THE ASTIN BULLETIN

INTERNATIONAL JOURNAL FOR
ACTUARIAL STUDIES
IN NON-LIFE INSURANCE AND RISK THEORY

VOL. 10, PART 2

MARCH 1979

CONTENTS

* * *

The Committee of Astin is not responsible for statements made or opinions expressed in the articles, criticisms and discussions published in the Bulletin.

Le Comité d'Astin rappelle que seul l'auteur de chaque publication est responsable des faits qu'il expose et des opinions qu'il exprime.

PRINTED IN THE NETHERLANDS

In Memoriam Paul Thyrion

Paul Thyrion s'est éteint à l'aube du 2 juin 1978. La nouvelle de sa mort s'est répandue rapidement dans le monde de l'assurance belge et parmi les nombreux amis qu'il comptait à l'étranger, frappant de consternation et d'émotion tous ceux qui l'avaient connu et qui appréciaient ses éminentes qualités.

Paul Thyrion avait 62 ans. Il était Ingénieur Civil de l'Ecole Royale Militaire de Belgique, Actuaire de l'Université Catholique de Louvain, Lieutenant-Colonel de réserve et Commandeur de l'Ordre de Léopold II. Il était entré à la Royale Belge en 1950 et en était devenu le Directeur Général en 1974. Commissaire à la Société Générale de Belgique, Administrateur de plusieurs sociétés, membre de la Commission des Assurances et du Conseil de Direction de l'Union Professionnelle des entreprises d'assurances, Paul Thyrion avait été Président de l'A.S.T.I.N. de 1968 à 1970 et Président de l'Association Royale des Actuaires Belges de 1971 à 1974. Il était également membre du Bureau de l'Association Actuarielle Internationale au sein duquel il occupait le poste de trésorier.

Les contributions actuarielles de Paul Thyrion, en particulier celles dans le domaine de la théorie collective du risque et dans celui des modèles "non-life", sont nombreuses, de grande valeur et universellement appréciées.

Mais Paul Thyrion n'avait pas que des qualités scientifiques, si éminentes soient-elles; il était aussi, tout simplement et dans toute l'acception du terme, un grand homme, doué d'une riche personnalité, droite, généreuse et noble.

Il n'était pas question pour lui que "Fair was foul and foul was fair": il ne pratiquait pas la confusion des valeurs.

Ses rapports avec les jeunes actuaires étaient toujours constructifs, sympathiques et stimulants; il les encourageait volontiers dans leurs recherches, n'hésitait jamais à consacrer une partie de son temps libre, rare et précieux, à l'examen de leurs travaux, et les conseillait toujours judicieusement.

On se rend mieux compte maintenant, avec le recul des jours, de l'énorme courage dont a dû faire preuve Paul Thyrion dans les dernières années de sa vie. On le disait quelque peu souffrant; en réalité, il était atteint d'un mal implacable dont lui seul, à part peut-être ses proches, connaissait l'existence. Jusqu'à ses derniers instants, il nous aura donné l'exemple d'une noblesse discrète et sereine.

J. Adam

LETTER TO THE EDITOR

Dear Sir:

I am delighted at the independent verification of my thesis in *Astin Bulletin*, **9**, 213, that the gamma distribution generally produces results nearer the truth of $F(x, t)$ than the so-called NP2, when both approximations are fitted by means of the first three moments of F.

In his contribution to the Astin Colloquium in Washington, D. C., T. Pentikäinen has reproduced 15 of the 24 comparisons I made in the first three and the fifth sets shown in my table thus obtaining 11 in favour of the gamma (a slight improvement over my 16 in 24!). He has added $48 - 15 = 33$ new results showing, in his table, that in 11 of them the gamma is superior to NP2 and that there is supposedly equality in 14 results. However, using Pentikäinen's own tabular values of $1 - F(x,t)$, 10 of the 14 "equalities" turn out to favour gamma and only three have the same value to the number of decimals shown.

Summarizing these results we have:

	Number of comparisons	Number in favour of gamma
Pentikàinen (Astin Colloquium):		
Taken from Seal	15	11
New	33	$11 + 10 + 1\frac{1}{2}$ (half of 3)
Seal (*Astin Bull.*):		
Not used in Pentikainen's extraction	9	5
Remainder (*viz.* fourth, sixth and seventh sets)	14	11
Total	71	$49\frac{1}{2}$

Several lines and groups of lines in this table produce a ratio of close to 70 % in favour of gamma.

Yours very truly,

HILARY L. SEAL

AUGUST 1977

The ASTIN Bulletin 10 (1979) 131-148

A NUMERICAL ILLUSTRATION OF OPTIMAL SEMILINEAR CREDIBILITY*

FL. DE VYLDER AND Y. BALLEGEER

INTRODUCTION

The homogeneous (in time) model of credibility theory is defined by a sequence $\Theta, X_1, X_2, \ldots$ of random variables, where for $\Theta = \theta$ fixed, the variables $X_1, X_2, \ldots$ are independent and equidistributed. The structure variable Θ may be interpreted as the parameter of a contract chosen at random in a fixed portfolio, the variable X_k as the total cost (or number) of the claims of the kth year of that contract.

Bühlmann's linear credibility premium of the year $t+1$ may be written in the form

$$f(X_1) + \ldots + f(X_t), \tag{1}$$

where f is a linear function. In optimal semilinear credibility, we look for an optimal f, not necessarily linear, such that (1) is closest to X_{t+1} in the least squares sense. In the first section we prove that this optimal f, denoted by f^*, is solution of an integral equation of *Fredholm* type, which reduces to a system of linear equations in the case of a finite portfolio. That is a portfolio in which Θ and X_k can assume only a finite number of values.

In the second section we see that the structure of such a portfolio is closely connected with the decomposition of a quadratic form in a sum of squares of linear forms.

In the last section we calculate numerically the optimal premium for a concrete portfolio in automobile insurance. We limit ourselves to the considera-tion of the number of claims. The optimal premium is compared with the usual linear premium. The difference is far from negligible.

As basic statistics we need the probabilities

$$p_{ij} = P(X_1 = i,\ X_2 = j)$$

In the third section we give a simple general solution to the subsidiary problem of adjusting the matrix p_{ij} of such probabilities.

1. THE FUNDAMENTAL RESULT

1.1. *Hypotheses. Notations. Definitions*

We consider a sequence $\Theta, X_1, X_2, \ldots$ of random variables such that for $\Theta = \theta$ fix-ed, the variables $X_1, X_2, \ldots$ are conditionally independent and equidistributed.

* Presented at the 12th ASTIN Colloqium, Portimão, October 1975.

All variables considered are supposed to have finite second order moments. The *risk premium* of each year is defined by

$$m_\Theta = E(X_1 \mid \Theta).$$

Here, and also hereafter in similar situations, the index 1 could be replaced by another one. The variables $X_1, X_2, \ldots$ are exchangeable in the sense of *De Finetti*. More generally, for each function f of one variable, we denote by f_Θ the random variable

$$f_\Theta = E(f(X_1) \mid \Theta)$$

Hereafter t will be a fixed positive integer. It is the number of years that we have already observed our portfolio. We have to make forecasts for the year $t+1$. Since t is fixed, the dependence on t is not always indicated in our notations.

1.2. *Lemma*

(I) For each couple f, g of functions of one variable:

(2) $$E(f(X_1)\, g(X_2)) = E(f_\Theta g(X_2)) = E(f(X_1) g_\Theta) = E(f_\Theta\, g_\Theta)$$

(II) For each function f of one variable and each function φ of t variables:

(3) $$E(\varphi(X_1, \ldots, X_t)\, f(X_{t+1})) = E(\varphi(X_1, \ldots, X_t)\, f_\Theta)$$

(III) For each function f of one variable:

(4) $$E(f(X_{t+1}) \mid X_1, X_2, \ldots, X_t) = E(f_\Theta \mid X_1, \ldots, X_t)$$

Demonstration.

(i) Using the conditional independence of X_1, X_2 for fixed Θ:

$$E(f(X_1)\, g(X_2)) = EE(f(X_1)\, g(X_2) \mid \Theta) = E(E(f(X_1) \mid \Theta)\, E(g(X_2) \mid \Theta)) = E(f_\Theta g_\Theta)$$

Also

$$E(f_\Theta g(X_2)) = EE(f_\Theta g(X_2) \mid \Theta) = E(f_\Theta E(g(X_2) \mid \Theta)) = E(f_\Theta g_\Theta)$$

and similarly

$$E(f(X_1)\, g_\Theta) = E(f_\Theta g_\Theta)$$

(ii) Writing

$$\varphi_\Theta = E(\varphi(X_1, \ldots, X_t) \mid \Theta),$$

we have in a similar way the more general result

$$E(\varphi(X_1, \ldots, X_t)\, f(X_{t+1})) = E(\varphi_\Theta f_\Theta) = E(\varphi(X_1, \ldots, X_t)\, f_\Theta)$$

(iii) From the conditional independence of $X_1, X_2, \ldots, X_{t+1}$, for fixed Θ, it follows that

$$f_\Theta = E(f(X_{t+1}) \mid \Theta) = E(f(X_{t+1}) \mid \Theta, X_1, \ldots, X_t)$$

Then, by applying the operator $E(. \mid X_1, \ldots, X_t)$ and using a general property of conditional expectations:

$$E(f_\Theta \mid X_1, \ldots, X_t) = E(E(f(X_{t+1}) \mid \Theta, X_1, \ldots, X_t) \mid X_1, \ldots, X_t) = E(f(X_{t+1}) \mid X_1, \ldots, X_t)$$

1.3. *Theorem*

Let f^* be a solution of

(5) $$E(X_2 \mid X_1) = f^*(X_1) + (t-1)\, E(f^*(X_2) \mid X_1)$$

Then, for every function f:

(6) $$E(m_\Theta - f^*(X_1) - \ldots - f^*(X_t))^2 \leqslant E(m_\Theta - f(X_1) - \ldots - f(X_t))^2$$

The mean square error in the approximation of m_Θ by $f^*(X_1) + \ldots + f^*(X_t)$ is given by

(7) $$E(m_\Theta - f^*(X_1) - \ldots - f^*(X_t))^2 = E(X_1X_2) - t\, E(X_1 f^*(X_2))$$

If g^* also satisfies

(8) $$E(X_2 \mid X_1) = g^*(X_1) + (t-1)\, E(g^*(X_2) \mid X_1),$$

then

(9) $$f^*(X_1) = g^*(X_1) \qquad \text{a.e.}$$

Demonstration.

Multiplying (5) by $f(X_1)$ and taking the mean value, we have

(10) $$E(f(X_1)\, X_2) = E(f(X_1)\, f^*(X_1)) + (t-1)\, E(f(X_1)\, f^*(X_2))$$

In particular, for $f = f^*$, we have

(11) $$E(f^*(X_1)\, X_2) = E(f^*(X_1))^2 + (t-1)\, E(f^*(X_1)\, f^*(X_2))$$

Using (2), we have for every f:

$$E(m_\Theta - f(X_1) - \ldots - f(X_t))^2 =$$
$$E(m_\Theta^2) - 2t\, E(m_\Theta f(X_1)) + E(f(X_1) + \ldots + f(X_t))^2 =$$
$$E(m_\Theta^2) - 2t\, E(m_\Theta f(X_1)) + t\, E\, f^2(X_1) + t(t-1)\, E(f(X_1)\, f(X_2)) =$$

(12) $$E(X_1X_2) - 2tE(f(X_1)\, X_2)) + t\, E\, f^2(X_1) + t(t-1)\, E(f(X_1)\, f(X_2))$$

Taking $f = f^*$ and using (11), we have

$$E(m_\Theta - f^*(X_1) - \ldots - f^*(X_t))^2 =$$
$$E(X_1 X_2) - 2t\, E(f^*(X_1)\, X_2) + t[E(f^*(X_1))^2 + (t-1)\, E(f^*(X_1)\, f^*(X_2))]$$
$$= E(X_1 X_2) - 2t\, E(f^*(X_1)\, X_2) + t\, E(f^*(X_1)\, X_2) =$$
$$E(X_1 X_2) - t\, E(f^*(X_1)\, X_2) \tag{13}$$

Since X_1 and X_2 are exchangeable, this proves (7). Neglecting a factor t, using (12) and (13), the difference between the second and the first member of (6) equals

$$d = E(f^*(X_1)\, X_2) - 2E(f(X_1)\, X_2) + E\, f^2(X_1) + (t-1)\, E(f(X_1)\, f(X_2))$$

Replacing the first two terms by their expression given by (10) and (11) and using (2), we have

$$\begin{aligned} d = {} & E(f^*(X_1))^2 && + (t-1)\, E(f^*(X_1)\, f^*(X_2)) \\ & - 2\, E(f(X_1)\, f^*(X_1)) && - 2(t-1)\, E(f(X_1)\, f^*(X_2)) \\ & + E(f(X_1))^2 && + (t-1)\, E(f(X_1)\, f(X_2)) = \end{aligned}$$
$$E(f^*(X_1) - f(X_1))^2 + (t-1)\, [E(f^*_\Theta)^2 - 2E(f_\Theta f^*_\Theta) + E(f_\Theta)^2] =$$
$$E(f^*(X_1) - f(X_1))^2 + (t-1)\, E(f^*_\Theta - f_\Theta)^2 \geqslant 0$$

This proves (6) and it only remains to show that (9) is true. Writing $h^* = f^* - g^*$, we have from (5) and (8):

$$0 = h^*(X_1) + (t-1)\, E(h^*(X_2) \mid X_1)$$

Multiplying this last relation by $h^*(X_1)$ and taking the mean value, we have

$$0 = E(h^*(X_1))^2 + (t-1)\, E(h^*(X_1)\, h^*(X_2))$$

or, by (2):

$$0 = E(h(^*X_1))^2 + (t-1)\, E(h^*_\Theta)^2$$

This implies

$$E(h^*(X_1))^2 = 0$$

and thus (9).

1.4. *Corollary*

Let f^* be solution of (5). Then, for each f:

$$E(X_{t+1} - f^*(X_1) - \ldots - f^*(X_t))^2 \leqslant E(X_{t+1} - f(X_1) - \ldots - f(X_t))^2 \tag{14}$$

Demonstration.

Using (3) it easily follows that for every function φ of t variables we have

$$E(X_{t+1} - \varphi(X_1, \ldots, X_t))^2 = E(X_{t+1} - m_\Theta)^2 + E(m_\Theta - \varphi(X_1, \ldots, X_t))^2$$

The difference between the members of (14) then is the same as that between the members of (6).

1.5. *Remark. Notation. Definition*

In DE VYLDER (1976), the fundamental relation (5) is derived in a geometrical way. In that paper the existence of f^* is proved.

The *optimal semilinear credibility premium* of the year $t+1$ is defined and denoted by

$$E^*(X_{t+1} \mid X_1, \ldots, X_t) = f^*(X_1) + \ldots + f^*(X_t), \tag{15}$$

where f^* is solution of (5).

1.6. *Theorem*

$$E\, E^*(X_{t+1} \mid X_1, \ldots, X_t) = E(X_{t+1}) \tag{16}$$

Demonstration.

Follows from (5) and (15) by taking the mean values.

1.7. *Determination of the Optimal Premium*

If the variables X_1 and X_2 have a joint density $p(x, y)$, then equation (5) becomes

$$\int y\, p(x, y)\, dy = f^*(x) \int p(x, y)\, dy + (t-1) \int f^*(y)\, p(x, y)\, dy \tag{17}$$

This is an integral equation of *Fredholm* type for the unknown function f^*.

If X_1 can only assume, with probability one, a finite number of values, say $0, 1, 2, \ldots, n$, then (5) becomes the linear system

$$\sum_{j=0}^{n} j p_{ij} = f_i^* \sum_{j=0}^{n} p_{ij} + (t-1) \sum_{j=0}^{n} f_j^* p_{ij} \quad (i = 0, \ldots, n), \tag{18}$$

where

$$p_{ij} = P(X_1 = i, X_2 = j), \tag{19}$$

$$f_i^* = f^*(i). \tag{20}$$

Equations (17) and (18) may serve as well for theoretical investigations as for the numerical computation of the optimal premium. Only the joint distribution of X_1 and X_2 is needed.

1.8. *The Linear Credibility Premium*

We shall denote the usual linear credibility premium of the year $t+1$ by

$$\bar{E}(X_{t+1} \mid X_1, \ldots, X_t) = (1-Z)\, E(X_1) + \frac{Z}{t} (X_1 + \ldots + X_t), \tag{21}$$

where

$$Z = \frac{t\, cov(X_1, X_2)}{var\, X_1 + (t-1)\, cov(X_1, X_2)} \tag{22}$$

The mean square error in the approximation of m_Θ by this premium equals

$$(1 - Z)\, cov(X_1, X_2). \tag{23}$$

By what precedes, it is never less than the mean square error in the approximation of m_Θ by the optimal premium, given by (7).

2. FINITE PORTFOLIOS AND QUADRATIC FORMS

2.1. *Hypotheses. Definition*

From now on we assume that the range of values of X_1 is a finite set of numbers say 0, 1, 2, ..., n.

We use the notation (19) for p_{ij} and set

$$p_i = P(X_1 = i) = \sum_{j=0}^{n} p_{ij} \quad (i = 0, 1, \ldots, n)$$

We denote by Q_p the quadratic form in the variables $x_0, x_1, \ldots, x_n$:

$$Q_p = \sum_{i,j=0}^{n} p_{ij}\, x_i x_j \tag{24}$$

(In the notation Q_p, p is of course not a numerical index, but a fixed symbol related to the notation p_{ij}.)

If Θ also can only assume a finite number of distinct values, say $\theta_0, \theta_1, \ldots, \theta_\nu$, we call the portfolio a *finite portfolio* and we write

$$u_\alpha = P(\Theta = \theta_\alpha), \tag{25}$$

$$p_{i/\alpha} = P(X_1 = i \mid \Theta = \theta_\alpha). \tag{26}$$

$$\begin{pmatrix} \alpha = 0, 1, \ldots, \nu \\ i = 0, 1, \ldots, n \end{pmatrix}$$

The numbers (25) and (26) completely describe our portfolio. For example:

$$p_{ijk\cdots} = P(X_1 = i, X_2 = j, X_3 = k, \ldots) = \sum_{\alpha=0}^{\nu} u_\alpha\, p_{i/\alpha}\, p_{j/\alpha}\, p_{k/\alpha} \cdots \tag{27}$$

Note that it is not assumed that the portfolio be finite in the following theorem.

2.2. *Theorem*

The $(n+1) \times (n+1)$ matrix $[p_{ij}]$ is semidefinite positive.

Demonstration.

For every function f of one variable, we have by (2):

$$E(f(X_1)\, f(X_2)) = E f_\Theta^2 \geqslant 0$$

Writing $f(i) = x_j$, this gives

$$Q_p = \sum_{i,j=0}^{n} p_{ij}\, x_i x_j \geqslant 0$$

for every value of $x_0, x_1, \ldots, x_n$

2.3. *Theorem*

Let $[q_{ij}]$ be an arbitrary $(n+1) \times (n+1)$ symmetric matrix with nonnegative elements adding up to unity. Define $q_i (i = 0, \ldots, n)$ by

$$q_i = \sum_{j=0}^{n} q_{ij}$$

Then, if one of the matrices $[q_{ij}]$ or $[q_{ij} - q_i q_j]$ is semidefinite positive, so is the other.

Demonstration.

Let Q_q and R_q be the quadratic forms

$$Q_q = \sum_{i,j=0}^{n} q_{ij}\, x_i x_j,$$

$$R_q = \sum_{i,j=0}^{n} (q_{ij} - q_i q_j)\, x_i x_j = Q_q - \left(\sum_{j=0}^{n} q_i\, x_i\right)^2$$

Then

$$Q_q = R_q + \left(\sum_{i=0}^{n} q_i\, x_i\right)^2$$

and if R_q is semidefinite positive, so is Q_q, à fortiori.

Conversely, let Q_q be semidefinite positive. Define the couple of random variables Y_1, Y_2 by

$$P(Y_1 = i, Y_2 = j) = q_{ij} \qquad (i, j = 0, 1, \ldots, n)$$

For every f we have, setting $f(i) = x_i$:

$$E(f(Y_1)\, f(Y_2)) = \sum_{i,j=0}^{n} f(i)\, f(j)\, q_{ij} = \sum_{i,j=0}^{n} q_{ij}\, x_i x_j \geqslant 0$$

since Q_q is semidefinite positive. In particular, for the function $f - Ef(Y_1) = f - Ef(Y_2)$, we have

$$E((f(Y_1) - Ef(Y_1))\, (f(Y_2) - Ef(Y_2))) \geqslant 0$$

or

$$R_q = \sum_{i,j=0}^{n} (q_{ij} - q_i q_j)\, x_i x_j \geqslant 0$$

2.4. *Theorem*

In the finite portfolio the form Q_p equals

$$Q_p = \sum_{\alpha=0}^{\nu} u_\alpha \left(\sum_{i=0}^{n} p_{i/\alpha} x_i \right)^2$$

Demonstration.

By (27):

$$Q_p = \sum_{i,j=0}^{n} p_{ij} x_i x_j = \sum_{\alpha=0}^{\nu} u_\alpha \sum_{i=0}^{n} p_{i/\alpha} x_i \sum_{j=0}^{n} p_{j/\alpha} x_j = \sum_{\alpha=0}^{\nu} u_\alpha \left(\sum_{i=0}^{n} p_{i/\alpha} x_i \right)^2$$

2.5. *Theorem*

Let $Q_q = \sum_{i,j=0}^{n} q_{ij} x_i x_j$ be a quadratic form with nonnegative symmetric coefficients q_{ij} adding up to unity. Then, to every decomposition

$$Q_q = \sum_{i,j=0}^{n} q_{ij} x_i x_j = \sum_{\alpha=0}^{\nu} \left(\sum_{i=0}^{n} a_{i\alpha} x_i \right)^2 \tag{28}$$

of Q_q in a sum of squares of linear forms with nonnegative coefficients $a_{i\alpha}$, there corresponds a finite portfolio for which

$$p_{ij} = q_{ij}, \tag{29}$$

$$u_\alpha = \left(\sum_{i=0}^{n} a_{i\alpha} \right)^2, \tag{30}$$

$$p_{i/\alpha} = a_{i\alpha} \Big/ \sum_{i=0}^{n} a_{i\alpha} \tag{31}$$

$$(i = 0, \ldots, n; \quad \alpha = 0, \ldots, \nu)$$

Demonstration.

We suppose of course that no linear form of the decomposition is the zero form.

Define u_α and $p_{i/\alpha}$ by (30) and (31). From (31) we have

$$\sum_{i=0}^{n} p_{i/\alpha} = 1 \qquad (\alpha = 0, \ldots, \nu).$$

By setting $x_0 = x_1 = \ldots = x_n = 1$ in (28), we have $\sum_{\alpha=0}^{\nu} u_\alpha = 1$

Also

$$q_{ij} = \sum_{\alpha=0}^{\nu} a_{i\alpha} a_{j\alpha} = \sum_{\alpha=0}^{\nu} u_\alpha p_{i/\alpha} p_{j/\alpha} = p_{ij}$$

by taking the coefficient of $x_i x_j$ in (28) and using (30) and (31).

2.6. *Remarks*

(I) Given the matrix $[p_{ij}]$, every possible finite portfolio for which (19) is valid thus results from a decomposition of Q_p in a sum of squares of linear forms with nonnegative coefficients. For all such possible portfolios, the credible premium (optimal or linear) will be the same.

(II) By 2.2., a necessary condition on a given matrix $[q_{ij}]$ to be the $[p_{ij}]$ matrix of some portfolio, finite or not, is that $[q_{ij}]$ be semidefinite positive.

(III) In the classical theory of decomposition of a quadratic form in a sum of squares of linear forms, the latter are generally independent and in number not larger than the dimension of the matrix of the quadratic form. For a decomposition giving rise to a portfolio, this is no longer needed. On the other side, we need linear forms with nonnegative coefficients, which is not the case in the classical theory.

(IV) As a simple illustration, we consider the form Q in two variables

$$Q = \frac{1}{29}(3x^2 + 12xy + 14y^2)$$

Among a lot of others, three possible decompositions are

$$Q = \frac{4}{29}\left(\frac{x}{2} + \frac{y}{2}\right)^2 + \frac{9}{29}\left(\frac{x}{3} + \frac{2y}{3}\right)^2 + \frac{16}{29}\left(\frac{x}{4} + \frac{3y}{4}\right)^2,$$

$$Q = \frac{27}{29}\left(\frac{x}{3} + \frac{2y}{3}\right)^2 + \frac{2}{29}\left(0\,x + 1\,y\right)^2,$$

$$Q = \frac{200}{203}\left(\frac{3x}{10} + \frac{7y}{10}\right)^2 + \frac{3}{203}\left(1\,x + 0y\right)^2$$

To these three decompositions correspond three different finite portfolios with same $[p_{ij}]$ matrix equal to

$$\begin{bmatrix} 3/29 & 6/29 \\ 6/29 & 14/29 \end{bmatrix}$$

For each of the three portfolios we would find the same optimal premium and the same linear credibility premium.

If we had a decomposition with only one square of a linear form, the two variables X_1 and X_2 should be independent. So the third decomposition shows that, in the present case, these variables are "nearly" independent.

3. ADJUSTMENT OF A $[p_{ij}]$ MATRIX

3.1. *The Problem*

In the next section, we apply the theory to a concrete portfolio in automobile insurance. We limit ourselves to the consideration of the number of claims. Then p_{ij} is the probability of i claims in one year, say the first, and j claims in another year, say the second, for a contract chosen at random in the portfolio.

Practically, the probability p_{ij} is estimated by an observed frequency q_{ij}. Except perhaps for estimates from very large samples, the matrix $[q_{ij}]$, of course symmetrized in the obvious way, does not fit in the theory because generally it is not semidefinite positive. So it must be transformed, as slightly as possible, in a usable matrix $[p_{ij}]$.

3.2. *Smoothing on a Fixed Ascending Diagonal*

Suppose, for a moment, that the parameter θ of each fixed contract is interpreted as the mean number of claims in one year, and that the arrivals are poissonnian. Then we should have

$$P(X_1 = i \mid \Theta = \theta) = e^{-\theta} \frac{\theta^i}{i!} \quad (i = 0, 1, 2, \ldots) \tag{32}$$

But since, for practical reasons, we do not consider a number of claims in one year greater than a fixed integer n, we replace (32) by

$$P(X_1 = i \mid \Theta = \theta) = c_{n,\theta}\, e^{-\theta} \frac{\theta^i}{i!} \quad (i = 0, 1, \ldots, n) \tag{33}$$

where $c_{n,\theta}$ is the suitable norming factor.

Denoting by $U(\theta)$ the structure function of the portfolio, we have, for a contract chosen at random

$$p_{ij} = \int_0^\infty c_{n,\theta}^2\, e^{-2\theta} \frac{\theta^{i+j}}{i!\,j!}\, dU(\theta) \qquad (i, j = 0, 1, \ldots, n)$$

For the probability of k ($k = 0, 1, \ldots, 2n$) claims in two years, we have then

$$_2p_k = \sum_{\substack{i,j=0\\ i+j=k}}^{n} p_{ij} = \left(\sum_{\substack{i,j=0\\ i+j=k}}^{n} \frac{1}{i!\,j!}\right) \int_0^\infty c_{n,\theta}^2\, e^{-2\theta}\, \theta^{i+j}\, dU(\theta) \tag{34}$$

So, for $i + j = k$ ($i, j = 0, 1, \ldots, n$), p_{ij} and $_2p_k$ are related by

$$p_{ij} = a_{ij}\, {}_2p_k \tag{35}$$

where

$$a_{ij} = \frac{\dfrac{1}{i!\,j!}}{\displaystyle\sum_{\substack{i,j=0\\ i+j=k}}^{n} \frac{1}{i!\,j!}}, \quad (i, j = 0, \ldots, n;\; i+j=k) \tag{36}$$

If we take

$$_2q_k = \sum_{\substack{i,j=0\\ i+j=k}}^{n} q_{ij} \tag{37}$$

and then use (35) with $_2p_k = {_2q_k}$, we have a first adjustment of the matrix $[q_{ij}]$. Since, for fixed k, the elements a_{ij} of (36) add up to unity, it is immediate that the sum of the elements of each ascending diagonal is the same in the initial and the adjusted matrix.

We reached (35), starting from a poissonnian hypothesis. Now we keep only (35) and abandon the poissonnian hypothesis, because this relation is in fact true in a more general situation. For example, if the factor $c_{n,0}^2\, e^{-2\theta}$ is replaced by another one not depending on i or j, then (35) remains true with a_{ij} given by (36).

3.3. *Extrapolation for the Last Ascending Diagonals*

For statistics deriving from small samples, the above method does not yet furnish a semidefinite positive $[p_{ij}]$ matrix. So a preliminary smoothing of the $_2q_k$'s is necessary.

If, again for one moment, we make the poissonnian hypothesis and do not neglect claims in number greater than n in one year, then we have

$$_2p_k = \int_0^\infty e^{-2\theta} \frac{(2\theta)^k}{k!}\, dU(\theta), \qquad (k = 0, 1, 2, \ldots) \tag{38}$$

Writing

$$r_k = k!\; {_2p_k} \qquad (k = 0, 1, 2, \ldots) \tag{39}$$

we have

$$r_k = \int e^{-2\theta}\, (2\theta)^k\, dU(\theta) \qquad (k = 0, 1, 2, \ldots)$$

From this relation it can be proved that

$$r_k^2 \leqslant r_{k-1}\, r_{k+1}, \qquad (k = 1, 2, \ldots) \tag{40}$$

and that equality for some k can only hold in a portfolio of homogeneous composition (that means: Θ = constant a.e.), in which case it holds for every

k. In the case of a binomial negative distribution for the total number of claims in a fixed period (here 2 years), which amounts to a gamma density for Θ, it can be verified that, for $k \to \infty$, we have

$$\frac{r_{k-1}\, r_{k+1}}{r_k} \to 1$$

These considerations suggest the following method of adjustment. We take

$$r_0 = 0!\; {}_2q_0,\; r_1 = 1!\; {}_2q_1,\; \ldots,\; r_{k_0} = k_0!\; {}_2q_{k_0}$$

and, from k_0 on, taken as large as possible, we set

$$(41) \qquad r_k = (1 + \varepsilon_{k,\alpha,\beta,\ldots})\, \frac{r^2_{k-1}}{r_{k-2}} \qquad (k \geqslant k_0 + 1)$$

where $\varepsilon_{k,\alpha,\beta,\ldots}$ is a positive quantity, decreasing with increasing k and containing parameters α, β, . . . to be determined in function of some requirements for the adjusted matrix. There is of course some arbitrariness in the choice of $\varepsilon_{k,\alpha,\beta,\ldots}$, but as we shall see in our numerical illustration of next section, this quantity, when properly chosen, introcuces only very small probabilities.

From the preceding discussion we only retain (41) and (39), because it is not difficult to see that (40) is valid in a more general situation than the poissonnian from which we started.

4. NUMERICAL ILLUSTRATION

4.1. *Basic Statistics*

The statistics used are those of Table 1.

TABLE 1: BASIC STATISTICS

$i \backslash j$	0	1	2	3	4	5
0	784	103	13	2	2	0
1	119	33	5	1	0	0
2	18	5	3	2	0	0
3	1	1	0	0	1	0
4	0	0	0	0	0	0
5	1	0	0	0	0	0

The number at the intersection of row i and column j in this table is the number of automobiles with i claims one year and j claims the following year among 1094 automobiles.

These statistics were established by *P. Thyrion* and used in THYRION (1972) and afterwards in DE VYLDER (1975).

On dividing by 1094 and symmetrizing, we obtain the matrix $[q_{ij}]$ of Table 2.

Most of our following numerical results were computed with a precision of 15 à 16 significant digits. Often, however, we reproduce the intermediate results with 3 significant digits only.

TABLE 2: NON ADJUSTED SYMMETRIZED MATRIX $[q_{ij}]$

		.717	.203	.0585	.0119	.00640	
							.00274
$i=0$	.717	.101	.0142	.00137	.000914	.000457	0
$i=1$	.101	.0302	.00457	.000914	0	0	.000914
$i=2$	.0142	.00457	.00274	.000914	0	0	0
$i=3$	.00137	.000914	.000914	0	.000457	0	0
$i=4$	.000914	0	0	.000457	0	0	0
$i=5$	.000457	0	0	0	0	0	
	.835	.137	.0224	.00366	.00137	.000457	
	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	

TABLE 3: ADJUSTED MATRIX $[p_{ij}]$

		.717	.203	.0585	.0119	.00493	
							.00261
$i=0$	.717	.101	.0146	.00149	.000308	.0000815	.00139
$i=1$	.101	.0293	.00446	.00123	.000408	.000134	.000676
$i=2$	.0146	.00446	.00185	.000815	.000335	.000127	000296
$i=3$	.00149	.00123	.000815	.000447	.000211	.0000909	.000116
$i=4$	.000308	.000408	.000335	.000211	.000114	.0000579	.0000410
$i=5$	.0000815	.000134	.000127	.0000909	.0000579	.0000410	
	.835	.137	.0222	.00428	.00143	.000532	
	$j=0$	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	

4.2. *Adjustment*

Our aim is to find a semidefinite positive matrix $(p_{ij}]$ as close as possible to the matrix $[q_{ij}]$.

Following the method explained in the preceding section, we take

$$
\begin{aligned}
{}_2p_0 &= q_{00} &&= .717\\
{}_2p_1 &= q_{01} + q_{10} &&= .203\\
{}_2p_2 &= q_{02} + q_{11} + q_{20} &&= .0585\\
{}_2p_3 &= q_{03} + q_{12} + q_{21} + q_{30} &&= .0119
\end{aligned}
$$

We tried of course to keep also for ${}_2p_4$ the observed corresponding frequency .00640, but this was unsuccessfull. From the above values, we have the value of r_0, r_1, r_2, r_3 by (39). We set

$$r_k = \left(1 + \frac{\alpha}{\beta^{k-4}}\right) \frac{r_{k-1}^2}{r_{k-2}} \qquad (k = 4, 5, \ldots, 10)$$

because we observed that a quantity $\varepsilon_{k,\alpha,\beta,\ldots}$ in (41) rapidly converging to zero gives a ${}_2p_4$ closer to .00640 than one converging more slowly to zero. From the values of the r_k ($k = 4, 5, \ldots, 10$) we deduce those of the p_k by (39) and choose α and β to satisfy

$$\sum_{k=0}^{10} {}_2p_k = 1 \tag{42}$$

From the values of the ${}_2p_k$ we then deduce those of the p_{ij} by (35).

For fixed β it is not difficult to determine α, with the required precision, from (42). So we still dispose of β. For a previously indicated reason, we try to take β as large as possible. Now, by calculating the characteristic values, we observed that for $\beta = 2$, we obtained a semidefinite positive matrix $[p_{ij}]$, while for $\beta = 4$, there appeared one negative characteristic value. We then tried the values $\beta = 2.1$, $\beta = 2.2$, $\ldots$, $\beta = 3.8$, $\beta = 3.9$ and found that for $\beta = 3$ all characteristic were still positive, while for $\beta = 3.1$ there appeared a negative one. In fact, for $\beta = 3$ there was a characteristic value so small that we preferred to take $\beta = 2.9$, although this was not essential. The corresponding value of α is $\alpha = 1.723\,569\,981\,730\,550$. The characteristic values of the adjusted $[p_{ij}]$ matrix are .732 .0151 .00154 .0000835 .0000096 .000000081. For the adjusted matrix, the mean value of the number of claims in one year is .202607, while for the original matrix it is .200640. Instead of (42), we could have used the relation making these mean values equal, but then, unless we introduced a new parameter, we would have had to change proportionally the now kept fixed quantities ${}_2p_0$, ${}_2p_1$, ${}_2p_2$, ${}_2p_3$. Since the difference between the two means is small in our actual adjustment, we keep it as it is.

A glance at Tables 2 and 3 is enough to be convinced of the quality of our adjustment, especially when one looks at the partial sums indicated in the margins.

A characteristic of our adjustment is that it used only the numbers ${}_2p_k$ and not the decomposition of such a number on the corresponding ascending diagonal. In other words, instead of Table 1, we used only the frequencies of k claims in two years. It seems that our method can be adapted for the case were the frequency of k claims in one year is the only statistical material.

4.3. *A Theorically Possible Portfolio Compatible with the $[p_{ij}]$ Matrix*

If we decompose the quadratic form Q_p by *Lagrange*'s method (successive completion of squares), taking the variables in the order $x_0, x_1, \ldots, x_5$, we find after some normalisations:

$$Q_p = \sum_{i,j=0}^{5} p_{ij}\, x_i\, x_j =$$

$$
\begin{array}{lllllll}
.972 & (.859x_0 + & .122x_1 + & .0175x_2 + & .00178x_3 + & .000369x_4 + & .0000983x_5)^2 \\
+ .0237 & & (.793x_1 + & .127\ x_2 + & .0545\ x_3 + & 0194\ \ x_4 + & .00652\ \ x_5)^2 \\
+ .00401 & & & (.540\ x_2 + & .287\ \ x_3 + & .125\ \ x_4 + & .0488\ \ x_5)^2 \\
+ .000285 & & & & (.392\ \ x_3 + & .373\ \ x_4 + & .235\ \ x_5)^2 \\
+ .000025 & & & & & (.32\ \ x_4 + & .68\ \ x_5)^2 \\
+\ 0000032 & & & & & & (\ \ x_5)^2
\end{array}
$$

As explained in section 2, this decomposition defines a portfolio for which the $[p_{ij}]$ matrix is our adjusted $[p_{ij}]$.

This portfolio does not serve in the sequel, but we calculated it to make sure that our adjusted $[p_{ij}]$ matrix is not a theorically impossible one.

4.4. *The Optimal Premium and the Linear Premium*

To make comparisons sensefull, these premiums are of course calculated both for the adjusted $[p_{ij}]$ matrix.

4.4.1. The optimal premium

From (18), we obtain, in table 4, the values of the f_i^* for the indicated values of $t+1$.

TABLE 4: COMPONENTS OF THE OPTIMAL PREMIUM
$E^*(X_{t+1} \mid X_1, \ldots, X_t) = f^*_{x_1} + f^*_{x_2} + \ldots + f^*_{x_t}$

$t+1$	f_0^*	f_1^*	f_2^*	f_3^*	f_4^*	f_5^*
2	.163922	.322485	.566282	1.285385	1.712988	2.060772
3	.070165	.201312	.385665	.938154	1.252583	1.495804
4	.041312	.154117	.301413	.748922	.993612	1.174104
5	.027911	.127399	.249519	.624949	822816	962363
6	.020394	.109677	.213655	.536605	.701129	.812356
7	015681	.096841	.187171	.470247	.609979	.700767
8	.012500	.087009	166728	.418507	539185	.614733
9	.010237	.079179	.150432	.377009	482654	.546539
10	.008562	.072763	.137116	.342977	.436504	.491274
20	002613	.041181	.073446	.179860	.219454	.238560
30	001290	.029042	.050507	.121616	144603	.155734
50	.000526	.018364	.031328	.073604	.084674	.091804
99	.000159	.009688	.016461	.037222	.040897	.046423
100	.000156	.009596	.016305	.036848	.040458	.045969

TABLE 5: PROBABILITY p_i OF i CLAIMS IN ONE YEAR

p_0	p_1	p_2	p_3	p_4	p_5
.834599	.136944	.022208	004283	.001434	.000532

From this table it follows, for example, that the optimal semilinear forecast of the number of claims in the 4th year, for a driver with 2, 2, 0 claims in the preceding years is

$$E^*(X_3|X_1=2, X_2=2, X_3=0) = f_2^*+f_2^*+f_0^* =$$
$$.301413+.301413+.041312 = .644138$$

To make a verification possible of relation (16) which amounts to

$$t\,E(f_{X_1}^*) = E(X_1)$$

or

$$t\sum_{i=0}^{5} p_i f_i^* = E(X_1)$$

where

$$E(X_1) = .202607$$

we give, in table 5, the values of p_i, the probability of i claims in one year, with a precision greater than in Table 3.

4.4.2. The linear premium

The credibility factor Z in (21), given in (22), is expressed in Table 6 for various values of $t+1$. Intermediate values computed from the not printed 15 digits precise $[p_{ij}]$ matrix are also indicated.

TABLE 6: CREDIBILITY FACTOR Z IN LINEAR FORECAST

$$E(X_{t+1}\mid X_1, \ldots, X_t) =$$
$$(1-Z)\,E(X_1) + Z/t\,(X_1+\ldots+X_t)$$

$t+1$	Z
2	.231545
3	.376024
4	.474773
5	.546537
6	.601048
7	.643859
8	.678373
9	.706788
10	.730590
20	.851300
30	.897310
50	.936566
99	.967244
100	.967564
$E(X_1)$	= .202607
$E(X_1^2)$	= .300577
$E(X_1X_2)$	= .101142
$\text{var}(X_1)$	= .259527
$\text{cov}(X_1, X_2)$	= .060092

The linear forecast for the above considered driver is

$$\bar{E}(X_3 \mid X_1 = 2, X_2 = 2, X_3 = 0) = (1 - Z)\, E(X_1) + Z\,(2 + 2 + 0)\,/\,3 = .739445$$

4.4.3. The mean quadratic errors

Table 7 gives, for different values o. $t+1$, the mean square error in the approximation of the risk premium m_Θ by the optimal premium and the linear premium. The formulae used are (7) and (23).

As expected, the optimal premium is always closer to m_Θ, and thus to X_{t+1}, than the linear premium.

TABLE 7: MEAN SQUARE ERROR FOR THE OPTIMAL AND THE LINEAR PREMIUM

$t+1$	Optimal	Linear
2	.0438	.0462
3	.0347	.0375
4	.0288	.0316
5	.0247	.0272
6	.0217	.0240
7	.0193	.0214
8	.0175	.0193
9	.0164	.0176
10	.0147	.0162
20	.00822	.00894
30	.00574	.00617
50	.00359	.00381
99	.00188	.00197
100	.00186	.00195

4.4.4. Comparative Tables

The values of the optimal premium and the linear one are given in Tables 8 and 9 for $t+1 = 2$ and $t+1 = 3$ respectively. As is seen, these values may differ very much, even for relatively small values of X_1, X_2. Consider, for example the case $X_1 = 0$, $X_2 = 3$ in Table 9.

TABLE 8: OPTIMAL AND LINEAR FORECAST FOR SECOND YEAR ($t+1 = 2$)

X_1	Optimal	Linear
0	.163922	.155694
1	.322485	.387239
2	.566282	.618784
3	1.285385	.850329
4	1.712988	1.081873
5	2.060772	1.313419

TABLE 9: OPTIMAL AND LINEAR FORECAST FOR THE THIRD YEAR ($t + 1 = 3$) [a]

$X_1 \setminus X_2$	0	1	2	3	4	5
0	140330 .126422	.271477 .314434	.455830 502446	1.008319 690458	1.322748 .878470	1.565969 1.066482
1	.271477 .314434	.402624 .502446	.586977 .690458	1 139466 .878470	1.453895 1.066482	1 697116 1 254494
2	.455830 .502446	.586977 .690458	771330 .878470	1.323819 1 066482	1.638248 1.254494	1.881469 1.442506
3	1.008319 690458	1.139466 .878470	1.323819 1.066482	1.876308 1 254494	2.190737 1.442506	2.433958 1.630518
4	1 322748 .878470	1 453895 1.066482	1.638248 1.254494	2 190737 1 442506	2.505166 1 630518	2.748387 1.818530
5	1.565969 1.066482	1.697116 1.254494	1.881469 1.442506	2.433958 1.630518	2.748387 1.818530	2 991608 2.006542

[a] The first number indicated is the optimal premium, the number beneath it, the linear one

In Table 9, the linear premium does of course not very on an ascending diagonal. This is not the case for the optimal premium. For example, 3 and 0 claims respectively in the first and the second year is much worse than 2 and 1 claim.

REFERENCES

BUHLMANN, H. (1967). Experience Rating and Credibility, *The Astin Bulletin*, **4**, 199-207.

BUHLMANN, H. (1970) *Mathematical Methods in Risk Theory*, Berlin: Springer-Verlag.

DE VYLDER, FL (1975). *Introduction aux théories actuarielles de crédibilité*, Bruxelles: Office des Assureurs de Belgique.

DE VYLDER, FL. (1976) Optimal Semilinear Credibility, *Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker*, **76**, 27-40.

JEWELL, W. S. (1973) Multi-Dimensional Credibility, University of California, Berkeley

THYRION, P. (1972). Quelques observations statistiques sur la variable "nombre de sinistres" en assurance automobile, *The Astin Bulletin*, **6**, 203-211.

The ASTIN Bulletin 10 (1979) 149-162

PROBABILITY OF RUIN UNDER INFLATIONARY CONDITIONS OR UNDER EXPERIENCE RATING

G. C. TAYLOR *

The effect of inflation of premium income and claims size distribution, but *not* of free reserves, on the probability of ruin of an insurer is studied.

An interesting similarity between this problem and the ruin problem in an experience-rated scheme is exhibited This similarity allows the deduction of parallel results for the two problems in later sections

It is shown that the probability of ruin is always increased when the (constant) inflation rate is increased.

The distribution of aggregate claims under inflationary conditions is described and used to calculate an upper bound on the ruin probability. Some numerical examples show that this bound is not always sharp enough to be practically useful. It is also shown, however, that this bound can be used to construct an approximation of the effect of inflation on ruin probability.

It is shown that if inflation occurs at a constant rate, then *ruin is certain*, irrespective of the smallness of that rate and of the largeness of initial free reserves and the safety margin in the premium. The corresponding result for experience-rated schemes is that a practical and "intuitively reasonable" experience-rating scheme leads eventually to certain ruin.

Finally, a simple modification of the techniques of the paper is made in order to bring investment income into account.

1. INTRODUCTION

The probability of ruin of a risk business has been studied under various conditions in the past, e.g. LUNDBERG (1909), CRAMÉR (1930, 1955), and others. Most of these studies have assumed that the risk process is either a stationary one or can be made stationary by means of a simple transformation.

Such models of the risk process do not include the case in which the phenomenon of inflation is causing the volume of premium income and of claims but not free reserves to vary in time. In current times, when rates of inflation in many countries have been, are and appear likely to remain for some time at high levels, it seems advisable to examine the impact of this feature on the solvency of the risk business in so far as this latter is described by the probability of ruin.

In carrying out this examination, it is noted that the operation of certain types of experience rating schemes is closely parallel to that of inflation on a "conventional" risk business, so that the methods foreshadowed in the preceding paragraph are also applicable to experience rated processes.

* The author gratefully acknowledges the use of facilities of the Swiss Reinsurance Company, Zurich, Switzerland in the preparation of this paper.

2. DESCRIPTION OF THE RISK PROCESS

We consider a risk process in which *premiums* received in the time-interval $[0, t]$ total $C(t)$ (the process begins at $t = 0$). Let $X(t)$ denote the *aggregation of claims occurring in the time-interval* $[0, t]$.

Suppose that $\{X(t),\ t \geqslant 0\}$ is a one-dimensional Markov process. Let $Z(t)$ denote the *free reserves at time t* and write x for $Z(0)$. Then

$$Z(t) = x + C(t) - X(t) \tag{1}$$

is also a one-dimensional Markov process.

Since $X(t)$ is the aggregation of claims up to time t, it is possible to write

$$X(t) = \sum_{i=1}^{N(t)} S_i,$$

where S_i is the random variable denoting the *size of the i-th claim* and $N(t)$ is the random variable denoting the *number of claims occurring in the time-interval* $[0, t]$.

Sometimes in the following sections, no further assumptions about the risk process will be made. At other times it will be necessary to place some restrictions on the random variables $N(t)$ and S_i.

3. ADDITION OF INFLATION TO THE RISK PROCESS

We now wish to superimpose an inflation process on the risk model described in Section 2. We suppose this process to be a deterministic one in that we assume the existence of a non-stochastic inflation factor $f(t)$ (> 0) at time t.

Premium volume at time t and also claims paid at time t are inflated by the factor $f(t)$ (assume $f(0) = 1$). Let C^*, X^* and Z^* represent the functions C, X and Z respectively after modification by the factor f. Then

$$C^*(t) = \int_0^t f(s)\, dC(s); \tag{2}$$

$$X^*(t) = \int_0^t f(s)\, dX(s)$$

$$= \sum_{i=1}^{N(t)} f(t_i)\, S_i, \tag{3}$$

where t_i is the epoch of the i-th claim;

$$Z^*(t) = x + C^*(t) - X^*(t). \tag{4}$$

Note that in (4) inflation is assumed to have no effect on free reserves. This is not unrealistic in the light of the experience of the last few years. In any case, this restriction is relaxed in Section 12.

4. THE RELATION BETWEEN INFLATION AND EXPERIENCE RATING

Consider a risk business subject to the same risk process as (1) except that each element of premium paid is modified by a refund or surcharge according to the difference between past premiums and past claims. Suppose that the precise form of this experience rating is such that the element of premium payable at time t is:

$$d\,C(t) = \{c - k[C(t) - X(t)]\}\,dt. \tag{5}$$

c being the *base rate* of premium payable, i.e. the premium rate when the experience follows its expected pattern exactly; and k being the *experience rating factor* at time t (normally, $0 < k < 1$).

It is easy to deduce from (5) that

$$C(t) = \frac{c}{k}[1 - e^{-kt}] + \sum_{i=1}^{N(t)} S_i[1 - e^{-k(t-t_i)}], \tag{6}$$

whence

$$Z(t) = x + e^{-kt}\int_0^t c\,e^{ks}\,ds - e^{-kt}\sum_{i=1}^{N(t)} S_i\,e^{kt_i},$$

or

$$\hat{Z}(t) = e^{kt}Z(t) = xe^{kt} + \int_0^t ce^{ks}\,ds - \sum_{i=1}^{N(t)} S_i\,e^{kt_i}. \tag{7}$$

From (2), (3) and (4) it can be seen that $\hat{Z}(t)$ represents *a "conventional" risk process subject to inflation at a continuous rate of k per unit time except that the initial free reserve also inflates at this rate instead of remaining constant as assumed in Section 3.*

In each of the following sections, this relation between a risk process in inflationary conditions and an experience-rated risk process permits the deduction of parallel results, although the emphasis is on the former in the section headings.

5. PROBABILITY OF RUIN IS NONDECREASING WITH INCREASING INFLATION

This result is proved by showing that any realization of $\{Z^*(t),\, t \geq 0\}$ leading to ruin also leads to ruin if the rate of inflation is increased.

Consider two Z^*-processes called Z_1^* and Z_2^* with associated inflation factors of f_1 and f_2 respectively. Suppose a particular realization of Z_1^* leads to ruin. Then for some t, we have

$$Z_1^*(t) < 0,\; Z_1^*(s) \geq 0 \text{ for } 0 \leq s < t. \tag{8}$$

Now, from (2), (3) and (4),

$$Z_2^*(t) - Z_1^*(t) = \int_0^t [f_2(s) - f_1(s)]\,dZ(s).$$

Integration by parts yields:

$$(9) \qquad Z_2^*(t) - Z_1^*(t) = g(t)Z_1^*(t) - \int_0^{t-0} Z_1^*(s)\, dg(s),$$

where $g(s) = f_2(s)/f_1(s) - 1$, and it has been assumed that this function is measurable. If $g(s)$ is a monotone nondecreasing function (recall that $g(0) = 0$), then $g(s) \geqslant 0$ for $s \geqslant 0$ and $dg(s) \geqslant 0$ for $s \geqslant 0$, and by (8) and (9)

$$(10) \qquad Z_2^*(t) - Z_1^*(t) \leqslant 0.$$

We may summarize the above in the following:

Result

If two Z^*-processes, Z_1^* and Z_2^*, are subject to measurable inflation factors of $f_1(t)$ and $f_2(t)$ such that the difference $f_2(t)/f_1(t)$ is nondecreasing with increasing t, then the probability of ruin (in finite or infinite time) is not less for the Z_2^*-process than for the Z_1^*-process.

Remarks

1. It is of course assumed that the initial reserves are the same in the Z_1^*- and Z_2^*-processses.
2. The result is entirely independent of the properties of the process Z. It includes, for example, cases where the claim number process is not Poisson, where sizes of different claims are not independent, etc.
3. The requirement that $f_2(t)/f_1(t)$ be monotone nondecreasing is easily seen to be equivalent to the requirement that the Z_2^*-inflation rate should always be not less than the Z_1^*-inflation rate in those cases where f_1 and f_2 are smooth and the term "inflation rate" therefore meaningful.

The situation for the experience-rated process $\hat{Z}$ is not so simple, However, in the case of zero initial reserves (i.e. $x = 0$), we see from a comparison of equations (4) and (7) that the $\hat{Z}$-process is exactly the same as a Z^*-process with $f(t) = \exp(kt)$. It follows, therefore, that, in this case, increasing k, the degree of experience rating, will *increase* the ruin probability.

We shall see further, in Section 11, that under experience rating the ultimate ($t = \infty$) probability of ruin is always 1.

That these results are not intuitive to some extent is clear from a paper by Seal (1969), in which he refers to the criticism that his simulated ruin probabilities (according to "conventional" risk processes) were too high. The suggestion is that in practice an insurer can use some kind of experience rating and, by basing premiums on past results, will be able to reduce the ruin probability.

The reasoning leading to this conclusion is probably somewhat along the following lines. There are two important classes of free reserve trajectory: that consisting of trajectories characterized by persistently light claims experience, and that characterized by persistently heavy claims experience. In the first case ruin does not occur whether experience-rated or not; in the second, premium rates are forced up by the poor experience, thus reducing the proportion of ruins.

The fallacy in such an argument is that it ignores the possibility of a light claims experience followed by a slightly heavier than usual experience. In this case the initial light experience forces premiums down so that the fund built up in this period is not particularly large, despite the absence of claims.

6. THE DISTRIBUTION OF AGGREGATE CLAIMS UNDER INFLATIONARY CONDITIONS

In this section we investigate the distribution of $X^*(t)$ under the more specific assumption that it is a compound Poisson variate, the claim number process having a Poisson parameter λ and the individual claim size distribution having d.f. $B(\cdot)$ at time zero. The method of obtaining the moment generating function (m.g.f.) of $X^*(t)$ is essentially that of ANDREWS and BRUNNSTROM (1976), though requiring some generalization since they take $B(\cdot)$ to be the d.f. of a single-point distribution.

Consider the time-interval $(jt/m, (j+1)\,t/m)$ where m is a very large positive integer and j is an integer between O and $(m-1)$. Because the length of this interval, t/m is small, the Poisson claim number process within it approximates a binomial process with parameters 1 and $\lambda t/m$. Therefore, the m.g.f. of aggregate claim amount in this small interval is:

$$(11) \qquad M_j^*(u) = \left(1 - \frac{\lambda t}{m}\right) + \frac{\lambda t}{m}\,\beta(uf(jt/m)) + O(m^{-2})$$

$$= 1 + \frac{\lambda t}{m}\,[\beta(u\,f\,(jt/m)) - 1] + O(m^{-2}),$$

where $\beta(u)$ is the m.g.f. associated with $B(\cdot)$. If the additional assumption of independence of sizes of different claims is made, then the cumulant generating function of $X^*(t)$ becomes:

$$(12) \qquad K^*(u, t) = \sum_{j=0}^{m-1} \log \left\{1 + \frac{\lambda t}{m}\,[\beta(u\,f(jt/m)) - 1] + O(m^{-2})\right\}$$

$$= \lambda \sum_{j=0}^{m-1} \frac{\beta(u\,f(jt/m)) - 1}{m/t} + O(m^{-1}).$$

Letting $m \to \infty$, we see that the c.g.f. becomes:

$$(13) \qquad K^*(u, t) = \lambda t \left[\frac{1}{t}\int_0^t \beta(u\,f(s))\,ds - 1\right].$$

From this it follows that the j-th cumulant of $X^*(t)$ is

$$\varkappa_j^*(t) = \lambda t\, \alpha_j \left\{\frac{1}{t} \int_0^t [f(s)]^j\, ds\right\}, \tag{14}$$

where α_j is the j-th order moment (about the origin) of the d.f. $B(\cdot)$ and the second factor on the right is the average value of $[f(s)]^j$ over $s \in [0, t]$.

Obviously, the m.g.f. of $X^*(t)$ is:

$$M^*(u, t) = \exp\left\{\lambda t \left[\frac{1}{t} \int_0^t \beta(uf(s))\, ds - 1\right]\right\} \tag{15}$$

In the most important special case, $f(s) = e^{ks}$, (13) and (14) can be put in a sometimes more convenient form. Equation (14) becomes:

$$\varkappa_j^*(t) = \lambda \alpha_j \left(e^{jkt} - 1\right) / jk, \tag{16}$$

whence

$$K^*(u, t) = (\lambda/k) \left[\sum_{j=1}^{\infty} \frac{\alpha_j}{j!} \frac{(u\, e^{kt})^j}{j} - \sum_{j=1}^{\infty} \frac{\alpha_j}{j!} \frac{u^j}{j}\right] \tag{17}$$

$$= (\lambda/k) \int_u^{u e^{kt}} \frac{\beta(v) - 1}{v}\, dv$$

7. AN UPPER BOUND ON THE PROBABILITY OF RUIN UNDER INFLATIONARY CONDITIONS

An upper bound on the ruin probability can be found using the method of GERBER (1973). Define $Y^*(t) = Z^*(t) - Z^*(0)$. Gerber shows that, if $\psi^*(x, t)$ is the probability of ruin before time t (in the model of Section 6), then

$$\psi^*(x, t) \leqslant \min_r e^{-rx} \max_{0 \leqslant s \leqslant t} E[\exp\{-r\, Y^*(s)\}], \tag{18}$$

where for the sake of simplicity we are now assuming that time has been so scaled that expected number of claims for unit time, i.e. λ, is equal to unity.

In our case this reduces to:

$$\psi^*(x, t) \leqslant \min_r e^{-rx} \max_{0 \leqslant s \leqslant t} \exp[-r\, C^*(s) + K^*(r, s)]. \tag{19}$$

Let us examine the square-bracketed term in (19). By (14), it is

$$s\left\{-r \cdot \frac{c}{s} \int_0^s f(u)\, du + \lambda \sum_{j=1}^{\infty} r^j \alpha_j / j\,! \cdot \frac{1}{s} \int_0^s [f(u)]^j\, du\right\}, \tag{20}$$

where c is premium income per unit time.

Since all claims are > 0, the α_j's are all > 0. Thus for large r, the higher powers of r dominate and expression (20) is positive and increasing. It also has a zero at $r = 0$. Differentiation (with s constant) shows that it has one turning point. Thus expression (20) is 0 at $r = 0$, becomes negative as r increases, and for s constant has a single real positive zero $\pi(s)$.

For $r > \pi(s)$ it is positive and increasing. In view of this, we can deduce from (18) that:

(21) $$\psi^*(x, t) \leqslant \min_{r \geqslant 0} e^{-rx} \max \{1, \exp [-r C^*(t) + K^*(r, t)]\},$$

since, for given $r \geqslant 0$, the maximum in (19) is 1 if $r \leqslant \pi(t)$, and is $-r C^*(t) + K^*(r, t)$ if $r \geqslant \pi(t)$. Note that, in (21) we consider only $r \geqslant 0$. This is because the maximum in (19) is always at least 1 (whether r is positive or negative), so that consideration of $r < 0$ tells us no more than that $\psi^*(x, t) \leqslant \exp(-rx)$ which is $\geqslant 1$ and can be improved upon by choosing $r = 0$ in (19). We can simplify (21) a little further by noting that the exponential term there is < 1 when $0 < r < \pi(t)$, and so

(22) $$\psi^*(x, t) \leqslant \min_{r \geqslant \pi(t)} \exp [-rx - r C^*(t) + K^*(r, t)].$$

where we recall that $r = \pi(t)$ is the unique real and positive solution of:

(23) $$- r C^*(t) + K^*(r, t) = 0.$$

The similarity between this result and Gerber's (19), both derived from (18) by very similar reasoning, is to be noted. The two formulas are easily seen to be identical if $f(t) = 1$ for all t.

Remark

Gerber (1973, p. 210) commented for the case $f(t) = 1$ that inequality (22) is rather sharp if t is not too small. It would follow then in our case of more general $f(t)$ that we could take the right side of inequality (22) as reasonable provided t is not too small and the rate of inflation underlying $f(t)$ is not too large.

In the case of an experience-rated scheme, the whole analysis goes through as before except that $Y^*(t)$ is replaced by:

$$\hat{Y}(t) = Y^*(t) + x(e^{kt} - 1).$$

Making this replacement and following through the previous working, we soon find that:

(24) $$\hat{\psi}(x, t) \leqslant \min_{r \geqslant \hat{\pi}(t)} \exp [-rxe^{kt} - r\hat{C}(t) + \hat{K}(r, t)]$$

where $\hat{C}$, $\hat{K}$ denote C^*, K^* with constant inflation rate k, and $r = \hat{\pi}(t)$ is the unique real and positive solution of the equation

$$(25) \qquad - r\,x(e^{kt}-1) - r\hat{C}(t) + \hat{K}(r, t) = 0.$$

8. AN APPROXIMATION

It would be useful to have on hand a simple approximation to the ratio $\psi_2^*(x, t) / \psi_1^*(x, t)$ where ψ_i^* is the ruin probability associated with inflation factor f_i. Table 1 in Section 10 shows that inequality (22) is not always as sharp as we would like, but that the ratio $\psi_2^*(x, t) / \psi_1^*(x, t)$ is usually approximated reasonably by the ratio of the upper bounds given by (22). At least this tends to be so in the "interesting" cases where probability of ruin is not too high.

This is demonstrated in Table 2 of Section 10.

9. INFLATION AND EXPERIENCE RATING COMBINED

There is no difficulty in combining an inflation factor of $f(t)$ and an experience rating factor of k. It is easily checked that reserves at time t are:

$$x + e^{-kt} \int_0^t c\, f(s)\, e^{ks}\, ds - e^{-kt} \sum_{i=1}^{N(t)} S_i\, f(t_i)\, e^{kt_i},$$

which leads us to consider the stochastic process,

$$\hat{Z}^*(t) = xe^{kt} + \hat{C}^*(t) - \hat{X}^*(t),$$

where $\hat{C}^*(t)$ and $\hat{X}^*(t)$ are the premium income and claims outgo respectively up to time t under the influence of an inflation factor of $\exp(kt)\, f(t)$.

10. NUMERICAL EXAMPLES

Consider the case in which the time-axis has been scaled in such a way that, in the absence of any inflation, the claim intensity is 1 per unit time. Suppose that money values have also been so scaled that (again in the absence of inflation) the distribution of individual claim size is $\chi^2_6/6$, i.e. m.g.f. is $(1 - r/3)^{-3}$. We shall assume constant rates of inflation, i.e. $f(t) = e^{kt}$, and consider the values $k = 0$, .05 and .15. Suppose that the basic premium income is 1.2 per unit time, thus allowing a safety margin of 20%. Then, by (22) and (23),

$$(26) \qquad \psi^*(x, t) \leq \min_{r \geq \pi(t)} \exp\left[- rx - 1.2\, r\, \frac{e^{kt}-1}{k} + K^*(r, t)\right],$$

where $r = \pi(t)$ is the real positive solution of

$$(27) \qquad - 1.2\, r\, \frac{e^{kt}-1}{k} + K^*(r, t) = 0.$$

In cases where the minimum in (26) is assumed for $r > \pi(t)$, the relevant value of r is that satisfying the equation:

$$- x - 1.2 \frac{e^{kt} - 1}{k} + \frac{1}{k} \left[\frac{\beta(re^{kt}) - 1}{r} - \frac{\beta(r) - 1}{r} \right] = 0$$

i.e.

$$(1 - \tfrac{1}{3} re^{kt})^{-3} - (1 - \tfrac{1}{3} r)^{-3} = r[kx + 1.2 (e^{kt} - 1)]. \tag{28}$$

Also

$$K^*(r, t) = \frac{1}{k} \int_r^{re^{kt}} \frac{(1 - \frac{1}{3} v)^{-3} - 1}{v} dv$$

$$= \frac{1}{3k} \int_r^{re^{kt}} [(1 - \tfrac{1}{3}v)^{-1} + (1 - \tfrac{1}{3}v)^{-2} + (1 - \tfrac{1}{3}v)^{-3}] dv$$

$$= \frac{1}{k} [- \log (1 - \tfrac{1}{3}v) + (1 - \tfrac{1}{3}v)^{-1} + \tfrac{1}{2}(1 - \tfrac{1}{3}v)^{-2}] \Big|_{v=r}^{re^{kt}}$$

We take initial reserves equal to 5 and, for each value of k, calculate for various t the upper bound (22) on $\psi^*(5, t)$ and the ratio of this bound to the corresponding bound on $\psi(5, t)$. The results are given in Tables 1 and 2 where the values of $\psi^*(x, t)$ obtained from a computer simulation are also given. The sample size for each simulated probability was 2400.

Similar calculations are made for the case of a negative exponential claim size distribution. Equation (28) is replaced by:

$$(1 - re^{kt})^{-1} - (1 - r)^{-1} = r [kx + 1.2 (e^{kt} - 1)].$$

i.e.

$$r = \tfrac{1}{2}(1 + e^{-kt}) [1 - \sqrt{1 - 4 (1 - 1/A) e^{-kt} (1 + e^{-kt})^{-2}}]$$

where

$$A = 1.2 + kx (e^{kt} - 1)^{-1}$$

Also,

$$K^*(r, t) = \frac{1}{k} \log \left\{ \frac{1 - r}{1 - re^{kt}} \right\}.$$

Tables 3 and 4 then summarize these calculations. Once again the results of a computer simulation (sample size again 2400) are given.

Several facts emerge from Tables 1 to 4.

TABLE 1: UPPER BOUND (22) ON $\psi^*(5, t)$ IN CASE OF A $\chi_6^2/6$ CLAIM SIZE DISTRIBUTION[a]

t	$k = 0$		$k = .05$		$k = .15$	
1	.021	(.0033)	.023	(.0038)	.031	(.0046)
2	.057	(.0096)	.071	(.014)	.105	(.024)
3	.094	(.023)	.122	(.031)	.194	(.048)
4	.126	(.032)	.169	(.055)	.283	(.091)
5	.154	(.054)	.212	(.074)	.364	(.150)
10	.235	(.098)	.360	(.172)	.631[b]	(.365)
25	.273	(.165)	.563[b]	(.383)	.944[b]	(.787)
∞	.273		1[b]		1[b]	

[a] Figures in parentheses are simulated ruin probabilities.
[b] Values based on $r = \pi(t)$.

TABLE 2: ESTIMATE OF RATIO $\psi^*(5, t)/\psi(5, t)$ BY THE RATIO OF THE CORRESPONDING UPPER BOUNDS (22) IN CASE OF A $\chi_6^2/6$ CLAIM SIZE DISTRIBUTION[a]

t	$k = .05$		$k = .15$	
1	1.10	(1.15)	1.48	(1.39)
2	1.25	(1.46)	1.84	(2.50)
3	1.30	(1.35)	2.06	(2.09)
4	1.34	(1.75)	2.25	(2.87)
5	1.38	(1.38)	2.36	(2.79)
10	1.53	(1.76)	2.69	(3.72)
25	2.06	(2.33)	3.46	(4.78)

[a] Figures in parentheses are taken from computer simulation.

TABLE 3: UPPER BOUND (22) ON $\psi^*(5, t)$ IN CASE OF A NEGATIVE EXPONENTIAL CLAIM SIZE DISTRIBUTION[a]

t	$k = 0$		$k = .05$		$k = .15$	
1	.108	(.009)	.117	(.011)	.136	(.017)
2	.182	(.035)	.205	(.041)	.258	(.053)
5	.311	(.096)	.379	(.121)	.529	(.200)
10	.397	(.158)	.520	(.233)	.883	(.436)

[a] Figures in parentheses are simulated ruin probabilities.

TABLE 4: ESTIMATE OF RATIO $\psi^*(5, t)/\psi(5, t)$ BY THE RATIO OF THE CORRESPONDING UPPER BOUNDS (22) IN CASE OF A NEGATIVE EXPONENTIAL CLAIM SIZE DISTRIBUTION[a]

t	$k = .05$		$k = .15$	
1	1.08	(1.22)	1.26	(1.89)
2	1.13	(1.17)	1.42	(1.51)
5	1.22	(1.26)	1.70	(2.08)
10	1.31	(1.47)	2.22	(2.76)

[a] Figures in parentheses are taken from computer simulation.

Firstly, in Tables 1 and 3 we see that the upper bound (22), even in the case of no inflation, is not as sharp as one might expect after a perusal of the calculations of GERBER (1973, p. 210). The bound does, however, improve with increasing t, whether inflation is present or not.

Secondly, for a given pair of inflation rates the ratio of upper bounds (22), as exemplified in Tables 2 and 4, can serve as a rough approximation to the ratio of the corresponding ruin probabilities, provided that these probabilities are not too large. Even though the simulated results of Tables 1 to 4 are based upon 2400 trials, the simulated low probabilities are still subject to random disturbance. However, for $k = .05$ in Table 2, the *average relative error* in the approximation to $\psi^*(5, t)/\psi(5, t)$ is 11%. The corresponding figure for $k = .15$ is 15%. If for $k = .15$, this error is calculated only on the basis of those t for which simulated probability is less than .2 (this corresponds to considering the values $t = 1, 2, 3, 4, 5$ for $k = .05$), then the average relative error is again only 10%.

In Table 4, the average relative error in the ratio for $t = 1, 2$ is 8% for $k = .05$. It is larger for $k = .15$ but mainly as a result of random error at $t = 1$ in the simulation.

Thirdly, as $\psi^*(5, t)$ increases with increasing t, the approximation to $\psi^*(5, t)/\psi(5, t)$ dealt with in Tables 2 and 4 becomes poorer.

In summary, it is fair to say that this approximation seems reasonable for $\psi^*(5, t) <$ about .2, but thereafter is rather dubious. However, the range $\psi^*(5, t) < .2$ is certainly the most interesting from a practical viewpoint.

11. EXPONENTIAL INFLATION MAKES ULTIMATE RUIN CERTAIN

The values of 1 given by (22) in the case $t = \infty$ are rather conspicuous in Table 1, and raise the question of whether ultimate ruin always occurs with probability 1 when inflation is present.

We consider here the case where there exists a constant $K > 0$ such that

$$\int_0^t f(s)\, ds \leqslant K f(t) \text{ for all } t. \tag{29}$$

For example, if there is a constant rate of inflation, i.e. the inflation factor is exponential, then (29) is satisfied. We also assume that the uninflated premium income is always received at a rate of c per unit time, and that individual claims in excess of cK occur with nonzero probability.

Under these conditions the rather discomforting answer to our question is that no matter how large the initial reserves, no matter how large the safety margin in premiums, no matter how small the rate of inflation (subject to (29)), the ultimate probability of ruin is always 1.

This is easily proved. Suppose that our assertion is untrue; then $\psi^*(x, t)$ approaches a limit (< 1) as $t \to \infty$.

Then

$$\left[\frac{d}{dt}\,\psi^*(x,t)\right] \Big/ \left[1-\psi^*(x,t)\right] \to 0 \text{ as } t \to \infty. \tag{30}$$

Now let $G(x, t, y)$ denote the probability that an insurer with initial reserve x will survive to time t and have reserves between 0 and y at that time.

If $B_t(.)$ denotes the d.f. of individual claim size at time t, then:

$$\lambda \int_0^\infty [1 - B_t(y)]\, d_yG(x, t, y) \;/\; \int_0^\infty d_yG(x, t, y) \to 0 \text{ as } t \to \infty \tag{31}$$

But reserves at time t are at most:

$$x + c\int_0^t f(s)\,ds \leqslant x + cKf(t) \tag{32}$$

By (31) and (32):

$$\lambda \int_0^\infty [1 - B_t(x + cKf(t))]\, d_yG(x, t, y) \;/\; \int_0^\infty d_yG(x, t, y)$$

$$\leqslant \lambda \int_0^\infty [1 - B_t(y)]\, d_yG(x, t, y) \;/\; \int_0^\infty d_yG(x, t, y)$$

$$\to 0 \text{ as } t \to \infty$$

i.e.

$$1 - B_t(x + cKf(t)) \to 0 \text{ as } t \to \infty \tag{33}$$

But, of course

$$B_t(z) = B(z \,/\, f(t)),$$

so that (33) becomes:

$$1 - B(cK + x \,/\, f(t)) \to 0 \text{ as } t \to \infty.$$

i.e.

$$1 - B(cK) = 0.$$

Since this contradicts our assumption that larger claims than cK (uninflated) *can* occur, our hypothesis of $\psi^*(x, t) < 1$ is false.

By an identical line of reasoning, we find that if individual claims in excess of $x + cK$ can occur in an experience rated scheme, then the probability of ultimate ruin is 1. This result was conjectured (though without any condition on the distribution of individual claim sizes) by Sidney Benjamin.

As was remarked in Section 5, this result is not entirely intuitive. However, it does become reasonable when one notes that (by formula (7)), the contribution to reserves at time t of all safety margins paid up to then is

$$(1+\eta)^{-1}\,\eta c\, e^{-kt}\int_0^t e^{ks}\,ds$$

$$= (\eta c \,/\, k)\,(1 - e^{-kt}) \,/\, (1+\eta)$$

where η is the proportion of risk premium taken as a safety margin. We see that accumulated safety margins converge to a finite limit with increasing t, i.e. the average safety margin per unit time tends to zero. In these circumstances, it is not surprising that $\hat{\psi}(x, \infty) = 1$.

This suggests that the experience rating formula (5) should be replaced by one which does not refund most safety margin. Perhaps, we could take

$$dC(t) = \left\{c - k\left[\frac{C(t)}{1+\eta} - X(t)\right]\right\} dt. \tag{34}$$

i.e. only the risk premium $C(t) / (1+\eta)$ is allowed for in the experience rating. Thus (34) can be rewritten as

$$dC(t) = \left\{\frac{\eta c}{1+\eta} + \left\{\frac{c}{1+\eta} - k\left[\frac{C(t)}{1+\eta} - X(t)\right]\right\}\right\} dt, \tag{35}$$

and we can see that a constant rate of safety margin $c\eta/(1+\eta)$ is maintained in addition to the experience rated risk premium.

However, there may be some sales difficulties with rating formula (34), since the proportion of the premium absorbed by the safety margin increases as the claims experience improves. One can well imagine the insured objecting to an increase in the relative safety margin being occasioned by a *favourable* experience.

12. ALLOWANCE FOR EARNINGS ON ASSETS

Of course, all of the preceding analysis has been made on the assumption that the free reserves of the insurer earn no interest. We now relax this assumption and suppose that interest is earned at a rate such that a unit invested at time zero accumulates to amount $A(t)$ at time t. Then the free reserves at time t under the operation of both interest and inflation are:

$$xA(t) + \int_0^t (f(s)\, A(t) \,/\, A(s))\, dC(s) - \int_0^t (f(s)\, A(t) \,/\, A(s))\, dX(s)$$

Discounting these free reserves back to time zero, we obtain

$$x + \int_0^t (f(s) \,/\, A(s))\, dC(s) - \int_0^t (f(s) \,/\, A(s))\, dX(s),$$

so that a process subject to an inflation factor $f(t)$ and an interest accumulation factor $A(t)$ is equivalent to a process with just an inflation factor of $f(t)/A(t)$. What matters, therefore, is whether rate of inflation is greater or less than the rate of interest. For example, if the difference between the force of inflation and the force of interest is constant and positive (be it ever so small), then the result of Section 11, viz. unit probability of ruin, still holds.

REFERENCES

ANDREWS, J. G. and R. R. L. BRUNNSTROM (1976). The discounted cost of a random sequence of failures, *Bulletin of the Institute of Mathematics and its Applications*, **12**, 275-279.

CRAMÉR, H. (1930). *On the mathematical theory of risk*, Skandia Jubilee Volume, Stockholm.

CRAMÉR, H. (1955). *Collective risk theory, a survey of the theory from the point of view of the theory of stochastic processes*, Skandia Jubilee Volume, Stockholm.

GERBER, H. U. (1973) Martingales in risk theory, *Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker*, **73**, 205-216.

LUNDBERG, F. (1909). Über die Theorie der Rückversicherung, *Transactions of the 6th International Congress of Actuaries*, **1**, 877-948.

SEAL, H. L. (1969) Simulation of the ruin potential of nonlife insurance companies, *Transactions of the Society of Actuaries*, **21**, 563-585.

The ASTIN Bulletin 10 (1978) 163-172

BAYESIANS LEARN WHILE WAITING *

WILLIAM S. JEWELL

In many estimation problems, incomplete as well as complete samples are available for Bayesian prediction. After developing the theory for a special, but useful family of distributions, examples are given in life testing, renewal risk processes, life contingencies, and the problem of estimating a defective distribution.

1. INTRODUCTION

In Bayesian prediction problems, one is interested in using observed values of a given process to update the prior knowledge about the process parameters, and thence to make better predictions about the process itself. Most of the theory concerns itself either with exact calculations using so-called natural-conjugate families of prior and likelihood distributions [1], or with best linear least-squares approximations, referred to in the actuarial literature as credibility theory [2]. However, both approaches consider only the use of complete data samples.

The purpose of this paper is to show that there are many situations in which *incomplete* observations also provide updating information, that is, Bayesians can learn while waiting for the finish of the sampling experiment. After developing the necessary theory and introducing the gamma-proportional-hazard family of distributions most appropriate for incomplete data formulations, examples are given from life testing, renewal risk processes, and life contingency reserving. It is shown in what sense an individual life (or cohort of similar lives) can learn about his (their) own remaining lifetime distributions with the passage of time. The paper concludes with the problem of estimating the parameters and the defect in a defective distribution.

2. MODEL

As is usual in Bayesian models, we assume that $\tilde{x}$, the random lifetime of interest, has a *likelihood distribution function*, $P(x \mid \theta)$, which depends upon an unknown random parameter $\tilde{\theta}$ which has a *prior distribution function*, $P(\theta)$. We use $P^c = 1 - P$ to denote the complementary distribution (or survival) function, and we assume that (continuous or discrete) densities exist, denoted by $p(x \mid \theta)$, $p(\theta)$ etc.

* An earlier version of this paper was presented at the 13th ASTIN Colloquium, Washington, D.C., May 1977.

[1] see AITCHINSON and DUNSMORE (1975).

[2] see JEWELL (1978).

The basic problem is to use observational data, sampled from the likelihood distribution with fixed, but unknown parameter, in Bayes' law to find the *posterior-to-data distribution* of the parameter, and thence to predict various moments and economic functions of the underlying lifetime process.

To illustrate the natural way in which incomplete samples arise, we consider a life-testing scheme in reliability, as in JEWELL (1977), in which:

1. N items, all with lifetimes drawn as samples from $P(x|\theta)$ with common and fixed θ, are put "on test" at epochs $\{t_i\}$, and removed from test at epochs $\{t_i + T_i\}$, $(i, 1, 2, \ldots, N)$;
2. C of these items (with indices in the set S) will have failed before removal with observed lifetimes $\{\tilde{x}_i = x_i \leq T_i\}$ $(i \in S)$;
3. The remaining lifetimes are not completely observed, since the items are still operating at removal, so it is known only that $\{\tilde{x}_i > T_i\}$ $(i \notin S)$.

Depending upon the experimental protocol, the $\{T_i\}$ may be fixed in advance, giving then a random C; or, C may be fixed in advance for a simultaneous test, giving a common, random time-on-test, T. Considering for a moment that the $\{T_i\}$ are fixed, and denoting the observed data by $D = \{y_1, y_2, \ldots, y_N; S\}$, where

$$\tilde{y}_i = \begin{cases} x_i & (i \in S) \\ T_i & (i \notin S) \end{cases}, \tag{2.1}$$

we can easily argue that the likehood density of this data set, given θ, is:

$$p(D \mid \theta) = \prod_{i \in S} p(x_i \mid \theta) \prod_{j \notin S} P^c(T_j \mid \theta). \tag{2.2}$$

Bayes' law then gives the predictive density for continued testing of items $j \notin S$, or for future experiments on other items with the same parameter value:

$$p(x \mid D) = \int p(x \mid \theta) \left[\frac{p(D \mid \theta)\, p(\theta)}{\int p(D \mid \theta')\, p(\theta')\, d\theta'} \right] d\theta. \tag{2.3}$$

The ratio in square brackets is the posterior-to-data parameter density, $p(\theta \mid D)$.

(2.2) is also useful for many other life testing protocols. Suppose that all items are put on test at the same epoch; the common testing interval T need not be fixed in advance, but may be a continuously-evaluated *stopping rule*, a possibly random decision to stop experimenting that depends upon the values $\{x_1, x_2, \ldots, x_C; S\}$ observed up to and including T, but not *directly* upon θ. In this case, the likelihood includes additional terms relating to the stopping rule that cancel out of the ratio in (2.3); the stopping rule is *non-informative*, and the likelihood kernel (2.2) is *sufficient for* θ. For instance, one could stop after the fifth failure, or at T equal to twice the first-observed complete lifetime.

3. THE PROPORTIONAL-HAZARD FAMILY

The calculation of (2.3) can, of course, be carried out by computer for any given prior and likelihood distributions. However, for model-building, it is desirable to use a family of distributions in which the calculations are especially tractable so that parametric behavior can be analyzed theoretically. Unfortunately, the Koopman-Pitman-Darmois exponential family of distributions so useful in credibility theory has no simple form for P^c; see JEWELL (1974).

However, a special case of the exponential family, the *proportional-hazard family*, has useful properties:

$$P^c(x \mid \theta) = e^{-\theta Q(x)}; \; p(x \mid \theta) = \theta q(x)\, e^{-\theta Q(x)}, \; (x \geq 0) \tag{3.1}$$

where $Q(x)$ is a monotone non-decreasing function ($Q(0) = 0$), and $q(x) = dQ(x)/dx$. We note:

1. $\theta Q(x)$ is the *cumulative hazard (failure) function*, making $q(x)$ a *unit- or prototype failure rate*;
2. If $\tilde{w}$ is a random variable with absolute failure rate, $q(w)$, and θ is an integer, the original lifetime, $\tilde{x}$, has a physical interpretation as

$$\tilde{x} = \min(\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_\theta);$$

3. This family includes the exponential, Weibull, and Gumbel (extreme-value) distributions.

The data likelihood (2.2) becomes:

$$p(D \mid \theta) = \prod_{i \in S} q(x_i)\, [\theta^C e^{-\theta TQT(D)}], \tag{3.2}$$

where TQT is a statistic,

$$TQT(D) = \sum_{i=1}^{N} Q(x_i) = \sum_{i \in S} Q(x_i) + \sum_{j \notin S} Q(T_j), \tag{3.3}$$

referred to in JEWELL (1977) as the *total-Q-on-test-statistic*, a generalization of the "total-time-on-test" concept of reliability life-testing. Note that if item k was already age S_k (and still working) when placed on test, then $Q(S_k)$ should be *subtracted* from the TQT.

A convenient natural conjugate prior for $\tilde{\theta}$ is the gamma density,

$$p(\theta) = p(\theta \mid C_0, Q_0) = \frac{Q_0 (Q_0\theta)^{C_0 - 1} e^{-\theta Q_0}}{\Gamma(C_0)}, \; (\theta \geq 0) \tag{3.4}$$

with hyperparameters C_0, Q_0; the usefulness of (3.4) in modelling uni-modal densities is well known. It is easy to see that Bayes' law then gives a posterior-to-data density of the parameter, $p(\theta \mid D)$, that is also gamma, with updated parameters:

$$C_0 \leftarrow C_0 + C;\ Q_0 \leftarrow Q_0 + TQT(D). \tag{3.5}$$

Furthermore, the updated means of $\tilde{\theta}$ and $\tilde{\theta}^{-1}$ obey the exact credibility formulae:

$$[E\{\tilde{\theta} \mid D\}]^{-1} = (1 - Z_1)\,[E\{\tilde{\theta}\}]^{-1} + Z_1 \left[\frac{C}{TQT(D)}\right]^{-1}, \tag{3.6}$$

$$E\{\tilde{\theta}^{-1} \mid D\} = (1 - Z_2)\,E\{\tilde{\theta}^{-1}\} + Z_2 \left[\frac{TQT(D)}{C}\right], \tag{3.7}$$

with credibility factors:

$$Z_1 = C/(C_0 + C);\ Z_2 = C/(C_0 - 1 + C). \tag{3.8}$$

The posterior-to-data variances are also easily obtained:

$$V\{\tilde{\theta} \mid D\} = \left[\frac{1}{C_0 + C}\right] [E\{\tilde{\theta} \mid D\}]^2, \tag{3.9}$$

$$V\{\tilde{\theta}^{-1} \mid D\} = \left[\frac{1}{C_0 - 2 + C}\right] [E\{\theta^{-1} \mid D\}]^2; \tag{3.10}$$

the first terms decrease with increasing C, and so, ultimately, with probability one, do the variances. This makes precise the difference between incomplete and complete samples; two different data sets could lead the to same mean forecast, but we would have more "confidence" in the result with the larger number of complete samples.

The terms in square brackets in (3.6) (3.7) are the classical maximum-likelihood estimators got from the term in square brackets in (3.2). If the experiment gives a large number of complete observations, relative to C_0, then the Bayesian and maximum-likelihood estimators coincide. However, for relatively incomplete tests, more weight is given the prior means, $E\{\tilde{\theta}\} = C_0/Q_0$, or $E\{\tilde{\theta}^{-1}\} = Q_0/(C_0 - 1)$.

Classical estimators are often obtained from Bayesian formulae when the prior knowledge becomes "diffuse"; in our model this corresponds to keeping $E\{\tilde{\theta}\}$ or $E\{\tilde{\theta}^{-1}\}$ fixed, and letting the corresponding variances (the prior uncertainty) increase without limit. From (3.9) (3.10) we see this corresponds to letting $C_0 \to 0$ or $C_0 \to 2$, respectively (with corresponding adjustments in Q_0). Thus, *with very uncertain prior knowledge*, we get:

$$E\{\tilde{\theta} \mid D\} = \left[\frac{C}{TQT(D)}\right], \tag{3.11}$$

$$(3.12) \qquad E\{\tilde{\theta}^{-1} \mid D\} = \frac{E\{\tilde{\theta}^{-1}\} + TQT(D)}{1+C}.$$

Thus, when estimating $\tilde{\theta}$, we place "full credibility" in the maximum-likelihood estimator, and ignore all prior information; but, when estimating $\tilde{\theta}^{-1}$, a Bayesian would always insist on keeping the prior mean as an initial data point, because the prior is still informative and proper in this case.

The *mixed*, or *predictive distribution* of $\tilde{x}$, averaged over all possible values of $\tilde{\theta}$, is:

$$(3.13) \qquad P^c(x \mid C_0, Q_0) = [Q_0 / (Q_0 + Q(x))]^{C_0}, \quad (x \geq 0)$$

with density

$$(3.14) \qquad p(x \mid C_0, Q_0) = (C_0 q(x) / Q_0)\,[Q_0 / (Q_0 + Q(x))]^{C_0+1},$$

a generalization of the shifted Pareto distribution. If the prototype failure function is Gumbel, we get exponential tails for large x in (3.13), while if the underlying failures are Weibull, we get the "more dangerous" algebraic tails. Posterior-to-the-data, predictive density is of the same form, but with updated parameters.

The cumulative hazard function of the mixed distribution is:

$$(3.15) \qquad R(x \mid C_0, Q_0) = -\ln P^c(x \mid C_0, Q_0) = C_0 \ln\,[1 + (Q(x) / Q_0)].$$

One can show that this mixing tends to decrease the rate of failure; in fact, the mixed population may have approximately constant or decreasing hazard rate, even with increasing $q(x)$.

Life testing applications are covered in more detail in JEWELL (1977), and the problem of model identification of the form of Q is also considered. We turn now to applications of these ideas in risk theory.

4. RENEWAL PROCESSES

In one model of the collective risk process, claims are assumed to follow a renewal process. If, during an exposure interval T, C events (accidents, claims, equipment failures, etc.) are observed, this means there are C complete interval samples $\{x_i\}$, and the final interval-in-progress, $T - \sum_{i=1}^{C} x_i$. If all intervals are sampled from (3.1) with fixed θ, the parameter updating becomes:

$$(4.1) \qquad C_0 \leftarrow C_0 + C;\; Q_0 \leftarrow Q_0 + \sum_{i=1}^{C} Q(x_i) + Q\left(T - \sum_{i=1}^{C} x_i\right).$$

Note that not only the random number of events in $(0, T]$, but also the actual lengths of the intervals provide information in the general case.

An important special case in risk processes occurs when $Q(x) = x$, leading to exponentially-distributed intervals, and a Poisson counting process, for each θ. However, here

$$(4.2) \qquad Q_0 \leftarrow Q_0 + \Sigma x_i + (T - \Sigma x_i) = Q_0 + T,$$

so we conclude that the Poisson process is special in that only the number of events in (o, T], not the epochs of events, provides predictive information!

5. INDIVIDUAL LEARNING ABOUT REMAINING LIFE

We turn now to the interesting question of whether or not a Bayesian can learn about his own remaining lifetime distribution function (rldf). For a mixed population with average tail distribution P^c,

$$(5.1) \qquad Pr\{\tilde{x} > T + u \mid \tilde{x} > T\} = \frac{P^c(T+u)}{P^c(T)} = P^c_T(u)$$

represents the fraction of those individual components alive (operating) at age T which will survive until age $T+u$.

However, for a single life component with known parameter θ, the appropriate rldf is:

$$(5.2) \qquad Pr\{\tilde{x} > T + u \mid \tilde{x} > T; \theta\} = \frac{P^c(T+u \mid \theta)}{P^c(T \mid \theta)} = P^c_T(u \mid \theta).$$

If we have to estimate this single life behavior as averaged over the population (i.e., without Bayesian learning), we get the prior expected rldf:

$$(5.3) \qquad E\{P^c_T(u \mid \tilde{\theta})\} = \int \frac{P^c(T+u \mid \theta)}{P^c(T \mid \theta)} p(\theta)d\theta,$$

which is clearly *not* identical with (5.1).

Now let us adopt the Bayesian point of view, and estimate the remaining life of a single individual who has lived to age T; since he is still alive, we have the single datum $D = \{x > T\}$, which must update the parameter density to:

$$(5.4) \qquad p(\theta \mid D) = \frac{P^c(T \mid \theta)p(\theta)}{\int P^c(T \mid \phi)p(\phi)d\phi} = \frac{P^c(T \mid \theta)p(\theta)}{P^c(T)}.$$

So the Bayesian-updated rldf will be

$$(5.5) \qquad E_{\tilde{\theta}/D}\{P^c_T(u \mid \theta)\} = \int \frac{P^c(T+u \mid \theta)\,p(\theta)}{P^c(T)} d\theta,$$

which is *exactly* the same as the population rldf in (5.1)! Stated another way, a single life (or component) cannot, on the average, gather any additional information about his remaining lifetime distribution by the mere passage of time, other than that given for the population as a whole—even though he can learn about his parameter! A surprising, but satisfying result.

6. COHORT LEARNING ABOUT REMAINING LIFE

This does not mean, however, that *several* incomplete samples cannot provide information about other lifetimes with the same θ, nor that a group of lives with the same θ cannot learn from the passage of time. Consider a cohort of N lives with the same parameter which are put "on test" at the same epoch. From Section 2, with $T_i = T$ for all i, we see that the data $D = \{x_i \leq T\ (i \in S);\ S\}$ changes (5.5) to:

$$(6.1)\quad E_{\tilde{\theta}|D}\{P^c(u \mid \tilde{\theta})\} = \int P^c(T+u \mid \theta)\ \frac{[P^c(T \mid \theta)]^{N-C-1}\ \Pi\ p(x_i \mid \theta)p(\theta)}{P^c(D)}\ d\theta,$$

where learning would clearly take place.

For the proportional-hazard family,

$$(6.2)\qquad P^c_T(u \mid \theta) = e^{-\theta[Q(T+u)-Q(T)]}.$$

If the prior at $T = 0$ is gamma with hyperparameters C_0 and Q_0, the posterior-to-data density of θ at T is gamma with hyperparameters $C_0 + C$ and $Q_0 + (N-C)Q(T) + \Sigma\, Q(x_i)$, giving finally the special cohort-experienced remaining-lifetime distribution function:

(6.3)

$$E_{\tilde{\theta}|D}\{P^c_T(u \mid \tilde{\theta})\} = \left[\frac{Q_0 + (N-C)\,Q(T) + \Sigma\, Q(x_i)}{Q_0 + (N-1-C)\,Q(T) + \Sigma\, Q(x_i) + Q(T+u)}\right]^{C_0+C}.$$

It is easy to see how learning vanishes when $N = 1$ and $C = 0$.

7. LIFE CONTINGENCIES AND RESERVES

To apply the results above, consider that we are determining the net single premium for a continuous life annuity of \$ 1/year, at force of interest δ, for an individual aged x. Given θ, this would be (we omit the usual overbar notation):

$$(7.1)\qquad a_x(\theta) = \int e^{-\delta u}\, P^c_x(u \mid \theta)\, du = \int e^{-\delta u - [Q(x+u)-Q(x)]}\, du.$$

Let us suppose that the prior on θ is gamma with hyperparameters C_1, Q_1 at the moment of underwriting (age x). The population-average annuity fair premium is then:

$$(7.2)\qquad a_x(C_1, Q_1) = Q_1^{C_1} \int e^{\delta-u}\, [Q_1 + Q_x(u)]^{-C_1}\, du,$$

where

$$(7.3)\qquad Q_x(u) = Q(x+u) - Q(x)$$

is the prototype cumulative failure function for the remaining life, beginning at age x.

Now, suppose we have insured a cohort of N lives aged x, all of whom have the same parameter, and let us follow the cohort for t additional years. During this time the data provided by the C expirations at additional ages $\{t_i\}$, together with the fact that $N-C$ lives are still in existence at age $x+t$, would update the hyperparameters to:

$$(7.4) \qquad C_2 = C_1 + C; \; Q_2(t) = Q_1 + (N-C)\, Q_x(t) + \sum_{i \in S} Q_x(t_i).$$

Although it is too late to change the premium, this additional knowledge could be useful in *adaptive modification* of the *reserves* on the $N-C$ outstanding policies; for a single-premium annuity of \$ 1/year still outstanding at age $x+t$, the correct adaptive reserve would be:

$$(7.5) \qquad {}_tV(a_x) = a_{x+t}(C_2, Q_2(t)).$$

We remind the reader that C_2 and $Q_2(t)$ will be random outcomes, depending upon actual cohort experience during ages $(x, x+t]$. Only when there is a single incomplete life under observation $(C_2 = C_1; Q_2(t) = Q_1 + Q_x(t))$ will no learning take place, and the reserves will follow the classic result for an average member of the mixed population:

$$(7.6) \qquad a_{x+t} = \int e^{-\delta u} \frac{p_x^c(t+u)}{p_x^c(t)}\, du = a_{x+t}(C_1, Q_1 + Q_x(t)).$$

A similar development could be given in terms of the net single premium for a life assurance of \$ 1, at force of interest δ, payable at the instant of death of an individual now aged x,

$$(7.7) \qquad A_x(\theta) = \int e^{-\delta u} p_x(u \mid \theta)\, du.$$

The appropriate formulae follow from the previous results by the universal relation $A_x = 1 - \delta a_x$.

It is of interest to follow through the actual stochastic behavior of a "learning reserve" of the type (7.5). First of all, we note that adaptive annuity reserves do not decrease as quickly as the corresponding a_{x+t}, for small t and $C = 0$, which can be seen from:

$$(7.8) \qquad \begin{aligned} \frac{da_{x+t}(C_2, Q_2(t))}{dt} &= \delta a_{x+t}(C_2, Q_2(t)) - 1 \\ &+ \left[\frac{C_2 q(x+t)}{Q_2(t)}\right] \{(N-C)\, a_{x+t}(C_2, Q_2(t)) \\ &- (N-C-1)\, a_{x+t}(C_2+1; Q_2(t))\} \end{aligned}$$

as compared with the well-known classical result

$$(7.9) \qquad \frac{da_{x+t}}{dt} = \delta a_{x+t} - 1 + \left[\frac{C_2 q(x+t)}{Q_1 + Q_x(t)}\right] a_{x+t}.$$

The term in square brackets is, of course, the failure rate at $x+t$ for the mixed population in the proportional-hazard family, *i.e.*, the derivative of (3.15). When the first and subsequent deaths occur, there is an instantaneous drop in (7.5), since C_2 increases by unity, but $Q_2(t)$ is continuous. In general, if fewer (more) lives than expected terminate during $(x, x+t]$, the reserves on the remaining lives are larger (smaller) than usual, since this indicates that the value of θ is smaller (larger) than average for this cohort. A complementary effect occurs for life assurance learning reserves.

It should be mentioned that a gamma-mixed proportional-hazard model should be used with care for human mortality. If, for example, the prototype failure rate is assumed to follow Makeham's law, $q(t) = A + Be^{\alpha t}$, we find that the mixed hazard rate (the derivative of (3.15)) is asymptotically constant, due to the failure-rate-decreasing properties of mixing! One would have to assume that, given θ, individuals follow a much stronger "wear-out" (say, Gumbel), in order to obtain a population Makeham-type law. It is interesting to speculate as to whether or not this occurs for closely-matched humans, where θ would have to include health, genetic, and environmental effects.

8. DEFECTIVE DISTRIBUTIONS

Component and human lives are finite, with probability one; however, defective distributions arise in a variety of other operational situations. Consider, for example, the estimation of the time it takes for a number of requests for bids, mailed survey responses, etc., to be returned. Some responses are received rather quickly; in other cases, an answer is never received.

A reasonable model for this situation would add an unknown *defect parameter*, ϕ, to the usual lifetime distribution, as follows:

$$(8.1)\qquad P^c(x \mid \theta, \phi) = \phi + (1-\phi)\, P^c(x \mid \theta);\; p(x \mid \theta, \phi) = (1-\phi)\, p(x \mid \theta).$$

ϕ is then the probability that the lifetime is "infinite".

Under the life testing scheme of Section 2, the likelihood of the date set D becomes:

$$(8.2)\qquad p(D \mid \theta, \phi) = \binom{N}{C} (1-\phi)^C \prod_{i\in S} p(x_i \mid \theta) \prod_{j\notin S} [\phi + (1-\phi)\, P^c(T_j \mid \theta)].$$

Assuming all the intervals T_j have common value T, we find the posterior-to-data density of θ and ϕ by a binomial expansion:

$$p(\theta, \phi \mid D) =$$

$$(8.3)\qquad K \sum_{j=0}^{N-C} \binom{N-C}{j} \phi^j (1-\phi)^{N-j} \prod_{i\in S} p(x_i \mid \theta)\, P^c(T \mid \theta)^{N-C-j}\, p(\theta, \phi)$$

where K is a normalizing constant to make $\int\int p(\theta, \phi \mid D)\, d\theta d\phi = 1$. To illustrate the calculations further, assume that the "honest" part of (8.1), $P^c(x \mid \theta)$, is from the proportional-hazard family (3.1), with gamma prior on θ (3.4). For simplicity, assume ϕ is, *a priori*, independent of θ, and has a beta prior density:

$$(8.4) \qquad p(\phi) = p(\phi \mid a_0, b_0) = B^{-1}(a_0, b_0)\, \phi^{a_0 - 1} (1-\phi)^{b_0 - 1} \; (0 \leq \phi \leq 1).$$

$B(a_0, b_0)$ is the beta function, $\Gamma(a_0)\,\Gamma(b_0) / \Gamma(a_0 + b_0)$. After straightforward calculations with these special forms, we find the mixed beta-gamma:

$$(8.5) \qquad p(\theta, \phi \mid D) = \sum_{j=0}^{N-C} \Pi_j(D)\, p(\phi \mid a_0 + j, b_0 + N - j)\, p(\theta \mid C_0 + C, Q_0 + TQT_j(D)),$$

where

$$(8.6) \qquad TQT_j(D) = \sum_{i \in S} Q(x_i) + (N - C - j)\, Q(T),$$

and the mixing probabilities are given by:

$$(8.7) \qquad \Pi_j(D) = K \cdot B(a_0 + j, b_0 + N - j)\, [Q_0 + TQT_j(D)]^{-(C_0 + C)},$$

where, again, K is a normalizing factor to make $\Sigma\, \Pi_j = 1$. It is important to note that, posterior-to-data, the estimates of θ and ϕ are dependent, unless all of the observations are complete. For estimating the mean defect, we have

$$(8.8) \qquad E\{\phi \mid D\} = \sum_{j=0}^{N-C} \Pi_j(D) \left[\frac{a_0 + j}{a_0 + b_0 + N}\right],$$

where we recognize the term in square brackets as the mean of ϕ, given only that we observe j defects out of N trials. For $N = 1$ and no failure:

$$(8.9) \qquad E\{\phi \mid D\} = \frac{a_0}{a_0 + b_0 + 1} \left[\frac{(a_0 + 1)\,[1 + Q(T)/Q_0]^{C_0} + b_0}{a_0\,[1 + Q(T)/Q_0]^{C_0} + b_0}\right],$$

which shows clearly how the mean defect increases from its original estimate of $a_0 / (a_0 + b_0)$ towards $(a_0 + 1) / (a_0 + b_0 + 1)$ as $T \to \infty$ with no failure. Of course, if the lifetime ever terminates, $E\{\phi \mid D\}$ jumps to $(b_0 + 1) / (a_0 + b_0 + 1)$. Other mixing models are given in JEWELL (1977).

REFERENCES

AITCHINSON, J. and J. R. DUNSMORE (1975). Statistical Prediction Analysis, Cambridge University Press, New York

JEWELL, W. S. (1974). "Exact Multidimensional Credibility", *Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker*, **74**, 193-214.

JEWELL, W S (1978). A Survey of Credibility Theory, *Scandinavian Actuarial Journal*, to appear.

JEWELL, W. S (1977). *Bayesian Life Testing Using the Total Q on Test*, chapter of Tsokos and Shimi (1977).

TSOKOS, C. P. and I. N. SHIMI (eds.) (1977). The Theory and Application of Reliability, Vol. 1, Academic Press, New York.

The ASTIN Bulletin, 10 (1979) 173-182

MODÈLES ADDITIFS ET NON ADDITIFS EN ACTUARIAT

PHILIPPE VINCKE

De nombreux modèles, en actuariat, se basent sur l'existence d'une fonction d'utilité additive Le but de cet article est de montrer que cette hypothèse enlève au problème traité son caractère dynamique et que la suppression de l'additivité conduit à une solution plus réaliste.

INTRODUCTION

Il arrive fréquemment, dans les applications, et notamment en actuariat, que la décision à prendre consiste en une suite de choix qui s'échelonnent dans le temps. Cette décision est donc représentée, non par une variable x mais par un vecteur $(x_1, x_2, \ldots, x_N)$ où chaque x_i correspond à une période différente, la politique s'étalant sur un horizon de N périodes. L'attitude généralement adoptée dans ce cas consiste à construire une fonction d'utilité additive du type

$$U(x_1, \ldots, x_N) = \sum_{i=1}^{N} U_i(x_i).$$

Néanmoins, le choix d'un tel modèle exige des conditions assez restrictives sur les préférences que cette fonction d'utilité est sensée représenter (en plus des conditions impliquées par l'existence de la fonction U).

Bien rares sont les travaux qui mettent en évidence les hypothèses qu'implique un tel modèle. Le but de cet article est de montrer, à l'aide d'un problème de la théorie du risque, que l'additivité de la fonction d'utilité enlève au problème traité son véritable caractère dynamique et que la suppression de cette hypothèse peut conduire à un modèle plus réaliste, fournissant une politique plus "raisonnable".

UTILITE ADDITIVE — INDEPENDANCE PREFERENTIELLE

Soit $\succ$ la relation représentant les préférences d'un individu dans l'ensemble des décisions $\{(x_1, x_2, \ldots, x_N)\}$. Soit U une fonction (supposée exister) telle que

$$(x_1, x_2, \ldots, x_N) \succ (x'_1, x'_2, \ldots, x'_N)$$

ssi

$$U(x_1, x_2, \ldots, x_N) > U(x'_1, x'_2, \ldots, x'_N).$$

L'hypothèse d'additivité consiste à supposer l'existence de fonctions U_i telles que

$$U(x_1, x_2, \ldots x_N) = \sum_{i=1}^{N} U_i(x_i);$$

elle implique, pour la relation $\succ$, des propriétés qui peuvent s'exprimer en termes d'indépendance préférentielle.

Un sous-ensemble E de l'ensemble $\{1, 2, \ldots, N\}$ est préférentiellement indépendant (ou indépendant au sens des préférences) dans $\{1, 2, \ldots, N\}$ ssi les préférences entre des politiques qui ne diffèrent que par les composantes correspondant à K sont indépendantes des autres composantes. Autrement dit, $E \subset \{1, 2, \ldots, N\}$ est préférentiellement indépendant dans $\{1, 2, \ldots, N\}$ ssi

(1) $$(x_1, x_2, \ldots, x_N) \succ (x'_1, \ldots x'_N)$$

et

(2) $$x_i = x'_i, \quad \forall\ i \in \{1, 2, \ldots, N\} \setminus E$$

entraînent

$$(y_1, y_2, \ldots, y_N) \succ (y'_1, \ldots y'_N)$$

où

(3) $$y_i = x_i, \ \forall\ i \in E,$$

(4) $$y'_i = x'_i, \ \forall\ i \in E,$$

(5) $$y_i = y'_i, \ \forall\ i \in \{1, 2, \ldots, N\} \setminus E$$

De manière intuitive, cela signifie que les préférences du décideur concernant le sous-ensemble E de périodes ne dépendent pas de ce qui s'est passé ou de ce qui pourra se passer au cours des autres périodes.

On peut montrer aisément que l'additivité de la fonction d'utilité nécessite l'indépendance préférentielle de tout sous-ensemble E de $\{1, 2, \ldots, N\}$. En effet:

$$(1) \Rightarrow \sum_{i=1}^{N} U_i(x_i) > \sum_{i=1}^{N} U_i(x'_i)$$

d'où, grâce à la relation (2)

$$\sum_{i \in E} U_i(x_i) > \sum_{i \in E} U_i(x'_i)$$

ou encore, par les relations (3) et (4)

$$\sum_{i \in E} U_i(y_i) > \sum_{i \in E} U_i(y'_i)$$

et, en vertu de la relation (5)

$$\sum_{i=1}^{N} U_i(y_i) > \sum_{i=1}^{N} U_i(y'_i)$$

par conséquent,

$$(y_1, y_2, \ldots, y_N) \succ (y'_1 \ldots, y'_N)$$

Sous certaines conditions concernant les espaces X_i auxquels appartiennent les x_i, l'indépendance préférentielle de tout sous-ensemble E de $\{1, 2, \ldots, N\}$ est aussi suffisante pour que la fonction d'utilité soit additive. Pour plus de précision sur ce sujet, nous renvoyons le lecteur aux travaux de DEBREU (1959), FISHBURN (1970), GORMAN (1968), KOOPMANS (1960), TING (1971), ...

L'indépendance préférentielle n'est cependant pas toujours une hypothèse très réaliste. Ainsi par exemple, en supposant qu'à chaque période le décideur ait intérêt à maximiser x_i, les fonctions U_i du modèle précédent seront croissantes. Par conséquent, parmi les deux vecteurs suivants, c'est le premier qui aura la préférence:

(1) (3, 5, 2, 2, 3, 4, 3)

(2) (1, 2, 2, 2, 3, 3, 3)

Ce résultat est naturel si ces 2 vecteurs représentent par exemple les profits réalisés par une entreprise suite à 2 investissements qui lui coûtent le même prix. Il est peut-être beaucoup moins naturel si les 2 vecteurs représentent les dividendes versés à des actionnaires à la fin de chaque année (une politique stable inspirant souvent une plus grande confiance) ou le pouvoir d'achat de la population pour 2 plans de développement différents (pour des raisons psychologiques ou politiques).

L'hypothèse de l'existence d'une fonction d'utilité additive supprime en fait le véritable caractère dynamique d'un problème puisqu'elle implique que les préférences du décideur relativement à chaque période soient indépendantes des autres périodes.

C'est ce que nous nous proposons d'illustrer ici au moyen d'un problème de la théorie du risque. Après avoir défini le problème et rappelé un modèle additif construit par FRISQUE (1974) pour le résoudre, nous présentons un modèle qui ne nécessite pas d'hypothèse d'indépendance préférentielle. La comparaison des deux modèles montre que la suppression de l'hypothèse d'additivité conduit à une politique plus stable et, à notre avis, plus réaliste.

LE PROBLÈME

Une compagnie d'assurances verse, à la fin de chaque période, des dividendes à ses actionnaires et fixe, pour la période suivante, la part de son portefeuille qu'elle engage dans un système de réassurance. Le problème consiste à déterminer une politique de "dividendes" optimale, le but de la compagnie étant de maximiser l'utilité moyenne des dividendes. Cette utilité sera sensée représenter les préférences des actionnaires. L'horizon considéré et le montant total des primes versées à la compagnie au cours de chaque période sont des données du problème. La distribution de probabilité caractérisant les sinistres à couvrir au cours de chaque période sera fixée par le modèle.

Notations

Soit

S_0 la réserve initiale,

N l'horizon considéré,

S_j la réserve à la fin de la période j ($j = 1, 2, \ldots, N$),

F_j la distribution de probabilité qui caractérise le montant total des sinistres pendant la période j ($j = 1, 2, \ldots, N$),

P_j le montant total des primes versées au cours de la période j ($j = 1, 2, \ldots, N$),

s_j le montant des dividendes versés aux actionnaires à la fin de la période j ($j = 1, 2, \ldots, N$),

k_{j-1} la partie de portefeuille gérée par la compagnie durant la période j ($j = 1, 2, \ldots, N$),

$U(s_1, \ldots, s_{1+k})$ l'utilité des dividendes $s_1 \ldots s_{1+k}$ ($k = 0, 1, \ldots, N-1$),

$U^{s_1 \ldots s_{j-1}}(s_j \ldots s_{j+k})$ l'utilité des dividendes $s_j \ldots s_{j+k}$ sachant que l'on a versé les dividendes $s_1 \ldots s_{j-1}$ au cours des périodes précédentes ($j = 2, \ldots, N$; $k = 0, 1, \ldots, N-j$).

Comme, $\forall\ k \in \{0, 1, \ldots N-1\}$:

$$U(s_1, \ldots, s_{1+k}) = U(s_1) + \sum_{l=1}^{k} U^{s_1 \ldots s_l}(s_{1+l})$$

les préférences des actionnaires sont entièrement caractérisées dès que l'on connaît $U(s_1)$ et $U^{s_1 \ldots s_{j-1}}(s_j)$, $\forall\ j \in \{2, \ldots, N\}$

Soit encore

$U^{s_1 \ldots s_{j-1}}[S_j]$ l'utilité moyenne optimale pour la période allant de la fin de la période j, avant le paiement de s_j, jusqu'à la fin de l'horizon, après le paiement de s_N, sachant que la réserve à la fin de la période j est S_j et que l'on a versé les dividendes $s_1 \ldots s_{j-1}$ au cours des périodes précédentes,

$U[S_1]$ l'utilité moyenne optimale pour la période allant de la fin de la première période, avant le paiement de s_1, jusqu'à la fin de l'horizon, après le paiement de s_N, sachant que la réserve à la fin de la première période est S_1,

$U[S_0]$ l'utilité moyenne optimale pour tout l'horizon considéré, sachant que la réserve initiale est S_0.

Le principe d'optimalité de la programmation dynamique permet d'écrire:

$$U^{s_1 \ldots s_{j-1}}[S_j] = \max \{U^{s_1 \ldots s_{j-1}}(s_j) + \int U^{s_1 \ldots s_j}[S_j - s_j + k_j(P_{j+1} - x)]\ dF_{j+1}(x)\}$$

sous les contraintes

$$\begin{cases} 0 \leqslant s_j \leqslant S_j \\ 0 \leqslant k_j \leqslant 1 \end{cases}$$

Modèle de Frisque

Dans ce modèle, la distribution des sinistres et les primes sont caractérisées, $\forall\ j$, par:

$$F_j(x) = \begin{cases} 0 & x < 0 \\ p & 0 \leqslant x < 2 \\ 1 & 2 \leqslant x \end{cases}$$

où $p > 1|2$, et $P_j = 1$.

D'autre part,

$$U(s_1) = s_1^{1/2}$$

et, $\forall\ j \in \{2, \ldots, N\}$:

$$U^{s_1 \ldots s_{j-1}}(s_j) = v^{j-1}\, s_j^{1/2}$$

Le facteur v n'a pas nécessairement un lien avec le taux d'intérêt. Il peut aussi exprimer "l'impatience" des actionnaires, c'est-à-dire "le degré qui caractérise la préférence d'un paiement récent sur un paiement ultérieur" [Borch (1968)]. Ainsi, par exemple, soit $v < 1$ et

$$U^{s_1 \ldots s_{j-1}}(s_j, s_{j+1}) = s_j + v s_{j+1}$$

Il vient

$$U^{s_1 \ldots s_{j-1}}(x, y) - U^{s_1 \ldots s_{j-1}}(y, x) = x(1-v) - y(1-v)$$

expression qui sera positive ssi $x > y$: le décideur préfère recevoir la plus grande somme en premier lieu. C'est la définition de l'impatience donnée par Koopmans (1960).

Le modèle de Frisque (1974) conduit à la solution suivante:

$$\begin{cases} s_j = \dfrac{S_j}{1 + K + K^2 + \ldots + K^{N-j}}, & j \in \{1, 2, \ldots, N\} \\ k_j = (S_j - s_j)\, \dfrac{p^2 - q^2}{p^2 + q^2}, & j \in \{1, 2, \ldots, N-1\} \\ k_0 = S_0\, \dfrac{p^2 - q^2}{p^2 + q^2} \\ U^{s_1 \ldots s_{j-1}}[S_j] = v^{j-1}\, S_j^{1/2} (1 + K + \ldots + K^{N-j})^{1/2}, & j \in \{2, \ldots N\} \\ U[S_1] = S_1^{1/2}\, (1 + K + \ldots + K^{N-1})^{1/2} \\ U[S_0] = S_0^{1/2}\, \sqrt{2}\, (p^2 + q^2)^{1/2}\, (1 + K + \ldots + K^{N-1})^{1/2} \end{cases}$$

où

$$K = 2v^2\, (p^2 + q^2)$$
$$q = 1 - p.$$

Modèle non additif

En vue de faciliter la comparaison de ce modèle avec celui de *Frisque*, nous définirons $F_j(x)$ et P_j de la même manière que ce dernier. D'autre part, supposons que

$$U(s_1) = s_1^{1/2},$$

et $\forall\ j \in \{2, \ldots, N\}$,

$$U^{s_1 \ldots s_{j-1}}(s_j) = v^{j-1}(j\, s_j - s_1 - s_2 - \ldots - s_{j-1})^{1/2}$$

La résolution du problème à l'aide de ce modèle a été conduite d'une manière analogue à celle de *Frisque*. Nous ne reproduisons pas ici les calculs, qui sont assez lourds. Le lecteur intéressé les trouvera faits en détails dans VINCKE (1977).

Ce modèle conduit à la solution suivante:

$$\left\{\begin{aligned}
& s_j = S_j \frac{j}{(1+L_j)B_j} - (s_1 + \ldots + s_{j-1}) \left[\frac{jL_{j-1}}{(1+L_j)B_j} - \frac{1}{j}\right], \quad j \in \{2, \ldots, N\} \\
& s_1 = \frac{S_1}{(1+L_1)B_1} \\
& k_j = [S_j - s_j - L_j(s_1 + \ldots + s_j)] \frac{p^2-q^2}{p^2+q^2}, \quad j \in \{1, 2, \ldots, N-1\} \\
& k_0 = S_0 \frac{p^2-q^2}{p^2+q^2} \\
& U^{s_1 \ldots s_{j-1}}[S_j] = v^{j-1} \left(\frac{B_j}{1+L_j}\right)^{1/2} [S_j - L_{j-1}[s_1 + \ldots + s_{j-1})]^{1/2}, j \in \{2, \ldots, N\} \\
& U[S_1] = \left(\frac{B_1}{1+L_1}\right)^{1/2} S_1^{1/2} \\
& U[S_0] = \sqrt{2}\,(p^2+q^2)^{1/2} \left(\frac{B_1}{1+L_1}\right)^{1/2} S_0^{1/2}
\end{aligned}\right.$$

où

$$\left\{\begin{aligned}
& L_j = L_{j+1} + \frac{1+L_{j+1}}{j+1}, \quad j \in \{1, 2, \ldots, N-1\} \\
& L_N = 0 \\
& B_j = j + 2v^2(p^2+q^2)\,(1+L_j) \frac{B_{j+1}}{1+L_{j+1}}, \quad j \in \{1, 2, \ldots, N-1\} \\
& B_N = N.
\end{aligned}\right.$$

COMPARAISON DES DEUX MODÈLES

Les tableaux qui suivent permettent de comparer les valeurs numériques auxquelles conduisent les deux modèles dans différentes situations. L'horizon choisi est de 5 périodes ($N=5$). La notation 01100 signifie qu'il n'y a pas sinistre au cours des première, quatrième et cinquième périodes et qu'il y a sinistre au cours des deuxième et troisième périodes. Les nombres qui apparaissent dans les tableaux doivent être multipliés par S_0.

TABLEAU 1 : $p=.7$; $q=.3$; $v=.2$

	01100		01111		11100		00000	
	Frisque	non additif	Frisque	non additif	Frisque	non additif	Frisque	non additif
k_0	.690	.690	.69	.69	.69	.69	.69	.69
S_1	1.690	1.690	1.69	1.69	.31	.31	1.69	1.69
s_1	1.610	.490	1.61	.49	.3	.09	1.61	.49
k_1	.050	.150	.05	.15	.009	.028	.05	.15
S_2	.030	1.050	.03	1.05	.001	.192	.13	1.35
s_2	.028	.272	.028	.272	.00095	.0488	.124	.4
k_2	.0009	.014	.0009	.014	.00003	.0032	.004	.032
S_3	.001	.76	.0011	.76	00002	.14	.01	.982
s_3	.001	.25	.001	.25	.000019	.046	.0095	.35
k_3	.00003	.009	.00003	.009	.0000006	.0018	.0003	.017
S_4	.000133	.519	.000067	.5	.0000016	.096	.00075	.65
s_4	.000127	.261	.000064	.255	.0000015	.05	.0007	.33
k_4	.000004	.0005	.000002	.00025	.000000048	.00027	.00002	.003
S_5	.00001	.25	.000001	.246	.000000148	.046	.00072	.323
s_5	.00001	.25	.000001	.246	.000000148	.046	.00072	.323

TABLEAU 2 : $p=.7$; $q=.3$; $v=.5$

	01100		01111		11100		00000	
	Frisque	non additif	Frisque	non additif	Frisque	non additif	Frisque	non additif
k_0	.69	.69	.69	.69	.69	.69	.69	.69
S_1	1.69	1.69	1.69	1.69	.31	.31	1.69	1.69
s_1	1.2	.177	1.2	.177	.22	.032	1.2	.177
k_1	.33	.79	.33	.79	.06	.145	.33	79
S	.16	.723	.16	.723	.03	.133	.82	2.3
s_2	.11	.162	.11	.162	.02	.03	.58	.48
k_2	.031	.16	.031	.16	.006	.027	.16	.8
S_3	.019	.4	.019	.4	.004	.076	.4	2.62
s_3	.014	14	.014	.14	.003	.026	.3	1.01
k_3	.0035	.016	.0035	.016	.00075	.0025	.075	.536
S_4	.0085	.276	0015	.244	.00175	.0525	.175	2.14
s_4	.0066	.14	.0011	.12	.00135	.026	.135	1 17
k_4	.0013	.01	.0002	.002	.00027	.002	.027	.33
S_5	.0032	.126	.0002	.122	.00067	.0285	.067	1.3
s_5	.0032	.126	.0002	.122	.00067	.0285	.067	1.3

TABLEAU 3: $p = .7$; $q = .3$; $v = .9$

	01100		01111		11100		00000	
	Frisque	non additif	Frisque	non additif	Frisque	non additif	Frisque	non additif
k_0	.69	.69	.69	.69	69	.69	.69	.69
S_1	1.69	1.69	1.69	1.69	.31	.31	1.69	1.69
s_1	.38	.0236	.38	.0236	.07	.0043	.38	.0236
k_1	.89	1.115	.89	1.115	.164	.2	.89	1.115
S_2	.42	.55	.42	.55	.076	.1	2.96	2.78
s_2	.115	042	.115	.042	.02	.0076	.811	.175
k_2	.21	.3	.21	.3	.038	.054	1.48	1.64
S_3	.095	.2	.095	.2	.018	.038	3.83	4.245
s_3	.033	.04	.033	.04	.0064	.0075	1.29	.624
k_3	.0422	.075	.0422	.075	.008	.014	1.61	2.2
S_4	.1	.235	.02	.085	.0196	.045	3.95	5.8
s_4	.053	.089	.01	.038	.01	.017	2	2.06
k_4	.034	.0819	.0068	.013	.0066	.0142	1.34	2.18
S_5	081	.228	.0032	.034	.016	.0422	3.29	5 92
s_5	.081	.228	.0032	.034	.016	.0422	3.29	5.92

TABLEAU 4: $p = .7$; $q = .3$; $v = 1.5$

	01100		01111		11100		00000	
	Frisque	non additif	Frisque	non additif	Frisque	non additif	Frisque	non additif
k_0	.69	.69	.69	.69	.69	.69	.69	69
S_1	1.69	1.69	1.69	1.69	.31	.31	1.69	1.69
s_1	.014	.0006	.014	.0006	.0026	.0001	.014	.0006
k_1	1.15	1.16	1.15	1.16	.21	.214	1.15	1.16
S_2	.526	.53	.526	.53	.097	.096	2.82	2.85
s_2	.005	.0026	.005	.0026	.001	.00045	.03	.0123
k_2	.357	.359	.357	.359	.066	.065	1.91	1.949
S_3	.164	.168	.164	.168	.03	.03	4.7	4.78
s_3	.02	.006	.02	.006	.0036	.001	.56	.144
k_3	.098	.1	.098	.1	.018	0196	2.82	3 146
S_4	.242	.26	.046	.062	.0444	.0486	6.96	7.782
s_4	.0726	.045	.0138	.0115	.0133	.0085	2	1.346
k_4	.116	.137	.022	.031	.02	.025	3 34	4.23
S_5	.28	.352	.01	.02	.05	.065	8.3	10.6
s_5	.28	.352	.01	.02	.05	.065	8.3	10.6

TABLEAU 5: $p = .9$; $q = .1$; $v = .5$

	01100		01111		11100		00000	
	Frisque	non additif	Frisque	non additif	Frisque	non additif	Frisque	non additif
k_0	975	.975	.975	.975	.975	.975	.975	.975
S_1	1.975	1 975	1 975	1.975	.025	.025	1.975	1.975
s_1	1.185	.15	1.185	.15	.015	.0019	1.185	.15
k_1	.77	1 48	.77	1.48	.00975	.01875	.77	1.48
S_2	.02	.345	.02	.345	.00025	.00435	1.56	3 3
s_2	.012	.082	.012	082	00015	.001	.936	.62
k_2	.0078	.034	.0078	.034	.0000975	.00036	.6	1.82
S_3	.0002	23	.0002	23	.0000025	.003	1.22	4 5
s_3	.000128	.076	.000126	.076	.0000015	.001	.77	1.43
k_3	.000072	.003	.000072	.003	.0000009	0	.44	1.9
S_4	.000146	.157	.000002	15	.0000019	.002	.89	4.97
s_4	.0001	.0733	.0000014	.07	.0000013	.00097	.63	2.54
k_4	.00004	.0025	.0000005	0	.0000005	.00001	.25	1.44
S_5	.000086	.08	.0000001	.08	.0000011	.001	.51	3.87
s_5	.000086	.08	.0000001	.08	.0000011	.001	.51	3.87

Ces tableaux permettent de voir:

— que les variations de dividendes d'une période à l'autre sont beaucoup moins fortes dans le modèle non additif que dans le modèle de *Frisque*: la suppression de l'hypothèse d'additivité (de l'indépendance préférentielle) conduit à une politique plus stable;

— que les dividendes sont plus petits dans le modèle 2 que dans le modèle 1 en début de politique et deviennent plus grands par la suite: le modèle non additif conduit à une politique plus prudente;

— que la part du portefeuille que l'on engage dans un système de réassurance est moins importante dans le modèle 2 que le modèle 1;

— que le facteur v joue un rôle prépondérant dans la façon de distribuer les dividendes: plus v est petit, plus les dividendes sont grands en début de politique;

— que les différences entre les modèles 1 et 2 sont d'autant plus grandes que v est petit: la forme de la fonction d'utilité dans le modèle non additif a un effet opposé à celui de v lorsque celui-ci est petit;

— que lorsque la probabilité de sinistre est petite (q petit), la politique de dividendes est fortement influencée par la réalisation d'un sinistre (comparaison des tableaux 2 et 5).

CONCLUSION

Comme nous l'avons dit précédemment, l'attitude généralement adoptée, lorsqu'on cherche une politique s'étalant sur plusieurs périodes, consiste à se baser sur une fonction d'utilité additive, ce qui exige des hypothèses très fortes et très restrictives. Le modèle non additif présenté ici n'a pas la prétention d'être le plus adéquat pour le problème posé mais il montre à notre avis que la suppression de l'additivité conduit à une solution plus raisonnable et plus réaliste.

De manière générale, nous pensons qu'une attention plus soutenue devrait être consacrée aux hypothèses qu'entraîne le choix d'un modèle mathématique en vue d'un problème d'actuariat, non seulement pour ce qui concerne les aspects purement actuariels du problème (distributions des sinistres, primes, ...) mais aussi pour la détermination des fonctions à maximiser, qui sont sensées représenter les préférences du décideur.

BIBLIOGRAPHIE

Borch, K. (1974). *The Mathematical Theory of Insurance*, Lexington Books.

Borch, K (1968). *The Economics of Uncertainty*. Princeton: Princeton University Press.

Debreu, G. (1959). Topological Methods in Cardinal Utility Theory, dans K. J. Arrow, S. Karlin et P. Suppes (Eds., 1960), *Mathematical Methods in the Social Sciences*, Stanford, California: Stanford University Press.

Fishburn, P. C (1970) *Utility Theory for Decision Making*, John Wiley & Sons Inc, New York.

Frisque, A. (1974). Dynamic Model of Insurance Company's Management, *Astin Bulletin*, **8**, 57-65.

Gorman, W. W. (1968). The Structure of Utility Functions, *Review of Economic Studies*, **35**, 367-390.

Koopmans, T. C. (1960). Stationary Ordinal Utility and Impatience, *Econometrica*, **28**, 287-309

Ting, M. (1971). Aggregation of attributes for multiattributed utility assessment, Technical Report 66, Operations Research Center, M.I.T.

Vincke, Ph. (1976). *Concept de quasi-ordre généralisé et théorèmes de représentation*, Thèse de doctorat, Université Libre de Bruxelles.

Vincke, Ph. (1977). Vers une généralisation des modèles de préférences utilisés en actuariat, Université Libre de Bruxelles.

The ASTIN Bulletin 10 (1979) 183-194

DYNAMIC PROGRAMMING, AN APPROACH FOR ANALYSING COMPETITION STRATEGIES

T. PENTIKÄINEN*

Stochastic-dynamic programming provides a technique for forecasting limits within which the insurance business will flow by a prefixed probability. The future development depends, among numerous other things, on management strategies, especially resources, which are planned for allocation in the acquisition of new business and for competition. This technique can be used to analyse different market situations. Various competitive measures and eventual counteractions by competitors can be assumed and simulated for the purpose. In this way the consequences of different strategies can be studied in order to find the most appropriate one. Our approach is similar to the well-known business games where teams play business in a simulated market The idea of applying dynamic programming to business games was suggested by Esa Hovinen (discussion at the Astin Colloquium in Washington in 1977).

1. STOCHASTIC-DYNAMIC PROGRAMMING

Stochastic-dynamic programming is a technique for making prognoses for the future development of the insurance business. When the initial state is known and necessary characteristics such as the volume of premiums, claim size distributions, expected number of claims, yield of interest, probable growth of

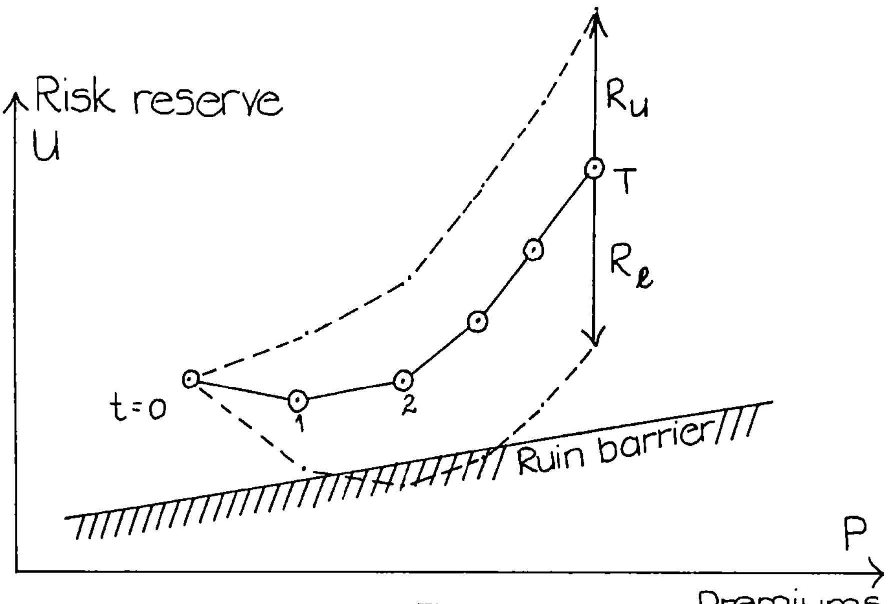

Fig. 1.

* Presented at the 14th ASTIN Colloquium, Taormina, October 1978.

the business, margins needed for operational cost, etc. are given or assumed, then it is possible step by step for future points in time $t = 1, 2, \ldots, T$ to make a prognosis for state variables such as premium volume P, risk reserve (free reserves) U, etc. Due to the stochastic character of the method, a distribution of each state variable is obtained for each time t. The probability of ruin is also obtained as a byproduct. Fig. 1 illustrates the idea. The mean value of the premium volume P and risk reserve U is calculated for $t = 1, 2, \ldots, T$. In addition, the limits of the stochastic flow of the business are estimated (upper limit R_u and lower limit R_l). When a ruin barrier is defined, the probability of ruin is obtained as a byproduct.

The flow of business and also the security limits R depend on the management strategy which the company is assumed to follow. Competition, especially, can be an important factor.

The dynamic programming approach is referred to in detail by the author in the papers listed in the bibliography.

2. BUSINESS MODELS FOR COMPETITIVE MEASURES

For model building it is necessary to know how the insurance market reacts to such competitive measures as changes in premiums, sales promotion efforts, etc. Obviously circumstances vary a great deal in different countries and even within a country, e.g. concerning the branch of insurance, perhaps concerning particular groups of insurance and clients, etc. It is well known that the degree of market saturation is one essential factor. The theory and technique for constructing market reaction models are developed for industrial and commercial practice. A good review is given by KOTLER (1975). These general approaches are clearly also applicable to insurance. Of course market reactions are mainly phenomena that can be ascertained only by collecting experience in real situations.

Two examples of the market reactions of the Finnish third party motor and motor vehicle insurance business are given in figs. 2 and 3. Company 1 reduced the premiums for third party motor insurance (fig. 2) by about 8 per cent and those for motor vehicle business (fig. 3) by about 15 per cent. The reduction was valid for one year, 1973-1974. The other companies followed suit, reducing motor vehicle (but not third party) rates to the same extent. Following this the companies again agreed on a joint level for rates. The reaction in the market share percentages can be clearly seen. Thanks to their different special groups of clients companies 3 and 5 were immune to the competitive action taken by company 1. Company 1 also carried out an advertising campaign whereas company 2 took some rationalisation measures which obviously temporarily reduced the volume of sales. Hence the changes in market shares were also due to reasons other than different rates, but this situation will not be analysed here. The reduction in third party motor rates was reflected in the market shares for motor vehicle insurance, too, even if the rates were not different.

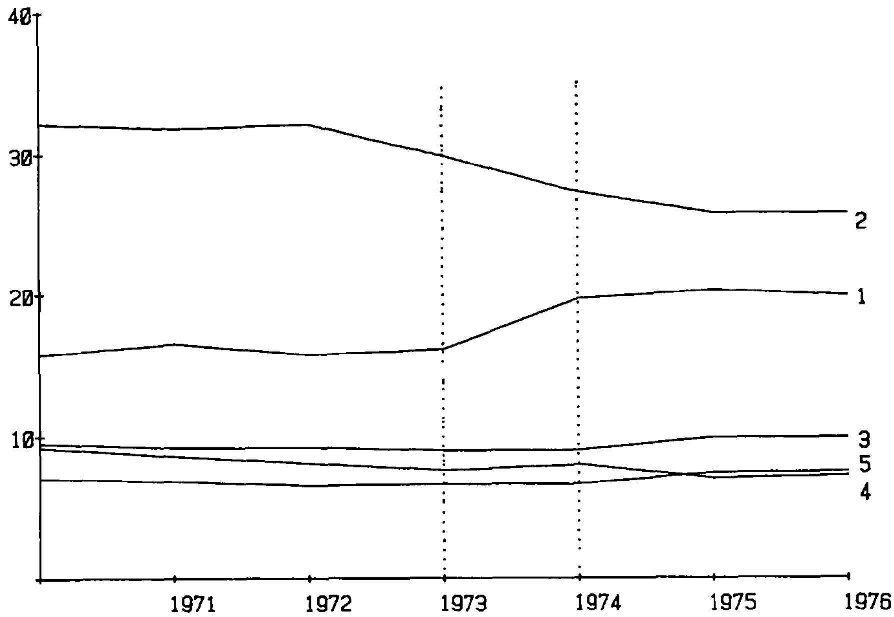

Fig. 2. Third party motor insurance Trend in market shares of the five largest companies as percentages of the whole market.

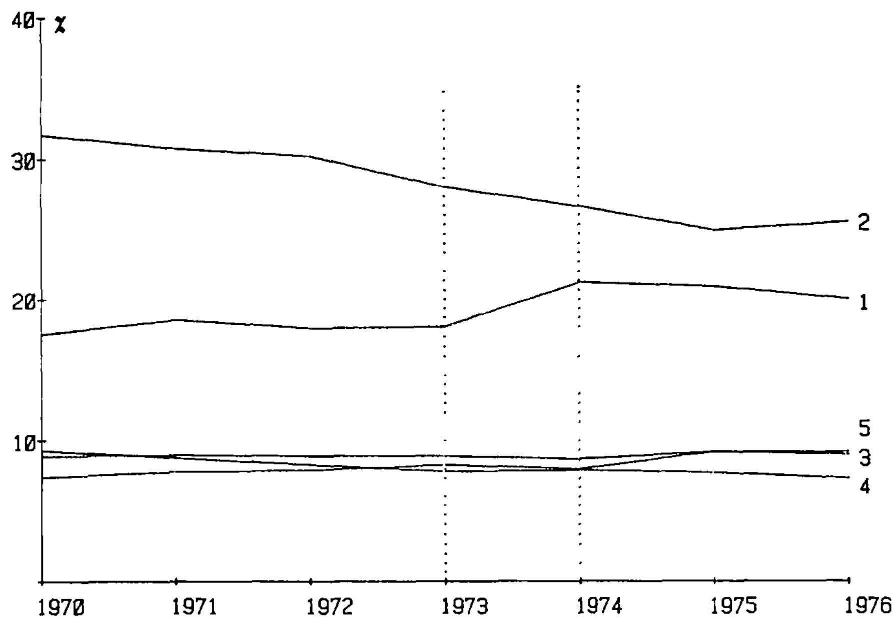

Fig. 3. Motor vehicle insurance Market shares as in fig. 2.

The return from competitive measures is described by what is called a sales response function. The return in our case is an increase in premium volume. The problems of how to find appropriate sales response functions will not be discussed here. A derivation of this kind is a standard excercise in economic theory (cf. KOTLER 1975). Our purpose is only to show how the dynamic programming technique can function if the sales response function and all other necessary initial facts are known.

To provide a simple illustration we assume that the sales response function is of the simple exponential form

$$P(t) = P(t-1) \cdot (1+g) \cdot (1-\pi(t))^{-p} \tag{1}$$

$P(t)$ is the premium volume for year t, g is the rate of natural growth of the business (level expected without competitive action) and $\pi(t)$ is the relative decrease in the premium rates, assumed to have been made in year t as a competitive action. p is the coefficient of elasticity (empirical data). The formula is a simplified version of formula (14) discussed in my paper (1978).

By partially differentiating formula (1) we obtain

$$\frac{\Delta P_{\pi}}{P} \approx p\pi \tag{2}$$

The relative sales response, i.e. the increase in premium volume due to π, is proportional to π, elasticity p being a proportionality coefficient.

In fact a reduction π in the premium rates has a double effect. On the one hand it promotes the sale of new business according to formula (2). On the other hand an amount πP is lost from the premium income (and at the same time, from the profit margin). This term πP should be subtracted from (1) to get the actual premium income. It is, however, convenient for the computation to use the unreduced premium volume P obtained from (1) and take the reduction πP into account as a loss of profit, as was done in the formulae represented in the paper mentioned above. This unreduced premium best demonstrates the effect of the competitive action. For this reason we have taken it as the variable P in the following figures. Because the competitive reductions $\pi(t)$ will be assumed to be only temporary in our examples, the final values $P(T)$ equal the actual premium incomes even if in the intermediate years the actual premium incomes deviate from P. Unreduced P also best represents the actual clientele.

From fig. 2 a value for p is got. It seems to be of the order of 2. We assume in the following $p = 1.5$.

It is obvious that the exponential sales response function is applicable only to an open market where saturation is not imminent. As a short time reaction it may also be more generally applicable, but if the premium reduction has a duration of several years, the sales function is probably more of the S form, as

Kotler claims. We omit this kind of analysis and use the simple form (1), because here we are only demonstrating the dynamic programming approach.

We have also simplified the example by assuming that the competitive reduction of premiums concerns the whole business of the company. Actually, of course, most non-life companies have many insurance branches and competition can be restricted only to part of the business. In principle the approach is also applicable to more complicated cases, but then the business must be divided into subsections, e.g. according to insurance branch. A simple example along this line was given by the author (1975).

Another simplification is the assumption that a premium reduction is the only competitive action. This is probably generally supported e.g. by an advertising campaign and other sales promotion efforts. Extension of the sales response function for this can also be found in the author's paper (1978).

3. A MULTI-UNIT COMPETITION MODEL

We are now going to deal with a market in which the leading companies are C_1, C_2 and C_3. In addition, a number of smaller companies operate in the same market. We assume that the latter have a joint tariff association and follow the same rates; hence we can "unite" them as a "fourth company", C_4, in our model.

In order to apply the model it is necessary to know, at least approximately, the initial state and a great number of parameters for each of the companies involved, in this case also as concerns competitors. In practice this may be difficult. However, at least in some countries the annual reports of the companies, the official statistics and other papers available can probably make it possible for a skilful analyst to gather numerous pieces of information and compile from them a picture on the state and resources of the competitors, at least when the analysis is continued for several years (collecting this kind of information may be a practice in many companies).

We apply the same formula (1) for all companies C_i, $i = 1, 2, 3, 4$. The premium reductions $\pi_i(t)$ which company i applies in year t are the decision variables of the model. Different competitive strategies are obtained by taking different values for these variables, i.e. the matrix $(\pi_i(t))$ where $i = 1, 2, 3, 4$ and $t = 1, 2, \ldots, T$, defines the total competitive strategy mixture.

The competitive effect can be expected to be proportional to the difference in premiums between companies, i.e. the cheaper the premiums a company i applies compared with the average level of the market, the more new business it can expect. Hence formula (1) must be amended by introducing the relative differences in the level of premiums as follows. The weighted average level of the premium reduction is

$$\bar{\pi}(t) = \frac{1}{P(t)} \sum_i P_i(t)\, \pi_i(t) \tag{3}$$

where $P(t) = \sum_i P_i(t)$ = the total volume of premiums on the market. Then the relative premium reduction for company i is

$$\pi_i'(t) = \pi_i(t) - \bar{\pi}(t) \tag{4}$$

This variable will replace π in formula (1). The loss of profit owing to the premium reduction must always be calculated on the basis of the absolute reduction π compared with the initial level $\pi = 0$. All companies have the same initial rate, i.e. $\pi_i(0) = 0$. Hence, if all companies reduce their premiums by the same relative amount $\pi_i = \bar{\pi}$, nobody will reap any benefit in the form of increased premium volume but, of course, all companies will suffer loss of profit due to reduced premiums. With some calculation formulae (2)-(4) show that generally changes in premiums $\Delta P_i(t)$ caused by any combination of variables $\pi_i(t)$ are

$$\sum_i \Delta P_i \approx 0. \tag{5}$$

This equation, where ΔP is again the change in *unreduced* premium income, is only approximately valid, because (2) is also an approximation obtained by a simple differentiation. A sales response of this type applies to saturated markets where competitive action mainly causes only an increase in market shares at the expense of the competitors.

In terms of the theory of games, we are dealing with an n-person multi-period zero-sum game in an oligopolistic market. The model can be extended to elastic markets, where a premium reduction increases the total demand for insurance. A factor $(1 - \bar{\pi}(t))^{-p'}$ must be attached to formula (1) for the purpose. This will be done in fig. 8.

Applying the formulae given above and those given in more detail by the author (1978), it is possible to compute the business flow for different mixed strategies $(\pi_i(t))$ $(i = 1, 2 \ldots; t = 1, 2, \ldots, T)$. The model can be programmed for a computer. The probability of ruin, the profits and losses and the final state of each company can be obtained as output for any strategy assumed. A good review can be obtained by arranging the main state variables, volume of premiums P and risk reserve U on a P,U-plane as in fig. 4. At the final point the number of the strategy is assigned (in fig. 4 only two strategies were applied). In our example $T = 5$ years. C_i indicates the company i. The lines (solid for company 1 and dotted for the others) from the initial point $P_i(0)$, $U_i(0)$ to the final point $P_i(T)$, $U_i(T)$ show the flow of the business as in fig. 1. A change $\pi_i(t-1) \rightarrow \pi_i(t)$ gives rise to a deviation from the normal flow $((\pi_i(t)) = 0)$ and also affects the other companies due to (5).

We are now ready to test the model by analysing the efforts and consequences of different strategies.

Strategy 1 was the "neutral" one, where no premium reductions were

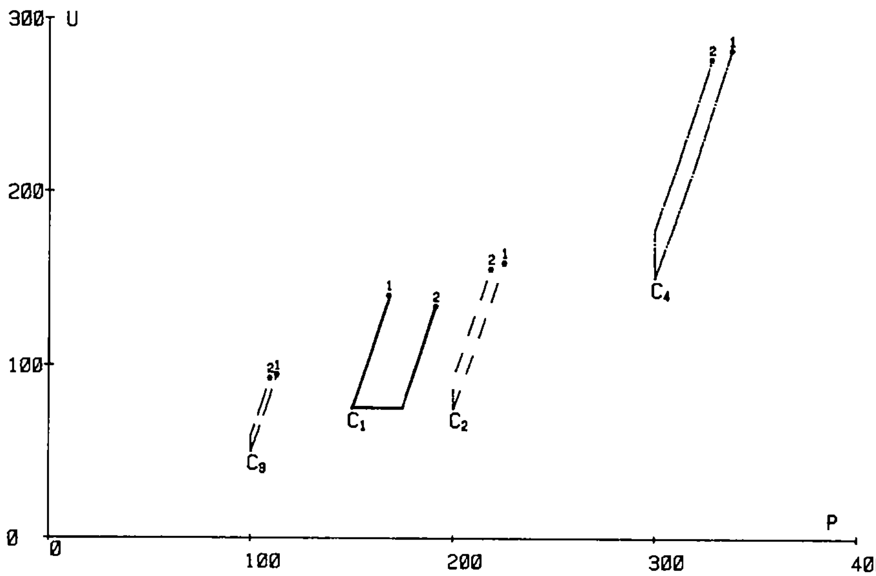

Fig. 4. Results obtained by different strategies. Units of P and U are some convenient multiple of the currency unit of the country (in our example 10^6 Fmk). Formulae and data as in example 1 in the author's paper (1978).

applied, i.e. all $\pi_i(t) = 0$. Due to the normal growth factor g in formula (1) and an assumed safety loading all companies get an increase in both premiums P and risk reserve U. Inflation can be treated separately, as we discussed in our paper (1978), hence it can be omitted in this connection, i.e. as a working hypothesis the monetary value is assumed to be constant.

Strategy 2 consists of an assumption that company C_1 reduces its premiums by 15% in one year $t = 1$ and the other companies do not react to it, i.e. their reductions are continually $= 0$. For $t > 1$ all companies again have joint rates ($\pi_i(t) = 0$). We see from fig. 4, how company 1 gains an increase in the volume of premiums whereas the competitors suffer a loss of premium incomes and in addition a small loss of profit, i.e. both $P_i(5)$ and $U_i(5)$ for $i = 2, 3, 4$ are somewhat smaller for strategy 2 than they were for strategy 1.

Deviating from the general practice in game theories we do not take maximising profit as a final objective of the company. Instead we assume here and in the following that company C_1 has an ambition to become the largest company in the market and surpass company C_2, which at the initial time point $t = 0$ is the largest. To this end the company experiments with different competitive reductions $\pi_1(1)$, which are applied for one year and then removed. The rest of the market does not take any counteractive measures (fig. 5). Because it is crucial how much the companies' resources can stand in reductions, an indicator for security, the probability of ruin, is introduced (cf. the

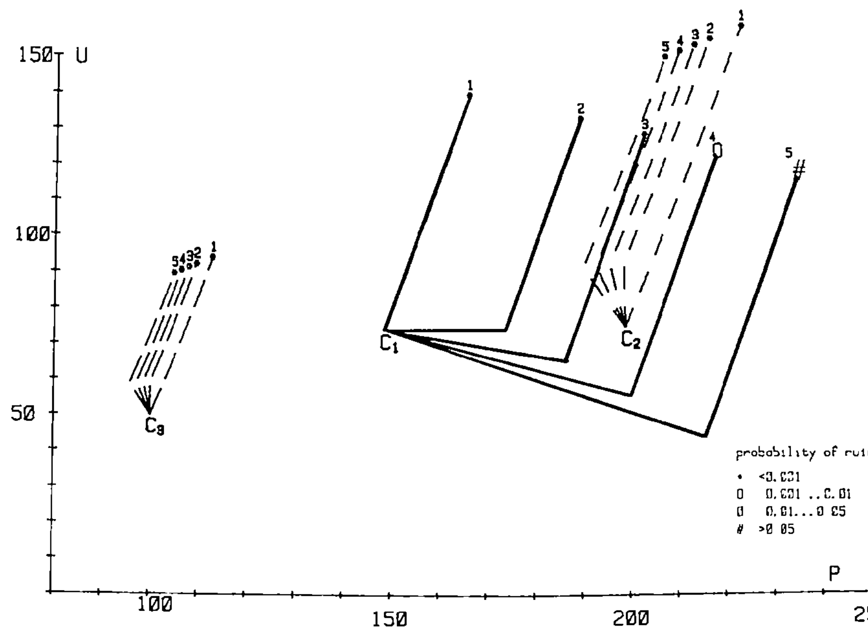

Fig. 5. Strategies of company C_1. The other companies do not take any counter-active measures.

author's paper, 1978). This is indicated by symbols in fig. 5 and in the following figures as it is shown in the right-hand corner of the picture. The reductions $\pi_1(1)$ for different strategies are as follows:

Strategy 1	$\pi = 0$
Strategy 2	$\pi = 0.1$
Strategy 3	$\pi = 0.15$
Strategy 4	$\pi = 0.20$
Strategy 5	$\pi = 0.25$

The results are given in fig. 5, where only companies C_1, C_2 and C_3 are noted. The probability of ruin for strategy 4 already begins to be alarming and for strategy 5 it is no longer acceptable. Hence it seems that strategy 3 is an acceptable choice.

Fig. 5 involves cases where the other companies do not take any counter-active measures. The analysis must be continued by studying different combinations of counteractions. That is done in fig. 6.

Strategy 1 is again neutral as in previous pictures, and strategy 2 is again the same as that in fig. 4, i.e. in the first year only company C_1 has reduction $\pi_1(1) = 0.15$ and the others have none. In strategy 3 all other companies respond to a premium reduction by making the very same reduction $\pi_i(2) =$

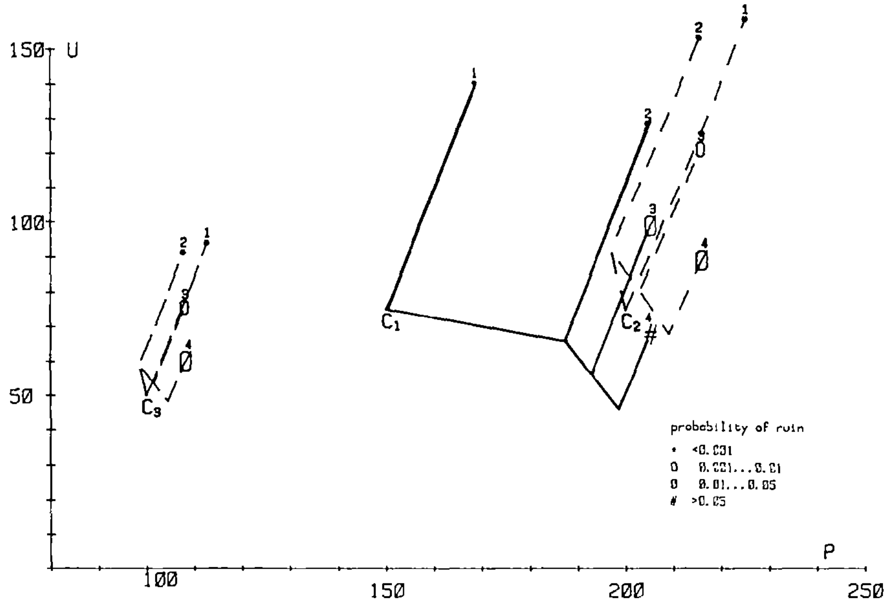

Fig. 6. Actions and counteractions.

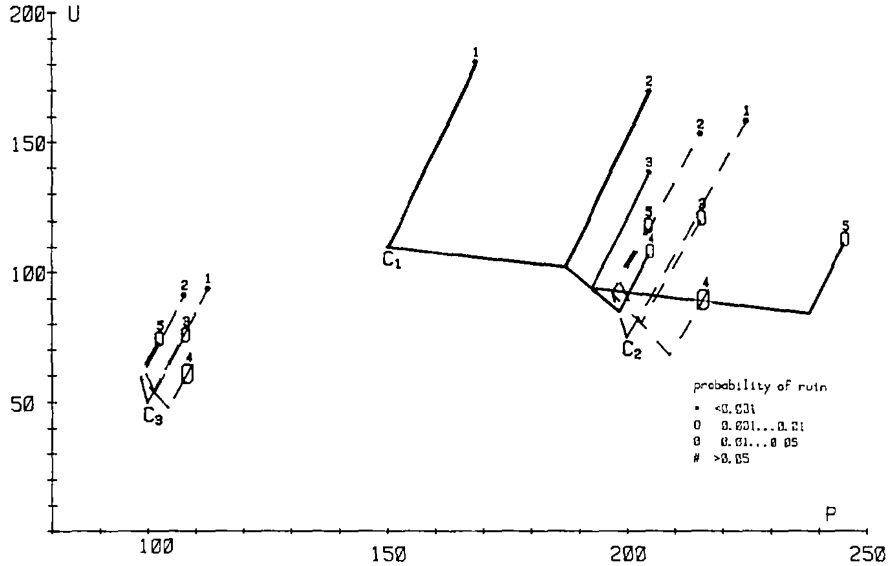

Fig. 7. Actions and counteractions; $U_1(0) = 110$.

0.15 ($i = 1, 2, 3$); hence all companies apply the same reduction in year $t = 2$. The result is, of course, a loss for all of them. It is interesting to observe that company C_1, due to losses, is already approaching a risky state, and more seriously than its competitors, as is shown by the symbols.

Strategy 4 assumes that the joint reduction will be continued for another year $t = 3$, but after that all companies will discontinue reductions. We see that the strategy puts company C_1 itself in difficulty, causing more serious losses for it than for its competitors.

We now present, as a further example, the same series of strategies but now assume that company C_1 has more initial risk reserves than it had in the preceding cases. Let $U_1(1) = 110$ million units, whereas in the preceding cases it was only 75. The very same strategies, 1-4, are now applied again (fig. 7). The better initial resources of company C_1 obviously first put a squeeze on the main competitors C_2 and C_3. If the objective of company C_1 is rootless growth, it can probably make use of its strong state (the relatively large risk reserve) for winning market shares from other companies, because these obviously cannot afford effective counteractions over a long time without losing their security. Hence we have still continued with a strategy alternative 5 where the other companies are compelled—for the sake of their increased losses—to remove their reductions for $t = 2$ whereas C_1 continues with them. Hence this strategy matrix is

$$\text{Strategy 5:} \quad (\pi_i(t)) = \begin{pmatrix} 0.15 & 0.15 & 0.15 & 0 & 0 \\ 0 & 0.15 & 0 & 0 & 0 \\ 0 & 0.15 & 0 & 0 & 0 \\ 0 & 0.15 & 0 & 0 & 0 \end{pmatrix}$$

We see how, as expected, C_1 reaches its goal, to be the largest in the market!

Finally we have experimented with a formula of elastic markets attaching another multiplicative factor $(1 - \bar{\pi})^{-p'}$ to (1). Hence an average reduction of rates $\bar{\pi}$ increases the total sum $P(t)$ of premiums by elasticity p'. We repeated the computations of fig. 4. The results are given in fig. 8.

Strategy 1 was again neutral $(\pi) = 0$, $p = 1.5$ and $p' = 0$. For strategies 2, 3 and 4 $\pi_1(1) = 0.1$ and all other $\pi_i(t) = 0$. In case 2 $p' = 0$, in case 3 $p' = 0.5$ and in case $4 = 1.0$. If $p = p' = 1.5$ then P and U of companies 2, 3 and 4 obtain approximately the same values as in case 1, i.e. the action of one company has no influence upon any other company. The market is perfectly elastic.

A further development of the situation obviously would lead us to well-known problems of the theory of n-person games in an oligopolistic market, such as possible collutions, equilibrium, etc. (cf. FRIEDMAN (1977)). Obviously the exponential sales response function (1) must also be amended and corrected according to accumulated experience if the competitive situation

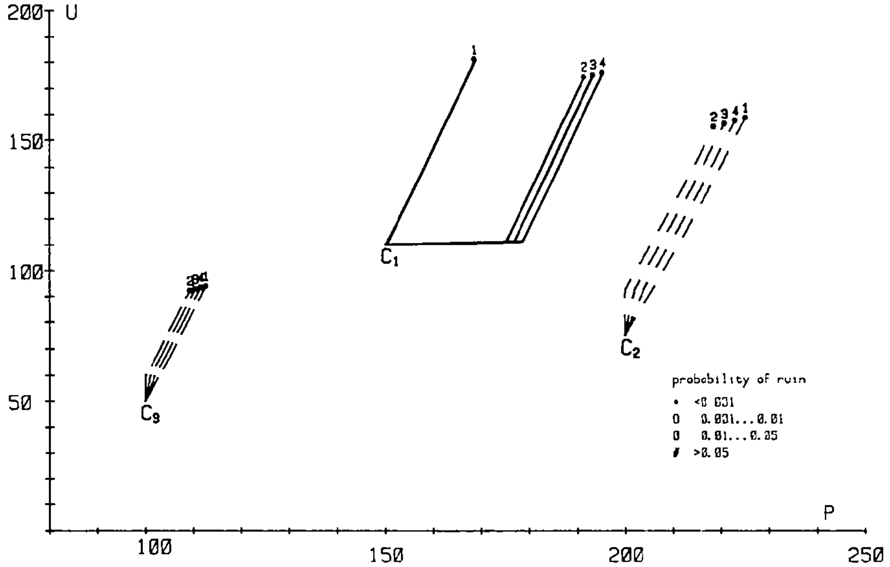

Fig. 8. Elastic market reactions.

continues for several years. Considerations like this are, however, already beyond the scope of this paper, which set out only to demonstrate how dynamic programming can be incorporated in the analysis of competitive strategies.

4. DISCUSSION

The idea outlined above can probably help in an estimation of the consequences of competitive measures and counteractions better than if this were done only using rules of thumb. One special merit of stochastic-dynamic programming is that it is able to give at least an approximation for the ruin probability, i.e. an estimation of the security.

Another merit of dynamic programming is its flexibility. Thanks to the simulation technique it is also able to operate rather complicated models without needing to narrow down the assumptions, as is often the case when other approaches are used. It is also possible to treat models providing multivariable utilities, in our example profit $(= U)$ and market share $(= P)$, whereas the conventional game theory mostly operates using only single variable utilities (profit). On the cther hand, it seems to be difficult to obtain elegant formulae for optimal strategies, equilibrium conditions, etc. as only data in tabular form or graphs can be obtained.

Probably "a play" by means of different strategies can help provide a better understanding of the structure and features of different alternatives of eventual policies. When the model is programmed for a computor the numerous alternatives can be plotted, as was illustrated in the preceding figures. The same program can also be used for playing a business game, where teams of participants are simultaneously "managing" companies C_1, C_2 ...

REFERENCES

BELLMAN, R. and R. KALABA (1965). *Dynamic programming and modern control theory.* Academic Press, New York and London.

FRIEDMAN, J. W. (1977) *Oligopoly and the theory of games.* North Holland Publishing Company, Amsterdam, New York and Oxford.

KOTLER, P. (1975). *Marketing decision making, a model building approach.* Holt, Rinehart and Winston, New York.

PENTIKÄINEN, T. (1975). A model of stochastic-dynamic prognosis. *Scandinavian Actuarial Journal,* **3** 29-53.

PENTIKÄINEN, T. (1976). Stochastic-Dynamic prognosis. *Transactions of the Congress of Actuaries,* Tokyo, 659-670.

PENTIKÄINEN, T. (1978). A solvency testing model building approach for business planning, *Scandinavian Actuarial Journal,* **6** 19-37.

The ASTIN Bulletin 10 (1979) 195-214

A NON SYMMETRICAL VALUE FOR GAMES WITHOUT TRANSFERABLE UTILITIES; APPLICATION TO REINSURANCE*

JEAN LEMAIRE

We define axiomatically a concept of value for games without transferable utilities, without introducing the usual symmetry axiom. The model—a generalization of a previous paper [6] extending Nash's bargaining problem—attempts to take into account the affinities between the players, defined by an a priori set of "distances". The general solution of all three- and four-person games is described, and various examples are discussed, like the classical "Me and my Aunt" and a reinsurance model.

Nous définissons de manière axiomatique un concept de valeur pour les jeux à utilités non-transférables, sans introduire l'axiome classique de symétrie. Le modèle — une généralisation d'un concept de valeur [6] étendant à plusieurs joueurs le problème de marchandage de Nash — tient compte des affinités entre les joueurs, données sous forme d'une matrice de "distances" a priori. Nous donnons la solution générale de tous les jeux à trois et quatre joueurs, et discutons plusieurs exemples classiques, dont le célèbre "Ma tante et moi" et le modèle de réassurance de Borch.

1. INTRODUCTION

In most of the value concepts of the cooperative theory of games [6], [10], [12], the authors have enforced a symmetry axiom: every symmetrical game has a symmetrical solution; that is, if the characteristic function of the game is symmetrical with respect to the bissecting line passing through the initial payoffs, the solution grants the same utility increase to each player. If this axiom seems innocuous (it is evident that the final payoff must not depend on a permutation, on a re-numbering of the players), it implies the implicit assumptions that the game is adequately represented by the characteristic function and that no element outside this function influences the behaviour of the participants and the results of the game. But everyday observations suggest that the players usually do not behave as one would expect from the abstract study of the game: some coalitions are formed more easily than others, two players that should coalize in order to make a profit do not unite because of personal antipathy, some persons are more likely to enter in a coalition with a given group than others, etc...; the characteristic function form of the game seems unable to forecast the coalitions that will effectively form, since it does not take into account the personal affinities between the players. For instance, the French Communist party, during the Fourth Republic consistently the largest party, never managed to enter into a government coalition, because no other party was ever willing to join it in a coalition.

* Presented at the 14th ASTIN Colloquium, Taormina, October 1978.

So the value—say the Shapley value, or any value computed on the basis of the characteristic function only—of this party is largely overestimated, since it does not consider the aversion of the other parties.

We shall in this paper develop a value concept that attempts to catch the notion of "affinities", by suppressing the symmetry axiom and introducing "distances" between players. It is a modification of our former [6] symmetrical value.

2. AXIOMS

Let $[N, v(C), \xi]$ be a game without transferable utilities (shortly a non-transferable game), where

— $N = \{1, \ldots, n\}$ is the set of the n players;
— $v(C)$ is the characteristic function, defined on all the non-void subsets C of N (the coalitions); the image of this function is a subset $v(C)$ of $E^{|C|}$, the Euclidean space of dimension $|C|$, such that $v(C)$ is non-empty, closed, convex and super-additive:

$$\forall\, C_a, C_b \subset N \ni - C_a \cap C_b = \phi,\ v(C_a \cup C_b) \supset v(C_a) \times v(C_b);$$

— ξ is the prospect space for the grand coalition N, i.e. the space delimited by the Pareto-optimal surface $v(N)$ and the hyperplanes perpendicular to the axes whose coordinates are the initial utilities of the players.

Let $[C, v(C'), \xi_C]$ be the subgame associated to the coalition C. The purpose of this paper is to define a value for such games. We shall assume that the players will sign a treaty

$$\bar{y}(N) = [y_1(N), \ldots, y_n(N)],$$

where $y_j(N)$ specifies the monetary payoff to player j. Since such a treaty usually involves side-payments (whose sum must be zero), the components of $\bar{y}(N)$ must satisfy a linear admissibility condition

$$y_1(N) + \ldots + y_n(N) = z \tag{1}$$

(the model can easily be extended to the games without side-payments. In that case the treaties have to mention the commodities owned or exchanged by each participant).

An example of a non-transferable game is the classical exchange of risks. Let the players be n insurance companies, of respective situations $[S_j, F_j(x_j)]$, where S_j is the initial surplus of company j, and $F_j(x_j)$ the distribution function of its total claim amount. Each company evaluates its situation by an utility function

$$U_j(x_j) = U_j[S_j, F_j(x_j)] = \int_0^\infty u_j(S_j - x_j)\, dF_j(x_j),$$

where $u_j(x)$ is the utility of a monetary amount x, with $u_j'(x) \geqslant 0$ and $u_j''(x) \leqslant 0$. The members of the pool will try to improve their situations by concluding a treaty of risk exchanges

$$\bar{y} = [y_1(x_1, \ldots, x_n), \ldots, y_n(x_1, \ldots, x_n)],$$

where $y_j(x_1, \ldots, x_n)$ is the amount that j has to pay if the claims for the different companies are respectively $x_1, \ldots, x_n$.

Since all the claims must be indemnified, the $y_j(x_1, \ldots, x_n)$ must satisfy the admissibility condition

$$(1') \qquad \sum_{j=1}^{n} y_j(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j = z$$

the total amount of all claims. After the signature of $\bar{y}$, the utility of j becomes

$$U_j(\bar{y}) = \int_\theta u_j[S_j - y_j(\bar{x})]\, dF(\bar{x}),$$

where θ is the positive orthant of E^n and $F(\bar{x})$ the n-dimensional distribution function of the claims $\bar{x} = (x_1, \ldots, x_n)$.

$\bar{y}$ is Pareto-optimal if there is no $\bar{y}'$ such that $U_j(\bar{y}') \geqslant U_j(\bar{y})\ \forall j$, with at least one strict inequality. Borch (see for instance [1]) has demonstrated that all the Pareto-optimal treaties are characterized by the following relations.

$$(2) \qquad k_j u_j'[S_j - y_j(\bar{x})] = k_1 u_1'[S_1 - y_1(\bar{x})] \qquad k_j \geqslant 0\ \forall j.$$

Let $K = \{k_1, \ldots, k_n\}$. The treaty is unique for given K, but there usually exists an infinity of K satisfying (1′) and (2).

It has been shown [5] that this reinsurance market is in fact a non-transferable game and that the problem of selecting an optimal set of constants k_j is identical to the determination of the value of the game. In [7] we have computed the Shapley value and the Nash-Lemaire value [6] of this game. Both values use the classical symmetry axiom. In the sequel, we shall extend axiomatically the latter value to the non-symmetrical case. We shall use four axioms.

Axiom 1: Linear invariance

The solution is not affected by a linear transformation performed on the utilities of the players.

Justification: Since utilities are only defined up to a linear transformation, it must obviously be the case for the solution.

Axiom 2: Strong Pareto-optimality

The solution depends on all the sub-treaties relative to all the sub-coalitions (with the exception of the sub-coalitions that form with probability zero—see section 4). Each sub-treaty (and the final treaty) must be Pareto-optimal and satisfy the admissibility condition.

Justification: The axiom expresses the fact that, during a negotiation, the bargaining strength of a player depends on the terms he obtained during the preceding discussions; a player will get more from his partners if he has signed a favourable treaty in a sub-coalition. We thus authorize the formation of any coalition during the bargaining process. Each one may negotiate with a disjoint group in order to unify. During this partial bargain, we suppose that each coalition acts as a single player: no one has the right to disavow his signature and quit his coalition in order to negotiate separately. We also assume that the grand coalition is formed step by step; at each step two coalitions only merge, so that N is obtained after $(n-1)$ steps [1]. Since the power of a player depends on all the already signed contracts, they must influence the final payoff. Each sub-treaty must of course be Pareto-optimal in the corresponding sub-game, and the admissibility condition must be satisfied.

Axion 3: Independence of irrelevant alternatives

During each negotiation between two coalitions, exclusion from the prospect space of possible payoffs other than the solution and the disagreement point (the utilities that the players get in case they cannot reach an agreement) does not affect the solution.

Justification: This axiom means that the solution, which by axiom 2 must lie on the upper boundary of the prospect space, only depends on the shape of this boundary in its neighbourhood, and not on distant points. This expresses a structure property of the bargaining process: during the negotiations, the set of the alternatives likely to be selected progressively reduces, so that at the end of the discussion, the solution must only compete with very close points, and not with propositions already eliminated during the prior stages of the bargaining.

Axiom 4: Partial symmetry

If, during a negotiation between two disjoint groups, the prospect space is symmetrical, so must be the treaty signed.

[1]) Those behavioural hypotheses are not very restrictive since the axiom considers *all* the grouping possibilities. For instance, we prohibit the simultaneous merging of three disjoint groups C_a, C_b, C_c But the solution will in particular study the grouping of C_a and C_b at one step and the adjunction of C_c during the next step. The two other cases (C_a and C_c unify first then absorb C_b, and C_b and C_c group and join C_a one step later) will also be considered. In the same fashion, some schemes of coalition forming where one player remains isolated until the final step, will intervene in the final treaty.

Justification: The classical symmetry axiom is weakened, since we only enforce it for the sets of two players or groups of players. It implies that the affinities between the players do not affect the discussions between two coalitions, which consist of a tough haggling between two groups trying to take as much advantage as they can from the situation. The affinities will intervene in the kind of coalitions that tend to form, in the propensity that some players have to start discussing with a particular group instead of another. In other words, the affinities influence the choice of the groups that enter negotiation, but not their negotiation itself. For example, the recent French political events demonstrate that the fact that the Communists and the Socialists have a strong affinity does not incite them to make concessions to each other: coalition forming and bargaining are two different things.

Therefore, we shall separate the computation of the value of a game in two distinct parts:

1. the coalition forming procedure, which consists of the determination of a set of probabilities $W = \{W_{C_a \dot{U} \overline{C_a}} \forall C \subset N, \forall C_a \subset C, \overline{C_a} = C \backslash C_a, C_a \neq \phi, \overline{C_a} \neq \phi\}$, interpreted as "weights associated to orders of formation of the coalitions $C = C_a \, U \, \overline{C_a}$";
2. the bargaining procedure, which attributes a payoff to each player, given the set W.

3. THE BARGAINING PROCEDURE: EXISTENCE AND UNICITY THEOREM

Let us denote $\bar{y}(C) = \bar{y}(x_i \mid i \in C)$ the treaty signed by a coalition C

and $U_i(C) = U_i[y_i(C)]$ the utility $i \in C$ derives from this signature.

Suppose that, at a given moment of the negotiation, a first group C_a of players has reached an agreement and signed a treaty $\bar{y}(C_a)$, allowing to each of its members an utility $U_i(C_a)$, while another group C_b (such that $C_a \cap C_b = \phi$) has concluded a treaty $\bar{y}(C_b)$, giving to each $j \in C_b$ an utility $U_j(C_b)$. Both groups meet in order to conclude a global treaty $\bar{y}(C_a \, \dot{U} \, C_b)$ (the symbol $\dot{U}$ has a slightly different meaning than the usual reunion sign. $C_a \, \dot{U} \, C_b$ means "C_a joins C_b". The $\cdot$ is placed to recall that the result not only depends on the set $C_a \, U \, C_b$, but also on the manner in which this coalition was formed, i.e. on C_a and C_b). If both coalitions cannot agree on a treaty $\bar{y}(C_a \, \dot{U} \, C_b)$, they necessarily return to the starting point of the negotiation, awarding to each player $U_i(C_a)$ (if $i \in C_a$) or $U_j(C_b)$ (if $j \in C_b$). For this reason, this point is called the disagreement point.

Lemma:

There exists one and only one treaty satisfying the axioms. It can be obtained by maximizing the expression

$$(3) \qquad \prod_{i \in C_a} [U_i(C_a \cup C_b) - U_i(C_a)] . \prod_{j \in C_b} [U_j(C_a \cup C_b) - U_j(C_b)],$$

providing each term of the product is non-negative.

Proof

The demonstration is a slight generalization of Nash's result [6]. Denote I the number of players of C_a $(0 < I < n)$ and L the cardinality of C_b $(0 < L \leqslant n - I)$. Number the players in such a way that the members of C_a occupy the indices 1 to I and the players of C_b the indices $I + 1$ to L. The vector

$$\bar{U}_d = [U_1(C_a), \ldots, U_I(C_a), U_{I+1}(C_b), \ldots, U_{I+L}(C_b)]$$

is the disagreement point of this negotiation. Let ψ be the maximum of (3). ψ is unique because of the convexity of $\xi_{C_a \cup C_b}$.

Suppose that ψ is distinct from $\bar{U}_d$ (otherwise the problem is trivial since the prospect space consists of a single point). We can subject all the players' utility functions to a linear transformation τ, by changing their origins so as to carry $\bar{U}_d$ to $\bar{U}_d^\tau = (0, \ldots, 0)$ and their units to carry ψ to $\psi^\tau = (1, \ldots, 1)$. Let $\xi^\tau_{C_a \cup C_b} = \tau(\xi_{C_a \cup C_b})$ be the image of $\xi_{C_a \cup C_b}$ by τ. $\xi^\tau_{C \cup C_b}$ is convex. ψ^τ is the unique point of tangency between $\xi^\tau_{C_a \cup C_b}$ and the hyperboloid whose equation is

$$\prod_{i=1}^{I+L} U_i = 1.$$

$\xi^\tau_{C_a \cup C_b}$ is even completely under the hyperplane H_1 of equation

$$\sum_{i=1}^{I+L} U_i = I + L.$$

In fact, if a point $P \in \xi^\tau_{C_a \cup C_b}$ was such that $\sum_{i=1}^{I+L} U_i > I + L$, it would be the same for any point of the segment $P\psi^\tau$ by convexity. Some of the points of this segment would be inside the hyperboloid, with thus $\prod_{i=1}^{I+L} U_i > 1$, contradicting the fact that ψ^τ maximizes $\prod_{i=1}^{I+L} U_i$.

Under H_1 we can construct a half hypersphere σ around ψ^τ with a radius sufficiently large as to include $\xi^\tau_{C_a \cup C_b}$. Consider first the game whose prospect space is limited by σ and H_1. This game is symmetrical, and ψ^τ is its solution by axioms 2 and 4. Axiom 3 allows us to withdraw all the points of $\sigma \backslash \xi^\tau_{C_a \cup C_b}$ without altering the solution. Finally through axiom 1 we can perform the inverse transformation

$$\xi_{C_a \cup C_b} = \tau^{-1} (\xi^\tau_{C_a \cup C_b})$$

and assert that $\psi = \tau^{-1} (\psi^\tau)$ is the optimal point.

Note that, as announced in the discussion of axiom 4, the negotiation between two groups of players is a "pure" bargaining, i.e. not influenced by affinities between players.

Theorem:

To each set of probabilities W can be associated one and only one treaty $\bar{y}(N)$ satisfying all the axioms. It can be obtained by the recursion.

$$
\begin{aligned}
& y_i(\{i\}) = x_i \\
& \vdots \\
& y_i(C) = \begin{cases} \sum\limits_{\substack{C_a \subset C \\ C_a \neq \phi}} W_{C_a \dot{U} \overline{C_a}}\, y_i(C_a \dot{U} \overline{C_a}) & i \in C \begin{cases} c = |C| \\ \forall C \ni 1 < c < n \\ \overline{C_a} = C \backslash C_a \end{cases} \\ 0 & i \notin C \end{cases} \\
& \vdots \\
& y_i(N) = \sum_{\substack{C_a \subset N \\ C_a \neq \phi}} W_{C_a \dot{U} \overline{C_a}}\, y_i(C_a \dot{U} \overline{C_a}) \qquad i = 1, \ldots, n.\ \overline{C_a} = N \backslash C_a,
\end{aligned}
$$

(4)

where, at each step, $\sum\limits_{C_a \subset C} W_{C_a \dot{U} \overline{C_a}} = 1$ and $W_{C_a \dot{U} C_a} > 0$, and $y_i(C_a \dot{U} \overline{C_a})$ is obtained by maximizing (3), with the disagreement point

$$
\begin{array}{ll}
U_i(C_a) & i \in C_a \\
U_j(\overline{C_a}) & j \in \overline{C_a}.
\end{array}
$$

Proof

1. Existence: It is sufficient to verify that $\bar{y}(N)$ satisfies all the axioms. This proof is straightforward.
2. Suppose that, for a given set $\{W_{C_a \dot{U} \overline{C_a}}\}$, there exist two different optimal solutions $\bar{y}(N)$ and $\bar{y}'(N)$, i.e. there exists at least an i such that $y_i(N) \neq y_i'(N)$.

We shall first show that the two solutions must differ in at least a partial treaty. In other words, it is impossible that $y_i(C_a \dot{U} \overline{C_a}) = y_i'(C_a \dot{U} \overline{C_a})$ for all $C_a \subset N$ and that $y_i(N) \neq y_i'(N)$. (4) expresses that the partial treaties $y_i(C_a \dot{U} \overline{C_a})$ are summarized by a weighted arithmetic mean. One could of course think of other parameters, like the geometric or the quadratic mean for instance, but the only parameter satisfying the admissibility condition is the weighted arithmetic mean

$$
y_i(N) = \sum_{\substack{C_a \subset N \\ C_a \neq \phi}} W^i_{C_a \dot{U} \overline{C_a}}\, y_i(C_a \dot{U} \overline{C_a}).
$$

We shall now show that the admissibility condition also implies that $W^i_{C_a \dot{U} \overline{C_a}} = W^1_{C_a \dot{U} \overline{C_a}} \forall i$. It is sufficient to prove it for $n = 3$. In this case, there are only three ways to form the grand coalition, which we shall note to simplify

$$\begin{aligned} A &= \{12\} \dot{U} \{3\} \\ B &= \{13\} \dot{U} \{2\} \\ C &= \{23\} U \{1\}. \end{aligned}$$

Thus
$$\begin{aligned} y_1(N) &= W^1_A\, y_1(A) + W^1_B\, y_1(B) + W^1_C\, y_1(C) \\ y_2(N) &= W^2_A\, y_2(A) + W^2_B\, y_2(B) + W^2_C\, y_2(C) \\ y_3(N) &= W^3_A\, y_3(A) + W^3_B\, y_3(B) + W^3_C\, y_3(C). \end{aligned}$$

(1) allows us to replace $y_1(A)$ by $z - y_2(A) - y_3(A)$, with similar relations for $y_1(B)$ and $y_1(C)$. We obtain

$$\begin{aligned} y_1(N) &= W^1_A\,[z - y_2(A) - y_3(A)] + W^1_B[z - y_2(B) - y_3(B)] + W^1_C[z - y_2(C) - \\ &\qquad - y_3(C)] \\ y_2(N) &= W^2_A\, y_2(A) + W^2_B\, y_2(B) + W^2_C\, y_2(C) \\ y_3(N) &= W^3_A\, y_3(A) + W^3_B\, y_3(B) + W^3_C\, y_3(C). \end{aligned}$$

Summing, and using (1), we get

$$\begin{aligned} z = {} & y_2(A)\,(W^2_A - W^1_A) + y_3(A)\,(W^3_A - W^1_A) + y_2(B)\,(W^2_B - W^1_B) + \\ & + y_3(B)\,(W^3_B - W^1_B) + y_2(C)\,(W^2_C - W^1_C) + y_3(C)\,(W^3_C - W^1_C) + \\ & + W^1_A z + W^1_B z + W^1_C z. \end{aligned}$$

Since the W's are the coefficients of a weighted arithmetic mean, $W^1_A + W^1_B + W^1_C = 1$, and the sum

$$\begin{aligned} & y_2(A)\,(W^2_A - W^1_A) + y_2(B)\,(W^2_B - W^1_B) + y_2(C)\,(W^2_C - W^1_C) \\ & + y_3(A)\,(W^3_A - W^1_A) + y_3(B)\,(W^3_B - W^1_B) + y_3(C)\,(W^3_C - W^1_C) \end{aligned}$$

must be identically equal to zero, $\forall\, y_2$ and y_3. Thus $W^i = W^1 \,\forall\, i$.

So there exists a coalition $C_a \subset N$ such that $y_i(C_a \dot{U} \overline{C_a}) \neq y'_i(C_a \dot{U} \overline{C_a})$. Since the solution of the maximization of (3) is unique, this result can only be explained by a difference of the disagreement points $y_i(C_a)$ and $y'_i(C_a)$. Suppose $U_i\,[y_i(C_a)] < U_i\,[y'_i(C_a)]$. There exists a player $j \in C_a$ such that $U_j\,[y_j(C_a)] > U_j\,[y'_j(C_a)]$, for otherwise $\bar{y}(C_a)$ would not be Pareto-optimal in the subgame $[C_a, v(C'_a), \xi_{C_a}]$.

The same argument can be repeated iteratively for the coalition C_a: there exists a $C_b \subset C_a$ such that $U_i\,[y_i(C_b)] < U_i\,[y'_i(C_b)]$. j must also belong to C_b (or another player j' such that $U_{j'}\,[y_{j'}, (C_b)] > U_{j'}\,[y'_{j'}, (C_b)]$), in fact, if j were a member of $C_a \backslash C_b$, $\bar{y}(C_b)$ would not be Pareto-optimal in $[C_b, v(C'_b), \xi_{C_b}]$ as $\bar{y}'(C_a/C_b)$ in $[C \backslash_a C_b, v(C'_b), \xi_{C_a \backslash C_b}]$ and axiom 2 would be violated.

So we can present a finite succession of coalitions

$$N \supset C_a \supset C_b \supset \ldots \supset C_f \supset \ldots \supset C_F$$

such that, for all $f < F$:

$$i, j \in C_f;$$
$$U_i[y_i(C_f)] < U_i[y_i'(C_f)];$$
$$U_j[y_j(C_f)] > U_j[y_j'(C_f)].$$

The last term C_F can only be the coalition formed by players i and j (otherwise we could have continued the process). There exists thus two treaties $\bar{y}(C_F)$ and $\bar{y}'(C_F)$, Pareto-optimal in $[\{ij\}, v(C), \xi_{\{ij\}}]$, i.e. such that

$$\max \{U_i[y_i(\{i,j\})] - U_i[y_i(\{i\})]\} \cdot \{U_j[y_j(\{i,j\})] - U_j[y_j(\{j\})]\}$$
$$= \max \{U_i[y_i'(\{i,j\})] - U_i[y_i'(\{i\})]\} \cdot \{U_j[y_j'(\{i,j\})] - U_j[y_j'(\{j\})]\}.$$

This contradicts the lemma, applied to the coalitions $C_a = \{i\}$ and $C_b = \{j\}$.

The solution is constructed by induction on the number of players of the coalitions: one must successively compute the value of all the two-player coalitions, then all the three-player sets, ... to end up finally with the grand coalition. The optimal treaty for a coalition C of c players is obtained by considering the set of its $2^{c-1} - 1$ (strict) sub-coalitions C_a for which there already exists a computed sub-treaty. For each C_a, one computes by (3) a treaty $\bar{y}[C_a \dot{U} (C \backslash C_a)]$. The utility granted to a player never diminishes when one or more partners are added to the coalition: (3) always provides a $U_i(C_a \dot{U} \overline{C_a})$ greater or equal than $U_i(C_a)$. The higher his disagreement point, the higher the utility awarded to a player. The procedure provides $2^{c-1} - 1$ (generally) different partial treaties, which are summed up by a weighted arithmetic mean. The fact that $W^i_{C_a \dot{U} \overline{C_a}}$ does not depend on i allows us to interpret those weights as "probabilities associated to orders of formation of the coalitions".

To sum up, the value concept takes into consideration all the possible orders of formation of the grand coalition, weighted by their respective probabilities; each player allies with other players or sets of players so that after $(n-1)$ junctions N is formed and a treaty concluded. All the grouping possibilities are considered, weighted, and account in the final solution.

For $n = 2$, the value coincides with the unweighted value [6], the Nash solution [8] and the Shapley value [12].

For $n = 3$, the value weights three different partial treaties $\bar{y}[\{12\} \dot{U} \{3\}]$, $\bar{y}[\{13\} \dot{U} \{2\}]$ and $\bar{y}[\{1\} \dot{U} \{23\}]$. Since the disagreement points are computed on the basis of coalitions of one or two persons, the partial treaties are the same as in the symmetrical value. The solution differs generally from the Shapley value.

For $n > 3$, however, the generalization is more than just "adding weights" to the partial treaties, since the disagreement points already take the affinities into account and favour the close partners.

Nothing was said up to now as far as the determination of the weights $W_{C_a \dot{U} \bar{C}_a}$ is concerned. This will be the subject of the next section.

4. FORMALIZATION OF THE AFFINITY CONCEPT: THE COALITION FORMING PROCEDURE

We suppose that the affinity between two players can be expressed by a non-negative number, d_{ij}, representing the "distance" (in a broad sense) between i and j: the larger the distance, the lesser the affinity between both players. $d_{ij} = \infty$ means that the antipathy between them is so strong that they will never join together a sub-coalition [2]). On the other hand, $d_{ij} = 0$ implies that the coalition $\{i, j\}$ will immediately form. This is a relatively uninteresting case, since it amounts to the same thing to consider $\{i, j\}$ as a single player. It is therefore not restrictive to suppose that the (symmetrical) matrix of the distances (the figures of the diagonal are irrelevant) does not contain more than one zero in each row or column (the reunion of three players in a single step is indeed not allowed, although the model could be easily adapted to this case, by introducing as a first stage the merging of the three players with probability one).

Define the "distance" between two coalitions C_a and C_b by

$$d_{C_a, C_b} = \frac{\sum_{i \in C_a} \sum_{j \in C_b} d_{ij}}{|C_a| \cdot |C_b|}.$$

The value of all the two-player coalitions can easily be computed by (3). Suppose, by induction, that we have already computed the solution for all the sets containing at most $(n-1)$ players. It only remains to calculate the value of the grand coalition.

A coalition configuration of order m (shortly a m-configuration) is a vector

$$C^m = (C_1, \ldots, C_m) \qquad \begin{array}{ll} C_a \cap C_b = \phi & a \neq b \\ \bigcup_{a=1}^{m} C_a = N & \\ C_a \neq \phi & \forall a, \end{array}$$

[2]) However, the hypotheses of the model imply that they will be forced to cooperate at the final step, since the grand coalition is bound to eventually form. This is a consequence of the fact that we required the value of a n-person game, a value that is useless if we know in advance that N will never form. But, as our theory also provides the value of all the $(n-1)$-person subgames, as well as the probabilities of formation of each subcoalition, no modification is required when one (or more) of the distances is infinite.

indicating the coalitions formed after step $(n-m)$. During a negotiation, m successively takes all the integer values, decreasing from n to 1. At the beginning, $n = m$, and $C^n = (\{1\}, \{2\}, \ldots, \{n\})$. After the final junction, $m = 1$ and $C^1 = (\{1 \ldots n\})$. For $1 < m < n$ there exists several different coalition configurations, denoted by C^m_x, C^m_y, ... Let M_m be the set of all the m-configurations. We shall denote $i \sim j$ if i and j belong to the same coalition of C^m, $i \not\sim j$ if they do not.

Each m-configuration C^m generates a number of descendants C^{m-1} obtained by joining two coalitions of C^m. Let D_1 be the set of all the descendants of C^m. Of course, two different m-configurations can produce the same descendant. Let W_{C^m} be the probability that C^m forms during the procedure, and $W_{C^{m-1} | C^m}$ the (conditional) probability that C^m generates C^{m-1}.

Naturally, this probability is zero if C^{m-1} cannot be a descendant of C^m.

We must associate to each distance matrix D a set W of probabilities $W_{C_a \dot{U} \overline{C_a}}$, defined $\forall\, C \subset N$, $\forall\, C_a \subset C \ni \overline{C_a} = C \backslash C_a$, $C_a \neq \phi$, $\overline{C_a} \neq \phi$.

$$D = \{d_{ij}\} \xrightarrow[R]{} W = \{W_{C_a \dot{U} \overline{C_a}}\}$$

Of course not any rule R that associates a set W to a matrix D is suitable for our problem. A rule will be said *coherent* if it satisfies the following conditions.

Condition 1 (Rules of probability calculus)

1.a. $W_{C^m} \geqslant 0 \qquad \forall\, C^m$

1.b. $\sum_{M_m} W_{C^m} = 1 \qquad m = 1, \ldots, n$

1.c. $\sum_{D_1} W_{C^{m-1} | C^m} = 1 \qquad \forall\, C^m$

1.d. $W_{C^{m-1}} = \sum_{M_m} W_{C^{m-1} | C^m} \cdot W_{C^m} \qquad \forall\, C^{m-1}$

Condition 2 (Relation between affinities and probabilities)

2.a. W_{C^m} is a non-increasing function of $d_{ij} \qquad \forall\, C^m \ni i \sim j$

W_{C^m} is a non-decreasing function of $d_{ij} \qquad \forall\, C^m \ni i \not\sim j$

2.b. $\lim\limits_{d_{ij} \to 0} W_{C^{n-1}} = 1 \qquad i \sim j$

2.c. $\lim\limits_{d_{ij} \to \infty} W_{C^m} = 0 \qquad \forall\, C^m, \quad i \sim j$

$\forall\, m \ni \quad 1 < m < n$

Condition 3 (Possible symmetry of two players)

3. If $d_{jl} = d_{il}$ $\forall\, l$, then $W_{C^m_x} = W_{C^m_y}$, where C^m_y is obtained from C^m_x by commuting i and j.

Condition 4 (Relations between successive configurations)

4. If $W_{C_x^m} > W_{C_y^m}$, then $W_{C_x^{m-1}} > W_{C_y^{m-1}}$ $\forall\, m$, if C_x^{m-1} is a descendant of C_x^m and if C_y^{m-1} is the descendant of C_y^m obtained through the same adjunction.

Condition 5 (Relations between configuration probabilities and weights)

5. $W_{C_a \cup \overline{C_a}} = W_{C^2}$, $\forall\, C_a$, where $C^2 = (C_a, \overline{C_a})$.

Condition 6 (Invariance with respect to a similarity)

6. W is not affected by a multiplication of the distances by a positive constant: if $d'_{ij} = kd_{ij}$ $\forall\, ij$, $W' = W$.

Note that any coherent rule determines a set W whose cardinality exceeds by far (for $n > 2$) the number of distances. It can be shown that $|D| = \frac{n(n-1)}{2} - 1$ and $|W| = \sum_{i=3}^{n} \binom{n}{i} (2^{i-1} - 2)$.

We obtain the following numbers for $3 \leqslant n \leqslant 10$.

n	Number of distances	Number of probabilities
3	2	2
4	5	14
5	9	64
6	14	244
7	20	846
8	27	2,778
9	35	8,828
10	44	27,488

There exists few coherent rules. In the sequel, we shall use the following rule

$$W_{C^{m-1} \mid C^m} = \frac{\dfrac{1}{d^2_{C_a, C_b}}}{\displaystyle\sum_{c=1}^{m} \sum_{d \neq c} \frac{1}{d^2_{C_c, C_d}}}$$

where $C^{m-1} = (C_1, \ldots, C_a \cup C_b, \ldots, C_m)$ is the descendant of $C^m = (C_1, \ldots, C_a, \ldots, C_b, \ldots, C_m)$. We thus suppose the attraction between two coalitions inversely proportional to the square of their distance.

5. RESOLUTION SCHEME OF ALL THREE-PERSON GAMES

1. Suppose three players, 1, 2 and 3, of initial utilities $U_1(\{1\})$, $U_2(\{2\})$ and $U_3(\{3\})$, and of affinities defined by the set (d_{12}, d_{13}, d_{23}). For the sake of simplicity, we shall in the sequel omit the braces, e.g. write 12 instead of $\{12\}$.

2. The maximization of the products

$$[U_1(12) - U_1(1)] \cdot [U_2(12) - U_2(2)]$$
$$[U_1(13) - U_1(1)] \cdot [U_3(13) - U_3(3)]$$
$$[U_2(23) - U_2(2)] \cdot [U_3(23) - U_3(3)]$$

provides the treaties

$$\bar{y}(12) = [y_1(12), y_2(12)]$$
$$\bar{y}(13) = [y_1(13), y_3(13)]$$
$$\bar{y}(23) = [y_2(23), y_3(23)].$$

3. Grand coalition

m	Configuration	Probability	
3	(1, 2, 3)		
2	(12, 3)	$W_{12,3} = A/d_{12}^2$	
	(13, 2)	$W_{13,2} = A/d_{13}^2$	where $A = \dfrac{1}{\dfrac{1}{d_{12}^2} + \dfrac{1}{d_{13}^2} + \dfrac{1}{d_{23}^2}}$
	(1, 23)	$W_{1,23} = A/d_{23}^2$	

m	Configuration	Probability	Treaty	Obtained by maximizing
1	(123)	$W_{12\dot{U}3} = W_{12,3}$	$\bar{y}(12\dot{U}3)$	$[U_1(123) - U_1(12)] \cdot [U_2(123) - U_2(12)] \cdot [U_3(123) - U_3(3)]$
		$W_{13\dot{U}2} = W_{13,2}$	$\bar{y}(13\dot{U}2)$	$[U_1(123) - U_1(13)] \cdot [U_2(123) - U_2(2)] \cdot [U_3(123) - U_3(13)]$
		$W_{1\dot{U}23} = W_{1,23}$	$\bar{y}(1\dot{U}23)$	$[U_1(123) - U_1(1)] \cdot [U_2(123) - U_2(23)] \cdot [U_3(123) - U_3(23)]$

Example 1. The constant-sum three-person game.

The characteristic function of this game is

$$v(\phi) = v(1) = v(2) = v(3) = 0$$

$$v(12) = v(13) = v(23) = v(123) = 1.$$

		Utilities	
1. Initial utilities	(.0,	.0,	.0)
2. 2-player coalitions			
$(1 \dot{U} 2)$	(.5,	.5,	.0)
$(1 \dot{U} 3)$	(.5,	.0,	.5)
$(2 \dot{U} 3)$	(.0,	.5,	.5)

3. Grand coalition. Distances: $d_{12} = 1$, $d_{13} = 2$, $d_{23} = 2.5$

Formation of N	Probability			
$(12\dot{U}3)$	$W_{12\dot{U}3} = W_{12,3} = .7092$	(.5,	.5,	.0)
$(13\dot{U}2)$	$W_{13\dot{U}2} = W_{13,2} = .1773$	(.5,	.0,	.5)
$(1\dot{U}23)$	$W_{1\dot{U}23} = W_{1,23} = .1135$	(.0,	.5,	.5)
	Value	(.4433,	.4113,	.1454)

We notice that 1 and 2 take a big advantage of their vicinity. Besides, the solution converges towards (.5, .5, .0) as d_{12} approaches 0. 1 becomes a little more than 2 because he is slightly nearer of 3.

Example 2. A pair of shoes.

"1 owns a left shoe. 2 and 3 are each in possession of a right shoe. The pair can be sold for 1 unit. How much is 1 entitled to?" This exemple is famous in game theory because important concepts like the core, the bargaining set, the kernel and the nucleolus completely fail to catch the threat possibilities of coalition (23) and lead to the paradoxical allotment (1,0,0). Moreover, the solution is the same if there are 999 left shoes and 1,000 right shoes: the situation becomes nearly symmetrical and the owners of right shoes still get nothing. The Shapley value, $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$, is certainly more intuitive, although it seems a bit too generous towards 1. Our unweighted value is $(\frac{4}{9}, \frac{5}{18}, \frac{5}{18})$.

The characteristic function is

$$v(\phi) = v(1) = v(2) = v(3) = v(23) = 0$$

$$v(12) = v(13) = v(123) = 1.$$

Using the same distances as in example 1, we obtain

Formation of N	Probability	Utility		
$(12\dot{U}3)$	$W_{12\dot{U}3} = .7092$	(.5,	.5,	.0)
$(13\dot{U}2)$	$W_{13\dot{U}2} = .1773$	(.5,	.0,	.5)
$(1\dot{U}23)$	$W_{1\dot{U}23} = .1135$	(.3333,	.3333,	.3333)
	Value	(.4811,	.3924,	.1265)

One notices that 2 makes the most out of his friendship with 1. The solution converges towards (.5, .5, 0) as $d_{12} \to 0$. The share of 1, always included in the interval [1/3, 1/2], diminishes when 2 and 3 feel more inclined to coalize before entering discussion with him. For the set ($d_{12} = 2$, $d_{13} = 2.5$, $d_{23} = 1$), for instance, the solution is (.3818, .3252, .2930). It tends to (1/3, 1/3, 1/3) as $d_{23} \to 0$.

Example 3. The reinsurance model.

As Gerber [3], [4] has shown that exponential utility functions possess very desirable properties for insurers, we shall suppose that

$$u_j(x) = \frac{1}{a_j}(1 - e^{-a_j x}) \qquad j = 1, \dots, n.$$

Solving equations (2), taking into account the admissibility condition (1′), leads to the solution

$$y_j(\bar{x}) = q_j z + y_j(0),$$

where

$$q_j = \frac{\frac{1}{a_j}}{\sum_{i=1}^{n} \frac{1}{a_i}}$$

and

$$y_j(0) = S_j - q_j \sum_{i=1}^{n} \left(S_i + \frac{1}{a_i} \operatorname{Log} \frac{k_i}{k_j}\right).$$

This is a familiar quota-share treaty, with quotas q_j and side-payments $y_j(0)$. As q_j does not depend on the constants k_j, the bargaining procedure will only have to determine the amount of the compensations $y_j(0)$.

Suppose that the three companies only differ by their attitude towards risk: $a_1 = .3$, $a_2 = .6$, $a_3 = .1$, while the other parameters are equal: the reserves equal to 10, and the total claim amounts are Γ-distributed, with a mean 1.2 and a variance 1.25.

The initial utilities are then

$$U_1(x_1) = 3.0778$$
$$U_2(x_2) = 1.6539$$
$$U_3(x_3) = 5.8242.$$

The treaties arising from the merging of two companies are

1. $\{1\} \dot{U} \{2\}$: Quotas $q_1 = 2/3$ Side payment $y_1(0) = -0.6778$

$q_2 = 1/3$

Utilities after reinsurance $U_1[\bar{y}(12)] = 3.1014$

$U_2[\bar{y}(12)] = 1.6560$;

2. $\{1\}\dot{U}\{3\}$: Quotas $q_1 = 1/4$ Side payment $y_1(0) = 0.7111$
$q_3 = 3/4$
Utilities after reinsurance $U_1[\bar{y}(13)] = 3.0856$
$U_3[\bar{y}(13)] = 5.8676$;

3. $\{2\}\dot{U}\{3\}$: Quotas $q_2 = .1429$ Side payment $y_2(0) = -1.2180$
$q_3 = .8571$
Utilities after reinsurance $U_2[\bar{y}(23)] = 1.6560$
$U_3[\bar{y}(23)] = 5.9599$.

Adding the third player leads to quotas $q_1 = 2/9$, $q_2 = 1/9$, $q_3 = 2/3$. 3, being the least risk averse, takes advantage of this to attract a large proportion of its partners' portfolios. As a compensation for its increased liabilities, it will naturally demand a high fixed sum. We obtain the following side payments and utilities.

	Side payments	Utilities
1. $\{12\}\dot{U}\{3\}$	$y_1(0) = .2127$	$U_1(\bar{y}) = 3.1065$
	$y_2(0) = 1.0844$	$U_2(\bar{y}) = 1.6565$
	$y_3(0) = -1.2971$	$U_3(\bar{y}) = 5.8565$
2. $\{13\}\dot{U}\{2\}$	$y_1(0) = .2882$	$U_1(\bar{y}) = 3.1013$
	$y_2(0) = 1.2576$	$U_2(\bar{y}) = 1.6554$
	$y_3(0) = -1.5458$	$U_3(\bar{y}) = 5.9583$
3. $\{1\}\dot{U}\{23\}$	$y_1(0) = .5356$	$U_1(\bar{y}) = 3.0834$
	$y_2(0) = 1.0890$	$U_2(\bar{y}) = 1.6565$
	$y_3(0) = -1.6264$	$U_3(\bar{y}) = 5.9897$.

The last company to enter the bargaining has a solid disadvantage.

With the set of distances $D_1 = (d_{12} = 1, d_{13} = 2, d_{23} = 2.5)$, the final solution is

$$y_1(0) = .2627 \qquad U_1(\bar{y}) = 3.1031$$
$$y_2(0) = 1.1156 \qquad U_2(\bar{y}) = 1.6565$$
$$y_3(0) = -1.3783 \qquad U_3(\bar{y}) = 5.8897$$

1 and 2 take advantage of their vicinity to pay as less as possible to 3. If we suppose that 1 and 3 are the closest friends, i.e. that $D_2 = (d_{12} = 2, d_{13} = 1, d_{23} = 2.5)$, the final treaty is

$$y_1(0) = .3029 \qquad U_1(\bar{y}) = 3.1003$$
$$y_2(0) = 1.2078 \qquad U_2(\bar{y}) = 1.6557$$
$$y_3(0) = -1.5107 \qquad U_3(\bar{y}) = 5.9438.$$

As the initial utilities correspond to side payments of ($y_1(0) = .6096$, $y_2(0) = 1.4659$, $y_3(0) = -1.2201$) the final solution achieves the same utility increase as a gain in capital of (.3469, .3503, .1582) for the set D_1, and of (.3067, .2581, .2906) for D_2.

6. RESOLUTION SCHEME OF ALL FOUR-PERSON GAMES

1. Treaties for all the sub-sets of two or three players: see § 5.
2. Treaty for the grand coalition. Distances (d_{12}, d_{13}, d_{13}, d_{23}, d_{24}, d_{34}).

m	Configuration	Probability	
4	(1, 2, 3, 4)		
3	(12, 3, 4)	$W_{12,3,4} = A/d^2_{12}$	
	(13, 2, 4)	$W_{13,2,4} = A/d^2_{13}$	with $A =$
	(14, 2, 3)	$W_{14,2,3} = A/d^2_{14}$	
	(1, 23, 4)	$W_{1,23,4} = A/d^2_{23}$	$\dfrac{1}{\dfrac{1}{d^2_{12}} + \dfrac{1}{d^2_{13}} + \dfrac{1}{d^2_{14}} + \dfrac{1}{d^2_{23}} + \dfrac{1}{d^2_{24}} + \dfrac{1}{d^2_{34}}}$
	(1, 24, 3)	$W_{1,24,3} = A/d^2_{24}$	
	(1, 2, 34)	$W_{1,2,34} = A/d^2_{34}$	

m	Parent	Descendant	Probability	
2	(12, 3, 4)	(123, 4)	$W_{123,4\mid 12,3,4} = B/d^2_{12,3}$	with $B =$
		(124, 3)	$W_{124,3\mid 12,3,4} = B/d^2_{12,4}$	$\left(\dfrac{1}{d^2_{12,3}} + \dfrac{1}{d^2_{12,4}} + \dfrac{1}{d^2_{34}}\right)^{-1}$
		(12, 34)	$W_{12,34\mid 12,3,4} = B/d^2_{34}$	
	(13, 2, 4)	(123, 4)	$W_{123,4\mid 13,2,4} = C/d^2_{13,2}$	with $C =$
		(134, 2)	$W_{134,2\mid 13,2,4} = C/d^2_{13,4}$	$\left(\dfrac{1}{d^2_{13,2}} + \dfrac{1}{d^2_{13,4}} + \dfrac{1}{d^2_{24}}\right)^{-1}$
		(13, 24)	$W_{13,24\mid 13,2,4} = C/d^2_{24}$	
	(14, 2, 3)	(124, 3)	$W_{124,3\mid 14,2,3} = D/d^2_{14,2}$	with $D =$
		(134, 2)	$W_{134,2\mid 14,2,3} = D/d^2_{14,3}$	$\left(\dfrac{1}{d^2_{14,2}} + \dfrac{1}{d^2_{14,3}} + \dfrac{1}{d^2_{23}}\right)^{-1}$
		(14, 23)	$W_{14,23\mid 14,2,3} = D/d^2_{23}$	
	(1, 23, 4)	(123, 4)	$W_{123,4\mid 1,23,4} = E/d^2_{12,3}$	with $E =$
		(1, 234)	$W_{1,234\mid 1,23,4} = E/d^2_{23,4}$	$\left(\dfrac{1}{d^2_{12,3}} + \dfrac{1}{d^2_{23,4}} + \dfrac{1}{d^2_{14}}\right)^{-1}$
		(14, 23)	$W_{14,23\mid 1,23,4} = E/d^2_{14}$	
	(1, 24, 3)	(124, 3)	$W_{124,3\mid 1,24,3} = F/d^2_{1,24}$	with $F =$
		(1, 234)	$W_{1,234\mid 1,24,3} = F/d^2_{24,3}$	$\left(\dfrac{1}{d^2_{1,24}} + \dfrac{1}{d^2_{24,3}} + \dfrac{1}{d^2_{13}}\right)^{-1}$
		(13, 24)	$W_{13,24\mid 1,24,3} = F/d^2_{13}$	
	(1, 2, 34)	(1, 234)	$W_{1,234\mid 1,2,34} = G/d^2_{2,34}$	with $G =$
		(134, 2)	$W_{134,2\mid 1,2,34} = G/d^2_{1,34}$	$\left(\dfrac{1}{d^2_{2,34}} + \dfrac{1}{d^2_{1,34}} + \dfrac{1}{d^2_{12}}\right)^{-1}$
		(12, 34)	$W_{12,34\mid 1,2,34} = G/d^2_{12}$	

m	Configuration	Treaty
1	$W_{123,4} = A/d^2_{12}\ B/d^2_{12,3} + A/d^2_{13}\ C/d^2_{13,2} + A/d^2_{23}\ E/d^2_{1,23} = W_{123,4}$	$\bar{y}(123\dot{U}4)$
	$W_{124,3} = A/d^2_{12}\ B/d^2_{12,4} + A/d^2_{14}\ D/d^2_{14,2} + A/d^2_{24}\ F/d^2_{1,24} = W_{124,3}$	$\bar{y}(124\dot{U}3)$
	$W_{134,2} = A/d^2_{13}\ C/d^2_{13,4} + A/d^2_{14}\ D/d^2_{14,3} + A/d^2_{34}\ G/d^2_{1,34} = W_{134,2}$	$\bar{y}(134\dot{U}2)$
	$W_{1,234} = A/d^2_{23}\ E/d^2_{23,4} + A/d^2_{24}\ F/d^2_{24,3} + A/d^2_{34}\ G/d^2_{2,34} = W_{1,234}$	$\bar{y}(1\dot{U}234)$
	$W_{12,34} = A/d^2_{12}\ B/d^2_{34} + A/d^2_{34}\ G/d^2_{12} = W_{12\dot{U}34}$	$\bar{y}(12\dot{U}34)$
	$W_{13,24} = A/d^2_{13}\ C/d^2_{24} + A/d^2_{24}\ F/d^2_{13} = W_{13\dot{U}24}$	$\bar{y}(13\dot{U}24)$
	$W_{14,23} = A/d^2_{14}\ D/d^2_{23} + A/d^2_{23}\ E/d^2_{14} = W_{14\dot{U}23}$	$\bar{y}(14\dot{U}23)$

Example 4. The homogeneous weighted majority game $(3; 2, 1, 1, 1)_h$.
This four-person game, a simplification of the game "Me and my Aunt" was studied by Owen [9] in his generalization of the Shapley value. The strongest player, 1, possesses two votes, while each of his opponents has only one. As three votes are required to win the game, the only winning coalitions are

(i) 1 and one, two or all three of his partners,
(ii) 234.

The game is however complicated by the fact that players 1 and 2 are parents; in fact, 1 is 2's aunt. Since we only want to study the influence of this relationship, we can set $d_{12} = 1$ and all the other distances equal to 2.

Coalition formation	Weight		Utility
$123\dot{U}4$	$W_{123,4} = .2527$		(.4722, .3889, .1389, .0)
$124U3$	$W_{124,3} = .2527$		(.4722, .3889, 0 , .1389)
$134\dot{U}2$	$W_{134,2} = .0774$		(.4444, .0 , 2778, .2778)
$1\dot{U}234$	$W_{1,234} = .0774$		(.0 , .3333, .3333, .3333)
$12\dot{U}34$	$W_{12,34} = .2222$		(.5 , .5 , .0 , .0)
$13\dot{U}24$	$W_{13,24} = .0588$		(.5 , .0 , .5 , .0)
$14\dot{U}23$	$W_{14,23} = .0588$		(.5 , .0 , .0 , 5)
		Value	(.4430, .3334, .1118, .1118)

The solution converges towards (.5 .5 , .0 , .0) when $d_{12} \to 0$. Owen's modified version of Shapley's value tends to (2/3, 1/3, 0, 0, 0) in this case (see discussion of § 7).

7. A FIVE-PERSON GAME

Example 5. Me and my Aunt.
This is the original game introduced by Davis and Maschler, perhaps the most celebrated game of the theory (see [2] for an interesting discussion of the game). It is in fact the homogeneous weighted majority game $(4; 3, 1, 1, 1, 1)_h$ with the addition that player 1 (my aunt) and player 2 (me) "in principle" agree to form a coalition.

The Shapley value is	(.6,	.1,	.1,	.1,	.1)
The kernel, the nucleolus and the Nash-Lemaire value agree on a	(3/7,	1/7,	1/7,	1/7,	1/7)
division proportional to the weights =	(.4286,	.1428,	.1428,	.1428,	.1428)

Most of the discussions among the game theorists in fact center on the words "in principle": the problem is phrased in an asymmetric fashion, whereas it is symmetric in terms of payoffs to coalitions. One way to capture into the model the preferences between 1 and 2 is to introduce some external feature, like our "affinities", independently of the characteristic function.

The computation of the weighted value, assuming that $d_{12} = 1$ and $d_{ij} = 2$ $\forall\ (ij) \neq (12)$ becomes rather lengthy. The solution is

$$(.4472, \quad .2849, \quad .0893, \quad .0893, \quad .0893)$$

and favour the nephew more than his aunt. The payoff vector converges towards (.5, .5, .0, .0, .0) when $d_{12} \rightarrow 0$, a division that we feel more intuitive than Owen's limiting value (.75, .25, .0, .0, .0). As a matter of fact, we think that, if 2 knows that his aunt feels compelled to agree with him and that the other players are consequently irrelevant, he should be able to "extract" $\frac{1}{2}$ from her. If the blood ties are strong enough, no other partnership is thinkable, and any threat of the aunt to negotiate with somebody else will not be credible: the asymmetry between 1 and 2 disappears and the equal division seems the only fair payoff.

Remark that the limit value does not depend on the particular choice of the rule R.

Note that the bargaining set for the configuration (12,345) grants player 1 a payoff in the interval [.50 .75] (it of course does not introduce any consanguinity in the problem). Our value thus stands at one end of this interval (the more generous towards the weaker player), Owen's generalization at the other end.

The different concepts of value attempt to be good predictors of the actual outcomes of negotiations. It is thus always interesting to compare the values with experimental data. "Me and my Aunt" has been effectively played 12 times under the direction of Selten and Schuster [11] (no preference relationship was introduced in the experiments). The game ended 8 times with a coalition between 1 and 2, with a payoff to 1 always inferior than .75. The division (.75, .25, .0, .0, .0) appeared twice during bargaining, but the stronger player was never able to protect his share and the coalition broke off. The average payoff was .4668 to 1, .1333 to the other players, a division that seems consistent with the predictions of the kernel and our unweighted value.

The facts that:

(i) the average gain of 1 was well under the figure predicted by the Shapley value;

(ii) even without affinities, 1 was never able to force a gain of .75,

naturally corroborates the idea that the Shapley value (or modified value) seems to be too generous towards the stronger players, by overlooking the threat possibilities of the weaker players.

It can besides be shown that, for $n > 2$, our value will always award more to the weaker players than Shapley's value. It is due to the fact that, if one accepts Shapley's axioms, the pivotal player becomes all of his admission value, while the axioms of § 2 have the effects by (3) of sharing this quantity between the members of the coalition according to their respective strengths.

REFERENCES

[1] Bühlmann, H. (1970). Mathematical methods in risk theory, Berlin.

[2] Davis, M. and Maschler M. (1965). The kernel of a cooperative game, *Naval Research Logistics Quarterly*, **12**, 223-259

[3] Gerber, H. (1974). On additive premium calculation principles, *Astin Bulletin*, **7**, 215-222.

[4] Gerber, H. (1974). On iterative premium calculation principles, *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, **74**, 163-172.

[5] Lemaire, J. (1973). Optimalité d'un contrat d'échange de risques entre assureurs, *Cahiers du C.E.R.O. Bruxelles*, **15**, 139-156.

[6] Lemaire, J. (1973). A new value for games without transferable utilities, *International Journal of Game Theory*, **2**, 205-213.

[7] Lemaire, J. (1977) Echange de risques entre assureurs et théorie des jeux, *Astin Bulletin*, **9**, 155-179.

[8] Nash, J. (1950). The bargaining problem, *Econometrica*, **18**, 155-162.

[9] Owen, G. (1971). Political games, *Naval Research Logistics Quarterly*, **18**, 345-355.

[10] Owen, G. (1972) Values of games without side payments, *International Journal of Game Theory*, **1**, 95-110.

[11] Selten, R. and K. Schuster (1968). *Psychological variables and coalition forming behavior*, Proc. of the conference of the IEA (Smolenice), London, 221-246.

[12] Shapley, L. S. (1953). A value for n-person games, *Annals of Maths. Studies*, **28**, 307-318, Princeton.

The ASTIN Bulletin 10 (1979) 215-222

OPTIMAL CLAIM DECISIONS FOR A BONUS-MALUS SYSTEM: A CONTINUOUS APPROACH *

NELSON DE PRIL

1. INTRODUCTION

For the premium calculation the insurer will split up his collectivity of risks into risk groups which are homogeneous with respect to some directly observable risk factors. All risks of such a risk group will be charged the same base premium. But it is clear that by such an a priori classification not all determined factors can be taken into consideration, so that there will still remain accident proneness differentials within a risk group. Since these differentials will be reflected in the course of time by the claim experience of each risk, the insurer can come to a fair tarification by adjusting, each period, the base premium according to the individual claim experience of the risk. Such a system in which earlier neglected risk factors are taken into account a posteriori is an individual experience rating system. Our main interest goes to the following side-effect of experience rating: since an unfavourable claim experience results in a premium increase, an experience rated policyholder is stimulated to self-insure small damages. This phenomenon is well know in connection with bonus-malus systems in motor-car insurance, which explains why it is called "bonus-hunger".

In the present paper a continuous time model for the bonus-malus system is set up which takes into account this hunger for bonus. An insured causing an accident will decide according to a certain decision rule whether to file a claim with his insurance company. The relevant information that he needs to make this decision is: his current risk class, the number of claims he has already filed during that period and the moment at which the decision is to be made. The decision of an insured causing at time t of period n an accident which amounts to L, can thus be thought of as being based on a decision rule of the following general form

$$L - L_n(i, k, t) \begin{cases} > 0 \text{ claim} \\ \leqslant 0 \text{ do not claim} \end{cases}$$

with $L_n(i, k, t)$ the amount that the actual accident must exceed in order to justify the filing of a claim, if the insured is at time t of period n in risk class i and has already filed k claims. The determination of the critical claim size

* An earlier version of this paper was presented at the 14th ASTIN Colloquium, Taormina, October 1978.

$L_n(i, k, t)$ should be made on economic grounds. Typical non-optimal critical claim sizes are some positive constant or the first year difference between the insurance premiums for filing and not filing a claim. The optimal value of $L_n(i, k, t)$ is clearly the one that minimizes the discounted expectation of the total future cost (premiums and self-defrayed claims) of the policyholder.

The problem of determining the optimal critical claim size was tackled in several papers under different restricting assumptions. In some of them an experience rating method is considered which avoids difficulties appearing by the general bonus-malus system; e.g. in De Leve and Weeda (1968) and Weeda (1975) a pure bonus system is considered in which the policyholder is classified according to the number of claimfree years since the last claim, so that the decision to file or not to file a claim exists only if no claim has been made during the same period. In other papers the general model in which the decision is to be taken is not satisfying; e.g. in the model of Lemaire (1976-77) a policyholder remains always insured which leads to a critical claim size that is independent of the period in which the accident takes place. Finally there are papers in which restrictions are made on the form of the decision rule itself; e.g. Martin-Löf (1973) supposes that the decision whether to file a claim has to be made at the end of the insurance period.

The most general approach to this problem was given by Haehling von Lanzenauer (1974), who considers a discrete time model and determines the optimal critical claim size by dynamic programming. However his formulation seems contradictory since he takes on the one side that the number of accidents is Poisson distributed and on the other side that in a short—but finite—time interval no more than one accident can occur. With a continuous model this problem will be avoided.

2. Description of the Model

We consider a risk group in which the accident proneness of a risk is represented by a risk parameter λ which is constant in time. We assume that the risks are independent so that we can restrict ourselves to the discussion of a single risk. We take a risk λ and assume that the number of accidents in each time interval of length t is Poisson distributed with mean λt. Further we introduce the following notations:

- $f_n(l)$ density function of a claim amount in period n.
- $F_n(l)$ the corresponding distribution function.
- w_n probability that the risk remains insured for the period n if it was insured for the period $n-1$. For the first period we have $w_1 = 1$, and if we introduce a last period N after which each risk has left the system with certainty, we have $w_{N+1} = 0$.
- δ force of interest.

We assume that the tarification in the risk group is based on a bonus-malus system that is determined in the following way.

— The length of an insurance period is 1.
— The number of classes in J.
— The class in which a risk is placed during the first period in the class s.
— The premium that a risk of class j has to pay at the beginning of period n to be insured for this period is $b_n(j)$ $(j = 1, \ldots, J;\ n = 1, \ldots, N)$.
— The transition rules are given in the form of probabilities $t_{ij}(k)$ $(i, j = 1, \ldots, J;\ k = 0, 1, \ldots)$ where $t_{ij}(k) = 1$ if a risk of class i moves to class j when k claims were filed in the past period, and $t_{ij}(k) = 0$ if such a risk goes to a class different from j. In order that the transition rules be complete and free of contradictions we must have: for each (i, k) there is one and only one j so that $t_{ij}(k) = 1$.

By convention the classes are numbered so that the highest premium corresponds to the class J. Then we have in a reasonable system that $t_{ij}(k) = t_{ij'}(k') = 1$, with $k > k'$, implies $j \geqslant j'$. By definition we call characteristic claim number of a class i the minimal number of filed claims that makes that a risk of this class will go to the highest class J for the following period. The characteristic claim numbers K_i are thus determined by

$$t_{iJ}(k) = \begin{cases} 0 & \text{for } k = 0, \ldots, K_i - 1 \\ 1 & \text{for } k = K_i, K_i + 1, \ldots \end{cases} \tag{2}$$

3. THE EXPECTATION OF THE TOTAL COST FOR THE POLICYHOLDER

We consider a risk λ who decides whether to file a claim according to a given decision rule of the form (1), where $L_n(i, k, t)$ is continuous in $0 \leqslant t < 1$. Let $A_n(i, k, t)$ represent the discounted expectation of all future cost (premiums and self-defrayed claims) if the risk is currently at time t of period n, belongs to risk class i, and has already filed k claims that period. According to the assumption that the number of claims is Poisson distributed, we have in a (short) time interval of length Δt that

$$\begin{aligned} A_n(i,k,t) &= (1 - \lambda\Delta t)\, e^{-\delta\Delta t} A_n(i,k,t+\Delta t) \\ &+ \lambda\Delta t\, F_n[L_n(i,k,t+\S)]\, \Big\{ e^{-\delta\S} \int_0^{L_n(i,k,t+\S)} l f_n[l \mid l \leqslant L_n(i,k,t+\S)]\, dl \\ &+ e^{-\delta\Delta t} A_n(i,k,t+\Delta t) \Big\} + \lambda\Delta t \{1 - F_n[L_n(i,k,t+\S')]\}\, e^{-\delta\Delta t} A_n(i,k+1,t+\Delta t) \\ &+ o(\Delta t) \end{aligned}$$

where $0 < \S, \S' < \Delta t$. Hereby $o(\Delta t)$ denotes a function $f(\Delta t)$ for which $\lim_{\Delta t \to 0} \dfrac{f(\Delta t)}{\Delta t} = 0$.

Dividing by Δt, we have

$$\frac{A_n(i,k,t+\Delta t) - A_n(i,k,t)}{\Delta t} = \frac{1}{\Delta t}[\delta\Delta t + o(\Delta t)]\, A_n(i,k,t+\Delta t)$$

$$+ \lambda\{1 - F_n[L_n(i,k,t+\S)]\}\, e^{-\delta\Delta t} A_n(i,k,t+\Delta t)$$

$$- \lambda\{1 - F_n[L_n(i,k,t+\S')]\}\, e^{-\delta\Delta t} A_n(i,k+1,t+\Delta t)$$

$$- \lambda e^{-\delta\S} \int_0^{L_n(i,k,t+\S)} l f_n(l)\, dl - \frac{o(\Delta t)}{\Delta t}$$

and by letting $\Delta t \to 0$, we obtain

(3)
$$\frac{dA_n(i,k,t)}{dt} = \delta A_n(i,k,t) + \lambda\{1 - F_n[L_n(i,k,t)]\}[A_n(i,k,t) - A_n(i,k+1,t)] - \lambda \int_0^{L_n(i,k,t)} l f_n(l)\, dl$$

Boundary conditions are found by considering the left-hand limit of $A_n(i, k, t)$ at $t = 1$.

(4a) $$A_n(i,k,1) = w_{n+1} \sum_{j=1}^{J} [b_{n+1}(j) + A_{n+1}(j,0,0)]\, t_{ij}(k) \text{ if } n = 1,\ldots,N-1$$

(4b) $$A_N(i,k,1) = 0$$

In (4a) we have taken into account the premium for the period $n+1$ of the unique class j determined by the class i and the number of claims k filed during period n. By means of the equations (3) and (4) every $A_n(i,k,t)$ can be determined recursively starting with $A_N(i,k,1) = 0$ for each (i,k). The recurrence differential equation (3) determines the evolution of $A_n(i,k,t)$ through period n and the formula (4a) gives the relation between the $A_n(i,k,t)$ for consecutive periods.

4. THE OPTIMAL CRITICAL CLAIM SIZE

A risk causing at time t of period n an accident which amounts to L has the disposal of two strategies. When he does not file a claim the present value at the moment of the accident of the expectation of his total cost is $L + A_n(i,k,t)$, where i is his current risk class and k is the number of claims he has already filed that period. When he files a claim the expectation of his total cost is $A_n(i,k+1,t)$. By definition the risk will make an optimal claim decision if the expectation of his total cost is minimized as a result of his decision making. The optimal critical claim size is thus

(5) $$L_n(i,k,t) = A_n(i,k+1,t) - A_n(i,k,t)$$

According to (3), derivation gives

$$\frac{dL_n(i,k,t)}{dt} = \delta L_n(i,k,t) + \lambda\{1 - F_n[L_n(i,k,t)]\} L_n(i,k,t)$$

$$- \lambda\{1 - F_n[L_n(i,k+1,t)]\} L_n(i,k+1,t) + \lambda \int_{L_n(i,k+1,t)}^{L_n(i,k,t)} l f_n(l)\, dl$$

and after partial integration we have

$$(6) \qquad \frac{dL_n(i,k,t)}{dt} = \delta L_n(i,k,t) + \lambda \int_{L_n(i,k+1,t)}^{L_n(i,k,t)} [1 - F_n(l)]\, dl$$

This recurrence differential equation determines the evolution of $L_n(i, k, t)$ through period n. The boundary conditions are obtained by taking the left-hand limit of $L_n(i, k, t)$ at $t = 1$. Hereby we can distinguish the following three cases.

a) $k = K_i, K_i + 1, \ldots$

According to the definition of characteristic claim number it follows from (4a) and (5) that

$$L_n(i,k,1) = w_{n+1} \sum_{j=1}^{J} [b_{n+1}(j) + A_{n+1}(j,0,0)]\,[t_{ij}(k+1) - t_{ij}(k)]$$

$$= 0 \qquad \text{for } k = K_i, K_i + 1, \ldots$$

and the solution of equation (6) reduces to

$$(7) \qquad L_n(i,k,t) = 0 \qquad \text{for } k = K_i, K_i + 1, \ldots$$

b) $n = N$

Using (4b) and (5) we find that

$$L_N(i, k, 1) = A_N(i, k+1, 1) - A_N(i, k, 1) = 0$$

so that the solution of (6) is

$$(8) \qquad L_N(i, k, t) = 0$$

The results (7) and (8) are intuitively appealing.

c) $k = 0, \ldots, K_i - 1$ *and* $n = 1, \ldots, N - 1$

According to (4a) a repeated use of (5) yields

$$L_n(i,k,1) = w_{n+1} \sum_{j=1}^{J} [b_{n+1}(j) + A_{n+1}(j,0,0)]\,[t_{ij}(k+1) - t_{ij}(k)]$$

$$= \ldots$$

$$= w_{n+1} \sum_{j=1}^{J} [b_{n+1}(j) + A_{n+1}(j,K_j,0) - L_{n+1}(j,K_j - 1,0)$$

$$- \ldots - L_{n+1}(j,0,0)]\,[t_{ij}(k+1) - t_{ij}(k)]$$

From (3) and (7) it follows that

$$\frac{dA_{n+1}(j, K_j, t)}{dt} = \delta A_{n+1}(j, K_j, t)$$

so that

$$A_{n+1}(j, K_j, t) = e^{-\delta(1-t)} A_{n+1}(j, K_j, 1)$$

and thus

$$\begin{aligned} A_{n+1}(j, K_j, 0) &= e^{-\delta} A_{n+1}(j, K_j, 1) \\ &= e^{-\delta} w_{n+2} \sum_{j=1}^{J} [b_{n+2}(j) + A_{n+2}(j, 0, 0)] \, t_{ij}(K_j) \\ &= e^{-\delta} w_{n+2} [b_{n+2}(J) + A_{n+2}(J, 0, 0)] \end{aligned}$$

This shows that $A_{n+1}(j, K_j, 0)$ is independent of j, so that we have for $k = 0, \ldots, K_i - 1$ and $n = 1, \ldots, N - 1$

(9) $$L_n(i, k, 1) = w_{n+1} \sum_{j=1}^{J} [b_{n+1}(j) - L_{n+1}(j, K_j - 1, 0) - \ldots - L_{n+1}(j, 0, 0)] \, [t_{ij}(k+1) - t_{ij}(k)]$$

In particular for $n = N - 1$ this formula reduces to

$$L_{N-1}(i, k, 1) = w_N \sum_{j=1}^{J} b_N(j) \, [t_{ij}(k+1) - t_{ij}(k)]$$

5. THE OPTIMAL CRITICAL CLAIM SIZE IN THE CASE OF EXPONENTIALLY DISTRIBUTED CLAIM AMOUNTS

We assume that the claim amounts are distributed according to: $F_n(l) = 1 - e^{-c_n l}$. Then equation (6) becomes

(10) $$\frac{dL_n(i, k, t)}{dt} = \delta L_n(i, k, t) + \frac{\lambda}{c_n} [e^{-c_n L_n(i, k+1, t)} - e^{-c_n L_n(i, k, t)}]$$

with given initial values $L_n(i, k, 1)$ for $k = 0, \ldots, K_i - 1$, and where $L_n(i, k, t) = 0$ for $k = K_i, K_i + 1, \ldots$.

We make the substitution

(11) $$L_n(i, k, t) = \frac{1}{c_n} \ln \frac{\phi_n(i, k, t)}{\phi_n(i, k+1, t)}$$

where we put $\phi_n(i, k, t) = 1$ for $k = K_i, K_i + 1, \ldots$

Substitution in equation (10) leads to

$$\frac{1}{\phi_n(i,k,t)} \frac{d\phi_n(i,k,t)}{dt} + \lambda \frac{\phi_n(i,k+1,t)}{\phi_n(i,k,t)} - \delta ln\phi_n(i,k,t)$$

$$= \frac{1}{\phi_n(i,k+1,t)} \frac{d\phi_n(i,k+1,t)}{dt} + \lambda \frac{\phi_n(i,k+2,t)}{\phi_n(i,k+1,t)} - \delta ln\phi_n(i,k+1,t)$$

$$= \dots$$

$$= \lambda$$

or

$$(12) \qquad \frac{d\phi_n(i,k,t)}{dt} = \lambda\phi_n(i,k,t) - \lambda\phi_n(i,k+1,t) + \delta\phi_n(i,k,t)\, ln\phi_n(i,k,t)$$

For given t we can compute the solutions $\phi_n(i,k,t)$, $k = 0, \dots, K_i - 1$, of (12) by successive approximations. We replace (12) by

$$(13) \qquad \begin{cases} \dfrac{d\phi_n^{(0)}(i,k,t)}{dt} = \lambda\phi_n^{(0)}(i,k,t) - \lambda\phi_n^{(0)}(i,k+1,t) \\[2ex] \dfrac{d\phi_n^{(\nu)}(i,k,t)}{dt} = [\lambda + \delta ln\phi_n^{(\nu-1)}(i,k,t)]\, \phi_n^{(\nu)}(i,k,t) - \lambda\phi_n^{(\nu)}(i,k+1,t) \\[1ex] \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \nu = 1, 2, \dots \end{cases}$$

These equations are of the form

$$(14) \qquad \frac{d\phi_n(i,k,t)}{dt} = (\lambda + a)\,\phi_n(i,k,t) - \lambda\phi_n(i,k+1,t)$$

and have as solution

$$(15) \qquad \phi_n(i,k,t) = \left(\frac{\lambda}{\lambda+a}\right)^{K_i-k} + e^{-(\lambda+a)(1-t)} \sum_{l=k}^{K_i-1} \frac{[\lambda(1-t)]^{l-k}}{(l-k)!} \left[\phi_n(i,l,1) - \left(\frac{\lambda}{\lambda+a}\right)^{K_i-l}\right]$$

where

$$(16) \qquad \begin{aligned} \phi_n(i,l,1) &= e^{c_n L_n(i,l,1)}\, \phi_n(i,l+1,1) \\ &= \dots \\ &= \exp\left[c_n \sum_{k=l}^{K_i-1} L_n(i,k,1)\right] \end{aligned}$$

The formulae (11), (15) and (16) determine the solution of equation (10). Together with (9) these formulae permit a policyholder to calculate his optimal critical claim size at each moment.

REFERENCES

DE LEVE, G. and P. J. WEEDA (1968) Driving with Markov-programming, *Astin Bulletin*, **5**, 62-86.

GRENANDER, U. (1957). Some Remarks on Bonus Systems in Automobile Insurance, *Skand. Aktuar.-Tidskr*, **40**, 180-197.

HAEHLING VON LANZENAUER, C. (1974). Optimal Claim Decisions by Policyholders in Automobile Insurance with Merit-Rating Structures, *Operations Research*, **22**, 979-990. Reprinted in Actuarial Research Clearing House 1974.2.

LEMAIRE, J. (1976). Driver versus Company: Optimal Behaviour of the Policyholder, *Scand. Actuarial J.*, **4**, 209-219.

LEMAIRE, J. (1977). La Soif du Bonus, *Astin Bulletin*, **9**, 181-190.

MARTIN-LÖF, A. (1973). A Method for Finding the Optimal Decision Rule for a Policy Holder of an Insurance with a Bonus System, *Skand. Aktuar-Tidskr.*, 23-29.

NORBERG, R (1975). Credibility Premium Plans which Make Allowance for Bonus Hunger, *Scand. Actuarial J.*, **3**, 73-86.

STRAUB, E (1969). Zur Theorie der Prämienstufensysteme, *Mitteilungen der Verein. Schweizer Versich. Math*, **69**, 75-85.

WEEDA, P J. (1975). Technical Aspects of the Iterative Solution of the Automobile Insurance Problem, Report of the Mathematical Centre BN27/75, Amsterdam.

The ASTIN Bulletin 10 (1979), 223-239

THE THEORY OF INSURANCE RISK PREMIUMS — A RE-EXAMINATION IN THE LIGHT OF RECENT DEVELOPMENTS IN CAPITAL MARKET THEORY

YEHUDA KAHANE *

1. INTRODUCTION

The premium calculation principle is one of the main objectives of study for actuaries. There seems to be full agreement among the leading theoreticians in the field that the insurance premium should reflect both the expected claims and certain loadings. This is true for policy, risk or portfolio. There are three types of positive loadings: a) a loading to cover commissions, administrative costs and claim-settlement expenses; b) a loading to cover some profit (a cost-plus approach); and c) a loading for the risk taken by the insurer when underwriting the policy. The administrative costs can be considered a part of "expected gross claims". Thus, the insurer's ratemaking decision depends on his ability to estimate expected claims (including costs) and on the selection of a fair risk loading.

The main concern in the literature is the appropriate measurement of the risk and the exact loading formula. BÜHLMANN [1970, ch. 5] and others identified four possible principles of risk loading, namely, the expected value principle, the standard deviation loading, the variance loading, and the loading according to the principle of constant utility. Various studies point to the advantages and disadvantages of these principles and also examine some additional loading forms—semi-variance, skewness, etc. (e.g., BÜHLMANN [1970], BENKTANDER [1971], BERGER [1972], BURNESS [1972], BERLINER [1974], BERLINER and BENKTANDER [1976], BOHMAN [1976], COOPER [1974], GERBER [1975] and others). Despite different preferences in choosing the appropriate loading calculation principle, all seem to agree that the risk loading must be positive, since, otherwise, the firm would just have to wait for its ruin, that is bound to come sooner or later, according to risk theory.

The purpose of this article is to re-examine the appropriate principle of premium calculation in light of the recent developments in the theory of finance and especially in the theory of capital market equilibrium. These developments may suggest a new point of view and raise a few questions regarding the loading rules.

* Senior Lecturer at the Faculty of Management, Tel Aviv University, Israel, and Academic Director, Erhard Center for Higher Studies and Research in Insurance, Tel Aviv University, Israel. The author wishes to acknowledge the very helpful discussions with Dr. B. Berliner on an earlier draft of this paper and the many remarks of the participants of the 14th ASTIN Colloquium in Taormina, October 1978.

The first question is related to whether or not, and how, investment income should be considered in premium calculation [1]). Some insurers and insurance regulators tend to disregard investment income altogether. They misinterpret, perhaps, earlier models in risk theory which concentrated on the insurance portfolio in isolation and disregarded the investments merely for the sake of simplicity. Other insurers, and especially in certain lines, deduct investment income through the calculation of the expected *present value* of the relevant cash flows (claims and expenses). This paper suggests that investment income should be considered in ratemaking, either through a present value calculation, or through a negative loading on expected claims.

Another problem which can be solved with the use of financial theory is related to the appropriate measurement of risk for ratemaking purposes. It is suggested that the traditional measures of riskiness of an individual risk (standard deviation, variance, etc.) be replaced by the "systematic" element of the variance and that the risk loading be proportional to this element.

It will be shown that, since the profit of the insurer is derived from both underwriting and investment incomes, the insurer might, under certain circumstances, even be willing to lose on his underwriting activities. The appropriate loading on the expected pure claims may therefore be negative, and this may offer a theoretical explanation for the willingness of some insurers to under-rate [2]). The exact conditions for a negative loading will be studied later and an explicit expression for the profit (loss) will be presented. And finally, it is suggested that risk loadings should be determined by capital market equilibrium and must therefore be *objective* and uniform for all insurers.

The main argument in the following analysis can be explained by viewing a very simple example: Assume an investment company which raises funds through the sale of bonds (debt) and invests its capital plus the external funds in an assets portfolio. The required return on the shareholder's investment reflects the risks of the investment portfolio and the financial leverage (debt) used. Notice that the shareholders derive an appropriate profit after the payment of a positive interest on the firm's debt. Now assume an insurer is similar to the investment company, except that it raises the additional funds as a by-product of the sale of insurance contracts, rather than through the use of regular debt instruments. According to QUIRIN and WATERS [1975], this is analogous to a firm which *charges* a positive interest rate from its creditors, rather than *paying* them for the use of their money. A positive underwriting profit on the insurance portfolio would mean that the insurer

[1]) This topic has attracted many economists and actuaries. A discussion and references to some sources may be found in BIGER and KAHANE [1978], PYLE [1971], QUIRIN and WATERS [1975] or in a book by COOPER [1974].

[2]) The traditional explanations for underrating are related to the attempt to preserve long-term connections with insureds, or to the lack of knowledge and experience (see BENKTANDER [1971]).

makes a higher overall rate of profit than the investment company. Although the analogy is imperfect and very simplistic it may still demonstrate that consistent underwriting profits violate capital market equilibrium.

Section 2 summarizes the developments in financial literature and the risk-return relationships in capital market equilibrium. This will be used in Section 3 to analyze the treatment of investment income in ratemaking and the implications of the financial theory for the measurement of underwriting risks and the loading factor to be used in ratemaking. Some reservations and a few concluding remarks are summarized in Section 4.

2. RISK RETURN RELATIONSHIPS AND CAPITAL MARKET EQUILIBRIUM

Assume that the insurance company competes for investors' funds in the capital market. The firms' profits must therefore compensate the existing and potential shareholders for the risks they assume through their investment. The insurers' profitability is affected by the premium formula, and thus the relationship between the required expected return and the risk level on the insurer's shares may serve as a key to the ratemaking formula (BORCH [1974, ch. 22]).

Fairly recent developments in financial theory suggest that exact relationships between the expected return and the risk must prevail in market's equilibrium. A brief summary of these developments follows prior to the discussion of the implications for ratemaking.

Risk and Diversification

The basic idea in portfolio theory, which has been suggested by the pioneering work of Markowitz [1952], is imbedded in the mathematical properties of the standard deviation. I.e., the standard deviation of a linear combination of stochastic variables is typically lower than the weighted sum of the individual standard deviations. Each individual risk is represented by a stochastic variable, which is assumed to be fully characterized by its expected value and standard deviation [3]). The expected value is taken as a measure of profitability, while the standard deviation is used as a measure of the risk. It can easily be seen that there would generally be some gain from holding diversified portfolios, since the standard deviation of the portfolio will be lower (i.e. less "risky") than that of an undiversified portfolio.

This can be demonstrated by considering two securities A and B (see fig. 1). All portfolios obtained by holding these securities in varying proportions are represented by a curve APB. The nature of this curve depends on the correlation between the random variables A and B. In the extreme case, where the securities are perfectly positively correlated, there would be no gain from

[3]) See a short discussion in the concluding remarks.

diversification (*AQB* in fig. 1). In the other extreme case, where all securities are perfectly negatively correlated, the investor would even be able to construct a portfolio with a positive expected return and zero standard deviation (i.e., a risk-free portfolio (*R* in fig. 1)), although it is composed of individual risky securities.

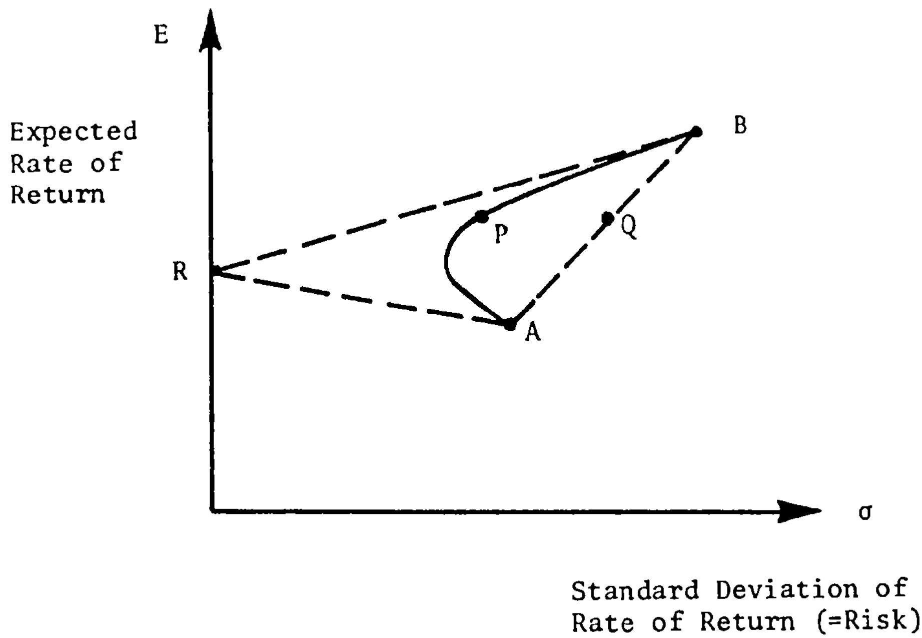

Fig. 1. The Effects of Diversification on the Portfolio's Expected Return and Risk.

Efficiency Frontier

Identifying the optimal portfolio is clearly not an easy task, since an infinite number of combinations of each pair of securities must be examined. The first step in the optimization is to calculate the efficient portfolio, which has the minimal standard deviation for a given level of expected value. This can be accomplished quite efficiently using the Quadratic Programming Technique (MARKOWITZ [1952]). Repeating the same process for all levels of expected value creates the *efficiency frontier* which is the locus of all portfolios having the lowest standard deviation at each level of expected value (curve *DEF* in fig. 2).

Knowing the efficiency frontier, the main problem is to select the *optimal* portfolio on that frontier. The traditional economic solution is based on the introduction of a set of indifference curves which represent the subjective trade-off between risk (standard deviation) and profitability (expected return). The optimal portfolio would be obtained at the tangency point between the

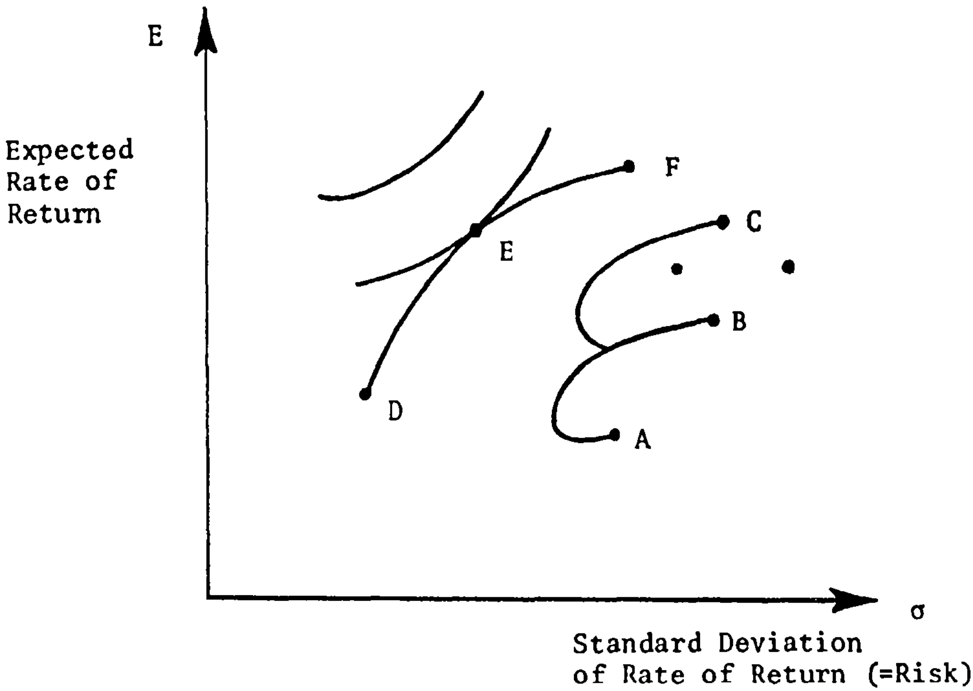

Fig. 2. Efficiency Frontier and the Optimal Portfolio.

highest possible indifference curve and the efficiency frontier (point E in fig. 2). This solution depends on the individual's subjective attitude toward risk reflected by the indifference curves and assumes a full knowledge of individual utilities.

The Capital Assets Pricing Model (CAPM)

The *CAPM* offers a new solution which does not depend on the individual's preferences and which is uniform for all investors. Its main assumption is the existence of a perfect capital market (i.e., there is a uniform interest rate at which each investor can borrow or lend any amount of money with no other transactions costs). The introduction of this interest rate, which is a risk-free security (R_f), causes dramatic changes in the efficiency frontier; combining a risky security, or portfolio, A with the risk-free security R_f will generate portfolios on the straight line R_fA (see fig. 3). The best combinations will lie on the ray R_fM which is tangent to the original efficiency frontier at M. Being on the section R_fM means that the investor lends part of his initial capital (purchases risk-free bonds). A portfolio represented by a point on ray R_fM but to the right of M is obtained by borrowing money at the risk-free rate and investing the capital and the borrowed funds in the risky portfolio M (i.e., by using "financial leverage"). The optimal portfolio is selected in two isolated stages. The first consists of finding the portfolio M of risky securities. In the second stage the desired mix of this portfolio with the risk-free asset is selected according to the tangency of $R_f M$ to the indifference curves.

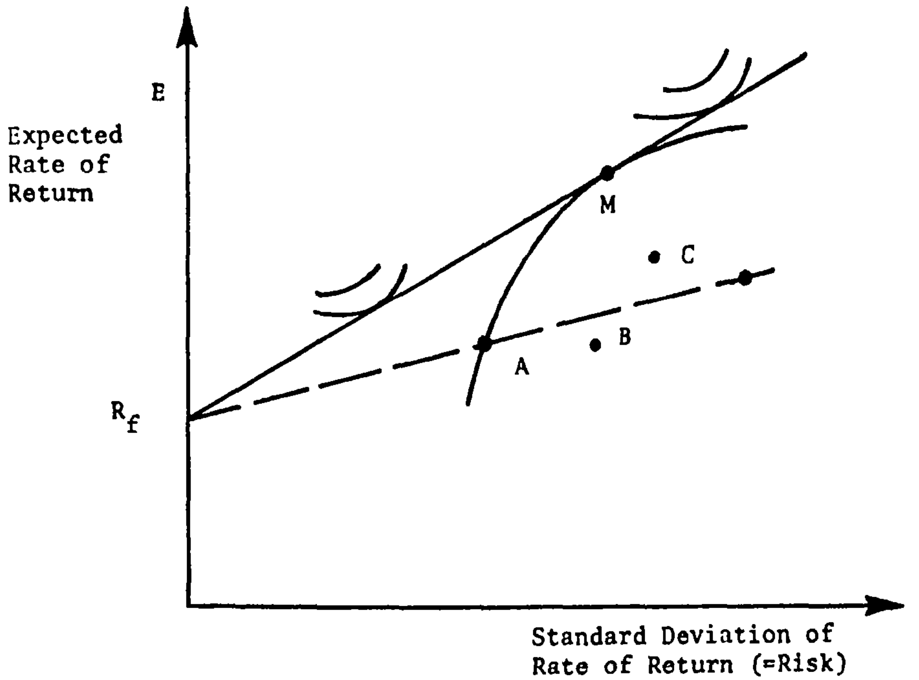

Fig. 3. Capital Asstes Pricing Model

The next step in the development of the $CAPM$ is based on the assumption that all investors have the same expectations concerning the means, standard deviations and covariances between all securities. Under a model of full agreement, all investors must hold the same portfolio composition of risky securities (point M). This portfolio is composed of all the risky ventures and is called the "market line" portfolio. The combinations of this portfolio with the risk-free interest rate, lie on a straight line called the "market line" which represents the risk-return relationship for *all* portfolios in the market. It is impossible to create a portfolio with a better performance which would be represented by a point above this capital market line. Any portfolio below this line would be inferior. The equation for the capital market line is

$$E_p = R_f + \frac{E_m - R_f}{\sigma_m} \sigma_p, \tag{1}$$

where E and σ denote expected value and standard deviation, respectively, and the subscripts p and m denote a portfolio and the market portfolio, respectively (SHARPE [1964], LINTNER [1965], MOSSIN [1966]).

Equation (1) represents the objective risk-return relationship for a portfolio in market equilibrium and can be interpreted as follows: The expected return on any investment portfolio equals the risk-free rate of interest plus a risk loading which is proportional to the standard deviation of the portfolio.

Under the *CAPM*, the appropriate risk measure for a portfolio of securities is the standard deviation and not its variance. This result stems from the basic assumption of the model and therefore cannot be used as an argument against the use of a risk loading proportional to the variance, which is recommended by some of the leading authorities in the Collective Risk Theory (BÜHLMANN [1970], BERLINER [1974], BOHMAN [1976], etc.).

Risk-Return Relationship for an Individual Risk

The capital market line is obtained through the holding of a combination of securities which are typically below it (like points A, B, C, in fig. 3). What is the appropriate risk-return relationship for the individual security? Further analysis of the *CAPM* showed that the expected return of each individual investment under equilibrium must satisfy the following equation

$$E_i = R_f + \frac{E_m - R_f}{\sigma_m^2} \sigma_{im}, \tag{2}$$

where the σ_{im} represents the *covariance* between the return on security i and the return on the market portfolio (The proof for these relationships is given by SHARPE [1970, pp. 86-90]). Equation (2) means that the expected return on the individual security equals the return on the risk-free asset plus a proportional risk loading. Unlike the relationship for a portfolio (equation (1)), the risk for an individual security is measured by σ_{im}, the covariance of the return on the security and the market portfolio. This suggests a new measure for the risk level of an individual security—the *systematic* risk element. A variation of this term, namely, σ_{im}/σ_m^2, is often used in financial literature for the same purpose and is called the "beta" coefficient.

The risk for an individual security, unlike the measure of risk for a portfolio (collective risk), is not measured by its standard deviation or variance. The full variance of the return on each security is split into two components: the systematic risk (representing the correlation with the market portfolio), and a non-systematic element (representing random fluctuations or noise). This is demonstrated by fig. 4, which shows the return on a hypothetical security i and the return on the market portfolio. The dots on this graph represent individual observations (periodic observations). The systematic element is captured by the slope of the regression line. The vertical deviations

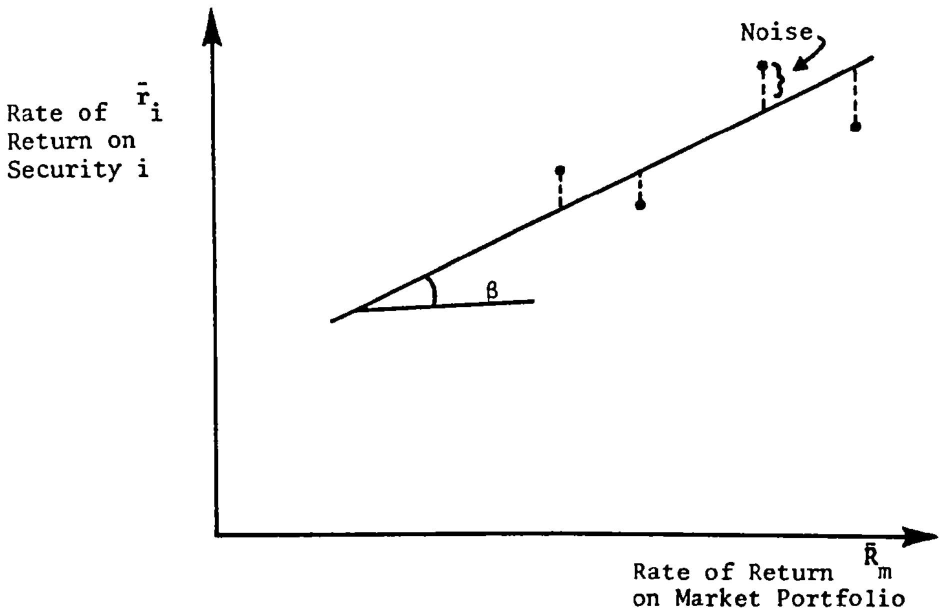

Fig. 4. Systematic and Non-Systematic Risk.

of the observed return from its conditional expected value represent a random noise.

The non-systematic element (the "noise") is excluded from the measurement of risk because it can be diversified away and eliminated to a great extent by holding appropriately diversified portfolios [4]). This results from the assumption that the random fluctuations of securities i and j are uncorrelated.

The return on securities fluctuates. Despite these fluctuations some securities may be regarded as risk-free where their rates of return have no consistent relationships with those of the market portfolio. In such a case their expected return must equal the risk-free rate of interest. Such securities are represented by lines with zero slope in fig. 5. Other securities may be represented by a slope of unity. Holding such securities has an effect similar to the holding of the market portfolio itself (despite their higher variance caused by random noise). Securities having slopes steeper than unity are "aggressive", i.e., they augment the fluctuation of the market and are therefore more risky than the market portfolio. Some securities may even have negative slopes, which means that they behave counter to the market portfolio. The expected return on these securities would be lower than the risk-free rate of interest since they have a risk reduction effect in a portfolio context.

[4]) See a quite similar idea in BERLINER [1974].

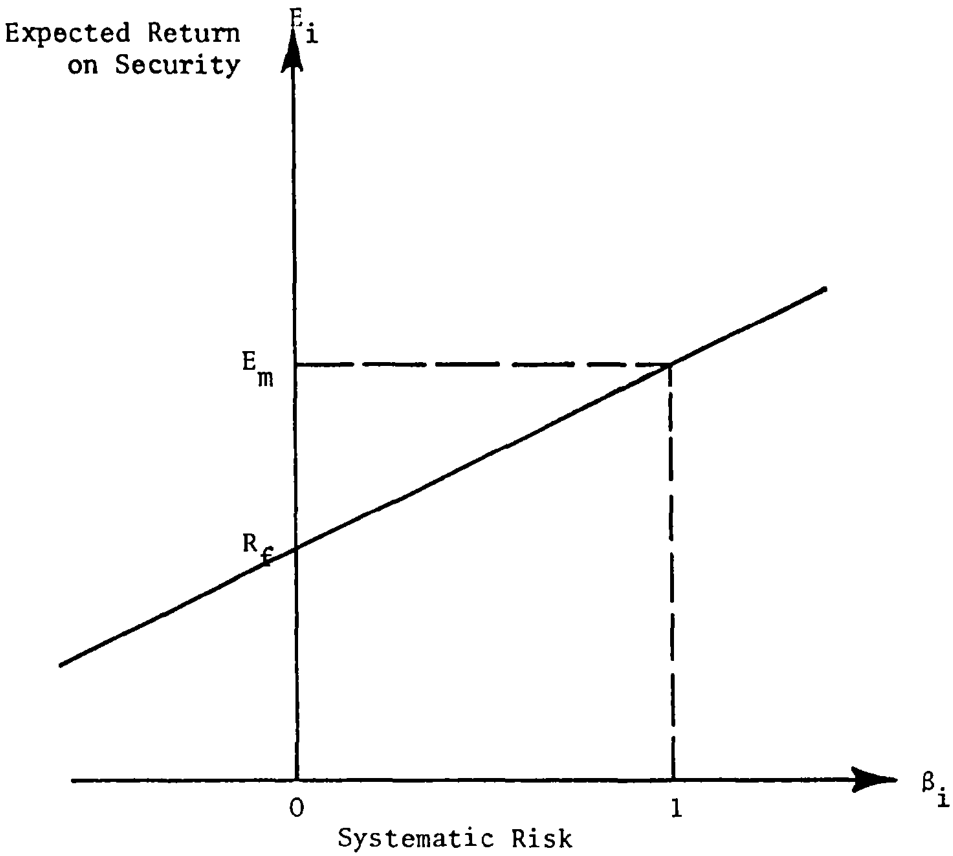

Fig. 5. The Risk-Return Relationship for Individual Risks.

3. IMPLICATIONS FOR INSURANCE RATEMAKING

The *CAPM* is obviously an over-simplified representation of financial markets in the real world. The model rests on the assumption that a security is completely described by a stationary probability distribution and that only the first two moments of the distribution are relevant. In addition, the model assumes uniform information among investors, identical investment planning horizons, and perfect capital markets with a risk-free rate of interest. Despite the over-simplifications, the model seems to capture some of the essential elements in real situations and has demonstrated a fairly good explanatory power in empirical tests [5]). Unfortunately, this model has hardly received the attention it deserves in actuarial literature. Among the few exceptions are the works by Borch [1974, ch. 9, 21, 22] and by Quirin et al. [1974].

[5]) There are a great number of empirical tests for the validity of the *CAPM*. A review of some of the tests can be found in Modigliani and Pogue [1974].

The potential of the *CAPM* for the analysis of the ratemaking issue is quite obvious. According to the *CAPM*, there should be an objective market price per unit of risk. This may suggest that the insurance risk loadings must be determined objectively, rather than through subjective considerations of the insurance company. It means that the loading should not depend on management attitude toward risk (i.e., its utility function). Moreover, the *CAPM* may be used to find the exact parameters for the risk loading.

The profit of the insurer is derived from two sources: its underwriting profits and its investment income. Thus, the ratemaking problem should be analyzed by considering the two income sources simultaneously. It will be shown that previous studies which simplified the analysis by examining the insurance portfolio in isolation (e.g., BENKTANDER [1971], BÜHLMANN [1970]) offered only a partial solution for the ratemaking problem.

Assume that the firm has m insurance activities (policies or lines). The firm collects $\$X_i$ in premiums for contract i and expects to make an underwriting *loss* (profit) of $X_i\tilde{r}_i$ dollars. $\tilde{r}_i$ is a stochastic variable representing the rate of underwriting loss in this line (a negative value will denote profit). The stochastic variables are clearly affected by the ratemaking formula in use (since it determines the expected rate of profit or loss through the profit loading).

Assume that the insurer holds an investment portfolio composed of n securities (assets). The amount invested in activity i is $\$X_i$ $(i = m + 1, \ldots, m + n)$, and the rate of return on this activity is a stochastic variable $\tilde{r}_i$. The total profit of the firm, $\tilde{Y}$, is

$$Y = \sum_{i=m+1}^{m+n} X_1\tilde{r}_i - \sum_{i=1}^{m} X_i\tilde{r}_i, \tag{3}$$

where the two summations in the right-hand term express the aggregate investment profit and the total underwriting loss (profit), respectively. Equation (3) can be expressed in terms of rate of return on equity, $\tilde{r}_y$, by dividing both sides of the equation by the equity capital K

$$\tilde{r}_y \equiv \frac{Y}{K} = \sum_{i=m+1}^{m+n} \frac{X_i}{K}\tilde{r}_i - \sum_{i=1}^{m} \frac{X_i}{K}\tilde{r}_i. \tag{4}$$

Let $x_i = \dfrac{X_i}{K}$ denote the premiums and investments relative to the capital. A subscript j can be added to the elements of the equation in order to relate it to a certain insurance company j

$$\tilde{r}_{yj} = \sum_{i=m+1}^{m+n} x_{ij}\tilde{r}_i - \sum_{i=1}^{m} x_{ij}\tilde{r}_i. \tag{5}$$

Note that r_i are assumed to be identical for all companies in the market.

Now make the brave assumption that the *accounting* rate of return on the firm's equity is equal to the *market* rate of return on the firm's shares [6]. Under this assumption, equation (5) also reflects the return on the firm's shares. The *CAPM* suggests that the expected return on firm j shares is related to its systematic risk β_j as follows

$$(6) \qquad E(\tilde{r}_{yj}) = R_f + (E_m - R_f)\,\beta_j .$$

Taking the expected value of equation (5) and substituting into (6) yields

$$(7) \qquad E(\tilde{r}_{yj}) = \sum_{i=m+1}^{m+n} X_{ij}\,E(\tilde{r}_i) - \sum_{i=1}^{m} X_{ij}\,E(\tilde{r}_i) = R_f + \beta_j\,[E_m - R_f].$$

Note that the systematic risk of a portfolio is a linear combination of the systematic risk elements of its components [7]. Therefore, the systematic risk of the insurance firm j is a weighted average of the systematic risk of all underwriting and investment activities, that is,

$$(8) \qquad \beta_j = \sum_{i=m+1}^{m+n} x_{ij}\beta_i - \sum_{i=1}^{m} x_{ij}\beta_i .$$

Substituting equation (8) into (7) and eliminating the subscript j for the simplicity of notation yields:

$$(9) \qquad \sum_{i=m+1}^{m+n} x_i\,E(\tilde{r}_i) - \sum_{i=1}^{m} x_i\,E(\tilde{r}_i) = R_f + \Big[\sum_{i=m+1}^{m+n} x_i\beta_i - \sum_{i=1}^{m} x_i\beta_i\Big]\,[E_m - R_f]$$

Since investment activities obey the same capital market equilibrium relationships, the expected return on every investment satisfies the equation

$$(10) \qquad E(\tilde{r}_i) = R_f + \beta_i\,(E_m - R_f) \qquad i = m+1, \ldots, m+n,$$

and the return on the entire investment portfolio is

$$(11) \qquad \sum_{i=m+1}^{m+n} x_i\,E(\tilde{r}_i) = \sum_{i=m+1}^{m+n} R_f x_i + \sum_{i=m+1}^{m+n} x_i\beta_i\,(E_m - R_f).$$

Subtracting (11) from (9) gives the expected underwriting profit which preserves the capital market equilibrium

$$(12) \qquad -\sum_{i=1}^{m} x_i\,E(\tilde{r}_i) = R_f\,\Big(1 - \sum_{i=m+1}^{m+n} x_i\Big) - \sum_{i=1}^{m} x_i\beta_i\,(E_m - R_f).$$

---

[6] The problem of consistency between accounting and market data, and especially the relationship between accounting and market betas, is studied in excellent papers by BEAVER and MANEGOLD [1975] and by BEAVER, KETTLER and SCHOLES [1970]. These papers give references to many earlier works on the subject

[7] Assume a portfolio z consisting of a linear combination of stochastic variables x and y $\tilde{z} = \tilde{a}x + b\tilde{y}$. The systematic risk of this portfolio (where $\tilde{m}$ denotes the return on the market portfolio) is

$$\beta_z = \operatorname{cov}(\tilde{z}, \tilde{m}) / \operatorname{var}(\tilde{m}) = [1 / \operatorname{var}(\tilde{m})] \cdot \operatorname{cov}[a\tilde{x} + b\tilde{y}, \tilde{m}] =$$
$$= [1 / \operatorname{var}(\tilde{m})] \cdot [a \cdot \operatorname{cov}(\tilde{x}, \tilde{m}) + b \cdot \operatorname{cov}(\tilde{y}, \tilde{m})] = a \cdot \beta_x + b \cdot \beta_y .$$

Assuming that each dollar of premium in insurance activity i generates g_i dollars of investment, the insurer's balance sheet equality (i.e., the requirement that assets equal the equity plus liabilities) is expressed as

$$(13) \qquad \sum_{i=m+1}^{m+n} x_i = 1 + \sum_{i=1}^{m} x_i g_i .$$

Substituting (13) into (12) and rearranging gives

$$(14) \qquad \sum_{i=1}^{m} x_i E(\tilde{r}_i) = R_f \sum_{i=1}^{m} x_i g_i + \sum_{i=1}^{m} x_i \beta_i [E_m - R_f].$$

This equation does not lead to a clear-cut statement about the expected rate of underwriting loss on each individual insurance activity. Given all β_i, E_m, R_f and the values of x_i, the equation is insufficient to determine a single set of $E(\tilde{r}_i)$ $(i = 1, \ldots, m)$ and there may be a large number of vectors that satisfy it (BIGER and KAHANE [1978]). However, one possible solution may be of special interest, since it resembles the *CAPM* relationship

$$(15) \qquad E(\tilde{r}_i) = R_f g_i + \beta_i [E_m - R_f], \qquad (i = 1, \ldots, m).$$

That is, on the average, the firm would be willing to lose on insurance activity i as much as g_i times the risk-free rate, plus a risk loading proportional to its systematic risk.

The Investment Income in Ratemaking

The intuitive solution in equation (15) is attractive, since it may have an interesting interpretation regarding the treatment of investment income in the ratemaking formula. The normative question of whether or not investment income should be considered in ratemaking was extensively discussed in the literature. However, this problem has seldom been examined under capital market equilibrium, and even in these cases it was studied under the simplified model where all insurance activities were aggregated and only one or two assets were assumed (PYLE [1971], QUIRIN and WATERS [1975]). According to equation (15), there is a negative loading $R_f g_i$ (recall that $E(\tilde{r}_i)$ represents expected underwriting *loss*) which represents the investment income and is indirectly generated through the insurance activity i.

Under the simplifying assumption made, the deduction should be proportional to an approximated value g_i, the funds generating coefficient. For example, if the activity generates one investment dollar for each dollar of premium but creates no systematic risk, the firm may be willing to underwrite this activity for an expected underwriting loss equivalent to the risk-free rate! On a line which generates more than one dollar of investments for each dollar of premiums (e.g., liability insurance) the firm is willing to lose even more.

A more accurate solution would probably be to deduct the investment income through the calculation of the present value of claim payment (similarly to the calculation of life insurance premiums). Note that this negative loading is proportional to the *risk-free* rate of interest, although the firm invests in a combination of risky assets. The investment risk is ignored in rate-making decision since it is accounted for by the risk premium element which is imbedded in the expected return on each risky asset, under capital market equilibrium.

Risk Loading in Ratemaking

The expected underwriting loss in equation (15) is also a function of the "risk" of the specific insurance contract. Equation (15) therefore may serve as a guide in determining the *risk loading*. Since the analysis concentrated on underwriting *losses*, a project with a *positive* beta has the desirable risk-reduction properties. Therefore, a positive beta would justify additional underwriting loss (over the negative investment income loading).

The risk loading is proportional to the systematic element of risk, beta, that reflects the contribution of an activity to the market portfolio. This means that the underwriting loss (profit) on an activity may fluctuate dramatically around its expected value (i.e., high variance) but nevertheless may be regarded as riskless by the shareholders of the firm. The risk loading is proportional to the beta, according to the objective price of a unit of risk. This price factor is given by the difference between the expected return on a market portfolio and the return on the riskless interest rate. This price is uniform for all investors.

Preliminary findings presented in a recent paper by BIGER and KAHANE [1978] suggest that underwriting profits are uncorrelated with the rate of return on the market portfolio (i.e., underwriting activities have no systematic risk). Thus, according to equation (15), the average underwriting *losses* should be approximately equivalent to the risk-free rate of interest (g_i is commonly close to 1), while for the liability insurance lines, which typically generate more funds because of the long claims settlement period (reflected by larger g_i), the losses must be even higher.

Rough empirical evidence in support of the ratemaking formula suggested in this paper can be obtained from aggregate statistical data of the insurance industry. Although the ratemaking formulas approved by regulators in most countries include a *positive* profit loading on net premiums, insurers often report underwriting losses. In view of the underwriting losses which insurers do complain about, and noting that the losses typically fluctuate in the range around the level of the risk-free rate of interest, the loading formula suggested here has some explanatory power. It appears as if competition has forced the rates to reach their equilibrium level, despite the regulatory formula.

4. CONCLUDING REMARKS

This paper examined some of the implications of recent developments in the financial literature and capital market equilibrium theory for the insurance ratemaking problem. In an early and almost unique study of this problem, Borch [1974, ch. 9] stated that no pareto optimal equilibrium can exist in (re)insurance market. We did not obtain a unique solution either, but we presented a possible solution that has a great intuitive appeal.

The ratemaking formula which is suggested in this paper has at least two types of loadings (on top of the "expected gross claims"): (a) a loading to reflect the investment income; and (b) a risk loading.

The Role of Investment Income

Earlier studies in risk theory concentrated on the statistical nature of the claims process in isolation. This simplifying assumption led, unfortunately, to the incomplete solution which ignored the ability of the firm to obtain an investment income as a result of its underwriting activity. The ratemaking formula suggested here includes a negative loading which is proportional to the amount of investment generated by the insurance activity. The relevant rate of interest is the risk-free rate of interest (even though the funds are typically invested in risky assets).

This may be considered an approximation to the deduction of the investment income through the calculation of *present values*. Such an approach is taken in the actuarial calculation of life insurance premiums but is often disregarded in non-life insurance.

Risk Loading

A second element in the ratemaking formula is a loading for the risk of the individual activity. Two new concepts are introduced: First, the risk level of an individual risk is measured by the systematic risk ("beta") rather than by the variance, standard deviation or other traditional measures. Secondly, most authorities in the field of risk theory concentrate on "internal" factors to determine the correct loading—those related to the nature of the individual activity or of the firm (e.g., management utility). It is argued in this paper that the appropriate loading is determined *objectively*, according to the market's price of a unit of risk, rather than through the subjective decision of the firm. Thus, the traditional thought that a small firm is "punished" by having to charge a higher loading (BENKTANDER [1971]) should be re-examined [8]).

[8]) Theoretically, the availability of reinsurance enables the small firms to transfer the excessive risks, as long as there is no discrimination in reinsurance rates.

Limitations

Despite its simplified assumption, the model sheds some light on the ratemaking formula problem. Some reservations and suggestions for future studies seem, however, to be required.

(a) The model is based on the assumption that insurers and investors know the correct parameters of the relevant distributions. The risk element resulting from statistical errors and incorrect (biased) estimation of the parameters has not been discussed here. Such an element may justify a special risk loading. Similar recognition should be given to the risk originated by non-stationary distributions.

(b) The level of aggregation affects the risk measure. The term "insurance activity" can be used in a narrow meaning (individual policy) or a broader sense (an insurance line). At higher levels of aggregation, the systematic risk approaches the standard deviation (since 'noise' is eliminated through diversification). Thus, when dealing with aggregated lines, the difference between the "beta loading" and the traditional loading, which is proportional to the standard deviation or variance, is very limited indeed [9]).

(c) All distributions were assumed to be characterized by the first two moments. This makes the model acceptable only for certain utility assumptions. It is not inevitable that loading factors which are related to higher moments should be analyzed under more sophisticated models. Thus, measures of asymmetry, like the skewness and semi-variance, may be needed in a loading formula (especially for risks with catastrophic nature—which are represented by extremely skewed distributions). Another shortcoming of the model is its limitation to a single period analysis so that it cannot handle diversification over a multi-period horizon—which may be needed for the risks with catastrophic nature.

(d) The analysis ignored the problem of inflation and growth. All parameters were assumed stable and in real terms. Non-zero inflation, for example, may cause some problems since riskless assets may become risky in real terms, and this may create difficulties with the *CAPM*. Also, since investment income often does not keep up with inflation, there may be a need for another element of positive loading. The problem of inflation is only partially handled in the model through the determination of the parameters.

The model suggested in this paper cannot be regarded as the final answer to the ratemaking problem in practice. There is still much room for further improvements through the development of models with more relaxed assumptions. Some adjustments will probably improve the explanatory power of the model. Among these, a possible suggestion is the analysis of the case

[9]) Some empirical evidence shows that the rate of return on shares in the stock market is related to both their "betas" and the standard deviations. (See a summary in MODIGLIANI and POGUE [1974]).

where investors have different planning horizons and may differ in their anticipation of the prospects of various securities. In addition, it would be worthwhile to examine the effects of other imperfections in the capital and reinsurance markets and the effects of possible differences between accounting and market data.

Despite these obvious shortcomings of the model, it contributes to a better understanding of and a new approach to the calculation of insurance rates.

REFERENCES

BEARD, R. E., T. PENTIKAINEN, and E. PESONEN. (1969). *Risk Theory*, Methuen &Co., London.

BEAVER, W., P. KETTLER, and M. SCHOLES. (1970). The Association between Market-Determined and Accounting-Determined Risk Measures, *Accounting Review*, 654-682.

BEAVER, W. and J. MANEGOLD. (1975). The Association Between Market-Determined and Accounting-Determined Measures of Systematic Risk: Some Further Evidence, *Journal of Financial and Quantitative Analysis*, **10**, 231-284.

BENKTANDER, G. (1971). Some Aspects of Reinsurance Profits and Loadings, *The ASTIN Bulletin*, **5**, 314-327.

BERGER, G. (1972). An Attempt on the Subject: Profit Goal and Security Loadings in Risk Insurance, *Transactions of the 19th International Congress of Actuaries*, **4**, 595-603.

BERLINER, B. (1974). Some Thoughts on (Re) insurance Loadings under a Ruin Criterion, *Scandinavian Actuarial Journal*, **2**, 76-80

BERLINER, B. (1977). A Risk Measure Alternative to the Variance, *The ASTIN Bulletin*, **9**, 42-58.

BERLINER, B. (1976). On the Choice of Risk Loadings, *Transactions of the 20th International Congress of Actuaries*, Tokyo.

BIGER, N. and Y. KAHANE. (1978). Risk Considerations in Insurance Ratemaking, *Journal of Risk and Insurance*, **45**, 121-131.

BOHMAN, H. (1976). Solvency and Profitability Standards, *Scandinavian Actuarial Journal*, **2**, 110-113.

BORCH, K. (1974). *The Mathematical Theory of Insurance*, Lexington Books.

BÜHLMANN, H. (1970). *The Mathematical Methods in Risk Theory*, Berlin: Springer Verlag.

BURNENS, E. (1972). Contingency Loadings in Life Insurance, *Transactions of the 19th International Congress of Actuaries*, **4**, 621-631.

Commonwealth of Massachusetts, Division of Insurance, Rate of Return and Profit Provision in Automobile Insurance, Unpublished Memorandom, (1977).

COOPER, R. W (1974). *Investment Return and P-L Insurance Ratemaking*, Homewood, Illinois, Irwin, Inc.

GERBER, H. (1974). On Additive Premium Calculation Principles, *The ASTIN Bulletin*, **7**, 215-222.

KAHANE, Y (1977). Ratemaking and Regulation in Property-Liability Insurance, *Quarterly Review of Economics and Business*, **17**, 97-112.

LINTNER, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *The Review of Economics and Statistics*, **47**, 13-37.

MARKOWITZ, H (1952). Portfolio Selection, *Journal of Finance*, **7**, 77-91.

MODIGLIANI, F. and J. POGUE. (1974). An Introduction to Risk and Return: Concepts and Evidence, *Financial Analysts' Journal*, 68-79 and 69-85.

MOSSIN, J. (1966) Equilibrium in a Capital Assets Market, *Econometrica*, **34**, 768-783.

PYLE, D. H. (1971). On the Theory of Financial Intermediation, *Journal of Finance*, **26**, 737-747.

QUIRIN, D. G., HALPERN, P. J., KALYMON, B. A., MATHEWSON, G. F. and W. R. WATERS. (1974). *Competition, Economic Efficiency and Profitability in the Canadian P-L Insurance Industry*, Toronto: Insurance Bureau of Canada.

QUIRIN, D G. and W. R. WATERS. (1975). Market Efficiency and the Cost of Capital: The Strange Case of Fire and Casualty Insurance Companies, *Journal of Finance*, **30**, 427-450.

SHARPE, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance*, **19**, 425-442.

SHARPE, W. F. (1970). *Portfolio Theory and Capital Markets*, New York, McGraw Hill.

INFORMATION FOR CONTRIBUTORS

All editorial matters should be sent to the Managing Editor:

P. ter Berg
c/o NVVA, Afdeling Statistiek en Onderzoek,
Boerhaavelaan 3,
2713 HA ZOETERMEER,
The Netherlands.

The following remarks are of special importance for manuscript preparing: Each paper should contain an introduction. As regards references, these should be reasonable complete; references in the text are given by the author's name in capitals followed by the year of publication between parentheses. At the end of the paper the references should be grouped alphabetically; for journal references give at least author(s), year, title, journal, volume and pages; for book references give at least author(s), year, title, city and publisher.
In mathematical expressions, authors are requested to minimize unusual or expensive typographical requirements. This may be achieved by using the slash in preference to built-up fractions and to write complicated exponentials in the form exp( ). Equation numbers must be at the left. Figures must be drawn in black ink on white paper in a form suitable for photographic reproduction with a lettering of uniform size and sufficiently large to be legible when reduced to final size. Figures should be designated by arabic numbers and must have a title. Any legends for figures must be typed on a single separate sheet rather than placed on the drawings. Tables should be numbered, should have a heading and should be prepared without vertical lines.
Material that has already been published or submitted for publication elsewhere, is not acceptable. So, a confirmation of the sole rights for the Bulletin is required.
It is advisable to minimize the size of the paper, maintaining the scientific quality and the clearness of the contents, in order to maximize the probability of acceptance.
No paper will be rejected until it has been reviewed by the Editorial Board.

* * *

SUBSCRIPTION

ASTIN (Actuarial Studies in Non-Life Insurance) is a section of the International Actuarial Association (IAA). Membership is open automatically to all IAA members and under certain conditions to non-members also. Applications for membership can be made through the National Correspondent or, in the case of countries not represented by a national correspondent, through the Secretary:

F. E Guaschi, Moorfields House, Moorfields, London EC2Y 9AL, England.

Members receive the Bulletin free of charge.

Non-members who want to receive the Bulletin as it comes out are invited to apply through the Treasurer:

J Adam, 53 Boulevard Emile Jacqmain, 1000 Bruxelles, Belgium.

Anyone interested in one or several items of the series of The Astin Bulletin may send his order directly to:

E J. Brill — Publisher, Oude Rijn 33a, Leiden, The Netherlands.

He will be invoiced accordingly.

COMMITTEE OF ASTIN

Chairman:	Giovanna FERRARA	Salerno
Vice-Chairman:	Joseph ADAM	Bruxelles
Past Chairman:	LEROY J. SIMON	Newark
Treasurer:	Joseph ADAM	Bruxelles
Secretary:	Francis E. GUASCHI	London
Editor:	Freek BOERMANS	Rotterdam
Members:	Harald BOHMAN	Saltsjöbaden
	Hans BÜHLMANN	Zürich
	Paul JOHANSEN	Copenhagen
	Erkki PESONEN	Helsinki
	Jean SOUSSELIER	Paris
	Jürgen STRAUSS	Munich
Honorary Chairman:	Edouard FRANCKX	Brussels

EDITORIAL BOARD

Members:	Joseph ADAM	Bruxelles
	Hans BÜHLMANN	Zürich
	Jan JUNG	Saltsjöbaden
	LEROY J. SIMON	Newark
Managing Editor:	Peter TER BERG	Zoetermeer

All editorial matters should be sent to the Managing Editor.
For further information see inside back-cover.