SOME UNIFYING REMARKS ON RISK LOAD

Philip E. Heckman

Some Unifying Remarks on Risk Load

Philip E. Heckman, Ph.D., ACAS, MAAA
Ernst & Young
Great Lakes Actuarial
Services Group

ABSTRACT

The purpose of this note is to point out the connections between the Marginal Surplus and Competitive Market Equilibrium approaches to calculating risk loads and to show that these methods incorporate and unify several other conceptual approaches to risk loading.

Some Unifying Remarks on Risk Load

The casualty actuarial literature has of late provided a forum for very active debate and discussion on the subject of risk load. In this note, we shall study two promising approaches to the problem and show that they are intimately connected and mutually illuminating. The two are the Marginal Surplus approach, as expounded by Rodney Kreps (1), and the Competitive Market Equilibrium approach of Glenn Meyers (2).

Marginal Surplus

This concept has been mentioned in several sources but treated concretely by Kreps in the context of reinsurance. The central idea is the following: any piece of business should be priced in such a way that , after deducting expected losses and expenses, there remains a contribution to the company's surplus which leaves the company in the same risk position as before the business was written. Aside from being the source of profit over the long run, this risk load satisfies the company's fiduciary obligation to maintain product quality - primarily, reasonable assurance of the ability to pay claims - for current and former policyholders and their potential beneficiaries.

Kreps quantifies this notion in a way that is fraught with implications. Suppose that we can treat the insurer's net worth as a random variable with finite second moment. Hence we can define a standard deviation, Σ, and express the insurer's actual surplus, S, as a multiple of this quantity:

$$S = Z\,\Sigma. \tag{1}$$

So far this is only a definition, but we may argue that countervailing pressures of the insurance and the capital markets will tend to confine Z within a fairly narrow equilibrium range.

To see why this is so, consider that Σ typifies the scale of variability on the distribution of aggregate net worth. The multiple, Z, then, should map directly to a value for the insurer's probability of ruin, the probability that present assets are insufficient to satisfy present liabilities so that present liabilities will have to be subsidized from future earnings. Such subsidies are routine in the public sector, but in the private sector they pose serious problems of equity and are often frowned upon. In particular the future earnings may prove insufficient to provide the subsidy to cover past mistakes, and actual insolvency may result. Hence there will be pressure from the insurance market (and from regulatory authorities) to keep the effective Z value reasonably high since the stability and reliability thus achieved is the central determinant of insurance product quality.

On the other hand, increasing Z to very high values will have little effect on the insurer's actual risk position; but it will be penalized by the capital markets since it ties up funds better utilized elsewhere. (Remember that the policyholder also participates in the capitalization of the insurer but receives actual equity only in mutual companies.) This is clear because each successive increment of Z has a smaller effect on the probability of ruin. Suppose that Z maps to a probability of .001 (For the normal distribution, Z $\cong$ 3.) This means that, of a thousand companies in a similar position, only one, on average, would prove deficient in runoff. Clearly there will be little reward from the market for cutting the probability further to .0001. There are not even 10,000 insurers in the entire market.

Let us suppose that we have an equilibrium Z value, a market consensus. It then makes sense to price new business in a way that keeps Z constant. That means we must examine the insurer's risk position before and after the transaction. Before the transaction, the standard deviation of the insurer's net worth is

$$\Sigma_0 = \Sigma. \tag{2}$$

After the transaction, it is

$$\Sigma_1 = \sqrt{\Sigma^2 + \sigma^2 + 2\rho\Sigma\sigma}, \tag{3}$$

where σ is the standard deviation of the net present value (NPV) of the accepted risk, and ρ is the coefficient of correlation between the risk's NPV and that of the existing book. With this notation, the required surplus contribution is

$$\begin{aligned} \Delta S &= Z\,(\Sigma_1 - \Sigma_0) = \\ &= Z\,\frac{2\rho\Sigma\sigma + \sigma^2}{\Sigma_1 + \Sigma_0}. \end{aligned} \tag{4}$$

In the limiting (and usual) case, where the added risk is only a small fraction of that of the existing book, (4) can be approximated

$$\Delta S \cong Z\rho\sigma, \tag{5}$$

where we have neglected terms of second and higher order in σ. Note that this is independent of the total standard deviation and depends only on the risk's standard deviation, its correlation with the rest of the book, and the product quality factor, Z. In the next section, we shall see this relation take on more concrete form in terms of explicit stochastic models.

Competitive Market Equilibrium Model

This approach was invented to deal with some long-standing problems of increased limit ratemaking where risk load has long been an explicit issue; but it addresses the general problem; and we shall review it in that context.

In addressing the problem of risk load, it is important to recognize that our deliberations are of little use unless we have a plausible way of estimating the relevant variances and covariances. This Meyers treats at considerable length in the context of the Collective Risk Model (3). This model treats aggregate loss payments as a random sum of variates from a known loss size distribution with a multiplicity drawn from a known claim count distribution. The conceptual framework is flexible enough to accommodate parameter uncertainty since the distribution parameters can themselves be drawn from specified prior distributions. This is very important because parameter uncertainty is the prime determinant of the risk load in a consistent and market-viable scheme. Along with the catastrophe hazard, it is a chief source of the correlation discussed in the previous section. Many of the causes of parameter uncertainty act market-wide and are a reflection of the climates, legal, political, seismic, and meteorological, which determine the fortunes of the industry as a whole.

Meyers' discourse on quantifying these risks is a solid demonstration that it is feasible to put actual numbers in the place of all the Greek letters and to reduce the problem to calculation. This capability is an extremely important one and can be expected eventually to have a profound effect on the industry and on the way it manages itself. We will not dwell here on the details but cite instead the form of the final result. This is expressed in matrix notation wherein the aggregate variance of an insurer's net worth can be written

$$\Sigma^2 = A + \mathbf{n}^T \cdot U + \mathbf{n}^T \cdot V \cdot \mathbf{n}, \qquad (6)$$

where
n is a vector of exposures in force by class,
U is a vector of positive elements giving process variance per unit of exposure,
V is a symmetric, positive definite matrix describing parameter and catastrophe risk,
A is a scalar - our addition - which quantifies all other sources of variation independent of exposures currently in force,
T as superscript denotes the transpose, which interchanges rows and columns.

Reference (2) allows for the case where V is singular, but we will not consider that here.

The next step in the derivation is to pose and solve the optimization problem to determine what combination of exposure in various lines and classes and at various limits a carrier in a given risk position will choose to write. There are two distinct, but equivalent, approaches to this problem. One is to maximize profit at a fixed level of risk, quantified as the variance given in (6). The other is to minimize risk at a fixed profit quota. We state the latter explicitly by way of illustration:

Minimize

$$A + \mathbf{n}^T \cdot U + \mathbf{n}^T \cdot V \cdot \mathbf{n} + \mu\ (P - \mathbf{n}^T \cdot \mathbf{R}) \qquad (7)$$

on **n** and μ, where

μ is the Lagrange Multiplier,

P is the target profit,

R is the vector of risk loads per unit of exposure, by class, supposed given.

Either approach leads to a spectrum of solutions on an "efficient frontier" in the space of profit vs. risk. The Lagrange Multiplier can be thought of as quantifying the relative importance of profit maximization and risk minimization, a tradeoff which must be decided by management.

At this stage, supposing that **R** is known, the solution can be expressed as

$$\mathbf{n} = \tfrac{1}{2} V^{-1} \cdot (\mu \mathbf{R} - U). \qquad (8)$$

Note that this is not guaranteed positive and will only be so for appropriate values of **R** and μ. The model, in fact, describes underwriting shutdown in the riskier classes as μ becomes smaller. The correct procedure when this happens is to exclude those classes ($n = 0$) and to solve the reduced problem for the remaining components of **n**.

To this point, **R** has been assumed known. The final step is to find what risk loads are needed to allow the market to clear. This is determined by equating industry supply with total market demand and solving for **R**. This is equivalent to using in place of **n**, the industry average exposure spectrum, $\bar{\mathbf{n}}$, and introducing an industry average profit requirement, $\bar{\mathbf{P}}$. The answer that emerges is

$$\mathbf{R} = \bar{P} \frac{U + 2V \cdot \bar{\mathbf{n}}}{\bar{\mathbf{n}}^T \cdot U + 2\bar{\mathbf{n}}^T \cdot V \cdot \bar{\mathbf{n}}} \qquad (9)$$

Note that $\bar{\mathbf{P}}$ remains to be determined and may, in fact, depend on $\bar{\mathbf{n}}$.

The model assumes tacitly that insurance pricing is supply-driven - that is to say, that capital committed to insurance enterprises is a scarce good. The existence of a slack market in the trough of every underwriting cycle gives testimony that this is not always the case. Because it is much easier to enter the insurance market than it is to withdraw, the presence of excess underwriting capacity has a profound effect on insurance pricing, very much like any other commodity. In fact, pricing behavior in the property/casualty industry bears a striking similarity to agricultural commodity pricing. In agriculture the excess capacity problem is "solved" by government subsidies and price supports. In the insurance industry, attempts at administered pricing have seldom had enduring success. Rather the *de facto* "solution" for excess capacity is, in effect, to pay insureds for accepting coverage until the excess capital has been dissipated and the market has moved back towards equilibrium.

Such behavior sounds bizarre; but it is, in fact, rational in the context of available information. The farmer deciding what and how much to plant and the developer deciding to initiate a new office development are in similar positions. The only remedy for underwriting cycles and the inefficiencies they cause is better and more timely information. The hope in introducing models such as the present one, with its statistical underpinnings, is to provide such information so that the industry's risk position and capital needs are definite quantities rather than vague notions. Insurance data are notoriously noisy - prone to large fluctuations and distortions, especially in the short term. Few people can be found in the industry who pay serious attention to monthly data, and few more who attend to quarterly results.

The reason for this is simple: to be interpretable, noisy data must be filtered. The design of the filter is all-important. The only noise filter in general use in insurance - and most other industries - is the device of averaging over a sufficiently long time interval. This imposes an unavoidable delay time and exposes the industry to the kinds of cycles that are observed. It is not widely recognized that the decisive advantage of statistical quality control methods is that they provide a real-time noise filter, allowing managers to discern the conditions which most urgently require action without waiting forever for the averages to settle down and without risking ill-advised interventions which will only amplify the noise and magnify the problem. The required information about process dispersion is very hard to come by. In insurance, it can only be obtained by a risk analysis: breaking the process down into components and reconstructing the variability without waiting around for things to happen. What Meyers has done here is a convincing first step in that direction. The ultimate goal is to make market capacity manageable in real time rather than in a feedback loop with a three-year time lag. If this is achieved, potential new entrants will not have to find out the hard way whether or not their contributions to the market were superfluous.

Connections

It is most illuminating to address the marginal surplus problem using the form for aggregate variance deduced in the context of the CME Model. Combining the two notations and using (1) and (6), we find for the needed surplus

$$S = Z\sqrt{A + \mathbf{n}^T \cdot \mathbf{U} + \mathbf{n}^T \cdot \mathbf{V} \cdot \mathbf{n}}\,. \tag{10}$$

Then the risk load vector giving the contribution to surplus needed to write an additional unit of exposure in any class is

$$\begin{aligned} \mathbf{R} &= \partial S / \partial \mathbf{n}^T \\ &= Z\frac{\frac{1}{2}\mathbf{U} + \mathbf{V} \cdot \mathbf{n}}{\sqrt{A + \mathbf{n}^T \cdot \mathbf{U} + \mathbf{n}^T \cdot \mathbf{V} \cdot \mathbf{n}}}. \end{aligned} \tag{11}$$

If we substitute the industry-average exposure spectrum, then (11) gives the same distribution of risk load by class as (9) as well as a way of re-expressing in terms of the other variables.

Even though it is seldom achieved in practice, the limit of large exposures is still instructive. The only terms that remain in this limit are those involving the systematic components of risk, parameter uncertainty and catastrophe hazard, as quantified in the matrix V:

$$\mathbf{R} \cong \frac{Z\,\mathbf{V} \cdot \mathbf{n}}{\sqrt{\mathbf{n}^T \cdot \mathbf{V} \cdot \mathbf{n}}} \tag{12}$$

In this limit, **R** depends only on how exposures are distributed among lines, classes, and limits and not on the total amount of exposure. This equation (12) is nothing more nor less than the expression of (5) in different notation and in the limit of large exposure.

To see this in closer detail, suppose that the exposure of the risk being insured is described by the vector **e** so that

$$\sigma^2 = \mathbf{e}^T \cdot \mathbf{U} + \mathbf{e}^T \cdot \mathbf{V} \cdot \mathbf{e}; \tag{13}$$

$$\Sigma^2 = A + \mathbf{n}^T \cdot \mathbf{U} + \mathbf{n}^T \cdot \mathbf{V} \cdot \mathbf{n}; \tag{14}$$

$$\begin{aligned} \Sigma^2 + \sigma^2 + 2\rho\Sigma\sigma &= A + (\mathbf{n}^T + \mathbf{e}^T) \cdot \mathbf{U} + \\ &\quad (\mathbf{n}^T + \mathbf{e}^T) \cdot \mathbf{V} \cdot (\mathbf{n} + \mathbf{e}); \end{aligned} \tag{15}$$

$$\begin{aligned} \Sigma^2 + \sigma^2 + 2\rho\Sigma\sigma &= A + \mathbf{n}^T \cdot \mathbf{U} + \mathbf{n}^T \cdot \mathbf{V} \cdot \mathbf{n} \\ &\quad + \mathbf{e}^T \cdot \mathbf{U} + \mathbf{e}^T \cdot \mathbf{V} \cdot \mathbf{e} \\ &\quad + 2\mathbf{n}^T \cdot \mathbf{V} \cdot \mathbf{e}. \end{aligned} \tag{15a}$$

From this we identify

$$\rho\sigma = \frac{\mathbf{n}^T \cdot \mathbf{V} \cdot \mathbf{e}}{\sqrt{A + \mathbf{n}^T \cdot \mathbf{U} + \mathbf{n}^T \cdot \mathbf{V} \cdot \mathbf{n}}} \,. \tag{16}$$

In the CME approach, the covariance identified above is essentially that of the individual risk with the overall market. This may be taken as a point of connection with the Capital Asset Pricing Model (CAPM), advocated by Feldblum (4) as a basis for risk loading. CAPM implies that assumption of "diversifiable" risk (process risk) cannot be compensated reliably in the marketplace whereas assumption of systematic risk (parameter uncertainty) can be and is regularly compensated. More accurately, compensation for assuming diversifiable risk is strongly dependent on market conditions. Furthermore, diversifiability is a relative notion, not an absolute one. If no actual opportunities for diversification exist, then diversifiability is merely an academic concept. Changing market conditions can have a radical effect on diversifiability of risk in insurance. For instance, the incipient market in insurance futures may prove to have just such an effect by broadening the opportunities for diversification.

Conclusions

At this point, it should be clear that we have a very striking convergence of different theoretical approaches, many of which have been thought to be mutually independent, or incompatible, or diametrically opposed. It may be helpful to draw up a thematic list of the concepts which have emerged in this inquiry and fitted themselves together like the pieces of a jigsaw puzzle:

- Marginal Cost Pricing
- Probability of Ruin
- Standard Deviation Principle
- Product Quality Assurance
- Variance Principle
- Market Equilibrium
- Constrained Optimization, Lagrange Multipliers, Efficient Frontiers
- Underwriting Cycles
- Capital Asset Pricing Model.

The only major concepts which have not surfaced are utility theory and the option pricing approach (5) although this does not deny their relevance. The point is that a viable theoretical approach to risk loading depends on the convergence of many ideas, the more the better. And anyone who says that the correct approach is one thing and not another is probably off the mark. The relevant question is how the pieces should fit together in a unified, convincing whole.

There is no doubting that the game is worth the candle because there is no lack of instances where a reliable and flexible pricing formula has significantly influenced market behavior - even to the point of creating markets where none existed previously. The most famous example of this is the Black-Scholes option pricing formula, which had enough descriptive power and conceptual plausibility to allow confident pricing of financial instruments which previously were traded sparsely and tentatively, if at all. This success, like success in any such endeavor, depended on gaining an understanding of the dispersion in the underlying process, on gaining control of the variability.

The insurance industry has traditionally been content to price to the mean and to rely on rules of thumb for governing variability - the Kenney Rule and the five percent underwriting margin are the best-known examples. The cost of this reliance is readily apparent because these rules of thumb leave yawning chasms of uncertainty in our knowledge of the industry's financial condition, its true underwriting capacity, and the quality of insurance products being offered to the public. Discounts are offered in slack markets with no reliable yardstick to gauge their financial consequences. This uncertainty is very costly indeed because it allows injection of excess capacity into a market that is already slack, although no one knows it yet. Further, capital committed to an insurance enterprise is not easily withdrawn. Capital injected into a slack market is likely as not to be consumed in subsidizing a superfluous, and practically irreversible market presence. Few entrepreneurs would fall into such a trap if they had adequate information; and, here, adequate information means quantitative estimates of risk, as well as cost, which accrue the uncertainties inhering in future events to the present and provide a basis for informed decision making. We have not merely to consider the variability of individual balance sheet items, but their covariability as well; we must learn to do accounting for variance if we are to bring the insurance process under control.

This is the goal, as we remarked earlier: real-time risk and capacity management. Without such a capability, the industry will remain locked in the predicament of having to relearn the same lessons every six years or so. The goal implies a challenge to the actuarial profession - for who will attain it if we do not? We must become as adept at characterizing, quantifying, and controlling variability as we have traditionally been at estimating expected values. If we do so, we will find ourselves doing both things better and more reliably.

Note also that, in discussing insurance risk, we have implied rather little, and said even less, about asset risk. Clearly, a "net present value" , and its variance, must involve the variability of assets as well as the countervailing liabilities. Only by considering both, can we get to the bottom line - and a useful result. Recent events in the industry have made it impossible to ignore asset risk. We may find ourselves broadening the purview of "actuaries of the third kind" to include the characterization and control of this bottom line variability.

The signs point hopefully to an early realization of these goals. The authors cited herein have done a great deal to hasten the day, and more yet if their works are examined together. It has been a pleasure to review their work and to underline its significance.

References

(1) Kreps, Rodney, "Reinsurer Risk Loads from Marginal Surplus Requirements", paper presented at the November, 1990, Meeting of the Casualty Actuarial Society, to be published in *Proceedings of the Casualty Actuarial Society.*

(2) Meyers, Glenn G., "The Competitive Market Equilibrium Risk load Formula", paper submitted to the Casualty Actuarial Society.

(3) See, *e.g.*, Heckman , Philip E., and Meyers, Glenn G., "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions", *Proceedings of the Casualty Actuarial Society*LXX, 1983, p. 22.

(4) Feldblum, Sholom, "Risk Loads for Insurers", paper presented to the May, 1990, Meeting of the Casualty Actuarial Society, to be published in *Proceedings of the Casualty Actuarial Society.*

(5) A useful review of this and other financial models is given in
Cummins, J. David, "Asset Pricing Models and Insurance Ratemaking", *ASTIN Bulletin* Vol. 20, No. 2, 1990.