

# ASTIN BULLETIN

A Journal of the International Actuarial Association

## CONTENTS

### EDITORS:

**Hans Bühlmann**  
Switzerland

**D. Harry Reid**  
United Kingdom

### CO-EDITORS:

**Alois Gisler**  
Switzerland

**David Wilkie**  
United Kingdom

### EDITORIAL BOARD:

**Björn Ajne**  
Sweden

**Marc Goovaerts**  
Belgium

**Jacques Janssen**  
Belgium

**William S. Jewell**  
USA

**Jean Lemaire**  
Belgium/USA

**Walther Neuhaus**  
Norway

**Jukka Rantala**  
Finland

**Axel Reich**  
Germany

**James A. Tilley**  
USA

### EDITORIAL AND ANNOUNCEMENTS

Editorial	1
The XXII ASTIN Colloquium Montreux	5

### INVITED PAPER

<b>J. LEMAIRE</b> Cooperative Game Theory and its Insurance Applications	17
---	----

### ARTICLES

<b>M. DE LOURDES CENTENO, O. SIMÕES</b> Combining Quota-Share and Excess of Loss Treaties on the Reinsurance of $n$ Independent Risks	41
<b>H. RAMLAU-HANSEN</b> Distribution of Surplus in Life Insurance	57

### WORKSHOP

<b>J. LEDOLTER, S. KLUGMAN, C. S. LEE</b> Credibility Models with Time-Varying Trend Components	73
<b>Th. MACK</b> A Simple Parametric Model for Rating Automobile Insurance or Estimating IBNR Claims Reserves	93
<b>R. SCHNIEPER</b> Separating True IBNR and IBNER Claims	111

### DISCUSSION PAPERS

<b>P. BROCKETT, M. GOOVAERTS, G. TAYLOR</b> The Schmitter Problem	129
<b>R. KAAS</b> The Schmitter Problem and a Related Problem A Partial Solution	133

### SHORT CONTRIBUTIONS

<b>C. M. RAMSAY</b> A Note of the Normal Power Approximation	147
Book Review	151
Letter to the Editors	153
Actuarial Vacancy	155

## **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* will publish 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.

*ASTIN BULLETIN* appears twice a year, each issue consisting of about 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 1700.

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

### EDITORIAL

#### *Whither AFIR?*

In an editorial in *ASTIN Bulletin* 17.2 in November 1987, Hans Bühlmann introduced Actuaries of the Third Kind. In *ASTIN Bulletin* 19.1 in April 1989, François Delavenne, Chairman of the newly formed AFIR section, described the formation and objectives of that section. These editorials, together with articles by Müller, Schweizer & Föllmer and Dhaene appeared in a special issue of *ASTIN Bulletin* (19S) in November 1989.

Since then the 1st AFIR International Colloquium has taken place in Paris in April 1990, and by the time this editorial is being read, the 2nd AFIR International Colloquium will have taken place in Brighton in April 1991. Sixty-four different papers were presented to the Paris Colloquium, and 82 will have been presented in Brighton. Some of these, and other articles, may make their way into the pages of *ASTIN Bulletin*.

What can we say so far about the way AFIR has developed? Most of the papers are derived from the general field of modern financial economics. They are based on statistical and mathematical approaches to the investment of institutional assets, and can clearly be differentiated from the many articles that are of interest to investment analysts around the world, dealing with the fortunes of particular companies, industries or national economies, and with the immediate prospects for share prices, interest rates or exchange rates.

Although many of the AFIR papers are of interest to financial economists generally, many also are of particular relevance to those actuaries concerned with insurance companies, pension funds and similar institutions that have non-tradable liabilities. It is in this area that AFIR can make its own special contribution.

Many of the papers have been primarily descriptive — what sort of model best describes a particular market, or does a particular market behave in line with some theoretically derived hypothesis? Others are prescriptive — how can theoretical ideas contribute to designing an optimal asset allocation strategy, or to appropriate methods for the calculation of premiums or the valuation of liabilities?

One can also classify the papers in a different way: do they relate to a general investment topic; to the asset side of a financial institution (asset-liability matching); or to the liability side of a financial institution? For actuaries in general, this last theme is perhaps the most interesting. The paper by Cummins in *ASTIN Bulletin* 20.2 describes some uses of the theoretical models of financial economics to justify particular methods of setting premiums, in a context in which premium rates need to be justified to a State Insurance Commissioner.

A similar application is the use of option pricing theory to calculate the values of pensions or annuities which increase in line with a consumer price index, subject to some upper limit, another is in the pricing of the guarantees inherent in a with profits life assurance as compared with a unit linked policy.

The two main themes for papers for the 2nd AFIR International Colloquium are: asset-liability matching; and interest rate models. A third theme that runs through a number of papers is the application of a stochastic model for investments other than the pure random walk model, in particular what has become known in Britain as the 'Wilkie investment model'.

I should like to suggest a number of areas towards which those interested in AFIR could apply their ingenuity. The first is asset-liability matching models, approached either through fixed interest matching, which requires an analysis of interest rate models and 'duration' measures, or through application of the portfolio selection approach. Although much has been done in this area, there remain many unsolved problems. What are appropriate asset allocation models which take account of liabilities emerging over the very many years that insurance companies and pension funds work in? What are the optimisation objectives of such an institution? It is not simply a matter of maximising terminal wealth or surplus at the end of a long period, since varying bonus rates on life policies or varying contribution rates for pension funds during the course of the period come into consideration. How does one allow for dynamic decision-making with such a long time horizon?

Further problems are: how does one allow for the existing assets held, and the potential costs of changing them? And how does one allow for the uncertainty that must exist in one's estimates of the probability distributions of returns on particular assets? One possible optimum portfolio is usually 100% in the asset that seems to promise the highest expected return, regardless of variability; but if there is uncertainty about one's estimate of that expected return, it is not advantageous to incur expense in pursuit of an uncertain marginal benefit, even for the risk-neutral investor.

In order to implement any asset allocation method, one must have some sort of model of the distribution of returns on the classes of investment under consideration. The statistical investigation of historic investment series seems to me to be the next major undertaking for members of AFIR. Very many investigations by financial economists have concentrated on the short term, gathering data at daily or weekly intervals for a small number of years. They have generally found that a random walk model of some kind fits the data reasonably well. Few investigations have considered the behaviour of such a series over long numbers of years, but those that have done so have generally discovered that the random walk model is an unsatisfactory description over the long run, and that a model that includes some sort of reversion of interest rates to a mean level or of share dividend yields or Price/Earnings ratios to a mean level is more satisfactory.

More gathering of long runs of data from a variety of different countries, and more statistical investigation of such data needs to be done. In an

international field, one would also like to see how exchange rates have behaved: randomly in the short run, and according to purchasing power parity in some way in the long run seems a plausible first hypothesis.

Consequential on these first two themes is what sort of equilibrium model results from an international economy, with multiple currencies, in which investors from different countries have different types of liabilities and different possible objectives. The classical Capital Asset Pricing Model (CAPM) assumes that all investors work in one currency (such as US dollars) and measure their utility as functions of wealth in that currency. But many international investors measure their wealth in Swiss francs, German marks or British pounds, and others measure in real terms (after allowing for price inflation) rather than in currency at all. What consequences does such a more elaborate structure have for the CAPM?

The final field of research I should like to propose to AFIR members relates to the liability side, building on the work of Cummins and others in relation to premium-rating, on the use of option pricing methods for valuing liabilities with inherent options included—anything of the form which pays the greater of A and B, or the lesser of A and B, includes an implied option—and the application of the methods of financial economists, whether through the CAPM or otherwise, to the question of the appropriate rate of return for discounting risky liabilities. This is of importance in the valuation of an insurance company where a realistic rather than a prudent valuation is required, for example for profit testing, estimating the value of a company for purchase or sale, or the consolidation of the accounts of insurance subsidiaries in a parent company which is not an insurance company.

All this sounds like plenty of work for the future. It is almost too late to produce a new paper for the 1992 International Congress of Actuaries, but it is hoped that there will be an AFIR International Colloquium in 1993 (location still to be decided), and the pages of the ASTIN Bulletin are available for those who would like a widespread and thoughtful international readership. Your offerings addressed to me or to one of the other editors please.

DAVID WILKIE



XXII ASTIN COLLOQUIUM  
MONTREUX, SWITZERLAND, 9th to 13th SEPTEMBER 1990

The Swiss organisers of the 22nd ASTIN Colloquium, mindful of the need to bring theory and practice closer together, arranged for the meetings to be held in the Casino in Montreux. Judging by the number of actuaries who at the end of the Colloquium departed for Geneva in second-class carriages, there is scope for further progress to be made.

With such an attractive setting as Montreux it was scarcely surprising that the attendance reached a new record level, with 256 actuaries from 23 countries and all five continents represented. The traditional ASTIN conviviality was well under way by the end of the reception with which we were welcomed on the Sunday evening, and anyone who did not make new friends during the days that followed can scarcely have been trying.

The business meetings began, naturally enough, with the opening ceremony, the highlight of which was an invited lecture by Peter Gmeiner, the First Secretary of the Swiss Insurance Association, on "The future European insurance market and the Swiss insurance industry". Part way through the Colloquium there was a second invited lecture by James W MacGinnitie on "Actuarial ethics and integrity". In view of the wide general interest of these lectures, which were not available in printed form, summaries of the lectures are appended to these notes.

### **Topic 1: Models of Finance**

Uncertainties abound in the world of investment, and most actuaries need to be concerned with financial risk whatever their field of work. Not surprising indeed, that ASTIN now has a sister group, AFIR, formed to consider financial risk. There is clearly an overlap between the two groups, as regards areas of interest. It was remarked by Philippe Maeder who, with Jean-Pierre Melchner, had prepared the summary of the papers under Topic 1, that there was scope for co-ordination between the two groups regarding topics for papers.

The four papers presented on Topic 1 confirmed the scope for applying models of finance to diverse areas of actuarial work. Philippe ARTZNER and Freddy DELBAEN consider credit insurance, and discuss the optimal time at which a borrower with default risk should prepay a risky fixed rate loan. Werner HURLIMANN considers the concept of a premium to cover the investment risk in life insurance. David SANDERS discusses a possible use for option pricing in the premium rating of stop loss and excess of loss reinsurance. Patrick BROCKETT and Yehuda KAHANE consider how a rational investor may choose between two investment opportunities.

**Topic 2: Experience rating**

Twelve papers were presented on this topic, including one paper transferred from Topic 3, and Alois Gisler presented a summary of them which he and René Schmieper had prepared.

The first group of papers is related to the determination of the pure risk premium and to the assessment of claims reserves. Alois GISLER and Peter REINHARD suggest that the problem of outliers in rating is best dealt with by a combination of credibility and robust statistics. GABRY et al. are also faced with outliers in a large volume of Dutch industrial fire insurance data which they are using to derive a set of risk premium rates. They use a pragmatic approach, applying a combination of top-slicing and credibility techniques. Erhard KREMER shows how to determine the necessary coefficients to make practical use of the exponential smoothing credibility estimator which he puts forward as an alternative to the credibility estimator with geometric weights. Ragnar NORBERG considers linear predictors and credibility estimators based on a continuous time model rather than a finite set of observations.

The final paper in the first group, by Thomas MACK, is alone in being unconnected with credibility theory. MACK reveals that the estimation of IBNR claims reserves is a special case of the analysis of cross classified data. He shows that, for example, the method of marginal totals for cross classified data leads to the chain ladder method for assessing reserves. The author advocates the use of an alternative model for the total claim amount, for both rating and reserving, based on the Gamma distribution.

The second group of papers relates to bonus-malus systems. Jean-Luc BESSON and Christian PARTRAT advocate the use of the Poisson-Gamma model for claim frequencies in motor insurance. They use a goodness-of-fit test to illustrate the superiority of this model, although Chresten Dengsoe suggested in the discussion that the test statistics put undue emphasis on the small number of policies with four or five claims. Hans GERBER explains the recent change in the bonus-malus system used in Switzerland. The new system imposes an increased penalty following a claim and is thereby an improved discriminator between low and high risks. Tormod Sande pointed out in the discussion that, even under the new system, high risk policies continue to pay on average substantially less than their share of premiums in the long term. This feature is common to all bonus-malus systems.

The third group of papers is devoted to the pricing of non-proportional reinsurance covers. Gunnar BENKTANDER advocates the use of a simple model to determine the extent to which the reinsurance risk premium for excess of loss cover in fire insurance is affected by varying the retention.

There are clearly immense practical difficulties in rating stop loss reinsurance cover. There is always the potential for over-generous claim settlement by the cedant at the expense of the reinsurer, especially if liability claims are covered. Reinsurers need all the help they can get if they are to make this form of cover available at affordable rates. In this regard, the two papers on stop loss cover are to be welcomed. Jozef TEUGELS and Bjorn SUNDT describe a scheme of



stop loss rating for motor fleets which takes account of the claims experience of the individual fleet. Lionel MOREAU also considers the rating of stop loss cover for motor fleets. He uses data from a large company over a five year period and, despite some shortcomings of the data, obtains a set of numerical results.

Reinstatement premiums are a common feature of non-proportional reinsurance cover, but there is little in the actuarial literature on the mathematical treatment of such premiums. Bjorn SUNDT discusses the pure premium and the loading needed for excess of loss cover with reinstatements. Reinstatement premium is a form of claims-dependent premium, and Stefan BERNEGGER considers the variance loading for excess of loss cover taking into account the influence of claims-dependent premiums. Since this paper addresses very much the same problems as that by SUNDT, it was presented under Topic 2 even though it was originally allocated to Topic 3.

### **Topic 3: Numerical methods**

Fifteen papers were presented on this topic and Erwin Straub presented a summary of them which he and André Dubey had prepared.

The first group of papers considers ruin probability and applications. Marc-Henri AMSLER uses the probability of ruin, the severity of ruin and the time of ruin in assessing the riskiness of an insurance portfolio. Examples are given relating to life assurance, and the results show the influence of different reinsurance programmes on the financial stability of the portfolio. François DUFRESNE, Hans GERBER and Elias SHIU show how classical risk theory, and in particular ruin theory, can be adapted when the gamma model is used to represent the aggregate claims process. Lourdes CENTENO provides an algorithm to calculate an optimum excess of loss retention, given certain assumptions regarding the calculation of the reinsurance premium. David DICKSON and Howard WATERS give an algorithm for approximating the finite time non-ruin probabilities for the classical risk model. The authors show that the algorithm can also be used to calculate infinite time non-ruin probabilities, and they address certain problems of numerical instability. Hans SCHMITTER derives an explicit expression for the ultimate ruin probability when the claim amount distribution is discrete with a finite number of steps.

The second group of papers considers the aggregate claims distribution. Marc GOOVAERTS and Robert KAAS give a recursive algorithm, using Panjer's formula, to compute the distribution function of a compound sum of claim numbers, when the number of summands follows a generalised Poisson distribution. Werner HURLIMANN proposes an approximation of the aggregate claims distribution by approximating the claim size distribution using the algebraic moment method. Thomas MULLER treats compound Poisson processes, their Panjer recursion and the effect of merging two or more portfolios. Some properties of compound Poisson processes are shown to be basic properties of the exponential power series. Bob ALTING VON GEUSAU proposes

a method to test the possibility of a trend over time in given data. In the Poisson case the distribution function of the proposed statistic can be calculated by means of the shovelboard approach, i.e. by making use of the fact that Poisson distributed variables, given their sum, are multinomially distributed. To aid our understanding, the author's presentation at the meeting included the display of a picture of a shovelboard, which is the basis of a well-known family pastime in the Netherlands. Erhard KREMER uses Fourier analysis to deal with the computation of the distribution function of total claims amounts where the ordered claims have been multiplied by given coefficients

The third group of papers relates to claims reserves. Teivo PENTIKAINEN and Jukka RANTALA analyse the three basic types of inaccuracies inherent in the estimation of claims reserves, model errors, parameter errors and stochastic errors. The authors simulate a claims process and analyse various estimation methods with regard to their sensitivity in respect of the three basic types of errors. Erwin KUMMERLI applies two formulae proposed by De Vylder and Kahane to run-off triangles for each of several classes of non-life business in a medium-sized company, and comments on the results. Hans EKHULT presents a program to calculate claims reserves in disability insurance as expected present values of future annuity payments.

The two remaining papers could not be allocated to any of the above three groups. Bruno KOLLER discusses spreadsheet programming languages and then shows how to use a spreadsheet to carry out Bayesian graduation, using an example from health insurance. Erhard KREMER applies the Cauchy-Schwarz inequality and derives an upper bound for the variance of the claims amount covered by stop-loss reinsurance.

During one of the working sessions on Topic 3 there was an impromptu debate on whether models or, alternatively, the observed data would normally provide the better indicator of future experience. Conflicting—and entertaining—views were expressed. The issue was finally clarified by Hans Bühlmann's comments that neither models nor data of the past will normally be in accord with the future experience, but that a model is constructed to try to reflect one's perception of what the future will hold.

### Speakers' Corner

Speakers' Corner is a well-established feature of ASTIN Colloquia, and provides an opportunity for members to make a contribution on the topic of their choice without the constraint of submitting a paper several months before the time of the colloquium.

Three of the papers in Speakers' Corner considered the probability of ruin and made the assumption of an underlying compound Poisson process. Richard VERRALL derives a sample re-use estimate of the probability of ruin, making use of the full bootstrap distribution and a saddlepoint approximation. Angela van HEERWAARDEN and Robert KAAS consider the concept of stop-loss order and develop a proof from which can be shown that the risk with higher

stop-loss premiums generates a higher ruin probability. Anna STEENACKERS and Marc GOOVAERTS obtain upper and lower bounds for stop-loss premiums and for ruin probabilities where certain features of the claim severity function are known.

Menachem BERG develops procedures for detecting possible trends in time non-homogeneous claim occurrence processes. Use is made of Bayesian revision procedures, and results for claim occurrence and claim size processes are combined to predict the total claim process. Udi MAKOV presents a sampling-resampling technique to assess the posterior distribution of a Bayesian credibility model for arbitrary likelihood function and prior distribution. It is explained that thereby the computational difficulties of evaluating integrals are overcome. Benedetto CONTI and Felix LAUHLI consider two classes of distribution functions which are regarded as important in non-proportional reinsurance work. Properties of these classes are set out and results are given following an analysis of the maximum likelihood estimator.

Bill JEWELL presents the third act of what has been described as a three-act play. The author advocates the formulation of the IBNR problem in continuous time and using a Bayesian approach. The paper points to the possibility of the working actuary of the future being able to predict *distributions* of numbers and amounts of IBNR claims.

Arthur RENSHAW shows how the existing range of actuarial graduation techniques can be considerably extended using generalised linear models. There is detailed discussion of how such models can be used to graduate the probabilities of death and the force of mortality.

Georg HARBITZ summarises the discussions which have taken place recently in Norway leading to the making of government regulations requiring appointed actuaries in general insurance companies as well as in life insurance companies. The detailed regulations are given by way of Appendix. These developments in Norway will be of interest in other countries where some statutory role for non-life actuaries is being considered.

### **Other Colloquium Events**

The ASTIN General Assembly took place after the coffee break on Wednesday morning. Alf Guldborg, President of the Swedish Actuarial Society, announced that the next Colloquium will be held in Stockholm in the summer of 1991, and welcomed members to complete a provisional registration form.

For the last few years there has been debate, sometimes heated, at ASTIN business meetings on the topic of the composition and system of election of the Committee. The Committee put forward some proposals at Montreux for changing the ASTIN rules and some alternative proposals were put forward by an ASTIN member. An interesting debate took place in which several members took part. Although contrary views were expressed, the discussion took place in a friendly atmosphere, as we would expect within a group such as ASTIN. The Committee's proposed changes to the rules were accepted by a majority.

decision and will be implemented. It is pleasing to note that the matter has finally been resolved.

Following the rule changes, the Committee will remain responsible for making nominations for Committee membership, and also it will still be possible for members to make further nominations at a General Assembly. However, the Committee, in making their nominations, are now charged with the responsibility of seeking a good balance of Committee membership as regards geographical spread, type of employment and research versus applied orientation. The Committee will give particular consideration to proposals through national actuarial organisations, but will reserve the right to make other nominations.

On Tuesday afternoon we boarded coaches for an enjoyable excursion into the Swiss countryside and mountains, including a visit to Gruyères.

We were privileged to meet on Wednesday evening for aperitifs inside the Château de Chillon, not normally available for private functions. This lakeside castle dates back to the 13th century, and narratives by well-known writers have contributed to its fame. Byron wrote "The Prisoner of Chillon", but we were not persuaded by the assertion of one eminent actuary that Byron had himself been imprisoned in the castle!

On this occasion the after-dinner speeches were delivered before the dinner began. Heralded by a fanfare of trumpets, the speakers included the retiring Chairman, Jean Lemaire, and his successor, Bjorn Ajne, who elegantly, entertainingly and appropriately referred to Jean's ability to make elegant, entertaining and appropriate speeches.

After the speeches we boarded the boat "La Suisse" for a cruise on Lake Geneva, with banquet and dancing. Needless to remark, the whole evening was superbly organised by our Swiss hosts and thoroughly enjoyed by the participants and accompanying persons.

After the final working session on Thursday morning, the brief closing ceremony took place. Bjorn Ajne announced the topics for papers for the 1991 Colloquium in Stockholm. The emphasis seemed to be very much on meeting modern challenges, the topics being The Use of Financial Theory in Insurance, High Tech Reinsurance and Modern Statistical Techniques.

It was no surprise that the Swiss organising committee, under the chairmanship of Robert Baumann and with Hans Gerber as head of the Scientific Committee, had done a most efficient job in organising all the aspects of the Colloquium. Our lack of surprise in no way diminished our gratitude to them. After making our farewells and leaving the Casino there was a final opportunity to take photographs of the flower-decked pathway by the lakeside, which had provided such pleasant morning and evening strolls in the sunshine each day. We look forward to meeting again in Sweden.

MARTYN BENNETT

**APPENDIX: INVITED LECTURES****Lecture: "The future European insurance market and the Swiss insurance industry" by Peter Gmeiner**

The speaker began by drawing attention to the insurance-mindedness of the Swiss, whose insurance premiums (life and non-life combined) in 1988 amounted to US\$ 2,324 per head of the population, about 60 per cent of this was life. In addition to the group life assurance provided by many employers for their staff, life assurance is widely used by individuals as a means of saving.

The Swiss approach to cartels is to allow them in principle but to seek to outlaw abuse; a fire insurance cartel had recently been prohibited. Agreements between insurers were seen as a means of avoiding the risk of insolvency. The market is closely regulated and insurance tariffs are in principle subject to approval. There are very few brokers operating in Switzerland, almost all the business being obtained through tied agents of the companies. Such brokers as there are have been active for only a few years, and in the major centres of population — mostly for industrial risks.

A feature of the Swiss insurance companies is their high capitalisation. The increase in the level of the stock market has enabled insurance companies to expand their capital in favourable conditions. Swiss companies transact a large amount of business outside their country, some Swiss companies started transacting foreign business when they were formed in the 19th century, and out of the total premium income of SF 70bn of the Swiss companies in 1988, SF 46bn related to foreign business.

The speaker then turned to the developments currently taking place in the EC and the influence they were likely to have on the conduct of insurance in Switzerland. He referred in particular to the intention within the EC to drop the examination and approval of insurance tariffs, perhaps with an exception with regard to compulsory insurance, and to the ending of insurance monopolies where they still exist — for example in some German states.

Mr Gmeiner then summarised the Swiss political aims and the options open to them. They would like to see European unity, of a kind which operated on the so-called principle of subsidiarity, with decision-making from bottom to top. They want to see a democratic Europe, with decisions taking account of local traditions. Switzerland would like to develop its policy of neutrality, in conjunction with the other neutral states: Austria, Sweden and Finland. Switzerland had already concluded a bilateral agreement with the EC on non-life insurance.

He ended by reviewing the reasons why, in his opinion, the Swiss insurers could face the future with confidence: they had a traditionally heavy commitment to foreign business and hence a long experience in handling it; the Swiss insurance companies are willing and able to adapt to new circumstances; they have great financial strength; and they are firmly rooted in the economically sound Swiss structure.

**Lecture: “Actuarial ethics and integrity” by James W. MacGinnitie**

The speaker began by referring to recent and current developments in North America, where the Society of Actuaries has introduced an admission course for new fellows, mainly on ethics, the Casualty Actuarial Society is developing a professionalism course for new associates and the Canadian Institute of Actuaries is also running courses on similar topics. He mentioned also the current discussions in Europe regarding the acceptability of actuarial reports and opinions across borders within the EC.

He next went on to summarise the features commonly associated with membership of a learned profession:

1. The members possessed expert knowledge not easily obtainable by the rest of the community or by clients
2. The members owned a technical language not easily understood by others.
3. It was difficult for outsiders to evaluate the quality of the advice they received, this being a matter essentially to be controlled by the profession
4. The member was in a position to be independent in a way that the client generally was not.
5. The members belonged to an élite group, had been subjected to a rigorous programme of study and were rewarded by such features as prestige, financial gain and camaraderie.

This all tended to lead to an unequal relationship between the professional and his or her client. It was fundamentally important that the member's special skills should be used in the best interests of the client, and that the client's interests should be placed ahead of the professional's interests.

The speaker listed a number of ethical issues facing actuaries, namely

The potential for abuse by the actuary of his or her position, and the need for the primacy of the interests of the client.

The actuary's responsibility to the public, especially in view of the increasing role of actuaries in the public arena

The development of codes of conduct.

The need for continuing education, to maintain the actuary's special knowledge and skills in current conditions.

He suggested that the testing of actuaries could be considered in three parts:

1. Knowledge of actuarial principles.
2. The ability to apply that knowledge in specific situations, such as to specific types of insurance.
3. Knowledge of specific legal and regulatory matters

There was increasingly a need to evaluate qualifications across national boundaries

Guidelines were required regarding the relationship between actuary and client (including business relationships). Most of the guidelines used in practice specified prohibitions, i.e. they set out what ought not to be done rather than what ought to be done, since the latter carried a much greater danger of leading to litigation.

He mentioned three key factors for a satisfactory relationship with the client : truth, confidence and consent.

The speaker then gave examples of the types of situations to be used as illustrations of potential ethical problems in the admission courses in the USA .

The danger of encouraging a client to agree to a liberal interpretation of regulations and hence lead him into an unsound course of action.

The dangers associated with inadequate data, inadequate time, or an inadequate budget

The difficulty of dealing with an unsavoury client, who wishes to do something that would be against the public interest, or even illegal.

The difficulty of deciding when to blow the whistle — at what point does the actuary have a liability to report illegal or unprofessional activity.

The need to see that errors that have been identified are corrected — one's own, or errors on the part of another actuary.

The decision as to who is the client — e.g the actuary's employer, or the person paying the fee, may not be the real client; for some purposes it may be considered appropriate to regard the members of a pension plan as the clients.

The speaker concluded with some comments about integrity He remarked that the actuarial profession had acquired a reputation for integrity, despite the fact that it had not specifically set out to select its members by reference to integrity, nor had it specifically trained for it. As examples of circumstances where there might be an especial need for integrity, he referred to pressures which might be placed on the actuary to:

1. reduce perceived margins in technical reserves;
2. increase the credibility adjustment following a good claims experience; and
3. replace advance funding by pay-as-you-go.

## LIST OF PAPERS

### Topic 1: Models of finance

ARTZNER P. and DELBAEN F

Credit Insurance with prepayment option

BROCKETT P. L. and KAHANE Y.

Risk, return, skewness and preference

HURLIMANN W.

Sur la couverture du risque financier dans l'actuariat

SANDERS D. E. A.

Option pricing and reinsurance

**Topic 2: Experience rating**

BENKTANDER G.

Exposure rating in fire — A simple model based on Swiss statistics

BERNEGGER S

Variance-loading in non-proportional reinsurance — An improved method for calculating the loading of NP-covers with the aid of personal computers

BESSON J.-L. and PARTRAT C.

Loi de Poisson — Inverse gaussienne et systèmes de bonus-malus

GABRY D G., KESTENS J C. A., LEENHEER F., PRINS H. J , ROOZENBOOM F., RUYGT F.A.M and WILLEMSE A.H

The estimation of net premiums for cover of fire and explosion risks

GERBER H. U.

The new bonus-malus system of Switzerland

GISLER A. and REINHARD P.

Robust credibility

KREMER E.

Practical exponential smoothing credibility

MACK T.

A simple parametric model for rating automobile insurance or estimating IBNR claims reserves

MOREAU L

Tarifification expérimentale d'un excédent de pertes en assurance automobile

NORBERG R.

Linear prediction and credibility in continuous time

SUNDT B

On excess of loss reinsurance with reinstatements

TEUGELS J. L. and SUNDT B.

A stop-loss experience rating scheme for fleets of cars

**Topic 3: Numerical methods**

ALTING VON GEUSAU B J.J

The shovelboard approach revisited

AMSLER M -H.

Risque de décès et risque de ruine — Réflexions sur la mesure du risque de ruine

CENTENO L.

An insight into the excess of loss retention limit

DICKSON D. C. M. and WATERS H. R.

Recursive calculation of survival probabilities



- DUFRESNE F., GERBER H. U. and SHIU E. S. W.  
Risk theory with the gamma process
- EKHULT H.  
Approximation of a function of two variables by a point matrix —  
Application to the termination function in disability insurance
- GOOVAERTS M. J. and KAAS R.  
Evaluating compound generalized Poisson distributions recursively
- HURLIMANN W.  
Algebraic moment approximations to compound Poisson distributions
- KOLLER B.  
Bayesian graduation on a spreadsheet
- KREMER E.  
Fourier methods for the claims amounts of largest claims reinsurance covers
- KREMER E.  
An elementary upper bound for the loading of a stop-loss cover
- KUMMERLI E.  
Verification de la réserve pour sinistres en suspens d'un assureur direct  
d'après les méthodes de F. de Vylder et de Y. Kahane
- MULLER T.  
The aggregate claims distribution in the individual model and others
- PENTIKAINEN T. and RANTALA J.  
On the evaluation of the run-off inaccuracies in claims reserving
- SCHMITTER H.  
The ruin probability of a discrete claims distribution with a finite number of steps

### Speakers' Corner

- BERG M. P.  
Trend detection and Bayesian prediction procedures for time nonhomogeneous claim processes
- CONTI B. and LAUCHLI F.  
The second class of G. Benktander and the class of E. Kremer: important properties and analysis of the likelihood function
- HARBITZ G.  
New law and regulations for appointed actuaries in Norway
- van HEERWAARDEN A. E. and KAAS R.  
Decomposition of ordered risks
- JEWELL W. S.  
Predicting IBNYR events and delays — III Collateral data

MAKOV U. E.

The implementation of Bayesian credibility models by a sampling-resampling technique

RENSHAW A. E.

Graduation by generalised linear modelling techniques

STEENACKERS A. and GOOVAERTS M. J.

Bounds on stop-loss premiums and ruin probabilities for given values of  $\mu$  and  $\sigma^2$

VERRALL R. J.

Bootstrapping, saddlepoint approximations and the probability of ruin

## INVITED PAPER

### COOPERATIVE GAME THEORY AND ITS INSURANCE APPLICATIONS

BY JEAN LEMAIRE

*Wharton School  
University of Pennsylvania, USA*

#### ABSTRACT

This survey paper presents the basic concepts of cooperative game theory, at an elementary level. Five examples, including three insurance applications, are progressively developed throughout the paper. The characteristic function, the core, the stable sets, the Shapley value, the Nash and Kalai-Smorodinsky solutions are defined and computed for the different examples.

#### 1. INTRODUCTION

Game theory is a collection of mathematical models to study situations of conflict and/or cooperation. It attempts to abstract out those elements that are common to many conflicting and/or cooperative encounters and to analyse these mathematically. Its goal is to explain, or to provide a normative guide for, rational behaviour of individuals confronted with strategic decisions or involved in social interaction. The theory is concerned with optimal strategic behaviour, equilibrium situations, stable outcomes, bargaining, coalition formation, equitable allocations, and similar concepts related to resolving group differences. The prevalence of competition in many human activities has made game theory a fundamental modeling approach in such diversified areas as economics, political science, operations research, and military planning.

In this survey paper, we will review the basic concepts of multiperson cooperative game theory, with insurance applications in mind. The reader is first invited to ponder the five following basic examples. Those examples will progressively be developed throughout the paper, to introduce and illustrate basic notions.

#### **Example 1. United Nations Security Council**

Fifteen nations belong to the United Nations Security Council: five permanent members (China, France, the United Kingdom, the Soviet Union, and the United States), and 10 nonpermanent members, on a rotating basis (in November 1990: Canada, Colombia, Cuba, Ethiopia, Finland, the Ivory Coast, Malaysia, Romania, Yemen, and Zaire). On substantive matters, including the investigation of a dispute and the application of sanctions,

decisions require an affirmative vote from at least nine members, including *all* five permanent members. If one permanent member votes against, a resolution does not pass. This is the famous "veto right" of the "big five," used hundreds of times since 1945. This veto right obviously gives each permanent member a much larger power than the nonpermanent members. But how much larger?

**Example 2.** Electoral representation in Nassau County [in LUCAS (1981)]

Nassau County, in the state of New York, has six municipalities, very unequal in population. The County Government is headed by a Board of six Supervisors, one from each municipality. In an effort to equalize citizen representation, Supervisors are given different numbers of votes. The following table shows the situation in 1964.

District	Population	%	No of Votes	%
Hempstead 1 } Hempstead 2 }	778,625	57.1	31	27.0
Oyster Bay	285,545	22.4	28	24.3
North Hempstead	213,335	16.7	21	18.3
Long Beach	25,654	2.0	2	1.7
Glen Cove	22,752	1.8	2	1.7
	1,275,801		115	

A simple majority of 58 out of 115 is needed to pass a measure. Do the citizens of North Hempstead and Oyster Bay have the same political power in their Government?

**Example 3.** Management of ASTIN money [LEMAIRE (1983)]

The Treasurer of ASTIN (player 1) wishes to invest the amount of 1,800,000 Belgian Francs on a short term (3 months) basis. In Belgium, the annual interest rate is a function of the sum invested.

Deposit	Annual Interest Rate
0-1,000,000	7.75%
1,000,000-3,000,000	10.25%
3,000,000-5,000,000	12%

The ASTIN Treasurer contacts the Treasurers of the International Actuarial Association (I.A.A. – player 2) and of the Brussels Association of Actuaries (A.A.Br. – player 3). I.A.A. agrees to deposit 900,000 francs in the common fund, A.A.Br. 300,000 francs. Hence the 3-million mark is reached and the

interest rate will be 12%. How should the interests be split among the three associations? The common practice in such situations is to award each participant in the fund the same percentage (12%). Shouldn't ASTIN however be entitled to a higher rate, on the grounds that it can achieve a yield of 10.25% on its own, and the others only 7.75%?  $\square$

**Example 4.** Managing retention groups [BORCH (1962)]

[For simplicity, several figures are rounded in this example]. Consider a group of  $n_1 = 100$  individuals. Each of them is exposed to a possible loss of 1, with a probability  $q_1 = 0.1$ . Assume these persons decide to form a risk retention group, a small insurance company, to cover themselves against that risk. The premium charged will be such that the ruin probability of the group is less than 0.001. Assuming that the risks are independent, and using the normal approximation of the binomial distribution, the group must have total funds equal to

$$P_1 = n_1 q_1 + 3 \sqrt{n_1 q_1 (1 - q_1)} = 10 + 9 = 19$$

Hence each person will pay, in addition to the net premium of 0.10, a safety loading of 0.09

Another group consists of  $n_2 = 100$  persons exposed to a loss of 1 with a probability  $q_2 = 0.2$ . If they form their own retention group under the same conditions, the total premium will be

$$P_2 = n_2 q_2 + 3 \sqrt{n_2 q_2 (1 - q_2)} = 20 + 12 = 32.$$

Assume now that the two groups decide to join and form one single company. In order to ensure that the ruin probability shall be less than 0.001, this new company must have funds amounting to

$$\begin{aligned} P_{12} &= n_1 q_1 + n_2 q_2 + 3 \sqrt{n_1 q_1 (1 - q_1) + n_2 q_2 (1 - q_2)} \\ &= 10 + 20 + 15 \\ &= 45. \end{aligned}$$

Since  $P_{12} = 45 < P_1 + P_2 = 51$ , the merger results in a decrease of 6 of the total safety loading. How should those savings be divided between the two groups? A traditional actuarial approach would probably consist in dividing the safety loading in proportion to the net premiums. This leads to premiums of 15 and 30, respectively. The fairness of this rule is certainly open to question, since it awards group 1 most of the gain accruing from the formation of a single company. In any case the rule is completely arbitrary  $\square$

**Example 5.** Risk exchange between two insurers

Insurance company  $C_1$  owns a portfolio of risks, with a mean claim amount of 5 and a variance of 4. Company  $C_2$ 's portfolio has a mean of 10 and a variance

of 8. The two companies decide to explore the possibility to conclude a risk exchange agreement. Assume only linear risk exchanges are considered. Denote by  $x_1$  and  $x_2$  the claim amounts before the exchange, and by  $y_1$  and  $y_2$  the claim amounts after the exchange. Then the most general form of a linear risk exchange is

$$\begin{aligned} y_1 &= (1-\alpha)x_1 + \beta x_2 + K \\ y_2 &= \alpha x_1 + (1-\beta)x_2 - K \end{aligned} \quad 0 \leq \alpha, \beta \leq 1$$

where  $K$  is a fixed (positive or negative) monetary amount. If  $K = 5\alpha - 10\beta$ , then  $E(y_1) = E(x_1) = 5$  and  $E(y_2) = E(x_2) = 10$ . So the exchange does not modify expected claims, and we only need to analyse variances. Assuming independence,

$$\begin{aligned} \text{Var}(y_1) &= 4(1-\alpha)^2 + 8\beta^2 \\ \text{Var}(y_2) &= 4\alpha^2 + 8(1-\beta)^2 \end{aligned}$$

If, for instance,  $\alpha = 0.2$  and  $\beta = 0.3$ ,  $\text{Var}(y_1) = 3.28 < 4$  and  $\text{Var}(y_2) = 4.08 < 8$ . Hence it is possible to improve the situation of both partners (if we assume, in this simple example, that companies evaluate their situation by means of the retained variance). Can we define "optimal" values of  $\alpha$  and  $\beta$ ?  $\square$

Those examples have several elements in common :

- Participants have some benefits to share (political power, savings, or money).
- This opportunity to divide benefits results from cooperation of all participants or a sub-group of participants.
- Individuals are free to engage in negotiations, bargaining, coalition formation.
- Participants have conflicting objectives; each wants to secure the largest part of the benefits for himself.

Cooperative game theory analyses those situations where participants' objectives are partially cooperative and partially conflicting. It is in the participants' interest to cooperate, in order to achieve the greatest possible total benefits. When it comes to sharing the benefits of cooperation, however, individuals have conflicting goals. Such situations are usually modeled as  $n$ -person cooperative games in characteristic function form, defined and illustrated in Section 2. Section 3 presents and discusses natural conditions, the individual and collective rationality conditions, that narrow the set of possible outcomes. Two concepts of solution are defined: the von Neumann-Morgenstern stable sets and the core. Section 4 is devoted to axiomatic approaches that aim at selecting a unique outcome. The main solution concept is here the Shapley value. Section 5 deals with two-person cooperative games without transferable utilities. The Nash and Kalai-Smorodinsky solution concepts are presented and applied to Example 5. A survey of some other solutions and concluding remarks are to be found in Sections 6 and 7.

## 2. CHARACTERISTIC FUNCTIONS

First, let us specify which situations will be considered in this paper, and some implicit assumptions.

- Participants are authorized to freely cooperate, negotiate, bargain, collude, make binding contracts with one another, form groups or subgroups, make threats, or even withdraw from the group
- All participants are fully informed about the rules of the game, the payoffs under each possible situation, all strategies available, ...
- Participants are negotiating about sharing a given commodity (such as money or political power) which is fully transferable between players and evaluated in the same way by everyone. This excludes for instance games where participants evaluate their position by means of a concave utility function; risk aversion is not considered. (In other words, it is assumed that all individuals have linear utility functions). For this reason, the class of games defined here is called "Cooperative games with transferable utilities." This major assumption will be relaxed in Section 5.

*Definition 1* · An  $n$ -person game in characteristic function form  $\Gamma$  is a pair  $[N, v]$ , where  $N = \{1, 2, \dots, n\}$  is a set of  $n$  players.  $v$  is a real valued characteristic function on  $2^N$ , the set of all subsets  $S$  of  $N$ .  $v$  assigns a real number  $v(S)$  to each subset  $S$  of  $N$ , and  $v(\emptyset) = 0$ .

Subsets  $S$  of  $N$  are called coalitions. The full set of players  $N$  is the grand coalition. Intuitively,  $v(S)$  measures the worth or power that coalition  $S$  can achieve when its members act together. Since cooperation creates savings, it is assumed that  $v$  is superadditive, i.e., that

$$v(S \cup T) \geq v(S) + v(T) \text{ for all } T, S \subset N \text{ such that } S \cap T = \emptyset$$

*Definition 2* Two  $n$ -person games  $\Gamma$  and  $\Gamma'$ , of respective characteristic functions  $v$  and  $v'$ , are said to be strategically equivalent if there exists numbers  $k > 0, c_1, \dots, c_n$  such that

$$v'(S) = kv(S) + \sum_{i \in S} c_i \quad \text{for all } S \subseteq N.$$

The switch from  $v$  to  $v'$  only amounts to changing the monetary units and awarding a subsidy  $c_i$  to each player. Fundamentally, this operation doesn't change anything. Hence we only need to study one game in each class of strategically equivalent games. Therefore games are often normalized by assuming that the worth of each player is zero, and that the worth of the grand coalition is 1 [In the sequel expressions such as  $v(\{1,3\})$  will be abbreviated as  $v(13)$ ].

$$v(i) = 0 \quad i = 1, \dots, n \quad v(N) = 1$$

**Example 1.** (UN Security Council). Since a motion either passes or doesn't, we can assign a worth of 1 to all winning coalitions, and 0 to all losing coalitions. The game can thus be described by the characteristic function

$$v(S) = 1 \quad \text{for all } S \text{ containing all five permanent members and at least 4 nonpermanent members}$$

$$v(S) = 0 \quad \text{for all other } S. \quad \square$$

Games such that  $v(S)$  can only be 0 or 1 are called simple games. One interesting class of simple games is the class of weighted majority games.

*Definition 3* A weighted majority game

$$\Gamma = [M; w_1, \dots, w_n],$$

where  $w_1, \dots, w_n$  are nonnegative real numbers and

$$M > \frac{1}{2} \sum_{i=1}^n w_i,$$

is the  $n$ -person cooperative game with characteristic function

$$v(S) = 1 \quad \text{if} \quad \sum_{i \in S} w_i \geq M$$

$$v(S) = 0 \quad \text{if} \quad \sum_{i \in S} w_i < M,$$

for all  $S \subseteq N$ .  $w_i$  is the power of player  $i$  (such as the number of shares held in a corporation)  $M$  is the required majority.

**Example 1.** It is easily verified that the UN Security Council's voting rule can be modelled as a weighted majority game. Each permanent member is awarded seven votes, each nonpermanent member one vote. The majority required to pass a motion is 39 votes. A motion can only pass if all five permanent members (35 votes) and at least four nonpermanent members (4 votes) are in favor. Without the adhesion of all permanent members, the majority of 39 votes cannot be reached.

$$\Gamma = [39; 7,7,7,7,7,1,1,1,1,1,1,1,1,1,1]$$

Does this mean that the power of each permanent member is seven times the power of nonpermanent members?  $\square$

**Example 2.** Nassau County's voting procedures form the weighted majority game [58, 31,31,28,21,2,2]. It clearly shows that numerical voting weights do not translate into political power. An inspection of all numerical possibilities reveals that the three least-populated municipalities have no voting power at



all. Their combined total of 25 votes is never enough to tip the scales. To pass a motion simply requires the adhesion of two of the three largest districts. So the assigned voting weights might just as well be (31,31,28,0,0,0), or (1,1,1,0,0,0). We need a better tool than the number of votes to evaluate participants' strengths.  $\square$

**Example 3.** (ASTIN money). Straightforward calculations lead to the total interest each coalition can secure

$$\begin{aligned} v(1) &= 46,125 \\ v(2) &= 17,437.5 \\ v(3) &= 5,812.5 \\ v(12) &= 69,187.5 \\ v(13) &= 53,812.5 \\ v(23) &= 30,750 \\ v(123) &= 90,000 \end{aligned} \quad \square$$

**Example 4.** (Retention groups) This example differs from the others in the sense that figures here represent costs (to minimise) and not earnings (to maximise). Instead of a superadditive characteristic function  $v(S)$ , a cost function  $c(S)$  is introduced. Scale economies make  $c(S)$  a subadditive function

$$c(S \cup T) \leq c(S) + c(T) \text{ for all } S, T \subseteq N \text{ such that } S \cap T = \emptyset$$

A "cost" game is equivalent to a "savings" game, of characteristic function

$$v(S) = \sum_{i \in S} c_i - c(S).$$

In the case of the example,  $c(S)$  is the premium paid by each coalition

$$\begin{aligned} c(1) &= 19 \\ c(2) &= 32 \\ c(12) &= 45 \end{aligned} \quad \square$$

### 3. VON NEUMANN-MORGENSTERN STABLE SETS AND THE CORE

**Example 3.** (ASTIN money) If they agree on a way to subdivide the profits of cooperation, the three Treasurers will have a total of 90,000 francs to share. Denote  $\alpha = (\alpha_1, \alpha_2, \alpha_3)$  the outcome (or payoff, or allocation). player  $i$  will receive the amount  $\alpha_i$ . Obviously, the ASTIN Treasurer will only accept an allocation that awards him at least 46,125 francs, the amount he can secure by himself. This is the individual rationality condition.  $\square$

*Definition 4* A payoff  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  is individually rational if  $\alpha_i \geq v(i)$   $i = 1, \dots, n$ .

*Definition 5* An imputation for a game  $\Gamma = (N, v)$  is a payoff  $\alpha = (\alpha_1, \dots, \alpha_n)$  such that

$$\alpha_i \geq v(i) \quad i = 1, \dots, n$$

$$\sum_{i=1}^n \alpha_i = v(N)$$

An imputation is an individually rational payoff that allocates the maximum amount (This condition is also called “efficiency” or “Pareto-optimality”).

**Example 3.** (ASTIN money) An imputation is any allocation such that

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= 90,000 \\ \alpha_1 &\geq 46,125 \\ \alpha_2 &\geq 17,437.5 \\ \alpha_3 &\geq 5,812.5 \quad \square \end{aligned}$$

**Example 4.** (Retention groups). In this cost example, an imputation is any set of premiums  $(\alpha_1, \alpha_2)$  such that

$$\begin{aligned} \alpha_1 + \alpha_2 &= 45 \\ \alpha_1 &\leq 19 \\ \alpha_2 &\leq 32 \end{aligned}$$

Let us now add a third group of  $n_3 = 120$  individual to this example, all subject to a loss of 1 with a probability  $q_3 = 0.3$ . A risk retention group with a ruin probability of .001 would require a total premium of

$$n_3 q_3 + 3 \sqrt{n_3 q_3 (1 - q_3)} = 36 + 15 = 51$$

If all three groups decide to merge to achieve a maximum reduction of the safety loading, the total premium will be

$$\begin{aligned} n_1 q_1 + n_2 q_2 + n_3 q_3 + 3 \sqrt{n_1 q_1 (1 - q_1) + n_2 q_2 (1 - q_2) + n_3 q_3 (1 - q_3)} \\ = 10 + 20 + 36 + 21 \\ = 87 \end{aligned}$$

In this case an imputation is a payoff  $(\alpha_1, \alpha_2, \alpha_3)$  such that

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= 87 \\ \alpha_1 &\leq 19 \\ \alpha_2 &\leq 32 \\ \alpha_3 &\leq 51 \end{aligned}$$

Are all those imputations acceptable to everybody? Consider the allocation (17, 31, 39). It is an imputation. It will however never be accepted by the first two groups. Indeed they are better off withdrawing from the grand coalition, forming coalition (12), and agreeing for instance on a payoff (15.5, 29.5). Player 3, the third group, cannot object to this secession since, left alone, he will be stuck to a premium of 51. He will be forced to make a concession during negotiations and accept a higher  $\alpha_3$ .  $\alpha_3$  needs to be at least 42 to prevent players 1 and 2 to secede. This is the collective rationality condition: no coalition should have an incentive to quit the grand coalition.  $\square$

*Definition 6.* A payoff  $(\alpha_1, \alpha_2, \dots, \alpha_n)$  is collectively rational if

$$\sum_{i \in S} \alpha_i \geq v(S) \quad \text{for all } S \subset N.$$

*Definition 7.* The core of the game is the set of all collectively rational payoffs.

The core of a game can be empty. When it is not, it usually consists of several, or an infinity, of points. It can also be defined using the notion of dominance.

*Definition 8.* Imputation  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$  dominates imputation  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$  with respect to coalition  $S$  if

(i)  $S \neq \emptyset$

(ii)  $\beta_i > \alpha_i$  for all  $i \in S$

(iii)  $v(S) \geq \sum_{i \in S} \beta_i$

So there exists a non-void set of players  $S$ , that all prefer  $\beta$  to  $\alpha$ , and that has the power to enforce this allocation.

*Definition 9.* Imputation  $\beta$  dominates imputation  $\alpha$  if there exists a coalition  $S$  such that  $\beta$  dominates  $\alpha$  with respect to  $S$ .

*Definition 7'.* The core is the set of all the undominated imputations.

Definitions 7 and 7' are equivalent.

**Example 4.** (Retention groups). The core is the set of all payoffs that allocate the total premium of 87, while satisfying the 3 individual and 3 collective rationality conditions.

$$\begin{aligned}
 \alpha_1 + \alpha_2 + \alpha_3 &= 87 \\
 \alpha_1 &\leq 19 \\
 \alpha_2 &\leq 32 \\
 \alpha_3 &\leq 51 \\
 \alpha_1 + \alpha_2 &\leq 45 \\
 \alpha_1 + \alpha_3 &\leq 63.5 \\
 \alpha_2 + \alpha_3 &\leq 75.3
 \end{aligned}$$

So the core enables us to find upper and lower bounds for the premiums

$$\begin{aligned}
 \alpha_1 + \alpha_2 + \alpha_3 &= 87 \\
 11.7 &\leq \alpha_1 \leq 19 \\
 23.5 &\leq \alpha_2 \leq 32 \\
 42 &\leq \alpha_3 \leq 51
 \end{aligned}$$

An allocation that violates any inequality leads to the secession of one or two groups.  $\square$

**Example 3.** (ASTIN money). The core consists of all payoffs such that

$$\begin{aligned}
 \alpha_1 + \alpha_2 + \alpha_3 &= 90,000 \\
 46,125 &\leq \alpha_1 \leq 59,250 \\
 17,437.5 &\leq \alpha_2 \leq 36,187.5 \\
 5,812.5 &\leq \alpha_3 \leq 20,812.5
 \end{aligned}
 \quad \square$$

Despite its intuitive appeal, the core was historically not the first concept that attempted to reduce the set of acceptable payoffs with rationality conditions. In their path-breaking work, VON NEUMANN and MORGENSTERN (1945) introduced the notion of stable sets

*Definition 10* A von Neumann-Morgenstern stable set of a game  $G = (N, v)$  is a set  $L$  of imputations that satisfy the two following conditions

- (i) (External stability) To each imputation  $\alpha \notin L$  corresponds an imputation  $\beta \in L$  that dominates  $\alpha$
- (ii) (Internal stability) No imputation of  $L$  dominates another imputation of  $L$ .

Stable sets are however usually very difficult to compute

The main drawback of the core and the stable sets seems to be that, in most cases, they contain an infinity of allocations. For instance, the core and the stable set of all 2-person games simply consist of all imputations. It would be preferable to be able to single out a unique, "fair" payoff for each game. This is what the Shapley value achieves

4 THE SHAPLEY VALUE

**Example 3.** (ASTIN money). Assume the ASTIN Treasurer decides to initiate the coalition formation process. Playing alone, he would make  $v(1) = 46,125$ . If player 2 decides to join, coalition (12) will make  $v(12) = 69,187.5$ . Assume player 1 agrees to award player 2 the entire benefits of cooperation; player 2 receives his entire admission value  $v(12) - v(1) = 23,062.5$ . Player 3 joins in a second stage, and increases the total gain to 90,000. If he is allowed to keep his entire admission value  $v(123) - v(12) = 20,812.5$ , we obtain the payoff

$$[46,125; \quad 23,062.5; \quad 20,812.5]$$

This allocation of course depends on the order of formation of the grand coalition. If player 1 joins first, then player 3, and finally player 2, and if everyone keeps his entire admission value, the following payoff results

$$[46,125, \quad 36,187.5, \quad 7,687.5]$$

The four other player permutations [(213), (231), (312), (321)] lead to the respective payoffs

$$[51,750; \quad 17,437.5; \quad 20,812.5]$$

$$[59,250; \quad 17,437.5; \quad 13,312.5]$$

$$[48,000; \quad 36,187.5; \quad 5,812.5]$$

$$[59,250, \quad 24,937.5, \quad 5,812.5]$$

Assume we now decide to take the average of those six payoffs, to obtain the final allocation

$$[51,750; \quad 25,875 \quad , \quad 12,375 \quad ]$$

We have in fact computed the Shapley value of the game, the expected admission value when all player permutations are equiprobable  $\square$

The Shapley value is the only outcome that satisfies the following set of three axioms [SHAPLEY, 1953])

*Axiom 1* (Symmetry). For all permutations  $\Pi$  of players such that  $v[\Pi(S)] = v(S)$  for all  $S$ ,  $\alpha_{\Pi(i)} = \alpha_i$ .

A symmetric problem has a symmetric solution. If there are two players that cannot be distinguished by the characteristic function, that contribute the same amount to each coalition, they should be awarded the same payoff. This axiom is sometimes also called anonymity, it implies that the selected allocation only depends on the characteristic function, and not, for instance, on the numbering of the players

*Axiom 2* (Dummy players). If, for a player  $i$ ,  $v(S) = v(S \setminus i) + v(i)$  for each coalition to which he can belong, then  $\alpha_i = v(i)$ .

A dummy player does not contribute any scale economy to any coalition. The worth of any coalition only increases by  $v(i)$  when he joins. Such an inessential player cannot claim to receive any share of the benefits of cooperation.

*Axiom 3* (Additivity). Let  $\Gamma = (N, v)$  and  $\Gamma' = (N, v')$  be two games, and  $\alpha(v)$  and  $\alpha'(v')$  their respective payoffs. Then  $\alpha(v+v') = \alpha(v) + \alpha(v')$  for all players.

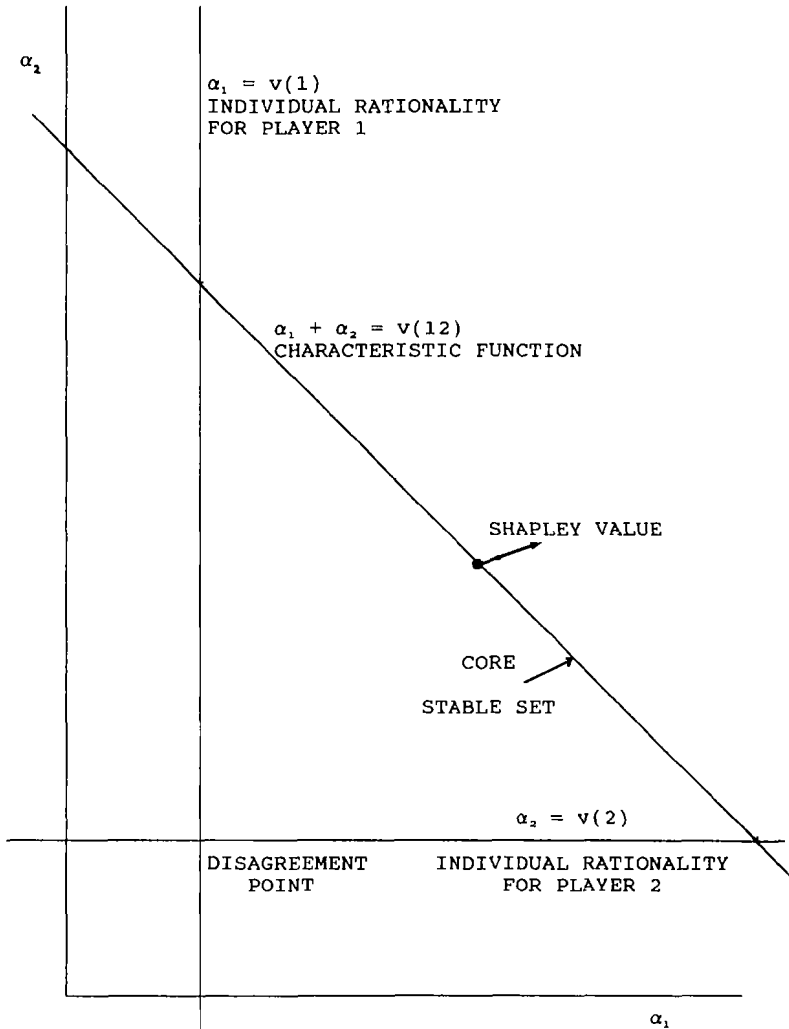


FIGURE 1 Two-person cooperative game with transferable utilities

Payoffs resulting from two distinct games should be added. While the first two axioms seem quite justified, the latter has been criticized. It rules out all interactions between the two games, for instance.

Shapley has shown that one and only one allocation satisfies the three axioms

$$\alpha_i = (1/n!) \sum_S (s-1)! (n-s)! [v(S) - v(S \setminus i)] \quad i = 1, \dots, n$$

where  $s$  is the number of members of a coalition  $S$

The Shapley value can be *interpreted* as the mathematical expectation of the admission value, when all orders of formation of the grand coalition are equiprobable. In computing the value, one can assume, for convenience, that all players enter the grand coalition one by one, each of them receiving the entire benefits he brings to the coalition formed just before him. All orders of formation of  $N$  are considered and intervene with the same weight  $1/n!$  in the computation. The combinatorial coefficient results from the fact that there are  $(s-1)!(n-s)!$  ways for a player to be the last to enter coalition  $S$ : the  $s-1$  other players of  $S$  and the  $n-s$  players of  $N \setminus S$  can be permuted without affecting  $i$ 's position

In a two-player game, the Shapley value is

$$\alpha_1 = (1/2) [v(12) + v(1) - v(2)]$$

$$\alpha_2 = (1/2) [v(12) + v(2) - v(1)]$$

It is the middle of the segment  $\alpha_1 + \alpha_2 = v(12)$ ,  $\alpha_1 \geq v(1)$ ,  $\alpha_2 \geq v(2)$ . This is illustrated in Figure 1.

**Example 1. (UN Security Council).** In a weighted majority game, the admission value of a player is either 0 or 1. One simply has to compute the probability that a player clinches victory for a motion. In the UN Security Council game, the power of a nonpermanent member  $i$  is the probability that he enters ninth in any coalition that already includes the five permanent members. It is

$$\alpha_i = \binom{8}{3} \underbrace{(5/15) (4/14) (3/13) (2/12) (1/11)}_{\text{all five permanent before } i} \underbrace{(9/10) (8/9) (7/8)}_{\substack{\text{3 of the} \\ \text{nonpermanent} \\ \text{before } i}} \downarrow (1/7)_{\substack{i \text{ then} \\ \text{enters}}}$$

$$= 0.1865\%$$

By symmetry, the power for each permanent member is

$$\alpha_i = 19.62\%$$

So permanent nations are 100 times more powerful than nonpermanent nations. [Note: in practice a permanent member may abstain without impair-

ing the validity of an affirmative vote. While this rule complicates the analysis of the game, it only changes the second decimal of the Shapley value].  $\square$

**Example 2.** (Nassau County) The Shapley value of the districts is  $(1/3, 1/3, 1/3, 0, 0, 0)$ . This analysis led the County authorities to change the voting rules by increasing the required majority from 58 to 63. There are now no more dummy players, and the new power indices are  $[0.283, 0.283, 0.217, 0.117, 0.050, 0.050]$ . This is certainly much closer to the original intention.  $\square$

**Example 4.** (Retention groups). In the two-company version of this game, the Shapley value is  $[16, 29]$ . In the three-company version, the value is  $[14.5, 26.9, 45.6]$ . The traditional pro rata approach leads to  $[13.2, 26.4, 47.4]$ . It does not take into account the savings each member brings to the grand coalition, or its threat possibilities. It is unfair to the third group, because it fails to give proper credit to the important reduction (10) of the total safety loading it brings to the grand coalition  $\square$

The Shapley value may lie outside the core. In the important subclass of convex games, however, it will always be in the core.

*Definition 11.* A game is convex if, for all  $S \subseteq T \subseteq N$ , for all  $i \notin T$ ,

$$v(T \cup i) - v(T) \geq v(S \cup i) - v(S).$$

A game is convex when it produces large economies of scale, a "snow-balling" effect makes it increasingly interesting to enter a coalition as its number of members increases. In particular, it is always preferable to be the last to enter the grand coalition  $N$ . The core of convex games is always non-void. Furthermore, it coincides with the unique von Neumann-Morgenstern stable set. It is a compact convex polyhedron, of dimension at most  $n-1$ . The Shapley value lies in the center of the core, in the sense that it is the center of gravity of the core's external points.

## 5. TWO-PERSON GAMES WITHOUT TRANSFERABLE UTILITIES

**Example 5.** (Risk exchange). As shown in the presentation of the example, selecting  $\alpha = 0.2$  and  $\beta = 0.3$  results in a decrease of  $\text{Var}(y_1)$  of 0.72, and a decrease of  $\text{Var}(y_2)$  of 3.92. This risk exchange treaty is represented as point 1 in Figure 2.

In this figure the axes measure the respective variance reductions,  $p_1$  and  $p_2$ . Point 2 corresponds to  $\alpha = \beta = 0.4$ . It dominates point 1, since it leads to a greater variance reduction for both companies. Point 3 is  $\alpha = 0.53, \beta = 0.47$ , it dominates points 1 and 2. It can be shown that no point can dominate point 3, and that all treaties such that  $\alpha + \beta = 1$  neither dominate nor are dominated by point 3. For instance, point 4 ( $\alpha = 0.7, \beta = 0.3$ ) will be preferred to point 3 by



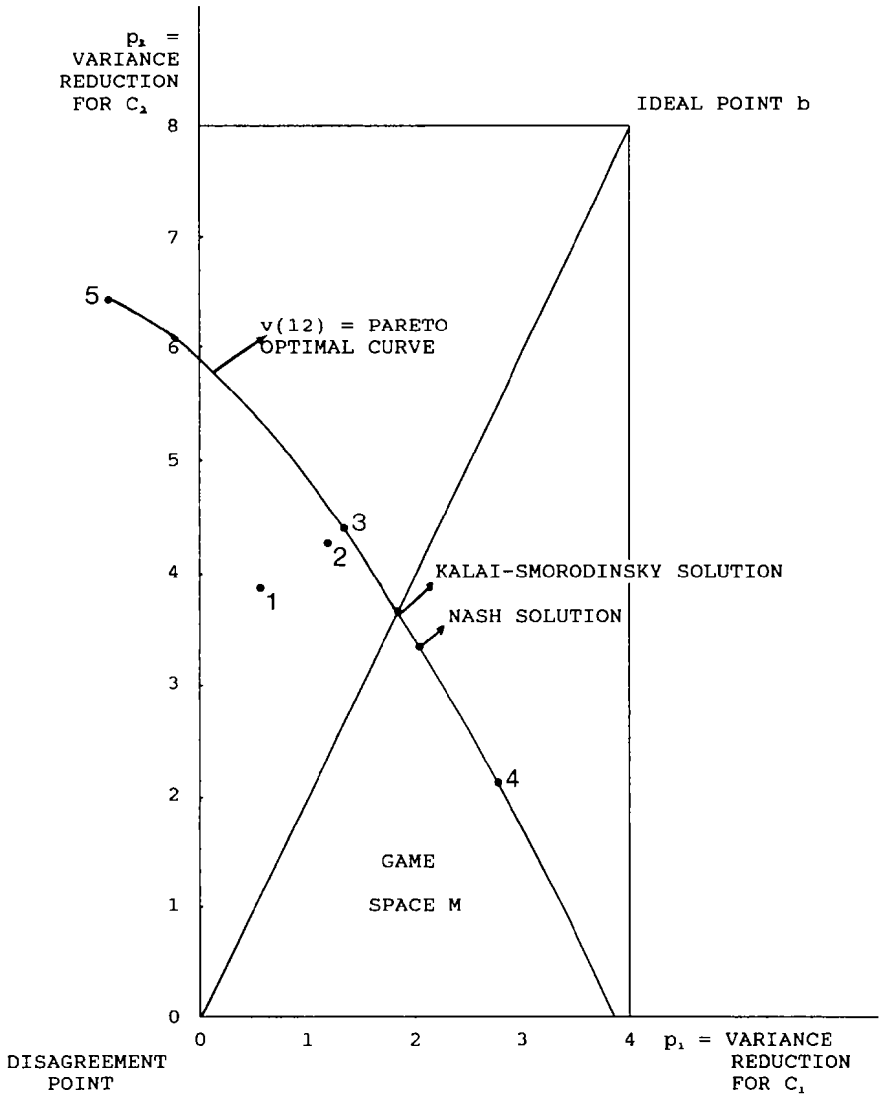


FIGURE 2 Two-person cooperative game without transferable utilities

$C_1$  However  $C_2$  will prefer point 3 to point 4 Hence neither point dominates the other The set of all treaties such that  $\alpha + \beta = 1$  forms curve  $v(12)$ , the Pareto-optimal surface. Points to the north-east of  $v(12)$  cannot be attained. All points to the south-west of  $v(12)$  correspond to a given selection of  $\alpha$  and  $\beta$ . The convex set of all attainable points, including the boundary  $v(12)$ , is called the game space  $M$ . That space is limited by the Pareto-optimal curve and the two axes. The axes represent the two individual rationality conditions: no

company will accept a treaty that results in a variance increase. For instance point 5 ( $\alpha = 0.35$ ,  $\beta = 0.65$ ) will not be accepted by  $C_1$ . While each point in the game space is attainable, it is in both companies' interest to cooperate to reach the Pareto-optimal curve. Any point that does not lie on the north-east boundary is dominated by a Pareto-optimal point. Once the curve is reached, however, the players' interests become conflicting.  $C_1$  will negotiate to reach a point as far east as possible, while  $C_2$  will attempt to move the final treaty north. If the players cannot reach an agreement, no risk exchange will take place. The disagreement point results in no variance reduction.

Hence all the elements of a two-player game are present in this simplified risk exchange example. In fact, Figure 2 closely resembles Figure 1, with an important difference: the Pareto-optimal set of treaties  $v(12)$  is a curve in Figure 2, while the characteristic function  $v(12)$  in Figure 1 is a straight line. This is due to the non-transferability of utilities in the risk exchange example. The players are "trading" variances, but an increase of 1 of  $\text{Var}(y_1)$  results in a decrease of  $\text{Var}(y_2)$  that is not equal to 1. Example 5 is a two-person cooperative game without transferable utility.  $\square$

*Definition 12.* A two-person cooperative game without transferable utilities is a couple  $(M, d)$ , where  $d = (d_1, d_2)$  is the disagreement point (the initial utilities of the players).  $M$ , the game space, is a convex compact set in the two-dimensional space  $E^2$  of the players' utilities; it represents all the payoffs that can be achieved.

Such a game is often called a two-person bargaining game. Let  $B$  be the set of all pairs  $(M, d)$ . Since no player will accept a final payoff that does not satisfy the individual rationality condition,  $M$  can be limited to the set of points  $(p_1, p_2)$  such that  $p_1 \geq d_1$  and  $p_2 \geq d_2$ . Our goal is to select a unique payoff in  $M$ .

*Definition 13.* A solution (or a value) is a rule that associates to each bargaining game a payoff in  $M$ . It is thus a mapping  $f: B \rightarrow E^2$  such that  $f(M, d)$  is a point  $p = (p_1, p_2)$  of  $M$  for all  $(M, d) \in B$ ;  $f_1(M, d) = p_1$  and  $f_2(M, d) = p_2$ .

The first solution concept for bargaining games was developed in 1950 by Nash. The Nash solution satisfies the four following axioms

*Axiom 1.* Independence of linear transformations

The solution cannot be affected by linear transformations performed on the players' utilities. For all  $(M, d)$  and all real numbers  $a_i > 0$  and  $b_i$ , let  $(M', d')$  be the game defined by  $d'_i = a_i d_i + b_i$  ( $i = 1, 2$ ) and  $M' = \{q \in E^2 \mid \exists p \in M \text{ such that } q_i = a_i p_i + b_i\}$ . Then  $f_i(M', d') = a_i f_i(M, d) + b_i$   $i = 1, 2$ .

This axiom is hard to argue with. It only reflects the information contained

in utility functions. Since utilities are only defined up to linear transformations, it should be the same for solutions.

### *Axiom 2. Symmetry*

All symmetric games have a symmetric solution. A game is symmetric if  $d_1 = d_2$  and  $(p_1, p_2) \in M \Rightarrow (p_2, p_1) \in M$ . The axiom requires that, in this case,  $f_1(M, d) = f_2(M, d)$ .

Like axiom 1, axiom 2 requires that the solution only depends on the information contained in the model. A permutation of the two players should not modify the solution, if they cannot be differentiated by the rules of the game. Two players with the same utility function and the same initial wealth should receive the same payoff if the game space is symmetric.

### *Axiom 3. Pareto-optimality*

The solution should be on the Pareto-optimal curve. For all  $(M, d) \in B$ , if  $p$  and  $q \in M$  are such that  $q_i > p_i$  ( $i = 1, 2$ ), then  $p$  cannot be the solution:  $f(M, d) \neq p$ .

### *Axiom 4. Independence of irrelevant alternatives*

The solution does not change if we remove from the game space any point other than the disagreement point and the solution itself. Let  $(M, d)$  and  $(M', d)$  be two games such that  $M'$  contains  $M$  and  $f(M', d)$  is an element of  $M$ . Then  $f(M, d) = f(M', d)$ .

This axiom formalizes the negotiation procedure. It requires that the solution, which by axiom 3 must lie on the upper boundary of the game space, depends on the shape of this boundary only in its neighbourhood, and not on distant points. It expresses the fact that, during negotiations, the set of the alternatives likely to be selected is progressively reduced. At the end, the solution only competes with very close points, and not with proposals already eliminated during the first phases of the discussion. Nash's axioms thus model a bargaining procedure that proceeds by narrowing down the set of acceptable points. Each player makes concessions until the final point is selected.

NASH (1950) has shown that one and only one point satisfies the four axioms. It is the point that maximizes the product of the two players' utility gains. Nash's solution is the function  $f$ , defined by  $f(M, d) = p$ , such that  $p \geq d$  and  $(p_1 - d_1)(p_2 - d_2) \geq (q_1 - d_1)(q_2 - d_2)$ , for all  $q \neq p \in M$ .

**Example 5.** (Risk exchange). In this example, the players' objective is to reduce the variance of their claims. Hence  $d = (0, 0)$ : if the companies cannot agree on a risk exchange treaty, they will keep their original portfolio, with no improvement. The players' variance reductions are

$$p_1 = 4 - 4(1 - \alpha)^2 - 8\beta^2$$

$$p_2 = 8 - 4\alpha^2 - 8(1 - \beta)^2$$

Maximising the product  $p_1 p_2$ , under the condition  $\alpha + \beta = 1$ , leads to the Nash solution

$$\alpha = 0.613$$

$$\beta = 0.387$$

$$p_1 = 2.203$$

$$p_2 = 3.491 \quad \square$$

Nash's axiom 4 has been criticised by KALAI and SMORODINSKY (1975), who proved that Nash's solution does not satisfy a monotonicity condition. Consider the two games represented in Figure 3. The space of game 1 is the four-sided figure whose vertices are at  $d$ ,  $A$ ,  $B$ ,  $D$ . The Nash solution is  $B$ . The space of game 2 is the figure whose vertices are at  $d$ ,  $A$ ,  $C$ ,  $D$ . From the second player's point of view, game 2 seems more attractive, since he stands to gain more if the first player's payoff is between  $E$  and  $D$ . So one would expect the second player's payoff to be larger in game 2. This is not the case, since the Nash solution of game 2 is  $C$ .

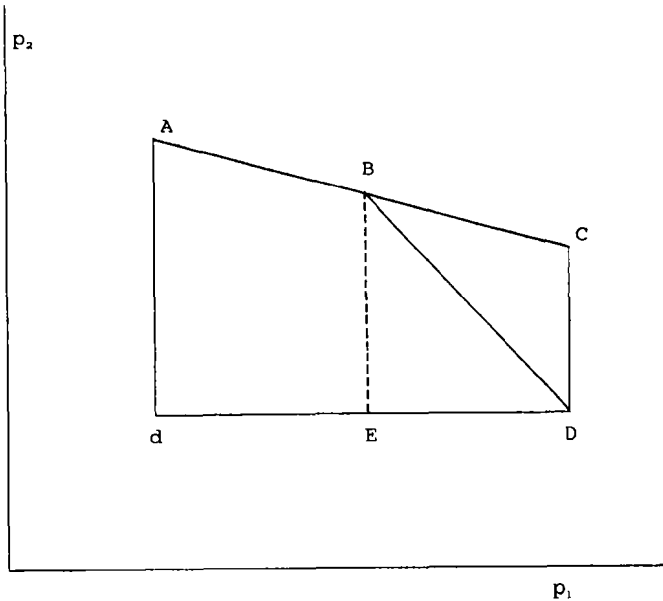


FIGURE 3 Non-monotonicity of Nash's solution

*Axiom 5. Monotonicity* Let  $b(M) = (b_1, b_2)$  the “ideal” point formed by the maximum possible payoffs (see Figure 2):  $b_i = \max \{p_i | (p_1, p_2) \in M\}$  ( $i = 1, 2$ ). If  $(M, d)$  and  $(M', d)$  are two games such that  $M$  contains  $M'$  and  $b(M) = b(M')$ , then  $f(M, d) \geq f(M', d)$ .

KALAI and SMORODINSKY have shown that one and only one point satisfies axioms 1, 2, 3, and 5. It is situated at the intersection of the Pareto-optimal curve and the straight line linking the disagreement point and the ideal point.

**Example 5.** It is easily verified that the equation of the Pareto-optimal curve is  $\sqrt{8-p_1} + \sqrt{4-p_2} = 12$ . Since the ideal point is (4,8), the line joining  $d$  and  $b$  has equation  $p_2 = 2p_1$ . Kalai-Smorodinsky’s solution point, at the intersection, is

$$\begin{aligned} \alpha &= 0.5858 \\ \beta &= 0.4142 \\ p_1 &= 1.9413 \\ p_2 &= 3.8821 \end{aligned}$$

It is slightly more favourable to player 2 than Nash’s solution. □

## 6. OTHER SOLUTION CONCEPTS – OVERVIEW OF LITERATURE

Stable sets and the core are the most important solution concepts of game theory that attempt to reduce the number of acceptable allocations by introducing intuitive conditions. Both notions however can be criticized.

Stable sets are difficult to compute. Some games have no stable sets. Some others have several. Moreover, the dominance relation is neither antisymmetric nor transitive. It is for instance possible that an imputation  $\beta$  dominates an imputation  $\alpha$  with respect to one coalition, while  $\alpha$  dominates  $\beta$  with respect to another coalition. Therefore an imputation inside a stable set may be dominated by an imputation outside.

The concept of core is appealing, because it satisfies very intuitive rationality conditions. However, there exists vast classes of games that have an empty core: the rationality conditions are conflicting. Moreover, several examples have been built for which the core provides a counter-intuitive payoff, as shown in Example 6.

### **Example 6.** A pair of shoes

Player 1 owns a left shoe. Players 2 and 3 each own a right shoe. A pair can be sold for \$ 100. How much should 1 receive if the pair is sold? Surprisingly, the core totally fails to catch the threat possibilities of coalition (23) and selects the paradoxical allotment (100, 0, 0). Any payoff that awards a positive amount to 2 or 3 is dominated, for instance (99, 1, 0) is dominated by (99.5, 0, 0.5).

Moreover, the paradox remains if we assume that there are 999 left shoes and 1000 right shoes. The game is now nearly symmetrical, but the owners of right shoes still receive nothing. The Shapley value is  $(66\frac{2}{3}, 16\frac{2}{3}, 16\frac{2}{3})$ , definitely a much better representation of the power of each player than the core.  $\square$

Many researchers feel that the core is too static a concept, that it does not take into account the real dynamics of the bargaining process. In addition, laboratory experiments consistently produce payoffs that lie outside the core. This led AUMANN and MASCHLER (1964) to define the bargaining set. This set explicitly recognizes the fact that a negotiation process is a multi-criteria situation. Players definitely attempt to maximise their payoff, but also try to enter into a "safe" or "stable" coalition. Very often, it is observed that players willingly give up some of their profits to join a coalition that they think has fewer chances to fall apart. This behaviour is modelled through a dynamic process of "threats" and "counter-threats." A payoff is then considered stable if all objections against it can be answered by counter-objections

**Example 7.** Consider the three-person game

$$\begin{aligned}v(1) &= v(2) = v(3) = 0 \\v(12) &= v(13) = 100 \\v(23) &= 50\end{aligned}$$

The core of this game is empty. For instance, the players will not agree on an allocation like  $[75, 25, 0]$ , because it is dominated by  $[76, 0, 24]$ . Bargaining set theory, on the other hand, claims that such a payoff is stable. If player 1 threatens 2 of a payoff  $[76, 0, 24]$ , this objection can be met with the counter-objection  $[0, 25, 25]$ . Player 2 shows that, without the help of player 1, he can protect his payoff of 25, while player 3 receives more in the counter-objection than in the objection. Similarly, objection  $[0, 27, 23]$  of player 2 against  $[75, 25, 0]$  can be counter-objectioned by  $[75, 0, 25]$ . So, if a proposal  $[75, 25, 0]$  arises during the bargaining process, it is probable that it will be selected as final payoff. Any objection, by either player 1 or player 2, can be countered by the other. On the other hand, a proposal like  $[80, 20, 0]$  is unstable. Player 2 can object that he and player 3 will get more in  $[0, 21, 29]$ . Player 1 has no counter-objection, because he cannot keep his 80 while offering player 3 at least 29

Thus, in addition to all undominated payoffs (the core), the bargaining set also contains all payoffs against which there exists objections, providing they can be met by counter-objections. The bargaining set for this example consists of the four points

$$\begin{aligned}&[ 0, 0, 0] \\&[75, 0, 25] \\&[75, 25, 0] \\&[ 0, 25, 25] \quad \square\end{aligned}$$

The bargaining set is never empty. It always contains the core. For more details, consult OWEN (1968, 1982) or AUMANN and MASCHLER (1964).

In 1965, DAVIS and MASCHLER defined the kernel of a game, a subset of the bargaining set. In 1969, SCHMEIDLER introduced the nucleolus, a unique payoff, included in the kernel. It is defined as the allocation that minimises successively the largest coalitional excesses

$$e(\alpha, S) = v(S) - \sum_{i \in S} \alpha_i$$

The excess is the difference between a payoff a coalition can achieve and the proposed allocation. Hence it measures the amount ("the size of the complaint") by which coalition  $S$  as a group falls short of its potential  $v(S)$  in allocation  $\alpha$ . If the excess is positive, the payoff is outside the core (and so the nucleolus exists even when the core is empty). If the excess is negative, the proposed allocation is acceptable, but the coalition nevertheless has interest in obtaining the smallest possible  $e(\alpha, S)$ . The nucleolus is the imputation that minimises (lexicographically) the maximal excess. Since it is as far away as possible of the rationality conditions, it lies in the middle of the core. It is computed by solving a finite sequence of linear programs. Variants of the nucleolus, like the proportional and the disruptive nucleolus, are surveyed among others in LEMAIRE (1983). The proportional nucleolus, for instance, results when the excesses are defined as

$$e(\alpha, S) = \left[ v(S) - \sum_{i \in S} \alpha_i \right] / v(S)$$

Since it consists of a single point, the nucleolus (also called the lexicographic center) provides an alternative to the Shapley value. The Shapley value has been subjected to some criticisms, mainly focussing on the additivity axiom and the fact that people joining a coalition receive their full admission value.

**Example 3.** (ASTIN money). The Shapley value, computed in Section 4, is

$$[51,750; \quad 25,875; \quad 12,375]$$

It awards an interest of 11.5% to ASTIN and I.A.A., and 16.5% to A.A.Br. This allocation is much too generous towards A.A.Br.'s Treasurer, who takes a great advantage from the fact that he is essential to reach the 3-million mark. His admission value is extremely high (in proportion to the funds supplied) when he comes in last. The nucleolus is

$$[52,687.5; \quad 24,937.5; \quad 12,375]$$

or, in percentages

$$[11.71; \quad 11.08; \quad 16.5]$$

It recognises the better bargaining position of ASTIN versus I A.A., but still favours A.A.Br. Both the Shapley value and the nucleolus, defined in an additive way, fail in this multiplicative problem. The proportional nucleolus suggests

$$[54,000; \quad 27,000, \quad 9,000]$$

or, in percentages,

$$[12; \quad 12; \quad 12],$$

thereby justifying common practice.  $\square$

Only the case of the two-person games without transferable utilities has been reviewed in Section 5. A book by ROTH (1980) is devoted entirely to this case. It provides a thorough analysis of Nash's and Kalai-Smorodinsky's solutions. The generalisation of those models to the  $n$ -person case has proved to be very difficult. In the two-person case, the disagreement point is well defined: if the players don't agree, they are left alone. In the  $n$ -person case, if a general agreement in the grand coalition cannot be reached, sub-coalitions may form. Also, some players may wish to explore other avenues, like possible business partners outside the closed circle of the  $n$  players. This is an objection against modeling market situations as non-transferable  $n$ -person games. Such games ignore external opportunities, such as competitive outside elements. See SHAPLEY (1964) and LEMAIRE (1974, 1979) for definitions of values in the  $n$ -player case.

Though somewhat dated by now, the book by LUCE and RAIFFA (1957) is still an excellent introduction to game theory and utility theory. It provides an insightful critical analysis of the most important concepts. An excellent book that surveys recent developments is OWEN (1968, 1982, especially the second edition). A booklet edited by LUCAS (1981) provides an interesting, simple, abundantly illustrated analysis of the basics of cooperative and non-cooperative game theory. Finally, the proceedings of a conference on applied game theory [BRAMS, SCHOTTER, SCHWODIAUER (1979)] provide a fascinating overview (from a strategic analysis of the Bible to the mating of crabs) of applications of the theory.

## 7 CONCLUSIONS

Game theory solutions have been effectively implemented in numerous situations. A few of those applications are

- allocating taxes among the divisions of McDonnell-Douglas Corporation
- subdividing renting costs of WATS telephone lines at Cornell University
- allocating tree logs after transportation between the Finnish pulp and paper companies
- sharing maintenance costs of the Houston medical library
- financing large water resource development projects in Tennessee
- sharing construction costs of multipurpose reservoirs in the United States



- subdividing costs of building an 80-kilometer water supply tunnel in Sweden
- setting landing fees at Birmingham Airport
- allotting water among agricultural communities in Japan
- subsidising public transportation in Bogota

Cooperative game theory deals with competition, cooperation, conflicts, negotiations, coalition formation, allocation of profits. Consequently one would expect numerous applications of the theory in insurance, where competitive and conflicting situations abound. It has definitely not been the case. The first article mentioning game theory in the *ASTIN Bulletin* was authored by BORCH (1960a). In subsequent papers, BORCH (1960b, 1963) progressively developed his celebrated risk exchange model, which in fact is an  $n$ -person cooperative game without transferable utilities. This model has further been developed by in the 1970s by Lemaire and several of his students [BATON and LEMAIRE (1981a, 1981b), BRIEGLEB and LEMAIRE (1982), LEMAIRE (1977, 1979)]. The *ASTIN Bulletin* has yet to find a third author attracted by game theory! It is hoped that this survey paper will contribute to disseminate some knowledge about the situations game theory models, so that the risk exchange model will not stand for a long time as its lone actuarial application.

## REFERENCES

- AUMANN, R and MASCHLER, M (1964) The Bargaining Set for Cooperative Games *Advances in Game Theory*, 443–476 *Annals of Math Studies* 52, Princeton University Press
- BATON, B and LEMAIRE, J (1981a) The Core of a Reinsurance Market *ASTIN Bulletin* 12, 57–71
- BATON, B and LEMAIRE, J (1981b) The Bargaining Set of a Reinsurance Market *ASTIN Bulletin* 12, 101–114
- BORCH, K (1960a) Reciprocal Reinsurance Treaties *ASTIN Bulletin* 1, 171–191
- BORCH, K (1960b) Reciprocal Reinsurance Treaties Seen as a Two-person Cooperative Game *Skandinavisk Aktuarietidskrift* 43, 29–58
- BORCH, K (1962) Application of Game Theory to Some Problems in Automobile Insurance *ASTIN Bulletin* 2, 208–221
- BORCH, K (1963) Recent Developments in Economic Theory and Their Application to Insurance *ASTIN Bulletin* 2, 322–341
- BRAMS, S, SCHOTTER, A and SCHWODIAUER, G, Eds (1979) *Applied Game Theory IHS-studies* 1, Physica-Verlag, 1979
- BRIEGLEB, D and LEMAIRE, J (1982) Calcul des primes et marchandage *ASTIN Bulletin* 13, 115–131
- DAVIS, M and MASCHLER, M (1965) The Kernel of a Cooperative Game *Naval Research Logistics Quarterly* 12, 223–259
- KALAI, E and SMORODINSKY, M (1975) Other Solutions to the Nash Bargaining Problem *Econometrica* 43, 513–518
- LEMAIRE, J (1974) A New Concept of Value for Games Without Transferable Utilities *International J. of Game Theory* 1, 205–213
- LEMAIRE, J (1977) Echange de risques entre assureurs et théorie des jeux *ASTIN Bulletin* 9, 155–180
- LEMAIRE, J (1979) A Non Symmetrical Value For Games Without Transferable Utilities Application to Reinsurance *ASTIN Bulletin* 10, 195–214
- LEMAIRE, J (1983) An Application of Game Theory Cost Allocation *ASTIN Bulletin* 14, 61–81

- LUCAS, W (1981) *Game Theory and Its Applications* *Proceedings of Symposia in Applied Maths, American Math Society Vol 24*, Providence, Rhode Island
- LUCE, R and RAIFFA, H (1957) *Games and Decisions* John Wiley, New York
- NASH, J (1950) The Bargaining Problem *Econometrica* **18**, 155–162
- OWEN, G (1968, 1982) *Game Theory* W Saunders, Philadelphia (first edition) Academic Press, New York (second edition)
- ROTH, A (1980) *Axiomatic Models of Bargaining* Springer Verlag, Berlin
- SHAPLEY, L (1953) A Value for  $n$ -person Games Contributions to the Theory of Games Vol II, 307–317 *Annals of Mathematics Studies* **28**, Princeton University Press
- SHAPLEY, L. (1964) Utility Comparisons and the Theory of Games *La décision, CNRS*, 251–263
- SCHMEIDLER, D (1969) The Nucleolus of a Characteristic Function Game *SIAM J. Applied Maths* , 1163–1170
- VON NEUMANN, J and MORGENSTERN, O (1944) *Theory of Games and Economic Behaviour* Princeton University Press

JEAN LEMAIRE

*Department of Insurance and Risk Management, Wharton School,  
3641 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6218.*

## ARTICLES

### COMBINING QUOTA-SHARE AND EXCESS OF LOSS TREATIES ON THE REINSURANCE OF $n$ INDEPENDENT RISKS

BY LOURDES CENTENO and ONOFRE SIMÕES

*ISEG, Universidade Técnica de Lisboa, Portugal*

#### ABSTRACT

In this paper, we seek to find the optimal retentions for an insurance company which intends to reinsure each of  $n$  risks belonging to its portfolio, by means of a pure quota-share treaty, a pure excess of loss treaty or any combination of the two. The criterion chosen to the selection of the optimal programme is the maximization of the adjustment coefficient, attending to the relationship existing between this coefficient and Lundberg's upper bound of the ruin probability.

#### 1. INTRODUCTION

Suppose that an insurance company seeks reinsurance for  $n$  independent risks (by a risk we mean a single policy or a group of policies—so we could speak of  $n$  independent lines of insurance), and has a choice between a pure quota-share treaty, an excess of loss treaty or any combination of the two, for any of the risks. The way this combination operates is as follows: first the quota share contract will apply, so that the insurer shall remain responsible for no more than its share—established by the contract—of any claim that may occur for that risk; afterwards, the excess of loss contract applies, so that, by no means, shall the insurer (of course considering only that part for which it remains responsible after the quota-share contract) pay more than a certain fixed amount of any claim that takes place.

The problem consists of determining the optimal retention limits for each risk, in each of the two forms of reinsurance. "Optimal" in the sense those limits maximize the adjustment coefficient and, therefore, minimize the upper bound to the ruin probability, supplied by Lundberg's inequality. This same criterion was also adopted by WATERS (1979) and CENTENO (1986) and, in a certain way, this work may be considered as a generalization of their results. Although this criterion does not by any means have to minimize the (analytically uncalculable) ruin probability, it is a good criterion if one wishes to give analytical results.

Surplus and stop loss treaties are not considered in this paper. WATERS (1983), derives sufficient conditions for the adjustment coefficient to be uni-model, for stop loss reinsurance.

For each  $i = 1, 2, \dots, n$ , let  $a_i$ , be the decision variable representing the quota-share retention on risk  $i$ ;  $M_i$ , the decision variable representing the excess of loss retention limit on risk  $i$ ;  $Y_i = \sum_{j=0}^{N_i} X_{ij}$ , with  $X_{i0} \equiv 0$ , the insurer's aggregate gross (of reinsurance) claims on risk  $i$ , in some fixed time interval, where  $N_i$  is the number of claims and  $\{X_{ij}\}_{j=1, \dots, N_i}$  are the individual claims;  $P_i$  the insurer's gross (of expenses and reinsurance) premium income with respect to risk  $i$  and  $e_i P_i$ , the amount used to cover the insurer's expenses with respect to the same risk.

After a combination of a quota-share with an excess of loss treaty the insurer will retain, from risk  $i$ ,  $Y_i(a_i, M_i) = \sum_{j=1}^{N_i} \min \{a_i X_{ij}, M_i\}$ , ( $i = 1, \dots, n$ ).

The choice of uniform  $a_1 = \dots = a_n$  and  $M_1 = \dots = M_n$ , which is generally made in practice, has been dealt with in CENTENO (1986). In this paper, therefore, retention limits which can, for instance, be set differently for portfolios of different classes of business are also dealt with

Let  $P_n(a_i, M_i)$  be the total reinsurance premium paid by the insurer, in respect to risk  $i$  (it is, naturally, the summation of the quota-share and excess of loss reinsurance premiums).

The problem which is to be solved is, then,

$$\begin{array}{ll} \text{Maximize} & R(\underline{a}, \underline{M}) \\ \text{sub. to:} & 0 \leq a_i \leq 1 \\ & M_i \geq 0 \\ & (i = 1, 2, \dots, n), \end{array}$$

where  $R(\underline{a}, \underline{M})$ , is the adjustment coefficient, defined, as it is known, as the unique positive root of

$$(1) \quad E \left[ \exp \left\{ R \sum_{i=1}^n Y_i(a_i, M_i) - R \sum_{i=1}^n [P_i(1 - e_i) - P_n(a_i, M_i)] \right\} \right] = 1.$$

Note that  $R(\underline{a}, \underline{M})$  is the adjustment coefficient (see BEARD, PENTIKAINEN and PESONEN (1984), p. 363) after taking account of the reinsurance arrangement.

## 2. ASSUMPTIONS AND PREMIMINARIES

$A_1$ :  $Y_i (i = 1, 2, \dots, n)$  are independent random variables;

For each  $i$  ( $i = 1, 2, \dots, n$ ):

$A_2$ :  $N_i (i = 1, 2, \dots, n)$  is a Poisson random variable with parameter  $\lambda_i$ ;

$A_3$ :  $\{X_{ij}\}_{j=1,2,\dots,N_i}$  are i.i.d. non-negative random variables, independent of  $N_i$ , and with common distribution function  $F_i$  such that

$$\begin{cases} F_i(x) = 0, & x \leq x_{i0} \\ 0 < F_i(x) < 1, & x > x_{i0}, \end{cases}$$

for some  $x_{i0} \geq 0$ ,

$A_4$ :  $\frac{d}{dx} F_i(x)$  exists and it is continuous everywhere;

$A_5$ : The m.g.f. of the random variables  $X_{ij}$ , exists in the  $(-\infty, Q_i]$  interval, for  $0 < Q_i \leq +\infty$  and

$$\lim_{t_i \rightarrow Q_i} E[e^{t_i X_{ij}}] = +\infty;$$

$A_6$ : The quota-share reinsurance premium is

$$(1 - a_i) P_i - c_i (1 - a_i) P_i = (1 - c_i) (1 - a_i) P_i,$$

where  $c_i (1 - a_i) P_i$ ,  $0 < c_i < 1$ , is the habitual commission paid by the quota-share reinsurer;

$A_7$ : The excess of loss reinsurance premium, which we denote  $P_i(a_i, M_i)$ , is calculated according to the expected value principle, i.e.,

$$P_i(a_i, M_i) = (1 + \alpha_i) \lambda_i \int_{M_i/a_i}^{+\infty} (a_i x - M_i) dF_i(x)$$

with  $\alpha_i > 0$ .

$A_8$ :  $e_i > c_i$ ;

$A_9$ :  $(1 - c_i) P_i - \lambda_i E[X_i] > 0$ , where  $E[X_i]$  denotes the expected value of  $X_{ij}$ ,  $j = 1, 2, \dots, N_i$ ;

$A_{10}$ :  $(1 - e_i) P_i < (1 + \alpha_i) \lambda_i E[X_i]$ ;

Finally, we assume that

$$A_{11}: \sum_{i=1}^n [(1 - e_i) P_i - \lambda_i E[X_i]] > 0.$$

From  $A_2$  and  $A_3$  it follows that  $Y_i$  and  $Y_i(a_i, M_i)$  have compound Poisson distributions. From  $A_6$  and  $A_7$  we can say that

$$(2) \quad P_u(a_i, M_i) = (1 - c_i) (1 - a_i) P_i + (1 + \alpha_i) \lambda_i \int_{M_i/a_i}^{+\infty} (a_i x - M_i) dF_i(x).$$

Assumption  $A_8$  is somewhat restrictive, but without it the insurer could reinsure the whole risk through a quota-share arrangement with a certain profit. The same applies to  $A_{10}$ , but with respect to the excess of loss reinsurance treaty  $A_9$  implies that the loading on the quota-share reinsurance premium is positive. At last,  $A_{11}$  assures the existence of a margin, necessary to cover eventual deviations from the expected losses, and also to pay the reinsurance costs.

Under assumptions  $A_1$ ,  $A_2$  and  $A_3$ ,  $R(\underline{a}, \underline{M})$  is the only positive root of

$$(3) \quad G(R; \underline{a}, \underline{M}) = 0,$$

where

$$(4) \quad G(R; \underline{a}, \underline{M}) = \sum_{i=1}^n \lambda_i \left[ \int_0^{M_i/a_i} e^{Ra_i x} dF_i(x) + e^{RM_i} [1 - F_i(M_i/a_i)] - 1 \right] - R \sum_{i=1}^n [(1 - e_i) P_i - P_n(a_i, M_i)]$$

(See BEARD, PENTIKAINEN and PESONEN (1984), p. 363, for the equivalence of (4) and (1).) Let  $E[W(\underline{a}, \underline{M})]$  denote the insurer's expected net profit, after reinsurance and expenses, i.e.,

$$(5) \quad E[W(\underline{a}, \underline{M})] = \sum_{i=1}^n \left[ (c_i - e_i) P_i + a_i [(1 - c_i) P_i - \lambda_i E[X_i]] - \lambda_i \alpha_i \int_{M_i/a_i}^{+\infty} (a_i x - M_i) dF_i(x) \right],$$

and let us define

$$\Upsilon = \{(\underline{a}, \underline{M}) : 0 \leq a_i \leq 1, M_i \geq 0 \text{ and } E[W(\underline{a}, \underline{M})] > 0\},$$

and

$$\Gamma = \{\underline{a} : 0 \leq a_i \leq 1, i = 1, 2, \dots, n \text{ and there exists at least one } \underline{M} \text{ such that } E[W(\underline{a}, \underline{M})] > 0\}.$$

Since

$$\frac{\partial}{\partial M_i} E[W(\underline{a}, \underline{M})] = \lambda_i \alpha_i (1 - F_i(M_i/a_i))$$

is non-negative, we can say that for fixed  $\underline{a}$ , the expected net profit will be maximum when  $M_i = +\infty$  ( $i = 1, 2, \dots, n$ ). Hence it is possible to specify  $\Gamma$  as being

$$(6) \quad \Gamma = \left\{ \underline{a} : \sum_{i=1}^n [(c_i - e_i) P_i + a_i [(1 - c_i) P_i - \lambda_i E[X_i]]] > 0 \right\}.$$

Let us denote  $\frac{\partial}{\partial R} G(R; \underline{a}, \underline{M})$  by  $D(R, \underline{a}, \underline{M})$  so that

$$(7) \quad D(R; \underline{a}, \underline{M}) = \sum_{i=1}^n \lambda_i \left[ \int_0^{M_i/a_i} a_i x e^{Ra_i x} dF_i(x) + M_i e^{RM_i} (1 - F_i(M_i/a_i)) \right] - \sum_{i=1}^n [(1 - e_i) P_i - P_{ii}(a_i, M_i)],$$

with  $G(R; \underline{a}, \underline{M})$  defined by (4).

The following lemma discusses the existence of the adjustment coefficient.

**Lemma 1:**

- (i)  $R(\underline{a}, \underline{M})$  exists, if and only if  $(\underline{a}, \underline{M}) \in \mathcal{T}$ ,
- (ii) For any  $(\underline{a}, \underline{M}) \in \mathcal{T}$ ,  $D(R; \underline{a}, \underline{M})$  is positive at  $R = R(\underline{a}, \underline{M})$ .

**Proof:**

- (i) By  $A_5$ , it is clear that for fixed  $(\underline{a}, \underline{M})$ ,  $G$  is defined for all  $R \in (-\infty, Q)$ ,

where

$$Q = \min \{\zeta_i\}$$

and

$$\zeta_i = \begin{cases} +\infty, & \text{if } M_i/a_i < +\infty \\ \frac{Q_i}{a_i}, & \text{if } M_i/a_i = +\infty \end{cases}$$

$$(i = 1, 2, \dots, n).$$

The first aspect to be considered, is that  $R = 0$  is a trivial solution of equation (3);

Secondly, we have that

$$\frac{\partial^2}{\partial R^2} G(R, \underline{a}, \underline{M}) = \sum_{i=1}^n \lambda_i \left[ \int_0^{M_i/a_i} (a_i x)^2 e^{Ra_i x} dF_i(x) + M_i^2 e^{RM_i} (1 - F_i(M_i/a_i)) \right],$$

it is non-negative,  $\forall (\underline{a}, \underline{M})$ , which means that  $G(R; \underline{a}, \underline{M})$  is a convex function of  $R$ ,

Third,

$$\begin{aligned} \lim_{R \rightarrow \mathcal{Q}} G(R; \underline{a}, \underline{M}) &= \sum_{i=1}^n \lim_{R \rightarrow \mathcal{Q}} \left\{ \lambda_i \int_0^{M_i/a_i} [e^{Ra_i x} - R(1+\alpha_i)a_i x] dF_i(x) + \right. \\ &\quad + \lambda_i [e^{RM_i} - R(1+\alpha_i)M_i] (1 - F_i(M_i/a_i)) - \lambda_i - \\ &\quad \left. - R[(c_i - e_i)P_i + a_i[(1 - c_i)P_i - (1+\alpha_i)\lambda_i E[X_i]]] \right\} \\ &= +\infty, \end{aligned}$$

by assumptions  $A_8$  and  $A_9$ .

Hence, as  $G(R; \underline{a}, \underline{M})$  equals zero when  $R$  is null,  $G(R; \underline{a}, \underline{M})$  is a convex function of  $R$ , and  $G(R; \underline{a}, \underline{M})$  tends to infinity when  $R$  tends to  $\mathcal{Q}$ , then, it will only exist such an  $R = \hat{R}(\underline{a}, \underline{M}) > 0$  which turns  $G(R, \underline{a}, \underline{M})$  to be null again, if and only if,

$$\left. \frac{\partial}{\partial R} G(R; \underline{a}, \underline{M}) \right|_{R=0} < 0.$$

To finish the prof, we only have to notice that

$$\left. \frac{\partial}{\partial R} G(R; \underline{a}, \underline{M}) \right|_{R=0} < 0 \Leftrightarrow E[W(\underline{a}, \underline{M})] > 0.$$

(ii) Immediate, given the proof of (i).  $\diamond$

The following lemma will be useful to the solution to our problem.

**Lemma 2:** For any  $\underline{a} \in \Gamma$  there exists a unique  $(\underline{a}, \underline{M}) \in \Upsilon$ , let it be  $(\underline{a}, \hat{\underline{M}})$ , such that

$$R(\underline{a}, \hat{\underline{M}}) = \frac{\ln(1+\alpha_i)}{\hat{M}_i}, \quad i = 1, 2, \dots, n.$$

**Proof:** Let us consider the set of points  $\underline{M}$  such that

$$\frac{\ln(1+\alpha_1)}{M_1} = \frac{\ln(1+\alpha_2)}{M_2} = \dots = \frac{\ln(1+\alpha_n)}{M_n} = \frac{1}{\tilde{M}}, \quad \tilde{M} > 0$$

and let us define

$$H(\underline{a}, \tilde{M}) = \tilde{M} G \left( \frac{1}{\tilde{M}}; \underline{a}, \tilde{M} \ln(1+\alpha_1), \dots, \tilde{M} \ln(1+\alpha_n) \right),$$



which is to say

$$\begin{aligned}
 H(\underline{a}, \tilde{M}) = & \sum_{i=1}^n \left\{ \lambda_i \int_0^{\tilde{M} \frac{\ln(1+\alpha_i)}{a_i}} \tilde{M} e^{\frac{1}{\tilde{M}} a_i x} dF_i(x) + \right. \\
 & + \lambda_i \tilde{M} (1+\alpha_i) \left[ 1 - F_i \left( \tilde{M} \frac{\ln(1+\alpha_i)}{a_i} \right) \right] - \tilde{M} \lambda_i - \\
 & - [(1-e_i) P_i - (1-c_i) (1-a_i) P_i] + \\
 & \left. + (1+\alpha_i) \lambda_i \int_{\tilde{M} \frac{\ln(1+\alpha_i)}{a_i}}^{+\infty} (a_i x - \tilde{M} \ln(1+\alpha_i)) dF_i(x) \right\}.
 \end{aligned}$$

Then

$$\begin{aligned}
 10 \lim_{\tilde{M} \rightarrow 0^+} H(\underline{a}, \tilde{M}) = & \sum_{i=1}^n - [(1-e_i) P_i - (1-c_i) (1-a_i) P_i - \\
 & - (1+\alpha_i) \lambda_i a_i E[X_i]] > 0,
 \end{aligned}$$

using  $A_8$  and  $A_{10}$ ,

$$\begin{aligned}
 2) \lim_{\tilde{M} \rightarrow +\infty} H(\underline{a}, \tilde{M}) = & - \sum_{i=1}^n \{ [(c_i - e_i) P_i + a_i [(1-c_i) P_i - \\
 & - \lambda_i E[X_i]] \} < 0;
 \end{aligned}$$

3) Differentiating  $H(\underline{a}, \tilde{M})$  twice with respect to  $\tilde{M}$  we obtain (see, for example, COURANT and JOHN (1974), p 77)

$$\frac{\partial^2}{\partial \tilde{M}^2} H(\underline{a}, \tilde{M}) = \sum_{i=1}^n \lambda_i \int_0^{\tilde{M} \frac{\ln(1+\alpha_i)}{a_i}} e^{\frac{1}{\tilde{M}} a_i x} \frac{(a_i x)^2}{\tilde{M}^3} dF_i(x) \geq 0.$$

Hence, for each  $\underline{a} \in \Gamma$  there exists a unique positive  $\tilde{M} = \tilde{M}(\underline{a})$  such that  $H(\underline{a}, \tilde{M}) = 0$  and it is clear from the definition of  $H(\underline{a}, \tilde{M})$  that

$$G \left( \frac{1}{\tilde{M}}; \underline{a}, \hat{M} \right) = 0,$$

where

$$\hat{M} = (\hat{M}_1, \hat{M}_2, \dots, \hat{M}_n) = (\tilde{M} \ln(1+\alpha_1), \tilde{M} \ln(1+\alpha_2), \dots, \tilde{M} \ln(1+\alpha_n)) \quad \diamond$$

This lemma implies that if we define

$$(8) \quad \hat{G}(\hat{R}; \underline{a}) = G(\hat{R}; \underline{a}, \hat{R}^{-1} \ln(1+\alpha_1), \hat{R}^{-1} \ln(1+\alpha_2), \dots, \hat{R}^{-1} \ln(1+\alpha_n)),$$

then  $\hat{G}(\hat{R}; \underline{a})$  has a unique positive root for each  $a \in \Gamma$ . Let us denote it  $\hat{R}(\underline{a})$ . It can be proved, using the Implicit Function Theorem (see for example COURANT and JOHN (1974), pp. 221-223), Part (2) of Lemma 1 and  $A_4$ , that  $R(\underline{a}, \underline{M})$ , for  $(\underline{a}, \underline{M}) \in \Upsilon$ , and  $\hat{R}(\underline{a})$ , for  $\underline{a} \in \Gamma$ , are twice differentiable.

### 3. THE SOLUTION TO THE PROBLEM

The following result provides the solution to our problem.

#### Result 1:

- (i) For a fixed value  $\underline{a} \in \Gamma$ , with  $a_i \neq 0$ ,  $\forall i = 1, 2, \dots, n$ ,  $R(\underline{a}, \underline{M})$  is a unimodal function of  $\underline{M}$ , and for any  $\underline{a} \in \Gamma$  its maximum value is  $\hat{R}(\underline{a})$
- (ii)  $\hat{R}(\underline{a})$  is a unimodal function of  $\underline{a}$ , for  $\underline{a} \in \Gamma$  and, at the point where it attains its maximum:

$$a) \quad a_i = 1 \text{ if and only if } \frac{\partial}{\partial a_i} \hat{R}(\underline{a})(a_i = 1) \geq 0,$$

or

$$b) \quad a_i \text{ is such that } \frac{\partial}{\partial a_i} \hat{R}(\underline{a}) = 0, \text{ if and only if}$$

$$\frac{\partial}{\partial a_i} \hat{R}(\underline{a})(a_i = 1) \leq 0, \quad i = 1, 2, \dots, n.$$

#### Proof:

- (i) The equation defining  $R(\underline{a}, \underline{M})$  for all  $(\underline{a}, \underline{M}) \in \Gamma$  is

$$(9) \quad G(R; \underline{a}, \underline{M}) = 0,$$

with  $G(R; \underline{a}, \underline{M})$  given by (4). Differentiating (9) with respect to  $M_i$  it can be

seen that  $\frac{\partial}{\partial M_i} R(\underline{a}, \underline{M}) = 0$  if and only if (using the Implicit Function

Theorem)

$$Re^{RM_i}(1 - F_i(M_i/a_i)) = R(1 + \alpha)(1 - F_i(M_i/a_i)).$$

So, using Lemma 2 we can say that for a fixed value of  $\underline{a} \in \Gamma$ , with  $a_i \neq 0$ ,  $\forall i = 1, 2, \dots, n$ , the only turning point of  $R(\underline{a}, \underline{M})$  is such that

$$(10) \quad M_i = R^{-1} \ln(1 + \alpha), \quad i = 1, 2, \dots, n.$$

Differentiating (9) twice with respect to  $M_i$  (using again the Implicit Function Theorem and (10)) we get

$$(11) \quad \left. \frac{\partial^2}{\partial M_i^2} R(\underline{a}, \underline{M}) \right|_{M_i = R^{-1} \ln(1 + \alpha_i)} = - \left. \frac{\lambda_i R^2 e^{RM_i} [1 - F_i(M_i/a_i)]}{D(R; \underline{a}, \underline{M})} \right|_{M_i = R^{-1} \ln(1 + \alpha_i)}$$

with  $D(R; \underline{a}, \underline{M})$  given by (7). We can see that each side of equation (11) is negative since  $D(R; \underline{a}, \underline{M})$  is positive by Lemma 1 (ii).

On the other hand,

$$\left. \frac{\partial^2}{\partial M_i \partial M_j} R(\underline{a}, \underline{M}) \right|_{\substack{M_i = R^{-1} \ln(1 + \alpha_i) \\ M_j = R^{-1} \ln(1 + \alpha_j)}} = 0, \quad j \neq i.$$

Hence we can conclude that for a fixed value  $\underline{a} \in \Gamma$  with  $a_i \neq 0$ ,  $\forall i = 1, 2, \dots, n$ ,  $R(\underline{a}, \underline{M})$  is a unimodal function of  $\underline{M}$ .

If  $\underline{a} \in \Gamma$  and  $a_k = 0$  for some  $k = 1, 2, \dots, n$ , then of course any value for the excess of loss retention limit of risk  $k$ , including  $M_k = R^{-1} \ln(1 + \alpha_k)$ , will provide the same value for the adjustment coefficient.

Then the maximum of  $R(\underline{a}, \underline{M})$  is attained at the point  $(\underline{a}, \underline{M})$  which is the unique point satisfying  $G(R; \underline{a}, \underline{M}) = 0$  and  $M_i = R^{-1} \ln(1 + \alpha_i)$ ,  $i = 1, 2, \dots, n$ , i.e., for a fixed  $\underline{a} \in \Gamma$ , the maximum of  $R(\underline{a}, \underline{M})$  is  $\hat{R}$  where  $\hat{R} = \hat{R}(\underline{a})$  is the only positive root of  $\hat{G}(\hat{R}; \underline{a}) = 0$ , with  $\hat{G}(\hat{R}; \underline{a})$  given by (8).

(ii) Differentiating

$$(12) \quad \hat{G}(\hat{R}; \underline{a}) = 0$$

with respect to  $a_i$ , we obtain

$$\begin{aligned} \frac{\partial}{\partial a_i} \hat{R}(\underline{a}) &= \frac{\hat{R}}{\hat{D}(\hat{R}; \underline{a})} \left[ (1 - c_i) P_i - (1 + \alpha_i) \lambda_i \int_{\frac{\ln(1 + \alpha_i)}{\hat{R} a_i}}^{+\infty} x dF_i(x) - \right. \\ &\quad \left. - \lambda_i \int_0^{\frac{\ln(1 + \alpha_i)}{\hat{R} a_i}} x e^{\hat{R} a_i x} dF_i(x) \right], \end{aligned}$$

where

$$\hat{D}(\hat{R}; \underline{a}) = D(\hat{R}; \underline{a}, \hat{R}^{-1} \ln(1 + \alpha_1), \hat{R}^{-1} \ln(1 + \alpha_2), \dots, \hat{R}^{-1} \ln(1 + \alpha_n)).$$

So,

$$(13) \quad \frac{\partial}{\partial a_i} \hat{R}(\underline{a}) = 0$$

if and only if

$$(14) \quad (1 - c_i) P_i = \lambda_i \int_0^{\frac{\ln(1 + \alpha_i)}{\hat{R} a_i}} x e^{\hat{R} a_i x} dF_i(x) + (1 + \alpha_i) \lambda_i \int_{\frac{\ln(1 + \alpha_i)}{\hat{R} a_i}}^{+\infty} x dF_i(x)$$

Differentiating (13) with respect to  $a_i$ , and using (14), we obtain

$$\frac{\partial^2}{\partial a_i^2} \hat{R}(\underline{a}) \Big|_{\frac{\partial}{\partial a_i} \hat{R}(\underline{a}) = 0} = - \frac{\lambda_i \int_0^{\frac{\ln(1+\alpha_i)}{Ra_i}} \hat{R}^2 x^2 e^{\hat{R}a_i x} dF_i(x)}{\hat{D}(\hat{R}; \underline{a})} \Big|_{\frac{\partial}{\partial a_i} \hat{R}(\underline{a}) = 0}$$

and

$$\left[ \frac{\partial^2}{\partial a_i a_j} \hat{R}(\underline{a}) \right] \left( \frac{\partial}{\partial a_i} \hat{R}(\underline{a}) = 0, \frac{\partial}{\partial a_j} \hat{R}(\underline{a}) = 0 \right) = 0, \text{ if } i \neq j.$$

This implies that there exists at most a point  $\underline{a} \in \Gamma$  such that (14) holds for  $i = 1, 2, \dots, n$ .

Noticing that

$$\lim_{a_i \rightarrow 0^+} \frac{\partial}{\partial a_i} \hat{R}(\underline{a}) = [(1-c_i)P_i - \lambda_i E[X_i]] \lim_{a_i \rightarrow 0^+} \frac{\hat{R}}{\hat{D}(\hat{R}, \underline{a})},$$

with  $\underline{a} \in \Gamma$ , is positive by  $A_9$  and Lemma 1 (ii), the proof is finished.  $\diamond$

To summarize, we can now conclude that the optimum programme of reinsurance, when a company is to reinsure  $n$  independent risks by a combination of the quota share and excess of loss forms of reinsurance, is the point  $(\underline{a}, \underline{M})$  which fulfils the following set of conditions:

$$\left\{ \begin{array}{l} \rightarrow M_i = \frac{\ln(1+\alpha_i)}{R}, \quad (i = 1, 2, \dots, n) \\ \rightarrow a_i: (1-c_i)P_i = \lambda_i \int_0^{\frac{\ln(1+\alpha_i)}{Ra_i}} x e^{Ra_i x} dF_i(x) + (1+\alpha_i)\lambda_i \int_{\frac{\ln(1+\alpha_i)}{Ra_i}}^{+\infty} x dF_i(x), \\ \quad \text{if } \left[ \frac{\partial R}{\partial a_i} \right] < 0 \text{ when } a_i = 1 \\ \rightarrow \text{or } a_i = 1, \text{ if } \left[ \frac{\partial R}{\partial a_i} \right] \geq 0 \text{ when } a_i = 1 \\ \quad (i = 1, 2, \dots, n) \\ \rightarrow G(R; \underline{a}, \underline{M}) = 0 \end{array} \right.$$

**Corollary 1:** If  $(1-c_i)P_i \geq \lambda_i(1+\alpha_i)E[X_i]$  for some  $i$  ( $i = 1, 2, \dots, n$ ), then the optimal arrangement is such that  $a_i = 1$ .

**Proof:** We only have to notice that in this case

$$\begin{aligned} (1 - c_i) P_i - \lambda_i \int_0^{R^{-1} \ln(1 + \alpha_i)} x e^{R x} dF_i(x) - (1 + \alpha_i) \lambda_i \int_{R^{-1} \ln(1 + \alpha_i)}^{\infty} x dF_i(x) &\geq \\ &\geq \lambda_i \int_0^{R^{-1} \ln(1 + \alpha_i)} x (1 + \alpha_i - e^{R x}) \geq 0. \quad \diamond \end{aligned}$$

Note that we can regard the quota-share reinsurance premium for risk  $i$  (see  $A_{\hat{\alpha}_i}$ ) as being calculated using the expected value principle with loading factor  $\hat{\alpha}_i$ , where

$$\hat{\alpha}_i = [(1 - c_i) - \lambda_i E[X_i]] / [\lambda_i E[X_i]].$$

Then, Corollary 1 implies that if  $\hat{\alpha}_i \geq \alpha_i$ , i.e. if quota-share is, in the obvious sense more expensive than excess of loss reinsurance, then excess of loss reinsurance is optimal. Excess of loss reinsurance was already proved to be the optimal form of reinsurance (see GERBER 1979), p 129), in the sense that it maximizes the adjustment coefficient, under the assumption that the loading coefficient is the same for the insurer and the reinsurer (which is not the case in our paper).

When the number of risks,  $n$ , is greater than one, the solution found for the problem, may not be the solution that we would obtain if the risks were considered separately. In other words, if we regard as optimal a set of retention limits that maximizes the adjustment coefficient, then what is optimal when each risk is considered individually may not be optimal when the risks are considered together, as we will see next.

In the result that follows,  $R(a_i, M_i)$  ( $i = 1, 2, \dots, n$ ) is, for fixed  $(a_i, M_i)$ , the adjustment coefficient associated to risk  $i$ , when this is considered on its own, defined as the unique positive root of

$$(15) \quad G_i(R_i; a_i, M_i) = 0,$$

where

$$(16) \quad G_i(R_i; a_i, M_i) = \lambda_i \left[ \int_0^{M_i/a_i} e^{R_i a_i x} dF_i(x) + e^{R_i M_i} [1 - F_i(M_i/a_i)] - 1 \right] - R_i [(1 - e_i) P_i - P_{ii}(a_i, M_i)]$$

if such a root exists, or zero otherwise <sup>1</sup>

**Result 2:** For fixed  $(\underline{a}, \underline{M}) \in \Upsilon$  we have

$$\min_{i=1, \dots, n} \{R_i(a_i, M_i)\} \leq R(\underline{a}, \underline{M}) \leq \max_{i=1, \dots, n} \{R_i(a_i, M_i)\}.$$

<sup>1</sup> The need to redefine  $R_i(a_i, M_i)$  comes from the fact that  $E[W(\underline{a}, \underline{M})] > 0$  does not imply that  $E[W_i(a_i, M_i)] > 0$ , for all  $i = 1, 2, \dots, n$

**Proof:**

Let

$$(17) \quad \min_{i=1, \dots, n} \{R_i(a_i, M_i)\} = R_k(a_k, M_k)$$

and

$$(18) \quad \max_{i=1, \dots, n} \{R_i(a_i, M_i)\} = R_l(a_l, M_l).$$

Then, considering the definition of  $R_i(a_i, M_i)$ ,  $\forall i = 1, 2, \dots, n$ , we have that

$$0 \leq R_k(a_k, M_k) \leq R_l(a_l, M_l),$$

and, on the other hand, having in mind the proof of Lemma 1, we know that

$$(19) \quad \begin{cases} G_i(R_i; a_i, M_i) < 0 & \text{if } 0 < R_i < R_i(a_i, M_i) \\ G_i(R_i; a_i, M_i) > 0 & \text{if } R_i > R_i(a_i, M_i) \end{cases}$$

for  $i = 1, 2, \dots, n$ .

From (19) and attending to (17) and (18) we have that

$$(20) \quad \sum_{i=1}^n G_i(R_k(a_k, M_k); a_i, M_i) \leq 0,$$

being zero if and only if  $R_k(a_k, M_k) = R_l(a_l, M_l)$ . Similiary

$$(21) \quad \sum_{i=1}^n G_i(R_l(a_l, M_l); a_i, M_i) \geq 0,$$

being zero if and only if  $R_k(a_k, M_k) = R_l(a_l, M_l)$ .

Then the result follows immediately, since  $R(\underline{a}, \underline{M})$  for  $(\underline{a}, \underline{M}) \in \Upsilon$  is the unique positive root of

$$(22) \quad \sum_{i=1}^n G_i(R; a_i, M_i) = 0 \quad \diamond$$

**Corollary 2:** If  $R_i(a_i, M_i)$  achieves its maximum value at  $(a_i, M_i) = (\hat{a}_i, \hat{M}_i)$ ,  $i = 1, 2, \dots, n$ , and if  $R(\underline{a}, \underline{M})$  achieves its maximum value at  $(\underline{a}, \underline{M}) = (\hat{\underline{a}}, \hat{\underline{M}})$ , then

$$\min_{i=1, \dots, n} \{R_i(\hat{a}_i, \hat{M}_i)\} \leq R(\hat{\underline{a}}, \hat{\underline{M}}) \leq \max_{i=1, \dots, n} \{R_i(\hat{a}_i, \hat{M}_i)\}$$

**Proof:** Attending to Result 2, to the definition of  $(\hat{a}, \hat{M})$  and to the definition of  $(\hat{a}_i, \hat{M}_i)$ ,  $i = 1, 2, \dots, n$ , then

$$\min_{i=1, \dots, n} \{R_i(\hat{a}_i, \hat{M}_i)\} \leq R(\hat{a}, \hat{M}) \leq R(\hat{a}, \hat{M}),$$

and

$$R(\hat{a}, \hat{M}) \leq \max_{i=1, \dots, n} \{R_i(\hat{a}_i, \hat{M}_i)\} \leq \max_{i=1, \dots, n} \{R_i(\hat{a}_i, \hat{M}_i)\}$$

which finishes the proof.  $\diamond$

4. EXAMPLE

Let  $n = 2$  and

$$G_1(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ 1 - \frac{1}{4} e^{-4x}(x+4), & \text{if } x > 0, \end{cases}$$

which corresponds to a  $\gamma \left(2, \frac{1}{4}\right)$ , and

$$G_2(x) = \begin{cases} 0, & \text{if } x \leq 1 \\ 1 - e^{-3(x-1)}, & \text{if } x > 1, \end{cases}$$

which is an exponential.

Let  $\lambda_1 = 2$ ,  $\lambda_2 = 10$ ,  $P_1 = 27$ ,  $P_2 = 23.5$ ,  $e_1 = e_2 = .35$ ,  $U_1 = 30$  and  $U_2 = 15$ . The expected profit, before any reinsurance arrangement takes place, is 3.491(6) (1.55 from risk 1 and 1.941(6) from risk 2),  $R$  is .02849 and, therefore, the upper bound given by Lundberg's inequality for the ruin probability, is 0.2774. Considering the two risks separately the adjustment coefficients are  $R_1 = 0.01487$  and  $R_2 = 0.1864$ , giving then upper bounds for the ruin probabilities of 0.6401 and 0.0610, for risks 1 and 2 respectively.

The optimal reinsurance programme was calculated assuming different values for  $\alpha_1$  and setting  $\alpha_2 = .3$ ,  $c_1 = c_2 = .25$ . The results can be seen on Table 1. Analysing Table 1, the main aspect that seems evident is that, as long as  $\alpha_1$  increases, a similar evolution is presented by ratio  $M_1/a_1$ , that is to say, the excess of loss form of reinsurance becomes less and less attractive.

Table 2 gives the same kind of information as Table 1, when treating the two risks separately. Note that  $R_1 < R < R_2$ . One way of explaining this occurrence may be the following: when the reinsurance problem is solved taking the risks together, there is a sort of a transfer of part of the income produced for the "less dangerous" (and, therefore "less needed" of reinsurance) risks, to

TABLE 1

$\alpha_1$	Optimal Retentions	Expected Net Profit	Adjustment Coefficient	Upper Bound by Lundberg's Inequality
0.3	$a_1 = 00.77$ $a_2 = 01.00$	1.4986	0.04300	0.1444
	$M_1 = 06.10$ $M_2 = 06.10$			
0.4	$a_1 = 00.57$ $a_2 = 01.00$	1.4177	0.03919	0.1714
	$M_1 = 08.59$ $M_2 = 06.69$			
0.5	$a_1 = 00.53$ $a_2 = 01.00$	1.3946	0.03827	0.1787
	$M_1 = 10.59$ $M_2 = 06.86$			
0.6	$a_1 = 00.52$ $a_2 = 01.00$	1.3846	0.03794	0.1814
	$M_1 = 12.39$ $M_2 = 06.92$			

TABLE 2

$\alpha_1$	Optimal Retentions	Expected Net Profit	Adjustment Coefficient	Lundberg's Inequality
0.3	$a_1 = 01.00$ $a_2 = 01.00$	$E[W_1] = 1.3317$ $E[W_2] = 1.5803$	$R_1 = 0.1552$	$\psi_1(30) \leq 6278$
	$M_1 = 16.90$ $M_2 = 01.34$	$E[W] = 2.9120$	$R_2 = 1959$	$\psi_2(15) \leq 0529$
0.4	$a_1 = 01.00$ $a_2 = 01.00$	$E[W_1] = 1.4583$ $E[W_2] = 1.5803$	$R_1 = .01508$	$\psi_1(30) \leq 6361$
	$M_1 = 22.31$ $M_2 = 01.34$	$E[W] = 3.0387$	$R_2 = 1959$	$\psi_2(15) \leq 0529$
0.5	$a_1 = 01.00$ $a_2 = 01.00$	$E[W_1] = 1.5101$ $E[W_2] = 1.5803$	$R_1 = 0.1495$	$\psi_1(30) \leq 6386$
	$M_1 = 27.12$ $M_2 = 01.34$	$E[W] = 3.0904$	$R_2 = 1959$	$\psi_2(15) \leq 0529$
0.6	$a_1 = 01.00$ $a_2 = 01.00$	$E[W_1] = 1.5322$ $E[W_2] = 1.5803$	$R_1 = 0.1490$	$\psi_1(30) \leq 6395$
	$M_1 = 31.54$ $M_2 = 01.34$	$E[W] = 3.1125$	$R_2 = 1959$	$\psi_2(15) \leq 0529$



subsidize the payment of the reinsurance of those potentially more risky. In this example such interaction implied a decrease in the joint expected net profit, but there are substantial benefits in the company's security, as a whole. Nothing of this can be achieved, if one insists on treating each risk separately

## REFERENCES

- ANDREADAKIS, M and WATERS, H R (1980) The Effect of Reinsurance on the Degree of Risk Associated with an Insurance Portfolio *ASTIN Bulletin* **11**, 119-135
- BEARD, R E, PENTIKAINEN, T and PESONEN, E (1984) *Risk Theory*, Chapman and Hall, London
- CENTENO, L (1986) Measuring the effects of reinsurance by the adjustment coefficient *Insurance Mathematics and Economics* **5**, 169-182
- COURANT, R and JOHN, F (1974) *Introduction to Calculus & Analysis, II* Wiley, London
- GERBER, H U (1979) *An Introduction to Mathematical Risk Theory* S S Huebner Foundation, Philadelphia
- WATERS, H R (1979) Excess of Loss Reinsurance Limits *Scandinavian Actuarial Journal*, 37-43
- WATERS, H R (1983) Some Mathematical Aspects of Reinsurance *Insurance Mathematics and Economics*, **2**, 17-26

MARIA DE LOURDES CENTENO

*Instituto Superior de Economia e Gestão, Rua Miguel Lupi, 20,  
1200 Lisboa, Portugal.*



# DISTRIBUTION OF SURPLUS IN LIFE INSURANCE

BY HENRIK RAMLAU-HANSEN

*Baltica Insurance Company Ltd., Ballerup, Denmark*

## ABSTRACT

This paper discusses distribution of surplus in life insurance within a general Markov chain framework. A conservative interest rate and a conservative set of transition intensities are used for reserving purposes whereas more realistic assumptions are used for the purpose of distributing surplus. The paper examines various actuarial aspects of distributing surplus through either cash bonuses, terminal bonuses or increased benefits. The results are illustrated by some examples.

## KEYWORDS

Distribution of surplus; bonus; with profits annuity policy; with profits disability policy.

## 1. INTRODUCTION

The traditional life policy is a participating policy with margins of safety built into the valuation elements to allow for protection for adverse deviations. Surplus or profit can, therefore, in most cases be expected to emerge over the life of a portfolio of business. A large proportion of the surplus is usually distributed to the policyholders as bonuses or dividends. This distribution of surplus may be carried out in various ways. One method provides cash payments or reduction of premiums as the surplus arises, or the accumulated value of the cash bonuses may be paid when the policy becomes a claim or expires. By this method, a separate savings account is attached to the policy and the surplus is credited to the account as it emerges. Another way of distributing surplus is through terminal bonuses paid only when the policy expires. By this method, only survivors get a share of the accumulated surplus. The third method, and perhaps the most widely used, is one in which the profit is distributed to the policyholders by means of increasing the insurance benefits. This method provides a gradual increase in the benefits granted under the policy.

It is believed that these three different ways of distributing surplus cover many of the methods used in practice. We shall in this paper discuss various actuarial aspects of the mentioned distribution methods. The idea is that the

surplus should be distributed to those policyholders who contributed to the profit. Moreover, the distribution should be equitable, and the actuarial present value of the surplus generated by a policy should equal the actuarial present value of the bonuses paid to that same policy.

The results are discussed within a general Markov chain framework where an insurance policy is modelled as a time-inhomogeneous Markov chain, see e.g. HOEM (1969, 1988). The paper is motivated by BERGER (1939), SVERDRUP (1969) and SIMONSEN (1970), who discussed some aspects of accumulation and distribution of surplus. Moreover, RAMLAU-HANSEN (1988) analysed the emergence of surplus using a general Markov chain and counting process framework.

## 2. THE MARKOV CHAIN MODEL

We shall in the following consider life insurance policies which can be modelled by time-inhomogeneous Markov chains with finite state spaces. Hence, let  $S(\cdot)$  denote the right-continuous sample path function of a time-inhomogeneous Markov chain with finite state space  $I$ , and assume that the process starts in a state  $l \in I$  at time 0. The transition probabilities are denoted by  $P_{ij}^0(s, t) = P(S(t) = j | S(s) = i)$ ,  $i, j \in I$ ,  $s \leq t$ , and the forces of transition  $\mu_{ij}^0(\cdot)$  are defined by

$$\mu_{ij}^0(t) = \lim_{h \rightarrow 0^+} P_{ij}^0(t, t+h)/h, \quad i, j \in I, \quad i \neq j.$$

The intensities are assumed to be integrable on compact intervals.

Consider an  $n$ -year insurance policy characterized by the following conditions:

1. While the policy stays in state  $i$ , premiums are paid continuously to the company at the rate  $\pi_i(\cdot)$ , i.e.  $\pi_i(t) dt$  is paid during  $[t, t+dt)$ . Annuity benefits received by the insured while in state  $i$  are denoted by  $b_i(\cdot)$ .
2. If the policy moves from state  $i$  to state  $j$  at time  $t$ , a lump sum benefit  $B_{ij}(t)$  is paid to the insured immediately after time  $t$ .
3. When the policy expires at time  $n$ , the insured receives an amount  $B_i(n)$  if the policy is in state  $i$  at the maturity date.

The quantities  $\pi_i(t)$ ,  $b_i(t)$ ,  $B_{ij}(t)$ , and  $B_i(n)$  are all assumed to be non-random. It should also be noted that we have restricted ourselves to continuous payment of premiums and annuities, benefits tied to transitions between different states and to maturity benefits. However, single premiums and other types of non-random payments can be incorporated easily. Note also that we have introduced different notation for premiums paid and annuity benefits received because the two types of payments are affected differently by surplus distribution. Moreover, we shall refer to the "standard" benefits ( $b_i(t)$ ,  $B_{ij}(t)$ ,  $B_i(n)$ ,  $i, j \in I$ ,  $i \neq j$ ) as one unit of benefits, because one of the distribution methods operates by increasing all benefits proportionally. Finally, expenses are not included explicitly but can be regarded as separate benefits.

It is assumed that the company is making its valuations on the basis of a constant force of interest  $\delta$  and a set of transition intensities  $\mu_{ij}(\cdot)$ . The basis  $(\delta, \mu_{ij}, i, j \in I, i \neq j)$  is often called the valuation basis of the first order, and we shall assume that the company is required to use this set of (conservative) assumptions in determining reserves and premiums. However, we shall assume that the actual force of interest is  $\delta^0$  ( $\delta^0 > \delta$ ) and that the actual behaviour of the Markov chain is governed by the intensities  $\mu_{ij}^0(\cdot)$ . The elements  $(\delta^0, \mu_{ij}^0, i, j \in I, i \neq j)$  are often called the second order basis, and we shall assume that surplus is distributed according to this set of (realistic) assumptions.

Given that the policy is in state  $i$  at time  $t$ , let  $V_i(t)$  denote the prospective premium reserve corresponding to the valuation basis of the first order. Moreover, let  $SP_i(t)$  be the single premium or the actuarial present value of one unit of future benefits, provided that the policy is in state  $i$  at time  $t$ . We shall also assume that the equivalence principle is followed, i.e.  $V_i(0) = 0$ . The reserve  $V_i(t)$  is given by

$$(2.1) \quad \begin{aligned} V_i(t) = & \sum_j \sum_{k \neq j} \int_t^n v^{u-t} P_{ij}(t, u) \mu_{jk}(u) B_{jk}(u) du \\ & + \sum_j \int_t^n v^{u-t} P_{ij}(t, u) [b_j(u) - \pi_j(u)] du \\ & + \sum_j v^{n-t} P_{ij}(t, n) B_j(n), \end{aligned}$$

where the  $P_{ij}(s, t)$ 's are the transition probabilities corresponding to the intensities  $\mu_{ij}(\cdot)$ . A similar expression holds for  $SP_i(t)$ ; just substitute 0 for  $\pi_j(u)$  in (2.1). It is well known, see e.g. HOEM (1969), that  $V_i(t)$  satisfies Thiele's differential equation

$$(2.2) \quad \frac{d}{dt} V_i(t) = \delta V_i(t) + \pi_i(t) - b_i(t) - \sum_{j \neq i} \mu_{ij}(t) R_{ij}(t),$$

where  $R_{ij}(t) = V_j(t) + B_{ij}(t) - V_i(t)$  denotes the amount at risk associated with a transition from state  $i$  to state  $j$  at time  $t$ . Similarly,  $SP_i(t)$  satisfies

$$(2.3) \quad \frac{d}{dt} SP_i(t) = \delta SP_i(t) - b_i(t) - \sum_{j \neq i} \mu_{ij}(t) [SP_j(t) + B_{ij}(t) - SP_i(t)].$$

### 3. ACCUMULATION OF SURPLUS

Assume in this section that no bonuses are paid and that the company just pays the promised benefits  $b_i(t)$ ,  $B_{ij}(t)$ , and  $B_i(n)$  in return for the premiums  $\pi_i(t)$ . The average surplus or profit realized over the term of the policy may then be derived in the following way. Assume that the policy is in state  $i$  at time  $t$  and that the amount  $V_i(t)$  has been reserved. Then during  $[t, t+dt)$  the actual interest earned is  $\delta^0 dt V_i(t)$ , the premiums and the annuity benefits are  $\pi_i(t) dt$

and  $b_i(t) dt$ , respectively, and the expected net loss due to transitions out of state  $i$  is  $\sum_{j \neq i} \mu_{ij}^0(t) dt R_j(t)$ . However, the reserve needed at time  $t + dt$ , assuming the policy is still in state  $i$ , is  $V_i(t + dt)$ , and hence the net profit becomes

$$\gamma_i(t) dt = (1 + \delta^0 dt) V_i(t) + \pi_i(t) dt - b_i(t) dt - \sum_{j \neq i} \mu_{ij}^0(t) dt R_j(t) - V_i(t + dt).$$

This leads to

$$\gamma_i(t) = \delta^0 V_i(t) + \pi_i(t) - b_i(t) - \sum_{j \neq i} \mu_{ij}^0(t) R_j(t) - \frac{d}{dt} V_i(t)$$

and using (2.2) we get

$$\begin{aligned} (3.1) \quad \gamma_i(t) &= (\delta^0 - \delta) V_i(t) + \sum_{j \neq i} (\mu_{ij}(t) - \mu_{ij}^0(t)) R_j(t) \\ &= \Delta \delta V_i(t) + \sum_{j \neq i} \Delta \mu_{ij}(t) R_j(t), \end{aligned}$$

introducing  $\Delta \delta = \delta^0 - \delta$  and  $\Delta \mu_{ij}(t) = \mu_{ij}(t) - \mu_{ij}^0(t)$ . Thus, assuming that the policy is in state  $i$  at time  $t$ , surplus accumulates at the rate  $\gamma_i(t)$ , which, according to (3.1), is the sum of the excess interest earnings and the profit or loss associated with transitions out of state  $i$ . The actuarial present value at time 0 of the total surplus accumulated over  $[0, t]$  during stays in the state  $i$  is given by

$$(3.2) \quad \Gamma_i(t) = \int_0^t e^{-\delta^0 s} P_{1i}^0(0, s) \gamma_i(s) ds,$$

and the present value of the total surplus accumulated over  $[0, t]$  is

$$(3.3) \quad \Gamma(t) = \sum_i \Gamma_i(t).$$

It should also be noted that

$$\begin{aligned} (3.4) \quad \Gamma(t) &= \sum_i \int_0^t e^{-\delta^0 s} P_{1i}^0(0, s) [\pi_i(s) - b_i(s)] ds \\ &\quad - \sum_i \sum_{j \neq i} \int_0^t e^{-\delta^0 s} P_{1i}^0(0, s) \mu_{ij}^0(s) B_j(s) ds \\ &\quad - \sum_i e^{-\delta^0 t} P_{1i}^0(0, t) V_i(t), \end{aligned}$$

and

$$\begin{aligned}
 (3.5) \quad \Gamma_i(t) &= \sum_{k \neq i} \int_0^t e^{-\delta^0 s} P_{1k}^0(0, s) \mu_{ki}^0(s) V_i(s) ds \\
 &+ \int_0^t e^{-\delta^0 s} P_{1i}^0(0, s) [\pi_i(s) - b_i(s)] ds \\
 &- \sum_{j \neq i} \int_0^t e^{-\delta^0 s} P_{1j}^0(0, s) \mu_{ij}^0(s) [B_{ij}(s) + V_j(s)] \\
 &- e^{-\delta^0 t} P_{1i}^0(0, t) V_i(t),
 \end{aligned}$$

see e.g. RAMLAU-HANSEN (1988) formulas (4.1) and (4.10). Hence,  $\Gamma(t)$  may be interpreted as the actuarial present value of the difference between the premiums received and the benefits and reserves that have to be provided. The gain  $\Gamma_i(t)$  may be interpreted similarly.

For a broader discussion of surplus accumulation and in particular various stochastic aspects, see RAMLAU-HANSEN (1988). However, note that in RAMLAU-HANSEN (1988)  $\Gamma(t)$  and  $\Gamma_i(t)$  are random variables and not actuarial values.

#### 4 DISTRIBUTION OF SURPLUS

##### 4.1. Cash bonuses

It was shown in the previous section that the surplus accumulates at the rate  $\gamma_i(t)$  in state  $i$  at time  $t$ . Hence, the surplus may be distributed by simply paying the policyholder an annuity  $\gamma_i(t)$  while the policy is in state  $i$ . These dividend payments may then supplement annuity benefits or partly offset premiums paid under the terms of the policy. The present value at time 0 of the total bonuses paid during  $[0, t]$  is

$$(4.1) \quad C(t) = \sum_i \int_0^t e^{-\delta^0 s} Y_i(s) \gamma_i(s) ds,$$

where  $Y_i(s) = 1$  if  $S(s) = i$  and 0 otherwise. Note that the amount  $C(t)$  is random, but  $EC(t) = \Gamma(t)$ . In practice, companies that pay cash bonuses do not pay the continuous annuities  $\gamma_i(t)$ , but they may distribute the surplus through annual instalments or by other means, cf. Section 5.1.

The amount  $C(t)$  may also be interpreted as the present value of the amount in a savings account attached to the insurance policy. During stays in state  $i$ , the account is then credited continuously at the rate  $\gamma_i(t)$ . Some companies do follow this procedure by deferring the payment of the cash bonus until the policy becomes a claim or expires. If the policy becomes a claim or expires at, say time  $t$ , then the amount  $\exp(\delta^0 t) C(t)$  is paid in addition to the policy

benefits. If two or more lump sum payments are possible under the policy, the surplus may be distributed through a series of payments.

It should be noted that the distribution of surplus through periodic payments allows all policyholders to share in the profit.

#### 4.2. Terminal bonuses

In this subsection we discuss a distribution method according to which the surplus is distributed to the policyholders only when the policies expire. No additional benefits are paid during the term of the policy, except at the maturity date. Hence, terminal bonuses may be used to enhance the maturity value of the policy.

It was shown in Section 3 that the actuarial present value of the total surplus accumulated during stays in a state  $i$  is  $\Gamma_i(n)$  given by (3.2). Hence, if this profit is to be distributed as a payment to those policyholders who are in state  $i$  at time  $n$ , each should receive

$$(4.2) \quad T_i(n) = \Gamma_i(n)/[e^{-\delta^n} P_{1i}^0(0, n)].$$

One might also limit the payment of bonuses to those survivors who are in the initial state at time  $n$ . Depending on the design of the policy, this practice may favour those policyholders who have not made any claims under the policy. In this situation, each of the survivors in state 1 should receive

$$(4.3) \quad T(n) = \Gamma(n)/[e^{-\delta^n} P_{11}^0(0, n)].$$

at time  $n$

However, it should be noted that by applying terminal bonuses only survivors are rewarded, and those who have died do not get a share of the profit, although they may actually have contributed to it. Hence, the method resembles in a way a tontine scheme, and this may explain why terminal bonuses are only used in connection with policies with a strong savings element.

#### 4.3. Increased benefits

In this section we assume that the surplus is used to increase the policy benefits. This is one of the most common ways of distributing surplus in practice. We shall assume that all benefits are increased proportionally so that the original relationship between the benefits is preserved. Hence, the surplus is used as a single premium to purchase additional units of benefits, cf. Section 2.

At issue, the net premium reserve is  $V_1(0) = 0$  and the policy provides the benefits  $b_i(s)$ ,  $B_j(s)$ , for  $s \geq 0$ , and  $B_i(n)$ . Let us now assume that the policy is in state  $i$  at time  $t$  and that the policy entered this state at some time  $t_i$ . Moreover, assume that past surplus has been used to buy  $D(t)$  units of additional benefits so that they are now promised to be  $b_j^*(s) = b_j(s)(1 + D(t))$ ,  $B_{jk}^*(s) = B_{jk}(s)(1 + D(t))$ , for  $s \geq t$ , and  $B_j^*(n) = B_j(n)(1 + D(t))$ . The rate of increase



of benefits at time  $u$  is denoted by  $d(u)$ , i.e.  $D(t) = \int_0^t d(u) du$ . It should be noted that  $D(\cdot)$  is actually a stochastic process since it is a function of the sample path of the Markov chain. At time 0,  $D(t)$  is unknown because the future course of the policy is unknown

Taking the increased benefits into account, the policy reserve is now

$$(4.4) \quad V_i^*(t) = V_i(t) + D(t) SP_i(t),$$

where both  $V_i(t)$  and  $SP_i(t)$  are calculated using the first order valuation basis, cf. (2.2)-(2.3). Hence, using arguments similar to the ones in Section 3, the average surplus that emerges at time  $t$  is given by the rate

$$\begin{aligned} \gamma_i^*(t) &= \Delta\delta V_i^*(t) + \sum_{j \neq i} \Delta\mu_y(t) [V_j^*(t) + B_y^*(t) - V_i^*(t)] \\ &= \Delta\delta V_i(t) + \sum_{j \neq i} \Delta\mu_y(t) [V_j(t) + B_y(t) - V_i(t)] \\ &\quad + D(t) \left\{ \Delta\delta SP_i(t) + \sum_{j \neq i} \Delta\mu_y(t) [SP_j(t) + B_y(t) - SP_i(t)] \right\} \end{aligned}$$

using (4.4). Thus,

$$(4.5) \quad \gamma_i^*(t) = \gamma_i(t) + D(t) \kappa_i(t)$$

if we introduce  $\kappa_i(t) = \Delta\delta SP_i(t) + \sum_{j \neq i} \Delta\mu_y(t) [SP_j(t) + B_y(t) - SP_i(t)]$ .

The surplus  $\gamma_i^*(t)$  is used to buy  $d(t)$  units of additional benefits at a cost of  $SP_i(t)$  per unit. Thus, we must have that

$$d(t) SP_i(t) = \gamma_i(t) + D(t) \kappa_i(t),$$

or

$$(4.6) \quad D'(t) = d(t) = q_i(t) + D(t) r_i(t),$$

where  $q_i(t) = \gamma_i(t)/SP_i(t)$  and  $r_i(t) = \kappa_i(t)/SP_i(t)$ . Equation (4.6) is a linear differential equation with solution

$$(4.7) \quad \begin{aligned} D(t) &= \int_{t_i}^t q_i(s) \exp \left( \int_s^t r_i(u) du \right) ds \\ &\quad + D(t_i) \exp \left( \int_{t_i}^t r_i(s) ds \right), \quad t \geq t_i, \end{aligned}$$

which yields, in a closed form, an expression for the total increase of the benefits due to the emerging surplus.

It should be noted that (4.7) holds only during the stay in state  $i$ . If the policy at some later time  $t_j$  moves to state  $j$  then a similar formula holds with  $t_j$  and  $j$  substituted for  $t_i$  and  $i$ , respectively. Thus, the rate of increase of benefits depends on the current state of the policy, but the policyholder should not expect any sudden changes in the benefits because  $D(\cdot)$  is a continuous function.

It should also be noted that in this section additional benefits are granted as the surplus is earned. In order to make this a prudent distribution method, it requires that at any time the future safety margins are sufficient to safeguard the company against any adverse experience. Moreover, since companies normally cannot reduce bonuses once they have been declared, it also requires surplus always to be positive, i.e.  $\gamma_i^*(t)$  has to be positive. If this is not the case, distribution of surplus will have to be deferred, and the method above will have to be modified.

If the original policy is a single premium policy, then  $V_i(t) = SP_i(t)$ ,  $\kappa_i(t) = \gamma_i(t)$ , and  $q_i(t) = r_i(t)$ . In this case, it follows from (4.7) that

$$(4.8) \quad 1 + D(t) = (1 + D(t_i)) \exp \left( \int_{t_i}^t r_i(u) du \right), \quad t \geq t_i.$$

Finally, we shall see that  $V_i^*(t)$  satisfies a second order differential equation although it was defined as a first order premium reserve, cf. (4.4). The reason is that the benefits are adjusted continuously. According to (4.4),

$$\frac{d}{dt} V_i^*(t) = \frac{d}{dt} V_i(t) + D'(t) SP_i(t) + D(t) \frac{d}{dt} SP_i(t),$$

and using (2.2)-(2.3) and (4.6) we get after some simple arithmetic the equation

$$\frac{d}{dt} V_i^*(t) = \delta^0 V_i^*(t) + \pi_i(t) - b_i^*(t) - \sum_{j \neq i} \mu_{ij}^0(t) [V_j^*(t) + B_j^*(t) - V_i^*(t)].$$

## 5. EXAMPLES

To illustrate some of the results, we shall consider two examples: A single-premium annuity policy and a disability policy. The first example focuses on ways of distributing interest surplus, whereas the other example is a discussion of surplus distribution in a three-state model. We have not included an example of a typical endowment policy, because we feel that the two other examples are more interesting.

### 5.1. An annuity policy

Let us consider a single-premium annuity policy where a benefit  $b$  is paid continuously throughout the life of an individual ( $x$ ). The first order premium

reserve is

$$V(t) = b \bar{a}_{x+t}, \quad t \geq 0,$$

using standard actuarial notation. We assume that the actual force of interest is a constant  $\delta^0 > \delta$  and that the interest earnings are the only source of surplus, i.e.  $\mu^0(\cdot) = \mu(\cdot)$ .

Then, according to (3.1), surplus is accumulated at the rate  $\gamma(t) = \Delta\delta V(t)$ , and we may, therefore, pay the insured the adjusted benefit

$$(5.1) \quad b_1(t) = b + \Delta\delta V(t).$$

Alternatively, (4.8) shows that the surplus may also be distributed by means of the increased benefits

$$(5.2) \quad b_2(t) = \exp[(\delta^0 - \delta)t] b.$$

It is interesting to note that (5.1) is typically a decreasing function of time/age, whereas (5.2) is increasing exponentially. Thus, the two formulas represent two completely different ways of distributing the same surplus.

In practice, however, it is not possible to adjust the benefits continuously as it is assumed in (5.1) and (5.2). In Denmark, for instance, pensions are adjusted only annually. Therefore, there is a need for more practical versions of (5.1) and (5.2). If, for example, the total surplus accumulated during year  $t$ ,  $t = 0, 1, \dots$ , has to be distributed through a level benefit  $b_3(t)$  payable continuously during year  $t$ , then  $b_3(t)$  has to be determined by

$$(5.3) \quad V(t) = b_3(t) \bar{a}_{x+t}^0 + v^0 p_{x+t} V(t+1), \quad t = 0, 1, \dots,$$

where the superscript "0" indicates that the values are based on  $\delta^0$ . Hence,  $b_3(0)$  is the level benefit that is paid continuously during year 0,  $b_3(1)$  is paid during year 1 etc. It follows from (5.3) that the series of benefits  $b_3(0)$ ,  $b_3(1), \dots$  serves the same purpose as the function  $b_1(\cdot)$ .

Similarly, the function  $b_2(\cdot)$  may be replaced by level annual benefits in the following way. Assume that the benefit is a level amount  $b_4(t)$  during year  $t$ . Then  $b_4(t+1)$  is determined by the equation

$$\begin{aligned} b_4(t) \int_t^{t+1} e^{-\delta^0(s-t)} {}_{s-t}p_{x+t} \Delta\delta \bar{a}_{x+s} ds + v^0 p_{x+t} b_4(t) \bar{a}_{x+t+1} \\ = v^0 p_{x+t} b_4(t+1) \bar{a}_{x+t+1}. \end{aligned}$$

Thus, we see that the surplus accumulated over the year is used to grant an increase of the benefit from  $b_4(t)$  to  $b_4(t+1)$ .

Table 1 gives examples for an annuity of 10,000 issued to a male aged 60. The valuation rate of interest is 4.5%,  $\delta = \log(1.045)$ , whereas the actual interest rate is assumed to be 8%, i.e.  $\delta^0 = \log(1.08)$ . Moreover, the mortality is  $\mu(t) = 0.0005 + 10^{0.038(x+t) - 4.12}$  which is the standard assumption used by Danish life companies

TABLE 1  
COMPARISON OF VARIOUS WAYS OF DISTRIBUTING SURPLUS FOR AN ANNUITY OF  
10,000 ISSUED TO A MALE AGED 60

Age $x+t$	$b_1(t)$	$b_2(t)$	$b_3(t)$	$b_4(t)$
60	13,885	10,000	13,835	10,000
61	13,784	10,335	13,734	10,350
62	13,682	10,681	13,632	10,713
63	13,580	11,039	13,529	11,089
64	13,477	11,409	13,426	11,479
65	13,373	11,791	13,322	11,884
70	12,853	13,902	12,803	14,141
75	12,345	16,391	12,297	16,861
80	11,869	19,326	11,825	20,161

The table highlights the difference between the payment schemes  $b_3(t)$  and  $b_4(t)$ . The calculations show that  $b_3(t)$  is larger than  $b_4(t)$  during the first 8 years after which  $b_4(t)$  exceeds  $b_3(t)$ . The distribution method that leads to  $b_4(t)$  is widely used in Denmark, primarily because it provides some protection against inflation. However, one might also argue that in years with low inflation, many retirees are presumably prepared to forfeit inflation protection in return for higher benefits while they are healthy and the quality of life is higher. Thus,  $b_3(t)$  should perhaps be recommended more widely than it has been until now.

## 5.2. A disability policy

We shall in this section consider an  $n$ -year disability policy issued on an able male aged  $x$ . The policy may be described by the three-state Markov model depicted in Figure 1. It is assumed that the policy provides a continuous

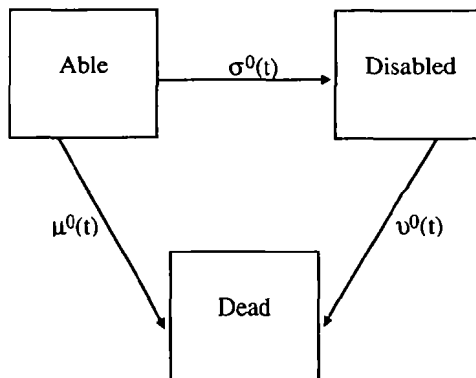


FIGURE 1 The disability model

annuity of 1 as long as the insured is disabled. Premiums are waived during disability, and it is assumed that the premium payments cease after  $m = n - 5$  years in order to avoid negative reserves close to maturity.

Danish companies assume in their valuations that the transition intensities are given by

$$\mu(t) = \nu(t) = 0.0005 + 10^{0.038(\tau+t)-4.12}$$

and

$$\sigma(t) = 0.0004 + 10^{0.060(\tau+t)-5.46}$$

The rate of interest is still assumed to be 4.5%, i.e.  $\delta = \log(1.045)$ . We shall study surplus distribution under the somewhat more realistic assumptions that the actual behaviour of the policy is governed by

$$\begin{aligned} \mu^0(t) &= \theta_1 \mu(t), \\ \sigma^0(t) &= \theta_2 \sigma(t), \end{aligned}$$

and

$$\nu^0(t) = \theta_3 \nu(t),$$

where  $\underline{\theta} = (\theta_1, \theta_2, \theta_3)$  is given below. Moreover, the actual rate of interest is also in this example 8%, i.e.  $\delta^0 = \log(1.08)$ .

The premium  $\pi$  and the first order reserves are given by

$$\pi = \bar{a}_{\chi}^{ai} \bar{n} / \bar{a}_{\chi}^{aa} \bar{m},$$

$$V_a(t) = \bar{a}_{\chi+t}^{ai} \bar{n-t} - \pi \bar{a}_{\chi+t}^{aa} \bar{m-t},$$

$$V_i(t) = \bar{a}_{\chi+t}^u \bar{n-t},$$

where  $\bar{a}_x^{ai} \bar{n} = \int_0^n v^s {}_s p_{\chi}^{ai} ds$ ,  $\bar{a}_x^{aa} \bar{n} = \int_0^n v^s {}_s p_{\chi}^{aa} ds$ ,  $\bar{a}_x^u \bar{n} = \bar{a}_x \bar{n} = \int_0^n v^s {}_s p_{\chi} ds$ ,

and where  ${}_s p_{\chi}^{aa} = \exp\left(-\int_0^s \mu(u) + \sigma(u) du\right)$ ,  ${}_s p_{\chi} = \exp\left(-\int_0^s \mu(u) du\right)$ ,

and  ${}_s p_{\chi}^{ai} = {}_s p_{\chi} - {}_s p_{\chi}^{aa}$ . The corresponding amounts at risk are  $R_{ai}(t) = V_i(t) - V_a(t)$ ,  $R_{ad}(t) = -V_a(t)$ , and  $R_{id}(t) = -V_i(t)$ . Here  $a$  denotes the state able,  $i$  the state disabled (invalid), and  $d$  the state dead.

According to (3.1), surplus accumulates at the rates

$$\begin{aligned} \gamma_a(t) &= \Delta\delta V_a(t) + \Delta\sigma(t) R_{ai}(t) + \Delta\mu(t) R_{ad}(t) \\ &= (\Delta\delta - \Delta\mu(t) - \Delta\sigma(t)) V_a(t) + \Delta\sigma(t) V_i(t), \end{aligned}$$

and

$$\begin{aligned} \gamma_i(t) &= \Delta\delta V_i(t) + \Delta\nu(t) R_{id}(t) \\ &= (\Delta\delta - \Delta\nu(t)) V_i(t) \end{aligned}$$

during stays in the states able and disabled, respectively. Here  $\Delta\delta = \delta^0 - \delta$ ,  $\Delta\mu(t) = \mu(t) - \mu^0(t)$ ,  $\Delta\sigma(t) = \sigma(t) - \sigma^0(t)$ , and  $\Delta v(t) = v(t) - v^0(t)$ . Hence, the present values at time 0 of the total accumulated surpluses are

$$(5.3) \quad \Gamma_a(n) = \int_0^n \exp(-\delta^0 s) {}_s p_x^{0aa} \gamma_a(s) ds,$$

$$(5.4) \quad \Gamma_i(n) = \int_0^n \exp(-\delta^0 s) {}_s p_x^{0ai} \gamma_i(s) ds,$$

and

$$(5.5) \quad \Gamma(n) = \Gamma_a(n) + \Gamma_i(n),$$

cf. (3.2). Here,  ${}_s p_x^{0aa}$  and  ${}_s p_x^{0ai}$  are second order values of  ${}_s p_x^{aa}$  and  ${}_s p_x^{ai}$ , respectively. The corresponding possible terminal bonuses  $T_a(n)$ ,  $T_i(n)$ , and  $T(n)$  are given by (4.2) and (4.3).

We have in Table 2 shown examples of (5.3)-(5.5) for policies with  $x+n = 65$  and  $x+m = 60$ . Moreover, it is assumed in these examples that  $\theta_1 = 0.7$ ,  $\theta_2 = 0.8$ , and  $\theta_3 = 1$  which are close to what currently is used by many Danish companies. The figures illustrate clearly the size of the surplus inherent in the policies. Take as an example the policy issued at age 30. Here the actuarial present value of the total surplus is 0.144 compared with the total value of the premium payments  $\pi \bar{a}_{x+m}^{aa}$  which equals 0.423. The surplus might be distributed through the terminal dividends given in Table 2. However, it is hard to argue that only paying 2.13 and 5.12 to the lives that are able and disabled at age 65 is an equitable way of distributing the profit. It is also difficult to justify that large amounts should be paid to the disabled lives who have already collected benefits under the terms of the policy

Table 3 shows for the example  $x = 30$  the possible benefits if the surplus is used to continuously increase the benefits. We have shown the rates of surplus accumulation  $\gamma_a^*(t)$  and  $\gamma_i^*(t)$ , cf. (4.5), together with  $1 + D_a(t)$  and  $1 + D_i(t)$ , respectively. Here  $1 + D_a(t)$  is the basic disability annuity that becomes payable if disability occurs at time  $t$ . This quantity and  $\gamma_a^*(t)$  have been calculated assuming that the policy has remained in the state able during  $[0, t]$ . Similarly,

TABLE 2

EXAMPLES OF PRESENT VALUES OF ACCUMULATED SURPLUSES AND POSSIBLE TERMINAL BONUSES FOR VARIOUS DISABILITY POLICIES WITH  $\underline{\theta} = (0.7, 0.8, 1)$

Issuc age	1000 $\pi$	$\Gamma_a(n)$	$\Gamma_i(n)$	$\Gamma(n)$	$T_a(n)$	$T_i(n)$	$T(n)$
20	19.0	0.086	0.037	0.123	3.97	9.29	5.65
30	26.8	0.101	0.043	0.144	2.13	5.12	3.03
40	40.8	0.110	0.049	0.159	1.05	2.77	1.51
50	65.5	0.103	0.040	0.143	0.43	1.13	0.60

TABLE 3  
 RATES OF SURPLUS ACCUMULATION AND SIZE OF INCREASED BENEFITS  
 AGE AT ISSUE  $x = 30$  AND  $\underline{\theta} = (0.7, 0.8, 1)$

Age $x+t$	$\gamma_a^*(t)$	$\gamma_i^*(t)$	$1 + D_a(t)$	$1 + D_i(t)$
30	0.002	0.560	1.00	1.00
40	0.012	0.654	1.14	1.39
50	0.029	0.655	1.51	1.93
60	0.045	0.381	2.68	2.69
61	0.044	0.324	2.97	2.78
62	0.041	0.258	3.38	2.87
63	0.036	0.183	4.02	2.97
64	0.027	0.098	5.38	3.07
64.5	0.019	0.050	7.17	3.12
65	0	0	$\infty$	3.17

$1 + D_i(t)$  is the annuity payable at time  $t$  and  $\gamma_i^*(t)$  measures the rate of surplus accumulation, provided that the insured became disabled just after time 0. It is interesting to note that (4.7) leads to

$$D_i(t) = \int_0^t q_i(s) \exp \left( \int_s^t r_i(u) du \right) ds = \exp \left( \int_0^t r_i(u) du \right) - 1$$

with  $q_i(s) = \gamma_i(s)/SP_i(s) = \Delta\delta - \Delta v(t)$ ,  $SP_i(t) = V_i(t)$ , and  $r_i(u) = q_i(u)$ . Hence,  $D_i(t)$  is in general easy to compute, and in the example in Table 3  $\Delta v(t) = 0$ , so  $1 + D_i(t) = \exp(\Delta\delta t)$ , cf. (5.2).

It is interesting to note that  $1 + D_a(t)$  and  $1 + D_i(t)$  increase at different rates. In particular, the sharp increase in  $1 + D_a(t)$  close to maturity should be noted. Actually, it is easily seen that  $1 + D_a(t) \rightarrow \infty$  as  $t \rightarrow n$ . It may be explained by the fact that close to maturity, the surplus is of the size  $O(h)$ ,  $h = n - t$ , whereas the price of providing additional benefits is  $\bar{a}_{\lambda+t}^{ai} \frac{1}{n-t} = O(h^2)$ . In practice, these excessive benefits should, of course, be avoided, and it may be achieved by shifting to a system with cash or deferred bonuses when the policy approaches maturity.

In Table 3,  $1 + D_i(t)$  yields the annuity at time  $t$  if the disability occurred at time 0. However, if disability occurs at some later time, say  $t_i$ , then it follows from (4.7) that the benefit at time  $t \geq t_i$  is given by

$$\begin{aligned} 1 + \tilde{D}_i(t) &= (1 + D_a(t_i)) \exp \left( \int_{t_i}^t r_i(u) du \right) \\ &= (1 + D_a(t_i)) (1 + D_i(t)) / (1 + D_i(t_i)). \end{aligned}$$

Thus, if for example disability occurs at age 40, then the initial annuity is 1.14, which after 10 years of disability will have risen to  $(1.14)(1.93)/1.39 = 1.58$ . It illustrates that the benefits while disabled depend on the duration of the disability.

TABLE 4  
 EXAMPLES OF DISABILITY ANNUITIES  $1 + D_i(t)$  IN THE SITUATIONS WHERE  $\theta_3 = 1, 2$  AND  $5$   
 AGE AT ISSUE  $x = 30$  AND  $(\theta_1, \theta_2) = (0.7, 0.8)$

Age $x+t$	$\theta_3 = 1$	$\theta_3 = 2$	$\theta_3 = 5$
30	1 00	1 00	1 00
40	1 39	1 42	1 52
50	1 93	2 07	2 53
60	2 69	3 18	5 27
65	3 17	4 11	9 01

TABLE 5  
 PRESENT VALUES OF ACCUMULATED SURPLUSES FOR DIFFERENT VALUES OF  $\underline{\theta}$   
 AGE AT ISSUE  $x = 30$

$\underline{\theta} = (\theta_1, \theta_2, \theta_3)$	$\Gamma_a(n)$	$\Gamma_i(n)$	$\Gamma(n)$
(0.7, 0.8, 1)	0 101	0 043	0 144
(0.7, 1, 1)	0 051	0 054	0 104
(0.7, 1, 2)	0 051	0 062	0 113
(0.7, 1, 5)	0 051	0 085	0 136
(0.7, 1, 10)	0 051	0 112	0 163

We have also shown in Table 4 the kind of disability annuities that can be offered if it is further taken into account that disabled lives often have a much higher mortality than able lives. We have shown examples of  $1 + D_i(t)$  in the situations where  $\theta_3 = 1, 2$ , and  $5$ . Otherwise, the assumptions are the same as in Table 3. It is clear that substantial mortality gains on the disabled lives might be used to increase the disability benefits further.

However, in some cases mortality gains on disabled lives would rather be used to offset unsatisfactory disability experience among able lives. In this way all get a share of the "favourable" mortality among disabled lives. To give an impression of to what extent an unfavourable value of  $\theta_2$  can be offset by a favourable value of  $\theta_3$ , we have shown in Table 5 some examples where  $\theta_2 = 0.8$  and  $1$ , and where  $\theta_3 = 1, 2, 5$ , and  $10$ . Hence, taking  $\underline{\theta} = (\theta_1, \theta_2, \theta_3) = (0.7, 0.8, 1)$  as our basis, it is seen that even  $\theta_3 = 5$  is not sufficient to eliminate the overall effect of  $\theta_2 = 1$ , whereas  $\theta_3 = 10$  more than compensates for the effect of  $\theta_2 = 1$ .

#### REFERENCES

- BERGER, A (1939) *Mathematik der Lebensversicherung*. Springer, Wien  
 HOEM, J M (1969) Markov chain models in life insurance *Blätter der Deutschen Gesellschaft für Versicherungsmathematik* IX, 97-107  
 HOEM, J M (1988) The versatility of the Markov chain as a tool in the mathematics of life insurance *Transactions of the 23rd international congress of actuaries*, vol R, 171-202



- RAMLAU-HANSEN, H (1988) The emergence of profit in life insurance *Insurance Mathematics and Economics* 7, 225–236
- SIMONSEN, W (1970) *Forsikringsmatematik*, hefte III Københavns Universitets Fond til tilvejebringelse af læremidler
- SVERDRUP, E (1969) Noen forsikringsmatematiske emner Statistical memoirs No 1, Institute of Mathematics, University of Oslo

HENRIK RAMLAU-HANSEN

*Baltica Insurance Company Ltd., Klausdalsbrovej 601, DK-2750 Ballerup, Denmark.*



## WORKSHOP

### CREDIBILITY MODELS WITH TIME-VARYING TREND COMPONENTS

BY JOHANNES LEDOLTER, STUART KLUGMAN, CHANG-SOO LEE

#### ABSTRACT

Traditional credibility models have treated the process generating the losses as stable over time, perhaps with a deterministic trend imposed. However, there is ample evidence that these processes are not stable over time. What is required is a method that allows for time-varying parameters in the process, yet still provides the shrinkage needed for sound ratemaking. In this paper we use an automobile insurance example to illustrate how this can be accomplished

#### KEYWORDS

Credibility; Kalman filter; shrinkage estimation; time series; trend components.

#### 1. INTRODUCTION

The goal of any ratemaking process is to estimate future claims on the basis of prior experience. The experience will be available for many groups over several time periods. It has been long known (MOWBRAY, 1914) that both statistical and business optimality is achieved by first estimating a rate for each group and then reducing the large values and increasing the small ones. Traditionally (e.g., BÜHLMANN and STRAUB, 1972) the initial estimates are sample means. Others (e.g. HACHEMEISTER, 1975) have recommended deterministic trend factors. Most all approaches that are currently used assume that the time series observations from a single group vary independently around a stable mean or trend.

Most time series, however, exhibit time-varying levels as well as autocorrelations among adjacent observations. The optimal forecasts for such series do not assign equal weights to all past observations, but discount the information according to their age; older observations get less weight. See BOX and JENKINS (1976) or ABRAHAM and LEDOLTER (1983) for a thorough discussion. Evidence for time-varying parameters was presented for automobile losses by BAILEY and SIMON (1959). A problem with most standard time series approaches, however, is that they are designed for making forecasts based on single series of relatively long lengths. Typical insurance problems contain many (sometimes hundreds) series of short (3-7 years) duration. Because these short series are occurring in a common external environment (e.g., of rising health care costs, automobile safety improvements, etc.) many of the features will be common to all of the series. The importance of both time series and

cross-sectional effects has also been noted in two recent econometrics papers by GARCIA-FERRER et al. (1987) and ZELLNER and HONG (1989) who use shrinkage methods to predict the economic growth rates of several countries.

The purpose of this paper is to bring together a dynamic model for the time-varying aspects of the problem and a shrinkage technique that takes account of the multiple group aspect. In Section 2 we review the credibility model with time-invariant parameters. In Section 3 we discuss univariate structural time-series models with time-varying trend and seasonal coefficients and we apply the shrinkage approach of Section 2 to the coefficients in the structural time series models. The final section illustrates this approach on actuarial data.

## 2. THE STANDARD CREDIBILITY MODEL

In all of the situations discussed in this paper the data consists of observations  $Y_t^{(i)}$ ,  $i = 1, \dots, k$ ,  $t = 1, \dots, n$  where  $k$  is the number of groups under consideration and  $n$  is the number of periods of observation. Typically, each value represents the amount paid in claims, divided by some measure of the size of the group,  $P_t^{(i)}$ . The objective is to forecast the value for a future period,  $Y_{n+1}^{(i)}$ , for each group.

A linear data generating model for the observations specifies

$$(2.1) \quad Y_t^{(i)} = \underline{x}_{it}' \underline{\beta}^{(i)} + e_t^{(i)} \quad e_t^{(i)} \sim N(0, \sigma^2/P_t^{(i)})$$

where  $e_t^{(i)}$ , for  $t = 1, \dots, n$  and  $i = 1, \dots, k$ , are independent and  $\underline{x}_{it}$  are  $p \times 1$  known design vectors, usually functions of  $t$ . Two well-known models take (1)  $p = 1$  and  $x_{it} = 1$  and (2)  $p = 2$  and  $\underline{x}_{it} = (1, t)'$ . The data generating model in (1) is part of a special case of the Bühlmann-Straub model (BÜHLMANN and STRAUB, 1972); the linear trend in (2) is part of the Hachemeister model (HACHEMEISTER, 1975). The factor  $P_t^{(i)}$  in (2.1) is a measure of the amount of data that produces the observation  $Y_t^{(i)}$ , which in most actuarial situations is an average of many observations. The forecast of  $Y_{n+1}^{(i)}$ , the observation at a future time period, is provided by the estimate of the mean  $E(Y_{n+1}^{(i)}) = \underline{x}_{i, n+1}' \underline{\beta}^{(i)}$ .

The standard credibility model also assumes that the coefficients  $\underline{\beta}^{(i)}$ , for  $i = 1, \dots, k$ , are independent realizations from a common distribution. That is,

$$(2.2) \quad \underline{\beta}^{(i)} = \underline{b} + \underline{a}^{(i)} \quad \text{where} \quad \underline{a}^{(i)} \sim N(0, \sigma^2 B).$$

Treating this second level distribution as a prior distribution, the Bayes shrinkage estimate of  $\underline{\beta}^{(i)}$  is given by

$$(2.3) \quad \tilde{\underline{\beta}}^{(i)} = Z_i \hat{\underline{\beta}}^{(i)} + (I - Z_i) \underline{b}$$

where

$$(2.4) \quad \hat{\underline{\beta}}^{(i)} = \left( \sum_t P_t^{(i)} \underline{x}_{it} \underline{x}_{it}' \right)^{-1} \sum_t P_t^{(i)} \underline{x}_{it} Y_t^{(i)}$$

is the weighted least squares estimate in group  $i$ ,

$$(2.5) \quad Z_i = B(B + V_i)^{-1},$$

and

$$(2.6) \quad V_i = \left( \sum_t P_t^{(i)} x_{it} x'_{it} \right)^{-1}.$$

A problem with this solution is that estimates of the quantities  $B$  and  $b$  must be obtained. A commonly accepted approach is to use the method of moments estimates that have been developed in variance components analysis (see SWAMY, 1971). However, there are a number of drawbacks with this approach. The estimates of  $B$  and  $Z_i$  are biased and, furthermore, the moment estimate of the scaled covariance matrix  $B$  need not be non-negative definite. These drawbacks can be overcome, in part, by either using the iterative estimation approach of DEVYLDER (1981, 1984), or a true Bayes approach instead of an empirical Bayes approach. The details of the Bayes analysis can be found in KLUGMAN (1987). DeVlyder proposes estimators  $\hat{B}$  and  $\hat{b}$  of  $B$  and  $b$  which depend via  $Z_i = B(B + V_i)^{-1}$  on the parameter  $B$  to be estimated. He suggests an iterative procedure where

$$(2.7) \quad \hat{b} = \left( \sum_i Z_i \right)^{-1} \sum_i Z_i \hat{\beta}^{(i)},$$

$$H = \sum_i Z_i (\hat{\beta}^{(i)} - \hat{b}) (\hat{\beta}^{(i)} - \hat{b})' / (k - 1),$$

$$\hat{B} = (H + H') / 2 \hat{\sigma}^2,$$

and

$$\hat{\sigma}^2 = \sum_i \sum_t P_t^{(i)} (Y_t^{(i)} - x'_{it} \hat{\beta}^{(i)})^2 / k(n - p).$$

The iterative procedure starts from an initial arbitrary non-negative definite symmetric matrix  $\hat{B}_0$ . It stops if, from one iteration to the next, the elements in  $\hat{B}$  do not change by more than a specified small quantity.

**Remark.** We can think of credibility models as consisting of two components. The first one in equation (2.1) models, for each group separately, the generation of the observations for given values of the coefficients  $\hat{\beta}^{(i)}$ ; we refer to this as the *data generating model*. The second component in equation (2.2) relates the parameters  $\hat{\beta}^{(i)}$  in the data generating model across the  $k$  groups; we refer to this as the *shrinkage component*. As mentioned above, a shortcoming of the traditional credibility model in equations (2.1) and (2.2) is that it does not

allow for time-varying coefficients. As a consequence the age of the observation does not enter into the analysis.

### 3. SHRINKAGE ESTIMATION IN MODELS WITH TIME-VARYING COEFFICIENTS

#### 3.1. Analysis of a single series

The following discussion concentrates on a single series (group) and, in order to simplify the presentation, we have omitted the group index  $i$ . In this paper we use structural time series models to incorporate time-varying coefficients into the data generating model. These models (see HARVEY and TODD, 1983; HARVEY, 1984) are of the form

$$(3.1) \quad Y_t = x_t' \beta_t + e_t, \quad e_t \sim N(0, \sigma^2/P_t)$$

$$\beta_t = T\beta_{t-1} + v_t, \quad v_t \sim N(0, \sigma^2 A).$$

As the notation indicates, the  $e_t$ 's are normal and independent with mean zero and variance  $\sigma^2/P_t$ , and the  $v_t$ 's are independent and multivariate normal with mean vector zero and covariance matrix  $\sigma^2 A$ . Furthermore  $e_t$  and  $v_t$  are mutually independent. Actuaries have used models of this type before. DE JONG and ZEHNWIRTH (1983), for example, use these models in the credibility context and show that the data generating equation of traditional credibility models can be formulated in this form. NEUHAUS (1987) applied this type of model to the prediction of number of policies, claim frequency and mean severity, and he discussed how to select the appropriate model and how to estimate its parameters. A recent application of these models in an insurance context is described by HARVEY and FERNANDEZ (1989) who combine a structural time series model for the size of claims with a model for the number of claims.

The simplest special case of the model in (3.1) assumes that  $p = 1$ ,  $x_t = 1$  and  $T = 1$ . This model allows the mean level  $\beta_t$  of the series to change over time according to a random walk,  $\beta_t = \beta_{t-1} + v_t$ . The exponentially weighted moving average forecasts that arise from this model (see ABRAHAM and LEDOLTER (1986), for example) are a special case of the recursive credibility model discussed by GERBER and JONES (1975) and its generalization by SUNDT (1981). If  $\text{Var}(v_t) = 0$ , implying that the coefficients  $\beta_t = \beta$  are time-invariant, then this model simplifies to the data generating equation of the Bühlmann-Straub model.

Another special case of interest is the model with a time-varying linear trend component where

$$(3.2) \quad x_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \beta_t = \begin{bmatrix} \beta_{0t} \\ \beta_{1t} \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

This model allows the slope  $\beta_{1t} = \beta_{1,t-1} + v_{2t}$  and the intercept  $\beta_{0t} = \beta_{0,t-1} + \beta_{1,t-1} + v_{1t}$  to change over time. With  $\lambda_1 = \lambda_2 = 0$  the model in (3.1) reduces to the data generating equation of the Hachemeister model.

If quarterly or monthly data are analyzed, it may be necessary to incorporate a seasonal component. A model with

$$(3.3) \quad \underline{x}_t = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \underline{\beta}_t = \begin{bmatrix} \beta_{0t} \\ \beta_{1t} \\ \gamma_t \\ \gamma_{t-1} \\ \gamma_{t-2} \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\underline{v}_t = \begin{bmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

can be used for quarterly data. The first two components in  $\underline{\beta}_t$  correspond to level and slope at time  $t$ . The last three components of  $\underline{\beta}_t$  correspond to additive seasonal factors. If the  $\lambda$ 's are zero, the model reduces to the Hachemeister linear trend model with seasonal indicators

The inference in structural time series models (3.1) is discussed in HARVEY and TODD (1983). The standard Kalman filter updating equations (see, for example, JAZWINSKI, 1970; MEINHOLD and SINGPURWALLA, 1983) are used to obtain  $\hat{\beta}_{n|n}$ , the estimate of  $\beta_n$  that is based on the observations  $Y_1, Y_2, \dots, Y_n$ . Furthermore, one can obtain its covariance matrix  $\sigma^2 G_{n|n}$ , predict future coefficients  $\beta_{n+i}$  from  $\hat{\beta}_{n+i|n} = T' \hat{\beta}_{n|n}$ , and future observations  $Y_{n+i}$  from  $\hat{Y}_n(i) = \underline{x}'_{n+i} T' \hat{\beta}_{n|n}$ .

Starting values are needed to initialize the Kalman filter recursions

$$(3.4) \quad \begin{aligned} \hat{\beta}_{t,t-1} &= T \hat{\beta}_{t-1,t-1} \\ \hat{\beta}_{t|t} &= \hat{\beta}_{t|t-1} + k_t (Y_t - \underline{x}'_t \hat{\beta}_{t|t-1}) \\ G_{t,t-1} &= T G_{t-1|t-1} T' + A \\ G_{t|t} &= G_{t|t-1} - k_t \underline{x}'_t G_{t|t-1} \\ k_t &= G_{t|t-1} \underline{x}_t (\underline{x}'_t G_{t|t-1} \underline{x}_t + P_t^{-1})^{-1}. \end{aligned}$$

For a single series in (3.1) we start these recursions with a  $p \times 1$  vector of zeros for  $\hat{\beta}_{0|0}$  and a diagonal matrix with very large diagonal elements for  $G_{0|0}$ . This

non-informative initialization reflects our ignorance about starting values in the absence of prior data. Other initialization approaches are possible (ANSLEY and KOHN, 1985; KOHN and ANSLEY, 1986, DEJONG, 1988), and their relationships are discussed in LEDOLTER, KLUGMAN and LEE (1989).

With a non-informative prior distribution the Kalman filter estimate  $\hat{\beta}_{n|n}$  is an unbiased estimator of the coefficient at time  $n$ ,  $\beta_n$ . The estimate is a weighted average of the  $n$  past observations. In general, older observations receive less weight if there is evidence that the coefficients are time-changing.

The Kalman filter updating equations, and therefore the estimate  $\hat{\beta}_{n|n}$  and the forecast  $\hat{Y}_n(l)$ , depend on the variance ratios  $A$  in equation (3.1). These parameters are estimated by maximum likelihood. The likelihood function of  $\sigma^2$  and  $A$  is obtained from the prediction error decomposition (SCHWEPPE, 1965). Assuming a non-informative initialization the log-likelihood function can be written as

$$(3.5) \quad l(\sigma^2, A; \text{data}) = c - \frac{n-p}{2} \log \sigma^2 - \frac{1}{2} \sum_{t=p+1}^n \log f_t \\ - \frac{1}{2\sigma^2} \sum_{t=p+1}^n (Y_t - x_t' \hat{\beta}_{t|t-1})^2 / f_t,$$

where  $Y_t - x_t' \hat{\beta}_{t|t-1}$  is the one-step-ahead prediction error at time  $t$ , and  $\sigma^2 f_t$  is its variance;  $\hat{\beta}_{t|t-1}$  and  $f_t = P_t^{-1} + x_t' G_{t|t-1} x_t$  can be obtained from the Kalman filter recursions. The maximization is simplified by the fact that one can concentrate the log-likelihood function with respect to  $\sigma^2$ ; the numerical maximization of the concentrated log-likelihood  $l_c(A, \text{data})$  needs to be carried out for elements in  $A$  only.

### 3.2. Analysis of multiple series and the introduction of shrinkage

So far we have discussed the analysis of a single series with time-varying coefficients. In insurance applications we not only have a single series, but we have  $n$  observations from  $k$  groups, and the estimation of  $A$  can be improved by incorporating information from the other groups. Here we assume that the  $A$  in the  $k$  groups are the same. As the value of  $n$  is usually small relative to  $k$ , it is not possible to estimate separate variance ratios for each series. Assuming independence across the  $k$  groups we can add the log-likelihood functions in (3.5) for the  $k$  groups and obtain estimates of a common  $A$  via numerical optimization. An estimate of the variance  $\sigma^2$  is obtained from

$$(3.6) \quad \hat{\sigma}^2 = \frac{1}{(n-p)k} \sum_{i=1}^k \sum_{t=p+1}^n (Y_t^{(i)} - x_{it}' \hat{\beta}_{t|t-1}^{(i)})^2 / f_t^{(i)}.$$

The estimate of  $A$  is used to carry out the Kalman filter recursions. This is done for each group separately, using a non-informative initialization. The resulting coefficient estimate  $\hat{\beta}_{n|n}^{(i)}$  provides us with an estimate of the parameter



at time  $n$ ,  $\beta_n^{(i)}$ ; its covariance matrix is given by  $\sigma^2 G_n^{(i)}$ . The estimate is a weighted average of the  $n$  observations. The estimate of  $A$  determines the weights in this average. Positive variance ratios in  $A$  imply that the importance of each observation in determining the estimate depends on its age. If the variance ratios are zero, then the Kalman filter estimates simplify to the usual regression estimates  $\hat{\beta}_n^{(i)}$  in equation (2.4).

So far there has been no shrinkage, as we have ignored the cross-sectional correlations. In order to effect shrinkage we introduce a second equation,

$$(3.7) \quad \beta_n^{(i)} = \underline{b}_n + \underline{g}_n^{(i)} \quad \text{where} \quad \underline{g}_n^{(i)} \sim N(0, \sigma^2 B_n).$$

This equation specifies that at time  $n$  the coefficient vectors in the structural time series model for the  $k$  groups vary independently around a common value  $\underline{b}_n$ . We combine this equation with the results from the  $k$  separate Kalman filters,

$$(3.8) \quad \hat{\beta}_{n|n}^{(i)} = \underline{\beta}_n^{(i)} + \underline{w}_n^{(i)} \quad \text{where} \quad \underline{w}_n^{(i)} \sim N(0, \sigma^2 G_{n|n}^{(i)})$$

are independent across groups. These two equations yield the standard two-stage credibility model in Section 2. The shrinkage estimate based on (3.7) and (3.8) is given by

$$(3.9) \quad \tilde{\beta}_n^{(i)} = Z_i \hat{\beta}_{n|n}^{(i)} + (I - Z_i) \underline{b}_n,$$

where  $Z_i = B_n(B_n + V_i)^{-1}$  and  $V_i = G_{n|n}^{(i)}$ . The results in Section 2 can be used to estimate  $\underline{b}_n$  and  $B_n$ . In our examples we have used deVylder's iterative approach discussed in Section 2.

### 3.3. Discussion

Adding this second equation to induce shrinkage is somewhat heuristic, but is needed as by itself the model in equation (3.1) does not incorporate cross-sectional correlations.

In theory, a cross-correlation structure can be introduced by specifying a certain covariance structure for the error terms in a multivariate version of the model in (3.1). However, it is usually quite difficult to identify the exact form of the cross-correlation structure, especially for the short time series which are typical with insurance data. We have avoided these modelling issues by introducing a heuristic shrinkage equation at the last available observation period.

Model-based approaches to shrinkage are clearly possible. One alternative to the above heuristic shrinkage approach is a model that introduces a shrinkage equation for the coefficient vector at the initial time period zero. That is, one assumes that  $\beta_0^{(i)} = \underline{b}_0 + \underline{a}_0^{(i)}$ , where the  $\underline{a}_0^{(i)}$ , for  $i = 1, \dots, k$ , are independent realizations from a normal distribution with mean vector zero and covariance matrix  $\sigma^2 B_0$ . This implies that at the initial time period the standard actuarial shrinkage model is valid. If the elements in  $A$  are zero, implying that the coefficients in the data generating model are time-invariant, this model and the

traditional credibility model are identical. For time-varying coefficient models we start from the standard actuarial shrinkage model at time zero, but assume that the coefficients for subsequent periods are subject to stochastic change. For the inference in this model one initializes the Kalman filter in each group by the same  $\hat{\beta}_{0|0}^{(i)} = \underline{b}_0$  and  $\mathcal{G}_{0|0}^{(i)} = B_0$ , treats  $\underline{b}_0$  and  $B_0$  as unknown parameters, and simultaneously obtains estimates of  $\lambda$ ,  $\underline{b}_0$  and  $B_0$ . This results in shrinkage of the Kalman filter estimates  $\hat{\beta}_{n|n}^{(i)}$  at time  $n$  towards the common initial mean  $\underline{b}_0$ . But even for modest positive values of  $\lambda$  this shrinkage effect disappears very quickly as  $n$  increases, and for moderate  $n$  there is hardly any shrinkage. It is for these reasons that we have rejected this alternative approach and have concentrated our discussion on the former, somewhat heuristic procedure.

Another model that introduces cross-sectional correlations is one that assumes that the  $k$  coefficients at time  $t$ ,  $\beta_i^{(i)}$ , for  $i = 1, \dots, k$ , vary independently around a common trend component  $\tilde{b}_t$ , which itself follows a structural time series model. LEE (1991) studies these common-trend type models in detail, and we hope to report on this work in a future paper.

The advantage of our admittedly heuristic method is that it is more general than the traditional credibility approach. It recognizes the fact that most time series exhibit changing levels, trends and seasonality, and it discounts previous observations when it determines their estimates. The difference between the two approaches is shown best in the case of the Bühlmann-Straub model. The traditional approach shrinks the sample means towards a common average, whereas our new approach shrinks exponentially weighted averages. Furthermore, it can be shown that for  $\lambda = 0$  our approach coincides with the solution in Section 2.

#### 4. EXAMPLES

In this section two examples are given, with the second one being analyzed in detail. These examples provide illustrations of situations in which models that combine time-varying and shrinkage aspects are likely to improve the results

##### 4.1. Worker's compensation

MEYERS (1984) studies yearly loss ratios under Worker's compensation insurance for 319 classes (occupation groups) and three years. A model without trend component is appropriate since these data are already adjusted for inflation. Meyers uses the Bühlmann-Straub model in his analysis. However, MEYERS and SCHENKER (1983) provide evidence that the loss ratios are not constant, but vary independently from year to year around a common mean. In the notation of our present paper

$$(4.1) \quad \underline{x}_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \underline{\beta}_t = \begin{bmatrix} \beta_{0,t} \\ \beta_{1,t} \end{bmatrix}, \quad T = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & 0 \end{bmatrix},$$

where  $\beta_{1,t}$  is an unchanging long-term average and  $\beta_{0,t}$  is the level in year  $t$ . An approach that combines this state-space model with shrinkage can be expected to improve the forecasts for future losses, as many of the 319 classes have very small sample sizes.

#### 4.2. Automobile bodily injury

The data for the second example are taken from the automobile insurance industry. Quarterly data on the amount (not adjusted for inflation) paid under the bodily injury component of automobile insurance policies (LOSS) and the number of cars covered by these policies (EXPOSURE) were obtained from 31 states. Only states without no-fault laws were included, as under no-fault

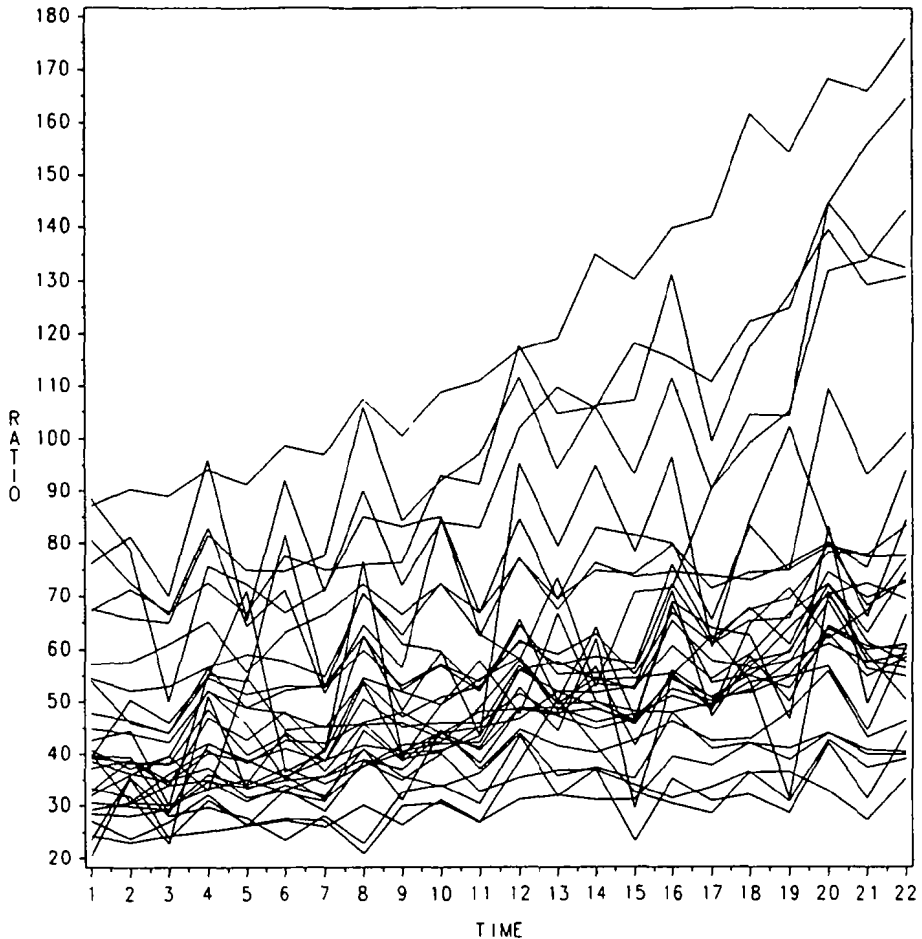


FIGURE 1 Multiple time series plot of the ratio  $R = \text{LOSS}/\text{EXPOSURE}$

laws many claims that would otherwise be covered by the liability portion of the insurance are paid under the bodily injury component. Data from the first quarter of 1983 to the second quarter of 1988 ( $n = 22$ ) are used in our analysis.

The ratio  $R_t^{(i)} = \text{LOSS}_t^{(i)} / \text{EXPOSURE}_t^{(i)}$ , where  $t = 1, \dots, 22$  (quarters) and  $i = 1, \dots, 31$  (states) is our dependent variable that needs to be predicted. The multiple time series plot of the ratios  $R_t^{(i)}$  in Figure 1 shows presence of seasonality and a need for a logarithmic transformation. The presence of seasonality is seen more clearly in Figure 2 where we have plotted estimates of the multiplicative seasonal indices for the 31 states. We use the following procedure to obtain the seasonal indices. For each univariate series we calculate centered yearly moving averages to estimate the trend component; we then obtain, for each time period  $t$ , an estimate of the seasonal factor from the ratio of the observation and the corresponding centered moving average, next, we average the seasonal factors for each quarter to obtain seasonal indices for the four quarters; finally, we normalize these indices so that they sum to four. The dot plot of these normalized seasonal indices in Figure 2 shows a seasonal pattern; in the fourth quarter the ratios  $R_t^{(i)}$  tend to be highest

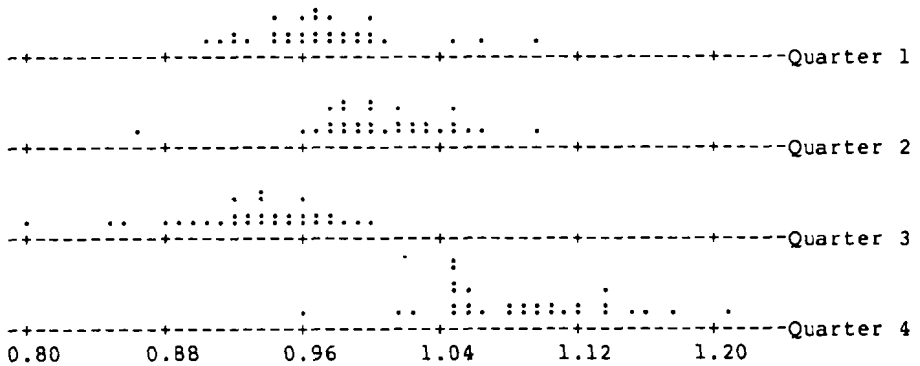


FIGURE 2 Dot plot of the seasonal indices for the 31 states

A multiple time series plot of the transformed observations,  $Y_t^{(i)} = \log R_t^{(i)}$ , is given in Figure 3. This plot indicates that a linear trend model with additive seasonal components provides a good description of the transformed observations.

In the standard actuarial model it is usually assumed that the variance of the error component is related to the exposure  $P_t^{(i)}$ ; that is,  $\text{Var}(e_t^{(i)}) = \sigma^2 / P_t^{(i)}$ . We now want to check whether this is a reasonable assumption. Since the exposures  $P_t^{(i)}$  do not change much over time, we calculate an average exposure  $\bar{P}^{(i)}$  for each state. Due to size differences among the states, these averages are quite different. Next, we adjust each time series  $Y_t^{(i)}$  for trend and seasonality and calculate an estimate of its variance. The residuals from a regression of  $Y_t^{(i)}$  on time  $t$  and additive seasonal indicators are used to calculate the variance

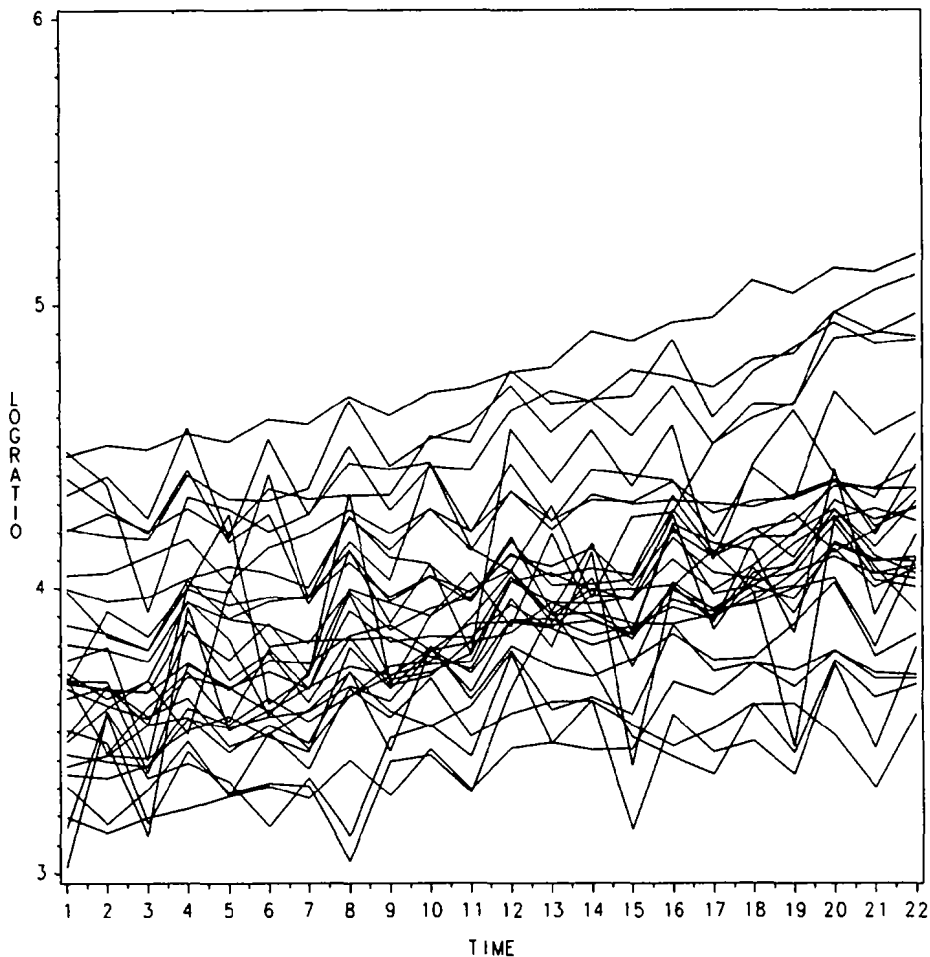


FIGURE 3 Multiple time series plot of  $Y = \log (\text{LOSS/EXPOSURE})$

estimate. In Figure 4 we plot the resulting mean square errors against the reciprocal of the average exposures. The linear relationship confirms that  $\text{Var}(e_t^{(i)}) = \sigma^2/P_t^{(i)}$  is a reasonable assumption

Based on this preliminary analysis we are led to consider the structural time series model with a linear trend and additive seasonal components,

$$(4.2) \quad \begin{aligned} Y_t^{(i)} &= \underline{x}_t' \beta_t^{(i)} + e_t^{(i)} & e_t^{(i)} &\sim N(0, \sigma^2/P_t^{(i)}) \\ \beta_t^{(i)} &= T\beta_{t-1}^{(i)} + \underline{v}_t^{(i)} & \underline{v}_t^{(i)} &\sim N(0, \sigma^2 A) \end{aligned}$$

where  $\underline{x}_t$ ,  $T$  and  $A$  are given in equation (3.3). Our model allows for time-varying coefficients and reduces to a linear trend regression model with quarterly indicators if  $\lambda_1 = \lambda_2 = \lambda_3 = 0$ .

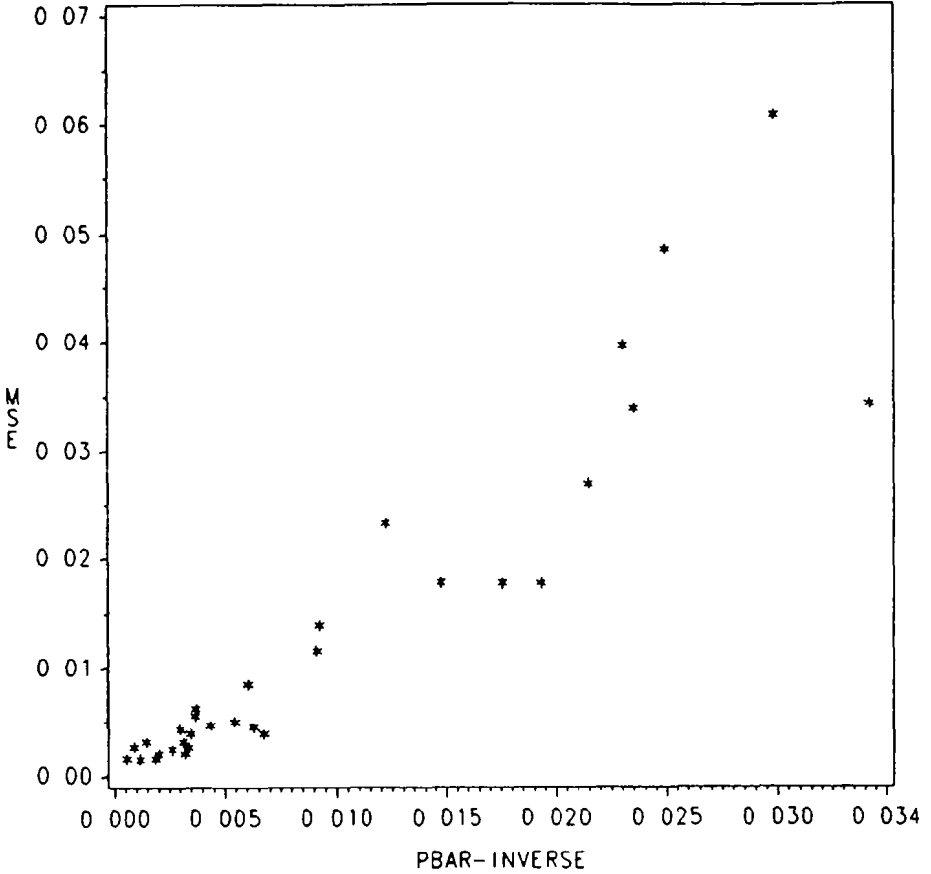


FIGURE 4 Plot of the mean square error from the linear trend regression with seasonal indicators against the reciprocal of average exposure,  $k = 31$  states

The maximum likelihood estimation approach in Section 3 is applied and, assuming independent groups, estimates of  $\sigma^2$  and the three variance ratios are obtained. It is found that  $\hat{\sigma}^2 = 3.8089 \cdot 10^{-3}$ ,  $\hat{\lambda}_1 = 0.0495$ ,  $\hat{\lambda}_2 = 0.0044$  and  $\hat{\lambda}_3 = 0.00008$ . The estimate  $\hat{\lambda}_3$  is close to zero and the log-likelihood deficiency (ratio),  $l_c(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3) - l_c(\hat{\lambda}_1, \hat{\lambda}_2, 0)$ , is quite small. This implies that the seasonal coefficients do not change much over time. Contours of the log-likelihood function of  $\lambda_1$  and  $\lambda_2$ , for  $\hat{\lambda}_3 = 0.00008$ , are plotted in Figure 5. This plot, as well as the large log-likelihood deficiency  $l_c(\hat{\lambda}_1, \hat{\lambda}_2, 0) - l_c(0, 0, 0) = 19.16$ , shows that a standard least squares approach which assumes time constant intercept and slope coefficients would be inappropriate.

In order to check the adequacy of the structural time series model in equation (4.2) we calculated the standardized one-step-ahead forecast errors for periods 6 through 22. Standardization of the forecast error by its standard error  $\hat{\sigma} f_t^{1/2}$  assures that its variance does not depend on time. We found that

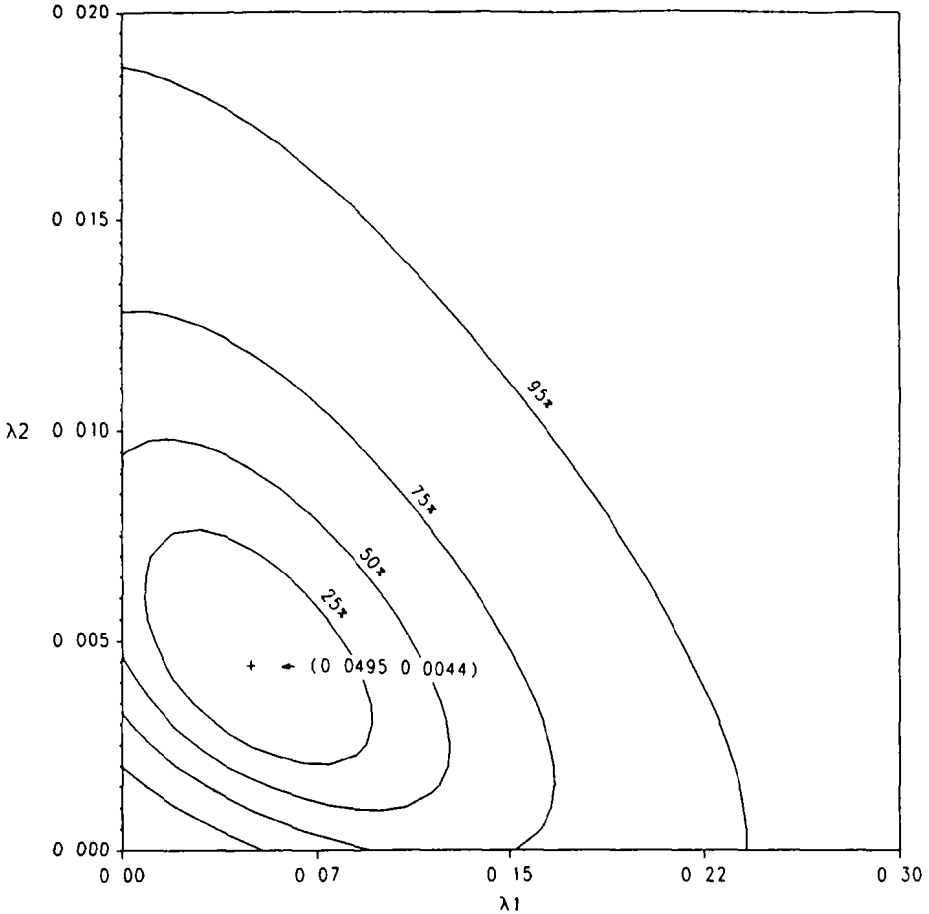


FIGURE 5 Contour plot of the concentrated log-likelihood function. The variance ratio that determines the variability in the seasonal component is set at its estimate  $\hat{\lambda}_3 = 0.00008$ . Probability coverage of the contours is obtained from the Chi-square approximation.

the standardized one-step-ahead forecast errors were serially uncorrelated for essentially all 31 groups.

The estimates of  $\hat{\lambda}_1$ ,  $\hat{\lambda}_2$  and  $\hat{\lambda}_3$  are used to calculate the estimates  $\hat{\beta}_{n|n}^{(i)}$ , for  $n = 22$  (the last available time period) and  $i = 1, \dots, 31$  (states). Dot diagrams of the  $k = 31$  estimates of intercept, slope and seasonal coefficients (only the first one is shown), together with their standard errors, are given in Figure 6. The standard errors are obtained from the diagonal elements in  $\hat{\sigma}^2 G_{n|n}^{(i)}$ .

We notice considerable variability among the  $k = 31$  intercept estimates. Furthermore, we find that the between group variability is much larger than the uncertainty that is associated with each estimate (that is, the within group variability as measured by the standard error of the estimate). This result indicates that there should be no or little benefit to shrink the intercept

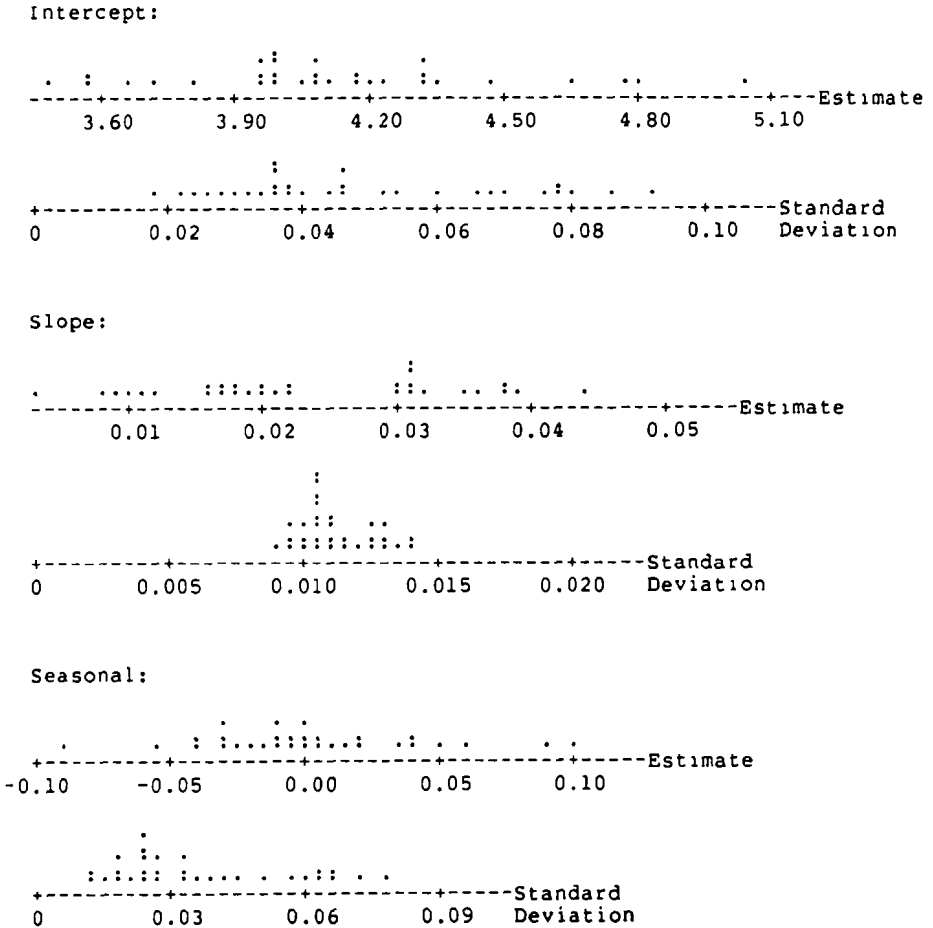


FIGURE 6 Dot plots of the estimates and their standard errors for the intercept, slope, and seasonal coefficients in model (4.2),  $k = 31$  states

estimates. The dot plots of the slope estimates and their standard errors show a different picture; the within group variability is quite large when compared with the variability between the slope estimates. These pictures suggest that shrinkage procedures should pool the slope estimates towards a common value. The same conclusion is reached for the seasonal factors (the third, fourth and fifth component of the beta vector). They, too, should be shrunk towards common means.

Next, we apply shrinkage and calculate the shrinkage estimate discussed in equation (3.9) of Section 3. That is, we compute

$$(4.3) \quad \tilde{\beta}_n^{(t)} = Z_t \hat{\beta}_{nn}^{(t)} + (I - Z_t) \hat{b}_n$$



where  $Z_i = B_n(B_n + V_i)^{-1}$  and  $V_i = G_{n|n}^{(i)}$ . DeVlyder's modification in (2.7) is used to estimate  $\hat{b}_n$  and  $B_n$ . The only minor difference is that we are using the maximum likelihood estimate  $\hat{\sigma}^2 = 3.8089 * 10^{-3}$  from the Kalman filter as the estimate of  $\sigma^2$ . In Figure 7 we compare the estimates before and after shrinkage. The graphs confirm what we had anticipated from the results in Figure 6. The slopes and seasonal components are shrunk towards their respective means, whereas the intercepts are essentially unchanged.

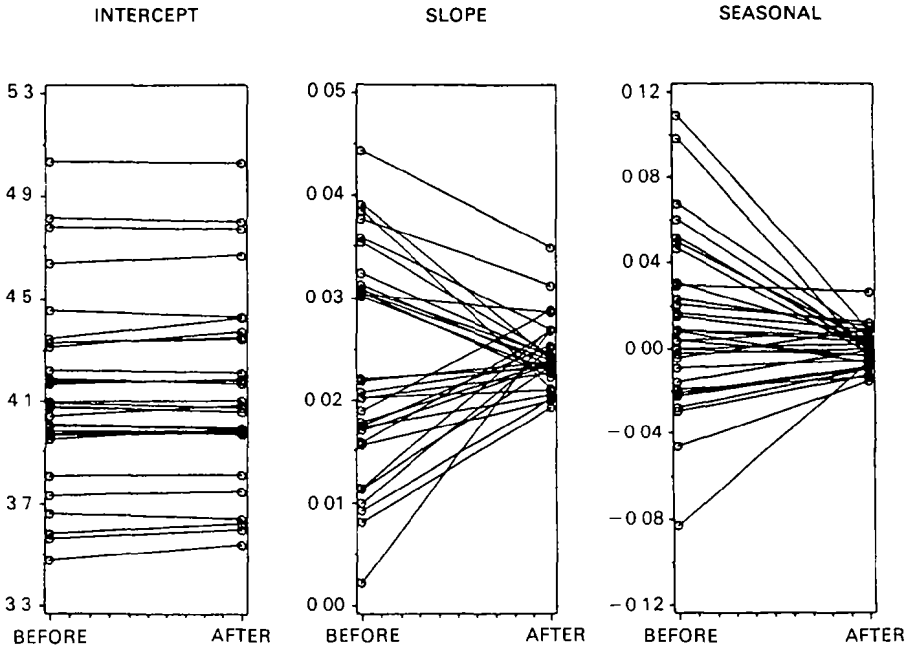


FIGURE 7 Intercept, slope and first seasonal coefficient estimate in model (4.2) before and after shrinkage,  $k = 31$  states

**Forecast comparisons**

The prediction of future values is a major reason for fitting models to data. We must now investigate whether the proposed new approach leads to forecast improvements. In particular, we address the following two questions:

- (1) Has shrinkage of the coefficients improved the forecasting performance of our time-varying trend component model? To address this issue we compare forecasts that are calculated from the shrinkage estimates  $\hat{\beta}_n^{(i)}$  in (4.3) [method 1] and forecasts that are calculated from the standard Kalman filter estimates  $\hat{\beta}_{n|n}^{(i)}$  [method 2].

- (2) Has our generalization of incorporating time-varying trend components helped the forecasting? To investigate this question we compare the forecasts that use the shrinkage estimates  $\hat{\beta}_n^{(i)}$  in (4.3) [method 1] with forecasts that are calculated from the shrinkage estimates in the standard regression model with constant linear trend and seasonal indicators [Hachemeister, method 3].

A true test of the forecast performance of a model is obtained by an out-of-sample comparison of forecasts and actual observations. Here we use the last four observations  $R_{19}^{(i)}$  through  $R_{22}^{(i)}$ , for  $i = 1, \dots, 31$ , as our hold-out sample. This is a reasonable choice as actuarial practice bases predictions of future premiums on about four to five years of past data. For each state we calculate four one-step-ahead forecast errors  $R_t - \hat{R}_{t-1}(1)$ ,  $t = 19, \dots, 22$ , where  $\hat{R}_t(1) = \exp[\hat{Y}_t(1)]$  is obtained by applying the inverse transformation to the forecast of the logarithmically transformed data. For each state separately, we then compute the mean square error MSE, the mean absolute deviation (error) MAD, and the mean absolute percent error MAPE. For each measure (MSE, MAD, MAPE) and for each method (methods 1 through 3) we calculate a weighted average that combines the information from the 31 states. The average exposures  $\bar{P}^{(i)}$ ,  $i = 1, \dots, 31$ , are used as weights. The results are given in Table 1. Table 1 also shows the results of a further refinement of method 1 (Kalman filter with shrinkage). In method 1R we shrink the last 4 components of the 5-dimensional coefficient vectors, but leave the first components (intercepts) unchanged.

TABLE 1  
WEIGHTED AVERAGES OF ACCURACY MEASURES. AVERAGE EXPOSURES ARE USED TO COMBINE  
THE INFORMATION FROM  $k = 31$  STATES

	Method 1 Kalman filter model (4.2) with shrinkage		Method 2 Kalman filter model (4.2) without shrinkage	Method 3 Hachemeister constant linear trend & seasonal indicator model with shrinkage
	I	I R		
MSE	32.28	31.88	39.24	38.02
MAD	3.75	3.70	4.20	4.32
MAPE	5.12	4.99	5.35	5.40

In addition to the comparison of the aggregate measures, we compare the measures for each state separately. We assign a score of 1 if in state  $i$  the first method leads to a lower MSE (MAD, MAPE) than the second. The proportion of states where method 1 outperforms method 2 (method 3) is given in Table 2.

TABLE 2  
PROPORTION OF STATES WHERE ONE METHOD OUTPERFORMS THE OTHER

Comparison	MSE	MAD	MAPE
Method 1 vs Method 2	58	55	58
Method 1 R vs Method 2	71	65	61
Method 1 vs Method 3	61	58	61
Method 1 R vs Method 3	65	55	55
Method 2 vs Method 3	55	52	55
Method 1 R vs Method 1	52	48	52

**Comments.** (i) For shrinkage methods we calculate the forecasts  $\hat{Y}_t(l)$  after shrinking the estimates that are obtained at time  $t$ . We carry out a new shrinkage if we go to another forecast origin. (ii) The Kalman filter methods 1 and 2 require estimates of the variance ratios  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$ . In order to avoid the numerical maximization of the log-likelihood for each forecast origin  $t$ , we use the estimates that are obtained from the complete data set ( $n = 22$ ). (iii) The transformation  $\hat{R}_t(l) = \exp[\hat{Y}_t(l)]$  results in the median of the predictive distribution of  $R_{t+l}$ . The mean of the predictive distribution can be obtained by incorporating the variance of the predictive distribution into the inverse transformation (see GRANGER and NEWBOLD, 1976). Because differences are usually relatively minor and because it is not obvious whether the mean of the posterior distribution is preferable to the median we have not pursued this adjustment.

### Interpretation of results

Table 1 shows that we can improve the one-step-ahead forecast performance if we allow the trend and the seasonal components to change over time. Comparing the results of the two shrinkage methods (methods 1 and 3) we find that the structural time series model in (4.2) leads to a 15.1, (16.1), 13.2 (14.4), and 5.2 (7.6) percent reduction in MSE, MAD, and MAPE, when it is compared to the Hachemeister model with fixed trend and seasonal components. The numbers in parentheses reflect the improvements if shrinkage is not applied to the intercepts in the structural time series model. Table 2 leads to a similar conclusion. The one-step-ahead forecasts from the structural time series model with shrinkage outperform the forecasts from the Hachemeister model in roughly 60 percent of the states (the proportion varies from 55 to 65 percent, depending on the accuracy measure that is used in the comparison).

Tables 1 and 2 also show that shrinkage of the coefficients improves the forecasts in the structural time series model (4.2). The size of the improvements that are due to shrinkage (method 1 vs method 2) is roughly the same as the one we obtain by allowing the trend and seasonal coefficients in the two

shrinkage methods to change over time (method 1 vs method 3) There is very little difference between the forecasts from the structural time series model without shrinkage and the Hachemeister shrinkage model with fixed trend and seasonal coefficients (method 2 vs method 3).

This example shows the feasibility of an approach that applies shrinkage to the coefficient estimates in structural time series models and illustrates its potential for forecast improvements GARCIA-FERRER et al. (1987) and ZELLNER and HONG (1989) reach a similar conclusion in their analysis of macroeconomic data. They find that individual country growth rate forecasts are improved by shrinking the forecasts to a common average However, their shrinkage methods are somewhat different from the ones considered in this paper. Furthermore, they apply shrinkage primarily to forecasts and not to estimates in time-varying coefficient models.

#### ACKNOWLEDGMENTS

STUART KLUGMAN'S research was in part supported by the Huebner Foundation. The authors wish to thank GLENN MEYERS and the Insurance Services Office for the data in Section 4. They also thank WILLIAM BELL, ANDREW HARVEY, WILLIAM WECKER and the referees for helpful comments.

#### REFERENCES

- ABRAHAM, B and LEDOLTER, J (1983) *Statistical Methods for Forecasting* New York Wiley
- ABRAHAM, B and LEDOLTER, J (1986) Forecast Functions Implied by ARIMA Models and Other Related Forecast Procedures *International Statistical Review* 54, 51-66
- ANSLEY, C and KOHN, R (1985) Estimation, Filtering, and Smoothing in State Space Models with Incompletely Specified Initial Conditions *Annals of Statistics* 13, 1286-1316
- BAILEY, R A. and SIMON, L J (1959) An Actuarial Note on the Credibility of Experience of a Single Private Passenger Car *Proceedings of the Casualty Actuarial Society* 46, 159-164
- BOX, G and JENKINS, G (1976) *Time Series Analysis, Forecasting, and Control* (2nd ed) San Francisco Holden-Day
- BÜHLMANN, H and STRAUB, E (1972) Credibility for Loss Ratios *Actuarial Research Clearing House*, Number 2
- DE JONG, P (1988) The Likelihood for a State Space Model *Biometrika* 75, 165-169
- DE JONG, P and ZEHNWIRTH, B (1983) Credibility Theory and the Kalman Filter *Insurance Mathematics and Economics* 2, 281-286
- DEVYLDER, F (1981) Practical Credibility Theory with Emphasis on Optimal Parameter Estimation *ASTIN Bulletin* 12, 115-131
- DEVYLDER, F (1984) Practical Models in Credibility Theory, Including Parameter Estimation In *Premium Calculation in Insurance*, DEVYLDER, F, GOOVAERTS, M and HAEZENDONCK, J, editors, Reidel
- GARCIA-FERRER, A, HIGHFIELD, R A, PALM, F and ZELLNER, A (1987) Macroeconomic Forecasting using Pooled International Data *Journal of Business and Economic Statistics* 5, 53-67
- GERBER, H and JONES, D (1975) Credibility Formulas of the Updating Type *Transactions of the Society of Actuaries* 27, 31-46
- GRANGER, C W J and NEWBOLD, P (1976) Forecasting Transformed Series *Journal of the Royal Statistical Society, Series B* 38, 189-203
- HACHEMEISTER, C (1975) Credibility for Regression Models with Applications to Trend In *Credibility Theory and Applications*, P KAHN, ed, New York Academic Press

- HARVEY, A (1984) A Unified View of Statistical Forecasting Procedures *Journal of Forecasting* 3, 245-283
- HARVEY, A and TODD, P (1983) Forecasting Economic Time Series with Structural and Box-Jenkins Models A Case Study *Journal of Business and Economic Statistics* 1, 299-315
- HARVEY, A and FERNANDES, C (1989) Time Series Models for Insurance Claims *Journal of the Institute of Actuaries* 116, Part 3, 513-528
- JAZWINSKI, A (1970) *Stochastic Processes and Filtering Theory* New York Academic Press
- KLUGMAN, S (1987) Credibility for Classification Ratemaking via the Hierarchical Linear Model *Proceedings of the Casualty Actuarial Society* 74, 272-321
- KOHN, R and ANSLEY, C (1986) Estimation, Prediction, and Interpolation for ARIMA Models with Missing Data *Journal of the American Statistical Association* 81, 751-761
- LEDOLTER, J, KLUGMAN, S and LEE, C S (1989) Credibility Models with Time-Varying Trend Components, Technical Report 159, Department of Statistics and Actuarial Science, University of Iowa
- LEE, C S (1991) Time Series Models for the Credibility Estimation of Insurance Premiums, unpublished Ph D dissertation (forthcoming), Department of Statistics and Actuarial Science, University of Iowa
- MEINHOLD, R and SINGPURWALLA, N (1983) Understanding the Kalman Filter *The American Statistician* 37, 123-127
- MEYERS, G (1984) Empirical Bayesian Credibility for Workers' Compensation Classification Ratemaking *Proceedings of the Casualty Actuarial Society* 71, 96-121
- MEYERS, G and SCHENKER, N (1983) Parameter Uncertainty in the Collective Risk Model *Proceedings of the Casualty Actuarial Society* 70, 111-143
- MOWBRAY, A (1914) How Extensive a Payroll is Necessary to Give a Dependable Pure Premium? *Proceedings of the Casualty Actuarial Society* 1, 24-30
- NEUHAUS, W (1987) Early Warning *Scandinavian Actuarial Journal*, 128-156
- SCHWEPPE, F (1965) Evaluation of Likelihood Functions for Gaussian Signals *IEEE Trans Inf Theory* 11, 61-70
- SUNDT, B (1981) Recursive Credibility Estimation *Scandinavian Actuarial Journal*, 3-21
- SWAMY, P A V B (1971) *Statistical Inference in Random Coefficient Regression Models* New York Springer
- ZELLNER, A and HONG, C (1989) Forecasting International Growth Rates Using Bayesian Shrinkage and other Procedures *Journal of Econometrics* 40, 183-202

JOHANNES LEDOLTER  
*Department of Statistics and Actuarial Science*  
*Department of Management Sciences*  
*The University of Iowa*  
*Iowa City, IA 52242*

STUART KLUGMAN  
*College of Business and Public Administration*  
*Drake University*  
*Des Moines, IA 50311*

CHANG-SOO LEE  
*Department of Statistics and Actuarial Science*  
*The University of Iowa*  
*Iowa City, IA 52242*



# A SIMPLE PARAMETRIC MODEL FOR RATING AUTOMOBILE INSURANCE OR ESTIMATING IBNR CLAIMS RESERVES

BY THOMAS MACK

*Munich Re, Munich, FRG*

## ABSTRACT

It is shown that there is a connection between rating in automobile insurance and the estimation of IBNR claims amounts because automobile insurance tariffs are mostly cross-classified by at least two variables (e.g. territory and driver class) and IBNR claims run-off triangles are always cross-classified by the two variables accident year and development year. Therefore, by translating the most well-known automobile rating methods into the claims reserving situation, some known and some unknown claims reserving methods are obtained. For instance, the automobile rating method of BAILEY and SIMON produces a new claims reserving method, whereas the model leading to the rating method called "marginal totals" produces the well-known IBNR claims estimation method called "chain ladder". A drawback of this model is the fact that it is designed for the number of claims and not for the total claims amount for which it is usually applied.

As an alternative for both, rating and claims reserving, we describe a simple but realistic parametric model for the total claims amount which is based on the Gamma distribution and has the advantage of providing the possibility of assessing the goodness-of-fit and calculating the estimation error. This method is not very well known in automobile insurance—although a satisfactory application is reported—and seems to be completely unknown in the field of claims reserving, although its execution is nearly as simple as that of the chain ladder method.

## KEYWORDS

Cross-classified data; (automobile & property) ratemaking; IBNR claims; Gamma model; maximum likelihood method.

## CONTENTS OF THE PAPER

1. A short overview of some automobile rating methods
2. Some methods of estimating IBNR claims reserves and their connection to automobile rating methods

3. A parametric model for rating automobile insurance or estimating IBNR claims reserves
  4. Statistical analysis of the model
  5. Improvement of the model in the case of known claims numbers
  6. Final remark
- App. A. Proof that the chain ladder method can be derived from the marginal totals conditions (and therefore is maximum likelihood in the Poisson case)
- App. B. Estimation of the (asymptotic) error variances
- Acknowledgement
- References

#### 1. A SHORT OVERVIEW OF SOME AUTOMOBILE RATING METHODS

In the automobile insurance tariffs of many countries several tariff variables are used, e.g. the horse-power class of the car, the bonus/malus (or no claims discount) class of the driver or the class of the territory where the car is principally garaged. In this way the portfolio of automobile insurance policies is cross-classified into a number of cells which are each supposed to be homogeneous, so that all policies of the same cell pay the same premium. For the sake of simplicity we will consider in the following only two tariff variables, which are subdivided into  $m$  and  $n$  classes respectively. When then have  $mn$  cells labelled  $(i, j)$ ,  $i = 1, \dots, m$ ,  $j = 1, \dots, n$ . Now let  $n_{ij}$  be the known number of insureds (policy years) of cell  $(i, j)$  and  $s_{ij}$  their observed total claims amount as realization of the random variable  $S_{ij}$ . For some of the cells,  $n_{ij}$  may be so small that it is not advisable to use  $s_{ij}$  as the only basis for the calculation of the net premium  $E(S_{ij})/n_{ij}$  of that cell. Therefore one searches for marginal parameters  $x_i$ ,  $i = 1, \dots, m$ , and  $y_j$ ,  $j = 1, \dots, n$ , with

either  $x_i y_j = E(S_{ij})/n_{ij}$  (multiplicative approach),

or  $x_i + y_j = E(S_{ij})/n_{ij}$  (additive approach).

This also reduces the number of figures needed to describe the tariff premiums from  $mn$  to  $m+n$ . In the following we only consider the multiplicative approach, but the methods described can easily be translated to the additive approach, too.

The problem of finding appropriate marginal parameters  $x_i$  and  $y_j$  is one of the classical problems of insurance mathematics. It has been known for a long time that the simple marginal averages

$$x_i = s_{i+} / n_{i+}$$

$$y_j = (s_{+j} / n_{+j}) / (s_{++} / n_{++})$$

(where a '+' indicates summation over the corresponding index) give a satisfactory approximation of  $E(S_{ij})/n_{ij}$  only if the tariff variables are independent. But generally this is not the case. Therefore, in the last 30 years several different methods have been proposed. We will now shortly review three of the



most well-known mainly following the description given by VAN EEGHEN/GREUP/NIJSSEN (1983) For a more comprehensive and more recent comparative analysis see JEE (1989)

The first breakthrough was achieved by BAILEY/SIMON (1960), who estimated  $x_i, y_j$  by minimizing

$$\begin{aligned} Q &= \sum_{i=1}^m \sum_{j=1}^n (s_{ij} - n_{ij} x_i y_j)^2 / (n_{ij} x_i y_j) \\ &= \sum_{i=1}^m \sum_{j=1}^n n_{ij} (s_{ij} / n_{ij} - x_i y_j)^2 / (x_i y_j), \end{aligned}$$

but their underlying assumption of  $Q$  having (up to a factor) the distribution of a chi-square will normally not be true (see VAN EEGHEN/GREUP/NIJSSEN 1983) Moreover, it can be shown (VAN EEGHEN/NIJSSEN/RUYGT 1982) that for the minimizing parameters  $x_i, y_j$  the inequalities

$$\begin{aligned} \sum_{j=1}^n n_{ij} x_i y_j &\geq \sum_{j=1}^n s_{ij}, \quad i = 1, \dots, m, \\ \sum_{i=1}^m n_{ij} x_i y_j &\geq \sum_{i=1}^m s_{ij}, \quad j = 1, \dots, n, \end{aligned}$$

hold, i.e. there results an overestimation of all marginal loss amounts (in the multiplicative case only).

Therefore BAILEY (1963) and later JUNG (1968) proposed estimating  $x_i, y_j$  directly from the intuitively appealing conditions

$$(1a) \quad \sum_{j=1}^n n_{ij} x_i y_j = \sum_{j=1}^n s_{ij}, \quad i = 1, \dots, m,$$

and

$$(1b) \quad \sum_{i=1}^m n_{ij} x_i y_j = \sum_{i=1}^m s_{ij}, \quad j = 1, \dots, n,$$

which can be solved iteratively: starting with, for example,  $y_j = 1$ , (1a) results in  $x_i = s_{i+} / n_{i+}$ , which is inserted in (1b) giving new  $y_j$  etc. The procedure converges quickly. This method has been called "marginal totals" If the random variables  $S_{ij}$  denote the number of claims instead of the total claims amount, then this method can be shown to be maximum likelihood under the assumption that all  $S_{ij}$  are independent and Poisson distributed with parameter  $n_{ij} x_i y_j$  (see VAN EEGHEN/GREUP/NIJSSEN 1983, p. 93). But for the more important case where  $S_{ij}$  is the total claims amount one has no model from which the equations above derive and thus, for example, a statistical test of the goodness-of-fit cannot be designed either

SANT (1980) proposed estimating  $x_i, y_j$  by the method of weighted least squares, i.e. by minimizing

$$\sum_{i=1}^m \sum_{j=1}^n (s_{ij} - n_{ij} x_i y_j)^2 / n_{ij} = \sum_{i=1}^m \sum_{j=1}^n n_{ij} (s_{ij} / n_{ij} - x_i y_j)^2$$

But the powerful tools of regression analysis like the  $R^2$ -statistic, the analysis of residuals and the estimation of the prediction error can only be applied rigorously if all  $S_{ij}$  are normally distributed with  $\text{Var}(S_{ij})$  proportional to  $n_{ij}$ . Both assumptions are not very realistic.

Using the additive approach, the weighted least squares method leads to the same equations for the marginal parameters  $x_i, y_j$  as the marginal totals method, which in this case is no longer the maximum likelihood estimator for Poisson distributed numbers of claims.

Altogether, in the case of  $S_{ij}$  being claims totals all three methods described above are only of a heuristic nature without an underlying realistic model.

## 2. SOME METHODS OF ESTIMATING IBNR CLAIMS RESERVES AND THEIR CONNECTION TO AUTOMOBILE RATING METHODS

We now turn to the problem of estimating IBNR claims reserves. For an overview see VAN EEGHEN (1981) or TAYLOR (1986). Here  $s_{ij}$  and  $S_{ij}$  respectively—we intentionally use the same symbols as before—denote the inflation-adjusted total amount of payments made in development year  $j, j = 1, \dots, n$ , for accidents occurred in accident year  $i, i = 1, \dots, m$ . If one works with incurred amounts,  $s_{ij}$  and  $S_{ij}$  denote the total amount of changes in valuation made in development year  $j$  on behalf of claims of accident year  $i$ . Working with incremental amounts we may assume that all  $S_{ij}$  are independent. Typically, one has  $n = m$  and  $s_{ij}$  is known for all  $i + j \leq m + 1$  (run-off triangle), and one is interested in estimating  $E(S_{ij})$  for  $i + j > m + 1$ . The known measure of exposure  $n_{ij}$  here normally only depends on the accident year  $i$ , i.e.  $n_{ij} = n_i$  (number of policies or number of claims reported in the first development year) or is even ignored (i.e.  $n_{ij} = 1$  for all  $i, j$ ).

One of the most important ways of treating the IBNR problem is to assume a multiplicative structure of the type

$$E(S_{ij}) = x_i y_j$$

and to estimate the parameters  $x_i, y_j$  from the triangle of known data. This way was used, for example, by DE VYLDER (1978), who estimated  $x_i, y_j$  by minimizing

$$\sum_{i,j} (s_{ij} - x_i y_j)^2$$

(where the summation is for all  $i, j$  where  $s_{ij}$  is known). This is exactly the same method as was used by SANT (1980) in the context of automobile insurance if

one puts all  $n_y = 1$  there. Analogously each method which estimates the marginal parameters  $x_i, y_j$  for cross-classified automobile insurance data can also be translated into a method for estimating the IBNR claims reserve. One only must take the different pattern of known data (triangle instead of rectangle) into account

Let us consider as further example the method of marginal totals.

Again working with  $n = m$  and  $n_y = 1$ , we get the conditions

$$(H_i) \quad \sum_j x_i y_j = \sum_j s_{ij}, \quad i = 1, \dots, m,$$

$$(V_j) \quad \sum_i x_i y_j = \sum_i s_{ij}, \quad j = 1, \dots, m,$$

where the summation is for those indices where the corresponding  $s_{ij}$  are known (i.e. in the case of a full triangle  $j$  runs from 1 to  $m+1-i$  and  $i$  from 1 to  $m+1-j$ ). The same equations are also obtained if one derives the maximum likelihood equations in the Poisson case.

Because of the triangular structure, the above equations can here be solved recursively: We start with the general observation that the solution of this type of problem is only unique up to a multiplicative constant  $c \neq 0$  because if  $x_i, y_j$  is a solution,  $x_i c, y_j/c$  is a solution as well. Therefore, without loss of generality we can put  $y_1 + \dots + y_m = 1$ . Then using equation  $(H_1)$  we have  $x_1 = s_{1+}$ . From equation  $(V_m)$  we get  $y_m = s_{1m}/x_1$ . Then  $(H_2)$  yields  $x_2$ ,  $(V_{m-1})$  yields  $y_{m-1}$  etc.

But it is also possible to derive a direct formula for the unknown mean claims amount  $E(S_{ij}) = x_i y_j$ . For  $h > m+1-i$  it can be shown (see KREMER 1985, p 133-136, or Appendix A where a shorter proof is given) that

$$x_i y_h = \left( \sum_{j=1}^{m+1-i} s_{ij} \right) \cdot f_{m+2-i} \cdot f_{m+3-i} \cdots f_{h-1} \cdot (f_h - 1)$$

where 
$$f_j = \left( \sum_{k=1}^{m+1-j} \sum_{l=1}^j s_{kl} \right) / \left( \sum_{k=1}^{m+1-j} \sum_{l=1}^{j-1} s_{kl} \right), \quad j = 2, \dots, m.$$

If one realizes that  $\sum_{l=1}^j s_{kl}$  is the accumulated claims amount of accident

year  $k$  known at the end of development year  $j$ , one sees that we have just obtained the well-known chain ladder method which is thus shown to be the same as the marginal totals method for  $n_y = 1$ . Furthermore, from the marginal totals conditions  $(H_i), (V_j)$  one easily sees that an incorporation (analogously to (1a) and (1b)) of the known exposure  $n_i$  into the estimation of the IBNR claims reserve can be dispensed with, as  $n_i$  can be amalgamated with the marginal parameter  $x_i$  (in the multiplicative approach only), whereas the

application of the chain ladder method to the claims ratios  $s_{ij}/n_i$  assumes a different model.

It is interesting to note that the analogue of the BAILEY-SIMON method seems to have never been published as a method for estimating the IBNR claims reserve.

Another interesting point is the fact that in the context of IBNR claims estimation only the multiplicative approach seems to have been used, although several applications to automobile rating indicate that there the additive approach might give a better fit (see e.g. CHANG/FAIRLEY 1979). A special feature of the additive approach is that it may lead to negative values  $E(S_{ij}) = x_i + y_j$ . This would make no sense in the ratemaking situation but in the case of claims reserving it can be very realistic (settlement gains)

Clearly, also in the context of claims reserving the least squares method and the marginal totals method (and, of course, the BAILEY-SIMON method) could be carried through with the additive approach, too, both producing an identical set of equations for  $x_i, y_j$  as has already been mentioned in the section on automobile rating

There is a natural connection between the multiplicative and the additive approaches because, through the log-transformation,

$$s_{ij}/n_{ij} \approx x_i y_j$$

becomes

$$\log (s_{ij}/n_{ij}) \approx \log (x_i) + \log (y_j).$$

This means that an estimate for  $E(S_{ij}/n_{ij})$  can be established by applying an additive approach to the log-transformed data  $\log (s_{ij}/n_{ij})$  and by transforming back the obtained solution  $\log (x_i), \log (y_j)$  using the exponential function. This was done by CHANG/FAIRLEY (1979) for automobile rating and by KREMER (1982) (see also ZEHNWIRTH 1989) for claims reserving (with  $n_{ij} = 1$ ). For the solution of the transformed (additive) problem, both used the method of (weighted) least squares (here giving the same result as the marginal totals method) in order to estimate the marginal parameters  $\log (x_i), \log (y_j)$ .

As ZEHNWIRTH (1989) points out, this procedure contains an implicit distributional assumption: In order to fulfill the conditions of normality and homoscedasticity for the least squares estimation of the parameters  $\log (x_i)$  and  $\log (y_j)$ , it has to be assumed that  $\log (S_{ij}/n_{ij})$  has a normal distribution with mean value  $\log (x_i) + \log (y_j)$  and a variance which is proportional to  $1/n_{ij}$ . This implies that  $S_{ij}/n_{ij}$  is assumed to have a lognormal distribution. CHANG/FAIRLEY and KREMER did not take this implicit distributional assumption into account. Therefore, they systematically underestimated  $E(S_{ij}/n_{ij})$  as they used  $x_i y_j = \exp (\log (x_i) + \log (y_j))$ , which is the median of the lognormal distribution whereas the expected value is  $x_i y_j \exp (\sigma_{ij}^2/2)$  with  $\sigma_{ij}^2 = \text{Var} (\log (S_{ij}/n_{ij}))$ . As stated above, we have homoscedasticity if we assume that  $\sigma_{ij}^2 = \sigma^2/n_{ij}$ , where  $\sigma^2$  can be estimated by

$$\sum_{i,j} n_{ij} (\log (s_{ij}/n_{ij}) - \log (x_i y_j))^2 / (c - m - n + 1),$$

which is just the expression to be minimized by the least squares method. Here  $c$  denotes the number of cells where  $s_{ij}$  is known.

Unfortunately, we have lost the multiplicative structure, as generally  $E(S_{ij}/n_{ij}) = x_i y_j \exp(\sigma_{ij}^2/2)$  cannot be cast into the form  $E(S_{ij}/n_{ij}) = \tilde{x}_i \tilde{y}_j$  anymore.

Whereas all the models discussed before have been shown to be only of a heuristic nature both in automobile rating and in claims reserving, the lognormal model relies on a parametric assumption for  $S_{ij}$ , and the instruments of regression analysis can be used to check this assumption against the data. In the next section another method is given which relies on a reasonable distributional model and therefore also allows the application of various important and useful statistical tools. This model has two advantages over the lognormal model. First, it is not just any model for  $S_{ij}$  but can be traced back to a micro-model for the total claims amount of each single insured unit and can therefore be expected to be realistic. Second, we can choose either the multiplicative or the additive structure for  $E(S_{ij}/n_{ij})$ , whereas the lognormal model yields neither of these structures.

### 3. A PARAMETRIC MODEL FOR RATING AUTOMOBILE INSURANCE OR ESTIMATING IBNR CLAIMS RESERVES

We use the same notations as before, i.e. we have  $mn$  cells labelled  $(i, j)$ , each with known measure of exposure  $n_{ij}$  (possibly independent of  $j$  in the case of claims reserving) and with total claims amount variable  $S_{ij}$  (realization  $s_{ij}$ ). In the case of claims reserving we know the realizations  $s_{ij}$  in the run-off triangle only. We now assume, following TER BERG (1980), that the total claims amount  $R_{ijk}$  of each unit  $k = 1, \dots, n_{ij}$  of cell  $(i, j)$  has a Gamma distribution with mean value  $m_{ij}$  (independent of  $k$ ) and shape parameter  $\alpha$  (independent of  $i, j, k$ ), i.e. with probability density function

$$f_{ij}(z) = \exp(-\alpha z/m_{ij}) z^{\alpha-1} (\alpha/m_{ij})^\alpha / \Gamma(\alpha)$$

(here the usual representation of the Gamma density has been reparametrized in order to implement the mean value  $m_{ij}$  directly as a parameter). Because in practice many units  $k$  will have a realization  $r_{ijk} = 0$  of  $R_{ijk}$ , the shape parameter  $\alpha$  has to be conceived of as smaller than 1 in order to attribute a high probability to the neighbourhood of  $z = 0$  (for instance, we have  $\text{prob}(R_{ijk} \leq m_{ij}/10) = 0.79$  for  $\alpha = 0.05$ ). Assuming that all  $n_{ij}$  units of cell  $(i, j)$  are independent, our distributional assumption implies that  $S_{ij} = R_{ij1} + R_{ij2} + \dots$  also has a Gamma distribution but with mean value  $n_{ij} m_{ij}$  and shape parameter  $n_{ij} \alpha$ . And this is the distribution we shall work with in the following, because we usually know only the realizations  $s_{ij}$  of  $S_{ij}$  and not those of  $R_{ijk}$ . The assumption that the shape parameter  $\alpha$  is the same for the units of all cells may seem questionable in some cases. But this should be detected by testing the goodness-of-fit (see next Section).

In the multiplicative approach we assume furthermore that  $m_y$  can be displayed in the form  $m_y = x_i y_j$  with unknown parameters  $x_i, y_j$ , which we shall estimate with the maximum likelihood method.

Assuming that all  $S_{ij}$  are independent, the likelihood function on the basis of the realizations  $s_{ij} > 0$  is given by

$$L = \prod_{i,j} \exp \left( - \frac{\alpha s_{ij}}{x_i y_j} \right) \left( \frac{\alpha s_{ij}}{x_i y_j} \right)^{n_{ij} \alpha} (s_{ij} \Gamma(n_{ij} \alpha)).$$

Therefore the loglikelihood function is

$$\log(L) = \sum \{ -\alpha s_{ij}/(x_i y_j) + n_{ij} \alpha \log(\alpha s_{ij}) - n_{ij} \alpha \log(x_i y_j) - \log(s_{ij} \Gamma(n_{ij} \alpha)) \}$$

(where the summation is for all  $i, j$  where  $s_{ij}$  is known). The maximum likelihood estimator are those values  $x_i, y_j, \alpha$  which maximize  $L$  or equivalently  $\log(L)$ . They are given by the equations

$$0 = \partial \log(L) / \partial x_i = \alpha \sum_j (s_{ij}/(x_i^2 y_j) - n_{ij}/x_i), \quad i = 1, \dots, m,$$

$$0 = \partial \log(L) / \partial y_j = \alpha \sum_i (s_{ij}/(x_i y_j^2) - n_{ij}/y_j), \quad j = 1, \dots, n,$$

which show that the last condition  $\partial \log(L) / \partial \alpha = 0$  is not needed for the calculation of the likelihood estimator for  $x_i, y_j$ , which can immediately be seen to be given by

$$(2) \quad \begin{cases} x_i = \frac{1}{n_{i+}} \sum_j \frac{s_{ij}}{y_j}, & i = 1, \dots, m, \\ y_j = \frac{1}{n_{+j}} \sum_i \frac{s_{ij}}{x_i}, & j = 1, \dots, n \end{cases}$$

These equations have a high intuitive appeal. For, considering the goal of approximating  $s_{ij}$  by  $n_{ij} x_i y_j$ , we see that this amounts to approximating  $s_{ij}/(n_{ij} y_j)$  by  $x_i$  and therefore the  $n_{ij}$ -weighted mean of  $s_{ij}/(n_{ij} y_j)$ ,  $j = 1, \dots, n$ , should be a reasonable estimator for  $x_i$ .

Also, equations (2) are not new. They have already been used by VAN EEGHEN/NIJSEN/RUYGT (1982). They call them the "direct method" and write (on page 111)

"This set of equations are a direct translation of the intuitive calculations presented ... by F. K. GREGORIUS. In fact, a solution is found when iteratively calculating the values  $x_i$  and  $y_j$  by means of the formulae given in (2) by letting  $y_j = 1$  ( $j = 1, \dots, n$ ) be the starting value. The procedure converges rapidly

We may rewrite (2) as

$$\sum_j n_{ij} x_i = \sum_j s_{ij} / y_j, \quad i = 1, \dots, m,$$

$$\sum_i n_{ij} y_j = \sum_i s_{ij} / x_i, \quad j = 1, \dots, n,$$

which is similar but not equivalent to (1a) and (1b).

As yet, we have not been able to find an argument why a 'satisfactory' solution should (approximately) satisfy (2)...

The method was more or less developed as a first try and we were surprised to see, that, once formalized, it produced practically the same results as the method of marginal totals."

So much for the quotation from VAN EEGHEN/NIJSSEN/RUYGT (1982).

One year later the Dutch actuaries found an argument for their method because the booklet of VAN EEGHEN, GREUP and NIJSSEN (1983) contains on page 109 a small hint saying that the assumption of a Gamma distribution for  $R_{ijk}$  would lead to the "direct method". But there, as in TER BERG (1980), a much more general regression model is considered, of which our simple cross-classified situation is just a special case. Moreover, these authors have concentrated on ratemaking, whereas we want to emphasize the applicability to claims reserving, too.

Finally, it is interesting to note that the likelihood equations for the additive approach

$$\sum_j (s_{ij}/(x_i + y_j)^2 - n_{ij}/(x_i + y_j)) = 0, \quad i = 1, \dots, m,$$

$$\sum_i (s_{ij}/(x_i + y_j)^2 - n_{ij}/(x_i + y_j)) = 0, \quad j = 1, \dots, n,$$

must be solved with the help of, for example, the NEWTON-RAPHSON numerical method. Moreover, these equations are different from those suggested by the "direct method":

$$x_i = \sum_j (s_{ij} - n_{ij} y_j) / n_{i+}, \quad i = 1, \dots, m,$$

$$y_j = \sum_i (s_{ij} - n_{ij} x_i) / n_{+j}, \quad j = 1, \dots, n.$$

#### 4. STATISTICAL ANALYSIS OF THE MODEL

This parametric approach with a realistic distributional assumption enables us to use many tools for the statistical analysis, as has been clearly set out by

ALBRECHT (1983), who describes the case  $\alpha = 1$  in considerable detail but again as a general regression model. Besides the consistent and (asymptotically) efficient estimation of the model parameters, we have the possibility of testing the significance of the tariff variables with the likelihood ratio test (see ALBRECHT (1983) for details), we can calculate the error variances of the parameter estimators and we can check the goodness-of-fit. We first consider the goodness-of-fit. According to our model,  $S_{ij}$  has a Gamma distribution with  $E(S_{ij}) = n_{ij} m_{ij}$  and  $\text{Var}(S_{ij}) = n_{ij} m_{ij}^2 / \alpha$ . The higher the shape parameter  $n_{ij} \alpha$  of this distribution, the closer it is to the normal distribution. If all  $S_{ij}$  are approximately normally distributed the statistic

$$\begin{aligned} \sum_{i,j} \frac{(S_{ij} - E(S_{ij}))^2}{\text{Var}(S_{ij})} &= \alpha \sum_{i,j} \frac{(S_{ij}/(x_i y_j) - n_{ij})^2}{n_{ij}} \\ &= \alpha \sum_{i,j} n_{ij} \left( \frac{S_{ij}}{n_{ij} x_i y_j} - 1 \right)^2 \end{aligned}$$

is, under the hypothesis of our model, approximately at chi-square with  $c - m - n$  degrees of freedom, where  $c$  is the number of cells where  $s_{ij}$  is known.

The special form of this statistic allows its application without having estimated  $\alpha$ . For this purpose we fix  $\alpha$  in such a way that the value of the statistic is just below the (say) 0.95-fractile of the chi-square distribution. If using this value of  $\alpha$  a normality condition like " $n_{ij} \alpha > 10$ " is fulfilled for nearly all cells, we may be satisfied with the goodness-of-fit of the model. But we have to realize that this goodness-of-fit test only checks the fit of aggregated figures and cannot test the distributional assumptions within the cells.

Applying this procedure to SANT'S (1980) collision data (126 cells) we get  $\alpha (<) = 0.021$  and the three lowest values of  $n_{ij} \alpha$  turn out to be 6.8, 9.4 and 11.5, so we may accept the multiplicative Gamma model. Using CHANG/FAIRLEY'S (1979) combined compulsory data (105 cells), we get  $\alpha (<) = 0.0094$  and have 9 cells where the resulting value of  $n_{ij} \alpha$  is lower than 10, the lowest being 4.5, so the fit is less satisfactory.

A simple formula for an estimator of  $\alpha$  is given by the method of moments, i.e. by equating the variances

$$\sum_{i,j} (s_{ij} - n_{ij} x_i y_j)^2 = \sum_{i,j} n_{ij} (x_i y_j)^2 / \alpha.$$

This yields  $\alpha = 0.014$  for Sant's data and  $\alpha = 0.0093$  for Chang/Fairley's data.

Strictly speaking we should use the likelihood estimator for  $\alpha$ . We then must solve the likelihood equation

$$0 = \partial \log(L) / \partial \alpha = \sum_{i,j} n_{ij} \{ \log(\alpha s_{ij}) - \log(x_i y_j) - \psi(n_{ij} \alpha) \}$$



Here equations (2) have been used to obtain  $\sum n_{ij} = \sum s_{ij}/(x_i y_j)$ .  $\psi(z) = \Gamma'(z)/\Gamma(z)$  denotes the digamma function, for which the asymptotic approximation

$$\psi(z) \approx \log(z) - (2z)^{-1} - z^{-2}/12$$

exists which even for arguments as low as  $z \geq 4$  is exact to 4 decimal places. This approximation yields as the solution of the likelihood equation

$$\alpha \approx (c + \sqrt{c^2 + ab})/a$$

with

$$a = 4 \sum_{i,j} n_{ij} \log(n_{ij} x_i y_j / s_{ij}) > 0,$$

$$b = \sum_{i,j} (3 n_{ij})^{-1}$$

$$c = \sum_{i,j} 1 = \text{number of cells where } s_{ij} \text{ is known}$$

Applied to Sant's data this yields  $\alpha \approx 0.0202$ . For Chang/Fairley's data we get  $\alpha \approx 0.0097$ . If we have some small exposures  $n_{ij}$  such that  $n_{ij} \alpha < 4$ , we should refine the approximation of the digamma function by using the recursion  $\psi(z) = \psi(z+1) - 1/z$  and by including more terms of the approximation series. Then a direct formula for  $\alpha$  cannot be given anymore. We must therefore solve the likelihood equation iteratively with the NEWTON-RAPHSON method.

Having estimated  $\alpha$ , we are also in the position to calculate the estimation error of the estimators for  $x_i$  and  $y_j$ . This is done in Appendix B.

According to the experience of the Dutch actuaries, the results of applying the "direct method" to automobile insurance data are rather close to the results obtained by the marginal totals method. Translated to the IBNR claims reserving problem this means that the "direct method" results will be similar to the chain ladder results. But with the "direct method" we can additionally make use of the aforementioned advantages. Moreover, the formulae provide the possibility of taking the exposure  $n_i$  of accident year  $i$  into account (which is different from the situation with the chain ladder). And perhaps the goodness-of-fit statistic or the size of the likelihood function gives an indication to answering the question "additive or multiplicative?" Because of these advantages of the parametric method we believe that before using a rather heuristic method like BAILEY/SIMON or chain ladder one should examine whether the parametric method fits the data.

#### 5. IMPROVEMENT OF THE MODEL IN THE CASE OF KNOWN CLAIMS NUMBERS

Especially in the claims reserving situation we will often have difficulties in finding an adequate measure  $n_{ij}$  of exposure

Therefore mostly  $n_{ij} = n_i$ , or even  $n_{ij} = 1$  is taken. However, this is not satisfactory because the exposure to further payments or changes in valuation varies in fact rather strongly over the development years. Therefore, a more meaningful measure of exposure will be the number  $t_{ij}$  of those claims of accident year  $i$  where there is a change in amount during development year  $j$ . These data  $t_{ij}$ ,  $i+j \leq m+1$ , are often available in practice.

Rating in property insurance presents a similar problem. There, even the risks of the same cell vary greatly with respect to their size, which is usually measured by the sum insured. Therefore, the number of risks is not a good measure for the exposure of a cell  $(i, j)$ , and the sum insured is taken instead. But then an assumption of our micro-model is not fulfilled anymore because the "units" of sum insured are not independent, as a single risk consists of several such units. We therefore must abandon our micro-model and try directly whether the Gamma model for  $S_{ij}$  with mean value  $E(S_{ij}) = n_{ij}x_i y_j$  and shape parameter  $n_{ij}\alpha$  fits the data if  $n_{ij}$  is the sum insured. The parameter  $\alpha$  then does not have a specific interpretation anymore. But if we know additionally the total number  $t_{ij}$  of claims of cell  $(i, j)$  we can apply the following stepwise approach which assumes a Gamma distribution (with shape parameter  $\alpha$ ) not for the total claims amount per risk unit but for the amount of each single claim. Of course, this procedure can also be applied in automobile ratemaking if the number  $t_{ij}$  of claims is available.

In these situations we should use  $t_{ij}$ —the corresponding random variable is denoted by  $T_{ij}$ —as an additional measure of exposure and adopt the following three-steps-approach, which follows the ideas of ALBRECHT (1983): In the first step we take the observed number  $t_{ij}$  of claims of cell  $(i, j)$  as the measure of exposure and assume that the size of each corresponding amount has a Gamma distribution with mean value  $m_{ij} = x_i y_j$  and shape parameter  $\alpha$ . Then we are in our original model (with  $n_{ij}$  replaced with  $t_{ij}$ ) leading to the direct method. This yields smoothed average claims amounts  $x_i y_j$ . In the second step we smooth the  $t_{ij}$  by assuming that all  $T_{ij}$  are independent of each other and that each  $T_{ij}$  has a Poisson distribution with parameter  $n_{ij}v_i w_j$  (here using the 'old' measure of exposure). Then the maximum likelihood estimator of  $v_i, w_j$  on basis of the realizations  $t_{ij}$  is given by the equations (1a) and (1b) with  $x_i, y_j, s_{ij}$  replaced with  $v_i, w_j, t_{ij}$  respectively. This yields smoothed numbers  $n_{ij}v_i w_j$  of claims. In the last step,  $E(S_{ij})$  is estimated by  $n_{ij}v_i w_j x_i y_j$  implying that in each cell the number of claims is independent of the average claims amount.

## 6 FINAL REMARK

In the context of this paper we should point out the following further connection between rating methods and claims reserving methods. Another important rating method which smoothes the claims experience of several tariff classes is the Bühlmann-Straub credibility model. It also uses a cross-classifying approach by the two dimensions 'tariff classes' and 'observation years'.

Therefore, one will presume that it could also be translated into a method for estimating IBNR claims reserves. But there is a difficulty because the Bühlmann-Straub model assumes that the average claims amount  $S_{ij}/n_{ij}$  of tariff class  $i$  has the same expected value over all years  $j$ , whereas in the run-off triangle the expected value of the average claims amount  $S_{ij}/n_{ij}$  of accident year  $i$  and development year  $j$  varies in a certain but unknown pattern over the development years. However, this difficulty can be overcome in such a way that the Bühlmann-Straub model can directly be used for claims reserving, too (see MACK 1990).

APPENDIX A

PROOF THAT THE CHAIN LADDER METHOD CAN BE DERIVED FROM THE MARGINAL TOTALS CONDITIONS (AND THEREFORE IS MAXIMUM LIKELIHOOD IN THE POISSON CASE)

We show that the chain ladder method

$$x_i y_h = \left( \sum_{j=1}^{m+1-i} s_{ij} \right) \cdot f_{m+2-i} \cdot f_{m+3-i} \cdots f_{h-1} \cdot (f_h - 1), \quad h > m + i - 1,$$

with

$$f_j = \left( \sum_{k=1}^{m+1-j} \sum_{l=1}^j s_{kl} \right) \bigg/ \left( \sum_{k=1}^{m+1-j} \sum_{l=1}^{j-1} s_{kl} \right), \quad j = 2, \dots, m,$$

can be deduced from the marginal conditions

$$(H_i) \quad \sum_{j=1}^{m+1-i} x_i y_j = \sum_{j=1}^{m+1-i} s_{ij}, \quad i = 1, \dots, m,$$

$$(V_j) \quad \sum_{i=1}^{m+1-j} x_i y_j = \sum_{i=1}^{m+1-j} s_{ij}, \quad j = 1, \dots, m.$$

Let  $c_{ij} = \sum_{l=1}^j x_i y_l$  and  $b_{ij} = \sum_{l=1}^j s_{il}$  ( $i + j \leq m + 1$ ) denote the expected and

the observed accumulated claims amount of accident year  $i$  at the end of development year  $j$  respectively. Then conditions  $(H_i)$  can be written shortly as  $c_{i, m+1-i} = b_{i, m+1-i}$ . For  $h > m + 1 - i$  we have

$$c_{ih} = c_{i, m+1-i} \cdot \frac{c_{i, m+2-i}}{c_{i, m+1-i}} \cdots \frac{c_{ih}}{c_{i, h-1}}.$$

Therefore

$$x_i y_h = c_{ih} - c_{i, h-1}$$

$$\begin{aligned}
&= c_{i, m+1-i} \frac{c_{i, m+2-i} \dots c_{i, h-1}}{c_{i, m+1-i} c_{i, h-2}} \left( \frac{c_{ih}}{c_{i, h-1}} - 1 \right) \\
&= \left( \sum_{j=1}^{m+1-i} s_{ij} \right) \cdot \frac{c_{i, m+2-i} \dots c_{i, h-1}}{c_{i, m+1-i} c_{i, h-2}} \cdot \left( \frac{c_{ih}}{c_{i, h-1}} - 1 \right)
\end{aligned}$$

and we have only to show that  $c_{ij}/c_{i, j-1} = f_j$ . Because of

$$\frac{c_{ij}}{c_{i, j-1}} = \frac{\sum_{l=1}^j y_l}{\sum_{l=1}^{j-1} y_l} = \frac{\left( \sum_{k=1}^{m+1-j} x_k \right) \left( \sum_{l=1}^j y_l \right)}{\left( \sum_{k=1}^{m+1-j} x_k \right) \left( \sum_{l=1}^{j-1} y_l \right)} = \frac{\sum_{k=1}^{m+1-j} c_{kj}}{\sum_{k=1}^{m+1-j} c_{k, j-1}}$$

and of

$$f_j = \left( \sum_{k=1}^{m+1-j} b_{kj} \right) \left/ \left( \sum_{k=1}^{m+1-j} b_{k, j-1} \right) \right.$$

it is enough to show that

$$(A_j) \quad \sum_{k=1}^{m+1-j} c_{kj} = \sum_{k=1}^{m+1-j} b_{kj}$$

and

$$(B_j) \quad \sum_{k=1}^{m+1-j} c_{k, j-1} = \sum_{k=1}^{m+1-j} b_{k, j-1}$$

hold for  $j = 2, \dots, m$ . We show this by recursion from  $j = m$  to  $j = 2$ .

$(A_m)$ , i.e.  $c_{1m} = b_{1m}$ , holds because of  $(H_1)$ .

$(B_j)$  follows from  $(A_j)$  and  $(V_j)$  as

$$\begin{aligned}
\sum_{k=1}^{m+1-j} c_{k, j-1} &= \sum_{k=1}^{m+1-j} (c_{kj} - x_k y_j) = \sum_{k=1}^{m+1-j} c_{kj} - \sum_{k=1}^{m+1-j} x_k y_j \\
&= \sum_{k=1}^{m+1-j} b_{kj} - \sum_{k=1}^{m+1-j} s_{kj} = \sum_{k=1}^{m+1-j} (b_{kj} - s_{kj}) = \sum_{k=1}^{m+1-j} b_{k, j-1}
\end{aligned}$$

Finally,  $(A_{j-1})$  follows from  $(B_j)$  and  $(H_{m+2-j})$  as

$$\begin{aligned}
\sum_{k=1}^{m+2-j} c_{k, j-1} &= \sum_{k=1}^{m+1-j} c_{k, j-1} + c_{m+2-j, j-1} \\
&= \sum_{k=1}^{m+1-j} b_{k, j-1} + b_{m+2-j, j-1} = \sum_{k=1}^{m+2-j} b_{k, j-1}
\end{aligned}$$

This completes the proof.

APPENDIX B  
ESTIMATION OF THE (ASYMPTOTIC) ERROR VARIANCES

We have estimated the marginal parameters  $x_i, y_j$  with

either  $x_i y_j = E(S_{ij}/n_{ij})$  (multiplicative approach)

or  $x_i + y_j = E(S_{ij}/n_{ij})$  (additive approach).

by the maximum likelihood method and now want to know how precise these estimates are, i.e. we want to calculate  $\text{Var}(X_i), \text{Var}(Y_j), \text{Var}(X_i, Y_j)$  or  $\text{Var}(X_i + Y_j)$  where  $X_i$  and  $Y_j$  denote the random variables corresponding to the estimators for  $x_i$  and  $y_j$  respectively. A standard result of maximum likelihood theory states that under certain regularity conditions which are fulfilled here, the following holds true: If a parameter vector  $\theta = (\theta_1, \dots, \theta_r)$  is estimated by the maximum likelihood method, the obtained estimator  $\hat{\theta}$  has asymptotically a normal distribution with mean value  $\theta$  and with a covariance matrix which is equal to the inverse of the information matrix

$$I(\theta) = E \left( - \frac{\partial^2 \log(L)}{\partial \theta_i \partial \theta_j} \right)_{i,j}$$

where  $L = L(\theta)$  is the likelihood function.

In our case we have  $\theta = (x_2, \dots, x_m, y_1, \dots, y_n)$  where we have omitted  $x_1$  without loss of generality in order to obtain a unique solution of the likelihood equations and have considered  $\alpha$  as being known (For the case of  $\alpha$  being included in  $\theta$ , TER BERG (1980) has shown that this does not change the calculation of  $\text{Var}(X_i), \text{Var}(Y_j)$  and  $\text{Cov}(X_i, Y_j)$ ). We now have

$$\begin{aligned} \text{Cov}(X_2, \dots, X_m, Y_1, \dots, Y_n) &\approx I(x_2, \dots, x_m, y_1, \dots, y_n)^{-1} =: I^{-1} \\ &\approx I(\hat{x}_2, \dots, \hat{x}_m, \hat{y}_1, \dots, \hat{y}_n)^{-1} =: \hat{I}^{-1} \end{aligned}$$

where  $\hat{x}_2, \dots, \hat{y}_n$  denote the estimated values of the true parameters  $x_2, \dots, y_n$ . From  $\hat{I}^{-1}$  we directly obtain asymptotic approximative values for  $\text{Var}(X_i), \text{Var}(Y_j)$  and  $\text{Cov}(X_i, Y_j)$ . This also gives immediately an approximation for

$$\text{Var}(X_i + Y_j) = \text{Var}(X_i) + 2 \text{Cov}(X_i, Y_j) + \text{Var}(Y_j)$$

which we want to know in the additive approach. In order to obtain  $\text{Var}(X_i, Y_j)$  for the multiplicative approach, we make use of a general theorem on the higher moments of normally distributed variables (see e.g. RICHTER 1966, p 369) to get

$$\begin{aligned} \text{Var}(X_i, Y_j) &\approx \text{Var}(X_i) \text{Var}(Y_j) + (\text{Cov}(X_i, Y_j))^2 + \text{Var}(X_i) (E(Y_j))^2 + \\ &\quad + 2 E(X_i) \text{Cov}(X_i, Y_j) E(Y_j) + (E(X_i))^2 \text{Var}(Y_j) \end{aligned}$$

(which holds exactly if  $X_i$  and  $Y_j$  are normally distributed) This can be calculated from  $\hat{I}^{-1}$  and from  $E(X_i) \approx \hat{x}_i, E(Y_j) \approx \hat{y}_j$ .

Therefore, the only thing left to do is the calculation of  $I$  and  $I^{-1}$ . Concentrating again on the multiplicative approach, the loglikelihood function is

$$\log(L) = - \sum_{i,j} (\alpha S_{ij}/(x_i y_j) + \alpha n_{ij} \log(x_i y_j) + g(\alpha, n_{ij}, S_{ij}))$$

and yields (using  $E(S_{ij}) = n_{ij} x_i y_j$  and the Kronecker symbol  $\delta_{ij}$  with  $\delta_{ij} = 1$  for  $i = j$ ,  $\delta_{ij} = 0$  otherwise)

$$\begin{aligned} A_{ik} &: = E \left( - \frac{\partial^2 \log(L)}{\partial x_i \partial x_k} \right) = \frac{\alpha n_{i+}}{x_i^2} \delta_{ik}, & 2 \leq i, k \leq m, \\ B_{ij} &: = E \left( - \frac{\partial^2 \log(L)}{\partial x_i \partial y_j} \right) = \frac{\alpha n_{ij}}{x_i y_j}, & 2 \leq i \leq m, 1 \leq j \leq n, \\ C_{lj} &: = E \left( - \frac{\partial^2 \log(L)}{\partial y_l \partial y_j} \right) = \frac{\alpha n_{+j}}{y_j^2} \delta_{lj}, & 1 \leq l, j \leq n, \end{aligned}$$

(where  $n_{+j}$  includes  $n_{1j}$ ). With the matrices  $A = (A_{ik})$ ,  $B = (B_{ij})$ ,  $C = (C_{lj})$  the information matrix  $I$  can be represented as partitioned matrix

$$I = \begin{pmatrix} A & B \\ B' & C \end{pmatrix}$$

where  $A$  and  $C$  are diagonal matrices.

Unfortunately, an explicit formula for the inverse matrix  $I^{-1}$  is not available. One therefore must apply a numerical inversion method. But the dimension of the inversion problem can be reduced with the help of the following result for the inverse of a partitioned matrix (which can be verified by calculating  $I^{-1}I$  and  $II^{-1}$ ):

$$\begin{aligned} I^{-1} &= \begin{pmatrix} D^{-1} & -D^{-1}BC^{-1} \\ -C^{-1}B'D^{-1} & C^{-1} + C^{-1}B'D^{-1}BC^{-1} \end{pmatrix} \\ &= \begin{pmatrix} A^{-1} + A^{-1}BF^{-1}B'A^{-1} & -A^{-1}BF^{-1} \\ -F^{-1}B'A^{-1} & F^{-1} \end{pmatrix} \end{aligned}$$

with

$$\begin{aligned} D &= A - BC^{-1}B', \\ F &= C - B'A^{-1}B. \end{aligned}$$

A straightforward calculation yields for the elements of  $D$  and  $F$

$$\begin{aligned} D_{ik} &= \alpha(\delta_{ik}n_{i+} - p_{ik})/(x_i x_k), & 2 \leq i, k \leq m, \\ F_{lj} &= \alpha(\delta_{lj}n_{+j} - q_{lj})/(y_l y_j), & 1 \leq l, j \leq n, \end{aligned}$$

with

$$p_{ik} = \sum_{j=1}^n \frac{n_{ij}n_{kj}}{n_{+j}}, \quad q_{lj} = \sum_{i=2}^m \frac{n_{il}n_{ij}}{n_{i+}}.$$

Therefore, only the smaller matrices  $D$  and  $F$  must be inverted in order to obtain  $I^{-1}$  and also  $\hat{I}^{-1}$ .

#### ACKNOWLEDGEMENT

I am very indebted to PETER ALBRECHT. I had an intensive exchange of letters with him about this 1983 paper, which was very helpful to me. In addition, he drew my attention to some weak points in an earlier version of this paper. I would also like to thank PETER TER BERG for his advice on the digamma function and ALOIS GISLER for encouraging me to go ahead with this paper.

#### REFERENCES

- ALBRECHT, P (1983) Parametric Multiple Regression Risk Models *Insurance Mathematics & Economics*, 49–66, 69–73 and 113–117
- BAILEY, R A (1963) Insurance Rates with Minimum Bias *Proceedings of the Casualty Actuarial Society*, 4–11
- BAILEY, R A and SIMON, L J (1960) Two Studies in Automobile Insurance Rate Making *ASTIN Bulletin* 1, 192–217
- CHANG, L and FAIRLEY, W B (1979) Pricing Automobile Insurance under Multivariate Classification of Risks Additive versus Multiplicative *Journal of Risk and Insurance*, 75–98
- DE VYLDER, F (1978) Estimation of IBNR Claims by Least Squares *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, 249–254
- JEE, B (1989) A Comparative Analysis of Alternative Pure Premium Models in the Automobile Risk Classification System *Journal of Risk and Insurance*, 434–459
- JUNG, J (1968) On Automobile Insurance Ratemaking *ASTIN Bulletin* 5, 41–48.
- KREMER, E (1982) IBNR-Claims and the Two-way Model of ANOVA *Scandinavian Actuarial Journal*, 47–55
- KREMER, E (1985) *Einführung in die Versicherungsmathematik* Vandenhoeck & Ruprecht, Göttingen
- MACK, Th (1990) Improved Estimation of IBNR Claims by Credibility Theory *Insurance Mathematics & Economics*, 51–57
- RICHTER, H (1966) *Wahrscheinlichkeitstheorie* Springer, Heidelberg
- SANT, D T (1980) Estimating Expected Losses in Auto Insurance *Journal of Risk and Insurance*, 133–151
- TAYLOR, G C (1986) *Claims Reserving in Non-life Insurance* North Holland, Amsterdam
- TER BERG, P (1980) On the loglinear Poisson and Gamma model *ASTIN Bulletin* 11, 35–40
- VAN EEGHEN, J (1981) *Loss Reserving Methods* Nationale-Nederlanden N V, Rotterdam
- VAN EEGHEN, J, GREUP, E K and NIJSSEN, J A (1983) *Rate Making* Nationale-Nederlanden N V, Rotterdam
- VAN EEGHEN, J, NIJSSEN, J A and RUYGT, F A M (1982) Interdependence of Risk Factors Application of Some Models New Motor Rating Structure in the Netherlands, 105–119 *ASTIN-groep*, Nederland
- ZEHNWIRTH, B (1989) *The Chain Ladder Technique — A Stochastic Model* Claims Reserving Manual Vol 2, 2–9 Institute of Actuaries, London

Dr THOMAS MACK

*Munchener Rückversicherungs-Gesellschaft, Königinstr. 107,  
D-8000 München 40, FRG.*





# SEPARATING TRUE IBNR AND IBNER CLAIMS<sup>1</sup>

BY R. SCHNIEPER  
*Winterthur, Switzerland*

## ABSTRACT

A simple model for IBNR claims is presented. Estimates for the loss reserves and for the ultimate claims rate are derived. Approximations to the mean square error of the estimators are produced. A more specific parametric model is suggested for the case that we deal with claim numbers instead of claim amounts. The general method is illustrated by a practical application to the pricing of a casualty excess of loss cover.

## 1. INTRODUCTION

The IBNR Method which we present in this paper has been developed in connection with the pricing of casualty excess of loss covers. The method can also be applied to loss reserving problems for long tail business, however it is best understood in connection with the practical problem which motivated its derivation.

A reinsurer has to quote a price for an excess of loss cover. The statistical information at hand are the revalued individual excess claims from different accident years as well as a revalued measure of the exposure pertaining to each accident year (e.g. the revalued premium income). The problems connected with the revaluation of the claims and of the measure of exposure are by no means trivial. We shall however assume that this revaluation can be performed in a satisfactory way and that our data have been corrected for premium and claims inflation. We shall call this revalued statistics the 'as if' statistics.

To price the cover we have to estimate the ultimate claims amount in the layer, i.e. to perform the IBNR correction. In this paper we present a simple method which requires only about twice the amount of computation of the chain-ladder method and which has the advantage of being practically unbiased. An additional advantage of the estimator defined below is that one can assess its precision. It is felt that these two properties are of special importance when pricing layers with high deductibles where data are scarce.

In the next section we present the general model. In the third section we restrict ourselves to claim numbers. In both these sections we illustrate the

<sup>1</sup> The Paper has been presented at the XX1th ASTIN Colloquium in New York under the title 'A Pragmatic IBNR Method'

theory with an extremely simple example. In the last section we apply our method to a practical problem.

2. THE GENERAL MODEL

2.1. Summary statistics

Most IBNR methods require only one summary statistics: the IBNR triangle. If we have the excess claims from  $n$  accident years, the IBNR triangle contains the following information:

acc year	dvpt year				Exposure
		1	2	$n$	
1		$X_{1,1}$	$X_{1,2}$	$X_{1,n}$	$E_1$
2		$X_{2,1}$	$X_{2,2}$	$X_{2,n-1}$	$E_2$
$n$		$X_{n,1}$			$E_n$

Where  $X_{i,j}$  is the total amount of excess claims from accident year  $i$  in development year  $j$ .

For our purposes we need a more detailed summary statistics which we now define. Let  $N_{i,j}$  denote the total claims amount pertaining to new excess claims, i.e. to claims which were not yet recorded as excess claims in development year  $j-1$ . This is the true IBNR component. Let  $D_j$  be the decrease in total claims amount between development year  $j-1$  and development year  $j$  with respect to claims already known as excess claims in development year  $j-1$ . This is the IBNER component (incurred but not enough reported claims).  $D_j$  may take negative values but cannot by definition be larger than  $X_{i,j-1}$ .

The following relations hold true between the  $X$ 's,  $N$ 's and  $D$ 's:

$$(2.1.1) \quad X_{i1} = N_{i1} \quad i = 1, \dots, n$$

$$(2.1.2) \quad X_{ij} = X_{i,j-1} - D_j + N_{ij} \quad i = 1, \dots, n \quad j = 2, \dots, n$$

Of course we only observe the variables for which  $i+j \leq n+1$ . We shall not as is usually done reduce the data to one IBNR triangle, the  $X$ -triangle, but we shall work with two triangles: the  $N$ -triangle of the genuine IBNR claims and the  $D$ -triangle of the IBNER claims

From (2.1.1) and (2.1.2) it is seen that the  $X$ -triangle can be derived from the  $N$ - and  $D$ -triangle.

To illustrate these definitions let us consider a very simple example.

EXAMPLE

There are 3 accident years. For each accident year we have the usual 'as if' statistics: revalued and developed individual excess claims as well as a revalued measure of exposure

	Claim number	Development year number		
		1	2	3
Accident year number 1	1	1	—	—
	2	2	2	1.5
	3	—	0.5	1.5
$E_1 = 20$	4	—	1	—
	5	—	1.5	2.5
	6	—	—	1
Accident year number 2	1	0.5	—	
	2	0.5	1.5	
	3	1.5	—	
$E_2 = 25$	4	—	0.5	
	5	—	2	
	6	—	1	
Accident year number 3	1	0.5		
	2	0.5		
	3	1		
$E_3 = 32$	4	1.5		
	5	2		

A claim demoted by '—' is a claim which has not yet reached the priority or which has dropped below the priority.

In our example the traditional IBNR triangle is:

X-triangle

		J		
		1	2	3
I	1	3	5	6.5
	2	2.5	5	
	3	5.5		

and the new statistics are

*N*-triangle

<i>i</i> \ <i>J</i>	1	2	3
1	3	3	1
2	2.5	3.5	
3	5.5		

*D*-triangle

<i>i</i> \ <i>J</i>	1	2	3
1	—	1	-0.5
2	—	1	
3	—		

**2.2. Assumptions**

Let  $H_k$  denote the set of those variables in the *N*- and *D*-triangle which are observed up to calendar year  $k$ .

$$H_k = \{N_y, D_y \mid i+j \leq k+1\}.$$

For the sake of convenience we also introduce

$$H_0 = \{0, \Omega\}.$$

$H_n$  is the set of all variables which have been observed so far.  $H_{i+j-2}$  is the history of the process up to the calendar year immediately preceding the emergence of  $N_y$  and  $D_y$ .

We make the following assumptions:

(A<sub>1</sub>) 
$$E[N_y \mid H_{i+j-2}] = E_i \lambda_j \quad i, j = 1, \dots, n$$

The expected IBNR claims amount does not depend on past history, it is the product of the exposure measure of the accident year with a factor depending on the development year only.

(A<sub>2</sub>) 
$$E[D_y \mid H_{i+j-2}] = X_{i,j-1} \delta_j \quad i = 1, \dots, n$$
  

$$i = 2, \dots, n$$

The expected decrease in IBNER claims amount is equal to the reported claims amount of the previous development year times a factor depending on the development year

We only observe those variables for which  $i+j \leq n+1$  but for the purpose of loss reserving and rating we shall need the assumptions to hold true for all  $i, j = 1, \dots, n$ .

If we knew whether individual claims are open or closed it might be preferable to replace the  $X_{i,j}$ 's in (A<sub>2</sub>) by the corresponding total claims amount pertaining to open claims

$$\begin{aligned}
 (A_3) \quad & \{N_{1j}, D_{1j} \mid j = 1, 2, \dots, n.\} \\
 & \vdots \\
 & \vdots \\
 & \vdots \\
 & \{N_{nj}, D_{nj} \mid j = 1, 2, \dots, n.\}
 \end{aligned}$$

are independent sets of random variables. i.e. random variables pertaining to different accident years are stochastically independent.

Assumptions (A<sub>1</sub>), (A<sub>2</sub>) and (A<sub>3</sub>), though they are quite general, are not always satisfied in praxis. In particular, as was remarked by one of the editors, a new claims manager arriving on the scene may have an impact across claims cohorts. In such a case assumption (A<sub>3</sub>) would of course no longer hold true. This I think, shows the limitations of all statistical models and methods used to assess loss reserves. When applying them to practical problems, we should always make sure that we have all the necessary information on the process generating the claims and that we take that information into account when choosing a statistical method to estimate the outstanding losses.

### 2.3. Pricing

We now focus our attention on the pricing problem, i.e. We want to estimate next year's expected excess claim amount  $E[X_{n+1,n}]$  or alternatively next year's expected ultimate claims rate

$$(2.2.1) \quad R = E \left[ \frac{X_{n+1,n}}{E_{n+1}} \right]$$

If the measure of exposure  $E_{n+1}$  is the premium income, then  $R$  is the expected ultimate burning cost. Assuming that (A<sub>1</sub>) and (A<sub>2</sub>) hold true for accident year  $n+1$ , one obtains straightforwardly:

$$\begin{aligned}
 (2.2.2) \quad R(\theta) = E \left[ \frac{X_{n+1,n}}{E_{n+1}} \right] &= \lambda_1 (1 - \delta_2) \cdot \dots \cdot (1 - \delta_n) + \\
 & \quad \lambda_2 (1 - \delta_3) \cdot \dots \cdot (1 - \delta_n) + \\
 & \quad \vdots \\
 & \quad \vdots \\
 & \quad \lambda_{n-1} (1 - \delta_n) + \\
 & \quad \lambda_n
 \end{aligned}$$

where

$$\theta = (\lambda_1, \dots, \lambda_n, \delta_2, \dots, \delta_n).$$

From (A<sub>1</sub>), (A<sub>2</sub>) and (A<sub>3</sub>) it follows that

$$(2.2.3) \quad \hat{\lambda}_j = \frac{\sum_{i=1}^{n+1-j} N_{ij}}{\sum_{i=1}^{n+1-j} E_i} \quad j = 1, \dots, n$$

and

$$(2.2.4) \quad \hat{\delta}_j = \frac{\sum_{i=1}^{n+1-j} D_{ij}}{\sum_{i=1}^{n+1-j} X_{i,j-1}} \quad j = 2, \dots, n$$

are biasfree estimates of the  $\lambda$ 's and  $\delta$ 's respectively.

$$(2.2.5) \quad R(\hat{\theta}) = \hat{\lambda}_1 (1 - \hat{\delta}_2) \cdot \dots \cdot (1 - \hat{\delta}_n) + \hat{\lambda}_2 (1 - \hat{\delta}_3) \cdot \dots \cdot (1 - \hat{\delta}_n) + \hat{\lambda}_n$$

is an estimate of the ultimate claims rate  $R$ . The individual estimates being biasfree and the correlation between the factors being 'small' because of (A<sub>3</sub>) the bias of  $R(\hat{\theta})$  can be neglected.

EXAMPLE (continued)

$$\hat{\lambda}_1 = \frac{11}{77} = 0.143 \quad \hat{\lambda}_2 = \frac{6.5}{45} = 0.144 \quad \hat{\lambda}_3 = \frac{1}{20} = 0.05$$

$$\hat{\delta}_2 = \frac{2}{5.5} = 0.364 \quad \hat{\delta}_3 = \frac{0.5}{5} = -0.1$$

$$\hat{R} = 0.100 + 0.159 + 0.050 = 0.309$$

#### 2.4. Loss reserving

The loss reserve for accident year  $t$  is

$$L_t = E[X_{t,n} \mid H_n]$$

Under assumption (A<sub>1</sub>) and (A<sub>2</sub>) it is easily seen that

$$\begin{aligned}
 (2.2.6) \quad L_i = & X_{i,n+1-i}(1-\delta_{n+2-i}) \dots (1-\delta_n) \\
 & + E_i[\lambda_{n+2-i}(1-\delta_{n+3-i}) \dots (1-\delta_n) \\
 & \quad + \lambda_{n+3-i}(1-\delta_{n+4-i}) \dots (1-\delta_n) \\
 & \quad \cdot \\
 & \quad \cdot \\
 & \quad \cdot \\
 & \quad + \lambda_{n-1}(1-\delta_n) \\
 & \quad + \lambda_n]
 \end{aligned}$$

i.e. the loss reserve consists in a component for IBNER claims and a component for IBNR claims the former depending on the claims observed so far and the latter on the exposure.

One obtains an estimate of  $L_i$  by replacing the parameters in (2.2.6) by their estimates (2.2.3) and (2.2.4) respectively.

EXAMPLE (continued)

Accident year $i$	$X_{i,n+1-i}$	$A_{n+1-i}$	IBNER <sub><math>i</math></sub>	$E_i$	IBNR <sub><math>i</math></sub>	$L_i$
1	6.5	1	6.5	20	0	6.5
2	5	1.1	5.5	25	1.25	6.75
3	5.5	0.700	3.85	32	6.67	10.52
	17					23.77

Where  $A_{n+1-i} = (1-\delta_{n+2-i}) \dots (1-\delta_n)$  is the IBNER correction factor.

To compute the loss reserves in practice we will of course use the original claims as opposed to the revalued claims used for pricing purposes; we will also have to choose a suitable measure of exposure.

It is interesting to compare (2.2.6) to the formulas for loss reserve provided by the chain-ladder method and by the Bornhuetter-Ferguson method respectively

The loss reserve for accident year  $i$  according to the chain-ladder method is:

$$(2.2.7) \quad L_i = X_{i,n+1-i} \cdot F_{n+1-i}$$

Where  $F_j$  is some factor pertaining to development year  $j$  (for details see for instance Nationale-Nederlanden [2]) The same quantity as estimated by the Bornhuetter-Ferguson method is:

$$(2.2.8) \quad L_i = X_{i,n+1-i} + E_i \cdot G_{n+1-i}$$

Where  $G_{n+1-i}$  is a factor which is applied to the exposure.

With a suitable notation we can rewrite (2.2.6) in the following way:

$$(2.2.9) \quad L_i = X_{i,n+1-i} \Delta_{n+1-i} + E_i A_{n+1-i}$$

It is seen that formally our estimator is a generalisation of both the chain-ladder and the Bornhuetter-Ferguson estimator:  $\Delta_{n+1-i} = F_{n+1-i}$  and  $A_{n+1-i} = 0$  gives the chain-ladder estimator whereas  $\Delta_{n+1-i} = 1$  and  $A_{n+1-i} = G_{n+1-i}$  gives the Bornhuetter-Ferguson estimator.

**2.5. Performance of the estimator**

We now want to assess the performance of  $R(\hat{\theta})$  defined in (2.2.5) In order to do so we need the following stronger assumptions.

$$(A_1) \quad E[N_y \mid H_{i+j-2}] = E_i \lambda_j \quad \text{Var} [N_y] = E_i \sigma_j^2$$

$$(A_2) \quad E[D_y \mid H_{i+j-2}] = X_{i,j-1} \delta_j \quad \text{Var} [D_y \mid H_n] = X_{i,j-1} \tau_j^2$$

Developing  $R(\hat{\theta})$  in a Taylor series, we obtain:

$$(2.3.1) \quad R(\hat{\theta}) \simeq R(\theta) + \sum_{i=1}^{2n-1} \frac{\delta R(\theta)}{\delta \theta_i} (\hat{\theta}_i - \theta_i)$$

(A<sub>3</sub>) implies that  $\hat{\theta}_i$  and  $\hat{\theta}_j$  are not strongly correlated for  $i \neq j$  hence

$$(2.3.2) \quad \text{mse} (R(\hat{\theta})) = E(R(\hat{\theta}) - R(\theta))^2 \simeq \sum_{i=1}^{2n-1} \left( \frac{\delta R(\theta)}{\delta \theta_i} \Big|_{\theta=\hat{\theta}} \right)^2 \text{Var} (\hat{\theta}_i)$$

where we have replaced the unknown quantities.

$$\frac{\delta R(\theta)}{\delta \theta_i}$$

by the approximations:

$$\frac{\delta R(\theta)}{\delta \theta_i} \Big|_{\theta=\hat{\theta}}$$

We still have to find approximations for the  $\text{Var} (\hat{\theta}_i)$  From (A<sub>1</sub>), (A<sub>2</sub>) and (A<sub>3</sub>) it follows that.

$$(2.3.3) \quad \text{Var} (\hat{\lambda}_j) = \frac{\sigma_j^2}{\sum_{i=1}^{n+1-j} E_i} \quad j = 1, \dots, n$$



$$(2.3.4) \quad \text{Var}(\hat{\delta}_j) = \frac{\tau_j^2}{\sum_{i=1}^{n+1-j} X_{i,j-1}} \quad j = 2, \dots, n$$

on the other hand we have the following biasfree estimators of  $\sigma_j^2$  and  $\tau_j^2$  respectively

$$(2.3.5) \quad \hat{\sigma}_j^2 = \frac{1}{n-j} \sum_{i=1}^{n+1-j} (N_y - \hat{\lambda}_j E_i)^2 \frac{1}{E_i} \quad j = 1, \dots, n-1$$

$$(2.3.6) \quad \hat{\tau}_j^2 = \frac{1}{n-j} \sum_{i=1}^{n+1-j} (D_y - \hat{\delta}_j X_{i,j-1})^2 \frac{1}{X_{i,j-1}} \quad j = 2, \dots, n-1$$

and if there are enough development years at hand we have:

$$\hat{\lambda}_n = 0 \quad \text{and} \quad \hat{\delta}_n = 0$$

and one may assume:

$$\sigma_n^2 = 0 \quad \text{and} \quad \tau_n^2 = 0.$$

Plugging the expressions given above into (2.3.2) we obtain an approximation for the mean square error of  $R(\hat{\theta})$

EXAMPLE (continued)

$$\frac{\delta R}{\delta \lambda_1} = (1 - \hat{\delta}_2)(1 - \hat{\delta}_3) = 0.700 \quad \frac{\delta R}{\delta \lambda_2} = (1 - \hat{\delta}_3) = 1.1 \quad \frac{\delta R}{\delta \lambda_3} = 1$$

$$\frac{\delta R}{\delta \delta_2} = -\hat{\lambda}_1(1 - \hat{\delta}_3) = -0.157 \quad \frac{\delta R}{\delta \delta_3} = -\hat{\lambda}_1(1 - \hat{\delta}_2) - \hat{\lambda}_2 = -0.235$$

$$\text{Var}(\hat{\lambda}_1) = \frac{\sigma_1^2}{E_1 + E_2 + E_3} = 48 \cdot 10^{-5} \quad \text{Var}(\hat{\lambda}_2) = \frac{\sigma_2^2}{E_1 + E_2} = 2 \cdot 10^{-5}$$

$$\text{Var}(\hat{\lambda}_3) = \frac{\sigma_3^2}{E_1} = 0$$

$$\text{Var}(\hat{\delta}_2) = \frac{\tau_2^2}{X_{11} + X_{21}} = 110 \cdot 10^{-5} \quad \text{Var}(\hat{\delta}_3) = \frac{\tau_3^2}{X_{12}} = 0$$

from which one obtains

$$\text{mse}^{1/2}(R(\hat{\theta})) = 0.017$$

Another possibility to evaluate (2.3.3) and (2.3.4) is to specify a parametric model. An example is given in the next section.

3. A MODEL FOR CLAIM NUMBERS

We use the same definitions as in section 2 with the difference that claim amounts are now replaced by claim numbers:  $X_{ij}$  denotes the number of excess claims from accident year  $i$  in development year  $j$ .  $D_{ij}$  is the decrease in total number of claims between development year  $j-1$  and development year  $j$  with respect to claims already known as excess claims in year  $j-1$ . ( $D_{ij}$  is a non-negative integer smaller or equal to  $X_{i,j-1}$ ).  $N_{ij}$  denotes the number of new excess claims pertaining to accident year  $i$  in development year  $j$ . Relations (2.1.1) and (2.1.2) hold true.

EXAMPLE (continued)

From the individual claims of the example of section 2 we obtain the following IBNR triangle for claim numbers.

*X*-triangle

		<i>J</i>		
		1	2	3
<i>i</i>	1	2	4	4
	2	3	4	
	3	5		

*N*-triangle

		<i>J</i>		
		1	2	3
<i>i</i>	1	2	3	1
	2	3	3	
	3	5		

*D*-triangle

		<i>J</i>		
		1	2	3
<i>i</i>	1	—	1	1
	2	—	2	
	3	—		

Under assumptions (A<sub>1</sub>) and (A<sub>2</sub>) relation (2.2.2) holds true.  $R(\theta)$  is now the expected ultimate claims frequency and  $\delta_j$  is the probability for an excess claim to drop below the priority between development year  $j-1$  and development year  $j$ .

The expressions given in (2.2.3) and (2.2.4) are biasfree estimates of the  $\lambda$ 's and  $\delta$ 's respectively. (2.2.5) gives an estimate of the ultimate claims frequency  $R(\theta)$ . The bias of the estimate  $R(\hat{\theta})$  can be neglected.

EXAMPLE (continued)

$$\hat{\lambda}_1 = \frac{10}{77} = 0.130, \quad \hat{\lambda}_2 = \frac{6}{45} = 0.133, \quad \hat{\lambda}_3 = \frac{1}{20} = 0.05$$

$$\hat{\delta}_2 = \frac{3}{5} = 0.6, \quad \hat{\delta}_3 = \frac{1}{4} = 0.25$$

$$R(\theta) = 0.189$$

The performance of  $R(\theta)$  can be assessed with (2.3.2).

We now make the following parametric assumptions:

$$(A_1'') \quad N_{ij} \mid H_{i+j-2} \sim \text{Poisson}(\lambda_j \cdot E_i)$$

$$(A_2'') \quad D_{ij} \mid H_{i+j-2} \sim \text{Binomial}(\delta_j, X_{i,j-1}).$$

It is easily seen that:

$$(A_i'') \Rightarrow (A_i') \Rightarrow (A_i) \quad i = 1, 2.$$

We also assume that  $(A_3)$  holds true. The log likelihoods of the parameters are:

$$(3.1) \quad l(\lambda_j) = - \left( \sum_{i=1}^{n+1-j} E_i \right) \lambda_j + \left( \sum_{i=1}^{n+1-j} N_{ij} \right) \log \lambda_j$$

$$(3.2) \quad l(\delta_j) = \left( \sum_{i=1}^{n+1-j} D_{ij} \right) \cdot \log(\delta_j) + \left( \sum_{i=1}^{n+1-j} X_{i,j-1} - \sum_{i=1}^{n+1-j} D_{ij} \right) \log(1 - \delta_j)$$

and it is seen, that the  $\hat{\lambda}_j$  and  $\hat{\delta}_j$  of (2.2.3) and (2.2.4) are *the maximum likelihood estimates* of the  $\lambda_j$ 's and  $\delta_j$ 's.

From the maximum likelihood theory we know that

$$\text{Var}(\hat{\lambda}_j) \rightarrow \left[ -E \left( \frac{\delta^2 l}{\delta^2 \lambda_j} \right) \right]^{-1} \quad \text{for} \quad \sum_{i=1}^{n+1-j} E_i \rightarrow \infty$$

we therefore use the following approximations:

$$(3.3) \quad \text{Var}(\hat{\lambda}_j) \simeq - \left( \frac{\delta^2 l}{\delta^2 \lambda_j} \Big|_{\lambda=\hat{\lambda}_j} \right)^{-1}$$

$$\text{Var}(\hat{\lambda}_j) \simeq \frac{\hat{\lambda}_j}{\sum_{i=1}^{n+1-j} E_i} \quad j = 1, \dots, n$$

analogously:

$$(3.4) \quad \text{Var}(\hat{\delta}_j) \simeq \frac{\hat{\delta}_j(1-\hat{\delta}_j)}{\sum_{i=1}^{n+1-j} X_{i,j-1}} \quad j = 2, \dots, n$$

and we obtain an approximation of the mean square error of  $R(\hat{\theta})$  by plugging (3.3) and (3.4) into (2.3.2)

EXAMPLE (continued)

$$\begin{aligned} \text{Var}(\hat{\lambda}_1) &= 17 \cdot 10^{-4} & \text{Var}(\hat{\lambda}_2) &= 30 \cdot 10^{-4} & \text{Var}(\hat{\lambda}_3) &= 25 \cdot 10^{-4} \\ \text{Var}(\hat{\delta}_2) &= 480 \cdot 10^{-4} & \text{Var}(\hat{\delta}_3) &= 469 \cdot 10^{-4} & & \end{aligned}$$

$$\frac{\delta R}{\delta \lambda_1} = 0.3 \quad \frac{\delta R}{\delta \lambda_2} = 0.75 \quad \frac{\delta R}{\delta \lambda_3} = 1$$

$$\frac{\delta R}{\delta \delta_2} = -0.097 \quad \frac{\delta R}{\delta \delta_3} = -0.185$$

$$\text{mse}^{1/2} [R(\hat{\theta})] = 0.080$$

#### 4. A PRACTICAL PRICING EXAMPLE

The following IBNR triangle ( $X$ -triangle) is borrowed from a practical motor third party liability excess of loss pricing problem:

acc year \ dvpt year								Exposure
	1	2	3	4	5	6	7	
1	7.5	28.9	52.6	84.5	80.1	76.9	79.5	10'224
2	1.6	14.8	32.1	39.6	55.0	60.0		12'752
3	13.8	42.4	36.3	53.3	96.5			14'875
4	2.9	14.0	32.5	46.9				17'365
5	2.9	9.8	52.7					19'410
6	1.9	29.4						17'617
7	19.1							18'129

The excess claims and the measure of exposure (premium of the whole portfolio) have been revalued. Based on these 'as if' statistics we want to estimate the ultimate burning cost

Using the chain-ladder method we obtain:

Accident year	Exposure	Total Claims Amount per dvpt year $n+1-i$	Cumulative Factor	Estimated Ultimate Claims Amount	Estimated Ultimate Burning Cost
1	10'224	79.5	1	79.5	0.78%
2	12'752	60	1.03	62	0.49%
3	14'875	96.5	1.05	101.1	0.68%
4	17'365	46.9	1.37	64	0.37%
5	19'410	52.7	2.00	105.3	0.54%
6	17'617	29.4	3.75	110.2	0.63%
7	18'129	19.1	17.07	326.0	1.80%
	110'372			848.3	0.77%

(For details on the chain-ladder method see for instance Nationale-Nederlanden [2]).

It is seen at once that the estimated ultimate burning cost pertaining to accident year 7 is much larger than the other estimated burning costs. This is due to a well known problem inherent to the chain-ladder method: the claims amount of the least developed accident year is multiplied with the largest cumulative factor providing thus a very imprecise estimate which can heavily influence the overall ultimate burning cost. This drawback of the chain-ladder method can easily be corrected by weighing the estimated ultimate burning costs of the individual accident years in a different way. Let  $F_j$  denote the cumulative factor provided by the chain-ladder method which is to be applied to the claims amount of development year  $j$ .  $X_j$ ,  $E_i$  and  $R$  denote respectively the total claims amount, the exposure and the ultimate burning cost as defined in section 2. The estimated ultimate burning cost pertaining to accident year  $i$  is then:

$$\frac{X_{i,n+1-i} \cdot F_{n+1-i}}{E_i}$$

The chain-ladder method weighs these estimates with  $E_i$ , the exposure of the corresponding accident year, thus giving the following overall estimated ultimate burning cost

$$R = \frac{\sum_{i=1}^n X_{i,n+1-i} \cdot F_{n+1-i}}{\sum_{i=1}^n E_i}$$

Instead of  $E_i$  we use the following weights:

$$\frac{E_i}{F_{n+1-i}}$$

which correspond to 'used exposure' and give less weight to less developed accident years.

We obtain the following overall estimated burning cost:

$$R = \frac{\sum_{i=1}^n X_{i,n+1-i}}{\sum_{i=1}^n \frac{E_i}{F_{n+1-i}}}$$

We have thus rederived a special case of the Cape Cod method [3], an IBNR method similar to the Bornhuetter-Ferguson method [1]. This method provides the following estimates:

Accident year	Exposure	Total Claims Amount as per dvpt year $n+1-i$	Cumulative Factor	'Used Exposure'	Estimated Ultimate Burning Cost
1	10'224	79.5	1	10'224	0.78%
2	12'752	60	1.03	12'335	0.49%
3	14'875	96.5	1.05	14'199	0.68%
4	17'365	46.9	1.37	12'697	0.37%
5	19'410	52.7	2.00	9'712	0.54%
6	17'617	29.4	3.75	4'698	0.63%
7	18'129	19.1	17.07	1'062	1.80%
		384.1		64'928	0.59%

We now consider the more detailed statistics of the  $N$ - and  $D$ -triangles. The statistics of new IBNR claims are:

acc year \ dvpt year	dvpt year						
	1	2	3	4	5	6	7
1	7.5	18.3	28.5	23.4	18.6	0.7	5.1
2	1.6	12.6	18.2	16.1	14.0	10.6	
3	13.8	22.7	4.0	12.4	12.1		
4	2.9	9.7	16.4	11.6			
5	2.9	6.9	37.1				
6	1.9	27.5					
7	19.1						

$N$ -triangle

The statistics of decreases in the claims amount are:

acc year \ dvpt year	2	3	4	5	6	7
1	-31	48	-85	230	39	25
2	-06	09	86	-14	56	
3	-59	101	-46	-311		
4	-14	-21	-28			
5	0	-58				
6	0					

*D*-triangle

The striking feature of these more detailed statistics is that even in development year 6 and 7 there is an important amount of new claims to the layer, however this fact is partly compensated by a decrease of the amount of already known excess claims and therefore the less detailed traditional IBNR statistics give the spurious impression that the total amount of excess claims is exactly known after six or seven development years which is obviously not the case in this example.

We now want to estimate the ultimate burning cost with our method. From (2.2.3) and (2.2.4) we obtain:

$j$	$\hat{\lambda}_j$	$\hat{\delta}_j$
1	0.45 $10^{-3}$	
2	1.06 $10^{-3}$	-0.359
3	1.40 $10^{-3}$	0.072
4	1.15 $10^{-3}$	-0.048
5	1.18 $10^{-3}$	-0.054
6	0.49 $10^{-3}$	0.070
7	0.50 $10^{-3}$	0.033

We see that the  $\hat{\lambda}$ 's reach a maximum in year 3 and decrease thereafter but it would be misleading to assume that  $\lambda_j = 0$  for  $j \geq 8$ .

Between the 1st and the 2nd development year there is an important increase of the known excess claims, after that the excess increase or decrease more or less randomly and the  $\hat{\delta}$ 's oscillate around zero.

By plugging the parameters into (2.2.5) we obtain the following estimate for the ultimate burning cost:

$$R(\hat{\theta}) = 0.61\%,$$

An estimate which is almost identical to the one obtained with the Cape Cod method.

Under assumptions  $(A_1)$ ,  $(A_2)$  and  $(A_3)$  we know that  $R(\hat{\theta})$  is a practically biasfree estimate of  $R(\theta)$ , whereas neither in the case of the chain-ladder estimate nor in the case of the Cape Cod estimate do we know anything about the bias of the estimator.

We now make the stronger assumptions  $(A'_1)$ ,  $(A'_2)$  and  $(A'_3)$  and we estimate  $\sigma_j$  and  $\tau_j$  according to (2.3.5) and (2.3.6).

$j$	$\hat{\sigma}_j$	$\hat{\tau}_j$
1	0 054	
2	0 074	0 387
3	0 109	1 269
4	0 079	1 177
5	0 056	3 460
6	0 057	0 303
7	0	0

The assumption  $\hat{\sigma}_7 = 0$  and  $\hat{\tau}_7 = 0$  is not very realistic, however it has little impact on the mean square error of  $R(\hat{\theta})$ . From (2.3.3) and (2.3.4) we now obtain the standard deviations of the estimators of our parameters.

$j$	$\sigma(\hat{\lambda}_j)$	$\sigma(\hat{\delta}_j)$
1	0 16 $10^{-3}$	
2	0 24 $10^{-3}$	0 070
3	0 40 $10^{-3}$	0 121
4	0 34 $10^{-3}$	0 095
5	0 29 $10^{-3}$	0 260
6	0 38 $10^{-3}$	0 026
7	0	0

We also need the following expressions:

$$\frac{\delta R}{\delta \lambda_1} = \Delta_1 = 1.253 \quad \frac{\delta R}{\delta \lambda_2} = \Delta_2 = 0.921 \quad \frac{\delta R}{\delta \lambda_3} = \Delta_3 = 0.993$$

$$\frac{\delta R}{\delta \lambda_4} = \Delta_4 = 0.948 \quad \frac{\delta R}{\delta \lambda_5} = \Delta_5 = 0.899 \quad \frac{\delta R}{\delta \lambda_6} = \Delta_6 = 0.967$$

$$\frac{\delta R}{\delta \lambda_7} = 1$$

$$\frac{\delta R}{\delta \delta_2} = -\lambda_1 \Delta_1 \frac{1}{1 - \delta_2} = -0.00041$$

$$\frac{\delta R}{\delta \delta_3} = -[\lambda_1 \Delta_1 + \lambda_2 \Delta_2] \frac{1}{1 - \delta_3} = -0.00166$$



$$\frac{\delta R}{\delta \delta_4} = -[\lambda_1 A_1 + \lambda_2 A_2 + \lambda_3 A_3] \frac{1}{1 - \delta_4} = -0.00279$$

$$\frac{\delta R}{\delta \delta_5} = -[\lambda_1 A_1 + \dots + \lambda_4 A_4] \frac{1}{1 - \delta_5} = -0.00381$$

$$\frac{\delta R}{\delta \delta_6} = -[\lambda_1 A_1 + \dots + \lambda_5 A_5] \frac{1}{1 - \delta_6} = -0.00546$$

$$\frac{\delta R}{\delta \delta_7} = -[\lambda_1 A_1 + \dots + \lambda_6 A_6] \frac{1}{1 - \delta_7} = -0.00574$$

From (2.3.2) we now obtain

$$\text{mse}^{1/2}(R(\hat{\theta})) = 0.13\%$$

Our method also provides a measure of the precision of the point estimator.

To summarize what we have obtained so far we can say that we have an estimate of the burning cost after seven development years (0.61%), this estimate is practically unbiased and reasonably precise since its standard deviation is (0.13%). Our detailed statistics have shown us that there are still some excess claims to be expected in the following development years, a fact which we would have overlooked if we had only used the usual IBNR statistics. To assess the impact of further development years on the ultimate burning cost we can use the experience of similar portfolios or some market statistics if that kind of data is available, if such is not the case we can extrapolate our estimates of the  $\lambda$ 's and of the  $\delta$ 's

Based on the analysis of the given portfolio, a realistic extrapolation would be:

$$\begin{aligned} \lambda_8 = \lambda_9 &= 0.5 \cdot 10^{-3} \\ \lambda_j &= 0 & j &= 10, 11 \\ \delta_j &= 0 & j &= 8, 9, \end{aligned}$$

Thus our estimate of the ultimate burning cost is

$$R = 0.71\%$$

#### REFERENCES

- [1] BORNHÜETTER, R. L. and FERGUSON, R. E. (1972) *The Actuary and IBNR* Proceedings of the Casualty Actuarial Society LIX, 181-195
- [2] *Loss Reserving Methods* Surveys of Actuarial Studies No 1, a publication of the Nationale-Nederlanden N.V., 1981
- [3] STRAUB, E. (1988) *Non-Life Insurance Mathematics* Springer Verlag

R. SCHNIEPER

'Winterthur' Swiss Insurance Company, Rudolfstr. 1, P.O. Box 286,  
CH-8401 Winterthur, Switzerland



## DISCUSSION PAPERS

### THE SCHMITTER PROBLEM

BY P. BROCKETT, M. GOOVAERTS, G TAYLOR

At the ASTIN Colloquium in Montreux, HANS SCHMITTER posed the following problem.

#### PROBLEM

Consider the class  $\mathcal{F}$  of distributions with range  $[0, b]$ , mean  $\mu$  and variance  $\sigma^2$ . Let  $\Psi_{\theta, F}(u)$  denote the probability of ultimate ruin under a compound Poisson claim process with given premium loading  $\theta$ , initial capital  $u$  and individual claim size  $d \cdot f \cdot F$ . For fixed  $\theta$  and  $\mu$ , which  $F \in \mathcal{F}$  maximizes  $\Psi_{\theta, F}(u)$  for a particular given  $u$ ? In particular, is  $F$  diatomic?

#### PRACTICAL BACKGROUND OF THE PROBLEM

H. SCHMITTER describes the following practical background in which the problem arises.

The problem of determining bounds for ruin probabilities arises when an insurer decides his reinsurance retentions in order to increase the stability of an account. He may not only choose between various forms of reinsurance (quota share, surplus, excess loss etc.) but he usually combines them in what is called a reinsurance program. When evaluating reinsurance programs he needs to compare their prices and the effectiveness of the protection they offer. The reinsurance price is the difference between the gross (i.e. before reinsurance) and the net (i.e. retained, after reinsurance) expected profit. The effectiveness of the protection, on the other hand, can be measured by the probability of ruin: the lower the probability of ruin of the retained account the more effective the reinsurance program. Computing ruin probabilities is often criticized as being pointless because their absolute values are said to be irrelevant. However, if two reinsurance programs both reduce the expected profit of the ceding company by the same amount the one leading to the smaller probability of ruin is likely to be preferable.

The ruin probability depends on the initial reserve (known to the ceding company), the security loading (defined as the expected retained profit, hence a function of the reinsurance program) and on the retained claim amount distribution. In practice, the latter is hardly ever known, apart from the maximum retained claim which is given by an excess loss deductible or a policy limit. At best we have to our disposal estimates of the expected value and the variance. An exact computation of the ruin probability is, therefore, not possible and one has to accept the determination of upper and lower bounds.

So far we do not even know the least upper bound in the case where the expected value, the variance and the maximum claim are known. Perhaps the answer to the above question is not an isolated problem but leads to further investigations and applications: Suppose that for several independent risks the expected profits, frequencies, expected values, variances and maximum claims are known. What is the least upper bound of the overall ruin probability for a given initial reserve? Is there a natural way of allocating parts of the initial reserve to the independent risks? A question often asked in practice.

#### DISCUSSION

At Montreux, GREG TAYLOR pointed out that  $F$  more dangerous than  $G$  in stop-loss order implies that  $\Psi_{\theta, F}(u) \geq \Psi_{\theta, G}(u)$  for all  $u$  (GOOVAERTS and DE VYLDER, 1984; TAYLOR, 1985)

Hence the problem is reduced to seeking an extremal distribution in  $\mathcal{F}$  in terms of stop-loss order. However an extremal distribution in terms of stop-loss order does not exist in class  $\mathcal{F}$ .

The problem was further discussed at the "1990 Risk Theory seminar at the Mathematisches Forschungsinstitut of the Federal Republic of Germany, in Oberwolfach"

MARC GOOVAERTS pointed out that an upper bound can be obtained by the criterion of danger which satisfies the range  $[0, b]$ ,  $\mu$  but not  $\sigma^2$  where now danger is defined as in BÜHLMANN et al (1977). One can deduce a distribution which is more dangerous than all of those belonging to the class of distributions with prescribed range, mean  $\mu$  but with a minimal variance, larger than  $\sigma^2$  in analogy to KAAS and GOOVAERTS (1986).

But only danger as well as first order stop-loss ordering will give rise to inequalities between ruin probabilities. If we have  $E(X) = E(Y)$  and  $E((X-t)_+) < E((Y-t)_+) \forall t$  then  $\Psi_{\theta, F_1}(u) \leq \Psi_{\theta, F_2}(u)$  uniformly for all  $\theta$  and  $u$ . The problem of finding  $\text{Sup}_{F_1 \in \mathcal{F}} E((X-t)_+)$  does not give rise to a uni-

form (in  $t$ ) extremal distribution.

It is solved by constructing a polynomial of degree two above  $(X-t)_+$  which is tangent to this function in 2 points. The abscissas of these points will be the mass points (a recent reference is e.g. GOOVAERTS et al., 1990). These results are known but they cannot be used to obtain an upper bound for the infinite time ruin probability because the extremal distribution depends on the value of  $t$ .

One finds the following solution: A risk  $X$  with spectrum  $(r, s)$  exists with mean  $\mu$  and variance  $\sigma^2$  if and only if  $s = r'$ , where  $r' = \mu + [\sigma^2/(\mu - r)]$

The following mass points of the extremal distributions are obtained:  $(0, 0')$  in case  $0 \leq t \leq 1/2 \cdot 0'$ ,  $(t + \sqrt{(\mu - t)^2 + \sigma^2}, t - \sqrt{(\mu - t)^2 + \sigma^2})$  in case  $1/2 \cdot 0' \leq t \leq 1/2(b + b')$  and  $(b, b')$  in case  $1/2(b + b') \leq t \leq b$ . This indicates that even for the simple extremal stop-loss problem no uniform extremal distribution exists. Also BROCKETT and COX (1985, 1986) present explicit solutions to the above problem when  $n = 1, 2$  or 3 moments are given using

Tchebycheff systems of functions. KEMPERMANN (1970) also solves this problem in general.

A problem closely related to the one stated by SCHMITTER and as intriguing is the following: consider  $S = X_1 + \dots + X_N$  under the classical assumptions and find  $\text{Sup}_{F \in \mathcal{F}} E((S-t)_+)$ .

This problem can be solved for the case  $F_X \in \mathcal{F}_1$  (= a set of distributions with given  $\mu$  and  $b$ ), (see BUHLMANN, GAGLIARDI, GERBER, and STRAUB, 1977) An attempt to solve the above problem ( $F_X \in \mathcal{F}$ ) has been presented by KAAS and GOOVAERTS (1984), cited above.

Also at Oberwolfach, P BROCKETT demonstrated that the  $F \in \mathcal{F}$  which minimizes the adjustment coefficient  $R$  of the claim process lies in the class  $D_2$  of diatomic distributions. Since  $\Psi_{\theta, F}(u) \sim \text{const. } e^{-Ru}$  for large  $u$ , this implies that the required  $F$  lies in  $D_2$  for sufficiently large  $u$ . It does not, however, identify  $F$  for smaller values of  $u$ . In fact, the extremal  $F$  for large  $u$  can be identified as follows:

$$\text{Mass } p = (b-\mu)^2/[\sigma^2 + (b-\mu)^2] \text{ at } \mu - \sigma^2/(b-\mu); \text{ and Mass } 1-p \text{ at } b.$$

Similar results can be obtained for maximizing the adjustment coefficient. These results can also be found in DE VYLDER, GOOVAERTS and HAEZENDONCK (1984), BROCKETT and COX (1983, 1986) and KEMPERMANN (1970, 1971).

GREG TAYLOR suggested that, to the extent that Schmitter's problem related to premium rating (as SCHMITTER had said it did), that problem was probably not the most relevant for solution. In practice, the assumption of unimodality of  $F$  would almost always be reasonable, and this additional restriction on  $F$  could be expected to decrease the upper bound on  $\Psi_{\theta, F}(u)$  substantially

Moreover, this additional condition does not add to the difficulty of the problem. The history of this goes back to VERBEEK (1977), who dealt with the extremal unimodal stop-loss premium with fixed mean and upper bound, and TAYLOR (1977) who extended the results to the context of an arbitrary finite number of linear constraints on the unimodal distribution. Much extension has subsequently been made by GOOVAERTS (and co-authors) and BROCKETT and COX

The relevant result for Schmitter's problem if unimodality is required is that the extremal distribution must lie in the class  $\mathcal{S}_3$  of step functions with 3 levels (with possible equality of 2 or 3 levels).

BROCKETT and COX (1985, 1986) demonstrate that the unimodal process lies in the class  $\mathcal{S}_2$ . As in the case where unimodality is not required, they give an explicit optimal solution to bounding the adjustment coefficient. They give the corresponding solution for an arbitrary finite number of linear constraints on  $F$ , and it is again true that his extremal distribution solves Schmitter's problem for sufficiently large  $u$ .

## REFERENCES

- BOWERS, N L Jr (1969) An upper bound on stop-loss net premium *Transactions of the Society of Actuaries* **21**
- BOWERS, N L Jr, GERBER, H U, HICKMAN, J E, JONES, D A and NESBITT, C J (1986) *Actuarial Mathematics* Society of Actuaries
- BÜHLMANN, H, GAGLIARDI, B, GERBER, H U and STRAUB, E (1977) Some inequalities for stop-loss premiums *ASTIN Bulletin* **9**, 169–177
- BROCKETT, P and COX, S Jr (1985) Insurance calculation using incomplete information *Scand Actuarial J.* 94–108
- BROCKETT, P and COX, S Jr (1985) Optimal Ruin Calculations using Partial Stochastic Information *Transactions of the Society of Actuaries*
- GOOVAERTS, M, DE VYLDER, F and HAEZENDONCK J (1984) *Insurance Premiums* North Holland
- GOOVAERTS, M, KAAS, R, VAN HEERWAERDEN, A E and BAUWELINCKX, T (1990) *Effective Actuarial Methods* North Holland
- JANSSEN, K, HAEZENDONCK, J and GOOVAERTS, M J (1986) Upper bounds on stop-loss premium in case of known moments up to the fourth order *Insurance Mathematics and Economics* **5**, 315–334
- KAAS, R (1987) *Bounds & Approximations for some Risk Theoretical Quantities* University of Amsterdam
- KAAS, R and GOOVAERTS, M J (1985) Bounds on distribution functions under integral constraints *Bulletin, Association Royale des Actuairev Belges* **79**, 45–60
- KAAS, R and GOOVAERTS, M J (1986) Bounds on Stop-loss premiums for compound distributions *ASTIN Bulletin XVI*, 13–17
- KARLIN, K and STUDDEN W J (1966) *Chebyshev systems with application in analysis and statistics* Interscience publishers
- KEMPERMANN, J H B (1970) On a class of moment problems *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability*, 101–126
- KEMPERMANN, J H B (1971) Moment problems with convexity conditions *Optimizing methods in statistics* Academic press, 115–178
- KREIN, M G (1951) The idea of P L CHEBYCHEV and A A MARKOV in the theory of limiting values of integrals and their further developments *Am Math Soc Transl Ser. 2*, **12**, 1–122
- MACK, T (1984) Calculation of the maximum stop-loss premium, given the first three moments of the loss distribution *Proceedings of the 4 countries Astin-Symposium*
- TAYLOR, G C (1985) A Heuristic Review of some Ruin Theory Results *ASTIN Bulletin* **15**, 73–88
- TAYLOR, G C (1977) Upper Bounds on Stop-Loss Premiums under Constraints on Claim Size Distributions *S A J.* 94–105
- DE VYLDER, F and GOOVAERTS, M (1982b) Analytical best upper bounds for stop-loss premiums *Insurance Mathematics and Economics* **1**, 3

P. BROCKETT, M. GOOVAERTS, G. TAYLOR  
*University of Texas at Austin, CBA-7.202 Austin Texas*  
*K.U. Leuven & Univ Amsterdam, de Bériotstraat 34, B-3000 Leuven*  
*Coopers & Lybrand, 580 George Street, Sydney, NSW 2001, Australia.*

# THE SCHMITTER PROBLEM AND A RELATED PROBLEM: A PARTIAL SOLUTION

BY R. KAAS

*University of Amsterdam, the Netherlands*

## ABSTRACT

At the 1990 ASTIN-colloquium, SCHMITTER posed the problem of finding the extreme values of the ultimate ruin probability  $\psi(u)$  in a risk process with initial capital  $u$ , fixed safety margin  $\theta$ , and mean  $\mu$  and variance  $\sigma^2$  of the individual claims. This note aims to give some more insight into this problem. Schmitter's conjecture that the maximizing individual claims distribution is always diatomic is disproved by a counterexample. It is shown that if one uses the distribution maximizing the upper bound  $e^{-Ru}$  to find a 'large' ruin probability among risks with range  $[0, b]$ , incorrect results are found if  $b$  is large or  $u$  small

The related problem of finding extreme values of stop-loss premiums for a compound Poisson ( $\lambda$ ) distribution with identical restrictions on the individual claims is analyzed by the same methods. The results obtained are very similar.

## 1. INTRODUCTION

In a paper presented at the ASTIN-colloquium 1990, HANS SCHMITTER gives a derivation of an exact algorithm to compute the value of the ultimate ruin probability  $\psi(u)$  for a compound Poisson ruin process with given premium income  $c$  per unit of time, and with claims having a finite number of mass points. In connection with this paper, he posed the following problem: given that the individual claims have mean  $\mu$  and variance  $\sigma^2$ , which claims distributions minimize and maximize the ruin probability for a given  $u$ ? A practical justification of the problem can be found in the paper by BROCKETT, GOOVAERTS and TAYLOR (1991), who also sum up the results of the discussion of this matter at the colloquia of Montreux and subsequently Oberwolfach.

In the classical ruin model, the non-ruin probability of a compound Poisson risk process can be shown to have a compound geometric distribution with geometric parameter depending only on the safety loading  $\theta$ , and with terms having a distribution function related to the stop-loss premiums of the individual claims.

In this note we also describe another problem, very similar to Schmitter's. Suppose a reinsurer has to determine a stop-loss premium for a risk with the following properties. the risk has a compound Poisson distribution with known

parameter  $\lambda$ , and the individual claims have known mean  $\mu$  and variance  $\sigma^2$ . To be able to quote a safe premium, the reinsurer tries to determine the claims distribution leading to the maximum value of the net stop-loss premium. Some work in this direction was done by KAAS and GOOVAERTS (1986) and STEENACKERS and GOOVAERTS (1990). See also GOOVAERTS et al. (1984)

A lower bound for both the ruin probability and the compound Poisson stop-loss premium under these restrictions is attained by the distribution concentrating all mass at  $\mu$ , see for instance GOOVAERTS et al (1990). This distribution is not actually an element of the set of feasible distributions, which is not a closed set. We will prove that both our functionals, ruin probabilities and compound Poisson stop-loss premiums, are continuous at this boundary point. Other functionals, like the variance, the skewness and the adjustment coefficient do not have this property. See Section 2.

In this paper we concentrate on the upper bounds, and indicate how one may find the diatomic claims distribution leading to the highest ruin probability using the algorithm mentioned above. The compound stop-loss premium can be computed by a very similar formula, based on special properties of the compound Poisson distribution. See Section 2. We found counterexamples for Schmitter's conjecture that the maximal ruin probability always is realized by a diatomic distribution. For the compound Poisson stop-loss premiums, the optimal diatomic distribution also was not always the overall maximum. See Section 3.

A useful heuristic approximation to the maximal ruin probability with diatomic claims is described in Section 4. It is based on maximization of the most important term of the geometric distribution. Our limited numerical experience shows that this solution leads to a ruin probability which is invariably close to the maximal diatomic ruin probability. For small  $\lambda$ , this same diatomic distribution also often leads to near-maximum compound Poisson stop-loss premiums.

One of the referees remarked that applying this heuristic approach one actually solves Schmitter's problem optimally for very small values of the initial capital. More precisely, if the initial capital/the retention is very small (less than  $\frac{1}{2}E[X^2]/E[X]$ ), the maximum ruin probability/compound stop-loss premium is attained for the diatomic distribution with 0 as a mass point.

In any case it can be shown that this heuristic solution is better than many other choices of the feasible distribution. If  $x_1$  and  $x_2$  are the mass points of the heuristically found feasible distribution, with  $x_1 < x_2$ , any distribution with least mass point larger than  $x_1$  leads to lower ruin probabilities and compound Poisson stop-loss premiums.

In Section 5 we impose one more restriction on the claims distribution, namely that the support is contained in an interval  $[0, b]$ . One might expect that the distribution with the largest value of the upper bound for the ruin probability  $e^{-Ru}$  also has a high probability of ruin. It can be shown that the adjustment coefficient  $R$  with the claims distribution is minimal for the diatomic distribution with  $b$  as one of its mass points. Then obviously  $e^{-Ru}$  is maximal. But if the maximum claim  $b$  is very large, the ruin probability with



this distribution is close to minimal rather than maximal. On the other hand, the adjustment coefficient  $R$  is maximal for the diatomic distribution with 0 as a mass point, but for small values of  $u$  this distribution has maximal ruin probability, in spite of the fact that it has minimal  $e^{-Ru}$ . So looking at the adjustment coefficient leads to the wrong answer, unless  $b$  is small and  $u$  is large, say for  $b \leq 2u - \mu$ , see the previous paragraph and Section 4.

In Section 2 it is shown that the third moment (skewness) of the compound Poisson distribution is maximal for the diatomic claims distribution with  $b$  as a mass point. So one may expect that for large retentions, this claims distribution leads to maximal stop-loss premiums. Also in Section 5 we will show that for small retentions the situation is reversed.

2. SOME THEORY AND NOTATION

In both problems we study, the issue is to find a maximum of a functional  $H_u$ , working on distribution functions  $F_X$  of random variables  $X$  in a certain set. More specifically, we may write both problems in the following form:

- (1) Maximize  $H_u[F_X]$
- subject to  $X$  is a non-negative random variable, with  $E[X] = \mu$ ,  
 $\text{Var}[X] = \sigma^2$

Here  $H_u[\ ]$  assigns to  $F_X$  either the ruin probability  $\psi(u)$  in a compound Poisson risk process with fixed safety loading  $\theta$  and initial capital  $u$ , or the stop-loss premium  $\pi_S(u)$  at retention  $u$  of a compound Poisson ( $\lambda$ ) distributed random variable  $S$ , both with individual claims distributed as  $X$ . In the remainder of this section we will give expressions for  $H_u[\ ]$  for both problems in case  $X$  has a finite range. Also, we will characterize the feasible random variables  $X$  having a two-point support. Finally, the theory of ordering of risks is applied to derive results on some integrals over  $H_u[\ ]$ .

Consider the classical actuarial ruin model, that is, assume a compound Poisson process with claims intensity  $\lambda$ , non-negative individual claims distributed as  $X$ , premium income per unit time  $c = (1 + \theta)\lambda E[X]$ , which means there is a safety loading  $\theta$  (assumed positive), and initial capital  $u$ . See for instance BOWERS et al. (1986, Chapter 12). Let the stochastic process  $N(t)$  denote the number of claims up to time  $t$ , and  $S(t) = X_1 + \dots + X_{N(t)}$  the accumulated claims until time  $t$ . Define the maximal aggregate loss as  $L = \max\{S(t) - ct \mid t \geq 0\}$ . The ultimate ruin probability  $\psi(u)$  denotes the probability that the insurer's surplus will ever become negative:

$$(2) \quad \psi(u) = P[\min\{u + ct - S(t) \mid t \geq 0\} < 0] = 1 - P[L \leq u].$$

Defining  $L_1, L_2, \dots$  as the amounts by which record lows in the insurer's surplus  $u + ct - S(t)$  are broken, and  $M$  to be the number of record lows in the surplus process, we may write

$$(3) \quad L = \sum_{i=1}^M L_i.$$

Then  $M$  has a geometric distribution with parameter  $\psi(0)$ . From Theorem III.2.2.3 in GOOVAERTS et al. (1990) we see that the geometric parameter  $\psi(0) = (1 + \theta)^{-1}$ , and the distribution function of the  $L_i$  equals

$$(4) \quad F_{L_i}(y) = 1 - \pi_X(y) / \pi_X(0),$$

where  $\pi_X(y) = E[(X - y)_+]$  denotes the net stop-loss premium for  $X$  at retention  $y$ , so  $\pi_X(0) = E[X]$

From (2) and (3) we obtain the following expression for the ruin probability

$$(5) \quad \psi(u) = P[L > u] = \frac{\theta}{1 + \theta} \sum_{m=0}^{\infty} \left\{ \frac{1}{1 + \theta} \right\}^m P[L_1 + \dots + L_m > u].$$

SCHMITTER (1990) gives the following expression for the ruin probability in case  $X$  has finite support  $\{x_1, x_2, \dots, x_m\}$ , with associated probabilities  $p_1, p_2, \dots, p_m$ :

$$(6) \quad \psi(u) = 1 - \frac{\theta}{1 + \theta} \sum_{k_1, k_2, \dots, k_m} (-z)^{k_1 + \dots + k_m} e^z \prod_{j=1}^m \frac{p_j^{k_j}}{k_j!},$$

$$\text{where } z = \frac{\lambda}{c} (u - k_1 x_1 - \dots - k_m x_m)_+.$$

Similar expressions can be found in GERBER (1990), SHIU (1989), and earlier TAKÁCS (1967). The indices  $k_j$  are assumed to range over  $0, 1, \dots$ . If all mass points  $x_j$  are strictly positive,  $j = 1, \dots, m$ , (6) is a sum with only a finite number of non-zero terms, so it leads to an easily programmed algorithm to compute  $\psi(u)$  for discrete claims distributions. If one of the mass points, say  $x_m$ , is equal to 0, carrying out the (infinite) summation over  $k_m$  in (6) leads to the same expression as (6) with  $m$  replaced by  $m - 1$ ,  $\lambda$  by  $\lambda(1 - p_m)$ , and  $p_j$  by  $p_j / (1 - p_m)$ ,  $j = 1, \dots, m - 1$ .

In Section III.5 of GOOVAERTS et al. (1990) we find that the distributions with mean  $\mu$  and variance  $\sigma^2$  that are diatomic with support  $\{x_1, x_2\}$ , for  $x_1 = \mu - \varepsilon$ , can be characterized by

$$(7) \quad \begin{aligned} x_1 &= \mu - \varepsilon, & x_2 &= \mu + \sigma^2 / \varepsilon, \\ p_1 &= P[X = x_1] = \sigma^2 / \{\sigma^2 + \varepsilon^2\}, & p_2 &= P[X = x_2] = 1 - p_1 \end{aligned}$$

For  $0 \leq x_1 < x_2 < \infty$ , we must have  $0 \leq x_1 < \mu$ , so  $0 < \varepsilon \leq \mu$ . Note that  $x_2$  increases with  $x_1$  for  $x_1 \in [0, \mu)$ .

Inserting (7) in (6) with  $m = 2$ , we see that  $\psi(u)$  is continuous for diatomic distributions as a function of  $\varepsilon$  at  $\varepsilon \downarrow 0$ . So there is a sequence of feasible diatomic distributions, whose ruin probabilities converge to the one of the claims distribution with  $P[X = \mu] = 1$ , or  $\varepsilon = 0$

The compound Poisson stop-loss premium can be written in the form

$$(8) \quad \pi_S(u) = \sum_{n=0}^{\infty} \lambda^n e^{-\lambda} / n! E[(X_1 + \dots + X_n - u)_+].$$

If the range of the claims is finite, there is an expression for the compound stop-loss premiums similar to (6). If  $S$  has a compound Poisson ( $\lambda$ ) distribution with individual claims distribution as in (6), and  $N_j$  counts the number of occurrences of claim size  $x_j$ , such that  $S = x_1 \cdot N_1 + \dots + x_m \cdot N_m$ , then it is well-known that the  $N_j$  are independent Poisson ( $\lambda p_j$ ) distributed random variables. So the stop-loss premium of  $S$  at retention  $u$  can be written as:

$$(9) \quad \begin{aligned} \pi_S(u) &= E[(S - u)_+] = E[S] - u + E[(u - S)_+] \\ &= E[S] - u + \sum_{k_1, k_2, \dots, k_m} e^{-\lambda} (u - k_1 x_1 - \dots - k_m x_m)_+ \prod_{j=1}^m \frac{(\lambda p_j)^{k_j}}{k_j!}. \end{aligned}$$

It is evident that  $\pi_S(0) = \lambda \mu$ ,  $\pi_S(\infty) = 0$ ,  $\psi(0) = (1 + \theta)^{-1}$  and  $\psi(\infty) = 0$  do not depend on the actual choice of the feasible distribution. We will show that this holds for the integrals over  $\pi_S(u)$  and  $\psi(u)$  as well; the weighted integrals over  $u \pi_S(u)$  and  $u \psi(u)$ , however, are minimal/maximal when the third moment of the individual claims is.

We will use the following identities, valid for non-negative random variables  $Y$  with  $E[Y^{j+2}] < \infty$ , and which can be proved by partial integration:

$$(10) \quad \begin{aligned} \int_0^{\infty} y^j \pi_Y(y) dy &= \int_0^{\infty} \frac{1}{j+1} y^{j+1} [1 - F_Y(y)] dy; \\ \int_0^{\infty} y^j [1 - F_Y(y)] dy &= \frac{1}{j+1} E[Y^{j+1}], \quad j \geq 0. \end{aligned}$$

Using (10) and familiar properties of moments of compound distributions, we may deduce for every feasible distribution of the individual claims:

$$(11) \quad \begin{aligned} \int_0^{\infty} \pi_S(u) du &= \frac{1}{2} E[S^2] = \frac{1}{2} \{ \text{Var}[S] + (E[S])^2 \} = \frac{1}{2} \{ \lambda(\sigma^2 + \mu^2) + \lambda^2 \mu^2 \}; \\ \int_0^{\infty} \psi(u) du &= E[L] = E[M] E[L_1] = \frac{1}{\theta} \int_0^{\infty} [1 - F_{L_1}(u)] du \\ &= \frac{1}{\theta \mu} \int_0^{\infty} \pi_X(u) du = \frac{1}{\theta \mu} E[\frac{1}{2} X^2] = \frac{\sigma^2 + \mu^2}{2 \theta \mu}. \end{aligned}$$

The following relations for weighted integrals hold

$$\begin{aligned}
 (12) \quad \int_0^\infty u\pi_S(u) du &= \frac{1}{6} E[S^3] = \frac{1}{6} E[(S - E[S] + E[S])^3] \\
 &= \frac{1}{6} \{ \lambda E[X^3] + 3 \lambda^2 \mu (\mu^2 + \sigma^2) + \lambda^3 \mu^3 \}; \\
 \int_0^\infty u\psi(u) du &= \frac{1}{2} E[L^2] = \frac{1}{2} E[E[L^2|M]] \\
 &= \frac{1}{2} E[M \cdot E[L_1^2] + M(M-1) (E[L_1])^2] \\
 &= \frac{1}{2} E[M] E[L_1^2] + \frac{1}{2} E[M(M-1)] (E[L_1])^2 \\
 &= \frac{1}{2\theta} E[L_1^2] + \frac{1}{\theta^2} (E[L_1])^2 = \frac{1}{6\theta\mu} E[X^3] + \frac{(\sigma^2 + \mu^2)^2}{4\theta^2\mu^2}.
 \end{aligned}$$

So the fatter the tail of the individual claims  $X$  (measured by their skewness, or what is the same since  $\mu$  and  $\sigma^2$  are given, by their third moment), the larger the integral over  $u\psi(u)$  and  $u\pi_S(u)$ .

In the theory of ordering of risks as described in GOOVAERTS et al. (1990), one compares stop-loss transforms or distribution functions of risks over the whole interval  $[0, \infty)$ . In our case it is sufficient if these functions are ordered only on the interval  $[0, u]$ . Suppose that for instance  $X$  has lower stop-loss premiums than  $Y$  on the interval  $[0, u]$ . If  $Z$  is another independent risk, we have

$$\begin{aligned}
 (13) \quad E[(X+Z-u)_+] &= \int_0^\infty E[(X+Z-u)_+ | Z = z] dF_Z(z) \\
 &= \int_0^\infty E[(X-(u-z))_+] dF_Z(z) \\
 &\leq \int_0^\infty E[(Y-(u-z))_+] dF_Z(z) = E[(Y+Z-u)_+].
 \end{aligned}$$

From this property we see directly that if  $X_1, X_2, \dots$  and  $Y_1, Y_2, \dots$  are sequences of independent risks distributed as  $X$  and  $Y$  respectively, and  $X$  has lower stop-loss premiums than  $Y$  on  $[0, u]$ , then we have  $E[(X_1 + \dots + X_m - u)_+] \leq E[(Y_1 + \dots + Y_m - u)_+]$  for all  $m = 1, 2, \dots$ . Using (8), we see that a compound Poisson distribution with  $X$  as claims distribution has a lower stop-loss premium in  $u$  than one with  $Y$ . Using (4) and (5), we see that ruin probabilities are lower as well.

### 3. MAXIMIZING THE FUNCTIONALS NUMERICALLY

It is easy to maximize the ruin probability numerically over the *diatomic* feasible distributions. This can be accomplished using algorithm (6), together with (7) to characterize the feasible diatomic distributions. It involves merely a

one-dimensional maximization over the interval  $x_1 \in [0, \mu]$ . To do this, one first computes (6) at a number of values of  $x_1$  to detect the interval in which the maximum is to be found, and subsequently uses a method like golden section search to determine the maximum more exactly. A reference for numerical techniques to compute a maximum of a function over an interval is PRESS et al. (1986). In Figure 1 we give graphs depicting the diatomic ruin probability  $\psi(u, x_1, x_2, p_1, p_2) = \psi(u, x_1)$  as a function of  $x_1 \in [0, \mu]$ , where  $x_1, x_2, p_1, p_2$  are related by (7). We took  $\mu = 3$ ,  $\sigma^2 = 1$ ,  $\theta = 0.5$ , and  $u = 1.5, 4.5$  and  $9$  respectively. In these graphs, the scale in the  $y$ -direction varies.

As announced, the ruin probability is minimal and continuous at  $x_1 \uparrow \mu$ . In Figure 1 we see that for small  $u$  ( $u = 1\frac{1}{2}$ ) the maximum ruin probability is found taking  $x_1 = 0$ . A close inspection reveals that the ruin probability does not depend on  $x_1$  if  $x_1 \geq u$ . Indeed in (6) one sees that the ruin probability does not (directly) depend on mass points larger than  $u$ . It also follows from (4) and (5). For large  $u$  ( $u = 9$ ),  $\psi(u)$  is very nearly constant for small to moderate values of  $x_1$ , then increases, and next decreases steeply to its minimal value at  $x_1 \uparrow \mu$ .

For intermediate  $u$  ( $u = 4.5$ ), the situation is rather unclear: there are some local maxima. For this specific situation we were able to find a three-point distribution with a larger ruin probability than the one corresponding to the maximizing diatomic distribution. In fact, for

$$\begin{aligned} x_1 &= 1.56592, & x_2 &= 2.67226, & x_3 &= 5.182086, \\ p_1 &= 0.071198, & p_2 &= 0.766835, & p_3 &= 0.161967 \end{aligned}$$

the ruin probability is 0.279271, which, although (probably) not the optimal solution, is higher than the maximal diatomic ruin probability 0.279185, found at  $x_1 = 2.5597$ ,  $x_2 = 5.2712$ .

Although we tried a lot of combinations of  $\mu, \sigma^2, \theta$  and  $u$ , we rarely found a randomly generated three-point distribution better than the best diatomic distribution; if we did, the difference was never substantial.

We did not try to optimize systematically over all three-point spectra. First, this is not a trivial task: if the number of mass points is  $m$ , the number of free variables equals  $2m - 3$ , being the number of support points  $x_j$  plus the number of probabilities  $p_j$ , minus the number of restrictions. So to find the maximal ruin probability over all three-point spectra involves solving a three-dimensional maximization, with borderline conditions  $p_j \geq 0$ . Second, even supposing we successfully optimized over three-point distributions, there is still no guarantee that for instance a 15-point support might not be better.

The fact that for small  $u$  the ruin probability is maximal at  $x_1 = 0$  can be explained as follows. By relation (11), one sees that neither  $\psi(0)$  and  $\psi(\infty)$ , nor  $\int \psi(u) du$  depend on  $x_1$ . By (12), however, we see that the weighted integral increases (linearly) with the third moment of the claims distribution. So the weighted integral is minimal for the diatomic distribution with  $x_1 = 0$ , which means that taking  $x_1 = 0$  gives the smallest integral over  $u\psi(u)$ . So at small values of  $u$ ,  $\psi(u)$  should be large for  $x_1 = 0$ . By similar reasoning, one

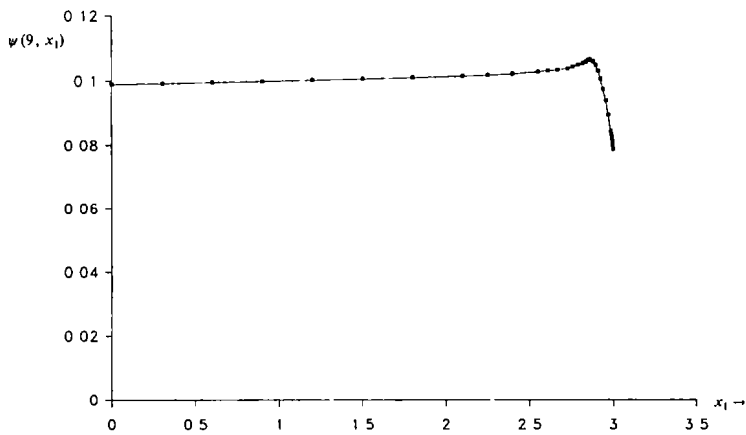
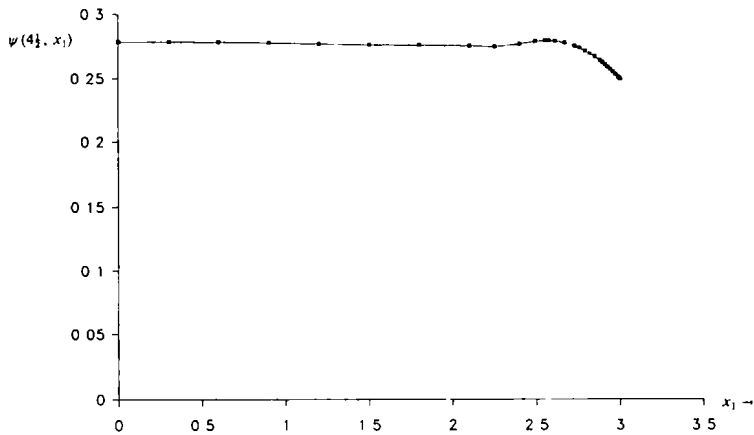
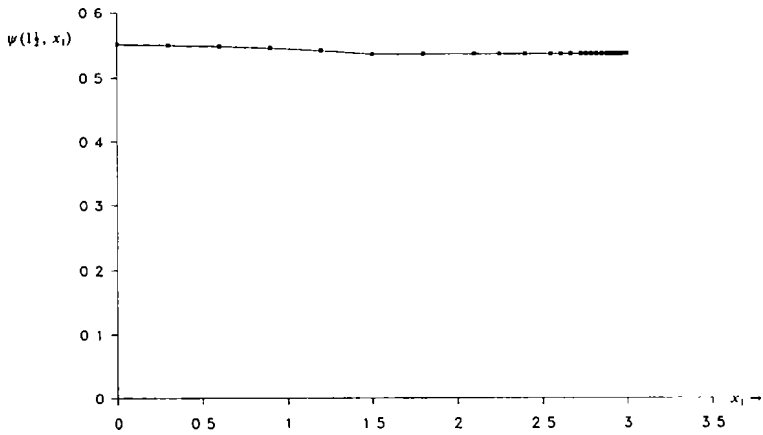


FIGURE 1  $\psi(u, x_1)$  as a function of  $x_1$ ,  $\mu = 3$ ,  $\sigma^2 = 1$ ,  $\theta = \frac{1}{2}$ ,  $u = 1\frac{1}{2}, 4\frac{1}{2}, 9$

explains that for large  $u$ , a large value of  $x_1$  leads to maximum  $\psi(u)$ . For too large values of  $x_1$ , we obtain low ruin probabilities (close to the minimal value), as explained in the following section.

For the same reasons, one can expect a similar pattern to arise in the case of compound Poisson stop-loss premiums. This is indeed the case: see Figure 2. In this figure, we took  $\lambda = 2$ ,  $\mu = 3$  and  $\sigma^2 = 1$ . At small  $u$  ( $u = 2$ ), the stop-loss premium is virtually constant over  $x_1$ , but it is maximal at  $x_1 = 0$ . At large  $u = 20$ , we see that the stop-loss premium is practically constant for  $x_1$  from 0 (where it equals 0.0109) to very close to  $\mu$ . Then it increases very steeply to its maximum value 0.0522, and for  $x_1 \uparrow \mu$ , it decreases continuously to its minimal value of 0.0088. For intermediate  $u = 7$ , with increasing  $x_1$ ,  $\pi_S(u)$  increases slightly and irregularly at first from 1.3373 to the maximal value 1.3954, and then for  $x_1 \uparrow \mu$ , it decreases again to its infimum 1.3008. For this case we found again an example where the maximal diatomic distribution was not a global maximum over all feasible claims distributions. The maximal diatomic distribution is at  $x_1 = 2\frac{2}{3}$ , where  $\pi_S(u) = 1.3954$ , but a larger stop-loss premium of 1.3995 is attained by the triatomic distribution

$$x_1 = 0, x_2 = 2.8, x_3 = 5.7143, p_1 = 0.0286, p_2 = 0.8754, p_3 = 0.0961.$$

In fact, as one of the referees pointed out, it can be proven that the diatomic distribution with  $x_1 = 0$  as a mass point is optimal for very small values of  $u$  ( $u \leq \frac{1}{2} E[X^2]/E[X]$ ). The proof goes as follows.

From Theorem III.5.2.3 of GOOVAERTS et al. (1990) we see that *uniformly* for all  $u \leq \frac{1}{2} E[X^2]/E[X] = \frac{1}{2}(\mu + \sigma^2/\mu)$ , the maximal stop-loss premium over the feasible distributions is attained for a random variable  $X_0$  having mass points 0 and  $\mu + \sigma^2/\mu$ , see (7). As a consequence of (13), we have immediately that if  $H$  is the distribution function of  $X_0$  and  $X$  is a feasible claim size, then  $F_X^{*n}$  has smaller stop-loss premium in  $u$  than  $H^{*n}$  for  $n = 2, 3, \dots$ , too. In view of (8), we have then found that  $H$  is the claims distribution maximizing the compound Poisson stop-loss premium, when the retention  $u \leq \frac{1}{2}(\mu + \sigma^2/\mu)$ .

Using (4), we can deduce by similar reasoning that this same claims distribution also maximizes not only  $P[L_1 > u]$  for  $u \leq \frac{1}{2}(\mu + \sigma^2/\mu)$ , but also  $P[L_1 + \dots + L_m > u]$  for all  $m = 2, 3, \dots$ , and thus maximizes the ruin probability (5).

So Schmitter's problem is solved for very small values of the initial capital  $u$ .

This result is confirmed in Figure 1 for  $u = 1\frac{1}{2}$ . But note that in Figure 2 for  $u = 2 > \frac{1}{2}(\mu + \sigma^2/\mu)$  still the distribution having mass point 0 led to the maximal compound Poisson stop-loss premium.

#### 4. AN APPROXIMATION FOR THE MAXIMIZING DIATOMIC DISTRIBUTION

Though we are as yet unable to solve the problem of maximizing  $\psi(u) = P[L > u]$  given  $\mu$  and  $\sigma^2$ , a problem we can solve is the maximization of  $P[L_1 > u]$ . We may expect  $P[L > u]$  to be large when  $P[L_1 > u]$  is, because the term with  $m = 1$  in (5) has the largest weight factor.

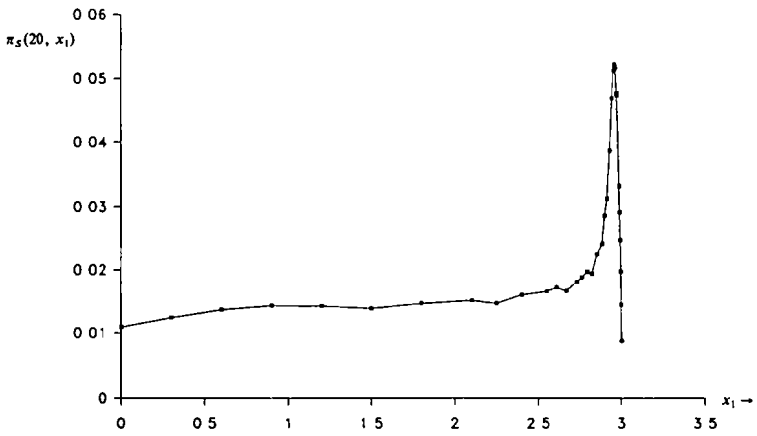
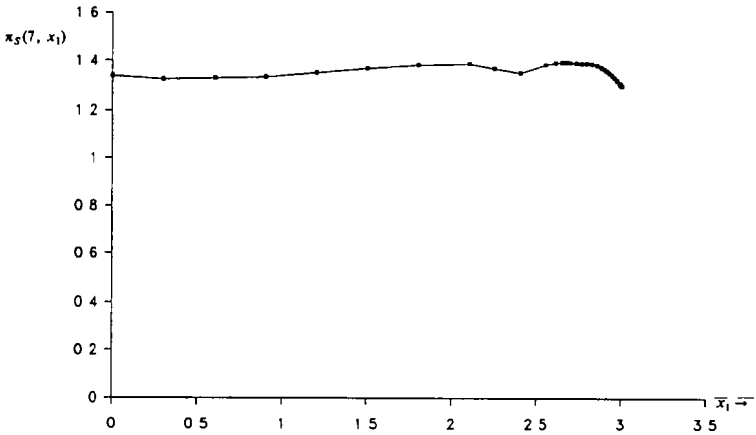
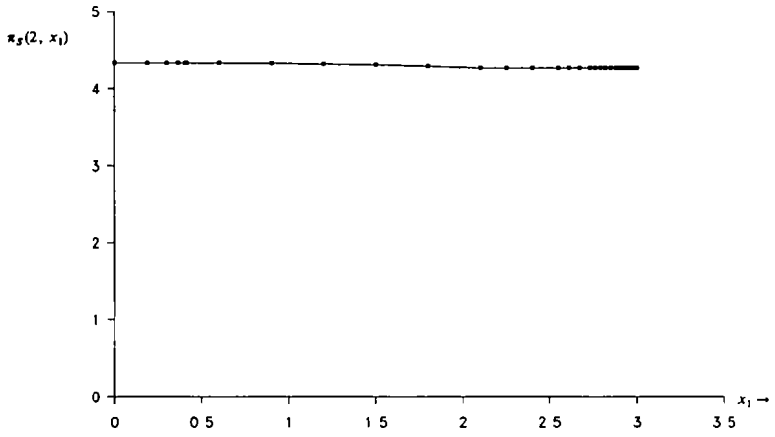


FIGURE 2  $\pi_S(u, x_1)$  as a function of  $x_1$ ,  $\mu = 3$ ,  $\sigma^2 = 1$ ,  $\theta = \frac{1}{2}$ ,  $u = 2, 7, 20$



In view of (4), and since  $\pi_X(0) = E[X] = \mu$  is given, to maximize  $P[L_1 > u]$  we just have to maximize  $\pi_X(u)$ , the stop-loss premium of  $X$ . The solution to this problem can for instance be found in GOOVAERTS et al. (1990), Theorems III.5.2.2 and 5.2.3. These theorems express that the maximal stop-loss premium for a (non-negative) risk  $X$  with mean  $\mu$  and variance  $\sigma^2$  at retention  $u$  is the diatomic distribution with smaller mass point  $x_1 = \max\{u-d, 0\}$ , where  $d = \{(\mu-u)^2 + \sigma^2\}^{1/2}$ . When  $\sigma$  is small with respect to  $|u-\mu|$ , we may write

$$(14) \quad (u-\mu)-d = (u-\mu-d) \frac{u-\mu+d}{u-\mu+d} = \frac{-\sigma^2}{u-\mu+d} \approx -\frac{1}{2}\sigma^2/(u-\mu).$$

So we may conclude that the diatomic distribution with the following mass points gives a ‘high’ ruin probability:

$$(15) \quad x_1 = \mu - \varepsilon, \text{ with } \varepsilon = \frac{\sigma^2}{u-\mu+d} \approx \frac{1}{2}\sigma^2/(u-\mu), \text{ so } x_2 = u+d \approx 2u-\mu$$

In the examples we tested, the diatomic distribution maximizing the ruin probability had  $x_1$  only slightly smaller than  $u-d$ . See Table 1.

Of course this same diatomic distribution maximizes the term with  $n = 1$  of the compound Poisson stop-loss premium (8). So one may expect this distribution to have a high stop-loss premium if the probability of just one claim is large, which is the case if  $\lambda$  is small. For large  $\lambda$ , however, this approximation will not be as useful.

Our heuristic procedure may not always lead to the optimal value, but it can be shown that it is better than many other choices. Suppose  $Z$  has distribution (15), and suppose  $Y$  is another feasible choice such that the least mass point of  $Y$  is larger than that of  $Z$ , which is  $u-d$ . We know that  $\pi_Z(t)$  is piecewise linear, with edges at  $u-d$  and  $u+d$ . Since  $Y$  has no mass below  $u-d$ , we have  $\pi_Y(u-d) = \pi_Z(u-d)$ . Also,  $\pi_Y(u) \leq \pi_Z(u)$  since  $\pi_Z(u)$  is maximal. So  $\pi_Y(t) \leq \pi_Z(t)$  for all  $t \leq u$ , which means that  $Y$  generates lower compound Poisson stop-loss premiums and ruin probabilities.

TABLE 1

VALUES OF  $\psi(u)$  FOR DIFFERENT VALUES OF THE HIGHER MASS POINT IN A DIATOMIC DISTRIBUTION

	$\mu = 1, \sigma^2 = 1, \theta = 1$			$\mu = 3, \sigma^2 = 1, \theta = 5$		
	$u = 1.5$	$u = 4.5$	$u = 9$	$u = 1.5$	$u = 4.5$	$u = 9$
$x_2 =$	$\psi(u) =$	$\psi(u) =$	$\psi(u) =$	$\psi(u) =$	$\psi(u) =$	$\psi(u) =$
$\infty$	102003	002315	000008	534796	248974	078779
$\mu + \sigma^2/\mu$	272504	039292	002315	550047	278350	098945
optimal	275023	081105	034151	550047	279190	106205
$u+d$	269824	078214	033632	534796	276506	101811
$2u-\mu$	272504	.078651	033659	550047	277596	101901
10	146348	071460	024767	534796	265714	106184
15	130637	055095	034151	534796	259498	101901
20	123125	044244	031936	534796	256613	097203

In particular, the diatomic solutions with support  $\{b, b'\}$  with  $b > u - d$  are apparently non-optimal.

##### 5. EXTREMAL VALUES OF THE ADJUSTMENT COEFFICIENT

Consider all claims distributions with mean  $\mu$ , variance  $\sigma^2$  and as an extra requirement, support contained in  $[0, b]$  for some  $b \geq \mu + \sigma^2/\mu$ . Just as we did in the previous section for  $P[L_1 > u]$ , one may tackle the problem of finding extremal ruin probabilities by using distributions leading to extremal values of related quantities like an approximation or an upper bound for the ruin probability. Here we use the upper bound  $e^{-Ru}$ , where the adjustment coefficient  $R$  is the positive solution to the equation

$$(16) \quad 1 + (1 + \theta)\mu r = E[e^{rX}].$$

Asymptotically, this upper bound can be used as an approximation, since  $\psi(u) e^{Ru}$  has a limit in  $(0, 1)$  for  $u \rightarrow \infty$ .

It can easily be shown that the diatomic distribution with mass points 0 and  $\mu + \sigma^2/\mu$  is minimal in second degree stop-loss order, while the one with mass points  $b$  and  $\mu - \sigma^2/(b - \mu)$  is maximal. See Theorem II 4.2.3 of GOOVAERTS et al. (1990). This implies that these special diatomic distributions have minimal and maximal moment generating functions on  $(0, \infty)$  in the class considered, and accordingly the corresponding adjustment coefficients (roots of (16)) are maximal and minimal respectively.

One would expect that the support  $\{\mu - \sigma^2/(b - \mu), b\}$ , with minimal adjustment coefficient, leads to large ruin probability, too. Taking  $b$  too large, however, so  $\mu - \sigma^2/(b - \mu)$  is very close to  $\mu$ , results in the opposite of what we wanted: the ruin probability of this distribution is very small rather than maximal. For  $b \rightarrow \infty$ , by (7) we see that the mgf  $E[e^{rX}] \rightarrow \infty$  for all  $r > 0$ , so then  $R \rightarrow 0$ , which gives us the trivial upper bound  $\psi(u) \leq 1$ . So we observe that for  $b \rightarrow \infty$ , the upper bound  $e^{-Ru}$  increases, while the ruin probability decreases. But if  $b$  is not too large, say such that  $\mu - \sigma^2/(b - \mu) \approx x_1$  as in (15), which means that  $b \approx 2u - \mu$ , this distribution does lead to a large ruin probability.

On the other hand we learn for instance from Figure 1 that for small  $u$ , the diatomic distribution with mass point  $x_1 = 0$  has maximal ruin probability, even though it gives the tightest upper bound  $e^{-Ru}$ .

It can be shown, too, that the compound Poisson distributions with these distributions for the individual claims are extremal in second degree stop-loss order. This means that they have minimal and maximal third moment, and since mean and standard deviation are fixed, also minimal and maximal coefficient of skewness. As proved at the end of Section 2, these same special spectra also generate the extreme values of  $\int u \pi_S(u) du$ . So one would be inclined to expect that they lead to high and low values of the compound Poisson stop-loss premium as well, but the same caveats as above apply here.

6. SOME FINAL REMARKS

To conclude, we comment on tables of some results for distributions with support  $\{\mu - \sigma^2/(b - \mu), b\}$  for different values of  $b$ . These distributions have minimal adjustment coefficient (maximal skewness) for all feasible distributions with support contained in  $[0, b]$ . They are compared to other distributions described above: the optimal diatomic distribution, the heuristical approximations to the optimum found by applying (15) and the distributions with only one positive mass point: support  $\{0, \mu + \sigma^2/\mu\}$  and  $\{\mu\}$ . The latter support is denoted by higher mass point  $\infty$ , where the mass on  $\infty$  is of course 0 (but contributes to  $\sigma^2$ ). Note that for  $u$  not too large and  $b = 20$ , the phenomenon described above indeed occurs. Even though we showed that looking at the minimal adjustment coefficient sometimes gives incorrect results, especially for large  $b$  or small  $u$ , we fear that this method will be used quite often.

Further note that for large  $u$  and  $\sigma^2$ , minimal and maximal ruin probability are widely apart. For  $\sigma^2$  small with respect to  $u$  and  $\mu$ , the ruin probability cannot vary enormously.

Table 2 gives some results for the compound Poisson stop-loss premiums. Note the meaningless results obtained by the wrong choice of  $b$  for large values of  $u$ , and also for small values of  $u$ .

An approach that we plan to follow in the near future is to try to optimize the compound Poisson stop-loss premium over the set of claim distributions with support  $\{0, \delta, 2\delta, \dots, n\delta\}$ . The more general problem is obtained taking limits for  $n \rightarrow \infty$  and  $\delta \downarrow 0$ . The restricted problem can be written in the form of the maximization of a non-linear criterion function with three linear constraints on the probabilities  $p_j = P[X = j\delta]$ , required to be non-negative

TABLE 2  
VALUES OF  $\pi_S(u)$  FOR DIFFERENT VALUES OF THE HIGHER MASS POINT  
IN A DIATOMIC DISTRIBUTION

$x_2 =$	$\mu = 3, \sigma^2 = 1, \lambda = 2$			$\mu = 3, \sigma^2 = 1, \lambda = 5$		
	$u = 2$ $\pi_S(u) =$	$u = 7$ $\pi_S(u) =$	$u = 20$ $\pi_S(u) =$	$u = 5$ $\pi_S(u) =$	$u = 20$ $\pi_S(u) =$	$u = 40$ $\pi_S(u) =$
$\infty$	4 270671	1 300816	0 008804	10 101076	1 004413	0 002488
$\mu + \sigma^2/\mu$	4 330598	1 337326	0 010879	10 138862	1 077055	0 003859
optimal	4 332192	1 395435	0 052178	10 138862	1 136463	0 058680
$u + d$	4 331675	1 374006	0 047330	10 105046	1 077758	0 049633
$2u - \mu$	4 324805	1 374694	0 047347	10 105033	1 077807	0 049638
5	4 270671	1 376488	0 014677	10 101069	1 105061	0 005110
10	4 270671	1 380493	0 022903	10 104438	1 113764	0 007883
15	4 270671	1 356405	0 034962	10 103393	1 124541	0 012330
20	4 270671	1 342594	0 047335	10 102812	1 116290	0 018726
25	4 270671	1 334135	0 052137	10 102458	1 103217	0 028545
30	4 270671	1 328482	0 051061	10 102223	1 091199	0 040868

for all  $j$ . By restricting to an arithmetic spectrum we are able to use Panjer's recursion instead of (9), the necessary partial derivatives can also be computed by a recursive scheme. The procedure can be generalized if more moments are known.

Of course, as the title of our paper indicates, maximization over the diatomic distributions only does not give a complete solution of either problem. We find, however, that by using this technique both problems are sufficiently solved for practical purposes. In the first place, our examples led us to the conviction that, although the optimal diatomic distribution is not always globally optimal, it is not much removed from this optimum. Second, in our opinion in practice one might judge the attractiveness of risks or risk processes with known mean and variance of the claims by the worst feasible diatomic distribution as well as by the overall worst feasible distribution.

#### ACKNOWLEDGMENT

The author wishes to thank ANGELA VAN HEERWAARDEN for some constructive suggestions, and MARC GOOVAERTS for stimulating discussions. Also, the valuable contributions of both referees are acknowledged.

#### REFERENCES

- BOWERS, N L, GERBER, H U, HICKMAN, J C, JONES, D A and NESBITT, C J (1986) *Actuarial Mathematics* Society of Actuaries, Itasca, Illinois
- BROCKETT, P L, GOOVAERTS, M J and TAYLOR, G C (1991) The Schmitter problem *ASTIN Bulletin* **21**, 129–132
- GERBER, H U (1990) From the convolution of uniform distributions to the probability of ruin *Mitteilungen der VSVM Heft 2/1989*, 283–292
- GOOVAERTS, M J, DE VYLDER, F and HAEZENDONCK, J (1984) *Insurance premiums* North-Holland, Amsterdam
- GOOVAERTS, M J, KAAS, R, VAN HEERWAARDEN, A E and BAUWELINCKX, T (1990) *Effective Actuarial Methods* North-Holland, Amsterdam
- KAAS, R and GOOVAERTS, M J (1986) Bounds on stop-loss premiums for compound distributions *ASTIN Bulletin* **XVI**, 1, 13–17
- PRESS, W H, FLANNERY, B P, TEUKOLSKY, S A and VETTERLING, W A (1986) *Numerical recipes, the art of scientific computing* Cambridge University Press
- SCHMITTER, H (1990) *The ruin probability of a discrete claims distribution with a finite number of steps* Paper presented at the XXII ASTIN-colloquium, Montreux, Switzerland
- SHIU, E S W (1989) Ruin probability by operational calculus *Insurance Mathematics and Economics* **8**, 243–249
- STEENACKERS, A and GOOVAERTS, M J (1990) *Bounds on stop-loss premiums and ruin probabilities for given values of the mean, variance and maximal value of the claim size* Paper presented at the XXII ASTIN-colloquium, Montreux, Switzerland
- TAKÁCS, L (1967) *Combinatorial Methods in the Theory of Stochastic Processes* Wiley, New York. Reprinted by Krieger, Huntington, NY (1977)

ROB KAAS

*Institute for Actuarial Science and Econometrics, University of Amsterdam,  
Jodenbreestraat 23, NL-1011 NH Amsterdam, the Netherlands*

## SHORT CONTRIBUTIONS

### A NOTE ON THE NORMAL POWER APPROXIMATION

BY COLIN M RAMSAY

*Actuarial Science*

*University of Nebraska – Lincoln, USA*

#### ABSTRACT

The normal power (NP) approximation essentially approximates the random variable  $X$  as the quadratic polynomial  $\tilde{X} \approx Y + \gamma(Y^2 - 1)/6$  where  $\tilde{X} = (X - \mu)/\sigma$  is the standardized variable,  $Y \sim N(0, 1)$ , and  $\mu, \sigma, \gamma$  are the mean, variance skewness of  $X$  respectively. The coefficients of this polynomial are not determined by equating the lower moments. It is shown that matching these moments does not improve the overall accuracy of the approximation.

#### 1. INTRODUCTION

Let  $X$  be the aggregate claims in one year,  $Z_k$  be the size of the  $k^{\text{th}}$  claim and  $N$  be the total number of claims, i.e.,

$$X = \sum_{k=1}^N Z_k,$$

with  $X = 0$  if  $N = 0$ . Let  $F(x)$  be the cumulative distribution function (cdf) of  $X$ . It is well known that  $F(x)$  is given by

$$F(x) = \sum_{k=0}^{\infty} p_k G^{**k}(x), \quad x \geq 0$$

where  $G(x)$  is the cdf of  $Z_k$ ,  $G^{**k}(x)$  is the  $k^{\text{th}}$  convolution of  $G$  with itself,  $G^{*0}(x) = 1$  for  $x > 0$ , and  $p_k = P[N = k]$

Direct evaluation of  $F(x)$  is possible only in very special cases, so approximations are needed. A simple and easy approximation to  $F(x)$ , is the normal power (NP) approximation. The essential idea of the NP approximation is to transform the standardized original variable  $\tilde{X} = (X - \mu)/\sigma$ , where  $\mu = E[X]$  and  $\sigma^2 = \text{Var}[X]$ , into a symmetric variable  $Y = v(\tilde{X})$ . In particular  $v$  is chosen so that  $Y$  is a standard normal variable or is nearly so. By inverting the Edgeworth expansion of the unknown cdf of  $\tilde{X}$  and using Newton's method (see BEARD et al. (1984), pp. 108–111), it can be proved that

$$(1) \quad \tilde{X} \approx Y + \frac{\gamma}{6}(Y^2 - 1)$$

where  $Y \sim N(0, 1)$  and  $\gamma$  is the skewness of  $X$ . This results in the NP approximation

$$(2) \quad F(x) \approx N \left( \frac{-3}{\gamma} + \sqrt{\frac{9}{\gamma^2} + 1 + \frac{6\tilde{x}}{\gamma}} \right),$$

where  $\tilde{x} = (x - \mu)/\sigma$ . This approximation is valid for  $\tilde{x} > 1$ , and is fairly accurate if  $0 \leq \gamma \leq 1$ , with the accuracy decreasing as  $\gamma$  increases.

## 2. THE MAIN RESULT

Since the inverse transform  $\nu^{-1}(Y)$  approximates  $\tilde{X}$ , one would expect the left hand side (LHS) and the right hand side (RHS) of equation (1) to have approximately equal moments. However this is not the case because

$$E \left[ \left( Y + \frac{\gamma}{6}(Y^2 - 1) \right)^k \right] = \begin{cases} 0 & \text{if } k = 1 \\ 1 + \gamma^2/18 & \text{if } k = 2 \\ \gamma + \gamma^3/27 & \text{if } k = 3 \end{cases}$$

while

$$E[\tilde{X}^k] = \begin{cases} 0 & \text{if } k = 1 \\ 1 & \text{if } k = 2 \\ \gamma & \text{if } k = 3. \end{cases}$$

If  $\gamma$  is small, the terms  $\gamma^2/18$  and  $\gamma^3/27$  can be neglected, giving an approximate equality between the first 3 moments of the LHS and RHS. On the other hand if  $\gamma$  is large, the variance and skewness of the RHS of equation (1) will be inflated, possibly leading to poorer approximations.

The important question at this point is this: can the accuracy of the NP approximation be improved by equating the first three moments of the LHS and RHS of equation (1)? To this end, consider the quadratic

$$(3) \quad \tilde{X} = aY + b(Y^2 - 1)$$

where  $a$  and  $b$  are real constants and, once again,  $Y \sim N(0, 1)$ . Matching the first three moments yield the following equations

$$\begin{aligned} 1 &= a^2 + 2b^2 \\ \gamma &= 6a^2b + 8b^3 \end{aligned}$$

These equations reduce to

$$(4) \quad \begin{aligned} a &= \sqrt{1 - 2b^2} \quad \text{for} \quad [-1/\sqrt{2} \leq b \leq 1/\sqrt{2}] \\ \gamma &= 6b - 4b^3. \end{aligned}$$

TABLE I

$\gamma = 0.08$			
$\bar{x}$	$F(\bar{x})$	NP	Adj-NP
1.0	0.1586	0.1587	0.1586
2.0	0.0249	0.0249	0.0249
3.0	0.0019	0.0019	0.0019
4.0	0.0001	0.0001	0.0001
$\gamma = 0.238$			
$\bar{x}$	$F(\bar{x})$	NP	Adj-NP
1.0	0.1579	0.1587	0.1583
2.0	0.0288	0.0289	0.0286
3.0	0.0031	0.0031	0.0030
4.0	0.0002	0.0002	0.0002
$\gamma = 0.593$			
$\bar{x}$	$F(\bar{x})$	NP	Adj-NP
1.0	0.1529	0.1587	0.1566
2.0	0.0362	0.0372	0.0362
3.0	0.0068	0.0064	0.0062
4.0	0.0011	0.0009	0.0008
$\gamma = 0.779$			
$\bar{x}$	$F(\bar{x})$	NP	Adj-NP
1.0	0.1526	0.1587	0.1553
2.0	0.0394	0.0411	0.0395
3.0	0.0084	0.0084	0.0080
4.0	0.0011	0.0014	0.0014
$\gamma = 1.082$			
$\bar{x}$	$F(\bar{x})$	NP	Adj-NP
1.0	0.1376	0.1587	0.1525
2.0	0.0376	0.0470	0.0440
3.0	0.0125	0.0119	0.0109
4.0	0.0042	0.0027	0.0024
$\gamma = 1.628$			
$\bar{x}$	$F(\bar{x})$	NP	Adj-NP
1.0	0.1280	0.1587	0.1452
2.0	0.0370	0.0562	0.0496
3.0	0.0140	0.0184	0.0159
4.0	0.0069	0.0057	0.0048

It is clear that for  $-2\sqrt{2} \leq \gamma \leq 2\sqrt{2}$ , equation (4) has exactly one root in the region  $-1/\sqrt{2} \leq b \leq 1/\sqrt{2}$ . Since the distribution of insurance claims are usually positively skewed, only the case where  $0 \leq \gamma \leq 2\sqrt{2}$  is considered.

For  $0 \leq \gamma \leq 2\sqrt{2}$ , let  $b_0$  be the unique root of equation (4) which lies in the region  $0 \leq b_0 \leq 1/\sqrt{2}$ , and let

$$(5) \quad a_0 = \sqrt{1 - 2b_0^2}.$$

Substituting the values into equation (4), the following approximation results

$$(6) \quad F(x) \approx N \left( \frac{-a_0}{2b_0} + \sqrt{\frac{a_0^2}{4b_0^2} + 1 + \frac{\tilde{x}}{b_0}} \right).$$

This approximation will be called the “adjusted” NP approximation

Table 1 shows the values produced by the traditional NP approximation (equation (2)) and by the adjusted NP approximation (equation (6)). The values of  $F$  and NP are taken from PENTIKAINEN (1987, pp. 32–34, cases 1, 3, 5, 6, 7, 8). Following PENTIKAINEN,  $F$  is actually  $1 - F$  (the right tail probability) for  $\tilde{x} > 0$ . From this table, it is clear that both NP approximations yield similar values. As a result, equation (2) must be viewed as being superior because it is easier to use, i.e., it requires fewer steps to derive this approximation

Finally, it should be noted that these approximations have not been properly calculated;  $F(x)$  should be approximated as follows:

$$(7) \quad \begin{aligned} F(x) &= P[\tilde{X} \leq \tilde{x}] \\ &\approx P[aY + b(Y^2 - 1) \leq \tilde{x}] \\ &= P[r_1 \leq Y \leq r_2] \quad \text{because } b > 0 \\ &= N(r_2) - N(r_1) \end{aligned}$$

where  $r_1 \leq r_2$  are the roots of the equation

$$(8) \quad \tilde{x} = ay + b(y^2 - 1),$$

with  $a = 1$ ,  $b = \gamma/6$  for the traditional NP approximation, and  $a = a_0$ ,  $b = b_0$  for the adjusted NP approximation. The approximation (7) will serve to increase the estimates of the right tail probabilities  $1 - F(x)$ . However, over the range of applicability of the NP approximations, i.e., for  $\tilde{x} > 1$  and  $0 \leq \gamma \leq 1$ , the extra term  $N(r_1)$  is insignificant.

#### REFERENCES

- BEARD, R E, PENTIKAINEN, T and PESONEN, E (1984) *Risk Theory* Chapman and Hall, London  
 PENTIKAINEN, T (1987) Approximate evaluation of the distribution of aggregate claims *ASTIN Bulletin* 17, 15–39

COLIN M RAMSAY

*Actuarial Science, 310 Burnett Hall, University of Nebraska – Lincoln,  
 Lincoln, NE 68588-0307, USA*



## BOOK REVIEW

M.J. GOOVAERTS, R. KAAS, A.E. VAN HEERWAARDEN, T. BAUWELINCKX (1990): *Effective Actuarial Methods*. Elsevier Science Publishers BV, Amsterdam, 316 pages, US\$ 92.25/DFL. 180.00

The 'Effective actuarial methods' comprise three separate essays on *Ordering of Risks* (Part 1), *Credibility Theory* (Part 2) and *IBNR Techniques* (Part 3). Via these topics the authors present material from actuarial science which is interesting, both from a mathematical and an applications point of view. The latter is highlighted by analyses based on real portfolio data using the software packages SLIC (stop-loss reinsurance), CRAC (credibility) and LORE (IBNR modelling).

In PART 1 a review of various orderings of risks, together with a discussion of the related algebraic properties, are given. Having these tools available, it is relatively easy to tackle specific problems in the collective risk model. These mainly are estimation and ordering of adjustment coefficients and ruin probabilities, but also results on optimal reinsurance are obtained. In many cases do these 'order' results allow for easier numerical calculations. After a rather trivial excursion into the realm of survival distributions, this first part closes with a discussion on incomplete information, i.e. situations where only moment conditions and/or shape information (like unimodality) of the relevant random variables are/is assumed. Think for instance of the construction of stop-loss premiums with  $n$  moments known.

PART 2 on Credibility Theory starts with a very readable introduction on 'what is credibility all about' before giving an overview of the various models and their analysis. The models included are those by BÜHLMANN, BÜHLMANN-STRAUB, the hierarchical one and regression type models. The material is presented in a well-documented, self-contained way which gives the reader a thorough insight into the basic theory. Proofs are given explicitly. Some interesting extensions of the 'classical theory' are given in Chapter VI. These comprise credibility formulae of the updating type together with results on covariance structures leading to such formulae. Furthermore, in a section on credibility for loaded premiums, it is shown how credibility estimators can be based on weighted loss functions; examples are Esscher and variance premiums. After some brief comments on multidimensional credibility, the authors spend some more time on semi-linear credibility where linear functions of *transformed* variables are considered as estimators. An interesting chapter on insurance applications of credibility theory, based on the CRAC-software package for two level, semi-linear hierarchical credibility ends this section of the book.

The final PART 3 contains an introduction of IBNR-techniques. These involve mechanical smoothing (where no underlying model is assumed), statistical methods (mostly of (auto)regression type) and credibility based methods (including Kalman filtering). Via the loss reserving software package LORE, the versatility of the methods presented is demonstrated on real data coming from:

- recuperation in credit insurance;
- loss-reserving for liability insurance for notaries;
- loss-reserving in automobile liability insurance, and
- ‘activity coefficients’ in a pension fund of physicians.

The overall material is well-balanced between the three parts with exercises adding to the course-book status. It is clear that having the software would add to the understanding of some of the material presented though this is by no means a necessity. One of the main attractions with respect to teaching lies in the fact that based on this one book, actuarial students will gain considerable insight into some of the specific techniques which are by now well-established as core material within modern actuarial science. I am convinced that many actuarial students, and indeed many researchers in the field, will find this text a very useful one to have on one’s bookshelf.

PAUL EMBRECHTS

## LETTER TO THE EDITORS

Dear Sir,

I am writing in connection with the Claims Reserving Manual, Volume 2, published by the Institute of Actuaries.

Volume 1 of this manual deals with arithmetic or deterministic methods, while Volume 2 covers more advanced methods involving probabilistic and statistical methods.

We are now considering further contributions to Volume 2 and would be very happy to receive articles written by members of ASTIN. The test for the inclusion of a method is that it has been found useful by a practitioner. The fact that a method may contain weaknesses from a theoretical point of view may be commented upon, but will not prevent its publication. Methods which have already been written up in journals are still eligible for inclusion in the manual, although the write-up should have a practical bias.

Contributions should be sent to

S. BENJAMIN esq ,  
Bacon and Woodrow,  
St Olaf House,  
London Bridge City,  
London SE1 2PE

Yours faithfully

S. BENJAMIN and R. VERRALL  
Editors, Claims Reserving Manual, Volume 2



## ACTUARIAL VACANCY

Faculty Position at the **University of Manitoba**  
Department of Actuarial & Management Sciences  
Faculty of Management

The Department of Actuarial & Management Sciences has an opening for a tenure-track appointment in actuarial science at the Assistant, Associate, or Full Professor level beginning July, 1991 or other mutually agreed date. Appointment as Department Head may be considered at a later date. Salary is competitive at all levels.

Qualifications include a Ph.D. in Actuarial Science or closely related area, or F.S.A. (F.C.I.A.) or equivalent. Candidates should have a strong interest in effective teaching, and evidence of research capability and interest in actuarial research. An appointment at senior levels requires an excellent research record in actuarial science. Industry experience is an asset and all candidates should have an interest in participating in an actuarial program within a management school context with a balanced emphasis on teaching and research. Primary duties will be teaching graduate and undergraduate courses in actuarial science and developing a research program in actuarial science.

The Faculty of Management offers actuarial education within a general management program at the undergraduate level. Students choosing the actuarial pattern graduate with a Bachelor of Commerce (Honours) degree with a major in actuarial science. Specialized actuarial education is offered in a Master's of Actuarial Science program in the Faculty of Management and joint undergraduate programs with the Faculty of Science. The Department of Actuarial and Management Sciences also houses the L.A.H. Warren Chair in Actuarial Science. A Ph.D. program in management science is scheduled to begin in 1993. There are presently three full-time faculty members in the actuarial area.

The University of Manitoba encourages applications from qualified women and men, including members of visible minorities, aboriginal people, and persons with disabilities and provides a smoke-free work environment. In accordance with Canadian immigration requirements, priority will be given to qualified Canadian citizens and permanent residents of Canada.

Applications will be accepted until April 15, 1991 or until the position is filled, and should be sent to:

Dr. JERRY GRAY, Associate Dean  
Faculty of Management  
University of Manitoba  
Winnipeg, MB R3T 2N2



## 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,  
Eagle Star House, 9 Aldgate High Street,  
London EC3N 1LD, 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

Björn AJNE	Sweden	Chairman
James N STANARD	USA	Vice-Chairman
G. Willem DE WIT	Netherlands	Secretary
Jean LEMAIRE	Belgium/USA	Treasurer
Hans BUHLMANN	Switzerland	Editor/IAA-Delegate
D. Harry REID	United Kingdom	Editor
Edward J. LEVAY	Israel	Member
Patrick PICARD	France	Member
Jukka RANTALA	Finland	Member
LeRoy J. SIMON	USA	Member/IAA-Delegate
Jürgen STRAUSS	Germany	Member
Gregory C TAYLOR	Australia	Member
Alois GISLER	Switzerland	Co-Editor
David WILKIE	United Kingdom	Co-Editor

Neither the COMMITTEE OF ASTIN nor CHUTRICK s a are responsible for statements made or opinions expressed in the articles, criticisms and discussions published in *ASTIN BULLETIN*

### **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. 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.
11. 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.