ABSTRACT OF THE DISCUSSION

Professor R. J. Verrall, Hon. F.I.A. (introducing the paper)**:** There have been many significant developments in recent years in the area of stochastic models for claims reserving, and this paper attempts to bring together enough of these to form a coherent approach to the subject. We hope that, by bringing the material together in this way, we have made it easier for people to learn and to understand about stochastic modelling, and we also hope that it will encourage the use of stochastic models in practice.

The paper begins by considering the chain-ladder technique, simply because it is so well known and widely used. This does not imply that we believe that the chain-ladder technique is suitable for all data sets, or even that it is the best model to use. It is, however, the best entry point to stochastic methods for a paper of this type. We present a number of different approaches to the chain-ladder technique, and concentrate on the predictive distributions, showing, for example, that the over-dispersed Poisson and the negative binomial models are different ways to obtain the same thing. Once in a stochastic framework, the way is open to consider many other models and to consider the data in greater detail. We cover some of the wide range of available models in this paper; space requirements prevented the inclusion of others, but we hope that the paper will help readers to understand and to use stochastic models in claims reserving, both those described in this papers, and others in the literature.

From an implementational point of view, we believe that the bootstrap method and Bayesian methods have a great deal to recommend them. In both cases, many of the computational difficulties which may be encountered are by-passed, and, in the case of Bayesian methods, it is surprisingly easy to obtain the predictive distribution. The predictive distribution is extremely important, and there are many practical issues, which we emphasise in Section 10, on which it sheds light. The Bayesian framework, and Markov chain Monte Carlo (MCMC) techniques in particular, also make it possible to incorporate many practical ideas into the stochastic framework, and we believe that the full impact of this approach has yet to be explored.

I now ask a number of, perhaps provocative, questions, which I hope will help stimulate the discussion:

— What would be considered to be a 'best estimate', and would all actuaries agree on this?
— Should prudence be allowed for in setting reserves, and should this be based on a predictive distribution? If not, what other consistent basis, which takes account of risk, could be used?
— How should an analysis of reserves be carried out when assessing the financial strength of a general insurance enterprise?
— Finally, can these questions be properly answered without the use of stochastic methods?

Mr D. E. A. Sanders, F.I.A. (opening the discussion)**:** The paper concentrates on theoretical aspects of stochastic loss reserves, and, it must be emphasised, in the main deals with the mechanical techniques, and, in the main, without considering judgement. The authors draw on many years of research and experience**,** and summarise their own and others' results giving appropriate references.

An important consideration in assessing claims reserves is to first establish the reason for undertaking the exercise. It may be, for example, a best estimate, a solvency assessment, a commutation, and so on. Each of these reserves will almost certainly be different. A reserving exercise is also equivalent to a pricing exercise. Indeed, reserve exercises result in financial transactions between different parties or generations of Lloyd's syndicates. It is for these different reserve requirements that we need to understand the dynamics and variability of the outcomes of our estimates. There are also other considerations, such as the stability of the outcome and the interpretation in the light of known, but uncertain, events.

The starting point of many practical reserving exercises, once we have reasonable data, is the chain-ladder method. The authors also start at the same point, and give an interpretation of the approach. One important consideration of the chain ladder is the implicit assumption that insurance claims are strongly correlated with the period in which the earth goes round the sun. For some classes this assumption holds — for others it does not.

The various stochastic approaches indicated by the authors have been well documented in actuarial literature.

One of the issues underlying many of the models is the need for increasing amounts (or non-negative increments), either individually or columnwise. In many lines of business this is not a constraint; paid claim amount development often gives a good measurement of the ultimate for personal lines business. However, for commercial and London Market business this gives real constraints with varied cash flows, reinsurance, salvage recoveries, and so on. Incurred claim developments often have negative incremental development, and there is difficulty in seeing how this may be treated in the theoretical background outlined by the authors. Other practitioners may wish to pick this up — as, indeed, did Thomas Mack when he reviewed his approach in the light of these issues in the Cambridge ASTIN meeting (Mack, 2000).

The use of overdispersed statistical models is becoming in vogue in actuarial circles, and the approach that the authors take is a very natural way of allowing for, and measuring, uncertainty and risk. These models will become more important as we move into fair value assessment and deflators.

I observe that the use of a recursive method, as in the overdispersed negative binomial model, effectively appears to work along the first diagonal only, and estimates beyond this diagonal will have process errors. The selection of a constant dispersion factor for all time may also not be appropriate.

In Sections 3 and 4 the authors consider other modelling distributions and parametric models. These include classic models, such as the Hoerl curve. These types of curves are often selected for their mathematical niceties, such as the ability to fit a linear model to a transform of the parameters, and not necessarily for the sound reason that the distribution is a good fit to the data. This invariably leads to several piecewise curves being used, associated with some form of smoothing for a better fit. The models may, in practice, have a danger of becoming too complex or over-parameterised. The authors have addressed this problem, with the result that the chain ladder and the Hoerl distribution may be considered as two extremes of a series of approaches. This makes me recall the work of Hilary Seal and others, where risk aversion and accident proneness can be measured in the context of lying somewhere between the two extremes of certain distributions, one of which is that famous negative binomial. Such results enhance our understanding of the issues.

In Section 6 the authors draw on the Bayesian approach and the Bornheutter-Ferguson (BF) method. The authors correctly imply that the BF reserve is a credibility approach. When establishing an actuarial reserve we apply a substantial amount of judgement. In practice, in addition to the paid claims, there are case estimates and incurred claims development, market knowledge, unusual claims to consider, premium rates and contract condition changes, and so on. The process of establishing a loss reserve is Bayesian in approach because, as actuaries, we modify the pure chain-ladder calculation to derive a selected ultimate. This is the natural framework in which we, knowingly or unknowingly, work. The selected ultimate that we apply to our reserving can be shown to be equivalent to the sum of a chain-ladder reserve, weighted by the chain-ladder's development percentage, and an unknown initial BF ratio, weighted so that the combined weight is unity. This has direct equivalence in the Bayesian estimation process. Furthermore, we can extend this concept by considering the BF initial estimate, not as a point, but as a distribution of estimates, with the fixed initial loss ratio as the mean of the distribution. I outlined this approach in a paper to the ASTIN Glasgow conference, with an indication of how judgement reduces the variability of the reserve estimates, using stochastic simulation.

Section 7 includes some detailed calculations. Here the authors show the equivalence of a number of methods, and how other approaches vary in deriving the estimates. I consider the results of the authors' calculations in the light of a reserving process. Actuaries tend to look at the ultimate position and the ultimate loss ratio in practice. They also need to talk to underwriters, who ask for a range of ultimate loss ratios so that they can compare their assessed results with those of the actuaries. See Table D1.

Table D1. Cumulative data

Year	1	2	3	4	5	6	7	8	9	10	Ultimate
1	5,012	8,269	10,907	11,805	13,539	16,181	18,009	18,608	18,662	18,834	18,834
2	106	4,285	5,396	10,666	13,782	15,599	15,496	16,169	16,704		16,858
3	3,410	8,992	13,873	16,141	18,735	22,214	22,863	23,466			24,083
4	5,655	11,555	15,766	21,266	23,425	26,083	27,067				28,703
5	1,092	9,565	15,836	22,169	25,955	26,180					28,927
6	1,513	6,445	11,702	12,935	15,852						19,501
7	557	4,020	10,946	12,314							17,749
8	1,351	6,947	13,112								24,019
9	3,133	5,395									16,045
10	2,063										18,402

Table D2. Incurred as % of chain ladder ultimate

Year	1	2	3	4	5	6	7	8	9	10
1	27%	44%	58%	63%	72%	86%	96%	99%	99%	100%
2	1%	25%	32%	63%	82%	93%	92%	96%	99%	
3	14%	37%	58%	67%	78%	92%	95%	97%		
4	20%	40%	55%	74%	82%	91%	94%			
5	4%	33%	55%	77%	90%	91%				
6	8%	33%	60%	66%	81%					
7	3%	23%	62%	69%						
8	6%	29%	55%							
9	20%	34%								
10	11%									

If we consider the year 3 results from the amounts derived by the authors, we can readily see that the calculated ultimates lie in a range \$23m to \$24.8m, with the exception of the reserves calculated by the log-Normal model (Table 19) — which appears to be a rogue assessment. Furthermore, the error is about 20% of the ultimate in most cases — the exceptions being the gamma and rogue log-Normal estimates. Looking at the chain-ladder losses expressed as a percentage of the chain-ladder ultimate, then, in development period 3, it can be seen from Table D2 that the underlying development percentage was in the range 55% - 60% of ultimate, with the second year from the top being an outsider at 32%. The current chain-ladder estimated ultimate of \$24m for year 8 is at the 55% level, and, accordingly, one should, in the absence of any information, be comfortable with that amount as a best estimate. The 20% error factor derived by the calculations of the authors thus seems high, as it is driven, in part, by the unusual development for year 2 period 3 for 32%, and removing this outlier would result in a substantial decrease in this factor. This is why it is vital to look at the data, and not to apply reserve techniques in a vacuum.

Now consider also the most difficult year for the authors — the last year with only one data point. Excluding the log-Normal tables (Tables 19 and 20), the calculated ultimates lie in a range \$16.5m - \$20.1m, with the chain-ladder estimate at \$18.4m. At this level, the current claims of \$2.1m represent 11% of selected ultimate within an historic range of 1% to 27%. The error estimate calculated by the authors ranges from 60% to 130% of the ultimate, which is consistent with the crude estimates that I have used. Based on this information alone, it would be very difficult to come up with a best estimate, as the variability potential is very real. One would, of necessity, have to apply judgement.

These two examples from the data which the authors used illustrate: firstly, the danger of running with a model without having a good look at the data, as the variability may be due to an irregular development or some other factor; and secondly, that the approach has a real advantage in making actuaries aware of the real uncertainty, and seek additional information to reduce that uncertainty to manageable levels before committing themselves to selecting the reserve. The underwriter, however, will not appreciate being informed that, although he thought that he wrote at a loss ratio of 100%, the range is 20% to 250%. Such an approach will possibly detract from the value of the actuary in the reserving process.

In Section 8 the authors introduce simulation techniques. In a practical sense this is a good way forward. It is fairly easy to emulate the underlying reserving model that produced the best estimates — including restraints and BF adjustments — on spreadsheets, and produce a range of outcomes that both represent model and data uncertainty.

The authors then discuss dynamic financial analysis (DFA) models, with an indication that the approach may be used. This is a natural consequence of introducing the stochastic process. The main issue of DFA claims is the cross-relationship of the various underwriting classes which is usually ignored in the reserving process. The other issue is that, for a practical DFA model, you certainly need a significant degree of judgement and simplification to manage many of these inter-relationships. The methods described by the authors in the previous sections deliberately do not use judgement, and therefore need to be adapted if they are to be of any real practical use. If the model incorporates too much variability, and does not recognise the necessary underwriting and management controls, the results will, at best, lack credibility, and could lead to inappropriate actions being taken by management.

In the final section the authors introduce to the discussion questions as to which model is best. My view is that the answer depends on the data, and that, to some extent, the data need to be used as part of the process to select the 'best' model. Others may have different views. The authors also raise the issue regarding best estimate. My view is that the best estimate is not viewed in isolation, but is conditional on the information and data available. More information leads to the refinement of the best estimate and the range. This information may be soft or hard, and two actuaries can clearly come up with different best estimates for the same business lines. This is a consequence of the view that reserving is, of necessity, a Bayesian process. Others, of course, may not share this view.

Mr K. P. Murphy, F.I.A.: This is an extremely comprehensive paper, with many new ideas. It considers an impressive array of statistical methods in the context of claims reserving. As the opener has shown, the authors have been very brave to include worked examples within the paper, which will generate a certain amount of critical attention in the future.

Professor Verrall posed a few questions in his introduction. I think that he answered his first question, about the definition of best estimate. The definition within actuarial guidance is that it is the expected value of the distribution of possible outcomes of the unpaid liabilities, which I think is a pretty good one. I think that most actuaries would tend to agree with the definition, but I am not sure that we would all agree with what is that best estimate.

The second question concerned prudential margins, and essentially asked whether reserves should be held at, or above, the best estimate, and, if above, what level of prudence is acceptable. That is something that, perhaps, we, as a profession, should be addressing. Obviously, it is key to have a predictive distribution of potential reserve outcomes, even if that is just an approximation to the true distribution. There are a couple of ways of defining a prudent level. One is to take a defined percentile of the predictive distribution (a value-at-risk type measure). The second way is to look at the expected level of liabilities, given that they are above a particular percentile (a tail value-at-risk). A slight cautionary approach with this, though, is that the distribution in the tail is quite key. There is a danger that the results that you get might not be entirely reasonable.

The third question dealt with financial strength assessment, which, I presume, stems from the Financial Services Authority (FSA) proposals on financial condition. Again, this is quite new, and there is very little guidance yet as to how to allow for it. I do not think that best practice has yet emerged. Financial condition is not just a function of reserving, but takes many risks all together. Reserve levels are a very important element.

This is a very timely paper, given that the new FSA proposals will be in force in a couple of years' time. The fourth question was about whether the stochastic framework is necessary for this. I do not see how it can be done without some sort of stochastic framework.

Mr J. W. T. Tanser, F.I.A.: I think that this is an excellent summary of the approaches to stochastic claims reserving. My experience of using stochastic reserving techniques suggests that there are two key issues that need to be borne closely in mind when interpreting the results.

The first issue is alluded to in ¶1.4, and concerns the use of judgement. The methods described in the paper generally assume the mechanical application of the given method with no application of judgement. In practice, one would hope to improve the quality of estimates by using judgement, and so the calculated error may overstate the actual variablility. However, it is possible to amend some of the methods described in the paper to reflect some aspects of the application of judgement. For example, if the actuary believes that a given development ratio in the historic triangle is truly exceptional, than he or she might limit the weight given to this factor when deciding upon the development pattern used. Similarly, when using the bootstrap technique, it is possible to reduce the weight given to truly exceptional residuals in the resampling process. In this way, the bootstrap technique can be made to mimic some of the aspects of 'actuarial judgement'.

The second issue relates to the appropriateness of the model underlying these techniques, which has been mentioned in the discussion, and is also in the paper in ¶10.8. The methods covered in the paper all attempt to reproduce the unadjusted chain ladder in some way. Therefore, the variability measures emerging assume that the chain ladder is an appropriate model. It is important that the data actually follow that. In practice, of course, no model is ever perfect, and, for this reason, stochastic techniques may tend to understate the true uncertainty. The extent of this model error will, of course, vary from case to case, and the actuary will tend to pick a model which will minimise his errors. However, this model error should not, generally, be assumed to be small.

For example, the chain-ladder technique is frequently used for classes where some event will affect all cohorts. Common examples are, for example, employer's liability or motor, where all accident years are exposed to changes in court awards. A stochastic technique, based upon the chain-ladder model, will not include an appropriate probability of such 'calendar year' effects, and so may underestimate the variability.

Mr P. Sharma (a visitor; Financial Services Authority)**:** As a non-actuary, I begin my contribution by saying that there is much in this paper that I do not understand, but that is probably to its credit rather than to its detriment!

One of the great concerns that there is in the modernisation of the insurance industry, and the modernisation of actuarial practice in the insurance industry, is that it is not enough that best practice is achieved in the biggest and the most able firms. What is needed is a diffusion of knowledge to all firms within the industry, and for stochastic approaches to be used much more widely throughout the industry. Any method, such as in this paper, which can cast light upon those approaches, which can popularise those approaches, which can educate more widely within the profession, and, indeed, outside the profession, is to be warmly welcomed.

Stochastic models are going to be absolutely fundamental to the reforms in accounting that the International Accounting Standards Board is introducing. The second thing is that we, at the FSA, want them to be absolutely fundamental to the reforms in regulation that we are introducing. However, I think that the technical challenges of making any approach that looks at the full probability distribution of losses and the technical difficulty of making these relevant to regulation are greater than they are for making them relevant to accounting. One speaker has already mentioned that the regulators are interested in the whole of the distribution, and especially the end of one of the two tails. One of the great challenges, one of the things that I hope this paper helps with, but I am not technically sufficiently able to say whether or not it does, is to get techniques that can give us meaningful information about the tail from what are, necessarily, very little data, and, what is more, that can do all of that in a way that is accessible, not just to the largest companies, but to small and medium-sized companies. I feel sure that, while this paper may not be the answer, it is a significant and worthwhile step in the direction of that answer. So, for that reason, if for no other, we, at the FSA, warmly welcome the paper.

Negotiations at the European level on the new approach to the replacement for the current solvency margin system are entering a fairly critical stage, and demonstrating to our European colleagues that stochastic models are practicable is going to be fundamental to how these negotiations develop. I shall be handing out several copies of the paper when I get to the European negotiations shortly.

Mr D. H. Craighead, F.I.A. (in written contribution that was read to the meeting)**:** This paper is very valuable for the insight that it provides into the statistical bases of claims reserving in general insurance. Actuaries need to have a clear understanding of the underlying statistical bases of the methods which they use. That is, however, only one side of the coin. The other side resides in the need to have a complete grasp of the problems that are encountered in practice, and the methods that are used to overcome those problems, insofar as they reflect on the statistical framework assumed.

I express grave doubts as to whether the chain-ladder method is at all suitable for London Market business. The difficulties encountered, in practice, are as follows, though there may well be others:

(1) Claims arising from large catastrophies, asbestosis and pollution must be separated out and handled quite differently, with their own patterns of development.
(2) Because of the nature of the underwriting, almost all business transacted, although divided into property (short-tail) and liability (long-tail), is actually, to some extent, a mixture of the two, and the ratio of the mix varies from year to year.
(3) The statistics produced, and on which estimates must be made, often contain errors to a major degree, usually corrected at a later stage of development.
(4) In liability business, there is a chronic understatement by cedants of outstanding claim amounts on the liability side; more extreme in the United States of America than in the United Kingdom or in the rest of Europe, but noticeable at both ends. Liability business tends to have an extremely long tail, and this feature presents considerable difficulty in determining the tail factor. It is not infrequent for the actuary attempting claims forecasting to have development figures for only six or seven years for the longest-developed year, and yet liability claims are not likely to be settled fully for some 15 to 20 years in the U.K., and for some 20 to 25 years in the U.S.A.
(5) There may be, and frequently is, a major change in underwriting policy because of changes in the underwriting team or in management decisions. Such a change may have the effect of altering later development patterns to a major extent.

This paper may be more valuable as a tool for understanding the statistical basis of the chain-ladder method than as an argument for its use in practice.

Dr N. A. Bruce, F.I.A.: This paper is an incredibly good summary of all the work to date, and I think that it will be an excellent grounding for those new to the subject. From an audit perspective, the idea of reserve ranges may, in the extreme, impact the auditor's role directly, due to the issue of 'going concern'. If a range of reserves includes the possibility of insolvency (as it probably will), at what level (or percentile!) does the auditor have to raise this possibility to management or to the regulators?

Following on from regulatory issues, the split of prediction error into its various components may provide some help in setting the theoretical basis of the phrases 'possible outcomes' and 'reasonable range of best estimates', as we have already heard. The latter may be described by the use of multiple models or parameters, implying that the overall range will be wider than that given by any one set of parameters within a model.

A possible method to help determine this parameter error for any given model is the reverse Monte Carlo. I am not sure whether it is well known in the profession, but I am aware of it from my background in condensed matter physics. This is effectively the stochastic equivalent of MLE estimation, and is useful for more complicated models, in particular those not based on standard probability distributions or even analytical distributions altogether. This method is able to plot the probability distribution of parameter sets for a model, allowing for outliers and other normal factors related to a finite data set. This will then be able to give an indication of the range of parameter sets that could give a plausible model for the past data that we actually have, and hence the variability of the reserve range overall. The variability will, of course, increase as the probability distribution flattens out around the MLE parameters.

To some degree, all models presented in the paper rely on past experience to project forward. This is not necessarily a good guide to the future, in particular for the longer-tail classes, because of calendar year effects. Such effects include court rulings, emergence of latent claims, etc. The most effective (and difficult) way to model these claims is to model from the ground up, working with the underlying processes of claims, and hence the analysis of any stochastic drivers that would impact the class (this would rely on some work both from past data and also a fair amount of actuarial judgement). One possible stage between the full ground up methodology and the quoted mechanical methods is to use stochastic inflation when projecting future claims. This would require past inflation to be stripped out of any triangles that you do use and explicit distributions to be used to model future calendar year effects. This may give the same best estimate as the current method, as the average future inflation may well be similar to that implicit in the CL method, but the variance is likely to be much higher for classes where large shocks in inflation are allowed for.

I now turn to communication of these results and the role of management, which, I agree with the authors, is an important area, which should be handled very carefully.

Eventually, the uncertainty in the reserving process should be allowed for in management information. This would affect the reserves predicted in future calendar years, and hence present management with a distribution of balance sheets for each year going forward. Thus, strategic decisions could be made based on the distribution of results going forward. I am, of course, talking about DFA, and its use in understanding the business and how it reacts to management decisions. This, I believe, is the way forward.

The paper provides the ideal starting point for some more sophisticated tools for determining reserves and their ranges. I believe that this is at one end of the spectrum of tools that ends with the full DFA, which, ideally, should be used as a fundamental management tool.

Mr S. Christofides (a visitor)**:** This is a substantial and comprehensive paper covering over 15 years of research in stochastic reserving methods, gathering all this research into a cohesive and structured paper, and making this subject more accessible to a wider audience.

The paper illustrates very clearly the difficulties of finding appropriate statistical models to fit insurance claims developments. In this context, the work to-date is primarily concerned with the choice of the underlying model and the appropriate error distribution. For those of us who are not professional statisticians, this is a difficult enough task before we are made aware of the existence of distributions such as the over-dispersed Poisson.

My own experience with stochastic reserving models is that, in most instances, the prediction errors estimated look too high in relation to what one expects in practice. The most likely reason for this is the limited amount of information that is generally used by these models. Much relevant information is ignored, even though it is often a significant factor in the actual reserve setting process. Bayesian methods and bootstrapping, as well as a move to quarterly or even monthly cohorts for shorter-tail classes, may help to bridge this gap. I am less convinced that modelling based on individual reported claims is ever going to provide a complete picture, although it may be helpful in managing any incurred, but not reported (IBNR).

Looking to the future, there are at least two major challenges before fair values can be estimated. Firstly, there is the economic impact on future claim payments to be evaluated.

Secondly, there is the whole tricky issue of consolidation across years, classes and territories to get to the overall figures to be reported. In other words, we need to understand both the systematic as well as the non-systematic components of the reserving risk. It is possible, for example, that, for large, well managed companies, the overall reserving risk reduces to just the systematic component, whereas, in the less well managed companies, it is the non-systematic risk which dominates. An ability to estimate market and operational risk is the key to making any substantial progress in this area. I am not aware of any stochastic reserving methods that tackle this issue at present.

In any case, as reserving risk is what is left as we unwind the underwriting risk, that is, in turn, just one element of the overall risk faced by the company, it is clear that a more complete model of the financial dynamics of the company is required if we are to make any significant progress. DFA is the current name for such models, as already indicated. The authors discuss DFA very briefly in Section 9, and suggest that this could be the subject of a future paper. I agree that this is the approach most likely to lead to success in the medium term. A good starting point here is the Casualty Actuarial Society's web site, which provides easy and free access to almost all the relevant DFA papers published over the last ten years.

Mr M. T. Malone, F.I.A.: It is useful to see, in one place, descriptions of the various approaches to estimating reserve variability, and it is even more useful to see the connections that have been revealed between these various models. I felt that I began to see the light in what was previously, perhaps, a dizzying array of various approaches. I wish that the paper had been around a few years ago, when I first started to try to understand what was going on with this topic.

I am a little sceptical, from a practitioner's standpoint, as to how much we can rely on estimates of reserve variability derived from a development triangle. I feel that, just because I use a chain-ladder based estimate in my work, it does not necessarily follow that I am wedded to the estimate of variability that would be produced by a model simply because it reproduces the same mean as I have found. I think that the authors touch on this point when they comment that it is, perhaps, futile to ask which of the various models that they describe is the best model, and I would suggest that it is probably not too helpful to get overly excited as to which model could be described as the true model underlying what is a chain-ladder based estimate, which, I get the feeling, was, perhaps, where much of the research was going in the past.

Looking forward, it seems to me that a more useful line to pursue is the DFA approach. I would feel more comfortable with modelling the drivers of future variability on this basis than with looking at the possibly limited variability captured in a few exposure years in a triangle, plus some model error.

I would also, perhaps, favour the DFA route as being one that was more readily able to allow for that future element of reserve variability which is actually a function of fluctuations in actuarial judgement. It seems to me that the impact on a company's published results will stem, not only from random variability of a stochastic nature, but also from decisions made by actuaries as part of the management.

Mr N. Shah, F.I.A.: I found this paper to be a very useful summary for the use of generalised linear modelling techniques for small data sets. My comments are primarily from a practitioner's point of view.

The triangulation data that these techniques have been applied to are just a consequence of history. They come from an era when computing power was expensive. Therefore, I question the value of actually applying such techniques to such limited data. Such sophisticated techniques may be more useful if applied to the underlying claims data, as has been alluded to by several speakers. In view of this, there is a danger that the results may be viewed as more scientific than they really are, and may be given more credibility than is truly justified for them.

In situations where we have mergers, commutation, and where we are pricing risks, where we need a robust view on variability, such methods and techniques are essential. However, they need

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

to be applied to more comprehensive data. That includes specific drivers of the claims and experience. Of course, this type of work is also essential to implementing any type of DFA. I commend the authors for bringing to the attention of practitioners the various studies that are out there, and how they may be applied to the data, and my vision for the future is that they will be applied to more comprehensive data rather than the limited amount that most reserving practitioners are forced to use.

Mr P. D. Johnson, F.I.A.: I join in the general welcome that has been given to this paper, not just for the reasons that have been given in terms of the way in which the paper appropriately focuses on the variability of our estimates of outstanding claims and brings together a lot of material that has been produced in recent years on this topic, but also because the paper illustrates to me very clearly the progress that the profession has made in this field since I first looked at outstanding claims over 35 years ago. In those days the data that we had were primitive by today's standards, and the same went for the methods that we employed, which I think reflected the state of the art at that time.

Since then, insurers generally have far better organised their record-keeping systems, which will now often generate the kind of data that can form the basis for papers such as this one. Since we lay claim to being a profession that knows how to handle uncertainty, we need to look at the shapes of distributions and the spread around the expected value. The more that we can reduce the uncertainty associated with our claim to be able to deal with uncertainty, the better. However, at the same time I do have some misgivings of the kind expressed by Mr Craighead, in his written contribution.

The authors have, to some extent, disarmed criticism by their disclaimers, in ¶¶1.3 and 1.4, where they say: "It is not the aim of this paper to present a panacea for claims reserving. There will be sets of data for which it is difficult to justify any standard model-based approach", and "for the most part, this paper considers claims reserving techniques *applied mechanically and without judgement* (once a particular technique has been selected)." I believe that we need to emphasise our awareness of the untidy features that are found in practice. The adjustments that we make to the raw data to reflect what is known about changes in underwriting practice or in the cover provided must be relevant to our measurement of uncertainty.

There are also influences external to the insurer, one of these being inflation. As someone who was engaged in claims reserving in the rollercoaster years of the 1970s, when, at one point, the rate of earnings inflation reached 36% p.a., I feel that we must always have this source of uncertainty in mind.

If we make probability statements about the variability in our estimates of the kind implied in the last sentence of ¶10.9, which says: "it would be of considerable value to know the estimated probability that outstanding liabilities might exceed the reserves held, which could be obtained if the predictive distribution can be estimated", then I think that we need to be very careful to state the caveats that apply to any such estimated probabilities. Any marked change in the rate of future inflation could have a profound effect on the figures that we have quoted.

Mr D. I. W. Reynolds, F.I.A.: I think that there may be more benefits of the Bayesian statistics approach than the authors bring out. The problems referred to earlier of dealing with small companies and large companies, whether it is consultants or whether it is people within the regulators who have access to a range of data, then the prior distribution under a Bayesian approach can take account of data that are in one's possession in relation to the market that the company is operating in and the claims reserves that one is trying to calculate. So, one can actually bring a broader range of data into one's prior distribution, and I think that this must be a strength of the method.

If we are looking ahead to how regulation may change, it should be noted that the Australians have already leapt ahead, as you would expect, in this area, and they are setting the standard that reserves should be set at the 75% probability point. Clearly, if regulation moves that way, there is no doubt that you will have to use stochastic methods of one sort or another.

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

I now raise two points which have not yet been discussed. One is the question of whether you are assessing gross reserves or net reserves. Given that companies may have quite different reinsurance portfolios, it would be quite difficult to take a net prior from a large company and to translate it to a small company. What are the views of those who have some practical experience and can talk on their views on gross or net reserving? I suspect that the answer is gross to work with claims, and then allow for the individual company's reinsurance to the extent that one can. The other point is on the practicality of using individual claims rather than triangles. Certainly from reading the paper, it sounds as though that is the way that the authors think things are going. Do others think so too?

Ms C. Cresswell (Institute Affiliate)**:** Fair value accounting is due to be with us, at least for listed companies, for the 2005 year-end, which is not far away. One of the things that we shall need to do is to establish fair value numbers to set risk margins. To do this, we need to think about the whole distribution of possible outcomes and, in particular, tail distributions, which feed into risk measures such as Butsic's expected value of policyholder's deficit measure.

The challenge for the research-minded actuary (or the actuarial-minded research statistician), drawing, no doubt, on both the modelling techniques described in the paper and enhanced, where appropriate, by the judicial use of actuarial judgement, referred to by earlier speakers, is to provide insight into the following questions:

— What size of risk margin do various classes of business deserve?
— Does writing larger volumes of business make the proportionate risk margin smaller; and, if smaller, typically how much smaller?
— How effective is reinsurance in mitigating risk?
— How do we combine risk margins between different classes of business?
— How do we arrive at some additive measure that accountants can use for segmental reporting so that their numbers still add up?

In summary, the time for thinking in terms of the distribution of possible outcomes rather than point estimates has arrived. The need for tools and techniques to help establish such numbers is self-evident. Papers such as the one before us, and research efforts which, I hope, will flow from them, will help the profession to address this issue in a constructive, comprehensive and coherent fashion.

Mr J. P. Ryan, F.I.A.: I shall, undoubtedly, be referring to this paper on many occasions in the future. It is an extremely valuable reference document, for which I, and many generations of actuaries in the future, will be indebted to the authors.

I now consider the comment about the parsimoniousness of the chain ladder, and also the comments made by Dr Barnett, Dr Dubossarsky and Dr Zehnwirth, in their written contribution. [See 'Written Contributions'.]

The authors referred to this, and it has been referred to on a number of other occasions, but it needs to be put into some degree of context. For a large motor insurer or some similar type of company, as a statement I think it is fairly valid. Changes in that type of organisation move relatively slowly. One tends to pick up trends in the data; the theoretical framework that the authors and some other speakers have referred to does, in fact, apply. Modelling can take place without difficulty using the techniques used in the paper, and their comments are perfectly valid, because it is impossible to change a very large organisation very rapidly. Therefore, for a large, fairly well balanced account, this is okay. This is probably also borne out by the comments of people like Professor Taylor. [See 'Written Contributions'.] My perception is that he is working with similar types of portfolios.

However, if you move away from those into some of the London Market companies, you come across the points that Mr Craighead wrote about, and I think that the issue is somewhat different. Each data point is almost completely separate, and therefore, from a practitioner's point of view, it is the knowledge of what is going on in the account that comes out of that that

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

makes the chain ladder very valuable. It is no longer a purely statistical exercise, because you do not have a coherent body of data. It fails every statistical test. Nevertheless, the practitioner can learn much from it. Take into account, say, somebody writing a book of D&O business over the last few years, last year-end was distorted by Y2K issues. There has been a downturn in the U.S. economy, which has had a big impact on D&O claims. Previous to that we had boom; stock market conditions had given rise to another set of claims. If we had had a change in claims reserving practices during that period, every single data point on the chain ladder would be different. A practitioner who understood how those affect things would have some idea of what was happening in the underlying business. In that case, the concept that one should not use the chain ladder or that a parsimonious model is a 'good thing' is nonsense in that type of environment, but makes perfect sense in something like a U.K. motor account. Therefore, I think that it is useful to have a health warning at the beginning of this type of paper.

The authors, with their smoothing from one type of model to another, have a powerful technique, but I think that it can go a little bit further, particularly when it comes to some of the implied criticisms from Dr Barnett *et al.* (See 'Written Contributions'.) I think that it is very important, but they are actually missing the point; they are only talking about large, stable portfolios.

I now move on to some of the other issues that Mr Christofides referred to. The authors allude to this at the end of the paper. The actual reserve variability of the account, as a whole, is very much less than the individual reserve estimate. It is an important point. Mr Christofides carefully pointed out that there is much more information than just comes in the model. Certainly, an exercise that I conducted in the past, while devising a risk-based capital system, was to get a whole range of people to re-estimate historic paid claims data. You actually got very much less variation in reserves for the portfolio as a whole, so I think that that bears out what is actually happening, and emphasises the point that the authors make about doing further exercises in this area.

One further point worth making, which Mr Shah referred to, is that modern computer techniques mean that we do not necessarily have such data constraints in the future. I, for one, have been exposed to new sets of data sources in the last few months, and I can assure you that cheaper computing power allows us to move into a much broader range of activity rather than just staying at the whole account level of the company.

Mr A. Kaufman (a visitor; Casualty Actuarial Society)**:** I am here on behalf of the Casualty Actuarial Society (CAS), to thank these authors for this important contribution to the literature.

The analysis of uncertainty in loss reserve estimates is an important issue for general insurance actuaries worldwide. It was one of major topics of targeted research identified by the CAS over ten years ago, when the CAS first started efforts at identifying key research areas. The authors have made important contributions to this area in the past, and we are very happy to see this additional contribution.

Mr M. H. Tripp, F.I.A.: Why do we throw away information? This question has already been hinted at, and needs reinforcing in the domain for thinking about in the future. I have never been keen on silos, and it is important to learn between disciplines. Looking at the life side of our profession, you realise that work like this takes place at policy level detail. If you look within the general insurance part of the actuarial profession, there is a body of thinking that has grown up around premium rating and a body of thinking that has grown up around reserving. Are we getting 'over-siloed'? Could aspects of the methodology and the thinking that has gone into using GLMs for premium rating be brought more into play when it comes to reserving, where, at present, we tend to use aggregated claims data? I wonder whether we are missing out on using information that is available from exposure descriptions and from the circumstances of individual claims. I know that the traditional response to this is that there is all too much variability, but, in attempts to remove heterogeneity from data and to try to find better for the future, I look for support in thinking this through.

Mr P. R. Archer-Lock, F.I.A.: I make a few comments on the practical application of bootstrapping methods. First, the traditional way of applying bootstrapping methods is to sample the residuals from the observed residuals set. I think that it is important to bear in mind, when you are doing this, whether the residuals set that you have actually observed is a fair and reasonable representation of the underlying distribution of residuals. With insurance data, which are often very skewed, when you are looking at small data sets, which are small triangles, it is quite important to bear in mind that the triangle of observed residuals may not be sufficient in terms of coming up with a distribution, particularly for the tail. So, in those respects, we find it useful to look at some other benchmark residuals, or at least some other distributions, to see whether any adjustments need to be made.

Secondly, the authors refer to the fact that we, as actuaries, often use Bornheutter-Ferguson (BF) or other credibility methods to try to input judgement into the reserving process, and to stabilise the results of the chain-ladder method in the more recent years. We have found it quite useful to apply bootstrapping style techniques to a combination of a chain-ladder and the BF method. When you are using this sort of technique, it is also important to bear in mind that the BF method, by introducing judgement, tends to reduce the variability seen in the answers coming out of the more recent year distributions, so, it is worthwhile considering how you can put back some extra variability to reflect the actual choice of a prior. Some methods that we have used, which have proved useful in practice on this, are to consider the range of historical ultimate loss ratios, maybe adjusted for premium rate movements, or to look again at some benchmark loss ratio distributions and build that extra uncertainty into the bootstrapping process for the prior loss ratio.

I disagree with one point made in the paper. In ¶10.11 the authors comment that: "If the aim of the report is to inform management for the purposes of setting reserves, highlighting the upside might be inadvisable." Although I think that I understand the sentiments of the authors in making that comment, I think that, if you are truly to inform management, it is important to consider the potential upsides as well as the downsides.

Mr A. J. Newman, F.I.A. (closing the discussion)**:** Almost everyone has welcomed this paper. Some speakers have captured our feelings. Mr Ryan said that it is an extremely invaluable repository for future generations, and that is certainly what I feel. Mr Malone said that it would have been most welcome a few years ago, when we were struggling with this. That was not a royal 'we', that was him and I and others who have said similar things.

A number of speakers have addressed the issue of model error, the past as a guide to the future, and the issue of whether modelling works and whether we are looking at the right models. Some speakers have noted that, if the data do not follow the chain-ladder distribution, then the true range is very much greater than that predicted by the method shown in the paper.

Professor Mack [see 'Written Contributions'] wrote about techniques or criteria for comparing or choosing models within the GLM framework, which his model presents and has clarified. I think that that was one of the requests of the authors towards the end of the paper, in their discussion: "Where should we go now? How would we choose between models?" Professor Mack seems to have an idea, so maybe we should ask him.

Other speakers also spoke about model error. Mr Craighead wrote that even the chain ladder, which so many of these things are based on, does not actually work. Mr Ryan agreed, but said that it is not really supposed to; rather, it is just a methodology by which we incorporate our external knowledge of what is going on.

A number of speakers noted that reserving, as practised here, is effectively a data smoothing exercise, and it incorporates information from a number of different areas. This is exemplified in the Bayesian paradigm that is presented and has been commented upon a number of times.

Dr Barnett, Dr Dubossarsky and Dr Zehnwirth wrote [see 'Written Contributions'] that the method presented has too many parameters. Countering that, Mr Shah, Mr Ryan and Mr Tripp said that the issue was not so much that we have too many parameters, but that we are not using enough data. Mr Tripp commented that, for life insurance we use individual policy data. I

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

was working in the pensions market when people moved across from using commutation functions to individual policy data. I have actually seen it happen, so, I think that it could certainly be a way of moving forward in the future, especially since, as has been noted, the pricing people tend to use much more data than the reserving people ever get given. We have had papers presented here before on the links that there could be between pricing and reserving in the London Market. There was also, once, a presentation, at the London Market Actuaries Group, of a practical system in use in the London Market. I cannot believe that it now no longer exists.

Dr Bruce said something to the effect that he welcomed the paper's reference to the fact that the range of future outcomes can be split between a process error and an estimation error, if you like. I certainly find that, in talking to auditors and to management, it is very helpful to be able to distinguish between numbers that could go in the books as range estimates and how bad the actual outcome could possibly be. As the authors note in ¶10.7, when we use these methods we get numbers that are very much worse than anyone would ever have imagined. We have heard ideas that, perhaps, we can reduce that variability by incorporating more information, for example, in using Bornheutter-Ferguson or just looking at chain ladders and incorporating data in that manner.

We have covered a great deal of material in the discussion. Some speakers have discussed accounting and regulation issues. (I expected a little more debate or opinion on this.) We have had the idea that this paper is welcome because it will, it is to be hoped, cause the best practice from the larger and bigger, best staffed and resourced firms to go down to those who are perhaps, perforce, sailing closer to the wind. We have had ideas about regulation and how we are going to need the tails if we are to move to systems which include risk margins; for example, those currently required in Australia, where you have to reserve to the 75th percentile. I was expecting something to be said on negative incremental claims. I was disappointed that nobody addressed the issue.

We have talked about the International Accounting Standards fair value requirements, which are coming in by the end of 2005. It is a subject that is dear to my own heart. Again, I would have thought that more speakers would talk about this. I think that that leaves us consultants more room to tell you about it and to be paid for it.

One speaker commented about communication. What the paper is doing is showing how to take information out of the data that are there before us. The next stage, as has been pointed out, is how to communicate that information onwards. We have had various suggestions that the one thing that management do understand is the balance sheet, and that the best way forward is a DFA model based on balance sheets, pushing those into the future and seeing how management reacts. So, we end up with communication.

Dr P. D. England (Institute Affiliate; replying)**:** We are grateful for the contributions to the discussion of this paper, both verbal and written, many of which are extremely pertinent. Time permits only a brief reply now, but we will make a full written reply.

The opener and Professor Taylor [see 'Written Contributions'] made some interesting points regarding the use in the paper of constant dispersion parameters. In fact, that restriction can be relaxed through joint modelling, as described in Renshaw (1994a).

Mr Murphy, Mr Sharma and Ms Cresswell mentioned extremes and taking care over the extremes of the predictive distribution. I agree totally that the model assumptions should be scrutinised carefully if the main interest is in the extremes.

Mr Tanser made the important point that weights can be very useful, in practice, to weight out influential points, a technique also mentioned by Mr Sanders.

I was very pleased by the comments provided by Mr Sharma. I think that education and diffusion of best practice are very important. I am always grateful when others provide a full written account of work which they present at conferences and seminars. It is obviously dangerous to submit your own work to scrutiny, but I think that it is good for the progress of knowledge. I also think that stochastic models are fundamental to the work that actuaries do, and I will come onto that later.

Dr Bruce (and others) mentioned DFA modelling, and how the methods described in the paper fit into DFA modelling, particularly when stochastic inflation rates are included, Again, I would support moving in that direction. I would only add that actuaries are quite good at building the structure of DFA models, but the difficult problem of obtaining parameter estimates and incorporating estimation error remains. In that respect, I think that a Bayesian framework offers an extremely useful way forward. I am pleased that Mr Reynolds concurs that a Bayesian approach may be beneficial.

Mr Christofides mentioned the very interesting areas of market risk and operational risk, both of which are included in the list of risks which must be considered under the FSA proposals, although the FSA consider the risks independently. Mr Christofides made the important point that reserving risk has to be considered alongside operational risk, and it is actually quite difficult to incorporate all the various risks in a cohesive framework.

Mr Malone introduced a useful dose of scepticism about estimates of variability. I think that it is right to be sceptical and to have an inquiring mind. As we said in the introduction to the paper, we do not consider the methods presented to be a panacea, but we hope that we have outlined some of the underlying theory. Mr Malone also mentioned DFA in the context of business drivers. What actuaries call drivers are really causal factors — what causes particular effects which you observe. This introduces the need for causal models within DFA, which is a challenging area.

Mr Archer-Lock mentioned the practical application of bootstrapping, and the idea of bootstrapping within a Bornhuetter-Ferguson framework. He has noticed that the variability in this case is low, and he would like to introduce extra variability (I think that he is the only person who wanted extra variability). Really, this ties in with the comments that we made in the paper that the BF technique as a Bayesian method assumed perfect prior information, and what Mr Archer-Lock is noticing is the effect of that assumption. We recommend using a Bayesian approach, but incorporating variability in the prior.

Turning to some of the written comments, Professor Mack makes the distinction between models assuming independent incrementals and recursive models. Rather than making a distinction, we believe that some of the models are simply different representations of the same thing, and we have tried to show this in the paper.

Mr Craighead seems to be making the point that some data sets, particularly in the London Market, are not amenable to analysis by the chain-ladder model. In making the point, he is agreeing with the sentiments of ¶1.3.

Dr Barnett, Dr Dubossarsky and Dr Zehnwirth appear to object to the simple model structures that we have imposed on the mean in some sections of the paper. The structures that we have used are for illustration purposes only, and help explain the similarities to traditional methods, or models which have been described previously. We start with chain-ladder type models, making the point, in ¶4.1, that it may be considered over-parameterised. We move on to the Hoerl curve, and state, in ¶4.2.4, how the predictor structure can be altered to improve the fit. We also consider smoothing models, which, by their very nature, will follow trends in the data, and deal naturally with model parsimony, trading between degrees of freedom and model fit. We discuss, in ¶¶10.7 and 10.8, the importance of using a correct model structure, and the implications of not doing so.

As far as we can see, rather than there being 'very much to strongly disagree with', Dr Barnett, Dr Dubossarsky and Dr Zehnwirth agree with points made in the paper. We strongly support the exploration of a variety of predictor structures, including structures with calendar year effects. Many such predictor structures can be investigated within the statistical modelling frameworks which we have described.

Several contributors have referred to additional papers which are likely to be relevant and valuable. We accept that there will be other references which we have not come across, or which others consider are worthy of greater attention. We have provided over 50 references, which we hope include the major contributions to this growing area of study, and provide a starting point for further investigation.

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

I would like to conclude by quoting from Chris Daykin's contribution to the discussion of Wilkie's (1995) paper on stochastic asset modelling:

> "I believe that stochastic modelling is of fundamental importance to our profession. How else can we seriously advise our clients and the wider public on the consequences of managing uncertainty in the different areas in which we work? It is important for all actuaries to come to grips with this type of modelling work, and we have much to learn about the different ways in which such modelling can be approached."

Reference

Wilkie, A.D. (1995). More on a stochastic asset model for actuarial use. *B.A.J.* **1**, 777-964.

The President (Mr P. N. S. Clark, F.I.A.): Thank you very much Dr England. I am delighted — and I am sure that you are — that there has been such a good attendance at this discussion. I am delighted that a number of younger members have contributed to the discussion, including at least one to whom I have fairly recently given a Fellowship certificate. I am encouraged by the degree of debate as well. Mr Tripp was right to highlight the problems of a silo mentality. You may have noticed a note passing along the table here. It was an encouraging development, because the pensions actuaries sitting either side of me are trying to work out now how to find a stochastic replacement for the minimum funding ratio. So, some good may come out of this paper in some very interesting directions.

What follows is clearly not going to be a deep technical comment for the reasons hereto mentioned, but I think that what this paper has done, and what the discussion has done, are to highlight that risk is one of the five key issues that the actuarial profession is looking at. Of course, risk is highlighted here in the paper and in general insurance in particular — three types of risk: investment risk, underwriting risk and reserving risk. Perhaps we spend a lot of time in other arenas talking about investment risk and underwriting risk, and it is right to focus for a time on reserving risk.

Of course, risk is not just a key theme for the actuarial profession; it is also a key theme for our regulator, and so I — and others here have already expressed it — am very pleased with the positive welcome that Mr Sharma gave to the paper. We are certainly keen, as a profession, to contribute to this ongoing debate.

There are various other things that we are looking at, and which were picked out by speakers. Clearly, it is vital to focus on, look at and discover the data. It is also vital to have the proper exercise of judgement in understanding both the data and the results, and then, finally, vital to have credibility. I was delighted that the closer recognised that I have a theme about communication. Communication and credibility are very much part and parcel of this, because it is so easy for actuaries to bring out a black box. There is so much here that could be described as a black box. Actuarial black boxes are not too much in vogue, unless they can be given credibility and unless we can communicate the results effectively. That is a challenge in all areas of actuarial work: in general insurance; in life insurance; in pensions; and in finance and investment. I was interested that Mr Sharma even wants that help as he tries to convince some of our continental European colleagues of the effectiveness and of the use of these kinds of models. So, there is a challenge there for us to have that credibility and to be able to communicate it effectively.

I ask you to show your appreciation to the authors for producing such a worthwhile paper and for stimulating such a positive debate.

WRITTEN CONTRIBUTIONS

Dr G. Barnett, Dr E. Dubossarsky and Dr B. Zehnwirth, A.I.A. [an abridged version of this contribution was read to the meeting]**:** We strongly support the paper's arguments for the

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

importance of a predictive distribution for reserves that incorporates both parameter risk and process variance. This predictive distribution plays a very important role in risk based capital calculations and dynamic financial analysis (DFA), especially as the reserving risk accounts for most of the company's capital requirements. However, there is much that we strongly disagree with, summarised in the following two paragraphs.

The paper's primary focus appears to be on the computation of a predictive distribution, based on one of a few arbitrary probabilistic models fitted to the paid losses, when the model does not capture the structure in the data. We argue that the focus should, instead, be on a probabilistic modelling framework that is used to identify a parsimonious model that captures accurately the salient features (structure) of the data. We disagree with many of the statistical arguments in the paper. Predictive distributions based on models that do not accurately reflect the structure in the data can be misleading and dangerous.

Paid loss development arrays satisfy certain axiomatic trend properties. See Course 7 of the Society of Actuaries (U.S.A.). The models considered in the paper ignore these trend properties, and therefore cannot capture the features (structure) of most real paid loss development arrays. For the example real data analysed in the paper, the chain-ladder models (e.g. log-linear Poisson) that are fitted are inadequate, and have no predictive power, as previously established in the literature. The critical issue of possibly changing superimposed or social inflation is explicitly omitted (¶10.15). It is vital to recognise that making multi-million dollar decisions, without knowing the superimposed inflation trends in the business, can lead to insurance company collapse. Equally, DFA based on such models can provide false indications. There are also omissions of previously published work; papers are cited, but highly relevant content ignored. In addition, there is a strong reliance on Wright (1990)'s Hoerl curve model with variance proportional to the mean. We believe that this model is flawed in several ways.

We distinguish between a probabilistic model and a probabilistic modelling framework. In fact, we believe that the principal reason for a probabilistic modelling framework is that it is only in such a framework that the adequacy and appropriateness of models can be properly assessed. Just because the model is probabilistic does not imply that the mean is right, let alone the variance. Compelling reasons for identifying (designing) an optimal probabilistic model for insurance data that captures the structure in the data are given by the celebrated American actuary Arthur Bailey (1942).

Zehnwirth (1994) explains the axiomatic trend structure satisfied by paid loss development arrays. This paper is used as a study guide by the Society of Actuaries (U.S.A. http://www.soa.org/research/lew_appendix_a.html) in its teaching programme. This paper is not cited, although it is fundamental to Barnett & Zehnwirth (1998), which is cited, but much of its content appears to be ignored. The identified optimal probabilistic model for a loss development array captures the trend structure and randomness in the data in such a way that a simulated triangle from the model has the same salient features as the real data. In other words, one should not be able to distinguish between the real data and the simulated data in respect of salient features. This is explained in Barnett & Zehnwirth (1998), and the subsequent revision, Barnett & Zehnwirth (2000). For a review of the latter by Fred Cripe (Chairperson, CAS Research Policy and Management Committee), see http://www.casact.org/pubs/actrev/may01/latest.htm .

The models discussed in the paper do not capture the features (or structure) of most real paid loss development arrays, and some of the models impose a structure that is extremely rare in real data. Most importantly, we can almost always distinguish between the real data and a triangle simulated from any one of the fitted models. If a model does not describe the essential features in the data with a parsimonious set of parameters, the statistics based on the model are meaningless; this even includes the estimate of the mean. It is also important that the parameters of the model have a simple interpretation.

Models that quantify *changing* superimposed inflation trends are explicitly omitted (see ¶10.15), as the paper comments: "Attempting to model claims inflation in each year individually is usually problematical, due to the number of parameters in the model and dependencies with the origin period and development period." If the models did not have so many parameters in the accident period and development period directions, there would be scope to introduce payment (or calendar) period parameters. A critical point here is that basing decisions on models that cannot capture major shifts in *superimposed* inflation trends is very dangerous. Not knowing the changing superimposed inflation trends in one's own business is a primary cause of insurance company failure or delayed major reserve upgrades. See the Zehnwirth paper on the HIH failure in a global context at http://www.insureware.com/Library/HIHDebacle.pdf .

On this very point, in respect of DFA ¶9.4 places emphasis on "...an economic scenario generator in a sensible way''. Firstly, one needs to quantify the impact on the business of past economic inflation. Secondly, *superimposed* inflation is often a much more important and critical issue than economic inflation for many lines of business (including workers' compensation). A sound DFA analysis cannot be conducted if the underlying model is not informative about what is going on in the business.

In ¶2.2.2 it is noted: "...and various extensions are suggested by Barnett & Zehnwirth (1998).'' A loss development array is analysed using ratios and related models, even though Venter (1998) and Barnett & Zehnwirth (1998, 2000) show that, for these very same data, ratios have no predictive power. If we graph incrementals in a development period versus cumulatives in the previous development period, we see that there is no correlation. A superior (but not best) model for these data is estimated by simply computing the mean of incrementals in each development period. The lack of predictive power demonstrated for the chain ladder also applies to the log-linear Poisson, as it and the ratio model give the same (out-of-sample) predictions. Venter (1998) provides tools for finding out whether ratios work for a given data set. Barnett & Zehnwirth (1998, 2000) extend Mack (1993, 1994a, 1994b) and Murphy (1994) type models, to provide tools to allow the actuary to determine whether averages of, or trends in, incrementals, in a development period, have more predictive power than ratios applied to cumulatives. Comments (¶2.1.1) regarding interchangeability of cumulatives and incrementals for modelling are not true in general, as changing superimposed inflation trends in incremental paid losses cannot be modelled from the cumulative data.

The log-linear Poisson model, Kremer's model (1982), Mack's model (1993) and the negative binomial model should all be abandoned, and no more needs to be written about them. These models have no structure effectively relating the different years. For the Kremer (1982) model and log-linear Poisson models, residuals versus accident years and development years exhibit no trends. The residuals in each period sum to zero. So, they have some structure! Accordingly, the structure in the data in the payment (calendar) period direction is hidden in the residual graphs of the other two directions. For the arithmetic average ratio method ($\delta = 2$ in Murphy, 1994; and Barnett & Zehnwirth, 1998, 2000), the residuals sum to zero in each development period.

Dannenburg, Kaas & Usman (1998) show that the chain-ladder technique amounts to simple arithmetic, giving the same incremental forecast, whether applied to an incremental triangle or to its transpose! One can regard development periods as accident periods and accident periods as development periods. Since chain ladder is equivalent to log-linear Poisson, that is, in turn, equivalent in structure to two-way ANOVA, it does not matter whether you cumulate the incrementals for each accident period across the development periods or cumulate the incrementals for each development period across the accident periods. The fact that the latter cumulative values are not particularly meaningful (also) indicates that the chain ladder is completely uninformative in respect of the structure in the data. The Mack model corresponding to cumulatives for each development period would condition on the first accident year (and would regard it as fixed!). Incidentally, contrary to popular belief, the Mack model has the same number of mean row parameters as the log-linear Poisson. This can be seen by writing out the full log likelihood as a sum of log conditional likelihoods (just as you do with an AR(1) process, say). It also has many additional (independent) variance parameters.

The paper relies heavily on the model presented in Wright (1990). (See Section 4.3.) In Wright (1990), a Hoerl curve model is derived with the variance proportional to the mean for the paid losses using risk theoretic arguments. Notwithstanding the fact that the derivation of the Hoerl curve contains a lacuna, the property that the variance is proportional to the mean violates the structure found in most, if not all, real paid loss development arrays. If the variance is proportional to the mean (on a dollar scale), then, on a log (percentage) scale, the variance is proportional to the reciprocal of the mean. Accordingly, for any data that are subject to an increasing inflation trend, the variance should reduce on a log scale. This is contrary to experience with actual paid data, where the variance is usually constant on a log scale, or occasionally increasing in later development years (i.e. when it is not constant, it is, if anything, moving in the opposite direction to that assumed in the model in Wright, 1990). Indeed, it does not even appear to be true for the data modelled in Wright (1990); see his Figure 8.2d. Moreover, we have found through experience (as early as 1988) that the Hoerl curve is not flexible enough to capture adequately most run-off patterns, and its shape parameters are not easily interpretable. England & Verrall (2001) and Wright (1990) claim that the Wright model can be fitted to negative data. This is unlikely to be the case if several contiguous cells have a mean that is negative, since the curve is always positive and the variance is proportional to the mean. For this reason, we believe that Wright (1990) Section 8.2 omits the 1978 accident year altogether, as the mean of the values in the last six development periods is −506! There are many other issues with Wright (1990). For example, in the data analysed therein, the mean and variance of the severity (individual claim size) distribution is obtained as fixed over the whole triangle (see Section 8.1, p 708). It is well known that this cannot be true for most, if not all, long tail liabilities. England & Verrall (2001) rely even more heavily on Wright (1990), and a comprehensive critique of Wright (1990) is in progress.

In ¶4.3.8, in comparing Zehnwirth (1985) with Wright (1990): "A useful critique of the differing assumptions concerning the distributional assumptions can be found in Appendix 4 of Wright (1990)." Notwithstanding that Zehnwirth (1985) has been superseded by Zehnwirth (1994) and many other papers, arguments in Appendix 4 of Wright (1990) are also flawed.

In ¶8.3.5, in relation to predictive distributions, we believe that the contribution of Barnett & Zehnwirth (1998, 2000) is totally misrepresented. In particular, it says: "it is not clear whether the two-stage procedure outlined in this section is adopted, or the estimation error alone is considered." The notion of a predictive distribution (incorporating both process variance and parameter estimation error) is discussed in Barnett & Zehnwirth (1998, 2000) and in Zehnwirth (1994). There is no need for the two-stage simulation procedure as outlined in the paper (in ¶8.3.2), since the predictive distribution for each cell can be computed analytically. The only reason why simulation is required is that the predictive distribution of *sums* of forecasts must be found numerically (e.g. via numerical convolution or via simulation), since it is not possible to do this algebraically. The paper by Dickson, Tedesco & Zehnwirth (1998), cited in Barnett & Zehnwirth (1998, 2000), emphasises the dangers of excluding parameter estimation error.

Paragraph 3.2.5 refers to the Kalman filter, and Verrall (1989): "The Kalman filter allows, for example, a way of smoothing the row parameters, instead of treating the rows as completely separate. In general, the Normal distribution is easier to use in this context, and there are many possibilities for imaginative use to be made of the existing theory." See also ¶4.2.3 for a similar comment. In fact, the Kalman filter, itself, is simply an efficient method of obtaining estimates of the state vector and its variance-covariance matrix, conditional on the data. One should not confuse a model with a weighted least squares algorithm for estimating its parameters. Moreover, it is not true that normality is necessarily required for the Kalman filter; see, for example, Zehnwirth's (1988) generalisation of the Kalman filter.

Verrall (1989) is another interesting case in point, where structure is imposed in a fitted model contrary to the structure in the data. The accident periods' individual level parameter estimates are increasing linearly (see his Figure 4.1), yet it is assumed by the empirical Bayes (Buhlman-Straub) model that the level parameters are random from the same distribution. It is also an example where the Kalman filter is used to smooth accident periods' individual level parameter estimates arbitrarily. (The variance in the system equation is chosen arbitrarily.) However, a single level parameter, with an additional single payment period parameter, would suffice here, and be far superior.

When there is a common trend in individual level parameter estimates, capturing that single trend with a simple trend parameter is superior to smoothing separate parameters for each year. For example, the smoothed values are biased toward the mean (a weighted mean of parameters), reflecting the middle years, which is especially bad for projection into the future. The more advantage is taken of the benefits of smoothing, the worse the forecasts are in the presence of a trend. The way to take proper advantage of smoothing is to model the main trends first.

Bootstrapping (Section 8.2) with a faulty model does not replicate the distributions in each cell in the triangle. Indeed, it can make an overparameterised model appear much better than it really is. We have already mentioned that log-linear Poisson and Kremer's (1982) models have residuals that sum to zero for each development period and each accident period. Bootstrapping these residuals is similar to bootstrapping from a sixth degree polynomial fitted to data with only a linear trend. If a parsimonious parametric model carries the same information as the data, there is no need for bootstrapping. The distributions can often be computed analytically.

Regarding the comment in ¶3.1: "the log-Normal distribution ... should never be used", the effect of the last sentence, in particular, seems to be a way to introduce a note of caution when none existed beforehand. There are good *a priori* reasons to believe that a log-Normal should be a good approximation to the distribution of a paid loss in a cell. (The log-Normal distributions for each cell are related by trend parameters.) Heuristically, it seems that the underlying generating process of the data is multiplicative; that is that there are numerous multiplicative influences in the process (for example, the economic influences on claim payments are mutiplicative, and the trends are consequently exponential), and invoking a central limit theorem leads to the log-Normal assumption, but it still has to be tested! Indeed, analysis of residuals indicates that the log-Normal is usually a good model, as long as the means are appropriately modelled and allowance is made for any heteroscedasticity in the development periods. Accordingly, the log transform is preferred for at least two reasons: to stabilise the variance, as the variance is generally proportional to the mean squared (as it will be in the presence of multiplicative influences); and to measure trends, which is an essential feature of the data.

Markov chain Monte Carlo is discussed in Section 8.4: "8.4.1 The method of simulation from the parameters outlined in the previous section is reasonable. Even then, the distribution may not be recognisable as a standard one. Markov chain Monte Carlo methods obviate this problem using simulation and numerical methods, and consider estimation and prediction at the same time within the same framework." and: "8.4.2 MCMC methods do not provide parameter estimates *per se*, but a simulated joint distribution of parameter estimates." In fact, MCMC provides simulations from the joint posterior distribution of parameters (not estimates), conditional on the data. From this, a posterior marginal mean or mode might be used as an estimate, if one was required, or an interval estimate could be produced.

Further, in ¶8.4.3 it says: "Unlike the methods based on GLMs (and Mack's method), which focus on modelling the mean, then investigate predictive distributions, MCMC methods provide the predictive distribution, from which summary statistics, such as the mean, can be calculated. As such, there are no extra steps to perform, or approximations to be made when considering the predictive distribution." Indeed, further sources of error can be included. For example, one may consider a framework where there is a flexible family of models with parameters that may be either included or not (i.e. set to zero), and in this way include model selection risk with process and parameter risk. (See, for example, Barnett, Kohn & Sheather, 1997.)

References

Bailey, A. (1942). Sampling theory in casualty insurance. *Proceedings of the Casualty Actuarial Society,* **XXIX**.

Barnett, G., Kohn, R., & Sheather, S. (1997). Bayesian estimation of an autoregressive model using Markov chain Monte Carlo. *Journal of Econometrics*, **74**, 237-254.

Barnett, G. & Zehnwirth, B. (2000). Best estimates for reserves. *Proceedings of the Casualty Actuarial Society*, **LXXVII**. 245-321.

Dannenburg, D.R., Kaas, R. & Usman, L.N. (1998). Gegeneraliseerde lineaire modellen voor IBNR-driehoeken. *Het Verzekerings Archief*, Jaargang **75**, 4e kwartaal 149-158.

Dickson, D.C.M., Tedesco, L. & Zehnwirth, D. (1998). Predictive aggregate claims distributions. *Journal of Risk and Insurance*, **65**, 689-709.

Venter, G. (1998). Testing assumptions of age-to-age factors. *Proceedings of the Casualty Actuarial Society,* **LXXXV**.

Zehnwirth, B. (1988). A generalization of the Kalman filter for models with state-dependent observation variance. *Journal of the American Statistical Association,* **83***,* 401, 164-167.

Zehnwirth, B. (1994). Probabilistic development factor models with applications to loss reserve variability, prediction intervals and risk based capital. Casualty Actuarial Society, Spring Forum.

Professor T. Mack, Hon.F.I.A. [an abridged version of this contribution was read to the meeting]**:** The authors have to be congratulated for giving a comprehensive review of stochastic reserving models for the chain-ladder and similar methods. Throughout, they do this on the background of generalised linear models (GLMs), and therefore the paper is particularly useful for actuaries with a good understanding of GLMs. For this reason, I would suggest to indicate this fact in the title of the paper. Actuaries with a rather classical frequentistical education in stochastics may have some difficulties, e.g. with an over-dispersed Poisson (ODP) variable for which it is unclear what its domain or its distribution is. This type of variable is used by the authors in their initial model.

Several other models are reviewed, and put into the framework of GLMs. Each time another model is presented, the practitioner usually will ask how he could decide which model to choose. Not much can be found in the paper as guidance regarding model selection; but this is, indeed, a difficult topic. A first idea can be found by precisely looking at the model assumptions. Reading between the lines, one will find that, essentially, two types of models are described:

(A) models assuming independent increments, like ODP (Section 2.3), log-Normal (Section 3.2), and Gamma (¶3.3); and

(B) models of a recursive structure, where the increments of each accident year are not independent any more, like negative binomial (Section 2.4), Normal approximation (Section 2.5) and Mack (Section 2.6).

The experienced actuary usually has an *a priori* opinion regarding independence of successive increments of his business. This independence is a strong assumption. It may be realistic in the case of pure IBNR claims, as in mortgage indemnity business, but will rarely hold for IBNER claims, as in motor business.

Looking at type A models, we see from Section 2.3 and ¶3.3 that the increments are assumed to be of a multiplicative cross-classified structure which can be linearised with a logarithmic link function. This leads to the GLM form:

$$E(C_{ij}) = m_{ij} = x_i y_j \quad \text{Var}(C_{ij}) = \phi V(m_{ij})$$

with variance function V, e.g. $V(m) = m^k$, where $k = 1$ gives ODP and $k = 2$ gives Gamma. Of course, these can easily be extended to the inverse Gaussian case $k = 3$ and the Normal case $k = 0$. The latter one, which was introduced by DeVylder in 1978, very simply solves the problem with the negative increments that one encounters with ODP, Gamma and inverse Gaussian.

All these models of type A are well known from the analysis of cross-classified data in rate-making. The fact that we now have only a triangle instead of a rectangle does not cause any problem regarding parameter estimation, except for the fact that we have fewer data. However, the fact that we have no measure of volume, i.e. assume the same volume in each cell, is a clear disadvantage in comparison to the rate-making situation. Taking the strong independence assumption and the missing volume together, we come to the conclusion that type A models do not seem to be good models. In addition, I have seen two unpublished papers from Australian actuaries that prove that the reserve estimates in the ODP model are biased.

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

These disadvantages of type A models are not encountered with type B models. These use the previous cumulative claims $D_{i,j-1}$ as a measure of volume for cell (i, j) and are therefore of a recursive structure, i.e. we work — in the very spirit of chain ladder — from one column to the next, assuming each time that the amounts of the previous column are given. Therefore, the increments of a fixed accident year are not independent any more. For this reason, one cannot formulate an overall GLM for the whole rectangle, we merely have a GLM in every column (development year). This formal difference has not been clearly set out in the paper. Moreover, for type B models we can exhibit the role of $D_{i,j-1}$ as volume more clearly if we consider the individual incremental development factors $F_{ij} = C_{ij}/D_{i,j-1}$. Then we can immediately apply the GLM models in the form:

$$\mathrm{E}(F_{ij}) = \lambda_j \quad \mathrm{Var}(F_{ij}) = \phi_j V(\lambda_j)/D_{i,j-1}.$$

Using e.g. the variance function $V(\lambda_j) = \lambda_j(\lambda_j - 1)$, we obtain the negative binomial model in a slightly more general form than the authors, who assume the same $\phi_j = \phi$ for all development years j. From their negative binomial model, the authors proceed in a somehow vague manner to what they call normal approximation to the negative binomial model. In our GLM formulation, we obtain this model directly by simply choosing $V(\lambda_j) = 1$. Further standard possibilities — not being treated by the authors — include another over-dispersed Poisson with $V(\lambda_j) = \lambda_j$ and the Gamma case $V(\lambda_j) = \lambda_j^2$, both being different from the corresponding cases of type A models. Especially attractive is the inverse Gaussian case $V(\lambda_j) = \lambda_j^3$, because here the maximum likelihood estimate of ϕ_j can be given explicitly:

$$\widehat{\phi}_j = \frac{1}{n+1-j}\sum_{i=1}^{n+1-j} D_{i,j-1}(F_{ij}^{-1} - \lambda_j^{-1}).$$

To me, it is not clear why the authors did not consider these latter models, especially because also these models, like all the above recursive GLM models, give the same reserve estimate as the chain ladder algorithm. Moreover, all calculations are simple, and can be done in a spreadsheet without use of GLM software. For example, for all recursive type B models the following simple and intuitive recursion for the mean squared error of prediction holds:

$$\mathrm{MSEP}(\hat{C}_{ij}) \approx \left(\mathrm{MSEP}(\hat{C}_{i,j-1})\right)^2 \hat{\lambda}_j^2 + \hat{D}_{i,j-1}^2\left(\mathrm{Var}(F_{ij}) + \mathrm{Var}(\hat{\lambda}_j)\right)$$

with starting value $\mathrm{MSEP}(C_{i,n+1-i}) = 0$, and

$$\mathrm{Var}(\hat{\lambda}_j) = \phi_j V(\lambda_j) \Big/ \sum_{i=1}^{n-j} D_{i,j-1}.$$

This recursion is taken from a paper by Mack which appeared in 1999 in the *ASTIN Bulletin*.

The authors did not make a clear distinction between a model and the way its parameters are estimated. Otherwise, they would have realised that their negative binomial model, as well as its Normal approximation, are both special cases of Mack's model, which assumes:

$$\mathrm{E}(F_{ij}) = \lambda_j \quad \mathrm{Var}(F_{ij}) = \sigma_j^2/D_{i,j-1}$$

and one clearly sees that it is identical to the above GLM formulation, except that the variance is slightly more general. Mack's model is distribution free, but — of course — each distribution-free model can be specialised further by assuming any specific distribution. Having selected a distribution, we have additional possibilities to estimate the parameters, e.g. maximum likelihood, which is employed by GLMs. If I had to choose a distribution here, I would prefer the Gamma or inverse Gaussian over such a strange thing like the overdispersed negative binomial.

One possibility to find the distribution which best fits the data is to take the one with has the

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

highest likelihood (as long as the same number of parameters is used). Again, this only works as long as classical distributions are used, and not over-dispersed ones.

Regarding the question of how to arrive at the full distribution of reserves, the authors have made a big step with the bootstrap approach. It enables us to assess the precision of approximative procedures. For example, we can fit a log-Normal distribution to the mean and standard deviation of the overall reserve of the ODP model, as given in Table 3, and then compare its percentiles to the bootstrap percentiles given in Table 34. In this case, we will find a good agreement between the bootstrap and the log-Normal approximation. In practise, we can do this once for each major business and then decide where we can content ourselves for the next years with the much simpler log-Normal approximation. My impression is that the difference between bootstrap and log-Normal approximation usually matters less than the impact of different reserving models.

My main conclusions are:

(1) The recursive type B models look much more reasonable than the cross-classified type A models.
(2) There are some more recursive GLM models than those considered in the paper, especially the inverse Gaussian model, which should be taken into consideration if one wants to choose a distribution and does not have negative increments or can ignore these.
(3) All recursive models are special cases of Mack's model, which can be chosen if one does not want to select a distribution or if negative increments cannot be ignored.

I am grateful to the authors for their interesting paper, which has stimulated me to put the various models into a GLM framework, which, indeed, helps to clarify similarities and dissimilarities.

Professor G. C. Taylor, F.I.A. [an abridged version of this contribution was read to the meeting]**:** I have read the paper with admiration. The authors cover much ground very concisely. A minor quibble might concern the almost total absence of adaptive filtering, with just passing reference to the Kalman filter. However, apart from this, there seems little to add to their very competent review. I now make a handful of observations, one technical concerning over-dispersed models, the remainder largely historical.

Over-dispersed models

These seem to me to be under-parameterised when applied to paid or incurred losses. I shall illustrate this by reference to the Poisson model, but the argument applies *a fortiori* to the negative binomial.

At the heart of the ODP model is the assumption that each cell's variance is a multiple of its mean, and that the *multiple is the same for every cell*. Routine calculation shows this to imply that, for individual claims, the coefficient of variation of claim size is inversely related to the mean claim size.

For the ODP model, the coefficient of variation of C_{ij} is:

$$v[C_{ij}] = (\varphi/m_{ij})^{1/2}. \tag{1}$$

Consider the case in which each cell is actually compound Poisson distributed:

$$C_{ij} = \Sigma_k X_{ijk} \quad k = 1, 2, \ldots, N_{ij} \tag{2}$$

with $\mathrm{E}[N_{ij}] = \lambda_{ij}$ and a_{ijr} denoting the r-th uncentralised moment of the X_{ijk}. Then:

$$m_{ij} = \lambda_{ij} a_{ij1} \tag{3}$$

$$\mathrm{Var}[C_{ij}] = \lambda_{ij} a_{ij2} \tag{4}$$

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

which yields (1) if one sets:

$$\varphi = a_{ij2}/a_{ij1} = a_{ij1}\{1 + v^2[X_{ijk}]\}. \tag{5}$$

Constant φ implies that the coefficient of variation of claim size is inversely related to the mean claim size.

To take a TPBI portfolio as an example, this means that, as one moves from low to high development years, with increasing average claim size, the distribution of claim size, relative to the average, becomes tighter. In my experience, this is the opposite of what one observes.

One can correct this situation by adding further parameters to enable the variance to depend on development year. Unfortunately, this would destroy the Poisson structure.

If one accepts the above TPBI argument, then (1) is under-parameterised. It requires correction to the following:

$$v[C_{ij}] = (\varphi/w_j m_{ij})^{1/2} \tag{6}$$

where the w_j effectively constitute a set of weights:

$$w_j = \varphi a_{j1}/a_{j2} \tag{7}$$

in the case where $a_{ijr} = a_{jr}$, independent of i.

Unfortunately, this translates back to:

$$\text{Var}[C_{ij}] = \varphi m_{ij}/w_j \tag{8}$$

which is similar to the model in ¶4.3.7, but no longer recognisable as Poisson related.

Historical comment

I thought that the paper under-represented the early development of stochastic models a little. For example, Dr Harry Reid's (1978) paper is not mentioned. I recognise that paper as not sitting easily within the theoretical framework with which the authors work, but, after all, it is, I think, the first paper on stochastic reserving.

I would nominate three specific areas in which the historical perspective presented in the paper appears unnaturally foreshortened. These are:

— the use of GLMs;
— credibility procedures; and
— bootstrapping.

A reader gets the impression from the paper that GLMs were unheard of in loss reserving until Renshaw's (1989) paper. It is evident from the paper of Taylor & Ashe (1983) that the software package GLIM was in use then (though that paper refers to only a trivial (normal) application).

GLMs are also discussed in Ashe's 1986 paper, where non-Normal applications are mentioned, and by Taylor (1988). Certainly, Frank Ashe and I were using GLMs in our loss reserving consulting from the early 1980s.

The interpretation and generalisation of Bornhuetter-Ferguson reserving in terms of credibility theory, in Section 6.3, is extremely elegant. The book of mine that the authors were kind enough to cite attempted to place this type of reserving, and some other related types (e.g. Cape Cod), in a credibility context. However, this was done at a rather superficial level. The authors here have fleshed it out, and added insight to the true meaning of these widely used methods.

I mention that the same book contains an entire chapter covering a larger collection of credibility methods of loss reserving, though with a focus somewhat different from the authors'.

That chapter provides several references and connections to the European literature not mentioned in the present paper.

The earliest reference to bootstrapping in the paper appears to be England & Verrall (1999). In fact, this technique was introduced into the loss reserving literature in 1986 by Ashe's paper. It is discussed in a little detail in Taylor's (1988) paper, also referred to earlier.

References

Ashe, F.R. (1986). An essay at measuring the variance of estimates of outstanding claim payments. *ASTIN Bulletin*, **16S**, 99-113.

Reid, D.H. (1978). Claim reserves in general insurance. *J.I.A.* **105**, 211-296.

Taylor, G.C. (1988). Regression models in claims analysis: theory. *Proceedings of the Casualty Actuarial Society*, **74**, 354-383.

The authors subsequently wrote: We would like to thank everyone who contributed to the discussion of this paper. We found the comments to be interesting and informative, and we are sure that the discussion adds greatly to the value of the paper.

As we said in ¶1.2, the purpose was to review some stochastic reserving methods and to explore some of the practical implications of their use. Subjects such as this can be very technical and sometimes difficult to explain, so we were very glad for the welcome the paper received and for the positive comments about its readability. Obviously, it was not possible to include everything that has been written on the subject (as mentioned in ¶1.7); several discussants have added comments on sources which were not included, but which they felt are relevant.

It is clear from the discussion that stochastic reserving has become an issue of great relevance because of the proposals for 'fair value accounting' and the forthcoming risk-based approach to regulation taken by the Financial Services Authority.

Although the methods described in the paper are not suitable for all types of data, it is our belief that the most sensible approach to understanding stochastic reserving methods is to start with the most familiar technique (the chain-ladder method), then to build on that to other methods that overcome its weaknesses.

Mr Sanders makes many good points in his opening to the discussion. We would agree with him that the stochastic approach has a real advantage in making actuaries aware of the uncertainty in a reserve estimate. Along with others (Messrs Tanser, Craighead, Bruce, Christophides, Johnson, Ryan), the opener points out that a method such as the chain-ladder technique cannot be applied blindly (if at all) to all sets of data that are likely to arise. Lloyd's and London Market data were cited as examples where there are likely to be large extrinsic influences in the data which will not be captured by the chain-ladder technique. The Bayesian approach was mentioned as a potentially useful way to allow for these influences, and we hope that further developments in this area will be forthcoming.

Messrs Christophides, Bruce, Malone and Shah and the opener all discuss dynamic financial analysis (DFA) as the natural next step on from stochastic reserving. Certainly, reserving risk should be considered when analysing insurance liabilities on a stochastic basis, and we are glad that the Australian regulators require risk margins for reserves to be set, based on a risk measure from a distribution of outstanding liabilities. We expect that a similar approach will be adopted more widely in other jurisdictions.

A number of discussants point out that the prediction errors may be larger than desirable practically. Reasons given for having lower values are that there is other information that should also be considered, and that the company would apply management controls to mitigate the risk. We would agree that this is an important issue, and we expect that a Bayesian approach, along the lines of that described in Section 6, might allow for the inclusion of this other information. However, a note of caution should be sounded about the degree to which prediction errors could be decreased, especially on an arbitrary basis. For example, Mr Tanser points out that stochastic models may understate the true uncertainty, due to model error.

We agree with Professor Mack that the approach taken in the paper reflects our background

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

in GLMs, and indeed we had a U.K. general insurance audience in mind when writing the paper, since they will be familiar with GLMs from premium rating for personal lines. We do not regard the distinction between the two classes defined by Mack, A and B, as being of particular importance. In fact, the same predictive distributions can be obtained from the over-dispersed Poisson (ODP) and the recursive negative binomial models. For us, the important aspect of this is the added understanding that can be gained by looking at the stochastic reserving methods in a number of different ways. For example, it allowed us to move to the Normal approximation, and thereby show how the results of Mack's model are related to those of the recursive negative binomial model, and hence to the ODP model (up to bias correction).

We agree with Professor Mack that there are other distributional assumptions that we could also have included, for example the inverse Gaussian model, and Professor Mack has suggested some interesting alternatives that can be investigated.

We do not agree that all recursive models are special cases of Mack's model. It is clear from his own description that his original 'distribution free' method considers the special case where, in his terminology, $V(\lambda_j) = 1$, which is the variance function for the normal distribution. The choice of variance function dictates the form of the residual required in the calculation of the dispersion parameters ϕ_j, and, again, Mack's original model uses a residual definition appropriate to the Normal distribution. This does not detract from the valuable contribution made by Professor Mack in his original model; we believe it simply provides clarification.

Mr Christophides makes the important point that we have concentrated on reserves for individual triangles. In fact, we believe that that is also true for most investigations of stochastic reserving which have appeared previously. As he points out, it would certainly be worthwhile investigating a set of triangles within a company to exploit any possible correlations when estimating the reserving risk. A similar point is made by Mr Reynolds, but in regard to market information in respect of other companies.

Drs Barnett, Dubossarsky and Zehnwirth welcome the use we make of the predictive distribution, but state that there is much they disagree with in the paper. We find it difficult to reconcile their comments with the content of the paper, but it would appear that they take particular exception to the emphasis given to the chain-ladder technique. We repeat that we view the chain-ladder technique as a starting point, not as the correct method in all circumstances, and we recommend that other models are also investigated, which will provide a greater understanding of the characteristics of the models and the underlying data. The range of models that we recommend include the models favoured by Barnett, Dubossarsky and Zehnwirth, but are certainly not limited to them. We are disappointed that these discussants are so vehemently opposed to our paper, whereas we would hope that it makes the models that they favour more accessible to practitioners. A response to all of their criticisms would fill many pages in this journal, and is unlikely to be of interest to most readers, but some comment seems necessary. Barnett, Dubossarsky and Zehnwirth state, more than once, that we: "rely heavily on the model presented in Wright (1990)." We do not; it is simply one of many models that we discuss. Barnett, Dubossarsky and Zehnwirth also suggest additional papers that they consider to be relevant, and clearly would have approved if we had given greater emphasis to some sources that were included. We repeat that it was not the purpose of the paper to provide a complete survey of the literature (¶1.7), although we did provide a reasonably representative list of references, which surpasses similar lists in most other papers in the literature by a considerable margin. Regarding the comment made by Barnett, Dubossarsky and Zehnwirth on ¶3.1, they have misquoted us when they quote: "the log-Normal distribution ... should never be used". What we actually say is: "This does not necessarily mean that the log-Normal distribution should never be used." It is our experience that models based on log-Normal distributions can be useful, but have considerable inherent difficulties. In particular, the question of negative incremental claims becomes critical. Perhaps of more importance is the inclusion of a component of the variance in the expected value on the untransformed scale (see equation 7.17), which can cause the results to be very unstable; the other models that we discuss do not have this inherent characteristic. Barnett, Dubossarsky and Zehnwirth's comments regarding bootstrapping are erroneous; the

https://doi.org/10.1017/S1357321700003810 Published online by Cambridge University Press

residuals for the log-linear Poisson model do not necessarily sum to zero across development period and accident period, and bootstrapping (as a method) does not make an over-parameterised model appear better than it really is (when the procedure is applied correctly, it replicates the estimation variance that would have been obtained analytically from the same model).

We agree with Professor Taylor that the Kalman filter could have received more attention, and we would encourage readers of the paper to use it as an introduction to the subject, and to explore the literature for useful additional methods that can be applied in this context, including dynamic models such as those using Kalman filter ideas. Along the same lines, it is certainly the case that modelling the dispersion parameter should also be considered, although this adds another layer of complexity to the approach. We welcome the additional references supplied by Professor Taylor, and are glad to acknowledge the significant contributions to the literature made by him and also by Mr Ashe.

Ms Cresswell summed up admirably some of the issues and questions that arise from the use of stochastic models, and of the proposed new accounting guidelines. We hope that the paper will help researchers begin to approach the issues raised.

Bootstrapping has been found by practitioners to be particularly attractive as a relatively easy way to obtain a predictive distribution. Several discussants have commented on this, and Mr Archer-Lock makes a number of very good points. Again, the issue of the 'correct' amount of variability arises, with the suggestion that it may be necessary to put back extra variability if a 'Bayesian' type approach is taken within a bootstrapping exercise. This is an important issue, and we look forward to further light being shed on it in the future.

Finally, we would like to thank the referees for their valuable comments, and the Institute of Actuaries for the opportunity to present our paper at a sessional meeting.