The Top-Down Approach to Calculation of the Insurance Premium

Wojciech Otto, dr hab.; Faculty of Economic Sciences, Warsaw University

1. Introduction

Commercial price of insurance coverage fluctuates due to changing market conditions. Premium calculation produces response to the question of an indifference price for the insurance company—selling coverage below this price means making expected loss, selling above this price means making excessive expected profit. Both situations often happen in practice, but they usually do not persist for a long time because of actions undertaken in response to them by actors of this game. So the premium calculation problem can be read as reflecting the supply side considerations.

The premium calculation requires to evaluate a number of components:

- expected value of claims due to the contract in question,
- risk loading and profit margin,
- margin for expenses (acquisition, settlement, and administration costs),
- solidarity component.

The first two components are the most specific to the insurance business, and the risk theory focuses on them almost solely. It does not mean that the margin for expenses is irrelevant. Often expense margin is set well above 25% of the commercial premium. However, typical problems with this component (allocation of administration costs to individual contracts, allocation of initial costs along the lifetime of a product etc.) are similar to those encountered in any commercial activity, no matter if we sell shoes, cars, or insurance policies. The solidarity component is much more specific to insurance—it can be encountered when the insurance coverage is offered at the same price to two groups of risks that systematically differ. A good example is when, by some reasons, life annuities are offered to males and females by the same premium. Then the premium for males contains the positive solidarity component and that for females the negative solidarity component. Thus equal premium leads to redistribution from overpriced males to underpriced females. Existence of such redistribution sometimes is enforced by law, and the aim to reconcile it with rules of market competition poses many challenging problems. However, they are beyond the scope of problems considered in this article.

Out of the first two components that are considered here, one is non-controversial and obvious—it is the expected value of claims. It is the other one —risk loading and profit margin—that is much less obvious, and a bulk of literature is devoted to various approaches to it. Sometimes risk loading and profit margin are treated separately—especially when various risk measures are directly applied to evaluate individual risks without taking the context of the whole portfolio of risks accepted by the insurance company into account. This approach is especially valid when we focus on the demand side.

The so called top-down approach explicitly states that the insurer is exposed to risk that stems from an individual contract only as far as the contract contributes to the risk of the whole portfolio. Thus the approach entails two basic steps. First, the risk loading formula for the whole portfolio is chosen on the basis of risk and return considerations on the whole company level. Then the risk loading has to be allocated to individual risks. The first step is the phase when the capital needed to ensure solvency is assessed, and so the required return on that capital is taken into consideration. This is why the second step is often called "capital allocation". Despite it, under this approach it means essentially the same as "risk loading allocation". A good example is the case of a large decentralized company where pricing (ratemaking, granting rebates etc.) is to some extent delegated to regional divisions, but the problem that remains to be solved by top management is to allocate capital (expected contribution to cover costs of capital) among these divisions.

The problem arises when the pricing criteria applied at company level lead to non-additive premium formulas. In particular, under a number of different sets of assumptions the pricing formula with risk loading proportional to the standard deviation is obtained. Despite the standard deviation, the principle applied to the portfolio does not lead directly to the balanced allocation of the risk load to individual risks, the well known *ad hoc* solution is to allocate the risk load proportionally to variances. The aim of the article is to show that the game theory provides a new justification to this old solution.

2. Pricing the portfolio by the standard deviation principle

The standard deviation principle applied to the premium for the whole portfolio reads:

$$\Pi(W) = \mu_w + \alpha\sigma_w \tag{1}$$

where $W = X_1 + X_2 + \ldots + X_n$ is the aggregate amount of claims for the portfolio that consists of n individual risks $X_1, X_2, \ldots, X_n$, and μ_W, σ_W are expectation and standard deviation of the portfolio.

The simplest justification of this principle is that it can be derived from the more general quantile principle:

$$\Pi(W) = F_W^{-1}(1-\varepsilon)$$

which means that we accept a situation when the premium will not suffice to cover claims with (presumably low) probability ε, and that the cumulative distribution function F can be well approximated by the normal distribution. It is important here to emphasize that the common additional assumption of independency of risks $X_1, X_2, \ldots, X_n$ is not necessary, as approximated normality can be achieved despite some weak dependencies, provided the number of risks n is sufficiently large. Then parameter α of formula (1) equals $q_{1-\varepsilon}$, the quantile of order $(1-\varepsilon)$ of the standard normal distribution.

A more sophisticated justification of the standard deviation principle takes into account the trade-off between risk and return on capital. The expected rate of return on insurance operations can be represented by the postulated equatation formula:

$$\Pi(W) - \mu_W = ru$$

where r denotes the risk premium rate (rate of return in excess of the risk free rate i_{RF}), and u denotes the amount of capital that serves to back the insurance risk (it is assumed that this part of capital is invested in risk-free assets). The willingness of shareholders to accept the insurance risk is expressed in the form of the following requirement:

$$\Pr\left\{u\left(1+i_{RF}\right)+\Pi(W)-W<u\left(1+i_{RF}+r-\eta\right)\right\}=\varepsilon$$

which means that shareholders admit that the rate of return could fall down by η from the expected rate $i_{RF}+r$ with probability ε. Assuming that W is normally distributed we obtain:

$$\Pi(W) = u_W + q_{1-\varepsilon}\sigma_W + u(r-\eta)$$

Solving both equations in respect of the capital and the premium we obtain:

$$u = \frac{q_{1-\varepsilon}}{\eta}\sigma_W$$

and:

$$\Pi(W) = \mu_W + \frac{rq_{1-\varepsilon}}{\eta}\sigma_W$$

that is another case of the formula (1), indeed.

The model can be made more realistic by taking into account that a part of the capital is sunk in the infrastructure of the company and so the whole return on this part has to be made on insurance operations. Just another enhancement can be obtained by deriving the risk premium rate from the capital market quotations. This can be justified by the assumption that shareholders are indifferent whether to invest in the insurance company or elsewhere in the capital market.

Just another justification of the standard deviation principle comes from the ruin theory. The well known approximation to the probability of ruin in the long run Ψ given the initial capital u is provided by the formula:

$$\Psi(u) = \exp\left(2\frac{\Pi(W) - \mu_W}{\sigma_W^2}\right)$$

For the predetermined level ψ of the ruin probability we obtain the premium formula:

$$\Pi(W) = \mu_W + \frac{-\ln\psi}{2u}\sigma_W^2$$

However, the model assumes that yearly increments of the company's capital come from the technical result $\Pi(W) - W$ only, leaving the cost of capital not covered (or, allowing to cover the cost implicitly by returns on investment of current capital in external assets). Supplementing the model by the assumption that the premium should also cover the constant yearly dividend paid out to shareholders (as a reward for being exposed to the risk of ruin), we obtain:

$$\Pi(W) = \mu_W + \frac{-\ln\psi}{2u}\sigma_W^2 + du$$

where d denotes the dividend rate. Now the premium loading contains two separate components—pure risk loading and the profit margin. Minimizing the premium we find the most efficient level of initial capital:

$$u_{opt} = \sigma_W\sqrt{\frac{-\ln\psi}{2d}}$$

And the corresponding premium formula:

$$\Pi(W) = \mu_W + \sigma_W\sqrt{-2d\ln\psi}$$

that is again a case of formula (1).

The last model follows the seminal paper by Bühlmann [1985], who also promoted using the term "top-down approach". The approach itself has been known much earlier.

3. Marginal premium formula and the balancing problem

Provided the whole portfolio is priced by formula (1), the indifference price of the additional risk X (independent of risks $X_1, X_2, \ldots, X_n$) is given by the marginal premium formula:

$$\Pi_{marg}(X) \doteq \Pi(W + X) - \Pi(W) = \mu_X + \alpha(\sigma_{W+X} - \sigma_W) \tag{2}$$

Assuming that standard deviations of the portfolio with and without the additional risk do not differ too much, we can make an approximation:

$$\sigma_{W+X} - \sigma_W = (\sigma_{W+X} - \sigma_W)\frac{\sigma_{W+X} + \sigma_W}{\sigma_{W+X} + \sigma_W} = \frac{\sigma^2_{W+X} - \sigma^2_W}{\sigma_{W+X} + \sigma_W} \approx \frac{\sigma^2_X}{2\sigma_W}$$

that leads to the marginal premium formula:

$$\Pi_{marg}(X) = \mu_X + \frac{\sigma}{2\sigma_W}\sigma^2_X \tag{3}$$

However, the sum of marginal premiums calculated for all risks from the portfolio does not suffice to cover the aggregate risk:

$$\Pi(W) - \sum_{i=1}^{n}\Pi_{marg}(X_i) = \frac{1}{2}\alpha\sigma_W$$

and the deficit equals exactly one half of the required loading for the portfolio.

The balancing problem is nothing strange—the excess of average cost over the marginal cost is a common effect of returns to scale, and in the case of an insurance company, it is essence of the diversification effect obtained by pooling risks. Nevertheless, the problem of allocating the missing half of the loading $\alpha\sigma_W$ remains. An intuitively appealing rule is to allocate the "common safety fund" proportionally to the individual contribution to the loading. This leads to the basic pricing formula:

$$\Pi_b(X) = \mu_X + \alpha\frac{\sigma^2_X}{\sigma_W} \tag{4}$$

However, the following alternative solution is also balanced:

$$\Pi_b(X) = \mu_X + \alpha\sigma_W\left(\frac{\sigma^2_X}{2\sigma_w} + \frac{c_{X,k}}{2c_{W,k}}\right)$$

despite it allocates the missing half of the safety loading proportionally to cumulants of order k, where $k = 2$ is just a special case. Any weighted average of different versions of the above formula is balanced as well, and we need to somehow justify our choice.

4. Justification of the basic premium formula (4)

In order to explain the solution proposed by Borch [1962], let us imagine that all risks are insured sequentially. Under the particular ordering of risks from the basic portfolio $\{X_1, \ldots, X_j, X_{j+1}, \ldots, X_n\}$ the risk X_{j+1} is priced as if the first j risks were insured already. Thus the corresponding marginal premium formula reads:

$$\Pi_{marg}(X_{j+1}) = \mu_{j+1} + \alpha\left[\sqrt{\sigma^2_1 + \ldots + \sigma^2_{j+1}} - \sqrt{\sigma^2_1 + \ldots + \sigma^2_j}\right]$$

Obviously the sum of premiums calculated this way for all risks in the portfolio is equal to the portfolio premium, but the loading is allocated very unequally, charging heavily those risks that have been insured at first, and only slightly those ones that have been insured the latest. Borch suggested setting the price of a given risk as an average of prices calculated according to this rule under all $n!$ possible orderings, showing that this rule is just equivalent to the equilibrium solution of an n-person game (derived first by Shapley in 1953). Borch recommended this solution as suitable to the problem of allocating the loading between few groups of risks rather than to a large number of individual risks. Indeed, one can hardly imagine calculations for large n. This would require averaging over 2^{n-1} possibly different values, that's how many subsets the set of remaining risks may have, when one of n risks is priced.

However, in a case of large n an approximate solution can be derived. Let us consider again the problem of pricing an additional risk X when each of its possible $(n + 1)$ positions in the sequence $\{X_1, \dots, X_j, X, X_{j+1}, \dots, X_n\}$ is treated as equally probable, and all $n!$ orderings of remaining risks are equally probable, too. Let us denote the set of basic risks preceding the risk X in a particular ordering by PRE. The derivation is based on the remark that under certain conditions the ratio of the sum of variances of preceding risks (elements of PRE) in the aggregate variance of all risks in the portfolio is approximately uniformly distributed over the unit interval:

$$\underset{u\in[0,1]}{\forall} \Pr\left(\sum_{j\in PRE}\sigma_j^2 \le u\sum_{j=1}^{n}\sigma_j^2\right) \approx u \tag{5}$$

Let us make use of this approximation now, leaving its justification to the next sections. According to the approximation we can write:

$$E\left(\sqrt{\sum_{j\in PRE}\sigma_j^2+\sigma_X^2}-\sqrt{\sum_{i\in PRE}\sigma_i^2}\right)\approx\int_0^1\left(\sqrt{u\sigma_W^2+\sigma_X^2}-\sqrt{u\sigma_W^2}\right)du$$

Using the abbreviated notation $c \doteq \sigma_X^2/\sigma_W^2$ we can express the above integral as:

$$\int_0^1\left(\sqrt{u\sigma_W^2+\sigma_X^2}-\sqrt{u\sigma_W^2}\right)du=\sigma_W\int_0^1\left(\sqrt{u+c}-\sqrt{u}\right)du$$

Multiplying and dividing the integrated function by c, and taking into account that c is small we get the result:

$$\sigma_W c\int_0^1\frac{\sqrt{u+c}-\sqrt{u}}{c}du\approx\sigma_W c\int_0^1\left(\frac{\partial}{\partial u}\sqrt{u}\right)du=\sigma_W c=\frac{\sigma_X^2}{\sigma_W}$$

that indeed leads to the balanced pricing formula (4).

The above derivation was proposed in the book written in Polish [2004] and another one written in English [2005, Chapter 20]. However, in both cases the approximation (5) has been presented as "intuitively appealing". Next sections of this article contain the rigorous justification in terms of the theorem on convergence in distribution and its refinements.

5. Convergence to the uniform distribution

Approximation (5) can be justified by the convergence in distribution of the share of risks preceding a given risk in the variance of the portfolio to the uniform distribution. The convergence requires that under an increasing number of risks in the portfolio the share of maximal variance in the aggregate variance of the whole portfolio vanishes:

$$\lim_{n \to \infty} \frac{\max\left\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2\right\}}{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2} = 0 \qquad (6)$$

In order to prove sufficiency of the above requirement some more precise definitions, notations and assumptions are needed. Risks will be treated as elements of the set (portfolio), with the share of their variances in the variance of the whole portfolio assigned to each of them. Notations and assumptions are summarized below:

(A) $E = \left\{e_1, e_2, \ldots, e_n\right\}$ is a basic set of elements;

(B) $y: E \to [0,1]$ is a function that assigns the real nonnegative number $y_j \doteq y(e_j)$ to each element of the basic set, such that $y_1 + y_2 + \ldots + y_n = 1$;

(C) $M = \max_{j=1,\ldots,n} \{y_1, y_2, \ldots, y_n\}$denotes the maximum of these numbers;

(D) $E^* = E \cup \{e^*\}$is a basic set E supplemented by the special element e^*;

(E) U is defined as a sum of y_j characterising these elements e_j that precede special element e^* for a given ordering of elements of set E^*;

(F) the probability function assigns to each ordering of the set E^* the probability $1/(n + 1)$.

Under assumptions (A)–(F) the random variable U has a discrete distribution on the support [0, 1] that is uniquely determined by the set of real numbers $\{y_1, y_2, \ldots, y_n\}$. Let us denote the corresponding cdf by F_U. Now the theorem can be formulated as follows:

T h e o r e m 1. Under assumptions (A)–(F) the cdf F_U is uniformly bounded from both sides:

$$\forall u \in [0,1) \qquad \frac{u}{1+M} \le F_U(u) \le \frac{u+M}{1+M}$$

and equals zero for $u < 0$, and one for $u \ge 1$.

In lights of the theorem convergence of M to zero means that the distribution converges to the uniform distribution. Convergence of M to zero implies that n tends to infinity, as obviously $M \ge 1$. For a given M the number n is unbounded from above, of course.

5.1. Proof of the Theorem

It is obvious that each particular ordering of set E^* corresponds to the inverted ordering. Hence, both variables U and $(1 - U)$ must have the same distribution. That is why the upper bound for $F_U(u)$ in the theorem equals the lower bound for $1 - F_U(1 - u)$ and vice-versa. So it suffices to prove that one of the bounds is true.

Below, the proof for the upper bound is presented. The proof is based on induction, and shows that the upper bound assumed to be correct for a random variable U determined by any admissible sequence $\{y_j\}$ containing $(n - 1)$ elements, implies that it is also correct for a variable U determined by any admissible sequence $\{y_j\}$ that contains n elements. The first step then, is to show that the upper bound is correct for $n = 1$.

1st step of the proof

For $n = 1$ the value y_1 equals one, and so $M = 1$. Variable U can take only values 0 or 1, each of them with probability 1/2. So the cdf $F_U(u)$ takes a value 1/2 on the interval $u \in [0, 1)$ and the value of 1 at the right endpoint $u = 1$. Thus the cdf obviously satisfies the inequality $F_U(u) \leq (u + 1)/2$ for all $u \in [0, 1]$.

The same distribution of variable U concerns also the case when set $\{y_1, y_2, \ldots, y_n\}$ contains only one positive element so that $y_j = M = 1$ for a given j. This is because in all orderings the value of the variable U depends only on the mutual positions of e_j and e^*. Thus we can restrict further considerations to the case when $n \geq 2$ and $M < 1$.

2nd step of the proof

For the purpose of this step of the proof let us denote by U the variable determined by a set of elements $\{e_1, e_2, \ldots, e_n\}$ and a corresponding set of values $\{y_1, y_2, \ldots, y_n\}$, and assume that $n \geq 2$ and $M < 1$. Let us also assume, for simplicity, that there are no larger values of y_j than the first one, so $y_1 + M$, and let us denote the maximum out of the subsequence $\{y_2, \ldots, y_n\}$ by M_1. Summarizing we have $0 < M_1 \leq M < 1$.

Let us also define the corresponding family of n random variables U_j, for each $j = 1, 2, \ldots, n$ being determined by the set $E_j \doteq E \setminus \{e_j\}$ and the same function y. Of course, assumptions (A)–(F) are satisfied for variables $U_j/(1 - y_j)$, and not for variables U_j themselves. So, assuming that the upper bound stated in the theorem is correct for all variables $U_j/(1 - y_j)$, means that for all $u \in [0, 1]$ the following inequalities hold:

$$\Pr\left(U_1 \leq u\right) \leq \frac{u + M_1}{1 + M_1 - M} \tag{7}$$

$$\Pr\left(U_j \leq u\right) \leq \frac{u + M}{1 + M - y_j} \qquad j = 2, 3, \ldots, n \tag{8}$$

The above upper bounds could be easily tightened, making use of simple remark that a probability is never greater than 1. However, for our purposes it suffices to refine the upper bound for the variable U_1 only. The refinement takes a form:

$$\Pr(U_1 \leq u) \leq u + M \tag{9}$$

For all $u > 1 - M$ the above inequality is correct as a probability is never greater than 1. In the case when $u \leq 1 - M$ its correctness could be deduced as follows:

- $u \leq 1 - M$ implies that: $-1 \leq -(M + u)$;
- but by definition $M \geq M_1$, so multiplying both sides of the last inequality by $(M - M_1)$ we obtain: $(M_1 - M) \leq (M_1 - M)(M + u)$,
- adding the term $M + u$ to both sides we obtain: $(M_1 - M) \leq (1 + M_1 - M)(M + u)$,
- finally, dividing both sides by $(1 + M_1 - M)$ and combining the result with (7) we obtain the confirmation that inequality (9) holds.

Now we can return to the distribution of variable U. Let us define in the space of all $n!$ orderings of elements of the set E the event A_j, that the element e_j has taken the last position in the ordering. The events $A_1, A_2, \ldots, A_n$ defined this way are separate, equally probable, and altogether they cover the entire space. Hence, we can write:

$$\Pr(U \leq u) = \frac{1}{n}\sum_{j=1}^{n}\Pr\left(U \leq u \middle| A_j\right) \tag{10}$$

However, conditional probabilities appearing on the RHS concerning events expressed in terms of values of variable U, could be expressed in terms of values of variables U_j as well. In order to do that, let us notice that for any given ordering of elements of set E there are $(n + 1)$ corresponding orderings of elements of the extended set that only differ by the position of special element e^*. When an arbitrary ordering of set E encompassed by event A_j takes place, then almost always (with one exception) variable U is identical to variable U_j. This exceptional case happens when special element e^* takes the last position, but then obviously $U = 1$. That is why we can write:

$$\forall u \in [0,1) \qquad \Pr\left(U \leq u \middle| A_j\right) = \frac{n}{n+1}\Pr\left(U_j \leq u\right) \qquad j = 1, 2, \ldots, n \tag{11}$$

Combining now (10) and (11) we obtain:

$$\forall u \in [0,1) \qquad \Pr(U \leq u) = \frac{n}{n+1}\sum_{j=1}^{n}\Pr\left[U_j \leq u\right] \tag{12}$$

Making use now of upper bounds (8) and (9) we conclude that:

$$\forall u \in [0,1) \qquad \Pr(U \leq u) \leq \frac{n}{n+1}\left\{M + u + \sum_{j=2}^{n}\frac{M+u}{M+1-y_j}\right\} \tag{13}$$

Considering into account that for any j the following equality holds:

$$\frac{M+u}{M+1-y_j}=\frac{M+u}{M+1}\left\{1+\frac{y_j}{M+1-y_j}\right\}$$

we can transform (13) to the form:

$$\forall u\in[0,1)\qquad \Pr(U\le u)\le\frac{M+u}{n+1}\left\{1+\frac{1}{M+1}\left[n-1+\sum_{j=2}^{n}\frac{y_j}{M+1-y_j}\right]\right\}\tag{14}$$

Now we can make two remarks:

- that each component of the sum appearing on RHS of inequality (14) can be bounded from above: $y_j/(M+1-y_j)\le y_j$, and:
- that the sum $(y_2+y_3+\ldots+y_n)$ equals $(1-M)$.

These remarks allow for simplification of the upper bound (14) to the form:

$$\forall u\in[0,1)\qquad \Pr(U\le u)\le\frac{M+u}{n+1}\left\{1+\frac{1}{M+1}\left[n-1+(1-M)\right]\right\}\tag{15}$$

Simple manipulations now allow for transforming the RHS to the desired form of the upper bound $(M+u)(M-1)$, and to extend this result to the right endpoint of interval [0, 1].

5.2. Are the bounds for the cdf tight enough?

The answer is that bounds stated in the theorem cannot be significantly tightened unless we impose additional restrictions on the sequence $\{y_1, y_2, \ldots, y_n\}$. This can be shown by considering the worst case when our portfolio consists of n risks with equal variances so that all numbers $y_j = 1/n$ for $j = 1, 2, \ldots, n$. Then the cdf takes on the interval the form:

$$\forall u\in\left[\frac{k-1}{n},\frac{k}{n}\right)\quad F_U(u)=\frac{k}{n+1}\qquad k=1,2,\ldots,n$$

On each subinterval the upper bound is attained at the left endpoint and the lower bound is almost attained near the right endpoint. Hence, we cannot tighten the bounds except by replacing the inequality symbol ≤ concerning the lower bound by the strict inequality symbol <. This is not a significant improvement, of course.

On the other hand, any additional restriction imposed on the sequence $\{y_1, y_2, \ldots, y_n\}$ may lead to more significant improvements of the bounds stated by the theorem. However, the variety of possible cases is unlimited, so there is a need to focus on these cases when the bounds stated by the theorem are unsatisfactory.

6. Bounding the premium loading

Let us inspect now how the theorem works when used to bound the premium loading. The loading for an additional risk X equals the expectation $E[h(U)]$, where the function h is defined as:

$$h(u) \doteq \sqrt{u\sigma_W^2 + \sigma_W^2} - \sqrt{u\sigma_W^2} \text{ (for simplicity we assume } \alpha = 1)$$

and the expectation is calculated in respect to the true distribution of variable U. Let us now denote the bounds for the cdf $F_U(u)$ stated by the theorem for $u \in [0, 1)$ by:

$$\underline{\underline{F}}(u) \doteq \frac{u}{1+M}, \text{ and } \overline{\overline{F}}(u) \doteq \frac{u+M}{1+M},$$

where both bounds reach 1 at the right endpoint $u = 1$. As cdf's $\underline{F}$, F_U and $\overline{F}$ are stochastically ordered and the function $h(u)$ is decreasing on interval $u \geq [0, 1]$, we can express the bounds for the premium loading as follows:

$$\int_{[0.1]} h(u)d\underline{F}(u) \leq \mathrm{E}(h(U)) \leq \int_{[0,1]} h(u)d\overline{F}(u).$$

In order to justify approximation of the loading by the portfolio loading σ_{W+X} multiplied by the share of the priced risk in the variance of the whole portfolio $\sigma_X^2/\sigma_{W+X}^2$, we should inspect whether the bounds divided by the expression σ_X^2/σ_{W+X} are close to unity. Denoting (as before) the ratio σ_X^2/σ_W^2 by c and executing all necessary calculations, we obtain:

$$\frac{\sigma_{W+X}}{\sigma_X^2}\int_{[0,1]} h(u)d\underline{F}(u) = \sqrt{1+c}\,\frac{M\left(\sqrt{1+c}-1\right)+\frac{2}{3}\left[(1+c)^{3/2}-1-c^{3/2}\right]}{c(1+M)} \tag{16}$$

for the lower bound and:

$$\frac{\sigma_{W+X}}{\sigma_X^2}\int_{[0,1]} h(u)d\overline{F}(u) = \sqrt{1+c}\,\frac{M\sqrt{c}+\frac{2}{3}\left[(1+c)^{3/2}-1-c^{3/2}\right]}{c(1+M)} \tag{17}$$

for the upper bound.

Now we can consider various scenarios.

i. The scenario when we price a single large risk on the background of the portfolio of very numerous small risks can be inspected by assuming fixed c and $M \to 0$. Then, both bounds tend to the same function $\frac{2\sqrt{1+c}}{3c}\left[(1+c)^{3/2}\right]-1-c^{3/2}$ that is close to 1 for a reasonably small c.
ii. The scenario when the size of the priced risk X is comparable to the size of the largest risk, so that $c = const \cdot M$. There is still no problem when we assume now that M tends to zero, as in this case both bounds tend to 1.
iii. The scenario when M is fixed and c tends to zero corresponds to the case when we allow for some large risks in the portfolio and try to price a risk that is incomparably smaller. In this case the lower bound tends to $\left(1+\frac{1}{2}M\right)(1+M)^{-1}$, that is still acceptable (at least for small M), but the upper bound tends to infinity, that is no longer acceptable.

Diverging the upper bound in scenario iii. is not due to poor bounding. Indeed, in the case of portfolio of n risks of equal size such that $Mn = 1$ and an additional risk characterized by c the exact result reads:

$$\frac{\sigma_{W+X}}{\sigma_X^2}\int_{[0,1]} h(u)dF_U(u) = \frac{\sqrt{1+c}}{c}\frac{1}{n+1}\left\{\sqrt{c}+\sum_{k=1}^{n}\left(\sqrt{\frac{k}{n}+c}-\sqrt{\frac{k}{n}}\right)\right\}$$

and for any fixed n and $c \to 0$ diverges to infinity.

However, the assumption that the company basically specializes in insuring large risks, and considers pricing only one incomparably small risk, is unrealistic. In practice, companies do have portfolios that might be composed of some large risks and a large number of small risks. Even when each small risk is incomparably smaller than the largest one, the aggregate variance of all small risks usually contributes substantially to the variance of the whole portfolio. In the next section a scenario of this kind is considered in details.

7. Pricing small risks in the presence of large risks

Our considerations are restricted here only to the special case of a portfolio composed both of large risks and small risks. For simplicity we assume that there are n large risks of the same share in the variance of the whole portfolio m such that their aggregate share $S = mn$ is less than 1. The lasting $(1 - S)$ is the share of the aggregate variance of a very large number of very small risks. Variance of each small risk is assumed to be infinitesimally smaller than the variance of a large risk.

Now the aggregate premium loading allocated to small risks can be bounded from above by the total portfolio loading less the lower bound for the aggregate loading allocated to large risks. More precisely, if we denote the ratio of the Shapley loading to the loading proportional to variance for the large risk by G and the analogous ratio for the aggregate of small risks by g, then the balancing equation holds: $(1 - S)g + SG = 1$. Thus, if we know that a lower bound $\underline{G}$ such that $G \geq \underline{G}$ exists, then the resulting upper bound $\bar{g}$ such that $g \leq \bar{g}$ equals:

$$\bar{g} = \frac{1 - S\underline{G}}{1 - S} \tag{18}$$

In order to derive the lower bound $\underline{G}$ we can apply formula (16) to price one of large risks on the background of the portfolio composed of lasting $(n - 1)$ large risks and all small risks. This means that we should make the following replacements on the RHS of (16):

- for $n = 1$ we replace c by $m(1 - m)^{-1}$ and M by zero; this leads to the exact result (as in this case lower and upper bounds coincide):

$$\underline{G} = G = \frac{2\left[1-(1-m)^{3/2}-m^{3/2}\right]}{3m(1-m)}$$

- for $n > 1$ we replace both c and M by $m(1-m)^{-1}$; then the lower bound for G reads:

$$\underline{G} = 1 - \sqrt{1-m} + \frac{2\left[1-(1-m)^{3/2} - m^{3/2}\right]}{3m}$$

Combining the last results with (18) and replacing m by (S/n) we obtain the upper bound for ratio g as a function of S and n:

$$\bar{g}(S,1) = g(S,1) = 1 + \frac{S}{1-S} + \frac{2\left[(1-S)^{3/2} + S^{3/2} - 1\right]}{3(1-S)^2} \text{ for } n = 1, \text{and:} \tag{19}$$

$$\bar{g}(S,n) = 1 + \frac{s}{1-s}\left\{\sqrt{1-\left(\frac{S}{n}\right)} + \frac{2n}{3S}\left[\left(1-\frac{S}{n}\right)^{3/2} + \left(\frac{S}{n}\right)^{3/2} - 1\right]\right\} \text{ for } n > 1 \tag{20}$$

Simple calculations (employing de Hospital's rule etc. to formulas (19) and (20)) lead to the conclusion that:

$$\lim_{S\to 1}\bar{g}(S,n) = +\infty \text{ for any fixed positive integer } n \tag{21}$$

The result confirms that in order to have an upper limit for the ratio of the Shapley loading to the loading proportional to variance for small risks, we have to assume that the share of small risks in the portfolio is non--neglectible.

In order to review in turn the case of fixed S and large n the approximation $\sqrt{1-S/n} \approx 1 - S/(2n)$ can be used. This yields the approximated upper bound:

$$\bar{g}(S,n) \approx 1 + \frac{2S^{3/2}}{3(1-S)}\frac{1}{\sqrt{n}}$$

where neglected terms are of order n^{-1}, $n^{-3/2}$, etc. This obviously means that:

$$\lim_{n\to\infty}\bar{g}(S,n) = 1 \text{ for any fixed } S \in (0,1) \tag{22}$$

The last result means that when we enlarge unlimitedly the heterogeneous portfolio by accepting new large as well as new small risks (but keeping the proportion S fixed), then the ratio tends to 1.

In order to illustrate how the upper bounds given by formulas (19) and (20) work in non-extreme cases, some results are calculated and presented in Table 1 below.

Table 1.

Upper bounds $\bar{g}(S,n)$ for the function $g(S,n)$

n	S = 0.25	S = 0.50	S = 0.75	S = 0.90
Exact values				
1	106.6%	121.9%	159.5%	236.3%

n	S = 0.25	S = 0.50	S = 0.75	S = 0.90
Upper bounds				
2	106.8%	126.5%	189.8%	380.5%
3	105.7%	122.8%	179.4%	353.0%
4	105.0%	120.3%	171.6%	329.6%
5	104.5%	118.5%	165.6%	311.3%
10	103.3%	113.6%	149.0%	259.1%
20	102.4%	109.9%	135.9%	217.0%
40	101.7%	107.1%	126.0%	184.9%
75	101.3%	105.3%	119.2%	163.0%
150	100.9%	103.8%	113.8%	145.1%
300	100.6%	102.7%	109.8%	132.2%
600	100.5%	101.9%	107.0%	122.9%
1200	100.3%	101.4%	105.0%	116.3%
2400	100.2%	101.0%	103.5%	111.5%
4800	100.2%	100.7%	102.5%	108.2%
9600	100.1%	100.5%	101.8%	105.8%

The content of the Table show that upper bounds given by (20) probably produce quite tight bounds when S is smaller than half. On the other hand, bounds for $S = 0.75$ and $S = 0.90$ seem to be heavily overestimated. This stems from the guess that for given S the ratio $g(S, n)$ should be a decreasing function of n. So the value $g(S, 2)$ should be smaller than $g(S, 1)$. Thus, the value $\bar{g}(S, 2)$ larger than $g(S, 1)$ for each case presented in the Table, seem to be an effect of overestimation of the true value by the bound (20). This overestimation is very small for $S = 0.25$, moderate for $S = 0.50$, quite substantial for $S = 0.75$ and very high for $S = 0.90$.

Leavingdoubts concerning the accuracy of bounding aside , we can try to draw practical conclusions that stem from the presented results. Let us assume for instance that we accept the approximation of the Shapley value by the variance principle if $g(S, n) > 110\%$. Results presented in Table 1 allow to conclude that the approximation is acceptable:

- in case of $S = 25\%$: for any n (so that any $m \le 25\%$ is acceptable);
- in case of $S = 50\%$: for $n \ge 20$ (so that $m \le 2.5\%$ is acceptable);
- in case of $S = 75\%$: for $n \ge 300$ (so that $m \le 0.25\%$ is acceptable);
- in case of $S = 90\%$: for $n \ge 4800$ (so that $m \le 0.019\%$ is acceptable).

8. Conclusions and extensions

Results presented in this article confirm that the variance principle applied to allocate the premium loading that has been set for the whole portfolio as proportional to the standard deviation can be justified as an approximation to the Shapley value. This result is also relevant for one reason that has not been exploited in the paper. Despite the sound interpretation, the

Shapley value has an obvious deficiency when used for pricing. This is because it does not remain unchanged under aggregation of risks—the Shapley value of a subportfolio of risks does not equal the sum of Shapley values of individual risks from this subportfolio. Perhaps the best approximation of the Shapley value that is free of this deficiency is the variance principle.

The approximation may appear to be poor when the portfolio contains both very large risks and very small risks, and when the share of small risks in the variance of the whole portfolio is small. However, when the portfolio contains some very large risks (say, m larger than one percent), then the distribution of the aggregate amount of claims significantly departs from the normal one. Thus the standard deviation principle is no longer adequate for setting the portfolio premium.

It is more difficult to defend the presented conclusion in the case when the share of small risks in the variance of the whole portfolio is small, 10% for instance. Then the maximum share of a large risk in variance m at which the approximation is still acceptable is less than 0.02%.

However, the requirements on the structure of the portfolio seem to be far too restrictive for two reasons:

- We have only studied in more detail the case when $S = mn$, whereas in practice collections of large risks are heterogeneous. It seems that for a given S and m the heterogeneity reflected by the number of large risks n greater than S/m will lead to more moderate ratio g of Shapley loading to the variance loading.
- Even in the case $S = mn$ we have derived upper bounds for the ratio g that seem to overestimate the true values, especially in the case of a relatively large S.

Both limitations of analysis presented in this paper are due to the aim of exploiting as much as possible Theorem 1, and to avoid going too far beyond. During the WEM Conference some preliminary results based on direct analysis of the mixed game (mixed of atoms and the ocean, as it is called in the language of the Game Theory) have been presented. More mature results have been presented few weeks later on the IME Conference in Leuven.

Literature

Borch Karl H., 1962, “Application of Game Theory to Some Problems in Automobile Insurance”, *ASTIN Bulletin* Vol. 2, No. 2, pp. 208–221.

Buhlmann Hans, 1985, “Premium Calculation from Top Down”, *ASTIN Bulletin* Vol. 15, No. 2.

Miśta Paweł, Wojciech Otto, 2005. “Premiums, Investments and Reinsurance”—Chapter 20, pp. 453–488, in *Statistical Tools for Finance and Insurance*, ed. Pavel Cizek, Wolfgang Haerdle and Rafał Weron, Springer Verlag.

Otto Wojciech, 2004, *Ubezpieczenia majątkowe*—Część I—*Teoria ryzyka*, WNT, Warsaw, Poland.

Shapley L.S., 1953, “A value for *n*-person games”, *Annals of Mathematical Studies*, Princeton, pp. 307–317.

Abstract

When the risk loading for the whole portfolio is set proportionally to the standard deviation, then the problem of coherent pricing of individual risks arises. Borch (1962), proposed a solution based on Shapley's value of the n-person game. However, the solution is suited only for small n, rather reflecting the game played by few companies that negotiate pooling their portfolios. Otto (2004) proposed an "intuitively appealing" approximation for the case of large n that leads to allocation of the risk loading proportionaly to variances. The paper is devoted to formally justify that the variance principle can be justified as an approximation to the Shapley's solution.