509	1.067
5	0.2679	1.05
6	0.2813	1.05
7	0.2959	1.05
8	0.3102	1.05
9	0.3257	1.05
10	0.3420	$\frac{LP_{4,10}}{LP_{4,1}} = 4.6$
Q_c	0.00464	
Q_m	0.0351	

The new LP_4-RSAL value is 0.099, and the new LP_4-elasticity value is 0.3183. We have therefore improved both values. Moreover, the LP_4-global asymptotic fairness takes the value 0.0351, much better than the same value for the Bayes scale (0.04036).

Notice that it is not necessary in this example to consider the right and left elasticities, since the continuous structure function is available. Nevertheless, we could obtain the same conclusions if we worked with both elasticities in the discrete case. For the Bayes scale, right and left elasticities (evaluated at 0.1212, near the mean value 0.1010) take the values 0.2243 and 0.1693, respectively. Then we should add to the linear program the new constraints

$$Right - Elasticity \ge 0.2243$$
$$Left - Elasticity \ge 0.1693$$

but we will obtain again as optimal solution the scale LP_4. And the new values of LP_4 right and left elasticities (evaluated at 0.1115) are now 0.3483 and 0.2742, respectively.

4. Finally, in the next table we remove the financial equilibrium condition keeping commercial, elasticity and RSAL constraints:

TABLE 5: BMS WITHOUT FINANCIAL EQUILIBRIUM CONDITION

i	LP_5	$\dfrac{LP_{5,i+1}}{LP_{5,i}}$
1	0.0777	1.5
2	0.1165	1.5
3	0.1748	1.5
4	0.2623	1.067
5	0.2801	1.05
6	0.2941	1.05
7	0.3088	1.05
8	0.3242	1.05
9	0.3404	1.05
10	0.3575	$\dfrac{LP_{5,10}}{LP_{5,1}} = 4.6$
Q_c	0.00476	
Q_m	0.0343	

The new LP_5-premiums have increased their values in approximately 4.5%, and lead to a positive profit for the insurance company. The global asymptotic fairness takes the best value (0.0343) of all the preceding cases. LP_5-RSAL and elasticity take the same values as in the scale LP_4.

8. CONCLUSIONS

In the preceding sections we have shown how linear programming methodology for BMS design is able to take into account a great number of theoretical properties and commercial requirements that could be interesting to manage in practical problems. Some of them can be mathematically represented by means of linear constraints. Examples of these properties are the financial equilibrium of the system, the monotonicity and variability of the premium scales, the selection of the initial class and even the improvement of BMS efficiency measures such as the RSAL and Elasticity. These properties (with the exception of financial equilibrium and premium monotonicity) are not taken into consideration by other classical methods for BMS design. Moreover, the linear programming methodology is based in an alternative measure (the global asymptotic fairness) which evaluates the global deviation between mean premiums and mean claim frequencies. The selection of appropriate weights in this fairness measure makes possible to account for other interesting properties, such as differentiate between policyholders' and insurers' extraordinary earnings, for instance. Finally, it is clear that this methodology allows a high degree of interaction between the BMS designer and the mathematical model. Indeed, when the initial aspirations of the BMS designer become unattainable, the use of this methodology usually allows to get a reasonable degree of compromise between them.

9. ACKNOWLEDGEMENTS

10. REFERENCES

BAIONE, F., LEVANTESI, S. and MENZIETTI, M. (2002) The Development of an Optimal Bonus-Malus System in a Competitive Market, *ASTIN Bulletin* **32(1)**, 159-169.

BOX, G.E.P. and TIAO, G.C. (1973) *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, MA.

DEGROOT, M.H. (1970) *Optimal Statistical Decisions*. McGraw-Hill, New York.

DE PRIL, N. (1978) The Efficiency of a Bonus-Malus System. *ASTIN Bulletin* **10(1)**, 59-72.

KEMENY, J.G. and SNELL, J.L. (1976) *Finite Markov Chains*. Springer-Verlag, Berlin.

HERAS, A., VILAR, J.L. and GIL, J.A. (2002) Asymptotic Fairness of Bonus-Malus Systems and Optimal Scales of Premiums. *The Geneva Papers on Risk and Insurance Theory* **27**, 61-82.

IGNIZIO, J.P. (1982) *Linear Programming in Single & Multiple Objective Systems*. Prentice-Hall, Englewood Cliffs.

LEE, P.M. (1989) *Bayesian Statistics: an Introduction*. Edward Arnold, London.

LEMAIRE, J. (1985) *Automobile Insurance. Actuarial Models*. Kluwer-Nijhoff Publishing, Dordrecht.

LEMAIRE, J. (1995) *Bonus-Malus Systems in Automobile Insurance*. Kluwer Academic Publishers, Dordrecht.

LEMAIRE, J. (1998) Bonus-Malus Systems: the European and Asian Approach to Merit-Rating. *North American Actuarial Journal* **2(1)**, 26-47.

LOIMARANTA, K. (1972) Some Asymptotic Properties of Bonus Systems. *ASTIN Bulletin* **6**, 233-245.

NORBERG, R. (1976) A Credibility Theory for Automobile Bonus Systems. *Scandinavian Actuarial Journal*, 92-107.

PESONEN, M. (1963) A Numerical Method of Finding a Suitable Bonus Scale. *ASTIN Bulletin* **2**, 102-108.

RAIFFA, H. and SCHLAIFER, R. (1961) *Applied Statistical Decision Theory*. The M.I.T. Press, Cambridge, MA.

ROMERO, C. (1991) *Handbook of Critical Issues in Goal Programming*. Pergamon Press, Oxford.

SMITH, J.Q. (1988) *Decision Analysis: a Bayesian Approach*. Chapman and Hall, London.

SAWARAGI, Y., NAKAYAMA, H. and TANINO, T. (1985) *Theory of Multiobjective Optimization*. Academic Press, New York.

VILAR, J.L. (2000) Arithmetization of Distributions and Linear Goal Programming. *Insurance: Mathematics and Economics* **27**, 113-122.

VERICO, P. (2002) Bonus-Malus Systems: "Lack of Transparency" and Adequacy Measure. *ASTIN Bulletin* **32(2)**, 315-318.

ANTONIO HERAS
Departamento de Economía Financiera y Contabilidad I
(Economía Financiera y Actuarial)
Facultad de Ciencias Económicas, Universidad Complutense de Madrid,
(Campus de Somosaguas), 28223 Pozuelo de Alarcón,
Madrid, Spain.
e-mail: aheras@ccee.ucm.es