Volume 24, No. 1 May 1994

ASTIN BULLETIN

A Journal of the International Actuarial Association

CONTENTS

Ceuterick

EDITORIAL POLICY

ASTIN BULLETIN started in 1958 as a journal providing an outlet for actuarial studies in non-life insurance Since then a well-established non-life methodology has resulted, which is also applicable to other fields of insurance For that reason *ASTIN BULLETIN* has always published papers written from any quantitative point of view—whether actuarial, econometric, engineering, mathematical, statistical, etc —attacking theoretical and applied problems in any field faced with elements of insurance and risk Since the foundation of the AFIR section of IAA, i e since 1988, *ASTIN BULLETIN* has opened its editorial policy to include any papers dealing with financial risk

ASTIN BULLETIN appears twice a year (May and November), each issue consisting of at least 80 pages

Details concerning submission of manuscripts are given on the inside back cover

MEMBERSHIP

ASTIN and AFIR are sections 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 a member of the Committee of ASTIN

Members of ASTIN receive *ASTIN BULLETIN* free of charge As a service of ASTIN to the newly founded section AFIR of IAA, members of AFIR also receive *ASTIN BULLETIN* free of charge

SUBSCRIPTION AND BACK ISSUES

ASTIN BULLETIN is published and printed for ASTIN by Ceuterick s a, Brusselsestraat 153, B-3000 Leuven, Belgium

All queries and communications concerning subscriptions, including claims and address changes, and concerning back issues should be sent to Ceuterick

The current subscription or back issue price per volume of 2 issues including postage is BEF 2 500

Back issues up to issue 10 (= up to publication year 1979) are available for half of the current subscription price

INDEX TO VOLUMES 1-20

The Cumulative Index to Volumes 1-20 is also published for ASTIN by Ceuterick at the above address and is available for the price of BEF 400

EDITORIAL AND ANNOUNCEMENTS

GUEST EDITORIAL

At the ASTIN Colloquium in Cambridge Willem de Wit has left the ASTIN Committee which he has served for many years. The editors have therefore invited him to write a Guest Editorial reflecting his thinking and his great experience from which ASTIN has profited over decades We are happy that he has accepted this task.

THE EDITORS

ACTUARIAL SCIENCE

PAST – PRESENT – FUTURE

The concept of insurance is very old, if tradition is to be believed. Already in ancient times we find traces of insurance business

Obviously we are talking about risks, and it is well known that already the Asipus collected data to describe risks and to point the way how to cope best with these risks. One could say that this was a very first start of risk management. Their description of risks was mainly based on concepts like certainty, trust and expertise (which we fundamentally still acknowledge today), while even religious considerations were taken into account. Good as well as bad results were recognized, but the concept of probability was unknown to them.

The roots of thoughts on probability we find at Aristoteles, but for the development of the concept of probability we have to wait until the 16th century, when Cardano (1565) wrote his ideas about that concept, while the final breakthrough was realised by Pascal and Fermat (1654) and Huygens (1657).

LIFE

Then the time was ripe for the first actuarial activity · John Graunt (1662), Johan de Witt (1671) and Edmond Halley (1694) made a mortality table. They used data which were taken from censuses. For the sale of annuities (among others) they used this historical material For the time being they had to manage with different, uncertain sources, until the 19th century when regular censuses started, from which up-to-date mortality tables could be derived.

In fact life insurance mathematics has always been very simple and the invention of commutation columns can be seen as the most important invention in life insurance mathematics. Recently the complexity has, also because of the use of computers, increased very much, on the one hand because of the application of stochastic techniques, on the other hand because of asset-liability-management.

NON-LIFE

For the non-life branch we had to wait somewhat longer Not only was there no data for a long time, but also the theory started later. If we consider the work of Lundberg (1909) as the start of actuarial activities in the field of non-life insurance, then we can't yet celebrate the centenary.

Still in 1940 the application of the theory of probabilitics was described as doubtful After the Second World War non-life insurance business became more important, and theories developed further, in which automobile insurance acted as a pioneer. However observational data still remained scanty and often one had to deal with limited samples. In contrast with life insurance, where only a simple two-dimensional development (mortality table) exists, non-life insurance moved quickly on to econometrical models, where numerous variables play a role. Still today one has to conclude that a number of fields of non-life insurance business are even yet very difficult to handle.

OBSERVATIONAL DATA

It appears very hard in many Countries to collect the adequate observational data and it is remarkable that sometimes for practical application one has to make use of old data, even sometimes from other countries

Slowly but surely this situation improves, but considering the coming of bigger and open markets, especially in Europe, it is very important to collect the adequate observational data It is also necessary to be sure that the data are mutually comparable.

FUTURE

In the succession of mortality tables actuaries notices soon a decline of mortality. First they tried to find an "explanation" for this, on the base of Newtonian determinism, but that failed. The consequences of a further decline of mortality are becoming more and more apparent, because of the recent shift from insurances of death risk to insurances with a long life risk Yet it appears impossible in any reliable way to make a good forecast of future mortality And that is just what we need. Also in non-life insurance this forecasting plays an important role Next to the process of inflation, one could also think of, for example, a changing risk structure, but also of changing legal and social views.

Now that developments in the world are happening rapidly, so also are risks, the basis of insurance, evolving at the same rapid pace. But all our observational data come from the past There are too few attempts that try to make a forecast for the future. If this is already being done, it is often not much more than the continuation of an observed trend from the past. Popper once wrote that a forecast in social science is in principle impossible (in my opinion we should conceive actuarial science as a social science) He warns particularly against assuming a continuation of what happened in the past into the future The elaboration of scenarios in which the parameters of the model can be changed in different ways seems to be an obvious alternative.

SOCIETY

Actuarial science becomes more and more socially involved. Because of the fact that in many countries the government is partly withdrawing from social insurance, private insurance companies are confronted with new problems, like affordability, obligatory acceptation, and so on, in short, with problems of solidarity.

WHERE TO?

Besides further development of the theory, in my opinion actuarial science in the future has especially to be working on

— collecting the adequate data, mutually comparable,
— developing scenarios for the future to forecast the consequences for the future,
— considering life, non-life and financial services as one and the same branche,
— social problems.

One must search for the greatest possible simplicity with regard to the above, both for our own practice as well as for those countries where the concept of insurance is not yet so far developed. After all they have to join in too

G. W. DE WIT

ARTICLES

RISK ALLOCATION IN CAPITAL MARKETS: PORTFOLIO INSURANCE, TACTICAL ASSET ALLOCATION AND COLLAR STRATEGIES

By ERIC CHEVALLIER AND HEINZ H MULLER

University of Zürich

ABSTRACT

The theory of risk exchange is applied on the allocation of financial risk in capital markets It is shown how the shape of individual payoff functions depends on risk tolerance and cautiousness. For the special case where the Neumann-Morgenstern utility functions of all individual investors belong to the HARA class and have non decreasing risk tolerance it is proved that generalized versions of "portfolio insurance", "tactical asset allocation" and "collars" are the only strategies occurring in price equilibrium.

KEYWORDS

Non linear risk sharing, price equilibrium; portfolio insurance.

1. INTRODUCTION

For quite a long time the MARKOWITZ (1952) approach and the Capital Asset Pricing Model (SHARPE, 1964; LINTNER, 1965) played a predominant role in financial economics In such a framework only linear risk allocations can occur. However, in 1973 BLACK and SCHOLES published their famous option pricing formula, which allows in particular for a replication of options by means of dynamic strategies. Options and their dynamic replication became increasingly popular. Nowadays, non linear investment strategies, such as portfolio insurance, tactical asset allocation and collars are widely used.

In actuarial science non linear risk allocations are a central issue in the reinsurance context. Already in 1960 Borch's theorem on Pareto efficient risk sharing was published. Later on, BUHLMANN (1984) proved the existence of a price density leading to a Pareto efficient risk allocation which is typically non linear. In BÜHLMANN (1980) and LIENHARD (1986) price densities were explicitly calculated for some special cases.

LELAND (1980) was the first who applied the actuarial results on non linear risk sharing in financial economics. By means of Borch's theorem he analysed the occurrence of portfolio insurance in the context of capital market equilibrium. MULLER (1990, 1991) applied Bühlmann's equilibrium model on the capital market and obtained some first results about the qualitative shape of risk allocations.

In this article total financial risk has to be allocated to n investors. Following the tradition of RUBINSTEIN (1976), BRENNAN (1979) and LELAND (1980) all investment decisions have to be made at one point of time. First, the main results of the theory of risk exchange are shortly summarized. Thereafter, different types of investment strategies such as portfolio insurance, tactical asset allocation and collars are explained in the context of risk exchange. It is shown how the type of investment strategy chosen by an individual investor depends on the risk tolerances of all investors and their sensitivity to wealth changes. Finally price equilibria are analysed in the special case where the Neumann-Morgenstern utility functions of all investors belong to the HARA class. Generalized versions of portfolio insurance, tactical asset allocation and collars are the only investment strategies which can occur.

2. THEORY OF RISK EXCHANGE

2.1. The model

As in RUBINSTEIN (1976), BRENNAN (1979) and LELAND (1980) trade takes only place at one point of time

There are n investors $i = 1, \dots, n$ with the following characteristics:

1) All investors have the same planning horizon and the same expectations. In particular, their expectations with respect to total financial wealth (aggregate market value of all financial assets in an economy) at the end of the planning period are given by a random variable $\tilde{W}$.
2) Moreover, the value of investor i's ($i = 1, \dots, n$) initial endowment at the end of the planning period can be represented by random variables $\tilde{X}_i$, s.t.

$$\tilde{X}_i \geq 0, \qquad \sum_{i=1}^{n} \tilde{X}_i = \tilde{W}, \text{ a.e.}^{1}.$$

3) Each investor $i = 1, \dots, n$ evaluates his wealth at the end of the planning period by a Neumann-Morgenstern utility function

$$u_i : R \to R, \qquad i = 1, \dots, n.$$

Hence, for the investors $i = 1, \dots, n$ with the initial risk allocation

$$(\tilde{X}_1, \dots, \tilde{X}_n)$$

[1] If initial endowments consist only of the market portfolio and a riskless asset, then

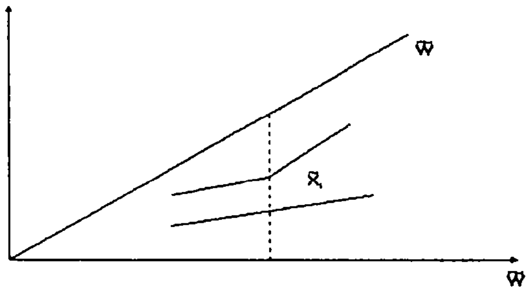
$$\tilde{X}_i = a_i + s_i \tilde{W}, \text{ a.e. } i = 1, \dots, n, \text{ with } \sum_{i=1}^{n} a_i = 0, \ \sum_{i=1}^{n} s_i = 1$$

holds

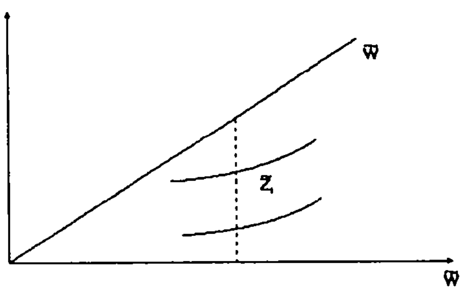
a reallocation of total financial risk

$$(\tilde{Z}_1, \ldots, \tilde{Z}_n) \qquad \text{with} \qquad \sum_{i=1}^{n} \tilde{Z}_i = \tilde{W}$$

has to be found.



Initial allocation of total financial wealth



Reallocation of total financial wealth

FIGURE 1

Obviously, this framework allows for the application of standard results in risk theory (e.g. BORCH (1960), BÜHLMANN (1980, 1984))

2.2. Theory of risk exchange: standard results

The following assumptions will be useful:

A.1. a) The random variable $\tilde{W}$ is given by the probability space $(R, \boldsymbol{B}, P)$, where $\boldsymbol{B}$ denotes the Borel-σ-algebra. There exist constants m, M with $0 < m < M < \infty$, such that

$$P[m \leq \tilde{W} \leq M] = 1 .$$

b) There exist measurable functions h_i, $i = 1, \ldots, n$ such that

$$\tilde{X}_i = h_i(\tilde{W}), \text{ a.e.}^2.$$

c) $E[\tilde{X}_i] > 0$.

A.2. The Neumann-Morgenstern utility functions u_i, $R \to R$, $i = 1, \ldots, n$ are twice differentiable and satisfy

$$u_i'(x) > 0, \quad u_i''(x) < 0 \quad \forall x.$$

A.3. The Neumann-Morgenstern utility functions u_i; $R \to R$, $i = 1, \ldots, n$ are three times continuously differentiable

Moreover, the following definitions are needed:

Definition 1: An n-tuple of random variables $(\tilde{Z}_1, \ldots, \tilde{Z}_n)$ is called a *feasible allocation* if it satisfies

$$\sum_{i=1}^{n} \tilde{Z}_i = \tilde{W}, \text{ a.e.}$$

Definition 2: A measurable function

$$\phi : [m, M] \to [0, \infty[$$

is called a *price density* if it satisfies

$$E[\phi(\tilde{W})] = 1.$$

Remark: Under a price density ϕ the value of a random variable $\tilde{Z}_i = f_i(\tilde{W})$ is given by

$$E[f_i(\tilde{W})\, \phi(\tilde{W})] = \int f_i(w)\, \phi(w)\, dP(w) \tag{1}$$

Definition 3: The tuple $\{\phi, (\tilde{Z}_1^*, \ldots, \tilde{Z}_n^*)\}$ consisting of a price density ϕ and a feasible allocation $(\tilde{Z}_1^*, \ldots, \tilde{Z}_n^*)$ is called a *price equilibrium* if for all $i = 1, \ldots, n$ $\tilde{Z}_i^*$ is the solution of

$$\max_{\tilde{Z}_i} E[u_i(\tilde{Z}_i)]$$

under

$$E[\tilde{Z}_i\, \phi(\tilde{W})] \leq E[\tilde{X}_i\, \phi(\tilde{W})].$$

[2] Assumption A 1 b is made for expository convenience As in BÜHLMANN (1984) one could define $\tilde{X}_1, \ldots, \tilde{X}_n$ as random variables on a common probability space (Ω, A, Π) Using the subsequent analysis one could show afterwards that $\tilde{W} = \sum_{i=1}^{n} \tilde{X}_i$ is a sufficient statistic for the problem under consideration

Definition 4: The feasible allocation $(\tilde{Z}_1^*, \ldots, \tilde{Z}_n^*)$ is called *Pareto efficient* if there exists no feasible allocation $(\tilde{Z}_1, \ldots, \tilde{Z}_n)$ satisfying

$$E[u_i(\tilde{Z}_i)] \geq E[u_i(\tilde{Z}_i^*)], \qquad i = 1, \ldots, n$$

with strict inequality for at least one $i \in \{1, \ldots, n\}$.

The standard results for this model can be summarized as follows (BÜHLMANN (1980, 1984)):

Theorem 1:

1) Under A.1., A 2., A.3.[3] there exists a price equilibrium $\{\phi, (\tilde{Z}_1^*, \ldots, \tilde{Z}_n^*)\}$.
2) Under A 1., A.2. each price equilibrium $\{\phi, (\tilde{Z}_1^*, \ldots, \tilde{Z}_n^*)\}$ has the following properties:
   a) The risk allocation $(\tilde{Z}_1^*, \ldots, \tilde{Z}_n^*)$ is Pareto efficient.
   b) There exist $\gamma_1, \ldots, \gamma_n \in (0, \infty)$ such that

$$u_i'(\tilde{Z}_n^*) = \gamma_i \phi(\tilde{W})^4, \text{ a e} \qquad i = 1, \ldots, n. \tag{2}$$

As an immediate consequence one obtains:

Corollary 1: Under A.1., A.2. for each price equilibrium $\{\phi, (\tilde{Z}_1^*, \ldots, \tilde{Z}_n^*)\}$ there exist measurable functions f_i such that

$$\tilde{Z}_i^* = f_i(\tilde{W}), \text{ a.e.}, \qquad i = 1, \ldots, n. \tag{3}$$

In the context of financial economics it is of particular interest to have some information about the shape of the functions $f_1, \ldots, f_n$ and ϕ. Some results on this issue are derived in the next section.

3. ANALYSIS OF PRICE EQUILIBRIA

3.1. Portfolio insurance, tactical asset allocation and collars

The term "portfolio insurance" is widely used for investment strategies where a reference portfolio is protected by a put option. Obviously such strategies lead to convex payoff functions. Therefore, LELAND (1980) introduced the term "general insurance policy" for convex payoff functions. In this article we use the following terminology:

Definition 5: An investment strategy leading to a twice differentiable payoff function

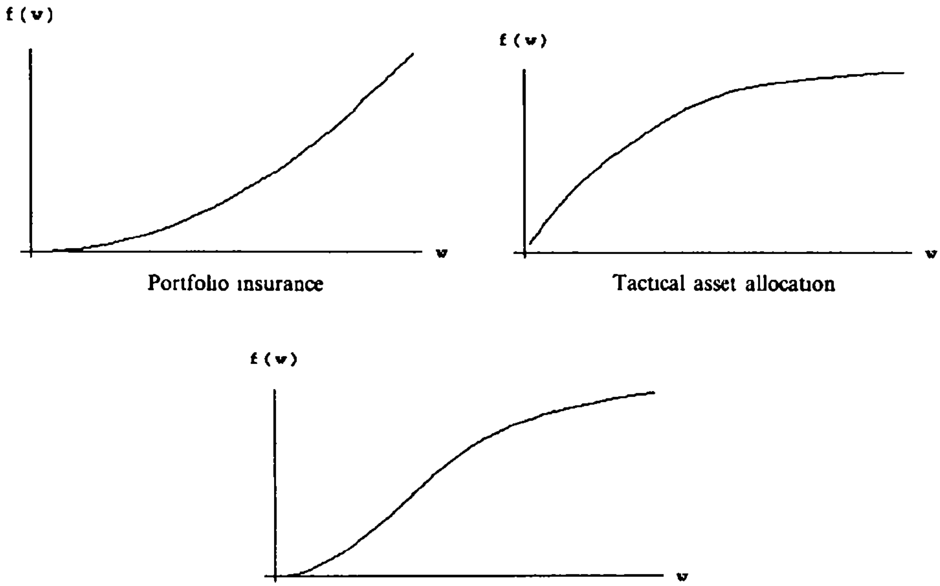
$$f: [m, M] \to R$$

[3] Instead of A3 BÜHLMANN (1984) assumes that the functions $\rho_i(x) = -\frac{u_i''(x)}{u_i'(x)}$ satisfy a Lipschitz condition In E Chevallier's forthcoming thesis assumptions A 1, A 2, A 3 will be relaxed

[4] See also BRENNAN/SOLANKI (1981)

is called

a) "Portfolio insurance" if $f''(w) > 0 \ \forall w \in [m, M]$.
b) "Tactical asset allocation" if $f''(w) < 0 \ w \in [m, M]$.
c) "Collar" if
$f''(w) > 0 \ \forall w \in [m, w_0)$ and
$f''(w) < 0 \ \forall w \in (w_0, M]$, where $m < w_0 < M$



Collar
FIGURE 2

Remarks:

1) Of course strategies with a continuous payoff function $f(w)$ can be approximated by buying and selling put and call options with different striking prices (see also LELAND (1980)).
2) The term "tactical asset allocation" is motivated by the widely used "buy low, sell high" strategies. The term "collar" is used for the popular investment policy where a reference portfolio is held, a put option is bought and a call option is sold.

3.2. Risk tolerance, cautiousness and properties of price equilibria

The following definition will be useful for the discussion of price equilibria.

Definition 6:

a) $\tau_i(x) = -\dfrac{u_i'(x)}{u_i''(x)}$ is called the *risk tolerance* of investor i.

b) $R_i(x) = \dfrac{d}{dx}\,\tau_i(x)$ is called the *cautiousness* of investor i[5].

Some well known characteristics of price equilibria can be formulated as follows:

Proposition 1: Under A.1., A.2. a price equilibrium $\{\phi, (f_1(\tilde{W}), \ldots, f_n(\tilde{W}))\}$ where ϕ and $f_1, \ldots, f_n$ are differentiable has the properties:

a) $\sum_{i=1}^{n} f_i'(w) = 1$,

b) $f_i'(w) = \dfrac{\tau_i(f_i(w))}{\sum_{j=1}^{n} \tau_j(f_j(w))} \in (0, 1)$,

c) $\phi(w) > 0, \quad \phi'(w) < 0$,

d) $\dfrac{\phi'(w)}{\phi(w)} = -\dfrac{1}{\sum_{j=1}^{n} \tau_j(f_j(w))}$.

Proof: e.g. BÜHLMANN (1984, p 16-17) or HUANG/LITZENBERGER (1986).

In order to decide whether a payoff function f_i corresponds to portfolio insurance, tactical asset allocation or a collar strategy its second derivative $f_i''(w)$ has to be known. The notion of a "representative investor" will considerably simplify the analysis of $f_i''(w)$.

Definition 7: Given a price equilibrium $\{\phi, (f_1(\tilde{W}), \ldots, f_n(\tilde{W}))\}$ a function v_m with

$$v_m'(w) = \phi(w)$$

is called Neumann-Morgenstern utility function of the representative investor.

Remark: The representative investor is a fictious individual representing the market. Under the conditions of Proposition 1 and differentiability assumptions the risk tolerance $\tau_m(w)$ and the cautiousness $R_m(w)$ of the representative

[5] Hence, the cautiousness R_i is a measure for the sensitivity of the risk tolerance $\tau_i(x)$ with respect to wealth changes

investor are given by

$$\tau_m(w) = -\frac{v'_m(w)}{v''_m(w)} = -\frac{\phi(w)}{\phi'(w)} = \Sigma_{j=1}^{n} \tau_j(f_j(w)), \tag{4}$$

$$R_m(w) = \frac{d}{dw}\tau_m(w) = \frac{\Sigma_{j=1}^{n} R_j(f_j(w))\,\tau_j(f_j(w))}{\Sigma_{j=1}^{n} \tau_j(f_j(w))}. \tag{5}$$

Hence, the risk tolerance $\tau_m(w)$ is the sum of individual risk tolerances, whereas the cautiousness $R_m(w)$ is a weighted mean of individual cautiousness terms.

Now the result on the second derivatives of the payoff functions f_i and the price density ϕ can be formulated as follows:

Theorem 2: Under A 1., A.2., A.3. a price equilibrium $\{\phi, (f_1(\tilde{W}), \ldots, f_n(\tilde{W}))\}$ where ϕ and $f_1, \ldots, f_n$, are twice differentiable has the properties:

a) $$\frac{f_i''(w)}{f_i'(w)} = \frac{1}{\tau_m(w)}\{R_i(f_i(w)) - R_m(w)\}, \qquad i = 1, \ldots, n,$$

b) $$\frac{\phi''(w)}{\phi'(w)} = -\frac{1}{\tau_m(w)}\{1 + R_m(w)\},$$

c) $$\frac{d}{dw}\ln\left(\frac{f_i'(w)}{f_j'(w)}\right) = \frac{R_i(f_i(w)) - R_j(f_j(w))}{\tau_m(w)}, \qquad i, j = 1, \ldots, n$$

Comments:

1) In particular Proposition 1 and Theorem 2 contain the key result

$$f_i'(w) = \frac{\tau_i(f_i(w))}{\tau_m(w)}, \qquad i = 1, \ldots, n, \tag{6}$$

$$\operatorname{sign}\{f_i''(w)\} = \operatorname{sign}\{R_i(f_i(w)) - R_m(w)\}, \qquad i = 1, \ldots, n \tag{7}$$

In other words, the slope of the payoff function f_i is given by the ratio of the risk tolerances $\tau_i(f_i(w))$ and $\tau_m(w)$, whereas the curvature of f_i is related to the difference of the cautiousness terms $R_i(f_i(w))$ and $R_m(w)$.

2) Theorem 2.a) leads to the following criteria

a) An investor $i \in \{1, \ldots, n\}$ chooses *portfolio insurance* if and only if

$$R_i(f_i(w)) > R_m(w) \qquad \forall w \tag{8}$$ [6]

b) An investor $i \in \{1, \ldots, n\}$ chooses *tactical asset allocation* if and only if

$$R_i(f_i(w)) < R_m(w) \qquad \forall w. \tag{9}$$

[6] LELAND (1980) derived a similar result in a less formal context

c) If an investor $i \in \{1, \ldots, n\}$ chooses a *collar strategy* there exists $w_0 \in (m, M)$ such that
   1) $R_i(f_i(w_0)) = R_m(w_0)$,
   2) $R_i(f_i(w)) - R_m(w)$ is strictly decreasing in w_0.

3) An easy calculation shows that under A.2., A 3. $u_i'''(f_i(w)) > 0$, $i = 1, \ldots, n$ implies $R_m(w) > -1$.

Therefore, one concludes from Theorem 2b):

$$\phi''(w) > 0 \text{ if } u_i'''(f_i(w)) > 0,\ i = 1, \ldots, n. \tag{10}$$

Proof of Theorem 2:

a) Differentiation of the formula in Proposition 1b)

$$f_i'(w) = \frac{\tau_i(f_i(w))}{\tau_m(w)}$$

leads to

$$f_i''(w) = \frac{R_i(f_i(w))}{\tau_m(w)} f_i'(w) - \frac{\tau_i(f_i(w))\, R_m(w)}{\tau_m^2(w)} \tag{11}$$

or

$$\frac{f_i''(w)}{f_i'(w)} = \frac{1}{\tau_m(w)} \{R_i(f_i(W)) - R_m(w)\}. \tag{12}$$

b) Differentiation of the formula in Proposition 1d)

$$\frac{\phi(w)}{\phi'(w)} = -\tau_m(w)$$

leads to

$$\frac{\phi'^2(w) - \phi(w)\,\phi''(w)}{\phi'^2(w)} = -R_m(w) \tag{13}$$

or

$$1 + R_m(w) = -\tau_m(w) \frac{\phi''(w)}{\phi'(w)}. \tag{14}$$

c) From Proposition 1b) one obtains

$$\ln\left(\frac{f_i'(w)}{f_j'(w)}\right) = \ln(\tau_i(f_i(w))) - \ln(\tau_j(f_j(w))) \tag{15}$$

and

$$\frac{d}{dw} \ln\left(\frac{f_i'(w)}{f_j'(w)}\right) = \frac{R_i(f_i(w))}{\tau_i(f_i(w))} \frac{\tau_i(f_i(w))}{\tau_m(w)} - \frac{R_j(f_j(w))}{\tau_j(f_j(w))} \frac{\tau_j(f_j(w))}{\tau_m(w)}. \tag{16}$$

4. ANALYSIS OF PRICE EQUILIBRIA FOR THE HARA CLASS

In Section 3 general properties of price equilibria were derived. Now we look at the special case where the risk tolerance functions $\tau_i(x)$ are linear. Assumption A.2. and A.3. are replaced by the assumption:

A.4. The Neumann-Morgenstern utility functions are increasing, concave and satisfy

a) $\tau_i(x) = a_i + R_i x > 0$, with $R_i \geq 0$, $i = 1, \ldots, n$,

b) Not all R_j identical.

Remarks:

1) Assumption A.4. allows for all Neumann-Morgenstern utility functions which belong to the HARA class and have a non negative cautiousness[7].
2) In the case where all R_j are identical the risk allocation is linear and a detailed analysis can be found in BÜHLMANN (1980) and LIENHARD (1986).
3) For $R_i > 0$ the Neumann-Morgenstern utility function u_i is only defined on the interval $\left(-\frac{a_i}{R_i}, \infty\right)$. Therefore, assumption A.2 is not satisfied and existence of a price equilibrium is not guaranteed by Theorem 1 However, it can be easily verified that Proposition 1 and Theorem 2 are still valid if Assumption A 2 and A 3. are replaced by A.4.

By restricting the analysis to the HARA class one obtains:

Lemma 1: Under A.1., A.4. a price equilibrium $\{\phi, (f_1(\tilde{W}), \ldots, f_n(\tilde{W}))\}$ where ϕ and $f_1, \ldots, f_n$ are differentiable[8] has the property:

$R_m(w)$ is strictly increasing.

Proof: A.4 and (5) lead to

$$(17) \qquad R_m(w)\,\tau_m(w) = \sum_{j=1}^{n} R_j \tau_j(f_j(w)),$$

$$(18) \qquad R_m'(w)\,\tau_m(w) + R_m^2(w) = \sum_{j=1}^{n} R_j^2 f_j'(w)^9.$$

[7] A negative cautiousness would lead to problems with satiation and an unrealistic investment behaviour (see ARROW (1965))

[8] If $\phi, f_1, \ldots, f_n$ are differentiable, then due to Proposition 1b), 1d) and A 4 they are also twice differentiable

[9] From the derivation of formula (18) it becomes obvious that Lemma 1 depends crucially on the assumption that each investor i has a constant cautiousness R_i (A 4 a)

Moreover, Proposition 1b), (5) and (18) imply

$$(19) \qquad R'_m(w) = \frac{\sum_{j=1}^{n} R_j^2 \tau_j(f_j(w))}{\tau_m^2(w)} - \frac{R_m(w) \sum_{j=1}^{n} R_j \tau_j(f_j(w))}{\tau_m^2(w)}$$

and the strict monotonicity of $R_m(w)$ follows from

$$\begin{aligned}
&\sum_{j=1}^{n} (R_j - R_m(w))\, R_j \tau_j(f_j(w)) \\
&\quad = \sum_{R_j > R_m(w)} (R_j - R_m(w))\, R_j \tau_j(f_j(w)) + \\
&\quad + \sum_{R_j < R_m(w)} (R_j - R_m(w))\, R_j \tau_j(f_j(w)) > \\
&\quad > \sum_{j=1}^{n} (R_j - R_m(w))\, R_m(w)\, \tau_j(f_j(w)) = 0.
\end{aligned}$$

Lemma 1 leads to the main result of this section

Theorem 3: Under A.1., A.4. a price equilibrium $\{\phi, (f_1(\tilde{W}), \ldots, f_n(\tilde{W}))\}$ where ϕ and $f_1, \ldots, f_n$ are differentiable has the properties:

a) The only investment strategies chosen by investors $i = 1, \ldots, n$ are
  - portfolio insurance,
  - tactical asset allocation,
  - collar strategy.

b) Investors $i \in \{1, \ldots, n\}$ with $R_i = \max_{j=1,\ldots,n} R_j$ choose portfolio insurance[10]

c) Investors $i \in \{1, \ldots, n\}$ with $R_i = \min_{j=1,\ldots,n} R_j$ choose tactical asset allocation[10]

Proof: Formula (5) implies

$$(20) \qquad \min_{j=1,\ldots,n} R_j < R_m(w) < \max_{j=1,\ldots,n} R_j \qquad \forall w \in [m, M].$$

Now, a), b) and c) follow immediately from Theorem 2a) and Lemma 1.
Some additional information about price equilibria in the HARA case is provided by the next result.

[10] See also MÜLLER (1990)

Proposition 2: Under the assumption of Theorem 3 one obtains

a) $$f_i'(w) = \frac{a_i + R_i\, f_i(w)}{a + \sum_{j=1}^{n} R_j f_j(w)}, \quad \text{where} \quad a = \sum_{j=1}^{n} a_j,$$

b) $$\frac{d}{dw} \ln \left(\frac{f_i'(w)}{f_j'(w)} \right) = \frac{R_i - R_j}{\tau_m(w)}.$$

Proof: Special case of Proposition 1b) and Theorem 2c).

Comments:

1) In particular Proposition 2b) implies

$$\frac{d}{dw} \left(\frac{f_i'(w)}{f_j'(w)} \right) \gtreqless 0 \Leftrightarrow R_i \gtreqless R_j \qquad i, j = 1, \ldots, n. \tag{21}$$

2) For sufficiently large values of w one can show $f_i(w) > 0$ for $i = 1, \ldots, n$ and Proposition 2 leads to the following inequalities

a') $$f_i'(w) > \frac{a_i + R_i f_i(m)}{a + w\ R_{\max}} \quad \text{with} \quad a = \sum_{j=1}^{n} a_j, \quad R_{\max} = \max_{j=1,\ldots,n} R_j,$$

b') $$\frac{d}{dw} \ln \left(\frac{f_i'(w)}{f_j'(w)} \right) > \frac{R_i - R_j}{a + w \cdot R_{\max}}, \quad \text{if} \quad R_i > R_j.$$

Finally, an example illustrates some typical properties of a price equilibrium in the HARA case.

Example:

— The random variable $\tilde{W}$ representing total financial wealth is uniformly distributed over [0.3, 20].

— There are $n = 3$ investors with risk tolerance functions

$$\tau_1(x) = 20 \cdot x,$$
$$\tau_2(x) = 2.5 \cdot x,$$
$$\tau_3(x) = x$$

and an initial risk allocation

$$(\tilde{X}_1, \tilde{X}_2, \tilde{X}_3) = (0.16 \cdot \tilde{W}, 0.35 \cdot \tilde{W}, 0.49 \cdot \tilde{W}).$$

The price equilibrium $\{\phi, (f_1(W), f_2(\tilde{W}), f_3(\tilde{W}))\}$ is illustrated in Figure 3.

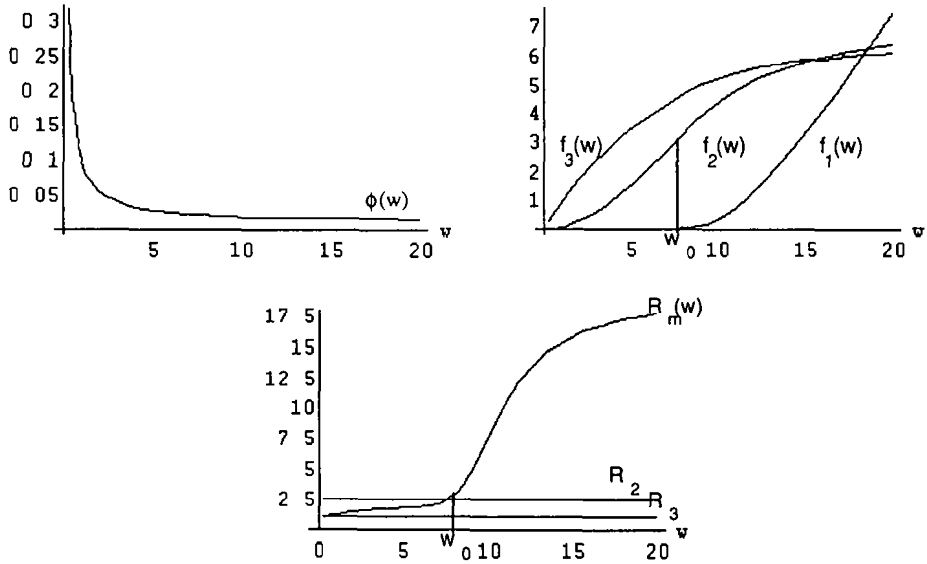


FIGURE 3

5. CONCLUSIONS

In this article the theory of risk exchange was applied to the allocation of financial risk Special emphasis was put on the shape of the payoff functions in price equilibrium. Under general conditions the role of risk tolerance and cautiousness was analysed. The notion of a representative investor was very useful for the interpretation of the results. Finally, in the HARA case a full characterization of all equilibrium payoff functions was possible.

ACKNOWLEDGEMENT

We are indebted to two anonymous referees for considerably improving our article.

REFERENCES

ARROW, K J (1965) *Aspects in the Theory of Risk Bearing* Helsinki

BLACK, F and SCHOLES, M (1973) The Pricing of Options and Corporate Liabilities *Journal of Political Economy* **81**, 637–659

BORCH, K (1960) The Safety loading of Reinsurance Premiums *Skandinavisk Aktuarietidskrift* **43**, 163–184

BRENNAN, M J (1979) The Pricing of Contingent Claims in Discrete Time Models. *The Journal of Finance* **XXXIV**, No 1, 53–68

BRENNAN, M J and SOLANKI, R (1981) Optimal Portfolio Insurance *Journal of Financial and Quantitative Analysis* **Vol. XVI**, No 3, September 1981

BÜHLMANN, H (1980) An Economic Premium Principle *ASTIN Bulletin* **11** (1), 52–60

BÜHLMANN, H (1984) The General Economic Premium Principle *ASTIN Bulletin* **14** (1), 13–21

HUANG, C and LITZENBERGER, R H (1988) *Foundations for Financial Economics* North Holland, New York, Amsterdam, London

LELAND, H E (1980) Who should buy portfolio insurance? *Journal of Finance* **XXXV**, No 2, 581–596

LINTNER, J, (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets *Review of Economics and Statistics*, February 1965, 13–37

LIENHARD, M (1986) Calculation of Price Equilibria for Utility Functions of the HARA Class *ASTIN Bulletin* **16**, 91–97

MARKOWITZ, H (1952) Portfolio selection. *Journal of Finance* **7**, 77–91

MÜLLER, H (1990) Price Equilibria and non linear Risk Allocations in Capital Markets *Mitteilungen der Schweizerischen Vereinigung der Versicherungsmathematiker*, Heft 1 1990, 115–128

MÜLLER, H (1991) Einsatz von Optionen Effizienzverbesserung durch nichtlineare Risikoallokationen *Schweiz Zeitschrift fur Volkswirtschaft und Statistik*, 1991, **Vol. 127** (3), 379–393

RUBINSTEIN, M (1976) The Valuation of Uncertain Income Streams and the Pricing of Option *Bell Journal of Economics and Management Science* **7**, 407–425

SHARPE, W F (1964) Capital asset prices a theory of market equilibrium under conditions of risk *Journal of Finance*, September 1964, 425–442

ERIC CHEVALLIER
Institute for Operations Research
University of Zürich
Moussonstr. 15
CH-8044 Zürich, Switzerland

HEINZ H. MULLER
IEW
University of Zürich
Blumlisalpstr. 10
CH-8006 Zürich, Switzerland

A RECURSIVE PROCEDURE FOR CALCULATION OF SOME COMPOUND DISTRIBUTIONS

BY OLE HESSELAGER

University of Copenhagen

ABSTRACT

We consider compound distributions where the counting distribution has the property that the ratio between successive probabilities may be written as the ratio of two polynominals. We derive a recursive algorithm for the compound distribution, which is more efficient than the one suggested by PANJER & WILLMOT (1982) and WILLMOT & PANJER (1987). We also derive a recursive algorithm for the moments of the compound distribution. Finally, we present an application of the recursion to the problem of calculating the probability of ruin in a particular mixed Poisson process.

KEYWORDS

Recursions; compound distributions; moments; probablity of ruin.

1. INTRODUCTION

Let

$$X = \sum_{i=1}^{N} Y_i$$

denote the aggregate claims amount where $X = 0$ if $N = 0$. It is assumed that the severities $Y_1, Y_2, \ldots$ are mutually independent and distributed on the non-negative integers with common probability function

$$f_y = \mathrm{P}(Y_i = y), \qquad y = 0, 1, \ldots \tag{1.1}$$

It is further assumed that N is stochastically independent of $Y_1, Y_2, \ldots$ with probability function

$$p_n = \mathrm{P}(N = n), \qquad n = 0, 1, \ldots$$

The compound distribution

$$g_x = \sum_{n=0}^{\infty} p_n f_x^{*n}, \tag{1.2}$$

where f^{*n} denotes the n-th convolution of f, can sometimes be calculated recursively. PANJER (1981) derived his by now famous recursive formula for the case where the counting probabilities p_n satisfy the recursive relation

$$p_n = \frac{an+b}{n} p_{n-1}, \qquad n = 1, 2, .. \tag{1.3}$$

SUNDT & JEWELL (1981) showed that (1.3) is satisfied by the Poisson, the binomial, and the negative binomial distributions, and no other. PANJER & WILLMOT (1982) went on to consider the class of counting distributions which satisfy a recursion

$$p_n = \frac{\sum_{i=0}^{k} a_i n^i}{\sum_{i=0}^{k} b_i n^i} p_{n-1}, \qquad n = 1, 2, \ldots, \tag{1.4}$$

for some k, and derived recursions for the compound distribution when $k = 1$ and $k = 2$. These recursions were further developed by WILLMOT & PANJER (1987). Recursions for a different extension of the class (1 3) can be found in SCHRÖTER (1990) and SUNDT (1992).

In the case of arbitrary k, it is clearly not possible to give a complete characterization of the class (1 4). ORD (1967) characterizes those distributions which satisfy a difference equation analogous to Pearson's differential equation, and also derives a recursive relation for the (factorial) moments. Also GULDBERG (1931) considered recursive calculation of moments for certain members of the class (1.4).

Important distributions satisfying (1.4), which are not already covered by (1 3), are the hypergeometric distribution ($k = 2$), the Polya-Eggenberger distribution ($k = 2$), the Waring distribution ($k = 1$), and the generalized Waring distribution ($k = 2$).

Note that the coefficients a_i and b_i appearing in (1.4) are only specified up to a multiplicative constant.

In this paper we consider the class (1.4) and derive a new recursion for the compound distribution (Section 2) The derivation is elementary, and is valid for arbitrary k. In Section 3 we derive a recursion for the moments of the compound distribution. In Section 4 the proposed recursive formula for the compound distribution is compared to that of WILLMOT & PANJER (1987) for $k = 1$ and $k = 2$, and is found to be more efficient. In Section 5 we present an application of the recursion to problem of calculating the probability of eventual ruin in a (particular) mixed Poisson process.

2. RECURSION FOR THE COMPOUND DISTRIBUTION

Assume that p_n satisfies (1.4) For $i = 0, \ldots, k$ we define the auxiliary functions

$$g_{i,x} = \sum_{n=0}^{\infty} n^i p_n f_x^{*n}, \qquad x = 0, 1, \ldots, \tag{2.1}$$

and note in particular that $g_{0,x}$ is the compound distribution (1.2). Let

$$\boldsymbol{g}_x = (g_{0,x}, \ldots, g_{k,x})', \tag{2.2}$$

and let m denote the smallest integer for which $f_m > 0$. Thus, $f_y = 0$ for $y = 0, \ldots, m-1$. The following result gives a recursion for the vector $\boldsymbol{g}_x$, and hence the compound distribution $g_{0,x}$.

Theorem 1: Assume that (1.4) holds true. With initial values

$$g_{i,0} = \sum_{n=0}^{\infty} p_n n^i f_0^n, \qquad i = 0, \ldots, k, \tag{2.3}$$

$$g_{i,x} = 0, \qquad i = 0, \ldots, k, \qquad x = 1, \ldots, m-1,$$

the compound distribution $g_x = g_{0,x}$ may be obtained by calculating $\boldsymbol{g}_x$ recursively as

$$\boldsymbol{g}_x = \boldsymbol{T}_x^{-1} \boldsymbol{t}_x, \qquad x \geq m \vee 1,$$

where

$$\boldsymbol{T}_x = \begin{pmatrix} 1 & -m/x & 0 & \cdots & 0 \\ 0 & 1 & -m/x & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 & -m/x \\ (b_0 - f_0 c_0) & (b_1 - f_0 c_1) & \cdots & (b_{k-1} - f_0 c_{k-1}) & (b_k - f_0 c_k) \end{pmatrix}, \tag{2.4}$$

and $\boldsymbol{t}_x = (t_{0,x}, \ldots, t_{k,x})'$ is given by

$$t_{i,x} = \frac{1}{f_m} \sum_{y=1}^{x} f_{m+y} \left\{ -\frac{m+y}{x} g_{i+1,x-y} + \frac{y-x}{x} g_{i,x-y} \right\}, \qquad i < k, \tag{2.5}$$

$$t_{k,x} = \sum_{y=m \vee 1}^{x} f_y \sum_{i=0}^{k} c_i g_{i,x-y}, \tag{2.6}$$

with $c_i = \sum_{j=i}^{k} \binom{j}{i} a_j$.

Remark 2.1. Note that T_x does not depend on the values of $g_{i,z}$, and that t_x can be calculated when g_z is known for all $z < x$. □

Proof: The expression (2.3) is obtained from the definition (2 1) of $g_{i,x}$ by noting that $f_0^{*n} = f_0^n$. Also the fact that $g_{i,x} = 0$ for $x = 1, \ldots, m-1$ is an immediate consequence of (2.1) since $f_x^{*n} = 0$ for $x = 1, \ldots, m-1$.

From De Pril (1985) we have the identity,

$$0 = \sum_{y=0}^{x} \left[(n+1)\frac{y}{x} - 1 \right] f_y f_{x-y}^{*n} \tag{2.7}$$

Multiplying (2.7) by $p_n n^i$ and summing over $n \geq 0$ yields

$$0 = \sum_{y=0}^{x} f_y \left\{ \frac{y}{x} g_{i+1,x-y} + \left(\frac{y}{x} - 1 \right) g_{i,x-y} \right\}. \tag{2.8}$$

By omitting terms corresponding to $y = 0, \ldots, m-1$ from the summation and substituting $x = x - m$, we obtain after a little rearrangement that

$$g_{i,x} - \frac{m}{x} g_{i+1,x} = t_{i,x}, \tag{2.9}$$

where $t_{i,x}$ is given by (2.5). From assumption (1.4) we obtain for $n \geq 1$ that

$$p_n \sum_{i=0}^{k} b_i n^i = p_{n-1} \sum_{i=0}^{k} a_i n^i = p_{n-1} \sum_{i=0}^{k} c_i (n-1)^i, \tag{2.10}$$

where

$$c_i = \sum_{j=i}^{k} \binom{j}{i} a_j. \tag{2.11}$$

Multiplying (2 10) by $f_x^{*n} = \Sigma_{y=0}^{x} f_y f_{x-y}^{*(n-1)}$ and summing over $n \geq 1$ yields for $x \geq 1$ the relation

$$\sum_{i=0}^{k} b_i g_{i,x} = \sum_{i=0}^{k} \sum_{y=0}^{x} f_y c_i g_{i,x-y}, \qquad x \geq 1. \tag{2.12}$$

By isolating terms involving $g_{i,x}$ on the left-hand side, we rewrite (2 12) as

$$\sum_{i=0}^{k} (b_i - f_0 c_i) g_{i,x} = t_{k,x}, \qquad x \geq 1, \tag{2.13}$$

where $t_{k,x}$ is given by (2.6). The linear equations $T_x g_x = t_x$, with T_x given by (2.4), now follow from (2.9) for $i = 0, \ldots, k-1$ and (2 13) QED

Remark 2.2. It is useful to consider separately the to cases where $m > 0$ ($f_0 = 0$) and $m = 0$ ($f_0 > 0$).

$m > 0$ When $f_0 = 0$ we note from (2.3) that $g_{0,0} = p_0$ and $g_{i,0} = 0$ for $i \geq 1$ Note also that the terms $f_0 c_i$ in the last row of T_x disappear in this case.

$m = 0$. The linear equations $T_x g_x = t_x$ are easily solved analytically in this case, and we obtain that

$$(2.14)\qquad g_{i,x} = \frac{1}{f_0} \sum_{y=1}^{x} f_y \left\{ \frac{y}{x} g_{i+1,x-y} + \frac{y-x}{x} g_{i,x-y} \right\}, \qquad i < k,$$

$$(2.15)\qquad g_{k,x} = \frac{1}{b_k - f_0 c_k} \left\{ \sum_{y=1}^{x} f_y \sum_{i=0}^{k} c_i g_{i,x-y} + \sum_{i=0}^{k-1} (f_0 c_i - b_i) g_{i,x} \right\}.$$

The initial values $g_{i,0}$ may be expressed in terms of the derivatives $\varphi^{(j)}(f_0)$, $j = 0, \ldots, k$, where $\varphi(\cdot)$ denotes the probability generating function of the counting distribution However, for the class (1.4) of counting distributions, there is in general no simple expression for $\varphi(\cdot)$.

□

Example 1: The Waring distribution arises as a mixed geometric distribution with a beta mixing function. If $\mathrm{P}(N = n|p) = (1-p)p^n$, and $p \sim$ Beta (α, β), then

$$p_n = \frac{\mathrm{B}(\alpha+n, \beta+1)}{\mathrm{B}(\alpha, \beta)},$$

and

$$p_n = \frac{n+\alpha-1}{n+\alpha+\beta} p_{n-1}.$$

This corresponds to (1.4) with $k = 1$ and

$$\begin{aligned} a_0 &= \alpha - 1 & a_1 &= 1 \\ b_0 &= \alpha + \beta & b_1 &= 1, \\ c_0 &= \alpha & c_1 &= 1 \end{aligned}$$

where c_i is obtained from (2 11). □

Example 2: For the hypergeometric distribution with parameters (s, D, S),

$$p_n = \frac{\binom{D}{n}\binom{S-D}{s-n}}{\binom{S}{s}},$$

it holds that

$$p_n = \frac{[n-(D+1)]\,[n-(s+1)]}{n[n+(S-D-s)]}\, p_{n-1},$$

which corresponds to (1.4) with $k = 2$ and

$$\begin{array}{lll} a_0 = (D+1)\,(s+1) & a_1 = -(D+s+2) & a_2 = 1 \\ b_0 = 0 & b_1 = S-D-s & b_2 = 1\,. \\ c_0 = Ds & c_1 = -(D+s) & c_2 = 1 \end{array}$$

□

Example 3: The Polya-Eggenberger (Negative Hypergeometric) distribution arises as a mixed binomial distribution with a beta mixing function. The probability function

$$p_n = \frac{\begin{pmatrix} \alpha+n-1 \\ n \end{pmatrix} \begin{pmatrix} \beta+M-n-1 \\ M-n \end{pmatrix}}{\begin{pmatrix} \alpha+\beta+M-1 \\ M \end{pmatrix}},$$

satisfies

$$p_n = \frac{[n-(M+1)]\,[n+(\alpha-1)]}{n[n-(M+\beta)]}\, p_{n-1},$$

which corresponds to (1 4) with $k = 2$ and

$$\begin{array}{lll} a_0 = -(M+1)\,(\alpha-1) & a_1 = -(M-\alpha+2) & a_2 = 1 \\ b_0 = 0 & b_1 = -(M+\beta) & b_2 = 1\,. \\ c_0 = -M\alpha & c_1 = (\alpha-M) & c_2 = 1 \end{array}$$

□

Example 4: The generalized Waring distribution arises as a mixed negative binomial distribution with a beta mixing function,

$$p_n = \frac{\Gamma(c+n)}{\Gamma(c)\,n!} \; \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} \; \frac{\Gamma(\alpha+n)\,\Gamma(\beta+c)}{\Gamma(\alpha+\beta+c+n)},$$

and

$$p_n = \frac{[n+(c-1)]\,[n+(\alpha-1)]}{n[n+(\alpha+\beta+c-1)]}\, p_{n-1}$$

This corresponds to (1.4) with $k = 2$ and

$$\begin{array}{lll} a_0 = (\alpha-1)(c-1) & a_1 = (\alpha+c-2) & a_2 = 1 \\ b_0 = 0 & b_1 = \alpha+\beta+c-1 & b_2 = 1\,. \\ c_0 = \alpha c & c_1 = \alpha+c & c_2 = 1 \end{array}$$

□

3. RECURSIVE CALCULATION OF MOMENTS

For the class (1 3) of counting distributions it was pointed out by DE PRIL (1986) that also the moments $m_s = \mathrm{E}X^s$, $s = 0, 1, \ldots,$ of the compound distribution can be calculated recursively in a simple manner. Expressions for the moments m_s are useful if one wants to calculate the NP- or Edgeworth approximation to the compound distribution as an alternative to the (exact) recursive method.

Let

$$\mu_s = \mathrm{E}Y_1^s$$

denote the s'th moment around the origin of the severity distribution, and define

$$m_{i,s} = \sum_{x=0}^{\infty} x^s g_{i,x}, \tag{3.1}$$

where $g_{i,x}$ is the auxiliary function (2.1). Note in particular that $m_s = m_{0,s}$ is the s'th moment of the compound distribution. The following result gives a recursion for the vector $(m_{0,s}, \ldots, m_{k,s})$, $s = 0, 1, \ldots,$ and hence the moments m_s

Theorem 2: Assume that (1 4) holds true. With initial values

$$m_{i,0} = \mathrm{E}N^i = \sum_{n=0}^{\infty} p_n n^i, \qquad i = 0, \ldots, k, \tag{3.2}$$

the moments $m_s = m_{0,s}$ of the compound distribution may be obtained by calculating $(m_{0,s}, \ldots, m_{k,s})$ recursively for $s = 1, 2, \ldots$ as

$$m_{i,s} = \sum_{j=0}^{s-1} \binom{s-1}{j} \mu_{s-j} m_{i+1,j} - \sum_{j=1}^{s-1} \binom{s-1}{j-1} \mu_{s-j} m_{i,j}, \quad i < k, \tag{3.3}$$

$$m_{k,s} = \frac{1}{b_k - c_k} \left[\sum_{i=0}^{k} c_i \sum_{j=0}^{s-1} \binom{s}{j} \mu_{s-j} m_{i,j} + \sum_{i=0}^{k-1} (c_i - b_i) m_{i,s} \right] \tag{3.4}$$

Remark 3.1. When $(m_{0,u}, \ldots, m_{k,u})$ is known for $u < s$, one calculates $m_{i,s}$ for $i = 0, \ldots, k-1$ from (3.3) and then $m_{k,s}$ from (3.4). □

Proof: According to (3.1) and (2.1) the initial values are given by

$$m_{i,0}=\sum_{x=0}^{\infty} g_{i,x}=\sum_{x=0}^{\infty}\sum_{n=0}^{\infty} n^i p_n f_x^{*n}=\sum_{n=0}^{\infty} p_n n^i .$$

To verify (3.3) we multiply (2.8) by x^s, $s \geq 1$, and sum over $x \geq 0$ to obtain

$$0=\sum_{x=0}^{\infty}\sum_{y=0}^{x} f_y\{y x^{s-1} g_{i+1,x-y}+(y x^{s-1}-x^s)\, g_{i,x-y}\}.$$

By changing the order of summation and using the binomial formula

$$x^s=\sum_{j=0}^{s}\binom{s}{j} y^{s-j}(x-y)^j, \tag{3 5}$$

(and the similar expression for x^{s-1}) it follows that

$$\begin{aligned}
0&=\sum_{y=0}^{\infty} f_y\left[y\sum_{j=0}^{s-1}\binom{s-1}{j} y^{s-1-j}(m_{i+1,j}+m_{i,j})-\sum_{j=0}^{s}\binom{s}{j} y^{s-j} m_{i,j}\right]\\
&=\sum_{j=0}^{s-1}\binom{s-1}{j}\mu_{s-j}(m_{i+1,j}+m_{i,j})-\sum_{j=0}^{s}\binom{s}{j}\mu_{s-j} m_{i,j}.
\end{aligned}$$

Equation (3.3) now follows by extracting the term corresponding to $j=s$ from the last sum and making use of the fact that

$$\binom{s}{j}=\binom{s-1}{j}+\binom{s-1}{j-1}.$$

To verify (3.4), multiply (2.12) by x^s, $s \geq 1$, and sum over $x \geq 0$ to obtain

$$\sum_{i=0}^{k} b_i m_{i,s}=\sum_{i=0}^{k}\sum_{x=0}^{\infty}\sum_{y=0}^{x} x^s f_y g_{i,x-y} c_i .$$

By changing the order of summation and using (3 5), it follows that

$$\begin{aligned}
\sum_{i=0}^{k} b_i m_{i,s}&=\sum_{i=0}^{k}\sum_{y=0}^{\infty}\sum_{j=0}^{s}\binom{s}{j} f_y c_i y^{s-j} m_{i,j}\\
&=\sum_{i=0}^{k} c_i\sum_{j=0}^{s}\binom{s}{j}\mu_{s-j} m_{i,j},
\end{aligned} \tag{3.6}$$

and (3.4) follows by solving (3.6) for $m_{k,s}$. QED

4. COMPARISON WITH THE RECURSIONS OF WILLMOT & PANJER (1987)

In PANJER & WILLMOT (1982) it is demonstrated how recursions for the compound distribution may be obtained by use of generating functions; in principle for arbitrary k when the counting distribution satisfies (1.4). Formulas for the cases $k = 1$ and $k = 2$ are found in WILLMOT & PANJER (1987). We cite the following recursive procedure:

Define the auxiliary function

$$q_0 = m, \tag{4.1}$$

$$q_x = \frac{(x+m) f_{x+m}}{f_m} - \sum_{y=1}^{x} \frac{f_{y+m}}{f_m} q_{x-y},$$

where m is the smallest integer such that $f_m > 0$, and also

$$t_0 = r-1,$$

$$t_x = \frac{(x+r)(x+r-1) f_{x+r}}{r f_r} - \sum_{y=1}^{x} \frac{(y+r) f_{y+r}}{r f_r} t_{x-y}, \tag{4.2}$$

where r is the smallest integer such that $r f_r > 0$.

For $k = 1$ the class (1.4) may be rewritten as

$$p_n = \frac{\beta(n-1)+\kappa}{\alpha n+1} p_{n-1}, \qquad n = 1, 2, \ldots,$$

and the compound distribution g_x satisfies the recursion

$$g_x = \frac{p_0 q_x + \sum_{y=1}^{x} [(\beta(y-x)+\kappa y) f_y - q_y] g_{x-y}}{x(\alpha - \beta f_0) + q_0}. \tag{4.3}$$

For $k = 2$ and $b_0 = 0$ we may rewrite (1.4) as,

$$p_n = \frac{\beta(n-1)(n-2)+\kappa(n-1)+\delta}{n(n-1)+\alpha n} p_{n-1}, \qquad n = 1, 2, \ldots$$

Define a new set of auxiliary functions,

$$u_x = \sum_{y=0}^{x} y q_{x-y} f_y, \qquad v_x = \sum_{y=0}^{x} t_{x-y} f_y, \tag{4.4}$$

and g_x can be calculated recursively as

$$g_x = \frac{\sum_{y=1}^{x} g_{x-y} k_{x,y}}{x[(x-t_0-1)(1-\beta f_0)+\alpha q_0]}, \tag{4.5}$$

where

$$(4.6) \qquad k_{x,y} = (x-y)\{t_y - \alpha q_y - \beta v_y + [\kappa y + \beta(x-y-1)] f_y\} + \delta u_y .$$

It is interesting to compare the recursions (4.1)-(4.6) to the one proposed in Theorem 1.

Each step in the proposed recursion involves $(k+1)$ summations of the type $\Sigma_{y=1}^{x} f_y h_{x,y}$ (for some function $h_{x,y}$). The number of computations involved with the calculation of g_x when $g_0, \ldots, g_{x-1}$ are known is therefore proportional to x, and the number of computations involved with g_x is of order x^2. In practice, the severity distribution f_y has finite support such that $f_y = 0$ for $y > y_{\max}$, say. In this case the sum $\Sigma_{y=1}^{x} f_y h_{x,y}$ involves only $y_{\max}$ non-zero terms, and the number of computations involved with g_x is of order x.

TABLE 1

COMPUTING TIME, *minutes seconds* TO OBTAIN g_x FOR $k = 2$ WHEN f_y HAS FINITE SUPPORT WITH $y_{\max} = 50$

x	$m > 0$	$m = 0$	Willmot & Panjer
200	0 04	0 04	0 07
400	0 09	0 08	0 22
600	0 14	0 13	0 45
800	0 20	0 19	1 16
1000	0 26	0 24	1 54
1200	0 32	0 30	2 41
1400	0 39	0 37	3 37
1600	0 46	0 44	4 43
1800	0 54	0 51	5 54
2000	1 02	0 59	7 18

Also the recursions (4.3) and (4.5) of WILLMOT & PANJER (1987) involve summations $\Sigma_{y=1}^{x}$. However, these sums do not simplify in the case where f_y has finite support, and the total number of computations is therefore of order x^2

Table 1 shows for $k = 2$ the total computing time as a function of x for the recursion of WILLMOT & PANJER (1987) and for the proposed recursion. For the latter, we have treated separately the two cases where $m > 0$ and $m = 0$ (see Remark 2 2). In the first case we have programmed the recursion as presented in Theorem 1, and the matrix T_x has been inverted using STSC APL standard facilities. In the latter case we have used the formulas (2.14) and (2 15) The computations were done on a 486,50 mHz PC. The severity distribution has been chosen such that $y_{\max} = 50$ It should be noted that the computing time does not depend on the actual choice of parameters for the counting distribution, and also not on the actual choice of severity distribution (except for the choice of $y_{\max}$) The results are also displayed in Figure 1, where the computing times (in seconds) are shown as a function of x It is seen that the

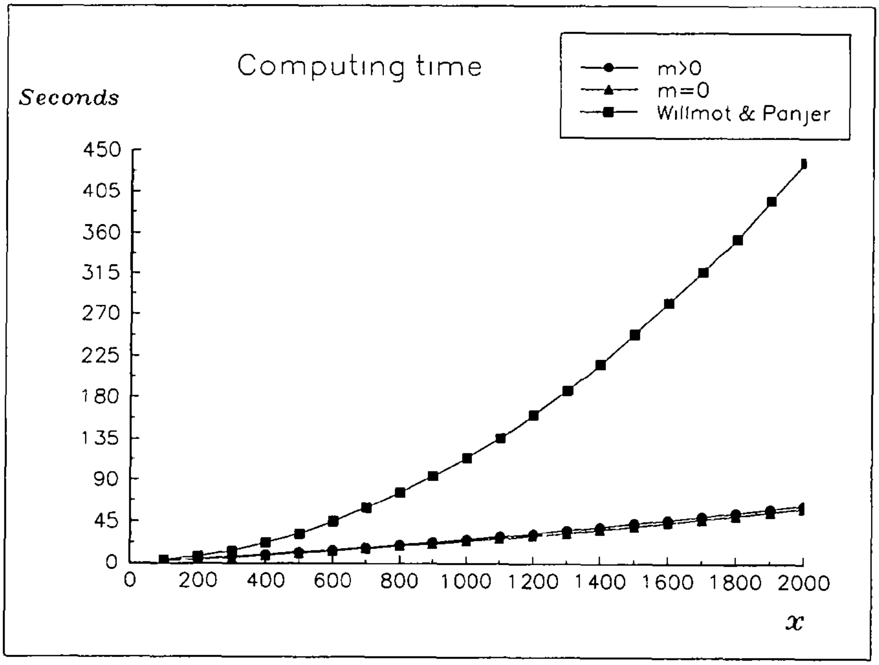


FIGURE 1 Computing time to obtain g_x for $k = 2$ when f_y has finite support with $y_{max} = 50$

total computing time is linear in x for the proposed recursion and quadratic for the recursion of WILLMOT & PANJER (1987).

With a hypergeometric counting distribution ($k = 2$) we have checked the recursions for numerical instabilities We consider two different severity distributions,

$$f_{1,y} = e^{-3y} / \sum_{y=0}^{20} e^{-3y}, \qquad y = 0, \ldots, 20.$$

$$f_{2,y} = 1/150, \qquad y = 0, \ldots, 149.$$

The distribution f_1 is very short-tailed with a high probability $f_0 = 0.2837$ of zero-claims. The second distribution f_2 is more heavy-tailed with a "large" average claim size $\mathrm{E}Y = 74.5$ For each of the severity distributions f_1 and f_2 we have calculated the compound distribution using a hypergeometric counting distribution with parameters (s, D, S) (see Example 2), where $D = S/4$ and $s = qS$, and (S, q) varies in the set $\{40, 100, 200\} \times \{0.25, 0.5, 0.75\}$. The corresponding average number of claims, $\mathrm{E}N = sD/S = qS/4$ is shown in Table 2. For the proposed recursion, $m > 0$, we have shifted the distributions f_1 and f_2 one step to the right, such that $m = 1$ in this case. The check for numerical instabilities was performed by simple graphical inspection. In Table 3 we have indicated by a * those cases where instabilities were found. All computations were continued until the 99.5% fractile of the compound distributions was reached.

TABLE 2

AVERAGE NUMBER OF CLAIMS, $\mathrm{E}N = qS/4$

	$q = 0.25$	$q = 0.5$	$q = 0.75$
$S = 40$	2.5	5	7.5
$S = 100$	6.25	12.5	18.75
$S = 200$	12.5	25	37.5

TABLE 3

NUMERICAL INSTABILITIES FOR COMPOUND HYPERGEOMETRIC DISTRIBUTIONS
INSTABILITIES ARE INDICATED BY A *

	Severity distribution f_1			Severity distribution f_2		
	$q = 0.25$	$q = 0.5$	$q = 0.75$	$q = 0.25$	$q = 0.5$	$q = 0.75$
	$m > 0$					
$S = 40$						
$S = 100$						
$S = 200$						
	$m = 0$					
$S = 40$						
$S = 100$				*		
$S = 200$	*	*		*	*	
	Willmot & Panjer					
$S = 40$						
$S = 100$					*	*
$S = 200$				*	*	*

It is noted that no instabilities were found for the proposed recursion in the case where $m > 0$. The recursion of WILLMOT & PANJER (1987) was unstable for the severity distribution f_2, when the average number of claims exceeds 10 (in this case). These instabilities can be attributed the accumulation of round-off errors. The proposed recursion, when $m > 0$, was unstable for "large" values of S and "small" values of q—irrespective of which severity distribution was used. An explanation for this instability can be found by examining the expression for $g_{k,x}$ in (2.15). This expression involves subtraction of terms $b_i g_{i,x}$, $i < k$, and subtraction (of equally large numbers) is known to increase the relative errors. For the hypergeometric distribution it holds that $b_0 = 0$ and $b_2 = 1$, whereas $b_1 = S-D-s$ (see Example 2). For the present combination of parameters it holds that $b_1 = S(0.75-q)$, which assumes its maximum when S is "large" and q is "small". In general, we would therefore expect that the proposed recursion is unstable for $m = 0$ when $S-D \gg s$ and stable when $S-D \approx s$.

It should be noted that all calculations were done with single precision, and that the results could (obviously) be improved by using double precision.

5. CALCULATION OF RUIN PROBABILITIES

Let

$$S(t) = \sum_{i=1}^{N(t)} Z_i,$$

where $N(t)$ denotes the number of claims incurred during $[0, t]$, and $Z_1, Z_2, \ldots,$ denote the corresponding claim amounts. The amounts Z_i are assumed to be independent of $N(t)$ and mutually independent with common distribution H. The average claim size is denoted by $\mu = \mathsf{E}Z_1$.

If premiums are paid continuously at a rate B pr. time unit, the maximal loss incurred is

$$L = \sup_{t \geq 0} \{S(t) - Bt\},$$

and the probability of ultimate ruin is

$$\psi(u) = \mathsf{P}(L > u),$$

where u denotes the initial capital. Assume that $B = (1+\theta)\lambda\mu$, where the relative safety loading θ is non-negative. It is a well known result (see e.g. BOWERS et al., 1986) that if $\{N(t)\}$ is a time-homogeneous Poisson process with claims rate λ, then

$$L \stackrel{D}{=} \sum_{i=0}^{M} L_i, \tag{5.1}$$

where M has a geometric distribution

$$\mathsf{P}(M = m) = (1-p)p^m, \quad p = \frac{1}{1+\theta}, \quad m = 0, 1, \ldots, \tag{5.2}$$

and $L_1, L_2, \ldots$ are mutually independent with common density

$$f(y) = (1 - H(y))/\mu. \tag{5.3}$$

PANJER (1986) suggested a discrete approximation to $f(y)$, and then to calculate $\psi(u)$ recursively by means of the Panjer-recursion, which is valid in the case of geometric counting distributions.

Consider now the case where $\{N(t)\}$, conditionally given $\Lambda = \lambda$, is a Poisson process with claims rate λ. Since, in this case,

$$(L|\Lambda = \lambda) \stackrel{D}{=} \sum_{i=0}^{M} L_i,$$

with M and L_i being distributed as before, it follows that

$$L \stackrel{D}{=} \sum_{i=0}^{M'} L_i,$$

where L_i still is distributed according to (5.3), and M' has a mixed geometric distribution. If we take a beta mixing function with parameters (α, β) for p appearing in (5.2), it follows that M' has the Waring distribution from Example 1. Using the same method as suggested by PANJER (1986) for discretizing the density (5.3), we may then apply Theorem 1 with $k = 1$ to obtain a recursive method for calculating $\psi(u)$.

Note, that if p is beta distributed with parameters (α, β), then the claims rate Λ is distributed as $(B/\mu)\,U$, where U is beta distributed with parameters (α, β).

REFERENCES

BOWERS, N L et al, (1986) *Actuarial Mathematics* The Society of Actuaries, Itasca, Illinois

DE PRIL, N (1985) Recursions for convolutions of arithmetic distributions *ASTIN Bulletin* **15**, 135–139

DE PRIL, N (1986) Moments of a class of compound distributions *Scand Actuarial J* **1986**, 117–120

GULDBERG, A (1931) On discontinuous frequency-functions and statistical series *Skandinavisk Aktuarietidskrift* **14**, 167–187

ORD, J K (1967) On a system of discrete distributions *Biometrika* **54**, 649–656

PANJER, H H (1981) Recursive evaluation of a family of compound distributions *ASTIN Bulletin* **11**, 22–26

PANJER, H H (1986) Direct calculation of ruin probabilities *The Journ of Risk and Insurance* **LIII**, 521–529

PANJER, H H and WILLMOT, G E (1982) Recursions for compound distributions *ASTIN Bulletin* **13**, 1–11

PANJER, H H and WILLMOT, G E (1986) Computational aspects of recursive evaluation of compound distributions *Insurance Math and Econ* **5**, 113 116

SCHRÖTER, K.J (1990) On a family of counting distributions and recursions for related compound distributions *Scand Actuarial J* **1990**, 161 174

SUNDT, B (1992) On some extensions of Panjer's class of counting distributions *ASTIN Bulletin* **22**, 61–80

SUNDT, B and JEWELL, W S (1981) Further results on recursive evaluation of compound distributions *ASTIN Bulletin* **11**, 27–39

WILLMOT, G E and PANJER, H H (1987) Difference equation approaches in evaluation of compound distributions *Insurance Math and Econ* **6**, 43–56

OLE HESSELAGER
Laboratory of Actuarial Mathematics, Universitetsparken 5,
University of Copenhagen, DK-2100 Copenhagen Ø.

SOME COMMENTS ON THE COMPOUND BINOMIAL MODEL

BY DAVID C.M. DICKSON*

The University of Melbourne

ABSTRACT

We show how ruin probabilities for the classical continuous time compound Poisson model can be approximated by ruin probabilities for a compound binomial model We also discuss ruin related results for a compound binomial model with geometric claim amounts

KEYWORDS

Ruin, compound binomial model; recursive calculation.

1. INTRODUCTION

GERBER (1988) presented some results for the compound binomial model which were analogues of results for the classical continuous time compound Poisson model These results were further discussed by SHIU (1989). WILLMOT (1992) presented some explicit results for ultimate ruin probabilities for the compound binomial model.

In this note we derive some known results for the compound binomial model using very elementary methods. We also present results for a binomial claim numbers/geometric claim amounts model which correspond to results for the classical continuous time Poisson/exponential model Our main purpose is to consider the conditions under which ultimate ruin probabilities for a compound binomial model give good approximations to ultimate ruin probabilities in the classical continuous time compound Poisson model

We start by considering some basic results for a general discrete time risk model.

2. A DISCRETE TIME RISK MODEL

We will consider a risk model with the following characteristics:

(a) X_i denotes the aggregate claim amount in the i-th time interval;
(b) $\{X_i\}_{i=1}^{\infty}$ is a sequence of independent and identically distributed random variables, each distributed on the non-negative integers,
(c) the insurer's premium income per unit time is 1,
(d) $E(X_i) < 1$.

* Part of this work was completed while the author was at Heriot-Watt University, Edinburgh

We will assume throughout that the insurer's initial surplus, denoted u, is an integer.

The insurer's surplus at time t ($t = 1, 2, 3, \ldots$) is denoted $Z(t)$ and given by

$$Z(t) = u + t - \sum_{i=1}^{t} X_i$$

The ultimate ruin probability for this model is defined by

$$\psi(u) = \Pr[Z(t) \leq 0 \quad \text{for some } t, \quad t = 1, 2, 3, \ldots]$$

This definition corresponds to that given by GERBER (1988) but differs from that used by SHIU (1989) and WILLMOT (1992). The reason for choosing this definition will become clear in Section 5 Note that ruin does not occur at time 0 if the initial surplus is zero. The survival probability is denoted $\delta(u)$ and $\delta(u) = 1 - \psi(u)$.

We define the severity of ruin function $G(u, y)$ for $u = 0, 1, 2, \ldots$ and $y = 1, 2, 3,$ by

$$G(u, y) = \Pr[T < \infty \quad \text{and} \quad Z(T) > -y]$$

where T is the discrete time of ruin and is defined by

$$\begin{aligned} T &= \min\{t \;\; Z(t) \leq 0, \quad t = 1, 2, 3, \ldots\} \\ &= \infty \quad \text{if} \quad Z(t) > 0 \quad \text{for} \quad t = 1, 2, 3, \ldots \end{aligned}$$

Thus $G(u, y)$ represents the probability that ruin occurs and that the deficit at the time of ruin is at most $y - 1$

We denote by $b(k)$ and $B(k)$ respectively the probability function and distribution function of X_i.

3 GENERAL RESULTS

Result 1: For $u = 1, 2, 3,$.

$$\delta(u) = \delta(0) + \sum_{k=1}^{u} \delta(k)[1 - B(u-k)] \tag{3.1}$$

Proof: By considering the possible aggregate claim amounts in the first time period we have that

$$\delta(0) = b(0)\,\delta(1)$$

and for $u = 2, 3, 4,$.

$$\delta(u-1) = b(0)\,\delta(u) + \sum_{j=1}^{u-1} \delta(j)\, b(u-j) \tag{3.2}$$

Hence, for $u=2, 3, 4, \ldots$

$$\sum_{k=0}^{u-1} \delta(k) = b(0) \sum_{k=1}^{u} \delta(k) + \sum_{k=2}^{u} \sum_{j=1}^{k-1} \delta(j)\, b(k-j)$$

$$= b(0) \sum_{k=1}^{u} \delta(k) + \sum_{k=1}^{u-1} \delta(k)\,[B(u-k) - b(0)]$$

$$= b(0)\,\delta(u) + \sum_{k=1}^{u-1} \delta(k)\, B(u-k)$$

Thus

$$b(0)\,\delta(u) = \delta(0) + \sum_{k=1}^{u-1} \delta(k)\,[1 - B(u-k)]$$

$$= \delta(u-1) - \sum_{k=1}^{u-1} \delta(k)\, b(u-k) \qquad \text{(by (3 2))}$$

so that $\delta(u-1) = \delta(0) + \sum_{k=1}^{u-1} \delta(k)\,[1 - B(u-1-k)]$

for $u=2, 3, 4,$, or equivalently,

$$\delta(u) = \delta(0) + \sum_{k=1}^{u} \delta(k)\,[1 - B(u-k)] \qquad \text{for} \qquad u = 1, 2, 3, \ldots$$

Result 2: The ruin probability from initial surplus zero is given by

$$\psi(0) = E(X_i) \tag{3.3}$$

Proof: For $y=0, 1, 2,$ define $g(0,y)$ to be the probability that ruin occurs from initial surplus zero and that the deficit at the time of ruin is y Note that when the initial surplus is $u\,(>0)$, $g(0, y)$ can be interpreted as the probability that the surplus falls below its initial level for the first time and by amount y When $y=0$, $g(0, y)$ gives the probability that the surplus returns to its initial level for the first time without previously having been below its initial level Using this interpretation we can write

$$\delta(u) = \delta(0) + \sum_{y=1}^{u} g(0, u-y)\,\delta(y) \tag{3 4}$$

The first term on the right hand side gives the probability that the surplus never falls below its initial level. For a fixed value of $y\,(<u)$, $g(0, u-y)\,\delta(y)$ gives the probability that the surplus falls below its initial level for the first time to y and that survival occurs from surplus level y. A similar interpretation applies when $y=u$.

Summing over y gives the probability that survival occurs and that the surplus process has not always been above its initial level.

By (3.1) we also have

$$\delta(u) = \delta(0) + \sum_{v=1}^{u} \delta(v)\,[1 - B(u-v)]$$

Since equations (3.1) and (3.4) hold for $u = 1, 2, 3, \ldots$, it follows that $g(0, y) = 1 - B(y)$. Equation (3.3) follows since

$$\psi(0) = \sum_{y=0}^{\infty} g(0, y)$$

If we write the premium income of 1 as $(1 + \theta)\, E(X_i)$, then

$$\psi(0) = 1/(1 + \theta) \tag{3.5}$$

as in the classical continuous time model.

We can easily obtain further ruin related results when the initial surplus is zero, starting with the joint distribution of the surplus prior to ruin and the deficit at ruin. We define a new function $f(u, x, y)$ for $x = 1, 2, 3, \ldots$ and $y = 0, 1, 2, \ldots$ as follows:

$$f(u, x, y) = \Pr\,[T < \infty,\, Z(T) = -y \quad \text{and} \quad Z(T-1) = x]$$

Thus $f(u, x, y)$ gives the probability that ruin occurs from initial surplus u, with a deficit of y at the time of ruin and a surplus of x one time unit prior to ruin. When $u = 0$, the function is defined for $x = 0, 1, 2, \ldots$, and $f(0, 0, y)$ simply gives the probability that ruin occurs at time 1 with a deficit of y at ruin. Thus $f(0, 0, y) = b(y + 1)$.

By considering the possible aggregate claim amounts in the first time period we can write

$$f(u, x, y) = \sum_{j=0}^{u} b(j) f(u + 1 - j, x, y) \quad \text{for} \quad u = 0, 1, 2, \ldots, x - 1, x + 1,$$

and

$$f(u, x, y) = \sum_{j=0}^{u} b(j) f(u + 1 - j, x, y) + b(x + y + 1) \quad \text{when} \quad u = x$$

Assuming that

$$\sum_{u=0}^{\infty} f(u, x, y) < \infty \tag{3.6}$$

we have that

$$\sum_{u=0}^{\infty} f(u,x,y) = \sum_{u=0}^{\infty} \sum_{j=0}^{u} b(j) f(u+1-j,x,y) + b(x+y+1)$$

$$= \sum_{u=1}^{\infty} f(u,x,y) \sum_{j=0}^{\infty} b(j) + b(x+y+1)$$

Hence

$$f(0,x,y) = b(x+y+1) \tag{3.7}$$

As an immediate consequence of this we have that

$$G(0,y) = \sum_{j=0}^{y-1} \sum_{x=0}^{\infty} b(x+j+1) = \sum_{j=0}^{y-1} [1-B(j)]$$

and

$$\psi(0) = \sum_{j=0}^{\infty} [1-B(j)]$$

Similarly

$$\Pr[T<\infty \quad \text{and} \quad Z(T-1) \le x-1 \mid u=0] = \sum_{j=0}^{x-1} \sum_{y=0}^{\infty} b(j+y+1) \tag{3.8}$$

$$= \sum_{j=0}^{x-1} [1-B(j)] = G(0,x)$$

We have not discussed the conditions under which (3.6) holds The most obvious situation when (3.6) holds is when Lundberg's inequality applies Formula (3.7) does however hold when the sum in (3.6) is infinite

The results presented above are in terms of a general distribution $B(k)$ However, they are in fact the same as results given by GERBER (1988) and SHIU (1989). This follows since the distribution of X_t can be expressed as a compound binomial distribution with binomial parameters 1 and $1-b(0)$ and probability function for individual claims $b(j)/[1-b(0)]$ for $j = 1, 2, 3, \ldots$.

4. THE BINOMIAL/GEOMETRIC MODEL

Throughout this section we assume that the distribution of the number of claims per unit time is binomial with parameters 1 and p, and the individual claim amount distribution is geometric with distribution function $P(x)$ and probability function

$$p(x) = (1-\alpha)\alpha^{x-1} \quad \text{for} \quad x = 1, 2, 3, \ldots$$

Then

$$B(k) = 1 - p\alpha^{k} \quad \text{for} \quad k = 0, 1, 2, \ldots$$

Since we have assumed that $E(X_t) < 1$, the parameters p and α must be such that $p/(1-\alpha) < 1$

We can rewrite equation (3.1) as

$$\psi(u) = \psi(0) - \sum_{k=1}^{u} [1 - \psi(k)] [1 - B(u-k)]$$

and inserting for $\psi(0)$ we have

$$\psi(u) = \sum_{k=1}^{u} \psi(k) [1 - B(u-k)] + \sum_{k=u}^{\infty} [1 - B(k)]$$

The continuous time compound Poisson analogue of this equation can be found in, for example, GERBER (1979).

Now insert for $B(k)$ to find that

$$\psi(u) = \sum_{k=1}^{u} \psi(k) p\alpha^{u-k} + \sum_{k=u}^{\infty} p\alpha^{k} \tag{4.1}$$

and

$$\psi(u+1) = \sum_{k=1}^{u+1} \psi(k) p\alpha^{u+1-k} + \sum_{k=u+1}^{\infty} p\alpha^{k} \tag{4.2}$$

If we multiply (4.1) by α, subtract from (4.2) and rearrange we find that

$$\psi(u+1) - \frac{\alpha}{1-p} \psi(u) = 0$$

The solution to this difference equation is

$$\psi(u) = c \left(\frac{\alpha}{1-p} \right)^{u}$$

from which it follows that $c = \psi(0)$. In fact, we can write $\psi(u) = \psi(0) \exp\{-Ru\}$, where R is the adjustment coefficient for this process. R is the unique positive number satisfying

$$E[\exp\{R(X_t - 1)\}] = 1$$

and it is an elementary exercise to show that for this model $\exp\{R\} = (1-p)/\alpha$ Thus we have a complete analogy with the form of the ruin probability for the Poisson/exponential model which can be written in exactly the same way (See, for example, GERBER (1979)). We note that this solution matches that given by WILLMOT (1992) for $\delta(u)$, allowing for different definitions of ruin/survival

We now extend the analogy to the severity of ruin. We can use the function $g(0, y)$ to write down an equation for $G(u, y)$ by considering the first occasion on which the surplus falls below (or returns to without previously having been below) its initial level

We have

$$G(u,y) = \sum_{k=u}^{u+y-1} g(0,k) + \sum_{k=0}^{u-1} g(0,k)\, G(u-k,y)$$

$$= \sum_{k=u}^{u+y-1} g(0,k) + \sum_{k=1}^{u} g(0,u-k)\, G(k,y)$$

Now insert $g(0,k) = 1 - B(k) = p\alpha^k$ to give

$$G(u,y) = \sum_{k=u}^{u+y-1} p\alpha^k + \sum_{k=1}^{u} p\alpha^{u-k} G(k,y)$$

and

$$G(u+1,y) = \sum_{k=u+1}^{u+y} p\alpha^k + \sum_{k=1}^{u+1} p\alpha^{u+1-k} G(k,y)$$

Using the same technique as before we find that

$$G(u+1,y) - \frac{\alpha}{1-p} G(u,y) = 0$$

and hence

$$G(u,y) = G(0,y)\left(\frac{\alpha}{1-p}\right)^u$$

Finally

$$G(0,y) = \sum_{k=0}^{y-1} g(0,k) = \sum_{k=0}^{y-1} p\alpha^k = p\,\frac{1-\alpha^y}{1-\alpha} \quad \text{for} \quad y = 1, 2, 3,$$

and so we can write

$$G(u,y) = (1-\alpha^y)\frac{p}{1-\alpha}\left(\frac{\alpha}{1-p}\right)^u = P(y)\,\psi(u)$$

Thus the form of $G(u,y)$ is identical to that for the Poisson/exponential model. (See, for example, DICKSON (1992)). However, unlike the Poisson/exponential model, the distribution of the deficit at the time of ruin is not identical to the individual claim amount distribution. The deficit is geometrically distributed with parameter α, but on $0, 1, 2, \dots$, since $G(u,y)/\psi(u)$ gives the probability that the deficit is less than or equal to $y-1$, given that ruin occurs, and so

$$\Pr[-Z(T) < y \mid T < \infty] = 1 - \alpha^y \quad \text{for} \quad y = 1, 2, 3, \dots$$

Let us now consider the situation when $u = 0$ further. We have already noted that the deficit at the time of ruin is geometrically distributed on $0, 1, 2, \dots$ with parameter α, and by (3 8) the distribution of the surplus at time $T-1$ is the same.

The conditional probability function of the deficit at T and of the surplus at $T-1$, conditioning on the event that ruin occurs, is

$$\tilde{g}(0,x)=(1-\alpha)\alpha^x, \quad x=0,1,2, \ldots$$

If we consider the conditional distribution of the surplus one time unit before ruin and of the deficit at ruin, conditioning on the event that ruin occurs, and again use a tilde to denote a conditional probability, then

$$\tilde{f}(0,x,y)=\frac{b(x+y+1)}{\psi(0)}=\frac{p(1-\alpha)\alpha^{x+y}}{p/(1-\alpha)}=\tilde{g}(0,x)\,\tilde{g}(0,y)$$

so that, conditionally, the surplus one time unit before ruin and the deficit at ruin are independent and identically distributed This situation also exists in the Poisson/exponential model where the surplus prior to ruin and deficit at ruin are independent, identically distributed variables, and the conditional distribution of the claim causing ruin is Gamma(2)

Finally, if we define the conditional probability function of the claim causing ruin as $h(0,z)$ for $z=1,2,3,\ldots$ then

$$h(0,z)=\sum_{x=0}^{z-1}\tilde{f}(0,x,z-x-1)=\sum_{x=0}^{z-1}(1-\alpha)^2\alpha^{z-1}=z(1-\alpha)^2\alpha^{z-1}$$

The conditional distribution of the claim causing ruin is thus negative binomial with parameters 2 and $1-\alpha$, shifted one unit to the right

5. CALCULATION OF RUIN PROBABILITIES

GERBER (1988) states that the compound binomial model can be used to approximate the continuous time compound Poisson model. In this section we investigate this statement by considering ultimate ruin probabilities

To calculate ruin probabilities for the compound binomial model, we will adapt the framework described by DICKSON and WATERS (1991, Sections 1 and 8) who use a discrete time compound Poisson model to approximate a classical continuous time compound Poisson model under which both the Poisson parameter and mean individual claim amount are 1. The characteristics of this model are as follows

(a) individual claim amounts are distributed on the non-negative integers with mean β, where $\beta(>1)$ is an integer;
(b) the Poisson parameter for the expected number of claims per unit time is $1/[(1+\theta)\beta]$;
(c) the premium income per unit time is 1.

We will replace this discrete compound Poisson model by a compound binomial model. We simply change (b), replacing the Poisson distribution by a binomial distribution with parameters 1 and $1/[(1+\theta)\beta]$. For reasons given by DICKSON and WATERS (1991) we can regard $\psi(\beta u)$ as an approximation to $\psi_c(u)$, the ultimate ruin probability for the continuous compound Poisson model. Note that the definition of $\psi(u)$ given in Section 2 corresponds to that used in this approximation. In effect all we are doing is approximating a discrete compound Poisson model (which approximates a continuous compound Poisson model) by a compound binomial model The approximation to the discrete compound Poisson model is reasonable for large values of β, since the Poisson distribution with parameter $1/[(1+\theta)\beta]$ is then very close to the approximating binomial distribution. For example, if $\beta=100$ and $\theta=0.1$, then the probability of more than one claim per unit time under the compound Poisson model is 0.00004. Note that there is one small difference between this formulation of the compound binomial model and that used by previous authors In this formulation, individual claim amounts are distributed on the non-negative integers rather than the positive integers The reason for this is simply that in order to approximate ruin probabilities in the classical continuous time compound Poisson model, we have to discretize the continuous individual claim amount distribution in that model In our first two examples, we will use the discretization proposed by DE VYLDER and GOOVAERTS (1988), which discretizes the distribution on the non-negative integers If we had chosen a discretization on the positive integers then our model would correspond to that used by previous authors

We will calculate ruin probabilities recursively from the formulae

$$\psi(1)=b(0)^{-1}[\psi(0)-1+B(0)] \tag{5.1}$$

and for $u=2, 3, 4, \ldots$

$$\psi(u)=b(0)^{-1}\left[\psi(u-1)-1+B(u-1)-\sum_{j=1}^{u-1} b(j)\,\psi(u-j)\right] \tag{5.2}$$

These formulae correspond to GERBER's (1988) formulae (5) and (6) In each of the following examples the premium loading factor, θ, is 10%

Example 1: Let the individual claim amount distribution in the continuous time model be exponential with mean 1. Then it is well known (see, for example, GERBER (1979)) that

$$\psi_c(u)=\frac{1}{1+\theta}\exp(-R_c u) \qquad \text{where} \qquad R_c=\theta/(1+\theta)$$

Table 1 shows exact and approximate values of $\psi_c(u)$ The approximate values are calculated from (3.5), (5.1) and (5.2). The legend for this table is as follows

(1) denotes the exact value of $\psi_c(u)$;
(2) denotes the approximate value when $\beta=50$,
(3) denotes the ratio of the value in (2) to that in (1);
(4) denotes the approximate value when $\beta=100$,
(5) denotes the ratio of the value in (4) to that in (1);
(6) denotes the approximate value when $\beta=200$;
(7) denotes the ratio of the value in (6) to that in (1).

TABLE 1

(SEE EXAMPLE 1 FOR DETAILS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$u=0$	0 9091	0 9091	1 0000	0 9091	1 0000	0 9091	1 0000
$u=2$	0 7580	0 7567	0 9983	0 7573	0 9992	0 7576	0 9996
$u=4$	0 6319	0 6299	0 9967	0 6309	0 9983	0 6314	0 9992
$u=6$	0 5269	0 5243	0 9950	0 5256	0 9975	0 5262	0 9988
$u=8$	0 4393	0 4364	0 9934	0 4378	0 9967	0 4386	0 9983
$u=10$	0 3663	0 3632	0 9917	0 3647	0 9959	0 3655	0 9979
$u=20$	0 1476	0 1451	0 9835	0 1463	0 9917	0 1470	0 9959
$u=40$	0 0240	0 0232	0 9673	0 0236	0 9836	0 0238	0 9918
$u=80$	0 0006	0 0006	0 9357	0 0006	0 9674	0 0006	0 9836

We note the following points about Table 1 ·

(a) When $u > 0$, the approximate values are less than the exact ones. This is to be expected since the compound binomial model excludes the possibility of multiple claims per unit time.

(b) As the value of β increases, the approximate values become closer to the exact ones. This is as expected for reasons given by DICKSON and WATERS (1991, Section 2)

(c) The larger the value of u, the poorer the approximation becomes.

(d) Even with a large value of β, the approximate values do not always agree with the exact values to four decimal places

Example 2: Let the individual claim amount distribution in the continuous time model be Pareto with parameters 2 and 1. Table 2 shows exact and approximate values of $\psi_c(u)$ (The exact values have been calculated using DICKSON and WATERS' (1991) algorithm and are "exact" at least to three decimal places) The legend for Table 2 is the same as for Table 1 The only additional comment that we would make about Table 2 is that, for the same magnitude of ruin probability, the approximate values are slightly closer to the exact values than in Example 1

TABLE 2

(SEE EXAMPLE 2 FOR DETAILS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$u=0$	0 9091	0 9091	1 0000	0 9091	1 0000	0 9091	1 0000
$u=2$	0 8102	0 8097	0 9994	0 8100	0 9997	0 8101	0 9998
$u=4$	0 7498	0 7491	0 9991	0 7494	0 9996	0 7496	0 9998
$u=6$	0 7021	0 7014	0 9990	0 7018	0 9995	0 7020	0 9997
$u=8$	0 6620	0 6613	0 9989	0 6617	0 9994	0 6619	0 9997
$u=10$	0 6271	0 6264	0 9988	0 6267	0 9994	0 6269	0 9997
$u=20$	0 4981	0 4974	0 9985	0 4978	0 9992	0 4980	0 9996
$u=40$	0 3479	0 3473	0 9982	0 3476	0 9991	0 3477	0 9995
$u=80$	0 2040	0 2036	0 9981	0 2038	0 9990	0 2039	0 9995

In Section 4 we discussed the binomial/geometric model as the discrete analogue of the Poisson/exponential model. In Example 3 we illustrate how ruin probabilities for the binomial/geometric model can be used to approximate those for the Poisson/exponential model. We have included this example purely for interest as the approach does not generalise to other compound Poisson models.

Example 3: We will use the same framework as in Examples 1 and 2, but will discretize the exponential individual claim amount distribution as a geometric distribution with mean β. This discretization is a reasonable one for large values of β since when β is large

$$P(x) = 1 - (1-\beta^{-1})^x \approx 1 - \exp\{-x/\beta\} \quad \text{for} \quad x = 0, 1, 2,$$

As noted in Section 4, for the geometric individual claim amount distribution,

$$\psi(\beta u) = \frac{1}{1+\theta}\exp(-R\beta u) \quad \text{where} \quad R = \log_e\left(\frac{(1+\theta)\beta - 1}{(1+\theta)(\beta-1)}\right)$$

It is easy to show that

$$\lim_{\beta\to\infty} \beta R = \frac{\theta}{1+\theta}$$

so that for large values of β, $\psi(\beta u)$ should give a good approximation to $\psi_c(u)$.

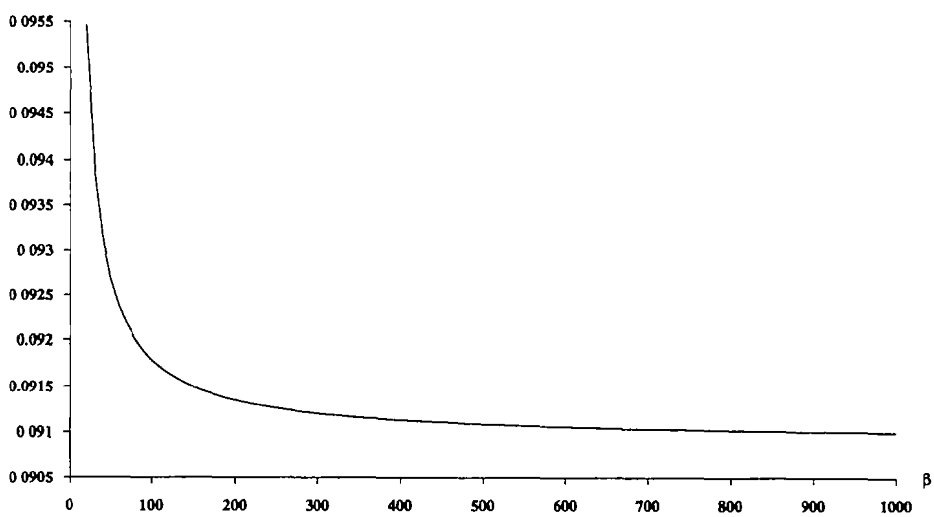


FIGURE 1 βR as a function of β when θ is 10%

Figure 1 shows the function βR (when θ is 10%) and Table 3 shows exact and approximate values of $\psi_c(u)$. The legend for Table 3 is as follows:

(1) denotes the exact value of $\psi_c(u)$;
(2) denotes the approximate value when $\beta = 100$;

(3) denotes the ratio of the value in (2) to that in (1),
(4) denotes the approximate value when $\beta = 1{,}000$;
(5) denotes the ratio of the value in (4) to that in (1),
(6) denotes the approximate value when $\beta = 10{,}000$,
(7) denotes the ratio of the value in (6) to that in (1)

TABLE 3

(SEE EXAMPLE 3 FOR DETAILS)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$u=0$	0.9091	0.9091	1.0000	0.9091	1.0000	0.9091	1.0000
$u=2$	0.7580	0.7566	0.9982	0.7578	0.9998	0.7579	1.0000
$u=4$	0.6319	0.6297	0.9965	0.6317	0.9997	0.6319	1.0000
$u=6$	0.5269	0.5241	0.9948	0.5266	0.9995	0.5269	0.9999
$u=8$	0.4393	0.4362	0.9930	0.4390	0.9993	0.4393	0.9999
$u=10$	0.3663	0.3631	0.9913	0.3659	0.9991	0.3662	0.9999
$u=20$	0.1476	0.1450	0.9826	0.1473	0.9983	0.1475	0.9998
$u=40$	0.0240	0.0231	0.9656	0.0239	0.9965	0.0239	0.9997
$u=80$	0.0006	0.0006	0.9323	0.0006	0.9931	0.0006	0.9993

Table 3 shows the same features as Tables 1 and 2. The great advantage of using the geometric discretization is that approximate values for $\psi_c(u)$ can be calculated from a formula. This allows us to use very large values for β, and shows that even with a large value of β (i.e. 10,000) the approximate values do not all match the exact ones to four decimal places. By contrast, if $b(x)$ and $B(x)$ in (5.1) and (5.2) are values from a compound Poisson distribution, then a relatively small value of β produces the same degree of accuracy. (See, for example, DICKSON and WATERS (1991, Table 5)).

We conclude that it is possible to successfully approximate ruin probabilities for the classical continuous time compound Poisson model by those for a compound binomial model. The main advantage in using the compound binomial model is that it is not necessary to perform recursive calculations to find the probability function $b(x)$ to use formulae (5.1) and (5.2). However, this advantage is outweighed by the fact that a large value of β is required when using the compound binomial model in order to obtain a good approximation to $\psi_c(u)$.

ACKNOWLEDGEMENT

I am grateful to the referees and editor for comments on the first draft of this paper.

REFERENCES

DE VYLDER, F and GOOVAERTS, M J (1988) Recursive calculation of finite-time ruin probabilities *Insurance Mathematics and Economics* **7**, 1–8

DICKSON, D C M and WATERS, H R (1991) Recursive calculation of survival probabilities *ASTIN Bulletin* **21**, 199–221

DICKSON, D C M (1992) On the distribution of the surplus prior to ruin *Insurance Mathematics and Economics* **11**, 191–207

GERBER, H U (1979) *An Introduction to Mathematical Risk Theory* S S Huebner Foundation Monograph Series No 8 Distributed by R Irwin, Homewood, IL

GERBER, H U (1988) Mathematical fun with the compound binomial process *ASTIN Bulletin* **18**, 161–168

SHIU, E S W (1989) The probability of eventual ruin in the compound binomial model *ASTIN Bulletin* **19**, 179–190

WILLMOT, G E (1992) Ruin probabilities in the compound binomial model *Insurance Mathematics and Economics* **12**, 133–142

DAVID C M. DICKSON
Centre for Actuarial Studies, Faculty of Economics and Commerce,
The University of Melbourne, Parkville,
Victoria 3052, Australia.

February 1993
Revised November 1993

LIMITING DISTRIBUTION OF THE PRESENT VALUE OF A PORTFOLIO

BY GARY PARKER

Simon Fraser University

ABSTRACT

An approximation of the distribution of the present value of the benefits of a portfolio of temporary insurance contracts is suggested for the case where the size of the portfolio tends to infinity. The model used is the one presented in PARKER (1922b) and involves random interest rates and future lifetimes Some justifications of the approximation are given. Illustrations for limiting portfolios of temporary insurance contracts are presented for an assumed Ornstein-Uhlenbeck process for the force of interest

KEYWORDS

Force of interest, Ornstein-Uhlenbeck process, Portfolio of policies; Present value function; Limiting distribution

1. INTRODUCTION

When considering random interest rates in actuarial functions, a question of particular interest is the distribution of the present value of a portfolio of policies Studying such distributions could be very useful in areas such as pricing, valuation, solvency analysis and reinsurance.

Some references which considered stochastic interest rates in actuarial functions are BOYLE (1976), WILKIE (1976), WATERS (1978), PANJER and BELLHOUSE (1980), DEVOLDER (1986), GIACOTTO (1986), DHAENE (1989), DUFRESNE (1988), BEEKMAN and FUELLING (1990), PARKER (1992b).

Recently, DUFRESNE (1990) derived the distribution of a perpetuity for i.i.d interest rates. FREES (1990) recursively expressed by an integral equation the distribution of a block of n-year annuities for i.i.d interest rates.

This paper, taken for the most part from the author's Ph.D thesis (PARKER (1992a)), presents an approximation of the limiting distribution, as the number of policies tend to infinity, of the average present value of the benefits for a specific type of portfolio of insurance contracts Although, theoretically, the approach may be used for any stochastic process for the interest rates, it is more convenient for Gaussian processes The approximation is justified by two correlation coefficients which happen to be relatively high mainly because of the definition of the present value function. Some illustrations of the distribution function of the present value of portfolios using the Ornstein-Uhlenbeck process are presented Finally, the moments of some approximate distributions are compared with the corresponding exact moments

2. A PORTFOLIO

Consider a portfolio of temporary insurance contracts, each with sum insured 1, issued to c lives insured aged x. Let $Z(c)$ be the random present value of the benefits of the portfolio

PARKER (1922b) used a definition of $Z(c)$ involving a summation over the c contracts of the portfolio. That is

$$Z(c) = \sum_{i=1}^{c} Z_i, \tag{2.1}$$

where Z_i is the present value of the benefit for the ith life insured of the portfolio. This definition is convenient for calculating the moments of $Z(c)$ because it is possible to simplify the expressions for these moments under the assumption that the future lifetimes of the c policyholders are mutually independent.

Another definition which is equivalent appears to be more appropriate for studying the limiting distribution of the random variable $Z(c)$.

Instead of summing over the c policies, one could consider summing the present value of the benefits in a given year over the n policy-years of the contract Algebraically, we have

$$Z(c) = \sum_{i=0}^{n-1} c_i \; e^{-y(i+1)}, \tag{2.2}$$

where

$$y(i+1) = \int_0^{i+1} \delta_s \, ds, \tag{2.3}$$

d_s is the force of interest at time s and c_i, $i = 0, 1, \ldots, n-1$ is the random variable denoting the number of policies where the death benefit is actually paid at time $i+1$. We let c_n be the number of lives insured surviving to the end of the term, n Note that the sum of the c_i's from i equal 0 to n is c, the total number of policies in the portfolio. Thus,

$$\sum_{i=0}^{n} c_i = c \tag{2 4}$$

When studying $Z(c)$, we will assume that the future lifetimes of the lives insured are mutually independent and independent of the forces of interest $\{\delta_s\}_{s \geq 0}$. In this case, the $\{c_i\}_{i=1}^{n}$ is multinominal We will also assume that the discounting of all the benefits for the policies in the portfolios is done with the same Gaussian forces of interest.

In the next section, we consider limiting portfolios, i.e portfolios where the number of contracts tends to infinity.

3. LIMITING DISTRIBUTION

Using (2.2), one could intuitively derive that the average cost per policy (defined as $Z(c)/c$) as the number of such policies tends to infinity would simply be a weighted average of the present value functions from year 1 to year n. The weights being the expected proportion of contracts payable in each year, i.e. ${}_{t|}q_x$ The probabilistic version of this intuition is presented in Theorem 1

Theorem 1: As c tends to infinity, the average cost per policy for a portfolio of n-year temporary insurance contracts tends in distribution to (see also proposition 5 of FREES (1990))

$$\zeta_n = \sum_{t=0}^{n-1} {}_{t|}q_x \; e^{-y(t+1)} \tag{3.1}$$

Proof: This result is true if

$$Z(c)/c - \zeta_n = \sum_{t=0}^{n-1} (c_t/c - {}_{t|}q_x) \; e^{-y(t+1)} \tag{3.2}$$

tends in probability to 0.

We use the well-known result that if X tends in probability to 0 and Y has finite mean and variance, then $X \; Y$ tends in probability to 0 (see, for example, CHUNG (1974, p 92)).

Here, c_t is binomial $(c, {}_{t|}q_x)$ so, $(c_t/c - {}_{t|}q_x)$ tends in probability to 0 for each t. And as $e^{-y(t+1)}$ is log-normally distributed with finite mean and variance, it follows that

$$\sum_{t=0}^{n-1} (c_t/c - {}_{t|}q_x) \; e^{-y(t+1)}$$

tends in probability to 0 □

Now, one could theoretically obtain the density function of ζ_n by integrating the joint density function of the $y(t)$'s over the appropriate domain. The expression would look like the following

$$f_{PZ_n}(z) = \int_{y_n} \cdots \int_{y_2} \int_{y_1} f_{\underline{Y}}(y_1, y_2, \ldots, y_n) \, dy_1 \, dy_2 \ldots dy_n \,, \tag{3.3}$$

where $\underline{Y} = (y(1), y(2), \ldots, y(n))$ and is multivariate normal

But this approach is not possible from a practical point of view as it is almost impossible to evaluate (3 3) even for n as small as 5 In the next section, however, we derive a recursive equation from which one can approximate the distribution of ζ_n.

4. APPROXIMATION

Since ζ_n is a summation over the policy-years, it is easy to break it down into the sum of ζ_{n-1} and a term for the nth policy year. The recursive equation for ζ_n is then given by:

$$\zeta_n = \sum_{t=0}^{n-1} {}_{t|}q_x \; e^{-y(t+1)} = \sum_{t=0}^{n-2} {}_{t|}q_x \; e^{-y(t+1)} + {}_{n-1|}q_x \; e^{-y(n)}$$

$$\zeta_n = \zeta_{n-1} + {}_{n-1|}q_x \; e^{-y(n)}. \tag{4.1}$$

Let z_t be a possible realization of ζ_t and y_j be a possible realization of $y(j)$

Let the function $g_n(z_n, y_n)$, a somewhat unusual function based on the distribution of ζ_n and the density function of $y(n)$, be defined as:

$$g_n(z_n, y_n) = P(\zeta_n \le z_n) \; f_{y(n)}(y_n | \zeta_n \le z_n), \tag{4.2}$$

or equivalently,

$$g_n(z_n, y_n) = f_{y(n)}(y_n) \quad P(\zeta_n \le z_n | y(n) = y_n). \tag{4.3}$$

From this last definition, it follows immediately that the distribution function of ζ_n is given by:

$$F_{\zeta_n}(z_n) = \int_{-\infty}^{\infty} g_n(z_n, y_n) \; dy_n, \tag{4.4}$$

where the function $g_n(z_n, y_n)$ may be calculated with a high degree of accuracy from the following recursive equation

$$g_n(z_n, y_n) \cong \int_{-\infty}^{\infty} f_{y(n)}(y_n | y(n-1) = y_{n-1}) \times \tag{4.5}$$

$$\times \; g_{n-1}(z_n - {}_{n-1|}q_x \; e^{-y_n}, y_{n-1}) \, dy_{n-1}$$

with the starting value:

$$g_1(z_1, y_1) = \begin{cases} \phi\left(\dfrac{y_1 - E[y(1)]}{V[y(1)]^{.5}}\right) & \text{if} \quad z_1 \ge q_x \; e^{-y_1} \\ 0 & \text{otherwise} \end{cases} \tag{4.6}$$

We use the notation $\phi(\,)$ to denote the probability density function of a zero mean and unit variance normal random variable. Note also that given that $y(n-1)$ equal y_{n-1}, $y(n)$ is normally distributed with mean

$$E[y(n)|y(n-1) = y_{n-1}] \tag{4.7}$$

$$= E[y(n)] + \frac{\operatorname{cov}(y(n), y(n-1))}{V[y(n)]} \quad \{y_{n-1} - E[y(n-1)]\}$$

and variance

$$(4.8)\qquad V[y(n)|y(n-1)=y_{n-1}] = V[y(n)] - \frac{\operatorname{cov}^2(y(n), y(n-1))}{V[y(n-1)]}$$

(see, for example, MORRISON (1990, p. 92))

To derive (4.5), we start by noting that from (4.1), we have that:

$$(4.9)\qquad P(\zeta_n \le z_n | y(n)=y_n) = P(\zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n} | y(n)=y_n)$$

Now using (4.2), (4.3) and (4.9), we have

$$(4.10)\qquad g_n(z_n, y_n) = P(\zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n}) \times \\ \times f_{y(n)}(y_n | \zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n})$$

The conditional probability density function of $y(n)$ in (4.10) may be written as: (MELSA and SAGE (1973, p. 98))

$$(4.11)\qquad f_{y(n)}(y_n|\zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n}) \\ = \int_{-\infty}^{\infty} f_{y(n)}(y_n|y(n-1)=y_{n-1}, \zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n}) \times \\ \times f_{y(n-1)}(y_{n-1}|\zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n})\, dy_{n-1}.$$

Equation (4.3) implies that

$$(4.12)\qquad f_{y(n-1)}(y_{n-1}|\zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n}) = \frac{g_{n-1}(z_n - {}_{n-1|}q_x\; e^{-y_n}, y_{n-1})}{P(\zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n})}$$

If we now make the following approximation (see the next section for some justifications)

$$(4.13)\qquad f_{y(n)}(y_n|y(n-1)=y_{n-1}, \zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n}) \cong \\ \cong f_{y(n)}(y_n|y(n-1)=y_{n-1}),$$

then equation (4.11) becomes

$$(4.14)\qquad f_{y(n)}(y_n|\zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n}) \cong \int_{-\infty}^{\infty} f_{y(n)}(y_n|y(n-1)=y_{n-1}) \times \\ \times \frac{g_{n-1}(z_n - {}_{n-1|}q_x\; e^{-y_n}, y_{n-1})}{P(\zeta_{n-1} \le z_n - {}_{n-1|}q_x\; e^{-y_n})}\, dy_{n-1}$$

Finally substituting this last expression (4.14) into (4.10), we obtain (4.5).

To obtain the starting value (4.6), we simply have to note that:

$$(4.15)\qquad \zeta_1 = q_x\; e^{-y(1)}$$

and that

(4.16)
$$g_1(z_1, y_1) = P(\zeta_1 \leq z_1 | y(1) = y_1) \;\; f_{y(1)}(y_1)$$
$$= P(\zeta_1 \leq z_1 | y(1) = y_1) \;\; \phi\left(\frac{y_1 - E[y(1)]}{V[y(1)]^{.5}}\right)$$

Then, since

(4.17)
$$\zeta_1 = q_x \;\; e^{-y_1} \quad \text{if} \quad y(1) = y_1,$$

we have that

(4.18)
$$P(\zeta_1 \leq z_1 | y(1) = y_1) = \begin{cases} 1 & \text{if} \quad z_1 \geq q_x \;\; e^{-y_1} \\ 0 & \text{otherwise} \end{cases}$$

Finally, by combining (4.18) and (4.16), we obtain (4.6). This completes the derivation of (4.5) and (4.6).

Before doing numerical evaluations of approximation (4.5), it is important to study in greater details and to justify the approximation (4.13) involved here. This is done in the next section.

5. JUSTIFICATIONS

Looking at the steps leading to (4.5), we note that the result is not exact due only to approximation (4.13) made in order to obtain a recursive equation involving only known quantities. This approximation may be justified theoretically by looking at two particular correlation coefficients, one of which validates the approximation for large values of n and the other for small values of n.

5.1 Correlation between $y(n)$ and $y(n-1)$

From the subject of multivariate analysis, we know that the approximation (4.13) will be acceptable if $y(n)$ and $y(n-1)$ are highly correlated (see, for example, MARDIA, KENT and BIBBY (1979, Section 6.5)). This is true since if they are highly correlated, knowing $y(n-1)$ would explain much of $y(n)$. Now if this is the case, introducing any other variable, correlated or not with $y(n)$, in the regression model to further explain $y(n)$ cannot improve the situation much.

Looking back at the definition of $y(n)$ (see (2.3)) it is clear that $y(n-1)$ and $y(n)$ must be highly correlated. Their correlation coefficient will be given by: (ROSS (1988, p. 280))

(5.1)
$$\rho(y(n), y(n-1)) = \frac{\operatorname{cov}(y(n), y(n-1))}{\{V[y(n)] \;\; V[y(n-1)]\}^{1/2}}.$$

Note that if the force of interest is modeled by a White Noise process, i.e.

(5.2)
$$\delta_t \sim N(\Delta, \sigma_w^2),$$

where it is understood that its integral, $y(t)$, is a Wiener process, it can be shown that, the expected value of $y(t)$ is

$$E[y(t)] = \Delta \quad t \tag{5.3}$$

and its autocovariance function is

$$\operatorname{cov}(y(s), y(t)) = \sigma_w^2 \quad \min(s, t) \tag{5.4}$$

If the force of interest is modeled by the following Ornstein-Uhlenbeck process.

$$d\delta_t = -\alpha(\delta_t - \delta)\, dt + \sigma \quad dW_t, \tag{5.5}$$

with initial value δ_0, then $y(t)$ has an expected value of

$$E[y(t)] = \delta \quad t + (\delta_0 - \delta) \left(\frac{1 - e^{-\alpha t}}{\alpha}\right) \tag{5.6}$$

and its autocovariance function is

$$\operatorname{cov}(y(s), y(t)) = \frac{\sigma^2}{\alpha^2} \min(s, t) + \tag{5.7}$$

$$+ \frac{\sigma^2}{2\alpha^3} [-2 + 2e^{-\alpha s} + 2e^{-\alpha t} - e^{-\alpha|t-s|} - e^{-\alpha(t+s)}]$$

(see, PARKER (1922b, equations 38 and 39))

The correlation coefficients between $y(n)$ and $y(n-1)$ for different values of n, when the force of interest is modeled by a White Noise (see (5.2)) and when it is modeled by an Ornstein-Uhlenbeck process (see (5.5)) with parameter $\alpha = .1$, 2 or 5 are presented in Table 1

TABLE 1

CORRELATION COEFFICIENT BETWEEN $y(n)$ AND $y(n-1)$
FORCE OF INTEREST AS WHITE NOISE AND ORNSTEIN-UHLENBECK PROCESSES

n	White Noise	Ornstein-Uhlenbeck		
		$\alpha = 1$	$\alpha = 2$	$\alpha = 5$
2	7071	8773	8707	8516
3	8165	9474	9423	9270
4	8660	9701	9659	9535
5	8944	9804	9769	9664
6	9129	9860	9829	9739
7	9258	9894	9867	9788
8	9354	9916	9891	9821
9	9428	9931	9909	9846
10	9487	9942	9922	9865
20	9747	9980	9969	9940
40	9874	9992	9987	9972
60	9916	9995	9991	9981

Results for the White Noise process are presented here because this process involves i.i.d. forces of interest, therefore, leading to the lowest correlation coefficients Results for the Ornstein-Uhlenbeck process are presented because it is the process used for illustration purposes in the next section.

Note that the correlation coefficient between $y(n)$ and $y(n-1)$ is not influenced by the parameter σ_w of the White Noise process. For the Ornstein-Uhlenbeck process, the parameter δ_0, δ and σ have no incidence on the correlation coefficients

Table 1 clearly shows that $y(n)$ and $y(n-1)$ are very highly correlated, especially for large values of n. Therefore, approximation (4.13) made to obtain the recursive equation (4.5) should be acceptable

Another correlation coefficient could also justify approximation (4 13), independently of the one discussed here This is the subject of the next section.

5.2. Correlation between $e^{-y(n)}$ and ζ_n

Again from the subject of multivariate analysis, we know that the approximation (4 13) would also be acceptable if $y(n-1)$ and ζ_{n-1} contained about the same useful information to explain $y(n)$ (see, for exemple, MARDIA, KENT and BIBBY (1979, Section 6 5)). This may be investigated by studying the correlation coefficients between $e^{-y(n-1)}$ and ζ_{n-1}

If $e^{-y(n)}$ and ζ_n are highly correlated, the approximation would be reasonable. The correlation coefficient between these two random variables is: (ROSS (1988, p. 280))

$$(5.8) \qquad \varrho(e^{-y(n)}, \zeta_n) = \frac{\operatorname{cov}(e^{-y(n)}, \zeta_n)}{\{V[e^{-y(n)}]\ V[\zeta_n]\}^{1/2}}\,.$$

Using (3.1), we obtain

$$(5.9) \quad \varrho(e^{-y(n)}, \zeta_n) = \frac{\sum_{i=0}^{n-1} {}_{i|}q_x \ \operatorname{cov}(e^{-y(n)}, e^{-y(i+1)})}{\left\{V[e^{-y(n)}] \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} {}_{i|}q_x \ {}_{j|}q_x \ \operatorname{cov}(e^{-y(i+1)}, e^{-y(j+1)})\right\}^{.5}},$$

where $\operatorname{cov}(e^{-y(i)}, e^{-y(j)})$ is given by

$$(5.10) \qquad \operatorname{cov}(e^{-y(i)}, e^{-y(j)}) = E[e^{-y(i)}\ e^{-y(j)}] - E[e^{-y(i)}]\ E[e^{-y(j)}]$$

Note that if the force of interest is Gaussian, the expected values involved in (5.10) are simply the expected values of lognormal variables (see PARKER (1992b, Section 6)).

The correlation coefficients between $e^{-y(n)}$ and ζ_n, for different values of n, when the force of interest is modeled by a White Noise or an Ornstein-Uhlenbeck process with particular parameters are presented in the following table. The mortality rates used are the male ultimate rates of the CA 1980-82 mortality table (COWARD (1988, pp. 227-231)).

TABLE 2

CORRELATION COEFFICIENT BETWEEN $e^{-Y(n)}$ AND ζ_n
FORCE OF INTEREST AS WHITE NOISE AND ORNSTEIN-UHLENBECK PROCESSES

n	White Noise $\Delta = 06$, $\sigma_w = 01$ $x = 30$	Ornstein-Uhlenbeck $\delta = 06$, $\delta_0 = 1$, $\alpha = 1$		
		$\sigma = 01$ $x = 30$	$\sigma = 02$ $x = 30$	$\sigma = 01$ $x = 50$
1	1 0000	1 0000	1 0000	1 0000
2	9447	9899	9899	9912
3	9199	9824	9824	9849
4	9064	9770	9770	9802
5	8980	9728	9727	9765
6	8925	9693	9692	9735
7	8890	9665	9663	9708
8	8868	9642	9638	9684
9	8856	9622	9617	9662
10	8851	9605	9599	9641
20	8969	9535	9518	9455
40	8999	9368	9321	8693
60	8486	8730	8494	—

Note that $\varrho(e^{-Y(1)}, \zeta_1)$ is 1 This implies that approximation (4.13) is exact for $n = 2$. The correlation coefficients of Table 2 suggest that the approximation should be good, especially for small values of n.

Combining the two conclusions drawn from the results presented in Table 1 and Table 2, we note that the approximation should be acceptable for all values of n

Now that approximation (4 5) appears to be justified, we may use it to find the distribution of ζ_n. Equations (4 4) and (4 5) may be computed by numerical integration or by some discretization method Although some methods are certainly more accurate than others, it is not our intention in this paper to discuss or compare the possible methods In the next section, we present some results obtained by an arbitrarily chosen discretization of (4.5)

6. ILLUSTRATIONS

Figure 1 illustrates the cumulative distribution function of ζ_n, $n = 5, 10, 15, 20$ and 25, the limiting average cost per policy for temporary insurance contracts issued at age 30 and with the force of interest modeled by a Ornstein-Uhlenbeck process with parameters $\delta = 06$, $\delta_0 = .1$, $\alpha = .1$ and $\sigma = .01$. The mortality rates are again the male ultimate rates of the CA 1980-82.

The range of possible values for ζ_5 is much shorter than the one for ζ_{25}. This is due to the fact that with a limiting portfolio, there is no fluctuation due to mortality, and therefore, all the possible variations in the random variable ζ_n are caused by the force of interest. When there are only five years of fluctuating force of interest involved, it is clear that the results will be less spread than when there are 25 years of fluctuating force of interest. Finally, it should be obvious why ζ_{25} takes larger values than ζ_5.

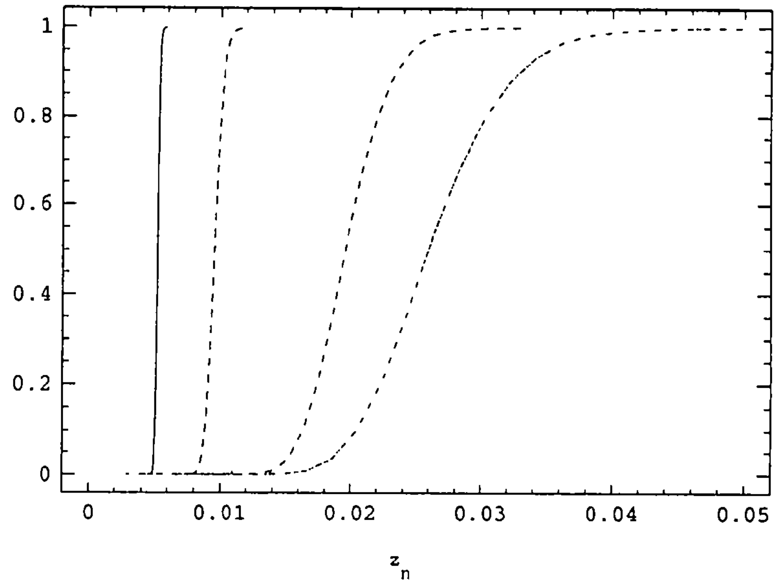


FIGURE 1 Cumulative distribution function of ζ_n
Temporary insurance policies issued at age 30, Ornstein-Uhlenbeck $\delta = 06$ $\delta_0 = 1$ $\alpha = 1$ $\sigma = 01$

——— 5 years
— — 10 years
15 years
— — 20 years
— 25 years

There is no doubt that the distribution of ζ_n provides very useful information in solvency problems. One may also be interested in using such information for pricing or valuation of a portfolio of insurance policies. In this regard, the relevant information is contained in the right tail of the distribution of ζ_n.

Table 3 contains some numerical values of the right tail of the distributions of ζ_5 and ζ_{25} illustrated in Figure 1

From Table 3, we know, for example, that a company charging a single premium of 005602 to each life insured of a very large portfolio of 5-year temporary contracts will meet its future liabilities with a probability of about 995.

TABLE 3

RIGHT TAIL OF THE APPROXIMATE DISTRIBUTION OF ζ_n, 5 AND 25 YEARS TEMPORARY INSURANCE ISSUED AT AGE 30, ORNSTEIN-UHLENBECK $\delta = 06$ $\delta_0 = 1$ $\alpha = 1$ $\sigma = 01$

5 years temporary		25 years temporary	
z_5	$F_{\zeta_5}(z_5)$	z_{25}	$F_{\zeta_{25}}(z_{25})$
005381	940609	036135	966095
005436	972183	038092	982494
005547	992830	040048	989498
005602	995229	042004	994551
005823	997927	049827	999505

7. VALIDATIONS

A validation of the results described above has been done by comparing the exact first three moments of ζ_n with its estimated first three moments from the approximate distribution.

A discretization of the variable ζ_n has been used to estimate the moments of the approximate distribution. Algebraically, the mth moment of ζ_n about the origin has been approximated by the following equation.

$$(7.1) \qquad \hat{E}[\zeta_n^m] \cong \sum_{i=0}^{h} \left(\frac{z_n[i] + z_n[i+1]}{2}\right)^m (F_{\zeta_n}(z_n[i+1]) - F_{\zeta_n}(z_n[i])),$$

where $z_n[i]$, $i = 1, 2, \ldots, h$ is the ith ordered value of ζ_n at which F_{ζ_n} was evaluated. For the illustrations presented above, h was chosen to be 25. To deal with the extremities of the distributions the following values were arbitrarily defined as.

$$(7.2) \qquad z_n[0] = z_n[1] - \left(\frac{z_n[2] - z_n[1]}{2}\right)$$

$$(7.3) \qquad z_n[h+1] = z_n[h] + \left(\frac{z_n[h] - z_n[h-1]}{2}\right)$$

$$(7.4) \qquad F_{\zeta_n}(z_n[0]) = 0$$

$$(7.5) \qquad F_{\zeta_n}(z_n[h+1]) = 1$$

The exact moments of ζ_n about the origin may be obtained by using the definition of ζ_n given by (3.1). Its mth moment about the origin is then given by

$$(7.6) \qquad E[\zeta_n^m] = E\left[\left(\sum_{i=0}^{n-1} {}_{i|}q_x \; e^{-y(i+1)}\right)^m\right].$$

Now, with m equal 1, the first moment is

$$(7.7) \qquad E[\zeta_n] = \sum_{i=0}^{n-1} E[{}_{i|}q_x \; e^{-y(i+1)}]$$

With m equal 2, the second moment is

$$(7.8) \qquad E[\zeta_n^2] = E\left[\left(\sum_{i=0}^{n-1} {}_{i|}q_x \; e^{-y(i+1)}\right)\left(\sum_{j=0}^{n-1} {}_{j|}q_x \; e^{-y(j+1)}\right)\right]$$

$$(7.9) \qquad = E\left[\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} {}_{i|}q_x \; {}_{j|}q_x \; e^{-y(i+1)-y(j+1)}\right]$$

$$(7.10) \qquad = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} {}_{i|}q_x \; {}_{j|}q_x \; E[e^{-y(i+1)-y(j+1)}].$$

With m equal 3, the third moment is

$$(7.11) \qquad E[\zeta_n^3] = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} {}_{i|}q_x \; {}_{j|}q_x \; {}_{k|}q_x \; E[e^{-y(i+1)-y(j+1)-y(k+1)}]$$

Note that the moments of ζ_n are exactly the limiting moments of the average cost per policy studied in PARKER (1992b)

Table 4 presents, for different terms of temporary insurance contracts issued at age 30, the exact moments of ζ_n, $E[\zeta_n^m]$, and the difference between the exact and the estimated moments (given by (7.1)), i.e. $E[\zeta_n^m] - \hat{E}[\zeta_n^m]$, for m equal 1, 2 and 3. The force of interest is modeled by an Ornstein-Uhlenbeck process with parameters $\delta = 06$, $\delta_0 = .1$, $\alpha = 1$ and $\sigma = .01$.

TABLE 4

COMPARISON OF EXACT AND APPROXIMATE MOMENTS OF ζ_n, n-YEAR TEMPORARY INSURANCE ISSUED AT AGE 30, ORNSTEIN-UHLENBECK $\delta = 06$ $\delta_0 = 1$ $\alpha = 1$ $\sigma = 01$

n	$E[\zeta_n^m]$			$E[\zeta_n^m] - \hat{E}[\zeta_n^m]$		
	$m = 1$ ($\times 10$)	$m = 2$ ($\times 100$)	$m = 3$ ($\times 1000$)	$m = 1$ ($\times 10$)	$m = 2$ ($\times 100$)	$m = 3$ ($\times 1000$)
1	01197	00014	00000	00000	00000	00000
2	02284	00052	00001	00000	00000	00000
3	03291	00108	00004	00000	00000	00000
4	04246	00180	00008	− 00001	00000	00000
5	05160	00266	00014	− 00003	00000	00000
10	09517	00909	00087	− 00017	− 00004	− 00001
15	14163	02023	00292	− 00031	− 00011	− 00003
20	19731	03964	00811	− 00041	− 00024	− 00009
25	26356	07167	02013	− 00054	− 00053	− 00030

Note that, in order to present more significant digits, the first moment has been multiplied by 10, the second moment multiplied by 100 and the third moment multiplied by 1000

From Table 4, we note that the exact and approximate first three moments of ζ_n agree to at least four, five and six decimal places respectively (for $n \leq 25$). This is excellent, especially if one considers that many approximations were involved before obtaining the estimated moments of ζ_n, $\hat{E}[\zeta_n]$.

Let the relative error for the mth moment of ζ_n be:

$$(7\ 12) \qquad \frac{|E[\zeta_n^m] - \hat{E}[\zeta_n^m]|}{E[\zeta_n^m]}$$

Then, for any term, n, the relative error on the expected value of ζ_n is about .2% or less. For its second moment, it is about .7% or less. And for its third moment, it is about 1.5% or less

The results for other parameters of the Ornstein-Uhlenbeck process and for other ages at issue, not illustrated here, were all excellent The maximum relative error observed, generally for the third moment, being about 3%. Although for the illustrations presented here, the error is always negative, for other situations it may be positive or even alternate over different ranges of values of the term, n. In all cases, however, the relative error is small.

From the justifications made in Section 5 and from the validations presented here, it appears that the approximation (4.13) suggested to obtain the resursive equation (4 5) has to be highly acceptable.

8 CONCLUSION

The results of this paper provides a way of approximating the distribution of limiting portfolios that is valid for any process for the force of interest as long as the conditional density function of $y(n)$ given $y(n-1)$ is known and expression (5.10) can be evaluated As indicated earlier, choosing a Gaussian process simplify things considerably

Although equation (4.5) might not be acceptable for any random variables, the very nature of the problem under consideration here, i.e. the present value of future benefits, has some particular properties which imply that the approximation is good The worse possible case for Gaussian interest rates is when they are independent, i e White Noise process Even in this case, the correlation resulting between consecutive present value functions is fairly high.

There is no doubt that knowing the distribution of the average cost per policy is useful for pricing, valuation, solvency and reinsurance The approximation suggested in this paper is certainly accurate enough for most situations one may encounter, it is more justifiable and less subjective than the testing of a limited number of scenarios and it avoids the extremely lengthy simulations required to obtain reasonable information about the tail of the distribution

ACKNOWLEDGEMENT

Comments from an anonymous referee are gratefully acknowledged.

REFERENCES

BEEKMAN J A and FUELLING C P (1990) Interest and Mortality Randomness in Some Annuities *Insurance Mathematics and Economics* **9**, 185–196

BOYLE P P (1976) Rates of Return as Random Variables *JRI* **XLIII**, 693–713

CHUNG K L (1974) *A Course in Probability Theory* Second edition, 365 pp, Academic Press, New York

COWARD L E (1988) *Mercer Handbook of Canadian Pension and Welfare Plans* 9th edition, 337 pp, CCH Canadian, Don Mills

DEVOLDER P (1986) Opérations Stochastiques de Capitalisation *ASTIN Bulletin* **16S**, S5–S30

DHAENE J (1989) Stochastic Interest Rates and Autoregressive Integrated Moving Average Processes *ASTIN Bulletin* **19**, 131–138

DUFRESNE D (1988) Moments of Pension Contributions and Fund Levels when Rates of Return are Random *Journal of the Institute of Actuaries* **115, part III**, 535–544

DUFRESNE D (1990) The Distribution of a Perpetuity, with Applications to Risk Theory and Pension funding *Scandinavian Actuarial Journal*, 39–79

FREES E W (1990) Stochastic Life Contingencies with Solvency Considerations *Transaction of the Society of Actuaries* **XLII**, 91–148

GIACOTTO C (1986) Stochastic Modelling of Interest Rates Actuarial vs Equilibrium Approach *Journal of Risk and Insurance* **53**, 435–453

MARDIA K V, KENT J T and BIBBY J M (1979) *Multivariate Analysis*, 463 pp, Academic Press London

MELSA J L and SAGE A P (1973) *An Introduction to Probability and Stochastic Processes*, 403 pp, Prentice-Hall, New Jersey

MORRISON D F (1990) *Multivariate Statistical Methods* 3rd edition, 586 pp, McGraw-Hill Inc, New York

PANJER H H and BELLHOUSE D R (1980) Stochastic Modelling of Interest Rates and Applications to Life Contingencies *Journal of Risk and Insurance* **47**, 91–110

PARKER G (1992a) An Application of Stochastic Interest Rates Models in Life Assurance, 229 pp, Ph D thesis, Heriot-Watt University

PARKER G (1992b) Moments of the present value of a portfolio of policies To appear in *Scandinavian Actuarial Journal*

ROSS S (1988) *A First Course in Probability* 3rd edition, 420 pp, MacMillan, New York

WATERS H R (1978) The Moments and Distributions of Actuarial Functions *Journal of the Institute of Actuaries* **105, Part 1**, 61–75

WILKIE A D (1976) The Rate of Interest as a Stochastic Process-Theory and Applications *Proc 20th International Congress of Actuaries, Tokyo* **1**, 325–338

GARY PARKER
Department of Mathematics and Statistics, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

WORKSHOP

BONUS MADE EASY [1]

BY JON HOLTAN

University of Oslo & Samvirke Insurance Company Ltd

ABSTRACT

The paper introduces an alternative approach to the traditional experience rating theory in automobile insurance The approach is based on a simple theory of how high deductibles financed by loans maintain the risk differentiation in an automobile insurance arrangement Thus the approach differs totally from the usual bonus-malus classes as well as from the credibility based experience rating ideas. The paper is of a theoretical nature and leads up to a mathematical description of how the approach may be optimalized within the framework of a risk model.

KEYWORDS

Bonus-malus systems; optimal deductibles financed by loans.

1. BACKGROUND

From a practical point of view it is well-known that the existing automobile bonus-malus systems possess several considerable disadvantages which are difficult, or even impossible, to handle within the traditional theory of experience rating. The aim of this paper is to introduce an alternative bonus-malus approach which, at least theoretically, eliminates the most important ones of these disadvantages.

2. CRITICISM OF EXISTING BONUS SYSTEMS

To motivate the new bonus-malus (B-M) approach it is appropriate to stress the usual criticism of the existing B-M systems. In particular, the existing systems are, among other things, based on two general characteristics:

(i) The claim amounts are omitted as a posterior tariff criterion
(ii) At any time the policyholders may leave an insurance company without any further financial commitments to the company.

These characteristics lead to three of the most considerable disadvantages:

(2.1) Regarding an occurred claim, the future loss of bonus will in many cases exceed the claim amount.

[1] An earlier version of this work has been presented at the ASTIN Colloquium, Stockholm 1991

(2.2) The systems create the possibility of malus evasion, that is, the possibility of the policyholders leaving the insurance company to avoid premium increase because of occurred claims.
(2.3) The systems stimulate a slide towards higher average discount rates in the insurance arrangements.

Because only the number of claims (and of course the discount rate) in an insurance period determines the premium in the following period, it follows that (2.1) is an immediate consequence of (i) In many cases (2.1) gives the policyholder a feeling of unfairness, especially if the loss of bonus is much higher than the occurred claim amount. A consequence of this is the well-known bonus hunger behaviour of the policyholders.

Disadvantage (2.2) is of course a consequence of (ii). Malus evaders let the remaining policyholders pay the bill for their (the evaders') claim costs. This has, at least in Norway, been a serious problem in the insurance industry, mainly because of an unsatisfactory exchange of bonus information between the insurance companies

Because all insurance arrangements attached to existing B-M systems are exposed to bonus hunger as well as malus evasion, it follows that (2.3) is a secondary consequence of (2.1) and (2.2). A higher average rate of discount is contrary to risk differentiation, which is the objective of all B-M systems. In an extreme situation the result might be that the great majority of the policyholders are at, or close to, the maximum rate of discount

A number of authors have focused on the disadvantages mentioned above, in particular the problem of bonus hunger – see e.g. NORBERG (1975), LEMAIRE (1985) (Chapter 18) and SUNDT (1989). The aim of these authors has not been to solve or eliminate the disadvantages, but rather to take them into the modelling account in connection with the mathematical optimalization of the B-M systems. However, to eliminate the disadvantages one probably has to leave the traditional framework of experience rating, and construct a bonus principle which is basically different. This is precisely the intention of this paper, and in Section 3 we will first introduce the alternative B-M idea, and thereafter place the idea into a mathematical description and notation. The alternative approach may be called a new premium system, and in Section 4 it is shown how the system may be optimalized within the framework of a risk model. In Section 5 some practical deficiencies of the system are discussed, and in Section 6 some concluding remarks are given.

3. AN ALTERNATIVE APPROACH TO EXISTING BONUS SYSTEMS

3.1. Preliminary aspects and assumptions

The fundamental principle of the existing B-M systems simply expresses that the higher the claim frequency of a policyholder, the higher the insurance costs that on average are charged to the policyholder However, this principle is also valid in an insurance arrangement *consisting of a high maximum deductible which is common to all policyholders* This follows from the simply fact that good drivers will pay fewer deductibles than bad drivers. Thus we may imagine a premium system where the costs of the incurred deductibles are defined as the malus (the loss of bonus) after a claim occurred. Within this framework it seems natural to assume an individual risk premium above the maximum deductible which is reflected by a priori tariff criteria, but not by a posteriori knowledge about the policyholders. This system defines a malus system rather than a bonus system. However, we may interpret the claim free driving bonus as *avoidance of deductibles*

Two questions are now appropriate:

(3.1) In what way do we determine the size of the maximum deductible?

To attain a suitable cost differentiation in the risk heterogeneous arrangement, the maximum deductible has to be relatively high, maybe as high as 2000-3000 US dollars (USD). This leads to question number 2

(3 2) How do we act when knowing that the average policyholder hardly manages (at least in Norway) to cash pay deductibles of more than about 1000 USD?

Let us first look at the latter problem. The new system solves problem (3.2) by giving the policyholders a possibility of financing the incurred deductibles by loans from the insurer. Moreover, this leads to the advantage of smoothing the "loss of bonus" (the deductible) over a period of time, precisely the way that the total loss of bonus is smoothed in the traditional systems.

Before commenting on problem (3.1), we shall illustrate the abovesketched premium system with a simple example: Let us assume that a policyholder has two occurred claims of respectively 5000 USD and 500 USD in periods number 3 and 9 during an insurance period of 15 years We also assume for simplicity that the deductible loans are ordinary term loans, and that the period of repayments is 5 years. Assume the maximum deductible to be, for instance, 2000 USD, and the premium for large claims above this maximum deductible to be 300 USD during the whole insurance period. Finally, the borrowing rate is assumed to be 10% in arrears. These assumptions lead to a sequence of payments for the policyholder shown in Figure 1 We note that the effect of the alternative system is not essentially different from the effect of a traditional B-M system; the insurance costs increase in the period(s) following an occurred claim. We also note that the loss of bonus is differentiated regarding the size of the claim amounts. Or to be more precise; the loss of bonus will never (except for the interest on the loan) exceed the claim amount, and *hence the bonus hunger effect is eliminated.* In theory the new system will not be exposed to malus evasion either, because the loan is repayed even if the insurance is terminated – see Section 5 for a further discussion on this. Hence, at least theoretically the new system eliminates the disadvantages (2.1), (2.2) and (2.3) in Section 2.

Return to problem (3.1). The solution of this problem ought to be linked to a mathematical optimalization of the system. In addition to problem (3.1), we have to decide a) the amortization form of the deductible loans, b) the length of the repayment period, and c) the rate of interest The conditions a), b) and c) are in practice given by the money market. Thus it may seem meaningless to find mathematical "optimal" lending conditions. However, these conditions will never be absolute, therefore it may be after all interesting to find optimal values at least for some of the conditions.

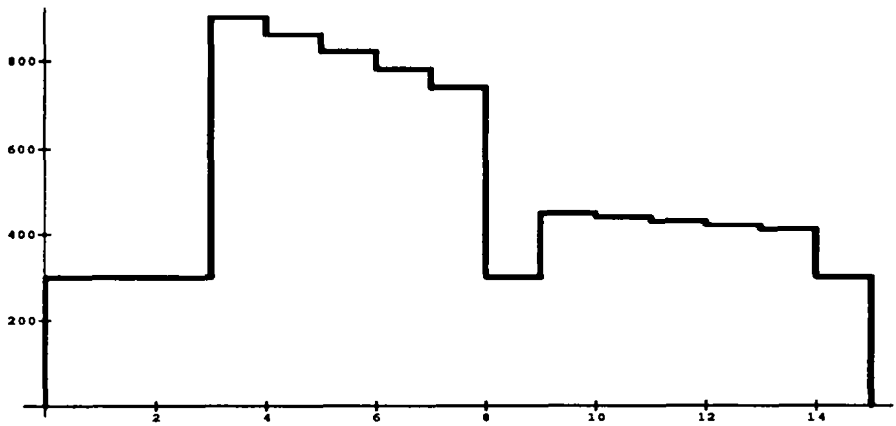


FIGURE 1 The payments for the policyholder over a period of 15 years

Now, stress item a), the amortization form of the loans In principle we ought to choose an amortization form which imitates the *traditional* influence of the premiums in the time periods following a claim. More precisely, an amortization form where the repayments are high during the first periods following a claim and then gradually fall. Moreover, this satisfies the desire of the policyholders to repay most of the claim costs shortly after the claim has occurred. Within *annuity loans* the repayments are exactly the same in the repayment period, while the repayments are not decreasing enough within ordinary *term loans*. Hence, these alternatives of the amortization form are ignored. However, there exists an alternative fulfilling all the mentioned properties, that is, the *exponential amortization form*. This form is also relatively handy in the mathematical computations

Before touching the mathematical description of the alternative system, one last assumption concerning the financing of the deductibles has to be made. In a practical application of the new system it is of course the policyholders who decide how much to pay cash, and how much to borrow Hence, a deductible is partially financed by a cash payment greater than or equal to zero, and partially by a sum borrowed from the insurer However, to simplify the mathematical analysis we assume the *entire* deductible of an occurred claim to be financed by a loan. This is an advantage because the costs are then smoothed over a period of time. In addition, a full-financing by loans is computationally easier to analyse.

3.2. Mathematical description

Assume the following mathematical description of the alternative system: Let Y_i; $i = 1, 2, \ldots$ be the values at time zero of the claim amounts of a policyholder that occurred at the time points T_i; $i = 1, 2, \ldots$, respectively. Let Z_i be the value at time zero of the amount payed by the policyholder of claim number i, and assume Z_i on the ordinary excess-of-loss form

$$Z_i = \min(Y_i, b), \tag{1}$$

where b is interpreted as the value at time zero of the common *maximum deductible* of all policyholders at time T_i.

Let π be the inflation discount intensity related to the values at time zero of the claim amounts Hence it follows that the *future nominal value* of Z_i at time T_i is $Z_i \exp(\pi T_i)$. Note besides that the deductible (b at time zero) is thought of as following the inflation intensity π.

Let $Z_i \exp(\pi T_i)$ be fully financed by a loan from the insurer. The loan is charged a rate of interest δ and continuously amortized by a stream of payment $\{r_i(s); s \geq 0\}$, where $s = 0$ refers to the time T_i of the claim occurrence.

The payment stream of loan number i has to satisfy (see e.g. GERBER (1990), Chapter 1)

$$Z_i \exp(\pi T_i) = \int_0^\infty v^s r_i(s)\, ds, \tag{2}$$

where $v^s = \exp(-\delta s)$ = the interest discount factor at time s.

Let $N(t)$ be the number of claims occurred in the time interval $(0, t]$. Then

$$r(t) = \sum_{i=1}^{N(t)} r_i(t - T_i) \tag{3}$$

is the *amortization rate* of the policyholder at time t.

Assume an exponential form of amortization, that is,

$$r_i(s) = B_i \exp(-\rho s). \tag{4}$$

B_i is here called "the initial amortization level", and may be interpreted as *interest + repayments* in the first repayment year. When the rate of interest δ is known, ρ expresses the *amortization profile* of the sums borrowed, that is, the obliquity of the repayments, or to which extent the repayments should be high in the beginning and then gradually decreasing.

From (2) and (4) we obtain

$$Z_i \exp(\pi T_i) = \int_0^\infty \exp(-\delta s)\, B_i \exp(-\rho s)\, ds$$

$$= \frac{B_i}{\delta + \rho},$$

or

$$B_i = Z_i \exp(\pi T_i)(\delta + \rho). \tag{5}$$

Formula (5) gives the relationship between ρ and "the initial amortization level" B_i when the rate of interest δ and the sum borrowed $Z_i \exp(\pi T_i)$ are known. In particular, we see that $\rho = 0$ (constant amortization) implies $B_i = \delta Z_i \exp(\pi T_i)$, which means solely repaying interest to infinity. Henceforth, we will assume $\rho \geq 0$.

From (4) and (5) we have

$$r_i(s) = Z_i(\delta+\rho)\exp(\pi T_i)\exp(-\rho s). \tag{6}$$

Therefore, from (3) we finally obtain the expression

$$r(t) = \sum_{i=1}^{N(t)} Z_i(\delta+\rho)\exp(\pi T_i - \rho(t-T_i)). \tag{7}$$

To obtain an impression of the effect of ρ, it may be suitable to take a closer look at the function (6). Under assumptions of $\delta = 10\%$ and $Z_i \exp(\pi T_i) = 1$, Figure 2 shows the stream of payments $r_i(s)$ for some specified values of ρ. Note that the higher ρ is, the higher the payments are during the first repayment period(s). In the case of $\rho = 0$, we see that only 10% interest of $Z_i \exp(\pi T_i) = 1$ is continuously payed

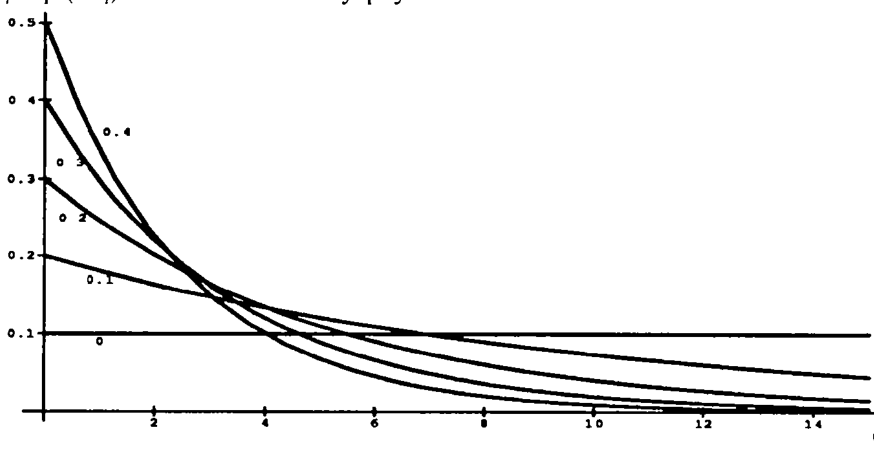


FIGURE 2 The stream of payments $\{r_i(s), s \geq 0\}$ when $\rho = \{0, 0.1, 0.2, 0.3, 0.4\}$

4. A MATHEMATICAL OPTIMALIZATION DESIGN

4.1. Model assumptions

To carry through an optimalization of the new system, a claim risk model has to be built. In this paper we assume the widely accepted negative binomial model, see e.g LEMAIRE (1991):

> The claim number process $\{N(t);\, t \geq 0\}$ of a policyholder is a homogeneous Poisson process given the claim intensity Θ. Let Θ follow a gamma distribution Gamma (α, β). Assume also the values at time zero $Y_1, Y_2, \ldots$ of the claim amounts to be independent and identically distributed (i.i.d.), and independent of $\{N(t);\, t \geq 0\}$ and of Θ.

Under these assumptions we also easily establish the values at time zero of the sums borrowed, $\{Z_i = \min(Y_i, b);\ i = 1, 2, \ldots\}$, to be i.i d. and independent of $\{N(t),\ t \geq 0\}$ and of Θ.

4.2. Choice of loss function

Within the risk model in subsection 4.1 and the mathematical description in subsection 3.2, we want to minimize an *expected loss function* to find some optimal parameter values of the system

The theoretical individual risk intensity of the policyholder at time t is easily evaluated as $Q(t) = \exp(\pi t)\Theta EY$. Now, the point is to estimate $Q(t)$ using a loss function which includes the amortization rate $r(t)$. In a real application of the system we have already indicated the suitability of a constant *individual* premium for all risks above the maximum deductible. For simplicity, we henceforth disregard this individual differentiation, and instead we assume a constant *collective* premium. Hence, let $p(t)$ be this premium of large claims at time t:

$$p(t) = \exp(\pi t)p = \exp(\pi t)\ E\Theta E(Y-Z), \tag{8}$$

where Y and Z are the values at time $t = 0$ of the random claim amount and the random sum borrowed, respectively Now, write

$$\Theta EY = \Theta EZ + \Theta E(Y-Z).$$

Then one can interpret $p(t)$ as an estimator of $\exp(\pi t)\ \Theta E(Y-Z)$. If we now just let $r(t)$ be an estimator of $\exp(\pi t)\Theta EZ$ and use the *traditional* expected quadratic loss function

$$E[p(t)+r(t)-Q(t)]^2,$$

we will in the first place obtain a loss expression dependent on the time t, which is not a desirable situation. In the second place $r(t)$ would not alone be a sufficiently good estimator of $\exp(\pi t)\Theta EZ$. Owing to the fact that the loss of bonus (the sums borrowed) is payed in arrears, the amortization rate $r(t)$ is too small during the first periods according to the true intensity $\exp(\pi t)\Theta EZ$.

However, to solve these problems we may construct a loss function which integrates the total cash flow of the policyholder over a period of time The actual loss function ought to reflect the total financing of a) the large claim risks and of b) all deductibles occurred in the actual optimalization period.

The following expected quadratic loss function takes care of the mentioned objections in a reasonable way.

$$E\left[\int_0^M v^t(p(t)+r(t))\,dt + v^M S(M) - \int_0^M v^t Q(t)\,dt\right]^2, \tag{9}$$

where

M = a restricted time horizon.

v^t = $\exp(-(\pi+\omega)t)$ = total discount factor at time t, with the *inflation* discount intensity π and a *mathematical weight* discount intensity ω $\exp(-\omega t)$ is hereby interpreted as a weight function; we see e.g. that $\omega = 0$ implies a uniform weight function over the time period $(0, M]$.

$p(t)$ = $\exp(\pi t)\, E\Theta E(Y-Z)$
= the large claim premium at time t

$r(t)$ = $\sum_{i=1}^{N(t)} Z_i(\delta+\rho)\exp(\pi T_i - \rho(t-T_i))$
= the amortization rate of the policyholder at time t.

$Q(t)$ = $\exp(\pi t)\,\Theta EY$
= the theoretical risk intensity at time t

$S(M)$ = $\sum_{i=1}^{N(M)} \int_M^\infty \exp(-\pi(t-M))\, r_i(t-T_i)\, dt$
= the value at time M of all future repayments caused by claims occurred in $(0, M]$.

Summary:

Loss function (9) may be interpreted as the expected quadratic deviation between a mathematical value at time zero of the actual cash flow of the policyholder and the corresponding mathematical value at time zero of the theoretical risk intensity of the policyholder over the time period $(0, M]$. Note that all raised loans during $(0, M]$ have to be repayed, and hence one has to include $v^M S(M)$ in the loss function.

4.3. Computation of the expected loss function

To minimize (9) analytically or numerically with respect to e.g. the system parameters δ, ρ and b, the function has to be of algebraic nature To obtain an algebraic form of (9) some statistical computations have to be made.

Let

$$Z(t) = \sum_{i=1}^{N(t)} Z_i \exp((\pi+\rho)T_i). \tag{10}$$

Then by (7)

$$r(t) = (\delta+\rho)\exp(-\rho t)\, Z(t), \tag{11}$$

and by simple algebra we obtain

$$v^M S(M) = \left(\frac{\delta+\rho}{\pi+\rho}\right)\exp(-(\pi+\omega+\rho)M)\, Z(M). \tag{12}$$

Introduce the annuity

$$\bar{a}_{\overline{M}|} = \int_0^M \exp(-\omega t)\, dt = \omega^{-1}(1-\exp(-\omega M)),$$

and the expression

$$\Psi = \int_0^M v^t r(t)\, dt = (\delta+\rho) \int_0^M \exp(-(\pi+\omega+\rho)t)\, Z(t)\, dt.$$

Then function (9) may be written as

$$\begin{aligned}(13) \qquad & E\Psi^2 + 2E[\Psi v^M S(M)] + E[v^M S(M)]^2 + \\ & + 2\bar{a}_{\overline{M}|} E[(\Psi + v^M S(M))(p - \Theta EY)] + \\ & + \bar{a}^2_{\overline{M}|}[p^2 - 2p(E\Theta)(EY) + E\Theta^2 (EY)^2].\end{aligned}$$

By (13) we have to find the 1.- and 2.-order moments of the $Z(t)$-process, that is $EZ(t)$ and $E[Z(s)\, Z(t)]$. However, the stochastic process $Z(t)$ does not have independent waiting times between steps, and hence the calculations become somewhat complex. We may however show that $Z(t)$ has the same distribution as

$$(14) \qquad Z^* = \sum_{i=1}^{N^*} Z_i \exp((\pi+\rho)\, U_i),$$

where

$$\begin{gathered}\text{given } \Theta = \theta,\ N^* \sim \text{Poisson}(\theta t), \\ U_1, \ldots, U_{N^*} \text{ are i.i.d.} \sim \text{Uniform } [0, t], \\ Z_1, \ldots, Z_{N^*} \text{ are i.i.d },\end{gathered}$$

and where N^*, the U_i's and the Z_i's are stochastically independent. This result was in general discovered by JUNG (1963); see also BÜHLMANN (1970), pp. 57-60 By standard statistical calculations we then obtain

$$(15) \qquad EZ(t) = E\Theta \frac{EZ}{(\pi+\rho)} [\exp((\pi+\rho)t) - 1],$$

and for $0 \le s \le t$

$$\begin{aligned}(16) \quad E[Z(s)\, Z(t)] = {} & E\Theta \frac{EZ^2}{2(\pi+\rho)} \text{k} [\exp(2(\pi+\rho)s) - 1] + \\ & + E\Theta^2 \frac{(EZ)^2}{(\pi+\rho)^2} [\exp((\pi+\rho)s) - 1]\, [\exp((\pi+\rho)t) - 1].\end{aligned}$$

To obtain an algebraic form of the expected loss function (13), one has to complete seven isolated computations. Below, these computations are noted as

$\psi_1, \ldots, \psi_7$ (remember the integral definition of Ψ):

(17) $\psi_1 = E\Psi^2$

(18) $\psi_2 = E[\Psi v^M S(M)]$

(19) $\psi_3 = E[v^M S(M)]^2$

(20) $\psi_4 = E(\Theta\Psi)$

(21) $\psi_5 = E[\Theta v^M S(M)]$

(22) $\psi_6 = E\Psi$

(23) $\psi_7 = E[v^M S(M)]$.

In this paper we restrict ourselves to indicate that (17)-(23) are easily calculated by use of standard statistical methods. The clue is here to use the expressions (15) and (16). Thus, for instance, we have

$$\psi_1 = E\Psi^2 = (\delta+\rho)^2 E\left[\int_0^M Z(t)\exp\left(-(\pi+\omega+\rho)t\right)dt\right]^2$$

$$= (\delta+\rho)^2 \int_0^M ds \int_s^M E[Z(s)\,Z(t)]\exp\left(-(\pi+\omega+\rho)(s+t)\right)dt\,.$$

Finally, we establish the expected loss function

$$E\left[\int_0^M v^t(p(t)+r(t))\,dt + v^M S(M) - \int_0^M v^t Q(t)\,dt\right]^2 \tag{24}$$

$$= \psi_1 + 2\psi_2 + \psi_3 + 2\,\bar{a}_{\overline{M}|}[p\psi_6 - EY\psi_4 + p\psi_7 - EY\psi_5] +$$
$$+ \;\bar{a}^2_{\overline{M}|}[p^2 - 2pE\Theta EY + E\Theta^2(EY)^2]\,.$$

4.4. Comments on the loss function

Under the model assumptions of subsection 4.1 we have

$$E\Theta = \alpha/\beta, \qquad E\Theta^2 = \alpha(\alpha+1)/\beta^2.$$

If the claim amount distribution is assumed known, the function (24) depends on eight unknown parameters. Two of them, α and β, can e.g. be estimated by the maximum likelihood estimators described by LEMAIRE (1985), Chapter 12 Further, it seems natural to keep the inflation intensity π, the mathematical weight intensity ω and the time horizon M constant (they might also be considered as random variables). Thus the actual optimalization (varying) parameters are the remaining *system parameters* δ, ρ and b.

In this connection, *analytical* optimal parameter solutions are in general difficult to find. However, *numerical* solutions are easily computed by a computer system, for example the mathematical software system Mathematica. Note that the maximum deductible b enters into the function (24) via the moments EZ and EZ^2 Thus, an approximating optimalization of b demands a statistical analysis of the claim amounts in a representative claim portfolio. Also the premium of large claims, $p(t)$, has to be estimated in association with a real claim portfolio

Note finally that the alternative premium system may be mathematically compared with traditional B-M systems via the expected loss function (9). Or to be more precise; within each of the traditional B-M systems one may construct an estimator to the estimand $\int_0^M v^t Q(t)\,dt$. By using these estimators in loss function (9), we are able to compare the expected losses of the traditional B-M systems with the expected loss of the alternative system, and hence find the best mathematically fitted system

4.5. The loss function for the special case M = ∞

To give some more information on the structure of the loss function, one may exhibit the function for the special case when the time horizon M tends to infinity Assume in this case that $\omega > 0$, which is in accordance with economic theory. When $M = \infty$, we see from (12), (16) and (19) that ψ_3 tends to zero. By (18), (21) and (23) then also ψ_2, ψ_5 and ψ_7 tend to zero. In formula (24) thus only ψ_1, ψ_4 and ψ_6 remain different from zero. Straightforward calculation gives

$$\psi_1 = \frac{1}{2\omega}\left(\frac{\delta+\rho}{\pi+\omega+\rho}\right)^2\left[E\Theta EZ^2 + \frac{2}{\omega}E\Theta^2(EZ)^2\right],$$

$$\psi_4 = \frac{1}{\omega}\left(\frac{\delta+\rho}{\pi+\omega+\rho}\right)E\Theta^2 EZ,$$

$$\psi_6 = \frac{1}{\omega}\left(\frac{\delta+\rho}{\pi+\omega+\rho}\right)E\Theta EZ$$

Inserting $p = E\Theta(EY-EZ)$ the loss function may then be put into the following form

$$x^2 A_1(b) - 2xA_2(b) + A_3(b),$$

with

$$x = \frac{\delta+\rho}{\pi+\omega+\rho} \tag{25}$$

$$A_1(b) = \frac{1}{2\omega}\left[E\Theta EZ^2 + \frac{2}{\omega}E\Theta^2(EZ)^2\right] \tag{26}$$

$$A_2(b) = \left(\frac{1}{\omega}\right)^2 EZ[(E\Theta)^2 EZ + \mathrm{Var}\,\Theta EY] \tag{27}$$

$$A_3(b) = \left(\frac{1}{\omega}\right)^2[(E\Theta)^2(EZ)^2 + \mathrm{Var}\,\Theta\,(EY)^2]. \tag{28}$$

The influence of the system parameters δ and ρ is contained in x, and thus is separated from that of the system parameter b.

For fixed b the loss function attains its minimum for

$$x = x(b) = A_2(b)/A_1(b), \tag{29}$$

and the minimum is

$$\min(b) = A_3(b) - A_2(b^2)/A_1(b). \tag{30}$$

Denoting the claim amount c.d.f by F, we have

$$EZ = \int_0^b [1 - F(y)]\, dy \tag{31}$$

$$EZ^2 = \int_0^{b^2} [1 - F(\sqrt{y})]\, dy. \tag{32}$$

Thus EZ and EZ^2 are continuous functions of b. If F is continuous, they are also differentiable. The same is then also true for $\min(b)$. Thus, for special choices of F it should not be difficult to minimize $\min(b)$ with respect to b, and thereby obtain a global minimum.

For the moment we content ourselves with the following remarks.
By (25) optimal values of δ and ρ for fixed b are related by

$$\delta(b) = [x(b) - 1]\,\rho(b) + (\pi + \omega)\, x(b).$$

Thus the interest intensity $\delta(b)$ is greater than, equal to or less than the market interest intensity $\pi + \omega$ according as $x(b)$ is greater than, equal to or less than one.

As b tends to infinity, EZ and EZ^2 tend to EY and EY^2 respectively. From (26)-(28) we see that

$$A_2(\infty) = A_3(\infty) = A_1(\infty) - \frac{1}{2\omega} E\Theta EY^2 = \left(\frac{1}{\omega}\right)^2 E\Theta^2 (EY)^2.$$

Thus by (29), $x(\infty) < 1$.

For b tending to zero, $A_1(b)$ will be of the order of magnitude b^2. $A_2(b)$ will be of the order of magnitude b, because of the second term within the paranthesis. Thus by (29), $x(0+) = \infty$. This means that there is (at least) one b with $x(b) = 1$. From (26)-(32) it can be shown that for such a b we will have $x'(b) < 0$ and $\min'(b) > 0$, if $F(y) > 0$ for $y > 0$. This proves that there is exactly one value of b with $x(b) = 1$ and that $x(b) > 1$ to the left of this point and $x(b) < 1$ to the right of it. Furthermore, $\min(b)$ has, at least locally, a minimum to the left of the point. This indicates that the optimal δ-value is greater than $\pi + \omega$, or, in other words, the interest intensity for the loan should be greater than the market interest intensity.

5. PRACTICAL SYSTEM DEFICIENCIES

In general it is often difficult, or even impossible, to eliminate deficiencies of an existing financial market system without generating other system deficiencies The automobile insurance B-M principle seems typically to be characterized by this two-sided effect, and hence it is not difficult to point out some general practical deficiencies of the alternative B-M approach. An obvious one is that a high common deductible necessarily involves a *lower total premium income* compared with traditional bonus systems, and thereby generates a lower insurance profit to the insurer. Another deficiency is the *credit risk* of the policyholders, or, more precisely, it is not certain that the policyholders are able to repay their deductible loans. Hence, the insurer has to, in one way or another, make conditions linked to the individual solvence security in order to meet possible losses One way of doing this is e g that the insurer demands the policyholders to save an amount of money in each insurance period to build up an individual risk reserve to cover (parts of) future incurred deductibles. A "claim risk account" with the insurer should, in regard to reduce the credit risk and to maximize the rate of interest on deposits, be closed for withdrawals during the insurance periods, except for financing incurred deductibles. Thus, the premium and claim costs of the policyholders will also have a more uniform dispersion during the insurance periods.

6. CONCLUDING REMARKS

In theory the alternative B-M approach eliminates the most important disadvantages of the existing B-M systems. A policyholder will for instance within the existing systems, unlike the alternative approach, often make a profit by asking a bank for a credit to cover an occurred claim cost, instead of reporting the claim to the insurer. This seems obvious, but can also under some specified conditions be explicitly shown by comparing the effective rate of interest on a banking credit with the "effective rate of interest" on the loss of insurance bonus. By constructing a B-M approach which eliminates bonus hunger, one also avoids mathematical risk modelling which includes assumptions about bonus hunger, as e.g. NORBERG (1975), LEMAIRE (1985) (Chapter 18) and SUNDT (1989) have built into their models.

On the other hand the alternative B-M approach contains, as pointed out in Section 5, some practical deficiencies like credit risk and lower premium income. The point is however that these deficiencies *are just relevant for the (existing) insurers, and not for the policyholders.* In other words; the alternative approach is less favourable to the existing insurers than to their customers. Thus, it seems conceivable that the traditional insurance industry at once will be rather sceptical about introducing the alternative B-M approach to the insurance market. It seems, however, more probable that the possible initiators in this connection will be the (future) financial institutions—or cooperations between institutions—which consist of a *superior* banking service and a *minor* (automobile) insurance service. In the first place these institutions are generally interested in introducing customer-friendly products to increase their market share and market profit in the insurance market. In the second place, and under these circumstances, they probably interpret the problem of lower premium income as of secondary importance, while they obviously have the best qualifications to handle the problem of credit risk. Finally, and in the third place, these institutions already have the general administrative device which the alternative B-M approach demands, or stated in its extreme form, an optimal combination of actuarial and banking knowledge and culture.

ACKNOWLEDGEMENTS

The present paper is based on a thesis aimed at obtaining the degree of cand.scient. at the University of Oslo in 1991. The author is grateful to professor WALTHER NEUHAUS for valuable suggestions and encouraging supervision, and to an anonymous referee for his suggestion of including subsection 4.5.

REFERENCES

BÜHLMANN, H (1970) *Mathematical methods in risk theory* Springer Verlag, Berlin

GERBER, H U (1990) *Life Insurance Mathematics* Springer Verlag

JUNG, J (1963) A theorem on compound Poisson processes with time-dependent change variables *Skandinavisk Aktuarietidsskrift* **46**, 95–130

LEMAIRE, J (1985) *Automobile Insurance Actuarial Models* Kluwer-Nijhoff, Boston

LEMAIRE, J (1991) Negative binomial or Poisson-Inverse Gaussian? *Paper presented at the ASTIN Colloquium in Stockholm 1991*

NORBERG, R (1975) Credibility Premium Plans which Make Allowance for Bonus Hunger *Scandinavian Actuarial Journal*, 73–86

SUNDT, B (1989) Bonus hunger and Credibility estimators with geometric weights. *Insurance Mathematics and Economics* **8**, 119–126

JON HOLTAN
Samvirke Insurance Company Ltd., P.O. Box 778 Sentrum,
N-0106 Oslo, Norway.

HIGH DEDUCTIBLES INSTEAD OF BONUS-MALUS. CAN IT WORK?

BY JEAN LEMAIRE AND HONGMIN ZI

Wharton School
University of Pennsylvania

ABSTRACT

HOLTAN (1994) suggests to replace traditional bonus-malus systems by a high deductible financed by a short-term loan. Practical consequences of this proposal are investigated here. Simulation is used to evaluate the efficiency of the Taiwanese Bonus-malus system and the variability of premiums of an average policyholder. Holtan's high deductible system is analysed under a compound Poisson assumption, with truncated exponential claims. It is shown that the introduction of a high deductible would increase the variability of payments and the efficiency of the rating system for most policyholders [1].

KEYWORDS

Motor insurance rating; bonus-malus systems; deductibles.

1. INTRODUCTION

Traditional merit-rating or bonus-malus systems (BMS) suffer from two major drawbacks

(i) The severe penalties needed to compensate no-claim discounts cannot be enforced, for commercial reasons. A continuous increase of the average discount follows, until the system reaches stationarity. This forces insurers to raise premiums annually After a few years, most policies cluster in the high-discount classes, and there is no significant premium differentiation between good and bad drivers.

(ii) Penalties after an accident at fault are independent of damages. This creates a bonus-hunger phenomenon, that induces policyholders to bear small claims themselves, in order to avoid future premium increases. In some cases, it is of the policyholder's interest to pay substantial amounts to their victims. This creates a feeling of unfairness, and encourages hit-and-run behaviour

[1] The authors would like to thank Messrs Ted Chung and Chen-Yeh Lai, who kindly provided detailed information about the Taiwanese merit-rating system and loss distributions

HOLTAN (1994) suggests an ingenious alternative to BMS rating. a high-deductible system (HDS). In this system, the premium would only provide coverage for the part of the losses in excess of a high deductible D. Policyholders who cannot afford to pay this amount could borrow it from the company, and reimburse this loan over a small number of years.

The implementation of a HDS could eliminate the two main drawbacks of BMS: the premium income would not decrease over time, and, since the penalty after a claim never exceeds the claim amount (except for interest on the loan), the hunger for bonus effect would be eliminated.

In this paper, we use simulation and a simple compound Poisson model to compare Holtan's proposal to the BMS in force in Taiwan, a system which is rather "tough" to policyholders (see LEMAIRE and ZI, 1994). It is shown that high deductibles improve the efficiency of the rating system, but increase the variability of the payments, as measured by the coefficient of variation. The Taiwanese BMS is analysed in Section 2. The HDS is studied in Section 3. Practical considerations are to be found in Section 4. Section 5 summarizes findings and suggest further research.

2. ANALYSIS OF THE TAIWANESE BMS

Our benchmark policyholder is a Taiwanese driver, whose annual number of claims is Poisson distributed, with a parameter $\lambda = 0.10$. At time 0, he enters the BMS described in Table 1, in class 4.

TABLE 1

TAIWANESE BONUS-MALUS SYSTEM

Class	Premium Level	Class after 0	1	2 claims	3	4	5+
9	150	3	5	6	7	8	9
8	140	3	5	6	7	8	9
7	130	3	5	6	7	8	9
6	120	3	5	6	7	8	9
5	110	3	5	6	7	8	9
4	100	3	5	6	7	8	9
3	80	2	5	6	7	8	9
2	65	1	5	6	7	8	9
1	50	1	5	6	7	8	9

Effects of inflation are removed by assuming that premiums, losses, deductibles, ..., escalate according to the same index

The evolution of the policyholder among the classes has been simulated for 30 years, the time it takes for system to reach a stationary state Figure 1 shows that the expected premium level constantly decreases over time, reaching a level

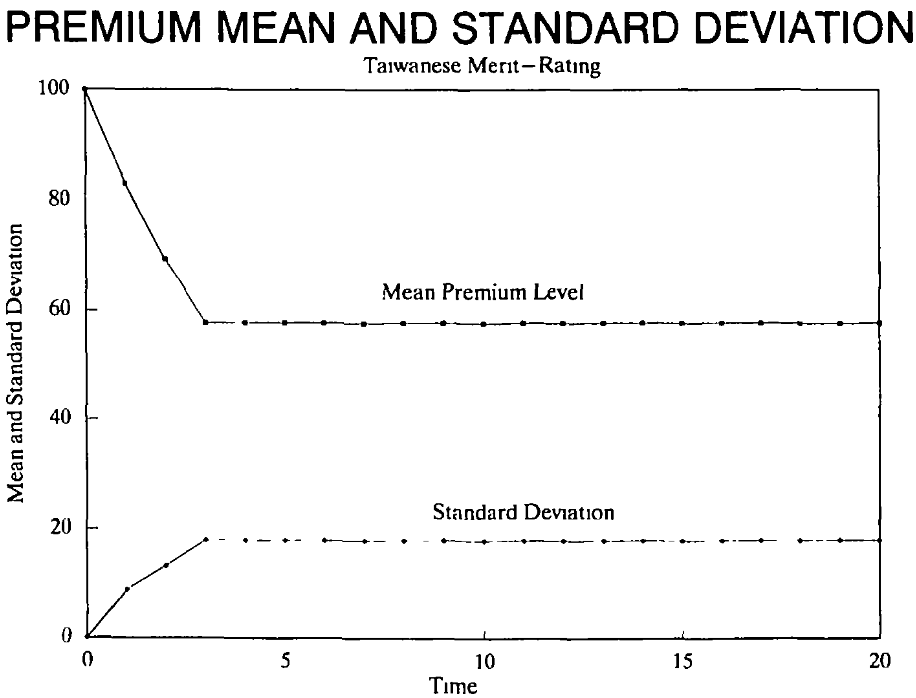


FIGURE 1

of 57.75 at time 30 [2]. The standard deviation of payments increases during the first 3 years, the time it takes for the best policyholders to reach class 1. Then it stabilizes around 17.89. As figures are expressed in premium levels in this section, and in dollars in Section 3, a dimension-less parameter has to be used for comparison purposes: the coefficient of variation (standard deviation divided by mean). For the benchmark Taiwanese driver, the coefficient of variation increases for 3 years, then stabilizes around 0.31 (see Fig. 2). Figure 3 shows the coefficient of variation as a function of λ, when the system is stationary.

Simulation was also used to compute the efficiency, the elasticity of the stationary premium with respect to the claim frequency. If $P(\lambda)$ denotes the stationary premium for a policyholder with a claim frequency λ, the efficiency curve $\varphi(\lambda)$ is defined as the relative increase of the premium, divided by the relative increase of the claim frequency (see LOIMARANTA, 1972, and LEMAIRE, 1985).

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

[2] The observed average premium level in Taiwan is higher than that, due to the constant flow of new policyholders entering the system in a high class However, since this note analyses two rating systems from a policyholder's point of view, new entries in the BMS are not considered

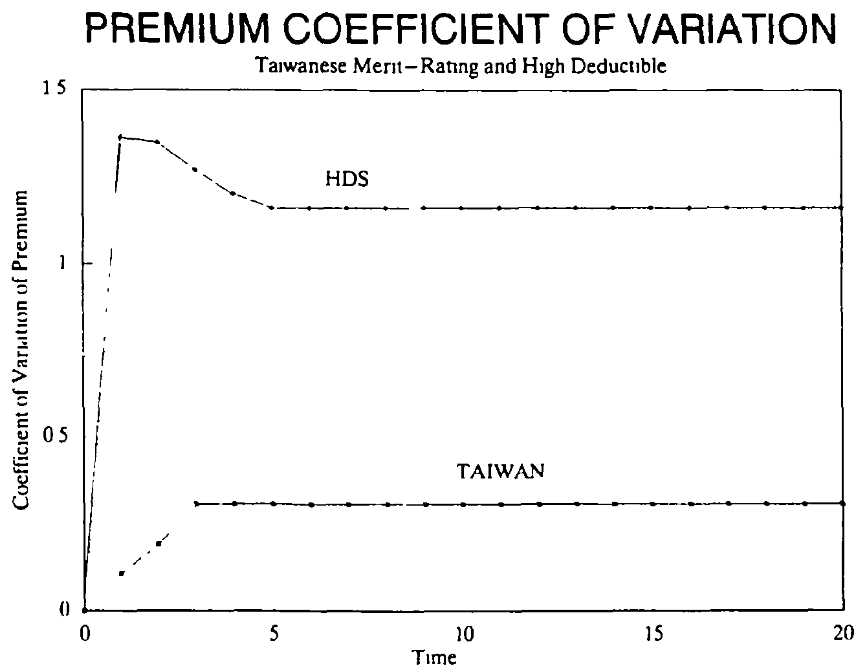


FIGURE 2

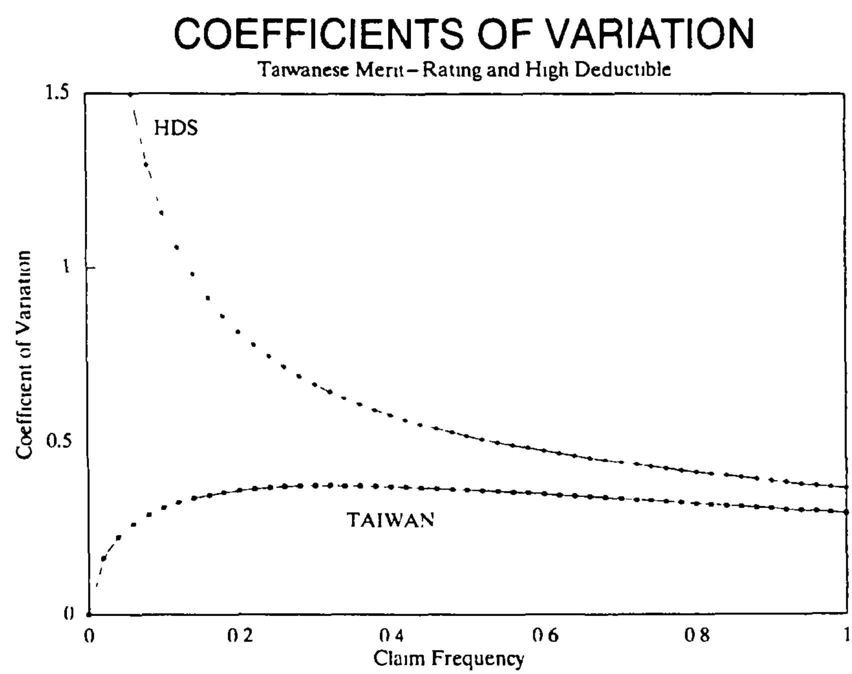


FIGURE 3

Ideally, the efficiency should be close to 1. In practice, the efficiency of most BMS in force around the world is much lower (LEMAIRE, 1988). For the Taiwanese BMS, the efficiency is very low for the most common values of λ ($\lambda \leq 0.10$); it peaks at 0.3 for claim frequencies in the [0.65 - 0.80] range (see Fig. 4). For $\lambda = 0.10$, $\varphi(0.10) = 0.1155$.

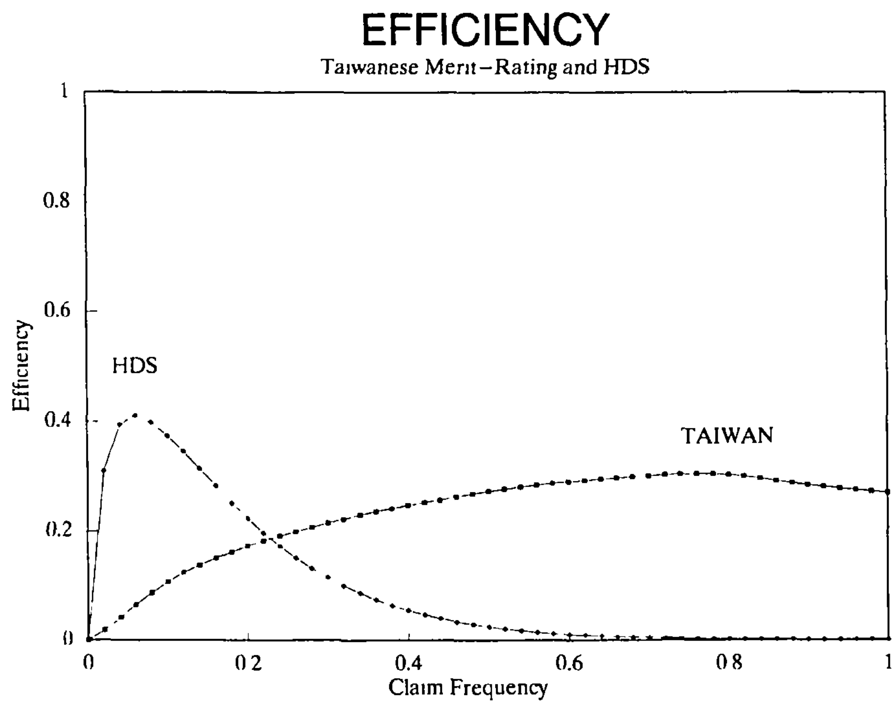


FIGURE 4

3. ANALYSIS OF THE HIGH-DEDUCTIBLE SYSTEM

Major assumptions for the HDS analysis are.

* Deductible: $D = \$3{,}000$
* Policyholders always borrow the entire loss amount L (up to \$3,000) from their insurer. Loans are reimbursed over a 5-year period, with decreasing amortization. A sum-of-the-digits principal repayment schedule is adopted: after a claim, 5/15 of the principal is repaid with the next annual premium, 4/15 the year after, ... All accidents occur in the middle of the year. The loan's interest rate is 3%, a low value since we assumed an inflation-free environment. This leads to the following payment schedule, for an accident that occurred at time $t - \frac{1}{2}$ and a loan $L = \min(D, \text{claim cost})$.

Time	Payment
t	3483 L
$t+1$	2867 L
$t+2$	2120 L
$t+3$	1393 L
$t+4$	0687 L
Total	1 0550 L

* The annual gross premium, without a deductible, is \$500. With 15% taxes, a 15% commission, and 10% operating expenses, the net premium is \$300.
* Claim amounts are exponentially distributed, with parameter $\mu = 1/3$ (using a \$1,000 currency unit).

As a consequence of these assumptions, the introduction of a \$3,000 deductible reduces the net premium to a *basic premium*

$$\lambda \int_D^\infty (x-D)\,\mu e^{-\mu x}\,dx = \frac{\lambda}{\mu}\,e^{-\mu D}$$

For the benchmark policyholder, the net premium is reduced from \$300 to \$110.36 = 0.1104.

Aggregate claims up to D form a compound Poisson process S, with a truncated exponential claim amount X. The first two moments of X are

$$E(X) = \int_0^D x\mu e^{-\mu x}\,dx + D\int_D^\infty \mu e^{-\mu x}\,dx$$

$$= \frac{1-\mu e^{-\mu x}}{\mu} = 1.8964$$

$$E(X^2) = \int_0^D x^2\mu e^{-\mu x}\,dx + D^2\int_D^\infty \mu e^{-\mu x}\,dx$$

$$= \frac{2}{\mu^2}(1-e^{-\mu D}) - \frac{2D}{\mu}\,e^{-\mu D} = 4.7563$$

For a compound Poisson process (see for example BOWERS et al., 1986, chapter 11),

$$E(S) = \lambda\,E(X) = (0.10)\,(1.8964) = 0\;1896$$
$$\mathrm{Var}\,(S) = \lambda\,E(X^2) = (0.10)\,(4.7563) = 0.4756$$

Disregarding all expenses, the expected payment for the first policy year consists only of the basic premium 0.1104. Expected payments (premium +

loan repayments) for the second year amount to

Basic premium + [(expected claim number) · (expected claim cost) · (0.3483 loan payment)]

$$= \frac{\lambda}{\mu} e^{-\mu D} + \lambda \frac{1-e^{-\mu D}}{\mu} (0.3483) = 0.1764$$

The variance of payments for the second year is

$$\mathrm{Var}\,(S) \cdot (0.3483)^2 = 0.0577 .$$

Expected payments for the third year are

Basic premium + [(expected claim number) · (expected claim cost) · (0 3483 of second-year loan + 0.2867 of first-year loan)] = 0 2308.

The variance is $\mathrm{Var}\,(S) \cdot (0.3483^2 + 0.2867^2) = 0.0988$.

The system reaches stationarity after five years. Expected payments for the sixth year are

Basic premium + [(expected claim number) · (expected claim cost) · (0.3483 of 5th-year loan + 0.2867 of 4th-year loan + 0.2120 of 3rd-year loan + 0.1393 of 2nd-year loan + 0.0687 of 1st-year loan)] = 0.31043.

Average stationary payments exceed the net premium of 0 3, since policyholders are constantly paying back loans. Expected payments, variances, and coefficients of variation are presented in Table 2. Figure 2 shows that, for a policyholder with $\lambda = 0.10$, the variability of payments is at all times much higher under the HDS than under the Taiwanese BMS. Figure 3 shows that, for all usual values of λ, the coefficient of variation is higher under the HDS.

TABLE 2

HDS EXPECTED PAYMENTS, VARIANCE, AND COEFFICIENT OF VARIATION

Time	Year	Expected Payments	Variance	Coef of variation
0	1	0 1104	0	0
1	2	0 1764	0 0577	1 3616
2	3	0 2308	0.0968	1 3481
3	4	0 2710	0 1182	1 2686
4	5	0 2974	0 1274	1 2002
5, 6, 7,	6 and after	0 3104	0 1296	1 1599

For the basic Compound Poisson process with exponential claims the coefficient of variation of losses is $\sqrt{2/\lambda} = 4.4721$, for $\lambda = 0.1$. The high-deductible system would reduce the coefficient of variation of policyholders' payments to 1.1599. Coefficients of variation in excess of 1 would probably be considered as too high by regulators and consumers. A reduction of payments variability can be achieved by

(i) spreading the loan reimbursements over more than five years, and/or
(ii) adopting a loan reimbursement schedule with level payments.

For instance, a five-year loan with equal payments of .2152 L would increase stationary expected payments to .3144, but reduce their variance to .1101. The coefficient of variation decreases to 1 0552, a 9.02% reduction. If the loan is spread out to 10 years, with equal payments of .1155 L, expected payments increase to 3331, their variance decreases to 0635, and the coefficient of variation drops to a more acceptable .7564.

Stationary payments for a policyholder with claim frequency λ amount to

$$P(\lambda) = 0\ 1104 + \frac{\lambda}{\mu}(1-e^{-\mu D})\ (1.055)$$

$$= 0.1104 + 0.3165(1-e^{-10\lambda})$$

if the basic premium[3] is set by the company at 0 1104 Consequently the efficiency is

$$\varphi(\lambda) = \frac{3.165\lambda e^{-10\lambda}}{0.1104+0\ 3165(1-e^{-10\lambda})}$$

Figure 4 shows that the efficiency of the HDS is higher than the efficiency of the Taiwanese BMS for the most common values of λ (under 0.22). For $\lambda = 0.10$, $\varphi(0.10) = 0.3751$. For the larger λ, the BMS is more efficient. Since most policyholders have a low λ, the computation of an average efficiency φ using any realistic structure function $u(\lambda)$

$$\varphi = \int_{\Lambda} \varphi(\lambda)\, u(\lambda)\, d\lambda$$

would provide a better efficiency for the HDS. $u(\lambda)$ is the density function of λ in the insurer's portfolio.

4. PRACTICAL CONSIDERATIONS

The implementation of a HDS instead of a BMS would lead to several practical problems:

1. Surcharges and discounts for other classification variables would need to be revised For instance, in many countries, inexperienced drivers have to pay

[3] In a definition of the efficiency from an insurer's point of view, the basic premium of 0 1104 would be replaced by $(\lambda/\mu)e^{-\mu D}$. From a policyholder's point of view, however, the basic premium is exogeneous, and not a function of his own λ

a hefty surcharge In addition, they also pay an implicit penalty, as they have to access the BMS at a level which is higher than the average stationary level. As this surcharge would disappear, explicit penalties for inexperience need to be reinforced.

2. The administration of a BMS is extremely inexpensive, and routinely handled by company computers. A HDS would lead to much higher expenses, since the insurer has to examine the credit worthiness of the policyholder before each annual period.
3. A bad (or unlucky) policyholder could face considerable debt and possibly personal bankruptcy. This is the kind of situation insurance is meant to avoid.
4. As a partial remedy for possible insolvencies, Holtan suggests to open an account for each policyholder. Each year, a specified amount would be set aside, to build up an individual risk reserve to cover future deductibles. Creating such accounts would eliminate the solvency problem for most experienced policyholders. However, it would do little to help young drivers, who not only form the group with the highest accident rate, but also the group with the worse credit rating. At most, policyholders could be induced to save the gross premium reduction created by the introduction of the deductible. In our benchmark situation, a \$ 3,000 deductible reduces the gross premium by \$ 190. So \$ 190 could be saved annually in the account. If the savings account accrue 3% (real) interest, it will take 13 years to save the amount of just one deductible.
5. With a HDS, many policyholders would in practice be prevented from switching to a new company after a claim, since the former insurer would demand a full reimbursement of the loan. This goes against current regulatory trends and creates an adverse selection process: claim-free policyholders would be free to leave a company, while policies with claims could not be eliminated from the portfolio and sent to the residual market
6. Taxes, commissions, and operating expenses have been disregarded in the preceding analysis. For simplicity, assume the operating expenses of the HDS are \$ 50, like in a BMS. It seems impossible to include these expenses in the loan reimbursement schedule. Commissions and taxes are not paid on deductibles. A policyholder, who has incurred a \$ 3,000 loss, will never accept to repay \$ 5,000, in order to provide \$ 750 to his broker, \$ 750 to his government, and \$ 500 to compensate the company for operating expenses. Since the broker, the government, and the insurer will not accept a decrease of their revenue, all of these expenses will need to be included in the basic premium, that covers losses above \$ 3,000. So the gross premium of a benchmark policyholder would be \$ 310 (\$ 110 net premium + \$ 200 expenses, tax and commission). 64.5% of the gross premium would be needed to cover expenses. While in practice such a high figure may be reached for some low-premium or high-deductible policies, it is certainly excessive for compulsory auto third party coverage

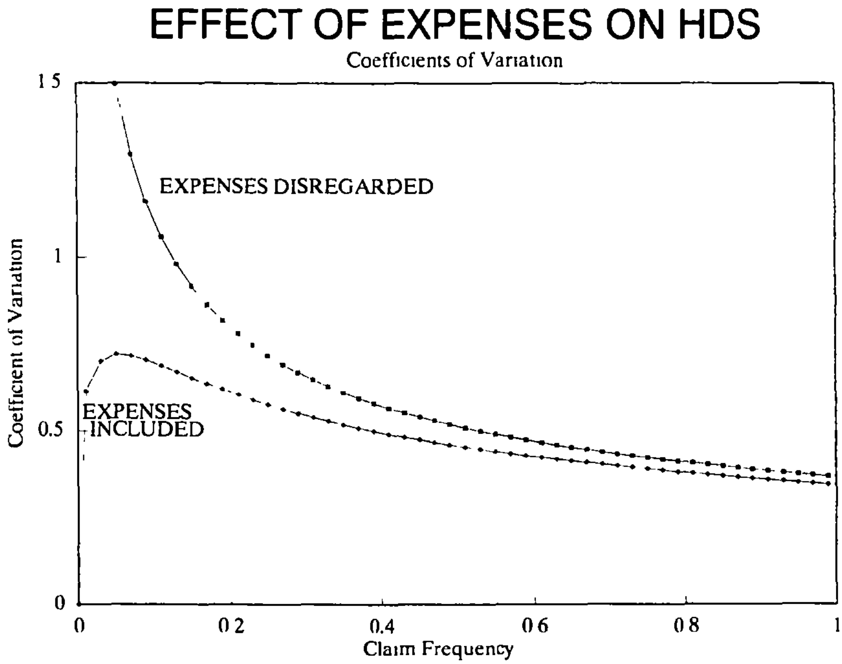


FIGURE 5

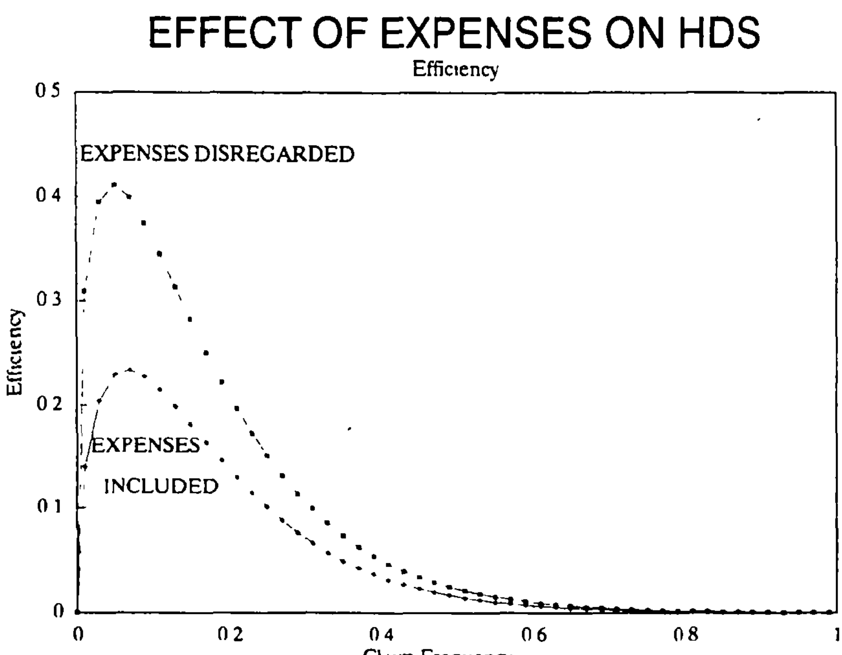


FIGURE 6

The inclusion of all expenses into the basic premium has another important consequence· a decrease of the efficiency and the payments coefficient of variation of the HDS. In a traditional BMS, expenses are proportional to the premium level, and bad drivers pay more commission, tax, and operating expenses. In a HDS, all policyholders contribute equally towards expenses. This reduces relative premium differentiation, and has a depressing effect on the efficiency curve and on the coefficient of variation of payments (see Fig. 5 and 6)

7. In the preceding analysis, the deductible has been set rather arbitrarily at \$3,000, following a suggestion by Holtan to set the deductible around the mean claim cost If the HDS is ever implemented, the value of the deductible will probably be decided by practical considerations, and not as the result of sophisticated modelling Holtan has presented a model, based on the minimisation of a quadratic expected utility function, that would provide an "optimal" deductible, after lengthy calculations. A simpler optimisation criterion coud be based on the efficiency. For instance, one could select the deductible in such a way as to maximise $\varphi(0.10)$. The first derivative (with respect to D) of $\varphi(0\ 10)$ is easily calculated, and a numerical procedure leads to an optimal deductible of \$2,941, very close to the value arbitrarily selected. Figure 7 compares the efficiency curve for various deductibles. It shows that $\varphi(0.10)$ is not an increasing function of D A very large D improves the efficiency for small λ's, but reduces $\varphi(0.10)$.

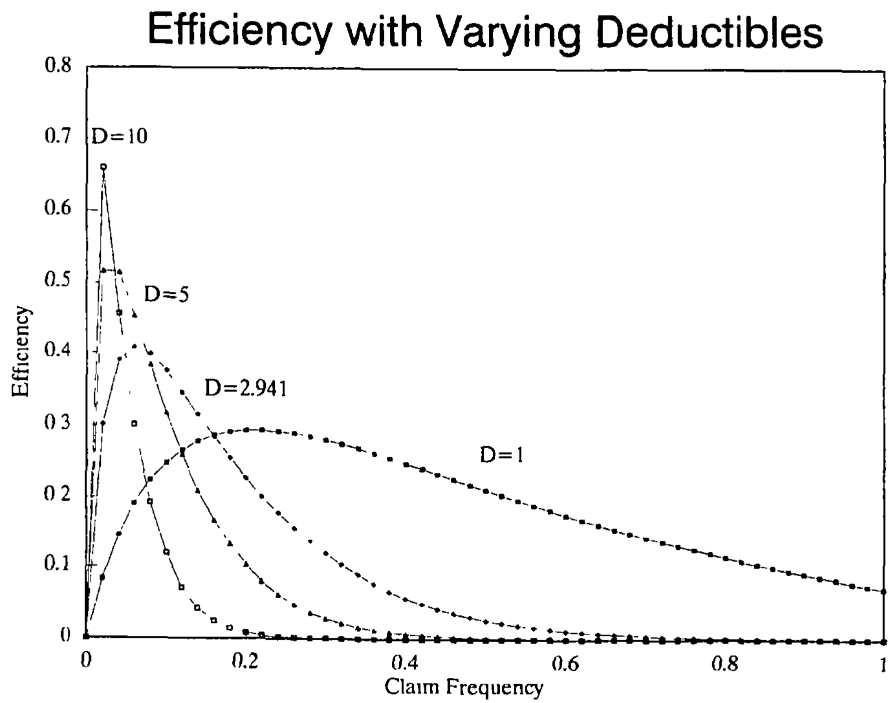


FIGURE 7

5. CONCLUSIONS

Compared to a traditional bonus-malus system, a high deductible system

1. reaches a steady state much faster;
2. increases premium income during early years;
3. has a higher efficiency for the most common values of the claim frequency; and
4. has a higher variability of payments for all policyholders.

Of course the first three points are in favour of the HDS, while point 4 is a very important drawback, that will probably prevent the application of a HDS in practice. Further research might be needed to improve Holtan's proposal For instance, one should investigate the impact of less severe forms of claim sharing than a straight deductible, such as proportional co-payments under D, or annual vs. per claim deductibles.

Finally, it should be pointed that a HDS would be a good application of the "bancassurance" concept, since both insurance (above the deductible) and banking (the loan under the deductible) expertise would be needed to manage the system. The banking segment of the industry would be induced to develop savings vehicles that would guarantee the repayment of the loans.

REFERENCES

BOWERS, N., GERBER, H., HICKMAN, J., JONES, D and NESBITT, C (1986) "Actuarial Mathematics" Society of Actuaries, Schaumburg, Illinois

HOLTAN, J (1994) "Bonus Made Easy" *ASTIN Bulletin* **24**

LEMAIRE, J (1985) "Automobile Insurance Actuarial Models" Kluwer Nyhoff Publishing Co, Boston, Massachussetts

LEMAIRE, J (1988) "A Comparative Analysis of Most European and Japanese Bonus-malus Systems" *Journal of Risk and Insurance* LV, 660-681

LEMAIRE, J and ZI, H (1994) "A Comparative Analysis of 29 Merit-Rating Systems" Submitted

LOMARANTA, K (1972) "Some Asymptotic Properties of Bonus Systems" *ASTIN Bulletin* **6**, 233–245

JEAN LEMAIRE AND HONGMIN ZI
Department of Insurance, and Risk Management,
Wharton School, 3641 Locust Walk,
University of Pennsylvania Philadelphia, PA 19104-6218, U.S A.

NOTE ON THE PAPERS BY J. HOLTAN AND BY J. LEMAIRE & H. ZI

According to the editorial rules of treating discussion situations in the ASTIN Bulletin the paper by J. LEMAIRE & H. ZI being somewhat a discussion on Holtan's paper was sent to the author of the original paper, who was given the opportunity to make an additional comment. The editors then received the following note by JON HOLTAN.

In this note I want to give some general comments on the papers by LEMAIRE & ZI (1994) and HOLTAN (1994)

Interpret henceforth a bonus-malus (BM) principle as consisting of two basic components:

(a) The BM design.
(b) The BM tariff parameters

Traditional actuarial literature has basically been preoccupied with component (b) Or more precisely, the tariff parameters of an *initial accepted* BM design have usually been mathematically optimalized within different criteria of succes like e g. high efficiency and financial balance. In my opinion, however, this strategy seems to be too narrow if the aim is to construct a BM principle which is *totally* optimalized in favour of both the insurer and the insured In our strive for maximizing BM advantages and minimizing BM disadvantages, actuarial BM research should instead simultaneously focus on both components (a) and (b). The construction of the High-Deductible System (HDS) in HOLTAN (1994) is an example of this strategy However, as pointed out in LEMAIRE & ZI (1994) (see Section 1 and 4) and HOLTAN (1994) (see Section 3, 5 and 6), a HDS compared with existing BM systems *both eliminates and generates* important disadvantages which are linked to component (a) Based on some mathematical model assumptions, LEMAIRE & ZI moreover concludes (see Section 3 and 5) that this two-sided conclusion is in principle also valid within some mathematical criteria of success linked to component (b) These complex, and perhaps confusing, conclusions make it difficult for us to decide whether to prefer the existing BM systems or the HDS However, the solution to this problem of decision seems to be naturally dependent on some strategic questions like: What kind of BM advantages and what kind of BM disadvantages will be the most important to focus on in the future automobile insurance market? In what way will new financial market structures and new electronic technology moderate the stated criticism of HDS, and hereby make room for creative insurance poducts like HDS? The answers to these questions are of course by now not obvious, and hence a continuous prospective assessment of the questions will probably be the most suitable way to proceed within the evaluating of HDS. In addition, and as mentioned in Section 5 in LEMAIRE & ZI (1994), the design of HDS may also be improved by further research For instance, a traditional BM system may be combined with a HDS such that all policyholders within the traditional system who attain a specific high rate of bonus discount are offered a separated (comprehensive insurance) HDS on a permanent basis. In the first place this modified HDS obviously moderates a great deal of the stated criticism of the pure HDS, while it in the second place gives the offered customers a customer-friendly choice between two different product alternatives.

In the immediate future the automobile insurance industry seems to meet market demands which are even more customer-orientated than today. Under the circumstances, and as intimated above, it seems to be a must for actuarial research within BM principles to be more orientated towards both the components (a) and (b). Or, in other words, more orientated towards an optimal combination of insurance market BM criteria and traditional actuarial BM methods.

REFERENCES

HOLTAN, J. (1994) Bonus Made Easy. *ASTIN Bulletin* **24**, (this volume)

LEMAIRE, J. and ZI, H. (1994) High Deductible instead of Bonus-Malus. Can it work? *ASTIN Bulletin* **24**, (this volume)

JON HOLTAN
Samvirke Insurance Co., P.O. Box 778 - Sentrum,
N-0106 Oslo, Norway.

ON THE EXACT CALCULATION OF THE AGGREGATE CLAIMS DISTRIBUTION IN THE INDIVIDUAL LIFE MODEL

BY KARL-HEINZ WALDMANN

Institut für Wirtschaftstheorie und Operations Research, Universität Karlsruhe

ABSTRACT

An iteration scheme is derived for calculating the aggregate claims distribution in the individual life model. The (exact) procedure is an efficient reformulation of De Pril's (1986) algorithm, considerably reducing both the number of arithmetic operations to be carried out and the number of data to be kept at each step of iteration. Scaling functions are used to stabilize the algorithm in case of a portfolio with a large number of policies Some numerical results are displayed to demonstrate the efficiency of the method.

KEYWORDS

Individual life model, aggregate claims distribution, De Pril algorithm.

1. INTRODUCTION

Consider a portfolio of m independent life insurance policies Suppose each policy to have an amount at risk $i \in I = \{1, \ldots, a\}$ and a mortality rate q_j with $j \in J = \{1, \ldots, b\}$. Let m_{ij} denote the number of all policies with amount at risk i and mortality rate q_j.

In the individual risk model the total amount of claims, S, is the sum $S = X_1 + \ldots + X_m$ of the m individual claims $X_1, \ldots, X_m$ produced by the policies. The distribution of S, $f(s) = P(S = s)$, referred to as the aggregate claims distribution, can be obtained by successively convoluting the m two-point distributions of the individual claims. Since the numerical calculation of an m-fold convolution is usually very time-consuming, numerous approximations can be found in the literature. See, e.g , BEARD, PENTIKAINEN and PESONEN (1984) for more details. The method derived in DE PRIL (1986) is a remarkable progress in computing the distribution of S exactly. Compared with Panjer's (1981) recursion formula, however, which can be thought of as the counterpart within the collective risk model, the computing time remains large (cf KUON, REICH and REIMERS (1987), DE PRIL (1988), REIMERS (1988))

In the present paper we shall reformulate the iteration scheme underlying the method of DE PRIL (1986). A (much) more efficient organization of the data will considerably reduce both the number of arithmetic operations to be carried out and the number of data to be kept at each step of interation. Further, we shall stabilize the algorithm by introducing a suitable scaling function. This scaling function will enable us to apply the algorithm to a portfolio with an essentially larger number of policies. Finally, some numerical results will be displayed to demonstrate the efficiency of the method

2. THE AGGREGATE CLAIMS DISTRIBUTION

For $j \in J$, we set $p_j = 1 - q_j$, $z_j = q_j/p_j$, $m_j = \Sigma_{i \in I} m_{ij}$, and $c = \Sigma_{i \in I} \Sigma_{j \in J}\, i\, m_{ij}$. Further, we use $[x]$ to denote the greatest integer less than or equal to x.

It has been shown in De Pril (1986) that the aggregate claims distribution can be computed recursively via

$$f(0) = \prod_{j=1}^{b} (p_j)^{m_j} \tag{1}$$

and for $s = 1, \ldots, c$

$$f(s) = \frac{1}{s} \sum_{i=1}^{\min(a,s)} \sum_{k=1}^{[s/i]} g(i,k) f(s-ki) \tag{2}$$

where

$$g(i,k) = (-1)^{k+1}\, i \sum_{j=1}^{b} m_{ij} z_j^k \tag{3}$$

Theorem 1: Equation (2) can be written as

$$f(s) = \frac{1}{s} \sum_{i=1}^{\min(a,s)} \sum_{j=1}^{b} i\, m_{ij}\, r(s,i,j) \tag{4}$$

where, for all $i \in I$, $j \in J$, $i \leq s$

$$r(s,i,j) = z_j \{ f(s-i) - r(s-i,i,j) \} \tag{5}$$

and $r(s,i,j) = 0$ otherwise.

Proof: Let

$$r(s,i,j) = \sum_{k=1}^{[s/i]} (-1)^{k+1} z_j^k f(s-ki)$$

Then, utilizing

$$\begin{aligned} r(s,i,j) &= z_j \left\{ f(s-i) - \sum_{k=2}^{[s/i]} (-1)^{(k-1)+1} z_j^{k-1} f(s-i-(k-1)i) \right\} \\ &= z_j \left\{ f(s-i) - \sum_{k=1}^{[(s-i)/i]} (-1)^{k+1} z_j^k f(s-i-ki) \right\} \\ &= z_j \{ f(s-i) - r(s-i,i,j) \} \end{aligned}$$

the assertion immediately follows from (2) □

Equations (4) and (5) can be thought of as an efficient reformulation of equation (2). The superiority results from

(a) a lower number of arithmetic operations to be carried out at each step of iteration
(b) arrays of smaller size to keep the data needed for further iterations

To specify (a), we first study equation (2). Fix (s, i, k). Then, having already computed $g(i, k-1)$, $g(i, k)$ can be obtained as the result of

$$\{i(-1)^k\} \sum_{j=1}^{b} (-z_j)\{m_{ij} z_j^{k-1}\}$$

which can be managed by $b+1$ multiplications and b additions. Two additional multiplications and one subtraction are necessary to compute $g(i, k) f(s-ki)$. Summing over k there is a need of $(b+3)[s/i]$ multiplications and $(b+1)[s/i]$ additions/subtractions.

On the other hand, by applying equations (4) and (5), for fixed (s, i, j), one multiplication and two subtractions are necessary to compute $r(s, i, j)$. Further, one additional multiplication is needed to obtain $\{i\, m_{ij}\}\, r(s, i, j)$ Summing over j, there is a need of $2b$ multiplications and $2b$ additions/subtractions.

Now let $\xi_m(s)$ (resp. $\xi_a(s)$) denote the number of multiplications (resp. additions/subtractions) to be saved by applying equations (4) and (5) in place of equation (2) at stage s of iteration. Then it easily follows that

$$\xi_m(s) = \sum_{i=1}^{\min(a, s)} \{(b+3)[s/i] - 2b\} \approx \{(b+3)\log(a+1)\}s - 2ab$$

$$\xi_a(s) = \sum_{i=1}^{\min(a, s)} \{(b+1)[s/i] - 2b\} \approx \{(b+1)\log(a+1)\}s - 2ab$$

where use has been made of $\log(a+1) < \Sigma_{i=1}^{a} 1/i < 1 + \log(a)$ (cf. e.g., Ross (1983)).

Now let us specify (b). To apply iteration scheme (2), an array with ac (resp. $c+1$) elements is needed to keep $g(i, k)$ (resp $f(s-ki)$) for further iterations. On the other hand, utilizing equations (4) and (5), an efficient implementation of $r(s, i, j)$ (resp. $f(s-i)$) needs an array with $a(a+1)b/2$ (resp. $a+1$) elements only

To illustrate the basic idea underlying the implementation of $r(s, i, j)$, observe (see Figure 1) that the $r(s, i, j)$ within the upper triangle (solid line) have to be kept at stage s, while at stage $s+1$ the $r(s, i, j)$ of the lower triangle (dashed line) have to be retained.

To manage these data in an efficient way, we rearrange the elements of the upper triangle in an array with $a(a+1)/2$ rows and b columns, and, switching to the lower triangle, we replace the entries of $(s-i, i, j)$ (not needed any longer) by the ones of (s, i, j) (to be kept for further use) and let the other entries unchanged.

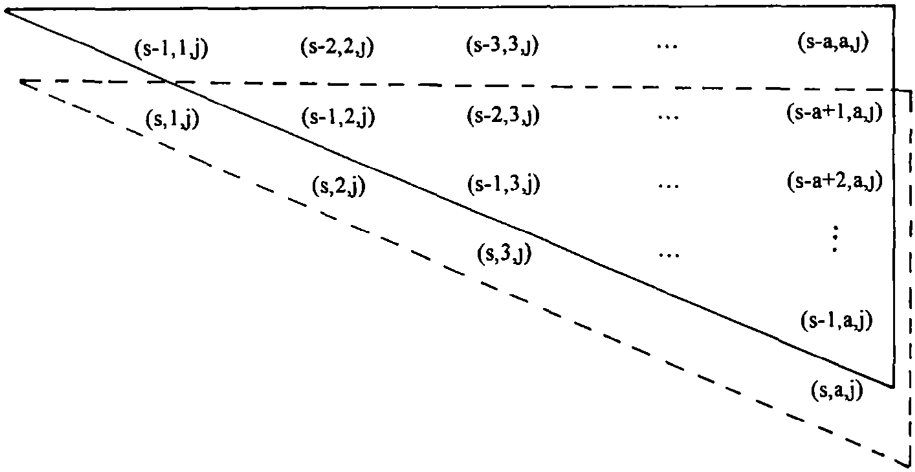


FIGURE 1 Actualization of the data

Formally, we introduce

$$v_i = i(i-1)/2+1$$
$$w_i = 0, \qquad i \in I$$

and actualize w_i at each step $s(s \geq 1)$ of iteration according to

$$w_i = \begin{cases} w_i+1, & \text{if} \quad w_i < i-1 \\ 0 & \text{otherwise} \end{cases}$$

Then w_i coincides with s modulo i and (v_i+w_i, j) is the position in the array, in which the entry of (s, i, j) can be found.

3. STABILIZATION OF THE ALGORITHM WITH RESPECT TO UNDERFLOW/OVERFLOW

Applying the algorithm to a portfolio with a large number of contracts, the initial value $f(0)$ is close to zero. This fact may cause an underflow followed by an abort or irregular running of the procedure.

To discuss this aspect in more detail, let ω and Ω denote the smallest and greatest numbers that can be represented on the computer to carry out the algorithm. Suppose $f(0) < \omega$. Then the algorithm stops with an underflow. On the other hand, by formally setting $f(0)$ equal to zero, the sequence $f(s)$ of iterates degenerates to a sequence that has all its elements equal to zero, which is not consistent with the property of being a probability mass function.

There are a variety of ways to overcome this difficulty. Three methods of different efficiency and/or applicability are to be stated as methods 1 to 3. There $f^*(s)$, $0 \leq s \leq c$, is used to denote the sequence of transformed iterates

Method 1: Suppose

$$f^*(s) = \gamma f(s), \qquad 0 \le s \le c$$

for some constant γ with $\omega < \gamma f(0) < \Omega$. Then the transformed iterates $f^*(s)$ can be obtained by formally starting (4) (resp. (2)) with $\gamma f(0)$ in place of $f(0)$. □

The use of a constant scaling function is the simplest way to stabilize the algorithm. A more refined method is to combine a constant scaling function with an exponential scaling function, which has been suggested by PANJER and WILLMOT (1986) within the collective risk theory.

Method 2: Suppose

$$f^*(s) = \gamma e^{-\alpha(s+\beta)} f(s), \qquad 0 \le s \le c$$

where α, β, γ are constants with $0 \le \alpha \le 0.5$, $\gamma > 0$, and

$$\beta = \sum_{i=1}^{a} \sum_{j=1}^{b} m_{ij} \log(p_j) \tag{6}$$

To compute $f^*(s)$, iteration scheme (4) has to be reformulated as

$$f^*(0) = \gamma e^{(1-\alpha)\beta}$$

$$f^*(s) = \frac{1}{s} \sum_{i=1}^{\min(a,s)} \sum_{j=1}^{b} i\, m_{ij}\, r^*(s, i, j), \qquad 1 \le s \le c$$

where, for all $i \in I$, $j \in J$

$$t(i,j) = z_j e^{-\alpha i}$$

$$r^*(s, i, j) = t(i,j)\{f^*(s-i) - r^*(s-i, i, j)\}, \qquad i \le s$$

and $r^*(s, i, j) = 0$ otherwise. □

Method 2 starts with a larger initial value as well as method 1 and additionally reduces the increase of the iterates. For large s, however, things may change and the transformation may lead to an ealier abort on account of an underflow. Our third method is one way to overcome this principal difficulty. It again starts with a larger initial value, reduces the increase of the iterates for $s \le E(S)$, and, additionally, reduces the decrease of the iterates for $s > E(S)$.

Method 3: Suppose

$$f^*(s) = \gamma e^{\alpha(s-\mu)^2} f(s), \qquad 0 \le s \le c$$

where

$$\alpha = -\beta/\mu^2$$

$$\mu = E(S) = \sum_{j=1}^{b} m_j q_j$$

and β as in (6) To compute $f^*(s)$, the modified iteration scheme reads

$$f^*(0) = \gamma$$

$$f^*(s) = \frac{1}{s} \sum_{i=1}^{\min(a,s)} \sum_{j=1}^{b} i\, m_{ij}\, r^*(s,i,j), \qquad 1 \leq s \leq c$$

where, for all $i \in I$, $j \in J$

$$t(s,i,j) = \begin{cases} z_j e^{\alpha i(2(s-\mu)-i)}, & i \leq s \leq 2a-1 \\ t(s-i,i,j)\, e^{2\alpha i^2}, & 2a \leq s \leq c \end{cases}$$

$$r^*(s,i,j) = t(s,i,j)\,\{f^*(s-i) - r^*(s-i,i,j)\}, \qquad i \leq s \leq c$$

and $r^*(s,i,j) = 0$ otherwise. □

It is not surprising that the last scaling function is superior to the other ones, since it is stimulated by the central limit theorem and thus best utilizes the asymptotic behavior of S as $m \to \infty$. Some numerical results to be given in the next section will illustrate the efficiency. We finally remark that $t(s,i,j)$ and $r^*(s,i,j)$ can be implemented in the same way as $r(s,i,j)$.

4. NUMERICAL RESULTS AND DISCUSSION

We consider as a starting point the portfolio discussed in GERBER (1979), p. 53.

q_j	m_{ij}				
0.03	2	3	1	2	—
0.04	—	1	2	2	1
0.05	—	2	4	2	2
0.06	—	2	2	2	1

Since the portfolio consists of 31 policies only, there is no need for a reformulation or stabilization of the algorithm We therefore expand the portfolio by considering km_{ij} policies in place of m_{ij} (for all $i \in I$ and $j \in J$).

Let $k = 5000$ (corresponding to 155 000 policies) to illustrate the numerical progress resulting from the application of equations (4) and (5) in place of equation (2). Then, being interested in computing the aggregate claims distribution up to the smallest c^* with $P(S > c^*) \leq 10^{-4}$, there is a saving of more than $4.4 * 10^9$ multiplications and a saving of more than $3.1 * 10^9$ additons/subtractions. Moreover, the arrays to be kept at each step of iteration can be reduced by 140 851 elements

The maximal k implying a stable algorithm has been determined on the basis of extended numbers (i e $\omega = 1.9 * 10^{-4951}$, $\Omega = 1.1 * 10^{4932}$). There stable means that the algorithm does not stop with an underflow or overflow and that both $|E'(S) - E''(S)|/E''(S) \leq 10^{-5}$ and $|\mathrm{Var}'(S)^{1/2} - \mathrm{Var}''(S)^{1/2}|/\mathrm{Var}''(S)^{1/2} \leq 10^{-5}$ hold, where $E'(S)$, $\mathrm{Var}'(S)$ are determined with help of the probability mass function of S and $E''(S)$, $\mathrm{Var}''(S)$ result from the moments of the individual claims and the properties of expectation and variance. The maximal k and the associated number of policies to be obtained in this way for $\gamma = 10^{4500}$ are displayed in Table 1.

TABLE 1

STABILITY OF THE ALGORITHMS UNDER CONSIDERATION ($\gamma = 10^{4500}$)

Method	maximal k	number of policies
Equations (4) and (5)	7 900	244 900
Method 1	15 100	468 100
Method 2 ($\alpha = 0\ 31$)	22 100	685 100
Method 3	80 100	2 483 100

Stability of our numerical results thus means stability with respect to the first two moments. For a more theoretial treatment of the numerical stability of recursive formulae the reader is referred to PANJER and WANG (1993).

ACKNOWLEDGEMENT

I would like to thank the referees for their detailed and helpful comments.

REFERENCES

BEARD, R E, PENTIKAINEN, T and PESONEN, E (1984) *Risk Theory* 3rd edition Chapman and Hall, London

DE PRIL, N (1986) On the exact computation of the aggregate claims distribution in the individual life model *ASTIN Bulletin* **16**, 109–112

DE PRIL, N (1988) Improved Approximations for the Aggregate Claims Distribution of a Life Insurance Portfolio *Scan Actuarial J* **1988**, 61–68

GERBER, H U (1979) *An Introduction to Mathematical Risk Theory* Huebner Foundation Monograph 8, Philadelphia

KUON, S, REICH, A and REIMERS, L (1987) Panjer vs Kornya vs De Pril A comparison from a practical point of view *ASTIN Bulletin* **17**, 183–191

PANJER, H H (1981) Recursive evaluation of a family of compound distributions *ASTIN Bulletin* **12**, 22–26

PANJER H H and WILLMOT, G E (1986) Computational aspects of recursive evaluation of compound distributions *Insurance Mathematics and Economics* **5**, 113-116
PANJER, H H and WANG, S (1993) On the Stability of Recursive Algorithms *ASTIN Bulletin*, to appear
REIMERS, L (1988) Letter to the Editor *ASTIN Bulletin* **18**, 113 114
ROSS, S M (1983) *Stochastic Processes* John Wiley, New York

Prof. Dr. KARL-HEINZ WALDMANN
Institut fur Wirtschaftstheorie und Operations Research,
Universität Karlsruhe, Postf. 6980, D-76128 Karlsruhe.

MODELLING MORTGAGE INSURANCE CLAIMS EXPERIENCE: A CASE STUDY

BY GREG TAYLOR

December 1991
Revised June 1993
A Coopers & Lybrand research report

ABSTRACT

Mortgage insurance indemnifies a mortage lender against loss on default by the borrower. The sequence of events leading to a claim under this type of insurance is relatively complex, depending not only on the credit worthiness of the borrower but also on a number of external economic factors.

Prominent among these external factors are the loan to valuation ratio of the insured loan, the disposable income of the borrower, and movements in property values A broad theoretical model of the functional dependencies of claim frequency and average claim size on these variables is established in Sections 6 and 7. Section 8 fits these models, extended by other "internal" variables such as the geographic location of the mortgaged property, to a real data set.

Section 9 compares the fitted model with the data, and finds an acceptable fit despite extreme fluctuations in the claims experience recorded in the data set.

KEYWORDS

Mortgage insurance, housing price index; loan to valuation ratio; regression.

1 INTRODUCTION

Mortgage insurance indemnifies a mortgage lender against loss on default by the borrower The typical sequence of events leading to the invocation of the indemnity is as follows.

The amount of the mortgage is repayable by a sequence of instalments, perhaps monthly, over a period of some years, up to perhaps 25 or in a few cases more. If a borrower fails to meet one or more of these instalments, arrears collection procedures will be instigated. If it appears that the borrower is experiencing financial difficulties which threaten his capacity to pay the scheduled instalments, the lender's initial response will usually be to attempt rehabilitation of the borrower, possibly by some form of rescheduling of the debt repayment.

In many cases this will render the borrower's difficulties temporary. In other less fortunate cases it will become clear that the borrower is quite unable to repay the debt. The lender will then force sale of the mortgaged property, and retain that part of the sale proceeds required to discharge the remaining debt. In the majority of sales, the proceeds will be sufficient for this purpose, but if they are not the mortgage insurance indemnity is invoked to reimburse the lender for the shortfall.

It is an elementary observation that inflation of property values reduces the call on mortgage insurance; the proceeds of property sales cover a greater proportion of the corresponding debts. It is also clear from the above description that a loan needs to go through several stages (healthy → in arrear → property under management → sale of property) before a mortgage insurance claim arises, and each of these stages involves some delay. As will be discussed in Section 3, each of them also depends on its own specific economic factors.

For these reasons, the underlying process generating mortgage insurance claims is complex and dependent on several variables which are exogenous to the insurance portfolio. Consequently, mortgage insurance run-off arrays, whether in terms of numbers or amounts of claims, exhibit very different characteristics from those of other lines of business. A striking example of this is given in Section 2.

These different characteristics necessitate rather different modelling techniques. The purpose of the present paper is to illustrate these techniques by means of a case study. Since this study is specific to a particular portfolio, it cannot be claimed that the modelling techniques illustrated are generally applicable. It is hoped, however, that they are fairly generally indicative of the **type** of modelling which needs to be attempted.

2. NUMERICAL EXAMPLE: PRELIMINARY DISCUSSION

The following data are given as an indication of the difficulties likely to arise if a mortgage insurance portfolio is subjected to conventional run-off analysis. More detail of the data on which this paper is based appears in Appendices E and G.

Year of loan advance	Number of claims, per 10,000 loan advances, emerging in development year (a)										
	0	1	2	3	4	5	6	7	8	9	10
1980					30	18	6	0	0	0	6
1981				116	42	31	5	0	0	0	
1982			54	27	45	36	13	13	4		
1983		25	20	20	23	9	0	3			
1984	0	13	24	55	35	5	0				
1985	1	21	134	68	15	6					
1986	0	17	30	4	2						
1987	3	1	0	2							
1988	0	0	5								
1989	0	0									
1990	0										

(a) **Development year** is defined as year of emergence of claim minus year of loan advance

Let the term relative claims frequency denote the number of claims per 10,000 loan advances. If C_{ij} denotes the relative claim frequency in development year j of year of advance i, and A_{ij} denotes the age-to-age factor:

$$A_{ij} = \sum_{k=0}^{j+1} C_{ik} \Big/ \sum_{k=0}^{j} C_{ik}, \tag{2.1}$$

then the following table of age-to-age factors is obtained.

Year of loan advance i	Age-to-Age factor in development year j = 1	2	3	4	5
1984	2.86	2.50	1.38	1.04	1.00
1985	7.12	1.44	1.07	1.03	
1986	2.71	1.08	1.05		
1987	1.00	1.50			

The great instability in these ago-to-age factors is evident in the sense of variability within a development year. The basic reason for the instability is clear from the first table. It is the apparent correlation between relative claim frequency and year of emergence of claim, i.e. with the number of the diagonal in the table. Such a data structure suggests application of the separation method (TAYLOR, 1977, 1986), with the model structure:

$$E[C_{ij}] = r_j \lambda_{i+j}. \tag{2.2}$$

The separation method yields the following parameter estimates.

j	$\hat{r}_j$	k	$\hat{\lambda}_k$
0	0.00		
1	0.06		
2	0.20		
3	0.22		
4	0.14	1984	366
5	0.11	1985	167
6	0.03	1986	195
7	0.03	1987	350
8	0.02	1988	196
9	0.00	1989	48
10	0.20	1990	29

This produces the following comparison between observed and fitted relative claim frequencies.

Year of loan advance	Observed and fitted (shown in bold type) relative claim frequency in development year																							
	0		1		2		3		4		5		6		7		8		9		10		Total	
1980									30	**52**	18	**18**	6	**6**	0	**9**	0	**3**	0	**0**	6	**6**	60	**94**
1981							116	**79**	42	**24**	31	**21**	5	**11**	0	**5**	0	**1**	0	**0**			195	**140**
1982					54	**72**	27	**36**	45	**28**	36	**38**	13	**6**	13	**1**	4	**0**					193	**181**
1983			25	**21**	20	**33**	20	**42**	23	**50**	9	**21**	0	**1**	3	**1**							101	**169**
1984	0	**1**	13	**9**	24	**38**	55	**76**	35	**28**	5	**5**	0	**1**									131	**159**
1985	1	**1**	21	**11**	134	**69**	68	**42**	15	**7**	6	**3**											245	**133**
1986	0	**1**	17	**20**	30	**38**	4	**10**	2	**4**													53	**73**
1987	3	**1**	1	**11**	0	**9**	2	**6**															6	**28**
1988	0	**1**	0	**3**	5	**6**																	5	**9**
1989	0	**0**	0	**2**																			0	**2**
1990	0	**0**																					0	**0**

The table indicates that the separation method achieves a reasonable fit No formal goodness-of-fit statistics are examined, because this model is later discarded. The difficulty is that, despite the reasonableness of the fit, the sequence of **escalation index numbers** λ_k is peculiar by normal standards Until some explanation of this peculiarity is found, it is impossible to produce any reliable projection of the sequence into future years.

One of the major objectives of subsequent sections of this paper will therefore be to obtain such an explanation. The discussion of this aspect of the modelling problem is taken up in Section 3.

3 THE PROCESS OF CLAIM OCCURRENCE

3.1. Major financial factors

As pointed out in Section 1, a loan must traverse several stages of financial deterioration before producing a mortgage insurance claim. These stages are subject to different financial influences Of these separate influences, two are of particular prominence·

(a) the onset of financial difficulties for the borrower, and
(b) in the event of forced sale, the extent to which the sale proceeds repay the outstanding loan.

These two factors are discussed in the following two sub-sections.

3.2. Onset of borrower's financial difficulties

Despite its importance in a borrower's budget, the mortgage payment instalment will nevertheless be to some extent a residual item in that budget. It will rank after tax and consumer expenditure on necessities (food, clothing, etc). In addition, most past loans have been of a type whereby the amount of instalment varies with variations in current day interest rates.

It appears, therefore, that a reasonable measure of the degree of financial pressure on mortgage borrowers would be provided by an estimate of the average residual income after allowance for tax, consumer expenditure and mortgage instalment. This residual income, called here the **home affordability index** (HAI), was constructed in the following form:

Home affordability index = average weekly gross household income
minus
tax
minus
consumer expenditure
minus
mortgage instalment,

expressed as a percentage of gross income

A baseline distribution of gross household income over these categories of expenditure was derived from a 1988/89 household expenditure survey (HES) conducted by the Australian Bureau of Statistics. The items of expenditure for this base year were adjusted to other years in various ways, indicated by the following table.

Item of income or expenditure	Adjustment from year to year according to
Gross household income	Average weekly earnings
Tax	Average weekly earnings (a)
Consumer expenditure	Consumer price index
Mortgage instalments	Average weekly earnings (b) Mortgage interest rates (b)

(a) Preliminary investigation indicated little variation in the effective average tax rate over the period concerned

(b) The average amount of a new loan was assumed to change in proportion with average weekly earnings These loans were assumed repayable over periods of 20 years, and the average mortgage instalment calculated on the basis of the most common interest rate charged in the year concerned in respect of the loan portfolio under analysis

The component time series used in the construction of the HAI (at year end) are set out as Appendix F.

The resulting HAI (at mid-year) is as set out in the following table.

The rather irregular progression of this index is seen in Appendix F to derive from quite reasonable component indexes Each of these components may be projected over future years, producing a rationally based projection of HAI. This situation may be contrasted with that which arises on application of "black box" estimates of past claims escalation, as in Section 2, and in which no guidance as to future escalation is available

Year	Home affordability index
1979	100 0
1980	104 8
1981	111 9
1982	101 7
1983	104 1
1984	128 9
1985	128 3
1986	101 7
1987	87 4
1988	90 6
1989	81 5
1990	81 2

3.3. Recovery of outstanding loan on forced sale

The HAI of Section 3.2 provides an indication of the likelihood that an individual borrower will experience financial difficulty in a particular year. However, such difficulty, while a necessary condition, **is not** sufficient for the emergence of a mortgage insurance claim. It is quite possible the borrower's difficulties are such as to force sale of the property, but that property values will be sufficient for the entirety of the outstanding loan amount to be recovered by the lender.

Whether or not this is the case will depend mainly on movements in property values between the date of advance of the loan and the date of the forced sale. In Sydney these movements may be estimated by reference to the **Housing Price Index** (HPI) computed and published by Residex Pty Limited. The following table was derived from that index with slight modification.

Year ended 30 June	Housing price index (Sydney) at mid-year (30/6/79 = 100)
1980	115 3
1981	145 1
1982	158 6
1983	158 4
1984	168 2
1985	177 2
1986	182 4
1987	191 5
1988	245 8
1989	363 5
1990	430 7

Evidently, the greater the increase in value of properties generally, the less the chance that forced sale of a particular property will lead to a loss to the mortgage lender.

3.4. Lags in claims process

While movements in the HAI (Section 3.2) and HPI (Section 3.3) have been identified as major variables in the frequency of mortgage insurance claims, it is to be expected that there will be a lag between cause and effect in each case.

Information from the company operating the mortgage insurance portfolio discussed in this paper was that, broadly:

(a) the average period between mortgage instalments falling in arrears and the property being taken under management (if indeed this latter occurred) was about 6 months; and
(b) the average period between taking a property under management and effecting its sale was also about 6 months.

On the basis of this information, it might be reasonable to expect lags of.

(a) 12 months between movements in the HAI and the consequent movement in claim frequency, and
(b) 6 months between a movement in the HPI and its consequent movement in claim frequency.

Thus, it has been assumed in subsequent modelling that a claim frequency experienced during year t is dependent upon:

(a) the value of the home affordability index at the **middle** of year $t-1$; and
(b) the value of the HPI at the **end** of year $t-1$.

Examination of alternatives suggested that this choice of lags provided about the best fit of model to data. Further detail on the incorporation of the HAI and HPI in the model is given in Section 6.2.

4. DATA

4.1. Variables affecting claims experience

Section 3 identified the HAI and HPI as likely to be major explanatory variables of **claim frequency** Other variables in this category include:

(a) the proportion of the original property value advanced by way of mortgage, i.e. the **loan to valuation ratio** (LVR);
(b) the geographic area of the mortgaged property (described in more detail in Section 4.2);
(c) the agreed term of the mortgage loan;

(d) the type of property mortgaged (e.g. new house, old unit, land only, etc.);
(e) the financial type of the loan (e.g. reducible loan with variable interest, interest only instalments with fixed interest rate, etc.)

In addition, it is likely that claims experience will vary with **development year**, even in the absence of movements in the HAI and HPI. This would reflect a process of natural selection operating on each year's mortgage advances, whereby the poorest risks succumb to financial pressures relatively early, and the remainder survive the mortgage term.

It is clear that the major variable affecting **claim size** will be the **size of the original loan**. In addition, the explanatory variables (a) to (e) of claim frequency potentially affect claim size also.

4.2. Form of data

As the tables of Section 2 indicate, claims experience relates to the period 1984 to 1990. In fact, the 1984 experience covers only 7 months of that year.

Data supplied in respect of these claims consisted of a claim by claim tabulation, recording in each case the relevant variables identified in Section 4.1:

(a) year of advance;
(b) amount of loan;
(c) value of property;
(d) geographic area of property;
(e) term of loan;
(f) type of property;
(g) financial type of loan;
(h) year of emergence of claim.

The tabulated geographic area was the postal code of the property. These codes were grouped into 14 broad urban and rural regions within the states of New South Wales and Australian Capital Territory, as follows:

> Metropolitan regions 1 to 5; Canberra (6); Newcastle (7); Wollongong (8), Central Coast (9); North Coast (10), South Coast (11); Blue Mountains (12), Southern Highlands (13), Other (14).

The exposure base for the study consisted of all loans advanced over the years 1980 to 1990 inclusive. These were recorded, loan by loan, according to the variables (a) to (g) listed above as potentially affecting claim frequency.

As the data described above constitute a unit record file, it is not practical to present the full detail here. It is not even practical to tabulate cells of data since there are 1499 exposure cells. However, Appendix G gives a tabulation of exposures and claims according to year of advance and development year. It is to be stressed that, while such a tabulation is interesting, it omits a great deal of the raw data.

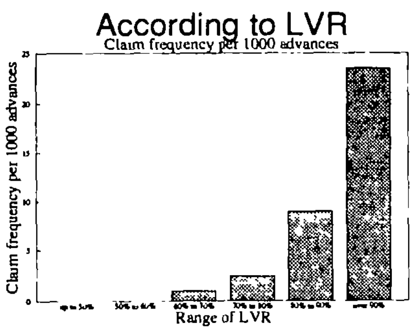
5. EXPLORATORY DATA ANALYSIS

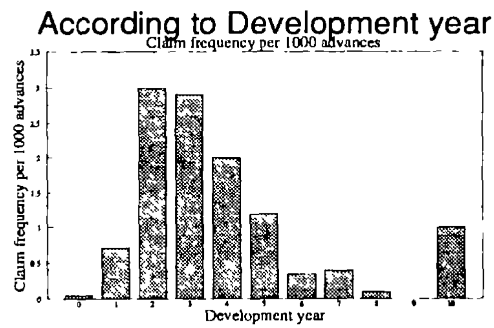
5.1. Claim frequency

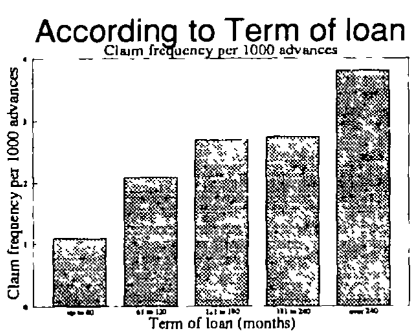
Section 4.1 identified a number of characteristics of individual loans (such as LVR, term of loan, etc.) which might have a bearing on the likelihood of those loans leading to claims These characteristics will be referred to here as **risk variables**.

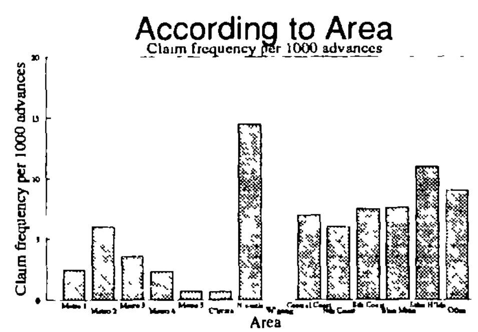
Initially, data concerning claim numbers were analysed according to the risk variables listed in Section 4.1. This provided initial guidance concerning the types of loans which were subject to high or low risk of default.

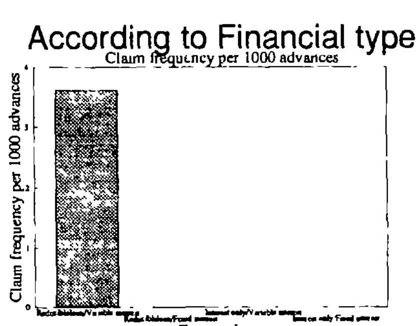
The results of this analysis are summarized in the following sequence of bar charts.

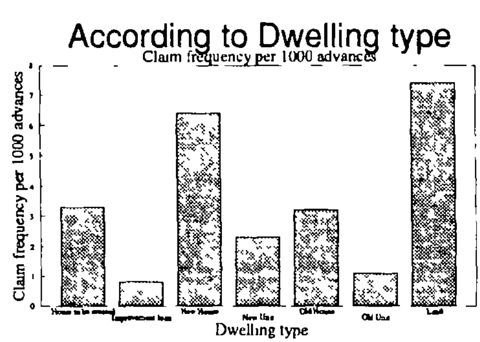












These charts raise the following possibilities:

(a) claim frequency peaks in the second, third and fourth years after the year of advance;
(b) claim frequency increases dramatically with increasing loan to valuation ratio (LVR);
(c) claim frequency increases significantly with increasing term of loan;
(d) certain geographic areas experience conspicuously higher or lower claim frequencies than average;
(e) defaults appear to be confined totally to reducible loans carrying a variable interest rate;
(f) claim frequency appears highest in relation to land, higher in relation to new properties than old, and lowest in relation to improvement loans

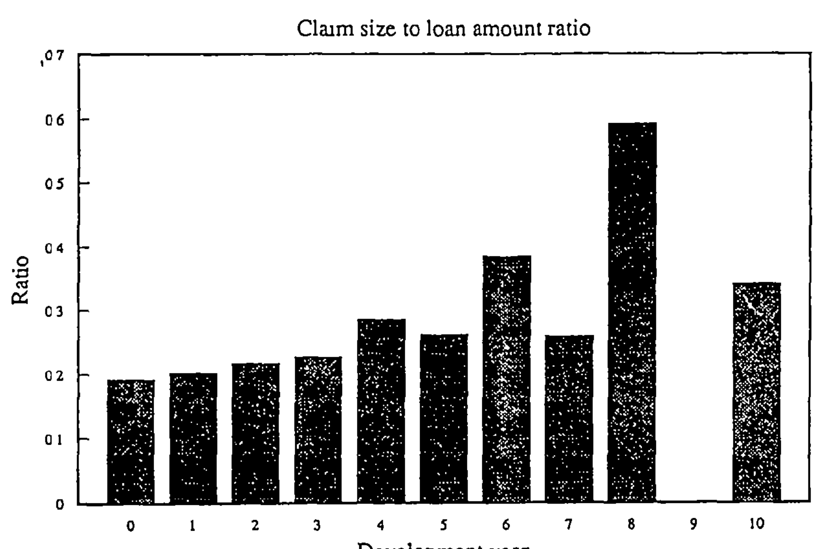
As stated, these are raised as **possibilities** only, rather than conclusions. Without further analysis, it would be impossible to determine whether all of these variables affect the default risk directly, or some of them are merely correlated with the genuinely operative risk variables.

For example, it might be the case that term of loan has no bearing on default risk, but appears to be relevant because LVR does have such a bearing and long terms are associated with high LVRs.

The question of possible correlation between risk variables is remarked upon further in Section 8.1.

5.2. Claim size

Initially, data concerning claim sizes were analysed according to the risk varibles listed in Section 4.1. This provided initial guidance concerning the



types of loans which led to larger or smaller losses when default occurred. The detailed results of this analysis are set out in Appendix D. The results indicate little variation in claim size with any of the risk variables except development year. The variation of claim size with development year is graphed in the preceding chart.

The chart suggests that the greater the time elapsed between advance of loan and default, the greater the **claim size to loan amount ratio**, i.e. the greater the loss on default expressed as a proportion of the original advance. This result is confirmed by formal regression analysis, as described in Section 8.2.

Since growth in property value generally increases with development year, this chart is consistent with the predicted form (7.2) of model.

6. FORM OF CLAIM FREQUENCY MODEL

6.1. General

In the following the basic units of tabulation of claims data will be referred to as **cells**. A cell will consist of an item of data associated with a particular combination of year of advance, development year, and any sub-set of the risk variables identified in Section 4.1.

It is reasonable that the total effect of risk variables on claim frequency should be **multiplicative**, i.e

(6.1) expected relative claim frequency = function (development year, HAI, HPI)
×
function (risk variables, e.g. LVR, geographic area, etc.).

The form of the first of the two functions on the right will be discussed in Section 6.2 As far as the second function is concerned, a reasonable first approximation would consist of the product of a factor in respect of each of the risk variables present. Equation (6.1) then becomes:

(6.2) expected relative claim frequency = function (development year, HAI, HPI)
×
factor dependent on LVR
×
factor dependent on geographic area
×
etc.

Interactions between the factors making up this product could be added if necessary.

Expected relative claim frequency (per loan advanced) is adjusted by a factor of 7/12 in all cells whose experience relates to 1984 This allows for the fact that the data include only 7 months' claims (Section 4 2).

Some of the risk variables identified in Section 4.1, e.g. financial type of loan, are categorical by nature. Others, e.g. LVR, are continuous by nature It was convenient for exploratory analysis of the data to convert all variables (i e. risk variables, not HAI and HPI) to categorical form Details appear in Section 5.1. The categorical form of data was retained in the final modelling, described in Section 8.1

6.2. Dependence on development year and economic variables

Preliminary analysis (Section 5.1) indicated that relative claim frequency, expressed as a function of development year, was generally consistent with the shape of a **Hoerl curve**. Appendix B provides a theoretical underpinning of this observation. Consequently, the model adopted for relative claim frequency in the absence of any other effects took the form:

$$\text{const} \times (j+\tfrac{1}{2})^{\alpha} \exp(-cj), \tag{6.3}$$

where j represents development year.

The modification of (6.3) by HAI and HPI raises some questions Consider HAI first.

As noted in Section 3.2, the HAI may be regarded as a measure of the average borrower's residual income after payment of mortgage instalment. An individual borrower will experience difficulties in payment of mortgage instalment if this residual income turns negative. The frequency with which this occurs in the event of movements of HAI will depend on the distribution of individual residual incomes, rather than just the average of this distribution represented by HAI. There is virtually no information available in respect of this distribution

There is, however, some evidence that individual gross incomes are subject to a Paretian distribution (MANDELBROT, 1960).

If a similar assumption is made about residual incomes after payment of mortgage instalment (i.e HAI), then Appendix A demonstrates that, to first approximation, logged claim frequency will contain a term linear in $R(i+j)/R(i)$, where i denotes year of advance, j development year, and $R(t)$ the HAI experienced in year t Allowance for the one year lag in the effect of HAI, as discussed in Section 3 4, modifies this term to $R(i+j-1)/R(i)$ (1 in the case $j=0$).

Because of the approximations leading to this result in Appendix A, an alternative linear term involving

$$\log[R(i+j-1)/R(i)] \quad \text{for} \quad j \geq 1;$$

or

$$0, \quad \text{for} \quad j=0, \tag{6.4}$$

was tried. This latter form produced a slightly better fitting regression than the unlogged ratio, and has been adopted henceforth. In fact, both alternatives produced quite similar results

Appendix B, particularly (B.10), demonstrates that, under seemingly reasonable assumptions about the accumulation of the amount of mortgage debt on default, and about property values on resale, claim frequency should also contain the following factor involving LVR and HPI:

$$L^{\nu}[H(i+j)/H(i)]^{-\nu}, \qquad \nu \text{ const.} > 0,$$

where L denotes LVR and $H(t)$ the HPI experienced in year t. In order to accommodate the lag in the effect of HPI discussed in Section 3 4, this last expression should be modified to the following.

$$L^{\nu}[H(i+j-\tfrac{1}{2})/H(i)]^{-\nu}, \qquad j \geq 1;$$

or

(6.5) $$L^{\nu}, \qquad j = 0,$$

where $H(t-\tfrac{1}{2})$ is interpreted a the HPI experienced at the **end** of year $t-1$.

Note that (6.5) indicates that claim frequency should include the **same power** of both LVR and HPI. However, this result was derived in Appendix B on the assumption that LVR affected the proportion of principal outstanding at default, but not the risk of default itself. In practice, it is likely that LVR is correlated with the ability of the borrower to meet financial commitments, in which case it intrinsically affects the risk of default For this reason, (6.5) should be generalized to the following:

$$L^{\lambda}[H(i+j-\tfrac{1}{2})/H(i)]^{-\nu}, \qquad j \geq 1;$$

or

(6.6) $$L^{\lambda}, \qquad j = 0.$$

Combination of (6.2) to (6.4) and (6.6) yields the following model:

(6.7) expected relative claim frequency in development year j of year advance i

$$\begin{aligned} &= \text{const.} \times (j+\tfrac{1}{2})^{\alpha} \exp(-cj) \\ &\quad \times L^{\lambda}[R(i+j-1)/R(i)]^{-\rho} \, [H(i+j-\tfrac{1}{2})/H(i)]^{-\nu} \\ &\quad \times \text{factor dependent on geographic area} \\ &\quad \times \text{etc. for } j \geq 1, \end{aligned}$$

and with the two square bracketed terms removed in the case $j = 0$.

Let $\mu(i,j)$ denote the expected relative claim frequency (6.7), and $E(i)$ the number of loans advanced in year i. Let $N(i,j)$ denote the number of claims emerging in development year j of year of advance i Then the claim frequency model adopted was:

(6.8) $$N(i,j) \sim \text{Poisson}\,[E(i)\,\mu(i,j)]$$

It should be noted that the precise form of dependency of relative claim frequency on LVR and HPI in (6.7) relies upon distributional assumptions made in Appendix B. If these assumptions were varied, the form of (6.7) would change. Consequently, an alternative to (6.7) is considered in Section 8.1, in which the terms involving LVR and HAI are replaced by:

$$\exp(\lambda L)\ \exp[-\nu\, H(i+j-½)/H(i)]\,.$$

This alternative model turns out to be inferior to (6.7).

7. FORM OF AVERAGE CLAIM SIZE MODEL

Appendix C shows that, on the same seemingly reasonable assumptions as in Appendix B (referred to in relation to the development of (6.5)), the average claim size in respect of loans advanced in year i should progress over development years according to the following parametric form:

$$E[Q(i,j)] = \text{const.} \times H(i+j)/H(i)\,, \tag{7.1}$$

where

$Q(i,j)$ = the claim ratio (i.e. ratio of claim size to original loan size) experienced in development year j of year of advance i;

$H(t)$ = HPI experienced during year t.

One may note the interesting effect whereby average claim size **increases** with development year even though outstanding principal is decreasing. Clearly this result derives from the assumptions made in Appendices B and C. Different assumptions would lead to a different parametric form in (7.1). However, an examination of the development of Appendix C indicates that the property of increasing $E[Q(i,j)]$ with $H(i+j)$ derives only from an assumption that the variable γ has a **decreasing failure rate**, where $\gamma = \alpha/\beta$ and

α = a random variable representing the factor by which outstanding principal has been enlarged after default by arrears of principal and interest and any other costs,

β = a random variable representing the factor by which the property value has been reduced by the forced nature of the sale and the associated expenses.

While there is no particular evidence concerning the failure rate of γ, it is interesting to note that the seemingly reasonable assumption of a Pareto distribution leads to the result (7.1) which is found in Section 8.2 to accord with experience, at least to the extent that the claim ratio trends upward with increasing property factor. However, because the Pareto assumption may be a little too specific, it is reasonable to widen the model (7.1) to the following:

$$Q(i,j) = a + b\, H(i+j)/H(i) + \text{error term}\,, \tag{7.2}$$

where approximately

$$\text{error term} \sim N(0, \sigma^2). \tag{7.3}$$

The appropriateness of this error term is discussed further in Section 8.2.

8. FITTING THE MODEL

8.1. Claim frequency

By (6.7) and (6.8),

$$\begin{aligned} \log E[N(i,j)] = {} & \log E(i) + \text{const.} + \alpha \log (j + ½) - cj \\ & + \lambda \log L - p \log [R(i+j-1)/R(i)] \\ & - \nu \log [H(i+j-½)/H(i)] \\ & + \text{term dependent on geographic area} \\ & + \text{etc.}, \; j \geq 1, \end{aligned} \tag{8.1}$$

with the two square bracketed terms on the right omitted for the case $j = 0$. This linear form, subject to the error structure (6.8), was fitted to the data using GLIM (Generalised Linear Interactive Modelling) (Royal Statistical Society, 1987). Various combinations of the potential explanatory variables listed in Section 4.1 were tried, and the main results are reported in the next table but one.

Geographic area		
Original coding (a)	First aggregation	Second aggregation
1, 4	Area 1	AREA 1
3	Area 3	
5	Area 4	
6	Area 5	
2	Area 2	AREA 2
7, 10-12	Area 6	AREA 3
9, 14	Area 7	
13	Area 9	
8	Area 8	AREA 4

(a) As set out in Section 4 2

The results of the trial regressions are displayed in the following table.

Variable	Coefficient in variable at left (a) in Regression No						
	1	2	3	4	5	6	7
Regression constant	−9 505	−12 18	−10 50	−9 848	−12 90	−5 776	−5 943
Development year	−1 093	− 1 143	− 1 218	−1 097	− 1 096	−1 119	−0 8536
Log (development year + ½)	4 908	5 066	4 558	4 906	4 903	5 076	4 505
LVR (d)	1 100	1 144	0 994	1 100	1 099		
Log (LVR)						8 93	8 413
Log (home affordability factor) (b)							−2 158
Property growth factor (c)	−3 039	− 3 070	− 2 036	−3 017	− 3 015		
Log (property growth factor)						−4 636	−5 658
Indicator variables (f)							
AREA 2				0 52	0 52	0 53	0 5131
AREA 3				0 87	0 87	0 87	0 8772
AREA 4				−5 24	− 5 24	−5 25	−7 254*
Area 2	0 60						
Area 3	0 16*						
Area 4	−0 35*						
Area 5	−0 26*						
Area 6	1 05						
Area 7	1 15						
Area 8	−5 33*						
Area 9	0 81						
60 ≤ Term < 120 months		3 74*					
120 ≤ Term < 180 months		2 95*					
180 ≤ Term < 240 months		2 00*					
240 ≤ Term		2 74*			3 06*		
Dwelling							
Improvements & increases			1 33*				
All other than improvements, increases & land only			0 64*				
Dwelling type missing			7 05*				
Deviance (e)	854	549	632	611	610	593	527

(a) Dependent variable in regression log (claim frequency), as in (8 1)
An asterisk attached to a coefficient in the table indicates that this coefficient differs from zero by less than 2 standard errors

(b) The home affordability factor is the ratio of values of HAI appearing in (8 1)

(c) The property growth factor is the ratio of values of HPI appearing in (8 1)

(d) The variable referred to here is in fact

$$10 \times LVR - 3\ 5$$

The variable log (LVR) uses the genuine LVR, though grouped in ranges of 10 percentage points width Each such range is represented by its mid-value

(e) Deviance is a measure of goodness of fit, related to the log likelihood ratio of the model A lower deviance implies a better fit

(f) The variables Area k and AREA m have already been described as 0-1 indicator variables The variables listed subsequently in the table are also of the 0-1 indicator type, taking the value 1 if the loan is subject to the risk variable displayed, 0 otherwise

By (6.8) and (8.1), the model is multivariate Poisson with multiplicative structure of the mean. GLIM fits this by maximum likelihood. Note that the logarithmic form of (8.1) is no more than a convenience of expression. It could equally have been written in its unlogged (multiplicative) form. In particular, (8.1) does **not** imply that the observations $N(i, j)$ (many of which are zero) are to logged

For the interpretation of this table, special reference should be made to geographic area of the mortgaged property. On the strength of the chart of Section 5.1, a number of areas, seemingly similar in claim frequency and/or physically contiguous, were aggregated. The areas at this initial level of aggregation were denoted by "Area k". These were 0-1 variables, taking the value 1 if the property lay in the relevant area, 0 otherwise.

Regression 1 in the table indicated that further aggregation was possible. The new variables resulting from this aggregation were denoted by "AREA m", and were 0-1 variables, each of which consisted of the sum of the relevant variables Area k. The key to the two aggregations is as shown in the previous table but one

It may be noted that the trial regressions included alternative versions of (8.1) in which the terms dependent on LVR and HPI were replaced by their respective unlogged forms, as discussed at the end of Section 6.2. These alternatives were, however, inferior to (8.1) in terms of fit.

Regression 7 provided the best fit of model to data, and was adopted as the final model. This final model, expressed in non-symbolic form, was as follows:

(8 2)

$$\begin{array}{ll}
\text{CLAIM FREQUENCY} = & 2\ 624\ (t+\tfrac{1}{2})^{4\ 505} \exp(-0\ 8536\ t) \\
\text{(per 1000 advances)} & \times \\
\text{IN DEVELOPMENT YEAR } t & (\text{LVR})^{8\ 413} \\
 & - \\
 & [(\text{HOME AFFORDABILITY FACTOR})^{2\ 158} \\
 & \times \\
 & (\text{PROPERTY GROWTH FACTOR})^{5\ 658}] \\
 & \times \\
 & \begin{cases} 1 \text{ if AREA } 1 \\ 1\ 670 \text{ if AREA } 2 \\ 2\ 404 \text{ if AREA } 3 \\ 0\ 0007 \text{ if AREA } 4 \end{cases}
\end{array}$$

where

HOME AFFORDABILITY FACTOR and PROPERTY GROWTH FACTOR are the ratios involving H and R respectively in (8.1).

The formula in the box indicates that claim frequency.

(a) moves sharply upward with increasing LVR,

(b) moves sharply downward as property values or disposable incomes after mortgage instalments increase;
(c) varies significantly by geographic area, exhibiting a particularly low value in the Wollongong district.

Because of correlations of the type discussed at the end of Section 5.1, not all of the risk variables exhibited a significant effect on claim frequency.

8.2. Average claim size

The form of the model was suggested in Section 7 as the following.

(7.2) $$Q(i,j) = a + b\,H(i+j)/H(i) + \text{error term},$$

where approximately

(7.3) $$\text{error term} \sim N(0, \sigma^2).$$

This model appears unnatural to the extent that the normal error term would permit claim sizes to be negative. This would be avoided by the inclusion of an error term which was by nature positive. An example would be a lognormal error term, as would be incorporated in an alternative model of the form:

(8.3) $$\log Q(i,j) = \log a + b \log [H(i+j)/H(i)] + \text{error term},$$

where

(8.4) $$\text{error term} \sim N(0, \sigma^2).$$

Equivalently,

(8.5) $$Q(i,j) = a[H(i+j)/H(i)]^b \times \text{error term},$$

where

(8.6) $$\text{error term} = \text{lognormal}\,(0, \sigma^2).$$

Note that both forms (7.2) and (8.5) accommodate the theoretical form (7.1)

Ordinary regression produced the following two alternative models.

Parameter	Unlogged model (a)	Logged model (b)
a	0 1622	0 1555
b	0 0494	0 3083
σ^2	0 0257	0 8676

(a) This is the model described by (7 2) and (7 3) Of the 425 observed claim ratios, 2 large values have been excluded as outliers
(b) This is the model described by (8 3) and (8 4)

In fact, neither of the two models considered in the preceding table produced an ideal fit to the data. Their respective residuals are tabulated in the following table.

Values of standardized residuals	Relative frequency of standardized residual in	
	Unlogged model	Logged model
	%	%
less than −3	0	1
−3 to −2	0	3
−2 to −1	12	8
−1 to 0	47	32
0 to 1	24	44
1 to 2	10	12
2 to 3	5	0
more than 3	1	0
Total	100	100

These two tabulations of standardized residuals are very much reflections of each other about the origin. While the unlogged model is somewhat skewed to the right, the logged model is about equally skewed to the left. This suggests that the correct model lies somewhere between normal and log normal. Such a model might be of the form (7.2), but with the error term strictly positive and skewed to the right but less so than log normal.

Note that the fitted values of claim ratios, according to the two alternative models, are:

(8.7)
$$EQ(i,j) = a + bH(i+j)/H(i) \text{ for unlogged model;}$$
(8.8)
$$= a[H(i+j)/H(i)]^b \exp(\tfrac{1}{2}\sigma^2) \text{ for logged model.}$$

In the event, (8.8) produced a rather heavy upward bias, about 18% in total, in fitted values of claim amount relative to observed amounts. The form of this comparison was exactly as reported in Section 9.2, but with the unlogged model used there replaced by the logged.

This result appears to indicate that the exponential scaling factor in (8.8) is not robust against the non-normality in the error term of (8.4), as was demonstrated in the above table of standarized residuals.

On the other hand, Section 9.2 indicates that the unlogged model provides an adequate fit, and accordingly it was adopted.

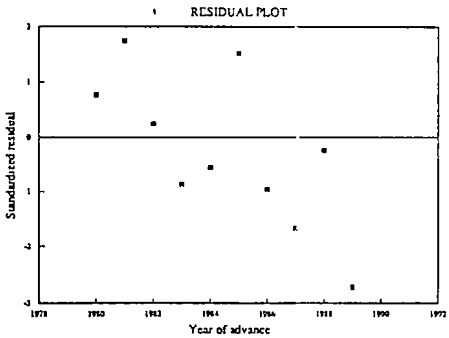
9 MODEL VERIFICATION

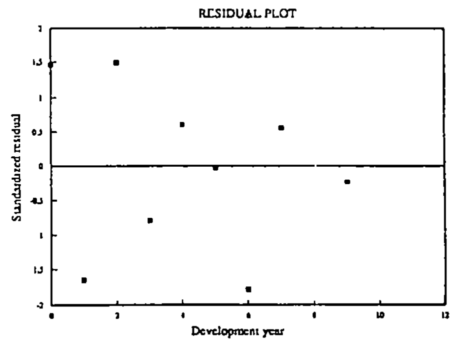
9.1. Claim frequency

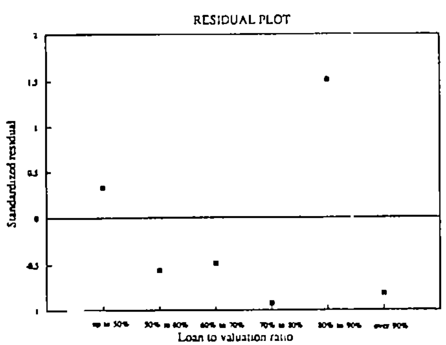
The model adopted in Section 8.1 has been used to compute standardized residuals according to several variables. The resulting residual plots appear below. Note that each residual relates to the aggregation of all experience at the value of the independent variable displayed. For example, the first residual in the first plot may be obtained from the second table of the present sub-section as:

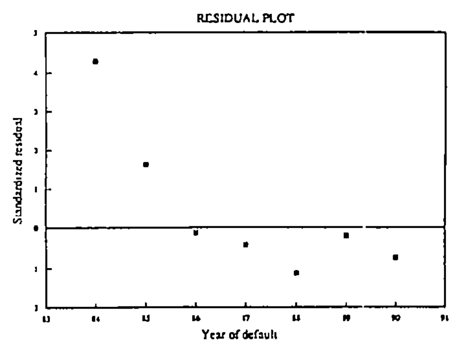
$$(8-6)/\sqrt{6} = 0.8$$

A plot of the residuals of all cells (taken over all explanatory variables) would not be helpful since the great majority of cells contain very small expectations.









These plots appear generally satisfactory in terms of magnitude, with the exception of year of default 1984. This one anomaly, in the relatively distant past, involves relatively few claims (see first table below) and is insufficient to invalidate the model.

The plot against year of advance contains a downward trend. If included in the model, year of advance appears as a highly significant explanatory variable; other things equal, claim frequency declines by 29% as between each year of advance and the next. Naturally, the effects of the other explanatory variables, particularly those which are time dependent, change.

While this model provides a superior fit to the data, the abstract nature of the year of advance effect is problematic It might be interpreted as a factor representing improvement in underwriting. However, in this case, the total improvement over the decade of underwriting would be almost 97%, which might strain credulity.

It seems more likely that year of advance is acting as a proxy for some other unidentified explanatory variable(s). When this variable is omitted from the model, its effect is largely captured by the other explanatory variables.

Moreover, an examination of the fitted numbers of claims (using the model which omits year of advance effect) against the data suggests that the apparent trend in the residuals may not be particularly meaningful (see second table below)

The following table displays the actual and model numbers of claims underlying the above plot of standardized residuals by experience year.

Period	Number of claims emerging	
	Actual	Model
1984 (7 months)	28	13
1985	32	24
1986	53	54
1987	168	174
1988	103	115
1989	21	22
1990	20	24
Total	425	425

The table illustrates the close agreement between actual and model numbers of claims for all experience years except 1984, despite the extreme fluctuations in numbers of claims.

More detailed information is given by the following table which tabulates experience and model simultaneously by year of advance and development year, and from which the above table may be derived.

Year of loan advance	Observed and fitted (shown in bold type) number of claims in development year											
	0	1	2	3	4	5	6	7	8	9	10	Total
1980					3 **1.8**	3 **1.5**	1 **1.2**	0 **1.2**	0 **0.3**	0 **0.0**	1 **0.0**	8 **6**
1981				13 **4.5**	8 **4.8**	6 **4.4**	1 **4.9**	0 **1.4**	0 **0.1**	0 **0.0**		28 **20**
1982			7 **4.9**	6 **7.6**	10 **8.7**	8 **11.4**	3 **3.5**	3 **0.3**	1 **0.1**			38 **37**
1983		5 **1.6**	7 **5.3**	7 **8.8**	8 **14.7**	3 **5.2**	0 **0.5**	1 **0.2**				31 **36**
1984	0 **0.1**	7 **4.3**	13 **15.5**	30 **37.7**	19 **16.8**	3 **1.8**	0 **0.8**					72 **77**
1985	1 **0.3**	16 **16.2**	104 **86.6**	53 **56.7**	12 **7.6**	5 **3.8**						191 **171**
1986	0 **0.2**	14 **17.1**	24 **24.6**	3 **4.8**	2 **3.1**							43 **50**
1987	3 **0.3**	1 **6.2**	0 **2.7**	2 **2.5**								6 **12**
1988	0 **0.4**	0 **2.7**	8 **5.6**									8 **9**
1989	0 **0.2**	0 **7.1**										0 **7**
1990	0 **0.3**											0 **0**

The following table presents these results in the same format as in Section 2, enabling comparison of the present set of results with those from the separation method

Year of loan advance	Observed and fitted (shown in bold type) relative claim frequency in development year											
	0	1	2	3	4	5	6	7	8	9	10	Total
1980					30 **18**	18 **9**	6 **7**	0 **7**	0 **2**	0 **0**	6 **0**	60 **43**
1981				116 **41**	42 **25**	31 **23**	5 **26**	0 **7**	0 **1**	0 **0**		195 **122**
1982			54 **38**	27 **34**	45 **39**	36 **51**	13 **16**	13 **1**	4 **0**			193 **179**
1983		25 **8**	20 **16**	20 **26**	23 **43**	9 **15**	0 **1**	3 **1**				101 **109**
1984	0 **0**	13 **8**	24 **28**	55 **69**	35 **31**	5 **3**	0 **1**					131 **140**
1985	1 **0**	21 **21**	134 **111**	68 **73**	15 **10**	6 **5**						245 **220**
1986	0 **0**	17 **21**	30 **30**	4 **6**	2 **4**							53 **62**
1987	3 **0**	1 **6**	0 **3**	2 **3**								6 **12**
1988	0 **0**	0 **2**	5 **3**									5 **5**
1989	0 **0**	0 **6**										0 **6**
1990	0 **0**											0 **0**

9.2. Average claim ratio

For each claim in the experience, a fitted value of its claim ratio was calculated according to (8 7) using the values of a and b tabulated in Section 8.2. Each of these claim ratios was multiplied by the associated amount of its loan, to produce a fitted claim size.

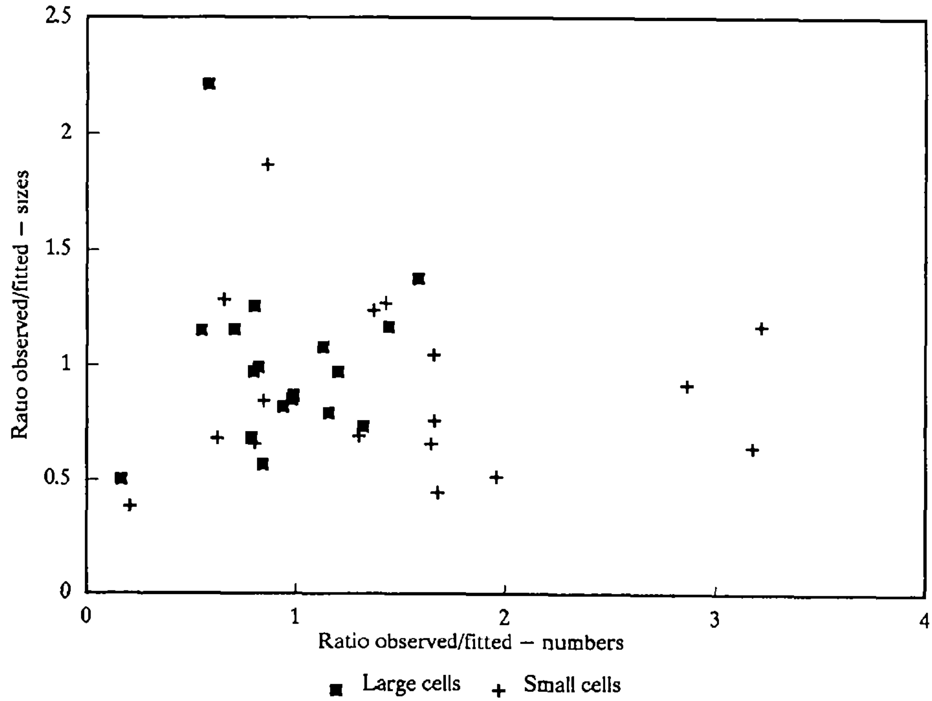
Observed and fitted claim sizes were then summarized in 2-way tabulations by year of advance and development year. These tabulations are displayed in Appendix E, and reduced to their corresponding 1-way tabulations below.

Year of advance	Amount of claims			Development year	Amount of claims		
	Observed	Fitted	Ratio Observed/fitted		Observed	Fitted	Ratio Observed/fitted
	\$ 000	\$ 000	%		\$ 000	\$ 000	%
1980	51	70	73	0	32	46	70
1981	294	312	94	1	425	471	90
1982	398	374	106	2	1750	1844	95
1983	354	323	110	3	1051	1133	93
1984	632	642	98	4	674	642	105
1985	1931	2063	94	5	321	301	107
1986	425	472	90	6	47	38	124
1987	46	69	67	7	31	35	88
1988	259	222	117	8	56	28	199
1989	0	0		9	0	0	
1990	0	0		10	1	7	14
Total	4388	4545	97		4388	4545	97

It should be particularly noted that the fitted amounts of claims, according to the above description are **conditional** upon the observed numbers of claims. This is a proper approach to examination of the fit of the average claim size model. Agreement between model and data appears satisfactory.

It is useful to carry out some check that the common dependence of the claim frequency and claim size models on the HPI does not lead to unwanted correlation between the two. That this does not in fact occur is indicated by the following scatter plot of the observed fitted ratios of average claim size against a similar ratio for number of claims.

Each point represents a particular combination of year of advance and development year. To give a simple indication of the significance of the plotted points, they are divided into "large cells" and "small cells". The former are those cells containing a fitted number of claims in excess of 5; otherwise the cell is "small".



9.3. Loan sizes associated with claims

While Section 9.2 models the claim size which will arise from a particular loan size **if a claim occurs**, it provides no indication of which loan sizes are likely to lead to claims

There is no particular reason to believe that the sizes of loans associated with claims will be representative of the entire portfolio of loans advanced. Indeed, the table below indicates that, on average, it is the larger loans that lead to claims.

Care is needed here, however, as the model of claim frequency in Section 9 1 conditions on LVR and other risk factors, for which average loan sizes may differ from the portfolio average, and so without further analysis it is not clear to what extent the inclusion of these factors in the model will effectively select average loan sizes above the portfolio average. This question is also examined in the following table.

Year of advance	As a percentage of portfolio average loan size		
	average loan size associated with past claims (a)		average loan size weighted by model numbers of future claims (b)
	%		%
1980	135	(8)	96
1981	144	(28)	102
1982	119	(38)	101
1983	116	(31)	102
1984	85	(72)	102
1985	95	(191)	102
1986	144	(43)	103
1987	97	(6)	100
1988	241	(8)	98
Average	109 (c)	(425)	102 (d)

(a) The numbers of claims on which the ratios are based are shown in parenthesis For each year of advance, the average size of loans associated with **recorded** claims has been calculated and related to the portfolio average (for that year of advance)

(b) For each combination of year of advance and risk variables, the average loan advanced and model claim frequency (according to the model of Section 8 1) are calculated The average loan advanced, weighted by model claim frequency, is then calculated for each year of advance

(c) Average of the entries in the column, weighted by numbers of claims shown in parenthesis

(d) Unweighted average of the entries in the column

The table suggests that the average loan size associated with claims of a particular cell for a particular year of advance is about 7% higher than the overall average loan size for the cell.

Thus, a forecast of future claim amount for a particular cell of development year j of year of advance i would be computed as·

$$1.07 \times \text{average loan size in year of advance } i \times \hat{N}(i,j)\,\hat{Q}(i,j),$$

where $\hat{N}(i,j)$, $\hat{Q}(i,j)$ are estimates of $N(i,j)$ and $Q(i,j)$ from Sections 9.1 and 9.2.

An alternative approach to the above would be to include loan size as an explanatory variable in the claim frequency model of Section 8 1. This might be awkward in practice, however, because it would increase very considerably the number of data cells entering into the regressions of Section 8.1.

10. CONCLUSION

Section 8 fits models to the claim frequency and claim ratio in the mortgage insurance portfolio examined. Section 9 verifies that these models provide a reasonable fit to the data.

The models therefore can be, and indeed have been, used to estimate the liability for claims still to emerge in respect of past years of loan advance. In order to carry out this estimation, one needs to project future values of the HAI and HPI. This in turn requires projection of incomes, tax rates, mortgage interest rates and growth in property values Projections such as these are, problems of substance in their own right, but are beyond the scope of the present paper

11. ACKNOWLEDGEMENT

I should like to acknowledge the computing assistance provided by Mr A.J Greenfield in the preparation of this paper.

APPENDIX A

DEPENDENCE OF CLAIM FREQUENCY ON HOME AFFORDABILITY INDEX

Let X denote the random variable representing the proportion of an individual's income required for tax, consumption and mortgage instalment. Assume this variable to be Pareto distributed, i.e. with p.d.f.

(A.1) $$f(x) = kx^{-\alpha-1}, \ k \text{ const.}$$

The borrower will experience financial difficulties if $X \geq 1$, which occurs with probability:

(A.2) $$P[X \geq 1] = kx^{-\alpha}/\alpha|_{x=1}$$

Now, suppose that X shifts by a factor of c to $X' = cX$. Then the probability (A.2) shifts to

(A.3) $$P[X' \geq 1] = P[X \geq 1/c] = kx^{-\alpha}/\alpha|_{x=1/c}$$

Comparison of (A.2) and (A.3) shows that the probability (A.2) has shifted by a factor of c^{α}. Now note that the scale shift of X to cX must shift the mean of X by a factor of c:

(A.4) $$E[X'] = cE[X]$$

Let

$$Y = 1 - X,$$

and note that

$$E[Y] \ \alpha \ \text{HAI}. \tag{A.5}$$

Then the factor by which HAI changes when X changes to X' is:

$$R = \{1 - E[X']\}/\{1 - E[X]\} \tag{A.6}$$
$$= (1 - c\mu)/(1 - \mu),$$

where

$$\mu = E[X].$$

Inversion of (A.6) yields:

$$c = [1 - R(1 - \mu)]/\mu. \tag{A.7}$$

Thus, the shift in HAI by a factor of R causes the frequency with which borrowers experience difficulties to shift by a factor of:

$$c^\alpha = \{[1 - R(1 - \mu)]/\mu\}^\alpha. \tag{A.8}$$

Now, it is convenient to analyse log (claim frequency), which will depend on log (frequency of borrower's difficulties), and (A.8) shows that this latter will depend on an additive term of:

$$\log c^\alpha = \alpha \log \{[1 - R(1 - \mu)]/\mu\}$$
$$\sim -\alpha R(1 - \mu) + \text{const.},$$

for small values of $(1 - \mu) R$.

Thus, to first approximation, the model of expected log (claim frequency) should include a linear term in R, the ratio by which HAI has changed since advance of the loan(s) in question.

APPENDIX B

DEPENDENCE OF CLAIM FREQUENCY ON HOUSING PRICE INDEX, LVR AND DEVELOPMENT YEAR

Consider a loan taken at time $t = 0$ Let $V(t)$ be the value of the associated property at time t, and $P(t)$ the amount of principal then outstanding. Then

$$V(t) = V(0)[H(t)/H(0)], \tag{B.1}$$

$$P(t) = P(0)f(t), \tag{B.2}$$

where

$H(t) =$ HPI at time t,

$f(t) =$ proportion of principal still to be repaid at time t

By (B.1) and (B.2),

$$P(t)/V(t) = Lf(t)\, H(0)/H(t), \tag{B.3}$$

where

(B.4) $$L = P(0)/V(0) = \text{loan to valuation ratio}.$$

Suppose that the borrower has encountered financial difficulties at some time $s < t$. At time t sale of the property is forced. At that point, the debt in respect of the loan will be $P(t)\,\alpha(t)$, where

$\alpha(t)$ = a random variable representing the factor by which outstanding principal has been enlarged by arrears of principal and interest and any other costs.

Similarly, the net proceeds of the sale of the property will be $V(t)\,\beta(t)$, where

$\beta(t)$ = a random variable representing the factor by which the property value has been reduced by the forced nature of the sale and the associated expenses

Then the ratio of outstanding debt to sale proceeds is.

(B.5) $$X(t) = \gamma(t)\,P(t)/V(t),$$

where

(B.6) $$\gamma(t) = \alpha(t)/\beta(t)$$

By (B.3) and (B.5),

(B.7) $$X(t) = L[H(t)/H(0)]^{-1} f(t)\,\gamma(t).$$

A claim will occur if $X(t) > 1$, i e if

(B.8) $$\gamma(t) > [H(t)/H(0)]\,[Lf(t)]^{-1}.$$

Now suppose that $\gamma(t)$ is Pareto distributed with d f.

(B.9) $$F(\gamma) = 1-(\gamma/a)^{-\nu}, \qquad \gamma > a,$$

assumed independent of t. Then, by (B 8), the probability of occurrence of a claim is:

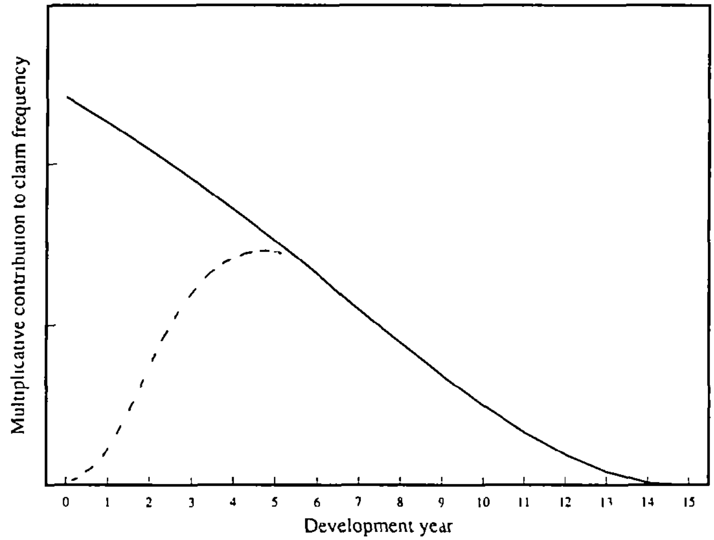
(B 10) $$P[X(t) > 1] = \{af(t)\,L[H(t)/H(0)]^{-1}\}^{\nu}.$$

Thus, expected claim frequency varies as a power of $L[H(t)/H(0)]^{-1}$. Note also that claim frequency for policies of a particular term n varies over development years t by a factor of

(B.11) $$[f(t)]^{\nu} \quad \alpha \quad [a_{\overline{n-t|}}]^{\nu},$$

which has the shape illustrated by the solid line in the following diagram

However, note the above assumption that the distribution of the factor $\gamma(t)$ is independent of t. While perhaps largely true, it will break down as $t \to 0$ as the screening procedures of the lender force claim frequency toward zero. Hence, the curve (B.11) of frequency over development year will be modified for small t in the manner indicated by the broken line in the diagram.



When allowance is made for the variety of original terms n, the dependence of claim frequency on development year is seen to be represented by a weighted average of curves of the type illustrated in the diagram.

APPENDIX C

DEPENDENCE OF AVERAGE CLAIM SIZE ON HOUSING PRICE INDEX

As noted just prior to (B.8), the financial difficulties of a borrower will lead to a claim if $X(t)$, as defined there, exceeds 1. In fact, by the same argument as led to that result, the amount of the claim will be

(C.1) $$A(t) = \alpha(t)\, P(t) - \beta(t)\, V(t) = \beta(t)\, V(t)\,[X(t) - 1].$$

Note that $\beta(t)$ and $\gamma(t)$ (and hence $X(t)$) will **not** be independent, even if $\alpha(t)$ and $\beta(t)$ are. For general random variables Y and Z, let μ_Y and μ_Z denote their means, v_Y and v_Z their coefficients of variation, and ρ_{YZ} their correlation. It is straightforward to demonstrate that:

(C.2) $$E[YZ] = \mu_Y \mu_Z (1 + \rho_{YZ} v_Y v_Z).$$

By (C.1) and (C.2),

(C.3) $$E[A(t)] = V(t)\, E[X(t) - 1]_+\, \mu_\beta (1 + \rho_{\beta X} v_\beta v_X),$$

where $E[Y]_+$ denotes $E[Y | Y > 0]$.

Now, by (B.5)

(C.4) $$E[X(t) - 1]_+ = E[\gamma(t) - V(t)/P(t)]_+\, P(t)/V(t).$$

By the Pareto assumption (B.9),

$$E[\gamma(t)-V(t)/P(t)]_+ = [V(t)/P(t)]\,\nu/(\nu-1), \tag{C.5}$$

whence (C.3) and (C.4) yield:

$$\begin{aligned} E[A(t)] &= V(t)\,\mu_\beta(1+\rho_{\beta X}v_\beta v_X)\,\nu/(\nu-1) \\ &\propto V(0)\,H(t)/H(0) \qquad \text{[by (B.1)]} \end{aligned} \tag{C.6}$$

if μ_β, v_β, v_X and $\rho_{\beta X}$ are the assumed independent of t.

Thus, the expected average claim size is directly proportional to property values, all other things equal. This has the interesting effect of causing average claim size in respect of a group of identical policies usually to **increase** with development year even though outstanding principal is decreasing.

APPENDIX D

EXPLORATORY ANALYSIS OF CLAIM SIZE

D1. Variation of claim ratio with loan to valuation ratio

Loan to valuation ratio	Number of claims	Claim to loan ratio		95% confidence limits (a)	
		Sample mean	Sample standard deviation	Lower	Upper
up to 50%	1	55.8%			
50 to 60%	1	56.9%			
60 to 70%	8	23.3%	13.7%	11.8%	34.8%
70 to 80%	36	23.9%	19.2%	17.4%	30.4%
80 to 90%	189	22.9%	18.4%	20.3%	25.6%
over 90%	191	23.5%	15.6%	21.3%	25.7%

(a) These are the symetric t-distribution confidence limits. Where the sample size is less than 2, the confidence limits do not exist.

D2. Variation of claim ratio with term

Term	Number of claims	Claim to loan ratio		95% confidence limits (a)	
		Sample mean	Sample standard deviation	Lower	Upper
months					
60 to 119	3	36.4%	14.1%	1.3%	71.4%
120 to 179	16	34.8%	29.8%	18.9%	50.7%
180 to 239	55	28.4%	20.2%	22.9%	33.9%
240 & more	352	22.0%	15.6%	20.4%	23.7%

(a) See Appendix D1

D3. Variation of claim ratio with area

Area	Number of claims	Claim to loan ratio		95% confidence limits (a)	
		Sample mean	Sample standard deviation	Lower	Upper
M1, M4	29	16.5%	11.7%	12.0%	20.9%
M2	63	21.2%	15.0%	17.5%	25.0%
M3	77	16.5%	12.6%	13.7%	19.4%
M5	5	25.8%	14.8%	7.5%	44.1%
Canberra	4	23.1%	13.0%	2.4%	43.8%
Coastal	100	24.6%	18.2%	21.0%	28.2%
Newcastle	32	31.7%	17.2%	25.6%	37.9%
Wollongong	0				
Other	116	27.5%	19.4%	23.9%	31.1%

(a) See Appendix D1

D4. Commentary

All pairs of confidence limits in Appendices D1 to D3 straddle the overall mean of 23.4% except in four cases. All four of these cases relate to area of residence, and are found in Appendix D3.

APPENDIX E

COMPARISON OF OBSERVED AND FITTED CLAIM AMOUNTS

The following are the amounts of claim **observed** in respect of each combination of year of advance and development year.

Year of advance	Amount of claims observed in development year										
	0	1	2	3	4	5	6	7	8	9	10
	$	$	$	$	$	$	$	$	$	$	$
1980					28522	13349	7873	0	0	0	1009
1981				115151	69711	105156	3724	0	0	0	
1982			71488	29799	102851	81026	35484	20827	56169		
1983		60085	71469	61801	85959	64416	0	10110			
1984	0	45337	68811	325411	180820	11766	0				
1985	9591	161743	1060021	474840	179612	44976					
1986	0	150351	219581	28174	26638						
1987	22882	7054	0	15810							
1988	0	0	258976								
1989	0	0									
1990	0										

The following are the amounts of claims **fitted** to each combination of year of advance and development year by the procedure described in Section 9.2.

Year of advance	Amount of claims fitted in development year										
	0	1	2	3	4	5	6	7	8	9	10
	$	$	$	$	$	$	$	$	$	$	$
1980					27287	25853	9332	0	0	0	7380
1981				125940	91833	84727	9687	0	0	0	
1982			56280	43406	129344	70032	19012	27658	28253		
1983		51324	96763	63585	74571	29094	0	7572			
1984	0	68421	121228	258339	167683	26301	0				
1985	14819	185929	1089849	576994	130423	64647					
1986	0	151670	258058	41149	20740						
1987	30697	13995	0	23866							
1988	0	0	221693								
1989	0	0									
1990	0										

Each cell in this table is of the form:

actual number of claims

×

fitted average claim size.

Hence comparison of the table with the previous one examines only variation of experience from model amounts of claim.

An alternative version of the preceding table consists of cells of the form:

fitted number of claims

×

fitted average claim size.

This table is as follows.

Year of advance	Amount of claims fitted in development year										
	0	1	2	3	4	5	6	7	8	9	10
	$	$	$	$	$	$	$	$	$	$	$
1980					16472	13202	11077	**0**	**0**	**0**	52
1981				44040	55444	61935	47805	**0**	**0**	**0**	
1982			39396	55278	111986	99883	22086	2637	2910		
1983		15962	73512	80326	136558	50459	**0**	1408			
1984	**0**	41551	144634	324560	148532	15693	**0**				
1985	4668	188718	907194	617384	82395	49662					
1986	**0**	185146	264079	66099	31805						
1987	3131	86881	**0**	29785							
1988	**0**	**0**	153966								
1989	**0**	**0**									
1990	**0**										

For cells in which where are no claims observed, the procedure of Section 9 2 does not produce a fitted average claim size These cells, **indicated in bold,** have been assigned a fitted amount of claims equal to zero

APPENDIX F

HOME AFFORDABILITY INDEX

Year (as at 31 December)	Economic indicators			Household expenditure					
	Average weekly earnings	Consumer price index	Mortgage interest rates (a)	Gross household income (b)	Tax (b)	Consumer expenditure (b)	Mortgage instalment (b)	Residual income: Amount	Residual income: As percentage of gross income
	$		p a	$ per week	$ per week	$ per week	$ per week	$ per week	
1978	224 35	82 4	11 50%	562 74	118 28	326 21	64 40	53 85	9 569%
1979	246 00	91 1	11 50%	617 05	129 70	360 65	70 61	56 08	9 089%
1980	278 25	100 0	12 00%	697 94	146 70	395 89	82 26	73 10	10 473%
1981	315 90	110 2	14 50%	792 38	166 55	436 27	107 18	82 39	10 397%
1982	346 70	123 4	15 50%	869 64	182 79	488 52	123 78	74 54	8 572%
1983	375 90	130 9	14 00%	942 88	198 19	518 22	124 22	102.26	10 846%
1984	405 40	136 0	13 50%	1016 88	213 74	538 41	130 41	134 33	13 210%
1985	428 20	147 5	15 00%	1074 07	225 76	583 93	149 07	115 30	10 735%
1986	450 85	161 4	15 50%	1130 88	237 70	638 96	160 96	93 25	8 246%
1987	477 70	173 7	14 50%	1198 23	251 86	687 66	162 07	96 64	8 066%
1988	521 65	187 7	14 25%	1308 47	275 03	743 08	174 68	115 68	8 841%
1989	560 75	203 0	17 25%	1406 55	295 64	803 65	217 77	89 48	6 362%
1990	600 68	213 0	15 50%	1506 69	316 69	843 24	214 46	132 30	8 781%

(a) The most common interest rates applying to loans in the mortgage insurance portfolio under analysis

(b) These four columns were derived in a consistent manner from the HES, as described in Section 3 2

APPENDIX G

DATA

The data described in Section 4.2 are summarized in the following table. This should be considered in conjunction with the qualification set out in the final paragraph of Section 4.2.

Year of advance	Number of loans advanced	Number of claims (a) recorded in development year										
		0	1	2	3	4	5	6	7	8	9	10
1980	1700					3	3	1	0	0	0	1
1981	1917				13	8	6	1	0	0	0	
1982	2231			7	6	10	8	3	3	1		
1983	3426		5	7	7	8	3	0	1			
1984	5496	0	7	13	30	19	3	0				
1985	7787	1	16	104	53	12	5					
1986	8077	0	14	24	3	2						
1987	9910	3	1	0	2							
1988	17646	0	0	8								
1989	11878	0	0									
1990	13614	0										

(a) Development year is defined as year of emergence of claim minus year of loan advance Claims emerging in 1984 represent the experience of only 7 months

PREMIUM RATING BY GEOGRAPHIC AREA USING SPATIAL MODELS

BY M. BOSKOV AND R. J. VERRALL

Department of Actuarial Science and Statistics
The City University

ABSTRACT

This paper gives a method for premium rating by postcode area The method is based on spatial models in a Bayesian framework and uses the Gibbs sampler for estimation. A summary of the theory of Bayesian spatial methods is given and the data which was analysed by TAYLOR (1989) is reanalysed An indication is given of the wide range of models within this class which would be suitable for insurance data. The aim of the paper is to introduce the models and to show how they can be utilised in an insurance setting.

KEYWORDS

Gibbs sampler; Postcodes; Premium rating; Spatial statistics

1. INTRODUCTION

The problem of accessing risk as a function of geographical area occurs in a number of fields, including insurance rating and epidemiology. The aim of the statistical analysis of the data is to assess the underlying variation in risk by area, usually postcode area. Two approaches can be taken. Either the raw data can be smoothed in order to remove as much random variation as possible, or the data can be used to allocate each postcode area to a rating category, allowing for the inherent random variation The example in this paper uses the first approach, although the methods can also be used for the second approach. The authors believe that the second approach may be more satisfactory if the data are in a suitable form.

The only previous paper, of which the authors are aware, which uses mathematical and statistical techniques for premium rating by postcode area is TAYLOR (1989). That paper used two-dimensional splines on a plane linked to the map of the region by a transformation chosen to match the features of the specific region. The present paper uses an entirely different approach, although some of the preprocessing aspects of the analysis wil be the same as those used by TAYLOR (1989). The example in Section 4 of this paper uses the data from TAYLOR (1989) As will become clear, there are disadvantages in using the data in the form available from that paper. The example is valid in that it applies a suitable model to the particular data set given However, the present authors believe that a slightly different model based on data for claim numbers and claim amounts separately could provide more informative results.

The methods described here are based on statistical methods for spatial data These methods have been developed for image restoration, often using data from satellites. However, the techniques can also be used for risk assessment in an insurance setting The aims of the analysis and some of the assumptions underlying the models differ from those in other applications, but the statistical and mathematical techniques are similar. The basis of the method is the use of a spatial probabilistic model in a Bayesian context. The Gibbs sampler is used to derive the posterior distribution from which inferences about the spatial structure of the data can be made These inferences can be used to assess the risk due to the geographic area. The basic philosophy is that there is an underlying "true" risk pattern over the whole region, and the data are a version of this pattern contaminated by random noise. The aim of the model is to reconstruct the "true" picture as far as possible. The analogy with image restoration is clear

The literature on spatial methods is large, and we mention just a few references which are particulary relevant to the work in this paper. The book by CRESSIE (1991) provides a useful overview and summary of the field. BESAG et al (1991) gives a summary of the Bayesian models and describe applications in archeology and epidemiology. The use of the Gibbs sampler was the subject of a discussion meeting at the Royal Statistical Society recently The papers and discussion are contained in part 1 of the Journal of the Royal Statistical Society, 1993 We would mention particulary GILKS et al. (1993) and SMITH and ROBERTS (1993).

The paper is set out as follows. Section 2 contains a specification of the spatial model. Section 3 describes the Gibbs sampler and simulation techniques which are used to estimate the posterior densities. Section 4 contains an example using the data from TAYLOR (1989) and the final section has the conclusions.

2. A BAYESIAN MODEL FOR SPATIAL DATA

The basis for any model for spatial data is that areas which are close together are more likely to be similar (in some sense) than areas which are far apart. In the context of image restoration, this would mean that adjacent areas would be likely to be the same, or similar, colour. In an insurance context, it means that we expect adjacent areas to be similar from the point of view of the underlying risk.

It is important to remember that we are interested in the true, underlying risk, and the data is just a sample providing an estimate of this risk In addition, we are considering only the risk due to geographical area. We will assume that the other factors have already been analysed, using (for example) a generalised linear model.

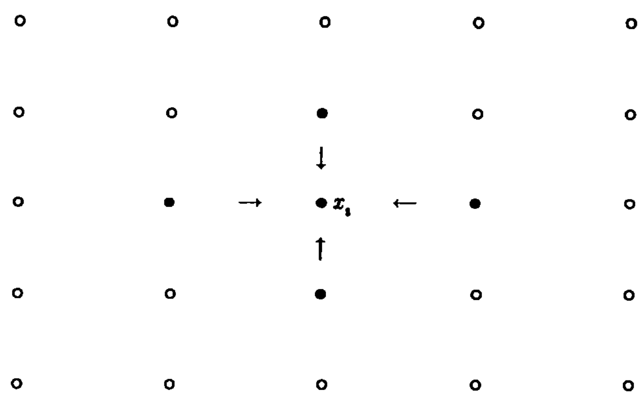
We assume that the geographical areas are numbered from 1 to n. Usually, the areas will correspond to postcode areas. Define x_i to be the true risk in area i and $\underline{x}$ to be the vector of risks over the whole region $\{x_i \,.\, i = 1, \ldots, n\}$. The joint prior distribution of $\underline{x}$ is not specified explicitly Instead, it is more useful to define the conditional densities

$$p_i(x_i \mid x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \qquad i = 1, \ldots, n. \tag{2 1}$$

This conditional density is the density of the risk at one location, given the risk at all the other locations. In reality, this will not depend on the risk at most of the other locations. This means that we can replace (2 1) by the conditional density of x_i, given the risk values in the neighbourhood of area i.

(2 2) $$p_i(x_i \mid \delta_i)$$

δ_i is defined as areas in the neighbourhood of the ith area For example, if we had an evenly spaced lattice, the prior distribution might be defined so that δ_i consisted of adjacent points One possibility is illustrated in the following diagram



In the insurance setting, δ_i can be interpreted as postcode areas which are adjacent to, or close, to, the ith area.

Suppose that the data observed are denoted by $\underline{y}$ with components $\{y_i \cdot i = 1, \ldots, n\}$. We use a simplified notation here, giving only the random variable y_i, and not the other (possibly non-random) information which may be in the data. The full likelihood may be found from

(2 3) $$f(\underline{y} \mid \underline{x}) = \prod_{i=1}^{n} f(y_i \mid x_i)$$

This assumes, as is reasonable, that the data are conditionally independent, given $\underline{x}$ The posterior density of $\underline{x}$, given $\underline{y}$, can be found using Bayes theorem:

(2 4) $$p(\underline{x} \mid \underline{y}) \propto f(\underline{y} \mid \underline{x})\, p(\underline{x})$$

The usual Bayesian estimate of $\underline{x}$ is the value of $\underline{x}$ which maximises the posterior density, the maximum *a posteriori* estimate. Of course, the most diffficult part of this maximisation is to actually determine the posterior density $p(\underline{x} \mid \underline{y})$ Although we have the conditional prior distributions given by (2 2), it is not straightforward to find the unconditional prior distribution and the posterior distribution. Instead, we exploit the conditional densities to obtain realisations from the posterior density. After obtaining a sufficient number of realisations, we may use the empirical density generated to find maximum *a posteriori* estimates. In other words, the estimation is based on a Monte Carlo method. The mechanics of this, which are based on a variant of the Metropolis algorithm called the Gibbs sampler (GEMAN and GEMAN (1984)) are given in Section 3.

x_i has been defined as the true risk in area i, and we now make the compounds of x_i more explicit. The risk level is assumed to be the sum of three components.

$$x_i = t_i + u_i + v_i \tag{2.5}$$

t_i is based on known factors. It is measured through covariates using, for example, a generalised linear model We shall assume that this component of the risk has already been removed from the data. In effect, we assume that the data have already been "standardised" to remove all variation which can be explained by the usual covariates, other than geographic location In the rest of this paper, t_i is therefore dropped from the specification of the model.

u_i represents a component with significant spatial structure.

v_i represents unexplained variation

It is the component u_i that is of interest in an analysis of the spatial structure of the data

We must now formulate the conditional prior distribution of $x_i | \delta_i$, (2.2), in terms of u_i and v_i. Henceforth t_i is ignored since it has already been removed from the data. It is reasonable to assume that u_i and v_i are independent Also, since there are no reasons to use any other distribution, we shall use a normal prior distribution for $\{v_i : i = 1, \ldots, n\}$

$$p(v_i) \propto \lambda^{-1/2} \exp\left(- \frac{1}{2\lambda} v_i^2 \right) \tag{2.6}$$

We have assumed that the risk at the ith region depends only on regions which are in the neighbourhood of the ith region It is also assumed that the prior conditional density of the spatial component, u_i, can be factorized into components representing the dependencies on each of the neighbouring regions and hence can be written as

$$p_i(u_i | u_1, u_2, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n) \propto \exp\left(- \sum_{j \in \delta_i} \phi(u_i - u_j) \right) \tag{2.7}$$

for some function ϕ. Note that the summation in (2 7) is only over j in δ_i

The function ϕ must reflect the fact that the spatial dependence will reduce as the distance between the regions increases It must therefore favour similar values for regions which are adjacent, and can be any even function. It could be preceded by a factor to allow for the precise proximity of the regions i and j. In this case, (2 7) is replaced by

$$p_i(u_i | u_1, u_2, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n) \propto \exp\left(- \sum_{j \in \delta_i} w_{ij} \phi(u_i - u_j) \right) \tag{2 8}$$

Possible choices for ϕ include

$$\phi(z) = \frac{z^2}{2\varkappa} \qquad \text{and} \qquad \phi(z) = \frac{|z|}{\varkappa}\,.$$

In this paper, we use the first of these possibilities. Thus,

$$(2.9) \qquad p_i(u_i | u_1, u_2, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n) \propto \exp\left(-\frac{1}{2\varkappa} \sum_{j \in \delta_i} (u_i - u_j)^2 \right)$$

The two scale parameters $\varkappa$ and λ, which determine the variances of u and v must also be given a prior distribution. A suitable choice for this prior distribution, which is close to the usual uninformative distribution but which avoids technical difficulties is

$$(2.10) \qquad \text{prior}(\varkappa, \lambda) \propto \exp\left(-\frac{\varepsilon}{2\varkappa} - \frac{\varepsilon}{2\lambda} \right)$$

where ε is a small positive constant, say 0.01. For a more detailed discussion of this choice, see BESAG et al (1991)

The conditional prior distribution for $x_i | \delta_i$, (2.2), can now be replaced by the prior distributions of $\underline{u}$, $\underline{v}$, $\varkappa$ and λ The posterior density of the parameters can be found as in (2.4), using Bayes theorem:

$$(2.11) \qquad p(\underline{u}, \underline{v}, \varkappa, \lambda | \underline{y}) \propto \prod_{i=1}^{n} f(y_i | x_i) \varkappa^{-n_i/2} \times$$

$$\times \exp\left(-\frac{1}{2\varkappa} \sum_{j \in \delta_i} (u_i - u_j)^2 \right) \lambda^{-1/2} \exp\left(-\frac{1}{2\lambda} v_i^2 \right) \text{prior}(\varkappa, \lambda)$$

where n_i is the cardinality of δ_i.

Note that the joint prior distribution of $\underline{u}$ has been obtained from the conditional prior densities, (2.9), using the derivation given in Section 2 of BESAG (1974) Various forms for $f(y_i | x_i)$ are appropriate for insurance data. In the example in Section 4, we use a normal distribution For data on claim numbers a Poisson distribution would be appropriate In the case of Poisson data, it is usual to assume that the mean of this distribution is $c_i e^{x_i}$, where c_i is the expected number of claims in region i ignoring the spatial effect. Then

$$(2.12) \qquad f(y_i | x_i) = \frac{\exp(-c_i e^{x_i})(c_i e^{x_i})^{y_i}}{y_i!}$$

A normal distribution for $f(y_i | x_i)$ is also useful in practice. The mean and variance of this distribution will depend on the application, and an example of this case is given in Section 4.

3. THE GIBBS SAMPLER

Having defined the Bayesian model, the remaing problem is to obtain maximum *a posteriori* estimates for the parameters The complexity, high dimensionality and multimodality of the problem rules out any normal optimization routines. However, it is possible to set up a Markov chain whose stationary distribution is consistent with the posterior distribution. One approach which produces such a Markov chain is called the Gibbs sampler. The principle of the Gibbs sampler is as follows

At each step in the chain the current value of each parameter is replaced by a new one which is chosen randomly from its distribution given all the other parameter values and the observed data Thus, in the terminology of Section 2, a value for x_i is sampled at random from the density

$$p_i(x_i \mid \delta_i, \underline{y}) \tag{3 1}$$

The values of the risk parameters in all regions other than i, including in δ_i, are assumed fixed at their current values in this step. This step involves sampling from each of the distributions subsumed into x_i: i.e. for u_i, v_i, $\varkappa$ and λ. Initial values of the parameters must be supplied.

Typically, the chain must be allowed to run for 1,000 steps before it will have converged to its stationary distribution, which can be used to find the maximum *a posteriori* estimates for the parameters. Once convergence has been obtained, a sample of every 10th step over the next 10,000 steps usually provides a reasonable estimate of the stationary distribution. This can be treated as a an empirical distribution from which the required estimates can be obtained in the usual way.

Note that the conditional posterior distributions which are required by the Gibbs sampler can be obtained in a straightforward manner. For example,

$$p(u_i \mid u_{-i}, \underline{v}, \varkappa, \lambda, \underline{y}) \propto f(y_i \mid x_i)\, p(u_i \mid u_{-i}, \varkappa) \tag{3.2}$$

where u_{-i} denotes all values in $\underline{u}$ except u_i.

For example, when the data have Poisson distributions and the posterior density is given by (2.11) and (2.12), then the marginal posterior of u_i is given by

$$p(u_i \mid u_{-i}, \underline{v}, \varkappa, \lambda, \underline{y}) \propto \exp\left(-c_i e^{u_i + v_i} + u_i y_i - \frac{n_i}{2\varkappa}(u_i - \bar{u}_i)^2\right) \tag{3.3}$$

where $\bar{u}_i$ is the mean value of u_i over δ_i Details of the marginal distributions of the other parameters in the case of Poisson data can be found in BESAG et al. (1991)

Once the marginal densities have been found and initial values of all the parameters supplied, the Gibbs sampler can be used to generate values of the parameters from the required posterior distribution. In effect, the procedure exploits the simpler conditional distributions to simulate the posterior distribution.

In some cases the random sampling does not present any problems. For example, when the data are normally distributed, the posterior distributions are also normal and the sampling procedure described above is fairly straightforward In other cases, the posterior distributions are more complicated and samples cannot be obtained by a direct method. Instead, a method such as adaptive rejection sampling must be used. It is very important that the sampling procedure and the computational approach are highly efficient in order to produce results reasonably fast. A particularly efficient from of rejection sampling is described by GILKS and WILD (1992). This form of sampling has to be used, for example, in the case of Poisson data. We now summarise the sampling process as described in greater detail in GILKS and WILD (1992).

Suppose a sample is required from the distribution whose density function is $f(x)$. For example, this density might take the form given in (3.3). The density, $f(x)$, need only be known up to a constant of integration. i.e. instead of knowing $f(x)$, we may only know $g(x)$ where

(3.4) $$g(x) = cf(x)$$

and c is an unknown constant.

It is necessary to define an envelope function $g_u(x)$ such that $g_u(x) \geq g(x)$ for all x, and a squeezing function $g_l(x)$ such that $g_l(x) \leq g(x)$ for all x. The procedure to obtain a sample from $f(x)$ is then as follows

Take a sample x^* from $g_u(x)$ and a sample w from $U(0, 1)$. Now use the squeezing function to test the value

if $w \leq \dfrac{g_l(x^*)}{g_u(x^*)}$ then accept x^*; if not then test

if $w \leq \dfrac{g(x^*)}{g_u(x^*)}$ then accept x^*; otherwise reject x^*

FIGURE 1

After each rejection of a sample value, the envelope and squeezing functions are redefined so as to reduce the probability of further rejection If the log density, $h(x) = \log(g(x))$ is considered, it can be seen that for the density (3.3), and for many others, $h''(x) < 0, \forall x$. It is therefore possible to define an envelope $h_u(x) = \log(g_u(x))$ where $h_u(x)$ is a piecewise linear function such that each line segment is a tangent to $h(x)$. Similarly, a piecewise linear function

$h_l(x) = \log(g_l(x))$ can be defined by chords meeting $h(x)$ at the same points as $h_u(x)$.

After each rejection of a value of x^*, this value is added to the set of points at which $h_l(x)$ and $h_u(x)$ meet $h(x)$ GILKS and WILD (1992) show that this provides an efficient method of generating samples for the Gibbs sampler

4. EXAMPLE

In order to illustrate the methods and to give an indication of the nature of the results, the data from TAYLOR (1989) are reanalysed in this section We would emphasise that this is really an illustration and does not represent a definitive rating conclusion. In particular, we would prefer to analyse claim numbers and claim amounts separately · see Section 5 for a more detailed discussion However, this example does enable the results to be compared with the method used by Taylor, which imposed a much greater degree of smoothness onto the results

The data relates to Household Contents Insurance in and around Sydney, Australia This region is divided into approximately 200 postcode areas The data have already been processed to remove the effects of all factors which can be modelled using generalised linear modelling techniques. All factors corresponding to t_i in (2.5) have been controlled out in order to make the data suitable for investigating the spatial effects Taylor also included a "rough fit of the 'geographic area effect'" in order to improve the fit of the other factors but this effect was, of course, not controlled out The final data used in this example consists of adjusted loss ratios.

The adjusted loss ratios are assumed to be normally distributed:

$$y_i | x_i \sim N\left(x_i, \frac{\alpha}{e_i}\right)$$

where e_i is the earned exposure in postcode area i,

and α is a constant, chosen as indicated below.

As noted in TALOR (1989), this normal approximation may be poor where e_i is small. However, in the model considered here, areas with low values of e_i will have a limited effect on the overall results. The constant α controls the amount of smoothing applied, as can be seen from the following maps. The maps show the values of the adjusted loss ratios divided into six bandes as follows·

A	Less than 0.5
B	0.5 to 0 7
C	0.7 to 0 9
D	0.9 to 1 1
E	1 1 to 1.3
F	Greater than 1.3

Map 1 shows the adjusted loss ratios of the raw data before the fitting of the spatial model. Maps 2 to 5 show those of the fits for various values of α. A value of 100 appears to be produce a similar level of smoothing to that achieved by Taylor and the overall pattern is very similar with areas of low risk in the south-east and north-east corners and a band of high risk just south of the river.

A referee has pointed out that a value of α of around 100 can be justified as follows

$$\alpha = \text{variance of loss ratio for a single risk}$$

If it is assumed that losses occur according to a Poisson process with rate θ and that the first and second moments of the distribution of the size of individual losses are μ_1 and μ_2, then

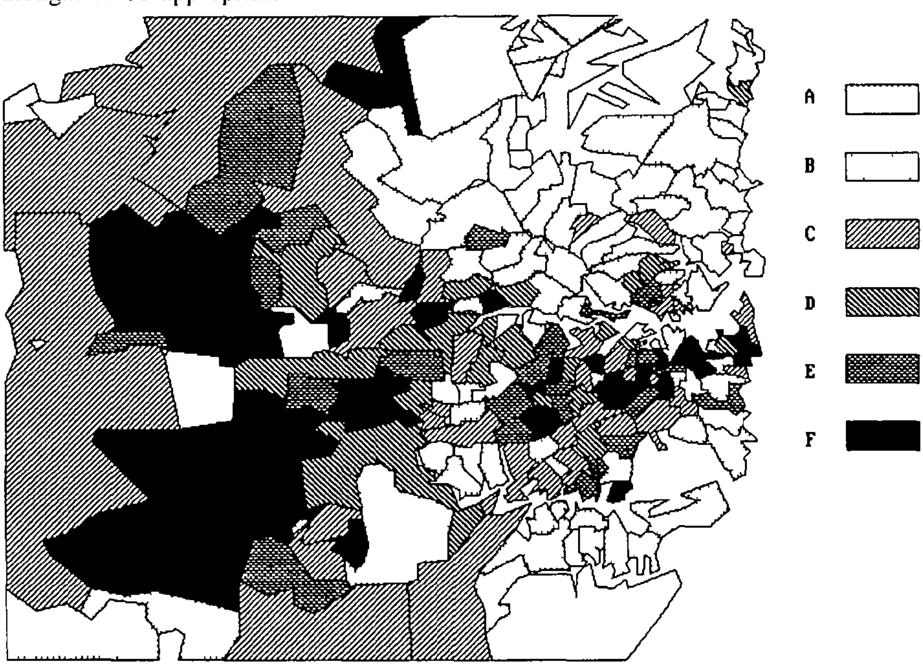
$$\alpha \simeq \frac{\theta\mu_2}{(\theta\mu_1)^2}$$

or

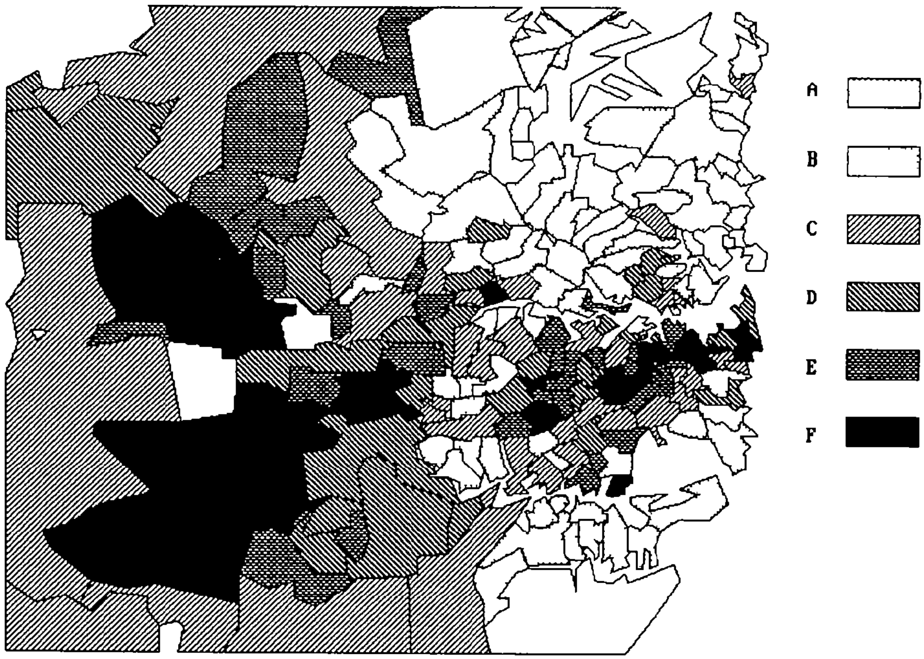
$$\alpha \simeq \frac{(1+r)^2}{\theta}$$

where r = coefficient of variance of claim size.

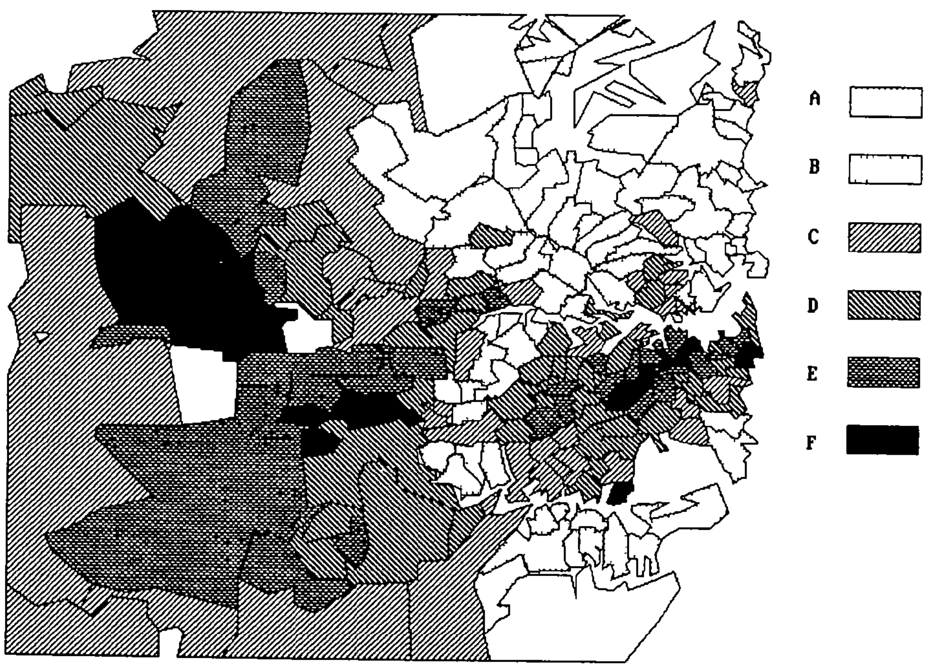
From the data the observed value of θ is approximately 0 1, so that $\alpha = 100$ corresponds to a value of r of 3 which seems reasonable. However, the choice of value for α should be a pragmatic one based on the level of smoothing which is thought to be appropriate



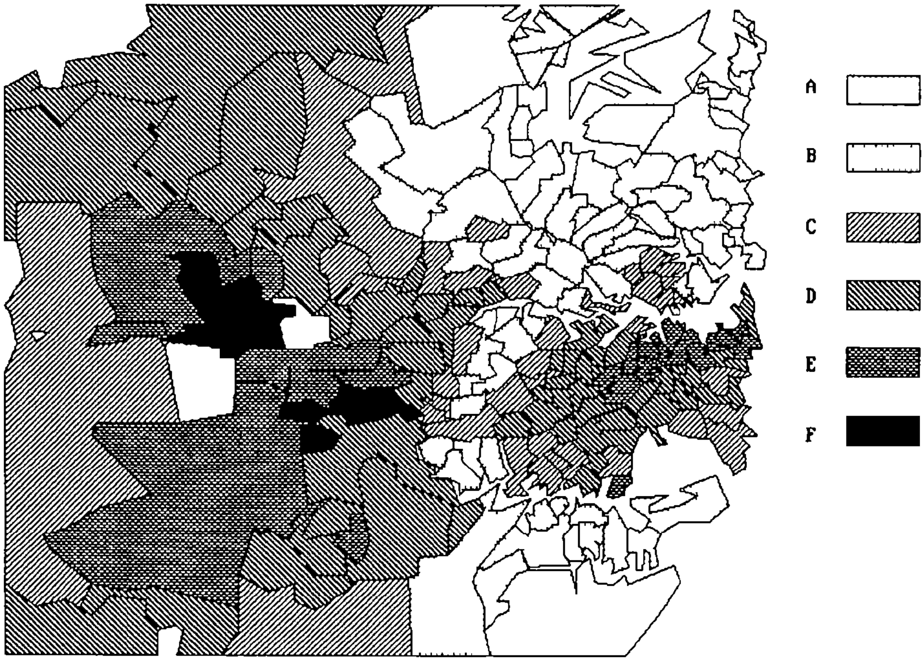
MAP 1 Raw Data



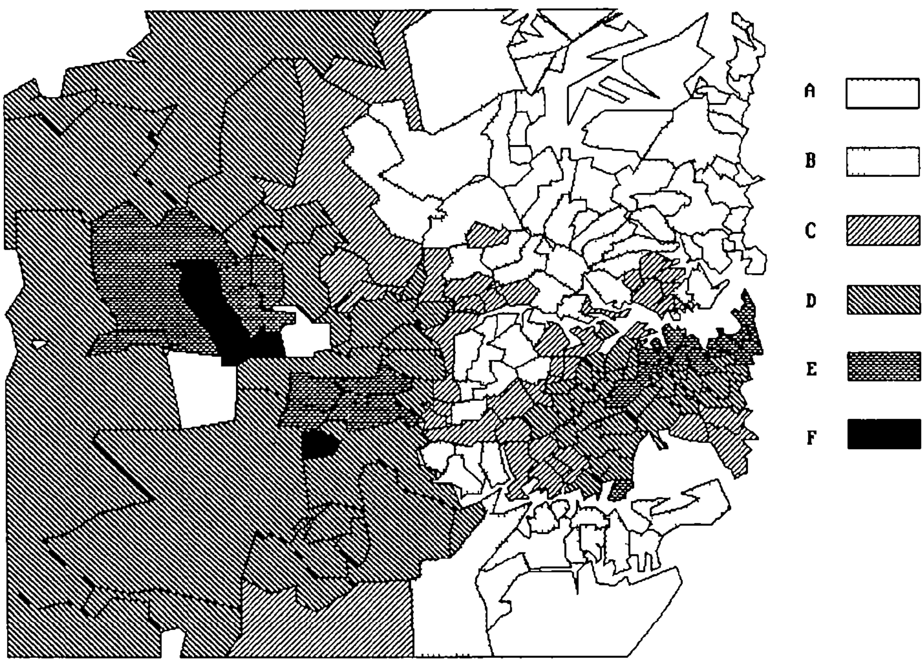
MAP 2 Smoothed values $\alpha = 20$



MAP 3 Smoothed values $\alpha = 50$



MAP 4 Smoothed values $\alpha = 100$



MAP 5 Smoothed values $\alpha = 200$

5 CONCLUSIONS

This paper has described how spatial statistical models can be used to analyse the geographic area effect in insurance data. The methods are applicable to all data in which there is a geographic area effect. The authors believe that the potential for these methods in insurance applications is great, and that they represent the best way of premium rating by postcode area.

The example has been approached from the point of view of smoothing the data over the postcode areas, using a continuous scale for the rating results These smoothed results have then been divided into bands for rating purposes An alternative approach would be to use a discrete scale for the results, corresponding to the required number of rating classes The spatial model would then be required to allocate each postcode area to one of the rating classes The use of this type of model is at present under investigation.

Unlike the method used in TAYLOR (1989) this method could easily be extended to an entire country rather than just one metropolitan area

It would be preferable to analyse the data for claim numbers and claim amounts separately This approach is already used to model claims experience with respect to other factors c.f RENSHAW (1993) Such a separation is particularly important in cases where claim severity has a long tailed distribution (e g. liability) where one large claim could dominate the loss ratio of a small area It may also prove to be the case that a simpler model using only a few of the factors is appropriate for claim severity while a more complicated model including spatial data can be used for the frequency. This involves more complex computations since the data would no longer be normally distributed. Again, this is under investigation and will be the subject of a subsequent paper.

ACKNOWLEDGEMENTS

We would like to thank GREG TAYLOR for supplying the data used in Section 4

The first named author is financially supported by Guardian Insurance

REFERENCES

BESAG, J (1974) Spatial Interaction and the Statistical Analysis of Lattice Systems (with discussion) *J Royal Statist Soc, Series B*, **Vol. 36**, No 2

BESAG, J, YORK, J and MOLLIE, A (1991) Bayesian Image Restoration, with Applications in Spatial Statistics (with discussion) *Ann Inst Statist Math*, **Vol. 43**, 1–59

CRESSIE, N (1991) *Statistics for Spatial Data* John Wiley and Sons, New York

GEMAN, S and GEMAN, D (1984) Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images *I E E E Trans Pattn Anal Mach Intell*, **Vol. 6**, 721–741

GILKS, W R, CLAYTON, D G, SPIEGELHALTER, D J, BEST, N G MCNEIL, A J, SHARPLES, L D and KIRBY, A J (1993) Modelling Complexity Applications of Gibbs Sampling in Medicine *J Royal Statist Soc, Series B*, **Vol. 55**, No 1

GILKS, W R and WILD, P (1992) Adaptive Rejection Sampling for Gibbs Sampling *Appl Statist*, **Vol. 41**, 337–348

RENSHAW, A E (1993) *Modelling the Claims Process in the Presence of Covariates* ASTIN Colloquium

SMITH, A F M and ROBERTS, G O (1993) Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods *J Royal Statist Soc, Series B*, **Vol. 55**, No 1

TAYLOR, G C (1989) Use of Spline Functions for Premium Rating by Geographic Area *ASTIN Bulletin* **Vol. 19**, No 1, 91–122

MICHAEL BOSKOV
Department of Actuarial Science and Statistics, The City University, Northampton Square, London EC1V 0HB, England.

SHORT CONTRIBUTIONS

MARTINGALES AND TAIL PROBABILITIES

BY HANS U. GERBER

At the twenty-eighth Actuarial Research Conference of the Society of Actuaries, WILLMOT and LIN (1993) presented a paper whose central result is a bound on the tail probability of a random sum. In the subsequent discussion, Professor Bühlmann raised the question, if this bound could be derived by martingale methods. The purpose of this note is to show how it can be done

We consider a random variable of the form

$$S = X_1 + \dots + X_N .$$

Here the random variables $N, X_1, X_2, \dots$ are independent, and the X_k's are assumed to be positive and identically distributed; their common distribution function is denoted by $F(x)$.

Let

$$p_k = \Pr(N = k), \qquad k = 0, 1, \dots$$

We assume the existence of a number ϕ, $0 < \phi < 1$, with

$$\Pr(N > k \mid N \geq k) \leq \phi \qquad \text{for} \qquad k = 1, 2, \dots \tag{1}$$

and a positive number r with

$$\phi \cdot \int_0^\infty e^{rx}\, dF(x) \leq 1 \tag{2}$$

(if $F(x)$ is sufficiently regular, we might choose the value of r for which equality holds). Then the result of Willmot and Lin is that

$$\Pr(S \geq x) \leq \frac{1 - p_0}{\phi} \cdot e^{-rx}$$

for any $x > 0$.

For the following proof we introduce

$$S_k = X_1 + \dots + X_k$$

and

$$Y_k = \begin{cases} e^{rS_k} & \text{if} \quad N \geq k \\ 0 & \text{if} \quad N < k . \end{cases}$$

We note the recursive relationship

$$Y_k = Z_k \cdot Y_{k-1}, \qquad k = 1, 2, \ldots$$

with

$$Z_k = \begin{cases} e^{rX_k} & \text{if} \quad N \geq k \\ 0 & \text{if} \quad N < k. \end{cases}$$

According to (1) and (2), the conditional expectation of Z_{k+1} (given $N \geq k$) is less than or equal to 1, which shows that the sequence $Y_1, Y_2, \ldots$ is a supermartingale

If we stop it at time

$$T = \min\{k : S_k \geq x \quad \text{or} \quad N < k\}$$

it follows that, given $N \geq 1$ and X_1,

$$Y_1 \geq E[Y_T | N \geq 1, X_1]$$

or

$$e^{rX_1} \geq E[e^{rS_T} 1_{\{S \geq x\}} | N \geq 1, X_1] \geq e^{rx} \Pr(S \geq x | N \geq 1, X_1).$$

Then we get

$$\begin{aligned} \Pr(S \geq x) &= (1-p_0) \cdot E[\Pr(S \geq x | N \geq 1, X_1)] \\ &\leq (1-p_0) \cdot E[e^{rX_1} \cdot e^{-rx}] \\ &\leq \frac{1-p_0}{\phi} e^{-rx}, \end{aligned}$$

which completes the proof.

REFERENCE

WILLMOT, G E and LIN X (1993) *Lundberg bounds on the tails of compound distributions* Research Report 93-14, Institute of Insurance and Pension Research University of Waterloo Forthcoming in the Journal of Applied Probability

HANS U. GERBER
Ecole des HEC, University of Lausanne, CH-1015 Lausanne.

FACULTY POSITION IN ACTUARIAL SCIENCE

The School of Actuarial Science invites applications for two tenure-track professoral appointments starting August 1994 The School is reponsible for teaching and research in actuarial science and the two new professors will join the seven faculty members currently in function.

The duties of a professor include teaching to undergraduate and graduate students, conducting active research in actuarial science, counseling students, supervising graduate students, and participating to the academic responsibilities of the School.

For the first available position, the qualifications requested of a candidate are

Hold (or be near completion of) a Ph D. preferably in actuarial science or related area, be well engaged in actuarial research, be a member of (or a candidate for membership in) a recognized association of actuaries.

For the second vacant post, is also admissible

A Fellow of any recognized association of Actuaries or the equivalent who can prove research production in actuarial science and capacity for supervising graduate students.

Applications must be sent in writing to the following address with a recent curriculum vitae enclosed.

Mr André Prémont, director
School of Actuarial Science
Alexandre-Vachon Building
Laval University
Sainte-Foy (Québec)
G1K 7P4

Laval University applies an equal opportunity program and dedicates half of its openings to women applicants In accordance with Canadian immigration requirements, this advertisement is directed, in the first instance, to Canadian citizens and permanent residents.

GUIDELINES TO AUTHORS

1. Papers for publication should be sent in quadruplicate to one of the Editors

   Hans Buhlmann,
   Mathematik, ETH-Zentrum,
   CH-8092 Zurich, Switzerland

   D Harry Reid,
   Eagle Star Insurance Company Ltd,
   The Grange, Bishop's Cleeve
   Cheltenham Glos GL52 4XX, United Kingdom

   or to one of the Co-Editors

   Alois Gisler,
   "Winterthur" Swiss Insurance Company,
   P O Box 357, CH-8401 Winterthur, Switzerland

   David Wilkie
   Messrs R Watson & Sons
   Watson House, London Rd, Reigate, Surrey RH2 9PQ, United Kingdom

   Submission of a paper is held to imply that it contains original unpublished work and is not being submitted for publication elsewhere
   Receipt of the paper will be confirmed and followed by a refereeing process, which will take about three months

2. Manuscripts should be typewritten on one side of the paper, double-spaced with wide margins The basic elements of the journal's style have been agreed by the Editors and Publishers and should be clear from checking a recent issue of *ASTIN BULLETIN* If variations are felt necessary they should be clearly indicated on the manuscript

3. Papers should be written in English or in French Authors intending to submit longer papers (e g exceeding 30 pages) are advised to consider splitting their contribution into two or more shorter contributions

4. The first page of each paper should start with the title, the name(s) of the author(s), and an abstract of the paper as well as some major keywords An institutional affiliation can be placed between the name(s) of the author(s) and the abstract

5. Footnotes should be avoided as far as possible

6. Upon acceptance of a paper, any figures should be drawn in black ink on white paper in a form suitable for photographic reproduction with lettering of uniform size and sufficiently large to be legible when reduced to the final size

7. References should be arranged alphabetically, and for the same author chronologically Use a, b, c, etc to separate publications of the same author in the same year For journal references give author(s), year, title, journal (in italics, cf point 9), volume (in boldface, cf point 9), and pages For book references give author(s), year, title (in italics), publisher, and city

   Examples

   BARLOW, R E and PROSCHAN, F (1975) *Mathematical Theory of Reliability and Life Testing* Holt, Rinehart, and Winston, New York
   JEWELL, W S (1975a) Model variations in credibility theory In *Credibility Theory and Applications* (ed P M KAHN), pp 193–244, Academic Press, New York
   JEWELL, W S (1975b) Regularity conditions for exact credibility *ASTIN Bulletin* **8**, 336–341

   References in the text are given by the author's name followed by the year of publication (and possibly a letter) in parentheses

8. The address of at least one of the authors should be typed following the references

Continued overleaf

COMMITTEE OF ASTIN

Guidelines to Authors *continued from inside back cover*

9. Italics (boldface) should be indicated by single (wavy) underlining. Mathematical symbols will automatically be set in italics, and need not be underlined unless there is a possibility of misinterpretation.
Information helping to avoid misinterpretation may be listed on a separate sheet entitled 'special instructions to the printer'. (Example of such an instruction: Greek letters are indicated with green and script letters with brown underlining, using double underlining for capitals and single underlining for lower case.)

10. Electronic Typesetting using Word Perfect 5.1 is available. Authors who wish to use this possibility should ask one of the editors for detailed instructions.

11. Authors will receive from the publisher two sets of page proofs together with the manuscript. One corrected set of proofs plus the manuscript should be returned to the publisher within one week. Authors may be charged for alterations to the original manuscript.

12. Authors will receive 50 offprints free of charge. Additional offprints may be ordered when returning corrected proofs. A scale of charges will be enclosed when the proofs are sent out.