

*Counting the uncountable: Quantification of risk in risk management and insurance*¹

Leigh Roberts, Victoria University of Wellington²

1. Introduction

Whatever interpretation is placed upon risk management, and there is an extraordinary variety of them, a common theme is the quantification of risk. Whether the risk in question is at the level of a company guarding against industrial risk, or a company in the financial sector involved with highly leveraged securities deals, or whether the risk is at the individual or societal catastrophic level, one seeks in some way to quantify the risk.

Just to define a risk precisely is hard enough, especially with the modern proactive approaches to risk management. These more refined risk management techniques emphasise the complementarity of risk and uncertainty, and take the view that uncertainty provides opportunity as well as risk. The emphasis has shifted away from easily identifiable "hard" risks, such as breakdown of physical goods and accidents, towards soft risks, such as reputational and human resource risks, and especially communication risks. The prevailing view in modern risk management is that one needs to put assets at risk to achieve one's aims (McNamee (1997)). One is not so much modelling risk, but uncertainty, and the environment impacting on the organisation at risk also needs to be modelled in some way. The extension of risk management into softer risks has made the task of modelling and quantifying risk that much harder.

It is a truism in management to say "what you can't measure you can't manage". If by "measure" is meant anything more than the broad comprehension and understanding of the problem at hand, the maxim as a generality is a nonsense. The rather pessimistic title of this paper may overstate matters somewhat; neverthe-less, effective measurement of risk is the exception rather than the rule.

The most basic paradigm for a model of riskiness is a categorisation into heavy/medium/light. This is often applied to the overall risk in question, but it is preferable to consider frequency and severity separately³, each subdivided into the three categories. Then one can consider perils, which would be very serious if they occurred, but are most unlikely to occur, separately from risks which occur much more often, but are less serious in nature.

Such broad categorisations are indeed models of risk, and may be as far as one can go or wishes to go for reasons of practicality or cost; but it is hardly what we have in mind here when we speak of quantification of risk.

By quantification we mean ascribing a number to a risk, however approximate that number may be. The quantification or measurement of a risk may take place

through a mathematical model, or it may simply be inferred from the workings of a market. Such a market would normally be an insurance market, *de jure* or *de facto*, i.e. there could be a formal insurance market in the class of risk, or there could be a pool operating for swapping risk.

By the word model we have in mind a mathematical process for quantifying risk. Such models will often encompass statistical equations or systems of differential or recursive equations, and the like; at least they will typically be expressed in mathematical terms. What in fact those models will produce is some sort of probability distribution of outcomes, either because the models are statistical in the first place, or because one can substitute different values of parameters into a deterministic model for a sensitivity test.

Producing a statistical distribution to describe risk is unfortunate. What the world demands, or at least what corporate players and many others demand, is a number, not a confidence interval or even worse an entire distribution. A further strong tendency in the modern world is to get someone else to deliver the number, so that one cannot be blamed when it turns out to be wrong: it is the age of the consultant.

With the constant clamour to quantify risk, it is easy to forget that one should only do so as long as it is cost effective. The task of knowing when to stop in the modelling effort is not made easier by modern society's exaggerated trust in numbers of whatever sort.

One indication of this trend is in the vast numbers of surveys constantly being undertaken. The subject matter of the surveys is usually of little consequence; and the surveyors have as a rule little idea of the

technicalities of survey design, and even less idea of the statistical skills needed to decode the results sensibly. Just as noticeable in recent years is the rise in popularity of market research: better to be politically correct to the *n*th degree, than to be seen to make decisions without consulting the populace, however ill-informed.

Perhaps the trend to specious numeracy is merely an indication of how very difficult it is to grapple with problems in modern society, or of how difficult it is to obtain any solutions. Paradoxically, if by mathematics one understands anything more conceptual than mere number crunching, such a trend takes place against a background of what seems to be an anti-mathematical bias amongst students and society generally.

And there are other manifestations of the trend. In the constant frenzy to demonstrate that one is achieving superior output with fewer resources, clutching at numbers which support one's argument is an art form widely practised in both private and public sectors. Most output in an increasingly service-based economy is essentially unquantifiable; but rather than recognise this basic truth, argument so often descends into the presentation of numbers of a trivial kind. A social welfare agency, for example, will cite how many people qualify for assistance, and calculate reductions in how long the beneficiaries wait on average in the queue to see an adviser. The postal authority will make a great play of noting that the average time for arrival of a first class letter is falling, and so on.

It is not that the data cited are totally irrelevant, merely that they are generally incidental to arguments about whether the organisation is actually achieving its stated objectives. Mission and strategy statements and the like are couched in language which

implies that outputs and outcomes⁴ are measurable since they are to be purchased, by internal transfers if not on an open market.

The emphasis on relatively trivial numbers because they can be established with certainty, and of course neglecting those figures which are not supportive of one's cause, is a form of risk management in itself. It is attempting to control the communication risk, but it is not quantification of risk.

Given the propensity to spurious quantification of modern society, how odd then that at the same time there is widespread dissatisfaction with the traditional measures of commercial value and profitability in national and company accounts. As regards the former, the existence of black or unofficial economies is by no means restricted to the developing world: several developed countries have a shadow economy amounting to about 20% of the official economy (Economist (1999a)).

As for the latter, it is increasingly recognised that the measures of equity capital, depreciation, operating profit, etc. in a company's accounts may be close to meaningless (see *i.a.* Economist (1999b)). This view is only enhanced when investment concerns apply huge leverage to bolster the returns on capital, in which case the tiny stock of residual shareholder worth is next to useless, save for being used as the denominator in calculating the incredibly high returns such organisations may achieve over their short but glorious careers before coming to grief. One is thinking particularly of hedge funds, which have nothing to do with hedging; they are investment companies, leveraged to a quite extraordinary degree, and apparently not subject to supervision of any kind (Greenspan (1998)).

In the best known case, that of the *Long Term Capital Management* (LTCM) fiasco, the size and leverage were such that the creditors felt it advisable, with the strong encouragement of the U.S. Federal Reserve, to bail LTCM out as soon as they hit real trouble in order not to panic the financial markets. Thus LTCM achieved incredibly high returns simply because of its tiny capital equity base. When the crunch came, as it must with such risky investment strategies, there was no downside, since they were almost immediately rescued.

Some of the original partners in LTCM are now trying to wrest control of the fund back from their creditors, arguing that their sudden fall from grace was nothing but a "thousand year storm" blip in the market (Economist, May (1999b)).

Such "thousand year" events impacting on financial markets do seem to crop up quite often: one thinks of the Mexican bail-out in 1994 and the Asian crisis in 1997/98, as well as the default on Russian government bonds in 1998 that was the proximate cause of LTCM's demise. Other well-known hedge funds have recently suffered substantial losses too (Martinson (1999)).

The whole sorry episode illustrates the need for government regulations which force such concerns to keep authorities informed. One could almost demand that companies indulging in this sort of horseplay mark to market⁵ daily and post the results on the internet, also daily. Such a move would not do much for commercial sensitivity, but the public has a right to protection *ex ante*. One of the main spurs to quantification of risk in the financial sector is the need to measure more accurately what these types of investment organisations are up to, assuming

that regulators or anyone else manage to prise any information out of them at all.

This paper first asks why it is that one needs to measure risk. The next section outlines the conditions to be satisfied for the effective modelling of risk. Then follows a discussion of the applications of mathematical models in insurance, with particular emphasis on the updating of models over time. Although couched in the language of insurance, the principles apply equally well to risk management more generally.

2. Why Measure Risk?

A first comment is the trite observation that if we wish to reduce risk, it would be desirable in some sense to measure it before and after. Quantification of risk would clearly be advantageous in assessing the efficacy of risk management programmes, and indeed in evaluating the performance of risk managers themselves (for the difficulties of measuring their performance, see *i.a.* Head & Horn, (1997)).

More fundamentally, however, either risk is increasing in our ever more vulnerable closely packed urban life styles, with heightened vulnerability to natural hazards and systems breakdowns of whatever sort, or we are becoming more aware of risk. As far as the financial sector is concerned, we are certainly more susceptible to malfunctioning markets with large numbers of players pursuing enormous gains at the expense of taking high risks, as discussed briefly in Section 1.

We consider measurement of risk in the context of a standard framework of risk management. Note that we are not considering *ex post* disaster management or damage limitation; we are instead playing "what if" games before disaster strikes, and

trying to position that process within a sensible framework for discussion.

The risk management template

Most risk management manuals utilise something like the following schema (Roberts, (1999), Head & Horn (1997), RMS (1999)):

1. For whom am I acting? What are their aims?
2. Identify risk.
3. Assess risk:
 - evaluate, quantify;
 - prioritise.
4. Deal with risk:
 - reduce or remove;
 - retain or transfer.
5. Consider overall position:
 - monitor, review;
 - communicate.

The need for quantification of risk is clear from steps 3 and 4 in the above table. We quantify risk in order:

- the more accurately to assess the importance of that risk and set priorities with conflicting uses of a limited budget;
- the more effectively to evaluate the options available under step 4;
- the more effectively to transfer risk; and
- the more accurately to estimate the burden of the retained risk.

In short, we basically want to measure risk in order to assess better how serious the risk is

on the one hand, and how best to deal with it on the other. In the former step the important elements are assessment and prioritisation; the quantification is really ancillary to those two. Similarly for the second part of the process, that of dealing with the risk, quantification is a step in the process of rendering the retention and transfer of risk effective. The priority is not to quantify risk; rather it is to assess a risk in a holistic fashion, viz. in the context of the overall position of the organisation bearing risk.

The five points listed above are discussed in order below.

For whom am I acting? What are their goals?

Identification of the party for whom one acts is not always clearcut. An example is afforded by the actuary for a proprietary life insurance company, who has responsibilities to the policyholders, the shareholders and the company board. Conflicts of interest arise under normal circumstances when insurance liabilities are valued and bonuses declared, simply because higher bonuses declared on traditional life insurance business mean that much less profit for the shareholders, and conservatism in valuing life insurance liabilities translates into delay in profit reaching the shareholders;⁶ but far more serious again are such conflicts under circumstances that are abnormal.

An example of the latter is the recent pensions mis-selling scandal in the UK. In the rush to privatise pensions in the latter half of the 1980s, the Thatcher government encouraged people to transfer out of employer pension schemes. A large number of financial advisers advised clients to forgo lucrative benefits in order to buy private pensions of lower benefit, thereby providing

themselves with very handsome commissions, see Lang (1995).

All of those individuals affected adversely are now to be restored to the position they were in under the previous arrangements. The retribution is incredibly complex, time-consuming and expensive, and the insurance companies are still being fined for tardiness in compliance. There was a dramatic slump in life insurance sales in the UK when the scandal emerged, the timing of which roughly coincided with the introduction of regulations providing for greater disclosure and other tightening of issuing conditions on UK life policies (Sigma (1999, p. 28)). The stigma remains, or at least the scandal still attracts substantial media interest (e.g. Levene, 1998)).

It is difficult enough to calculate the amounts needed for retribution. The real crunch however arises when it comes to deciding who will pay: the policyholders, through the surplus accumulated by past generations of policyholders within the life insurance fund (the *estate*), or the shareholders. The life actuary is placed in an invidious position indeed, on which Ferris (1999) has a good discussion.

Nor can it be taken for granted that the aims of a commercial organisation can be summed up as maximising profit. That may or may not be so: in any case any such aim is vacuous without a specified time frame. Such generalised goals, however lofty in nature, are not necessarily of much operational utility.

The commercial world generally, and the financial world in particular, seems more interested in high growth than in maximising profitability, whatever lip service is paid to the latter. This is especially the case for

banks and insurers for which the connection between the fee paid and the service provided is indirect. The service provided by insurance, for instance, is not the claims paid but the protection given.

A clearer example is perhaps afforded by the traders in financial markets who are much more strongly motivated by skewed incentive structures than by any thought of maximising long term profit for their employer. A good example is the demise of Barings, for which see, *i.a.*, Economist (1995), Kuprianov (1995), Hoon et al. (1995).

Identify risk

As already noted, identification of the risks faced by an organisation is far from trivial. Some typical means of facilitating identification are given in Head & Horn (1997) and RMS (1999).

Any given risky situation impinges on several parties. As an example, the case of the US Air Force jet ploughing into the gondola containing skiers in the Italian Alps in January 1998 had several interested parties: the US Air Force, the US government, the Italian government, the local populace, families of victims, the local authorities and those running local medical facilities, *inter alia*. This situation is one of disaster limitation *ex post* rather than risk management *ex ante*, which is really what the table at the beginning of Section 2 refers to. But the point remains: for any risky or catastrophic situation there are many concerned parties, and the risks faced by these various organisations are different, even though it is the one catastrophic event under consideration.

Risk can be greatly dependent on the environment. The distributional channel chosen by an insurer, for instance, strongly

affects the quality of the risks accepted into the portfolio. Those direct writing companies relying on obtaining business over the telephone, or through newspaper advertisements or the internet, will procure a very different class of customer from that obtained by companies relying on agents or brokers for their policy proposals.

Assess risk

In the assessment, evaluation and prioritisation of risk, some degree of quantification of the risk in question is clearly desirable. But it is important to realise that there can easily be decreasing returns to scale in modelling risks, even when it is possible to measure them.

The damage to Wellington from a major earthquake, for example, is estimated as anything between \$NZ 5 and 20 billion. A more precise figure may be useful to the EQC,⁷ who would like to know how likely the damage is to exceed their current fund, which is about \$NZ 3.5 billion (including reinsurance cover). In particular, a rough idea of the shortfall of the fund when the big one strikes may help in setting up contingent lines of assistance from the financial sector or other governments.

The EQC Fund used to be kept separate from other government funds, and the issue arose as to how much it would in fact be worth in the event of a large earthquake striking Wellington, since much of it was invested in NZ government bonds. These days the EQC Fund is little more than an entry in the Government accounts, and is effectively subsumed into general government funds. The question becomes one of seeing how far a large earthquake impinges on the value of government funds, domestic and offshore, and how quickly these values bounce back after the event. If recent events are a reliable

guide, markets bounce back remarkably quickly after a serious event; the Kobe earthquake in January 1995 provides a good illustration, since the Japanese stock market had largely recovered within a few days.

The sum mooted for the total cost of an earthquake in Wellington is however immaterial for the civil authorities, who will try to organise emergency plans in the same way regardless. Any cost estimate is sufficiently large that the importance of the risk, and in particular prioritisation of the risk, is fairly clear.

As another illustration, consider an insurer faced with a policy proposal. There may be little point in underwriting the policy too thoroughly, or modelling that line of business too closely, since the insurer's room to manoeuvre may be limited by the rigidity of the market. Should the premium be larger than that charged by competitors they will find little business; should it be smaller they may increase their business turnover rapidly, but probably only by taking on poor risks.

The transfer of risk

Transfer of risk takes place through insurance markets, either *de jure* or *de facto*, as noted in Section 1. If one has a good idea of how to price the risk, one can better evaluate the purchase of insurance or the determination of a price to tender. Participants in pools set up to exchange risks, such as the electronic catastrophe risk exchange CATEX, want a reliable estimate of their own risk as well as of the risk assumed from another in exchange.

That said, there can definitely be decreasing returns to scale in modelling risk in an attempt to quantify it. There is also such a thing as spurious quantification, viz.

assigning to a risk a quantity which, even if moderately accurate, may not be of any consequence. An obvious example is a detailed statistical analysis of an insurance claims process which forms only a small part of the insurance company's portfolio.

Quantification is not an end in itself: it should be indulged in only in so far as the marginal benefit exceeds its marginal cost.

Overall position

The final step is to monitor the situation and revise the whole process as necessary. A vital element is communication: the communication requirements within the organisation central to the risk management process itself, communicating the results of the risk management process to the media, and communication considered as a risk to be managed, on all of which McNamee (1997) has a good discussion.

3. Prerequisites for Good Modelling

Meaningful quantification of risk depends on good modelling of a risky or uncertain situation. In this section we discuss those characteristics of a risky situation which are desirable for effective modelling.

The importance of a clear definition of the risk is discussed first, followed by an outline of the need for good data. This latter is subdivided into a section stressing statistical aspects, with some emphasis on risk factors and proxy variables, and a further section emphasising insurance aspects, particularly disaggregation and units of exposure to risk. While this dichotomy is convenient, the separation of material into these two parts is slightly fuzzy at the edges.

There follow discussions of the desirability of stability in the situation modelled, the possibility of having a good theoretical model of the risk, the materiality of the risk being modelled in the overall risky or uncertain situation faced, and the appropriate use of software.

Clear delineation of the risk

It is normal insurance practice to distinguish between an object insured, for example a house or boat, and the peril, such as storm, volcano, earthquake, fire, etc.

The separate modelling of severity and frequency is desirable because the factors operating on these two aspects of a risk are likely to be different. Inflation for instance will affect claim severity or magnitude, but will not impinge directly on the number of claims expected. On the other hand, an increase in the number of cars on the roads will impact on the number of accidents, but presumably not on the average claim size.

There should be minimal ambiguity in cover, for instance from government regulations, from overlay of covers, or from loose policy wording. Freeman & Kunreuther (1997, ch. 7) point out that even when one would intrinsically expect a risk to be insurable - viz. the claims level should be reasonably predictable, with not overmuch adverse selection or moral hazard⁸ - there may be institutional reasons for a formal insurance market not to exist.

They cite the examples of underground storage tanks and lead-based paints as environmental risks in the US which have become uninsurable. In the former case there was political pressure to set up alternative state guarantee funds, vitiating the development of the insurance market. In the case of lead-based paints, ambiguous

government regulation renders the risk uninsurable. They stress the need for well-specified standards and regulations as prerequisites for the existence of an insurance market.

Good data

The central requirement is to have available a sufficient volume of reliable past data. One prerequisite for this is that the underlying risky situation be moderately stable (see below).

Data on exposure should ideally be monitored on a policy year basis, commencing from policy inception. On the other hand, claims should be monitored on a development year basis, i.e. commencing from the date on which the claim was incurred. Both of these are in fact often measured on a calendar year basis. In any case, the date on which a claim is incurred is often only vaguely known. An example is given by asbestosis claims, which arose in many cases some years after the insurers had thought the claims incurred in a given policy year had completely run off (see *i.a.* Sigma (1995)). There are delays endemic in the insurance markets, between the insurer receiving notification of a claim and entry onto the computer system, for instance, and especially between primary insurer and reinsurer.

Ideally there should be a wide range of risk factors, viz. factors which impinge on the risk, and which can be used for modelling purposes. Such factors will typically differ between frequency and severity; for health insurance, they are generally assumed to differ between inception rate, duration of illness and severity.

The goal is for the risk factors to explain a high proportion of variation in the endogenous variable modelled (a high R^2 in

standard regression models), without needing too many risk factors for this good fit, since parsimonious models are easier to interpret and lead to sharper coefficient estimates (lower standard errors).

Even when the risk factors are not themselves measurable, there may be at hand proxy variables that one can use instead for modelling. Requirements here are that the risk factor and proxy variable are closely correlated, and that the proxy variable is not subject to large measurement error.

There should ideally be a close connection between rating and risk factors. Rating factors are those used by insurers for rating purposes, i.e. calculating premiums, but may be chosen for marketing reasons rather than for having anything to do with the risk.

Disaggregation

One can operate on a global level; but if one's understanding of the situation extends no more deeply, the model is questionable as soon as underlying factors change, since risk factors are being subsumed in the modelling methodology. On the other hand, to disaggregate too far makes the modelling exercise pointless, for stitching the lot together will lead to errors. It is a fine art to know how many layers of modelling to use in a given situation, and the answer depends of course on the quality and quantity of the data available, which in turn may well depend on how much it costs to get the data.

Exposure to risk

The exposure (to risk) unit is the finest meaningful subdivision of data, and will usually be grouped into larger cells for modelling. Using a well-defined exposure unit enhances understanding of the risk process. More fundamentally it leads to a

consistent basis on which to price, measure claims rates and evaluate the relative sizes of different policies and portfolios over time.

The simplest example is that of automobile insurance. The exposure unit is usually the vehicle year, and the basic cell for statistical purposes may for instance gather together drivers or owners of a particular age, driving a medium-sized car and living in a given suburb. These factors comprise the risk factors being used in the statistical analysis; should the model be used for rating purposes, they transmute into rating factors. The other prime example of an insurance line with a good measure of exposure is life insurance, for which it is natural to use the life year. Property classes are less likely to possess a reasonable measure of exposure, especially the commercial lines, with very heterogeneous portfolios. It is common to take exposure in these classes as the sum insured for household risks, or the PML or EML (probable maximum loss, expected maximum loss) for commercial property.

For classes with no upper limit to claims and no obvious exposure unit, such as liability and catastrophe lines, the premium is often used as the measure of exposure. In theory, however, the premium should be derived from the exposure. The premium rate is properly expressed as a number of dollars per unit of exposure, and the size of the portfolio is properly the total exposure of that portfolio.

While equating exposure to premium is akin to calculating output as the cost of inputs, it does give some measure of consistency to quantification of risk over time and between portfolios, often in the absence of any real alternative. Roberts, (1995b) has further discussion on this point.

Statistical cells

For modelling purposes, the ideal is to have independent units or cells of risk with the same distribution, with large numbers of such cells in each policy in an insurance portfolio. This makes for consistency of rating over time and from one policy to another within the portfolio. Ideally each of these cells will be large enough to apply the law of large numbers (LLN) and the central limit theorem (CLT)⁹, and sufficiently small for the exposure units therein to have uniform characteristics (risk factors).

Stability

It is clearly unreasonable to expect perfectly steady state behaviour in the real world. By stability we understand that there can be change, but it must be relatively smooth and predictable, or at least able to be sensibly modelled in some way. A simple example is that of inflation, which can be removed from overt consideration by modelling claim severity in real terms.

A counter-example could be where compulsory wearing of seat belts is introduced, vitiating at a stroke the collection of data on claims rates over the period before the change. A second example is afforded by sudden discontinuities in stock exchange price behaviour over time, say when the Russian government defaulted on its debt in September 1998, causing the problems with LTCM (see section 1).

Ex post, fitting a model can easily incorporate discontinuities by judicious use of dummy variables. But it is surely an exaggeration to claim on this basis that one is modelling risk, since there is no predictive power for such sudden lurches in the market, at least from mathematics. There is simply no way in which mathematicians, no matter

how inspired, can predict severely discontinuous behaviour; generally speaking mathematical models are predicated on smoothly continuous change.

Theoretical models

When past data concerning the risk to be modelled is sparse or unreliable, one has to rely on scientific models of the risk process, and possibly incorporate the theoretical model into a statistical or simulation model.

The archetypal example is that of modelling earthquakes. While there is a surprising amount of data from history about big events, and much recent information about generally smaller events, there is far from enough data accurately to model the occurrence of large "one in a thousand years" types of earthquakes by purely statistical means. Similar problems arise in modelling other natural hazards, but they are less pronounced than for earthquakes.

It does not help that several of the theoretical models concerning earthquakes seem to be contradictory. "Stress release models", for example, posit that when there is an earthquake, stress within the earth's mantle is released, thereby rendering another large event less likely, at least for a long time (Lu, Harte and Bebbington (1999)). Other models, rather more intuitively, model the probability that a large event will spawn another one in the near future; this class of models include the so-called ETAS¹⁰ models, see Ogata (1998).

In fact the differences between these classes of models are more apparent than real. The stress release models are modelling big events over long time intervals, while the ETAS models concern themselves with aftershocks following events of all sizes in the short term. It would be preferable to

combine the attributes of both classes of models for a more complete model.

One class of the more statistically orientated types of earthquake models rely on predicting large events from patterns of fore-shocks, but that raises the problem of differentiating after-shocks arising from the last big event from fore-shocks presaging the next large event. Some of the shocks associated with big events are extremely large in their own right, and the task of picking out the major events from the catalogues of earthquakes is a challenging one.

While the whole area of modelling earthquakes may be fraught with difficulty, the underlying problem is simple enough to state. Modern instruments are capable of increasingly precise measurement of magnitude, acceleration and other technical aspects of earthquakes, but recent good data allow us to model only relatively small events. To extrapolate models calibrated from small events to the large ones that we want to model is to some extent a guessing game. See Vere-Jones (1995) and Roberts (1995a); for a discussion of approaches of the civil authorities to earthquakes, see Vere-Jones, Harte and Kozuch (1998).

A flicker of light illuminating the gloom is that relative risk may be more easily treated than absolute risk. Again to choose earthquake cover as an example, it is clear that Wellington premiums should be far larger than Auckland premiums. Indeed, data for total earthquake damage over the last 150 years in both places could give one a rough idea of comparative premium levels to use. For the volcano peril, it is equally clear that Auckland premiums should exceed Wellington premiums.

Materiality

Materiality (or importance) of the risk modelled could be interpreted in at least two ways. Firstly, claims modelled in a particular line of insurance may be relatively unimportant in that the profit from that line depends more on the investment performance obtained on the premiums, since they are received before claims are paid out, and on the level of expenses. More generally, the insurance line modelled may form only a small part of the overall company portfolio of risks.

The second factor is rigidity in the insurance market, in the sense that if one charges premiums markedly higher than the market norm one will sell little business, whereas if one lowers the premium one may well be expanding the portfolio by attracting poorer risks. Thus the leeway to charge differential premiums is limited.

Those two points notwithstanding, it may still be worthwhile modelling claims risk at the micro level, partly because one obtains a better understanding of whether the line is profitable or not in the long term; and partly because modelling the risk alerts one to the possibility of differential rating in the future, perhaps obtaining a niche marketing position. See Section 4 for a wider discussion of the general advantages of mathematical modelling of a portfolio of risks.

Use of appropriate software

It is frequently the case that complex mathematical models are used inappropriately just because the software for fitting them is readily available. The inappropriateness may stem from the

inherent unsuitability of the model for the situation being modelled, the data not being of sufficient quality or quantity for proper use of the model, or the fact that the model is applied without sufficient knowledge of what the underlying mathematics is meant to do, and what the output means.

In this connection one thinks perhaps most readily of multivariate statistical models. Principal component analysis and discriminant analysis are widely used models of this type; and the latter in particular has formed the basis of predicting insolvency (the quaintly named "distress") of companies in the finance sector on the basis of ratios formed from their accounts (profitability, leverage, ratio of insurance liabilities to premiums outstanding for an insurer, etc).

The fruitful use of multivariate statistics models however presupposes a reasonable grasp of statistical theory on the one hand, and a good knowledge of linear algebra on the other. Inputting the data and pressing the button to produce the output may or may not take long, but interpretation is not so easy for those inexperienced in using such models.

4. Modelling Risks: Rating and Updating

In this section we consider modelling risk and updating the model over time. The standpoint we adopt is the insurer's, in that we consider premium rating and updating, and use insurance terminology. Nevertheless the discussion applies to general modelling of risk and uncertainty.

A discussion of market driven rating and retrospective rating is followed by a brief description of what we mean by mathematical models. After listing general advantages of mathematical modelling of risk, we consider the types of modelling which may be appropriate for different

insurance classes. Finally we move on to updating models, considering successively credibility theory and Bayesian methodology.

Premium rating

Insurance premiums are expressed as functions of rating factors, which need not coincide with risk factors. Race may for example be a factor in HIV deaths, but cannot be used as a rating factor because of cultural sensitivity. As a further example, risk factors in car insurance may be age and sex of the driver, but the closest available proxy variables may be the age and sex of the owner of the car. In any case rating factors may be chosen for purely marketing reasons.

In principle, rating can be effected by either following market rates or using a model of the claims process, in which explanatory variables are risk factors or their proxies. In practice of course neither of these methods would be applied in isolation. One cannot afford to rate entirely by a statistical model without some regard for what the competitors are up to, and in like vein even those who follow the market usually do so with a fair feeling for the situation and how it might develop, which is tantamount to modelling in a simple sense.

Market driven rating

Supposing that there is a market for transferring the risk in question, one adheres closely to the market premium rates. The usual argument for this is that deviation will either mean loss of business if the premium is higher than the market, or an underwriting loss if the premium is lower, as explained in Section 3.

Retrospective rating

Should the previous year's underwriting result have been worse than expected, or at least worse than the results of other players in the market, the insurer may attempt to rectify matters through retrospective rating.

This usually assumes one of two forms:

- profit sharing, typically applied in group life insurance, or
- experience rating, the clearest example of which is the no claims discount commonly used in car insurance.

The former allocates the past year's profit between the insurer and the policyholder (generally an employer), while the latter reduces next year's premiums for a good claims record.

Taken too far, these procedures vitiate the essence of insurance, which is that claims are supposed to be variable for the individual policyholder. Moreover too strong a reliance on retrospective rating implies that one is charging too much, which may lose business in the medium term. Both forms of retrospective rating are however consistent with a market driven rating strategy, since both are widely used in the insurance markets mentioned.

Modelling the claims process

In this paper, a mathematical model means primarily one of the following three types. The classification is somewhat rough, in that it is perfectly feasible to have random inputs into a simulation of a risk process, and simulation is often used with statistical models. However, this classification will suffice for our purposes.

(1) Statistical models

These typically assume the form

$$Observation = signal + noise$$

of which the archetype is linear regression type models. The model is driven by the random noise, which is usually taken to have simple properties of independence, etc. The purpose of fitting the model is to remove the random noise from the observation to reveal the signal.

(2) Theoretical models

A scientific model will tend to incorporate any available past data in a statistical fashion, so that the difference between "statistical" and "theoretical" models should not be overstressed. By the latter we have in mind particularly natural hazards models, especially those for earthquakes (see Section 3). This paper is mainly concerned with statistical modelling, which largely boils down to doing something meaningful with whatever past data is available.

(3) Deterministic simulation

The dominant model in quantifying risk is deterministic simulation to model future cash flows, with "what if" games or "scenario tests" being played on spread sheets. The intent is frequently to set up putative balance sheets and company accounts over the future, with a view to investigating profitability, capital adequacy, etc. This type of model is used not so much to model the "micro" risk process under consideration, but rather to translate that result into its effect on the whole organisation faced with risk. Taking the additional step, from the organisation at risk to the environment in which that organisation operates, is generally less amenable to mathematical analysis.

Discounted cash flow models of this type form the cornerstone of work in the traditional actuarial fields of life insurance and pensions. For modelling profit from an insurance policy, for instance, a life actuary will input withdrawal rates, mortality rates, interest rates and expenses, etc. into projections of future experience. Basic problems are to decide the extent to which conservatism is to be built into the parameters of the model, and the choice of interest rate to be used for discounting.

A further problem lies in the interpretation of the output, and arises from having a large number of parameters in the model. All one can really do is attempt to isolate those inputs which seem especially significant for the value of the output. In modelling insurance policies, for example, the lapse rates assumed are crucial, since it takes time for the policy to recoup initial expenses; for modelling pensions liabilities neither the rate of salary increase nor the interest rate earned on funds is so vital, but the difference between them has a large effect on the answer.

Another major type of mathematical model is systems of differential or difference equations. In general these are more common in the physical than in the social sciences, although one subclass of general application is *Dynamical Systems*, which includes chaotic and fractal models. While this type of model is rapidly becoming of wider significance in dealing with risk, space precludes its treatment in this paper.

Advantages of modelling claims

It may be cheaper in the short term to rely on the market to set premium rates, especially when the insurance class concerned is only a minor part of the portfolio. Nevertheless, refusing to contemplate deviations from what

the market does would not seem a fruitful way to proceed in the long term.

Advantages of modelling the claims arising under a risky portfolio include:

- an enhanced understanding of the portfolio;
- an indication of whether too much or too little is being charged, giving some feel for the real profit or loss from that line;
- the possibility of differential rating in the future, perhaps leading to a niche marketing position;
- scientific incorporation of new information into the model, typically the latest year's results, or developments in related markets overseas. Incorporating new information can be done by simply rerunning the model with the latest data available; but we also have in mind the formal incorporation of data in the updating of a prior distribution of a parameter to the posterior distribution. This is known as Bayesian analysis and is discussed later in this section.
- strengthening the case for approving premium rates. In many states of the USA, rates need to be approved by the Insurance Commissioner. Claims models commonly used for this purpose include credibility theory (see below) and the capital asset pricing model.¹¹

Statistical modelling in insurance

Whether or not it is worthwhile for an insurer to model the claims process depends strongly on the class of business. Paradoxically enough, those classes which are most amenable to statistical modelling are often those where it may not be worthwhile to bother modelling at the micro level.

Short-tailed insurance lines

Having a short tail in this context means that claims run off quickly, so that within 1 or 2 years of the exposure year all claims can be considered as having been paid in full. The insurance lines most amenable to statistical modelling are short-tailed lines with good measures of exposure, particularly household and automobile, although excluding liability claims arising from car accidents. Since claims are relatively steady, maximising investment income and minimising expenses attributable to those classes tend to be more important in determining insurance profit, and detailed modelling of the claims process may be an unnecessary luxury. That said, a simple statistical model is often used to estimate outstanding claims for reserving purposes, and especially for determining the IBNR claims.¹²

Liability Insurance

Liability lines are long-tailed, and difficult to model. Inflation of court settlements can differ substantially from ordinary inflation, and payments are often made desultorily over many years until a claim is finally settled. There may be some rudimentary modelling of future cash flows, and reserving may take the form of present values so produced. However, past data is both heterogeneous and little guide to the future, so that overly statistical models are out of place.

Catastrophe Insurance

The statistical modelling for catastrophe classes (earthquakes, storms, etc.) may be rudimentary, largely because of scarcity of data. Nevertheless some statistical input into the model is probably needed for catastrophe

class rating, although such models are less useful in assessing overall risk than in differentiating between prices charged to particular groups (see the discussion in Section 3).

The basic problem with insurance of natural hazards, and other cases which combine a rare event with enormous potential damages, is that insurance is suitable for averaging over a portfolio at a given time, not for averaging over time. One may calculate a premium on the basis that the event in question will occur once in a hundred years, but the insurer is in no position to accumulate reserves over many years, even if the odds given are accurate. Accounting conventions decree that one year's business is fully dealt with at the end of the year, and there are no government incentives to accumulate reserves of any magnitude. Even if an insurer were to attempt to build up the required reserves, it would very smartly render itself liable to a takeover attempt. Jaffee and Russell (1997), Roberts and Rollo (1993) and Roberts (1995b) discuss this point further.

Health Insurance

Health insurance is highly problematic for modelling, even in comparison with liability and catastrophe classes. Accident insurance has at least the promising feature that one can usually decide where the accident took place, so that one can often know whether it is the responsibility of the employer or the insurer. Creeping illness, or disease or injury of gradual process, is far harder to attribute to a cause, and a real difficulty for determining liability.

A basic problem for this whole area of risk is extreme adverse selection, in that people who know they are already sick will propose for

sickness cover. Once cover has commenced, health insurance combines the worst features of general insurance (subjectivity, as to whether or not a claim has been incurred, and also in determining the claim amount) and life insurance (long-run nature of some of the policy covers).

The above points notwithstanding, modelling sickness and accident lines is probably essential for the insurer writing these classes. The business generally is very susceptible to institutional factors, such as varying underwriting standards applied to proposals, and the energy expended on managing troublesome long-run claims to a speedy conclusion. The validity of industry-wide statistics to an individual insurer is dubious.

Overmuch reliance on statistical models may not be warranted. Models tend to be simulation models of future cash flows, inputting inception rates, withdrawal or lapse rates, duration factors, severity (lump sum or income), interest rates, etc. All of these need to be guesstimated from past experience or industry sources.

Updating the model

Just as important as the initial modelling framework is the provision made for updating the model over time. Should actual claims experience differ much from that expected, the insurer must ask whether the discrepancy is *random*, in which case one does not want to adjust the "correct" premium rates; or *systematic*, in which case one does want to change the rates.

Credibility Theory

One way to have a foot in both camps is to use credibility theory, which takes the updated premium P to be a weighted average of collateral information and one's own data:

$$P = z P_0 + (1 - z) P_1 \quad (1)$$

Here P_0 denotes the premium estimated from one's own past experience, P_1 the premium estimated from collateral sources, such as other players in the local market or from related markets overseas, and z the credibility factor, i.e. the credibility of one's own data.

The equation (1) begs the question of how to estimate the credibility factor z . The formula used from the early decades of this century by American actuaries, principally to justify rate increases to insurance commissioners, is

$$z = \frac{n}{n + K} \quad (2)$$

in which n is the portfolio size (number of policies) and K is a constant.

The actuaries calculated a minimum size of portfolio for full credibility ($z = 1$) by arguments similar to those used to find the minimum sample size to achieve a given level of precision in statistical estimation of parameters (e.g. Hossack, Pollard and Zehnwirth (1983)). They chose K so that for n exceeding this minimum portfolio size the credibility was effectively unity.

Deeper theoretical justification for what the American actuaries had been doing for so long was not forthcoming until the early 1960s, when Buhlmann and Straub produced formulae similar to (2) using Bayesian methodology (Buhlmann (1967)).

Bayesian updating

Updating risk estimates from a model consists in general of simply running the model again with the additional year's results. But when a parameter is assumed to have a statistical distribution, Bayes'

Theorem can be used to update that distribution over time as more information is incorporated.

The parameter chosen to be random is usually of central importance, such as mean claim size or claims rate (number of claims per unit of exposure).

Bayes' Theorem states that

$$\text{posterior density} \propto \text{likelihood function} \\ \times \text{prior density}$$

As a simple example, consider a portfolio of n policies, each of which has a probability q of having at least one claim in a year. The probability function is the binomial

$$\binom{n}{x} q^x (1-q)^{n-x} \quad (3)$$

The likelihood function is the same function, but *ex post*, i.e. after the realised value x is known. Thus Bayes' Theorem for (3) boils down to

$$\text{posterior density} \propto q^x (1-q)^{n-x} \\ \times \text{prior density} \quad (4)$$

The mathematics works out very neatly when there is a conjugate prior to the chosen likelihood. The conjugate prior for the binomial likelihood, for example, is the beta distribution, with density function a constant times $q^a (1-q)^b$. Updating the beta prior distribution also leads to a beta posterior distribution, which in turn is updated into a further beta distribution, as will be evident from (4), and so on.

More general approaches to updating via Bayes' Theorem deal with the mean and variance of the modelled distribution rather than the full likelihood function, and utilise

least squares to obtain a formula like (2). See *i.a.* Waters (1987), Klugman (1992) and Kahn (1975).

The Bayesian approach is a scientific means of updating a mathematical model over time, with less chance of over-reacting to one year's heavy underwriting loss. Other advantages include the flexibility to incorporate collateral information of whatever kind, and also the fact that over time the insurer's own data becomes fully credible.

There are also strong disadvantages. The principal objection is the extreme vulnerability to the subjective judgement of the modeller, in that a vague or flat prior will produce a very different posterior distribution from that produced by a sharp or spiky prior. Otherwise stated, the modeller with firm *a priori* views will influence the posterior distribution strongly. Further problems include finding a good model for the situation modelled from which to derive the likelihood function. Bayesian analysis is really only worthwhile if applied for several successive years, which presupposes that the underlying situation being modelled is not changing rapidly: see the discussion on stability in Section 3.

5. Conclusion

While it may be only relatively rarely that one can measure risk in any precise sense, it is important to try to quantify risk, at least to the extent that it is cost effective. It is however vital to realise the limitations of what one can do in this direction. Risk factors may be unknown or unmeasurable, and sophisticated software and copious data do not guarantee a good model, especially in the hands of an inexperienced operator. Above all there is far too much reliance on

the few mathematically tractable models available, especially when they are simple enough to be taught at less than advanced levels.

aspect of reality very well, or when the underlying mathematics works out neatly.

However important it may be to quantify risk, it is vital to retain a healthy scepticism towards any model. This is far from easy when one has invested much time in building the model, especially when it mirrors one

Footnotes:

- ¹ A debt is owed David Harte for helpful comments on this paper. Responsibility for statements and opinions expressed herein remains of course with the author.
- ² Leigh Roberts is Director of the Financial Mathematics programme at Victoria University of Wellington.
- ³ The distinction between frequency and severity is important, since the influences on each tend to be separate, on which we comment in Section 3. Frequency is often labelled as likelihood and severity as consequences (RMS (1999)).
- ⁴ In the NZ public service, output tends to be interpreted in a narrow sense, in that there is the expectation that it can be measured, while outcome is less amenable to measurement. The Crown, for instance, in the person of the minister purchases output from the public service, be it papers giving policy advice, or construction of a building, etc. While the output can be measured, by default by the cost of the inputs provided, the outcomes for the public are not so easily quantified.
- ⁵ Marking to market means valuing assets and liabilities at market value when a market exists, and otherwise valuing contracts as if they were being terminated, or the position closed out, at that time.
- ⁶ The above comments apply to traditional endowment assurance policies. More modern unit linked or unbundled policies have premiums paid into a unitised fund held in the policyholder's name, and the above points do not apply to the same extent.
- ⁷ Earthquake Commission, formerly the Earthquake and War Damage Commission.

-
- ⁸ *Adverse selection* is the tendency for insurance to be purchased or continued by poorer than average risks. *Moral hazard* is used somewhat loosely in different contexts, but is properly defined in insurance as policyholders changing their behaviour after they become insured, *because* they are insured.
- ⁹ The LLN asserts that the average of a large sample will be close to the population mean or average. The CLT says that the sample average, suitably scaled, is approximately normally distributed.
- ¹⁰ Epidemic Type Aftershock Sequence.
- ¹¹ Each line of insurance has its own β value, giving rise to the market view of the right rate of return from that line, and one prices accordingly.
- ¹² Incurred But Not Reported claims. When accounts are drawn up at calendar year end, typical examples are claims incurred in December and not notified to the insurer until January.
-

References

Buhlmann, H. (1967). Experience rating and credibility, *Astin Bulletin*, 4, 199-207.

Economist (1995). The Bank that disappeared, *The Economist*, 4-10 Mar.

Economist (1999a). Black hole, *The Economist*, p. 63, 28 Aug-3 Sept.

Economist (1999b). Caveat emptor, *The Economist*, pp. 85-86, 29 May-4 June.

Economist (1999c). Think of a number, *The Economist*, p. 81, 11-17 Sept.

Ferris, S. (1999). Who pays? The aftermath of the UK pensions mis-selling scandal, *The Actuary (Australia)*, 1 (50).

Freeman, P.K. and Kunreuther, H. (1997). *Managing environmental risk through insurance*, Kluwer Academic Press.

Greenspan, A. (1998). Statements to the Congress: Alan Greenspan, Oct 1, 1998, *Federal Reserve Bulletin*, 84 (12): 1046-1050.

Head, G.L. and Horn, S. (1997). *Essentials of Risk Management*, 3rd ed. Insurance Institute of America, 2 vols.

-
- Hoon, L.S., Yu, D., Sargent, S. and Shimomura, K. (1995). The nine days that sank Barings, *Asiamoney*, 6 (3): 15-21.
- Hossack, I.B., Pollard, J.H. and Zehnwirth, B. (1983). *Introductory statistics with application in general insurance*, CUP.
- Jaffee, D. M. and Russell, T. (1997). Catastrophe insurance, capital markets, and uninsurable risks, *Journal of Risk and Insurance*, 64 (2): 205-230.
- Kahn, P.M. (ed), (1975). *Credibility Theory and Applications*, Academic Press.
- Klugman, S. (1992). *Bayesian Statistics in Actuarial Science, with Emphasis on Credibility*, Huebner International Series on Risk, Insurance and Economic Security, Kluwer Academic Publishers.
- Kuprianov, A. (1995). Derivatives debacles, *Economic Quarterly* 81 (3): 1-30. Federal Reserve Bank of Richmond.
- Lang, J. (1995). An Australian abroad, *Actuary Australia* 1 (28): 16.
- Levene, T. (1998). Prudential in new pension sales row, *Guardian Weekly*, 16 Aug, p.9.
- Lu, C.S., Harte, D.S. and Bebbington, M. (1999). A linked stress release model for historical earthquakes from Japan, *Earth, Planets and Space* 51. (forthcoming).
- Martinson, J. (1999). Soros loses \$700m. in failed gamble on internet shares, *Guardian Weekly*, 19 Aug, p.14.
- McNamee (1997). Risk management today and tomorrow. Paper presented to State Services Commission, NZ, April. www.mc2consulting.com/risknz.htm.
- Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics* 50 (2): 379-402.
- RMS (1999). *Risk Management: Australian/New Zealand Standard*, 4360/1999, Standards Association of Australia.
- Roberts, L.A. (1995a). On the theory and changing practice of insurance, *Australian Institute of Actuaries' Quarterly Journal*, pp. 26-35.
- Roberts, L.A. (1995b). Risk management of catastrophes in New Zealand, *Australian Institute of Actuaries' Quarterly Journal*, pp. 23-46.
- Roberts, L.A. (1999). Recent developments in risk management and insurance of natural hazards, *Australian Actuarial Journal* 5 (1): 173-184.
-

-
- Roberts, L.A. and Rollo, C.A. (1993). Risk management of natural hazards, *Australian Institute of Actuaries' Quarterly Journal* pp. 2-13.
- Sigma (1995). The London market, *Occasional Publication of Swiss Reinsurance Company*, No.2/1995.
- Sigma (1998). Life and health insurance markets benefit from reforms in state pension and health systems, *Occasional publication of Swiss Reinsurance Company*, No.6/1998.
- Vere-Jones, D. (1995). Forecasting earthquakes and earthquake risk, *International Journal of Forecasting* 11: 503-538.
- Vere-Jones, D., Harte, D.S. and Kozuch, M. (1998), Operational requirements for an earthquake forecasting programme in New Zealand, *Bulletin of the New Zealand National Society for Earthquake Engineering* 31 (3): 194-205.
- Waters, H.R. (1987), *An introduction to credibility theory*, Institute of Actuaries Education Service, London.
-

BOOK NEWS

Jerry Mushin, Senior Lecturer at Victoria University of Wellington, has recently had the third edition of his textbook, *Income, Interest and Prices*, published by Dunmore Press. Like the first two editions, this is aimed at first-year undergraduates, but is also of interest to other readers seeking insight into current macroeconomic issues but without previous background in economics or advanced knowledge of mathematics. Aside from some new graphs, tables and flowcharts, the main improvements over the second edition are the addition of a new overview chapter called "The Complete Model" and a list of algebraic notation at the end of the book. An edition especially revised for the British market is due to be published later this year.