Institute of Actuaries of Australia

Reserving Judgement

Considerations Relevant to Insurance Liability Assessment Under GPS210

Andrew Houltram BSc(Ma), FIAA

Presented to the Institute of Actuaries of Australia
XIV General Insurance Seminar 2003
9-12 November 2003

The Institute of Actuaries of Australia
Level 7 Challis House 4 Martin Place
Sydney NSW Australia 2000
Telephone: +61 2 9233 3466 Facsimile: +61 2 9233 3446
Email: insact@actuaries.asn.au Website: www.actuaries.asn.au

RESRVING JUDGEMENT

Considerations Relevant to Insurance Liability Assessment Under GPS210

CONTENTS

Page

1 Introduction ... 2

2. The Central Estimate ... 5
- 2.1 The Central Estimate as the Mean ... 5
- 2.2 Random Samples from Skewed Distributions and Removal of Outliers ... 6
- 2.3 Qualitative Advice ... 8
- 2.4 Data Verification ... 8
- 2.5 Judgement ... 13

3. Stand-Alone Risk Margins ... 14
- 3.1 Order of Determination ... 14
- 3.2 Risk Margin Components – Process Variability and Estimation Error ... 14
- 3.3 Quantitative Methods used to assess Risk Margins ... 19
- 3.4 Testing Performance of Analytical Methods Against known (artificial) Processes ... 23

4. Correlations and Diversification Benefit ... 37
- 4.1 'Direct' and 'Indirect' Sources of Correlation ... 37
- 4.2 Empirical Analysis that would support Diversification Benefit Assessment (and why it can't be done) 40
- 4.3 Existing Guidance to Actuaries assessing Correlation and Diversification ... 42
- 4.4 Exploring the Meaning of Correlation Coefficients ... 46
- 4.5 Summary and Conclusions ... 50

5. Conclusions ... 52

Bibliography ... 53
Acknowledgements ... 53

APPENDICES

A. Checks of Chain Ladder Fit
B. Mack Method
C. Description of Bootstrapping Procedure
D. Overview of Modelling Zehnwirth's PTF Family with the ICRFS-Plus Package
E. Detail of Two Illustrative Claims Generating Processes
F. Comparison of 'True' Distribution of possible Claims Outcomes for the Artificial Claims Generating Processes with Assessment by different Quantitative Methods
G. Correlation Scatter Plots

1. Introduction

Actuarial assessment of Insurance Liabilities involves a combination of model fitting and judgement. On occasion, the actuary's aim may simply be to identify a model that fits past experience well, and then to use it to infer properties of projected claims run-off. However, the task is generally less straightforward. For a variety of reasons, claim development patterns may not be stable over time. A wide range of items that are subject to change can influence them. Some examples include:

- Business mix
- Exposure
- Policy conditions
- Legislative change
- Propensity to claim
- Claims management
- Judicial change
- Price change
- Technological change
- Insurer staff turnover of key personnel
- Change in reinsurance arrangements

Actuaries performing Insurance Liability assessments must form a view about the extent to which such changes have occurred in the past, what their effect has been (and will be) on claims run-off, and whether further changes are possible through the run-off period.

Even where none of these influences can be identified as playing a role, it is not uncommon to see judgement made that past experience is unrepresentative of likely future experience. This can be due to items such as:

- Presence or absence of catastrophe events
- Presence or absence of large claims
- Presence of unusually high or low superimposed inflation

An example of this sort of judgement relates to the removal of 'outliers' from the claims history to prevent them from exerting 'undue influence' on the valuation result. Distinguishing outliers worthy of this treatment, from genuine features of claims experience that capture information about likely future experience, is one of the more important assessments associated with Insurance Liability estimation.

Another judgement area relates to fitting a tail to the run-off when there is an insufficient history to base projections on experience.

Actuaries often form their view of likely future experience with input from the insurers' underwriters, claims managers and senior management. Valuable insights concerning likely claim cost progression can be gained from these sources. However, there can be a danger associated with putting weight on advice that lowers the Insurance Liability estimate if the advice can't be objectively verified, particularly if it comes from those with a vested interest in the result.

Notwithstanding that it is the actuary who is responsible for the Insurance Liability estimate, it is not uncommon for projection assumptions to be influenced by views sourced from others. The extent to which the Insurance Liability estimate incorporates these qualitative influences can depend on:

- The vigour with which the sources make their point of view.
- The readiness or otherwise of the actuary to incorporate them in their estimate.
- The strength and quality of the line of reasoning.
- The ability of the actuary to verify the qualitative advice.
- The track record of other qualitative advice from similar sources.

It is quite proper that the Approved Actuary absorbs and assesses qualitative advice, determines whether environmental or other changes can be identified which will impact run-off patterns, and forms a view about whether past experience is representative of likely future experience. The alternative of mechanically projecting current run-off based on past patterns will rarely produce a better result.

In addition to quantitative analysis, <u>actuarial judgement</u> forms a vital component of Insurance Liability assessment. This paper explores some of the implications of the judgement application to the assessment of:

- Central estimates:
- Risk margins, and
- Diversification benefits.

Application of judgement can increase the risk of inadvertently introducing bias into the central estimate assessment, it can contribute to violation of assumptions that underlie quantitative techniques used to assess risk margins, and it can affect correlation between adequacy of central estimates across business classes. These effects are important, and need to be considered as part of Insurance Liability assessment.

Risk margins provide protection against variation in actual claims outcomes away from the central estimate. The sources of this variation can be categorised as those relating to:

- Intrinsic variability about the 'true' mean (*process variability*); and
- Risk that our central estimate may be different from the true mean (*estimation error*).

One reason the second component arises, is that claims history can be considered as a set of observations that is a random sample from of an underlying distribution. Sampling error means that claim projection parameters based on the historical observations, will be different from the true value. In Section 2, I aim to illustrate that the skewed nature of claims cost distributions mean that it is more likely that claim projection parameters will be understated than overstated if based on a historical average.

A second reason why *estimation error* arises is that it is introduced by real and/or assessed changes in portfolio conditions necessitating the application of judgement. This is discussed in Section 3, where I describe some of the forms this risk can take. Many of the forms have impact that does not decrease in relative size as portfolio size increases. By definition, they are therefore systemic. In Section 4, I argue that they can be a significant source of correlation between adequacy of insurance liability estimates, and that they therefore act to reduce diversification benefits.

Reserving judgement is a common feature of central estimate assessment. Further, just as mechanical application of an actuarial method to assess central estimates will rarely produce the most suitable result, so too mechanical application of various methodologies and formulae to arrive at risk margin and diversification benefit allowances will rarely provide a good result. Judgement is just as important an element of risk margin and diversification assessment as it is for the central estimate.

In Section 3, I explore of some quantitative risk margin assessment methods; noting sources of claim outcome variability from the central estimate that they do and do not capture, and some of the problems that result when the methods are used even though the assumptions that they rely aren't satisfied.

In Section 4, I explore issues pertinent to correlation and diversification benefit assessment, and include observations on the results set out in the Bateup & Reed and Collings & White papers.

This paper does not provide *the answer* to how Insurance Liabilities can be objectively assessed in practice. In my view judgement remains a valuable component that should add reliability to the result. There seem to be features of general insurance liabilities and assessment methods that increase risk of understatement of central estimates and risk margins and risk overstatement of diversification benefits. An aim of this paper is to draw attention to these effects. Being aware of them is an important step toward limiting the likelihood of them leading to Insurance Liability mis-statement.

For the majority of this paper, my discussion more directly relates to Outstanding Claims Liabilities than Premium Liabilities, though much of what is said is applicable to both. I have also largely ignored what might be called the '*half coefficient of variation rule*'. This requires those companies where the Approved Actuary judges the distribution of possible claims outcomes to be particularly variable and skew to maintain their reserves at a level targeting a higher probability of adequacy than the 75% level that generally applies.

Insurance Liability assessment under GPS210 is a complex topic. Through this paper, I hope to contribute to the development of actuarial understanding and practice in this important area by drawing attention to considerations that in my view are important, but haven't been widely discussed in the literature on the topic to date.

2 The Central Estimate

2.1 The Central Estimate as the Mean

Prior to GPS210, a range of interpretations of the term '*central estimate*' was possible. Both GPS210 and PS300 now define it more precisely as the expected value of the liabilities. For this purpose, expected value is defined as the statistical mean (ie if our assessment of all the possible values of the liability that could be realised is expressed as a statistical distribution, the central estimate is the mean of that distribution).

For most insurance classes, this probabilistic distribution is generally accepted to be right-skewed. If outcomes are better than anticipated, there is generally less scope for them to be much better than anticipated than there is for the outcomes to be much more adverse.

The definition of the central estimate carries several implications that create practical difficulties for the Approved Actuary:

- For typical long-tail claims outcome distributions, the APRA central estimate definition will sit significantly higher than the most likely outcome. Where the valuation is conducted outside of a statistical framework, there is natural tendency to set projection assumptions toward the most likely outcome. This can give rise to a risk of central estimate understatement. Where the tendency is resisted, communication issues can arise since projection assumptions that target a mean can look conservative to those who don't fully appreciate properties of skewed distributions, or the nature of APRA's central estimate definition.
- In contrast to other measures of centrality such as the median or the mode, assessment of the mean requires (implicit or explicit) formation of a view about the entire probabilistic distribution of possible outcomes. This is a result of the skewed nature of claims cost outcome distributions. In particular, the small possibility that outcomes could be highly adverse needs to be taken into account. Outcomes over, say the 95^{th} or 98^{th} percentile, have the potential to significantly influence the position of the mean, but the actuary will almost certainly be in a position where the allowance must be subjective.
- A number of scenarios may require assessment to properly reflect the mean of the assessed distribution of outcomes. For example, at the time of writing, amendments are under consideration that would, with retrospective effect, improve benefits for injured workers under the statutory workers' compensation scheme in Western Australia. Subject to materiality considerations, a full assessment of the central estimate under the GPS 210 and PS 300 definition would seem to require the Actuary to form a view regarding:
    - The likelihood of the amendments being passed (in full or altered form), and
    - The implications under each scenario for the distribution of possible claims outcomes.

- The mean is an arithmetic concept, rather than one that carries obvious physical meaning readily grasped by all. By contrast, median (being the amount equally likely to be too great or too small to cover the liability that ultimately emerges) is more readily understood. Communication of the meaning of the term 'central estimate' can be difficult.
- In selecting projection assumptions, it will more often than not be the case that projection assumptions estimating the mean will need to be above the average of past experience. This is because the mean of a small sample from a right-skewed distribution is more likely than not to be below the mean of the underlying distribution.

If the central estimate represents a mean then, because the distribution of possible claims cost outcomes is right skewed, when successive actuarial valuations are compared, 'prior year releases' should be more common than 'prior year strengthenings.' Despite the theoretical position, I tend to come across more strengthenings than releases in my audit support role.

2.2 Random Samples from Skewed Distributions and Removal of Outliers

The point that, more often than not, it will be the case that projection assumptions estimating the mean will need to be above the average of past experience, can also be worth bearing in mind when performing a comparison of actual versus expected experience over short periods, or for particular development periods, where outcomes can be considered as random drawings from a skewed distribution.

One should expect (*colloquial meaning*) actual outcomes to be below expected (*statistical meaning*). This may seem a pedantic point, but in my view it is worth having as part of one's mindset to protect against the natural tendency in a non-statistical analysis to select assumptions that reflect the most likely outcome. Our statutory obligations require us to target a mean.

The following example provides an illustration.

The table below lists 10 simulations of five observations of a lognormal distribution with mean 8,000, and standard deviation 6,400. The mean of the 5 observations from each simulation is shown in the row named 'Avg'.

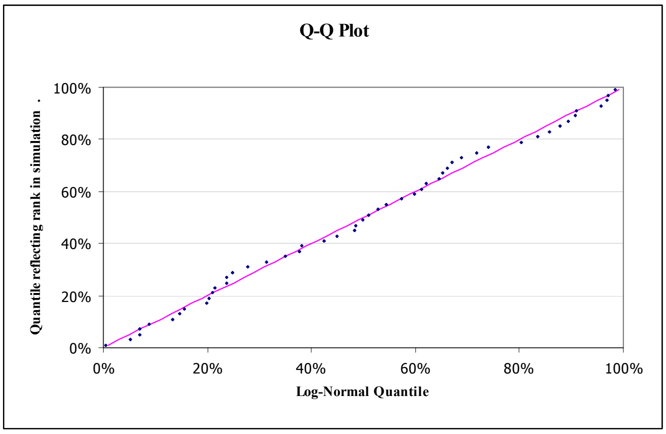
It can be instructive to put oneself in the position of an actuary looking at these as, say, a recent history of PPCI values for a given development period, and seeking to arrive at a projection assumption. I have marked in bold red those observations that I, and I suspect other actuaries, might be tempted to regard as 'outliers' that should be removed from the analysis so as not exert 'undue influence'. The row labelled 'Avg*' is the mean of the observations for each simulation after exclusion of the 'outliers'.

Simulated Values from Lognormal
Mean = 8,000 Standard Deviation 6,400

	Sim 01	Sim 02	Sim 03	Sim 04	Sim 05	Sim 06	Sim 07	Sim 08	Sim 09	Sim 10
Obs1	6,349	13,232	5,018	3,440	3,065	4,769	8,826	**28,840**	15,021	6,553
Obs2	4,126	4,446	1,994	3,484	9,378	6,750	3,573	8,214	16,060	**20,893**
Obs3	3,771	6,069	9,819	7,739	3,769	3,875	7,633	2,851	2,215	7,428
Obs4	1,005	11,400	3,540	**15,909**	5,721	5,050	2,424	8,115	6,242	8,523
Obs5	**23,499**	14,203	2,990	6,083	12,418	8,390	7,121	5,460	23,242	2,210
Avg	7,750	9,870	4,672	7,331	6,870	5,767	5,915	10,696	12,556	9,121
Avg*	3,813	9,870	4,672	5,187	6,870	5,767	5,915	6,160	12,556	6,179

Mean of the Avg amounts = 8,055 Mean of the Avg* amounts = 6,699

It is possible to check that the 50 simulated observations are reasonably representative of the underlying distribution by means of a Q-Q plot. This plot compares the quantile for each observation that reflects its rank in the fifty observation points, with its quantile position with respect to the particular lognormal distribution.



Because the points lie reasonably close to the y=x line in the Q-Q plot, it shows that the sample of fifty observations is not 'unrepresentative' of the underlying lognormal distribution.

One might argue whether a 'typical' actuary presented with the history as shown would include or exclude the outliers, or treat them in some special manner. The point of the illustration is not really affected by this. Its aim is to illustrate:

- It is more likely than not that the mean of a sample from a skewed distribution will understate the true mean of the underlying distribution. (In six of the ten simulations shown, the average of the sample was less than 8,000 – over 25,000 simulations, the sample average was below 8,000 57.1% of the time)
- Not adjusting assumptions to be heavier than recent experience, where there is an absence of outliers in the claims history, can inadvertently introduce bias. (In the ten simulations shown; simulations 02, 03, 05, 06, 07, and 09 might be regarded as having no outliers. In two-thirds of these cases, the sample mean is below the true mean)
- One must be careful when deciding to exclude outliers, in the belief that they are unrepresentative, that inadvertent bias does not result. (The mean of the Avg* amounts is significantly lower than the true mean of the underlying distribution)

Because actuaries deal with skewed distributions, it follows that if the environment is such that the underlying claims process is stable (but still subject to random fluctuation around the underlying stable process), a claim projection that is supposed to target the mean, that bases projection assumptions on an average of recent experience will more often than not underestimate it.

Risk of bias increases if the actuary begins to judge that there are aberrant 'large' observations present in the claims history that are unlikely to be representative of future experience. While there may be occasions when this is genuinely the case, if the claims process can be likened to drawing a random sample from a skewed distribution, it will be far more common that large observations will be under-represented. Therefore, in targeting the central estimate, it should be more common to add a loading to the average of past, than to make adjustments to limit the influence of large 'outliers.'

2.3 Qualitative Advice

Qualitative advice from insurance company claims, underwriting and management staff can also add to risk of central estimate understatement. Except from actuarial peers, it is my impression that it is rare for the actuary to be cautioned by insurance company staff, that too little allowance has been incorporated in the estimates for tail development or large claims, or that development patterns will be longer than previously experienced. The advice is almost always given to help the actuary understand why a lower Insurance Liability estimate is appropriate.

Typically, a company will seek continuous improvement to its underwriting practices and claims management. It is therefore quite natural that qualitative company advice will tend to paint the picture that future experience will be more favourable than the past.

In assessing this advice, the actuary needs to recognise situations where it is provided by those with a vested interest in the result. While one should not be dismissive of the advice in these cases, it is important that it be objectively verified before being given any weight. The Insurance Liability estimate is the responsibility of the Approved Actuary, and the Approved Actuary alone.

Risk of underestimation can be limited by:

- Not giving weight to advice about changes in likely claims development patterns until they begin to be reflected in claims experience, or can be verified as likely in some objective way.
- Ensuring that any qualitative advice is supported by verifiable examples before it is incorporated in estimates. For instance, if a change in policy conditions is cited as a reason why claims experience will improve, the actuary can request examples of renewed policies where the change has occurred, and examples of claims that have occurred in the past that would be impacted by the change. This is an important component of the general principle that actuaries should verify the data that inputs their Insurance Liability assessments.

2.4 Data Verification

Discussion about data verification falls outside the general theme and scope of this paper. However, because it is crucially important to the reliability of actuarial estimates, it is a related topic, and it is pertinent to make some observations regarding practice in this component of actuarial investigation.

Veracity of data has a significant impact on the reliability of the actuary's estimates. The lower the level of data scrutiny, the greater the possibility that unidentified data issues will slip through and undermine the actuarial review.

Codification of the actuary's responsibilities regarding data verification can be found in PS300 (covering all actuarial estimates of outstanding claims and premium liabilities for any entities involved in general insurance activities), GN353 (mandatory for valuations under the APRA prudential standard GPS210), and APRA's guidance note GGN210.1 (indicating APRA's intent with regard to requirements associated with the written advice that the Board of an Insurer must receive from the Approved Actuary on the valuation of its Insurance Liabilities in accordance with GPS 210).

Each contains statements about the responsibility an actuary has to verify their data. Excerpts from each are set out below:

> *PS 300 (Paragraph 22)*
>
> *It is the actuary's responsibility to ensure that the data utilised are appropriate and sufficient for the valuation. The actuary **should**, where possible**, take reasonable steps** to verify the overall consistency of the valuation data with the insurer's financial records.*
>
> *Guidance Note 353 (Paragraph 8)*
>
> *The actuary **should take reasonable steps** to verify the consistency, completeness and reliability of the data collated, against the company's financial records. The actuary **should** discuss the completeness, accuracy, and reliability of the data with the company's auditor (refer to GN 551 'Actuaries and Auditors'). The actuary **should** include in the written report on the valuation of the liabilities a description of the measures taken to investigate the validity of the data, and should outline the results of those data checks.*
>
> *GGN 210.1[1] (Paragraph 6)*
>
> *It is the role of the Approved Actuary to make it clear that he or she may require access to and information from, management, underwriters, other employees of the company and the company's auditors. Approved Actuaries **must** however**, take full responsibility** for their advice and reports and must therefore be satisfied as to the validity of the information provided to them or work undertaken for them.*

APRA's guidance is clear and unequivocal about where responsibility for data validation lies. By contrast, the Institute's standards and guidance seem to be open to a wider range of interpretation, and uses less emphatic language. In practice, considerable differences are encountered in the emphasis placed by actuaries on data validation.

The Actuary is best placed to investigate data veracity (or more precisely – direct and assess it), knowing how the data will be used, in the valuation, and the implications that would follow if the data were incorrect. . If the actuary does not ensure the necessary data checks are performed, properly interpret and explain the implications of the results of the checks, and explain the implications of any restrictions on data verification, it is most unlikely that anyone else will.

Notwithstanding the importance of data verification, and the requirements of the relevant standards and guidance notes, it is not uncommon to see actuarial Insurance Liability reports routinely including paragraphs along the following lines, with no accompanying contextual explanation:

1. *We have relied on information provided by Insurance Company XYZ. While independent verification was not undertaken, we have checked the information for reasonableness and consistency. Reliance was placed on, but not limited to the accuracy of this information. The reader of this report is relying on Insurance Company XYZ, not <Approved Actuary's name, or the firm they work for> for the accuracy and reliability of the data.*

or

2. *In developing this Report Consulting Firm X has relied upon historical data and other quantitative and qualitative information supplied by Insurance Company XYZ without audit or independent verification. We have however reviewed this information for reasonableness and consistency. The accuracy of our results is dependent on the accuracy and completeness of these underlying data. Therefore any material discrepancies discovered in this data by Insurance Company XYZ should be reported to us and the report amended accordingly.*

---

[1]APRA's intention stated in the introduction to GGN210.1 was that it would be replaced by the standards of the IAAust as they were developed, but at the time of writing this paper GGN 210.1 is still found on APRA's website, and it has not been repealed.

The actuary's responsibility with respect to data veracity should be considered in light of how it contributes to the objectives and requirements of GPS 210. Describing these, GPS 210 states:

> *"The Board of an insurer that is required to have an Approved Actuary must obtain written advice from the Approved Actuary on the valuation of its insurance liabilities. The requirement is designed to aid Boards to perform their duties by ensuring they are adequately informed."*

It is difficult to accept that the objectives of GPS 210 are met if the Board is presented with a report containing a set of disclaimers, unaccompanied by a proper explanation, stating that data has not been verified as accurate and complete, and that inaccurate or incomplete data could render the Insurance Liability assessment unreliable.

There will unavoidably be some reliance on the Insurance Company for data veracity for non-financial fields (such as date claim finalised) since a census check of the field is never feasible. It should nonetheless be possible to provide a good explanation of the checks that have been done to establish enough confidence in the field to support its use, the results of those checks, and what has not been checked (and why). My interpretation of GN353 is that, for GPS 210 assessments, it is mandatory that data verification is the subject of a significant degree of effort that is documented in the report. However, iIn practice there seems to be a wide range of interpretation of the degree of rigour required of both the data checks and their documentation.

An important component of data verification is to tie the financial aspects of the data back to the General Ledger. It may not necessarily be the Approved Actuary who performs this check, but it is their responsibility to see that it is done, and to review the results of that check. This is because, regardless of the requirements of the professional standards and guidance notes, it is the Insurance Liability assessment for which the actuary is responsible, that risks being affected by unidentified data problems.

It should take little effort to include a table of numbers in the valuation report, disclosing the result of the reconciliation of valuation data to the General Ledger. This is more complete disclosure of the reconciliation result than a generic materiality statement.

A common problem for the actuary is that, in order to adhere to profit reporting timeframes, data will be extracted with an effective date prior to the effective date at which ledger figures are audited. In such instances, it should still be possible to check the actuarial data against a ledger running total.

It should also be possible to re-extract the actuarial data at the effective audit date, and in a form that allows the original extract to be recreated. Most actuaries would then:

- Check that the original actuarial data is consistent with the re-extract taken at the effective audit date.
- For financial fields, investigate whether the actuarial data is consistent with the corresponding audited ledger fields. This includes gross fields, and those relating to reinsurance and other recoveries.

- Investigate what manual adjustments will be made to the ledger to arrive at figures for the financial statements, and determine whether these should impact the actuarial valuation. Examples might include:
  - Reversal of known claims processing errors
  - Transactions associated with reinsurance arrangements that cannot be fully automated in the products system; such as aggregate covers, stability clauses, and facultative arrangements
  - Unclosed business estimates

It should not be taken for granted that system calculations will always be correct. Where a system calculation inputs the General Ledger, and also the actuarial data, verifying that they are consistent will not constitute a full check of the field. For example, earned and unearned premium are worthy of independent checks, particularly if there are policies with atypical features such as non-annual policy periods, or premium payment by instalment is possible.

Actuarial analysis of Insurance Liabilities typically relies on a wide range of non-financial fields that are not subject to external audit (but may be subject to internal audit controls). These fields must generally be verified outside of the external audit framework. Common examples include:

- Accident date
- Date of claim report
- Date of claim finalisation
- Claim status
- Portfolio subdivisions such as policy type that are at a finer level of granularity than is identifiable in the General Ledger.
- Payment-type codes
- Claim-type codes (such as material damage or personal injury in a public liability portfolio)

Understanding and checking such fields requires that the actuarial data be at a detailed claim-by-claim level. A census check of all the data will never be feasible, but it ought to be possible to walk through examples of claims with a claims staff member (sometimes a processor can be better than someone in a more supervisory or managerial role) to check your interpretation of the data. It is possible to do this by comparing the actuarial data with information that claims staff can read from claim system screens, but at times it can be worthwhile going the further step of checking interpretation against the physical claims file.

Such checks might reveal features such as:

- Payment-type codes not having the veracity you thought.
- Claims that reopen, having the original finalisation date wiped (this can mean retrospective changes to claim finalisation triangles that will impact valuation methods such as PPCF)
- Activities associated with claims handling that do not get captured on the claims system
- Portfolio segments administered differently to the general body of claims

- The meaning of claim counts (for example a single event giving rise to say, a claim against a building and a contents policy might be recorded under a single claim number, and not be captured as in the claim count for both the buildings and the contents valuation data)
- How errors are corrected
- Changes to claims administration processes
- Treatment of bulk policy and claims items; and the meaning of various fields in them (for instance, the date coded in the field 'accident date' may not carry its normal meaning)

It is important that such checks are repeated at intervals. Ideally this should be at each Insurance Liability review.

It can be worthwhile reviewing the most recent external audit closing report and management letter as part of the process of determining whether there are data veracity or control issues that may impact the Insurance Liability estimates. Similarly, the Approved Actuary should review relevant internal audit and technical audit reports which may have considered transactional processing issues, and the reliability of information contained in the product system.

Where a change to business processes could have the potential to impact the valuation, it is important to check whether any have occurred. Among other things, this can include matters such as how claims with little supporting information are recorded (where case reserves will be notional), when claims are regarded as finalised, and how case reserves are set. Turnover in key staff commonly precipitates such changes.

One specific item that can be worthwhile checking is whether any bordereaux are present in the claims and policy data. These arrangements are often processed under a single claim or policy number, but as there are many underlying claims and policies underlying a bordereau they will typically require special treatment.

Understanding the treatment of AASB1023 accounts where companies might argue that elements of the Standard are open to interpretation can also be important if an affected item impacts the Insurance Liability valuation.

In summary, to meet the aims and objectives of GPS 210, the Approved Actuary needs to conduct sufficient investigation to justify their reliance on each field used in the valuation, and to verify that fields carry the meaning attributed to them. This applies to all fields including financial and non-financial items.

The potential impact of unidentified data problems or misinterpretation does not diminish with increasing portfolio size. Risk associated with unidentified data problems and mis-interpretation is therefore systemic.

2.5 Judgement

Generally, reliable estimates of Insurance Liabilities rely to some degree on stability in past claim trends to project future claims outcomes. This allows patterns observed in well-developed accident years to be used as a basis to estimate the run-of of payments relating to more recent accidents.

Because so many aspects of the environment that affects claim cost development and run-off are subject to change over time, it is the rule rather than the exception that judgement is required to allow for anticipated changes that might follow changes to claims management, underwriting, business mix and the external environment.

Where judgement impacts are material, it is important that the possibility the judgement could turn out to be erroneous is taken into account. It is appropriate to consider what the impact on the risk margin should be, if the central estimate gives heavy weight to judgement that run-off will be more favourable than is evidenced by the history. The risk margin and the central estimate should not be regarded as independent of one another. This is because, the purpose of the risk margin is to add to the central estimate, with the aim of arriving at an amount that provides a 75% likelihood of covering the claims cost that arise. As the interest is in the total quantity, *Central Estimate + Risk Margin,* more so than its components it follows that there should be some interaction between the positions taken with respect to each.

Application of judgement introduces scope for inadvertent bias in the sense that claims process variability will occur centralised around some number other than your central estimate. The risk is unavoidable, and needs to be recognised when considering risk margins. Similarly, if there is consistent application of judgement across a number of classes, implications for correlation between adequacy of the liability estimates across the classes also needs to be considered. These items are discussed through the remainder of the paper.

Sometimes, the point is made that judgement reduces variability in how outcomes will differ from central estimates, compared with how this difference would track if judgement were not applied. I would agree that this should be the case, otherwise it would follow that judgement was detracting from the claims estimation exercise. However, I would also comment that an identified need to exercise judgement is an acknowledgement that there are 'un-modelled' sources of variation. It should therefore follow, that when one's attention turns to risk margins, the risk that claims outcomes are subject to un-modelled sources of variability is taken into account.

3 Stand - Alone Risk Margins

3.1 Order of Determination

The next step taken by most actuaries after central estimate assessment is assessment of the risk margin required to take the overall estimate to a 75% likelihood of adequacy.

I expect statisticians would be bemused by this order of determination, since risk margin assessment involves estimating where a single percentile in the distribution lies. The mean on the other hand is a weighted average of all percentiles. It seems to make sense that the 75^{th} percentile ought to (implicitly or explicitly) be an input to the central estimate determination[2]. Nonetheless, it is currently standard actuarial practice is to assess risk margins based on analysis that is conducted after a central estimate has been determined.

The order of determination could be an indication that, in practice, the central estimate is sometimes set based on what seems to be a reasonable scenario, a scenario representing a median outcome, or possibly even a mode. If so, this would be a problem, since, for general insurance liabilities, likely outcomes and the median would be lower than the mean of the full probabilistic distribution, due to its skewed nature.

3.2 Risk Margin Components – Process Variability and Estimation Error

It can be useful to think of the risk margin as providing protection against two components of variability in outcomes compared to the central estimate.

1. Risk that our central estimate may not be accurate[3] (referred to here as '*estimation error*')
2. Intrinsic variability around the 'true' mean (referred to here as '*process variability*')

At the time of writing, paragraph 22 of GPS 210 states that "*the risk margin is the component of the value of the Insurance Liabilities that relates to inherent uncertainty in the central estimate*." This statement directly acknowledges the first source of variability, but even if we were 100% confident that we had 'correctly' determined the central estimate, a risk margin would be required to protect against the *process variability* around it. In order to arrive at an assessment of the Insurance Liabilities that aims to provide a 75% likelihood of sufficiency, consideration of the both sources of variability is required. I would suggest that this is unfortunate wording in GPS210 rather than a deliberate attempt to exclude *process variability*.

[2] Approximate methods that might avoid this requirement are lost once one recognises that the distribution of possible claims outcomes is skew.
[3] In the sense that process variability acts around some other number (being the 'true' mean) rather than our assessment of the central estimate.

3.2.1 Sources of Estimation Error

In actuarial nirvana, where the environment is stable, the nature of the claims process (including the nature of random elements), is perfectly understood, and data is error free, the two components should be closely related. In this fantasyland, the outstanding claims estimation model will reflect the perfectly understood claims process. However, since selected model parameters will be based on a claims history that is a small sample of observations subject to *process variability*, fitted parameters will most likely differ from the 'true' ones underlying the process. Therefore, as well as being the generator of variability around the true mean, *process variability* also generates *estimation error*.

In practice, *estimation error* can arise from other sources as well. It is my view that these sources are generally more important. The most significant category of other sources is associated with application of judgement in its various forms. Another important source of estimation error can be unidentified data errors.

As a hypothetical example, of *estimation error* that is unrelated to *process variability*, an actuary might judge that the length of the claim reporting tail for a public liability portfolio will be reduced compared to past experience, say due to the impact of various tort law reforms. Because the actual impact may be different from what the actuary anticipates, this risk represents a potential source of *estimation error* that is not associated with *process variability*.

Real and/or assessed changes in portfolio conditions necessitating the application of judgement introduce risk of *estimation error* that is not associated with *process variability*. Important features of these sources of this risk include:

- They are particular to each valuation, since different actuaries will form their own view regarding the likely impact of changing circumstances.
- 'Industry analysis' is unlikely to help quantify them.
- They are not generally captured by common analytic techniques that actuaries use to help evaluate uncertainty and set risk margins
- They are a source of systemic risk, in the sense that their relative impact does not diminish with increasing portfolio size.
- Because judgement is a feature of most valuations, this source of estimation error is also a feature of most valuations.

Sources of *estimation error* that are not associated with *process variability* will generally increase the probabilistic spread of how outcomes might vary around the assessed central point. Hence, for central estimates, there is potential for upside and downside risk to offset. However, considering implications for risk margins, it is only the downside potential of *estimation error* that is important. For the risk margin, there is no offset from upside potential.

Analytic techniques most commonly used to set risk margins, such as the Mack method and bootstrapping, treat *estimation error* as arising solely due to *process variability*. The risk that model parameters underlying central estimates might differ from the 'true' parameters is treated by them as arising because model parameters are determined based on a claims history which is a small sample of observations that are subject to *process variability*. Therefore, when one is using such analytic techniques, it is important to bear in mind that they will not capture all sources of variability that need to be taken into account in risk margin assessment. One can argue that a risk margin derived using them should be considered a minimum (or possibly – below the minimum).

The importance of the sources of error not related to *process variability* depends on:

- Whether or not a portfolio's circumstances have changed so that history forms a less helpful basis from which to assess Insurance Liabilities.
- Whether or not a portfolio's circumstances ought to be judged to have changed so that history forms a less helpful basis from which to assess Insurance Liabilities.
- The judgement position taken.
- How closely the actuarial models mirror the true underlying claims process (which will always be unknown).
- The rigour accompanying data verification.

3.2.2 Systemic and Independent Components of Variability in how Outcomes Compare to Actuarial Estimates

The variability of ultimate interest in Insurance Liability assessment is variability in the extent to which actual outcomes depart from the actuarial estimate of the liability. When examining the sources of this variability, it can be useful to try to categorise them as systemic or independent to help understand the extent to which risk margins should decrease with increasing portfolio size.

Systemic variability	Can be defined as variability that does not diminish as a proportion of the liability as portfolio size increases.
Independent variability	Can be defined as variability that diminishes as a proportion of the liability as portfolio size increases.

Independent sources of variability relate to intrinsic variability of the claims experience of individual risks, that do not carry implications for likely claims costs for other risks in the portfolio. The chance occurrence or non-occurrence of 'non-catastrophe' events that give rise to claims is an example of a source of independent variability.

Systemic sources of variability can be categorised into two components:

1. *Drivers of claims cost that may affect multiple policies simultaneously.*

   For example, in the case of premium liability assessment, for a home portfolio without much geographical diversification, risk of storm damage could be considered systemic. Inflation is another commonly cited example that would affect cost across many claims in long tail portfolios.

2. *Variation in the adequacy of outstanding claims estimates that results from the Insurance Liability estimation process*.

   For example, a judgement that prior poor performance will turn around, say due to improvements in claims handling might prove to be ill-founded. The risk that this could happen generates a risk that outstanding claims estimates might prove inadequate. This risk doesn't diminish with the size of the portfolio, so is systemic. However, it is not associated with intrinsic variability in claims experience. The risk arises due to the valuation process, and the chance that the actuary's judgement might be wrong.

   Sources of the second component of systemic risk include:

   - Adoption of an inappropriate (ill-fitting) projection model.
   - Optimism that recent poor performance is not indicative of likely future performance. (where this proves to be ill founded)
   - Judgement overlays in parameters selection.
   - Data that is incomplete, inaccurate, or insufficient.
   - Not recognising items such as changes in:
     - Policy conditions
     - Exposure
     - Legislation
     - Judicial interpretation
     - Case estimation practice
     - Claim management practice
     - Business Mix
   - Not recognising implications of the skewed nature of Insurance Liabilities on appropriate parameter selection.

Compared to a small insurer, the greater wealth of data on which a large insurer's Insurance Liability estimates are based should mean that the component of *estimation error* arising due to *process variability* is relatively small. However, the other sources of estimation error are much less likely to diminish with increasing portfolio size. Risk arising from e*stimation error* that is not associated with *process variability* is systemic in nature.

Because data comparing claims outcomes with previous projections isn't publicly available, it isn't possible to determine empirically how quickly systemic sources of variability come to dominate as portfolio size increases. However, it will be faster than if we were only concerned with the first of the two systemic variability components.

3.2.3 PS300 Categorisation of Sources Of Variability

Another way of categorising sources of uncertainty influencing risk margin assessment is along the lines of the breakdown provided in PS300. The Professional Standard makes specific reference to the following categories. With the exception of a specific reference to process error, each can be regarded as relating both to *process variability* and *estimation error*.

i. **Model Selection Error** - Because the models are a simplification of the claims process, it might be that none of the various models used is an entirely accurate representation of reality.

ii. **Parameter Error** - Because there are components of randomness in the claims process, it is not possible to estimate the parameters of that process with complete precision even if complete confidence were felt in the nature of the model.

iii. **Parameter Evolution Error** – This is described as deriving from the inclusion in a model as constants any parameters that are in fact subject to change over time. The change may be smooth over time, or may come in the form of sudden 'shocks.'

iv. **Process Error** - Even if the parameters could be estimated with precision, it would not be possible to predict outstanding claims with the same precision because of the random component in future experience.

Other important sources of uncertainty include:

v. **Input Data Error** - Any erroneous data will similarly have introduced uncertainties into the estimate of those parameters

vi. **Judgement Error** – At each Insurance Liability assessment, the actuary must form a view about the possibility that systemic (ie non-random) changes may occur in claims experience. Because actuarial judgement is not perfect, there is always a risk associated with the possibility that judgement calls might be wrong.

Each source requires consideration when assessing risk margins, and to avoid risk margin understatement, it is important to ensure that none are left unaddressed.

It can be worthwhile reviewing the methods that have been used to assess the central estimate and the risk margin under each of these categories to reduce the chance that sources of variability have been missed. It is also possible that valuation processes might increase some sources of uncertainty.

For example, it is not uncommon for Insurance Liability estimates to be based on data extracted from claims and policy administration systems earlier than the effective assessment date in order to meet profit reporting and other deadlines. Typically, this requires grossing-up of the data inputting the valuation to reflect what the actuary expects claims experience to the effective assessment date to look like.

Compared to the case where data is extracted at the effective balance date, grossing-up of data extracted earlier introduces an additional source of uncertainty. In terms of the previous list, it is a mix of process error, and input data error.

Possibly more importantly, it interferes with the actuary's ability to go through the normal control cycle process of comparing experience with previous projections. Hence, where grossed up data is used for analysis, the actuary should assess the degree to which uncertainty is increased, and take this into account in the risk margin assessment.

3.3 Quantitative Methods Used to Assess Risk Margins

3.3.1 Brief description of three methods

Actuarial literature describes a range of quantitative assessment techniques that can be used to estimate predictive distributions for claims outcomes. There is increasing interest in this active field of research. In this paper, I have focussed on three that I see used most often in practice. The three techniques are:

- The Mack Method
- Bootstrapping
- Modelling based on Zehnwirth's Probabilistic Trend Family ('PTF') as implemented in the statistical claims reserving tool ICRFS -Plus™ ("ICRFS")

Each technique carries (explicit or implicit) assumptions that need to be met before meaning can be attached to its results. For these methods to produce meaningful results and for them to be relied upon, it is important that the assumptions on which they are based are tested and verified. However, in my audit support work, I often come across situations where these methods are relied upon even though their underlying assumptions are not met.

As a minimum, it would seem necessary to check that a well fitting model, having the form assumed by the method can be found. This well-fitting model needs to give a similar result to the adopted central estimate. If such a model can't be found it seems illogical to expect sound conclusions to be drawn from the quantitative risk margin assessment method.

If judgement (as opposed to simple model fitting) is applied as a significant input to determine the central estimate, it is less likely that such a model will be found. In such an instance, just as judgement forms the basis of the central estimate, so to it will need to form the basis of the risk margin assessment based on estimates of process variability, and the risk that the judgement could be wrong.

A brief description of the three methods and some comments on their underlying assumptions follow. The interested reader is referred to the references in the bibliography for more comprehensive descriptions.

Mack

Mack (1994) has developed a formula describing the standard error of prediction of outstanding claims estimates, where those estimates have been determined based on a particular mechanical application of the chain ladder.

The particular mechanical application is where the development factor F_t that, when multiplied by the cumulative claims cost to development period t, gives the expected value for cumulative claims cost to development period t+1 is determined as the ratio of:

(#) - The sum across accident years of cumulative claims cost to development period t+1, to

- The corresponding sum of cumulative claims cost to development period t.

This is a standard application of the chain-ladder, but is different to an application that selects assumptions based on experience over the last few transaction years rather than the full history triangle.

Superficially, the chain ladder model appears to be very simple. However, in his 1994 paper Mack highlights that its use effectively makes a number of significant assumptions about the claim cost development process.

His formula is derived based on the premise that these assumptions are all met. In his paper, he warns that the assumptions are quite strong, and won't be met by all claims triangles. Helpfully, his paper includes a series of semi-formal tests which one should verify as met before applying his formula. Examples of these tests are set out in Appendix A. In practice, it is important that tests are conducted, and are found to provide a satisfactory result before this method is used.

He argues that an implicit assumption made when applying the chain ladder, is that the estimator of the 'true' underlying development factor is unbiased. Other ways of selecting development factors (for instance, the simple average of the individual development factors for each accident year) would give rise to the same expected value for the estimator. To answer the question: "Why is the particular estimator favoured?", he recalls that amongst several unbiased estimators, preference should be given to the one with the smallest variance. He concludes that we must therefore be assuming that the estimators (#) described previously are the unbiased estimators with least variance.

Based on implications that follow from this conclusion, and other assumptions that he shows are implied by use of the chain ladder, he derives a formula for the standard error of prediction.

Further commentary on the Mack method is set out in Appendix B.

The method assumes that a well fitting[4] chain ladder method describes the central estimate. One implication that follows from this is that where it would not be reasonable to set the central estimate using the chain ladder approach, it would not be reasonable to rely upon the Mack method to help determine the standard error of prediction and hence risk margins. Generally, actuaries would not regard the chain ladder as a suitable approach to set the central estimate for long tail insurance classes.

[4] Well fitting in the sense that the claims history meets the specified tests.

For practical application, the mean and standard error of prediction need to be supplemented with a distributional assumption (this contrasts with the following two methods that estimate the full predictive distribution directly). Mack suggests lognormal, but there is no compelling reason why other forms, such as gamma or log gamma ought not be regarded as equally appropriate candidates.

Due to its relatively thick tail, the 75th percentile of a lognormal distribution commonly sits a little below the 75th percentile of alternative right skewed distributions with the same mean and variance.

Bootstrapping

Bootstrapping derives estimates of the probability distribution of a statistic by repeated re-calculation of it, using random samples drawn with replacement from the original data. Because the sampling is with replacement, at each repetition some items in the data set are selected two or more times, and others are not selected at all. When this is repeated a large number of times, we get '*pseudo-calculations'* of the statistic that, provided certain conditions are met, should have a similar distributional form to its underlying probability distribution.

In general insurance claims reserving, bootstrapping is typically performed to derive results by re-sampling residuals rather than the data itself. (This is because the data can't be reasonably assumed to be identically distributed across accident and development years, but residuals <or some suitably defined function of them> can be). This application of bootstrapping effectively assumes that the underlying process is of known form, and that repeated re-sampling of the observed residuals around this process captures information about variability characteristics. [5]

The interested reader is directed to Efron and Tibshirani (1993) for more detail regarding bootstrapping procedures.

Bootstrapping is often presented as a straightforward and easy to apply method for assessing distributional characteristics of a process based on sample data. However, it is a tool that requires care to ensure reasonable results. For instance, sometimes it is necessary to adjust for bias in parameter estimates, and to ensure that the residual definition is appropriate. Appropriate residuals are not always observed values less fitted values (see England & Verrall (1999)).

One must recognise that bootstrapping fails when its assumptions are violated. For example if:

- The residuals are not independent
- The residuals (or function of them) are not identically distributed
- The assumed underlying process is different from the actual process

These violations arise as a result of the chosen reserving model not fitting the data well. In practice, the run-off process is complex, and a well fitting model may be elusive.

The bootstrapping procedure used for the illustrations in this paper is described in Appendix C. It is one of the methods presented by Bonnard et al (1998). I have illustrated bootstrapping with this approach, since it is one I have come across in practice. Depending on the circumstances, one might argue that alternative structures might have greater technical merit.

[5] An important underlying assumption is that we have a representative sample of all the possible distinct population values present in our data (Rubin, 1981)

Zehnwirth's Probabilistic Trend Family ('PTF') as implemented in ICRFS-Plus

This assessment method is based on the assumption that incremental payments have heteroscedastic lognormal distributions, following a piecewise loglinear structure[6] in the accident, development and calendar years.

The ICRFS interface incorporates a number of visual and statistical tests of fit, and statistics that, when properly utilised, protect against inclusion of non-significant parameters in the selected model. For the reader who is unfamiliar with the ICRFS interface, Appendix D provides an example that illustrates the model fitting process. For more details about the underlying model, the reader is referred to Barnett & Zehnwirth (2000).

In practice, ICRFS is an interactive modelling tool. Judgement and knowledge of the underlying business play a key role in the model fitting process. However, the models presented in this paper have been fitted using an automatic model fitting functionality that is built into the program.

The starting model was fully parameterised in the accident, development and calendar year directions.

The model structure was determined by backward elimination of insignificant parameters, with significance determined by T-ratios, using the in-built ICRFS model optimisation method. First, insignificant changes in trends and levels were removed. Insignificant trends were then set to zero, and, finally, smoothing was done on accident year parameters.

This process is open to a number of sound qualitative and technical criticisms, and this automatic process should not be used to fit models in practice. However, for the purpose of this paper, the approach has the advantage of allowing the interested reader to reproduce the results. Examples can also be found where this process gives unreasonable results, and drawing attention to these may be of benefit to novice users of ICRFS. My aim is not to pass any opinion on the ICRFS reserving package or PTF models, and I would stress that no reader should expect to draw general conclusions about the reliability of the models fitted using the ICRFS package based on the presentations in this paper.

The ICRFS reserving tool includes functionality it terms 'PALD' that can be used to assess the predictive distribution of claims outcomes consistent with the selected model. PALD samples from the given model, which defines a joint distribution of correlated lower triangle cells, each lognormally distributed with its own mean, standard deviation and correlation with other cells. Each individual sample is a complete lower triangle, which in turn yields an outstanding claims liability total. Repeated sampling of values of outstanding claims totals forms the distribution of interest.

ICRFS allows selective inclusion of parameter uncertainty in forecasting and estimation. Parameter uncertainty was included in this exercise because it is the full predictive distribution that is of practical interest in setting reserves.

[6] Piecewise loglinear structure means that the triangle of logs of incremental payments is described by piecewise linear trends in accident and calendar years, and smoothed levels across the accident years. The log of the incremental payments are normally distributed, with means determined by the sum of accident, calendar and development levels for each cell.

There are limitations to the extent that the PALD output, as I have described it, should be expected to generate a reasonable picture of the predictive distribution of claims outcomes. The more significant limitations are:

- No direct account is taken of the possibility that future changes could occur to the parameters of the piecewise loglinear structure. Particularly when a number of past changes are incorporated into the fitted model, the assessed variability of claims cost outcomes risks being understated as a result of this.
- No direct account is taken of changes in rates of claim finalisation. Without adjustment, if payment levels are increasing because finalisation rates have increased, under a PTF/PALD assessment, this can contribute toward overstatement of the amount required to secure liabilities with a given likelihood of adequacy. The reverse is true if payment levels have fallen because finalisation rates have been reduced. Any adjustment would be subjective, and would carry risks associated with subjective adjustment.
- For property classes, if the claims history includes spikes relating to events (such as hailstorms), without a subjective adjustment, the fitted PTF model and PALD estimates will, at least in part, treat the spike as indicating variability around a fitted model. This is fine in the case of premium liability estimates, but for outstanding claims, the actuary should know at the time of the valuation, whether or not such an event has recently occurred and needs to be reserved for. This can contribute overstatement of the likely variability in outstanding claims outcomes.

The items in the list are also limitations associated with the other two methods.

3.4 Testing performance of Analytical Methods against known (artificial) Processes

3.4.1 Introduction

In this section, the three methods are examined in two artificial test settings. In each setting, the claims process follows a series of predetermined rules that incorporate a random element.

The first test setting involves a claims generating process that satisfies many of the assumptions associated with chain-ladder assessments. The second involves a claims generating process that does not satisfy these assumptions.

The aim of this exercise is to explore how well each method performs under the two claims generating processes. A priori, one would expect that where the assumptions made by an analysis technique are met, it should perform reasonably well. Where they are not met, one would expect less satisfying results.

For each process, a series of development triangles of incremental claim payments is generated. The three assessment techniques are then applied to estimate the probabilistic distribution of claims outcomes.

In practice, we never have perfect knowledge of the claims generating process, so we can never test whether our estimate of claims outcome distributions follows the 'true' distribution. However, in the test settings, claims history has been simulated from a predetermined process. I have modelled this process as continuing into the future. Therefore, the 'true' probabilistic distribution of claims outcomes is known. The distribution assessed under the three methods can therefore be compared with the 'true' distribution.

My aim is to explore the general effect of model specification error. By this I mean exploring the effect of using methods to assess the full probabilistic distribution of claims outcomes and stand-alone risk margins when the assumptions that underlie the assessment method are not satisfied. It is not possible to test every combination of claims process, and risk margin assessment technique, and I would not expect this exploration to allow general conclusions to be drawn about whether any particular technique is superior to others.

3.4.2 Important Cautionary Note on Interpretation of the Results

For each of the 'true' distributions, the probabilistic distribution presented represents the '*process variability*' associated with the mechanism that I have set to generate claims outcomes, spread around the true mean. The process underlying claims cost generation is stable, and finishes at a known and fixed development period.

Where I have presented risk margin required to get from true mean to the 75th percentile of the 'true' distribution, it is a margin required to cover *process variability*. For the illustrations and statistics labelled 'true', it is known with certainty that there is no possibility of *estimation error*.

For the presentations of the quantitative assessment methods, there is a possibility (which can be seen in the presentations) that the assessed mean could be different from its true value. Therefore the risk margin assessed by each of the quantitative methods needs to cover both *process variability* and *estimation error*. Hence, the assessed risk margins ought to be higher than the 'true' margin required to cover *process variability* only.

It is also worth noting that in the illustrations of the quantitative methods I have applied them mechanically. In practice, actuaries using the illustrated methods are likely to consider whether:

- Outliers exist or are under-represented
- Judgemental overlays should be made to calculated results.
- In the future the claims generating process might depart from the past process.
- Historically there have been shifts in the claims generating process.

Additionally, in the modelling environment there is no pressure from anyone with a vested interest in the result, and the modeller does not feel constrained by any concern about deviating from the previous valuation assumptions too much.

In summary, there **are** **no** sources of '*estimation error'* that a risk margin derived from the analyses needs to take into account, other than that those deriving from *process variability* and inappropriate specification of the form of the model.

In practice however, a risk margin must provide protection against all sources of variation in outcomes from the central estimate.

3.4.3 The Test Processes

The full detail of the claims generating process that I have set for this test is set out in Appendix E. The main criteria that I have tried to meet in setting the process are that for:

Process 1 - Chain Ladder assumptions are generally satisfied.

Process 2 - Chain Ladder assumptions are generally not satisfied.

Process 1

For Process 1, claims are modelled so that for a given accident year, the mean value of the cumulative claims cost at the end development period k, is a multiple, F_{k-1}, of the observed cumulative claims cost to the end of development period k-1. Incremental claims cost for a development period includes a random element that follows a lognormal distribution; the precise nature of which depends on the expected value for the incremental claims cost.

For the test process, the development factors F_k are:

Process 1 – Chain Ladder Factors									
Development Period (j)	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
$\mathbf{F_{i,j}}$	3.000	1.650	1.300	1.200	1.080	1.060	1.040	1.020	1.005

To start the process off for each accident year, the development Period 0 claims cost is modelled as an observation of a lognormal random variable with mean 20,000, and standard deviation 10,000.

For later development periods, the random component has standard deviation equal to 60% of the expected incremental value. The lognormal distribution that the random element is modelled to follow is translated. Before translation, its expected value is 80% of the expected incremental claims cost. The translation shifts the expected value for the random element to be zero.

It is difficult to imagine what underlying features of a set of risks would lead to this sort of claims process being followed in practice. Nonetheless, the chain-ladder effectively assumes that this is the sort of process that claim costs follow.

Process 2

For Process 2, claims cost is generated as a random selection from a series of lognormal distributions. The distributions are a function of development year, but are independent of accident year. For instance, claims in development period 0 are modelled as lognormal random variables with mean 12,000 and standard deviation 3,600; claims in development period 1 are modelled as lognormal random variables with mean 24,000 and standard deviation 8,400. Details for the other development periods are shown in Appendix E.

True probabilistic distribution of Claims Outcomes

By simulation, a claims history can be repeatedly generated under each process. The 'true' probabilistic distribution of possible claims outcomes, associated with the outstanding claims liability for any history triangle is difficult to determine analytically, but can be quite easily determined by repeated simulation.

The nature of Process 1 is that this 'true' distribution is dependent on what is observed for the history triangle (more precisely; it depends on the observation for the last diagonal). By contrast, the 'true' distribution for Process 2 is not contingent at all on the particular observation for the claims history.

This difference has an implication for the way results can be presented.

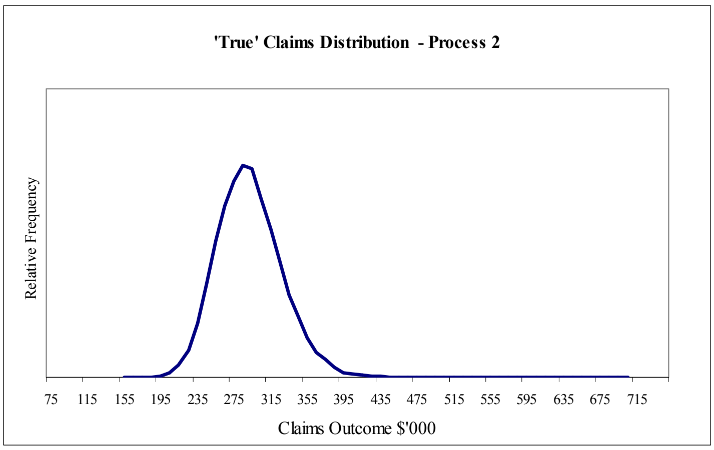
For Process 1

The 'true' distribution for outstanding claims costs is different for each history triangle. So each time a history triangle is re-generated by simulation, there will be a different assessment of the distribution by the Mack method, bootstrapping and the PTF model. There will also be a different 'true' distribution.

For Process 2

The 'true' distribution for outstanding claims costs is not dependent on what is observed for each history triangle. So each time a history triangle is re-generated by simulation, although there will be a different assessment of the distribution by the Mack method, bootstrapping and the PTF model; the 'true' distribution of outstanding claims liability will not change.

For Process 2, the total outstanding claims liability is the sum of a large number of lognormal distributions. Its distribution is straightforward to determine by simulation. The 'true' distribution of possible claim outcomes was determined based on 25,000 simulations using the Excel add-in package @Risk.© It is graphed below.



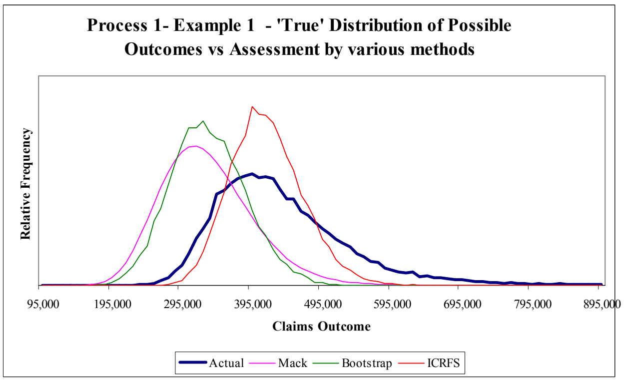
The distribution has:

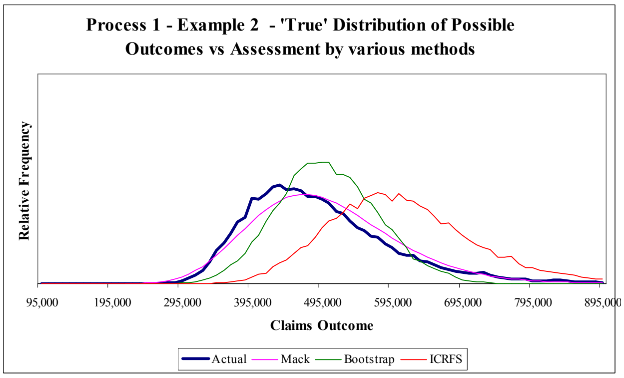
- Mean $292,000
- Standard deviation $35,100
- Coefficient of variation 12.00%
- 75th percentile $314,000
- ‘True’ required risk margin as a percentage of the ‘true’ mean to give rise to a 75% likelihood of adequacy of 7.45%

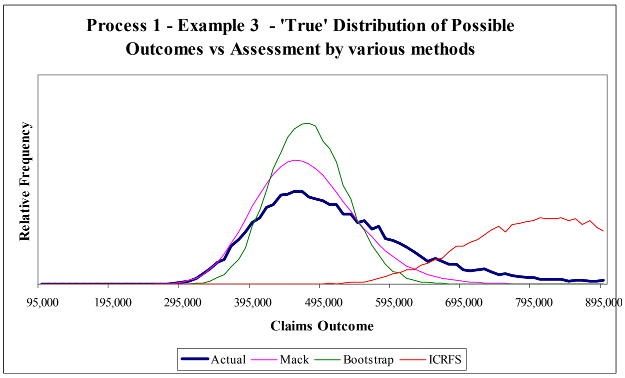
3.4.4 Process 1 Results

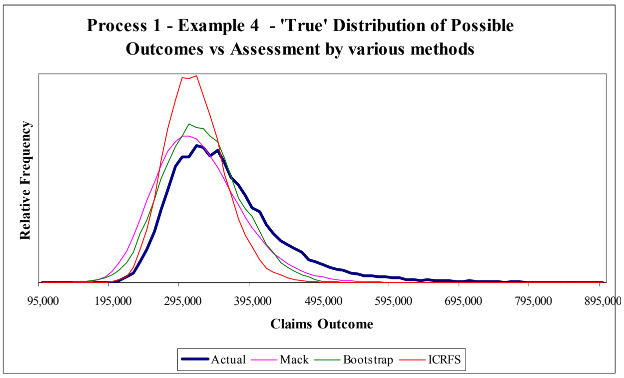
Process 1 follows chain ladder assumptions, so ‘a priori’ one would expect that methods used to assess the full distribution of possible claims outcomes, that rely on chain-ladder assumptions being satisfied are likely to perform reasonably well.

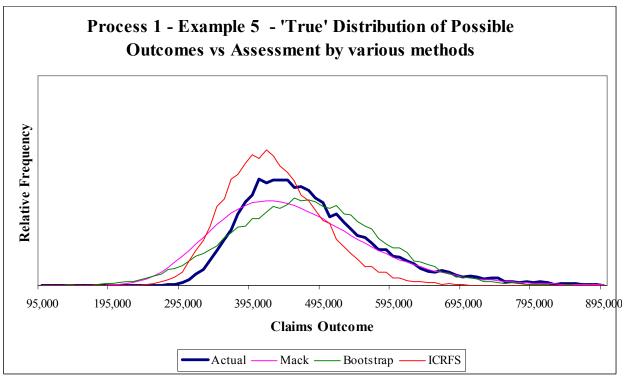
The results of simulations from six randomly generated history triangles are set out graphically below. The history triangles associated with each simulation (and six others) are provided in Appendix E. Appendix F shows graphs for all history triangles.

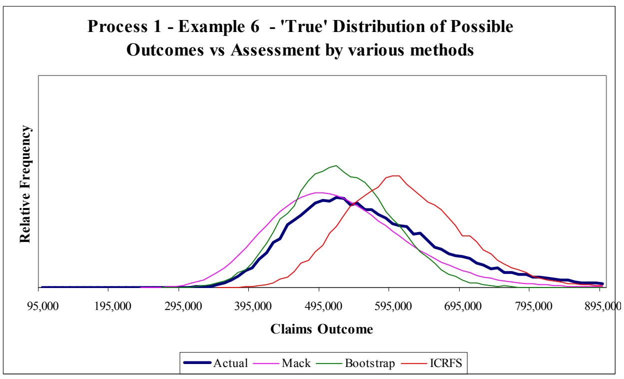












Key statistics from the twelve tests of Process 1 are set out in the next two tables.

Results Summary – Process 1 $'000								
Example	Mean				75th Percentile			
Number	True	Bootstrap	Mack	ICRFS	True	Bootstrap	Mack	ICRFS
1	439	331	332	416	485	370	368	449
2	495	500	503	609	548	562	551	672
3	519	474	475	854	576	519	510	939
4	358	319	322	314	395	356	358	341
5	482	456	471	426	531	520	539	469
6	572	518	522	612	632	578	570	666
7	483	363	364	387	536	398	393	419
8	393	419	422	413	431	465	467	452
9	521	356	361	230	582	405	400	264
10	521	506	510	504	574	554	558	590
11	403	464	466	445	450	518	507	499
12	410	401	403	336	459	455	441	377

Process 1 Risk Margin to get from True/Assessed Mean To True/Assessed 75th Percentile				
Example No.	**True**	**Bootstrap**	**Mack**	**ICRFS**
1	10.5%	11.8%	10.8%	7.9%
2	10.7%	12.4%	9.5%	10.3%
3	11.0%	9.5%	7.4%	10.0%
4	10.3%	11.6%	11.2%	8.6%
5	10.2%	14.0%	14.4%	10.1%
6	10.5%	11.6%	9.2%	8.8%
7	11.0%	9.6%	8.0%	8.3%
8	9.7%	11.0%	10.7%	9.4%
9	11.7%	13.8%	10.8%	14.8%
10	10.2%	9.5%	9.4%	17.1%
11	11.7%	11.6%	8.8%	12.1%
12	12.0%	13.5%	9.4%	12.2%

The following observations can be made:

- The Mack and Bootstrap assessment of the mean is almost exactly the same for each simulation (simulation number five being the exception). This is not surprising as each of them is trying to fit a chain ladder to the same set of triangles.
- The Mack and Bootstrap central estimates sit below the true mean more often than they sit above it (9 times out of 12). Investigating this aspect further, with 25,000 simulations, just over 75% of the time, the outstanding claims estimate derived from applying chain ladder factors based on the sample claims triangles gives a lower result than that calculated based on the 'true' underlying chain ladder factors. This is an example of the effect of the phenomenon discussed in Section 2.2, that more often than not, the mean of a sample from a skewed distribution will understate the true mean.

- For this particular example (ignoring the central estimate understatement), in percentage terms, bootstrap assessment of the required risk margin was generally above the 'true' requirement needed to cover *process variability*, and the Mack assessment was generally below it. This means that the Mack assessment did not come up with a high enough margin even to cover process variability. In practice, we would want a risk margin to protect against *process variability* and *estimation error*.

- Though in percentage terms, the bootstrap assessment of the required risk margin sits above the 'true' requirement, it is not enough to make up for the central estimate understatement. In absolute terms, the assessed 75th percentile was below the 'true' 75th percentile in 9 of the 12 examples.

- Visually, the bootstrapping procedure and the Mack method seem to do a reasonable job of capturing the general shape of the 'true' distribution. Notwithstanding the previous comments, the assessed distribution is generally in 'roughly' the right location.

- For examples 1, 7 and 9, the tables and the charts in the appendices the Mack and bootstrap assessments of the distribution are particularly distant from the 'true' distribution. However, when tests as described in Appendix A are performed on these triangles, the assumption that the triangles are following the chain-ladder assumptions would be rejected.[7] This illustrates the importance of checking fit before drawing conclusions.

- The ICRFS assessment of the outstanding claims cost distribution is often quite different to Mack and Bootstrapping. This is not surprising as the underlying form of the model is quite different.

- While the ICRFS assessment of the 75th percentile point often lies some distance from the 'true' location; for the twelve example triangles, it has the merit of being above the true point 50% of the time and below it 50% of the time.

- The PTF result in example 3 is worthy of comment. It is a long way from the true distribution. This seems to have arisen as a result of the backward elimination process which spuriously found:
  - a sustained increasing accident year trend
  - compensated for by a calendar year trend, which
  - ceased and became zero in the last calendar year.

  It is not difficult to find a model fitted using ICRFS in a less mechanical way that has a distribution much closer to the true distribution. Notwithstanding that its statistical fit is not unreasonable, the features of the model derived by backwards elimination would most likely be rejected on qualitative grounds. This serves as an example of the danger associated with trying to fit the PTF models with ICRFS in the simplified way that has been done for the purpose of contrasting with the Mack method and bootstrapping in this paper. In practice, a qualitative assessment is required to check that model makes sense, even if it provides a good theoretical fit.

[7] The check is whether chain ladder assumptions about the assessed development factors are being followed. Even though the generating process is a chain ladder, it is about different development factors than would be estimated based on a sample history.

3.4.5 Process 2 Results

Process 2 does not follow chain ladder assumptions, and so 'a priori' one would expect that methods used to assess the full distribution of possible claims outcomes that rely on chain-ladder assumptions being satisfied, are likely to perform less well under this process.

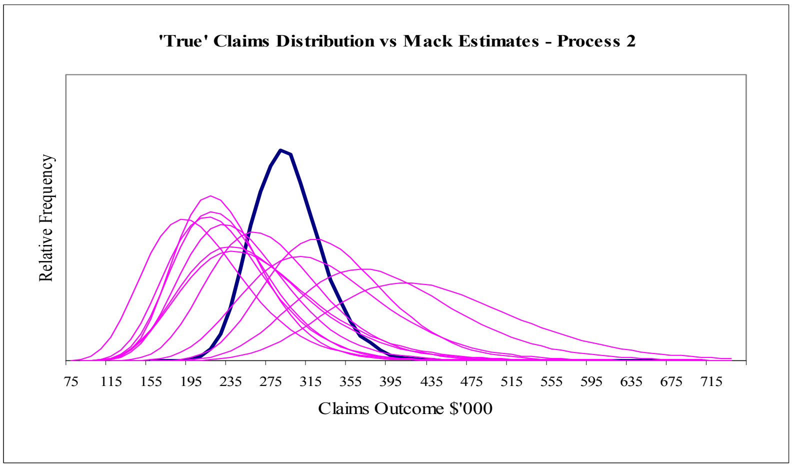
Although the process doesn't follow chain ladder assumptions (and tests suggested in Appendix A would generally clearly show this to be the case for the claims history triangles), triangles of claim development factors calculated from history triangles generated by this process often don't '*look*' too bad. However (with one exception) each triangle would fail the tests described in Appendix A.

In practice, I have seen chain ladder results given weight with claims development triangles that appear to follow chain ladder assumptions less closely.

Mack

Each example history triangle has an associated outstanding claims cost distribution estimate assessed according to the Mack method.

Twelve claims history triangles generated from Process 2 are set out in Appendix E. Mack method calculations were performed for each triangle. The resultant assessment of the distribution of possible claims outcomes for each of the twelve triangles is set out in the chart below, and compared with the 'true' distribution.



Points I note are:

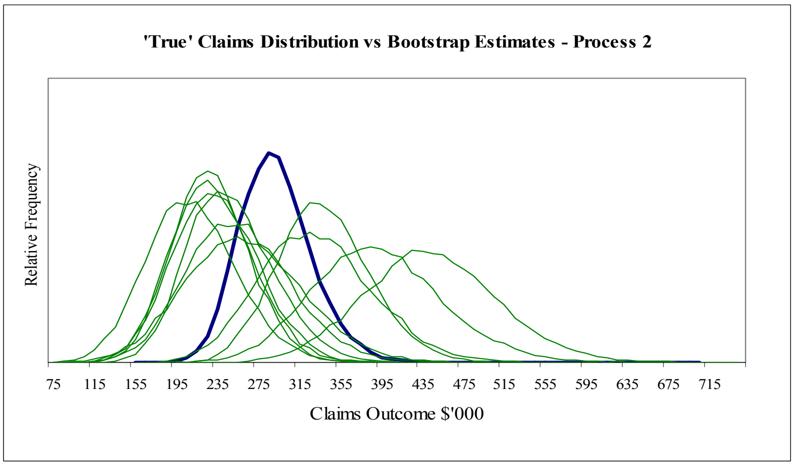
- The assessed distribution can be very different depending on the history triangle that arose from the process.
- Based on the twelve examples, the distribution of the distributions appears right-skewed.
- The higher the 'chain-ladder' estimate of the mean, the greater the Mack estimate of the variability around the assessed mean.
- Unsurprisingly, the spread of each distribution is wider than the 'true' spread. The reader is re-directed to the earlier cautionary note on interpretation of the results. The 'true' spread relates to process variability only. However, in practice, uncertainty in outcomes arises from process variability and estimation error.

It is worth considering whether triangles that give rise to the curves furthest to the right contain observations that, in practice, would be given reduced weight as 'outliers.' Removing, (or giving less than full weight to) 'large' observations would narrow the Mack curves and shift them to the left.

Bootstrapping

The same example history triangles, has an associated outstanding claims cost distribution estimate assessed according to the Bootstrap method as described in Appendix C.

The resultant estimates of the distribution of possible claims outcomes for each of the twelve triangles is set out in the chart below, and compared with the 'true' distribution.

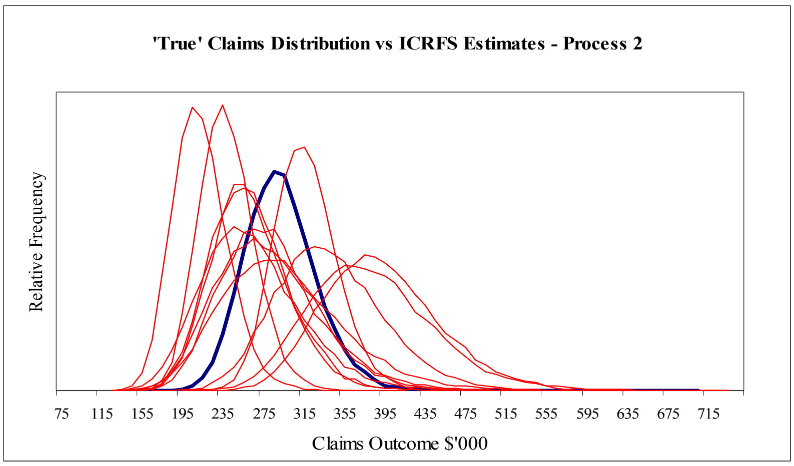


General observations associated with the bootstrapping results are not qualitatively different to the observations for the Mack assessment. The assessed distributions are typically a little less wide than under Mack. More comments on the results are set out after the tables on page 33.

PTF as assessed by ICRFS

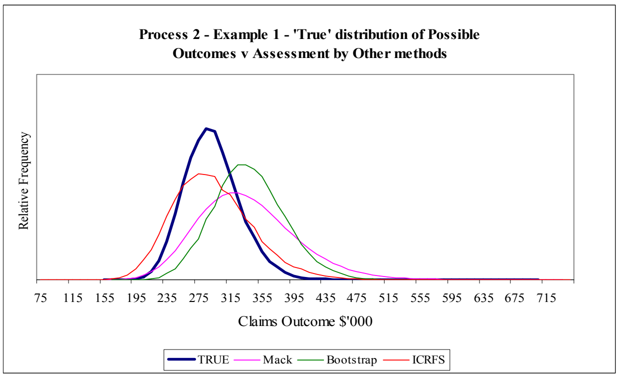
A PTF model was fit to the same history triangles using the reserving package ICRFS.

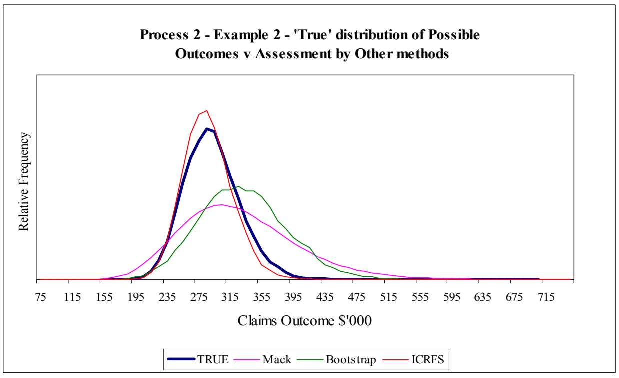
The assessment of the distribution of possible claims outcomes for each of the twelve triangles is set out in the chart below, and again compared with the 'true' distribution.

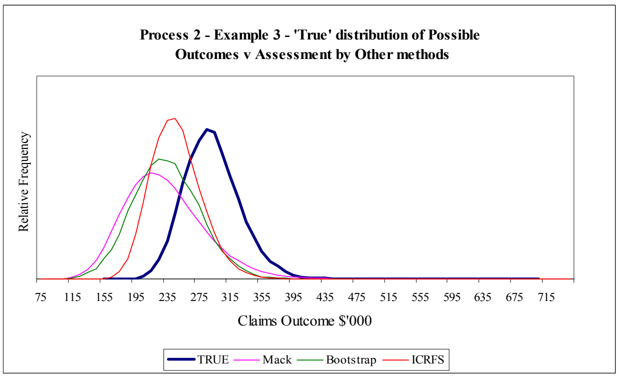


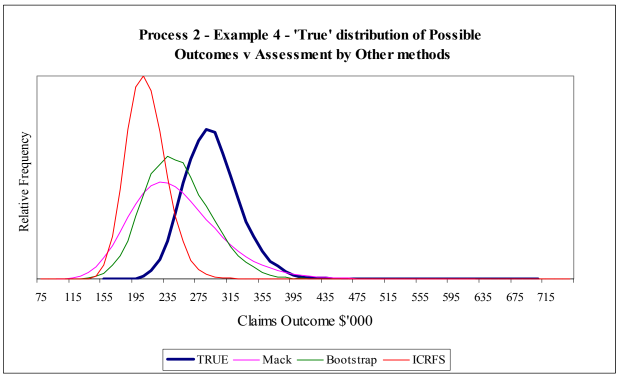
Under this process, the ICRFS models come noticeably closer than Mack or bootstrapping to capturing features of the 'true' distribution. As the models are not restricted to a chain-ladder, and the process doesn't satisfy chain-ladder assumptions, one would expect the PTF models to outperform bootstrapping and Mack in this situation. More comments on the results are again set out after the tables on page 33.

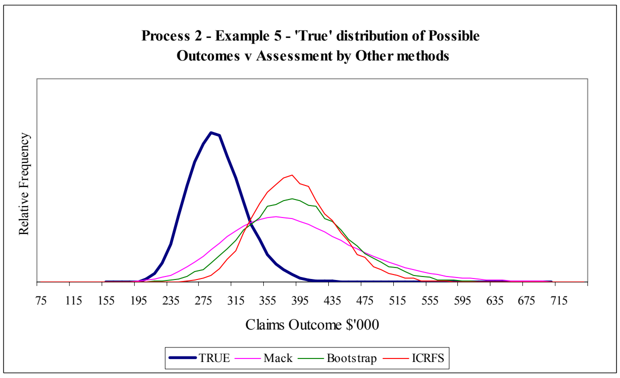
A second presentation of the results for Process 2 is set out below, that is consistent with the presentation for Process 1. Six charts are shown here, and a total of twelve (including the six here) are shown in Appendix F. The set of charts confirms the earlier impression that the ICRFS PTF model's additional flexibility allows it to better model this (non-chain ladder) process. In 8 of the 12 graphs presented in appendix F for Process 2, the visual impression is that the ICRFS model captures the true distribution much better than the other two methods.

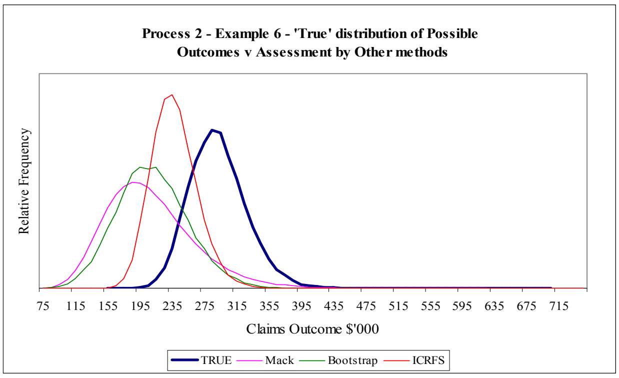












The next two tables gather key statistics from each of the models for the twelve examples.

Results Summary – Process 2 $'000								
Example Number	Mean				75th Percentile			
	True	Bootstrap	Mack	ICRFS	True	Bootstrap	Mack	ICRFS
1	292	339	336	292	314	368	374	323
2	292	332	327	286	314	367	370	306
3	292	234	232	246	314	262	262	266
4	292	248	246	209	314	275	279	225
5	292	391	390	390	314	432	439	421
6	292	210	208	237	314	239	240	255
7	292	229	228	266	314	253	255	288
8	292	443	440	312	314	486	498	338
9	292	261	259	310	314	299	299	337
10	292	278	276	269	314	308	311	306
11	292	230	228	137	314	257	260	158
12	292	257	264	265	314	289	305	289

Process 2 Risk Margin to get from True/Assessed Mean To True/Assessed 75th Percentile				
Example No.	**True**	**Bootstrap**	**Mack**	**ICRFS**
1	7.5%	8.6%	11.3%	10.6%
2	7.5%	10.5%	13.1%	7.0%
3	7.5%	12.0%	12.9%	8.1%
4	7.5%	10.9%	13.4%	7.7%
5	7.5%	10.5%	12.6%	7.9%
6	7.5%	13.8%	15.4%	7.6%
7	7.5%	10.5%	11.8%	8.3%
8	7.5%	9.7%	13.2%	8.3%
9	7.5%	14.6%	15.4%	8.7%
10	7.5%	10.8%	12.7%	13.8%
11	7.5%	11.7%	14.0%	15.3%
12	7.5%	12.5%	15.5%	9.1%

Observations include:

- As for Process 1, there is an impact associated with sampling from a skewed distribution. 8 out of the 12 triangles having the central estimate understated under Mack and Bootstrapping.
- 8 out of the 12 examples have the 75th percentile underestimated by the Bootstrap, Mack and ICRFS assessments. Even though in percentage terms, the assessed risk margin is higher than the 'true' margin required to protect against process variability, the estimation error is such that the 'additional' risk margin under each of the methods is not large enough to compensate for central estimate understatement.

- Testing whether chain-ladder assumptions are met by the triangles, should in practice lead to rejection of chain ladder (and hence the Mack method, and the particular application of bootstrapping illustrated) in 11 of the 12 illustrations. (Triangle 1 being the exception)
- The ICRFS fit for triangle 11 is a rather spectacular understatement of outcomes that are possible based on the underlying claims generating process. Examining this, the understatement has arisen because, by chance a negative trend is detected across transaction years in the observed history triangle that is assessed as statistically significant. I detected the same trend, when the PTF model was assessed manually (at –16% per period). In practice, it is unlikely that an actuary would project an indefinite continuation of such strong negative 'superimposed' inflation.
- Visually, the bootstrapping procedure and the Mack method do not do as good a job at a capturing the general shape of the 'true' distribution as they do for Process 1. The ICRFS assessment doesn't do a very good job either, but generally it is closer than the other assessments.

Unconditional Likelihood of Mack Method 75th Percentile covering Insurance Liabilities - Process 2

In practice, we would have only one triangle of claims history on which to base our analysis. Because the 'distribution of the distributions' assessed by the Mack method is skewed, one might pose the question, what is the unconditional likelihood that the 75th percentile point assessed under the Mack method will in fact prove adequate to cover the claims cost that emerges.

This would be a much harder question to answer for Process 1 (where the distribution of possible outcomes depends on the history triangle). However, for Process 2, it is possible to simulate many different history triangles that could have been observed, given the nature of the claims generating process I have set (each of which is associated with the same 'true' distribution for outstanding claims). The result of the Mack method assessment can be determined for each history triangle. More general observations about how closely the Mack Method assessment mirrors the underlying distribution of possible outcomes for this process can then be made.

Because of the formulaic nature of the Mack Method, it is possible to set up an automated process to take a very large number of simulations whereby claims history triangles are generated by Process 2. For each simulation, the Mack method can be used in an automated fashion to set the central estimate, and to arrive at the risk margin targeting a 75% likelihood of sufficiency.

To assess the unconditional likelihood, 10,000 simulations were performed, where a history triangle was generated, and based on each of these, assessed the mean and 75th percentile of the distribution of possible claim outcomes using the Mack Method. In this set of simulations:

- The distribution of the distributions was confirmed as right-skewed.
- 41% of the time, the location of the 75th percentile was understated
- Based on analysis of a single random triangle (and in practice there will be only one history triangle per valuation class) the likelihood that the 75th percentile judged by the Mack analysis would be sufficient to cover the claims cost that actually emerges was 71%.

  This was calculated follows:

  Let M_i = 75th percentile of the distribution of possible claims outcomes as assessed by the Mack Method for simulation i.

  X = the random variable describing the true outstanding claims liability (ie a single observation drawn from the distribution function described earlier).

  Then, the likelihood that a 75th percentile as determined by the Mack method will be sufficient to cover the cost of claims that emerges was calculated as:

$$\sum_{all.10,000.simulations} \Pr(X > M_i)/10{,}000$$

  The unconditional likelihood in this instance seems to below the desired level of 75%.

3.4.6 Conclusions

The main conclusion that I think can be drawn from the illustrations presented in this section are that one should not blindly apply the results of a mechanical, if statistical, analysis of variability. Most actuaries would agree that such an approach is rarely appropriate to set central estimates, but it seems to be quite common for risk margins to be set rather mechanically, based on something like the results of application of the Mack method, bootstrapping, or guidance provided in existing papers.

Risk margin assessment requires as much judgemental input as the central estimate. Where a quantitative method is used as an input, it is important to check that the assumptions underlying the method are verified, and a reasonableness check performed on the model structure. An attraction of the ICRFS platform is that it isn't possible to construct your model without simultaneously generating tests of fit. Where other quantitative methods are used, tests of fit must be also be generated in order to check whether meaningful conclusions can be drawn from the analysis.

The task of risk margin assessment is not complete without considering risk of estimation error that is not captured by the quantitative analysis, and adjusting analytic results to take them into account. I would argue that these risks are present in the majority of actuarial valuations. The risks are associated with the potential sources of *estimation error* that are not associated with *process variability.* They include items relating to the application of judgement, unidentified data issues, and data misinterpretation.

An observation that I did not anticipate before going through this exploration exercise was the strong effect associated with bias arising from setting projection parameters based on an average of past observations, if the underlying distributions are right skewed.

4 Correlations and Diversification Benefit

Where a company writes risks across a number of classes, it seems reasonable to anticipate that the values assessed to give a 75% likelihood of sufficiency for each class, when summed, will give an amount that has a greater than 75% likelihood of covering the company's total Insurance Liabilities. The reduction from this sum to arrive at the amount assessed to have a 75% chance of covering the company's total Insurance Liabilities is called the diversification benefit.

A diversification benefit can arise because there may be differences in the extent to which claims run-off experience deviates from the central estimate for different classes.

4.1 'Direct' and 'Indirect' Sources of Correlation

Correlation, between how the amount ultimately paid compares to the central estimate for different classes can arise in two categories, that I have termed Direct and Indirect correlation, and defined below:

1. *Direct Correlation* - arising due to innate connections between cost drivers for different classes. One way to think of this is as cost drivers that are outside of the insurer's or actuary's control once the business has been underwritten.
2. *Indirect Correlation* - connection between the adequacy of central estimates arising from other sources.

If we were concerned only in the correlation between claims outcomes in an absolute sense, direct correlations would be our only interest. However, because we are concerned about correlations between the adequacy of central estimates, the indirect correlations also come into play.

Examples illustrate the distinction between direct and indirect correlation.

4.1.1 Direct Correlation

Direct correlation can itself be divided into two categories:

<u>Items affecting the events giving rise to a claim in the first place</u>.

An example of this might be the chance of a weather event such as a hailstorm occurring. To a degree, claims cost in say, home buildings and commercial fire classes will be associated because costs in both are likely be higher if such an event occurs.

Items in this category are most important for correlation of Premium Liability outcomes. They should affect outstanding claims liabilities less, as whether or not such an event has occurred should be known and factored into the Outstanding Claims Liability estimate before the actuarial advice is finalised.

Environmental Items affecting the way claims liabilities run-off

An example of this is the tort law reforms that have been implemented across different jurisdictions that, with some retrospective effect, have affected claims run-off patterns.

A second example relates to the effect of the business cycle. For instance, at times of economic downturn, claims under Directors & Officers policies tend to become more frequent. Under the same economic conditions, return to work prospects for injured workers might be poorer, leading to increased workers' compensation costs.

Direct correlation arises from sources such as:

- The tendency for economic cycles to have impact across a range of business classes
- The tendency for the weather to impact across a range of business classes

Direct correlation can be described without the need to refer to the central estimate, but direct correlation nonetheless has implications for the chance that estimates in different classes could be over-adequate together, or inadequate together.

It ought to be possible to assess (past) direct correlation between classes by analysing industry statistics.

4.1.2 Indirect Correlation

Indirect correlation refers to sources of association between the adequacy of outstanding claims cost across different classes that arises from sources other than innate connections between cost drivers.

An important source of, indirect correlation can be thought of as the *Actuary- Effect*. Many sources of indirect correlation are related to sources of *estimation error* that are not related to *process variability*, and which are systemic in nature. They have been described earlier and include:

- Reliance on case estimates
- Parameter selections that don't recognise that for skewed distributions the average of past experience is likely to lie below the true mean (estimates not recognising that the mean is greater than the median)
- Assumptions regarding the frequency of large claims (one of the more common and more crucial examples of the previous category)
- Weight given to qualitative company advice regarding portfolio or claims management changes without sufficient scrutiny
- Judgement overlays in parameter selection

For example, hypothetical Insurer XYZ might assert that its case estimation practice has been reviewed, and management of litigated claims is moved to a specialist area that aims to set more realistic estimates on these claims, and to settle the claims more quickly. In this hypothetical example, the Approved Actuary will need to form a judgement regarding:

- How to investigate and confirm the extent of these changes
- How to incorporate allowance for this change by altering claim development assumptions and assumptions about rates of claim finalisation rate.

It is unlikely that the judgement calls will be prove to be completely correct in hindsight when claims have run-off. Consistent application of judgement increases the likelihood that if one class has been underestimated, other classes could be underestimated also.

Another source of indirect correlation relates to changes in claims management procedures implemented across a range of classes that might affect their cost. This could be planned ;in which case there may be an '*Actuary Effect*' (as the actuary determines if and how projection assumptions should be altered) and also a chance that 'true' cost changes, could occur (the insurer would be hoping predominantly in the direction of savings). On the other hand it could be unplanned. This might happen if a key claims manager resigned who had control over management of claims across a number of classes.

Some actuaries have expressed the view that reserving strength is the subject of a cycle mirroring the Insurance cycle. (McCarthy & Trahair (1999)). The view is that as insurance markets go through a cycle of hardening and softening, so too reserving levels seem to go through a cycle. If this hypothesis is true, it could be another source of correlation impacting adequacy of reserves across a range of classes.

4.1.3 Industry Analysis and Correlation Assessment

Direct correlation effects are likely to be assessable with industry analysis. An unfortunate property of indirect correlation is that its assessment is not really assisted by Industry analysis.

Industry analysis consistent with the correlations of interest would involve appraisal, for a given Company, of the extent to which the estimated outstanding claims liabilities for different classes turns out to be too high or too low compared to the amount that is eventually paid. Public data allowing an appraisal of this correlation for different companies does not exist. If such data did exist, I suggest that it would reveal quite different correlations for different companies. The reason for this is that the forces in play that give rise to *indirect correlations* will be specific to individual companies.

For instance, for two companies, A and B, with similar portfolios, one would not expect to be able to draw reasonable conclusions about whether the actuary for Company A has overestimated Insurance Liabilities across a number of classes; based on a knowledge of that the Actuary for Company B did, if the Company B experience was say, in part a result of the actuary overestimating the impact of a change in case reserving methodology.

4.1.4 Factors Associated With the Strength of Indirect Correlations

There are instances where indirect sources of correlation might be expected to be strong. The instances can be characterised as those cases where the actuary applies judgement as part of the assessment of central estimates and risk margins with some element of consistency across classes. I would argue that it is reasonable to expect a given actuary's judgement to be applied consistently across classes most of the time, especially for the particular case that is relevant to our interest, which is at a single valuation date, and for a single company.

As well as judgement, there is also the matter of influence to consider. This includes both overt and more subtle instances. For example:

- The influence of the actuary's own past experience.
- The influence of being cognisant of a situation where the Insurance Liability assessment might give rise to a reported capital adequacy position that is weak (either because the actuary knows without being told, or has this pointed out to them by company management). Some actuaries will have experienced across the board pressure on their estimates where critical thresholds, such as loan covenant triggers loom.
- The influence of not wanting to move too much from the previous basis in successive valuations.
- Attitudes to the extent of the necessity to allow explicitly for superimposed inflation.
- How motivated company management is to assist the actuary to arrive at their claim projection assumptions, and the actuary's reaction to such assistance.

It is my view that there will be many instances where indirect sources of correlation will be stronger than the direct sources. Perhaps this will even generally be the case. While they can't be quantified, indirect correlation sources still need to be recognised when assumptions are set to quantify diversification benefits.

4.2 Empirical Analysis that would support Diversification Benefit Assessment (And why it can't be done).

Assessment of diversification benefit is problematic. Ultimately, like most aspects of Insurance Liability assessment, judgement inevitably plays a significant role.

Even though pragmatically, objective quantitative analysis that would support a diversification benefit can't be performed, it can still be useful to think through what the analysis would look like, to help conceptualise the problem.

The illustration is in respect of two classes, but could be readily extended to a greater number. To remove complications and remain focussed on the assessment of diversification benefits, the illustration assumes that business volumes and the general environment remain constant over time.

Ideally, we would be in possession of the central estimate set at a large number of valuations for the two classes. Some notation is helpful. For the valuation performed in year i, let the central estimate for class A be A_i, and the central estimate for class B, be B_i.

We would also like to be in possession of information about the amount that was actually paid that the estimates A_i and B_i, aimed to cover (ie we would like claims run-off in respect of the past valuations is complete). Let us call these values A_i^* and B_i^*

If we had N such valuations, standalone risk margins M_A and M_B, could be empirically set[8] by backsolving for M_A and M_B so that

$A_i^* > A_i \times (1+M_A)$ in only 25% of cases; and

$B_i^* > B_i \times (1+M_B)$ in only 25% of cases.

If the diversification benefit is expressed as a percentage reduction to the risk margin, we seek the value for '*d*' that:

$A_i^* + B_i^* > A_i \times (1+M_A(1\text{-}d)) + B_i \times (1+M_B\,(1\text{-}d))$ in only 25% of cases.

d would be obtained empirically by backsolving.

The empirical risk margin and diversification benefit allowance determined in this way would take into account any central estimate bias, the sources of variability relevant to risk margin assessment, and the correlations.

However, there are many (obvious) reasons why such an empirical study can't be done, including:

- A large number of observations will be necessary. We will never be in a position to have enough data to look at the problem this way – particularly once it is recognised that we would like to be looking at the particular company under consideration.
- Changes can occur to the business under consideration that renders the past less relevant. So even if we were in possession of a long history of how outcomes compared with prior estimates, we would not be in a possession of all of the information that we would like to have available, that would be of genuine assistance:

  eg. - Change to business volume

  - Change in the characteristics of the risks covered within a class (underwriting or benefit changes; changes to claims handling; changes to propensity to claim)

- Changes can occur to the valuation approach.

  eg. - The actuary may 'learn' from past experience.

  - There may be a change in which actuary is doing the work.

  - There may be improvements to the valuation approach

Notwithstanding that we will only ever have a small set of data to review, it would be still be possible to review the track record of whether, when Insurance Liability estimates prove to be too low in one class, there are generally offsets from other classes where Insurance Liability estimates have proved to be too high.

[8] provided we were satisfied we had a large enough number N, and that there would be no changes over the future run-off period that had implications for risk margins.

Such a review can be done at annual intervals for short-tail classes. The position with long tail classes is more complicated, since outstanding claims at any valuation will not have fully run off by the next.

What can be done based on data analysis is very limited. This presents a serious problem for actuaries who need to assess diversification benefit. Any assessment will therefore be subjective.

The allowance will be the conclusions drawn from a thought experiment that:

- Frames the question we aim to solve by setting the diversification allowance.
- Considers all of the relevant sources of correlation.
- Converts this assessment into a diversification allowance.

A reasonableness check of the assessment will then be required.

4.3 Existing Guidance to Actuaries assessing Correlation and Diversification

The two papers providing assistance to actuaries responsible for GPS210 assessments provide the following descriptions of their approach to assessing correlations between adequacy of Insurance Liability estimates for different business classes:

Bateup & Reed:

> *"To allow for diversification across lines of business, we subjectively selected an assumed correlation matrix between the net outstanding claims liabilities of certain lines of business. The selected correlation matrix was based on our market knowledge and input received from senior actuaries practicing in the industry."*

Collings & White:

> *"Our qualitative assessment consisted of a survey of all of the experienced general insurance actuaries in our firm. Each was asked to identify the degree of correlation between outstanding claims liabilities for ten different classes of business as either high, medium or low (and positive or negative). Once the survey results were compiled, several meetings were held, during which survey results were discussed for each pairing of classes. Participants put forward their views as to the causes and extent of correlation and a consensus view was formed."*

It is no surprise that the correlation coefficients incorporated in the advice set out in the papers are judgement based. In my view there is no practical alternative to this. What I do find surprising is the notion that any general guidance would be considered applicable to any particular company. Direct sources of correlation can be assessed across different companies, and conclusions and benchmarks determined. However, it seems to me that indirect sources of correlation will typically be more important. These sources of correlation are specific to individual company circumstances and are not amenable to 'industry analysis.'

While the authors of the two papers may not have intended that their analysis be used as the definitive basis for diversification benefit assessment, considerable weight seems to be given to the suggestions set out in their papers.

From an audit perspective, an issue arises because the results set out in each paper results in a diversification benefit that, for companies with several classes of business, commonly more than halves the standalone risk margin. This is often a highly material reduction to the balance sheet liabilities and its support is worth examining closely.

Compared to the pre GPS210 environment, based on what I have seen, standalone risk margins do not seem to have risen, but diversification benefit recognition has become much more common. Typically the justification presented in written reports references the Bateup & Reed or Collings & White papers. It is my suspicion that, overall, net of diversification benefit allowance, there would be many instances where risk margins have fallen compared to the pre-GPS210 environment. It is also my suspicion that this is not an effect that APRA would have intended the new Prudential Standard to have.

Examining the Assumptions in the Bateup & Reed and Collings & White papers

Although ultimately, all sources of diversification are of interest, in this paper I have focussed on that between Outstanding Claims Liability adequacy across different classes. This is firmer ground for actuaries than Premium Liability assessment; which is a new and developing area. However, I expect many of the points that I to be equally applicable to Premium Liability assessments.

Both papers present results based on an assumed correlation matrix describing the association between the likelihood of adequacy of Insurance Liability assessments for different classes. The assumed matrices are set out below.

Bateup &Reed

The table below shows the assumed correlation matrix applicable to the total variance (ie with correlations between systemic components reduced to allow for the fact that some components of variability are independent) in order to estimate the variance of the aggregate net outstanding claims liability for an insurer with multiple lines of business.

	Liability	CTP	W'Comp	Prof Indemnity	Inwards Re	Fire/ISR	Motor	House	Other
Liability									
CTP	25								
W'Comp	25	35							
Prof Indemnity	25	25	25						
Inwards Re	25	25	25	25					
Fire/ISR	0	0	0	0	5				
Domestic Motor	0	25	0	0	5	10			
H'holders	0	0	0	0	5	10	20		
Other	0	0	0	0	5	5	10	10	

In effect, the matrix is set for an industry 'average size' portfolio, which (in $2001) is set as

'Industry Average' Portfolio[9]	
Class	**2001$m**
Liability	123
CTP	313
Workers' Comp	88
PI	57
Fire/ISR	22
Domestic Motor	26
Householders	25
Other	42
Inwards Re	52
Total	**749**

Collings & White

I find this paper less clear about whether the matrix it presents, is intended to relate to the systemic component of variability only. Alternatively, it may have been reduced in a manner consistent with the Bateup & Reed presentation, in recognition that some sources of variability are independent. If the presentation is after this reduction has been made, I don't find the paper clear about the size of portfolio that the authors intend it to be appropriate for.

For this reason, it isn't clear to me whether the Collings & White presentation is directly comparable with that of Bateup & Reed. If the presentation is in respect of the systemic component, without a reduction for the independent component of variability, then the values in the matrix would need to be reduced before making direct comparison with the Bateup & Reed presentation. This detail may be important for those wanting to fully understand the assumptions underlying the paper's recommendations, but it does not affect the general thrust of the comments that I have regarding the two sets of assumptions.

The Collings &White correlation matrix is presented below:

	Liability	**CTP**	**W'Comp**	**Prof Indemnity**	**Marine**	**Fire/ISR**	**Motor**	**House**	**Other**
Liability									
CTP	**25**								
W'Comp	**25**	**25**							
Prof Indemnity	**30**	**10**	**15**						
Marine	**0**	**0**	**0**	**0**					
Fire/ISR	**0**	**0**	**0**	**0**	**10**				
Motor	**0**	**0**	**0**	**0**	**10**	**20**			
H'holders	**0**	**0**	**0**	**0**	**10**	**40**	**20**		
Other	**0**	**0**	**0**	**0**	**10**	**20**	**30**	**20**	

[9] Bateup & Reed (2001), Appendix D.

Comments on the Matrices

The immediately striking thing about each matrix is the large number of zero elements. For these class combinations, this would seem to dismiss the possibility adequacy of Insurance liability estimates in these classes could be related due to things such as:

- Reliance on case estimates.
- The way large claims allowances are assessed, given they will be based on a small sample from a highly skewed distribution.
- Parameter selections that might inadvertently reflect a median or mode, more so than a mean.
- Consistent application of judgement across a range of classes.

The zeros suggest that the indirect sources of correlation I have contended are important earlier in this paper, were either not considered, or were considered but assessed to be unimportant.

In my opinion, it is quite difficult to accept this as a reasonable default position.

It is easier to accept that the matrices include allowance for direct sources of correlation. Looking at which entries are higher than others, intuitively they seem to make sense. Where a common factor like the weather, or the position in the business cycle would be expected to have an effect on claims cost for two classes the entry in each matrix is higher.

A complicating factor in assessing whether this forms a reasonable basis from which to estimate diversification in the adequacy of Insurance Liability estimates, is that; at least in terms of past events, the weather conditions will have been known. The actuary would know whether the recent past experience included a storm that would be expected to increase the cost of claims in both ISR and home classes. This knowledge should be an input to the Insurance Liability central estimates. However, the actuary will still need to estimate the cost associated with the event, and it is possible that adequacy of these estimates in different classes may be correlated. Similarly, the position in the business cycle will typically be known when the Insurance Liability assessment is set. With this source, the correlation of interest will more commonly arise where there is an unexpected movement along a business cycle path.

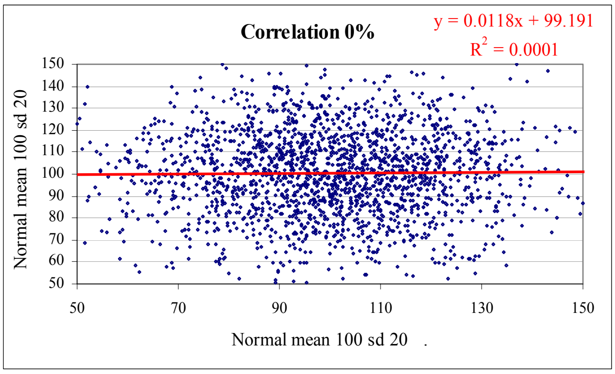
This line of reasoning adds to my impression that it is the indirect sources of correlation that are most important to consider when assessing diversification benefits.

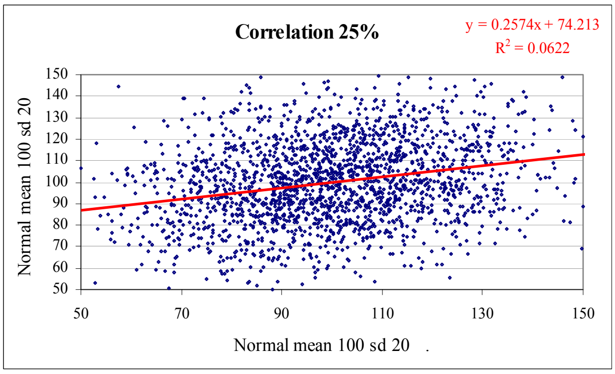
Assessment of the non-zero entries (and any suggestion about alternative values describing the ‘zero’ relationships) requires appreciation of the meaning of correlation coefficients of different magnitude.

4.4 Exploring the Meaning of Correlation Coefficients

The strength of linear association described by a correlation coefficient of a given magnitude is not something that is necessarily easy to grasp naturally. To assist, I have presented a series of charts that aim to illustrate its meaning.

The first chart is a scatter-plot generated from 2,000 random observations of a bivariate normal random variable (X_1, X_2), where the marginal distributions for X_1 and X_2 each have mean 100, and standard deviation 20, and where the correlation coefficient of X_1 and X_2 is 0.00. The random observations have been generated using the Excel add-in package @Risk.© A second chart shows the equivalent plot, where the correlation coefficient of X_1 and X_2 is 0.25. [10] In each chart, each marker represents observation of a single pair. The red line is the least squares line of best fit drawn through the scatter plot.





A number of observations can be made from examination of the charts:

- Visually, the degree of association implied by a 25% correlation coefficient seems quite weak to me.
- For the special case where X_1 and X_2 have equal standard deviation, the slope of the least squares line of best fit equals the correlation coefficient. This provides a degree of tangible interpretation of the meaning of correlation co-efficient.[11]
- The coefficient of determination of the least squares line of best fit, r^2, is the square of the correlation coefficient. In a sense, the square of the correlation coefficient is a measure of the proportion of the variance of the variable X_2 that is not 'explained by' variation in the outcome of X_1.

  ie If $\overline{X_2}$ equals the mean of the values of X_2 in the scatter plot. Then, if we define for each ordered pair (X_1, X_2), a quantity X_2', being the point on the least squares line of best fit corresponding to X_1, then it can be shown that:

$$\sum(X_2-\overline{X_2})^2=\sum(X_2'-\overline{X_2})^2+\sum(X_2-X_2')^2$$

(Total)   (explained by model) + (not explained by model)

---

[10] A co-efficient of variation of 25% is the Bateup & Reed 'benchmark' for public liability, workers compensation and professional indemnity portfolios outstanding claims liabilities of approximately $134m in current vales as at 2001.

[11] In the more general case, where the standard deviation of the marginal probability distribution functions are σ_1,σ_2, with $\sigma_1\neq\sigma_2$, the slope of the line of best fit is $\rho\times\sigma_1/\sigma_2$.

In our situation dealing with the degree of association between outcomes for two insurance classes, '*explained by model*' represents variation in X_2 that is anticipated due to variation in X_1. r^2 represents that portion of the total variance, that is '*explained*' by (or is associated with) movement in X_1

ie $$r^2 = \sum (X_2' - \overline{X_2})^2 / \sum (X_2 - \overline{X_2})^2$$

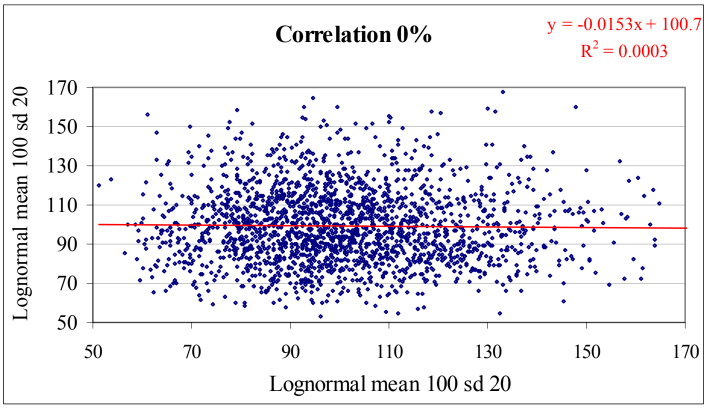
The strength of association implied by correlation coefficients of different magnitude can be illustrated by means of a series of scatter plots. Four examples are set out below, with values reflecting those in the correlation matrices. A larger selection is set out in Appendix G.

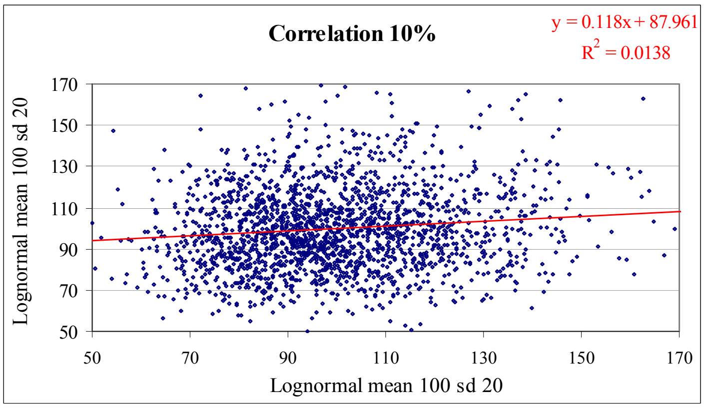
The scatter plots are generated by simulation using the excel add-in package @Risk. Each plot represents observations of a pair of lognormal random variables with mean 100 and standard deviation 20, associated with a correlation coefficient of the stated magnitude. Each dot represents a single simulation.

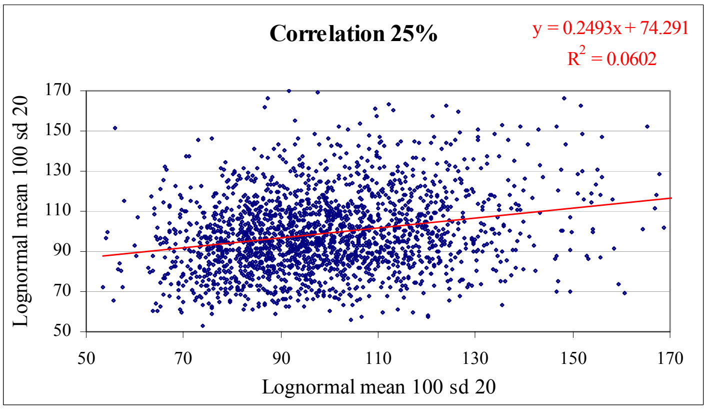
2,000 simulated observations are shown in each plot. The least squares line of best fit, and its associated r^2 value is shown in each chart.

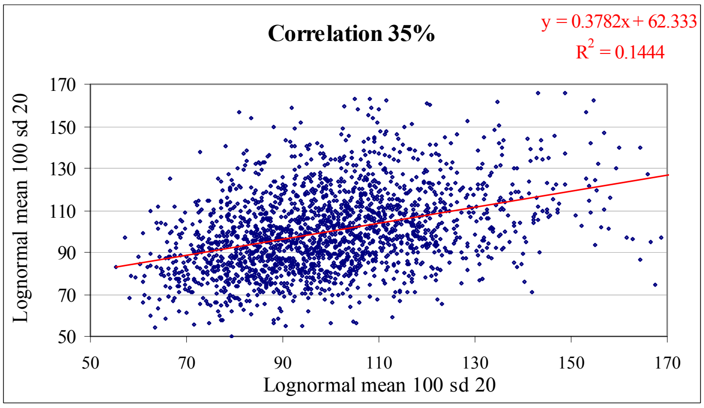
My aim in presenting this series is to help actuaries who are trying to select assumptions for correlation between the likelihood of adequacy of outstanding claims liability estimates, and to allow some assessment of the correlation matrices set out in the two papers.

One can interpret the charts as expressing the distribution of outcomes that might be possible for two classes each with a central estimate for the liabilities of 100, and standard deviation 20, where the correlation between the likelihood of adequacy of the estimate is expressed as a correlation coefficient of the stated magnitude.









In my view the strength of relationship associated with a correlation coefficients of 10% and 25% (which are the most common non-zero entry in the assumed correlation matrices) is quite weak. Without the assistance of the least squares line of best fit, visually they aren't easy to distinguish from the zero correlation scatter plot.

Looking at the scatter plots, it is not difficult to imagine that diversification benefits supporting a halving of standalone risk margins could be assessed to arise from having liabilities spread across a range of classes, if one accepts that the correlation between adequacy of insurance liability estimates is described these charts. However, it is my view that likelihoods of adequacy for a given insurer would be considerably more closely related.

Correlation Coefficients and Conditional Probabilities

A second way of exploring the meaning of correlation coefficients involves examining conditional probability distributions. For simplicity, this presentation is based on the bivariate normal distribution.

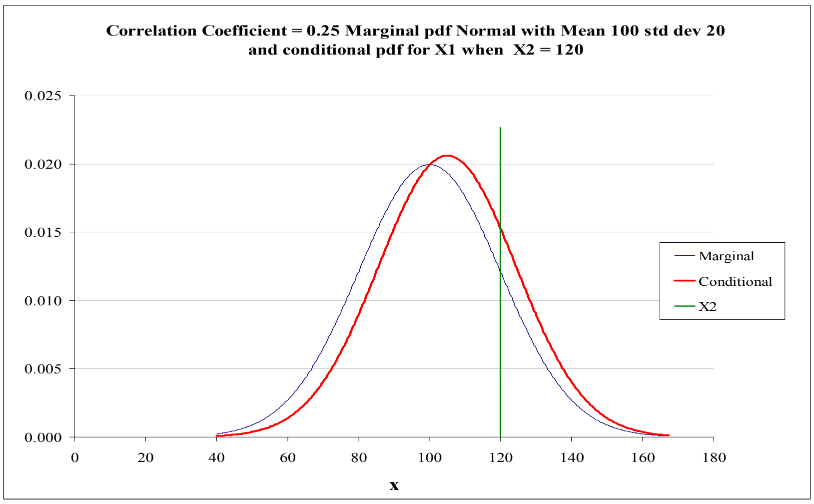
Given the random variables X_1 with mean μ_1 and standard deviation σ_1, and random variable X_2 with mean μ_2 and standard deviation σ_2; if the correlation coefficient of X_1 and X_2 is ρ and X_1 and X_2 jointly have the bivariate normal distribution, then the conditional density of X_1 given X_2 is a normal distribution with:

Mean	$\mu_1 + \rho \times (\sigma_1 / \sigma_2) \times (X_2 - \mu_2)$; and
Standard deviation	$\sigma_1 \times \sqrt{(1-\rho^2)}$

The following set of charts show for two normally distributed random variables X_1 and X_2 each with mean 100 and standard deviation 20, and a given correlation coefficient:

1. The original marginal probability density function for X_1 (thin blue line)
2. For a given observation of X_2 (shown by the green vertical line)
3. The conditional probability density function for X_1 (the thicker red line)

This can be easily set up on a spreadsheet, so as to explore various combinations of marginal pdfs. In practice, it is uncommon for actuaries' to assume that claims cost outcomes are drawn from a normal distribution. Nonetheless, this can still be a helpful way to explore the meaning of correlation coefficients of different magnitude.

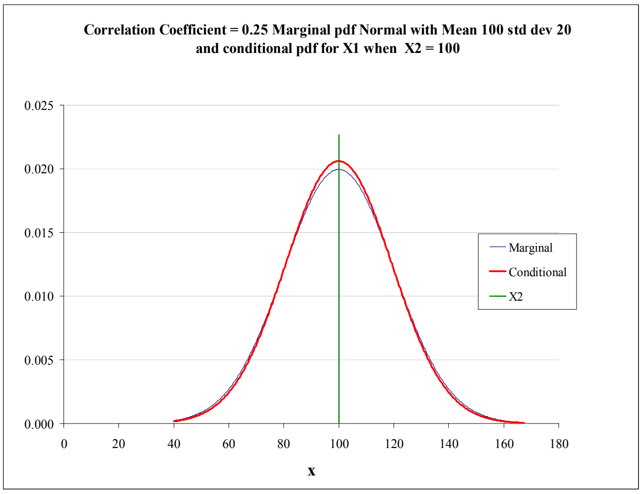


The graph shows that, if originally we estimate that claims outcomes (for each of two classes) will be drawn from a distribution represented by the blue line, then if the outcome for one class is as shown by the green weight, then our revised estimate of the outcome for the related class is that it will be drawn from a distribution represented by the red line.

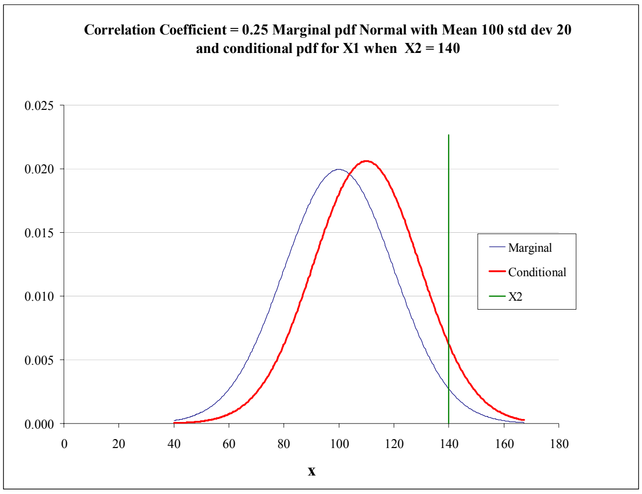
In the case where the original distributions are normal with mean 100 and standard deviation 20, then if the correlation coefficient is 0.25, and the outcome in one class is 120, then our revised estimate is that the outcome for the second class will be drawn from a distribution with mean 105, and standard deviation 19.4.

Another way to look at the strength of association is to note that, after observing the outcome of 120 for the second class, the assessed probability that outcomes in the first class will exceed 100 increases from 0.50 to 0.60.

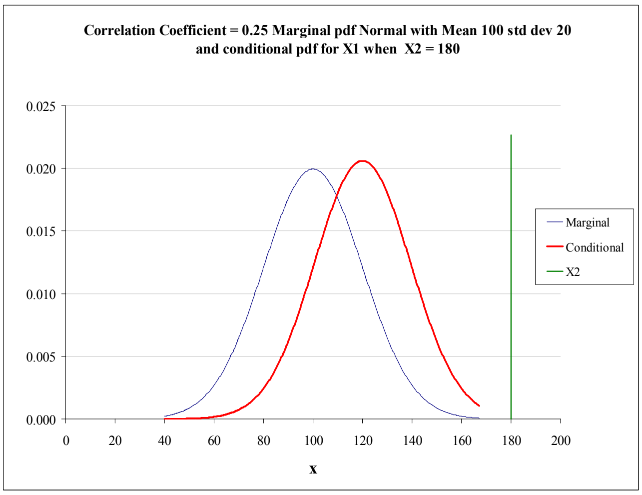
Similar charts are presented below, that look at how the conditional probability distribution moves, if the outcome in one class takes on different values.



$X_2 = 100$
Conditional probability for X_1 has:
Mean = 100; standard deviation 19.4



$X_2 = 140$
Conditional probability for X_1 has:
Mean = 110; standard deviation 19.4



$X_2 = 180$
Conditional probability for X_1 has:
Mean = 120; standard deviation 19.4

These charts appear to confirm the impression gained from the scatterplots, that a correlation coefficient of 0.25 is not a strong degree of association. Revised estimates of the outcome distribution for X_1 are not greatly influenced by the outcomes we might see for the correlated variable X_2.

4.5 Summary and Conclusions

I would like to be presenting an analytic solution to how diversification benefits should be objectively assessed. However, the sources of correlation that will act to reduce diversification benefits are:

- Particular to individual Insurer circumstances.
- Particular to the different circumstances that surround reviews at different points in time.
- Associated with judgement, which will vary over time, but will have some degree of consistency at any given review.
- Not greatly assisted by industry analysis.

In my view, from an analytic perspective, the problem is an intractable one. The best that one can do is to go through the thought process of identifying sources of correlation that may be in play. This includes those that I have categorised as ‘direct’ and those that I have categorised as ‘indirect,’ and think through how confident we can be about their effects. This will include taking a step back and reviewing how each aspect of the review was conducted and what the risk points are with respect to items such as

- Data veracity
- The influence of company staff that have sought to assist our understanding
- Where we have applied judgement
- The risk that the approach to parameter selection risks bias

It is worth reviewing whether, if prior year strengthenings have been required in a one class in the past, have there tended to be offsets from releases in other classes.

In a framework where correlation matrices are used as the basis for the diversification benefit assessment, it is necessary for the actuary to reach a position where they are confident they understand the strength of association described by a given correlation coefficient.

Ultimately, the position is similar, if less familiar, than the position with respect to the central estimate and the risk margin, with judgement inevitably being required and the end result being reviewed for reasonableness.

The current papers guiding actuaries through diversification benefit assessment, result in a calculated diversification benefit that often reduces risk margins by 50% for an insurer writing across a range of classes. I would question whether this should be regarded as a reasonable default position given that we can’t do work to verify with a reasonable degree of certainty that a benefit of this magnitude exists.

This section has presented the items that I think need to be considered in arriving at an assessment of what might impact adequacy of estimates across a range of insurance classes. Many of them are not directly related to the nature of the risk that is insured, or correlations between cost drivers. The other sources are ones related to the valuation processes and the influences on them. Both are important, and in my view, although it isn't obvious which are the stronger, I suspect that for outstanding claims it is the indirect sources. For premium liabilities direct correlation sources will probably play a bigger role.

In such an environment it doesn't seem reasonable to have as the starting position, that diversification benefits that will halve risk margins exist. In audit support work, I have sometimes been challenged that if I don't think diversification benefits of this order of magnitude exist, I need to demonstrate this. I find it very difficult to accept the notion that if a view is taken that high diversification benefits exist, but can't be demonstrated, it is up to me to show they are not there.

APRA's explicit recognition of diversification benefits in the prudential standards has had the effect of providing a new point at which (in some companies) management can pressure their actuaries to reduce their provision recommendations. I would question whether this helps the GPS210 aim of ensuring the board's of general insurers receive reliable advice about the extent of their Insurance Liabilities.

5 Observations & Conclusions

The more significant conclusions that I have reached regarding GPS210 Insurance Liability estimates are:

- The distinction between variability in claims outcomes in an absolute sense (*process variability*) and variability in the gap between claims outcomes and the actuarial estimates (a mixture of *process variability* and *estimation error*) is very important. Failure to recognise this difference will generally result in a risk margin that is too low to satisfy its aim.
- Quantitative methods used in practice to assess risk margins do not capture an important relevant component of variability. In the paper, I have called this '*estimation error not associated with process variability'*. For the reasons set out in this paper, I expect that sources of this variability are generally:
  - Significant
  - Systemic in nature
  - Correlated across Insurance classes
- Diversification benefits derived from the formulae set out in the Bateup & Reed and Collings & White guidance to actuaries are highly material. If we can't do work to verify with a reasonable degree of certainty that a benefit of this order of magnitude exists, it doesn't seem prudent, or even reasonable, to take the credit for it.
- The nature of the insurance claims process; in particular the skewed nature of possible claims outcomes, gives rise to risk of central estimate understatement if one is not aware of the properties of samples from skewed distributions.
- Checking model fit is important, both for central estimate determination and risk margin assessment.
- Insurers will always be aiming to improve their underwriting and claims management practices. It is unsurprising that they are keen for their actuaries to take the improvements into account before they are evidenced in claims outcomes. It is difficult to justify taking such advice into account without being able to either verify it, or see it clearly evidenced in claims data. Doing so risks underestimation.
- It is not hard to see how an Insurance Liability assessment can be at risk of a triple whammy effect of central estimate understatement, risk margin understatement, and diversification benefit overstatement. If this happens, the overall Insurance Liability estimate could fall well below the level required to provide a 75% likelihood of adequacy.
- The Actuarial Control Cycle can assist in reviewing risk margins. If in utilising the control cycle to review prior year claim cost estimates, what are supposed to be 1 in 4 year events seem to be more frequent than that, this could indicate a problem.

Bibliography

Barnett G, and Zehnwirth, B (2000) "Best Estimates for Reserves" as published in the Proceedings of the CAS Volume LXXXVII, Numbers 166 & 167.

Bateup R and Reed I (2001) "Research and Data Analysis Relevant to the Development of Standards and Guidelines on Liability Valuation for General Insurance". Presented to the Institute of Actuaries of Australia XIII General Insurance Seminar.

Bonnard R, Greenwood M, Greybe S, (1998) "Bootstrapping Reserve Estimates." Presented to the Actuarial Society of South Africa.

Collings S and White G (2001) "APRA Risk Margin Analysis", Presented to the Institute of Actuaries of Australia XIII General Insurance Seminar.

Efron, B and Tibshirani R J, (1993) An Introduction to the Bootstrap. Chapman & Hall, London.

England, P D. & Verrall, R.J. (1999) "Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving." Insurance : Mathematics and Economics, 25, 281-293.

England, P D & Verrall R J (2001) "Stochastic Claims Reserving in General Insurance." British Actuarial Journal, 8 , III, 443-544.

England, P D (2002) Addendum to "Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving." Insurance: Mathematics and Economics, 31, 461-466.

Mack, T (1994) "Measuring the Variability of Chain Ladder Reserve Estimates" Casualty Actuarial Society, Spring Forum.

McCarthy P & Trahair G (1999) – "Lack of Industry Profitability and Other Stories" Presented to the Institute of Actuaries of Australia XII General Insurance Seminar.

Mooney, C. Z., & Duval, R. D. (1993). Bootstrapping: A nonparametric approach to statistical inference. Newbury Park, CA: Sage Publications.

Rubin, D B (1981) "The Bayesian Bootstrap" Annals of Statistics 9: 130-134.

Zehnwirth B, Barnett G "Claims Reserving – Should ratios be used" (unpublished).[12]

Acknowledgements

The views and opinions expressed in this paper are solely mine and may not reflect the views of my employer.

I am grateful for assistance of Dr Eugene Dubossarky and Warrick Gard during the course of preparing this paper, whose comments helped clarify my thinking on many of the subject matters discussed.

[12] Though unpublished at the time this paper was prepared, I understand it will be published on the Insureware website in the near future. The webpage reference will be: http://www.insureware.com/library/articles.shtml.

Appendix A

Checks on Chain Ladder Fit

This appendix sets out three tests that can be performed to verify whether assumptions underlying chain-ladder projections are met. In my view, these tests should be performed before chain-ladder techniques are used to project claims experience, and before analytic techniques are used to provide input to Insurance Liability risk margin assessments, that assume a well fitting chain ladder model fits the claims history well. The tests set out here have been based on those described by Mack (1994), and considering views expressed in Zehnwirth & Barnett's unpublished paper "*Claims Reserving – Should Ratios be Used.*"

Assumption 1 – Form of Model

The Chain Ladder method estimates ultimate incurred claim costs for a given accident year i as the product of development factors for development periods j+1 to n, applied to the cumulative incurred claim costs at development period j. The development factors are assumed to be independent of accident year.

According to the model, the incremental claims cost in development year k, should vary around a constant multiple of the cumulative claims cost to the end of development period k-1.

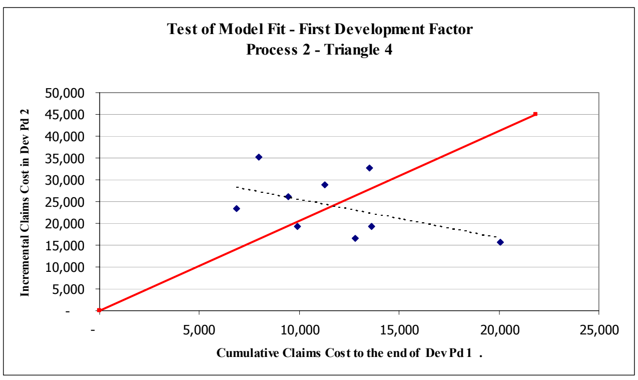
That is ; if Y = incremental claims cost in development period k; and
X = cumulative claims cost to development period k-1; and
F_k = the assumed development factor taking cumulative claims cost at the end of development period k-1 to cumulative claims cost at the end of development period k.

The pairs of observations (X,Y) should fall randomly about a line with equation

$$Y = (F_k - 1)\,X$$

It is possible to check graphically whether this relationship seems to hold at each development period.

An example from Process 2 is shown below. The chain ladder estimate for the first development factor is 3.06. This means that the estimate of the incremental claims cost in development period 1, is 2.06 times the (cumulative) developed cost to the end of development period 0. The red line has equation y = 2.06x. Observed points do not seem to vary randomly about this line (the least squares line of best fit is shown by the dashed line and actually has negative slope). In practice, a chain ladder model would not be accepted as fitting this data.



The form of the model should be assessed this way for each development period (where there are a reasonable number of observation points to assess).

Assumption 2 (No calendar year effect)

The chain ladder method assumes that development factors are independent of accident year. In practice, calendar year effects such as changes in claims handling, case reserving, or superimposed inflation can lead to violation of this assumption.

Qualitative tests include checking whether there had been changes with a calendar year effect such as changes to claims handling procedures, changes in case estimates setting procedures, and large movements in inflation rates.

A more formal test is suggested in Mack (1994). It is illustrated by example below (for Process 2 triangle 4). The test progresses as follows:

Individual Chain ladder factors are determined, and at each development period, the median found.

Accident Period	Development Period: Ci1	Ci2	Ci3	Ci4	Ci5	Ci6	Ci7	Ci8	Ci9	Ci10
1		2.421	1.216	1.238	1.401	1.050	1.030	1.036	1.008	1.009
2		2.954	1.366	1.239	1.166	1.060	1.089	1.054	1.020	
3		1.778	1.585	1.103	1.062	1.117	1.028	1.026		
4		2.302	1.350	1.137	1.134	1.310	1.035			
5		5.398	1.538	1.189	1.215	1.063				
6		3.776	1.349	1.201	1.082					
7		3.563	1.332	1.359						
8		3.420	1.403							
9		4.406								
10										
Median		3.420	1.358	1.201	1.150	1.063	1.032	1.036	1.014	1.009

For each development period, the observed development factor is described as 'L' if it is above the median, and 'S' if it is below (and * if it is equal to the median).

	Ci1	Ci2	Ci3	Ci4	Ci5	Ci6	Ci7	Ci8	Ci9	Ci10
1		S	S	L	L	S	S	*	S	*
2		S	L	L	L	S	L	L	L	
3		S	L	S	S	L	S	S		
4		S	S	S	S	L	L			
5		L	L	S	L	*				
6		L	S	*	S					
7		L	S	L						
8		*	L							
9		L								
10										

If there are no calendar year effects, each diagonal should have a roughly equal number of 'L' and 'S' values.

To test this, count the number of 'L' and 'S' factors in each diagonal.

For example take the last diagonal,

Number of 'L' factors = L(10) = 5
Number of 'S' factors = S(10) = 3

Z9 = min(L(10),S(10)) = 3 (the 9 indicating there are 9 observations to consider)

If Z9 is significantly less than (L(10)+S(10))/2, then there is evidence of a calendar year effect.

Mack (1994) describes an overall test that determines a test statistic based on all the calendar years, that can be tested to determine whether it is likely to have arisen by chance alone.

The design of the test is based on the hypothesis that each development factor (that does not give rise to an *) has a 50% chance of being an 'L' and a 50% chance of being an 'S'.

For the example diagonal L(10) + S(10) = 8

The probability that L(10) = m is $\dfrac{8!}{(8-m)!m!}\Big/2^8$

So that the probability that m = 4 is 0.273438

A table can be built up based on this to determine the probabilities of various values for Z9, Z8, Z7 and so on

For example, the probability that Z5 = 0 is 0.0625
Z5 = 1 is 0.3125; and
Z5 = 2 is 0.6250

The overall test concerns Z = Z2 + Z3 + …Z9; based on the first two moments of the distribution for Z under the null hypothesis that there is no calendar year effect.

The reader is referred to Mack (1994) for full details of the test.

Assumption 3 – No correlation between successive development factors

The chain ladder method assumes that across an accident year, the development factor experienced at a given development period is not affected by whether preceding factors have been particularly high or low.

Mack (1994) suggests that this assumption can be tested based on a test known as Spearman's Rank Correlation coefficient.

The general nature of the test is illustrated here, based on the same triangle as previously.

Considering the first and second development period development factors, we have eight pairs, that are set out in the next table. The development factors are ranked for each development period in turn, from 1 for the lowest to 8 for the highest. Other calculations for the test are shown in the table.

Accident Period	F_0	F_1	A=Rank F_0	B=Rank F_1	C=A-B	C^2
1	2.421129	1.215519	3	1	**2**	4
2	2.954018	1.366152	4	5	**-1**	1
3	1.777683	1.585133	1	8	**-7**	49
4	2.301957	1.350070	2	4	**-2**	4
5	5.398467	1.537619	8	7	**1**	1
6	3.775807	1.349386	7	3	**4**	16
7	3.562837	1.331617	6	2	**4**	16
8	3.419639	1.403059	5	6	**-1**	1

The Spearman Rank Correlation coefficient is defined by the statistic

$$r^* = 1 - 6\sum \frac{C^2}{N(N^2-1)}$$

where N is the number of pairs (ie here N= 8)

If successive development factors are uncorrelated, the expected value of r* is 0, and its variance is 1/(N-1) . For the example, the hypothesis that the successive development factors are uncorrelated would not be rejected.

Mack suggests an adaptation of this test that results in a single test statistic that can be used to test whether, across the entire history triangle, successive development factors are uncorrelated.

For full details of the suggested test, the interested reader is referred to Mack (1994).

Appendix B

Mack Method

What I have termed the Mack method in this paper is my interpretation of an assessment technique described by Thomas Mack in his 1994 paper "*Measuring the Variability of Chain Ladder Reserve Estimates*."

In the paper Mack explores the assumptions that implicitly sit behind the chain-ladder algorithm, where development factors $F_{k,}$ that take cumulative cost from development year k to k+1 are estimated as

$$F_k = \sum_{i=1}^{n-k} D_{i,j+1} \Big/ \sum_{i=1}^{n-k} D_{i,j}$$

Where

$D_{i,,j}$ represents the cumulative cost to the end of development year j from accident year I; and

n represents the total number of accident and development years in the claims triangle (which is also assumed to be the last development year in which claims cost will emerge).

Mack advances the point that using the chain-ladder algorithm, with development factors assessed in this way implicitly assumes that:

1. The claims process follows this model, and
2. The selected estimator for F_k is the unbiased estimator of minimum variance.

Mack highlights that this model therefore asserts:

- $E[D_{i,k+1}|D_{i,k}, D_{i,k-1},\ldots D_{i,1}] = F_k \times D_{i,k}$
  - => Same 'underlying' age-to-age' development factor exists for all accident years, but with random variation around this point
  - => For a given accident year, at a given stage of development, it is only the most recent observed cumulative claims cost outcome that carries predictive power regarding future development amounts. The chain-ladder is 'memoryless' in the sense that, for a particular accident year, the path taken to reach $D_{i,k}$ carries no predictive power – only the value $D_{i,k}$ does.
  - => No calendar year effect (such as would likely be violated if there was a change in claims management practice).
- Development factors are independent of the accident year
- Successive development factors are uncorrelated (ie after a rather high value of $D_{i,k}/ D_{i,k-1}$, the expected size of the next development factor $D_{i,k+1}/ D_{i,k}$ is the same as after a rather low value of $D_{i,k}/ D_{i,k-1}$)
- $Var(D_{i,k+1}/ D_{i,k}|\ D_{i,1}\ D_{i,2\ldots}\ D_{i,k}) = \alpha_k^2\ / D_{i,k}$ where α_k is a factor that may depend on k, but not i. [13]

[13] For details on the reasoning behind this assertion, the reader is directed to Mack's 1994 paper, in which he presents this as an implication of asserting that F_k is the estimator of the true value of the development factor that has minimum variance.

Mack asserts that these assumptions are strong and cannot be taken as met by every run-off triangle. He also suggests tests that can be used to verify these assumptions. Where the tests reveal that the implicit assumptions are not satisfied, he makes the point that the chain-ladder method should not be applied. It would also follow that, in this circumstance, this method should not be used to assess variability.

Mack then derives a formula to estimate the standard deviation of the outstanding claims liability estimate. The formula is not straightforward to describe, and the interested reader is directed to Mack's paper for its details, and example calculations.

I would draw attention to the following:

- The assessment of variability from the central estimate assumes that the chain-ladder method as specified has been used to arrive at the central estimate. It would not be reasonable to attach meaning to the Mack Method result for variability if a different approach was used to set the central estimate.
- Chain-ladder factors are assumed to be derived in a mechanical fashion as specified earlier (and this is OK, because before using the Mack method, tests will have been conducted to confirm that the method for deriving the chain-ladder factors is reasonable). However, if judgemental adjustments are applied such as:
  - Removal of 'outliers'
  - Incorporating an observed trend
  - Only using the last few diagonals rather than the complete history triangle to derive assumptions,

  then it would follow that use of the Mack method would be invalidated.
- If any of the assumptions underlying the assessment method are not satisfied, then in my view, meaning should not be attributed to the Mack method assessment of variability.
- The result of Mack's formula is a chain-ladder central estimate and an estimate of the standard deviation of possible claims outcomes. For practical use for GPS 210 assessments, an assumption is required for the distributional form of the variation in outcomes from the chain-ladder central estimate. Mack suggests a lognormal distribution be assumed, where a distributional form is required. No compelling reason is provided to support this suggestion; save that it has an intuitively appealing shape. Other standard distributional forms with a similarly appealing shape can be found that would imply a higher value for the 75th percentile, for any given mean and standard deviation.

I would also draw attention to the following point made by Mack in his paper indicating that care must be taken when employing the chain ladder method:

"The well known weak points of the chain-ladder method should not be concealed: These are the fact that the estimators of the last two or three factors F_n, F_{n-1},.. F_{n-2} rely on very few observations and the fact that the known claims amount $D_{n,1}$ of the last accident year (sometimes $D_{n-1,2}$ too) forms a very uncertain basis for the projection to ultimate. This is most clearly seen if $D_{n,1}$ happens to be zero. Then we have $D_{i,n} = 0$, (the estimate of the outstanding claims liability) $R_n = 0$, and (the standard error of the outstanding claims liability estimate) s.e.(R_n) = 0 which obviously makes no sense….

Thus, even if the statistical instruments developed do not reject the applicability of the chain ladder method, the result must be judged by an actuary and/or underwriter who knows the business under consideration. Even then, unexpected future changes can make all estimations obsolete."

Appendix C

Description of Bootstrapping Procedure

This appendix describes the bootstrapping procedure used to estimate the distribution of possible claims outcomes set out in Section 3.3. The process is described more fully by Bonnard R, Greenwood B and Greybe S (1998) in their paper "*Bootstrapping reserve estimates*". The description in this appendix is made with reference to the first example triangle generated from Process 2, as defined in the main body of this paper.

Incremental and cumulative payments for the first example triangle are:

Incremental

	1	2	3	4	5	6	7	8	9	10
1	8,691	16,412	13,021	14,220	10,261	6,349	9,286	2,500	3,111	3,242
2	12,284	15,302	18,410	9,155	7,240	4,126	4,756	1,943	3,981	
3	12,044	25,216	9,428	17,196	8,932	3,771	3,665	2,774		
4	7,823	14,241	9,843	11,990	11,487	1,005	1,367			
5	9,678	21,875	16,528	23,593	15,624	23,499				
6	8,520	12,163	34,000	13,779	11,429					
7	14,529	41,605	21,048	20,067						
8	10,190	18,336	13,488							
9	14,092	23,813								
10	9,823									

Cumulative

	1	2	3	4	5	6	7	8	9	10
1	8,691	25,103	38,124	52,344	62,605	68,954	78,240	80,739	83,850	87,092
2	12,284	27,586	45,996	55,151	62,391	66,517	71,273	73,217	77,198	
3	12,044	37,260	46,688	63,884	72,817	76,588	80,252	83,027		
4	7,823	22,064	31,908	43,897	55,384	56,389	57,756			
5	9,678	31,553	48,081	71,674	87,298	110,797				
6	8,520	20,684	54,684	68,463	79,892					
7	14,529	56,134	77,182	97,250						
8	10,190	28,526	42,014							
9	14,092	37,905								
10	9,823									

D1	D2	D3	D4	D5	D6	D7	D8	D9
2.931	1.545	1.321	1.183	1.114	1.071	1.031	1.046	1.039
E1	E2	E3	E4	E5	E6	E7	E8	E9
2.931	4.530	5.984	7.078	7.884	8.444	8.709	9.110	9.462

Chain Ladder factor D_t is determined as the ratio of the sum of payments associated with development year t+1, divided by the sum of payments associated with development year t.

Eg 2.931 = (25,103 + 27,586 +…+37,905)/(8,691+12,284 + …+14,092)

The E_t factors represent cumulative chain-ladder factors

4.530 = 1.545 x 2.931
5.984 = 1.321 x 4.530, and so on.

The chain-ladder central estimate is then determined based on the assessed development factors in the usual way. For example, across accident year 10:

Cumulative payments to the end of development period 2 = 2.931 x 9,823 = 28,794
Cumulative payments to the end of development period 3 = 1,545 x 28,794 = 44,499; and so on.

This process gives rise to the following projection for cumulative payments:

	1	2	3	4	5	6	7	8	9	10
1	8,691	25,103	38,124	52,344	62,605	68,954	78,240	80,739	83,850	87,092
2	12,284	27,586	45,996	55,151	62,391	66,517	71,273	73,217	77,198	80,182
3	12,044	37,260	46,688	63,884	72,817	76,588	80,252	83,027	86,852	90,209
4	7,823	22,064	31,908	43,897	55,384	56,389	57,756	59,570	62,314	64,723
5	9,678	31,553	48,081	71,674	87,298	110,797	118,669	122,397	128,035	132,985
6	8,520	20,684	54,684	68,463	79,892	88,984	95,307	98,300	102,829	106,804
7	14,529	56,134	77,182	97,250	115,028	128,119	137,222	141,532	148,052	153,776
8	10,190	28,526	42,014	55,502	65,648	73,119	78,315	80,775	84,496	87,762
9	14,092	37,905	58,580	77,385	91,532	101,949	109,192	112,622	117,810	122,365
10	9,823	28,794	44,499	58,784	69,530	77,443	82,945	85,551	89,492	92,951

A pseudo cumulative 'central estimate' triangle is derived by assuming the last development period estimate of cumulative payments reflects the central estimate value, and working backwards based on the chain ladder factors. The resulting 'pseudo' cumulative central triangle is:

1	9,204	26,978	41,694	55,078	65,147	72,561	77,716	80,158	83,850	87,092
2	8,474	24,838	38,386	50,708	59,978	66,804	71,551	73,798	77,198	
3	9,534	27,944	43,186	57,049	67,479	75,158	80,498	83,027		
4	6,840	20,049	30,985	40,932	48,415	53,924	57,756			
5	14,054	41,195	63,664	84,102	99,476	110,797				
6	11,287	33,085	51,130	67,544	79,892					
7	16,251	47,635	73,617	97,250						
8	9,275	27,186	42,014							
9	12,932	37,905								
10	9,823									

For instance, across accident year 1 from the left:

'Pseudo' cumulative payments to the end of development period 1 = 87,092 / 9.462 = 9,204
'Pseudo' cumulative payments to the end of development period 2 = 9,204 x 2.931 = 26,978; and so on

By taking differences between successive cumulative values, the 'pseudo' cumulative central triangle can readily be transformed into a pseudo incremental central payment triangle:

1	9,204	17,774	14,715	13,384	10,069	7,414	5,156	2,441	3,693	3,242
2	8,474	16,364	13,548	12,323	9,270	6,826	4,747	2,247	3,400	
3	9,534	18,411	15,242	13,863	10,429	7,679	5,340	2,529		
4	6,840	13,209	10,936	9,947	7,483	5,510	3,831			
5	14,054	27,141	22,469	20,437	15,375	11,321				
6	11,287	21,797	18,046	16,414	12,348					
7	16,251	31,384	25,982	23,632						
8	9,275	17,911	14,828							
9	12,932	24,973								
10	9,823									

For instance, across accident year 1, pseudo-triangle incremental payments in development period 2 are determined as 26,978 – 9,204 = 17,774.

If $C_{i,,j}$ represents actual incremental payments for accident year i and development year j, and $C_{i,j}^*$ represents the 'corresponding central ('pseudo') incremental payments; then if we define $\varepsilon_{i,j}$ values such that $C_{i,j} = C_{i,j}^* + \varepsilon_{i,j}$; the $\varepsilon_{i,j}$ values can be considered as observed residuals that are indicative of the variability associated with the chain-ladder process generating claim payments.

The triangle of $\varepsilon_{i,j}$ residual values is set out below

1	-513	-1,362	-1,694	836	192	-1,065	4,130	58	-582	0
2	3,810	-1,062	4,862	-3,167	-2,030	-2,700	10	-304	582	
3	2,511	6,805	-5,814	3,333	-1,497	-3,908	-1,676	246		
4	983	1,032	-1,092	2,043	4,004	-4,505	-2,464			
5	-4,376	-5,265	-5,942	3,155	250	12,178				
6	-2,767	-9,634	15,954	-2,635	-918					
7	-1,723	10,222	-4,934	-3,565						
8	915	425	-1,340							
9	1,161	-1,161								
10	0									

For example $\varepsilon_{1,1} = 8{,}691 – 9{,}204 = -513$

$\varepsilon_{1,2} = 16{,}412 – 17{,}774 = -1{,}362$ and so on.

The estimate of the standard deviation of the residuals associated with development year j is determined as

$$\sigma_j = \frac{1}{n-j} \sum_{i=1}^{\min(n-1,n-j+1)} e_{i,j}^2$$

where n = the total number of development periods (here n = 10)

For example, the estimate of the standard deviation of the residuals for development year 1 is

$1/9 \times \{(-513)^2 + (3{,}810)^2 + \quad (0)^2\} = 2{,}450$

Similarly, the estimate of the standard deviation of the residuals for development year 2 is 5,883.

Standardised residuals are defined as $\varepsilon_{i,j}^{sd} = \frac{\varepsilon_{i,j}}{\sigma_j}$

The triangle of $\varepsilon_{i,j}^{sd}$ values is set out below

	1	2	3	4	5	6	7	8	9	10
1	-0.209	-0.232	-0.231	0.275	0.089	-0.154	1.405	0.209	-0.707	0.000
2	1.555	-0.180	0.662	-1.040	-0.940	-0.389	0.003	-1.088	0.707	
3	1.025	1.157	-0.792	1.095	-0.693	-0.564	-0.570	0.879		
4	0.401	0.175	-0.149	0.671	1.853	-0.650	-0.838			
5	-1.786	-0.895	-0.809	1.036	0.116	1.756				
6	-1.130	-1.638	2.173	-0.865	-0.425					
7	-0.703	1.737	-0.672	-1.171						
8	0.374	0.072	-0.183							
9	0.474	-0.197								
10	0.000									

For instance, $\varepsilon_{1,1}^{sd} = -513/2{,}450 = -0.209$

Bootstrapping proceeds by generating a series of 'pseudo' triangles of incremental payments, where the payment for accident year i development year j is $C_{i,j}^*$ plus a residual term. The residual term is generated by pooling the standardised residuals and resampling them with replacement. Each time that a resampling of residuals takes place, the residuals are 'unstandardised' by multiplying by σ_j.

For example, for one bootstrap loop, the sample of resampled residuals might be

	1	2	3	4	5	6	7	8	9	10
1	0.374	-0.672	0.072	1.555	0.474	0.662	0.474	-0.180	1.405	0.474
2	0.116	0.209	-0.570	0.401	-0.231	0.275	-0.180	-0.865	0.116	
3	- 0.838	-0.389	-0.183	-0.838	-0.703	-0.149	0.662	-0.231		
4	- 0.672	-0.570	-0.149	0.275	0.474	1.737	0.209			
5	- 0.707	1.555	-1.088	-0.197	1.157	-0.838				
6	1.756	-0.564	-0.183	0.116	-0.149					
7	- 0.672	0.662	0.662	1.405						
8	- 1.786	-0.838	1.853							
9	- 0.232	-0.809								
10	- 0.183									

Because the re-sampling is with replacement, it is possible for a particular resampled residual to occur more than once. For instance, the value –0.183 appears twice.

The corresponding 'unstandardised' triangle of residuals is:

	1	2	3	4	5	6	7	8	9	10
1	915	-3,953	531	4,735	1,023	4,591	1,393	- 50	1,155	-
2	283	1,231	- 4,185	1,222	- 498	1,904	- 531	- 242	95	
3	- 2,053	-2,291	- 1,340	- 2,552	- 1,519	- 1,032	1,947	- 65		
4	- 1,646	-3,352	- 1,092	836	1,023	12,046	615			
5	- 1,732	9,150	- 7,990	- 601	2,499	- 5,811				
6	4,303	-3,316	- 1,340	352	- 321					
7	- 1,646	3,895	4,862	4,277						
8	- 4,376	-4,931	13,610							
9	- 567	-4,760								
10	- 447									

For example the resampled residual for accident year 1, development period 1 is
0.374 x 2,450 = 915

The resampled residual triangle is added to the pseudo incremental central payment triangle to arrive at a 'pseudo' resampled triangle. The pseudo triangle associated with this bootstrap loop is:

	1	2	3	4	5	6	7	8	9	10
1	10,119	13,822	15,246	18,120	11,092	12,005	6,549	2,391	4,848	3,242
2	8,757	17,595	9,363	13,544	8,772	8,730	4,216	2,006	3,495	
3	7,480	16,120	13,902	11,312	8,910	6,648	7,287	2,464		
4	5,194	9,857	9,843	10,783	8,506	17,556	4,447			
5	12,322	36,291	14,479	19,837	17,874	5,510				
6	15,590	18,481	16,705	16,766	12,026					
7	14,606	35,279	30,844	27,909						
8	4,899	12,980	28,438							
9	12,365	20,213								
10	9,376									

For example, the entry for accident year 1, and development year 1 is 9,204 + 915 = 10,119

Based on the standard chain-ladder methodology, a set of chain-ladder factors can be derived from the pseudo and applied to give a 'pseudo' central estimate projection. The cumulative 'pseudo' triangle, and the associated chain ladder factors corresponding to the example bootstrap loop triangle are shown below:

10,119	23,941	39,187	57,306	68,399	80,403	86,952	89,343	94,191	97,432
8,757	26,352	35,715	49,260	58,031	66,761	70,977	72,982	76,477	
7,480	23,600	37,502	48,813	57,723	64,371	71,658	74,122		
5,194	15,051	24,894	35,677	44,183	61,740	66,186			
12,322	48,613	63,092	82,929	100,802	106,312				
15,590	34,071	50,776	67,542	79,568					
14,606	49,885	80,729	108,638						
4,899	17,879	46,317							
12,365	32,577								
9,376									

D1	D2	D3	D4	D5	D6	D7	D8	D9
2.978	1.580	1.356	1.197	1.153	1.082	1.030	1.051	1.034

The corresponding central estimate projection for the example bootstrap loop is set out below, along with the associated central estimate of the outstanding claims liability.

	1	2	3	4	5	6	7	8	9	10	
1	10,119	23,941	39,187	57,306	68,399	80,403	86,952	89,343	94,191	97,432	0
2	8,757	26,352	35,715	49,260	58,031	66,761	70,977	72,982	76,477	79,109	2,632
3	7,480	23,600	37,502	48,813	57,723	64,371	71,658	74,122	77,932	80,614	6,491
4	5,194	15,051	24,894	35,677	44,183	61,740	66,186	68,164	71,667	74,134	7,947
5	12,322	48,613	63,092	82,929	100,802	106,312	115,065	118,503	124,593	128,881	22,569
6	15,590	34,071	50,776	67,542	79,568	91,764	99,319	102,287	107,544	111,245	31,676
7	14,606	49,885	80,729	108,638	130,008	149,934	162,278	167,127	175,717	181,764	73,126
8	4,899	17,879	46,317	62,822	75,180	86,703	93,841	96,645	101,612	105,109	58,792
9	12,365	32,577	51,469	69,810	83,542	96,346	104,278	107,394	112,914	116,800	84,222
10	9,376	27,921	44,112	59,831	71,600	82,574	89,372	92,043	96,773	100,104	90,728
											378,183

Process variability is then modelled around the bootstrap central estimate. At this stage of the simulation, the appropriate process distribution should be used if it is known. Here, process variability is modelled in a similar way to which the pseudo history triangles are created. The standardised residuals are resampled with replacement, and are then 'unstandardised'

An example resampled standardised set of residuals is set out below:

	1	2	3	4	5	6	7	8	9	10
1										
2										0.209
3									-0.149	1.756
4								-0.838	-0.707	0.474
5							-1.638	1.025	0.671	-0.650
6						-0.425	-1.786	-1.171	-0.180	-0.231
7					-1.130	-0.693	0.003	-0.425	1.025	-0.232
8				-0.792	1.756	-0.154	0.116	-1.638	-0.149	-0.209
9			-0.838	-0.895	-1.638	1.555	0.879	-0.838	-1.638	2.173
10		-0.940	-0.149	0.662	0.401	-0.231	-0.209	0.879	-0.564	1.737

The corresponding 'unstandardised' residuals are set out below. For example, the unstandardised residual corresponding to –0.940 in accident year 10 development year 2 is –0.940 x 5,883 = -4,137.

The total of the process variability residuals for each accident year is shown to the right of the triangle, with the grand total of –17,265 shown at the bottom of the column.

	1	2	3	4	5	6	7	8	9	10	
1											
2										-	-
3									1,445	-	1,445
4								-50	-126	-	-177
5							-2,379	-59	582	-	-1,856
6						-1,251	1,099	-158	1,787	-	1,477
7					-3,538	-1,368	4,130	393	-736	-	-1,119
8				270	-426	4,903	-537	-59	-582	-	3,570
9			-5,164	-3,167	3,754	-1,600	-5,253	246	900	-	-10,285
10		-4,137	-7,990	3,155	-1,933	-	-452	-242	1,279	-	-10,320
											-17,265

The value of the outstanding claims liability associated with this bootstrap cycle is 378,183 – 17,265 = 360,918

This cycle is repeated many times to build up a picture of the distribution of possible claims outcomes as assessed by this method.

Appendix D

Overview of Modelling Zehnwirth's PTF Family with the ICRFS-Plus Package

This appendix provides a brief introduction to how estimates of the outstanding claims liability are made using Zehnwirth's probabilistic trend family ('PTF') model, as it is implemented in the statistical reserving tool ICRFS-Plus^{TM}. ('ICRFS'). The appendix is aimed at those unfamiliar with the model, or the ICRFS reserving tool. For a definitive description, the interested reader is referred to Barnett and Zehnwirth (2000), and the ICRFS manuals.

The PTF models assume that incremental claims cost follows heteroscedastic lognormal distributions that follow a piecewise loglinear structure.

i.e.

$$y(w,d) = \alpha_w + \sum_{j=1}^{d} \gamma_j + \sum_{t=2}^{w+d} \iota_t + \varepsilon$$

Where $p(w,d)$ Denotes the incremental cost in the loss development array corresponding to accident period w and development period d

$y(w,d)$ = Log($p(w,d)$)

$w+d$ = Payment period

α_w = A parameter capturing the effect of the relative claims level associated with a given accident period w (typically an exposure measure)

γ_j = A parameter capturing the effect of a change in trend in the claims level across development periods at a given development period j.

ι_t = A parameter capturing the effect of a change in trend across transaction periods at a given transaction period t.

ε = The term capturing the modelled variability about the trends. It has a normal distribution with mean 0 and variance σ^2 (after adjustment to remove heteroscedasticity)

Modelling with the PTF family involves fitting a piecewise linear set of normal regression equations to the logarithms of the observations.

ICRFS provides a range of diagnostics which determine whether a given model provides a reasonable fit to the claims data, that would imply it is suitable for projecting the run-off of the outstanding claims liability.

Typically, data would be inflation adjusted prior to commencing the modelling, and exposure information would be entered directly as a valuation input. Modelling progresses through an iterative cycle: of

- *Model identification* (selecting where you think parameters indicating trend changes may be warranted)
- *Model estimation* (where ICRFS calculates model parameters consistent with your selections)
- *Testing* model assumptions by:
  - Reviewing residual plots to check that the residuals around the model are approximately normally distributed and of equal variability; and that no systematic pattern exists.
  - Reviewing the statistical significance of parameters (and removing those that are nor statistically significant)
  - Reviewing goodness of fit measures (and considering parsimony)
- Performing *validation* and stability analyses

Once a satisfactory model has been identified, a separate functionality that is built into ICRFS called PALD can be used to determine the probabilistic distribution of possible claims outcomes that is consistent with the model.

Modelling Process for the Illustrations in this Paper

In practice, there should be substantial input to model construction. For example, based on knowledge of the business and the environment, it is likely that the actuary will have an expectation regarding points at which trend changes will be identified. It is also important that the final model is sensible from a qualitative perspective. (For instance a model that did not include run-off in the far tail would not be accepted, no matter how well it fit the data).

However, for the purpose of illustration in this paper, modelling was performed more mechanically, using a functionality that is built into the program.

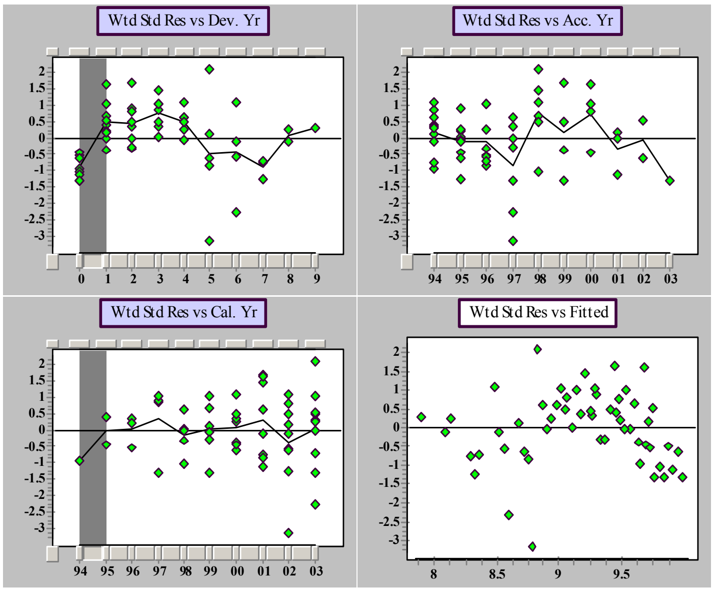
Model structure was determined by backward elimination of insignificant parameters, with significance determined by T-ratios using the in-built ICRFS model optimisation method. First, insignificant changes in trends and levels were removed. Insignificant trends were then set to zero, and finally smoothing was done on accident year parameters.

Reasons for adopting this approach, notwithstanding that it is open to sound theoretical objections, and some further comments on it are set out in the main body of the paper.

Modelling Process for general use of ICRFS to fit PTF models

Model Identification

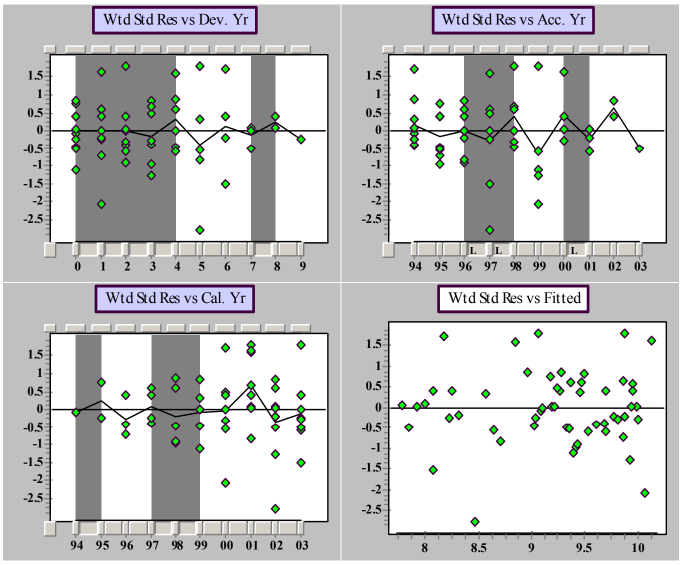
Standardised residual plots across development, accident and payment periods are examined visually to identify any non-randomness. An equivalent plot charting the standardised residuals by size of the fitted value is also available for review. In the example below, the fit is unsatisfactory. For example across development year, residuals are predominantly positive at early development periods and negative toward later development periods.



Based on where trends in residual patterns appear to start, and any knowledge of the underlying business, periods are identified, at which additional parameters could be considered.

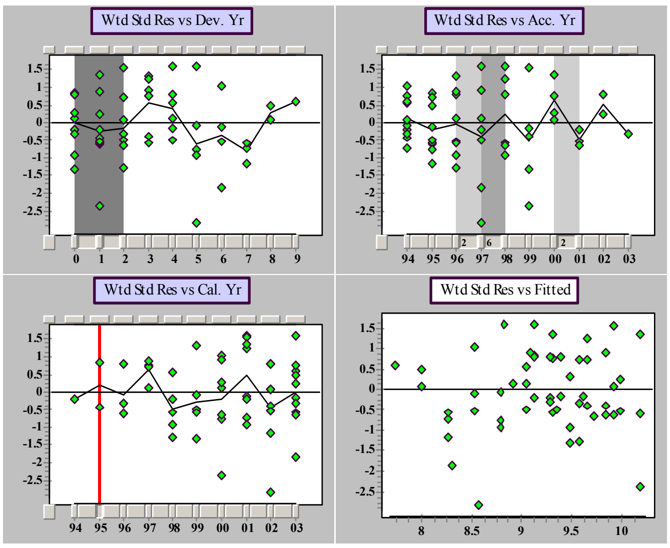
Model Estimation

ICRFS calculates parameter estimates based on the user's selection for where they think parameters should be considered, and replots the residuals. An example is shown below of what the residual plots might look like after this is done. Visually, the plots appear more satisfactory, with far fewer departures from the assumption that residuals are independent observations of a standard normal random variable.



At this point, some of the incorporated parameters may not be statistically significant. ICRFS provides a series of T-test results that check for statistical significance. Generally, parameters would be removed and the model re-estimated, if these tests indicated non-significance.

A screen shot showing how the residual plots might look after this process has been completed is set out below. At face value, the plots look less satisfactory than the previous plots, but the underlying model is better in the sense that it does not arrive at a 'better looking' fit by including non-significant parameters.



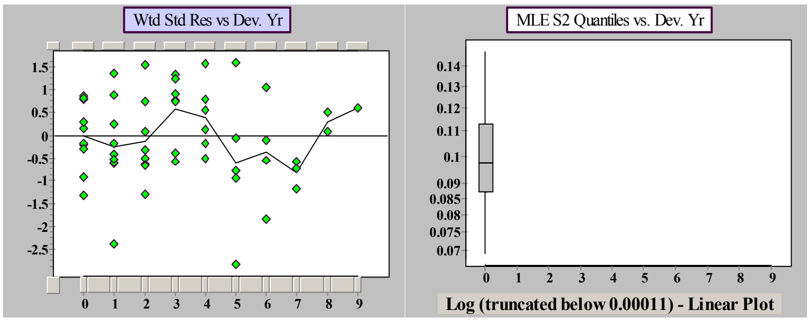
Heteroscedasticity Adjustment

A key assumption in the model is that the (log) residuals are independent, identically distributed normal random variables. Among other things, this means that the residuals should have constant variability around the fitted trends.

The ICRFS interface includes graphical tools that allow heteroscedasticity that would violate this assumption to be identified and adjusted for. The constant variation assumption is often violated because payments display a higher degree of volatility in the tail. Adjustment parameters can be included in the model to allow for changes in the variance. With the ICRFS interface, it is easiest to make this adjustment by development period.

The process by which this is done is similar to that described for the other model parameters. Screenshots set out below show some of the visual aids in ICRFS that help with this process. The residuals by development period are shown on the left. The aim is to identify periods at which there is a change in residual spread. The right hand chart shows the estimated variance consistent with your model.

At this point the model assumes constant variance across development periods. Examination of the residual plot doesn't provide obvious evidence of heteroscedasticity, and in this instance, adjustment for heteroscedasticity doesn't seem to be warranted.



Competing well fitting models (Parsimony)

In a choice among competing models, other things being equal the simplest is preferable.

Under-parameterisation risks not capturing important features of the underlying claims process. Over-parameterisation, on the other hand risks fitting what is actually just a random component of experience. An over-parameterised model will lead to high prediction errors and is therefore undesirable.

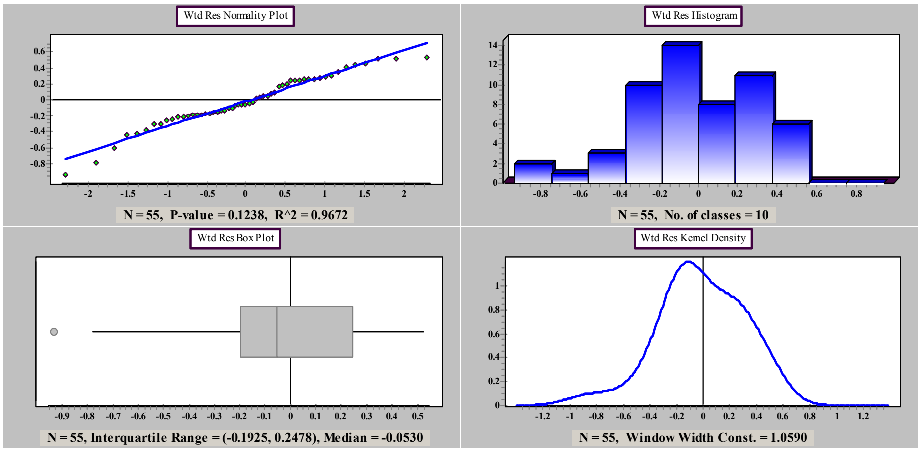
ICRFS provides a series of statistics that can be used to allow parsimony considerations to be taken into account when selecting a model. Some of these include:

- Akaike's information criterion (a statistic that increases with the number of fitted parameters, and decreases with log likelihood)
- Baye's information criterion (similar to Akaike's information criterion, but with a different way of setting the parameter penalty)
- SSPE – sum of squares prediction error (which looks at 'one-step ahead' prediction errors)

Generally, low values for SSPE, AIC, and BIC are preferable to high ones. However, other aspects of model testing and fitting such as the significance of selected parameters, and the fit of the distributional assumptions mean that these criteria are not the be all and end all when deciding on a final model.

Normality Test

The model assumes that the standardised residuals have a normal distribution. This assumption can be tested. A screenshot of the component of ICRFS that illustrates this check is provided below. In this example, the likelihood that a claims history that followed the assumptions of the model will give rise to a pattern of standardised residuals less normal than what is observed is about 1 in 8. The fitted model is not great based on this test, but would not be rejected at a 5% significance level.



Qualitative Overview

An important step in the modelling process is that it makes sense in terms of known features of the underlying business. An apparently well fitting model would not be accepted if it failed a qualitative review of sensibility.

Model Validation

This step analyses the change in the level and variability of the central estimate if the last diagonal in the triangle is removed, and parameters are re-estimated based on the remainder of the data triangle. The projection is then performed based on the new parameter estimates. This step could be repeated by removing the diagonals one by one (until the remaining data is too sparse for the results to be meaningful).

Generally speaking, a model that displays high stability would be preferred over a model that produces high variability under the validation procedure.

Predictive distribution associated with the fitted model

ICRFS includes functionality it terms 'PALD' that can be used to assess the predictive distribution of claims outcomes consistent with the selected model. PALD samples from the given model, which defines a joint distribution of correlated lower triangle cells, each lognormally distributed with its own mean, standard deviation and correlation with other cells. Each individual sample is a complete lower triangle, which in turn yields an outstanding claims liability total. Repeated sampling of values of outstanding claims totals forms the distribution of interest.

ICRFS allows selective inclusion of parameter uncertainty in forecasting and estimation. Parameter uncertainty is usually included in this exercise because it is the full predictive distribution that is of practical interest in setting reserves.

The result presentation described as the 'kernel' would usually be the preferred distribution to be used as the assessed predictive distribution associated with the fitted model.

Appendix E

Detail of the Assessed Claims Generating Processes

Description of Claims Generating Process 1

The first example assesses a claim triangle generated from a process that aims to reflect many of the assumptions that underlie the chain-ladder assessment method.

The process is defined as follows:

Let $C_{i,j}$ = the incremental payment arising from accident year i, and paid in development year j.

Let $D_{i,j}$ = the cumulative payments arising up to the end of development year j from accident year i (ie $D_{i,j} = C_{i,1} + C_{i,2} + \ldots + C_{i,j}$).

Let $F_{i,j}$ = the chain ladder factor that when multiplied by Di,j gives the expected value of $D_{i,j+1}$.

Let $E[D_{i,j+1}|D_{i,j}]$ equal the expected value of $D_{i,j+1}$ given the observation $D_{i,j}$.

$C_{i,1}$ = $D_{i,1}$ is a random drawing from a lognormal distribution with mean 30,000 and standard deviation 20,000.

$E[D_{i,j+1}|D_{i,j}] = F_{i,j} \text{ x } D_{i,j}$

and

$D_{i,j+1}|D_{i,j} = F_{i,j} \text{ x } D_{i,j} + \varepsilon_{i,j+1}$

So $C_{i,j+1} = D_{i,j+1} - D_{i,j}$

I have set the 'error' term $\varepsilon_{i,j+1}$ associated with $D_{i,j+1}$ as a random observation from a translated lognormal distribution. Before translation, I have set the lognormal distribution to have:

- Mean equal to 80%[14] of $E[C_{i,j+1}|D_{i,j}]$ and
- Standard deviation equal to 60% x $E[C_{i,j+1}|D_{i,j}]$.

This distribution is translated so the expected value of $\varepsilon_{i,j+1}$ is zero. The aim in setting the error term is that it has:

- Constant coefficient of variation over all development periods where a claim projection will be required, and
- Standard deviation that is proportional to the expected value (ie proportional to $E[C_{i,j+1}|D_{i,j}]$.

[14] Readers who try to determine the purpose of this factor should note that in my view it neither adds to nor detracts from the analysis. It is just a remanent from modelling work that I did as exploration associated with this paper.

The true underlying chain-ladder factors $F_{i,j}$ that I have applied are:

Process 1 – Chain Ladder Factors									
Development Period (j)	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
$F_{i,j}$	3.000	1.650	1.300	1.200	1.080	1.060	1.040	1.020	1.005

A feature of this process is that observations at later development periods are contingent on the cumulative experience for the accident year to the prior development period.

Example Incremental Triangle Generated by Process 1

An example of a development triangle of incremental claim payments, the associated cumulative claim payment triangle, and triangle of paid claim development factors for a single simulation generated from the first process is set out below.

Example Incremental triangle Generated By Process 1

	Development Period									
Accident Period	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	22,273	20,778	16,189	19,297	13,508	18,324	7,221	1,785	1,021	499
2	24,963	65,139	26,986	47,491	24,474	5,968	6,770	5,660	2,362	
3	15,495	49,186	44,560	19,004	20,602	14,620	6,412	6,453		
4	14,965	14,346	10,400	13,644	6,803	3,912	4,228			
5	24,794	34,038	36,031	45,033	20,385	12,014				
6	10,497	40,382	26,287	11,640	18,634					
7	31,723	51,537	26,320	25,016						
8	23,096	33,450	20,109							
9	28,574	23,687								
10	16,552									

Cumulative triangle

	Development Period									
Accident Period	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	22,273	43,051	59,240	78,537	92,045	110,369	117,589	119,374	120,395	120,894
2	24,963	90,102	117,087	164,579	189,053	195,021	201,791	207,451	209,814	
3	15,495	64,681	109,241	128,245	148,847	163,467	169,879	176,332		
4	14,965	29,311	39,711	53,355	60,158	64,069	68,297			
5	24,794	58,832	94,863	139,896	160,281	172,295				
6	10,497	50,879	77,166	88,806	107,440					
7	31,723	83,260	109,580	134,595						
8	23,096	56,546	76,656							
9	28,574	52,261								
10	16,552									

Chain Ladder Factors

	Development Period									
Accident Period	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1		1.933	1.376	1.326	1.172	1.199	1.065	1.015	1.009	1.004
2		3.609	1.300	1.406	1.149	1.032	1.035	1.028	1.011	
3		4.174	1.689	1.174	1.161	1.098	1.039	1.038		
4		1.959	1.355	1.344	1.127	1.065	1.066			
5		2.373	1.612	1.475	1.146	1.075				
6		4.847	1.517	1.151	1.210					
7		2.625	1.316	1.228						
8		2.448	1.356							
9		1.829								
10										
		2.693	1.434	1.298	1.160	1.084	1.046	1.028	1.010	1.004

The claims generating process that I have set comes across as being quite complicated. The following example relating to accident period 5 in the above set of triangles should help to explain it.

> $C_{5,1} = D_{5,1}$ (development period 1 in the incremental and cumulative triangle) = 24,794. This is a random observation from a log-normal distribution with mean 20,000 and standard deviation 10,000.
>
> $C_{5,2}$ is 34,038 (the corresponding value in the second triangle is $D_{5,2}$ is 58,832 = 34,038 +24,794)
> This value is generated as follows:

- The expected value for $D_{5,2}$ given the observation of $D_{5,1}$ is 3.00 x 24,794 = 74,382.
- The corresponding expected value for $C_{5,2}$ is 74,382 – 24,794 = 49,588.
- $C_{5,2}$ is 49,588 + a random element
- The random element associated with $C_{5,2}$ is a random observation from a lognormal distribution with
  - mean 80% x 49,588 = 39,670
  - standard deviation 60% x 49,588 = 29,753

  translated by subtracting 39,670 so that the expected value of the random element is zero.

For any given claims history triangle, the process completely describes the probabilistic distribution of possible claims outcomes. Though it would be very complicated to describe analytically, it is straightforward to determine the outcome distribution by simulation.

The distribution of possible claims outcomes is dependent on the last diagonal of the simulated claims history triangle. Consequently, each example of a claims history that is generated based by this simulation is associated with its own 'true' distribution of possible outcomes. This is an unavoidable consequence of constructing a claims process that follows chain-ladder assumptions. In contrast, the second process described below has a 'true' distribution of possible claims outcomes that is independent of the particular outcome of a given simulation.

Description of Claims Generating Process 2

The second process generates payments as a random selection from a series of lognormal distributions. The distributions are a function of development year, but are independent of accident year. Outcomes for a given development year are modelled as independent from other development years.

The process, and an example incremental and cumulative triangle are described below. I have also illustrated the individual incurred claim development factors associated with the example.

Process 2

If X is the lognormal random variable representing the claims outcome for a particular accident and development year, then, if μ and σ refer to the mean and the standard deviation of the normally distributed random variable Y such that $Y = \log X$

Then, if $C_{i,j}$ = the incremental payment amounts arising from accident year i, and paid in development year j

and $Y_{i,j} = \log (C_{i,j})$

Then under the claims generating process I have defined, $Y_{i,j}$ values are independent random variables drawn from a normal distribution with mean μ_j and standard deviation σ_j, where these parameters have the values set out in the next table. The table also shows the mean, standard deviation, and coefficient of variation for the resultant lognormally distributed $C_{i,j}$ variables.

Process 2					
Development Period j	**$C_{i,j}$ Mean**	**$C_{i,j}$ Std Deviation**	**$C_{i,j}$ CV[15] %**	μ_j	σ_j
0	12,000	3,600	30	9.3496	0.29356
1	24,000	8,400	35	10.0820	0.33994
2	20,000	9,000	45	9.8113	0.42942
3	16,000	8,800	55	9.5482	0.51409
4	10,000	7,000	70	9.0110	0.63149
5	8,000	6,400	80	8.7398	0.70335
6	5,000	4,500	90	8.2205	0.77208
7	4,000	4,000	100	7.9475	0.83255
8	3,000	3,300	110	7.6099	0.89050
9	2,000	2,400	120	7.1549	0.94446

A feature of this process is that, across an accident year, what has occurred in previous development periods carries no explanatory power. This is in sharp contrast to one of the chain-ladder assumptions, namely that what has occurred up until the immediately preceding development period for the accident year is the *only* explanatory variable that should be use to project claims run-off experience.

An example of a development triangle of incremental claim payments, the associated cumulative claim payment triangle, and triangle of paid claim development factors for a single simulation generated from the process is set out below.

[15] CV = coefficient of variation

Example Incremental triangle Generated By Process 2

Accident Period	Development Period 0	1	2	3	4	5	6	7	8	9
1	8,691	16,412	13,021	14,220	10,261	6,349	9,286	2,500	3,111	3,242
2	12,284	15,302	18,410	9,155	7,240	4,126	4,756	1,943	3,981	
3	12,044	25,216	9,428	17,196	8,932	3,771	3,665	2,774		
4	7,823	14,241	9,843	11,990	11,487	1,005	1,367			
5	9,678	21,875	16,528	23,593	15,624	23,499				
6	8,520	12,163	34,000	13,779	11,429					
7	14,529	41,605	21,048	20,067						
8	10,190	18,336	13,488							
9	14,092	23,813								
10	9,823									

Cumulative triangle

Accident Period	Development Period 0	1	2	3	4	5	6	7	8	9
1	8,691	25,103	38,124	52,344	62,605	68,954	78,240	80,739	83,850	87,092
2	12,284	27,586	45,996	55,151	62,391	66,517	71,273	73,217	77,198	
3	12,044	37,260	46,688	63,884	72,817	76,588	80,252	83,027		
4	7,823	22,064	31,908	43,897	55,384	56,389	57,756			
5	9,678	31,553	48,081	71,674	87,298	110,797				
6	8,520	20,684	54,684	68,463	79,892					
7	14,529	56,134	77,182	97,250						
8	10,190	28,526	42,014							
9	14,092	37,905								
10	9,823									
Chain Ladder Factors		2.931	1.545	1.321	1.183	1.114	1.071	1.031	1.046	1.039

Chain Ladder Factors

Accident Period	Development Period 0	1	2	3	4	5	6	7	8	9
1		2.888	1.519	1.373	1.196	1.101	1.135	1.032	1.039	1.039
2		2.246	1.667	1.199	1.131	1.066	1.072	1.027	1.054	
3		3.094	1.253	1.368	1.140	1.052	1.048	1.035		
4		2.820	1.446	1.376	1.262	1.018	1.024			
5		3.260	1.524	1.491	1.218	1.269				
6		2.428	2.644	1.252	1.167					
7		3.864	1.375	1.260						
8		2.799	1.473							
9		2.690								
10										
Avg		2.931	1.545	1.321	1.183	1.114	1.071	1.031	1.046	1.039

Process 1 Triangles

The 12 randomly generated triangles for Process 1 that were used to provide the illustrations set out in Section 3.4 and Appendix F are set out below.

	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	22,273	43,051	59,240	78,537	92,045	110,369	117,589	119,374	120,395	120,894
2	24,963	90,102	117,087	164,579	189,053	195,021	201,791	207,451	209,814	
3	15,495	64,681	109,241	128,245	148,847	163,467	169,879	176,332		
4	14,965	29,311	39,711	53,355	60,158	64,069	68,297			
5	24,794	58,832	94,863	139,896	160,281	172,295				
6	10,497	50,879	77,166	88,806	107,440					
7	31,723	83,260	109,580	134,595						
8	23,096	56,546	76,656							
9	28,574	52,261								
10	16,552									

	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	18,092	47,250	69,722	84,648	124,648	135,084	144,028	147,950	150,546	151,186
2	15,449	59,084	94,451	122,327	138,478	143,393	149,807	155,801	157,682	
3	17,013	39,613	76,848	115,979	147,358	152,761	161,981	172,712		
4	7,878	20,980	55,218	65,282	94,257	102,751	107,560			
5	20,739	71,403	96,482	122,426	139,730	144,700				
6	26,474	92,462	141,527	206,516	241,107					
7	14,138	28,616	43,491	54,332						
8	19,913	111,528	157,014							
9	14,258	43,118								
10	17,930									

	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	26,554	74,786	142,171	166,726	183,752	191,146	207,786	213,922	216,419	218,248
2	11,434	33,776	60,868	72,340	90,470	101,030	105,249	108,362	109,238	
3	8,384	18,549	26,902	31,591	34,856	36,747	37,581	39,787		
4	20,065	49,760	94,616	117,055	138,359	143,596	151,045			
5	21,070	71,520	114,100	146,384	175,917	187,410				
6	10,920	28,435	37,214	45,480	52,985					
7	23,461	61,365	101,737	143,908						
8	14,196	50,699	120,024							
9	13,052	57,991								
10	20,298									

	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	10,559	31,740	43,258	65,466	75,603	79,286	81,781	83,490	84,870	85,090
2	18,909	36,996	50,021	60,039	71,606	75,453	79,043	80,853	83,606	
3	8,031	15,172	25,482	29,608	31,973	34,358	38,068	38,853		
4	17,943	46,280	68,557	109,263	130,888	148,598	163,736			
5	23,517	50,262	91,002	114,765	155,718	163,634				
6	12,659	49,864	103,753	121,226	147,917					
7	34,778	73,242	114,034	130,968						
8	14,580	30,892	41,001							
9	15,209	31,251								
10	15,062									

	Development Period									
	0	1	2	3	4	5	6	7	8	9
Accident Period										
1	14,597	48,835	90,952	133,702	144,483	149,618	155,364	158,474	159,898	160,641
2	26,727	56,127	80,333	102,246	115,923	124,101	129,665	135,647	137,051	
3	12,741	28,799	51,340	66,864	75,298	77,781	81,663	85,297		
4	12,197	66,940	116,085	189,632	238,067	256,107	268,433			
5	36,379	106,431	211,090	250,988	372,767	397,362				
6	33,095	66,724	149,868	169,640	195,716					
7	40,445	111,841	177,024	243,329						
8	15,626	33,902	49,074							
9	8,905	23,925								
10	18,767									

	Development Period									
	0	1	2	3	4	5	6	7	8	9
Accident Period										
1	30,749	104,555	139,080	167,694	192,689	211,288	234,565	243,460	249,334	250,342
2	18,744	59,501	84,859	104,948	123,952	131,196	138,104	146,359	150,178	
3	23,811	45,603	66,784	83,474	96,641	109,340	119,395	123,575		
4	21,178	72,806	132,125	164,317	182,358	193,882	198,915			
5	17,629	79,708	131,514	164,120	177,588	185,545				
6	19,800	39,850	74,538	95,155	123,458					
7	27,280	74,116	167,166	198,330						
8	31,000	65,706	145,182							
9	26,647	52,680								
10	19,750									

	Development Period									
	0	1	2	3	4	5	6	7	8	9
Accident Period										
1	11,005	32,460	55,443	68,105	76,930	79,764	84,554	87,710	89,183	89,741
2	10,432	30,309	43,826	52,996	60,117	66,110	70,261	73,734	74,699	
3	19,228	61,920	84,437	118,352	145,302	154,077	164,585	169,942		
4	40,071	71,818	126,239	142,520	160,824	166,083	175,594			
5	8,329	27,434	36,556	48,689	59,806	62,913				
6	37,588	80,964	113,339	144,595	168,517					
7	11,999	35,752	45,883	73,890						
8	23,454	51,597	82,190							
9	27,396	60,946								
10	21,713									

	Development Period									
	0	1	2	3	4	5	6	7	8	9
Accident Period										
1	30,027	58,453	89,482	108,252	130,976	145,411	156,487	162,450	170,427	171,350
2	16,320	38,990	77,352	98,668	117,965	135,545	140,616	145,952	153,063	
3	24,984	70,812	98,899	145,326	159,692	169,778	182,444	188,610		
4	26,200	102,351	133,953	153,001	224,085	246,565	265,858			
5	16,310	32,113	57,183	67,751	84,668	88,957				
6	39,655	96,230	178,247	212,986	237,964					
7	23,570	80,923	133,350	153,207						
8	23,901	59,181	93,499							
9	9,206	22,467								
10	12,042									

	Development Period									
	0	1	2	3	4	5	6	7	8	9
Accident Period										
1	21,459	37,959	53,232	68,733	75,379	77,898	83,438	86,260	87,342	88,033
2	13,413	62,947	117,920	132,902	158,028	174,205	181,453	184,958	187,133	
3	10,030	20,093	26,076	35,911	40,141	42,145	43,347	44,274		
4	10,252	20,436	27,127	38,538	42,810	47,128	50,843			
5	10,869	40,349	58,783	89,550	104,979	108,563				
6	25,400	64,845	121,136	136,303	154,882					
7	27,057	50,276	68,630	85,204						
8	34,823	83,471	105,230							
9	7,668	22,226								
10	33,614									

	Development Period									
	0	1	2	3	4	5	6	7	8	9
Accident Period										
1	27,065	96,241	170,449	195,590	257,528	275,620	293,746	308,764	313,286	314,285
2	16,269	47,968	69,520	85,012	94,851	98,843	112,554	116,633	118,588	
3	15,453	32,787	47,746	58,844	73,465	80,528	87,157	91,957		
4	9,651	23,938	41,717	50,970	62,785	66,558	69,342			
5	26,489	60,088	80,783	98,872	114,270	128,292				
6	43,466	124,880	256,085	293,536	374,755					
7	18,846	65,799	95,494	108,400						
8	28,046	74,456	96,261							
9	12,896	46,984								
10	21,073									

	Development Period									
	0	1	2	3	4	5	6	7	8	9
Accident Period										
1	10,923	28,544	75,238	102,369	123,789	144,047	148,454	159,817	166,593	167,136
2	25,914	52,445	99,599	117,910	157,857	165,692	170,279	174,458	177,744	
3	8,645	16,953	43,953	53,800	63,236	66,684	70,244	73,782		
4	22,646	51,479	78,913	104,398	113,686	123,392	134,922			
5	15,299	29,336	74,028	111,802	138,914	149,549				
6	16,950	41,545	81,656	105,221	116,345					
7	13,331	25,252	37,354	45,241						
8	11,161	25,658	39,702							
9	41,814	99,348								
10	9,030									

	Development Period									
	0	1	2	3	4	5	6	7	8	9
Accident Period										
1	16,697	31,161	43,337	61,181	69,640	71,720	76,771	79,048	79,687	80,147
2	28,457	118,337	160,489	211,952	244,756	258,547	270,508	278,961	286,332	
3	14,933	41,507	66,642	82,306	97,349	114,176	131,567	138,149		
4	11,601	36,229	73,656	90,505	98,207	105,530	114,580			
5	20,654	53,272	82,907	110,969	137,609	163,694				
6	12,412	25,369	63,378	78,373	84,263					
7	14,490	41,016	57,118	79,141						
8	8,092	15,099	28,252							
9	11,336	45,482								
10	23,572									

Process 2 Triangles

The 12 randomly generated triangles for Process 2 that were used to provide the illustrations set out in Section 3.4 and Appendix F are set out below.

Accident Period	Development Period 0	1	2	3	4	5	6	7	8	9
1	8,691	16,412	13,021	14,220	10,261	6,349	9,286	2,500	3,111	3,242
2	12,284	15,302	18,410	9,155	7,240	4,126	4,756	1,943	3,981	
3	12,044	25,216	9,428	17,196	8,932	3,771	3,665	2,774		
4	7,823	14,241	9,843	11,990	11,487	1,005	1,367			
5	9,678	21,875	16,528	23,593	15,624	23,499				
6	8,520	12,163	34,000	13,779	11,429					
7	14,529	41,605	21,048	20,067						
8	10,190	18,336	13,488							
9	14,092	23,813								
10	9,823									

Accident Period	Development Period 0	1	2	3	4	5	6	7	8	9
1	14,830	17,616	13,948	6,997	11,450	13,232	4,892	2,727	1,633	2,488
2	10,734	19,125	30,631	12,025	8,334	4,446	6,637	2,725	1,335	
3	19,368	20,105	11,387	8,033	22,040	6,069	8,566	6,454		
4	12,030	15,352	9,834	23,183	6,326	11,400	2,804			
5	9,237	34,893	14,039	6,528	7,838	14,203				
6	10,626	22,307	10,266	8,500	12,045					
7	12,831	24,092	7,212	15,174						
8	14,359	68,890	9,929							
9	10,511	41,499								
10	9,528									

Accident Period	Development Period 0	1	2	3	4	5	6	7	8	9
1	12,060	19,940	20,094	15,194	5,845	5,018	7,790	907	871	2,248
2	16,418	22,623	11,792	12,253	6,070	1,994	10,876	3,510	2,677	
3	10,902	21,078	28,910	49,157	7,198	9,819	2,239	5,222		
4	10,841	29,304	20,521	8,500	6,019	3,540	2,256			
5	13,251	18,121	16,560	20,789	9,620	2,990				
6	13,501	27,358	11,883	18,234	2,645					
7	13,002	14,957	24,758	23,870						
8	10,302	26,168	18,826							
9	15,435	12,383								
10	10,291									

Accident Period	Development Period 0	1	2	3	4	5	6	7	8	9
1	13,612	19,344	7,103	9,533	19,869	3,440	2,178	2,685	645	715
2	9,902	19,349	10,711	9,546	8,242	3,484	5,448	3,598	1,374	
3	20,090	15,624	20,898	5,854	3,845	7,739	2,104	1,971		
4	12,810	16,678	10,323	5,464	6,058	15,909	2,336			
5	7,991	35,149	23,193	12,538	16,974	6,083				
6	9,454	26,243	12,472	9,666	4,760					
7	11,258	28,854	13,302	19,173						
8	13,530	32,737	18,648							
9	6,840	23,298								
10	14,080									

	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	17,561	28,235	18,938	19,186	9,787	3,065	19,422	8,809	3,095	7,514
2	14,987	22,528	26,876	12,298	31,167	9,378	616	3,663	14,436	
3	14,074	17,807	34,889	17,917	8,528	3,769	5,033	5,173		
4	10,216	18,081	14,930	17,924	15,503	5,721	5,062			
5	20,950	19,333	16,914	10,605	12,280	12,418				
6	12,121	15,402	17,422	20,108	5,689					
7	14,694	19,098	27,274	15,466						
8	13,515	18,280	29,821							
9	14,704	19,485								
10	12,785									

	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	10,879	14,332	15,936	5,442	12,177	4,769	8,843	2,532	663	1,324
2	19,587	26,117	11,705	16,743	8,046	6,750	3,613	734	1,420	
3	9,673	14,001	19,404	10,090	9,306	3,875	6,322	8,917		
4	7,955	53,969	10,229	10,565	7,671	5,050	2,712			
5	8,504	34,286	23,242	16,984	9,063	8,390				
6	13,869	20,078	21,559	14,966	11,811					
7	9,625	27,293	12,104	6,387						
8	13,108	25,484	13,544							
9	10,166	19,221								
10	8,217									

	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	9,729	23,602	13,692	25,965	13,568	8,826	5,007	1,030	2,418	238
2	20,569	27,654	18,498	26,611	7,193	3,573	3,065	2,972	274	
3	8,989	23,342	24,692	16,336	5,839	7,633	6,007	5,781		
4	12,382	27,269	18,620	7,788	20,688	2,424	5,923			
5	12,472	32,451	14,801	20,797	2,793	7,121				
6	9,354	24,548	26,600	20,705	3,741					
7	9,642	14,060	23,212	12,253						
8	20,358	30,039	21,419							
9	10,337	20,533								
10	10,604									

	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	10,809	34,594	10,186	12,933	14,660	28,840	6,899	9,682	2,881	6,759
2	12,025	23,215	15,868	13,782	4,138	8,214	4,406	4,021	7,302	
3	13,188	18,034	9,190	12,412	10,786	2,851	5,858	1,391		
4	9,221	18,480	18,948	8,832	11,855	8,115	3,473			
5	12,006	17,859	21,751	8,960	37,176	5,460				
6	15,797	20,155	21,105	31,340	9,881					
7	14,377	17,822	14,684	15,558						
8	7,932	24,303	18,601							
9	13,338	23,290								
10	18,772									

	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	17,729	27,176	14,977	19,398	9,430	15,021	2,857	1,814	2,159	1,970
2	17,894	13,936	11,889	13,079	12,172	16,060	7,668	4,791	509	
3	10,058	20,298	23,357	61,027	8,797	2,215	6,539	4,409		
4	10,101	22,174	25,766	13,657	8,581	6,242	5,030			
5	10,343	34,048	24,472	8,623	16,004	23,242				
6	12,057	26,230	17,197	15,701	6,792					
7	11,971	27,802	12,214	38,605						
8	7,968	10,627	9,633							
9	8,340	29,114								
10	7,433									

	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	12,949	25,650	31,926	18,842	7,384	6,553	1,431	1,271	3,866	1,208
2	14,304	34,948	36,874	13,256	3,525	20,893	8,369	3,725	8,554	
3	10,090	25,464	19,575	5,009	20,827	7,428	4,822	751		
4	9,247	15,118	12,967	16,666	2,194	8,523	8,473			
5	8,255	19,954	25,695	9,681	8,254	2,210				
6	16,059	43,970	15,926	29,047	25,999					
7	17,840	23,050	24,780	21,578						
8	16,238	17,379	5,632							
9	15,800	23,385								
10	9,627									

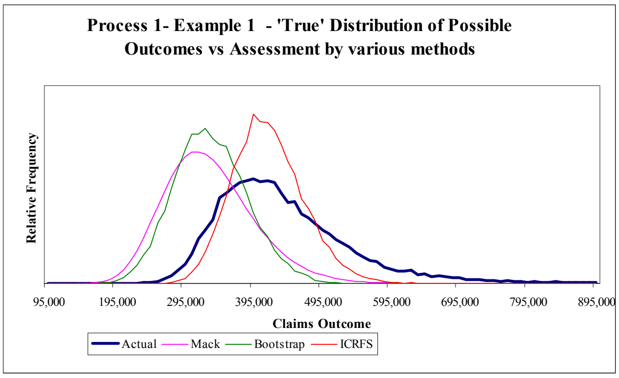
	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	12,504	24,492	16,782	11,715	3,347	12,430	3,936	4,663	2,397	1,642
2	7,985	30,371	13,655	10,190	16,897	12,737	1,875	1,880	4,062	
3	15,559	15,953	6,198	12,493	8,257	5,353	958	575		
4	16,795	35,166	36,280	13,522	16,619	8,902	1,875			
5	14,196	42,549	22,579	17,920	6,517	4,311				
6	13,136	22,438	13,509	19,771	2,239					
7	11,206	15,339	17,724	30,728						
8	9,496	10,452	22,319							
9	6,583	25,016								
10	10,174									

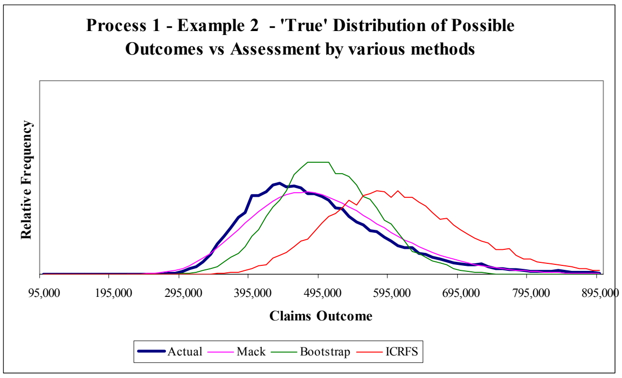
	Development Period									
Accident Period	0	1	2	3	4	5	6	7	8	9
1	16,120	27,101	18,251	12,432	5,468	5,792	8,755	2,678	770	2,773
2	9,083	20,766	7,292	3,993	28,380	3,170	1,597	8,411	3,061	
3	10,686	19,169	12,163	19,202	1,338	6,317	10,419	3,988		
4	9,039	25,622	15,423	12,368	7,278	6,671	2,694			
5	9,668	44,791	20,903	7,677	4,290	2,058				
6	7,041	25,951	9,530	29,095	10,223					
7	7,619	27,876	29,664	11,999						
8	9,112	24,742	18,685							
9	10,838	17,970								
10	9,150									

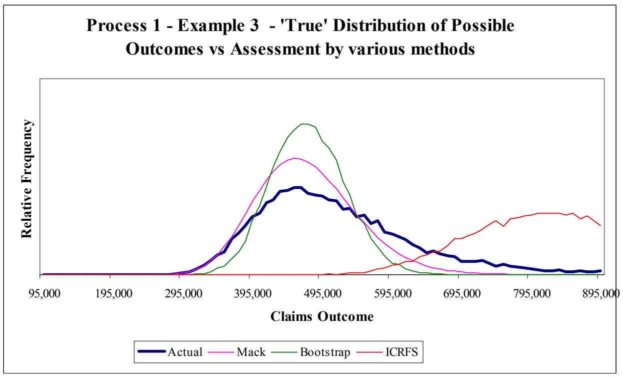
Appendix F

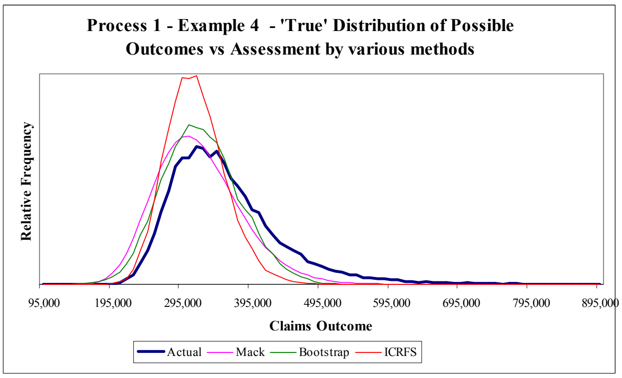
Comparison of ‘True’ Distribution of possible claims outcomes for the Artificial Claims Generating Processes with assessment by different quantitative methods

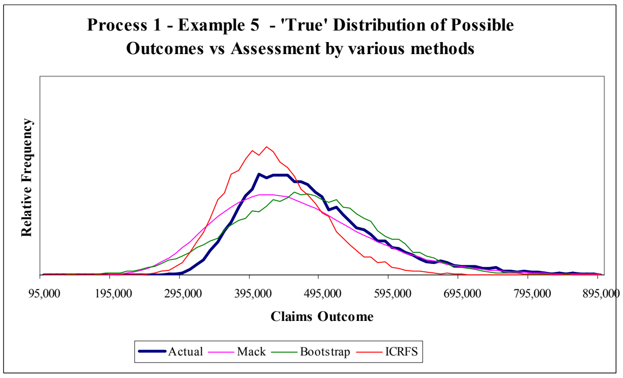
Process 1 – Chain-Ladder Process Followed

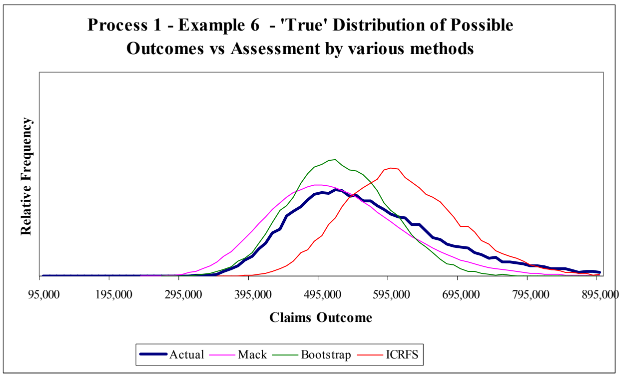


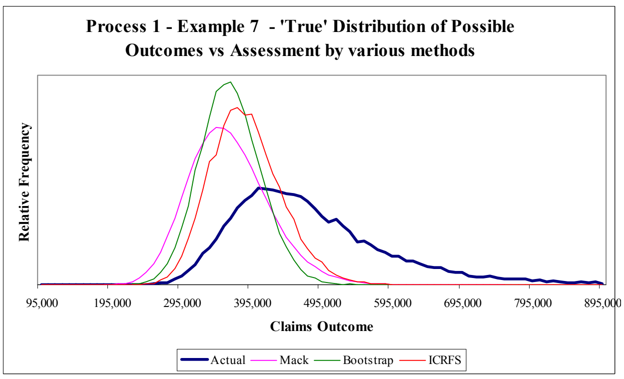


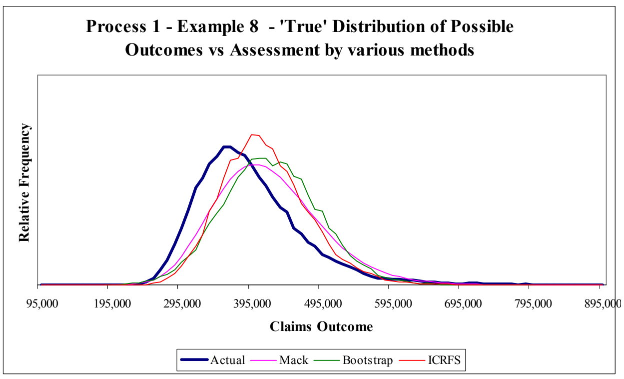


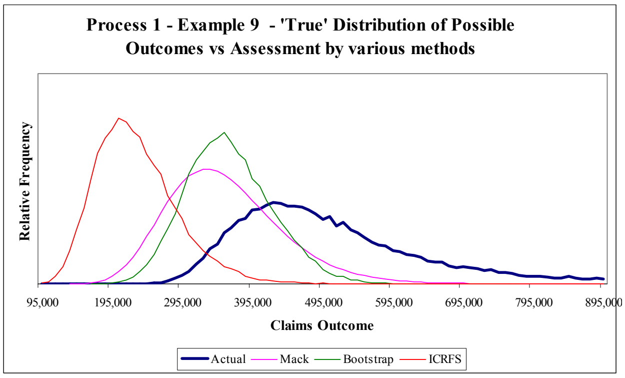


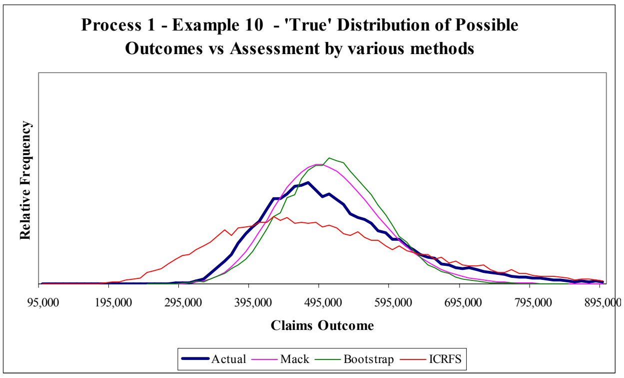


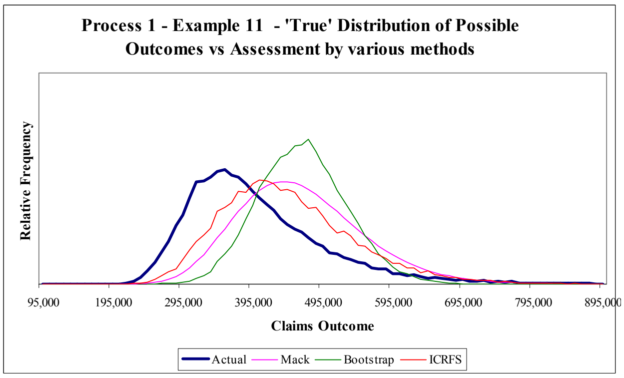


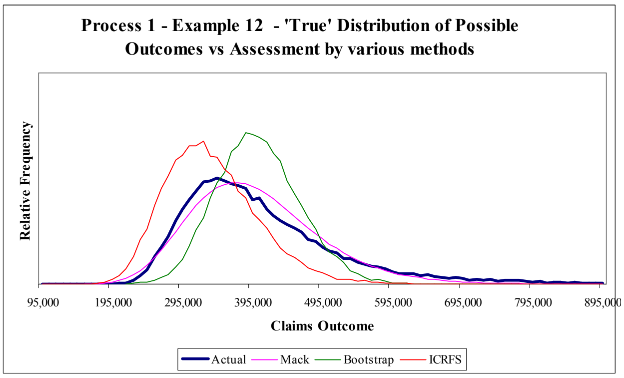





Process 1 - Example 9 - 'True' Distribution of Possible Outcomes vs Assessment by various methods
Relative Frequency
95,000
195,000
295,000
395,000
495,000
595,000
695,000
795,000
895,000
Claims Outcome
Actual
Mack
Bootstrap
ICRFS

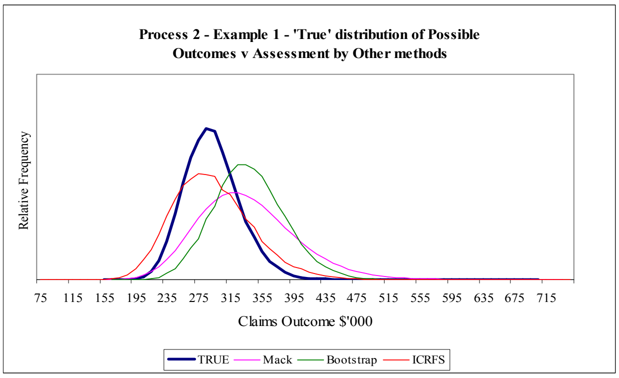

Process 1 - Example 10 - 'True' Distribution of Possible Outcomes vs Assessment by various methods
Relative Frequency
95,000
195,000
295,000
395,000
495,000
595,000
695,000
795,000
895,000
Claims Outcome
Actual
Mack
Bootstrap
ICRFS

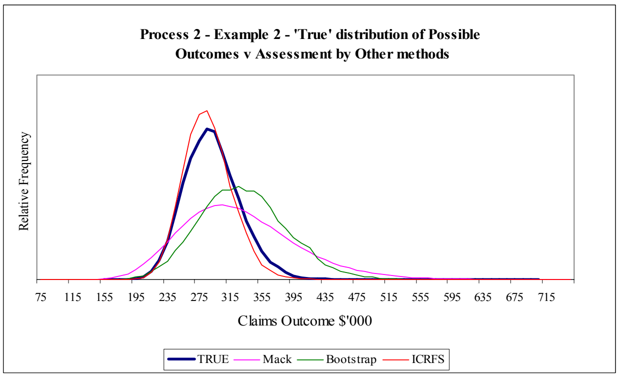

Process 1 - Example 11 - 'True' Distribution of Possible Outcomes vs Assessment by various methods
Relative Frequency
95,000
195,000
295,000
395,000
495,000
595,000
695,000
795,000
895,000
Claims Outcome
Actual
Mack
Bootstrap
ICRFS

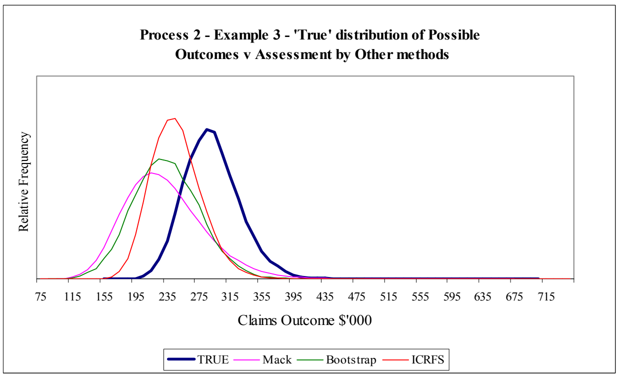

Process 1 - Example 12 - 'True' Distribution of Possible Outcomes vs Assessment by various methods
Relative Frequency
95,000
195,000
295,000
395,000
495,000
595,000
695,000
795,000
895,000
Claims Outcome
Actual
Mack
Bootstrap
ICRFS

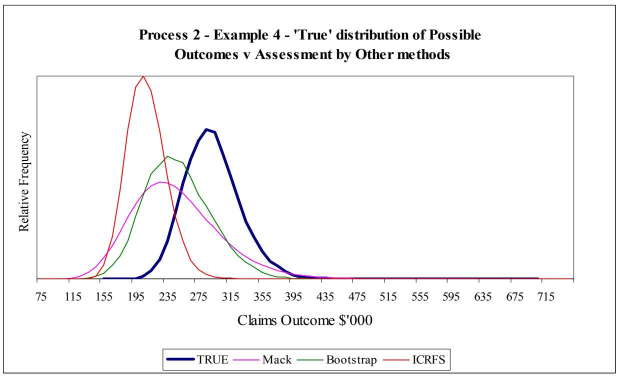
Process 2 (Chain-Ladder Process not followed)

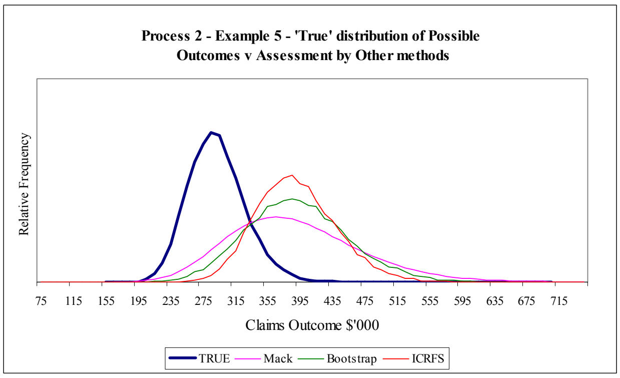
Comparison of ‘True’ Distribution of possible claims outcomes with assessment by different methods

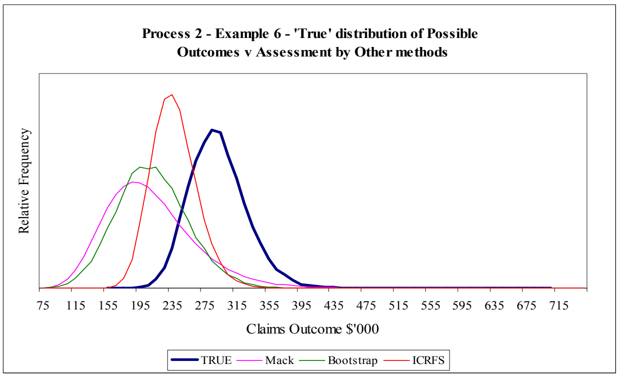


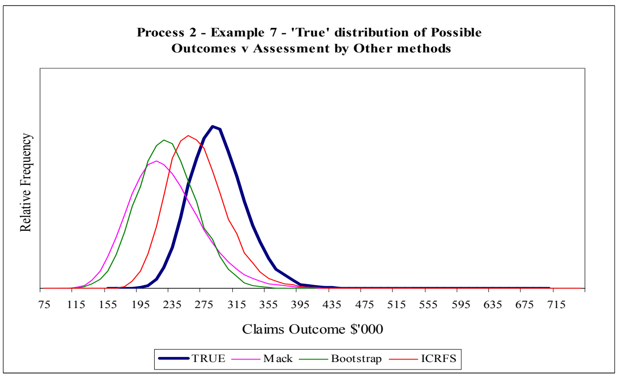


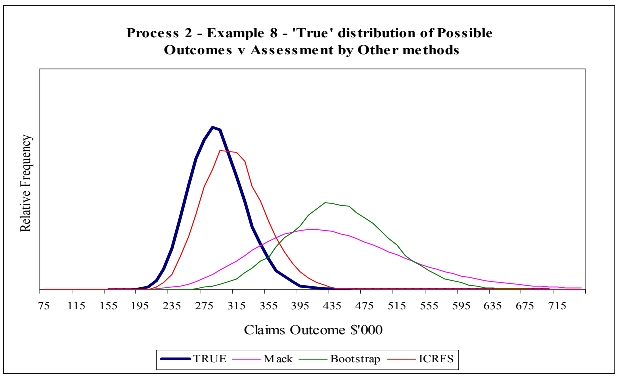


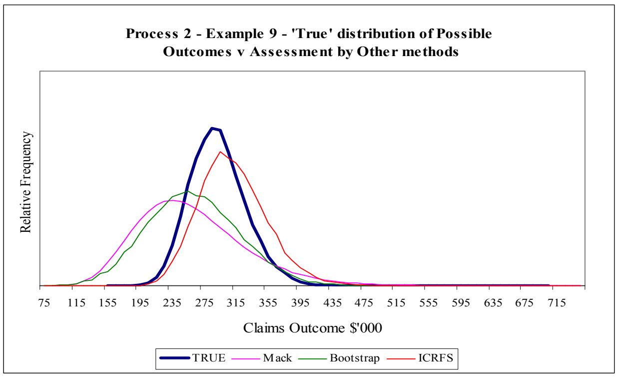


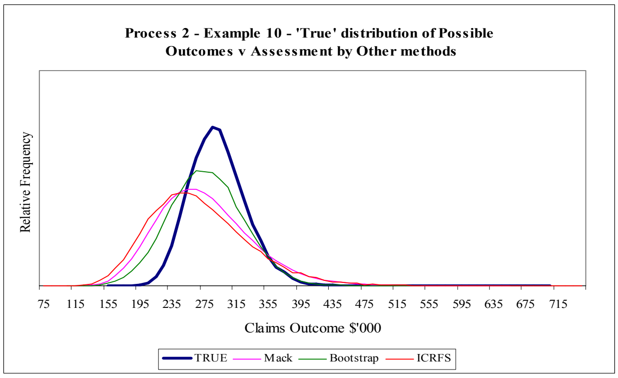


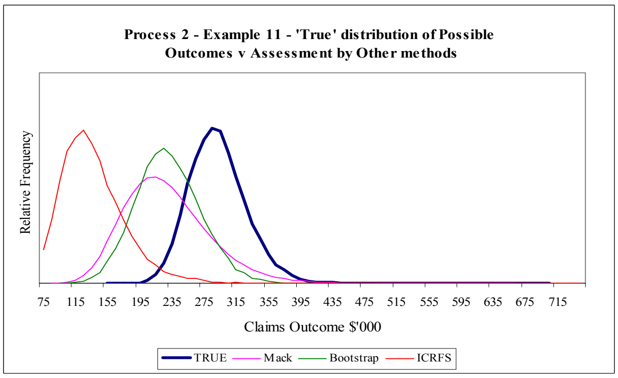


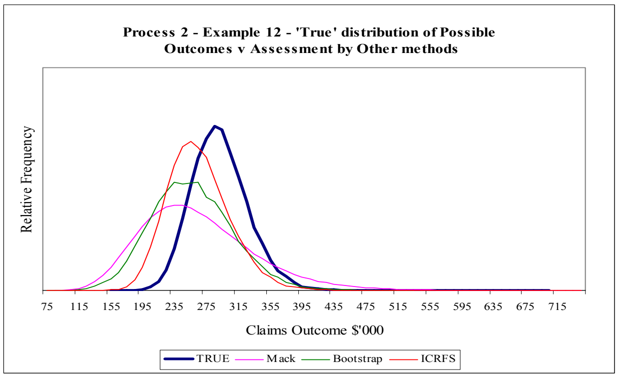





Process 2 - Example 9 - 'True' distribution of Possible Outcomes v Assessment by Other methods
Relative Frequency
75 115 155 195 235 275 315 355 395 435 475 515 555 595 635 675 715
Claims Outcome $'000
TRUE
Mack
Bootstrap
ICRFS


Process 2 - Example 10 - 'True' distribution of Possible Outcomes v Assessment by Other methods
Relative Frequency
75 115 155 195 235 275 315 355 395 435 475 515 555 595 635 675 715
Claims Outcome $'000
TRUE
Mack
Bootstrap
ICRFS


Process 2 - Example 11 - 'True' distribution of Possible Outcomes v Assessment by Other methods
Relative Frequency
75 115 155 195 235 275 315 355 395 435 475 515 555 595 635 675 715
Claims Outcome $'000
TRUE
Mack
Bootstrap
ICRFS


Process 2 - Example 12 - 'True' distribution of Possible Outcomes v Assessment by Other methods
Relative Frequency
75 115 155 195 235 275 315 355 395 435 475 515 555 595 635 675 715
Claims Outcome $'000
TRUE
Mack
Bootstrap
ICRFS

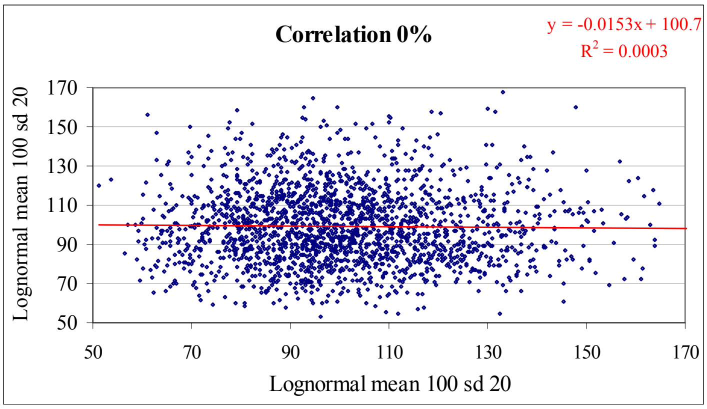
Appendix G

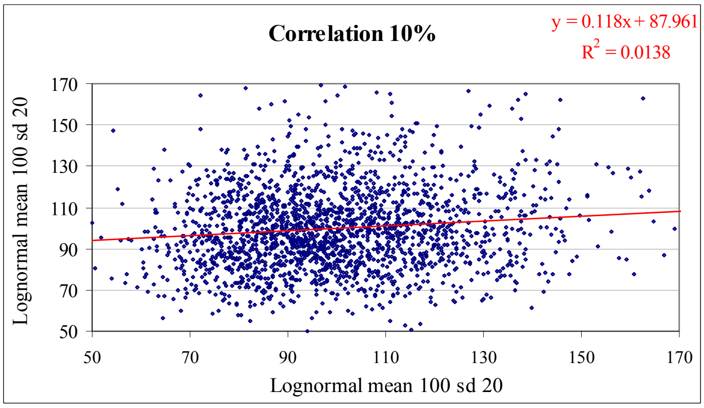
Correlation Scatter Plots

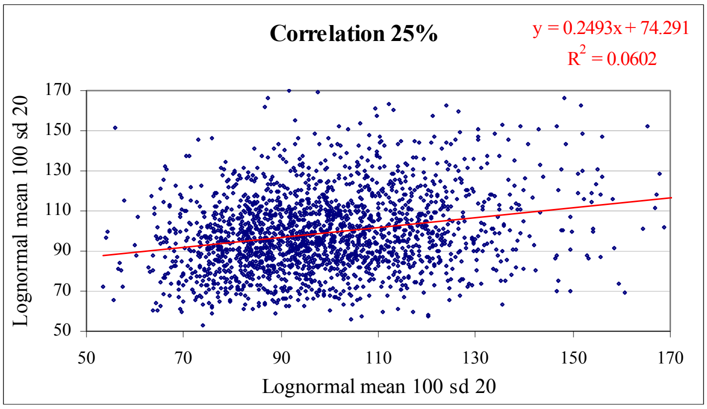
The scatter plots set out in this Appendix have been generated by simulations using the Excel add-in package @risk©. Each plot represents observations of a pair of lognormal random variables with mean 100 and standard deviation 20, associated by correlation coefficients of varying magnitude.

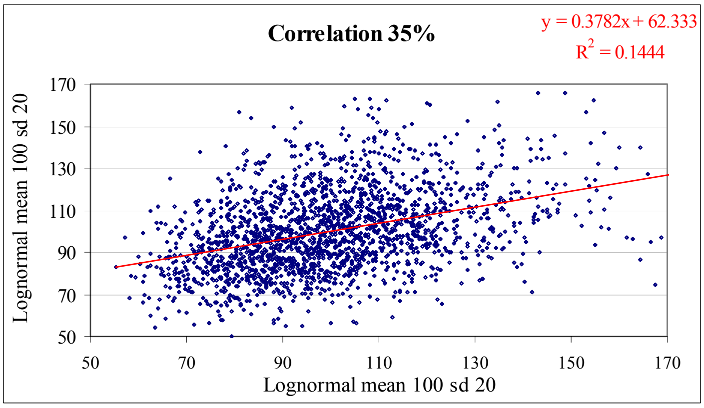
Each dot represents a single simulation. 2,000 simulated observations are shown in each plot. The least squares line of best fit, and its associated r^2 value is also included in each chart.

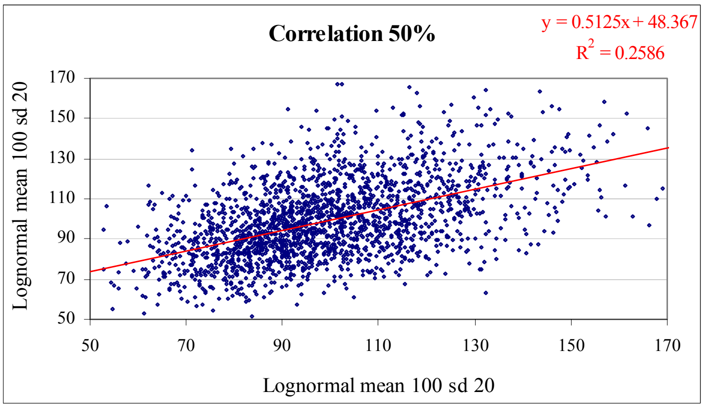
The series of charts aim to provide a tangible representation of the physical meaning associated with varying degrees of association. They may assist actuaries selecting assumptions for the correlation between the likelihood of adequacy of outstanding claims liability estimates.

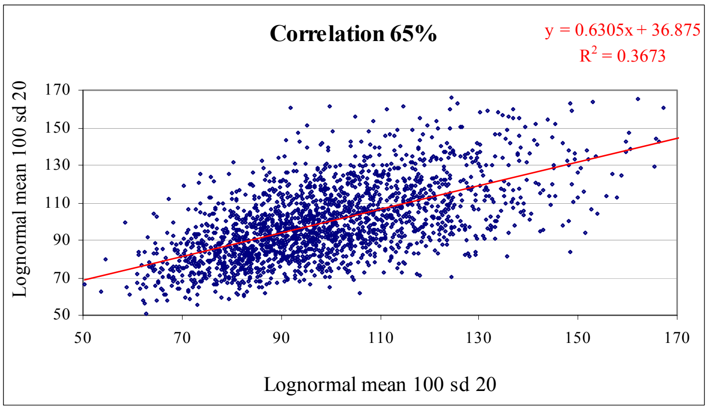


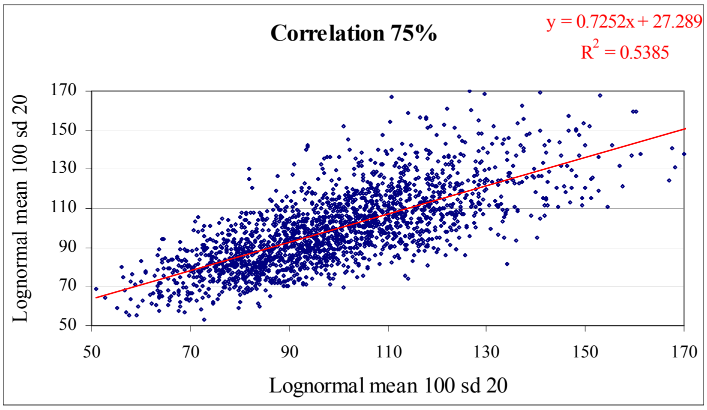




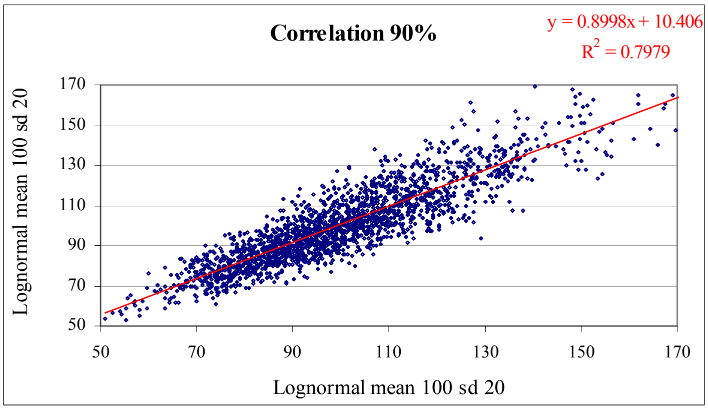




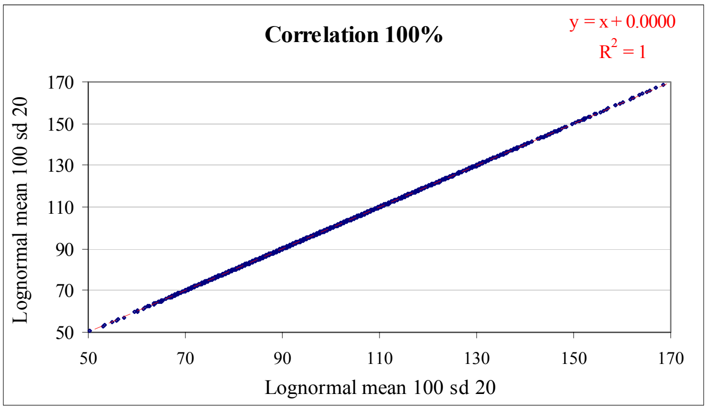




Correlation 75%
y = 0.7252x + 27.289
R2 = 0.5385
Lognormal mean 100 sd 20
Lognormal mean 100 sd 20



Correlation 90%
y = 0.8998x + 10.406
R2 = 0.7979
Lognormal mean 100 sd 20
Lognormal mean 100 sd 20



Correlation 100%
y = x + 0.0000
R2 = 1
Lognormal mean 100 sd 20
Lognormal mean 100 sd 20