Institute of Actuaries of Australia

Pricing Alternative forms of Commercial Insurance cover

Prepared by Andrew Harford

Presented to the Institute of Actuaries of Australia
Biennial Convention 23-26 September 2007
Christchurch, New Zealand

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: actuaries@actuaries.asn.au Website: www.actuaries.asn.au

ABSTRACT

In Australia, it is becoming increasingly common for companies to implement various forms of risk management in an effort to reduce insurance claims – and consequently realize savings on insurance costs. For example, they are implementing occupational health and safety measures, driver training programs and utilizing the services of risk engineers. Commercial insurers are responding to this trend by offering alternative forms of insurance cover. The alternative covers pass the cost savings realized from improved risk management back to the client, while still ensuring the client is adequately protected from unexpected large losses.

This paper discusses the approaches that we have developed to price policies offering alternative types of cover, as well as other approaches that were considered - including those suggested by the literature and from discussions with colleagues. The paper emphasizes methods that are practical to implement. The four types of alternative cover types considered are:

- Aggregates
- Excess of Loss policies (XOL)
- Burners
- Policies with Claims Experience Discounts (CED)

The approaches discussed are at a general level, but there is some discussion of specific issues that have been encountered while pricing risks in various products, such as Workers compensation, Fleet and ISR. This paper does not focus on the relative merits of the various alternatives and does not explore the market for alternative cover types but concentrates mainly on pricing methodology.

Key words: *Commercial Pricing Burner Aggregate XOL CED Alternative*

1. Introduction

1.1 Challenges that pricing alternative risks present to the actuary

As the size of the market for alternative cover types grows it seems timely to produce a paper that brings together the literature on the subject. This paper is not intended to be an exhaustive review of the subject, but it aims to highlight many of the issues that the actuary faces. It attempts to provide a theoretical framework to aid the pricing of these risks, but the actuary still faces many practical challenges.

In many cases, the amount of data available is limited or irrelevant and the actuary has to resort to more approximate means. There are many issues on which there is still no consensus among practitioners. Determining an appropriate loading for profit is one of the most difficult parts of the pricing process.

The actuary often needs to rely on the input of other professionals, such as underwriters, to come up with a suitable premium. The actuary also faces the challenge of explaining their conclusions to underwriters. In many cases a simpler, more approximate method is preferred to ease the process of understanding the results as well as the process of communicating these results to underwriters.

Despite these difficulties the actuary has much to add to the process of assessing an appropriate premium. Deriving a premium for the alternative forms of cover often depends on assessing the variability as well as expected value of the claims. Actuaries are well placed to assess this variability and can be of great assistance to the underwriter in setting the premium. Some of the issues surrounding the pricing of these risks can go against an underwriter's intuition because they depend on the probability of events occurring, and the more extreme events are not always reflected in past data. The actuary can assist in explaining these concepts to underwriters.

1.2 Structure of paper

Section 2 begins by defining the type of policies that are being considered in this paper. This is done in order to avoid any ambiguities that could arise by using terms that can have alternative meanings in a different context. Section 3 then goes on to describe three different high level approaches to pricing these types of policies.

Section 4 considers the issues relating to the construction of a suitable total claims cost distribution. For all the pricing approaches considered in this paper, it is necessary to construct detailed claims distributions. Certain features of the claims distribution are more important than others (eg. getting a good fit for the tail) depending on the alternative cover and class that are being considered. The complexity of the distribution depends on the line of business being considered and the amount of information that is available from the client. Some of the approaches discussed in the literature are described in this section and there is some commentary provided on the ones that were found to be the most suitable.

Sections 5-9 proceed by discussing the issues surrounding the pricing of each component of the premium for Aggregates and XOLs. The paper concludes with Section 10, which describes the approach taken to pricing risks with retrospective premium adjustments (ie. CEDs and Burners).

2. Definitions

Many of the terms used in this paper sound similar to terms used in reinsurance and personal lines insurance, but they have a different meaning in a commercial insurance context. For example, the stop loss clause described in this paper is very different from a reinsurance stop loss contract. Definitions of the types of policies being examined are given below to ensure there are no ambiguities in the use of terminology. Throughout this paper we will use the notation $(X)_+$ to denote Max$(X,0)$

Conventional policies

This refers to the standard type of policy that covers the cost of all claims of a certain type. Policies with low excesses can be considered to be conventional if the excess only removes a small fraction of the total expected claims cost. If the insured has N claims in the period covered and $X_1, X_2, .. ,X_N$ are the cost of the individual claims covered by this policy, the insurer pays:

$$T=X_1+ X_2+ \ldots +X_N$$

Aggregates

If the total claims cost for the insured period exceeds a pre-specified limit, the insurance company pays the policyholder the difference between the total claims cost and the pre-specified aggregate limit. Thus, if A is the aggregate limit, the insurer pays:

$$(X_1+ X_2+ \ldots +X_N - A)_+$$

These policies may include a *stop loss* clause that protects the client from individual large claims. The insurance company pays the excess cost of a claim over the stop loss limit on all the claims that exceed the limit. The cost of large claims below the stop loss limit then counts towards the assessment of the aggregate claim cost. If the stop loss limit is S the insurer pays:

$$(X_1-S)_+ +(X_2-S)_+ +\ldots+ (X_N-S)_+ +$$
$$([\mathrm{Min}(X_1,S) + \mathrm{Min}(X_2,S) +\ldots+ \mathrm{Min}(X_N,S)]-A)_+$$

Example:

Suppose a client has an aggregate limit of $1,000,000 and a stop loss limit of $500,000. If they have three claims costing $750,000, $300,000 and $300,000 then the first claim hits the stop loss, so the insurer pays $250,000 on this claim. The remaining $500,000 of the claim is added to the other two claims to get $1,100,000, which is $100,000 above the aggregate limit. The insurer therefore pays $250,000+$100,000=$350,000 to the insured.

The insurance contract will specify if the stop loss clause applies on a *per event* basis or a *per claim* basis. The formula above uses X_i to represent the cost of individual claims, so this formula would be used when the stop loss clause is applied on a per claim basis. If the stop loss applies on a per event basis we can use the same formula to calculate the cost to the insurer, but the X_i will represent the cost of claims per event (which may consist of one or more claims) rather than the cost of individual claims.

The presence of a stop loss clause on a per claim basis means that the insurer pays more to the insured whenever there is a large individual claim. However, if the sum of the small claims (claims under the stop loss limit) is greater than the aggregate limit then the stop loss clause is irrelevant. If we have a large claim in this situation it has to be paid in full anyway because claims are already above the aggregate limit. Therefore, as the aggregate limit gets lower it increases the likelihood that the sum of the small claims will reach the aggregate, so that the value of the stop loss clause to the insured becomes lower.

Excess of Loss (XOL)

If the insured has a claim that exceeds a certain limit, the insurer pays the insured the difference between the limit and the cost of the claim. XOL policies can be considered to be policies with very high excesses that are set at such a level that only the largest claims will be expected to exceed the limit. If the limit is *B* the insurer pays:

$$(X_1-B)_+ + (X_2-B)_+ + .. + (X_N-B)_+$$

Burners

A burner is a policy where the premium paid depends directly on the claims experience. For a policy covering one year, the policyholder pays a deposit premium at the start of the year followed by an adjustment premium at the end of the year depending on claims experience. The adjustments may continue for several years – the number of years usually depends on how long it takes for claims to run off.

A notional premium is calculated each year by applying an adjustment factor (which allows for expenses, profit and commission etc…) and an IBNR factor (to allow for claims development) to the incurred claims experience at the end of the year. The adjustment premium paid is the difference between the notional premium and the premium paid to date. A burner also specifies the minimum and maximum amount of premium the policyholder has to pay.

The insurer may get its required expenses and profit from a margin in the adjustment factor, or the burner may be structured so that the client pays an *insurance charge* at policy inception to cover the insurer's expenses and profit. The insurance charge can also cover any difference between the expected cost of claims above the maximum less the expected benefit to the insurer when the claims go below the minimum.

The equations for calculating the adjustment premium from the insured are given below. Suppose the burner has a minimum premium of Y, a maximum premium of Z, a deposit premium of D, an adjustment factor of F, an insurance charge of C, an IBNR factor in year i of I_i, and incurred claims of X_i at the end of year i. The notional premium at the end of year i is $F\, I_i\, X_i$, which represents the cost of claims multiplied by the adjustment factor and the IBNR factor, and this amount is subject to the maximum and minimum The premium adjustments in each year are the difference between the notional premiums for the current year (subject to the maximum and minimum) less the premium paid to date. Therefore the adjustment premiums are:

$$\begin{array}{ll} i=0: & C+D \\ i=1: & G(F\, I_1\, X_1) - D \\ i>1: & G(F\, I_i X_i) - G(F\, I_{i-1}\, X_{i-1}) \end{array}$$

Where:

$$G(x) = \begin{cases} Y & x \le Y \\ x & Y < x < Z \\ Z & x \ge Z \end{cases}$$

The function G is applied to the notional premium so that client never pays more than Z or less than Y.

Example:

Suppose a burner has no insurance charge, a minimum of $500,000, a maximum of $900,000, a deposit premium of $500,000, an adjustment factor of 100/80 and IBNR factors of 1.5 in year 1 and 1.2 in year 2. Suppose the client's claims were $300,000 at the end of year 1 and $650,000 at the end of year 2.

The client pays the deposit premium of $500,000 at the start of the contract. At the end of year 1 the notional premium for the contract is $300,000 X 100/80 X 1.5 = $562,500, which is $62,500 more than the premium already paid. Therefore the client pays an extra $62,500 to the insurer. At the end of year 2 the notional premium for the contract is $975,000 = $650,000 X 100/80 X 1.2. This amount is greater than the burner's maximum of $900,000, so the premium is capped at this amount. The client has already paid the insurer $562,500, therefore they must pay an extra $334,500=$900,000-$562,500 to the insurer.

Claims experience discount (CED)

CEDs give the policyholder a retrospective refund of premium for better than expected claims experience. The expected claims cost is calculated as a specified percentage of premium (usually reflecting the expected loss ratio). This is compared to the actual cost of claims occurring in the period. The policyholder is reimbursed with a percentage of the difference between expected and actual claims, when actual claims are better than expected. The amount returned is usually limited to a maximum percentage of premium. In long tail lines of business the CED may be assessed at a point in time before all the claims are settled. In this case an IBNR adjustment is applied to the claims experience at a specified time after the policy's inception to allow for subsequent claim development– ie. the total claims cost is multiplied by an IBNR factor. There is generally only one adjustment made to the premium and this is usually made after one or two years. The exact timing of the return premium will be specified in the insurance contract.

If we charge a premium of P at time 0 and return a proportion, Π, of the amount by which claims are better than expected, subject to a maximum percentage of premium, J, then at time i we would return the following amount to the insured:

$$\text{Min}\,(\Pi\,(LP - I_i X_i)_+\ ,\ JP)$$

Where:
I_i is the IBNR adjustment factor at time i
X_i is the incurred cost of claims at the end of year i
L is the loss ratio applied to the premium to get the 'expected claims'

Example:

Suppose the insurer offers its client a CED policy with a loss ratio of 70%, and an IBNR factor of 1.1. They return 50% of the amount by which claims are better than expected at the end of year 1 subject to a maximum of 20% of premium. Say the premium was $900,000 and claims amounted to $500,000 at the end of year 1. The expected claims on this policy are $900,000 X 0.7 = $630,000. If we adjust the claims for IBNR we get $500,000 X 1.1 = $550,000. This means that claims are $80,000 =$630,000-$500,000 better than expected. The insurer returns 50% of this, $40,000, to the client. Note that $40,000 is less than 20% of $900,000, so the maximum does not apply in this case.

3. General Approach

This paper considers three approaches to pricing alternative covers depending on the amount of information we have at our disposal and the alternative type of policy we are considering. Commission and taxes are not considered in this paper, as these are expected to be proportional to the premium, and paid at roughly the same time as the premium.

The alternative cover types can be considered in two groups:

i) Aggregates and XOLs
ii) Burners and CEDs

In Aggregates and XOL policies the policyholder pays a fixed premium and the insurer only covers part of the total cost of claims. In contrast, the insurer pays the total cost of all claims in Burners and CEDs, but the premium varies depending on the claim experience. A different approach to pricing is often required for the two groups of alternative policies.

3.1 Calculating the premium from first principles

This method can be used to price aggregates and XOL policies when we have sufficient detailed information about the client and we are not provided with a premium for the conventional type of policy. We build up the premium by estimating the cost of the individual components that make up the premium, namely:

Expected claims cost
Expenses
Reinsurance
Investment income
Profit

The cost of providing each component of the premium is calculated by considering the underlying principles involved. For example, the expected cost of claims in an XOL policy is calculated by constructing a suitable total claims cost distribution to estimate the expected cost of claims exceeding the excess limit. Determining the expected claims cost and the allowance for profit and reinsurance are the most difficult issues to resolve using this approach.

3.2 Adjusting the conventional premium to reflect difference in cover

We rarely have access to sufficient data on the client to be able to build a detailed model of future expected claims. This may be because the client is currently with another insurer who only provides us with an aggregated loss history by policy year. It might also arise if there has been a major shift in the client’s exposure to claims, so that past data is no longer relevant. For example, the number of claims reported by the client might be expected to reduce following a change in risk management practices. We have to resort to more approximate methods to estimate the premium in this case

If a conventional quote has been prepared by an underwriter and the client wishes to consider the possibility of using an aggregate or excess of loss policy, it can be useful to use the underwriter’s quote as a starting point for our calculations. In a commercial insurance environment the underwriter will gather information from the client about its exposure to various risks to come up with a quote for a conventional policy. For example, for a motor fleet risk, they will receive details of all the vehicles currently in the fleet. They usually also receive a history of claims for the client, although this might only be a summary of claims by policy year rather than a detailed listing of individual claims. They might also receive reports from risk engineers commenting on the quality of the risk – eg. commenting on the quality of client's risk management practices and exposure to certain types of claims. The underwriter gathers all this information together to make an assessment of an appropriate premium for the client, often with the aid of a model provided by the actuary. They will make some subjective allowances in the premium to allow for changes in the client's exposure to various risks over time. The quote produced by the underwriter will include estimates of the expected number of claims from the client, the expected average claim size as well as an allowance for expenses, reinsurance costs, profit and investment income. This produces a *technical premium* for the risk (a premium which achieves a target return on capital). This premium may need to be adjusted to take market conditions into account – a *market premium*.

We can use the underwriter's assessment of the cost of each component of the technical premium for a conventional policy as a starting point for our calculations and modify the cost of each component to refect the reduced cover provided by the alternative risk. So, for example, when we try to estimate the profit component of an aggregate policy we would use the profit margin on the conventional policy as a starting point and consider how the level of capital needed for a conventional and aggregate policy differ. The aggregate will require less capital in dollar terms because it always pays less on claims so a smaller dollar amount is required as profit (see Section 9 for more details). We consider each component of the conventional premium in turn and modify it to reflect the differences in cover.

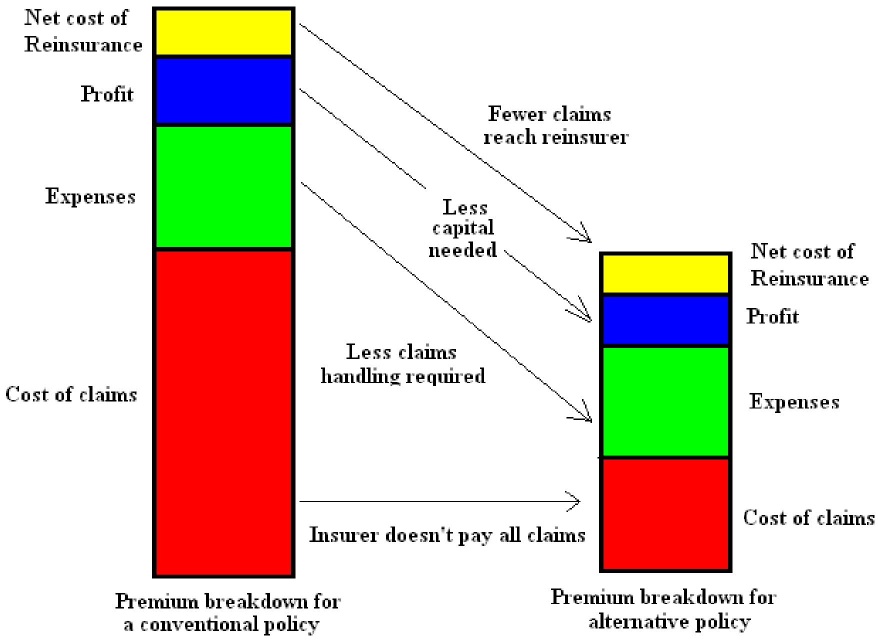

Figure 1: Adjusting the conventional premium for differences in cover

The underwriter will usually have made an assessment of the expected claim cost for the conventional policy and will have derived a premium by dividing this figure by a 'target loss ratio'. The target loss ratio will have implicit assumptions about loadings for reinsurance, expenses, investment income and profit. These loadings can also be adjusted reflect the reduced cover provided by the alternative risk. This approach contrasts to the 'first principles' approach that only considers the historical data at hand and ignores the underwriter's assessment of risk. The underwriter is often better positioned to more accurately price the conventional policy because their quote will take changes in exposure and risk management factors into account, which may not always be evident in the historical data. It is important to use the underwriter's quote before it has been adjusted for market conditions (if available) so that we can assess an appropriate premium for the alternative risk that achieves the required return on capital.

In the 'first principles' method we produce a quote for an alternative cover by independently assessing the cost of each component of the premium and ignoring the underwriter's assessments. This will ignore potentially useful information. The 'first principles' method can produce results that are inconsistent with the underwriter's assessment of the conventional premium. For example, if the limit on an aggregate policy is set to 0 then it will provide the same cover as a conventional policy. If we adopt a 'first principles' approach to pricing this aggregate we will end up a premium that is not equal to the premium set for the conventional policy because we have made assumptions which are not consistent with the underwriters assumptions. Therefore it is usually preferable to use the approach of modifying the underwriter's quote when it is available. However, if the underwriter hasn't produced a quote for the conventional policy we have no choice except to resort to pricing from 'first principles'.

The approach of adjusting the underwriter's quote to reflect the difference in cover has the advantage of ensuring consistency between the premium for the conventional and alternative policy. It can make some of the problems encountered in the pricing of alternative covers more tractable. It is often necessary to use this approach when the amount of data about the client that is at our disposal is limited. However, the main limitation of this approach is the confidence we place on the accuracy of the premium set for the conventional policies and how it is split it into its component parts. This approach has not been explored before in the literature, so this paper will examine the advantages and disadvantages of this approach in some detail.

3.3 Discounting future expected cash flows

The insurer pays for all the claims under a burner and CED arrangement, so these policies offer exactly the same claims coverage as a conventional policy. The only difference between a conventional policy and these two types of policy is that there may be one or more premium adjustments (depending on claims experience) after the initial premium has been paid. Therefore, one approach to pricing these risks is to discount the expected premium flows from the initial premium and subsequent premium adjustments and ensure that the discounted cash flows are equal to the conventional premium.

In theory we could adopt a 'first principles' approach to pricing Burners and CEDs. We could discount all the cash flows of the policy – premium flows as well as the payment of claims, expenses and the cost of capital (ie.profit). This is the approach that is described by Meyers (1986), but this approach suffers the same disadvantages as pricing from 'first principles' - it ignores the underwriter's knowledge of the risk and may produce an answer that is inconsistent with the underwriter's assessment. Burners and CEDs are nearly always offered as an option to a policyholder, so a conventional premium is almost always available. Therefore, in this paper, we only concentrate on discounting the expected cash flows from premium adjustments.

4 Aggregate claim distributions

For XOL policies we need to construct a distribution for individual claims to assess the expected cost of claims over the limit. For all the other alternative cover types we need to construct an aggregate claims distribution - a distribution for the client's total claim cost. In this paper we will use the term *aggregate claim distribution* to refer to the distribution of a client's total claim cost, which should not to be confused with the similar sounding term 'aggregate policy' that was defined in Section 2. The term *total claim cost distribution* may also be used to describe this distribution.

We can use an aggregate claim distribution to assess the future premium flows from a burner or CED because these depend on the client's claim experience. The aggregate claim distribution can also be used to assess the expected cost of claims above the aggregate limit for an aggregate policy. An aggregate claim distribution is derived from underlying assumptions about the distribution of claim size and claim counts:

If N is a random variable representing the number of claims per year and X_1, X_2,…, X_N are random variables representing the claim sizes of the N claims then the aggregate claim cost is

$$T = X_1 + X_2 + \ldots + X_N$$

Several approaches to deriving the aggregate claims distribution from the underlying claim number and size distributions are discussed in the paper by Heckman-Meyer (1983). Some methods, such as Panjer recursion, Panjer (1981), can be used to efficiently construct the aggregate claims distribution, however the Monte Carlo method is by far the easiest to understand and implement as well as the most flexible. It is reasonably accurate and with modern computing power its lack of efficiency is usually not a major issue. In this paper we concentrate on using the Monte Carlo method.

In the Monte Carlo method, we make some assumptions about the claim number and claim size distributions and then simulate a large number of sample values for the aggregate claim cost using the underlying distribution assumptions. The distribution formed by the simulated total costs approaches the aggregate claim distribution as the number of simulations increases. Typically, if we expect a few hundred claims on a policy we can simulate tens of thousands of possible scenarios for the total claims cost in a matter of minutes. One major advantage of the Monte Carlo approach to finding the aggregate claims distribution is its flexibility - we can model different types of claims separately and add them up to get the aggregate cost. For example, in motor fleets, it makes sense to have separate claim number and claim severity distributions for different types of vehicles. The Monte Carlo method also allows us to calculate other quantities of interest as part of the simulation. For example we can calculate the expected reinsurance recoveries by calculating these for each separate simulation.

The main focus of this paper is to discuss issues in pricing alternative risks so we will not concentrate too much on how aggregate claim distributions are constructed, as this topic is covered more adequately in other sources (see the References Section). Instead, we will concentrate on the issues faced in the construction of the distribution that are specific to assessing the cost of alternative cover types. First we consider how to select assumptions for the distribution of claim numbers and individual claim sizes.

4.1 Claim numbers

The first consideration in choosing an appropriate claim number distribution is to look at the amount of claims data we have available. If we have access to a large volume of individual claims data for the client it will be possible to use this data to test various distribution assumptions. In many cases, the clients individual claims history will not be available or there will not be a sufficient volume of claims to test distribution assumptions. In these cases more approximate methods are needed.

The Poisson distribution is a commonly used distribution for claim numbers. However, because the mean of the Poisson distribution is equal to its variance, it may fail to give sufficient variability to the number of claims that would be seen in practice. The Negative Binomial can be used as an alternative to increase the variability of claim numbers. A survey of potential candidate claim number distributions is given in Wang (1998), and Klugman et al (2004).

If we have sufficient data we can test various distribution assumptions. Klugman et al (2004) also surveys a number of useful methods to test the goodness of fit for different distributions. After selecting an appropriate form for the distribution, we need to select appropriate parameters. If we don't have a quote for a conventional policy, we will need to adopt the 'first principles' approach and form our own opinion of the expected number of claims under the policy. This involves looking at trends in historical data and using actuarial and/or underwriting judgement to allow for known changes such as changes in exposure, policy conditions and the external claims environment.

If we don't have detailed historical claims data we need to choose distribution assumptions by looking at our existing portfolio. If the claim type we are modelling is not subject to accumulation risks the Poisson distribution may be appropriate. If claims follow a Poisson process, the inter-arrival times for successive claims should follow an exponential distribution. When we looked at the inter-arrival times of a particular claim type for large individual clients they did appear to follow an exponential distribution. Figure 2 shows a Quantile-Quantile plot (Q-Q plot) of actual versus expected inter-arrival times for one particular client. It plots the percentage of arrival times between claims we would have expected to see below a given value if our fitted distribution was correct against the actual percentage observed below the same value. This procedure was carried out for a large number of clients and the exponential assumption appeared to fit well for all the clients. Therefore in cases where individual claims data was unavailable we assumed that the claims followed a Poisson process.

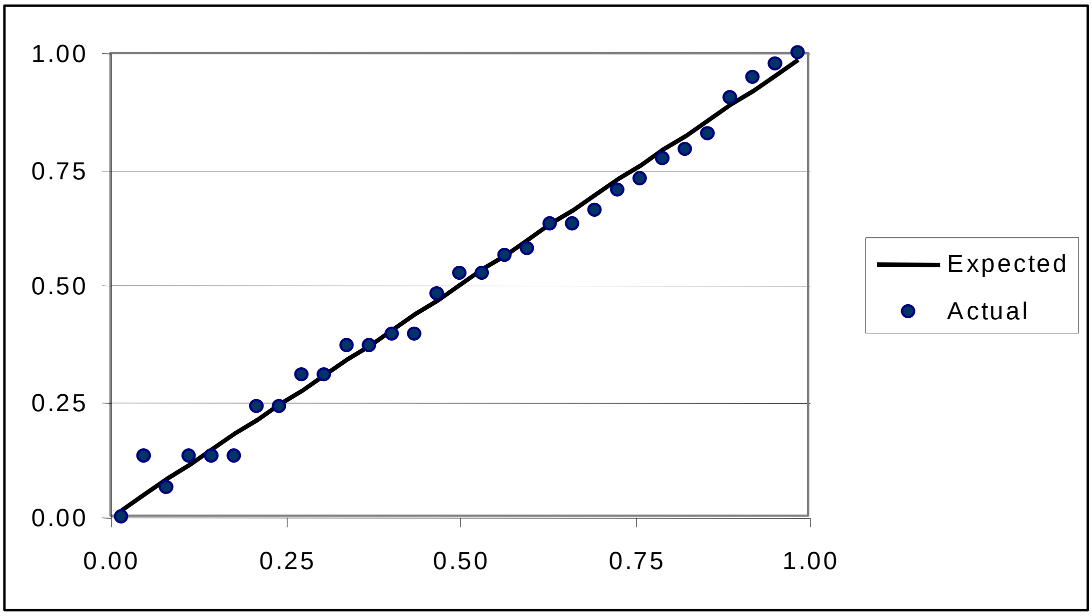

Figure 2: Q-Q plot for actual claim inter-arrival times for one particular client versus a fitted exponential distribution

If we are adopting the approach of adjusting the underwriter's quote to reflect differences in cover we might want to consider using the frequency assumptions underlying this quote. In many cases the underwriter will have made an assessment of the expected claim frequency for different claim types when putting the quote together and these are often useful to help set parameters for the claim number distributions. We can set the expected value of our distribution to be equal to the underwriter's assessment of claim numbers, as long as there is no bias in their selection. The selections may be biased if the underwriter uses the claim numbers for other reasons, such as assessing the expected claims handling costs.

4.2 Claim severity

If a large volume of individual claims data is available for the client it may be possible to fit a claim size distribution to the data. For most alternative covers it is important to ensure that the tail of the distribution fits adequately – the largest claims will be the ones that are most likely to cause the alternative policies to hit their various limits. It can also be important to ensure that the distribution fits well at the smaller end, particularly for CEDs and Burners, because they provide returns in premium that could be underestimated otherwise.

If we are adopting the approach of adjusting the underwriter's quote to reflect differences in cover we will want to ensure that the distributions we have used are consistent with the average claim size and frequency assumptions used by the underwriter. In particular, the expected value of our aggregate claims distribution should equal the underwriter's estimate of the expected claim cost for the conventional policy. This often means that the average claim size of our individual claims distributions should be set to equal the average claim size chosen by the underwriter.

In the case where we have limited information we may be able to use the data for the whole portfolio to help select appropriate distribution assumptions. We can fit various parametric distributions to the claims in a portfolio and then scale them to suit a particular client. So, for example, if we find that a lognormal distribution fits well to our whole portfolio of claims, we might consider using a lognormal distribution for the claim severity distribution for an individual client. When we do this we need to take into account the fact that the client may have higher or lower average claim size than the rest of the portfolio. We can allow for this by adjusting the parameters of the distribution so that the mean of the distribution is equal to the underwriter's assessment of the average claim size.

Even if we set the average claim size of our individual claims distributions equal to the average claim size chosen by the underwriter, this does not tell us how we should set the variance or other parameters of our claim distribution. We can get some of this information by looking at how the mean and variance tend to be related for other similar clients. When we investigated our largest clients for a particular claim type we noticed that the lognormal distribution looked like it was appropriate. We fitted μ and σ parameters for each of the clients using the method of moments and the value of sigma seemed to be independent of the value of μ (see Figure 3). It appears reasonable to conclude that σ is fixed for all values of μ. This is consistent with an assumption that the variance increases in proportion to the mean squared, or equivalently that the coefficient of variance is constant

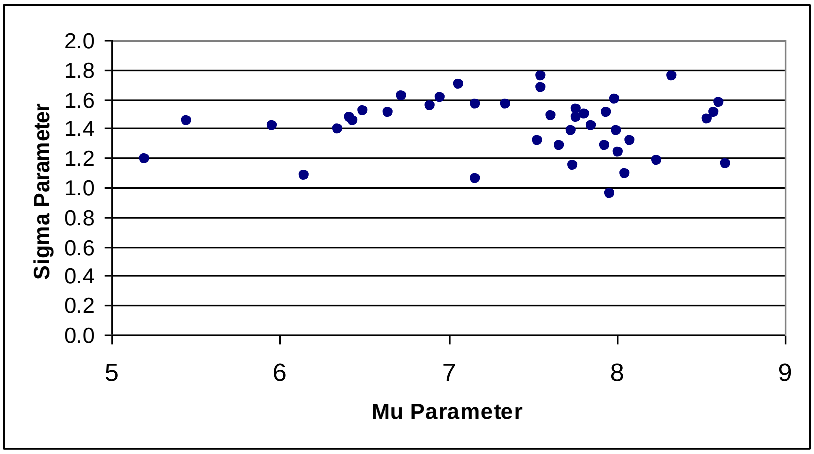

Figure 3: Fitted values of μ and σ for the claims distribution of the largest clients in one particular class of business

Often it is not possible to find any standard distribution that fits the data well at all claim sizes. We may consider fitting different distributions to different claim types and to make assumptions about the relative mix of different types of claim. It might also be necessary to fit different distributions to fit different claim sizes.

For example, we might wish to model nil claims and claims that only involve the payment of investigation fees or assessor's fees separately from other claims, because these claims are typically small and reasonably common. They usually produce a bad fit at the small end of a parametric claims distribution if they are not excluded. Modelling these correctly can be important if we adopt an underwriter's assessment of claim frequency and total expected cost, because it will make the average size of non-nil or non-assessor claims much higher. This in turn typically increases the variance of the resulting aggregate distribution.

In motor fleet, we have found that different vehicle classes have different claim size distributions. Some of this is driven by the relative sums insured in the different classes, but it is also caused by the different type and severity of claims that different vehicles can have. It is desirable to have a claims size distribution that varies by vehicle class and sum insured.

4.3 Miscellaneous issues

The miscellaneous issues considered in this section are methods for improving the accuracy of our model of the total claims cost, or to allow for the uncertainty in our distribution fitting procedures. They are technical niceties that can greatly complicate the distribution fitting process and they can be ignored or allowed for more subjectively if the increased accuracy they provide is thought to be immaterial. They are included here for the sake of completeness.

4.3.1 Accumulations

We need to consider if we want to allow for accumulation risks in our model. Including a model for accumulations can make the model much more complex. However, if we assume that all the claims are independent we could potentially underestimate the cost of the alternative cover, since the limits of the alternative could easily be breached by multiple claims arising from one event. It can be important to model accumulations if we are pricing an aggregate with a 'per event' stop loss.

Some types of accumulations will arise from well-studied types of catastrophes such as cyclones and earthquakes. Models often exist for these types of accumulation and these could be incorporated into our model if the client had sufficiently high exposure to the risk. However, many other forms of accumulations exist that may be unique to a specific industry or client. One example that commonly arises in motor fleet is the presence of road trains in a fleet. A road train typically consists of a Prime Mover hauling three or more Semi-Trailers. If we treat a road train's prime mover and semi-trailers as independent risks we would underestimate the variability in the claims distribution. This can be an issue if we are pricing a fleet with a large number of road trains. We face similar issues in ISR with multiple properties in close proximity, and in workers compensation where a large number of workers are concentrated in one location.

If we adopt a Monte Carlo method of simulating claims, we have the flexibility to create any model of accumulation risks that we like. However, it will often be difficult, if not impossible, to verify the accuracy of any model we construct for accumulation risks – since past data will often be limited or irrelevant and the client's exposure to accumulations may be unique. It may be more practical to ignore accumulations (unless it is a common type of accumulation where modelling is possible) and to make a subjective adjustment to allow for this deficiency.

4.3.2 Allowance for Claims Development

If an alternative cover type involves premium adjustments it can be important to model how claims develop over time. On long tail lines the final premium adjustment may be made at a point in time before all the claims have been settled. The distribution we require in this case is not the aggregate distribution of the ultimate cost of claims, but the aggregate distribution of incurred claims at a point in time, since this is what the adjustment premium will be based on. The distribution of the total claims incurred at a point in time before all claims are settled will often have a distribution that has a higher variance and a lower mean than the aggregate distribution of total ultimate claim costs.

One approach that could be suggested involves simulating the ultimate claims cost for individual claims and then applying a stochastic development pattern to these claims that depends on the ultimate size of the claims. This approach may be invalidated if the development of claims has changed over time, or it is expected to change in the near future (for example due to changes in legislation). Developing a stochastic model for future claims development will greatly complicate the process of producing a quote for an alternative cover type, so it may simpler to ignore it and allow for it more subjectively

If we want to develop a stochastic model for claim development it may be preferable to simulate the development of individual claims separately, rather than model the development of the total cost of claims. Large claims will develop very differently from small claims, so if during the course of a particular simulation we generate a large claim for the client we will want it to follow the pattern generally observed for large claims – ie. a potentially long reporting delay where the value of the claims is 0, then it gradually increases in size and depending of the level of conservatism in the case estimate setting procedures it may be overstated until it finally settles.

One proposed model is as follows - If C is the ultimate cost of the claim, we could model the claim size at time T, C_T, as:

$$C_T = D_{C,T}\,C$$

Where $D_{C,T}$ is a random variable that depends on the size of the claim and it represents the ratio of incurred cost of a claim to its ultimate cost at the point in time. For medium sized and large claims $D_{C,T}$ will have some non zero probability, $P_0(C,T)$ of being 0 for low values of T – this represents the proportion of claims that are not reported by time T and this proportion would usually be expected to increase as the claim size increases but it decreases with time. D_T will also have some non zero probability, $P_1(C,T)$, of being 1 – this represents the proportion of claims that are settled by time T and this proportion decreases as the claims size increases, but increases with time. The values of $P_0(C,T)$ and $P_1(C,T)$ can be estimated from historical data at the time period of interest (eg. after the end of each year). We may wish to band claims of similar size together to make these estimates. For other claims that are reported but not settled, the distribution of $D_{C,T}$ needs to be estimated from historical experience. We found that the gamma distribution tends to provide a good fir for this variable. Different types of claims may need to be separated when carrying out the process of fitting distributions if they have very different claim development characteristics. Note that $D_{C,T+1}$ and $D_{C,T}$ are not independent random variables – so a model needs to be created describing their dependence relationship (eg. through the use of a copula).

This model does not work for claims with an ultimate cost of 0 (or other very small claims). These claims may have been large at some point in time and then settled at no cost. A separate model is needed for these based on historical experience of how these claims tend to develop.

If we have sufficient data we could model the variability of claims development based on the client's historical experience. However, this approach requires caution. If the client is currently with another insurer with very different claims estimation procedures to the insurer who is underwriting the risk, the fitted development distributions may not be appropriate.

It is usually preferable to derive development distributions by examining the whole portfolio of the insurer writing the policy rather than relying on the client's own claims history. This is because the client is unlikely to have had enough claims for us to have confidence in the fitted distribution.

4.3.3 Model and parameter error

Inevitably there will be some error in the selection of appropriate claims distributions and appropriate parameters for the distributions. Wang (1997) describes a 'proportional hazards' transform that provides a loading to allow for the uncertainty surrounding the fitting of the claim distribution:

Suppose T is our fitted aggregate claim distribution. Let $f(x)$ be the PDF of T and $F(x)$ be the CDF of T. If $S(x)=1-F(x)$ it can be shown, using integration by parts, that:

$$\boldsymbol{E}[T] = \int_0^\infty x f(x)dx = \int_0^\infty S(x)dx$$

Wang's proportional hazard transform consists of a mapping $V(x)=[S(x)]^r$ where $0 \leq r \leq 1$ and the proportional hazards mean is:

$$H_r[T] = \int_0^\infty [S(x)]^r\, dx$$

The value of r is subjectively chosen by the actuary to represent the level of uncertainty in the fitted distribution. When $r=1$ this means there is no model uncertainty. Smaller values of r are associated with increasing levels of uncertainty around model selection. This transform has many useful properties that Wang describes in his paper. It also has the effect of applying greater loadings to distributions with fatter tails (ie. have higher probability of large claims).

The proportional hazards transform tell us how we can adjust a distribution to subjectively allow for model and parameter uncertainty. If we want to calculate the value of some quantity of interest that depends on the aggregate claim distribution (such as the expected cost of claims above a limit) we will want to estimate this from our Monte Carlo simulation. In a typical Monte Carlo simulation of an aggregate claim distribution we assume that each simulated scenario we generate is an equally likely possibility for the total claims cost. On this basis we can calculate a value of interest (like the cost of claims over a limit) by calculating it in each individual scenario and then average it across all the scenarios. However, if we wish to apply a proportional hazards transform we can no longer consider each scenario to be an equally likely possibility. The transform essentially modifies the distribution so that larger values become more common – these are the claims that are subject to the greatest uncertainty and are exposed to the greatest risk of underestimation.

For example, the table below shows the ten deciles of the lognormal distribution with parameters μ=10, σ=2. Let F be the CDF of this distribution. If we carry out a Monte Carlo simulation each of these would be considered to be equally likely scenarios for the distribution. However, if we apply a proportional hazards transform with r=0.8, we produce a new distribution with a CDF of G, where these scenarios are no longer equally spaced. A value of 285,815 only represents the 84.2 percentile of the transformed distribution rather than the 90th percentile. If we want to calculate some value of interest on the transformed distribution then we can't simply calculate the value in each scenario and average it across all scenarios. Instead we need to use 'importance sampling' – see Herzog et al (2002). This involves calculating a weighted average of the quantity of interest over all the scenarios where the weights represent the derivative of the function $1-(1-p)^r$ at the percentile we are examining. In this case this weight equals $r(1-p)^{r-1}$. So in this example, the derivative of $1-(1-p)^{0.8}$ at p=0.9 is $0.8(1-0.9)^{0.8-1}$

x	F(x)	S(x)	$S(x)^{0.8}$	$G(x)$
1,697	0.1	0.9	0.919	0.081
4,092	0.2	0.8	0.837	0.163
7,717	0.3	0.7	0.752	0.248
13,271	0.4	0.6	0.665	0.335
22,026	0.5	0.5	0.574	0.426
36,559	0.6	0.4	0.480	0.520
62,869	0.7	0.3	0.382	0.618
118,568	0.8	0.2	0.276	0.724
285,815	0.9	0.1	0.158	0.842

Generalizing this, if we were simulating 1,000 equally likely scenarios, we would sort the scenarios in order of increasing size and assume that the first scenario represented the 0.0005 percentile, the second represents the 0.0015 percentile and so on until the 0.9995 percentile. Suppose we wanted to calculate the cost of claims over some threshold B, ie.$(X\text{-}B)_+$, for the transformed distribution with a transform parameter value of r. We would calculate $(X\text{-}B)_+$ for each scenario and multiply it by $r(1\text{-}p)^{r-1}$ where p is the percentile value associated with the scenario. We add all of these values together and divide this by the sum of $r(1\text{-}p)^{r-1}$ for all the values of p that we have used to represent the simulation percentiles.

If we are applying a first principles approach to pricing we can use these ideas to take parameter and model error into account. However, if we are considering the approach of modifying the underwriter's quote we need to be aware that an allowance for this type of error will already be present in the quote for the conventional policy. The profit margin in the underwriter's quote should already contain an allowance for the impact that inadequate estimation of the risk can have on capital.

The method of discounting premium adjustments for Burners and CEDs described in 2.4 does not allow any margin for model and parameter uncertainty. We can explicitly allow for these by modifying our distribution using the proportional hazards transform.

5. Expected cost of claims

Once a suitable aggregate claim distribution has been constructed, the first step in the calculation of the premium for the alternative risk is to calculate the expected cost of claims. If we have simulated the aggregate claim distribution by Monte Carlo method then we can calculate the cost of claims exceeding the aggregate limit in each simulation and get the average cost over all simulations. In mathematical notation, If $f(x)$ is the PDF of the aggregate claims distribution, the expected cost of claims subject to the aggregate limit, A, is given by:

$$\alpha = \int_A^\infty (x - A)\, f(x)dx = \int_0^\infty (x - A)_+\, f(x)dx$$

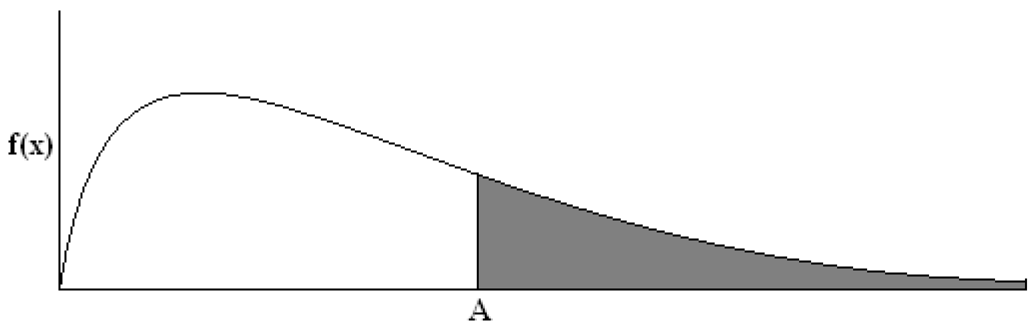

Figure 4: Claims paid under an Aggregate policy

The value of α can be estimated in a Monte Carlo simulation by calculating the total claims cost for each simulation, then calculating $(x\text{-}A)_+$ and averaging this value across all the simulations. If there is a stop loss clause in the aggregate contract, we can calculate the cost of claims paid by the insurer for each simulated scenario (using the formula given in the aggregate policy definition from Section 2). This can then be averaged across all the scenarios. The ability to easily calculate the cost of an aggregate with a stop loss for each scenario in a simulation is another reason why the Monte Carlo approach is preferred. It is not easy to calculate the cost of this contract if you construct the aggregate claim distribution from a method like Panjer recursion because the cost to the insurer doesn't just depend on the total claim cost, it also depends on what the individual claims costs are.

An XOL policy (as defined in Section 2) can be considered to be a policy with a high excess. Therefore we don't need to construct an aggregate claim distribution if we are estimating the expected claim cost for an XOL policy. We only need to construct a claim distribution for individual claims and calculate the expected value of claims subject to the XOL limit. We then multiply this figure by the expected number of claims from the ground up.

For example, if individual claims have a lognormal distribution with parameters μ and σ, and the XOL limit is B then the expected value of an individual claim subject to the limit is equal to the expected value of $(X - B)_+$. The formula for this is:

$$\boldsymbol{E}[(X - B)_+] = (1 - \Phi(\beta - \sigma))\, e^{\mu + \sigma^2/2} - (1 - \Phi(\beta))B$$

Where Φ is the standard normal CDF and $\beta = (ln(B) - \mu)/\sigma$

For more complicated distributions, or distributions involving hazard transforms, it may be necessary to evaluate $\boldsymbol{E}[(X - B)_+]$ through simulation. We simulate a large number of claim values from the selected distribution, evaluate $(X\text{-}B)_+$ and average this over all the simulations.

6. Reinsurance

The cost of reinsurance for alternative covers should usually be lower than the cost we would apply to a conventional policy. This reflects the reduced claims coverage provided by the alternative policy. The proportion of the conventional reinsurance cost that should be charged will depend on the nature of the reinsurance cover as well as the nature of the alternative cover. In this paper we only consider two types of non-proportional treaty reinsurance covers that are typically in place – individual XOL and Catastrophe reinsurance. Unless we are using a 'first principles' approach to pricing the alternative risk, we assume that the reinsurance cost component of the conventional premium can be broken down between the different types of reinsurance used.

If the cost of reinsurance for a conventional policy is not large it is probably not necessary to have a sophisticated approach for working out the cost of reinsurance for an aggregate or XOL policy. The methods outlined below are more useful when the cost of reinsurance forms a significant proportion of the premium.

6.1 XOL Reinsurance cost for an XOL policy

To reduce the confusion in terminology that could arise in this section, whenever the term 'XOL policy' is used in this section it refers to the alternative cover provided by the insurer to its client as defined in Section 2. When the term 'XOL reinsurance' is used it refers to an individual XOL treaty between the insurer and a reinsurer.

If we use the approach of modifying the underwriter's quote for a conventional policy to reflect differences in cover we will have an estimate of how much of the premium is needed from the policyholder to cover the conventional policy's net cost of XOL reinsurance (the cost of reinsurance net of the expected recoveries) – let's call this amount *R*. We assume that this cost is allocated to a policy in proportion to its expected cost of reinsurance recoveries. If we write the policy as an XOL policy then a claim needs to be bigger from the ground up before we start getting recoveries from the reinsurer. This is because the insurer does not have to cover the cost of claims below the insureds XOL policy limit. If we allocate the cost of reinsurance based on the expected level of recoveries on the policy (which seems like a reasonable approach), we can estimate the cost of reinsurance for the XOL policy by multiplying *R* by the ratio of the expected recoveries from the conventional versus the XOL policy. In order to do this we need an individual claim size distribution for claims from the ground up – or at least a claim distribution for claims above the XOL limit on the policy. We can then assess the relative cost of claims net of reinsurance recoveries under a conventional policy against the net cost of claims under the XOL policy.

Suppose the insurer has XOL reinsurance treaty in place that has a lower limit of *L* and an upper limit of *U*. If the insurer writes a conventional policy with no excess then the insurer will recover $(\mathrm{Min}(X-L,U-L))_+=(X-L)_+-(X-U)_+$ from the reinsurer for a claim of size *X* from the ground up. If the insurer writes the policy as an XOL policy with a limit of *B*, then a claim must be bigger than $B+L$ from the ground up before we can get any recovery from the reinsurer. The reinsurance recovery in this case is $\mathrm{Min}((X-B-L)_+,U-L) =(X-B-L)_+-(X-B-U)_+$ for a claim of size *X* from the ground up under an XOL policy with an excess point of *S*. The ratio of the expected reinsurance recoveries on the XOL policy versus conventional is given by:

$$\frac{\int_0^\infty ((x-B-L)_+ - (x-B-U)_+)f(x)dx}{\int_0^\infty ((x-L)_+ - (x-U)_+)f(x)dx} = \frac{\boldsymbol{E}[(x-B-L)_+]-\boldsymbol{E}[(x-B-U)_+]}{\boldsymbol{E}[(x-L)_+]-\boldsymbol{E}[(x-U)_+]}$$

where *f(x)* is the PDF for the individual claims distribution (for claims from the ground up). If this ratio is applied to *R*, the XOL reinsurance cost component of the quote for the conventional premium, this should give us the cost of reinsurance for the XOL policy. The value of *R* can be estimated using Monte Carlo simulation, or it can be computed exactly if there is a formula to calculate the value of $\mathbf{E}[(X-L)_+]$. For example, the formula to calculate $\mathbf{E}[(X-L)_+]$ for a lognormal distribution was given in section 5.

This approach is valid if the reinsurance contract has no reinstatements. If the reinsurance contract provides reinstatements, it is assumed that the expected reinstatement costs (if any) have been included as a component of the XOL reinsurance cost allocated to the conventional policy. The same ratio will be applicable to the expected reinstatement costs. Therefore, it is valid to apply the ratio shown above to the total XOL reinsurance charge.

If we have adopted the 'first principles' approach to estimating the cost of reinsurance we could still use the formula above, except we would first need to assess the value of *R*. This can be estimated by calculating the total cost of claims from the ground up and applying the reinsurance loading used in the target loss ratio assumptions.

6.2 XOL Reinsurance cost for an aggregate policy

When assessing the cost of XOL reinsurance allocated to an aggregate policy again we use the approach of assessing the relative level of recoveries from the conventional policy versus the aggregate policy. A large claim under an aggregate agreement would only be covered by an XOL reinsurance agreement if the claim triggered a payment under the aggregate policy. The size of the claim as well as the sum of all the other claims and how they compare to the aggregate limit and XOL limits need to be considered.

For example, suppose an aggregate policy had a limit of $3M, the insurer had an XOL reinsurance policy covering $5M XS $5M, and the policyholder had a large claim of $6M and its other claims for the year amounted to $2M. The large claim results in a $5M claim payment to the insured. In this case, there will be no reinsurance recovery for the claim.

More generally, suppose the XOL reinsurance policy has a lower limit of L and an upper limit of U, and the aggregate has an aggregate limit of A with no stop loss. Suppose the policy has a single large claim costing C_L (a claim that costs more than the lower reinsurance limit) and that the cost of the small claims is C_S. If the small claims add up to something greater than A then the insurer pays the entire cost of the large claim. However, if the small claims add up to something less than A, there is some 'spare capacity' on the aggregate policy to meet the cost of the large claim. The amount of spare capacity on the aggregate policy to meet the large claim is $(A\text{-}C_S)_+$. The insurer pays a total of $(C_L - (A - C_S))_+$ to the insured on this policy, but the expected payout from the insurer to the insured in respect to the large claim *only* is:

$$(C_L - (A - C_S)_+)_+$$

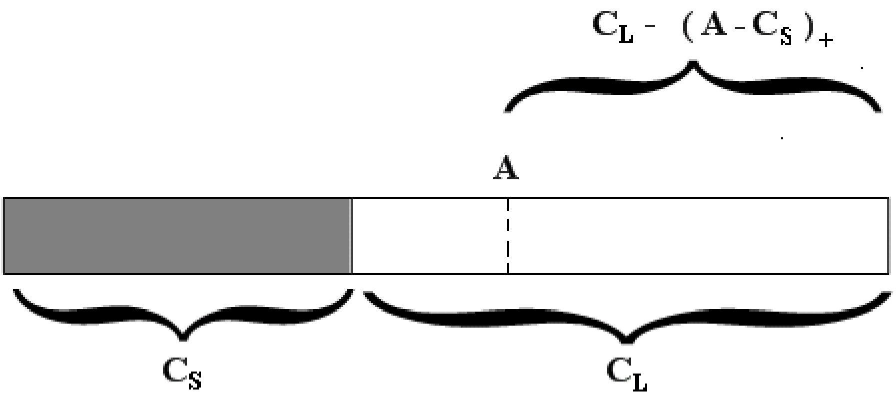

Figure 5: The part of the large claim that is payable by the insurer

If this amount is greater than L the insurer will be able to make a recovery from its reinsurer up to a maximum of $U\text{-}L$. The insurer cannot claim a recovery from its reinsurer if it does not pay at least L on the individual claim (after applying the aggregate limit). The expected payment from the reinsurer under the XOL reinsurance contract in respect to the large claim is therefore:

$$\text{Min}\,(\,(\,(\boldsymbol{C_L} - (\boldsymbol{A}\text{-}\boldsymbol{C_S})_+)_+ - L)_+\,,\, U\text{-}L)$$

If a per claim stop loss limit of S were in place then we would need to consider the cost of the claim exceeding the stop loss, as well as the spare capacity on the aggregate limit. If the stop loss is a smaller amount than the 'spare capacity' this results in a larger payout from the insurer to the insured. If the stop loss is a higher amount, then the payment from the insurer to the insured are identical to the case where there is no stop loss. The expected payout under the XOL reinsurance contract in respect to the large claim would be:

$$\text{Min}\,(\,(C_L - \text{Min}(S, (A\text{-}C_s)_+)_+)\ \text{-}L\,,\, U\text{-}L)$$

As you can tell, these equations look complicated when there is only one large claim to consider. Calculating the reinsurance recoveries from the multiple large claims on the same policy can get tricky. This is usually improbable for most aggregate policies where the aggregate limit is well below the XOL reinsurance limit. In the case where multiple claims hit the XOL reinsurance limit, the claims can be ranked from largest to smallest. The insurer pays the full cost for largest claims if the smaller claims cause the aggregate to be breached, and gets the same recoveries from the reinsurer on these that they would if they written the policy as a conventional policy.

When we carry out our claims simulation to estimate the expected claims costs for the aggregate, we can estimate the expected recoveries on the XOL reinsurance contract at the same time. We can then calculate the ratio of the expected cost of claims under the XOL for the aggregate policy relative to the conventional policy, as before. We then apply this ratio to the total cost of reinsurance on the conventional policy to work out the cost of the reinsurance component of the premium for the aggregate policy.

6.3 Catastrophe Reinsurance cost for an alternative policy

On an alternative policy type, a recovery from Catastrophe reinsurance can only be made if:

- Claims arise as a result of a catastrophe
- The claims are above the limits of the policy
- The total cost of catastrophe claims from all policies, conventional or otherwise, exceeds the catastrophe limit, L, up to a limit U-L

.
If we wanted to estimate the expected recoveries from catastrophe reinsurance this would involve constructing a model for catastrophes. This would vastly complicate the modelling process. Unless the catastrophe loading forms a large component of the premium, a subjective approach to setting this rate would usually be carried out.

7. Expense allowance

Expenses usually form a large proportion of the premium for an alternative cover. However, the expenses charged in dollar terms will usually be similar to the expenses charged under a conventional policy. In this paper we will consider underwriting and claims expenses separately. We assume that all expenses (such as overheads) not directly related to claims and underwriting can be allocated to one of these two functions of the insurer.

Claims handling expenses may or may not be charged to the customer depending on whether or not the insurer handles the processing of all their claims. A client may be using an aggregate or XOL policy as an alternative to self-insurance and may decide to handle all its own claims. This is particularly common for comprehensive cover in motor fleets. In this case only a nominal expense for claims will need to be charged to allow for administration of any claims over the limit that arise under the policy. If all claims are being handled by the insurer, it seems reasonable that the claims expenses under the alternative policy should be at least as much as the claims expenses on the conventional policy. Indeed, there may be additional administration involved on the alternative policy that needs to be charged for.

Underwriting expenses should be similar for both the conventional and alternative arrangement. In both cases, the amount of work necessary to assess and process the risk will approximately be the same. We consider it reasonable to charge the same underwriting expense on both risks, although there may be some adjustment of the expenses relating to overheads.

8. Investment income

Claims made under aggregate or XOL policies are usually paid later than claims under a conventional policy. In short tail lines, such as property and motor, the insurer will pay claims on a conventional policy soon after they are reported. For an aggregate policy the insurer has to wait until some time after the policy expires before it can assess whether the total claims cost has gone over the aggregate limit. This means that the insurer usually pays claims later under an aggregate policy than a conventional policy. Delays can also arise under XOLs and aggregates because the large claims that cause the limits to be breached take longer to settle. These delays mean that the percentage allowance for investment return needs to be increased to reflect the difference in timing of claim payments.

The difference in investment income between different types of policies may not be material on short tail lines of business, but it can be more significant in long tail lines. For long tail XOL policies, a simple model of claim duration versus claim size (from the ground up) should suffice to estimate the income component.

We have not had much experience writing aggregates on long tail classes, but the most likely causes of breaching the aggregate limit would be the occurrence of large claims, catastrophes or the underestimation of the risk. Therefore the delay should be intermediate between the delay on an XOL and a conventional policy.

9. Profit allowance (Cost of Capital)

In order to determine a suitable profit allowance, we assume that a rate of return on capital has been set by the company's management. Therefore, we can set the required profit margin by determining the amount of capital that needs to be allocated to the alternative risk.

In order to establish an appropriate capital allocation it will help to step back and consider the purpose of capital. At its simplest level, capital is the difference between a company's assets and liabilities and is used as a buffer to prevent insolvency. Capital will fluctuate as the company's assets and liabilities fluctuate. Some reasons might be:

- Changing value of Investments
- Mis-estimation of risks
- Catastrophes or unusual large claims

It seems reasonable for shareholders to require compensation from policyholders for the fluctuations that arise in the company's liabilities, and perhaps even some of its assets, since the liabilities arise out of the process of writing insurance. If the company invests in risky assets it does not seem appropriate to require higher profit margins from policyholders for these additional risks. Most fluctuations that lead to insolvency will not be limited to one policy, and are usually systematic across all policies (eg. all policies being underpriced for the same reason). Therefore, it can be argued that the cost of capital should be allocated between policyholders in a way that somehow depends on their total claims cost distribution.

In the literature on this subject there does not appear to be any single approach that has universal agreement between practitioners. Some of the approaches suggested by the GISG reinsurance working party (1999), Clark (1996), Mango (2003) and Wang (1997) are:

1. The profit margin is set as some percentage of the standard deviation of the cost of claims

2. The capital is set at a level so that all claims under the contract can be paid from the premium received plus this capital with some pre-defined probability (eg.it is enough to pay claims in 95% of cases)

3. The capital is defined to be the Value At Risk (VaR), calculated as the amount that losses exceed the premium multiplied by the probability of this occurring.

4. The profit margin is set by applying the proportional hazards transform (PH transform) described in Section 4.3.3

5. Capital consumption – each policy has access rights to the entire capital of the insurer. The cost of capital is treated like an overhead expense where each policy pays for the right to access capital depending on the magnitude and likelihood of their potential calls. This amount is calculated by running simulations.

Some of the features of these approaches are:

1. The standard deviation approach ignores parameter risk and considers upside as well as downside risk. Arguably the insurer is not concerned by lower than expected claims (although these cases will help boost capital). It is also difficult to pick an appropriate percentage of the standard deviation to apply as the loading.

2. Setting a capital level to an amount that can pay all claims with a certain probability ignores diversification benefits from writing multiple risks. Also, it is difficult to know which pre-defined probability is the most appropriate.

3. The VaR method may produce a lower profit margin than other methods (see the GISG reinsurance working party paper, 1999)

4. The proportional hazards transform provides profit loadings that charge more for heavy tailed distributions – which is consistent with market observations. The *r*-value for the proportional hazards transform has to be chosen subjectively.

5. The capital consumption framework has many similarities to the VaR concept, except it is not a capital allocation methodology. The capital consumption method considers the timing of capital flows and can take inflation and investment risk into account. In short tail lines the profit loadings from this method will be proportional to loadings produced by VaR.

These methods suggest ways of allocating capital or choosing profit margins that are useful if we are using a ‘first principles’ approach to pricing. However, the answers produced by the first two methods are heavily dependent on the choice of arbitrary standard deviation percentages or probabilities of sufficiency and there is no clear logic to help decide what these should be.

Many of these problems can be circumvented when we consider how these approaches fit into the framework of adjusting the underwriter’s quote to allow for differences in cover. The underwriter will already have a profit margin built into the premium for a conventional policy. First we figure out how this margin should be adjusted for each of the methods above, and then we will evaluate the suitability of each of the methods in this framework. We will consider this for short tail lines only and therefore restrict ourselves to considering the first four methods only.

The required adjustment to the profit margin can be expressed in the form of a ratio of the capital required for the alternative risk type versus the capital required for the conventional policy. This ratio can be applied to the dollar amount of the profit margin built into the conventional policy to get a profit margin for the alternative policy. If $f(x)$ is the PDF of the total claim cost distribution, T, and $F(X)$ is the CDF of T, and $\boldsymbol{E}[T]$ is the expected value of T then the 'profit loading adjustment factor' for an aggregate policy with limit A and no stop loss clauses under each method are:

1.

$$\frac{\sqrt{\int_0^\infty ((x-A)_+ - \alpha)^2 f(x)dx}}{\sqrt{\int_0^\infty (x-\boldsymbol{E}[T])^2 f(x)dx}},$$

where $\alpha = \int_0^\infty (x-A)_+ f(x)dx$

2.

$$\frac{F^{-1}(q)-\alpha}{F^{-1}(q)-\boldsymbol{E}[T]},$$

some probability q, 0≤q≤1

3.

$$\frac{\int_0^\infty ((x-A)_+ - \alpha)_+ f(x)dx}{\int_0^\infty (x-\boldsymbol{E}[T])_+ f(x)dx} = \frac{\int_{\alpha+A}^\infty (x-A-\alpha) f(x)dx}{\int_{\boldsymbol{E}[T]}^\infty (x-\boldsymbol{E}[T]) f(x)dx}$$

4.

$$\frac{\int_A^\infty S(x)^r dx}{\int_0^\infty S(x)^r dx},$$

Where $S(x)$=1-$F(x)$ and 0≤r≤1

We can estimate the value of these ratios from the simulations for the aggregate distribution.

Given that the profit margin is set to achieve a certain return on capital, each of the methods above will require the company to hold different levels of capital for the alternative risk. In order to evaluate which ratio is the most appropriate one to use we need some criteria by which we can establish its suitability. The method chosen was as follows:

We establish the level of capital we need to ensure that the company is solvent with some probability. If we write the risk on a conventional basis we can work out the extra level of capital that is needed to ensure that the company is still solvent with the same level of probability, call this extra capital C. However, if we write this policy as an aggregate with aggregate limit A, the extra level of capital needed to ensure solvency at the required level will be a smaller amount, C_L. The ratio of C_L to C provides a benchmark by which we can assess the other methods. This is valid on the basis that the profit margin for the conventional and alternative risks are set in order to achieve a certain return on capital. The limitations of this benchmark are explored later.

Figure 6 shows a comparison of the ratios obtained for each method for a typical aggregate written by a hypothetical company holding a portfolio of other risks. Values of the ratio are shown for different aggregate limits. The 'aggregate cost as percentage of conventional cost' shows the ratio of the expected cost of aggregate claims (as per Section 5) divided by the expected cost of claims the conventional policy for different aggregate limits. This is essentially equivalent to applying the same percentage profit margin to the conventional and alternative risk. The line 'Benchmark - Equal probability of insolvency' shows the benchmark ratio described in the last section.

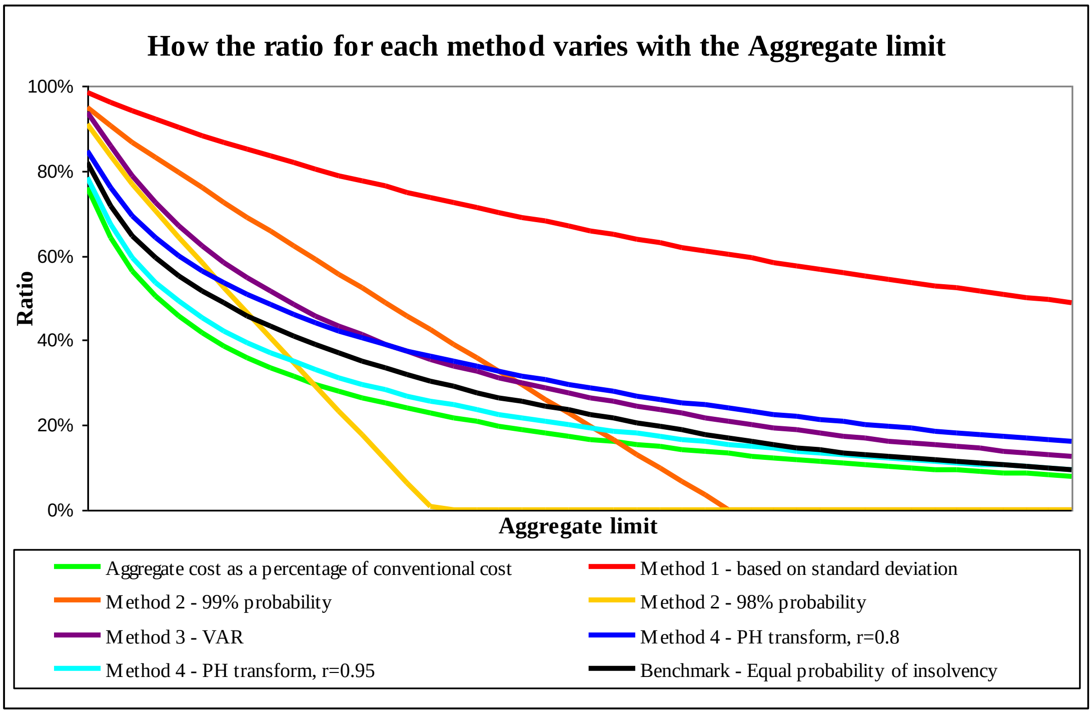

Figure 6: Comparison of various methods of adjusting profit margin

The method based on standard deviations tends to imply a high level of capitalisation and may lead to profit margins that are too large. Method 2 produces a ratio of 0 when the limit exceeds the pre-defined probability and often produces a capitalisation that is too high when the aggregate limit is low. The fixed percentage profit margin (or equivalently the ratio described by the line 'aggregate cost as percentage of conventional cost') produces a level of capitalisation that is too low. The VaR method and proportional hazards method appear to provide a better approximation to the benchmark ratio that the other methods in most of the scenarios that were tested.

The method of adjusting the profit margin of the conventional policy to allow for differences in cover means that we avoid answering the difficult question of how this margin should be determined for a conventional risk. If we have a portfolio of conventional policies then the diversification benefits of writing the whole portfolio are taken into account when setting the margin for one policy. Therefore, if we have a substantial portfolio of alternative risks we can't ignore the diversification benefits that writing alternative risks will add to our portfolio.

The method chosen for setting the benchmark ratio only considers the *marginal capital* required to write an additional risk. There has been a lot written on this subject in the literature and this approach is generally considered too simplistic. Mango (1997) explains how to derive a capital allocation in a game-theoretical framework. The profit margin for the conventional policy should reflect the average 'extra capital' that is needed if we consider what happens when the risks are written in a different order (ie consider the extra capital needed for each order in which risks could be written and average this over all combinations). We could construct a better benchmark by considering the average 'extra capital' needed for the conventional policy versus the average 'extra capital' needed for the aggregate policy if we wrote all the risks in its portfolio in a different order. This will tend to increase the value of the benchmark ratio, and is roughly equivalent to ratio obtained by getting the original benchmark for a portfolio half the size of the existing portfolio - this is because most of the combinations we will average across will have our risk written in the middle.

The benchmark also ignores correlation between risks in the same portfolio, particularly in the tail of distribution. Tail dependence between risks will tend to increase the level of capital required. In conclusion, the VaR method and proportional hazards method appear to be the most appropriate methods for adjusting the profit margin of a conventional policy for differences in risk, but they may produce margins that are too low if there is a lot of tail dependence between risks.

10. Retrospective premium adjustments

10.1 Burners

Meyers (1986) describes a method of pricing burners from first principles – he discounts future expected cash flows from premium adjustments and subtracts the discounted cost of claims and expenses to determine an operating profit. The terms of the burner are then refined to achieve a pre-determined operating profit.

Under the assumption that we already know the price for a conventional policy and we want the same level of profitability for the burner, we can simplify this method by just considering the discounted expected cash flows from the initial premium and premium adjustments and ensuring that they are at least equal to the conventional premium. An allowance for the extra administration expenses associated with burners may also be desired. This is similar to the method adopted by Bender (1994), however in his paper he is not describing a pricing methodology, but is trying to determine a loss ratio for a portfolio of policies with retrospective premium adjustments.

We need to choose a discount rate that reflects the credit risk and other risks (eg. investment risk) associated with the delay in receiving premium adjustments from the insured. The interest rate we use to discount expected cash flows should reflect the credit rating of the client we are offering the burner to. Alternatively we might prefer to discount using a risk free rate and make an explicit assessment of the credit risk associated with not receiving future premium adjustments from the insured due to insolvency.

Loss of investment income may be immaterial on short tail lines and on burners that use appropriate IBNR factors – so the selection of the correct interest rate may not be the most important issue in pricing the burner. Even if the IBNR factors for a burner in a long tail line do not reflect the true level of IBNR and the lost investment income is more significant, the uncertainty surrounding the selection of an appropriate total claims cost distribution is likely to be an issue which is much greater in magnitude.

Using the notation used in the definition of a burner in Section 2, the expected value of $G(FI_i X_i)$ is:

$$Y \int_{-\infty}^{Y/FI_i} f(x)dx + FI_i \int_{Z/FI_i}^{Y/FI_i} xf(x)dx + Z \int_{Z/FI_i}^{\infty} f(x)dx$$

where $f(x)$ is the PDF of X_i, the total reported claim cost in year i.

If we are using Monte Carlo simulation we can calculate the value of the notional premium, FI_iX_i, directly for each simulation then calculate the adjustment premiums this generates and discount these values (subject to the maximum and minimum premiums) to get the NPV of the total expected premium flows. If the simulation generates a return in premium to the client in some simulations, there will be no credit risk associated with this, so these premium flows can be discounted at a risk free rate.

We can adjust the parameters of the burner (the values of C, F, I_i, Y and Z) to get the NPV of the expected premium flows to be equal to the premium of the conventional policy. It is also possible to test the sensitivity of the various parameters to come up with a more robust premium.

If a model is used which doesn't include a random component for the development of clams (especially in the final year) the expected cash flows may not be correctly assessed. However, the uncertainty surrounding the selection of an appropriate total claims cost distribution is usually a bigger issue.

10.2 Claims experience discounts

We can follow the same approach of calculating discounted cash flows when it comes to pricing policies with a claim experience discount. The return in premium is usually calculated at one point in time – so that there is a single (or no) premium adjustment. Using the notation of the definition of a CED from Section 2, the return premium in year i is:

$$\text{Min}(\ \Pi\,(L\,P - \ I_i X_i\,)_+\ ,\ JP)$$

At time i (i might be one or two years depending on the class of business).

The expected value of this cash flow is:

$$\left(LP - I_i \int_{JP/I_i}^{LP/I_i} x f(x)\,dx \right)\Pi$$

Where $f(x)$ is the PDF of X_i, the total cost of claims incurred by time i. If we are estimating this by Monte Carlo simulation we need to calculate Min($\Pi\,(L\,P - I_i\,X_i\,)_+$, JP) for each value of X_i that is simulated and average this amount across all simulations.

Again, we can price this risk by ensuring that initial premium less the expected discounted value of this cash flow is equal to the conventional premium. We can discount the return premium at the risk free rate because there is no credit risk associated with this return of premium.

There are similar issues surrounding the allowance for a random component in the development of claims in CEDs as we had with burners. However, because we are only concerned with claims being better than expected, getting the correct variance of the distribution is more important for CEDs than it is for burners (skewness is more important for burners). In addition, a final CED payment is often assessed at an earlier stage of development than it would be for a burner. Therefore, it could be argued that including a random component for the development of claims is more important for CEDs. Not including a random development factor could understate the expected return in premium to the insured.

References

Papers and books discussing the construction of aggregate claim distributions

Heckman P, Meyers G, 1983, *The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Distributions*, Proceedings of the Casualty Actuarial Society 1983 Vol: LXX pp. 22-61

Panjer H, 1981, *Recursive Evaluation of a family of Compound Distributions*. ASTIN Bulletin 12, 22-26.

Herzog T & Lord G, 2002, *Applications of Monte Carlo Methods to Finance and Insurance*, ACTEX publishing

Klugman S, Panjer H & Willmot G, 2004, *Loss Models: From Data to Decisions, 2nd Edition*, Wiley series in probability and statistics

Wang S, 1998, *Aggregation of Correlated Risk Portfolios: Models and Algorithms*, Proceedings of the Casualty Actuarial Society 1998 Vol: LXXXV pp. 848-939

Other papers

Bender R, 1994, *Aggregate Retrospective Premium Ratio as a Function of the Aggregate Incurred Loss Ratio*, Proceedings of the Casualty Actuarial Society 1994 Vol: LXXXI pp. 36-74

Meyers G, 1986, *The cash flows of a retrospective rating plan*, Proceedings of the Casualty Actuarial Society 1986 Vol: LXXIII pp. 113-128

Meyers G, 1980, *An Analysis of Retrospective Rating*, Proceedings of the Casualty Actuarial Society 1980 Vol: LXVII pp. 110-143

GISG Reinsurance Pricing Working Party, 1999, UK General Insurance Conference: http://www.actuaries.org.uk/files/pdf/library/proceedings/gen_ins/1999gic/reinsprice.pdf

Clark D, 1996, *Basics of Reinsurance Pricing*, CAS study note: http://www.casact.org/library/studynotes/clark6.pdf

Mango D, 2003, *Capital Consumption: An Alternative Methodology for Pricing Reinsurance*, 2003 ASTIN Colloquium

Mango D, 1997, *An Application of Game Theory: Property Catastrophe Risk Load*, Casualty Actuarial Society Forum 1997: Spring pp. 31-51

Wang S, 1997, *Implementation of PH-Transforms in Ratemaking*, Discussion paper for 1997 CAS ratemaking Seminar