On optimal monotone Bonus-Malus systems where the premiums depend both on the number and on the severity of the claims

Heikki Bonsdorff

Abstract

Bonus-Malus systems depending both on the number and on the severity of claims are dealt with (cf. Bonsdorff, 2005). We consider the existence and uniqueness of an optimal BMS satisfying a monotonicity condition, guaranteeing that the premiums do not increase after a claims-free year. Existence is considered under several optimality criteria, analogous to those in the classic Bonus-Malus systems. Focusing on one of these criteria, we show that under a general condition the optimal solution is essentially unique. An algorithm for determining the optimal solution is presented. As a special case the results are obtained also for the classic case.

Keywords: Optimal Bonus-Malus system, Severity of claims, General Markov chains, Tihonov theorem, Convex optimization, Direct methods in the calculus of variations.

1. Introduction

The aim of this paper is to consider optimal Bonus-Malus systems (BMS), where the premiums depend, besides on the number of the claims, also on the severity of the claims. When studying the optimality, we restrict ourselves to the situation where the premiums are monotone with respect to the bonus classes (called bonus coefficients here) in order to exclude the possibility that after a claims-free year the premium increases. The existence of such premiums will be shown, under general conditions, with respect to criteria which are analogous to common criteria in the classic case. Concentrating on one of these criteria, we consider the uniqueness of the optimal solution and the calculation of the solution. We will show that under certain natural conditions the optimal solution is essentially unique. We also present an algorithm for the calculation of the solution. As a special case we get, as corollaries, corresponding results for the classic case.

By the classic case we mean a BMS with a finite number K of bonus classes, the transitions depending on the number of claims of the previous year. The claims number process of an individual policy is assumed to be Poisson(λ). Under assumption that the BMS has a superbonus class the Markov chain of the transitions between the bonus classes possesses a unique invariant probability distribution (π_i), $i = 1, \dots, K$, to which the n'th step transition probabilities converge.

Assuming that the parameter λ obeys a probability distribution in the insurer's portfolio, the optimality of the premiums can be considered.

Norberg (1976) considers the optimality criterion

$$(1.1) \qquad \int \Big(\sum_{i=1}^{K} (y_i - \lambda)^2 \pi_i(\lambda) \Big) dU(\lambda) = \min(y_1, \dots, y_K),$$

where $y_i,\ i = 1, \dots, K$ are the premiums of a policy in class i, $(\pi_i(\lambda)),\ i = 1, \dots, K$ is the invariant probability distribution associated to the parameter λ and $U(\lambda)$ is the probability distribution of the parameters λ in the portfolio. Norberg (1976) solves the optimal solution $(y_1, \dots, y_K)$ under condition (1.1.) In a pioneering paper, Pesonen (1962) considered the optimality problem attaining the solution in a special case.

Also other criteria are used in the classic case. Verico (2002) suggests the criterion

$$(1.2) \qquad \int \Big(\sum_{i=1}^{K} (y_i \pi_i(\lambda) - \lambda \Big)^2 dU(\lambda) = \min(y_1, \dots, y_K).$$

Instead of the quadratic deviation in (1.1) and (1.2), also absolute deviation is used, see, e.g., Heras et al. (2004). Borgan et al. (1981) used a criterion which takes into account the state of the BMS after a finite time from the starting point. For other criteria, see Lemaire (1995).

Gilde and Sundt (1989) show that in a Bonus-Malus system previously used in Norway, the optimal solution of premiums is such that in certain situations the premium increases after a claims-free year. The Norwegian system was constructed in a "natural way", especially so that after a claims-free year there occurs a transition to a "bigger" (better) class and after at least one claim to a "smaller" (worse) class. However, when moving after a claims-free year from class 3 to class 4, the premium increases in the optimal bonus scale. See Gilde and Sundt (1989), Fig. 1. As Gilde and Sundt (1989) notice this kind of bonus scale can not be accepted by the policyholders. In order to avoid the described phenomenon, they restricted the admissible solutions to the case where the premiums increase linearly with respect to the number of the classes.

In a BMS planned in a "natural way", the increase of the premiums after a claims-free year can be eliminated by restricting the admissible solutions to those which are non-decreasing with respect to the number of the classes, i.e., requiring that the premiums $y_1, \dots, y_K$ satisfy

$$(1.3) \qquad y_1 \le y_2 \le \dots \le y_K.$$

(In this paper we use the order of classes opposite to that of Gilde and Sundt, 1989.) Of course, condition (1.3) is a milder restriction than that of Gilde and Sundt (1989). Baione et al. (2002) and Heras et al. (2004) consider optimality under condition (1.3). Heras et al. (2004) present how linear programming can be used to solve the optimality for a criterion based on absolute deviation.

In this paper we consider optimal premiums in the case where the transitions depend both on the number and on the severity of the claims, and where the set of bonus classes, called *bonus coefficients* here, is an interval $[a, b]$, $0 < a < 1 < b$, say $a = 0.3$, $b = 1.5$, where 1 presents the premium before bonus or malus. See Bonsdorff (2005) for a general background for this paper.

In the following we sketch some ideas how a BMS of this type can be constructed. The endpoints a and b of the interval could be chosen so that they are suitable for a competitive market. When constructing the transition rules, the designer has an initial thought how a claims/a non-claims year ought to influence the premiums. On this basis the designer defines the transition rules on $[a, b]$. Of course, the transition rules ought to be constructed in a logical way. The minimum requirement is that

(1.4) after a claims-free year the transition is to the left.

These transition rules define the BMS with initial premiums. Alternative premiums $f(x)$, $x \in [a, b]$, can be achieved via a function $f : [a, b] \to [a, b]$. If f is non-decreasing, it is impossible that after a claims-free year the premium increases when condition (1.4) is satisfied.

Given a BMS with transition rules, the optimality of the premiums $f(x)$, $x \in [a, b]$, is studied throughout this paper with the restriction that f is a non-decreasing function on $[a, b]$.

Note that as in the classic case, also in this case, the optimality is studied with respect to given transition rules (which in this case can be interpreted to induce initial premiums as mentioned above).

The claims process of a single policy is assumed to be compound Poisson. Under general conditions the transitions between bonus coefficients establish a general (homogeneous) Markov chain with state space $[a, b]$, see Bonsdorff (2005). For a general reference to Markov chains, the reader is referred to Nummelin (1984). The compound Poisson process of a single policy is determined by the parameter $u = (v, \lambda)$, characteristic of the policy where v is the parameter of the distribution of the severity of an individual claim and λ the parameter of the Poisson process. The criterion (1.2) has in this case the form

$$\min_{f \text{ non-decreasing}} \int \left(\int f d\pi_u - E_u \right)^2 dU(u), \tag{1.5}$$

where π_u is the invariant probability measure associated to the parameter u, E_u the corresponding expected value of the yearly claims amount and U the distribution of the parameter u in the insurer's portfolio. Correspondingly, the criterion (1.1) has the form

$$\min_{f \text{ non–decreasing}} \int \left(\int (f - E_u)^2 d\pi_u \right) dU(u). \tag{1.6}$$

We will show the existence of an optimal solution of (1.5) and that of (1.6) under general conditions. We also prove the existence with respect to some other optimality criteria. As a corollary we get corresponding results also for the classical case. In each case we call the optimal solution *optimal premiums scale* with respect to the optimality criterion in question.

In practice there might be a need to require also other properties concerning the premiums than monotonicity. This means that the admissible optimal premium scale is determined by a criterion with additional requirements besides monotonicity. We will show how the existence of an optimal solution with respect to additional requirements can be proven in certain cases.

We consider the uniqueness of the optimal premium scale under criterion (1.6) by means of convex optimization. It turns out that under general conditions an essential uniqueness can be reached.

We will also develop an algorithm for the evaluation of the optimal premium scale under criterion (1.6). The treatment is based on direct methods in the calculus of variations and convex optimization.

In Section 2 we present the basic assumptions which will be valid for the rest of the paper.

In Section 3 we show that an optimal solution exists under several optimality criteria. The proofs are based on the Tihonov theorem, which we recall in Section 3.

In Section 4 we consider, under criterion (1.6), uniqueness and in Section 5 determining of the optimal premium scale.

The main results are formulated under regularity conditions concerning the invariant probability measures π_u. In Section 6 we show that these conditions are valid when certain natural conditions are fulfilled.

In the Appendix we give a reduced example in the classic case, where the optimal solution of (1.1) is not monotone, and show how the optimal solution under the monotonicity requirement can be found.

2. Basic assumptions

In this Section we present the basic assumptions which are supposed to be valid throughout this paper and will not be repeated later.

We denote the σ-algebra of Borel sets in R^n by $\mathcal{B}$ and the restriction of the Lebesgue measure to $\mathcal{B}$ by l. The *measurability* of functions on R^n means measurability with respect to $\mathcal{B}$. Correspondingly, we say that a probability measure μ on $[a, b]$ is *absolutely continuous*, if it is absolutely continuous with respect to l, i.e. if $l = 0$, then $\mu = 0$. For a measurable function f and a probability measure v we often write vf instead of $\int f dv$.

Basic assumptions for the Bonus-Malus system and for the claims process of an individual policy: *Let the number of claims be a Poisson process with intensity* λ *and let the claims be i.i.d. and independent of the claims number process, i.e. the claims process is compound Poisson. The set of the bonus coefficients is* $[a, b],\ 0 < a < b$.

The transitions from a bonus coefficient to another are determined as follows:

$$X_o \text{ is the initial coefficient,}$$
$$(2.1) \qquad X_n = g_k(X_{n-1}, Y_{n-1}), \ n = 1, 2, \ldots$$

where X_n is the bonus coefficient in year n, Y_{n-1} the total amount of claims in year $n-1$, k the number of claims in year $n-1$, $k = 0, 1, 2, \ldots$ and g_k are measurable functions from $[a,b] \times R_+$ (g_0 from $[a,b] \times \{0\}$) to $[a,b]$, where R_+ denotes the positive real axis. According to (1.4) it will be assumed that $g_0(X_{n-1}, 0) < X_{n-1}$, when $a < X_{n-1}$ and $g_0(a, 0) = a$.

Further we assume that there exists a number n_0 such that starting from any $x \in [a,b]$, after $n \geq n_0$ claims-free years the policy is in the coefficient a. We call a the superbonus coefficient and the last-mentioned assumption the superbonus assumption.

It follows from the assumptions above that the sequence (X_n) is a Markov chain with transition probability kernel

$$(2.2) \qquad P(x, A) = \sum_{k=0}^{\infty} p_k \int_{B_k} \mu_k(dy),$$

where $P(x, A)$ is the probability that the chain transfers from the bonus coefficient x to the Borel set A in one step, $B_k = \{y \mid g_k(x, y) \in A\}$, p_k is the probability of k claims and μ_k are the conditional probability distributions of the total amount of the claims in one year given k claims. It follows from Theorem 3.1 of Bonsdorff (2005) that the Markov chain (X_n) has a unique invariant probability measure π and that the n-step transition probabilities $P^n(x, A)$ converge to $\pi(A)$ for all $x \in [a,b]$ and all Borel sets $A \subset [a,b]$. The n-step transition probabilities are defined as follows

$$P^n(x, A) = \int P(x, dz) P^{n-1}(z, A),$$

$x \in [a,b]$, $A \subset [a,b]$, $A \in \mathcal{B}$. Note the notation typical of Markov chains where in integrals the measure lies before the integrand.

Basic assumptions for the insurance portfolio: *We assume that the size of an individual claim of a single policy is distributed due to the distribution F_v, where v is the policy's characteristic parameter in a distribution family $\{F_v\}$. Let the parameter v obey the distribution V and let the Poisson parameter be distributed according to the distribution W. The parameter $u = (v, \lambda)$ characterizes a single policy. Let the joint distribution of V and W, i.e. the distribution of u, be U. With each u are associated the corresponding transition probability $P_u(x, A)$ and its limit (invariant) probability measure π_u. We assume that the functions $u \mapsto P_u^n(x, A)$ are measurable for all $x \in [a,b]$ and for all Borel sets $A \subset [a,b]$.*

Denote by E_u the expectation of the total amount of claims of a policy with the parameter u. We assume that the functions $u \mapsto E_u$ are measurable and that

$$(2.3) \qquad \int E_u^2 du = M < \infty.$$

We call the set of all parameters u the parameter space.
Finally, we assume that

$$\text{(2.4)}\quad \begin{array}{l} \textit{with exception of a finite set } H = \{b = x_0, \ldots, x_m = a,\ x_i \in [a,b]\}, \\ \textit{the probability measures } \pi_u \textit{ are absolutely continuous and that} \\ \pi_u(x_i) > 0 \text{ for all } u \in U \text{ and } x_i \in H. \end{array}$$

Let us comment the last assumption. It follows from the superbonus assumption that $\pi_u(a) > 0$ for all u, see Bonsdorff (2005), p. 315. If the BMS has the property that for all u,

(2.5) after exceeding a certain total amount of claims a certain number n_o of consecutive years, the policy starting from any $x \in [a, b]$ is in b,

also $\pi_u(b) > 0$ for all u. This can be seen similarly as $\pi_u(a) > 0$ in Bonsdorff (2005). Further, it follows from the superbonus assumption that there exists a finite number of points

$$x_0 = b > x_1 > x_2 > \ldots > x_{m_0} = a \tag{2.6}$$

such that after claims-free years the policy, starting from b, transfers from x_j to x_{j+1}, $j = 0, \ldots, m_0 - 1$ and from x_{m_0} to x_{m_0}. The set $D = \{x_0, \ldots, x_{m_0}\}$ will be called the *skeleton* of the BMS. Again, we can verify that if the BMS has the property (2.5), $\pi_u > 0$ at all the skeleton points. Therefore, in a BMS designed in a natural way, π_u is positive at the skeleton points for all u.

However, later in Section 6 we will show that the probability measures π_u are absolutely continuous outside a finite set under certain natural conditions.

3. Existence of an optimal monotone premium scale

In this Section we prove the existence of an optimal monotone premium scale under certain optimality criteria. Technically we do it so that we first prove the result for criterion (1.5) and then attain the other results as corollaries, especially for criterion (1.6), on which we will focus later in this paper. As a corollary, we get the results in the classic case. We also show how the existence of an optimal solution can be proven in certain cases with respect to additional requirements besides monotonicity.

The proofs are based on the Tihonov theorem, which we recall in the following. The reader is referred e.g. to Kelley (1955) for a thorough treatment of general topology. Let $(T_\alpha, \mathcal{A}_\alpha)$ be topological spaces, where T_α is a set and $\mathcal{A}_\alpha$ its topology, $\alpha \in J$, J being an arbitrary index set.

The *product topology* $\mathcal{D}$ is defined for the cartesian product $\prod_{\alpha \in J} T_\alpha$ as follows: $\mathcal{D}$ is the topology which has the base

$$\mathcal{D}' = \{ \prod_{\alpha \in J} V_\alpha \text{ where } V_\alpha \in \mathcal{A}_\alpha \text{ for all } \alpha \text{ and } V_\alpha = T_\alpha \text{ for all } \alpha$$
$$\text{with at most a finite number of exceptions} \}.$$

Hence $\mathcal{D}$ consists of the unions of the sets $E \subset \mathcal{D}'$.

The Tihonov theorem states that if every $(T_\alpha, \mathcal{A}_\alpha)$ is compact, then $(\prod_{\alpha \in J} T_\alpha, \mathcal{D})$ is compact.

If $T_\alpha = T$ for all α, $\prod_{\alpha \in J} T_\alpha = T^J =$ the set of all functions $J \to T$. In the sequel we consider the case where for all α, $(T_\alpha, \mathcal{A}_\alpha) = [a, b]$ equipped with the ordinary topology (relative to $[a, b]$) and $J = [a, b]$. Hence, by the Tihonov theorem the set of all functions $[a, b] \to [a, b]$ is compact with respect to the product topology.

Let G be the set of all functions $[a, b] \to [a, b]$ and $F = \{f \in G : f \text{ non-decreasing}\}$. As mentioned, we first consider criterion (1.5). Let A be the functional defined on F as follows (cf. 1.5)

$$Af = \int \Big(\int f d\pi_u - E_u \Big)^2 du = \int (\pi_u f - E_u)^2 du, \tag{3.1}$$

where du means integration with respect to U.

Throughout this paper by a minimum we mean absolute minimum, if not otherwise stated. Accordingly, if Q is a class of functions and α is a functional defined on Q and $f \in Q$ is such that $\alpha(f) \le \alpha(g)$ for all $g \in Q$, we say that α gets its minimum in Q at f.

We will now prove the following fundamental existence theorem for an optimal solution by showing that the class F of non-decreasing functions is compact and the functional A is continuous with respect to $\mathcal{D}$. Note that compactness of a topological space and continuity of a function on it are somewhat opposing properties. Roughly speaking, the less open sets in the topology, the easier/more possible to show compactness and the more difficult/impossible to show continuity. The product topology with "small" amount of open sets has been chosen in order to reach compactness. Also continuity can be proven but the proof is not straightforward with respect to the product topology. Continuity will play an important role also later in this paper. Therefore, it is separately mentioned in the following Theorem, as well as in Corollary 3.1.

Theorem 3.1: *The functional $A : F \to R$ is continuous with respect to $\mathcal{D}$ and gets its minimum in F, where F is the class of non-decreasing functions $[a, b] \to [a, b]$.*

Proof. We prove the assertion in the following steps:

1° F is compact

2° A is continuous in F.

Thus the assertion follows, since a continuous real-valued function on a compact set attains its minimum.

1° We show that F is compact. By the Tihonov theorem G is compact with respect to the product topology. We will show that F is closed, and hence compact as a closed subset of a compact set. It suffices to show that every point of F^c is an interior point. Let $f \in F^c$. Then there exist $x_1, x_2 \in [a, b]$ such that $x_1 < x_2,\ f(x_1) > f(x_2)$. Denote $\varepsilon = f(x_1) - f(x_2)$. Then the neighbourhood of f

$$U = \{\xi \in G \mid \xi(x_1) > f(x_1) - \varepsilon/2,\ \xi(x_2) < f(x_2) + \varepsilon/2\} \subset F^c.$$

2° We now turn to the functional A. Let $f \in F$. To begin with, we show that the integral

$$\int (\pi_u f - E_u)^2 du \tag{3.2}$$

exists.

First we note that the functions $u \mapsto (\pi_u f - E_u)^2$ are measurable. The measurability of the functions $u \mapsto E_u$ follows from the basic assumptions. Further, it follows from the assumptions that the functions $u \mapsto \pi_u(A)$ are measurable for all $A \in \mathcal{B},\ A \subset [a, b]$. It is then easy to verify that the functions $u \mapsto \pi_u f$ are measurable.

Secondly, we notice that the integral (3.2) is finite. We get

$$\int (\pi_u f - E_u)^2 du = \int (\pi_u f)^2 du - 2 \int E_u \pi_u f du + \int E_u^2 du. \tag{3.3}$$

The first term of the right-hand side is $\leq b^2$, the second ≤ 0, and the third one is finite by assumption.

Now we turn to the continuity of the functional A from F to R_+, where F is equipped with the product topology. By (3.3) it suffices to show that the functionals

$$A_1 f = \int (\pi_u f)^2 du \tag{3.4}$$

and

$$A_2 f = \int E_u \pi_u f du \tag{3.5}$$

are continuous. We first consider case A_1.

Let $f \in F,\ \varepsilon > 0$. We have to show that there exists a neighbourhood $U \subset F$ of f such that $|A_1 f - A_1 \xi| < \varepsilon$ for all $\xi \in U$. To begin with, we construct two sequences of step functions which approximate the function f.

For each integer n, divide the interval $[a, b]$ into 2^n subintervals by the points

$$x_k = a + k\frac{b-a}{2^n},\ 0 \le k \le 2^n. \tag{3.6}$$

For each n we define the functions h_n, k_n as follows: if $x \in [x_k, x_{k+1}), h_n(x) = f(x_k),\ k_n(x) = f(x_{k+1}),\ k = 0, \ldots, 2^n - 1, h_n(b) = f(x_{2^n-1}), k_n(b) = f(x_{2^n}) = f(b)$. It follows from the monotonicity of f that

$$\begin{gathered} h_n \le f \le k_n, \\ h_n \le h_{n+1},\ k_{n+1} \le k_n \end{gathered} \tag{3.7}$$

and that f has at most denumerable discontinuity points. It is easy to see that $h_n \to f,\ k_n \to f$ at all continuity points of f. Hence the convergence is almost everywhere (a.e.) with respect to l.

In the following, to begin with, we restrict ourselves to the set $H^c = [a, b] - H$ where H is the finite exception set (cf. the basic assumptions). We denote the restrictions of the probability measures π_u to H^c by π_u^*. Since by assumption π_u is absolutely continuous for each u in H^c, $h_n \uparrow f,\ k_n \downarrow f$ π_u^*-a.e. for each u. Thus

$$\pi_u^* h_n \uparrow \pi_u^* f,\ \pi_u^* k_n \downarrow \pi_u^* f \tag{3.8}$$

by the Lebesgue monotone convergence theorem, for all u.

We will show that

$$\int \pi_u^* h_n du \uparrow \int \pi_u^* f du,\ \int \pi_u^* k_n du \downarrow \int \pi_u^* f du.$$

Let $\delta > 0$ be arbitrary. By the Egoroff theorem there exists a set E_δ in the parameter space (cf. the basic assumptions) such that $U(E_\delta) < \delta$ and that the convergences (3.8) of the sequences of measurable functions of u, $(\pi_u^* h_n)$ and $(\pi_u^* k_n)$ are uniform in E_δ^c. Hence, there exists n_1^δ such that when $n \ge n_1^\delta$, then for all $u \in E_\delta^c$

$$|\pi_u^* h_n - \pi_u^* f| < \delta,$$

whence

$$\int |\pi_u^* h_n - \pi_u^* f| du \le \delta + (b-a)\delta = c\delta,$$

where $c = (1 + b - a)$. Correspondingly, there exists n_2^δ such that when $n \ge n_2^\delta$

$$\int |\pi_u^* k_n - \pi_u^* f| du \le c\delta.$$

Thus, when $n \ge n^\delta = \max(n_1^\delta, n_2^\delta)$

$$\int \pi_u^* k_n du - \int \pi_u^* h_n du \le 2c\delta. \tag{3.9}$$

We define the neighbourhood of f, U_{n^δ} as follows: $U_{n^\delta} = U'_{n^\delta} \cap V$ where

$$(3.10) \qquad \begin{aligned} U'_{n^\delta} = \{\xi \in F \mid\ & h_{n^\delta} - \delta < \xi < k_{n^\delta} + \delta \text{ at points } x_0, \dots, x_{2^{n^\delta}}, \\ & \text{where the condition } h_{n^\delta} - \delta < \xi \text{ is substituted} \\ & \text{by condition } a \le \xi \text{ at points where } h_{n^\delta} - \delta < a \\ & \text{and condition } \xi < k_{n^\delta} + \delta \text{ is substituted by} \\ & \text{condition } \xi \le b \text{ at points where } k_{n^\delta} + \delta > b\} \end{aligned}$$

and

$$(3.11) \qquad \begin{aligned} V =\{\xi \in F \mid\ & f - \delta < \xi < f + \delta \text{ at points } z \in H, \text{ where} \\ & \text{the conditions with respect to } a \text{ and } b \\ & \text{are adjusted similarly as in } U'_{n^\delta}\}. \end{aligned}$$

Clearly, U_{n^δ} is a neighbourhood of f. By definition, the inequalities

$$(3.12) \qquad h_{n^\delta} - \delta < \xi < k_{n^\delta} + \delta$$

hold true for all $\xi \in U_{n^\delta}$ at all points $x_0, \dots, x_{2^{n_\delta}}$. Next we will show that it follows from the monotonicity of the functions ξ that the inequalities (3.12) hold true for all points $x \in [a, b]$. Let $x \in [a, b]$, then $x \in [x_k, x_{k+1}]$ for some k. We have $\xi(x_k) > h_{n^\delta}(x_k) - \delta$. If $\xi(x) \le h_{n^\delta}(x_k) - \delta$, ξ would not be monotone. Thus $\xi(x) > h_{n^\delta}(x_k) - \delta = h_{n^\delta}(x) - \delta$. Similarly, $\xi(x) < k_{n^\delta}(x) + \delta$.

Let now $\xi \in U_{n^\delta}$. By (3.7) and (3.12) we have

$$(3.13) \qquad |f - \xi| \le k_{n^\delta} - h_{n^\delta} + 2\delta.$$

Hence, for all u

$$\pi^*_u |f - \xi| \le \pi^*_u k_{n^\delta} - \pi^*_u h_{n^\delta} + 2\delta,$$

whence by (3.9)

$$(3.14) \qquad \int \pi^* |f - \xi| du \le \int \pi^*_u k_{n^\delta} du - \int \pi^*_u h_{n^\delta} du + 2\delta \le 2(c+1)\delta.$$

Thus, by (3.14) and (3.11)

$$(3.15) \qquad \begin{aligned} & \Big| \int \pi_u f du - \int \pi_u \xi du \Big| \le \int |\pi_u f - \pi_u \xi| du \le \int \pi_u |f - \xi| du \\ & = \int \pi^*_u |f - \xi| du + \int_H \pi_u |f - \xi| du \le 2(c+1)\delta + \delta = (2c+3)\delta \\ & = m\delta, \text{ where } m \text{ denotes } (2c+3). \end{aligned}$$

By (3.15) we get

$$\Big|\int(\pi_u f)^2du-\int(\pi_u\xi)^2du\Big|\le\int(\pi_u f+\pi_u\xi)|\pi_u f-\pi_u\xi|du$$
$$\le 2bm\delta<\varepsilon,\ \text{when } \delta \text{ is chosen to be } <(2bm)^{-1}\varepsilon.$$

This completes the proof of the continuity of the functional A_1. Consider now the functional $A_2 f=\int E_u\pi_u f du$. Let $\xi\in U_{n^\delta}$. By the Schwarz's inequality

$$\Big|\int E_u\pi_u f du-\int E_u\pi_u\xi du\Big|\le\int E_u|\pi_u f-\pi_u\xi|du$$
$$\le\Big(\int E_u^2du\Big)^{1/2}\Big(\int|\pi_u f-\pi_u\xi|^2du\Big)^{1/2}.$$

The first factor is $M^{1/2}$ by the basic assumptions. Consider the second factor. Let E be the set of the parameter space where $|\pi_u f-\pi_u\xi|\ge 1$. By (3.15), $U(E)\le m\delta$. Hence, by (3.15)

$$\int|\pi_u f-\pi_u\xi|^2du=\int_E|\pi_u f-\pi_u\xi|^2du+\int_{E^c}|\pi_u f-\pi_u\xi|^2du$$
$$\le(b-a)^2m\delta+\int_{E^c}|\pi_u f-\pi_u\xi|du\le[(b-a)^2+1]m\delta.$$

Hence

$|A_2f-A_2\xi|\le(M[(b-a)^2+1]m)^{1/2}\delta^{1/2}<\varepsilon$ when δ is chosen to be $<(M[(b-a)^2+1]m)^{-1}\varepsilon^2$.

This completes the proof of Theorem 3.1.

In the following we present some corollaries of Theorem 3.1 with different optimality criteria. In Corollary 3.1 the criterion is analogous to that of (1.1), whereas in Corollary 3.2 a criterion based on absolute deviation is used.

Corollary 3.1: *The functional $B:F\to R$ (cf. 1.6)*

$$Bf=\int\pi_u(f-E_u)^2du \tag{3.16}$$

is continuous with respect to the product topology $\mathcal{D}$ and gets its minimum in F, where F is the class of non-decreasing functions $[a,b]\to[a,b]$.

Proof. Similarly as in the proof of Theorem 3.1, we see that the integral in (3.16) exists. We have

$$Bf=\int\pi_u f^2du-2\int E_u\pi_u f du+\int E_u^2du.$$

At the end of the proof of Theorem 3.1 it was shown that the functional $A_2 f = \int E_u \pi_u f du$ is continuous. It remains to show the continuity of the functional $f \mapsto \int \pi_u f^2 du$.

Let $\xi \in U_{n_\delta}$. We get

$$\Big| \int \pi_u f^2 du - \int \pi_u \xi^2 du \Big| \leq \int \pi_u (f+\xi)|f-\xi| du \leq 2b \int \pi_u |f-\xi| du.$$

Hence the assertion follows from (3.15).

Corollary 3.2: *The functional $Cf = \int |\pi_u f - E_u| du$ gets its minimum in F. The result holds true with the weaker assumption $\int E_u du < \infty$ than* (2.3).

Proof. Note that $\int |\pi_u f - E_u| du$ is the distance ρ of $\pi_u f$ and E_u in the L^1-metric. By the triangle inequality we get

$$Cf - C\xi = \rho(\pi_u f, E_u) - \rho(\pi_u \xi, E_u) \leq \rho(\pi_u f, \pi_u \xi).$$

Similarly $C\xi - Cf \leq \rho(\pi_u f, \pi_u \xi)$. Hence $|Cf - C\xi| \leq \int |\pi_u f - \pi_u \xi| du$. The assertion follows from (3.15).

For the rest of this paper we will mainly focus on the functional B. In practical situations there may naturally arise a need for additional restrictions on premiums than that of monotonicity. For example, the growth of the premium on the interval $[a, b]$ should be limited, cf. Heras et al. (2004). Recall that in Theorem 3.1 and in the Corollaries above the existence of an optimal solution was proven with respect to the set $F = \{f : [a, b] \to [a, b] : f \text{ non-decreasing}\}$. The additional requirement means that instead of F, one should prove the existence of an optimal solution in a subset F' of F. If F' is closed, it is compact since F is compact. Accordingly, we have the following Corollary.

Corollary 3.3: *Let $F' \subset F$ be closed. Then the functional B gets its minimum in F'.*

We give two examples how Corollary 3.3 can be applied.

1° We set an additional requirement which limits the growth of the premiums on $[a, b]$. Let c_1 and c_2 be constants, $c_1 \geq 0$, $c_2 > 1$. Let

$$\begin{aligned} F' = &F \cap \{f : [a, b] \to [a, b] : f(x_2) - f(x_1) \leq c_1 + c_2(x_2 - x_1) \\ &\text{for all } x_1, x_2 \in [a, b],\ x_1 < x_2\}. \end{aligned}$$

One can easily verify that F' is closed and thus compact.

2° In Theorem 3.1 it was not assumed that

$$f(a) = a,\ f(b) = b. \tag{3.17}$$

However, the assertion of Theorem 3.1 remains true if the additional assumption (3.17) is made, since the corresponding subset F' of F is closed, as one can immediately verify.

In the proof of Theorem 3.1 and its Corollaries the invariance of the probability measures π_u has not been used. The only things used are that π_u and U are probability measures, the functions $u \mapsto \pi_u(A)$ and $u \mapsto E_u$ are measurable, $\int E_u^2 du < \infty$ and that with exception of a finite set $H' \subset [a, b]$ the probability measures π_u are absolutely continuous. Accordingly, we have the following Corollary, where the optimality criterion is closely connected to that of Borgan et al. (1981).

Corollary 3.4: *Denote by ν_u the probability measures*

$$\nu_u = a_o \pi_u + \sum_{i=1}^{n_0} a_i P_u^i(x, \cdot), \tag{3.18}$$

where $n_0 \geq 1,\ a_i \geq 0,\ i = 0, 1, \ldots, n_0,\ \sum_{i=0}^{n_0} a_i = 1$, the starting point $x \in [a, b]$ and P_u^i are the i-step transition probabilities associated with the parameter u.

If the i-step transition probabilities $P_u^i(x, \cdot)$ are absolutely continuous outside a finite exception set H', then Theorem 3.1 and Corollaries 3.1–3.3 hold true when the probability measures π_u are substituted by those of ν_u.

Proof. The measurability of the functions $u \mapsto \nu_u(A)$ follows from the basic assumptions for the insurance portfolio.

In Section 6 we give examples where the last-mentioned assumption of Corollary 3.4 holds true.

The classical BMS is a special case of our treatment, where the single claims are equal to 1 with probability 1 and there are only a finite number of bonus coefficients (classes). We have the following Corollary.

Corollary 3.5: *We make the following standard assumptions. The* BMS *has a finite number s of bonus classes. The* BMS *possesses a superbonus class, say s, i.e., starting from any class, after a sufficient number of consecutive claims-free years, the policy is in class s. The transitions between bonus classes depend on the class of the preceding period and on the number of claims of the preceding period only. For each policy the number of claims is Poisson distributed with an intensity λ, characteristic of the policy. The parameter λ is assumed to obey the probability distribution U in the insurance portfolio, and $\int \lambda^2 dU(\lambda)$ is assumed to be finite.*

Under the assumptions above, the functional

$$B = \int \Big(\sum_{i=1}^{K} (y_i - \lambda)^2 \pi_i(\lambda) \Big) dU(\lambda)$$

attains its minimum in F, where

$$F = \{(y_1, \ldots, y_K) \in R^K,\ a \leq y_1 \leq y_2 \leq \ldots \leq y_K \leq b\}$$

and $\{\pi_1(\lambda), \ldots, \pi_s(\lambda)\}$ is the invariant limit distribution of the Markov chain associated with the BMS.

Corresponding Corollaries associated with Theorem 3.1 and Corollaries 3.2 and 3.4 hold true, as well.

Remark 3.1: As observed above, Corollary 3.5 and those associated with it follow as special cases from earlier results of this paper. However, they can be proven directly essentially more easily. Let us take as an example the proof of Corollary 3.5 in the classic case. We have to show the compactness of F and the continuity of the functional $B : F \to R$

$$B(y_1, \dots, y_K) = \int \Big(\sum_{i=1}^{K} (y_i - \lambda)^2 \pi_i(\lambda) \Big) dU(\lambda).$$

In this case F is trivially compact as a closed and bounded subset of R^K; thus the Tihonov theorem is not needed. The proof of the continuity of B is straightforward and is left to the reader.

4. Uniqueness

In this Section we consider the uniqueness of the minimum of the functional B in F, cf. Corollary 3.1. A complete uniqueness cannot be attained, because the value of the premium function f can be changed at single points such that the value of B does not change. Further, if the BMS is such that there are, roughly speaking, areas in $[a, b]$ which the BMS does not visit, the probability measures π_u vanish on such areas. Accordingly, the premium function can be defined arbitrarily in such areas without any effect on the value of the functional. However, if we assume that the BMS "sufficiently" visits "big" sets of $[a, b]$, we can attain an essential uniqueness. We shall also deal with the classic case.

Our treatment is based on convex optimization. In the following we introduce some concepts and results of convex optimization in vector space context, see, e.g., Luenberger (1969) for a general reference to the topic. Let L be a vector space. A set $C \subset L$ is said to be *convex set* if $x_1, x_2 \in C$ implies that all points of form $\alpha x_1 + (1 - \alpha)x_2,\ 0 \le \alpha \le 1$ are in C. A real-valued functional J defined on a convex set C of L is said to be *convex functional* if

$$J(\alpha x_1 + (1 - \alpha)x_2) \le \alpha J(x_1) + (1 - \alpha)J(x_2) \tag{4.1}$$

for all $x_1, x_2 \in C$ and all $\alpha,\ 0 < \alpha < 1$. If (4.1) holds as a strict inequality whenever $x_1 \ne x_2$, J is called *strictly convex.* If a strictly convex functional J attains its minimum x_0 in a convex set C, i.e., $J(x_0) \le J(x)$ for all $x \in C$, the minimum is unique. The well-known result above is an immediate consequence of Proposition 2, p. 216 of Luenberger (1969), which states that if J is convex, the set $S = \{x : x \in C,\ J(x) \le J(x_0)\}$ is convex. In fact, let J be strictly convex. If $J(x_1) = J(x_0)$ such that $x_1 \in C,\ x_1 \ne x_0$, we have $\alpha x_1 + (1 - \alpha)x_0 \in S$, whence $J(\alpha x_1 + (1 - \alpha)x_0) = J(x_0)$. On the other hand, $\alpha J(x_1) + (1 - \alpha)J(x_0) = J(x_0)$ which contradicts the strict convexity.

Let us define the measure σ on $[a,b]$ as follows: In the restriction to H^c, $\sigma = l$ and $\sigma(x_i) = 1,\ i = 0,\dots,m$ when $x_i \in H$, cf. (2.4). Let L be $L^\infty([a,b],\sigma)$, briefly L^∞, i.e., the equivalence classes of bounded functions $[a,b] \to R$ where functions in the same class differ from each other only in a σ–null set. We denote by $\overline{f}$ the class represented by f and by $\overline{F}$ the set $\overline{F} = \{\overline{f} : \overline{f}$ has a representant $f \in F\}$. It is easy to verify that $\overline{F}$ is convex.

If B attains its minimum in F at the function f, we say that f is a *minimal function of B in F*, cf. Corollary 3.1.

Let $E \in \mathcal{B}$. If $P_u^n(x,E) > 0$ for some n, we denote $L_u(x,E) > 0$. We will show that if for sufficiently many u, $L_u(a,E) > 0$ for all l–positive sets E, we have an essential uniqueness.

Theorem 4.1. *Assume that there exists a U–positive set U' in the parameter space such that for all $u \in U'$*

$$L_u(a,E) > 0 \textit{ for all } l\textit{–positive } E \in \mathcal{B},\ E \subset [a,b]. \tag{4.2}$$

Then

(i) *the minimal functions of B in F have a common set Γ of continuity points. The corresponding set of discontinuity points Γ^c is at most denumerable,*

(ii) *if f and g are two minimal functions of B in F, then $f = g$ in $\Gamma \cup H$, where H is the finite exception set defined in (2.4),*

(iii) *there exists a unique minimal function of B in F which is right continuous in H^c,*

(iv) *if a minimal function of B in F is continuous in H^c, it is the unique minimal function.*

Proof. We first show that the minimal functions f and g differ from each other at most in an l–null set $\Omega \subset H^c$. We define $\overline{B}$ in L^∞ as follows $B\overline{f} = Bf$. One easily verifies that $\overline{B}$ is well defined. It follows from Corollary 3.1 that $\overline{B}$ gets a minimum in $\overline{F}$. We will show that $\overline{B}$ is strictly convex on L^∞. Consequently the minimum of $\overline{B}$ in $\overline{F}$ is unique, since $\overline{F}$ is convex. Hence two minimal functions can differ at most in an l–null set $\Omega \subset H^c$.

Let $f_1, f_2 \in F,\ 0 < \alpha < 1$. By a straightforward calculation we get

$$\begin{aligned} &B(\alpha f_1 + (1-\alpha)f_2) - (\alpha Bf_1 + (1-\alpha)Bf_2) \\ &= \int \pi_u(\alpha f_1 + (1-\alpha)f_2 - E_u)^2 du - \alpha \int \pi_u(f_1 - E_u)^2 du \\ &- (1-\alpha)\int \pi_u(f_2 - E_u)^2 du = \alpha(\alpha-1)\int \pi_u(f_1 - f_2)^2 du \le 0. \end{aligned}$$

As a consequence, we get

$$\overline{B}(\alpha\overline{f}_1 + (1-\alpha)\overline{f}_2) - (\alpha B\overline{f}_1 + (1-\alpha)B\overline{f}_2) = \alpha(\alpha-1)\int \pi_u(f_1 - f_2)^2 du \le 0 \tag{4.3}$$

for all $\overline{f}_1, \overline{f}_2 \in \overline{B}$. Hence $\overline{B}$ (as well as B) is convex. In (4.3) equality holds if and only if

$$\int \pi_u (f_1 - f_2)^2 du = 0. \tag{4.4}$$

We will show that from (4.4) follows that $\overline{f}_1 = \overline{f}_2$, which implies the strict convexity of $\overline{B}$ on L^∞. Let now (4.4) hold true for $f_1, f_2 \in F$. Note first that if $f_1(x_i) \neq f_2(x_i)$ for some $x_i \in H$, it follows from the assumption (2.4) that $\int \pi_u (f_1 - f_2)^2 du > 0$. Hence (4.4) implies that $f_1 = f_2$ in H. Denote $E = \{f_1 \neq f_2\}$. Let us make an antithesis that $l(E) > 0$. Let $u \in U'$. Then $L_u(a, E) > 0$. By the superbonus assumption $L_u(x, a) > 0$ for all $x \in [a, b]$. Hence

$$L_u(x, E) > 0 \text{ for all } x \in [a, b]. \tag{4.5}$$

The chain P_u is uniformly ergodic (see Bonsdorff, 2005). Hence, by Corollary, p. 34 of Orey (1971), it follows from (4.5) that $\pi_u(E) > 0$, whence $\pi_u(f_1 - f_2)^2 > 0$. Since U' is U–positive, $\int \pi_u (f_1 - f_2)^2 du > 0$, which contradicts (4.4). Thus $\overline{B}$ is strictly convex on $\overline{F}$ and has a unique minimum in $\overline{F}$. Hence two minimal functions f and g differ from each other at most in an l–null set $\subset H^c$. Consequently, $f = g$ in H.

We will now prove the assertion by means of the result above. Let f and g be minimal functions. As noted above, $f = g$ in H, especially $f(a) = g(a),\ f(b) = g(b)$. Denote the set of the continuity points of f by Γ. Since f is monotone, Γ^c is at most denumerable. In order to prove (i) and (ii) it must be shown that g is continuous at x if and only if $x \in \Gamma$, and that in Γ $f = g$.

Let $x_0 \in (a, b)$ be a continuity point of f. Since f and g differ from each other at most in an l–null set, it can be chosen sequences (x_n) and (y_n) so that $x_n \uparrow x_0$ and $y_n \downarrow x_0$ and $f(x_n) = g(x_n),\ f(y_n) = g(y_n)$. Since f is continuous at x_0, $\lim f(x_n) = \lim f(y_n) = f(x_0)$. Thus

$$\lim g(x_n) = \lim g(y_n) = f(x_0). \tag{4.6}$$

Since g is monotone, it follows from (4.6) that $g(x_0) = f(x_0)$. Further, from the monotonicity of g follows that there exist $\lim g(x_0-),\ \lim g(x_0+)$. It follows from (4.6) that $\lim g(x_0-) = \lim g(x_0+) = f(x_0) = g(x_0)$. Hence g is continuous at x_0. Since $f(a) = g(a),\ f(b) = g(b)$, similarly as above we see that if f is continuous at a or b, then also g is continuous at a or b, respectively. Conversely, if g is continuous at y_0, then f is continuous at y_0, and $f(y_0) = g(y_0)$. Thus we have shown (i) and (ii).

Let now f be a minimal function of B in F. Since f is monotone, there exists at each point x the right-hand side limit $f(x+)$. Define $g = f$ in $\Gamma \cup H$ and $g(x) = \lim f(x+)$ in $(\Gamma \cup H)^c$. Clearly, $g \in F$. Since the probability measures π_u are absolutely continuous in H^c, g is a minimal function of B in F and, evidently, the unique in H^c right continuous minimal function of B in F.

The last part of the assertion is obvious.

In Section 6 we give examples where assumption (4.2) of Theorem 4.1 holds true, see Theorem 6.1.

We turn to the classic case. If the BMS is such that there are classes which can not be reached from the superbonus class a, the invariant measures π_λ vanish at those classes, and the uniqueness of B cannot be attained with respect to them. An example of such a BMS is given in Bonsdorff (1992), p. 222 where the BMS never returns to its initial class. If the BMS does not posses such transient bonus classes, complete uniqueness can be attained, cf. Theorem 4.2 below. If the BMS has transient bonus classes, uniqueness can be achieved for the recurrent bonus classes by dropping the transient bonus classes and considering the remaining irreducible Markov chain.

Theorem 4.2. *Let in the classical case for some* $\lambda > 0$

$$L_\lambda(a,i) > 0 \text{ for all } i = 1,\dots,K. \tag{4.7}$$

Then B gets its minimum in F at a unique premium scale $(y_1,\dots,y_K)$, where B and F are defined in Corollary 3.5.

Proof. We will show that B is strictly convex in R^K. Let $f_1 = (f_1^1,\dots,f_1^K)$, $f_2 = (f_2^1,\dots,f_2^K) \in F$. In order to show the strict convexity, it is sufficient to show (c.f. 4.4) that from

$$\int \Big(\sum_{i=1}^{K}(f_1^i - f_2^i)^2\pi_i(\lambda)\Big)dU(\lambda) = 0 \tag{4.8}$$

follows that

$$f_1^i = f_2^i \text{ for all } i = 1,\dots,K. \tag{4.9}$$

It is easy to see that if (4.7) holds true for some λ, it holds for all $\lambda > 0$. Using this fact one can see similarly as in the proof of Theorem 4.1 that (4.8) implies (4.9).

5. Numerical calculation of the optimal solution

In this Section we consider how the optimal solution for the functional B in F can be calculated. Minimizing of B is in fact a problem of calculus of variations. We treat the question by means of direct methods in the calculus of variations, see, e.g., Gelfand and Fomin (1963), pp. 192–193. *Throughout this Section it is assumed that assumption (4.2) of Theorem 4.1 holds true.*

We will construct a sequence of step functions ξ_n such that $B\xi_n$ converges to the optimal solution. For each n divide the interval $[a,b]$ into $2^n m$ subintervals as follows: The partition D_0 consists of the exception set $H = \{x_0,\dots,x_m\}$ (see the basic assumptions for the insurance portfolio). The partition D_n consists of D_{n-1}

and of the midpoints of the subintervals of D_{n-1}. We denote the points of D_n by $a = x_0, x_1, \ldots, x_{2^n m} = b$.

For each n, we define a subset F_n of F as follows: F_n is the set of non–decreasing step functions $[a, b] \to [a, b]$ which are constant on the intervals $[x_j, x_{j+1})$, $i = 0, \ldots, 2^n m - 1$, i.e.,

$$F_n = \{f : [a, b] \to [a, b] \text{ such that } f(x) = y_j \text{ when } \\ x \in [x_j, x_{j+1}),\ j = 0, \ldots, 2^n m - 1,\ f(b) = y_{2^n m} \text{ where } y_j \le y_{j+1}, \\ j = 0, \ldots, 2^n m - 1\}. \tag{5.1}$$

It is easy to verify that F_n is compact for each n. Since B is continuous by Corollary 3.1, there exists for each n a function $\xi_n \in F_n$ such that

$$B\xi_n \le B\xi \text{ for all } \xi \in F_n. \tag{5.2}$$

One easily verifies that F_n is convex. Next we show that B is strictly convex in F_n. Let $f_1, f_2 \in F_n$. It follows from the proof of Theorem 4.1 that (4.4) implies that $f_1 = f_2$ in H and that f_1 and f_2 can differ from each other at most in an l–null set in H^c. Evidently, this implies $f_1 = f_2$. Consequently, B is strictly convex in F_n and the minimal function of B in F_n is unique. Hence the sequence (ξ_n) is uniquely determined.

By the Helly compactness theorem, see Ewing (1985), p. 183, F is sequentially compact in terms of pointwise convergence, i.e., each sequence of F has at least one subsequence which converges pointwise to a function of the class F. Hence the sequence (ξ_n) has a subsequence (ξ_{n_k}) which converges pointwise to limit function $\xi_0 \in F$. Pointwise convergence is equivalent to convergence with respect to the product topology $\mathcal{D}$. Since B is continuous with respect to $\mathcal{D}$, we have

$$\lim B\xi_{n_k} = B(\lim \xi_{n_k}) = B\xi_0.$$

The sequence $(B\xi_n)$ is monotonically non-increasing because $D_{n+1} \subset D_n$, whence also $(B\xi_n)$ converges, and consequently $\lim B\xi_n = B\xi_o$. Thus we have

$$B\xi_n \downarrow B\xi_0 \tag{5.3}$$

and

$$B\xi_0 = \lim B\xi_n \le B\xi_n \text{ for all } n. \tag{5.4}$$

Theorem 5.1. *Let assumption (4.2) of Theorem 4.1 be valid. Then for all* $\xi \in F$, $B\xi_0 \le B\xi$, *where* $\xi_0 \in F$ *is the above-defined limit function of a convergent subsequence* (ξ_{n_k}) *of the uniquely determined sequence* (ξ_n) *defined in (5.2). The sequence* $(B\xi_n)$ *converges monotonically non–increasing to* $B\xi_0$.

Proof. By Corollary 3.1 there exists $g \in F$ such that for all $f \in F$, $Bg \le Bf$. Hence $Bg \le B\xi_0$. We will show that $B\xi_0 \le Bg$. For each n, we define the function

$\overline{h}_n$ on $[a,b]$ as follows: On each interval $[x_j, x_{j+1})$ of the above-defined partition D_n, $j = 0, \ldots, 2^n m - 1$, $\overline{h}_n(x) = g(x_j)$, $\overline{h}_n(b) = g(b)$. Clearly, $\overline{h}_n \in F_n$. Since H is included to every partition D_n, $\overline{h}_n = g$ on H. Similarly as in the proof of Theorem 3.1, we see that $\overline{h}_n \to g$, with exception at most in a denumerable set $E \subset [a,b]$. Clearly, $E \subset H^c$. Since π_u is absolutely continuous in H^c for all u, $\pi_u(E) = 0$ for all u. Hence $\overline{h}_n \to g$ in E^c, where $\pi_u(E) = 0$ for all $u \in U$.

Let $\overline{h}_n' = \overline{h}_n$ on E^c, $\overline{h}_n' = 0$ on E and let $g' = g$ on E^c, $g' = 0$ on E. Since E is a π_u–null set for all u, $B\overline{h}_n' = Bh_n$ and $Bg' = Bg$. The sequence $(\overline{h}_n')$ converges pointwise to g'. Since B is continuous, $\lim B\overline{h}_n' = Bg'$, and consequently $\lim B\overline{h}_n = Bg$.

Let $\varepsilon > 0$. Then there exists n_0 such that $B\overline{h}_n < Bg + \varepsilon$, when $n \geq n_0$. By (5.2) and (5.4), $B\xi_0 \leq B\overline{h}_n$ since $\overline{h}_n \in F_n$. Thus $B\xi_0 \leq Bg + \varepsilon$, where $\varepsilon > 0$ is arbitrary. Hence $B\xi_0 = Bg$. The last part of the assertion follows from (5.3).

We have reduced the problem to minimize B in the set F_n instead of F. The functions of the class F_n can be identified with vectors $(y_0, \ldots, y_n)$ satisfying the monotonicity condition in (5.1). Here we have written n instead of $2^n m$. Accordingly, we have the following problem. Minimize

$$B(y_0, \ldots, y_n) = \int \Big(\sum_{i=0}^{n} (y_i - E_u)^2 \pi_u(I_i) \Big) dU(u), \tag{5.5}$$

where $I_i = [x_i, x_{i+1})$, $i = 0, \ldots, n-1$, $I_n = b$ (cf. 5.1), with constraints

$$a \leq y_0 \leq y_1 \leq \cdots \leq y_n \leq b. \tag{5.6}$$

Note the similarity of (5.5) to the problem of determining the optimal BMS in the classic case under constraints (5.6). Observe, however, that in our case U is in general at least two-dimensional.

In what follows we will consider the case where U is a discrete distribution concentrated on a finite number k of points. In this case the problem (5.5) can be written in the form: Minimize

$$B(y_1, \ldots, y_n) = \sum_{j=1}^{k} \Big(\sum_{i=1}^{n} (y_i - E_{u_j})^2 \pi_{u_j}(I_i) \Big) p_j \tag{5.7}$$

under constraints (5.6), where $p_j > 0$, $j = 1, \ldots, k$, $\sum_{i=1}^{k} p_j = 1$ denote the probabilities of the distribution U at points u_j, $j = 1, \ldots, k$. For each i and j, $\pi_{u_j}(I_i)$ can be evaluated by a simulation method, see Section 4 of Bonsdorff (2005), cf. also Nummelin (2002) and Robert and Casella (2000). As stated earlier in this Section, the optimal solution of (5.7) $(y_0', \ldots, y_n')$ is unique. Using the differentiability and the strict convexity of B in (5.7), the unique optimal solution $(y_1', \ldots, y_n')$ under constraints (5.6) can easily be found. We illustrate this by an example in the classic case in the Appendix (see also the examination of the classic case below). The consideration in the general case (5.7) is similar. The optimal solution of (5.7) can also be found by means of existing software packages for convex optimization, see e.g. Ben–Tal and Nemirovski (2001).

Let us assume that U is a discrete distribution concentrated on a finite number of points. In view of the considerations of this Section we can now present an algorithm for the evaluation of the optimal solution of the minimizing problem for the functional B in F.

Algorithm:

1^o Form a sequence of partitions D_n as described above. The partitions define the function classes F_n, see (5.1).

2^o Let n be fixed. Solve the unique $\xi_n \in F_n$ such that $B\xi_n \le B\xi$ for all $\xi \in F_n$ in the following steps

a) Evaluate the values of $\pi_{u_j}(I_i)$, see (5.7), by the simulation method described in Section 4 of Bonsdorff (2005).

b) Then solve the convex optimization problem (5.7) as illustrated in the Appendix or with help of existing software packages. Denote $(y_0', \dots, y_n')$ the unique optimal solution.

3^o Calculate $B\xi_n = B(y_0', \dots, y_n')$ for growing values of n. The sequence $(B\xi_n)$ converges monotonically non-increasing to $B\xi_0$ where ξ_0 is an essentially unique optimal solution of B in F, cf. Theorems 4.1 and 5.1. Terminate the process for some n_0 and use ξ_{n_0} for an estimate for the optimal solution.

Remark 5.1. From computational point of view, the rapid growth of the number of subintervals of the partitions D_n might be problematic. Alternatively, the partitions can be defined so that the number of subintervals does not grow fast, e.g., as follows: Let $\hat{D}_0 = H$, $\hat{D}_{n+1} = \hat{D}_n +$ the midpoint of the longest subinterval of $\hat{D}_n$. If the longest interval is not unique, choose that one with the smallest left endpoint x_j. The sequence $(\hat{D}_n)$ fulfils the essential features of (D_n) : 1^o $H \subset \hat{D}_n$ for all n, 2^o $\hat{D}_n \subset \hat{D}_{n+1}$, 3^o the length of the longest subinterval converges to zero. Similarly as above, the sequence $(\hat{D}_n)$ of partitions induces function class $\hat{F}_n$ corresponding to F_n, and further, a uniquely determined sequence $(\hat{\xi}_n)$ of minimal functions of B in $\hat{F}_n$ having a subsequence $(\hat{\xi}_{n_k})$ which converges pointwise to $\hat{\xi}_0$, where $\hat{\xi}_0$ is a minimal function of B in F. Correspondingly as above, $B\hat{\xi}_n \downarrow B\hat{\xi}_0 = B\xi_0$.

Let us now consider the classic case. We assume that, besides the assumptions of Corollary 3.6, assumption (4.7) of Theorem 4.2 holds true. The minimization problem can be written as follows: Minimize

$$B(y_1, \dots, y_K) = \sum_{j=1}^{k} \Big(\sum_{i=1}^{K} (y_i - \lambda_j)^2 \pi_j(i) \Big) p_j, \tag{5.8}$$

with constraints

$$a \le y_0 \le y_1 \le \dots \le y_K \le b, \tag{5.9}$$

where K is the number of bonus classes, $\pi_j = (\pi_j(1), \ldots, \pi_j(K))$ is the invariant probability distribution of the transition probability matrix P_{λ_j} associated with the parameter $\lambda_j,\ j = 1, \ldots, k$ and p_j's are as in (5.7).

For each j the distribution π_j can be calculated, e.g., by norming the left eigenvector associated with the eigenvalue 1 of the transition probability matrix P_{λ_j}. Then the unique optimal solution of (5.8) can be calculated as illustrated in the Appendix or by means of existing software packages for convex optimization problems.

See Heras et al. (2004) for solving the optimality by means of a method based on linear programming for the functional C in the classical case.

6. Absolute continuity of π_u and ν_u outside a finite set

In this section we will show that the conditions (2.4), (4.2) and the assumptions of Corollary (3.4) hold true under certain natural conditions, when the BMS has smooth and monotone transition rules. Following the usual practice, we say that a random variable is *continuously distributed* if its distribution function has a continuous derivative.

The following result is mentioned to serve as an example, and it does not aim to be exhaustive.

In this Section we write m instead of m_0 for the index appearing in the definition of the skeleton, see (2.6).

We make the following assumptions concerning the distribution of the individual claims (6.1), transfers after a year when claims have happened (6.2) and transfers after a claims-free year (6.3).

(6.1) *For all policies (i.e., for all $u \in U$) the size of an individual claim is continuously distributed with a positive density function on R_+.*

(6.2) *After a number $k \geq 1$ of claims, the policy transfers from $x \in [a, b)$ to the right, determined by the functions g_k in (2.1). For each x and k there exists $y_{k,x} > 0$ such that if $y \geq y_{k,x}$, then $g_k(x, y) = x_j(k, x)$ where $x_j(k, x)$ is some of the skeleton points $> x$ depending on k and x. For all $x \in [a, b)$ and $k \geq 1$ $\lim g_k(x, 0+) = x$ and the functions $g(y) = g_k(x, y)$ have a continuous derivative $g' > 0$ on the interval $[0, y_{k,x})$ and are continuous at $y_{k,x}$, where $g(0)$ is defined to be equal to $\lim g_k(x, 0+) = x$. Further, we assume that $g_k(b, y) = b$ for all $y > 0$, for all $k \geq 1$.*

(6.3) *We assume that either of the following conditions, 1° or 2°, is valid.*

1° *After a claims-free year, the policy transfers from $x \in (x_{j+1}, x_j]$ to x_{j+1}, and if $x = a$, it transfers to a, where the x_j's are the skeleton points, cf. (2.6).*

2° *Let ψ be a continuously differentiable function from $[x_{m-1}, b]$ onto $[a, x_1]$ with $\psi' < 0$ on $[x_{m-1}, b]$, cf. (2.6). After a claims-free year the policy transfers from $x \in [x_{m-1}, b]$ to $\psi(x)$ and from $x \in [a, x_{m-1}]$ to a.*

Theorem 6.1: *Under assumptions (6.1), (6.2) and (6.3), π_u is absolutely continuous outside the finite set D, the skeleton of the BMS, for all $u \in U$. For* *all* $u \in U$ *and* $x_i \in D$, $\pi_u(x_i) > 0$. *Condition (4.2) of Theorem 4.1 is valid. In addition,* ν_u *is absolutely continuous outside a finite set.*

Proof. First we observe that $\pi_u(x_i) > 0$ for all $u \in U,\ x_i \in D$. In fact, it follows from assumption (6.2) that (2.5) holds for all u, whence $\pi_u(x_i) > 0$ for all $x_i \in D$, cf. the discussion after the basic assumptions for the insurance portfolio.

In the following, we will show that π_u is absolutely continuous outside D in both cases, 1° and 2°.

Let $u \in U$ be arbitrary. To begin with, we deal with the event that at least one claim has occurred in the year in question. Note first that the total amount of claims in one year Y is continuously distributed. In fact, Y can be expressed as follows:

$$Y = Z_1 + \sum_{k=2}^{n(\lambda)} Z_k,$$

where the Z_i's are individual claims and $n(\lambda)$ is the random number of the claims. (We have excluded the trivial case $n(\lambda) = 1$.) Thus Y is continuously distributed as a convolution of two variables, from which at least one is continuously distributed.

Next we will show that, for fixed x and $k \geq 1$, the random variable $g(Y)$, cf. (2.1) and (6.2), is continuously distributed on $[x, b]$ with the exception of the point $g(y_{k,x}) \in D$. Let $F_{g(Y)}$ be the distribution function of $g(Y)$. Let $x \leq t \leq g(y_{k,x})$. Then we have $F_{g(Y)}(t) = F_Y(g^{-1}(t))$, where F_Y is the distribution function of Y and g^{-1} is the inverse function of g. Thus by (6.1) and (6.2),

$$\text{(6.4)} \qquad \begin{aligned} \frac{dF_{g(Y)}(t)}{dt} &= F_Y'(g^{-1}(t))\frac{1}{g'(t)} > 0, \ \text{when } x \leq t < g(y_{k,x}) \text{ and} \\ &= 0, \ \text{when } t > g(y_{k,x}). \end{aligned}$$

Hence $F_{g(Y)}$ is continuously differentiable on the interval $[x, g(y_{k,x}))$, and on $(g(y_{k,x}), b]$ if $g(y_{k,x}) < b$. Accordingly, the distribution of $g(Y)$ is absolutely continuous on $[x, b]$ outside the point $g(y_{k,x}) \in D$. Thus, if $k \geq 1$ and A is an l–null set such that $y_{k,x} \notin A$, then $P(g(Y) \in A) = P(g_k(x, y) \in A) = 0$. Consequently, from all $x \in [a, b]$ the probability distributions of the transition to the right, i.e., given at least one claim, are absolutely continuous outside the set D.

We will now consider the absolute continuity in case 1°. Let $x \in [a, b],\ C \subset [a, b] - D,\ C \in \mathcal{B},\ l(C) = 0$. We have for all n

$$\text{(6.5)} \qquad P_u^n(x, C) = \int P_u^{n-1}(x, dz) P_u(z, C),$$

where P_u^n is the n-step transition probability. If there is at least one claim in the $(n-1)'$th year, $P_u(z, C) = 0$ for all z, since $P_u(z, \cdot)$, given at least one claim, is absolutely continuous outside D. Thus $P_u^n(x, C) = 0$ for all u and all x.

On the other hand, if the $(n-1)$'th year is claims-free, $P_u(z, C) = 0$ for all z by the construction. Hence $P_u^n(x, C) = 0$ for all n and all u, whence $\pi_u(C) = \lim_n P_u^n(x, C) = 0$. Consequently, π_u is absolutely continuous in D^c.

We turn to case 2°. Let $x \in [a,b]$, $C \subset [a,b]-D$ be a Borel-null set and $u \in U$ be arbitrary. We will show that $P_u^n(x,C)=0$ when n is big enough. Assume that $n \geq m$, where m is the subindex in (2.6) in the definition of the skeleton, denoted m_0 in (2.6). Denote the event $\{X_n \in C \mid X_0 = x\}$ by E. It follows from the definition of the skeleton that E can be divided into subevents $EB_0, \ldots, EB_k$, where

(6.6)
$$
\begin{aligned}
&B_0 = \{\text{at least one claim in year } n-1\},\\
&B_j = \{\text{years } n-1,\ldots,n-j \text{ claims free, at least one claim in year}\\
&n-j-1\},\ j=1,\ldots,m-1.
\end{aligned}
$$

As in case 1°, $P_u(EB_0)=0$. Consider the event B_j, $1 \leq j \leq m-1$. Note that it follows from the definition of the skeleton that $\psi(x_i)=x_{i+1},\ i=0,\ldots,m-1$, where $x_i \in D$, and that from (6.3), 2° follows that ψ is one-to-one from $[x_{m-1},b]$ to $[a,x_1]$. Hence it follows from the construction that $X_n \in C$ if and only if $X_{n-j} \in (\psi^{-1})^j(C)$ and, in addition, that $(\psi^{-1})^j(C) \subset D^c$.

The function $(\psi^{-1})^j$ is continuously differentiable on the closed interval $[a,x_1]$. Hence the derivative of $(\psi^{-1})^j$ is limited on $[a,x_1]$. This implies that $l[(\psi^{-1})^j(C)] = 0$. We have

$$
\begin{aligned}
P_u(EB_j) &= P_u(B_j)P_u(E|B_j) \leq P_u(E|B_j)\\
&= \int P_u^{n-j-1}(x,dz)p_u(z,(\psi^{-1})^j(C),
\end{aligned}
$$

where $p_u(z,(\psi^{-1})^j(C))$ is the probability that the chain moves from z to $(\psi^{-1})^j(C)$, given at least one claim in year $n-j-1$. Since there exists at least one claim in year $n-j-1$, $(\psi^{-1})^j(C) \subset D^c$ and $l[(\psi^{-1})^j(C)]=0$, it follows that $p_u(z,(\psi^{-1})(C))=0$ for all z. Consequently, $P_u(EB_j)=0$ for all j, whence $P_u^n(x,C)=0$ for all x and all u when $n \geq m$. Similarly as in case 1°, we get that the probability measures π_u are absolutely continuous in D^c for all u.

We turn to the absolute continuity of ν_u outside a finite set, with starting point $x \in [a,b]$, see (3.18). The proof in case 1° is similar to that for π_u. We turn to case 2°. We choose the finite exception set to be (cf. Corollary 3.4)

$$
H' = D \cup \{\psi^n(x) \mid n=1,\ldots,m-1\}.
$$

We have shown that when $n \geq m$, the measures π_u and $P_u^n(x,\cdot)$ are absolutely continuous outside $D \subset H'$. Thus it suffices to show that the probability measures $P_u^n(x,\cdot)$ are absolutely continuous outside H' when $n<m$.

Let $n<m$ and C a null set, $C \subset [a,b]-H'$. We have the corresponding partition for the event $\{X_n \in C|X_0=x\}$ as in (6.6) with the "additional" event $B' = \{\text{years } 1,\ldots,n \text{ are claims–free}\}$. It remains to notice that $P_u^n(x,C)$, given B', is equal to zero. This is clear since $P_u^n(x,\psi^n(x))$, given B', is equal to 1.

It follows from (6.2) and (6.4) that condition (4.2) is valid. This completes the proof.

Appendix

We give an example in the classic case where the unconstrained minimizing problem (see 5.8) gives a non-monotonic solution and show how the problem can be solved when the monotonicity condition (5.9) is required. Our example is a reduced version of an example in Gilde and Sundt (1989). It does not aim to be realistic but only to serve as an illustration of the phenomenon and of the solution of the problem.

Before going to the example, we briefly consider the minimizing problem (5.8) under constraints (5.9) generally. Let assumption 4.7 be valid. We have

$$B(y_1, \ldots, y_K) = \sum_{j=1}^{k}\Big(\sum_{i=1}^{K}(y_i - \lambda_j)^2 \pi_j(i)\Big)p_j = \sum_{i=1}^{K}\Big(\sum_{j=1}^{k}(y_i - \lambda_j)^2 \pi_j(i) p_j\Big).$$

The functions

$$f_i(y_i) = \sum_{j=1}^{k}(y_i - \lambda_j)^2 \pi_j(i) p_j, \quad i = 1, \ldots, K$$

are differentiable and strictly convex and get a unique minimum at the point

$$(A1) \qquad y_i^0 = \frac{\sum_{j=1}^{k} \lambda_j \pi_j(i) p_j}{\sum_{j=1}^{k} \pi_j(i) p_j},$$

where $f_i' = \dfrac{\partial B}{\partial y_i} = 0$. Consequently,

$$B = B(y_1, \ldots, y_K) = \sum_{i=1}^{K} f_i(y_i)$$

gets its unique minimum in R^K at $\overline{y}_0 = (y_1^0, \ldots, y_K^0)$. This can be proven also by probabilistic arguments, cf. Pesonen (1963) and Norberg (1976). Let F be the set of the vectors of R^K which satisfy constraints (5.9). If $\overline{y}_0 \in F$, it is the unique optimal solution in F (and in any case the unique optimal solution of the unconstrained minimizing problem).

If $\overline{y}_0 \notin F$, the existing unique optimal solution in F can be found by decomposing B to lower-dimensional differentiable strictly convex functions. We will demonstrate this in a special case by means of the following Example.

Example: We describe the transition mechanism of the BMS by means of the table below. The table should be read as follows: $a_{i,m} = l$ means that l claims cause a transition from class i to class m, and $a_{i,m} = l+$ that a number of claims $\geq l$ cause transition from class i to class m. Note that in this Example we have followed a common convention in the classic case that after claims-free years the policy moves "to the right", not to the left as in the general case in this paper.

Table

	1	2	3	4	5
1	1+			0	
2	1+			0	
3	1+			0	
4	2+	1			0
5	2+		1		0

$$= \begin{bmatrix} a_{1,1}, \dots, a_{1,5} \\ \\ \\ \\ a_{5,1}, \dots, a_{5,5} \end{bmatrix}$$

We assume that

(A2) there are two claims intensities $\lambda_1 = 0.1$ and $\lambda_2 = 0.7$ with associated probabilities $p_1 = 0.8,\ p_2 = 0.2,$ respectively.

We consider the problem minimizing B under constraints

$$(A3) \qquad \lambda_1 \leq y_5 \leq \cdots \leq y_1 \leq \lambda_2$$

and denote the set of vectors $(y_1, \dots, y_5)$ fulfilling (A3) by F.

The claims intensities λ_1 and λ_2 associated with the table above induce Markov chains P_{λ_1} and P_{λ_2}, respectively. Solving the corresponding invariant probability distributions gives

$$\begin{aligned} (A4) \qquad & \pi_1(1) = 0.0133,\ \pi_1(2) = 0.0078,\ \pi_1(3) = 0.0741,\ \pi_1(4) = 0.0861, \\ & \pi_1(5) = 0.8187 \\ & \pi_2(1) = 0.3308,\ \pi_2(2) = 0.0869,\ \pi_2(3) = 0.0857,\ \pi_2(4) = 0.2500, \\ & \pi_2(5) = 0.2466. \end{aligned}$$

Thus we get by (A1), (A2) and (A4)

$$(A5) \qquad \overline{y}_0 = (0.6169, 0.5398, 0.2346, 0.3523, 0.1420).$$

We notice that $y_4^0 > y_3^0$, whence $\overline{y}_0 \notin F$, cf. (A3). We will now seek for the existing unique vector $(y_1', \dots, y_5')$ where the strictly convex continuous function B gets its minimum in the convex compact set F.

We make the following decomposition

$$(A6) \qquad B(y_1, \dots, y_5) = \overline{B}(y_1, y_2, y_5) + \tilde{B}(y_3, y_4),$$

where $\overline{B} = \sum_{j=1}^{2} \Big(\sum_{i=1,2,5} (y_i - \lambda_j)^2 \pi_j(i) \Big) p_j$ and $\tilde{B} = \sum_{j=1}^{2} \Big(\sum_{i=3,4} (y_i - \lambda_j)^2 \pi_j(i) \Big) p_j$.
Both $\tilde{B}$ and $\overline{B}$ are differentiable and strictly convex. Let us consider $\tilde{B}$. Since $\tilde{B}$ is continuous and strictly convex, it gets unique minimum in the compact convex set

$$\tilde{F} = \{(y_3, y_4) \in R^2 : \alpha_1 \le y_4 \le y_3 \le \alpha_2\}.$$

Since $y_3^0 < y_4^0$, the point (y_3^0, y_4^0) where the partial derivatives vanish, does not belong to $\tilde{F}$. Hence $\tilde{B}$ cannot get its minimum in $\tilde{F}$ at an inner point of $\tilde{B}$. Accordingly, $\tilde{B}$ gets its minimum in $\tilde{F}$ on the border of $\tilde{F}$, i.e., on one of the lines $y_3 = \alpha_2,\ y_4 = \alpha_1,\ y_3 = y_4$. The restriction of $\tilde{B}$ to the line $y_3 = \alpha_2$ has a minimum at (0.7,0.3523), to the line $y_4 = \alpha_1$ at (0.2346,0.1) and to the line $y_3 = y_4$ at (0.3063,0.3063). Comparison of the values of $\tilde{B}$ at the points above shows that $\tilde{B}$ is smallest at (0.3063,0.3063), whence $\tilde{B}$ gets its unique minimum in $\tilde{F}$ at this point. Further, $\overline{B}$ gets its unique minimum in R^3 at $(0.6169, 0.5398, 0.1420)$, cf. (A5). Let $\hat{y} = (0.6169, 0.5398, 0.3063, 0.3063, 0.1420)$. Clearly, $\hat{y} \in F$, and by (A6) B gets at $\hat{y}$ its unique minimum in F.

References

Baione, F., Levantesi, S. and Menzietti, M., 2002, The development of an optimal Bonus-Malus system in a competitive market. ASTIN Bulletin, **32**, 159–169.

Ben–Tal, A. and Nemirovski, A., 2001, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications* (MPS–SIAM Series on Optimization, Philadelphia).

Bonsdorff, H., 1992, On the convergence rate of Bonus-Malus systems. *ASTIN Bulletin*, **22**, 217–223.

Bonsdorff, H., 2005, On asymptotic properties of Bonus-Malus systems based on the number and on the size of the claims. Scandinavian Actuarial Journal, **2005**, 309–320.

Borgan, Ø, Hoem, J.M. and Norberg, R., 1981, A nonasymptotic criterion for the evaluation of automobile Bonus systems. Scandinavian Actuarial Journal, **1981**, 165–178.

Ewing, G. M., 1969, *Calculus of Variations with Applications* (Dover Publications, New York).

Gelfand, I. M. and Fomin, S.V., 1963, *Calculus of Variations* (Prentice–Hall, Englewood Cliffs, New Jersey).

Gilde, V. and Sundt, B., 1989, On Bonus systems with credibility scales. Scandinavian Actuarial Journal, **1989**, 13–22.

Heras, A., Gil, J.A., Garcia-Pineda, P. and Vilar, J.L., 2004, An application of linear programming to Bonus Malus system design. ASTIN Bulletin, **34**, 435–456.

Kelley, J.L., 1955, *General Topology* (Princeton: D. van Nostrand Company, Inc.)

Lemaire, J., 1995, *Bonus-Malus systems in automobile insurance.* (Kluwer Academic Publishers, Boston).

Luenberger, D.G., 1969, *Optimization by Vector Space Methods* (John Wiley and Sons, New York).

Norberg, R., 1976, A credibility theory for automobile bonus systems. Scandinavian Actuarial Journal, **1976**, 92–107.

Nummelin, E., 1984, *General Irreducible Markov Chains and Non-negative Operators.* In Cambridge Tracts in Mathematics, **83** (Cambridge University Press, Cambridge.)

Orey, S., 1971, *Lecture Notes on Limit Theorems for Markov Chain Transition Probabilities* (London: Van Nostrand Reinhold Company).

Pesonen, E., 1962, A numerical method of finding a suitable bonus scale. ASTIN Bulletin, **2**, 102–108.

Robert, C.P. and Casella, G., 2000, *Monte Carlo Statistical Methods* (New York: Springer).

Verico, P., 2002, Bonus-Malus systems: "Lack of transparency" and adequacy measure, ASTIN Bulletin, **32**, 315–318.

Address for correspondence:
Heikki Bonsdorff
Mellstenintie 9 C 1
02170 Espoo
Finland
E–mail: heikki.bonsdorff@kolumbus.fi