

DESIGN OF OPTIMAL BONUS-MALUS SYSTEMS WITH A FREQUENCY
AND A SEVERITY COMPONENT ON AN INDIVIDUAL BASIS IN
AUTOMOBILE INSURANCE

BY

NICHOLAS E. FRANGOS* AND SPYRIDON D. VRONTOS*

ABSTRACT

The majority of optimal Bonus-Malus Systems (BMS) presented up to now in the actuarial literature assign to each policyholder a premium based on the number of his accidents. In this way a policyholder who had an accident with a small size of loss is penalized unfairly in the same way with a policyholder who had an accident with a big size of loss. Motivated by this, we develop in this paper, the design of optimal BMS with both a frequency and a severity component. The optimal BMS designed are based both on the number of accidents of each policyholder and on the size of loss (severity) for each accident incurred. Optimality is obtained by minimizing the insurer's risk. Furthermore we incorporate in the above design of optimal BMS the important a priori information we have for each policyholder. Thus we propose a generalised BMS that takes into consideration simultaneously the individual's characteristics, the number of his accidents and the exact level of severity for each accident.

KEYWORDS

Optimal BMS, claim frequency, claim severity, quadratic loss function, a priori classification criteria, a posteriori classification criteria.

1. INTRODUCTION

BMS penalize the policyholders responsible for one or more claims by a premium surcharge (malus) and reward the policyholders who had a claim free year by awarding discount of the premium (bonus). In this way BMS

* Department of Statistics, Athens University of Economics and Business, Patission 76, 10434, Athens, Greece. E-mail for correspondence nef@aueb.gr and svrontos@aueb.gr

This work has been partially supported by 96SYN 3-19 on "Design of Optimal Bonus-Malus Systems in Automobile Insurance" and the General Secretariat of Research and Technology of Greece. The authors would like to thank the referees for their valuable comments.

encourage policyholders to drive carefully and estimate the unknown risk of each policyholder to have an accident.

A BMS is called optimal if it is: 1. financially balanced for the insurer, that is the total amount of bonuses is equal to the total amount of maluses. 2. Fair for the policyholder, that is each policyholder pays a premium proportional to the risk that he imposes to the pool. Optimal BMS can be divided in two categories: those based only on the a posteriori classification criteria and those based both on the a priori and the a posteriori classification criteria. As a posteriori classification criteria are considered the number of accidents of the policyholder and the severity of each accident. As a priori classification criteria are considered the variables whose their values are known before the policyholder starts to drive, such as characteristics of the driver and the automobile. The majority of BMS designed is based on the number of accidents disregarding their severity. Thus first let us consider the design of optimal BMS based only on the a posteriori claim frequency component.

1.1. BMS based on the a posteriori claim frequency component

Lemaire (1995) developed the design of an optimal BMS based on the number of claims of each policyholder, following a game-theoretic framework introduced by Bichsel (1964) and Bühlmann (1964). Each policyholder has to pay a premium proportional to his own unknown claim frequency. The use of the estimate of the claim frequency instead of the true unknown claim frequency will incur a loss to the insurer. The optimal estimate of the policyholder's claim frequency is the one that minimizes the loss incurred. Lemaire (1995) considered, among other BMS, the optimal BMS obtained using the quadratic error loss function, the expected value premium calculation principle and the Negative Binomial as the claim frequency distribution. Tremblay (1992) considered the design of an optimal BMS using the quadratic error loss function, the zero-utility premium calculation principle and the Poisson-Inverse Gaussian as the claim frequency distribution. Coene and Doray (1996) developed a method of obtaining a financially balanced BMS by minimizing a quadratic function of the difference between the premium for an optimal BMS with an infinite number of classes, weighted by the stationary probability of being in a certain class and by imposing various constraints on the system. Walhin and Paris (1997) obtained an optimal BMS using as the claim frequency distribution the Hofmann's distribution, which encompasses the Negative Binomial and the Poisson-Inverse Gaussian, and also using as a claim frequency distribution a finite Poisson mixture. As we see, all the BMS mentioned above take under consideration only the number of claims of each policyholder disregarding their severity.

1.2. BMS based on the a priori and the a posteriori claim frequency component

The models mentioned above are function of time and of past number of accidents and do not take into consideration the characteristics of each individual.

In this way as mentioned in Dionne and Vanasse (1989), the premiums do not vary simultaneously with other variables that affect the claim frequency distribution. The most interesting example is the age variable. Suppose that age has a negative effect on the expected number of claims, it would imply that insurance premiums should decrease with age. Premium tables derived from BMS based only on the a posteriori criteria, even though are a function of time, do not allow for a variation of age, even though age is a statistically significant variable.

Dionne and Vanasse (1989, 1992) presented a BMS that integrates a priori and a posteriori information on an individual basis. This BMS is derived as a function of the years that the policyholder is in the portfolio, of the number of accidents and of the individual characteristics which are significant for the number of accidents. Picech (1994) and Sigalotti (1994) derived a BMS that incorporates the a posteriori and the a priori classification criteria, with the engine power as the single a priori rating variable. Sigalotti developed a recursive procedure to compute the sequence of increasing equilibrium premiums needed to balance out premiums income and expenditures compensating for the premium decrease created by the BMS transition rules. Picech developed a heuristic method to build a BMS that approximates the optimal merit-rating system. Taylor (1997) developed the setting of a Bonus-Malus scale where some rating factors are used to recognize the differentiation of underlying claim frequency by experience, but only to the extent that this differentiation is not recognized within base premiums. Pinquet (1998) developed the design of optimal BMS from different types of claims, such as claims at fault and claims not at fault.

1.3. Allowance for the severity in BMS

In the models briefly described above the size of loss that each accident incurred is not considered in the design of the BMS. Policyholders with the same number of accidents pay the same malus, irrespectively of the size of loss of their accidents. In this sense the BMS designed in the above way are unfair for the policyholders who had an accident with a small size of loss. Actually as Lemaire (1995) is pointing out all BMS in force throughout the world, with the exception of Korea, are penalizing the number of accidents without taking the severity of such claims into account. In the BMS enforced in Korea the policyholders who had a bodily injury claim pay higher maluses, depending on how severe the accident was, than the policyholders who had a property damage claim. The BMS designed to take severity into consideration include those from Picard (1976) and Pinquet (1997). Picard generalized the Negative Binomial model in order to take into account the subdivision of claims into two categories, small and large losses. In order to separate large from small losses, two options could be used: 1. The losses under a limiting amount are regarded as small and the remainder as large. 2. Subdivision of accidents in those that caused property damage and those that cause bodily injury, penalizing more severely the policyholders who had a bodily injury accident. Pinquet (1997) designed an optimal BMS which makes allowance

for the severity of the claims in the following way: starting from a rating model based on the analysis of number of claims and of costs of claims, two heterogeneity components are added. They represent unobserved factors that are relevant for the explanation of the severity variables. The costs of claims are supposed to follow gamma or lognormal distribution. The rating factors, as well as the heterogeneity components are included in the scale parameter of the distribution. Considering that the heterogeneity also follows a gamma or lognormal distribution, a credibility expression is obtained which provides a predictor for the average cost of claim for the following period.

Our first contribution in this paper is the development of an optimal BMS that takes into account the number of claims of each policyholder and the exact size of loss that these claims incurred. We assumed that the number of claims is distributed according the Negative Binomial distribution and the losses of the claims are distributed according the Pareto distribution, and we have expanded the frame that Lemaire (1995) used to design an optimal BMS based on the number of claims. Applying Bayes' theorem we find the posterior distribution of the mean claim frequency and the posterior distribution of the mean claim size given the information we have about the claim frequency history and the claim size history for each policyholder for the time period he is in the portfolio. For more on this subject we refer to Vrontos (1998).

Our second contribution is the development of a generalized BMS that integrates the a priori and the a posteriori information on a individual basis. In this generalized BMS the premium will be a function of the years that the policyholder is in the portfolio, of his number of accidents, of the size of loss that each of these accidents incurred, and of the significant a priori rating variables for the number of accidents and for the size of loss that each of these claims incurred. We will do this by expanding the frame developed by Dionne and Vanasse (1989, 1992).

Pinquet (1997) is starting from a rating model and then he is adding the heterogeneity components. We design first an optimal BMS based only on the a posteriori classification criteria and then we generalize it in order to take under consideration both the a priori and the a posteriori classification criteria.

2. DESIGN OF OPTIMAL BMS WITH A FREQUENCY AND A SEVERITY COMPONENT BASED ON THE A POSTERIORI CRITERIA

It is assumed that the number of claims of each policyholder is independent from the severity of each claim in order to deal with the frequency and the severity component separately.

2.1. Frequency component

For the frequency component we will use the same structure used by Lemaire (1995). The portfolio is considered to be heterogeneous and all policyholders have constant but unequal underlying risks to have an accident. Consider that

the number of claims k , given the parameter λ , is distributed according to Poisson(λ),

$$P_i(k|\lambda) = \frac{e^{-\lambda} \lambda^k}{k!},$$

$k = 0, 1, 2, 3, \dots$ and $\lambda > 0$ and λ is denoting the different underlying risk of each policyholder to have an accident. Let us assume for the structure function that $\lambda \sim \text{gamma}(a, \tau)$ and λ has a probability density function of the form:

$$u(\lambda) = \frac{\lambda^{a-1} \tau^a \exp(-\tau\lambda)}{\Gamma(a)}, \quad \lambda > 0, a > 0, \tau > 0$$

with mean $E(\lambda) = a/\tau$ and variance $\text{Var}(\lambda) = a/\tau^2$. Then it can be proved that the unconditional distribution of the number of claims k will be Negative Binomial (a, τ), with probability density function

$$P(k) = \binom{k+a-1}{k} \left(\frac{\tau}{1+\tau}\right)^a \left(\frac{1}{1+\tau}\right)^k,$$

mean equal to $E(k) = a/\tau$ and variance equal to $\text{Var}(k) = (a/\tau) (1 + 1/\tau)$. The variance of the Negative Binomial exceeds its mean, a desirable property which is common for all mixtures of Poisson distribution and allows us to deal with data that present overdispersion.

Let us denote as $K = \sum_{i=1}^t k_i$ the total number of claims that a policyholder had in t years, where k_i is the number of claims that the policyholder had in the year i , $i = 1, \dots, t$. We apply the Bayes' theorem and we obtain the posterior structure function of λ for a policyholder or a group of policyholders with claim history k_1, \dots, k_t , denoted as $u(\lambda|k_1, \dots, k_t)$. It is that

$$u(\lambda|k_1, \dots, k_t) = \frac{(\tau + t)^{K+a} \lambda^{K+a-1} e^{-(t+\tau)\lambda}}{\Gamma(a+K)},$$

which is the probability density function of a $\text{gamma}(a+K, t+\tau)$. Using the quadratic error loss function the optimal choice of λ_{t+1} for a policyholder with claim history k_1, \dots, k_t will be the mean of the posterior structure function, that is

$$\lambda_{t+1}(k_1, \dots, k_t) = \frac{a+K}{t+\tau} = \bar{\lambda} \left(\frac{a+K}{a+\bar{\lambda}} \right), \quad \text{where } \bar{\lambda} = \frac{a}{\tau}. \quad (1)$$

From the above it is clear that the occurrence of K accidents in t years just necessitates an update of the parameters of gamma, from a and τ to $a+K$ and $t+\tau$ respectively and the gamma is said to have the important property of the stability of the structure function as the gamma is a conjugate family for the Poisson likelihood.

2.2. Severity component

Let us consider now the severity component. Let x be the size of the claim of each insured. We consider as y the mean claim size for each insured and we assume that the conditional distribution of the size of each claim given the mean claim size, $x|y$, for each policyholder is the one parameter exponential distribution with parameter y , and has a probability density function given by

$$f(x|y) = \frac{1}{y} \cdot e^{-\frac{x}{y}}$$

for $x > 0$ and $y > 0$. The mean of the exponential is $E(x|y) = y$ and the variance is $var(x|y) = y^2$. The mean claim size y is not the same for all the policyholders but it takes different values so it is natural our prior belief for y to be expressed in the form of a distribution. Consider that the prior distribution of the mean claim size y is Inverse Gamma with parameters s and m and probability density function, see for example Hogg and Klugman (1984) given by

$$g(y) = \frac{\frac{1}{m} \cdot e^{-\frac{m}{y}}}{\left(\frac{y}{m}\right)^{s+1} \cdot \Gamma(s)}.$$

The expected value of the mean claim size y will be:

$$E(y) = \frac{m}{s-1}.$$

The unconditional distribution of the claim size x will be equal to:

$$\begin{aligned} P(X = x) &= \int_0^\infty f(x|y) \cdot g(y) dy = \\ &= s \cdot m^s \cdot (x + m)^{-s-1} \end{aligned}$$

which is the probability density of the Pareto distribution with parameters s and m . Thus, one way to generate the Pareto distribution is the following: if it is for the size of each claim given the mean claim size $x|y$ that $x|y \sim Exponential(y)$ and for the mean claim size y of each policyholder that $y \sim Inverse\ Gamma(s, m)$ then it is for the unconditional distribution of the claim size x in the portfolio that $x \sim Pareto(s, m)$. In this way, the relatively tame exponential distribution gets transformed in the heavy-tailed Pareto distribution and instead of using the exponential distribution which is often inappropriate for the modelling of claim severity we are using the Pareto distribution which is often a good candidate for modelling the claim severity. Taking the mean claim size y distributed according the Inverse Gamma, we incorporate in the model the heterogeneity that characterizes the severity of the claims of different policyholders. We should note here that such a generation of the Pareto distribution does not appear for the first time in the actuarial literature. Such a use can be found for example in Herzog (1996). To the best of our knowledge it is the first time it is used in the design of an optimal BMS.

In order to design an optimal BMS that will take into account the size of loss of each claim, we have to find the posterior distribution of the mean claim size y given the information we have about the claim size history for each policyholder for the time period he is in the portfolio. Consider that the policyholder is in the portfolio for t years and that the number of claims he had in the year i is denoted with k_i , by $K = \sum_{i=1}^t k_i$ is denoted the total number of claims he has, and by x_k is denoted the claim amount for the k claim. Then the information we have for his claim size history will be in the form of a vector x_1, x_2, \dots, x_k and the total claim amount for the specific policyholder over the t years that he is in the portfolio will be equal to $\sum_{k=1}^K x_k$. Applying Bayes' theorem we find the posterior distribution of the mean claim size y given the claims size history of the policyholder x_1, \dots, x_k and it is that:

$$g(y|x_1, \dots, x_k) = \frac{\frac{1}{(m + \sum_{k=1}^K x_k)} \cdot e^{-\frac{m + \sum_{k=1}^K x_k}{y}}}{\left(\frac{y}{(m + \sum_{k=1}^K x_k)}\right)^{K+s+1} \cdot \Gamma(K+s)}$$

which is the probability density function of the Inverse Gamma($s + K, m + \sum_{k=1}^K x_k$). This means that the occurrence of K claims in t years with aggregate claim amount equal to $\sum_{k=1}^K x_k$ just necessitates for the distribution of the mean claim size an update of the parameters of the Inverse Gamma from s and m to $s + K$ and $m + \sum_{k=1}^K x_k$ respectively and the Inverse Gamma distribution is said to have the important property of being conjugate with the exponential likelihood. The mean of the posterior distribution of the mean claim size will be:

$$E(x|y) = \frac{m + \sum_{k=1}^K x_k}{s + K - 1} \quad (2)$$

and the predictive distribution of the size of the claim of each insured x will be also a member of the Pareto family.

2.3. Calculation of the Premium according the Net Premium Principle

As shown, the expected number of claims $\lambda_{t+1}(k_1, \dots, k_t)$ for a policyholder or a group of policyholders who in t years of observation have produced K claims with total claim amount equal to $\sum_{k=1}^K x_k$ is given by (1) and the expected claim severity $y_{t+1}(x_1, \dots, x_k)$ is given by (2).

Thus, the net premium that must be paid from that specific group of policyholders will be equal to the product of $\lambda_{t+1}(k_1, \dots, k_t)$ and $y_{t+1}(x_1, \dots, x_K)$, i.e. it will be equal to

$$Premium = \frac{a + K}{t + \tau} \cdot \frac{m + \sum_{k=1}^K x_k}{s + K - 1} \quad (3)$$

In order to find the premium that must be paid we have to know:

1. the parameters of the Negative Binomial distribution a and τ , (see Lemaire (1995) for the estimation of the parameters of the Negative Binomial)
2. the parameters of the Pareto distribution s and m (see Hogg and Klugman (1984) for the estimation of the parameters of the Pareto distribution)
3. the number of years t that the policyholder is under observation,
4. his number of claims K and
5. his total claim amount $\sum_{k=1}^K x_k$.

All of these can be obtained easily and taking under consideration that the negative binomial is often used as a claim frequency distribution and the Pareto as a claim severity distribution this enlarges the applicability of the model.

2.4. Properties of the Optimal BMS with a Frequency and a Severity Component

1. The system is fair as each insured pays a premium proportional to his claim frequency and his claim severity, taking into account, through the Bayes' theorem, all the information available for the time that he is in our portfolio both for the number of his claims and the loss that these claims incur. We use the exact loss x_k that is incurred from each claim in order to have a differentiation in the premium for policyholders with the same number of claims, not just a scaling with the average claim severity of the portfolio.
2. The system is financially balanced. Every single year the average of all premiums collected from all policyholders remains constant and equal to

$$P = \frac{a}{\tau} \frac{m}{s-1} \quad (4)$$

In order to prove this it is enough to show, considering that the claim frequency and the claim severity are independent components, that:

$$E_{\Lambda}[\Lambda] = E[E[\lambda | k_1, \dots, k_t]] = \frac{a}{\tau}$$

and that

$$E_Y[Y] = E[E[y | x_1, \dots, x_k]] = \frac{m}{s-1}.$$

A proof of the first can be found in Lemaire (1995), and of the second in Vrontos (1998).

3. In the beginning all the policyholders are paying the same premium which is equal to (4).
4. The more accidents caused and the more the size of loss that each claim incurred the higher the premium.
5. The premium always decreases when no accidents are caused.
6. The drivers who had a claim with small loss will have one more reason to report the claim as they will know that the size of the claim will be taken into consideration and they will not have to pay the same premium with somebody which had an accident with a big loss. In this way the phenomenon of bonus hunger will have a decrease and the estimate of the actual claim frequency will be more accurate.
7. The severity component is introduced in the design of a BMS which from a practical point of view is more crucial than the number of claims for the insurer since it is the component that determines the expenses of the insurer from the accidents and thus the premium that must be paid.
8. The estimator of the mean of severity may not be robust and therefore it is prone to be affected by variation. For practical use a more robust estimator could be used. (i.e. cutting of the data, M-estimator).

3. DESIGN OF OPTIMAL BMS WITH A FREQUENCY AND A SEVERITY COMPONENT BASED BOTH ON THE A PRIORI AND THE A POSTERIORI CRITERIA

Dionne and Vanasse (1989, 1992) presented a BMS that integrates risk classification and experience rating based on the number of claims of each policyholder. This BMS is derived as a function of the years that the policyholder is in the portfolio, of the number of accidents and of the significant – for the number of accidents – individual characteristics. We extend this model by introducing the severity component. We propose a generalized BMS that integrates a priori and a posteriori information on an individual basis based both on the frequency and the severity component. This generalized BMS will be derived as a function of the years that the policyholder is in the portfolio, of the number of accidents, of the exact size of loss that each one of these accidents incurred, and of the significant individual characteristics for the number of accidents and for the severity of the accidents. Some of the a priori rating variables that could be used are the age, the sex and the place of residence of the policyholder, the age, the type and the cubic capacity of the car, etc. As already said one of the reasons for the development of a generalized model which integrates a priori and a posteriori information is that premiums should vary simultaneously with the variables that affect the distribution of the number of claims and the size of loss distribution.

The premiums of the generalized BMS will be derived using the following multiplicative tariff formula:

$$Premium = GBM_F * GBM_S \quad (5)$$

where GBM_F denotes the generalized BMS obtained when only the frequency component is used and GBM_S denotes the generalized BMS obtained when only the severity component is used.

3.1. Frequency Component

The generalized BMS obtained with the frequency component GBM_F will be developed according to Dionne and Vanasse (1989, 1992). Consider an individual i with an experience of t periods. Assume that the number of claims of the individual i for period j , denoted as K_i^j , follows the Poisson distribution with parameter λ_i^j , and K_i^j are independent. The expected number of claims of the individual i for period j is then denoted by λ_i^j and consider that it is a function of the vector of h individual's characteristics, denoted as $c_i^j = (c_{i,1}^j, \dots, c_{i,h}^j)$, which represent different a priori rating variables. Specifically assume that $\lambda_i^j = \exp(c_i^j \beta^j)$, where β^j is the vector of the coefficients. The non-negativity of λ_i^j is implied from the exponential function. The probability specification becomes

$$P(K_i^j = k) = \frac{e^{-\exp(c_i^j \beta^j)} (\exp(c_i^j \beta^j))^k}{k!}.$$

In this model we assume that the h individual characteristics provide enough information for determining the expected number of claims. The vector of the parameters β^j can be obtained by maximum likelihood methods, see Hausmann, Hall and Griliches (1984) for an application. However, if one assumes that the a priori rating variables do not contain all the significant information for the expected number of claims then a random variable ε_i has to be introduced into the regression component. As Gourieroux, Montfort and Trognon (1984a), (1984b) suggested, we can write

$$\begin{aligned} \lambda_i^j &= \exp(c_i^j \beta^j + \varepsilon_i) = \\ &= \exp(c_i^j \beta^j) u_i, \end{aligned}$$

where $u_i = \exp(\varepsilon_i)$, yielding a random λ_i^j . If we assume that u_i follows a gamma distribution with $E(u_i) = 1$ and $Var(u_i) = 1/a$, the probability specification becomes

$$P(K_i^j = k) = \frac{\Gamma(k+a)}{k! \Gamma(a)} \left[\frac{\exp(c_i^j \beta^j)}{a} \right]^k \left[1 + \frac{\exp(c_i^j \beta^j)}{a} \right]^{-(k+a)},$$

which is a negative binomial distribution with parameters a and $\exp(c_i^j \beta^j)$. It can be shown that the above parameterization does not affect the results if there is a constant term in the regression. We choose $E(u_i) = 1$ in order to have $E(\varepsilon_i) = 0$. Then

$$E(K_i^j) = \exp(c_i^j \beta^j) \text{ and } Var(K_i^j) = \exp(c_i^j \beta^j) \left[1 + \frac{\exp(c_i^j \beta^j)}{a} \right].$$

The interesting reader can see for more on the Negative Binomial regression Lawless (1987). The insurer needs to calculate the best estimator of the expected number of accidents at period $t + 1$ using the information from past experience for the claim frequency over t periods and of known individual characteristics

over the $t+1$ periods. Let us denote this estimator as $\hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^{t+1})$. Using the Bayes theorem one finds that the posterior structure function for a policyholder with K_i^1, \dots, K_i^t claim history and c_i^1, \dots, c_i^{t+1} characteristics is gamma with updated parameters $(a + \sum_{j=1}^t K_i^j, \frac{a}{\exp(c_i^j \beta^j)} + t)$. Using the classical quadratic loss function one can find that the optimal estimator given the observation of K_i^1, \dots, K_i^t and C_i^1, \dots, C_i^{t+1} , is equal to:

$$\begin{aligned} & \hat{\lambda}_i^{t+1}(K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^{t+1}) \\ &= \int_0^\infty \lambda_i^{t+1}(K_i^{t+1}, u_i) f(\lambda_i^{t+1} | K_i^1, \dots, K_i^t; c_i^1, \dots, c_i^t) d\lambda_i^{t+1} = \\ &= \frac{1}{t} \sum_{j=1}^t \exp(c_i^j \beta^j) \left[\frac{a + \sum_{j=1}^t K_i^j}{a + t \exp(c_i^j \beta^j)} \right], \end{aligned}$$

where $\sum_{j=1}^t K_i^j$ denotes the total number of claims of policyholder i in t periods. When $t = 0$, $\hat{\lambda}_i^1 \equiv \exp(C_i^1 \beta^1)$ which implies that only a priori rating is used in the first period. Moreover when the regression component is limited to a constant β_0 , one obtains the well-known univariate without regression component model, see Lemaire (1995), Ferreira (1974).

Now we will deal with the generalized bonus-malus factor obtained when the severity component is used. It will be developed in the following way.

3.2. Severity Component

Consider an individual i with an experience of t periods. Assume that the number of claims of the individual i for period j is denoted as K_i^j , the total number of claims of the individual i is denoted as K and by $X_{i,k}^j$ is denoted the loss incurred from his claim k for the period j . Then, the information we have for his claim size history will be in the form of a vector $X_{i,1}, X_{i,2}, \dots, X_{i,K}$, and the total claim amount for the specific policyholder over the t periods that he is in the portfolio will be equal to $\sum_{k=1}^K X_{i,k}$. We assume that $X_{i,k}^j$ follows an exponential distribution with parameter y_i^j . The parameter y_i^j denotes the mean or the expected claim severity of a policyholder i in period j . As we have already said, all policyholders do not have the same expected claim severity, their cost for the insurer is different and thus it is fair each policyholder to pay a premium proportional to his mean claim severity. Consider that the expected claim severity is a function of the vector of the h individual's characteristics, denoted as $d_i^j = (d_{i,1}^j, \dots, d_{i,h}^j)$, which represent different a priori rating variables. Specifically assume that $y_i^j = \exp(d_i^j \gamma^j)$, where γ is the vector of the

coefficients. The non-negativity of y_i^j is implied from the exponential function. The probability specification becomes

$$P(X_{i,k}^j = x) = \frac{1}{\exp(d_i^j \gamma^j)} \cdot e^{-\frac{x}{\exp(d_i^j \gamma^j)}}.$$

In this model we assume that the h individual characteristics provide enough information for determining the expected claim severity. However if one assumes that the a priori rating variables do not contain all the significant information for the expected claim severity then a random variable z_i has to be introduced into the regression component. Thus we can write

$$\begin{aligned} y_i^j &= \exp(d_i^j \gamma^j + z_i) = \\ &= \exp(d_i^j \gamma^j) w_i, \end{aligned}$$

where $w_i = \exp(z_i)$, yielding a random y_i^j . If we assume that w_i follows an inverse gamma($s, s-1$) distribution with

$$E(w_i) = 1 \text{ and } Var(w_i) = \frac{1}{s-2}, \quad s > 2,$$

then y_i^j follows inverse gamma($s, (s-1)\exp(c_i^j \gamma^j)$) and the probability specification for $X_{i,k}^j$ becomes

$$P(X_{i,k}^j = x) = s \cdot [(s-1)\exp(d_i^j \gamma^j)]^s \cdot (x + (s-1)\exp(d_i^j \gamma^j))^{-s-1}$$

which is a Pareto distribution with parameters s and $(s-1)\exp(c_i^j \gamma^j)$. It can be shown that the above parameterization does not affect the results if there is a constant term in the regression. We choose $E(w_i) = 1$ in order to have $E(z_i) = 0$. We also have

$$E(X_{i,k}^j) = \exp(d_i^j \gamma^j) \text{ and } Var(X_{i,k}^j) = \frac{[(s-1)\exp(d_i^j \gamma^j)]^2}{s-1} \left(\frac{2}{s-2} - \frac{1}{s-1} \right).$$

The insurer needs to calculate the best estimator of the expected claim severity at period $t + 1$ using the information from past experience for the claim severity over t periods and of known individual characteristics over the $t + 1$ periods. Let us denote this estimator as $\hat{y}_i^{t+1}(X_{i,1}, \dots, X_{i,K}; d_i^1, \dots, d_i^{t+1})$. Using the Bayes theorem the posterior distribution of the mean claim severity for a policyholder with claim sizes $X_{i,1}, \dots, X_{i,K}$ in t periods and characteristics d_i^1, \dots, d_i^{t+1} is inverse gamma with the following updated parameters:

$$IG(s + K, (s-1)\exp(d_i^j \gamma^j) + \sum_{k=1}^K X_{i,k}).$$

Using the classical quadratic loss function one can find that the optimal estimator of the mean claim severity for the period $t + 1$ given the observation of

$X_{i,1}, \dots, X_{i,K}$ and d_i^1, \dots, d_i^{t+1} , is the mean of the posterior inverse gamma and thus it is equal to

$$\begin{aligned} \hat{y}_i^{t+1}(X_{i,1}, \dots, X_{i,K}; d_i^1, \dots, d_i^{t+1}) &= \\ &= \int_0^\infty y_i^{t+1}(X_i^{t+1}, w_i) f(y_i^{t+1} | X_{i,1}, \dots, X_{i,K}; d_i^1, \dots, d_i^{t+1}) dy_i^{t+1} = \\ &= \frac{(s-1)^{\frac{1}{t}} \sum_{j=1}^t \exp(d_i^j \gamma^j) + \sum_{k=1}^K X_{i,k}}{s + K - 1} \end{aligned}$$

When $t = 0$, which implies that only a priori rating is used in the first period it is $\hat{y}_i^1 = \exp(d_i^1 \gamma^1)$.

3.3. Calculation of the premiums of the Generalized BMS

Now we are able to compute the premiums of the generalized optimal BMS based both on the frequency and the severity component. As we said the premiums of the generalized optimal BMS will be given from the product of the generalized BMS based on the frequency component and of the generalized BMS based on the severity component. Thus it will be

$$\begin{aligned} \text{Premium} &= \text{GBM}_F * \text{GBM}_S = \\ &= \frac{1}{t} \sum_{j=1}^t \exp(c_i^j \beta^j) \left[\frac{a + \sum_{j=1}^t K_i^j}{a + t \exp(c_i^j \beta^j)} \right] \frac{(s-1)^{\frac{1}{t}} \sum_{j=1}^t \exp(d_i^j \gamma^j) + \sum_{k=1}^K X_{i,k}}{s + K - 1}. \quad (6) \end{aligned}$$

3.4. Properties of the Generalized BMS

1. It is fair since it takes into account the number of claims, the significant a priori rating variables for the number of claims, the claim severity and the significant a priori rating variables for the claim severity for each policyholder.
2. It is financially balanced for the insurer. Each year the average premium will be equal to

$$P = \exp(c_i^{t+1} \beta^{t+1}) \exp(d_i^{t+1} \gamma^{t+1}) \quad (7)$$

In order to prove the above equation and assuming that claim frequency and the claim severity component are independent it is sufficient to show that

$$E \left[\hat{\lambda}_i^{t+1} (K_i^1, \dots; K_i^t; c_i^1, \dots, c_i^{t+1}) \right] = \exp(c_i^{t+1} \beta^{t+1})$$

and that

$$E \left[\hat{y}_i^{t+1} (K_i^1, \dots, K_i^t; d_i^1, \dots, d_i^{t+1}) \right] = \exp(c_i^{t+1} \gamma^{t+1}).$$

3. All the properties we mentioned for generalized BMS without the a priori rating variables hold for this BMS as well. In the beginning all the policyholders with the same characteristics are paying the same premium which is equal to (7).
4. The more accidents are caused and the more the size of loss that each claim incurred the higher is the premium.
5. The premium always decreases when no accidents are caused.
6. This generalized BMS could lead to a decrease of the phenomenon of bonus hunger.
7. The severity component, which is more crucial than the number of claims for the insurer, is introduced in the design of the generalized BMS.
8. The premiums vary simultaneously with the variables that affect the distribution of the number of claims and the size of loss distribution.

3.5. Estimation

The premiums will be calculated according (6). We have to know the number of the years t that the policyholder is in the portfolio, his total number of accidents in t years and his aggregate claim amount in t years.

For the frequency component of the generalized BMS we have to estimate the parameters of the negative binomial regression model, that is the dispersion parameter a and the vector β . This can be done using the maximum likelihood method. For more on the negative binomial regression the interested reader can see Lawless (1987), Gourieroux, Montfort and Trognon (1984a) and Gourieroux, Montfort and Trognon (1984b).

For the severity component of the generalized BMS we have to estimate s and γ^j . We will achieve this using the quasi-likelihood and according to Renshaw (1994). Renshaw is using the generalized linear models as a modelling tool for the study of the claim process in the presence of covariates. He is giving special attention to the variety of probability distributions that are available and to the parameter estimation and model fitting techniques that can be used for the claim frequency and the claim severity process based on the concepts of quasi-likelihood and extended quasi-likelihood.

Following Renshaw (1994) consider the following scheme. The mean claim severity is denoted by y_u , categorized over a set of units u . The data take the form (u, k_u, x_u) where x_u denotes the claim average in cell u based on n_u claims. Independence of n_u and x_u is assumed. The units $u \equiv (i_1, i_2, \dots)$ are a cross-classified grid of cells defined for preselected levels of appropriate covariates, often rating factors. Denoting the underlying expected claim severity in cell u by μ_u and assuming the independence of individual claim amounts, the cell means are modelled as the responses of a GLM with $E(x_u) = \mu_u$ and $Var(X_u) = \phi V(\mu_u)/n_u$. Covariates defined on $\{u\}$ enter through a linear predictor, linked

to the mean μ_u . For those unfamiliar with the generalized linear models we refer to the classical text of McCullagh and Nelder (1989). In McCullagh and Nelder (1983, 1989) a re-analysis of the celebrated car insurance data of Baxter, Coutts and Ross (1979), based on independent gamma distributed claim amounts can be found.

Let us focus now on the Pareto distribution with parameters s and $(s-1)\exp(c_i^j \gamma^j)$ and density

$$P(X_{i,k}^j = x) = s \cdot [(s-1)\exp(d_i^j \gamma^j)]^s \cdot (x + (s-1)\exp(d_i^j \gamma^j))^{-s-1}$$

We have that

$$E(X_{i,k}^j) = \exp(d_i^j \gamma^j) \text{ and } Var(X_{i,k}^j) = \frac{[(s-1)\exp(d_i^j \gamma^j)]^2}{s-1} \left(\frac{2}{s-2} - \frac{1}{s-1} \right), \quad s > 2.$$

Introducing the reparameterisation:

$$\mu = \exp(d_i^j \gamma^j) \text{ and } \varphi = \frac{s}{s-2},$$

a 1:1 mapping $(s, (s-1)\exp(c_i^j \gamma^j)) \rightarrow (\mu, \varphi)$ with domain $R_{>2} \times R_{>0}$ and image set $R_{>2} \times R_{>1}$ implies that we can construct a GLM based on independent Pareto distributed claim amounts for which the mean responses, X_u , satisfy mean $\mu_u = E(X_u)$, variance function $V(\mu_u) = \mu_u^2$, scale parameter $\varphi > 1$ and weights n_u so that $Var(X_u) = \varphi V(\mu_u) / n_u$. Apart from the mild extra constraint on the scale parameter, these details are identical to those of the GLM based on independent gamma responses and the two different modelling assumptions lead to essentially identical GLMs. They differ only in the parameter estimation method. In the case of gamma response we use maximum likelihood method and in the case of Pareto response we use maximum quasi-likelihood.

4. APPLICATION

4.1. Description of the Data

The models discussed are applied in a data set that one Greek insurance company provided us. The data set consists of 46420 policyholders. The mean of the claim frequency is 0.0808 and the variance is 0.10767. The a priori rating variables were age and sex of the driver, BM class and the horsepower of the car. The drivers were divided in three categories according their age. Those aged between 28-45, those between 46-55 and those aged between 18-27 or higher than 55. The drivers were also divided in three categories according the horsepower of their car. Those who had a car with a horsepower between 0-33, between 34-66 and between 67-99. The drivers were also divided in three categories according their BM class. The current Greek BMS has 16 classes, from 5 to 20. The malus zone includes classes from 12 to 20, the bonus zone includes classes 5 to 8 and the neutral zone includes classes from 9 to 11.

We fitted the Negative Binomial distribution on the number of claims and the Pareto distribution on the claim sizes. We will find the premiums determined from the optimal BMS based on the a posteriori frequency component, the premiums determined from the optimal BMS based on the a posteriori frequency and severity component and the premiums determined from the optimal BMS with a frequency and a severity component based both on the a priori and the a posteriori criteria.

4.2. Optimal BMS based on the a posteriori frequency component

We apply the Negative Binomial distribution. The maximum likelihood estimators of the parameters are $\hat{a} = 0.228$ and $\hat{\tau} = 2.825$. We will find first the optimal BMS based only on the frequency component following Lemaire (1995). The BMS will be defined from (1) and is presented in Table 1. This optimal BMS can be considered generous with good drivers and strict with bad drivers. For example, the bonuses given for the first claim free year are 26% of the basic premium. Drivers who have one accident over the first year will have to pay a malus of 298% of the basic premium.

TABLE 1.
OPTIMAL BMS BASED ON THE A POSTERIORI FREQUENCY COMPONENT

<i>Year</i>	<i>Number of claims</i>						
	t	0	1	2	3	4	5
0		100					
1		74	398	722	1046	1370	1693
2		59	315	572	829	1086	1342
3		48	261	474	687	899	1112
4		41	223	404	586	768	949
5		36	194	353	511	669	828
6		32	172	313	453	594	734
7		29	155	281	407	533	659

4.3. Optimal BMS based on the a posteriori frequency and severity component

Let us see the implementation of an optimal BMS based both on the frequency and the severity component. We fit the Pareto distribution to the claim sizes and we find the maximum likelihood estimates of s and m . It is $\hat{s} = 2.382$ and $\hat{m} = 493927.087$. In order to find the premium that must be paid we have to know the age of the policy, the number of claims he has done in these years and the aggregate claim amount. The steps that must be followed in order to find this optimal BMS are:

1. We find the age of the policy t .
2. We find the total number of claims k that the policyholder has done in t years.

3. We find the aggregate claim amount for the policyholder, $\sum_{k=1}^K x_k$
4. We compute the premiums using (3).
5. We go to the table with the specific total claim amount and we find the premium that corresponds to k claims in t years of observation.

The Bonus-Malus System determined in the above way is presented in the following tables. Here we will illustrate only the cases that the aggregate claim amount of a policyholder is equal to 250000 drs, and 1000000 drs. It is obvious that we can use the above formula with any value that the aggregate claim amount can have. We use these values of the aggregate claim amount for brevity. In the following tables we will use the actual values, the premiums are not divided with the premium when $t = 0$, as it will be interesting to see the variation of the premiums paid for various number of claims and claim sizes in comparison not with the premium paid when $t = 0$ but with the specific claim sizes. This is the basic advantage of this BMS in comparison with the one that takes under consideration only the frequency component, the differentiation according the severity of the claim. Of course the percentage change in the premium after on or more claims could be also interesting.

Let's see an example in order to understand better how such BMS work. In Table 3 we can see the premiums that must be paid for various number of claims when the age of the policy is up to 7 years. For example a policyholder with one accident of claim size 250000 drs in the first year of observation will pay 100259 drs (see Table 2). If the second year of observation he has an accident with claim size 750000 drs, then, a surcharge will be enforced and he will have to pay 203964 drs, which is the premium for two accidents of aggregate claim amount 1000000 drs in two years of observation (see Table 3). If in the third year he does not have an accident, he will have a reduction in the premium because he had a claim free year and he will pay 168947 drs, which is the premium for two accidents of aggregate claim amount 1000000 drs in three years of observation (see Table 3).

TABLE 2.

OPTIMAL BMS BASED ON THE A POSTERIORI FREQUENCY AND SEVERITY COMPONENT –
TOTAL CLAIM SIZE OF 250000.

<i>Year</i>	<i>Number of claims</i>						
	<i>t</i>	0	1	2	3	4	5
0		28841					
1		21300	100259	128122	143269	152788	159323
2		16886	79479	101567	113575	121121	126302
3		13987	65834	84130	94076	100327	104618
4		11937	56188	71803	80292	85626	89289
5		10412	49007	62627	70031	74683	77878
6		9232	43454	55530	62095	66220	69053
7		8292	39031	49878	55775	59480	62025

TABLE 3.

OPTIMAL BMS BASED ON THE A POSTERIORI FREQUENCY AND SEVERITY COMPONENT –
TOTAL CLAIM SIZE OF 1000000.

<i>Year</i>	<i>Number of claims</i>						
	t	0	1	2	3	4	5
0		28841					
1		21300	201336	257290	287708	306823	319947
2		16886	159607	203964	228077	243230	253634
3		13987	132206	168947	188921	201472	210091
4		11937	112834	144192	161239	171952	179307
5		10412	98414	125765	140633	149976	156392
6		9232	87262	111513	124697	132982	138670
7		8292	78380	100163	112005	119446	124556

It is obvious that this optimal BMS allows the discrimination of the premium with respect to the severity of the claims. Table 4 shows the premiums that must be paid when the policyholder is observed for the first year of his presence in the portfolio, his number of accidents range from 1 to 5 and the aggregate claim amount of his accidents ranges from 250000 to 4000000 dr. A policyholder who had one claim with claim size 250000 will have to pay a premium of 100259 drs, a policyholder who had one claim with claim size 500000 will have to pay a premium of 133951 drs and a policyholder who had one claim with claim size 1000000 will have to pay a premium of 201336 drs.

TABLE 4.

COMPARISON OF PREMIUMS FOR VARIOUS NUMBER OF CLAIMS AND CLAIM SIZES
IN THE FIRST YEAR OF OBSERVATION.

<i>Claim Size</i>	<i>Number of claims</i>				
	1	2	3	4	5
250000	100259	181903	263547	345191	426835
500000	133951	243032	352113	461194	570275
1000000	201336	365291	529246	693201	857155
2000000	336106	609808	883511	1157213	1430915
3000000	470876	854326	1237775	1621225	2004675
4000000	605646	1098843	1592040	2085237	2578434

For more on such a system the interesting reader can see Vrontos (1998).

4.4. Generalized optimal BMS with a frequency and a severity component based both on the a priori and the a posteriori classification criteria

Let us calculate now the premiums of the generalized optimal BMS based both on the frequency and the severity component when both the a priori and the a posteriori rating variables are used. As we said the premiums of the generalized optimal BMS will be given from the product of the generalized BMS based on the frequency component, GBM_F , and of the generalized BMS based on the severity component, GBM_S .

Implementing the negative binomial regression model we estimate the dispersion parameter a and the vector β of the significant a priori rating variables for the number of claims. We found that many a priori rating variables are significant for the number of claims. These are the BM class, the age and the sex of the driver and the interaction between age and sex. In the multivariate model $\hat{a} = 47.96$ is larger, than in the univariate negative binomial model where we had $\hat{a} = 0.228$. This result indicates that part of the variance is explained by the a priori rating variables in the multivariate model. The estimates of the vector β can be found in the appendix. The parameters of GBM_S , that is the parameter of the Pareto s , and the vector parameter γ of the significant for the claim severity a priori rating variables d_i^j , are found using the quasi-likelihood method. The significant a priori characteristics for the claim severity are the age and the sex of the driver, the BM class, the horsepower of the car, the interaction between age and sex and the interaction between age and class. The premiums are calculated using (6). Below we can see the premiums for different categories of policyholders.

Let us examine two groups of policyholders which have the following common characteristics. They belong in the malus zone, their car's horsepower is between 67 and 99, and their age is between 28 and 45. If the policyholder is a man he will have to pay the following premiums after one or more accidents of total claim amount 500000 in the first year.

TABLE 5.

MEN, AGE 28-45, MALUS-ZONE, HORSEPOWER 67-99.

Year	Number of claims					
	0	1	2	3	4	5
t	0	1	2	3	4	5
0	100413					
1	39609	279852	357626	399906	426474	444718
2	24670	174304	222745	249079	265627	276990
3	17914	126568	161743	180865	192881	201132
4	14063	99358	126970	141981	151414	157891
5	11574	81777	104503	116858	124622	129953
6	9834	69482	88792	99289	105886	110415
7	8549	60401	77187	86313	92047	95985

If the policyholder is a woman with the above characteristics she will have to pay the following premiums after one or more accidents of total claim amount 500000 in the first year.

TABLE 6.
WOMEN, AGE 28-45, MALUS-ZONE, HORSEPOWER 67-99.

<i>Year</i>	<i>Number of claims</i>						
	t	0	1	2	3	4	5
0		75096					
1		28632	248946	318131	355742	379376	395605
2		17688	153791	196531	219766	234367	244392
3		12797	111263	142184	158994	169557	176810
4		10024	87160	111383	124551	132826	138508
5		8240	71641	91551	102374	109176	113846
6		6994	60813	77713	86901	92674	96639
7		6076	52828	67510	75491	80506	83950

We notice that men have to pay higher premiums than women. We saw an example of premiums obtained with generalized optimal BMS with a frequency and a severity component based both on the a priori and the a posteriori classification criteria. Other combinations of a priori characteristics could be used and also higher total claim amounts.

It is interesting to compare this BMS with the one obtained when the only the a posteriori frequency and severity component are used. Using this BMS we saw from Table 4 that a policyholder with one accident with claim size of 500000 drs in one year has to pay 133951 drs. Using the generalized optimal BMS with a frequency and a severity component based both on the a priori and the a posteriori classification criteria, a man, age 28-45, who belongs to the malus zone, with a car with horsepower between 67-99 for one accident of claim size 500000 drs in one year will has to pay 279852 drs, while a woman, age 28-45, who belongs to the malus zone, with a car with horsepower between 67-99 for one accident of claim size 500000 drs in one year will has to pay 248946 drs. This system is more fair since it considers all the important a priori and a posteriori information for each policyholder both for the frequency and the severity component in order to estimate his risk to have an accident and thus it permits the differentiation of the premiums for various number of claims and for various claim amounts based on the expected claim frequency and expected claim severity of each policyholder as these are estimated both from the a priori and the a posteriori classification criteria.

5. CONCLUSIONS

We developed in this paper the design of an optimal BMS based both on the a posteriori frequency and the a posteriori severity component. We did this by fitting the Negative Binomial distribution in the claim frequency and the Pareto distribution on the claim severity, extending the – classical in the BMS literature – model of Lemaire (1995) which used the Negative Binomial distribution. The optimal BMS obtained has all the attractive properties of the optimal BMS designed by Lemaire, furthermore it allows the differentiation of the premiums according to the claim severity and in this way it is more fair for the policyholders and it is obtained in a very natural context according to our opinion.

Moreover, we developed the design of a generalized optimal BMS with a frequency and a severity component based both on the a priori and the a posteriori classification criteria extending the model developed by Dionne and Vanasse (1989, 1992) which was based only on the frequency component. The BMS obtained has all the attractive properties of the one obtained by Dionne and Vanasse (1989, 1992) and furthermore it allows the differentiation of the premiums utilizing the severity component in a very natural context. This generalized BMS takes into consideration simultaneously the important individual's characteristics for the claim frequency, the important individual's characteristics for the claim severity, the claim frequency and the claim severity of each accident for each policyholder.

An interesting topic for further research could be the extension of the two above BMS for different claim frequency and claim severity distributions.

REFERENCES

- BAXTER, L.A., COUTTS S.M. and ROSS G.A.F. (1979) *Applications of Linear Models in Motor Insurance*. 21st International Congress of Actuaries.
- BICHSEL, F. (1964) Erfahrung-Tarifierung in der Motorfahrzeug-haftpflichtversicherung, *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, 119-129.
- BUHLMANN, H. (1964) Optimale Prämienstufensysteme, *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker* 193-213.
- COENE, G. and DORAY, L.G. (1996) A Financially Balanced Bonus-Malus System. *Astin Bulletin* **26**, 107-115.
- DIONNE, G. and VANASSE, C. (1989) A generalization of actuarial automobile insurance rating models: the negative binomial distribution with a regression component, *Astin Bulletin* **19**, 199-212.
- DIONNE, G. and VANASSE, C. (1992) Automobile insurance ratemaking in the presence of asymmetrical information, *Journal of Applied Econometrics* **7**, 149-165.
- FERREIRA, J. (1974) The long term effect of merit-rating plans on individual motorists, *Operations Research*, **22**, 954-978.
- GOURIEROUX C., MONTFORT A. and TROGNON A. (1984a) Pseudo maximum likelihood methods: theory, *Econometrica*, **52**, 681-700.
- GOURIEROUX C., MONTFORT A. and TROGNON A. (1984b) Pseudo maximum likelihood methods: application to Poisson models, *Econometrica*, **52**, 701-720.
- HAUSMANN J.A., HALL B.H. and GRILICHES Z. (1984) Econometric models for count data with an application to the patents-R&D relationship, *Econometrica*, **46**, 1251-1271.

- HERZOG, T. (1996) *Introduction to Credibility Theory*. Actex Publications, Winstead.
- HOGG, R.V. and KLUGMAN, S.A. (1984) *Loss Distributions*, John Wiley & Sons, New York.
- MCCULLAGH, P. and NELDER, J.A. (1983, 1989) *Generalized Linear Models*. Chapman & Hall.
- LAWLESS, J.F. (1987) Negative Binomial Distribution and Mixed Poisson Regression, *Canadian Journal of Statistics*, **15**, 3, 209-225.
- LEMAIRE, J. (1995) *Bonus-Malus Systems in Automobile Insurance*, Kluwer Academic Publishers, Massachusetts.
- PICARD, P. (1976) Generalisation de l'étude sur la survenance des sinistres en assurance automobile, *Bulletin Trimestriel de l'Institut des Actuaire Français*, 204-267.
- PICECH, L. (1994) The Merit-Rating Factor in a Multiplicating Rate-Making model. *Astin Colloquium*, Cannes.
- PINQUET (1997) Allowance for Costs of Claims in Bonus-Malus Systems, *Astin Bulletin*, **27**, 33-57.
- PINQUET J. (1998) Designing Optimal Bonus-Malus Systems From Different Types of Claims. *Astin Bulletin*, **28**, 205-220.
- RENSHAW A.E. (1994) Modelling The Claims Process in the Presence of Covariates. *Astin Bulletin*, **24**, 265-285.
- SIGALOTTI (1994) Equilibrium Premiums in a Bonus-Malus System, *Astin Colloquium*, Cannes.
- TAYLOR G. (1997) Setting A Bonus-Malus Scale in the Presence of Other Rating Factors. *Astin Bulletin*, **27**, 319-327.
- TREMBLAY, L. (1992) Using the Poisson Inverse Gaussian in Bonus-Malus Systems, *Astin Bulletin* **22**, 97-106.
- VRONTOS S. (1998) *Design of an Optimal Bonus-Malus System in Automobile Insurance*. Msc. Thesis, Department of Statistics, Athens University of Economics and Business, ISBN: 960-7929-21-7.
- WALHIN J.F. and PARIS J. (1999) Using Mixed Poisson distributions in connection with Bonus-Malus Systems. *Astin Bulletin* **29**, 81-99.

NICHOLAS E. FRANGOS AND SPYRIDON D. VRONTOS
Department of Statistics
Athens University of Economics and Business
Patission 76, 10434, Athens, Greece
E-mail: nef@aueb.gr and svrontos@aueb.gr