Variance in Claim Reserving

Simon Brickman (Chairman)
Caroline Barlow
Anthony Boulter
Andrew English
Lloyd Furber
David Ibeson
Julian Lowe
Richard Pater
David Tomlinson

Summary

This paper is educational. There has been considerable development of statistical techniques for predicting claim payments over recent years, which has yet to be assimilated by the profession and put into practice by practitioners.

We want to spread the knowledge of these techniques, to dispel some of the mystique, and to give some examples which demonstrate how they work in practice. These techniques do not replace existing methodologies, but serve to enrich the actuary's tool box.

So that readers can form a view of the success of existing methods in the past we include a review of the variance of the actual out-turn from the reserves of some UK insurance companies over the past ten or so years, to which we add some thoughts on the factors which may have contributed to the variances.

We would like to encourage a healthy scepticism of "black box" techniques and some of the pitfalls for the unwary are presented as a warning against using them without an understanding of the limitations. For example we believe that the use of the term '*Confidence intervals*' is to be discouraged since we think it conveys a false impression of the modelling process, which applies to past data. The circumstances that will apply in the future can not be known at the moment, so the model is emphatically not a crystal ball.

A bibliography is included for the reader who has been encouraged to pursue the subject further.

Terms of Reference

- *(Briefly) identify areas of application for the claim reserve variance*
- *Review past reserve variance by company and class using DTI Returns*
- *Review selected Statistical Reserving Methods*
- *Compare Methods using real life historic data sets covering medium and long tail business.*
- *Health warning! Why methods can fail, with illustrations.*
- *Provide selected bibliography for further reading on Statistical Methods*

The views expressed in this paper represent the consensus of the members of the working party in their personal capacity and do not necessarily reflect the views of their employers or every individual.

Index

An asterisk (✱) next to the title indicates a mathematical section which may be skipped without prejudicing understanding of the principles concerned. Parts of sections in the body of the text are also marked in such a fashion.

1. Introduction

or

Why this paper?

The motivation for this paper is a desire to educate the profession. A number of statistical techniques have developed over recent years that purport to assist in determining claim reserves, and we want to bring these to the attention of a wider audience.

These techniques have slowly been gaining in popularity in recent years. This is mainly due to the increasing speed and low cost with which microcomputers can handle the heavy computational load necessary to carry out the calculations.

However, the mathematics is heavy, and this factor alone is enough to put off many practitioners. We believe that this should not be seen as an obstacle. The paper aims to pull together the underlying principles to show the factors that these methods have in common and, by comparing and contrasting their results, to bring out their particular features.

We hope this paper will give readers a better understanding of how these techniques work in practice, and give sufficient explanation of their strengths and limitations so that they may judge when their use may be appropriate.

The title of 'Claim Variance' has been used somewhat liberally. The paper is concerned with the difference between actual out-turn of claims and the predictions used to decide upon reserves before the claims have been paid. The main thrust of the paper looks at statistical methods which provide a measure of the variance of the estimate made for the claim reserve. The accountant's use of variance, in the sense of "difference from expected", is employed in consideration of the insurance industry's track record in the field of claim reserving. Section 2 deals with a review of DTI Returns over the past ten years or so. There is some suspicion that companies may manipulate their reserves, and one or two popular theories of systematic bias have been tested.

Section 3 introduces the subject of 'stochastic modelling'. A scientific approach to statistical claim reserving, in the form of a formal modelling structure and the use of the statistical diagnostics, is established. This section provides a checklist of essential statistics for the modeller.

Section 4 briefly considers the value of the claim variance in terms of how it can be used and applied.

Section 5 sets out concerns that the doubting actuary may have about using statistical methods. The reader may care to judge whether the paper adequately addresses these concerns.

Section 6 starts by posing a real life claim reserving problem in the form of genuine UK medium and long tail business.

Outlines are provided of three statistical reserving techniques, within the framework of Generalized Linear Models (Log-Incremental Payments, Log-Incremental II, and Operational Time) . Brief details of Generalized Linear Models are given in an appendix. A statistical "add-on", Bootstrapping, which can be used to provide information on the variance of all types of reserving methods, is also discussed.

These techniques are then used on the real test data to demonstrate how they cope in practice. Results are compared with the known outcome (these were not revealed to the operators in advance of their analysis!). The lessons of the study are then discussed so that readers may appreciate the differences between the techniques. Included in this section are some details on the practical limitations of each of the techniques.

Section 7 summarises the particular features of each of the methods.

Section 8 briefly considers the value of the statistical methodology, and why statistical methods may fail to produce the correct forecast.

Section 9 winds up with an extensive list of further reading which will be of particular use to readers who wish to have a better understanding of the mathematical bases that underlie the reserving techniques discussed. References to other methods are also given.

A glossary of terms used in this paper is given in appendix 2.

This paper will have achieved its objective if the reader is left with a better understanding of the application of statistical methods to the subject of claim reserving, and feels sufficiently confident to try out the techniques.

2. Review of Past Reserving Adequacy

or

How Good Are We?

2.1. Introduction

It would be comforting to feel that more complicated claims reserve modelling is unnecessary because insurers already reserve accurately. Is this the case?

The main tranche of data available to test the current state of affairs is provided by the DTI returns. These we have used to examine two aspects of the question:

- *Biases:* Do the swings and roundabouts balance out? Or is there consistent under or over reserving? Is there any evidence that reserving practice is influenced by underwriting result?

- *Variability*: How close are initial reserves to ultimate payments? How much does it vary across different companies? How quickly do estimate ranges settle towards the ultimate? What difference is there between classes of business?

Needless to say, the data is not ideal. Since this investigation is only to provide a backdrop to the main work of the report, it has not been possible to do more than apply a broad brush approach. As such the 'biases' investigation has not distinguished between classes of business (although the 'variability' analysis has). In both cases there has been a degree of stratification by size of company as follows:

Total incurred claims in 1991	Number of Companies
< £10 million	9
£10m to £100m	17
£100m to £1000m	16
> £1000 million	7

It is accepted that much of the data has been aggregated despite considerable heterogeneity. However the conclusions drawn are very broad in nature, and it is not felt that they are threatened by the crudity of the approach. Moreover, looking at the aggregated position is appropriate if management decisions regarding reserve strength are taken at an overall portfolio level rather than at an individual class level.

2.2 Biases in Reserving

The main thrust of the investigation has been to test reserve adequacy against underwriting results. The method has been to postulate a variety of hypotheses (for instance that surpluses in reserves revealed in Form 23 of the DTI return will be correlated with the previous year's underwriting result). In general, data has been plotted in monetary terms for each level of stratification, and the points quartered in such a way that, if no correlation exists, equal numbers of points would be found in each quadrant. Statistical significance was tested using a simple contingency table/chi squared approach.

The significance levels for the chi squared distribution with one degree of freedom are:

Chi squared Value	Significance Level
3.84	95%
5.24	97.5%
6.64	99%
10.83	99.9%

The returns for the years 1987 to 1991 inclusive have been used, and each company has provided a data point for each year. In order to use all the data together for each approach, scaled versions have been produced by dividing through by Net Premium Income.

Various hypotheses have been investigated:

2.2.1 Hypothesis - Claims reserves will be boosted in years when underwriting results are good, and weakened when u/w results are poor.

How Tested: Savings on estimates recorded in year t+1 were plotted against underwriting results in year t.

Logic: If the overall company underwriting result is healthy, the company will for the sake of prudence take the opportunity to bolster reserves. As such, when the bulk of claims comes in - in the next financial year - there will be a higher than usual release of reserves.

Chi squareds:

Company size	Chi squared	For or against hypothesis
< £10m	7.1	For
£10m to £100	0.00	Not sig.
£100m to £1000m	1.0	Not sig.
> £1000m	2.3	Not sig.
All companies	6.6	For

Conclusion: There is a weak level of significance here (given the large number of chi squareds being carried out), but there seems some evidence to support the hypothesis.

It needs to be noted, however, that if you spend your good underwriting result on boosting claims reserves, that will in itself worsen the result. That thought process leads on to testing the hypothesis in a different way.

2.2.2 Hypothesis - Claims reserves will be boosted in years when underwriting results are good, and weakened when u/w results are poor.

How tested: Savings on estimates recorded in year t+1 were plotted against underwriting results in year t. The latter were, however, adjusted by the amount of savings accruing in year t+1. The adjustment comprised simply adding those savings to the u/w result.

Logic: The underwriting result in year t would have been better if the savings liquidated in year t+1 had been taken up front. The reserves are boosted in proportion to the prevailing 'feel good' factor in year t - which is governed by the results **prior** to boosting.

There was some discussion in the working party as to whether the add-back to the u/w result should be limited to the increase in savings over the previous year (which is, after all, the true distortion to the result). In the end, however, it was felt that the psychological impetus is applied to senior management through the operation of the actual incurred losses - leading to the underwriting result. It is then that decisions might be made to sideline profit into the reserves.

Results: A full set of charts for this approach are attached as Appendix 3, but the results are summarised below.

Chi squareds:

Company size	Chi squared	For or against hypothesis
< £10m	7.1	For
£10m to £100	8.5	For
£100m to £1000m	6.3	For
> £1000m	5.1	For
All companies	36.0	For

Conclusion: For each stratum, and overwhelmingly for all companies taken together, there appears to be very strong support for the hypothesis.

Effectively we are adding one variable to the other prior to testing for correlation. The justification for this seemingly dubious procedure is the assumption that the variable had previously been deducted, by the process of sidelining profit from year t to year t+1. Thus the data is being corrected rather than distorted. If this assumption is incorrect, clearly the process builds in correlation, and throws doubt on the validity of the conclusion above.

2.2.3 Hypothesis - The level of IBNR set up is related to the level of underwriting profit.

How tested: In this case data in monetary terms was not felt to be appropriate. IBNR was expressed as a percentage of total claims reserves, and plotted against the underwriting result in the same year as a percentage of Net Premium Income.

Logic: If a company has a good underwriting result in a particular year, it will feel able to bolster/rebuild its IBNR reserve. Since it is more easily manoeuvred than case estimated claims reserves, it will therefore rise as a percentage of total claims reserves.

Chi squareds:

Company size	Chi squared	For or against hypothesis
< £10m	5.8	Against (!)
£10m to £100	2.3	Not sig.
£100m to £1000m	1.8	Not sig.
> £1000m	2.9	Not sig.
All companies	0.3	Not sig.

Conclusion: The perverse result for small companies probably illustrates the problem that there are two contrary forces acting. On the one hand a big IBNR will reduce the underwriting result. On the other hand a bad underwriting result reduces the desire for a large IBNR. Perhaps for small companies the former is a bigger force than the latter.

Overall however it is interesting that there does not seem to be an IBNR effect to match the Claims Reserve effect. IBNR reserving does not seem to be correlated with underwriting result.

2.2.4 It was suggested that work done in America showed some correlation between reserving levels and **changes** in underwriting results. Thus, as performance **improves** insurers, take the opportunity to boost reserves, and *vice versa*. Analysis showed some weak correlation for larger companies, but this did not seem a particularly fruitful line of approach.

2.3. Variability Of Reserves

Form 33 of the DTI Returns permits an analysis of reserving accuracy over different accident years.

The form details for each accident year, on a gross basis, paid, outstanding and IBNR claim amounts over successive financial years.

The data is split by risk-group and it enables claim triangles to be constructed that show the movement of ultimate claim estimates as the accident year develops.

The analysis has been centred on the seven major composites, i.e. the "over £1,000m total incurred claims" group.

For each of the seven companies ultimate claim triangles have been produced for each of the one-year accounting classes, i.e.

- Accident & Health
- Motor
- Property Damage
- Pecuniary Loss
- General Liability

By using the accounting class, the data is split broadly into the different types of business, which should ensure consistency for all companies. However it is acknowledged that the differing mix of risk-groups within each class does introduce a level of heterogeneity.

The graphs shown in Appendix 4 plot the development of the total claim estimate as a percentage of the 1991 year end value for each company and for accident years 1981-1986 inclusive.

The line on each graph represents the arithmetic average of all points for each development year (NB development year 1 is the year of origin).

The graphs exhibit the following characteristics :

How close are the initial estimates to the ultimate value?

The General Liability and Pecuniary Loss classes have a large spread in the initial estimate.

The spread is reduced for the Motor and Accident & Health classes, with the Property Damage class showing the smallest variability.

In terms of the average, the most noticeable trend is the initial over estimation for the Property and Pecuniary Loss classes.

How quick is the convergence?

As one would expect, the long-tail classes show the slowest convergence.

The Accident class shows no consistent tendency to either under or over estimate.

The Motor class shows a tendency to under-reserve for development years 2 onwards.

The Property Damage class shows consistent over-reserving for all development years.

The Pecuniary Loss class shows consistent over-reserving for all development years.

The General Liability class shows consistent under-reserving for all development years.

The observed variability does not of course directly reflect the accuracy of actuarial reserving methods - rather it depends on the nature of the reserves used for the DTI returns.

Results for smaller companies

A similar analysis was performed on a selection of smaller "£100 - £1,000m total incurred" companies.

The results were consistent with the above but, as might be expected, the spread of estimates tended to be greater.

The exception was the General Liability class, where claims tended to be over-reserved at early years.

2.4. Final Observations

2.4.1 It is clear that companies do strengthen reserves when they feel they can afford it, and that they draw them down when they need to. This amounts to the operation of implicit equalisation reserves, and is obviously prudent business practice. It does, however, imply that the reserving process is not even aiming at absolute accuracy (or accuracy plus a margin).

2.4.2 For the majority of reserves we (reserve modellers) are not under scrutiny here. Most of the claims reserves are still based on case estimating; which could clearly be improved upon.

2.4.3 It is interesting to note that the IBNR reserves, which are most likely to be influenced by modelling, are the least distorted by underwriting performance. It may not be too optimistic to suggest that, whilst they are far from accurate, modelled approaches are providing unbiased estimates.

2.4.4 Some consistent biases are evident, with the strong implication that reserving could be improved. These are most pronounced in the consistent over-reserving of the short-tail Pecuniary Loss and Property Damage accounting classes.

3. Stochastic Modelling Background

or

What are Stochastic Methods?

Most claims reserving methods are based on some assumptions about the underlying shape of the run-off. The assumptions usually define a mathematical model of the run-off. The difference between stochastic and non-stochastic methods is that in stochastic methods the mathematical model is not confined to the underlying pattern: the variation of the data around the underlying pattern is also modelled. The stochastic approach offers three main benefits:

(a) The influence of each data-point in determining the fitted model should depend on the amount of random variation in that data-point: figures with large random components should have relatively little influence.

(b) The reliability of the fitted model, and the likely magnitude of random variation in future payments can be estimated. This enables `standard errors' indicating the reliability of predictions to be calculated.

(c) Statistical tests may be applied to the modelling process to verify any assumptions and gain understanding of the variability of the claims process.

3.1. Types Of Stochastic Models

There are three basic types of stochastic models depending on the data to which they are applied:

(a) Models that are applied to aggregate data, that is, a run off triangle of amounts paid or incurred.

(b) Models that also involve triangles of numbers of claims to enable more accurate estimation of average costs, frequency and inflation.

(c) Models that are applied to a database of individual settlement amounts.

Models of types (a) and (b) may be used to derive reserves and standard errors; with approximate techniques being necessary to estimate the distribution of the aggregate reserve. Models of type (c) may be able to estimate more accurately the aggregate reserve distribution.

3.2. Modelling Results

Models may be used to derive estimates of future cash flows or incremental incurred claims, total reserves for origin or calendar periods and any associated standard errors or probability distributions.

Models usually produce a set of fitted values which may be compared with the actual data to derive residuals of the fitted model. These form a useful basis for testing modelling assumptions and examining the nature of the claims development process.

3.3. Basis For Stochastic Modelling

Models usually fall into one of two categories:

(a) Ad-hoc models where no assumptions of the underlying process are made, the data is modelled using any shape and variance structure that happens to fit past development.

(b) Models derived from an underlying theory of the claims process. These models start with a set of assumptions that are then refined and calibrated to the data, or else if not appropriate, alternative models may be suggested.

There are occasions when each of the above approaches is more appropriate. However, when modelling a small data set, for example less than 10 years of annual development, there are few data points in the tail to construct a model of the variance. In these cases, a prior view of the variance structure can be helpful.

3.4. Stochastic Modelling Misconceptions

There are three common misconceptions with regard to statistical approaches, which should be dispelled.

i) "The 'optimal' statistical model is the best for producing forecasts"

The 'optimal' statistical model **may not** be the best for producing forecasts. The 'optimal' statistical model may tell us that there is instability in trends in more recent payment years. Judgement about future trends could then be based on analysis of other data types, e.g. claim numbers closed.

ii) "The model represents explicitly the underlying claims generating process."

There are essentially two approaches to formulating the initial model. It may be generated because it is believed that the model represents the underlying processes at work. Alternatively the model may be developed in terms of simple components which fit the observed experience.

In either case the model should be proved by checking that all assumptions inherent in the model are supported by the data.

iii) "A stochastic model (as opposed to a deterministic model) is always useful."

A model contains information or assumptions. If the assumptions contained in the model are not supported by the data then the model is not useful.

3.5. Testing Models - Diagnostics

3.5.1. Residual Plots

For each data point we have an observed and a fitted value. The difference between these is defined as the residual error. If divided by an amount proportional to the estimated variance for the point, the residual is known as a standardised residual. If the residuals are assumed to be normal, then the standardised residual is a normalised residual.

The model assumes $E(r) = 0$ and $Var(r) =$ constant

Hence we can plot residuals against origin (or underwriting period), development and payment (or calendar) periods. If $E(r)$ is not equal to zero, it may show up as a systematic error in the residuals. If the plot against development period looks non random, then it is likely that the assumed shape of the run off is inappropriate. If the plot against origin period looks non random, then it is likely that the assumed level for some origin year(s) is wrong, and if the plot against calendar period looks non random, then it is likely that the assumed inflation model is wrong.

However, the above residual plots may still appear reasonable even though there are systematic errors in the fitted model. These may often be detected by examining the triangle of residuals, for example, plotting positive residuals in one colour, negative residuals in another and by setting the brightness in accordance with the magnitude.

Even if the E(r) is zero, the Var(r) may not be constant. For example, the residuals may be fanning outwards or inwards with development. In this case, the model for the variance may be wrong. It is important to correct for this as the model for the variance determines the amount of weight that each point is given in fitting the model and hence the fitted pattern. If the residuals appear to be fanning in with development, then the data points in the tail may be given too little weight compared with earlier values. Conversely, if the residuals are fanning out with development, then the data in the tail will be given too much weight and predicted variances in the tail will be too low.

If the residuals are standardised rather than normalised, note that they may still be skew. Where Generalized linear models are being used, this will be allowed for in the modelling.

When modelling small data sets, for example, 6 years of annual development data, residual plots have to be used with care since it is easy to see patterns in the residuals and end up with an over-parameterised model with unrealistically low standard errors. It is often helpful to gain experience of random residuals by creating triangles of random normally distributed numbers and examining the plots. Most people will see patterns in small sets of random residuals!

An example of a residual plot is shown in 6.2.2 as part of the analysis used for the log-incremental payments technique.

3.5.2. Statistical Tests

The use of F, t and other tests is helpful in deciding on the number of parameters to use in a model.

One approach is to use the standard "GLIM" type of analysis using F tests. This approach starts with a possibly over parameterised model and then fits models with subsets of the original parameters. An F statistic may be constructed and used to check that the reduction in parameters doesn't introduce significant extra residual variability.

Standard errors of fitted parameters may be checked to test whether the fitted parameters are significantly different from zero, and hence whether or not the parameters should be included in the model. They may also check whether parameters are significantly different from each other (if they represent different levels of the same parameter set).

Where Kalman filters are used, some sort of parameter counting method is needed to allow for the dynamic nature of the model.

Where models are being compared that have a different structure, then statistics such as the AIC or BIC may be useful.

3.5.3. Validation

It is possible to refit models ignoring the most recent 1, 2 or 3 years data and compare the results with current estimates. If the estimates are stable then the model may give a reasonable estimate for future development. If the estimates are unstable compared with the estimated standard errors, then the model may be unreliable for predicting the future.

It is usually helpful to model different data types (for example paid and incurred) and to apply different methods. If the answers are similar, then the model are likely to be more reliable than if they are inconsistent, in which case judgement will be required to eliminate the inappropriate models.

It is often helpful to examine plots of the actual development together with the fitted models so that graphs of incremental and cumulative development appear reasonable for the data set being modelled.

3.6. Modelling Error

Whatever model is used the final outcome will inevitably differ from that estimated. The standard error can be useful in indicating the size of the likely error of an estimate. However there are potentially other sources of error.

Taylor [10] sets out the components of the prediction error as follows

specification error	arising from the initial specification of the model - typically this will be due to assuming linearity which does not exist
selection error	due to incorrect selection of the predictors
estimation error	due to the fact that the estimated parameters are still only random variables
statistical error	reflecting the inherent random noise in the process

The standard error is the sum of the estimation and statistical errors, i.e. the parameter uncertainty and the residual variation. Hence the model is still subject to unmeasured specification and selection error.

This subject is further considered in the section dealing with the concerns of the Doubting Actuary.

4. Application of the Claim Variance
or
What use is it?

Essentially this can be answered by considering the converse. If we don't know the claim variability (or at least have a feel for the sensitivity of the reserves established) how can we form an opinion as to the adequacy of the Claim Reserve? After all, the requirement for adequacy may carry with it the implication of "with margins for caution". As the required size of any margins can only be judged by reference to the variability of the reserve the importance of these statistics should be readily apparent.

Other possible uses include.

(a) Assessment of reserve adequacy may be applied in the context of both absolute (as above) and also relative terms comparing origin years and lines of business.

(b) Basis of allocation of capital. Again, both in absolute and relative terms.

(c) Basis for comparing modelling of different data sets, for example, paid/incurred/average costs and numbers.

(d) Basis for discussions with the DTI or Inland Revenue. Quantifying the uncertainty in estimated reserves can be helpful in these circumstances.

5. What Does the Doubting Actuary Require?

or

What value statistical Techniques?

5.1. Benchmarks

In order to illustrate the questions we answer in this paper we have constructed a "Doubting actuary", whose concerns are set out in the rest of this section. The criteria by which we would like this paper to be judged are our successes at answering the various issues raised below.

5.2. The "Doubting Actuary"

I am the Doubting Actuary. I am responsible for advising on appropriate levels of provisions for a wide range of types of general insurance. As well as needing to arrive at a "best estimate" of claims from business already written - not necessarily an easy task - I would also like as much information about possible differences between the eventual out-turns and my current estimates. I want this information

(a) for assistance in monitoring my own performance
(b) so that I can decide whether differences between my estimates and those of my colleagues are material
(c) so that I can advise on the range of possible outcomes.

In my own mind, I have several ranges which need to be taken into account:-

- Range A includes any point estimate and consists of those values which I regard as equally valid. If the eventual out-turn is within this range I will regard myself as having been a good predictor, and I would be prepared to support provisions anywhere within the range

- Range B surrounds Range A, and consists of estimates which I do not regard as unreasonable but which I would not myself be prepared to recommend

- Range C surrounds Range B and consists of estimates which may be achieved in practice but which I do not consider as reasonable for current provisions. If the eventual out-turn is within Range C I shall say "That is not what I expected but it is not a major surprise"

- Range D surrounds Range C and consists of results which are possible but which would be a major surprise

- Range E surrounds Range D and consists of results which I currently believe to be impossible.

These concepts are described in qualitative rather than quantitative language and the boundaries between the ranges themselves are often intuitive points which I would be pushed to define other than by saying that they are my subjective impressions.

I would welcome anything upon which I felt able to rely to help me refine these concepts.

Over time I have had the benefit of various presentations on various statistical methods which might address the problem. Some of the presenters have had a financial interest in promoting their methods and some have not.

In general the methods have three components:

(a) a statistical model
(b) a way of fitting the model to past data - i.e. choosing parameters
(c) a justification for the belief that the model will predict future claims experience.

So far as (a) is concerned, I can normally follow the work done by others and, provided it is published, and scrutinised by others with statistical expertise, I am happy to take it on trust.

(b) is usually a way of solving a large number of simultaneous equations with the benefit of a computer. I learnt how to solve simultaneous equations at the age of ten so I have no difficulties with this stage and I probably underestimate the effort and skill which were needed to produce the method of solution.

I do have difficulties with (c). I consider that it is my responsibility to determine whether the model is actually appropriate for making predictions about future claim payments; and it is the point upon which proponents of various statistical methods seem weakest.

Using models for predictions requires:

(a) that the model describes behaviour in the future, whether or not it has done so in the past.

(b) that the parameters have been correctly determined.

To put this concept into probability terminology, suppose that X is the event that claims exceed some specified level. Then:

$$P(X) = P(X/A_1)\,p(A_1) + P(X/A_2)\,P(A_2) + P(X/A_3)\,P(A_3) + P(X/A_4)\,P(A_4)$$

where: A_1 is the event that future claims will be in accord with the model and the parameters have been correctly determined,

A_2 is the event that future claims will be in accord with the model but the parameters determined from past experience are not appropriate for the future.

A_3 is the event that the model was correct for past behaviour but is no longer so

A_4 is the event that the model was never correct and the apparent fitting to the past experience is illusory.[1]

We can probably put a numerical value on P (X/A_1) but would anyone have any idea of the values to be assigned to the other possibilities?

My doubts about the applicability of whatever model is under discussion stem from several sources:

(a) We know that it is possible to go through the motions and assign values to parameters but this does not necessarily mean that the model was appropriate in the past, let alone the future. Far more work seems to have been done assuming that some model will be appropriate than in assessing whether it is so; and in demonstrating this fact in a convincing way. Unless supporters of models can meet me on this matters I am going to be rather unconvinced and unwilling

[1]In the language of Taylor [1] A_2,A_3,A_4 correspond to a belief that the specification error particularly, and to a lessor extent selection error - both of which cannot be measured - are likely to be significant in practice, thereby invalidating the use of standard errors as a measure for the reliability of a model.

to rely upon the results. I would also like some information about the consequences of the model fitting only approximately.

(b) Many models seem to be dependent upon assumptions which are not in fact true. Usually it is assumed that incremental claims can be described by a parameter and an error term, the error being independent. In my conception of the claims process, a fixed but unknown number of claims occur during the period of exposure and there is then a variable period for each claim until it is paid (or reported). Since a claim can be paid (or reported) in only one period, I would expect the error terms to be negatively correlated. When incremental claims are taken to be log normal, the method ought not to be applied to both quarterly and annual data. I have read that the sum of a log normal distributions is not log normal and I deduce that quarterly and annual claims can not both be log normal.

(c) If an astrologer produced a model based upon the movements of the stars which "explained" the past claims experience, would I be happy to use his model for predictions of the future? Surely not! If his predictions always turned out to be true, would I change my mind? Probably, eventually after some period; it would be perverse not to accept that his model seemed to work but unless I understood the mechanism which turned star movements into claims I would always be concerned that the predictions would fail at the next attempt. Clearly one does not have to emulate, or even understand, the complicated processes which produce claims at various times; they may combine to fit some straightforward statistical model, but the validity of assumptions which appear not to be met in practice will need to be explained.

(d) Since I am not an expert on statistics, my opinions of the reliability of proposed methods will be influenced by my assessment of their proponents. Inaccuracy on matters which I can check will make me doubtful about things which I can not. Examples of statements which undermine my confidence are:

   (i) the method will work on all classes and types of business. Look at the list of classes of business for which I have a responsibility! Is Property Catastrophe Reinsurance to be treated in the same way as Personal Motor?

   (ii) working on incremental paid claims will give more reliable answers than incorporating information about outstandings. True perhaps for personal lines. My aviation account, which is exposed to (compared with other types of insurance) a relatively small number of large claims, can be predicted reasonably well by using a chain ladder on incurred claims. At short durations little has been paid and it is difficult to draw useful conclusion. Loss ratios vary wildly from year to year from, say 20% to 300%! For some casualty accounts nothing at all may be paid for the first couple of years.

(iii) statistics should be adjusted for exposure and inflation. Yes, if it can be done. What does one do in the case of, say, excess of loss reinsurance of Employers' Liability?

(iv) anyone who uses the chain ladder is an idiot. I have a vested interested here, in that pride won't let me agree to this statement! Anyone who believes that any responsible actuary is going to use the results of a naive chain ladder blindly, doesn't know much about what actuaries do. My own methods would call for an examination of the data, possibly the removal of large claims, and then, if it seemed the right thing to do, calculation of the linked ratios implied by the past experience which would then be adjusted in the light of all collateral information, including trends in the experience, because my job is to guess what they will be in the future not to say what they were in the past.

I look forward to reading the rest of the paper in the hope that it will assist me to do my job better.

6. Statistical Techniques & Application to Real Data

or

What Are These New Methods and How Good Are They?

The statistical techniques presented here are not intended to cover the full range of possible methods. Rather they should be seen as representative of a range of methods that are currently attracting considerable interest, as witnessed by the plethora of papers being produced on the subject (some are given in the bibliography.)

The first three use general linear models to obtain the parameters and should be seen as frameworks for deriving an ultimate model for the data in the sense that the user must interpret the results of a particular model and then exercise his judgement, based on the diagnostics, as to whether the model

a) can be improved by more or fewer parameters;

b) provides a fair representation of past data; and

c) is suitable for the future.

With regard to c) it is vitally important that the actuary considers whether the parameters fitted to the past data need adjustment in the light of known or predicted developments in the outside world.

Certain mathematical sections may be skipped without prejudicing understanding of the principles concerned. These are denoted by an asterisk (*) next to the title, or in the left hand margin of the section, and apply to the remainder of the section.

6.1. Data Description

Three classes were prepared for the various reserving techniques, comprising two long tail classes and one medium. Six sets of data were constructed initially, which is why here and in the rest of the paper they are referred to as classes one, three and five.

The data for each class was for ten accident years developing quarterly. The data triangles provided were:

Paid Claims net of salvage and subrogation
Number of Closed Claims
Case Estimates
Number of Open Claims

The salvage and subrogation could have been considered separately, but it was decided to stick to examining the paid claims net of salvage, as it is more common for data to be held at this level.

All the data was net of a typical reinsurance program, with no particular features of merit, that remained pretty much consistent over the period considered. The data is gross of any special facultative or stop-loss type reinsurance arrangements.

The reserving methods were used to predict ultimate claims and the claims expected to be paid in the next three years, for each accident year and for the class as a whole, giving a range one Standard Error either side. Updated triangles were then provided, giving the position three years down the line. The predictions of payments over the next three years could then be compared to the actual outcome. The payments over three years represented up to 80% of the reserve as a whole for the three classes, so good estimates for this figure should augur well for the reliability of the reserve estimate.

Class 1 - Long

This class of business is very long tail indeed, with a modest number of claims still being reported some thirteen years after the original accident year. After ten years development, for every three claims closed two new claims are reported and of the total number of claims reported for a given accident year, as many as one in five hundred are still outstanding ten or more years after the accident year. Around 95% of claims are paid in the first year, these claims being around one hundredth of the average loss paid ten years later. This is typical of employer's liability business. The average payments and case estimates do not change markedly after ten years of development. It seems likely that the payment and reporting of claims after year ten will continue for a considerable time into the future at pretty much the level they are currently. Modest growth was experienced in the first few years under consideration; the class was then drastically curtailed. Each of the last five accident years had approximately half the number of claims of each of the first five; the last five accident years could therefore be expected to behave somewhat differently and this is evident in some of the patterns that emerge - for example having a slower reporting pattern.

Initially high inflation of claims payments of 10% or so in the initial calendar years gradually decreased over the period considered to around 5% in the recent years.

Class 3 - Medium

Whilst certainly not having the long tail features of class one this does still have a few claims reported some ten years or more after the accident year. For this class though, for every four claims that are closed after year ten, only around one claim is reported and only around one in ten thousand of the total claims are still outstanding after ten years.

Around 80% of claims are paid in the first two years, these claims being around one tenth of the average loss paid after ten years of development. The payments and case estimates in the tail are fairly erratic, being distorted by the occasional very large claim. This is typical of property business.

This class grew markedly in the first two years, growing 50% as measured by the ultimate number of claims, then gradually declining thereafter to pretty much its' original level. Most accident years show fairly stable payment and reporting patterns.

Inflation was again at a high level, in excess of 10% initially, then dropped away to low levels of a few percent before starting to increase again.

Class 5 - Long

As for class one, claims are still being paid and reported well into the tail of this class with around three claims being reported for every five claims closed after year ten and one in five hundred claims still being open after ten years of development.

Around 80% of claims are paid in the first two years, these claims being around one hundredth of the average loss paid after year ten. The tail is again erratic with some very large claims indeed. This is typical of general liability / bodily injury business.

The class grew rapidly, doubling in size in the first five years then contracted rapidly, reducing by 50% to its' original level. The patterns show the turmoil of these changes and exhibit some marked changes in reporting and payment patterns, with speeding up closely followed by slowing down!

Inflation was more consistent across the years, dipping then rising again, but not exceeding the ten percent levels experienced by the other classes initially.

6.2. Regression model based on Log-Incremental payments

6.2.1. Overview of method

The first method presented is the natural extension of the chain-ladder model. Statistics are introduced by reference to the error term. A mathematically simple structure for the error terms is assumed, namely that the errors in the logarithms of the incremental claim payments are independently identically distributed (i.i.d) normal random variables.

The method in its simplest form is very general, and is likely to suffer the same problems encountered by the straightforward chain-ladder.

Use of logarithms for the log-Incremental method does have some theoretical objections but nevertheless has been found to work well in practice.[16]

In practice these models are relatively easy to develop. The recommended methodology set out in Section 3 ensures the modeller tests the validity of the key error structure assumption, and systematically identifies any problems thrown up by the data.

The statistical significance of the modelled fit is generally improved by reducing the number of parameters. To this end simplified models for the pattern of claims run-off within each accident year (the *development year* axis) are introduced.

6.2.1.1. Introduction

The method applied in section 6.2.2 fits models to the various sets of data - in this case the triangles of past claim numbers and payments. This modelling can be viewed at two levels. At the first level, the models are specific examples of more general types of model - Generalised Linear Models. At another level, they are an extension and refinement of a more basic type of model - that is what is sometimes called the Stochastic Chain Ladder (see [6] for a more in depth comparison with the Chain-Ladder).

To put the method into context, a brief description of the more general class of models, and the more basic type of model is given, before the specific method used is described.

6.2.1.2. Stochastic Chain-ladder

There is a considerable body of work describing models which are the stochastic equivalent of the Chain-ladder [6]. These models are broadly the same as the deterministic (i.e. without a random component) Chain-ladder, except that:

- the model is explicitly defined;
- the parameters of the model are estimated using a statistical technique to obtain a "best fit", rather than choosing factors "by eye" or by a simple calculation;
- finally, explicit assumptions are made about the systematic and random elements of the claims process, so estimates of the variance of the reserve estimates can be produced, rather than just a point estimate.

The model of the claims process is described as:

$$\mathrm{Log}(P_{ij}) = a(i)+b(j)+E_{ij}$$

where P_{ij} are the incremental claim payments in Accident Year i at development period j, a(i) and b(j) are the parameters fitted by the model and E_{ij} is an Error term. The E_{ij} are assumed to be independent and identically normally distributed with mean zero and variance σ^2.

The assumption that the Eij are identically distributed across the whole triangle is an area of the model open to debate. The payments in the early years of development are likely to be larger than in the tail, and so one might expect the error term to reflect this. Conversely however, it is the tail where a few large claims can have a particularly distorting affect. The simple error structure can be refined, as is the case in some other models.

For an n x n triangle, there are n parameters a(i) and n-1 parameters b(j) (we set b(0)=0, so that there is a unique solution for a(i) and b(j)). In total then we have 2n-1 parameters. This has the same number of parameters as the basic Chain-ladder, which implicitly assumes a "level" for each of the n Accident years, corresponding to the a(i), and a development year effect, corresponding to the n-1 b(j) parameters the model fits.

* For convenience, denote $Y_{ij} = \mathrm{Log}(P_{ij})$ and let $\theta_{ij} = E(P_{ij})$. Under our assumptions above, the a(i) and b(j) can be estimated by a simple regression procedure to produce Maximum Likelihood Estimates of the Y_{ij}, $MLE(Y_{ij})$, in our Log-space. We then have to produce estimates back in our payment-space. It can then be shown that the

Maximum Likelihood Estimate of θ_{ij}, $MLE(\theta_{ij})$, back in our payment-space, is given by:

$$MLE(\theta_{ij}) = \exp(MLE(Y_{ij}) + \tfrac{1}{2} Var(MLE(Y_{ij})))$$

It can also be shown that the Standard Error, SE, of the estimates of θ_{ij} is given by:

$$SE(MLE(\theta_{ij})) = MLE(\theta_{ij}) \times (\exp(Var(MLE(Y_{ij})) - 1)^{1/2}$$

Thus, from our simple regression estimates of Y_{ij} in the Log-space, we are able to back out estimates of the expected values of the actual payments and their standard error. The practical details of how one may perform the regression using spreadsheet regression functions, and make the transformation back into the payment space are not set out in this paper. The interested reader may refer to the IOA Reserving Manual Volume II [1].

There is a slight hitch in this method of producing estimates of expected payments, in that the Maximum Likelihood Estimates of θ_{ij} can be shown to be biased - that is:

$$E (MLE (\theta_{ij})) > \theta_{ij}$$

The MLE of θ_{ij} is asymptotically unbiased, that is, as the sample size gets larger, the MLE gets closer to the true value of the θ_{ij} . For small sample sizes however, as can be the case with reserving data, the bias may be significant. Section 6.5.2. gives an example of how Bootstrapping can be used to quantify the extent of this bias.

Alternatively, a different, unbiased estimate of θ_{ij} can be made. Finney [3] showed that an unbiased estimate of θ_{ij}, say θ_{ij}', can be constructed as follows:

$$\theta_{ij}' = \exp(MLE(Y_{ij})) \times g_m(\tfrac{1}{2} s^2), \text{ where:}$$

$$g_m(t) = \sum_{k=0}^{\infty} m^k(m+2k)/(m(m+2)....(m+2k))(m/(m+1))^k t^k/k!, \text{ and:}$$

$$s^2 = n/(n-1) \times Var (MLE(Y_{ij}))$$

where:

n = the number of data points ; and

m = the number of degrees of freedom of the model i.e. n less the number of parameters

Similar adjustments can be made to produce unbiased estimates for the variance of our expected payments. In practice, especially if working in a spreadsheet environment, the extra effort needed to produce unbiased estimates is usually deemed to be disproportionate to the extra accuracy gained.

One can extend the estimation of variances of payments from one payment to several - either all the future payments for a given Accident year, say, or the total of all future payments for all Accident years. This enables one to produce an estimate of the Standard Error for the reserve for a given Accident Year, or the Standard Error for the reserve for all Accident years combined.

The SE's for these combinations of payments are obtained from the standard identity $Var(A+B) = Var(A) + Var(B) + 2 \times Cov(A,B)$. In this context, this translates to:

$$Var(MLE(\theta_{ij}+\theta_{kl}))=Var(MLE(\theta_{ij})) + Var(MLE(\theta_{kl})) + 2\times Cov(MLE(\theta_{ij}),MLE(\theta_{kl}))$$

where it can be shown that:

$$Cov(MLE(\theta_{ij}),MLE(\theta_{kl}))=MLE(\theta_{ij}) \times MLE(\theta_{kl}) \times \exp(Cov(MLE(Y_{ij}), MLE(Y_{kl}))-1)$$

Whilst some of the algebra above looks daunting, the various calculations for the expected payments and their variances all boil down to the simple regression estimates of a(i) and b(j), hence Y_{ij}, in our Log-space, which are then suitably transformed to re-cast them as estimates back in our original payment space.

6.2.1.3. Regression model based on Log-Incremental payments

The model described in 6.2.1.2. is similar to the basic Chain-ladder, with the addition of a random component, and uses the same number of parameters in defining the model. There is no reason, however, why the model should slavishly follow the structure of the basic Chain-ladder model.

The basic nxn Chain-ladder may be considered over-parameterised, as it fits 2n-1 parameters to a data triangle containing $n\times(n+1)/2$ points. When modelling, one wants to strike a balance between models that adhere too tightly to the data, by having too many parameters, and models which over-simplify a complex process with too few parameters to accurately reflect elements of the underlying process. One may also want to use the model to make predictions beyond the range of the data - in this case for development periods beyond the latest development period in the data. These considerations lead to refinements of the model, which may describe the data with fewer parameters, which are more stable with little or no loss to the standard of the "fit" of the model, and which enable one to make predictions beyond the latest available development year.

The refinements to the model broadly address two questions. Firstly, whether the model can equally well be fitted without the use of all n of the a(i) parameters - the Accident year level parameters. Secondly whether the data exhibits some sort of pattern which can be fitted by a curve, or combination of curves, to describe the b(j) parameters, rather than have n-1 separate parameters to describe each development period. The process of making these two refinements, as it pertains to the actual Working Party database, is described in 6.2.2. An overview of the nature of the refinements and how they are implemented is given below.

Considering the first point, there are methods, as indicated by Renshaw [9], of "scientifically" partitioning Accident years in some optimal fashion. Again, in practice, the extra sophistication is often thought not to reap rewards commensurate with the effort involved. One can, however, visually inspect the data to see if given Accident years are of a similar "level" of payment. In doing this, one may try and first normalise the data in some sense - for example, by adjusting the payments by an Exposure measure, such as the number of claims, or adjusting the data by an inflation index. One may then find that the accident year levels fall into a small number of groups. For example, if in a 10x10 triangle, the business written appeared to be of a different nature in the first five years compared to the last five years, the model for the a(i)'s may be of the form:

a(i) = A, for i =1,2,....5, and,

a(i) = B, for i = 6,7,....10.

A and B are constants, as are C,D,E and F below.

Revisions of the model may be examined by using the simple regression method on the Yij's, back in our Log-space. This will produce estimates of all the parameters involved and produce statistics as to the overall fit of the model, as well as to the significance of the individual parameters. When examining the fit of a model, the Residuals (the differences between fitted and actual values) are also examined, to observe whether the model exhibits any unwanted features. For a good model, we would expect that the Residuals are suitably "random", that is they do not exhibit any systematic pattern. The

Residuals can be examined to see how they vary by Accident Year, Development Year or Payment Year. Again, the interested reader is referred to the IOA Reserving Manual [1] for further information on Residuals and their characteristics and the actual process of fitting a model.

The second, and more fruitful, refinement one may make is to model the development parameters in Log-space, b(j), by a curve, or a combination of curves. Frequently one may find that given classes of business have their own particular "shape" in the first few years of development but adhere closely to a curve, or combination of curves thereafter. This can be determined by visually examining the data, to see what sort of families of curve suggest themselves. The same process of fitting the model to the data by regression, then examining the statistics regarding the fit of the model and the Residuals is gone through, as indicated above.

Typically, the revised model may take the form of:

$b(j) = C$, for $j = 1$

$b(j) = D$, for $j = 2$

$b(j) = f(j)$, for $j > 2$

There are a variety of types of curve, f(j), which may be tried. These include:

<u>Exponential</u>

$f(j) = E \times (j-2)$, for $j > 2$

In Log-space, this is just a straight line, indicating that the claims payments die away exponentially over time, after year 2.

<u>Power Curve</u>

$f(j) = F \times \log(j-1)$, for $j > 2$

This assumes that the claims payments decay according to a power curve, after year 2.

Hoerl Curve

$f(j) = E \times (j-2) + F \times \log(j-1)$, for $j > 2$

This can be seen to be a combination of the first two curves. It indicates that the decay of the claims payments is a combination of exponential and power curve. This curve proves to be an effective fit for a wide variety of classes of business.

One can extend the range of possible curves by fitting different curves to different sections of the development - for example an exponential decay with parameter E1 for $2 < j < 7$ and exponential decay with another parameter, E2, for $j > 6$. Also, there is no reason why other curves should not be fitted - if part of the development in Log-space looks like a quadratic curve, go and see what the fit and the Residuals look like !

A further possible extension is to include a calendar year effect. As an alternative to stripping out inflation before the model is applied, one can explicitly model inflation. The basic model would have the additional term:

$$\text{Log}(P_{ij}) = a(i)+b(j)+Y\times(i+j)+E_{ij}$$

where $Y = \log(1+u)$, and;
u = annual rate of inflation across the whole triangle

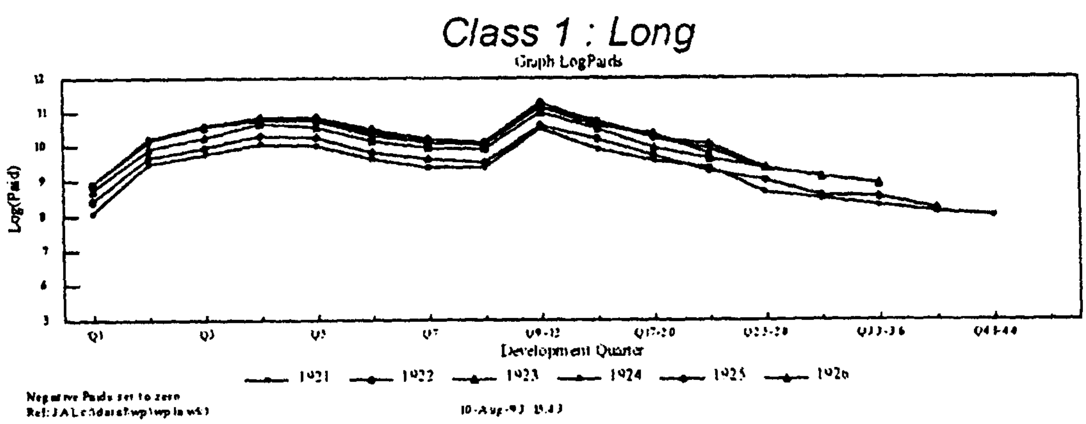
For all the curves above, it is just a matter of re-formulating the model and using regression in Log-space to fit different combinations of parameters to the data. The process of examining different curves and the various statistics for each, for the Working Party Database is described in 6.2.2.

6.2.2. Data Analysis

The following gives some general comments about the procedures that were performed in looking at the three classes. Rather than give a blow by blow account for each class, various stages of the modelling and diagnostic process will be illustrated for a sample class, rather than for all three classes. At the end of the section the final model arrived at for each class will be summarised and any particular features of the modelling as it applied to that class noted.

Looking at the data

We are modelling the logs of incremental payments, so the first step is to look at graphs of the log-incremental payments. Due to the limits of matrix manipulation referred to in section 6.2.3, we are restricted to looking at quarterly data at the start of the development and annual data thereafter, or the matrices that need manipulating become too unwieldy. The choice here was to look at quarterly data for the first eight quarters, and a variety of curves thereafter. The graphs to examine were therefore of the logs of the first eight quarters' incremental payments and the incremental annual payments thereafter, an example of which is given below:



We can justify looking at a combination of quarterly and annual payments as the total payments in each of the first eight quarters are of the same order of magnitude as the subsequent annual payments, so this does not cause us undue problems with our assumptions regarding the Residuals being uniform across the triangle - indeed it could be said to be a positive feature.

Other methods partition the development period (in this case ten years) into different time periods: this is just doing the same. It should be noted that we are not assuming anywhere that the sum of four lognormally distributed variables is also lognormal, although even if we were, such simplifying assumptions are made in many other areas. For example in option pricing one assumes that a stock price is lognormally distributed when applying the Black-Scholes model, and the same model is used when examining a stock index comprising many stocks. In option pricing models, just as in reserving and many other models, the model is not "reality", it is just a useful representation of reality. As long as one is aware of strengths and weaknesses in the assumptions, one can happily make simplifying assumptions.

Another advantage of grouping the quarterly payments into annual for later developments is that it reduces or largely eliminates the problem of negative claims. All of the classes had negative quarterly payments in the later stages of development, but none of the annual payments did.

Whilst it would be nice to be able to fit a model with an unlimited number of quarterly development points, as can be seen from the graph above, this does not present a severe set-back. For the vast majority of classes, the first eight quarters or so are likely to have their own particular shape, and any curve that is to be fitted to the data after that point could equally well be fitted to the annual as the quarterly data. Because the first eight quarters for the three classes did have their own particular characteristics, the model for three classes was just a piece-wise linear section for the first eight quarterly payments followed by different curves. For the example above it can be seen that these first eight points do have a very consistent shape. The development thereafter looks suspiciously like two straight lines (because this is a graph of log-payments, this means the actual payments decay according to two exponential curves), see the section on Class 1 for more details.

Fitting the parameters

Having examined the graphs of the log-payments to choose a model, we need to fit the model to the data using regression. If we have n past data points and p parameters, this means constructing an nxp matrix defining the model for each data point, along the lines set out in section 6.2.1.4. This matrix (called the Design matrix) is then our "X" range when performing regression using a standard spreadsheet package, the "Y" range being a column of the data points themselves. This process and the various diagnostic tests mentioned below are further described in the IOA Reserving Manual Volume II [1], so they will not be expounded on in any detail here.

The standard regression output of most spreadsheets then outputs the fitted parameters and their standard errors (SE's), along with an overall model variance figure and an R-squared statistic, which gives an indication as to how good the fit of the model is. Ideally we would like the model variance to be as low as possible, and the R-squared to be as near to one as possible. One cannot just focus on these numbers however, as one does not want to over-parameterise the model and it is important to look at the Residuals as described subsequently. One can also look at the T-ratio of the various parameters - the parameter divided by its' SE. If the parameter is significantly different from zero, we would expect the absolute value of the T-ratio to be bigger than two.

A typical section of regression output from the model is reproduced below:

Class 1 : Long Design 5 - Unique level first eight quarters,Hoerl curve therafter

Regression Output:								
Constant		0						
Std Err of Y Est		0.090						
R Squared		0.989						
No. of Observations		112						
Degrees of Freedom		92						
	a1	a2	a3	a4	a5	a6	a7	a8
X Coefficient(s)	11.522	11.861	12.045	12.090	12.034	11.435	10.940	10.768
Std Err of Coef.	0.095	0.097	0.098	0.099	0.098	0.097	0.097	0.097
T-Ratio	120.794	122.129	122.675	122.664	122.467	117.310	113.227	111.382

Grouping the Accident year levels

The T-ratios for the accident levels (a1,....a8 are illustrated above) do not in themselves mean anything other than that accident year has payments significantly different from zero - no great surprise !! To see if one can group the accident year levels, the accident year parameters are visually examined. If some of the parameters look similar, say we thought a3=a4=a5 above, we could amend the Design matrix so that instead of separate levels a4 and a5, those accident year levels were described as a3+C, and a3+D, where C and D are just two constants. The regression could then be performed again. This time, if the T-ratios for C and D are greater than two, it means that the accident year parameters for a4 and a5 are significantly different from a3. If not, then those accident year levels can be taken to have a common level. More rigorous treatments of this stage of the modelling process can be made (although they require rather more than a simple spreadsheet and a few minutes of time!) and are referred to in section 6.2.1.4.

For the classes examined, the most likely grouping appeared to be between some or all of the first five accident years and some or all of the remaining five accident years. Various combinations were tried, but the groupings were at best of two accident years with a tentative level of significance. This is not altogether surprising - the business has been changing and inflationary forces have been increasing the level of claims payments.

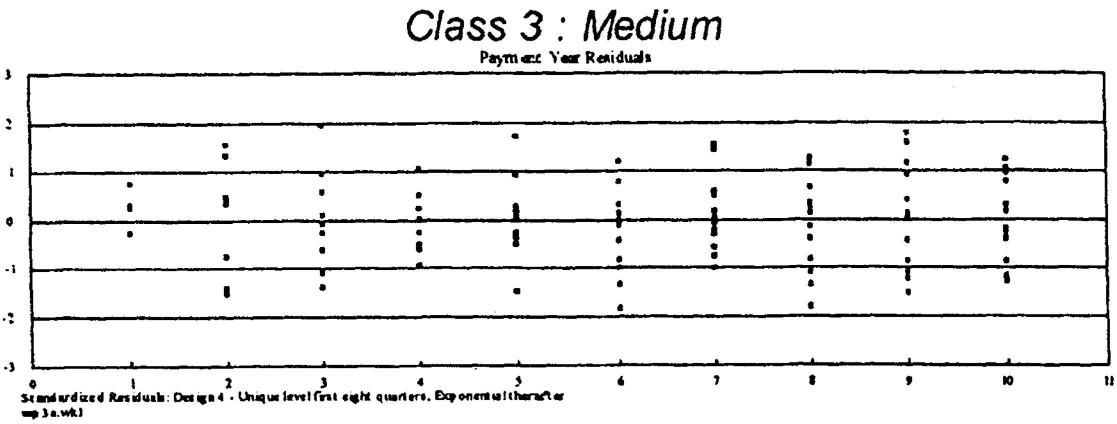
To try and increase our chances of reducing the number of accident year parameters, the data triangles were scaled by one or both of an inflation index across the payment direction (i.e. the diagonals) and an exposure index in the accident direction (i.e. down the columns). The inflation index was arrived at by a simple separation-type process - in practice this could draw on outside knowledge of the classes of business concerned. The exposure index was taken to be the ultimate number of claims, as determined by a simple chain ladder method.

For each of the classes, a slightly less parameterised model could now be found, either by grouping accident year levels for 1923-1926 or 1928-31. Again the significance was tentative and the overall fit of the model was not drastically improved, or in some cases worsened. The future payments would have to be adjusted by an estimated future inflation index adding further uncertainty. Given the marginal improvement in the model and the additional uncertainty arising from the exposure and inflation indices, it was decided to leave each of the classes being modelled by ten separate accident year parameters.

Looking at the Residuals for different curves

The three types of curve outlined in section 6.2.1.4 were examined for each class. The only variations that suggested themselves were to try a double-exponential curve for Class 1 with various points at which the two curves joined. For a given curve, various plots of the Residuals were examined, as well as the various statistics outlined above. Once the basic Design matrix for a model had been set up, it could easily be applied to all three classes, so each curve that was looked at for one class was also looked at for all the other classes too.

The Residuals examined took two basic forms, with variations on each. The first was a standard scatter plot, showing the distribution of the residuals in the three "directions" implicit in the triangle - development, accident year and payment. We expect roughly one in twenty Residuals to be outside the range -2 to +2. We also expect the Residuals to be suitably random. Any non-randomness may indicate defects in the model, or suggest refinements to the model. The Residuals for class five and examination of the data for that class generally suggested refinements to the model other than using the directly calculated parameters. An example of this type of Residual output is given below:

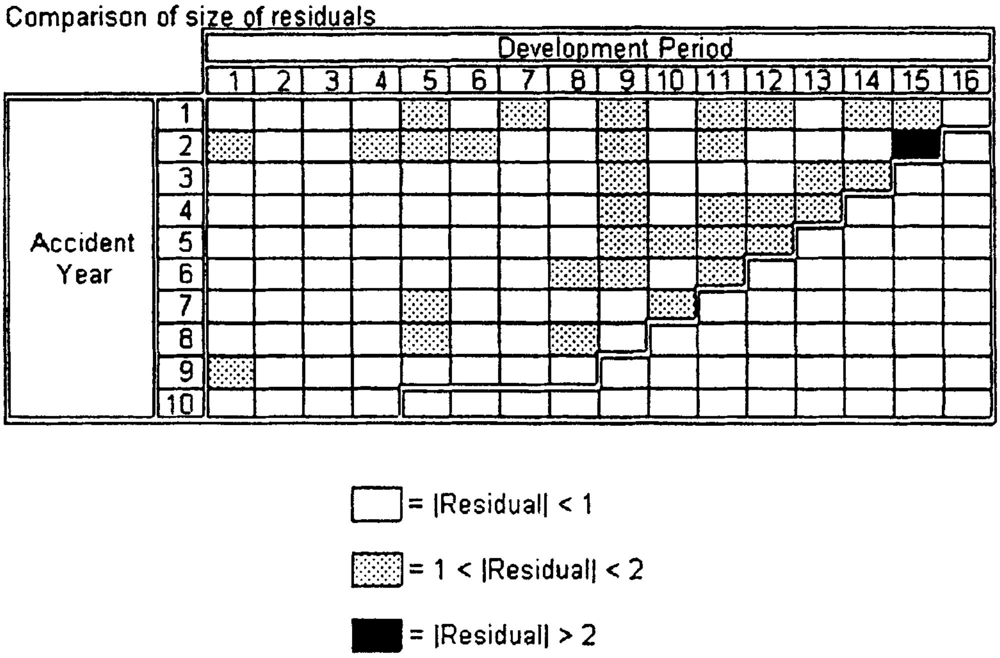


If one particular point looks way off-beam compared to the surrounding points, its' effect on the model can be ascertained and if necessary that point can be removed from the regression in arriving at the model. Care needs to be taken in doing this, so that salient features of the data are not obliterated without good reason. The author of this text prefers to give the data "the benefit of the doubt" rather than rush to remove outlying Residuals wherever they may pop up, unless there are strong reasons to the contrary.

A further aid to modelling was the following type of Residual plot:

Class 3 : Medium

Standardized Residuals: Design 4 - Unique level first eight quarters, Exponential therafter
Comparison of size of residuals



Another such Residual plot showed the sign of the Residuals, either positive or negative. This distribution of the size of the Residuals in a "triangular plot" such as that shown above gives a useful visual summary of the model and can often bring out features of the fit of the model not instantly obvious from the statistics or scatter plots of the Residuals.

The other Residual plot showing the sign of the Residuals can also be useful but needs to be treated with a little more care. One particularly large outlying Residual can "pull" the parameters towards it, so that all the Residuals along, say, the particular accident year in which the Residual resides are pulled towards it, producing a row of Residuals which tend to be of one sign or the other, possibly by only a small margin.

Validation

Once one has arrived at a model one is happy with, one can and should try and see what the model would have predicted had it been applied a number of years ago and compare this to the known outcome. Only for class five did this, and other checks, lead to the model being refined slightly.

The Tail

One advantage of modelling claim payments by curves is that the curves can be extended beyond the triangle of data to give an estimate of the tail. As with any projection beyond known data, this involves something of a leap of faith that the trends in the first ten years will continue thereafter. For some classes this is clearly not the case - particularly for business such as that exhibited by class one, where the payments decay up to a point but then continue at pretty much the same level as claims continue to be reported at pretty much the same level as they are closed.

A disadvantage of some of the curves used in this type of modelling is that the exponential component representing claims payments decaying over say years seven to ten, tends to die away too rapidly thereafter, and the model does not add greatly to the estimation of the tail. At this stage, one can review the projected tail payments relative to the known case estimates and the rates at which claims continue to be closed and reported.

The tail calculation chosen was an extension of the curves in all three cases, but with the decay parameter modified to reflect, where appropriate, the expected slow down in the decay of paid claims by reference to the case estimates and rates of closure and reporting in the later stages of development. The payments were extended by a further twenty years. The Standard Error of the reserve could only be calculated for the reserve up to year thirteen. This SE as a percentage was applied to the total reserve including the tail - because of the extra uncertainty in this area of the calculation, this SE is likely to be understated.

For classes such as class one, the payments after thirteen years are a few thousand a year and the reserve may be of the order of twenty times this amount, representing payments continuing at the same sort of level for several more decades. To put these problems into perspective, for practical internal purposes such claims are likely to be discounted, and hence their significance reduced. They may well be included implicitly in special reserves, not specific to a given class or accident year, held in respect of latent claims or for claims where the future development at such advanced stages cannot accurately be quantified with any degree of certainty.

Class 1 - the model

The initial graph of the log-incremental payments suggested that curves after year two consisting of two exponential curves (i.e. two straight lines in the log-space) should be examined. Although providing a good fit, none of these pairs of exponential curves provided as good a fit as a Hoerl curve.

The Hoerl curve Residuals had a few spiky areas, but these were not felt to be sufficiently material to alter the basic model. A Power curve with a reduced number of accident year levels provided equally good fit statistics and higher levels of significance for the parameters, but the Residuals were clearly unsatisfactory.

The tail calculation is a weakness for this type of model, as for later accident years the tail is many times greater than the reserve up to development year thirteen to which the model was extended (practical considerations again limited the number of years one could extend the analysis in a spreadsheet). The total tail calculated is of the same order as the rest of the total reserve for this class.

The final model, to year thirteen, consisted of twenty parameters:

a1-a10	Ten Accident year level parameters
b1-b8	Eight initial quarterly payment parameters in the first two years
A,B	Two Hoerl curve parameters for development after year two

A further parameter was the adjustment to the decay rate in the tail.

This compares to, say, the forty-nine parameters one would have used in the basic chain-ladder for such a ten by forty triangle.

This model had a very low Standard Error and a high R-squared of 0.989. This reflects the very consistent incremental payments over the bulk of the triangle, and we would expect to be able to make good predictions of the overall payments in the forthcoming years for such a model. The model may not fare so well when considering the tail of the distribution.

Class 3 - the model

Both the Exponential and the Hoerl curve provided good levels of fit to this data, but not all the Hoerl curve parameters were significant and so the Exponential curve was preferred.

The final model had nineteen parameters:

a1-a10	Ten Accident year level parameters
b1-b8	Eight initial quarterly payment parameters in the first two years
s	One Exponential parameter for development after year two

A further parameter was the adjustment to the decay rate in the tail.

This model had a fairly low Standard Error and an R-squared of 0.987. Again this reflects fairly consistent incremental payments over the bulk of the triangle. Some of the tail projections for the older accident years only projected payments slightly above or below the case estimates. This feature was left in the final reserve estimates as the total Incurred (paid plus case estimates) position showed little or no development in the final years. From the follow up data three years on, two of the last four accident years in fact showed a decline in the total Incurred position over this period.

Class 5 - the model

This class proved the hardest to fit. Whereas classes one and three showed either rapid growth then gradual decline or gradual growth then sharp decline, this class exhibited both rapid growth and decline and accompanying mayhem in the payment, closure and reporting of claims. The class, even if stable, is of a volatile nature being long tail and the subject of some very large claims indeed.

This Hoerl family of curves is the only one that seemed to fit the class adequately. This led to the potential model with the lowest number of parameters, as several accident year levels and a group of the first eight quarterly, payment parameters were amenable to being grouped together leaving an eleven parameter model. This had some slightly unsatisfactory Residuals and seemed too few a number of parameters to describe a large number of data points.

The twenty parameter model based on the Hoerl curve was chosen, but with slight refinements to the two Hoerl parameters to reflect the apparent change in payment patterns. The initial parameters, fitted by reference to the entire triangle, produced payments that came through too quickly in the forthcoming years then died away too rapidly, as inferred from the ebbs and flows of the payment and reporting patterns across the accident years. The final model still probably suffers from the drawbacks described in class one as to the too rapid decaying of the payments in the tail.

The final model had the same number of parameters as class one.

Revisiting the models in three years time

For each class the original model was re-fitted to the updated data. Depending on the difference between experience and expectation, and the fit to the latest data, the original parameters were kept or the new ones adopted, in formulating a revised estimate of the ultimate claims.

Class 1

Overall the predicted payments were very close to the actual payments. Although the band for predicted payments was only 2.5% either side, the total payments were within this band. The later accident years actual payments were on the higher sides of expectation however, as suspected. The re-fitted model was very similar to the initial model, but boosting the payments slightly. These new parameters were adopted, with a slightly "thickened" tail decay parameter. This was the only class where the revised estimate of ultimate claims was slightly higher than previously.

Class 3

Again overall the actual total payments were in the middle of the predicted payments band. The model parameters were virtually unchanged from the original model and the slightly revised parameters were adopted. The tail calculation was now changed to a default of the case estimates if greater for the last few accident years - there are clearly a few large claims outstanding and the best estimate of these seems to be the case estimates themselves. The Incurred position for the older accident years had remained at or about the same level as three years previously, and in some cases decreased, so this seemed a satisfactory compromise. The estimates of ultimate claims remained virtually unchanged for this class.

Class 5

The fitted parameters had moved towards the refined Hoerl curve parameters originally adopted. Given the varying and variable forces acting on this class, the model seemed to have produced good estimates of the outcome over the three years as a whole and the original parameters were retained. With some reservations, the original tail parameters were also retained.

The model as applied to the latest data produced the same levels of overall ultimate claims, whilst increasing some of the latest accident years and decreasing some of the earlier ones.

6.2.3. Practical Limitations

All the analysis in section 6.2 was performed using a Lotus spreadsheet. There are various limitations of spreadsheets generally when it comes to matrix manipulation or performing regression analysis.

Part of the process of arriving at the variance-covariance matrix referred to in section 6.2.1.4 involves manipulating matrices. In most common spreadsheets there is a limit to the size of matrix that can be multiplied or inverted. In Lotus or Excel this is about an 80x80 matrix. This limits the number of future values the model can project to 80 points. Clearly for a ten accident year triangle, this is quite a considerable practical constraint. However, it still allowed the models fitted to be projected to development year thirteen, which is adequate for most purposes.

When performing the regression, there is a limit in the standard Excel regression facility to the number of dependent variables of seventeen. For this reason Lotus was chosen in preference, as the limit is seventy-five. The regression could of course have been performed using matrices in Excel.

6.3. Log-Incremental Claims II

6.3.1. Overview of method

The basis of this extension is the previous section where we considered the process of modelling the log incremental claims. We are reminded that the stochastic framework does not restrict itself to cash flows and these approaches can be applied to other triangles of claim data. The extensions are:

1) The inclusion of an additional parameter to reflect the trends between every two contiguous payment years.

2) The relaxation of the assumption that all claim payments are i.i.d. normally distributed random variables. Thus the variances are allowed to differ by development period. This extension may be desirable because variances often change over development periods.

3) Varying parameter modelling. In the preceding descriptions it was commented that b(0) is set to 0 to reduce equations to be solved. Through the use of varying parameter modelling this problem of multicollinearity (more unknowns than equations) is reduced. Varying parameter modelling is where we are able to include relationships between the parameters. Thus we have dynamic parameters by using exponential smoothing or credibility weighting. To solve the subsequent equations we use generalised least squares. The Kalman filter is a generalised least squares algorithm. [16]

The model does not purport to represent the underlying claim generating processes. This approach is advocated on the grounds that the multitude of variables involved in generating the claims are invariably complex and to attempt to model all the underlying processes might lead to an inefficient, impractical and potentially incorrect forecast.

Dr Ben Zehnwirth has developed a commercially available computer package called ICRFS (Interactive Claims Reserving and Forecasting System) which embodies these techniques. It has been used to analyse the test data. [17]

6.3.2. Data Analysis

In general, the same techniques were used in fitting the parameters as outlined in section 6.2.2 for log-linear regression, the main exception being that annual data only was used. In practice the modelling process would include full interaction with a reserving specialist who has business knowledge of the accounts. In this artificial environment no information other than the actual triangles was provided.

Although for each of Class 1, Class 3 and Class 5 a number of loss development arrays were available we only analysed the incremental payments and the closed claim counts. The latter were analysed for the purpose of determining whether .

1 Any instability in trends in the incremental payments can be "explained" by changes in the speed of settling claims.

2. The claim counts are more or less predictable than the incremental payments.

Accident years

More informed decisions about the future could be made if accident years exposure bases were made available. The three incremental payment arrays present similar accident year trends. That is the trends tend to increase until 1926 at which point they decrease until 1928/1929. These changing trends may well be explained by changing exposures.

Payment years

The three trends obtained for each model during the modelling process are diverse ranging from 0% for Class 1 to 10.6% for Class 5 (even after adjusting for accident year trends). An interesting result was that apart from Class 5 where there appeared to be a slight increase in inflation in the later years the payment year trends appeared stable. A different picture was obtained when looking at the data gross of salvage and subrogation but this was outside of the scope of the working party's analysis.

Development Period factors

From inspection a model utilising probabilistic development factors was adopted rather than a smooth curve model.

Closed claim arrays

Closed claim arrays were analysed in order to determine whether instability of trends in payments are 'caused' by changes in speed of settlement. From the analysis this did not appear so. Moreover, the closed claim arrays appeared less stable than the corresponding incremental payments.

Projecting beyond the triangle

In order to estimate the ultimate claims it is necessary to assume a pattern of claim payments beyond the confines of the triangle. In the absence of additional information it was decided simply to continue the latest probabilistic development factor into the future and to assume the same inflation factor into the future. This is consistent with an assumption of long tail for Class 1, medium tail for Class 3 and long tail for Class 5.

Validation

An integral part of the modelling process is validating and testing the model for stability.

This is performed by assigning zero weight to.

1) The last payment year 1931

2) The last two payment years 1931 and 1930.

i.e. we investigate whether the model would forecast the distributions of (incremental) payments for the last two payment years, had we used the model structure at year end 1929 and moreover tested the estimates of outstanding payments for stability.

The validation analysis also aids in determining the most appropriate assumptions for the future.

Validation: Class 1

The validation table is as follows.

Years Included In Estimation	Future Payment Year Trend %	Forecast To The End Of Triangle
1922-1931	0 ± 0	$367{,}243 \pm 9{,}976$
1922-1930	0 ± 0	$362{,}301 \pm 15{,}180$
1922-1929	0 ± 0	$303{,}256 \pm 28{,}639$

The validation for 1922-1929 illustrates that we would not have been able to forecast the tail of the triangle (up to the development year 9) from the development pattern up to year 7. This may be expected and does not cause concern but should be noted for the purpose of estimating the tail beyond development year 9.

The standard error for the payment year trend was set to zero. This was done after examining the test statistics which indicated that the parameter for payment year trends was not significant. This view was taken as a result of the modelling process and would require investigation.

The payment year trend was assumed from the payment year 1923.

We have estimated a base development year trend along development years 7-9 of -21.05% ± 2.57%. So for the future we are assuming that the mean base development year trend is -21.05% and the standard deviation of the trend is 2.57%.

The total number of parameters used was 7.5 (the fraction is a consequence of as a result of using dynamic parameters through the use of the Kalman filter).

Validation: Class 3

The validation table is as follows:

Years Included In Estimation	Future Payment Year Trend %	Forecast To The End Of Triangle
1922-1931	2.44 ± 1.27	806,550 ± 26,972
1922-1930	1.73 ± 1.40	772,795 ± 29,778
1922-1929	1.74 ± 1.70	776,243 ± 41,279

The poor validation was a result of slight changing payment year trends. In a practical environment the decision as to which trend is more appropriate for projection purposes is required. We were not in a position to justify any change in the model.

The payment year trend was assumed from the payment year 1923.

We have estimated a base development year trend along development years 5-9 of -70.85% ± 2.10%. So for the future we are assuming that the mean base development year trend is -70.85% and the standard deviation of the trend is 2.10%. This would be in accordance with the business being medium term.

The total number of parameters used was 6.8 (as a result of using dynamic parameters)

Validation: Class 5

The validation table is as follows:

Years Included In Estimation	Future Payment Year Trend %	Forecast To The End Of Triangle
1922-1931	10.74 ± 2.49	875,767 ± 57,665
1922-1930	12.08 ± 3.19	874,564 ± 90,675
1922-1929	19.88 ± 4.12	1,081,002 ±174,262

The validation result when we removed two payment years was a result of a change in the inflationary trend (as can be seen). In a practical environment the decision as to which trend is more appropriate for projection purposes is required. Again we were not in a position to justify any change in the model.

The payment year trend was assumed from the payment year 1923.

We have estimated a base development year trend along development years 6-9 of -39.03% ± 3.55%. So for the future we are assuming that the mean base development year trend is -39.03% and the standard deviation of the trend is 3.55%.

During the modelling process it was discovered that the development pattern for development years 0 to 1 for accident years 1922 to 1927 exhibited a different pattern from those of 1928 onwards. The decision was taken that these observations would be weighted out because the underlying development pattern had changed and the results presented are from the model created using these weightings. Without this allowance the model would have been very unstable.

The total number of parameters used was 6.0 (as a result of using dynamic parameters).

Comparison with next three years payments

Normally this process would be carried out annually and in a practical environment where our analysis indicates we would modify the model to incorporate later years information.

On initial inspection Class 1 and Class 3 appeared to forecast reasonably well in comparison with the actual payments. Class 5 did not appear to compare well and this was due to the inflation parameter, which was identified as a problem during the validation process.

6.3.3. Practical Limitations

In theory one could perform the calculations in a spreadsheet, as indicated in 6.2.3 for the log-Incremental method. However use of dynamic linear modelling, such as the Kalman filter, introduces a level of complexity which would be difficult to program and run in a reasonable timeframe. Therefore in practical terms commercial packages are required.

6.4 An Operational Time Stochastic Model

6.4.1 Description Of Method

This model has been developed by Tom Wright [18]. The model attempts to (partially) represent the underlying claim generating process. It starts with the premise that the cost of settling claims depends on the order in which they settle. Typically, for example, later settled liability claims cost more.

The method therefore develops a model of the claim settlement cost, as a function of the relative proportion of claims settled.

The method is likely to be of greatest use in circumstances where the greatest source of variation in predicting ultimate claim cost is due to the individual claim costs e.g. in motor bodily injury.

Ultimate numbers of claims are required, and timing of cash flows are derived from a given settlement pattern for claim numbers. The method gives the individual expected cost of each claim.

The concept of operational time can trace its origins back to the model developed by Harry Reid in 1979 [19], which was done in a non stochastic framework. The concepts were further developed by Taylor [20], [21].

6.4.1.1. Data

There are two base data triangles required, one containing claim numbers and a second containing loss amounts. Generally, these may take one of three forms:

	Claim Number Triangle	Claim Amount Triangle
a)	The number of claims closed	total of all payments on claims closed with part payments assigned to the development period of closure
b)	The total number of payments, including part payments	usual paid loss triangle, with each part payment assigned to the development period in which it was made.

c)The number of claims closed	usual paid loss triangle, with each part payment assigned to the development period in which it was made.

Data in formats a) and b) are equivalent from a modelling point of view. Format c) requires a more detailed model which may make effective use of an additional triangle, namely one containing numbers of claims outstanding. Format c) is the one usually encountered in practice.

6.4.1.2. Notation

Throughout the rest of this note, the following notation is used:

Subscripts:	w	Year of Origin
	d	Development period
	τ	Operational time
Triangles:	$N_{w,d}$	Number of claims closed
	$Y_{w,d}$	Paid loss amounts
	$X_{w,d}$	Random variable of individual claim amounts
	$S_{w,d}$	Observed average claim amounts (that is, $Y_{w,d} / N_{w,d}$)
	$\tau_{w,d}$	Average operational times

Estimated ultimate number of claims:	M_w
Mean claim amount in real terms:	m_τ

6.4.1.3. Operational Time

Operational time (τ) is defined as the proportion of all claims closed to date. Thus for each origin year, operational time starts at 0 and increases ultimately to 1. Transformation into operational time eliminates the need to model settlement rates. Use of operational time overcomes a major problem with stochastic modelling in development time. It is often the case that large claims take longer to settle than small claims, for this reason we model m as a function of τ. When modelling in development time, because the time to settlement for an individual claim is uncertain, the appropriate claim size distribution for that claim is also uncertain. Whilst it is not difficult to calculate the expected value of projected future claim payments, the calculation of standard errors is extremely complex (except in the special case where the claim size distribution does not vary with delay).

6.4.1.4. Claim Numbers

The first step in the modelling process is to estimate the ultimate number of claims M_W and their standard errors. Where triangles are compiled on a notification year basis, the number of claims is known, that is, it is equal to the number reported. Where data is analysed on an accident or underwriting year basis, then the expected ultimate number must be estimated by another method. The possible methods to obtain these estimates and their standard errors are not dealt with in this paper.

6.4.1.5. Initial Assumptions

In order to clarify the explanation of the modelling process, we make some initial assumptions which will be relaxed later. All these assumptions may be tested by use of residual plots and other diagnostic tests, they are not general restrictions on the validity of the model

(i) The expected claim size in real terms m_τ is the same for all years of origin, that is, m_τ does not depend on w.

(ii) The coefficient of variation (φ) of individual claim amounts is the same for all operational times, that is:

$$Var(X_\tau) = \varphi^2 . m_\tau^2$$

(iii) The data $Y_{w,d}$ is not affected by inflation

(iv) The standard error of the ultimate number of claims is zero

(v) Part payments are not present in the data triangles, that is, the data is of type a) or b) as set out in section 3.

6.4.1.6. Modelling Under Initial Assumptions

A triangle of average operational times may be calculated as:

$$\tau_{w,d} = (N_{w,1} + N_{w,2} + \dots + N_{w,d-1} + ½ N_{w,d}) / M_w$$

A triangle of average claim amounts may be calculated as:

$$S_{w,d} = Y_{w,d} / N_{w,d}$$

In order to project future claim payments, we need a model for m_τ. This is achieved by fitting models to the sample means $S_{w,d}$. To fit these models we need expressions for the mean and variance which may be derived from the initial assumptions.

$$E(S_{w,\tau}) = m_\tau$$

$$Var(S_{w,\tau}) = \varphi^2.m_\tau^2 / N_{w,d}$$

It is not necessary to have any further knowledge about the distribution of S in order to fit models of generalized linear form. Use of this form allows great flexibility in the model for m_τ. Use of a log link function (see Appendix 1) and a variety of terms in the linear predictor enables the following example models to be tested:

	Model	Terms in the Linear Predictor			
(i)	$m_\tau = \exp(\beta_0 + \beta_1 .\tau + \beta_2 . \log(\tau))$	1,	τ,	$\log(\tau)$	
(ii)	$m_\tau = \exp(\beta_0 + \beta_1 .\tau + \beta_2 . \tau^2)$	1,	τ,	τ^2	
(iii)	$m_\tau = \exp(\beta_0 + \beta_1 .\tau + \beta_2 .\tau^2 + \beta_3 .\tau^4)$	1,	τ,	τ^2,	τ^4

These and other models may be fitted to the observed data points $(S_{w,d}, \tau_{w,d})$, the fitted models extending over the range (0, 1)

6.4.1.6.1. Model Zero

The modelling process starts with the fitting of a deliberately over parameterised model (model zero) which consists of a piece-wise exponential function of τ. The number of sub intervals can be chosen to make the model as flexible as desired. This model may be used to test assumption (ii), the variance assumption, and also to quantify the amount of random variation in the data. This enables subsequent F tests to determine the best model. The variance assumption is tested by examining plots of standardised residuals against operational time. If m_τ has been fitted using model zero, then the variance of the standardised residuals, $Var((S_{w,\tau} - m_\tau).\sqrt{N_{w,\tau}} / m_\tau)$ equals φ^2, which does not depend on τ. Hence if the pattern of the residuals does not vary with operational time, then the variance assumption may be reasonable. If this is the case, and residual plots against origin and payment periods also look reasonable, then the fitted model zero may be used to quantify the random variation inherent in the data.

6.4.1.6.2. Model Selection

Once model zero has been validated, other models for m_τ may be fitted and the residuals checked for trends against operational time. F tests can be used to help find the best model by identifying those models with the best compromise between a) relatively few, and more accurately estimated, parameters, and b) residual variation which is not much greater than the purely random variation identified in model zero.

6.4.1.6.3. Prediction

If a suitable model can be found, then the expected value of each future claim can be obtained by evaluating the fitted value of m_τ. The variance of each future claim may be obtained by evaluating $\varphi^2.m_\tau^2$ using the estimated values for φ and m_τ. Assuming the amounts of future claims are statistically independent, the mean and variance of the total may be calculated, augmenting the resulting variance to allow for estimation error in the fitted means m_τ .

6.4.1.7. Relaxing The Initial Assumptions (✱)

(i) The expected claim size in real terms m_τ is the same for all years of origin, that is, m_τ does not depend on w.

This assumption may be relaxed by allowing the β_0 parameter to vary with origin year. In practice, at most, only two or three levels of this parameter are required for most data triangles. Whilst this allows different groups of origin years to have different levels of m_τ, it is still assumed that the pattern is the same.

(ii) The coefficient of variation (φ) of individual claim amounts is the same for all operational times, that is:

$$Var(X_\tau) = \varphi^2.m_\tau^2$$

This may be replaced by:

$$Var(X_\tau) = \varphi^2.m_\tau^\alpha \quad \text{for some } \alpha$$

This allows for the coefficient of variation of individual claims to depend on the mean claim size. If examination of the residual plots against operational time for model zero with $\alpha=2$ suggests that the variance is decreasing, then the model may be refitted using a smaller value.

(iii) The data $Y_{w,d}$ has been adjusted to remove the effects of inflation

An overall rate of inflation may be simultaneously estimated as part of the modelling process by inclusion of an extra parameter. If i represents the annual force of inflation and p represents the number of development periods per year, then the example models in 6.4.1.6 above, become:

Model	Terms in the Linear Predictor
(i) $m_\tau = \exp(i.(w+d/p) + \beta_0 + \beta_1 .\tau + \beta_2 . \log(\tau))$	$w+d/p$ 1, τ, $\log(\tau)$
(ii) $m_\tau = \exp(i.(w+d/p) + \beta_0 + \beta_1 .\tau + \beta_2 . \tau^2)$	$w+d/p$ 1, τ, τ^2

When fitting these models, the parameters estimated are (i, β_0, β_1, β_2). Incorporating future claim inflation in the projections involves additional calculations to quantify the variation due to uncertainty in the future rate of claim inflation and uncertainty in the real time scale of the run off.

(iv) The standard error of the ultimate number of claims is zero

The estimated ultimate numbers of claims M_W are used for two purposes: a) for calculating the triangle of operational times, and, b) in calculating estimates from the fitted model. Provided the estimates M_W are unbiased and not highly correlated, because the model is fitted to the whole triangle simultaneously, most of the variability from source a) is already taken into account in the fitted scale parameter and any additional variability can reasonably be ignored. The additional variability arising from source b) can be quantified for each origin year in terms of a standard error u:

$$u = \left(\tau_1 . m_1 + \frac{\hat{\mu}}{\hat{M}} \right) . v$$

where: $\hat{\mu}$ is the expected total of future payments for the origin year, calculated by summing m for each expected future claim

$\hat{M}$ is the estimated ultimate number of claims for the origin year

τ_l is the latest operational time for the origin year

m_l is the fitted mean value corresponding to τ_l

v is the standard error of the estimate

The expression in brackets above, is a weighted average of the fitted value at time τ_l, (m_l) and the mean value of future claims: $\hat{\mu} / (\hat{M} - N_l) = a$. That is, the expression in brackets equals $\tau_l . m_l + (1 - \tau_l) . a$

(v) Part payments are not present in the data triangles, that is, the data is of type a) or b) as set out in section 3.

Where data is of type c), the model may be extended to allow for part payments as outlined below

6.4.1.8. Modelling Part Payments (*)

If m_τ represents the average cost of closed claims, but the observed data contains part payments, then the data $Y_{w,d}$ has been increased by the amounts of these part payments. This extra amount may be expressed as the number of part payments multiplied by an average cost. If we express the number of part payments as a constant proportion (c_1) of the number of claims outstanding; and the average amount as a proportion (c_2) of the average cost of closed claims (m_τ), we have an expression for the additional amount arising from part payments, that is:

$$\text{Expected Number} \times \text{Mean Amount} = (c_1 \,.\, L) \,.\, (c_2 \,.\, m_\tau)$$

Where L is the average number of outstanding claims corresponding to $Y_{w,d}$. Expressing this as an average amount per closed claim, and combining the constants c_1 and c_2 into a single value, c, we have the amount derived from part payments per closed claim equal to:

$$c. (L/N). m_\tau$$

The constant c represents the expected part payment per outstanding claim as a percentage of the average cost of claims closed. Thus, expressing the ratio $L_{w,d}$ / $N_{w,d}$ as $R_{w,d}$ we have:

$$E(S_{w,d}) = (1 + c.\, R_{w,d}) \,.\, m_\tau$$

The constant c is usually small, typically around 0.1. This is because the number of part payments per outstanding claim is usually small (say 0.2), and the average cost of those payments is often less than the average cost of closing payments (say 0.5); hence multiplying these two factors together produces a small value for the c parameter. Approximating ($1 + c.\, R_{w,d}$) as $\exp(c.R_{w,d})$, this model can simply be built into the model and the c parameter estimated from the data as part of the fitting process; making use of a revised model for $Var(S_{w,\tau})$, namely:

$$Var(S_{w,\tau}) = \varphi^2.m_\tau{}^2 / (\exp(c.R_{w,d}) \,.\, N_{w,d})$$

Returning to our example models used earlier, we now have :

Model	**Terms in the Linear Predictor**
(i) $m_\tau = \exp(c.R_{w,d} + i.(w+d/p) + \beta_0 + \beta_1 . \tau + \beta_2 . \log(\tau))$	$R_{w,d}$, $w+d/p$, 1, τ, $\log(\tau)$
(ii) $m_\tau = \exp(c.R_{w,d} + i.(w+d/p) + \beta_0 + \beta_1 . \tau + \beta_2 . \tau^2)$	$R_{w,d}$, $w+d/p$, 1, τ, τ^2

The vector of parameters estimated becomes: $(c, i, \beta_0, \beta_1, \beta_2)$.

For some lines of business, it is unlikely that the rate at which part payments are made, or their average costs as a percentage of closed average cost, remains constant across operational time. This sort of change is accommodated within the same sort of model described above; the effect is usually to make m_τ increase less rapidly, or even decrease, as operational time approaches 1.

6.4.2. Data Analysis

General Points

- Modelling in operational time is designed to give accurate estimates of the expected cost of each future claim. The timing of those claims settlements is an input to the model and is used to apply the effects of future inflation.

- In all three data sets, the data was modelled quarterly giving 220 data points in each triangle.

- No information was available as to the lines of business, claim types or actual accident periods. This limited the reliability of the analysis. In particular, no information was provided on the possible impact of part payments in the data. For this reason, the modelling was carried out ignoring the possibility of part payments. Including part payment parameters in the modelling may have improved the reliability of the estimates, particularly for data set 5.

- All three data sets have the feature that over 50% of claims (by number) are settled in the first year and at significantly lower average cost than subsequent claims. Improved models may have been obtained by fitting to development after the first 4 quarters, which could have reduced the numbers of parameters required to fit the models reliably.

- Operational time (and most other) models are more reliable where different claim types are analysed separately. Data set 1 appears to be subject to latent claims which may not be fully reflected in the settlements to date as paid and incurred development do not appear to be totally consistent.

Data Set 1

There is significant development of claim numbers in the tail, and these claims are generally settled at high cost. The estimated reserves are dependent on the accuracy of the claim number estimates.

Numbers of claims reported were modelled using a number of methods and an ultimate selected for each accident year.

Model zero was fitted using one level parameter for all origin years and seven intervals in operational time. Examination of residuals led to a total of 6 origin year parameters being fitted. This is an unusually large number for models of this sort which suggests fundamental changes are occurring in the average claim sizes in this data set. The origin year groups and operational time intervals fitted were:

Origin Year	Groups	Operational Time Intervals
1922	1	$0.000 \leq \tau < 0.179$
1923	2	$0.179 \leq \tau < 0.363$
1924	3	$0.363 \leq \tau < 0.556$
1925	3	$0.556 \leq \tau < 0.791$
1926	3	$0.791 \leq \tau < 0.894$
1927	4	$0.894 \leq \tau < 0.948$
1928	5	$0.948 \leq \tau \leq 1.000$
1929	6	
1930	6	
1931	6	

The first graph shows the fitted model zero and the data points. It is clear that over 60% of the claims are settled in the first development year, but the average cost of those claims is around 1/20 of the average cost of claims settled in the tail. This suggests that the data may contain a mix of claim types possibly with a changing mix for different accident years.

The final fitted model replaced the 7 straight lines of model zero with 6 polynomial coefficients. The relatively large number of parameters was needed to cope with the sudden increase in average cost at operational time 0.8. The 13 fitted parameters are shown below together with their standard errors. This model has 207 degrees of freedom and is plotted in the second graph.

		Estimate	S.Error
Past Force Of Inflation:		0.090	0.008
Origin Year Group	1 :	0.059	0.088
" " "	2 :	0.154	0.080
" " "	3 :	0.314	0.065
" " "	4 :	0.171	0.055
" " "	5 :	-0.155	0.049
" " "	6 :	-0.348	0.037
Predictor Term τ	:	4.365	0.485
" " τ^2	:	-13.219	1.869
" " τ^4	:	62.930	8.816
" " τ^6	:	-309.011	42.377
" " τ^7	:	428.698	59.060
" " τ^8	:	-169.903	24.140

The residual plots against origin year show a greater spread for 1929 and 1930 than for earlier years. This may be indicating that the pattern of average costs is changing. However, changes in the claims settled in the first two years may not be a reliable indicator for changes in claims settled in the tail.

The results from the modelling are set out below:

	Total Future Payments					
Origin Year	Expected Amount	Parameter Uncertainty	Inflation Variation	Severity Variation	Claim No. Variation	Error Of Prediction
1922	24,014	1,609	523	6,175	2,769	6,975
1923	48,525	3,000	1,182	9,308	5,399	11,233
1924	64,818	3,575	1,712	11,794	7,023	14,288
1925	72,538	3,952	2,025	12,575	7,684	15,391
1926	91,065	4,822	2,630	14,096	9,348	17,783
1927	56,457	2,975	1,643	10,268	5,613	12,186
1928	54,093	2,713	1,554	8,352	4,995	10,221
1929	62,483	2,869	1,716	7,836	5,117	9,938
1930	96,894	3,547	2,506	9,124	6,110	11,808
1931	137,779	3,729	3,162	9,641	9,082	14,118
Total	708,666	29,136	18,652	32,147	20,858	51,626

Comparison With Data Three Years Later

The model over estimated the cash flows in the following three years, mainly as a result of the approximate estimates of the run off of claim numbers. However, the estimated ultimate for the oldest accident year is just below the incurred claim figure 3 years later. This suggests that the overall reserve estimated may not be as overestimated as the 3 year payment figure. Even three years later, the ultimates for this class of business are very uncertain. Furthermore, there appear to be some latent claim types that may not be fully reflected in the payments to date. Modelling aggregate payments using other models gives estimated ultimates below the level of current outstandings, but a reliable prediction of the next three years payments. The operational time model estimates ultimates allowing for significant IBNR, but spreads the reserve according to settlements and so overestimates the next three year payments. The table below shows the difference in actual payments compared with those estimated as a multiple of the estimated standard error.

Year	Actual Payments
1922	-1.4
1923	-2.4
1924	-2 1
1925	-1.9
1926	-2.3
1927	-1.1
1928	-0.8
1929	-2.0
1930	-1.5
1931	-0.6
Total	-3.8

Data Set 3

The development of claim numbers in the tail is relatively small, no new claims being expected for the oldest origin year. The estimated reserves are dependent on the accuracy of the claim number estimates and so the estimates for this data set should be more reliable that for the first data set.

Numbers of claims reported were modelled using a number of methods and an ultimate selected for each accident year.

Model zero was fitted using one level parameter for all origin years and seven intervals in operational time. Examination of residuals led to a second origin year parameter being fitted for 1925 and 1926 as these years appear to have a higher average cost. The origin year groups and operational time intervals fitted were:

Origin Year	Groups	Operational Time Intervals
1922	1	$0.000 \le \tau < 0.219$
1923	1	$0.219 \le \tau < 0.407$
1924	1	$0.407 \le \tau < 0.585$
1925	2	$0.585 \le \tau < 0.807$
1926	2	$0.807 \le \tau < 0.894$
1927	1	$0.894 \le \tau < 0.949$
1928	1	$0.949 \le \tau \le 1.000$
1929	1	
1930	1	
1931	1	

The first graph shows the fitted model zero and the data points. It appears that over 80% of the claims are settled in the five development quarters, but the average cost of those claims is significantly lower than the average cost of claims settled in the tail. This suggests that the data may contain a mix of claim types.

The final fitted model replaced the 7 straight lines of model zero with 6 polynomial coefficients. The relatively large number of parameters was needed to cope with the sudden increase in average cost at operational time 0.8. The 9 fitted parameters are shown below together with their standard errors. This model has 210 degree of freedom since one negative increment was excluded from the fit. The model is plotted in the second graph.

		Estimate	S.Error
Past Force Of Inflation	:	0.022	0.002
Origin Year Group 1	:	0.694	0.084
" " " 2	:	0.798	0.084
Predictor Term τ	:	-4.024	2.331
" " τ^2	:	41.996	19.206
" " $\tau 3$	:	-181.670	68.465
" " $\tau 4$	:	384.976	119.126
" " $\tau 5$	:	-391.610	99.629
" " τ^6	:	153.070	32.046

The residual plots are generally well behaved for this data set and the results from the modelling are set out below:

	Total Future Payments					
Origin Year	Expected Amount	Parameter Uncertainty	Inflation Variation	Severity Variation	Claim No. Variation	Error Of Prediction
1922	1,957	88	1	1,971	0	1,973
1923	2,902	130	4	2,405	306	2,428
1924	5,497	246	12	3,321	548	3,375
1925	12,584	565	32	5,281	828	5,375
1926	22,057	979	60	6,942	1,376	7,144
1927	31,953	1,361	94	7,812	1,934	8,163
1928	60,817	2,422	187	10,368	4,068	11,399
1929	129,331	4,420	423	13,942	7,608	16,491
1930	231,498	6,069	838	16,525	11,659	21,132
1931	422,476	7,318	1,473	18,303	20,808	28,700
Total	921,074	22,832	3,125	32,676	25,496	47,422

Comparison With Data Three Years Later

The next three years development indicated that the model underestimated 1927 accidents and overestimated those from 1929 and 1930.

The projected cash flows over the following three years are based on an approximate model for the numbers settled and hence not too much notice should be taken of the significance of theses results. The table below shows the difference in actual payments compared with those estimated as a multiple of the estimated standard error.

Year	Actual Payments
1922	-0.8
1923	-0.1
1924	0.0
1925	-0.7
1926	-0.1
1927	2.3
1928	-0.7
1929	-1.6
1930	-2.4
1931	-0.5
Total	-1.8

Data Set 5

The development of claim numbers in the tail is relatively small, with only a few new claims being expected for the oldest origin year. Numbers of claims reported were modelled using a number of methods and an ultimate selected for each accident year.

Model zero was fitted using one level parameter for all origin years and 13 intervals in operational time. Examination of residuals led to no additional origin year parameters being fitted. The origin year groups and operational time intervals fitted were:

Origin Year	Groups	Operational Time Intervals
1922	1	$0.000 \leq \tau < 0.130$
1923	1	$0.130 \leq \tau < 0.234$
1924	1	$0.234 \leq \tau < 0.301$
1925	1	$0.301 \leq \tau < 0.407$
1926	1	$0.407 \leq \tau < 0.467$
1927	1	$0.467 \leq \tau < 0.597$
1928	1	$0.597 \leq \tau < 0.652$
1929	1	$0.652 \leq \tau < 0.724$
1930	1	$0.724 \leq \tau < 0.788$
1931	1	$0.788 \leq \tau < 0.845$
		$0.845 \leq \tau < 0.897$
		$0.897 \leq \tau < 0.948$
		$0.948 \leq \tau \leq 1.000$

The first graph shows the fitted model zero and the data points. It appears that over 60% of the claims are settled in the five development quarters, but the average cost of those claims is significantly lower than the average cost of claims settled in the tail. This suggests that the data may contain a mix of claim types.

The final fitted model replaced the 13 straight lines of model zero with 4 polynomial coefficients. The 6 fitted parameters are shown below together with their standard errors. This model has 214 degrees of freedom and is plotted in the second graph.

		Estimate	S.Error
Past Force Of Inflation	:	0.087	0.005
Origin Year Group 1	:	1.167	0.036
Predictor Term τ^3	:	7.229	1.655
" " τ^5	:	-76.016	16.372
" " τ^6	:	140.315	28.033
" " τ^7	:	-67.937	13.385

The residual plots are generally well behaved for this data set and the results from the modelling are set out below:

	Total Future Payments					
Origin Year	Expected Amount	Parameter Uncertainty	Inflation Variation	Severity Variation	Claim No. Variation	Error Of Prediction
1922	8,849	924	83	5,659	1,923	6,049
1923	17,371	1,795	181	7,996	2,579	8,593
1924	32,037	3,253	377	10,964	3,863	12,077
1925	66,921	6,412	882	15,900	5,437	18,008
1926	95,928	8,662	1,383	19,071	6,779	22,059
1927	122,238	8,768	1,982	20,924	6,186	23,598
1928	97,086	6,271	1,681	18,445	5,781	20,391
1929	124,181	6,775	2,361	20,376	6,353	22,517
1930	166,521	7,987	3,519	23,236	8,232	26,151
1931	219,703	9,532	4,968	26,101	15,593	32,248
Total	950,833	59,212	17,417	56,983	22,886	87,065

Comparison With Data Three Years Later

The next three years development indicated that the model overestimated the development in the tail.

The projected cash flows over the following three years are based on an approximate model for the numbers settled and hence not too much notice should be taken of the significance of theses results. The table below shows the difference in actual payments compared with those estimated as a multiple of the estimated standard error.

Year	Actual Payments
1922	-0.3
1923	-0.8
1924	-0.9
1925	-0.9
1926	-0.6
1927	-1.1
1928	-0.7
1929	0.8
1930	1.0
1931	0.3
Total	-0.4

To investigate the cause of the overestimation in the tail, the model was refitted to the data three years later and two possible causes identified. Firstly, inflation over the three year period was significantly lower than the average of past inflation. Secondly, the original model assumed that the average cost of claims increased significantly as operational time approaches 1. In fact, many claims in the tail were settled at low cost so that the model of average cost possible should have been decreasing in the extreme tail.

The original model was also refitted allowing for a part payment parameter. This produced more accurate estimates of the development in the tail. However, it is unwise to fit this parameter unless one has knowledge about the nature of the part payment process for the particular data set being analysed.

6.4.3. Practical Limitations

Model depends on a reliable model for claim numbers.

Whilst these models may be fitted in the GLIM package [7], they are time consuming to develop. Otherwise a commercial package could be purchased.

6.5. Bootstrapping

6.5.1. Overview of Method

This method does not by itself generate claim reserves. Rather it is a mechanism for measuring the uncertainty of a particular method e.g. chain-ladder. It utilises the errors between the particular method and actual observed data to generate alternative 'actual' data. The method is then repeatedly used to generate a distribution of claim reserves.

It can be applied to any particular chosen deterministic method which works directly on the claims data. It does not work for methods which require an external information feed e.g. Bornhuetter-Ferguson.

The method is conceptually a good all-purpose way of measuring the uncertainty of any chosen deterministic method. However, unlike some of the previous methods described it does not provide a detailed breakdown of the sources of error, and requires a lot of computer time to generate the results. One area where it does score, however, is in enabling the modeller to assess how much of the variability in reserves arises from the statistical error (the random "noise" of the claims process) and how much comes from the reserving method itself. This information is not provided by any of the other methods.

6.5.1.1. Introduction

Because Bootstrapping in a reserving context is relatively unknown, this section gives a general explanation of what the technique involves, and then goes on to examine two reserving models to which it can be applied.

6.5.1.2. What is it generally?

Given a sample of data A, from an unknown distribution B, Bootstrapping is a technique for obtaining information about a random variable C(A,B) by re-sampling the observed data A in an appropriate way.

6.5.1.3. What are A,B,C in a reserving context?

Consider an estimate of outstanding claims. A triangle of paid claims, say, is taken ("A"). The claims have some unknown distribution ("B"). A model is fitted to the data, which produces estimates of past (fitted) and future claim payments. The future claim payments, or Reserve, is a random variable ("C(A,B)").

The model, or reserving method, can be a simple model, such as the basic Chain-ladder method, or a more complicated method, such as a Regression model based on Log-Incremental payments.

6.5.1.4. What does Bootstrapping add to basic Chain-ladder methods?

In the reserving example above, Bootstrapping lets us produce an estimate of the variance of the Reserve, C(A,B). The basic method only gives us a point estimate of the Reserve, with no indication of the extent to which we expect the Reserve to vary either side of this expected value. The analysis in 6.5.2 will show how we can obtain a graph of the distribution of the reserve, with accompanying estimates of variance, from Bootstrapping a basic Chain-ladder reserving method.

Reserving methods, whether basic or more sophisticated, make implicit or explicit assumptions about the claims process in fitting a model to it. By providing information about the fit of a model, Bootstrapping lets the modeller assess how appropriate these assumptions are.

6.5.1.5. Can Bootstrapping add anything to more sophisticated models?

Yes. Take for example the Regression model based on Log-Incremental payments from the IOA claims reserving manual. The model of the claim process is described as:

$$\text{Log}(P_{ij}) = a(i)+b(j)+E_{ij}$$

where Pij are the claim payments in Accident Year i at development period j, a(i) and b(j) are the parameters fitted by the model and Eij is an Error term.

The method produces Maximum Likelihood Estimates for the expected values of claims payments, MLE(E(Pij)). But the MLEs are biased, that is:

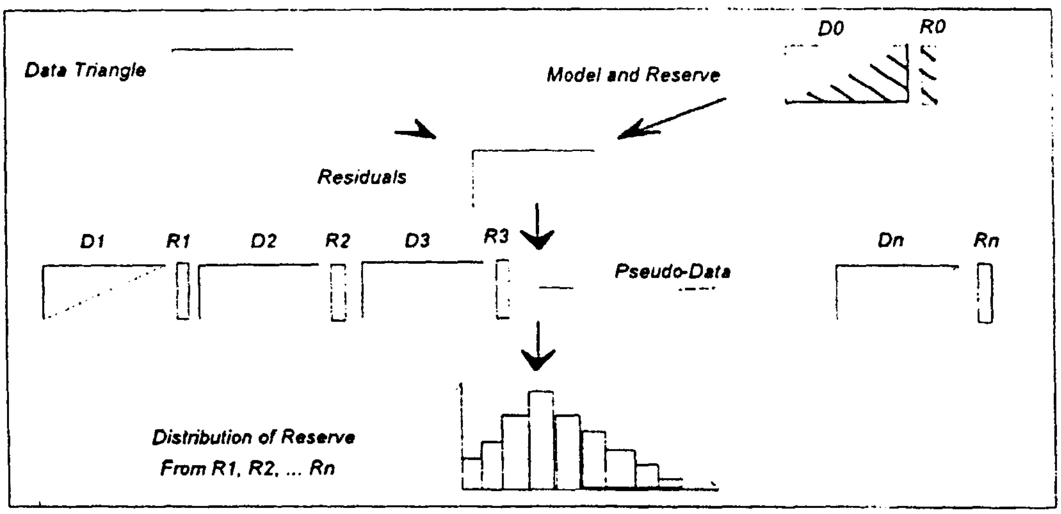
$$E(MLE(E(P_{ij}))) > E(P_{ij})$$

The MLEs are asymptotically unbiased, i.e. as the sample size gets larger, E(MLE(E(Pij))) gets nearer to E(Pij). However, for "small" sample sizes, as is usually the case with reserving data, the E(MLE(E(Pij))) may be considerably different from E(Pij). Bootstrapping lets you make an estimate of this bias and may provide a better estimate than a traditional asymptotic estimate (e.g. a maximum likelihood estimator) with only a small sample size. The examination of bias was the original impetus for looking at Bootstrapping (or more generally, the Jackknife).

Some more sophisticated models produce estimates of the variance of the projected reserve. Bootstrapping can give the modeller an indication of an extent to which the model variance is a result of the underlying "noise" in the data (Statistical error) or due to uncertainty in the modelling process itself - such as mis-specifying the model, or the fact that the estimates of the parameters of a model are themselves random variables and contribute a degree of uncertainty to the predicted reserve.

6.5.1.6. Is it as easy as A,B,C?

Yes and No! Consider the basic Chain-ladder model. The model produces fitted values for past claim payments, which are different from the actual claim payments. These differences between fitted and actual values are called the Residuals. Given "n" such Residuals each can be given a "mass", or probability of 1/n, to produce an empirical distribution for that set of Residuals. A random selection from this empirical distribution is chosen, which generates a new set of data, called Pseudo-data. This process can be repeated many times to produce a large collection of such sets of Pseudo-data. For each set of Pseudo-data the reserving model can be applied and a Reserve estimate produced - a Pseudo-Reserve. If enough sets of Pseudo-data are produced we can produce an estimate of the distribution of the Pseudo-Reserve and infer details about its distribution, such as its variance. The approach is outlined graphically below:



We can also go one step further, and produce a set of Pseudo-Data for all future values in a similar way. These future Pseudo-Data represent "reality"; they are the simulated completion of the rectangle. The Pseudo-Reserve, however, is still calculated from the first "half" of the rectangle and attempts to "fill in" the rest, using, in this case, the chosen reserving method.

We expect the Pseudo-Reserve Standard Error to be made up of an error due to the randomness of the underlying data (often called "Statistical Error", as indicated by the Standard Error of the future Pseudo-Data) plus other error terms due to the specification of the model and the ability of the model to fit the correct parameters, even if the model were correct. Comparing the relative sizes of the Standard Errors of the future Pseudo-Data and the Standard Errors of the Pseudo-Reserve, gives an indication of the extent to which the variability in reserve estimates is due to the underlying noise of the data, as opposed to variance introduced by the process of estimating the reserve.

6.5.1.7. A small amount of theory

To be able to justify inferring results from Bootstrapping, the Residuals should be independent and identically distributed (there is no requirement for them to be normally distributed). If one looks at cumulative data, the residuals are unlikely to be independent, so the method tends to be applied to incremental data. Some sophisticated reserving methods also make assumptions about Residuals, but tend to make the more restrictive assumption that they are independent, identically and normally distributed.

6.5.1.8. What can Bootstrapping be applied to?

Any reserving method that can be performed automatically in a spreadsheet is amenable to Bootstrapping methods. Once Bootstrapping has been set up for one reserving method, it can quite easily be extended to another. For part of the analysis that follows in 6.5.2 the Add-In @Risk is used in conjunction with a Lotus spreadsheet.

@Risk lets one enter random variables in a spreadsheet cell. @Risk then effectively recalculates the spreadsheet as many times as required, each time picking a value from the chosen random distributions and collating statistics regarding chosen cells in the spreadsheet that are functions of the random variables. In this case, the random distribution is just re-sampling the triangle of Residuals. @Risk is available for a few hundred pounds and can be added to many standard spreadsheet packages, such as Lotus, Excel or Symphony.

Whilst the results for more basic models should be treated with caution, for any method that can be performed in a spreadsheet, Bootstrapping provides a first estimate of the variance of reserves which is certainly better than no estimate at all. Very little extra effort is needed, once one has a clear idea of what one is trying to do! The method boils down to recreating lots of sets of triangles, and then performing one's reserving method on those new Pseudo-Data triangles. The reserving methods are usually such that they can just as easily be applied to the new triangles of Pseudo-Data as to the original data, so the bulk of the effort goes into re-sampling the Residuals to come up with lots of sets of Pseudo-Data.

To illustrate the technique, section 6.5.2 examines the application of Bootstrapping to the basic Chain-ladder method and to Regression models based on Log-Incremental payments. For the first method information is obtained about the distribution of the reserve that was not available from the original method alone. For the second method, we can examine the extent of the Bias in the model and the breakdown of the Standard Errors produced by the model between Statistical Error ("noise" in the data) and additional sources of error introduced by the process of reserving.

6.5.2. Data Analysis

We have seen above that Bootstrapping is not itself a reserving method, it is just a technique for enhancing the information available regarding existing reserving methods. This section will not dwell therefore on the results obtained for particular classes, but will illustrate some of the information that Bootstrapping provides and comment on how the techniques were applied in practice.

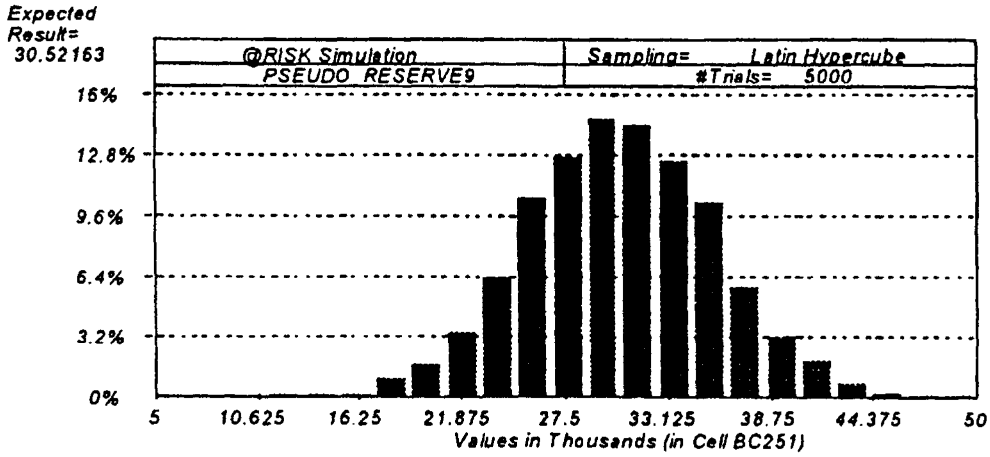
Bootstrapping the Basic Chain-ladder

The analysis was performed in a Lotus spreadsheet for each class. A volume-weighted quarterly chain-ladder was performed. A volume weighted approach has the benefit that all the Residuals for each accident year add up to zero, and hence the Residuals in total add up to zero. Having completed the square, both for the future payments and the fitted past payments, the difference between the fitted and actual incremental claims was calculated - this is our set of Residuals.

To construct a set of Pseudo-Data, which is the heart of the technique, lots of sets of Residuals need to be calculated and added to the original set of fitted past data. This was done using the Lotus Add-in @Risk, as described in 6.5.1.8. From each set of Pseudo-Data, the revised cumulative triangle was constructed and the chain-ladder applied to produce a reserve estimate - the Pseudo-Reserve.

The Add-in @Risk performs all these simulations by adding a Residual, picked at random, from the set of Residuals to each original past fitted data (the choice is different for each point but can replicate the choice for other points in the triangle). The new Pseudo-Reserve is calculated (by calculating the spreadsheet within @Risk), and the results collated so the distribution of the Pseudo-Reserve and statistics of interest can be examined.

A sample of the distribution output is given below. It shows the distribution of the reserve estimate calculated by the basic chain ladder on 5,000 sets of Pseudo-data for Class one, Accident year 1926:



The graph shows the extent to which the reserve estimate of the basic chain-ladder varies either side of the expected value, assuming that the variation of the Pseudo-Data is typical of the variation of the claims process underlying the data for Class one. Such graphs and accompanying statistics could be provided for any reserving method that can be set up in a spreadsheet to provide a point estimate of reserves by a simple recalculation of the spreadsheet for a given set of data. Clearly this is information of interest and an improvement on a single point estimate of the reserves.

The Bootstrapping output includes estimates of the Standard Error of the reserve estimate for each accident year and for the reserve estimate for all accident years combined. These were perhaps not surprisingly higher than the SE's as calculated by the Log-linear regression methods indicated in section 6.2, but they were only different by a factor of about two, so could not be said to be unreasonable.

As well as the distribution of the reserve, Bootstrapping can give this user an idea as to what goes into the variability of the claims estimate, as indicated in section 6.5.1.6. This is done by comparing the variability of the Pseudo-reserve estimate with the variability of the Pseudo-data for those future payments. We expect the reserving method to add uncertainty to the underlying claims process, and so by comparing the two, we can examine the components of the reserve SE.

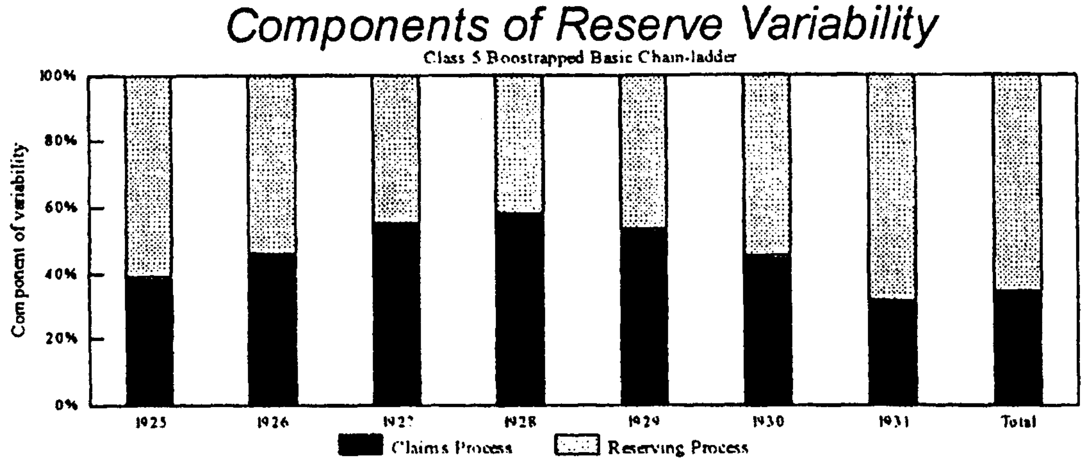
The following table shows the Pseudo Reserve and future Pseudo-Data to development year nine for Class five. We observe that the Pseudo-Reserve SE is greater than the Pseudo-Data SE - this gives us a measure of the variability that the reserving process is adding relative to the implicit variability of the claims process:

Bootstrapped Basic Chain-ladder

Class 5

	Accident Year	Reserve to year nine	SE / Reserve	Future Pseudo Data to year nine	Proportion of reserve error due to claims process	Proportion of reserve error due to reserving process
Predicted	1922					
SE						
Predicted	1923					
SE						
Predicted	1924	12,352		12,303		
SE				3,216		
Predicted	1925	37,603		37,581		
SE		11,562	31%	4,536	39%	61%
Predicted	1926	73,585		73,547		
SE		12,283	17%	5,633	46%	54%
Predicted	1927	71,944		71,882		
SE		11,663	16%	6,487	56%	44%
Predicted	1928	69,919		69,915		
SE		12,375	18%	7,214	58%	42%
Predicted	1929	92,312		92,289		
SE		14,753	16%	7,936	54%	46%
Predicted	1930	123,674		123,578		
SE		18,527	15%	8,490	46%	54%
Predicted	1931	163,576		163,369		
SE		28,923	18%	9,233	32%	68%
	Total	644,965		644,463		
		55,485	9%	19,075	34%	66%

The results are displayed graphically as follows:



The split of the reserving variability worked best for class five, which was itself the most variable class. For other classes for some accident years, the variability of the reserve was largely or entirely attributed to the claims process, which is clearly unrealistic, but perhaps a function of the over-parameterised nature of the basic chain-ladder.

A weakness in all the above is the assumption that the Residuals are uniform across the triangle. Clearly the payments in the initial years are substantially larger than the payments in the later years, so the Residuals in the earlier years, if added to the fitted payments in the later years, may be imparting an undue amount of variability to that section of the Pseudo-Data. That said, however, it is the tail of the triangle where a few large claims may have a particularly distorting effect, so that is not to say that the larger Residuals applied to the smaller payments is completely unrealistic.

The problem of applying one set of Residuals to the entire triangle can be overcome by scaling the Residuals, or partitioning them into several sets - an early and a late development set, for example. This was briefly examined and did not materially affect the results and is not considered further.

Bootstrapping the Log-Linear Regression method

The previous section described how Bootstrapping could enhance the information gleaned from basic chain-ladder reserving methods. This section describes how the techniques can be used to look at more sophisticated methods.

One use of the technique is to look at the components of reserve variability in the same fashion as that done for the chain-ladder. To do this, the log-payments were bootstrapped. That is to say the Residuals were taken to be the set of differences between the actual and fitted log-payments.

The Add-in @Risk could not be used for Bootstrapping this method, as the results could not be obtained by a simple calculation of the spreadsheet. Instead a simple macro was constructed to loop through choosing from the set of Residuals, forming the new set of Pseudo-Data, performing the regression on the Pseudo-Data, calculating the expected reserve from the revised set of regression parameters, and collating the results.

The results are summarised below, in a similar fashion to that shown previously:

Bootstrapped Log-Linear

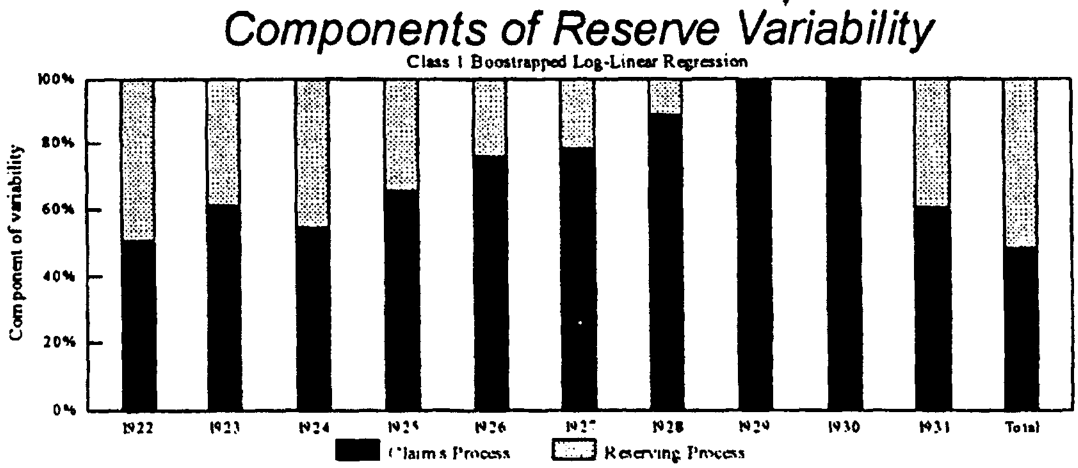
Class 1

	Accident Year	Reserve to year thirteen	SE / Reserve	Future Pseudo Data to year thirteen	Proportion of reserve error due to claims process	Proportion of reserve error due to reserving process
Predicted	1922	6,537		6,475		
SE		564	9%	286	51%	49%
Predicted	1923	14,116		14,097		
SE		1,062	8%	652	61%	39%
Predicted	1924	24,670		24,717		
SE		1,607	7%	878	55%	45%
Predicted	1925	36,610		36,617		
SE		2,157	6%	1,423	66%	34%
Predicted	1926	48,448		48,184		
SE		2,352	5%	1,790	76%	24%
Predicted	1927	37,244		37,090		
SE		1,481	4%	1,160	78%	22%
Predicted	1928	31,853		32,066		
SE		1,085	3%	966	89%	11%
Predicted	1929	38,826		39,450		
SE		1,338	3%	1,326	99%	1%
Predicted	1930	73,619		73,660		
SE		2,683	4%	2,714	100%	0%
Predicted	1931	121,166		120,743		
SE		5,349	4%	3,245	61%	39%
	Total	433,090		433,099		
		12,361	3%	5,993	48%	52%

The SEs are of a similar size to those predicted by the Log-linear regression method in 6.2.2. The components of reserve variability are now less for the reserving process than for the claims process - although the reserving process element is still understated.

It is interesting to compare these results with those for the basic chain-ladder, which had considerably more uncertainty being added by the reserving process.

The results are again displayed graphically below:



The other area which Bootstrapping can be useful in investigating the bias of methods such as Log-linear regression. Section 6.2.1.2 and 6.5.1.5 described how the process of transforming from estimated log-payments to estimated payments introduced bias into the reserving process for this model.

How big do we expect this bias to be? Section 6.2.1.2 detailed an adjustment to the biased payments, involving a daunting-looking infinite series. The series can be simplified to a few terms, for sizes of n (number of observations), p (number of parameters) and small model variance, of the orders of magnitude we are usually dealing with in a reserving context. This reduces to:

Corrected Payments / Uncorrected payments $\approx m / (m+1)$

where $m = n - p$

In other words our uncorrected payments are about 1/m too big, or around 1% in the context of the models looked at in section 6.2.

To examine the bias, we need to Bootstrap the actual v. fitted payments, rather than the actual v. fitted log-payments, as was done previously. This was done again using a macro in a Lotus spreadsheet. The results are summarised below:

Bootstrapped Log-Linear Regression (bootstrapping the actual payments)

Class 1

	Accident Year	Pseudo Reserve to year nine	Pseudo Future Data to year nine	Ratio Pseudo Reserve to Pseudo Future
Predicted SE	1922			
Predicted SE	1923			
Predicted SE	1924	8,118	7,731	105%
Predicted SE	1925	19,664	18,833	104%
Predicted SE	1926	32,679	31,627	103%
Predicted SE	1927	28,579	27,944	102%
Predicted SE	1928	26,559	26,239	101%
Predicted SE	1929	34,405	34,362	100%
Predicted SE	1930	68,605	67,841	101%
Predicted SE	1931	116,819	114,965	102%
	Total	335,428	329,680	102%

[Note that the Pseudo-Reserve and Pseudo-data figures for the above table only go up to year nine, whereas the previous Bootstrapping table projected numbers up to year thirteen, so the two sets of numbers are not directly comparable.]

At first sight the results look as if they support the expected size of the bias. However, the treatment of negative claims comes back to haunt the Log-linear regression method. We have picked the Residuals from the entire triangle and applied them to all the actual fitted payments. This means that payments in the tail have some large negative payments. The model coped with this by setting such payments equal to the actual original fitted payment - this in itself introduced bias in to the reserving process.

This bias will have caused some or all of the effects above. Time and materiality prevented the working party looking into this further.

6.5.3. Practical Limitations

The concepts of Bootstrapping are quite straightforward, the main limitation is the speed at which computers can perform the many calculations required. The heart of the Bootstrapping process is the construction of many sets of Pseudo-Data and the collating of the Pseudo-Reserves calculated from each. This is a very intensive number crunching process and needs a powerful PC if the calculations are to be carried out in a realistic time frame. The Bootstrapping calculations were performed on a speed-doubled 486 PC, operating at 66Mhz.

The Bootstrapping of the Basic Chain-ladder reserving method was carried out in a Lotus version 2.01 spreadsheet with the Add-In @Risk. The spreadsheet comprised fifteen 10×40 triangles (quarterly data was used), that is 6,000 individual calculations. The re-sampling of the residuals involved a recalculation of a 10×40 triangle of past and future payments, each cell being a random sample from 216 residuals. The add-in @Risk provided the facility to enter the triangle of re-sampled residuals as random variables in cells, and then collated the results for some 60 cells of interest within the spreadsheet. This is clearly quite a number intensive spreadsheet! For each class 5,000 simulations were performed (i.e. 30 million individual calculations), which took about ten hours. Similar work has been done for smaller, annual, triangles for which around 40,000 simulations can be performed in about ten hours. Though time-consuming, the calculations are not beyond the scope of today's more powerful PCs.

The Bootstrapping of the Log-Linear method was not performed with the assistance of @Risk. This was because the results could not be produced for each new set of Pseudo-Data simply by recalculating the spreadsheet, because part of the process involved performing regression on the Log-Payments each time. Instead a simple macro was set up to run through a loop, each time selecting a new set of Pseudo-Data, performing the regression, calculating the reserves and storing the results in a table. This was equally time-consuming, and again about 5,000 simulations could be performed in ten hours.

The time constraints are clearly quite large, but, especially when annual payments are being considered, need not pose a problem. Although it is helpful to have packages such as @Risk to collate simulation results, the production of statistics regarding Pseudo-Reserves and Pseudo-Data generally can still be performed in any straightforward spreadsheet in a reasonable time frame.

6.6. Summary of Results

The following tables bring together the results of using the three stochastic methods, and Bootstrapping using the Basic Chain-Ladder on the three classes of data described in 6.1.

The modellers were asked to predict the payments over the next three years, and an ultimate reserve. The tables show these predictions together with the actual payments made over the three years, and the latest estimate of reserves required, *excluding IBNR*.

It is interesting to compare the accuracy of the predicted payments, and whether the ultimate reserve looks reasonable in the light of three years development. However the main purpose of this *entirely unrealistic exercise* was to draw out the practical aspects of these methods, and readers should not draw any conclusions as to the value of these methods purely on the strength of the results shown here. In particular the results are subject to a great deal of subjective judgement as the modeller is required to make a number of assumptions in interacting with the models. In practice these would be made on the basis of much additional information, some of which may be prompted by skilled analysis of the diagnostics from using one or more of these statistical methods.

Class 1 Comparison of models versus actual results

Accident Year	Actual Figures: Case Estimates at 12/1931	Payments 1932 - 34	Case Estimates at 12/1934	Payments + Case (2)+(3)	Projected Results: Operational Time approach: Range for Payments: From	To		Range for Reserve: From	To	Log-Incremental II: Range for Payments: From	To		Range for Reserve: From	To
	(1)	(2)	(3)	(4)										
1922	20,202	8,507	17,116	25,623	10,236	18,617		17,039	30,989	6,744	8,878		13,250	19,136
1923	27,682	11,861	26,534	38,395	20,659	33,106		37,292	59,759	11,896	15,168		23,676	32,848
1924	41,001	18,557	31,279	49,836	26,792	41,944		50,530	79,105	17,905	22,205		36,155	48,091
1925	40,404	22,947	32,306	55,253	29,835	45,906		57,147	87,930	23,469	28,519		48,246	61,386
1926	45,077	26,213	33,422	59,635	38,485	57,163		73,282	108,849	26,448	31,966		55,745	67,761
1927	38,133	23,300	25,520	48,820	23,968	37,163		44,271	68,643	20,872	24,974		42,355	49,533
1928	31,952	25,966	19,682	45,648	24,804	36,361		43,872	64,314	19,754	23,870		37,771	42,959
1929	32,325	25,597	15,173	40,770	31,578	43,523		52,545	72,421	25,358	29,644		43,709	48,671
1930	49,044	50,350	20,864	71,214	54,404	69,504		85,086	108,703	47,558	54,564		76,228	84,006
1931	66,289	90,282	32,784	123,066	85,968	105,598		123,661	151,897	80,096	90,552		119,270	130,142
Total	393,109	303,580	254,680	558,260	387,369	448,243		657,040	760,292	294,671	315,569		504,002	576,936

Note: The SE for the payments figures in the Log-Incremental II method are upper bounds because the information was only readily available by separate payment year

Note:
= Actual payments within one SE
= Actual payments between one and two SE's
= Actual payments outside two SE's

Class 1 Comparison of models versus actual results

Accident Year	Actual Figures: Case Estimates at 12/1931	Payments 1932 - 34	Case Estimates at 12/1934	Payments + Case (2)+(3)	Projected Results: Log-Linear Regression: Range for Payments: From	To		Range for Reserve: From	To	Bootstrapped Basic Chain-ladder: Range for Payments: From	To		Range for Reserve: From	To
	(1)	(2)	(3)	(4)										
1922	20,202	8,507	17,116	25,623	5,793	7,154		22,925	28,314	6,668	6,668		26,455	26,455
1923	27,682	11,861	26,534	38,395	10,608	12,724		37,061	44,706	5,996	16,908		41,818	41,818
1924	41,001	18,557	31,279	49,836	16,763	19,625		52,193	61,578	14,017	25,211		24,051	96,949
1925	40,404	22,947	32,306	55,253	23,252	26,752		65,149	75,292	20,739	31,715		53,877	93,815
1926	45,077	26,213	33,422	59,635	29,563	33,690		75,170	85,333	25,225	35,841		66,721	94,309
1927	39,133	23,300	25,520	48,820	22,220	25,268		51,546	57,748	17,918	27,060		45,198	84,030
1928	31,952	25,966	19,682	45,648	19,082	21,746		40,383	44,921	16,930	24,754		37,527	51,947
1929	32,325	25,597	15,173	40,770	23,758	27,132		45,541	50,593	22,213	30,271		43,283	55,639
1930	49,044	50,350	20,864	71,214	46,809	53,700		79,968	89,320	46,097	54,141		79,374	90,998
1931	66,289	90,282	32,784	123,066	82,676	91,481		124,523	139,814	79,177	90,087		121,659	137,391
Total	393,109	303,580	254,680	558,260	293,023	306,774		614,797	657,278	282,015	315,621		610,336	682,978

Note:

- = Actual payments within one SE
- = Actual payments between one and two SE's
- = Actual payments outside two SE's

Class 3 Comparison of models versus actual results

Accident Year	Actual Figures: Case Estimates at 12/1931	Actual Figures: Payments 1932 - 34	Actual Figures: Case Estimates at 12/1934	Actual Figures: Payments + Case (2)+(3)	Projected Results – Operational Time approach: Range for Payments From	Range for Payments To	Range for Reserve From	Range for Reserve To	Projected Results – Log-Incremental II: Range for Payments From	Range for Payments To	Range for Reserve From	Range for Reserve To
	(1)	(2)	(3)	(4)								
1922	1,816	459	473	932	(15)	3,929	(15)	3,930	645	9[illegible]1	769	1,005
1923	2,750	2,661	1,261	3,922	474	5,330	474	5,330	1,615	2,2[illegible]3	1,934	2,472
1924	7,243	5,540	1,254	6,794	2,122	8,872	2,122	8,872	4,114	5,5[illegible]4	4,947	6,165
1925	14,191	8,646	5,550	14,196	7,018	17,484	7,209	17,959	8,240	11,0[illegible]6	10,166	12,376
1926	16,564	19,341	2,850	21,991	13,435	28,307	14,913	29,202	17,175	22,6[illegible]7	20,814	24,934
1927	32,450	45,012	4,485	49,497	21,238	35,812	23,790	40,116	24,518	32,2[illegible]8	29,792	35,354
1928	39,722	47,015	6,708	53,723	43,683	63,835	49,418	72,216	46,942	57,0[illegible]2	56,241	63,167
1929	70,300	89,569	6,880	96,449	97,462	125,950	112,840	145,823	90,916	106,2[illegible]2	108,909	119,493
1930	110,561	151,294	23,570	174,864	176,671	212,165	210,367	252,630	154,745	177,3[illegible]5	189,136	206,448
1931	169,734	341,912	42,066	383,978	329,044	377,008	393,777	451,176	285,371	328,[illegible]93	353,279	388,823
Total	465,331	711,449	94,897	806,346	743,552	824,272	873,652	968,496	659,330	718,3[illegible]2	790,778	845,444

Note: The SE for the payments figures in the Log-Incremental II method are upper bounds because the information was only readily available by separate payment year

Note:
- = Actual payments within one SE
- = Actual payments between one and two SE's
- = Actual payments outside two SE's

Class 3 Comparison of models versus actual results

Accident Year	Actual Figures: Case Estimates at 12/1931 (1)	Payments 1932 - 34 (2)	Case Estimates at 12/1934 (3)	Payments + Case (2)+(3) (4)	Projected Results: Log-Linear Regression: Range for Payments From	To		Range for Reserve From	To		Bootstrapped Basic Chain-ladder: Range for Payments From	To		Range for Reserve From	To
1922	1,816	459	473	932	825	1,025		1,123	1,394		1,130	1,130	-	1,130	1,130
1923	2,750	2,661	1,261	3,922	2,010	2,574		2,574	3,150		2,802	2,802	-	2,801	2,801
1924	7,243	5,540	1,254	6,794	5,029	6,146		6,225	7,535		4,621	4,621	-	5,605	5,605
1925	14,191	8,646	5,550	14,196	9,839	11,948		11,956	14,345		10,363	10,363	-	12,382	12,382
1926	18,564	19,341	2,650	21,991	17,402	21,025		20,943	24,950		17,186	17,188	-	20,611	20,611
1927	32,450	45,012	4,485	49,497	24,705	29,750		29,590	35,070		10,389	43,763		12,071	50,223
1928	39,722	47,015	6,708	53,723	40,113	48,263		47,938	56,668		32,194	62,580		38,139	71,973
1929	70,300	89,569	6,880	96,449	78,838	95,087		94,135	111,377		80,978	108,918		93,332	125,338
1930	110,561	151,294	23,570	174,864	152,874	185,665		182,483	217,153		146,063	172,805		172,753	204,403
1931	169,734	341,912	42,066	383,978	332,724	384,969		389,434	460,681		328,886	356,414		394,524	428,766
Total	465,331	711,449	94,897	806,346	699,677	751,035		816,072	902,649		641,059	774,135		754,617	922,001

Note:

= Actual payments within one SE

= Actual payments between one and two SE's

= Actual payments outside two SE's

Class 5 Comparison of models versus actual results

	Actual Figures				Projected Results									
					Operational Time approach					Log-Incremental II				
Accident Year	Case Estimates at 12/1931	Payments 1932 - 34	Case Estimates at 12/1934	Payments + Case (2)+(3)	Range for Payments			Range for Reserve		Range for Payments			Range for Reserve	
					From	To		From	To	From	To		From	To
	(1)	(2)	(3)	(4)										
1922	5,723	5,896	1,256	7,152	2,266	12,061		2,799	14,898	7,491	10,775		12,790	19,072
1923	7,180	7,983	1,361	9,344	6,968	19,878		9,017	25,724	12,914	17,962		22,276	31,716
1924	16,403	15,949	2,083	18,032	15,044	31,316		20,792	43,282	31,569	42,899		54,941	75,533
1925	35,338	34,808	9,443	44,251	34,162	54,368		51,647	82,196	47,298	62,510		83,404	109,128
1926	53,140	52,080	14,173	66,253	47,902	68,924		78,667	113,190	70,237	91,877		125,759	158,541
1927	50,018	57,477	8,539	66,016	58,428	76,550		105,827	138,649	60,540	80,532		110,515	136,673
1928	44,632	45,777	11,718	57,495	43,405	59,380		81,997	112,174	51,358	67,540		97,850	117,026
1929	56,121	68,028	21,937	89,965	51,767	69,572		105,959	142,403	66,826	87,982		135,249	161,315
1930	62,857	85,215	21,476	106,691	60,447	84,700		138,696	194,345	76,677	101,559		175,566	212,208
1931	44,726	95,741	49,423	145,164	62,723	112,676		157,133	282,273	82,160	109,916		219,063	270,615
Total	376,138	468,954	141,409	610,363	441,743	530,795		863,768	1,037,898	551,592	629,028		1,065,062	1,264,180

Note: The SE for the payments figures in the Log-Incremental II method are upper bounds because the information was only readily available by separate payment year

Note:
- = Actual payments within one SE
- = Actual payments between one and two SE's
- = Actual payments outside two SE's

Class 5 Comparison of models versus actual results

	Actual Figures				Projected Results									
					Log-Linear Regression					Bootstrapped Basic Chain-ladder				
Accident Year	Case Estimates at 12/1931	Payments 1932 - 34	Case Estimates at 12/1934	Payments + Case (2)+(3)	Range for Payments			Range for Reserve		Range for Payments			Range for Reserve	
					From	To		From	To	From	To		From	To
	(1)	(2)	(3)	(4)										
1922	5,723	5,896	1,256	7,152	2,358	3,669		3,190	4,963	2,959	2,959	-	3,917	3,917
1923	7,180	7,983	1,361	9,344	6,490	9,505		8,563	12,592	12,035	12,035	-	14,471	14,471
1924	16,403	15,949	2,083	18,032	17,703	24,687		23,196	32,430	18,365	41,017		36,151	36,151
1925	35,338	34,808	9,443	44,251	33,507	45,123		44,232	59,254	41,773	65,957		45,024	85,006
1926	53,140	52,080	14,173	66,253	53,125	70,162		71,519	92,758	61,244	85,810		85,382	119,598
1927	50,018	57,477	8,539	66,016	46,689	61,280		64,834	82,367	52,165	73,455		74,983	103,997
1928	44,632	45,777	11,718	57,495	42,849	56,284		62,205	78,152	42,591	64,177		67,448	96,458
1929	56,121	68,028	21,937	89,965	50,155	65,867		77,849	97,114	50,708	74,072		87,423	120,681
1930	62,857	85,215	21,476	106,691	66,808	87,771		116,542	144,689	58,576	83,590		115,967	156,833
1931	44,726	95,741	49,423	145,164	79,944	96,032		154,869	197,871	64,915	95,875		146,265	209,099
Total	376,138	468,954	141,409	610,363	437,637	482,369		670,463	758,723	461,536	542,740		741,799	881,441

Note:

= Actual payments within one SE

= Actual payments between one and two SE's

= Actual payments outside two SE's

7. Features of the Statistical Reserving Techniques

or

When Can they be Used?

Features	Modelling Method			
	Log-Incremental Claims	**Log-Incremental Claims II**	**Operational Time**	**Bootstrapping**
Model Developed to represent underlying Claims Process	✗	✗	✓	✗
Separation of Different Sources of Variability	✓	✓	✓	(✓)
Allows Projection Beyond the Range of the Available Data	✓	✓	✓	✗
Independent of Distributional Assumptions	✗	✗	✓	✓
Handles Negative Claims Adequately	✗	✗	✗	✓
Requires Numbers as well as Amounts	✗	✗	✓	✗
Predicts future Cash Flow	✓	✓	✗	(✓)

Note: ( ) indicates basic approach can be extended to provide feature

8. Benchmarking Statistical Reserving Techniques

or

How should we judge the success of a reserving technique?

It is important to reiterate that the use of a particular reserving approach should not be judged solely on the accuracy of its prediction, nor whether its standard error range encompassed the ultimate result. Clearly these are important measures, but the benefits to the practitioner of a particular reserving methodology embrace wider considerations. These include

> *Do the diagnostics assist the practitioner to identify features in the data requiring further investigation?*
>
> *Is the method robust from year to year?*
>
> *Does the method lend itself to a ready analysis of the changes from one valuation to the next? (an "Analysis of Surplus")*
>
> *Does the method aid the lay manager in making business decisions?*

8.1 Why Methods Don't Always Predict the Future Accurately

Some of the more common reasons for a method failing to predict the future accurately are:

1. Future is not like the past.

2. Assumptions are wrong.

   Note: The simple chain ladder also contains lots of implicit assumptions which users should be aware of - Thomas Mack's paper [6] is well worth studying.

3. Paid and incurred may be giving different pictures.

4. May have ignored problems with the data / modelled the wrong data.

9. Other Methods & Further Reading
or
Where Now?

9.1. Further Reading

The numbers refer to references made in the relevant sections of the paper. Otherwise papers are shown for the purposes of further reading.

[1]CHRISTOFIDES S.(1990) "Regression models based on Log-Incremental payments", Institute of Actuaries Reserving Manual Volume II

[2]EFRON B. (1979) "Bootstrap Methods, Another Look at the Jackknife" (the 1977 Rietz Lecture), The Annals of Statistics Vol 7 No 1, pages 1-26

[3]FINNEY D.J. (1941) "On the distribution of a Variate whose logarithm is Normally distributed", JRSS Supplement 7, pages 155-61

[4]FREEMAN D.A. & PETERS S.C. (1984) "Bootstrapping a Regression Equation: Some Empirical Results", Journal of the American Statistical Association Vol 79 No 385, pages 97-106

[5]HARRISON P.J. & STEPHENS C.F. (1976) "Bayesian Forecasting", Journal of the Royal Statistical Society (B) 38

[6]MACK T. (1993) "Measuring the Variability of Chain Ladder Reserve Estimates", Casualty Actuarial Society

[7]MCCULLAGH P. & NELDER J.A. (1983) "Generalized Linear Models", Chapman & Hall

[8]MURPHY D.M. (1993) "Unbiased Loss Development Factors", Casualty Actuarial Society

[9]RENSHAW A.E. (1989) "Chain-ladder and Interactive modelling", JIA 116 Part III

[10]TAYLOR G.C. "Regression Models in Claims Analysis I: Theory"

[11]VERRALL R.J. (1989) "A State-Space Representation of the Chain-Ladder Linear Model", JIA 116 Part III, pages 589-611

The chain ladder model parameters are fitted using a Bayesian approach

[12]VERRALL R.J. (1991) "Chain-Ladder and Maximum Likelihood", JIA 118 Part III, pages 489-499

[13]VERRALL R.J. (1991) "On the Unbiased Estimation of Reserves from Loglinear Models", Insurance: Mathematics and Economics, Vol 10 No. 1, Pages 75-80

[14]VERRALL R.J. (1993) "Statistical Methods for the Chain-Ladder Technique", City University

[15]VERRALL R.J. (1993) "Negative Incremental Claims: Chain-Ladder and Linear Models", JIA 120 Part I, pages 171-185,

The practical problem of handling negative incremental claims when modelling with a log incremental model is addressed. A statistical estimate of the "correct" adjustment is given.

[16]WRIGHT T.S. (1990) "A Stochastic Method for Claims Reserving in General Insurance", JIA 117 Part III, pages 677-733

[17]ZEHNWIRTH B. (1990) "Probabilistic Development Factor Models with applications to loss reserve variability, prediction intervals and risk based capital", CAS Loss Reserve Seminar 1990

[18]WRIGHT T.S. (1992) "Stochastic Claims Reserving when Past Claim Numbers are Known", PCAS 1992

[19]REID D.H. (1978) "Claim Reserves in General Insurance", JIA 105

[20]TAYLOR G.C. (1981) "Speed of Finalization of Claims and Claims Run-Off Analysis", ASTIN Bulletin 12, pages 81-100

[21]TAYLOR G.C. (1983) "An Invariance Principle for the Analysis of Non-Life Insurance Claims", JIA 110, pages 205-242

Appendix 1: Generalized Linear Models ("GLMs")

When we construct a model, we are trying to find a mathematical structure which describes features of a set of data. In a stochastic model, we assume that the model has a systematic element and a random element. GLMs comprise models whose systematic and random components have a certain structure.

Consider modelling a set of data as a realisation of a set of random variables $(Y_1, Y_2, \ldots Y_n) = \underline{Y}$, with expected values $(\mu_1, \mu_2, \ldots \mu_n) = \underline{\mu} = E(\underline{Y})$. GLMs are then characterised by three features:

1. The Y_i's have a distribution belonging to the exponential family. This includes Normal, Poisson, Binomial, Gamma and χ^2 distributions amongst others.

2. There is a set of factors affecting the model through what is called a Linear Predictor, $\underline{\eta}$, where $(\eta_1, \eta_2, \ldots \eta_n) = \underline{\eta}$ and:

$\eta_i = \Sigma x_{ij} \beta_j$ for $j=1, \ldots p$, where x_{ij} are a series of factors affecting the model.

3. The linear predictor, $\underline{\eta}$, is connected to the model by what is called a Link function, such that:

$$\underline{\eta} = g(\underline{\mu})$$

g has to satisfy certain conditions, such as being monotonic and differentiable.

A simple example of a GLM may make the above characteristics clearer. Consider the classical linear model (otherwise known as the General Linear Model):

$$y_i = \sum_{j=1}^{p} x_{ij} \beta_j + \varepsilon_i, \text{ for } i = 1, 2, \ldots n, \text{ where } \varepsilon_i \text{ is } N(0, \sigma^2).$$

This is a particular case of a GLM where:

1. The distribution is Normal.

2. The predictor is $\eta_i = \sum_{j=1}^{p} x_{ij} \beta_j$, for i=1,2,.....n.

3. The Link function is "identity", that is $\underline{\eta} = g(\underline{\mu}) = \underline{\mu}$.

GLMs cover a variety of familiar techniques, such as regression, analysis of variance, analysis of contingency tables and so on. With the range of possible distributions and Link functions, they also provide a much wider set of models which can be applied in a reserving context, premium rating or a variety of other actuarial applications.

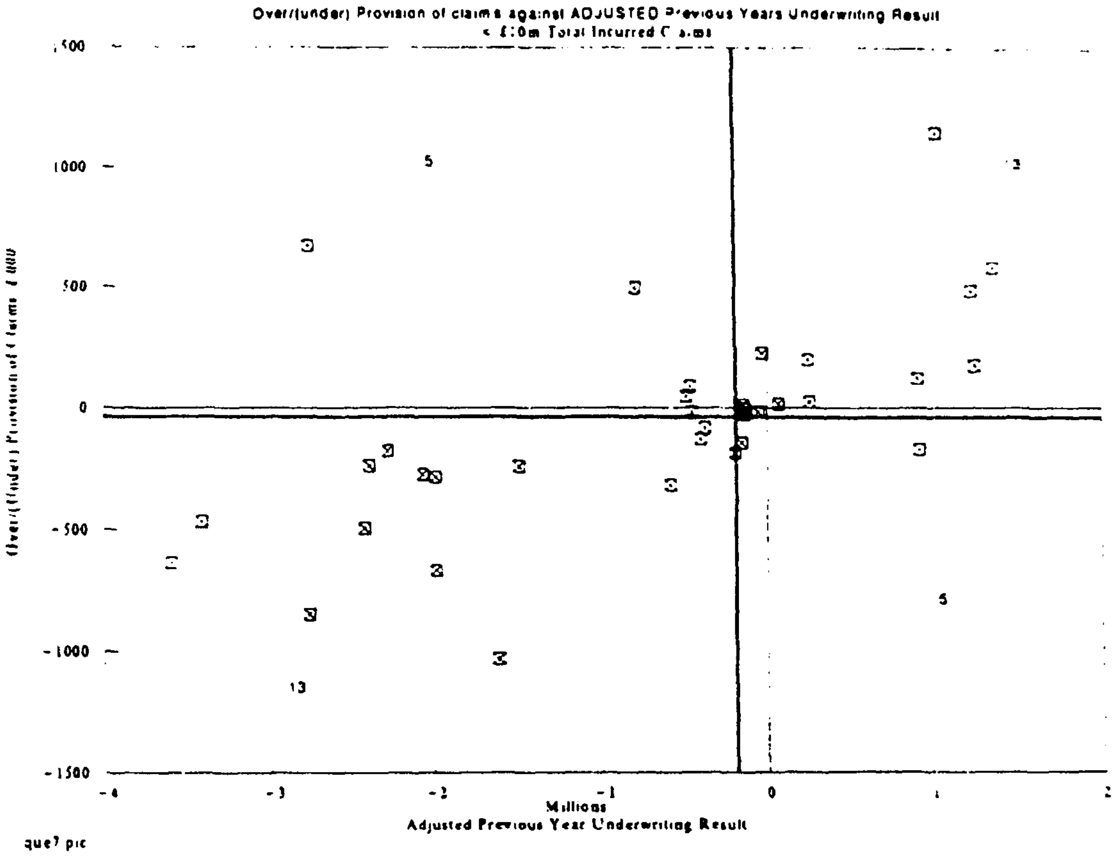
Appendix 2: Glossary

Below is a brief list of terms commonly used in the statistical techniques described in this paper, and their associated definitions

Akike Information Criterion (AIC)	A discriminatory statistic to indicate the level of significance of the number of parameters used in a model relative to the number of degrees of freedom available. The measure contains a penalty factor if too many parameters are used.
Bayes Information Criterion (BIC)	A variant on AIC, which is generally believed to be more powerful.
General Linear Models	special case of the Generalized Linear Model with the Link function set equal to the identity. This corresponds to a linear model with a normal error structure.
Generalized Linear Models	Linear Model applied to data transformed by the link function
Kalman Filter	In the claims reserving context, this filter is used as a smoothing algorithm. It may be thought of as the recursive use of Bayes theorem. With conventional approaches to development pattern modelling, one must assume either that all years have the same development pattern, or else that individual years, or groups of years are independent. Use of the Kalman filter allows the fitted pattern to change or adapt smoothly across years.
Log-Space	the original claims data is transformed by taking logarithms of the incremental payments, and regression is then performed on this. The data is now said to be in "Log-Space".
Operational Time	Operational time (τ) is defined as the proportion of all claims closed to date.

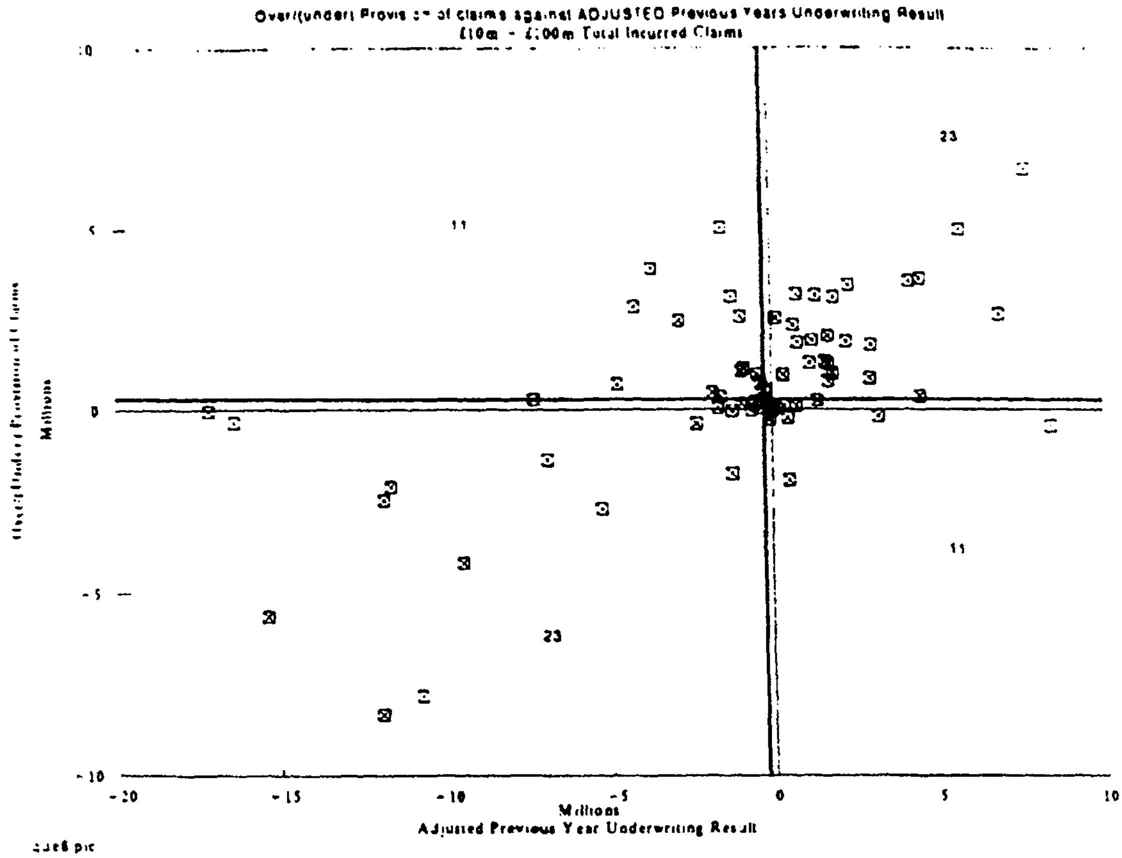
Standard Deviation	root mean square deviation of a statistic
Standard Error	standard deviation of the estimate, allowing for parameter uncertainty

Appendix 3: Charts of the over/(under) Provision of Claims against Adjusted Previous Year's Underwriting Result



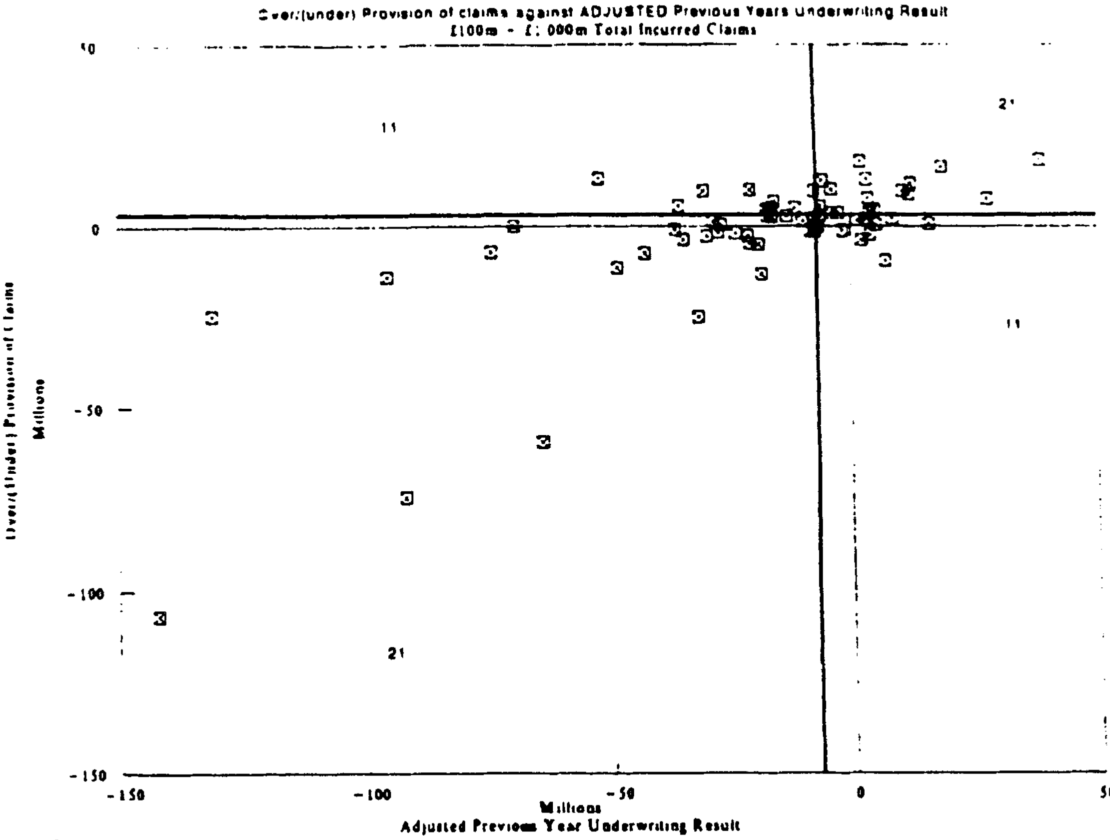
Note: ADJUSTMENT has simply added over/(under) provision emerging in year t+1 back into the underwriting result in year t.

Chi Squared = 7.1



Note: ADJUSTMENT has simply added over/(under) provision emerging in year t+1 back into the underwriting result in year t.

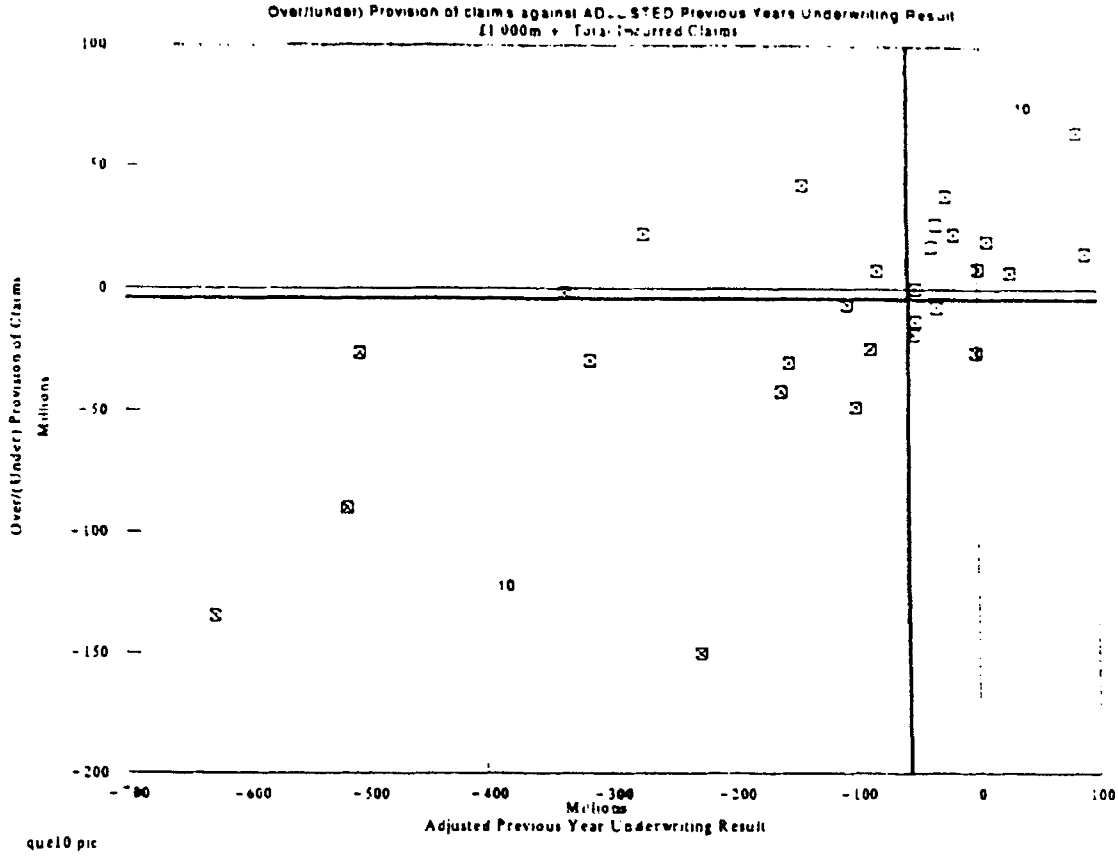
Chi Squared = 8.5



aue9.pic

Note: ADJUSTMENT has simply added over/(under) provision emerging in year t+1
back into the underwriting result in year t.

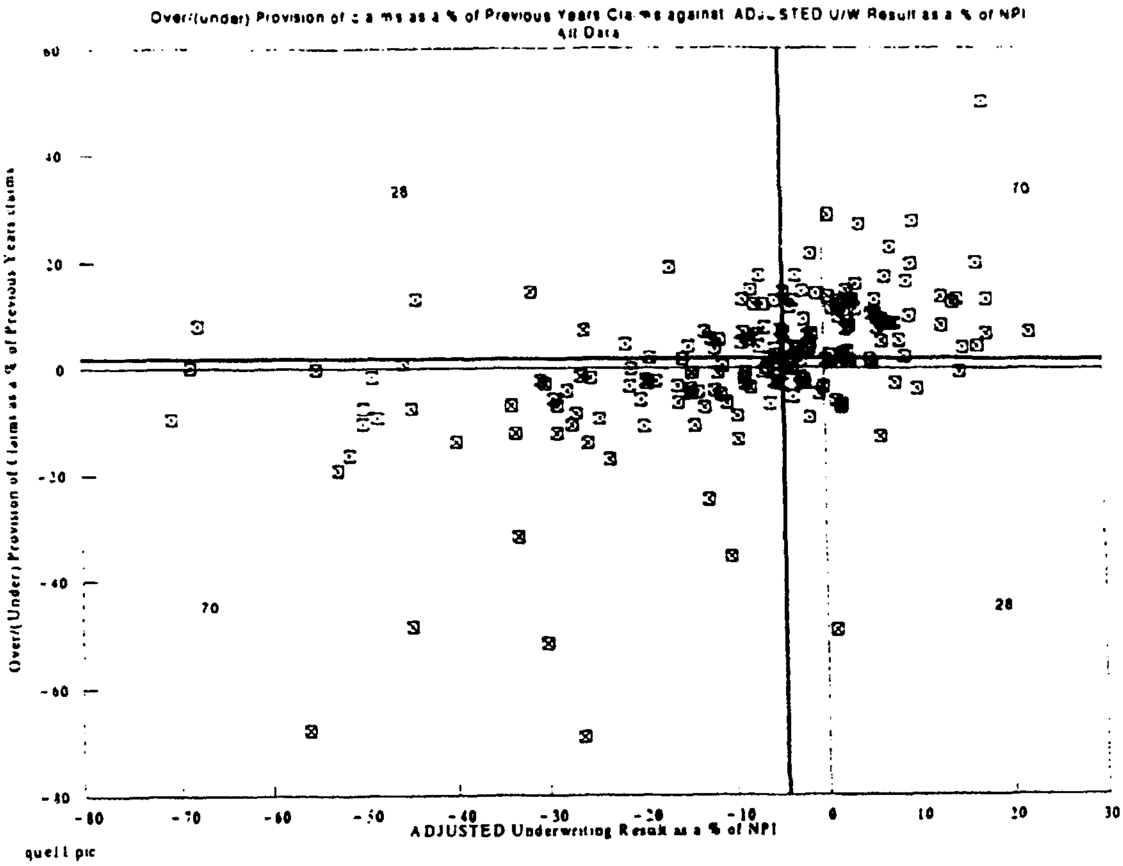
Chi Squared = 6.25



que10 pic

Note: ADJUSTED has simply added over/(under) provision emerging in year t+1
back into the underwriting result in year t.

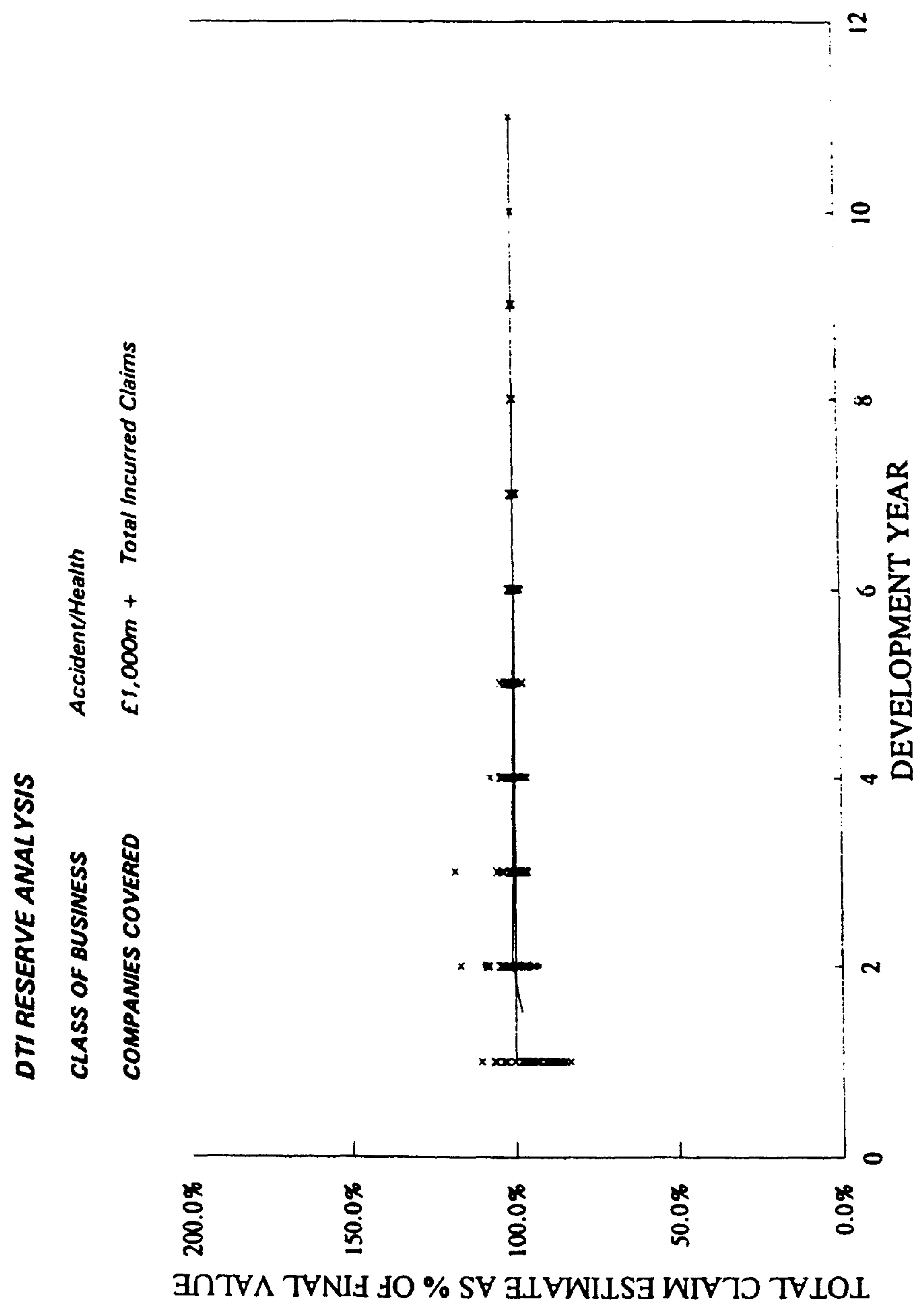
Chi Squared = 5.1

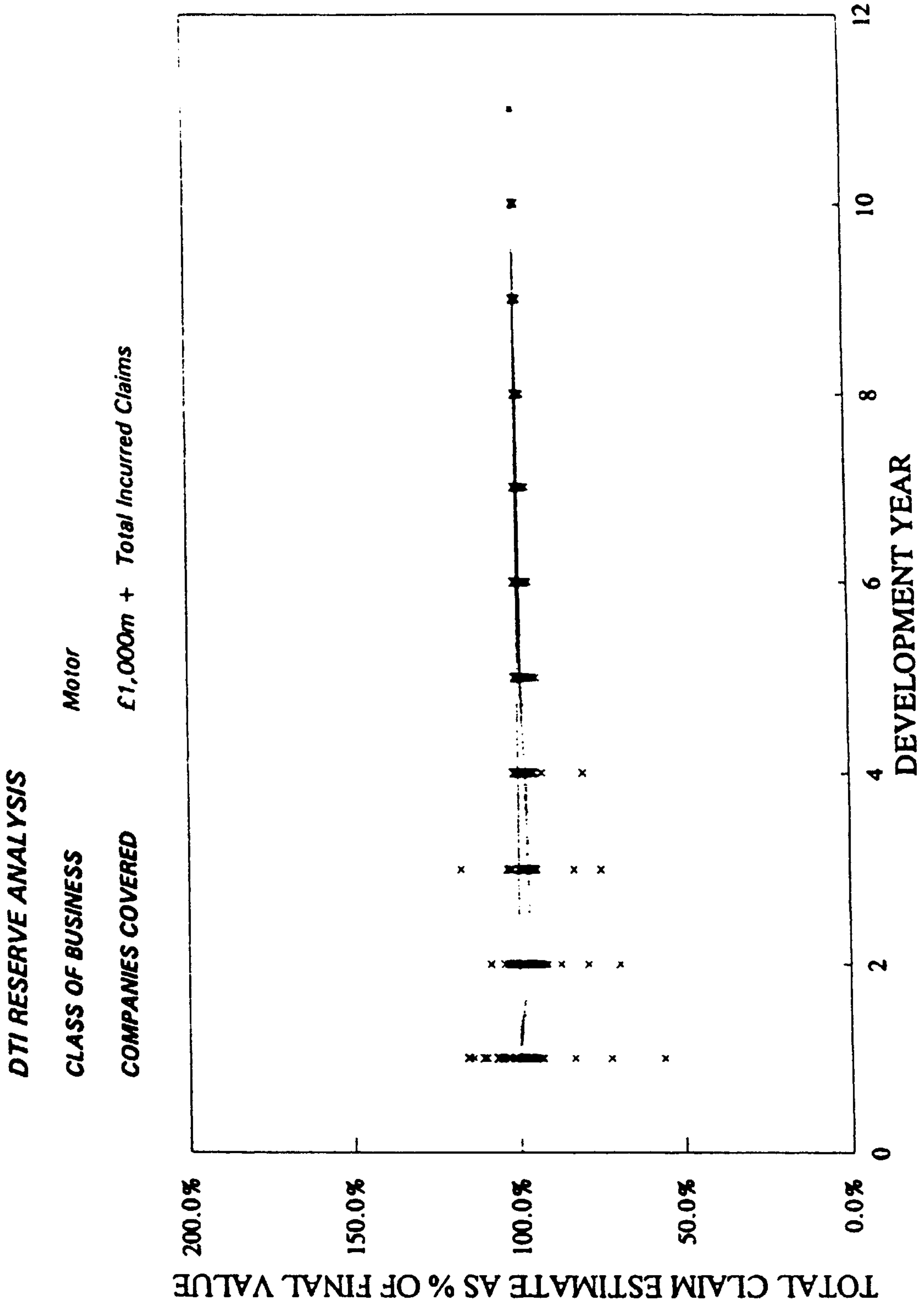


Note: ADJUSTED has simply added over/(under) provision emerging in year t+1
back into the underwriting result in year t.

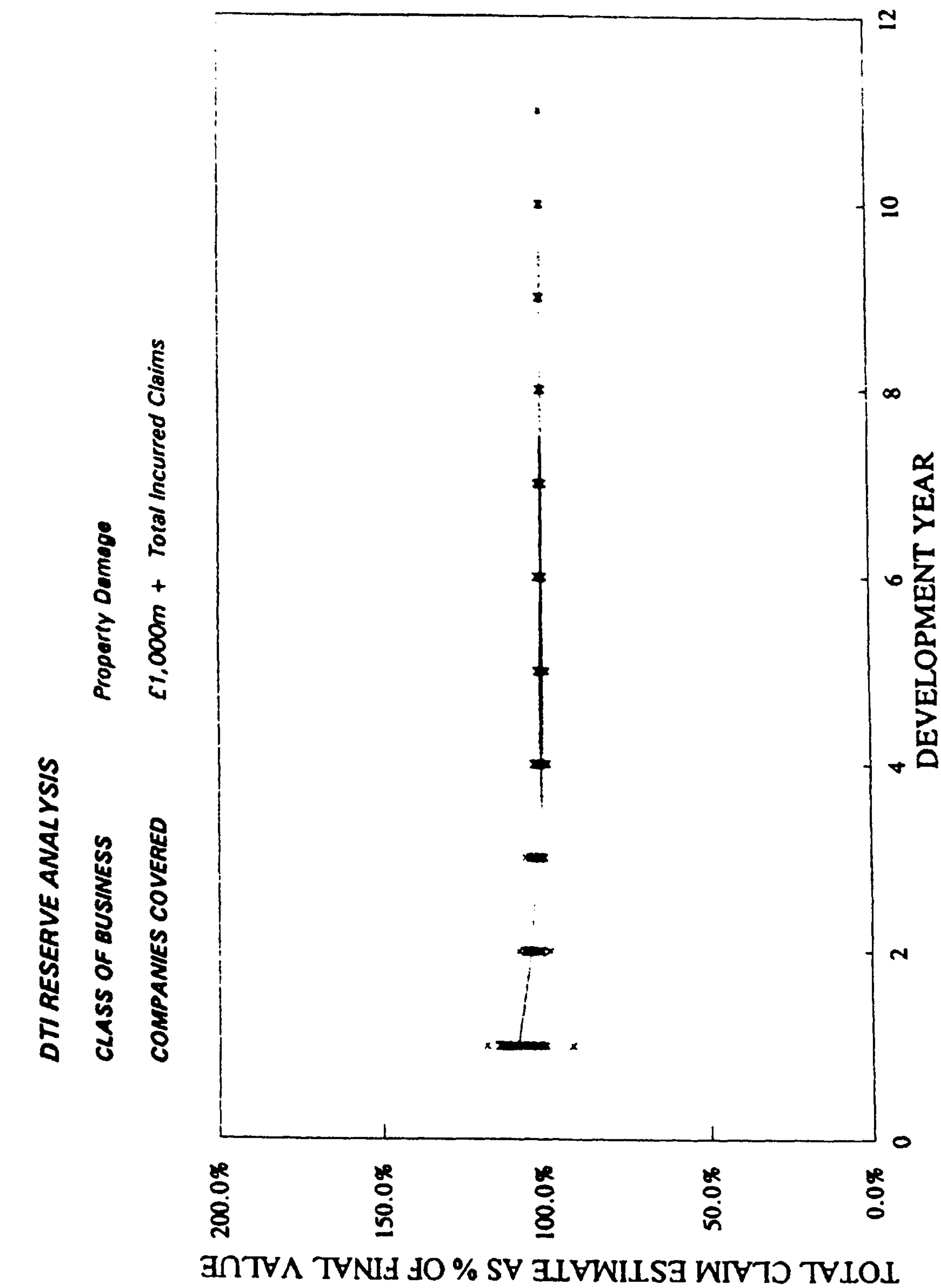
Chi Squared = 36.0

Appendix 4: Graphs of the development of the Total Claims Estimate, grouped by Accounting Class and Size of Business

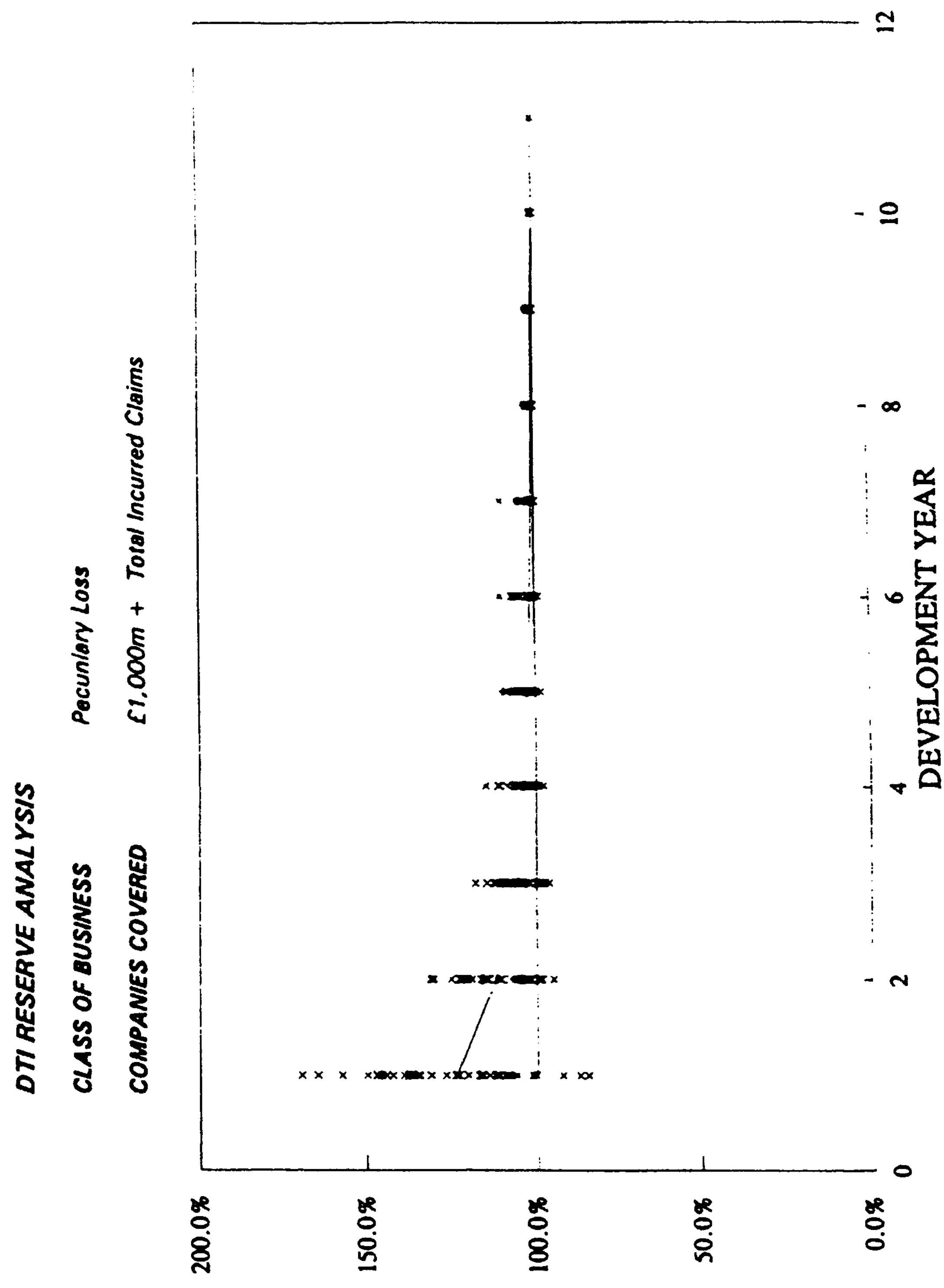


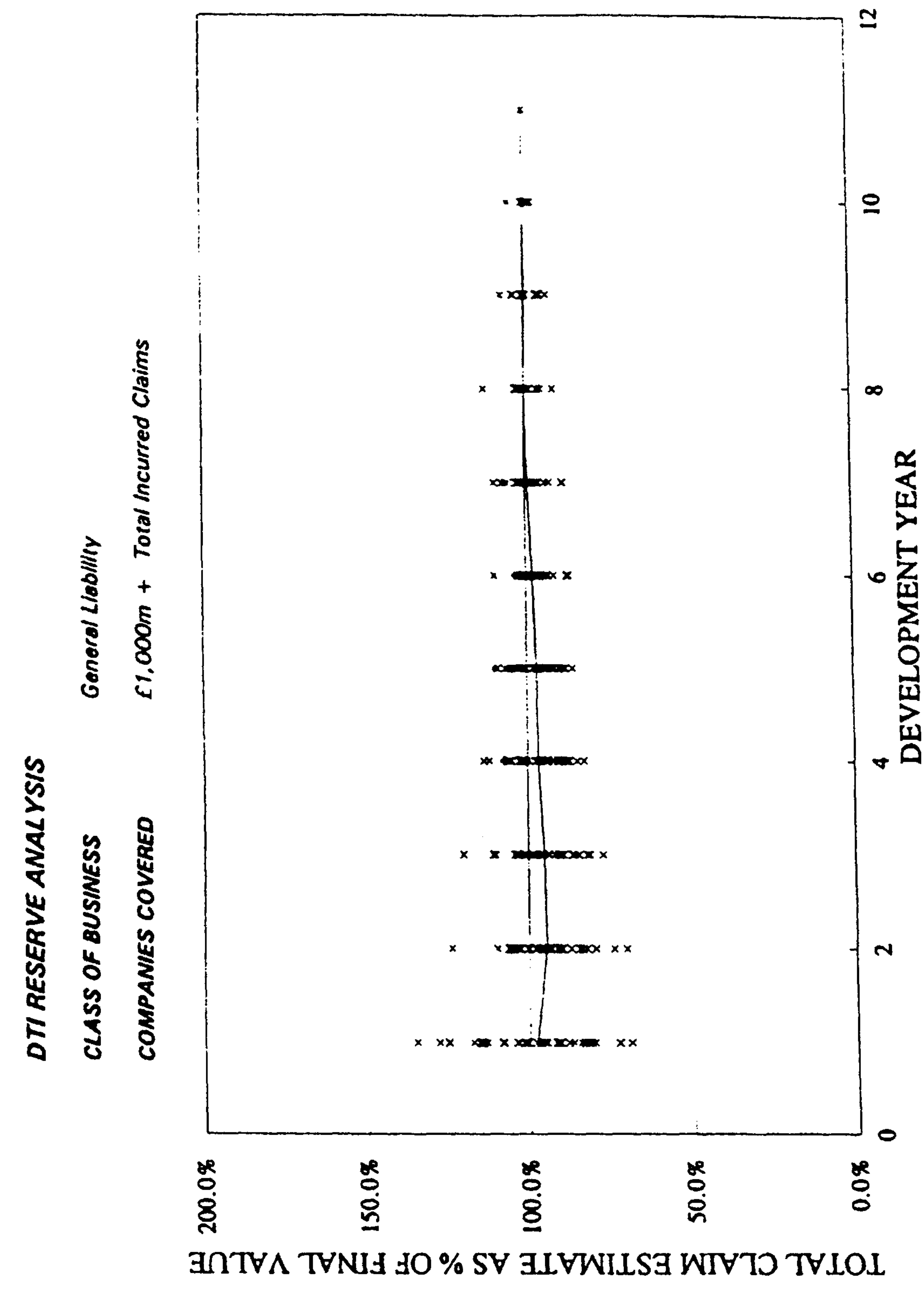


DTI RESERVE ANALYSIS
CLASS OF BUSINESS Motor
COMPANIES COVERED £1,000m + Total Incurred Claims
TOTAL CLAIM ESTIMATE AS % OF FINAL VALUE
200.0%
150.0%
100.0%
50.0%
0.0%
0
2
4
6
8
10
12
DEVELOPMENT YEAR

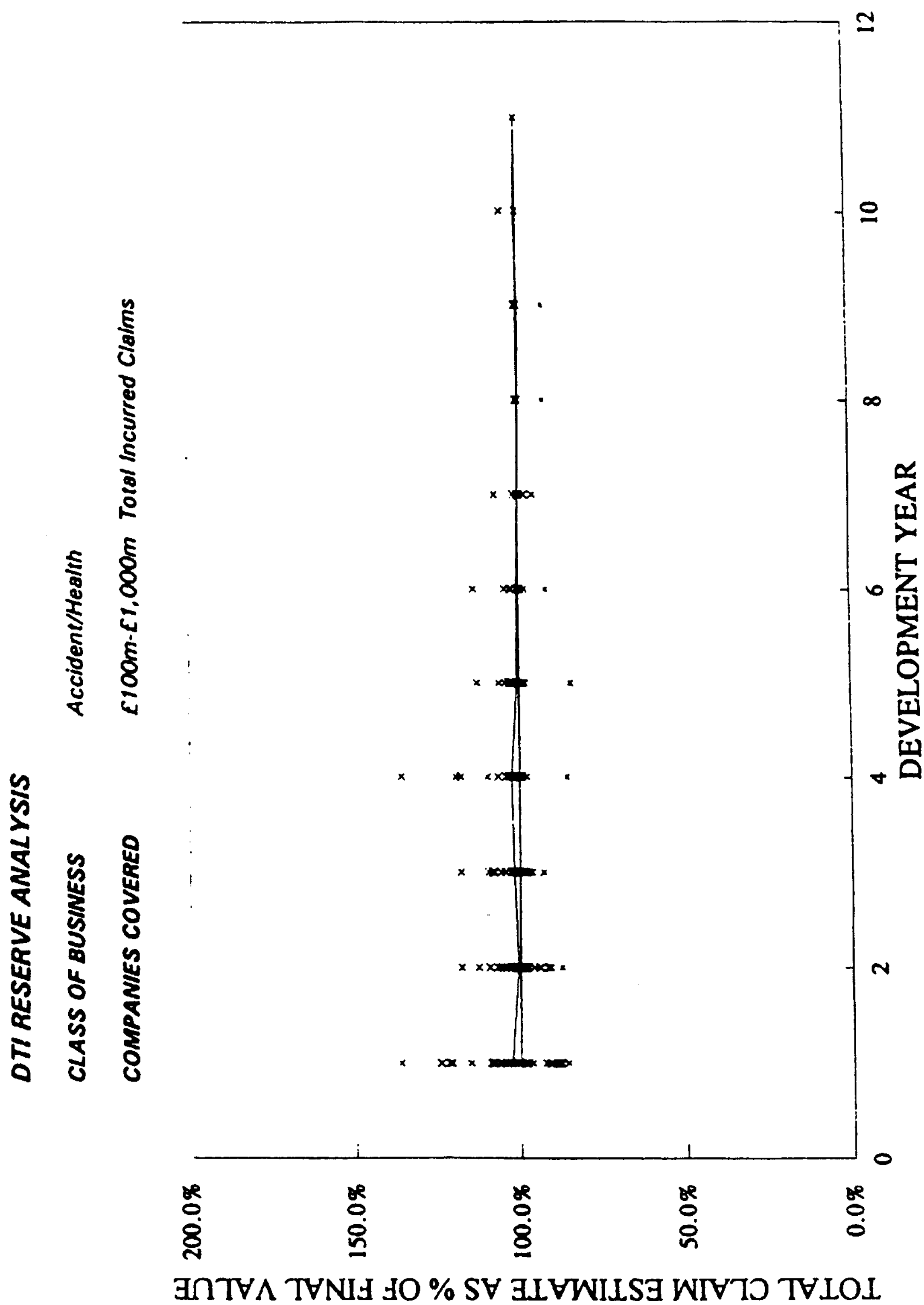


Variance in Claim Reserving Page112


DTI RESERVE ANALYSIS
CLASS OF BUSINESS
Pecuniary Loss
COMPANIES COVERED
£1,000m + Total Incurred Claims
200.0%
150.0%
100.0%
50.0%
0.0%
0
2
4
6
8
10
12
DEVELOPMENT YEAR
TOTAL CLAIM ESTIMATE AS % OF FINAL VALUE



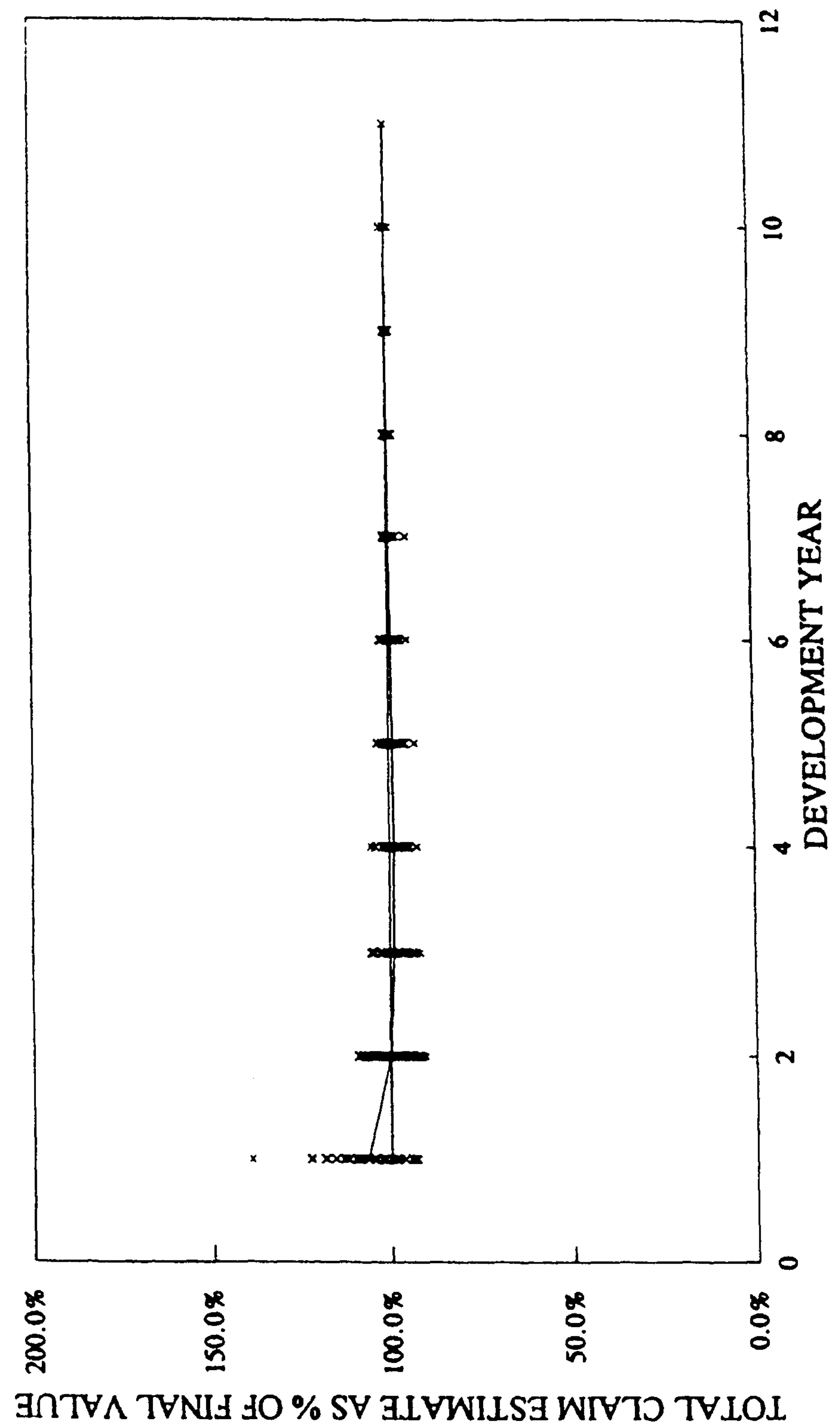
Variance in Claim Reserving Page114


DTI RESERVE ANALYSIS
CLASS OF BUSINESS Accident/Health
COMPANIES COVERED £100m-£1,000m Total Incurred Claims
TOTAL CLAIM ESTIMATE AS % OF FINAL VALUE
200.0%
150.0%
100.0%
50.0%
0.0%
0
2
4
6
8
10
12
DEVELOPMENT YEAR

DTI RESERVE ANALYSIS

CLASS OF BUSINESS *Motor*

COMPANIES COVERED *£100m-£1,000m Total Incurred Claims*

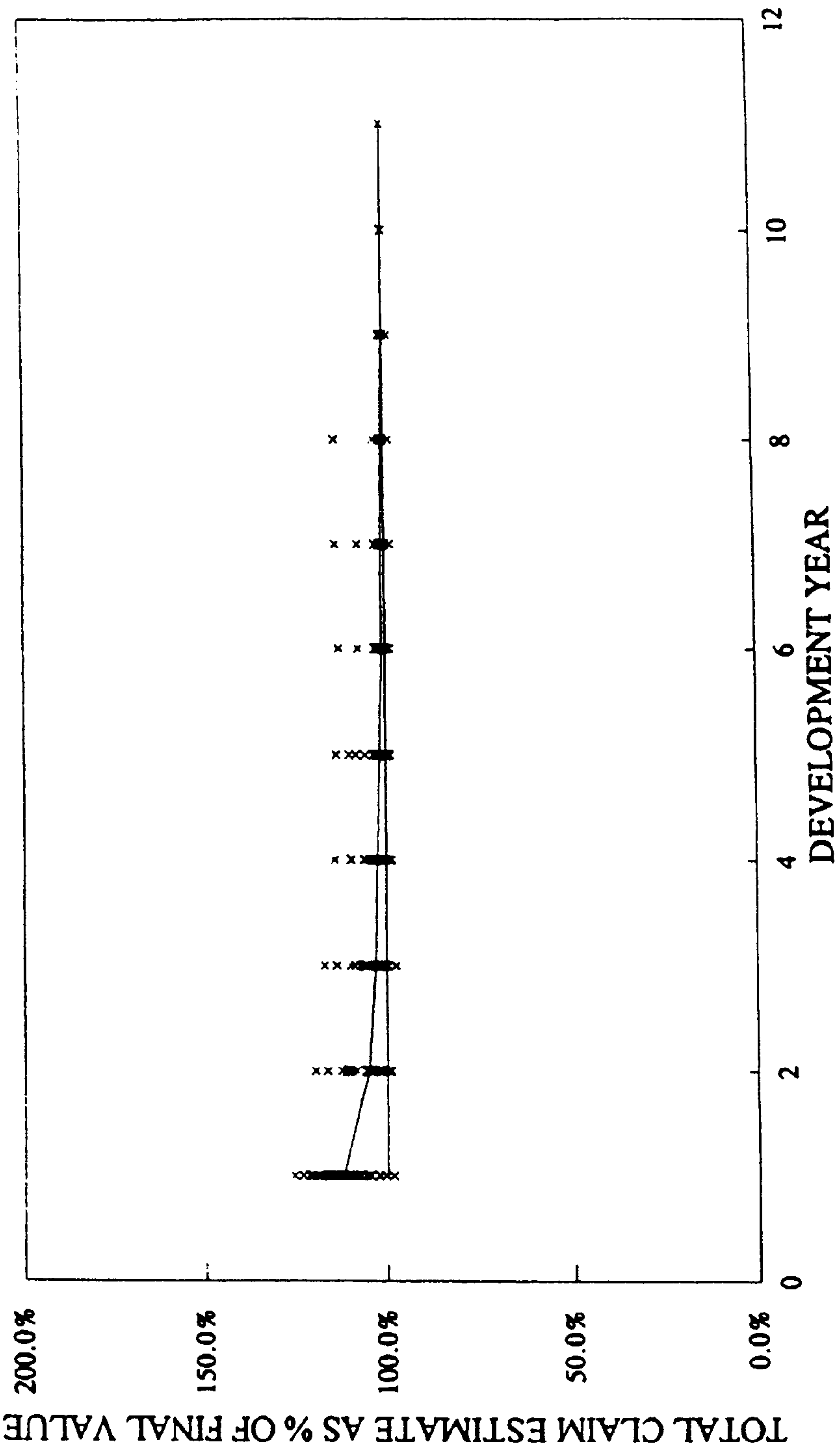


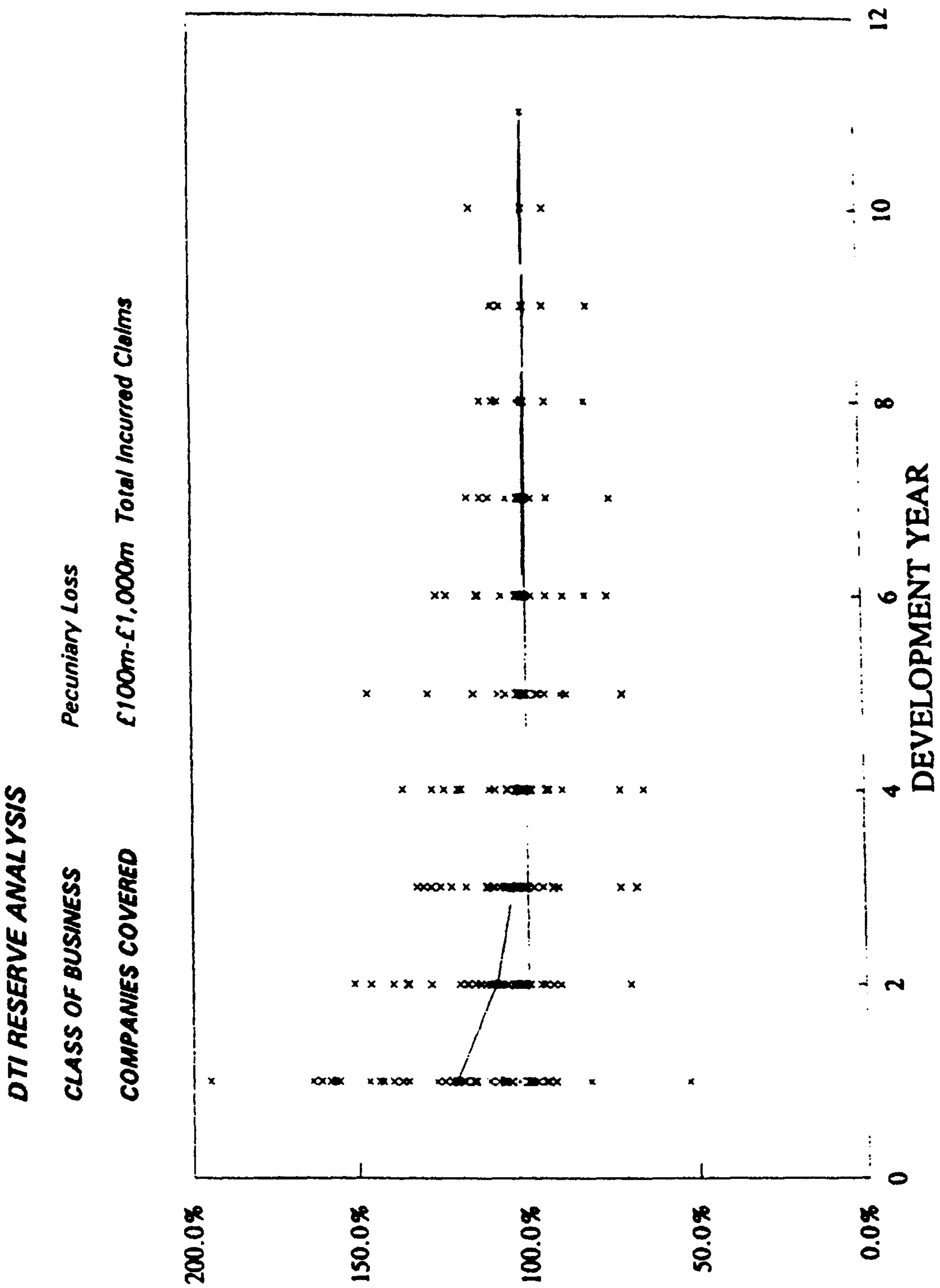
Variance in Claim Reserving Page116

DTI RESERVE ANALYSIS

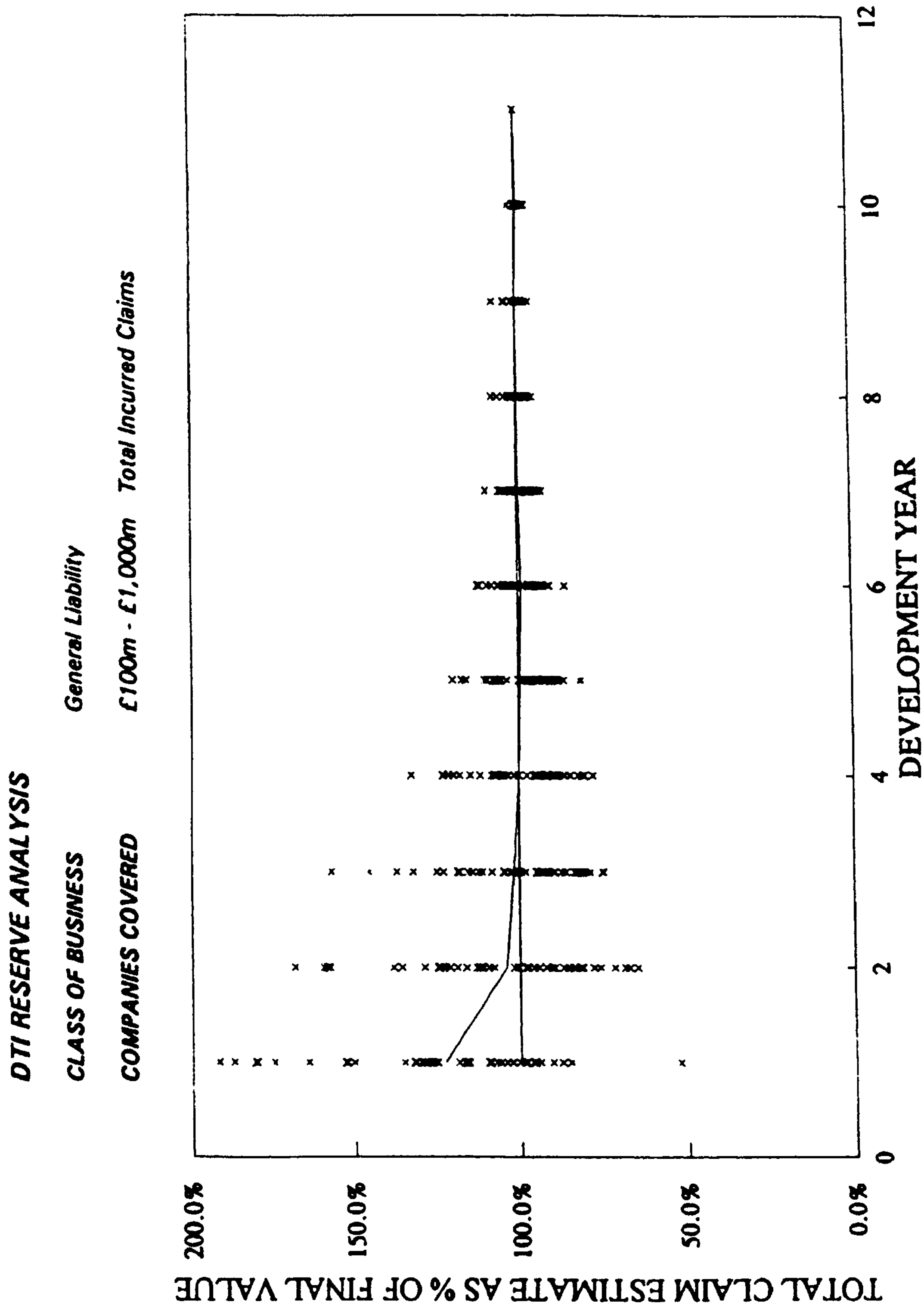
CLASS OF BUSINESS Property Damage

COMPANIES COVERED £100m-£1,000m Total Incurred Claims

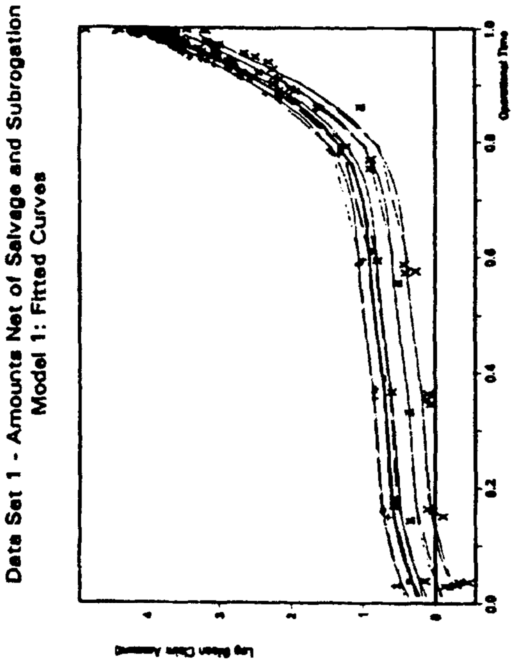


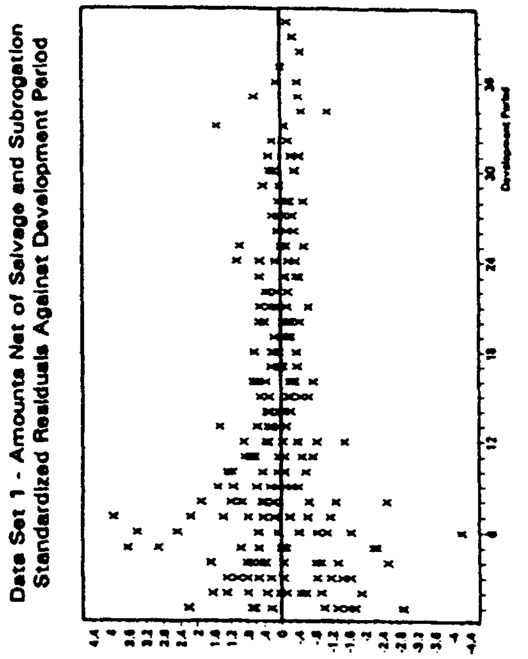


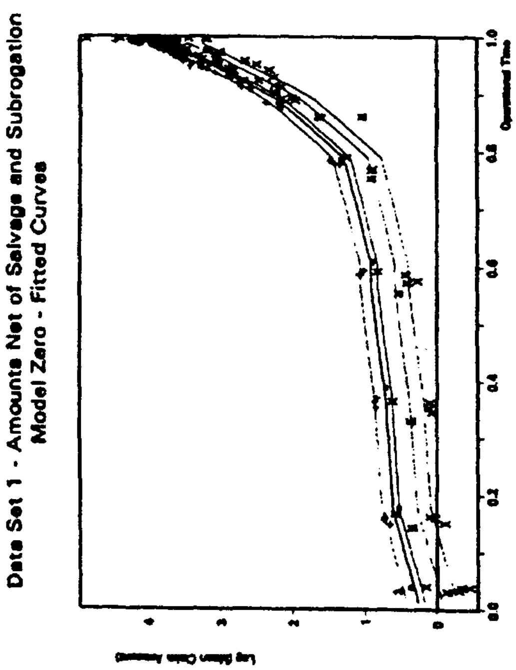
Variance in Claim Reserving Page118

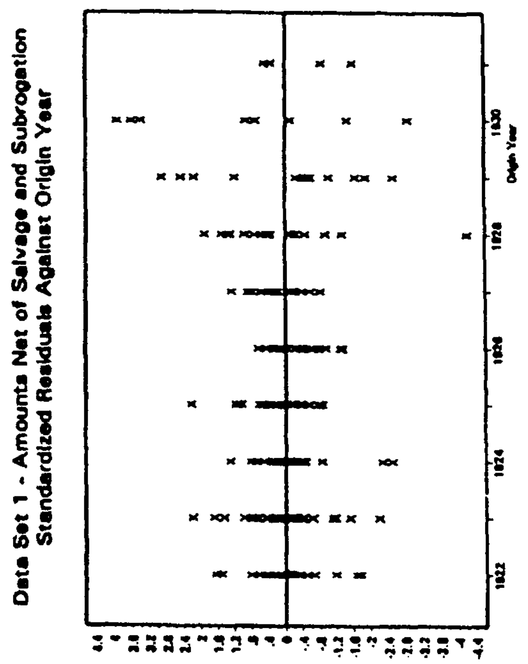

DTI RESERVE ANALYSIS
CLASS OF BUSINESS General Liability
COMPANIES COVERED £100m - £1,000m Total Incurred Claims
TOTAL CLAIM ESTIMATE AS % OF FINAL VALUE
200.0%
150.0%
100.0%
50.0%
0.0%
DEVELOPMENT YEAR
0
2
4
6
8
10
12

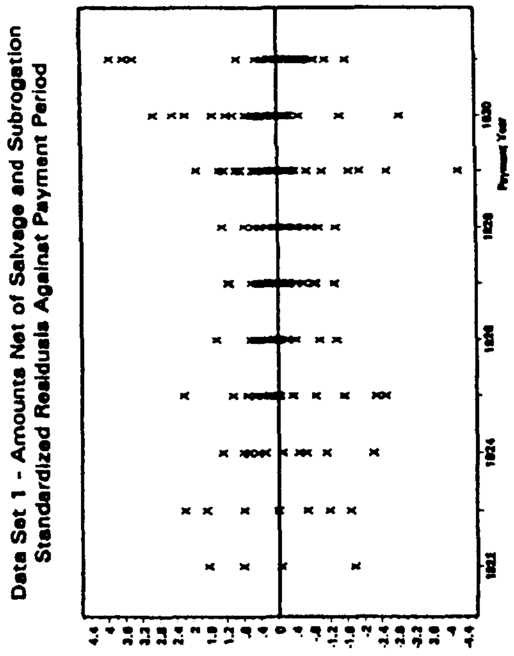
Appendix 5: Charts of Analysis For Operational Time

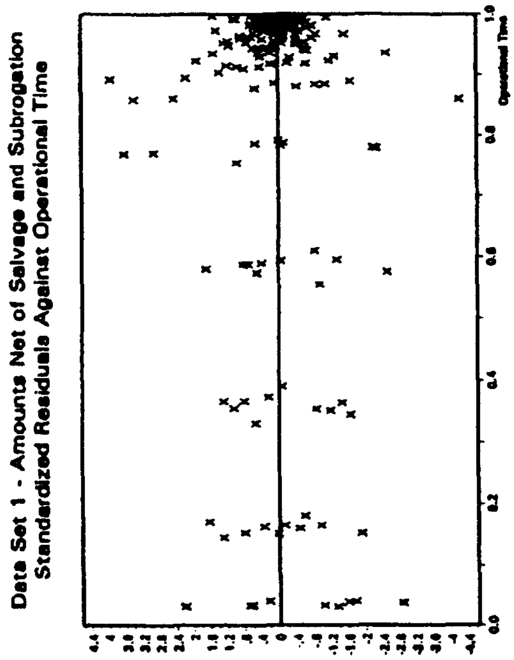




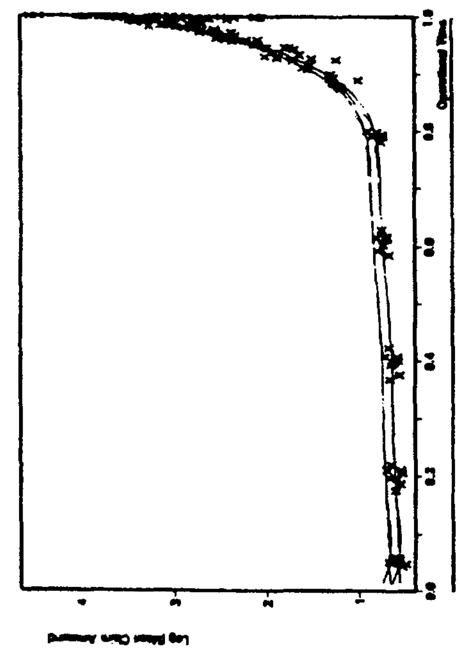




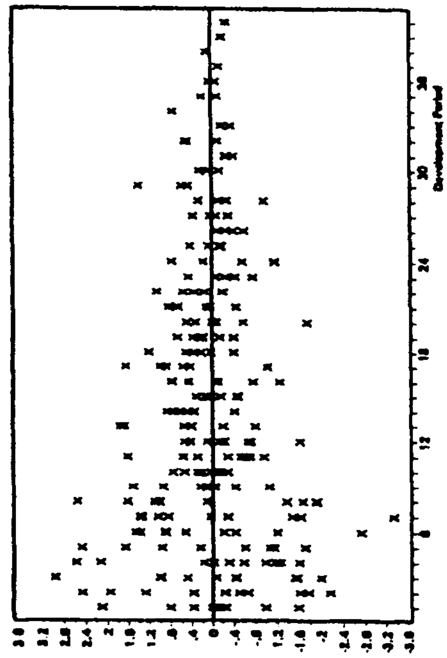

Data Set 1 - Amounts Net of Salvage and Subrogation
Standardized Residuals Against Payment Period
Payment Year


Data Set 1 - Amounts Net of Salvage and Subrogation
Standardized Residuals Against Operational Time
Operational Time

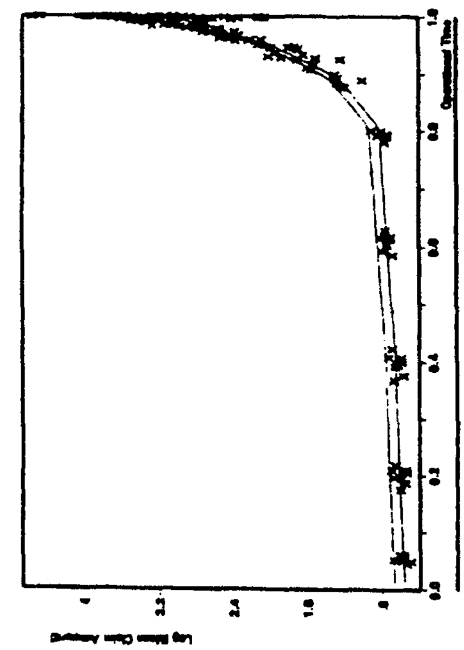
Data Set 3 - Amounts Net of Salvage & Subrogation
Model 1: Fitted Curves



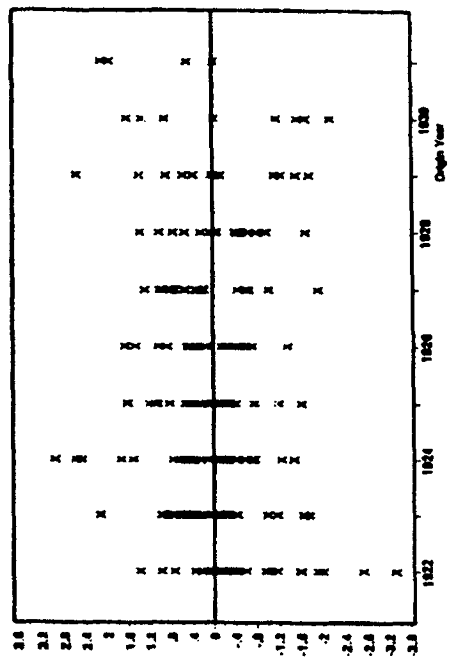
Data Set 3 - Amounts Net of Salvage & Subrogation
Standardized Residuals Against Development Period



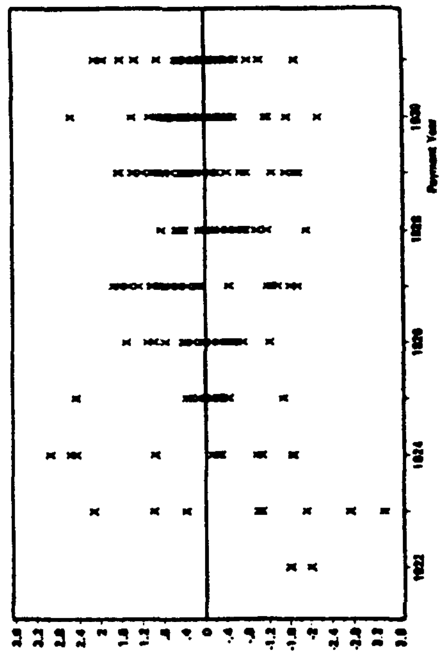
Data Set 3 - Amounts Net of Salvage & Subrogation
Model Zero - Fitted Curves



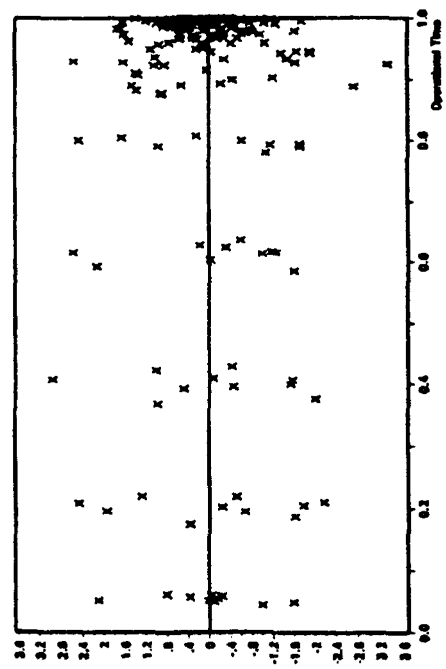
Data Set 3 - Amounts Net of Salvage & Subrogation
Standardized Residuals Against Origin Year

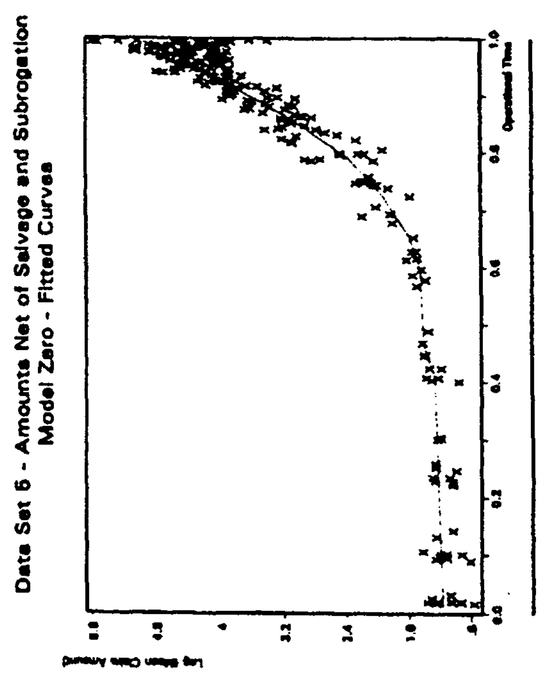


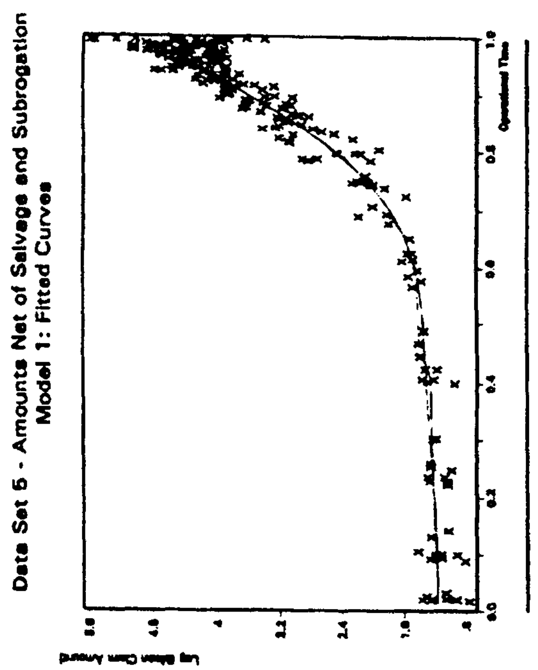
Data Set 3 - Amounts Net of Salvage & Subrogation
Standardized Residuals Against Payment Period

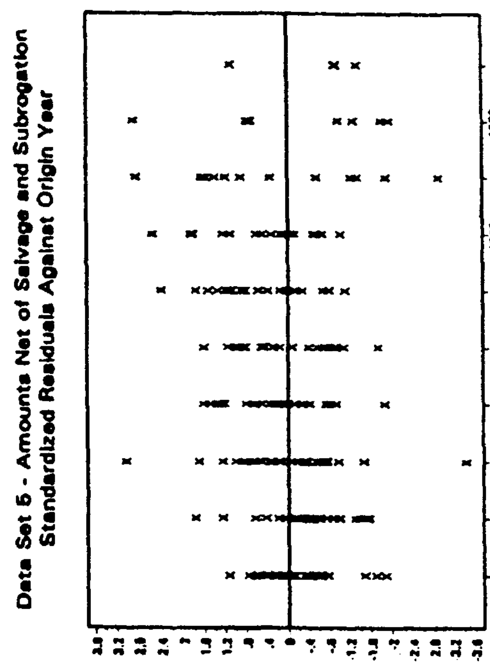


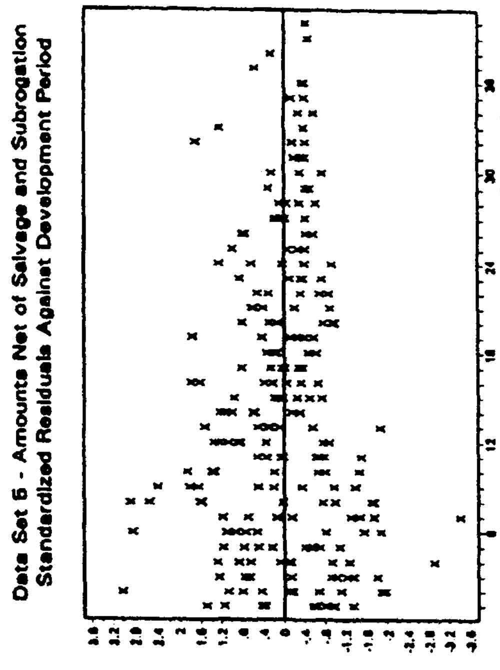
Data Set 3 - Amounts Net of Salvage & Subrogation
Standardized Residuals Against Operational Time

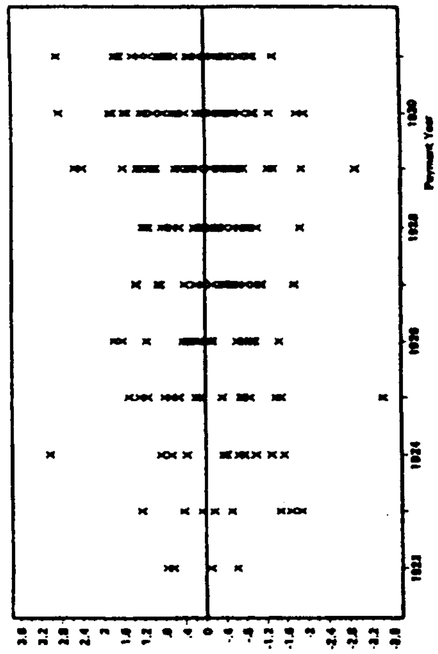



Data Set 5 - Amounts Net of Salvage and Subrogation
Model Zero - Fitted Curves

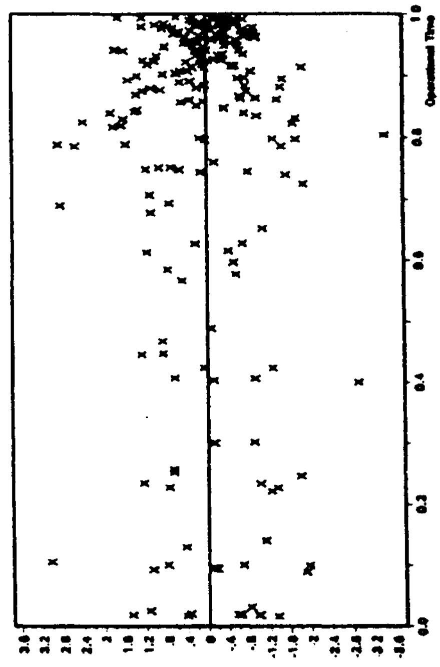

Data Set 5 - Amounts Net of Salvage and Subrogation
Model 1: Fitted Curves


Data Set 5 - Amounts Net of Salvage and Subrogation
Standardized Residuals Against Origin Year


Data Set 5 - Amounts Net of Salvage and Subrogation
Standardized Residuals Against Development Period



Data Set 5 - Amounts Net of Salvage and Subrogation
Standardized Residuals Against Payment Period



Data Set 5 - Amounts Net of Salvage and Subrogation
Standardized Residuals Against Operational Time