

Defence against the Dark Arts

A Practical Handbook for Reserving Actuaries

Paper submitted for GIRO 2008

Author: Lis Gibson

Date: 25 July 2008

Defence against the Dark Arts

A Practical Handbook for Reserving Actuaries

Contents

1. Introduction	3
2. The Seven Deadly Sins of Reserving	4
3. Actuarial Reserving Personality Types	7
4. A Guide to Duelling – aka Peer Review	11
5. Customised Defence Against the Dark Arts	15

1. Introduction

Purpose

This short note has been written with two aims in mind; firstly to provoke thought and debate about how we each use judgement in reserving, and secondly to look again at some of the reserving pitfalls we can encounter, in the context of our own personal styles. By understanding ourselves better as individuals, it is to be hoped that we will better appreciate the traps we are most likely to fall into, and better equip ourselves through self-challenge and peer review to produce sound reserving judgements.

Scope

I have gathered a selection of common pitfalls and, for the sake of making this potentially dry subject a bit more fun, have endeavoured to link these pitfalls to the seven deadly sins. This does not mean that there are only seven pitfalls – or even seven broad categories – and you will be pleased to read that I have resisted the urge to try to force everything to begin with the same letter – but hopefully there will be enough food for thought.

I have then developed a 3 dimensional model for actuarial reserving personality type which uses three criteria to segment actuaries into 8 core reserving personality types. This is slightly tongue in cheek, but only slightly, and I hope readers will find the model a useful tool in understanding their own habits and dispositions in reserving, as well as those of others. I have discussed which deadly sins the different personality types might be most prey to, and I hope that in the GIRO workshop session we will be able to develop and debate this discussion.

Next I have developed a mathematical model to simulate the behaviours of the 8 core reserving personality types when faced with a reserving exercise. I have also developed a model for the interaction between actuaries in challenge and peer review and have modelled the full range of interactions available. The results are summarised in this note, with some observations and conclusions, and the model itself can be found, in an Excel file, on the Institute website, should anyone wish to play with it.

Finally I have made a few observations on how the individual can take some of these themes and develop their own customised defence mechanisms, tailored to their own style and habits, making the best use of peer review and challenge.

2. The Seven Deadly Sins of Reserving

SALIGIA

The seven deadly sins, also known as the capital vices or cardinal sins, are a classification of vices that were originally used in early Christian teachings to educate and instruct followers concerning (immoral) fallen man's tendency to sin. The Roman Catholic Church divided sin into two principal categories: "venial", which are relatively minor, and could be forgiven through any sacramentals or sacraments of the church, and the more severe "capital" or mortal sin. Mortal sins destroyed the life of grace, and created the threat of eternal damnation unless either absolved through the sacrament of confession, or forgiven through perfect contrition on the part of the penitent. Beginning in the early 14th century, the popularity of the seven deadly sins as a theme among European artists of the time eventually helped to ingrain them in many areas of Christian culture and Christian consciousness in general throughout the world. One means of such ingrain was the creation of the mnemonic "SALIGIA" based on the first letters in Latin of the seven deadly sins: superbia, avaritia, luxuria, invidia, gula, ira, acedia. In English we translate these as Pride, Greed, Lust, Envy, Gluttony, Wrath and Sloth.

Deadly Sin 1: Over-Confidence

"I know what I am doing" - Pride

We can sometimes become overly confident in the quality of the data, the efficacy of our chosen method or our own personal experience or judgement. We are probably able to recite, or at least dredge up from our memory, the key limitations of our favourite reserving methods, and I won't repeat them here, but are these in the forefront of our mind when we are reserving? Do we persuade ourselves that the answer is "right" once we can see the lines converge on the development graph, or do we challenge ourselves on what we are really assuming when we adopt a particular method and whether we have good enough reason to make such assumptions?

Deadly Sin 2: Over-Granularisation

"I want 140 reserving classes, each in 3 currencies, then take out problem accounts, large losses and Cats" – Greed

The technical problem here is obvious – by subdividing to the n th degree we can end up by taking out all of the kicks and bad performance and leaving the expectation of beautiful and well behaved loss experience in the future. Many data sets are already too small to provide much help on the extreme tail of the distribution of all possible ultimate claims outcomes which we are seeking to find the mean for – throwing away examples of unusual behaviour will result in an even more distorted analysis.

Knowing the issue, though, is only part of it. It can be hard to resist the drive for increasing granularisation, especially when there is disappointingly poor loss experience. There is a tendency to want to identify the culprits and declare the matter resolved. We need to not only be aware of this risk but to actively challenge ourselves about it in practical terms.

Deadly Sin 3: Love

“Roses are red, violets are blue, I’ll drop my reserves if you’re asking me to” – Lust

OK, so this deadly sin was always going to be a bit of a reach, but there is a serious point here. Sometimes we can be too eager to please – to build a relationship internally in our company or externally with a client. Of course, we are all aware of this danger but in practice it can be difficult to diagnose when there has become a problem.

One way in which this issue might manifest itself is in a process of **Limbo Reserving**. One actuary produces an estimate which under challenge is reduced, more data is provided and the estimate is reduced further, more data, further reduction and so on. At each point the incremental change may be small, so that while the data supporting it is perhaps weak, the change is agreed to for the sake of good relations. Once one change is “banked” the process can begin again with new data and a new target reduction. Both sides can be behaving with integrity from their perspective – with the provider of data seeing their role as correcting the actuary’s incomplete understanding, but the result can be that actuary is manoeuvred away from a position in which they are genuinely comfortable professionally. This process can interact powerfully with over-granularisation.

One tip to guard against this is to have an implicit or explicit Limbo Metre, which keeps a record of the number of iterations and triggers self-challenge or peer review if the revised reserve estimate deviates more than a certain amount from the earliest estimate.

Deadly Sin 4: Inappropriate Benchmarks

“I wish my reserving class was like yours” – Envy

We can apply benchmarks implicitly or explicitly. Explicit benchmarks are often used when we are short of data and we make use of them on the grounds that we feel we have no alternative. In such circumstances it can be fairly easy to maintain a healthy degree of scepticism about their appropriateness, although we may still imbue the resulting calculations with an undue status simply because they are our work product.

Implicit benchmarks can be trickier to manage. In some ways the application of judgement based on experience can be thought of as a form of implicit benchmarking, and used with care this can be a valuable tool on the actuaries’ belt. However, used unconsciously we can distort our analysis by the use of implicit benchmarks, encouraging data or analysis to behave as we expect it to. Being conscious of our use

of implicit benchmarks, and giving careful consideration to how and when they were formed, will help us to use them in a positive manner.

Deadly Sin 5: Scattergun

“Let’s do paid, incurred, numbers, amounts paid and settled, closed, open, re-opened, stochastic, case reserve strength, exposure method, BF, ELR and the dartboard method” – Gluttony

Now I’m not saying that using lots of methods is a bad idea. I would encourage actuaries to make good use of the data they have and to use multiple methods to cast light on different aspects of the data and the trends it displays, as well as for diagnostic purposes. Indeed, if the range of methods being used explores all aspects of the data then choosing an estimate based on an average of many methods, each applied in a reasonable manner, can have statistical merit in choosing a mean.

That said, producing lots of rather different results without diagnosing the reason for the differences may not be the most effective way to produce a meaningful best estimate. Using methods which you believe to be appropriate, suitable to the data and which you can interpret mindful of their limitations, is more important than volume.

Deadly Sin 6: Resistance to Challenge

“I’m right and you’re wrong” - Wrath

If I were a psychologist perhaps I would be tempted to say that the third and sixth sins are two sides of the same basic “insecurity” coin, so it is fortunate that I am not.

It is as important that we are able to challenge ourselves and be challenged by others, as it is that we are able to avoid being steamrollered. In both cases, we need to come to a position where we know what judgements we are making and we are clear why we are making them, and are confident that we have the data or information to support those judgements. Whether the judgements were originally our own or those of another should not impact their acceptability to us, but rather our view of their merits.

Deadly Sin 7: Anchoring Error

“What did we do last time?” – Sloth

Anchoring error is everywhere. It’s our irrational preference for last time’s judgement over this time’s, even when we usually have more data this time than last time. It’s hard to shake it off. In the age of models and spreadsheets it can become embedded into our reserving process at the micro level. We can see it in peer review and challenge, when there is a reluctance to change a judgement, and perhaps a reluctance of the peer reviewer to suggest a change; a kind of anchoring error by proxy. There is good material available on this phenomenon already which I will not repeat here, but I would encourage reserving actuaries to be alive to this issue.

3. Actuarial Reserving Personality Types

I have developed three criteria below: disposition, experience and persuasion. I have assumed that in reserving actuaries are attempting to estimate the true mean, and that, for any given class and dataset, they each have a distribution of best estimates with a mean and standard deviation.

Disposition – Optimistic or Pessimistic (O/P)

There is a scale on which we all lie which determines our general disposition – although this may vary depending on which class of business we are reserving for and potentially on other factors. For simplicity, I have assumed that we are all hard-wired into a disposition and that we carry this tendency at all times. I assume that optimistic actuaries tend to produce reserving estimates below the true mean for each class, with the amount of error being random and related to the general degree of uncertainty for the class in question and the extent of the optimism of the actuary; similarly for pessimism.

The optimist might be more likely to get trapped by Limbo Reserving, whereas the pessimist might be resistant to challenge.

Experience – Novice or Veteran (N/V)

Our experience, particularly through multiple market cycles, seeing surprising improvements, surprising deteriorations, and tracking the same portfolio over a long period of time, can teach us more about reserving than simple methodology. We can think of this as the creation of a vast personal benchmark database. For simplicity I have assumed that we are each equally experienced or inexperienced in all classes of business. I have assumed that veterans are able to produce reserve estimate with lower standard deviations, whereas novices produce more volatile estimates.

Persuasion – Malleable or Stubborn (M/S)

This criterion is all about the reaction of the actuary to challenge. A malleable actuary will tend to respond to challenge by adopting the point of view of the challenger, regardless of whether they have a better argument, depending on his degree of malleability. The stubborn actuary will similarly stick to his own position, regardless of the merits of any challenge, again depending on the degree of his stubbornness.

The persuasion criteria relates directly to the deadly sins of “love” and “resistance to challenge”.

The Eight Archetypal Reserving Actuaries

By combining the 2 extremes for each of the 3 criteria we can generate 8 archetypal reserving actuaries. You might find it helpful to think about which type of actuary you are:

Actuary A: OVM: This experienced actuary has an optimistic outlook and is willing to change their mind.

Actuary B: OVS: This experienced actuary has an optimistic outlook and is resistant to challenge.

Actuary C: ONM: This novice actuary has an optimistic outlook and is willing to change their mind.

Actuary D: ONS: This novice actuary has an optimistic outlook and is resistant to challenge.

Actuary E: PVM: This experienced actuary has a pessimistic outlook and is willing to change their mind.

Actuary F: PVS: This experienced actuary has a pessimistic outlook and is resistant to challenge.

Actuary G: PNM: This novice actuary has a pessimistic outlook and is willing to change their mind.

Actuary H: PNS: This novice actuary has a pessimistic outlook and is resistant to challenge.

Modelling the Judgements of the Archetypal Actuaries

I have developed a model to look at and compare the reserving judgements of the 8 archetypal actuaries. The model is contained in an Excel spreadsheet and can be found on the Institute website.

I have made up a portfolio of 10 fictitious reserving classes each with a “true mean” reserve which represents the mean of all possible reserve outcomes, and each with an uncertainty level – which is higher for more volatile classes and is expressed as a percentage. The uncertainty level applied to the true mean reserve generates what I call an uncertainty amount and can be thought of as a kind of standard deviation about the true mean. The uncertainty level varies from 2% to 20% of the mean reserve for my classes and the total true mean across the ten classes is 100,000. The full data can be found in the spreadsheet model.

For each of the three criteria I have developed a scale which determines the degree to which an attribute is present in a given Actuary and have set assumptions for each of the archetypal actuaries.

For each actuary and for each class the reserve estimate they produce is the result of a random process. I have assumed that the distribution of estimates they might produce is normal, with a mean and standard deviation driven by the actuary's disposition and experience, combined with the uncertainty level for the class of business, as follows:

Mean for Actuary A for Class n =

$$\text{True Mean (n)} * \{ 1 + [\text{Uncertainty Level (n)} * \text{Disposition (A)}] \}$$

The disposition factor is negative for optimists, resulting in a mean below the true mean, related to the uncertainty level, and the converse is true for pessimists. For actuaries intermediate between optimism and pessimism the factor is zero and the actuary's mean is equal to the true mean for the class of business.

Standard Deviation for Actuary A for Class n =

$$\text{True Mean (n)} * \text{Uncertainty Level (n)} * [1 + \text{Experience (A)}]$$

The experience factor is negative for veterans (minimum is -50%) resulting in a standard deviation below the uncertainty amount of the class in question, and for novices it is positive (maximum is 50%) resulting in an increased standard deviation. For actuaries intermediate between novice and veteran the experience factor is zero and the standard deviation for the actuary is equal to the uncertainty amount for the class of business.

The table below shows the assumptions I have made for the factors for each of the archetypal actuaries.

Table 1: Archetypal Actuaries

Actuary	Type	Disposition	Experience	Persuasion
A	OVM	-50%	-25%	-50%
B	OVS	-50%	-50%	100%
C	ONM	-100%	50%	-75%
D	ONS	-100%	25%	75%
E	PVM	50%	-25%	-50%
F	PVS	50%	-20%	100%
G	PNM	100%	50%	-75%
H	PNS	100%	25%	75%

I have used the above equations, the table of factors and normal distributions based on 1000 simulations to generate the reserving judgements of the 8 archetypal actuaries for my 10 fictitious classes of business. The full results can be found in the Excel model. For each actuary I have recorded the mean of their 1000 simulations for each

of the ten classes, and in the aggregate, as well as the standard deviation for each class and for the aggregate.

The table below shows the mean and standard deviation for each actuary for their aggregate reserve estimate across the 10 classes of business from the 1000 simulations I performed.

Table 2: Mean and Standard Deviation by Actuary

Actuary	Type	Mean	SD
A	OVM	96,276	3,724
B	OVS	96,213	3,787
C	ONM	92,486	7,514
D	ONS	92,507	7,493
E	PVM	103,769	3,769
F	PVS	103,821	3,821
G	PNM	107,613	7,613
H	PNS	107,586	7,586

The optimistic actuaries are producing means below the 100,000 true mean, and the pessimistic actuaries are higher. The novices have higher standard deviations than the veterans. The differences between the outcomes for the matching pairs of malleable and stubborn actuaries (A&B, C&D, E&F, G&H), is due to the randomness of the simulation process.

In the section which follows I will consider what happens when these actuaries start to interact with each other.

4. A Guide to Duelling – aka Peer Review

In the discussion which follows I am going to contrast what I define as “challenge” with what I define as “peer review”. For the purposes of this note I define challenge as the process by which say Actuary B challenges Actuary A to revisit his estimates in the light of the estimates of Actuary B. The outcome is a revised position for Actuary A, not a consensus, and the amount by which Actuary A is willing to move his estimates in response to Actuary B is determined by the Persuasion factor of Actuary A.

Challenge

The most stubborn actuary has a persuasion factor of 100% and he will not change any of his estimates under challenge. The most malleable actuary has a persuasion factor of -100% and he will abandon his own reserve estimates and adopt those of his challenger fully. An actuary with an exactly intermediate position between stubborn and malleable has a persuasion factor of 0% and would adopt the midpoint of his own estimate and that of his challenger. The outcomes for other percentages are achieved by pro rata.

Perfect Peer Review

I have introduced the concept of a perfect peer review. In this Actuary B peer reviews Actuary A. In this perfect peer review the two actuaries are able to share their judgements and together adopt the better of the two estimates for each class, that is, the one nearer to the true mean reserve. The outcome is a consensus and so the perfect peer review process, unlike challenge, is commutative by which I mean that the outcome for B peer reviewing A is the same as for A peer reviewing B.

Of course the perfect peer review is not possible in real life as we are all human beings, and so I have also developed the ideal of an experience based peer review.

(Experience Based) Peer Review

Here Actuary B peer reviews Actuary A. The outcome of the experienced based peer review of A by B lies between that of the challenge of A by B and the perfect peer review of A by B, and its position between these two outcomes depends on the combined experience of the two actuaries. Where both actuaries are extreme veterans, with a combined experience level of 100%, the outcome is the perfect peer review. For two extreme novices with a combined experience level of -100%, the outcome is the challenge of A by B. For actuary pairs with a combined experience level of zero, the outcome is the midpoint between the perfect peer review and the challenge of A by B, and all other permutations are by pro rata.

My premise in this definition is to suppose that the experience of both reviewer and the actuary being reviewed will input into the success of the exchange of judgements and understanding.

Using these definitions and my simulations of the reserving estimates of the 8 archetypal actuaries I have simulated the challenge, perfect peer review and experience based peer review of each pairing, by class of business.

I have defined the absolute error amount to be the absolute value of the difference between an actuary's mean and the true mean reserve for a given class of business, and for the aggregate of the ten classes. To examine the effectiveness of the three forms of interaction I have looked at the improvement in absolute error due to the process, and the improvement in standard deviation due to the process. In both cases reductions are "good" so result in positive improvements, expressed as a percentage of the original value.

The tables below summarise the results. The primary actuary is the actuary carrying out the reserving exercise initially and the secondary actuary is the reviewer.

Table 3: Improvements in Actuarial Estimates due to Challenge

Challenge Outcomes: Improvement in Mean								
Primary Actuary	Secondary Actuary							
	A	B	C	D	E	F	G	H
A	0%	1%	-63%	-72%	73%	72%	14%	8%
B	0%	0%	0%	0%	0%	0%	0%	0%
C	45%	44%	2%	2%	69%	68%	21%	22%
D	6%	6%	0%	0%	19%	19%	25%	25%
E	55%	51%	-31%	-24%	0%	2%	-81%	-80%
F	0%	0%	0%	0%	0%	0%	0%	0%
G	69%	68%	24%	26%	43%	44%	0%	-1%
H	19%	19%	25%	25%	6%	6%	0%	0%

Challenge Outcomes: Improvement in Standard Deviation								
Primary Actuary	Secondary Actuary							
	A	B	C	D	E	F	G	H
A	27%	45%	-28%	-30%	28%	23%	-29%	-9%
B	0%	0%	0%	0%	0%	0%	0%	0%
C	55%	69%	16%	28%	56%	52%	11%	28%
D	12%	12%	10%	12%	12%	12%	11%	12%
E	19%	43%	-49%	-31%	23%	17%	-54%	-24%
F	0%	0%	0%	0%	0%	0%	0%	0%
G	54%	68%	9%	22%	55%	52%	11%	24%
H	12%	13%	11%	12%	12%	12%	11%	12%

Actuary A, a veteran optimist, can improve his mean outcome when challenged by a pessimist, but is susceptible to damage to the quality of his work when interacting with novices, especially fellow optimists. Similarly Actuary E, a veteran pessimist, can benefit from challenge from optimists, but can be led astray by novices, especially fellow pessimists.

Actuaries B and F are supremely stubborn and so impervious to challenge.

Actuaries C and D are novice optimists and their estimates have sufficient room for improvement that they benefit from challenge from any of the others, with C benefiting more due to his greater malleability.

Similarly actuaries G and H, the novice pessimists, can benefit from challenge by any of the others, with the more malleable G benefiting more.

Table 4: Improvements in Actuarial Estimates due to Perfect Peer Review

Perfect Peer Review Outcomes: Improvement in Mean								
Primary Actuary	Secondary Actuary							
	A	B	C	D	E	F	G	H
A	54%	49%	38%	38%	98%	96%	61%	66%
B	50%	49%	31%	32%	79%	76%	43%	48%
C	71%	66%	57%	54%	79%	81%	97%	90%
D	69%	66%	53%	53%	84%	87%	90%	100%
E	97%	79%	60%	69%	58%	57%	40%	37%
F	97%	73%	62%	74%	55%	57%	40%	40%
G	79%	70%	100%	92%	69%	69%	54%	54%
H	85%	75%	92%	99%	70%	70%	53%	53%

Perfect Peer Review Outcomes: Improvement in Standard Deviation								
Primary Actuary	Secondary Actuary							
	A	B	C	D	E	F	G	H
A	32%	45%	17%	15%	28%	25%	11%	11%
B	11%	26%	2%	2%	-2%	1%	-2%	1%
C	57%	68%	35%	40%	57%	52%	24%	31%
D	51%	60%	25%	30%	45%	41%	16%	23%
E	24%	34%	6%	4%	33%	30%	15%	17%
F	32%	42%	12%	12%	34%	34%	16%	18%
G	55%	64%	24%	28%	58%	54%	32%	37%
H	44%	57%	21%	19%	50%	47%	25%	31%

All of the actuaries benefit from perfect peer review in almost all combinations both in terms of mean and standard deviation.

(Note: while the outcome of the perfect peer review is symmetrical between the primary and secondary actuary, the table of percentage improvements is not because these are measured against the starting point for the primary actuary).

Table 5: Improvements in Actuarial Estimates due to Experience Peer Review

Experience Peer Review Outcomes: Improvement in Mean								
Primary Actuary	Secondary Actuary							
	A	B	C	D	E	F	G	H
A	41%	43%	-25%	-17%	88%	89%	61%	71%
B	44%	49%	15%	20%	69%	65%	21%	30%
C	55%	55%	2%	9%	72%	73%	21%	30%
D	38%	44%	7%	13%	68%	64%	33%	44%
E	91%	75%	33%	53%	44%	41%	-32%	-17%
F	75%	92%	22%	35%	40%	40%	14%	19%
G	73%	69%	24%	34%	53%	53%	0%	6%
H	67%	85%	34%	44%	38%	36%	7%	13%

Experience Peer Review Outcomes: Improvement in Standard Deviation								
Primary Actuary	Secondary Actuary							
	A	B	C	D	E	F	G	H
A	37%	47%	3%	9%	41%	39%	13%	28%
B	16%	26%	9%	11%	9%	12%	13%	17%
C	59%	71%	16%	33%	63%	59%	11%	36%
D	42%	52%	16%	23%	50%	48%	21%	31%
E	37%	41%	1%	21%	37%	34%	-12%	13%
F	43%	50%	17%	23%	36%	35%	14%	18%
G	61%	72%	9%	30%	59%	56%	11%	29%
H	50%	61%	20%	30%	43%	39%	16%	24%

In almost all cases the actuaries' outcomes improve as a result of this experience related peer review. The exceptions are Actuaries A and E, the veteran malleable actuaries who can be led astray by their novice fellow optimists and pessimists respectively.

5. Customised Defence Against the Dark Arts

Know Thyself

Which archetypal actuary best described you? Why not rate yourself on the three scales and create your own persona. Feel free to play with the model to see how your personality type might interact with others. See whose reviews benefit you most and which could actually damage you.

Which of the deadly sins are you most prey to? What steps can you take to incorporate some defences into your normal working practices?

Actively Improve Your Peer Review

Whilst my modelling here is in part a bit of fun, one thing it does illustrate clearly is the benefit which can be gained from peer review, and the greater benefit which can be gained from high quality peer review. We should practice, both being reviewed and reviewing, and get better at identifying the key issues and drivers of difference and exchanging experience and judgement. No doubt developing in these areas will help our general communication skills in relation to reserving, but they will also help the quality and robustness of our work.