May 2013 N°24

SCOR Papers

Does risk diversification always work? The answer through simple modelling

By **Marc Busse**[1], **Michel Dacorogna**[1] **and Marie Kratz**[2]

[1]*SCOR SE, Switzerland*

[2]*ESSEC Business School Paris - CREAR, France*

Abstract

With a simple example of throwing a dice, we show how to price an insurance policy. We further study how this price decreases when many similar policies are sold. The diversification benefits increase with the number of policies and similarly the risk loading of the premium required for the risk decreases tending to zero. This is true as long as the risks are completely independent. However, when introducing in addition a biased dice played by a crooked croupier, a non-diversifiable risk does appear. Indeed, we can show analytically that, with the biased dice, there exists an additional term in the variance, which does not decrease with the number of policies in the portfolio and leads to a limit to diversification. We propose and study analytically three cases of introducing the non-diversifiable risk. For each of them, the behavior of the risk loading based on the underlying risk process is examined and a numerical illustration is provided. Then the results are discussed in view of the risk loading. Such a modelling could be used to study particular investment choices under uncertainty.

SCOR Global Risk Center

1 Introduction

A risk is the potential that a chosen action, including the choice of inaction, will lead to uncertainty concerning the future outcome. This notion also implies that our choice can have an influence on the future outcome. Risk and uncertainty are often used indistinctly, but in economics the tradition wants to distinguish between the two concepts. In insurance, "Risk" may be defined as the randomness with knowable probabilities (measurable uncertainty), while "Uncertainty" is the randomness with unknowable probabilities (unmeasurable uncertainty). Insurances are the providers of a way of reducing the risk, but stay away of uncertainty. To have a better understanding of how insurances price risks and what problems they are facing, we present here a simple example, where the risk is well quantifiable (see also [1]).
We assume an insurance customer approaches a company with the aim to insure a risk modeled by the throwing of a die (measurable uncertainty) as follows:

- the customer must pay 10 EUR every time a die displays a 6 and nothing otherwise
- moreover he throws the die 6 times

The customer would like to know the price an insurance company is going to ask him to cover such risk.
We have to consider if the customer is the only one asking for such an insurance or if there are other people asking for the same product. Of course, in reality, there is always an amount of persons that need to be covered for the same risk. Here, we start by calculating the risk premium for one customer (which means one insurance policy). Then, we generalize the approach for several customers, which means considering a portfolio containing several policies. This allows us to explore the effect of diversification on the price of risk. We will examine various numbers of policies in the portfolio to see how the risk is reduced, when increasing the size of the portfolio.
The idea is that the more policies we have, the more diversified is the portfolio of policies, thus the risk and the cost of capital are diminished. In this paper, we will examine how diversification works and what can be its limitation.

2 Fair Game – a first probabilistic model

Consider a fair game, when the die is unbiased and it is equally probable to get one of the six faces of the die.

One policy case Let X be a Bernoulli random variable (rv) defined on a probability space $(\Omega, \mathscr{A}, \boldsymbol{P})$ representing the loss obtained when throwing the unbiased die, i.e when obtaining a "6".

$$X = \begin{cases} 1 & \text{with probability} \quad p = 1/6 \\ 0 & \text{with probability} \quad 1-p \end{cases}$$

Recall that $\boldsymbol{E}(X) = p$ and $var(X) = p(1-p)$.
The loss amount when playing once is modeled by lX, with $l = 10$ EUR (in our example).

Note that all rv's introduced in this paper will be defined on the same probability space $(\Omega, \mathcal{A}, \boldsymbol{P})$.
Let $(X_i, i = 1, \ldots, n)$ be a n-sample with parent rv X, corresponding to the sequence of throws when playing n times (independent games). In our example $n = 6$. The number of losses after n games is modeled by $S_n = \sum_{i=1}^{n} X_i$, a binomial distribution $\mathcal{B}(n, p)$. Recall that

$$\boldsymbol{P}[S_n = k] = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \cdots, n, \tag{1}$$

$\boldsymbol{E}(S_n) = \sum_{i=1}^{n} \boldsymbol{E}(X_i) = n\boldsymbol{E}(X) = np$ and that, by independence,

$$var(S_n) = \sum_{i=1}^{n} var(X_i) = n\ var(X) = np(1-p)$$

Table 1: The loss distribution for a fair die with $\mathrm{n} = 6$ throws.

number of "6" k	Loss per game $l\ X(\omega)$	Probability Mass $\boldsymbol{P}[S_n = k]$	Cdf $\boldsymbol{P}[S_n \le k]$
0	0	33.490%	33.490%
1	10	40.188%	73.678%
2	20	20.094%	93.771%
3	30	5.358%	99.130%
4	40	0.804%	99.934%
5	50	0.064%	99.998%
6	60	0.002%	100.000%

From now on, we consider the case $n = 6$ (but keep the notation n for 6) and we denote the total loss amount by $L := l \times S_n$.
We are interested in knowing the risk premium the insurance will ask to the customer if he buys the insurance policy.
In Table 1, we show both the probability of losses as well as the cumulative probability of losses according to (1). The expected total loss amount is given by $\boldsymbol{E}(L) = l\boldsymbol{E}(S_n) = 10$. We see that there is a 26.32% probability ($\boldsymbol{P}[S_n > 10] = 1 - \boldsymbol{P}[S_n \le 10]$) that the company will turn out paying more than the expectation. Thus, we cannot simply ask the expected loss as premium. Insurance companies are supposed to guarantee the payment of a claim up to a high probability. We thus need to include some capital to guarantee this payment at a certain high probability. This is the subject of the next sections.

The case of a portfolio of N policies We consider now the case where the insurance company holds a portfolio of N identical and independent policies. The generalization to this case is immediate. It simply means to consider our example with Nn possible outcomes, so it could be modeled by a Bernoulli rv for each throw when playing Nn times. The total number of losses in the portfolio of N policies with n throws can then be represented by the new rv $Z = \sum_{i=1}^{Nn} X_i$, which is $\mathscr{B}(Nn, p)$ distributed. The total loss amount of the portfolio becomes $L = lZ$, with cdf denoted by F_L.

3 Computation of the risk premium

3.1 The technical premium

One policy case For any risk incurring a loss L, we can define, as in [1], the technical premium, P, that needs to be paid, as:

$$P = \boldsymbol{E}(L) + \eta K + e \quad \text{with} \tag{2}$$

η: the return expected by shareholders before tax
K: the capital assigned to this risk
e: the expenses incurred by the insurer to handle this case.
We will assume in this example that the expenses are a small portion of the expected loss

$$e = a\boldsymbol{E}(L) \quad \text{with } 0 < a << 1$$

which transforms the premium, using the definition of L, as

$$P = nl(1+a)\boldsymbol{E}(X) + \eta K \tag{3}$$

This example illustrates the case of one policy with n possible outcomes.
From now on, we will consider a portfolio of N independent policies (or contracts).

The case of N policies The premium for one policy within the portfolio can be deduced from (3) as

$$P = \frac{Nnl(1+a)\boldsymbol{E}(X) + \eta K_N}{N} = nl(1+a)\boldsymbol{E}(X) + \frac{K_N}{N}\eta$$

where K_N is the capital assigned to the entire portfolio.
Notice, as mentioned above, that L corresponds then to the total loss amount lZ. We will name the risk incurred by the portfolio by L as well, from which we can deduce the capital K_N assigned to the entire portfolio.

3.2 Risk Measures

First we have to point out the role of capital for an insurance company. It ensures that the company can pay its liability even in the worst case. In our case the capital will cover the risk with 99% probability. For this, we need to define the capital we have to put behind the deal. We are going to

use a risk measure for this. We will consider two standard risk measures, the Value-at-Risk (VaR) and the tail Value-at-Risk (TVaR). Let us remind the definitions of these quantities (see e.g.[3]).
The Value-at-Risk with a confidence level α is defined for a risk L by

$$\mathrm{VaR}_\alpha(L) = \inf\{q \in \mathbb{R} : \boldsymbol{P}(L > q) \le 1-\alpha\} = \inf\{q \in \mathbb{R} : F_L(q) \ge \alpha\} \tag{4}$$

where q is the level of loss that corresponds to a VaR_α (simply the quantile of L of order α).
The tail Value-at-Risk at a confidence level α of a continuous L satisfies

$$\mathrm{TVaR}_\alpha(L) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(L)\mathrm{d}u$$

In our case, it can be approximated by a discrete sum, which may be seen as the average over all losses larger than VaR_α:

$$\mathrm{TVaR}_\alpha(L) = \frac{1}{1-\alpha}\sum_{u_i \ge \alpha}^{1} q_{u_i}(L)\Delta u_i \tag{5}$$

where $q_{u_i}(L) = \mathrm{VaR}_{u_i}(L)$ and $\Delta u_i \equiv u_i - u_{i-1}$ corresponds to the probability mass of the particular quantile q_{u_i}.
These two kinds of risk measure, which we call, in full generality, ρ, allow us to define the capital needed to ensure payment of the claim up to a certain confidence level. We then define the risk-adjusted-capital K as a function of the risk measure ρ associated to the risk L as

$$K = \rho(L) - \boldsymbol{E}(L) \tag{6}$$

since the risk is defined as the deviation from the expectation.
We will choose in the rest of the study a threshold of 99% for α.
We could have also defined K as $K = \rho(L) - \boldsymbol{E}(L) - P$ since the premiums can serve to pay the losses. It would change the premium defined in (2) as follows, we call the new one $\hat{P}$:

$$\hat{P} = \frac{1+a-\eta}{1+\eta}\,\boldsymbol{E}(L) + \frac{\eta}{1+\eta}\rho(L)$$

Such an alternative definition would reduce the capital but does not change fundamentally the results of the study.

3.3 Cost of Capital and Risk Loading

An insurance is a company in which we can invest. Therefore the shareholders that have invested a certain amount of capital in the company expect a return on investment. So the insurance firm has to make sure that the investors receive their dividends, which corresponds to the cost of capital the insurance company must charge on its premium. In this example, we fix the percentage of cost of capital η, before taxes, at 15%. It will give the shareholders a return on investment after taxes of approximately 10%, when considering a standard tax rate of 30%.

Let R denote the risk loading per policy, defined as the cost of the risk-adjusted-capital per policy. Using (6), R can be expressed as a function of the risk measure ρ, namely

$$R = \eta \, \frac{K}{N} = \eta \left(\frac{\rho(L)}{N} - nl\boldsymbol{E}(X) \right) \tag{7}$$

The risk measure ρ can here be either VaR or $TVaR$ (see (4) and (5)).
We revisit our numerical example with $n = 6$, $l = 10$ and $\boldsymbol{E}(X) = 1$, $\alpha = 99\%$ and $\eta = 0.15$; for a discussion on the choice of the values of η and e, we refer to [1]. We compute the cost of capital given in (7) for an increasing number N of policies in the portfolio. The results are displayed in Table 2 for both VaR and TVaR.

Table 2: The Risk loading per policy as a function of the number of policies in the portfolio.

Number N of Policies	Risk Loading R with $\rho = VaR(\alpha)$	$\rho = TVaR(\alpha)$
1	3.000	3.226
5	1.500	1.644
10	1.050	1.164
50	0.450	0.510
100	0.330	0.372
1'000	0.102	0.116
10'000	0.032	0.037

When considering a large number N of policies, the binomial distribution of Z, and so of L, can be replaced by the normal distribution $\mathcal{N}(Nnp, Nnp(1-p))$ (for $Nn \geq 30$ and p not close to 0, nor 1; e.g. $np > 5$ and $n(1-p) > 5$) using the Central Limit Theorem (CLT). The VaR of order α of Z can then be deduced from the αth-quantile, q_α, of the standard normal distribution $\mathcal{N}(0,1)$, as:

$$VaR_\alpha(Z) = \sqrt{Nnp(1-p)}\; q_\alpha + Nnp.$$

Thus the risk loading R will become, in the case of ρ being VaR:

$$R = \eta \times \sqrt{\frac{nlp(1-p)}{N}}\; q_\alpha$$

ever smaller as a function of N.

We can see in Table 2 that the risk loading drops practically by a factor 100 for a portfolio of 10'000 policies. It is worth noting that the risk loading with $TVaR$ is always slightly higher than with VaR for the same threshold, as TVaR goes beyond VaR in the tail of the distribution.
For a portfolio consisting of one policy ($N = 1$), the risk loading represents 30% of the loss expectation ($\boldsymbol{E}(L) = 10$ in our example). This is uneconomical: nobody would pay such a price for covering a risk. This issue can be solved by bundling many of those risks in a portfolio. Looking at the results in Table 2, it appears that it is the right way to proceed, as the risk almost disappears if we bundle enough risks together. With 10'000 policies the risk loading is almost 100 times lower than with one policy. This observation fascinated Andrej Kolmogorov, the discoverer of the CLT. Unfortunately, as often in life, there is a limit to this. We are going to explore it in the next section.

4 Limits to diversification

In order to explore those limits, we are going to modify our model introducing a new risk represented by a biased die. Let us assume that our policyholder plays in a casino, where it is a croupier that throws the die. If the croupier is a crooked one and cheats on the players, then our "fair" game would become unfair and the probability of paying 10 at each throw is no more going to be 1/6. Suppose that the die is biased and gives a higher probability to throw a 6. There are various ways of biasing the game.

4.1 A first approach with a biased probability to obtain a six – Case 1

A first approach, and the simplest one, is to have a die with a biased probability $q = (1+b)\frac{1}{6}$ of getting a 6 (with the bias $0 \le b \le 10\%$). Then we proceed as previously and can notice that it leads to an increase of the risk premium but with no dramatic change, as the CLT still applies, every throw being independent of each other; so the diversification effect will be taken in full.
The expectation of Z is $\boldsymbol{E}[Z] = Nnq$, and the variance of Z, denoted var_1, is $var_1(Z) = Nnq(1-q)$.
For one contract, the loss expectation is then

$$\frac{1}{N}\boldsymbol{E}[L] = \frac{l}{N}\,\boldsymbol{E}[Z] \;=\; lnq \tag{8}$$

and the loss variance

$$\frac{1}{N^2}\,var_1(L) \;=\; \frac{l^2}{N^2}\,var_1(Z) \;=\; l^2\,\frac{n}{N}\,q(1-q). \tag{9}$$

We see that for one contract the variance will decrease as the number of contracts increases and its theoretical risk is not much different than the one described above with the unbiased die.
Numerical application
Our numerical example uses $N = 1, 10, 20, 50, 100, 1'000$ and $10'000$ with a bias $b = 0.1\%, 1\%, 5\%$ and 10%; we still consider $n = 6$, $\alpha = 99\%$ and $\eta = 0.15$.

Table 3: The Risk loading per policy as a function of the bias of the biased die in the portfolio using VaR and TVaR risk measures, with $\alpha = 99\%$. The full diversification still applies as the bias is small

Risk measure ρ	Number N of Policies	$b = 0$ Fair Game	$b = 0.1\%$	$b = 1.0\%$	$b = 5.0\%$	$b = 10.0\%$
		Risk Loading R				
VaR						
	1	3.000	2.999	2.985	2.925	2.850
	5	1.500	1.499	1.485	1.425	1.650
	10	1.050	1.049	1.035	0.975	1.050
	50	0.450	0.479	0.465	0.465	0.480
	100	0.330	0.329	0.330	0.330	0.330
	1'000	0.102	0.102	0.102	0.103	0.105
	10'000	0.032	0.032	0.032	0.033	0.033
TVaR						
	1	3.226	3.227	3.216	4.562	4.501
	5	1.644	1.643	1.632	1.581	1.791
	10	1.164	1.163	1.151	1.236	1.171
	50	0.510	0.534	0.524	0.525	0.541
	100	0.372	0.371	0.373	0.374	0.375
	1'000	0.116	0.116	0.116	0.118	0.120
	10'000	0.037	0.037	0.037	0.037	0.038
$E(\mathbf{L})/\mathbf{N}$		*10.00*	*10.01*	*10.10*	*10.50*	*11.00*

As can be seen in Table 3, the effect of biasing the die does only slightly change the effect of diversification. It affects the risk loading making it even smaller with the VaR because of the change in expectation. We recall that the risk loading is computed from the difference between the risk measure and the expectation (see (6) and (7)). With TVaR, it hardly moves for the large number of policies. To see an effect on the risk loading, one needs to put up to 10% bias and the effect is an increase of only 3%, from 0.037 to 0.038 !

4.2 Another approach with a systematic risk

Let us design a different form of bias, with a much higher and interesting impact: a systematic risk. The croupier will make from time to time all or a big portion of the players loose *at the same time*. We include this effect in order to have a systematic risk (non-diversifiable) in our portfolio. We can model the behavior of the crooked croupier as if he would play with 2 dices, one unbiased (with loss probability $p = 1/6$) and another biased (with loss probability $q >> p$). Let U be the rv modelling the behavior of the croupier. The croupier will use the biased die with probability $\tilde{p}$ and the unbiased one with probability $1-\tilde{p}$. Hence the distribution of the rv U is Bernoulli $\mathscr{B}(\tilde{p})$ such that $U = 1$ when the chosen die is the biased one and $U = 0$ otherwise.
Let us introduce the rv X_{ij} (for $i = 1,\dots,n$ and $j = 1,\dots,N$) representing the number of losses of the jth player at the ith game.

4.2.1 A first way to use a biased die – Case 2

Assuming independence among all the X_{ij}'s ($i = 1,\dots,n,\ j = 1,\dots,N$) and U, we can interpret the total number of losses Z after n games by the N players, as the proportion of losses when playing with one die, plus the complementary proportion of losses when playing with the other die.
Its distribution can then be written, for $k \in \boldsymbol{N}$, as

$$\begin{aligned}
\boldsymbol{P}(Z=k) &= \boldsymbol{P}[\{\omega\in\Omega : Z(\omega)=k, U(\omega)=1\}] + P[\{\omega\in\Omega : Z(\omega)=k, U(\omega)=0\}] \\
&= \boldsymbol{P}[(Z=k) \mid (U=1)]\,\boldsymbol{P}(U=1) + \boldsymbol{P}[(Z=k) \mid (U=0)]\,\boldsymbol{P}(U=0) \\
&= \tilde{p}\ \boldsymbol{P}[(Z=k) \mid (U=1)]\ +\ (1-\tilde{p})\ \boldsymbol{P}[(Z=k) \mid (U=0)]
\end{aligned}$$

The conditional variables, $Z_q := Z|\ (U=1)$ and $Z_p := Z|\ (U=0)$, are distributed as $\mathscr{B}(Nn,q)$ and $\mathscr{B}(Nn,p)$ with mass probability distributions denoted by f_{Z_q} and f_{Z_p} respectively. The mass probability distribution f_Z of Z appears as a mixture of f_{Z_q} and f_{Z_p} (see e.g. [2])

$$f_Z = \tilde{p}\ f_{Z_q}\ +\ (1-\tilde{p})\ f_{Z_p} \quad \text{with } Z_q \sim \mathscr{B}(Nn,q) \text{ and } Z_p \sim \mathscr{B}(Nn,p) \tag{10}$$

Note that looking at the $(X_{ij}, i=1,\dots,n,\ j=1,\dots,N)$ as $(\tilde{X}_i, i=1,\dots,Nn)$, all $\tilde{X}_i,\ i=1,...,Nn$ are distributed as Bernoulli with probability p or q, depending on which die has been chosen. The total number of losses Z can be written as $Z = \sum_{i=1}^{n}\sum_{j=1}^{N} X_{ij} = \sum_{i=1}^{Nn} \tilde{X}_i$, with independent but not identically distributed $\tilde{X}_i$.
The expected loss amount for the portfolio is given by

$$\begin{aligned}
\boldsymbol{E}[L] = l \times \boldsymbol{E}(Z) &= l \times \Big(\tilde{p}\ \boldsymbol{E}(Z_q)\ +\ (1-\tilde{p})\ \boldsymbol{E}(Z_p)\Big) \\
&= Nnl\,\Big(\tilde{p}\ q\ +\ (1-\tilde{p})\ p\Big)
\end{aligned}$$

whereas for each policy, it is

$$\frac{l}{N}\boldsymbol{E}[L]\ =\ l\ n\,\Big(\tilde{p}\ q\ +\ (1-\tilde{p})\ p\Big) \tag{11}$$

from which we deduce the risk loading defined in (6) and (7).

Table 4: The Risk loading per policy in Case 2 as a function of the bias of the probability of using the biased die for the portfolio using VaR and TVaR measures with $\alpha = 99\%$. The biased die has a probability of giving a loss $\mathbf{q} = 50\%$. We see the effect of the non-diversifiable risk.

Risk measure ρ	Number N of Policies	Risk Loading R $\tilde{p} = 0$ Fair Game	$\tilde{p} = 0.1\%$	$\tilde{p} = 1.0\%$	$\tilde{p} = 5.0\%$	$\tilde{p} = 10.0\%$
VaR						
	1	3.000	2.997	4.469	4.346	5.693
	5	1.500	1.497	2.070	3.450	3.900
	10	1.050	1.047	1.770	3.300	3.450
	50	0.450	0.477	1.410	3.060	3.030
	100	0.330	0.327	1.605	3.000	2.940
	1'000	0.102	0.101	2.549	2.900	2.775
	10'000	0.032	0.029	2.837	2.866	2.724
TVaR						
	1	3.226	3.232	4.711	4.755	5.899
	5	1.644	1.707	2.956	3.823	4.146
	10	1.164	1.266	2.973	3.578	3.665
	50	0.510	0.760	2.970	3.196	3.141
	100	0.372	0.596	2.970	3.098	3.020
	1'000	0.116	0.396	2.970	2.931	2.802
	10'000	0.037	0.323	2.970	2.876	2.732
$\boldsymbol{E}(\mathbf{L})/\mathbf{N}$		*10.00*	*10.02*	*10.20*	*11.00*	*12.00*

Let us compute the variance of Z, named $var_2(Z)$. We can write (for more details, see the Appendix)

$$\boldsymbol{E}(Z^2) = Nn\left[\tilde{p}\; q\big(1-q+Nnq\big) \;+\; (1-\tilde{p})\; p\big(1-p+Nnp\big)\right]$$

which, combined with (11), provides

$$\begin{aligned} var_2(Z) &= \boldsymbol{E}(Z^2) - \boldsymbol{E}^2(Z) \\ &= Nn\left[q(1-q)\tilde{p} + p(1-p)(1-\tilde{p}) + Nn(q-p)^2\tilde{p}(1-\tilde{p})\right] \end{aligned}$$

from which we deduce the variance for the loss of one contract as $\frac{1}{N^2}\ var_2(L) = \frac{l^2}{N^2}\ var_2(Z)$, i.e.

$$\frac{1}{N^2}\ var_2(L) = \frac{l^2 n}{N}\Big(q(1-q)\tilde{p} + p(1-p)(1-\tilde{p})\Big)\ +\ l^2 n^2 (q-p)^2 \tilde{p}(1-\tilde{p}) \tag{12}$$

Notice that in the variance for one contract, the first term will decrease as the number of contracts increases, but not the second one. It does not depend on N and thus represents the non-diversified part of the risk.

In Table 4, we see that this new method has much more effect than in Case 1, in particular when the probability of having the biased die played is high. The interesting point is that from $\tilde{p} = 1\%$, the risk loading hardly changes when there is a large number of policies (starting at $N = 1000$) in the portfolio. This is true for both VaR and TVaR. The non-diversified term dominates the risk. The case $N = 10'000$ and $\tilde{p} = 0.1\%$ is interesting. There is a big difference between the risk loading of Var and TVaR. We notice that for 10'000 policies, the risk loading is multiplied by 10 in the case of TVaR and hardly moves in the case of VaR ! This effect is also seen for the same $\tilde{p}$ and less number of policies but to a lower extend.

4.2.2 Another way – Case 3

Now proceed as follows. For each game i, we choose one of the 2 dices, then the N policy holders play this game with this die. If Z_i denotes the number of losses for the N players at game i, then, by independence between the players, Z_i is distributed as a Binomial $\mathcal{B}(N, p_i)$ with $p_i = p$ or q, depending on the die that has be chosen for this game. There will be a certain number, say j, of games for which the number of losses of the N players are distributed as $\mathcal{B}(N, q)$. Therefore the total number of losses for those games, $\sum_{i=1}^{j} Z_i$, is distributed as $\mathcal{B}(jN, q)$, whereas for the $n-j$ remaining games, the corresponding total number of losses is distributed as $\mathcal{B}((n-j)N, p)$. We deduce that the distribution of the total number of losses Z for the n games can be expressed, for $k = 0, ..., nN$, as

$$\boldsymbol{P}(Z=k) \quad = \quad \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j\ (1-\tilde{p})^{n-j}\ \boldsymbol{P}\big[Z_q^{(j)} + Z_p^{(n-j)} = k\big] \tag{13}$$

$$\text{with} \qquad Z_q^{(j)} \sim \mathcal{B}(jN, q) \quad \text{and} \quad Z_p^{(j)} \sim \mathcal{B}((n-j)N, p) \tag{14}$$

from which we deduce (see the Appendix for details)

$$\boldsymbol{E}(Z) = Nn\big((q-p)\tilde{p} + p\big)$$

and, for one contract,

$$\frac{1}{N}\boldsymbol{E}(L) = \frac{l}{N}\boldsymbol{E}(Z) = nl\big(\tilde{p}\ q\ + (1-\tilde{p})\ p\big) \tag{15}$$

which is equal to the expectation (11) obtained with the previous method.

We can evaluate the variance of L by first computing $\boldsymbol{E}(Z^2)$ (see the Appendix for details). We have

$$\boldsymbol{E}(Z^2) = Nn\Big(q(1-q)\tilde{p} + p(1-p)(1-\tilde{p})\Big) +\ N^2 n^2\Big(p(1-\tilde{p}) + q\tilde{p}\Big)^2 + N^2 n (q-p)^2 \tilde{p}(1-\tilde{p})$$

from which we deduce the variance var_3 of Z as

$$\begin{aligned} var_3(Z) &= \boldsymbol{E}[Z^2] - \left(\boldsymbol{E}[Z]\right)^2 \\ &= Nn\left[q(1-q)\tilde{p} + p(1-p)(1-\tilde{p}) + N(q-p)^2\tilde{p}(1-\tilde{p})\right] \end{aligned}$$

which is different from the variance $var_2(Z)$ obtained with the previous method.

Table 5: The Risk loading per policy in Case 3 as a function of the bias of the probability of using the biased die for the portfolio using VaR and TVaR measures with $\alpha = 99\%$ and 10 million Monte-Carlo simulations. The biased die has a probability of giving a loss $q = 50\%$. We see the effect of the undiversifiable risk.

Risk measure ρ	Number N of Policies	Risk Loading R $\tilde{p} = 0$ Fair Game	$\tilde{p} = 0.1\%$	$\tilde{p} = 1.0\%$	$\tilde{p} = 5.0\%$	$\tilde{p} = 10.0\%$
VaR						
	1	3.000	2.997	2.969	4.350	4.200
	5	1.500	1.497	1.470	1.650	1.800
	10	1.050	1.047	1.170	1.350	1.500
	50	0.450	0.477	0.690	0.990	1.200
	100	0.330	0.357	0.615	0.945	1.170
	1'000	0.102	0.112	0.517	0.882	1.186
	10'000	0.032	0.033	0.485	0.860	1.196
	100'000	0.010	0.008	0.475	0.853	1.199
TVaR						
	1	3.226	3.232	4.485	4.515	4.448
	5	1.644	1.792	1.870	2.056	2.226
	10	1.164	1.252	1.342	1.604	1.804
	50	0.510	0.588	0.824	1.183	1.408
	100	0.375	0.473	0.740	1.118	1.358
	1'000	0.116	0.348	0.605	1.013	1.295
	10'000	0.037	0.313	0.563	0.981	1.276
	100'000	0.012	0.301	0.550	0.970	1.269
$\boldsymbol{E}(\mathbf{L})/\mathbf{N}$		*10.00*	*10.02*	*10.20*	*11.00*	*12.00*

Now for one contract we obtain:

$$\frac{1}{N^2} var_3(L) = \frac{l^2}{N^2} var_3(Z) = \frac{l^2 n}{N}\Big(q(1-q)\tilde{p} + p(1-p)(1-\tilde{p})\Big) \; + \; l^2 n \; (q-p)^2 \tilde{p}(1-\tilde{p}) \tag{16}$$

Notice that the last term is only multiplied by n and not n^2 as in the previous case. Nevertheless, as in Case 2, it will not decrease with the number of policies as it is also not divided by N. This last term is not diversified by the number of policies. It looks alike the one of (12), however its effect will be smaller than in var_2. With this method we have also achieved to produce a process with a non-diversified risk.

Let us revisit our numerical example. In this case, we cannot, contrary to Cases 1 and 2, directly use an expression of the distributions. We have to go through Monte-Carlo simulations. We have for each throw to choose if the crooked croupier picks the biased die or the fair die. Over 6 throws, the chances that a biased die is chosen is quasi 0, if the probability of such a choice is 0.1%. To get enough of these cases, we need to redo the operations many times, and then average over all the simulations. The results shown in Table 5 are obtained with 10 million simulations. We ran it also with 1 and 20 million simulations to check the convergence. It converges well as can be seen in Table 6.

The results shown in Table 5 follow what we expect. The diversification due to the total number of policies is more effective in this case than in case 2, but we still experience a part which is not diversifiable. The case $N = 10'000$ and $\tilde{p} = 0.1\%$ follows a similar behavior as in case 2 : the risk loading is multiplied by 10 in the case of TVaR while it is almost not different than in the unbiased case for VaR ! In this case, because we use Monte Carlo simulations, we have also computed the case with 100'000 policies. It is interesting to note that, as expected, the risk loading for the unbiased case continues to decrease. In the completely unbiased case it decreases by $\sqrt{10}$ as expected. However, except for $\tilde{p} = 0.1\%$ in the VaR case, it stops really decreasing for the biased cases. We have reached then the non-diversifiable part of the risk. The case $p = 0.1\%$ is interesting because we see here the limitation of the VaR as a risk measure. Although we know that there is a part of the risk that is non-diversifiable, VaR does not catch it really when $N = 10'000$ or $100'000$ while tVaR does not decrease significantly between $10'000$ and $100'000$ reflecting the fact that the risk cannot be completely diversified away.

Table 6: The Risk loading for $N = 100$ in Case 3 as a function of the number of Monte Carlos simulations and the bias of the probability of using the biased die for the portfolio using VaR and TVaR measures with $\alpha = 99\%$. The biased die has a probability of giving a loss $q = 50\%$. We see that the number of simulations has very little impact.

Risk measure ρ	Number N of Policies	Risk Loading R $\tilde{p} = 0$ Fair Game	$\tilde{p} = 0.1\%$	$\tilde{p} = 1.0\%$	$\tilde{p} = 5.0\%$	$\tilde{p} = 10.0\%$
VaR						
	1 million	0.330	0.357	0.615	0.945	1.170
	10 million	0.330	0.357	0.615	0.945	1.155
	20 million	0.330	0.357	0.615	0.945	1.170
TVaR						
	1 million	0.375	0.476	0.738	1.115	1.358
	10 million	0.374	0.472	0.739	1.117	1.357
	20 million	0.375	0.473	0.740	1.118	1.358
$E(\mathbf{L})/\mathbf{N}$		*10.00*	*10.02*	*10.20*	*11.00*	*12.00*

To explore the convergence of the simulations, we present in Table 6 the results obtained for $N = 100$ and for various number of simulations. For this number of policies, the number of simulation has no influence. Obviously, with a lower number of policies, the number of simulations plays a more important role as one would expect, while for a higher number of policies, it is insensitive to the number of simulations above 1 million.

4.2.3 Discussion – comparison of the methods

In Table 7, we present a summary of the expectation and the variance obtained in our three cases. In the first case, we see that the variance will decrease with increasing N, while both other cases contain a term in the variance that does not depend on N. Those two cases are the ones containing a systemic risk that is not diversified. Note that $var_2(Z)$ contains a non-diversified part which corresponds to n times the non-diversified part of $var_3(Z)$; we have

$$var_2(Z) - var_3(Z) = N^2 n(n-1)\ (q-p)^2 \tilde{p}(1-\tilde{p})$$

thus $var_2(Z)$ will be always larger than $var_3(Z)$ (for $n > 1$). This is consistent with the numerical results in Tables 4 and 5, where we see that the effect of increasing the number of policies is more important in the later case.
Case 3 is the most interesting because it shows both the effect of diversification and the effect of the non-diversifiable term in a more obvious way. This example is thus more suitable to be used to explore other properties.

Table 7: Summary of the analytical results (expectation and variance per policy) for the 3 cases of biased games presented in this study

	Expectation $\frac{1}{N}\boldsymbol{E}(L)$	Variance $\frac{1}{N^2}var(L)$
case 1	$ln\ q$	$l^2 n\ \frac{1}{N}\ q(1-q)$
case 2	$ln\left(\tilde{p}\ q\ +\ (1-\tilde{p})\ p\right)$	$\frac{l^2 n}{N}\left(q(1-q)\tilde{p}+p(1-p)(1-\tilde{p})\right)+l^2n^2(q-p)^2\tilde{p}(1-\tilde{p})$
case 3	$ln\left(\tilde{p}\ q\ +\ (1-\tilde{p})\ p\right)$	$\frac{l^2 n}{N}\left(q(1-q)\tilde{p}+p(1-p)(1-\tilde{p})\right)+l^2n\ (q-p)^2\tilde{p}(1-\tilde{p})$

5 Conclusion

In this study, we have shown the effect of diversification on the pricing of insurance risk through a simple modeling. Illustrative examples have been constructed introducing in different ways a component of the risk that cannot be diversified. To this aim, we have used a biased die in the game and studied simple models, well suited for the understanding of the problem. Those models allow for a straightforward analytical evaluation of the impact of the non-diversified part.
In real life, insurers have to pay special attention to the effects that can weaken the diversification benefits. For instance, in the case of motor insurance, the appearance of a hail storm will introduce a "bias" in the usual risk of accident due to a cause independent of the car drivers, which will hit a big number of cars at the same time and thus cannot be diversified among the various policies. There are other examples in life insurance for instance with pandemic or mortality trend that would affect the entire portfolio and cannot be diversified away. Special care must be given to those risks as they will affect greatly the risk loading of the premium as can be seen in our examples.
These examples might also find applications for real cases. In particular, Case 3 may be used in conjunction with risk appetite. It will be the subject of a following paper.

Acknowledgement We would like to thank Philip Keller who first came up with the idea and the model of the crooked croupier.

References

[1] M.M. DACOROGNA, C. HUMMEL, Alea jacta est, an illustrative example of pricing risk. *SCOR Technical Newsletter*, SCOR Global P&C (2008).

[2] B.S. EVERITT, D.J. HAND, *Finite Mixture Distributions*. London Chapman & Hall (1981).

[3] A. MCNEIL, R. FREY, P. EMBRECHTS, *Quantitative Risk Management*. Princeton (2005).

Appendix

For completeness, we are going to give the details of the straightforward computation of the first two moments of Z, in Case 2 and Case 3.

A Case 2 – computation of the variance of Z

In case 2, the computation of the expectation is given in the text. We concentrate here on the computation of the variance of Z, named $var_2(Z)$. For this we first compute $\boldsymbol{E}(Z^2)$

$$
\begin{aligned}
\boldsymbol{E}(Z^2) &= \sum_{k=0}^{Nn} k^2 \boldsymbol{P}(Z=k) \\
&= \tilde{p} \sum_{k=0}^{Nn} k^2 \boldsymbol{P}(Z_q = k) \; + \; (1-\tilde{p}) \sum_{k=0}^{Nn} k^2 \boldsymbol{P}(Z_p = k) \\
&= \tilde{p}\Big(var(Z_q) + \boldsymbol{E}^2(Z_q)\Big) \; + \; (1-\tilde{p})\Big(var(Z_p) + \boldsymbol{E}^2(Z_p)\Big) \\
&= Nn\big[\tilde{p}\; q\big(1-q+Nnq\big) \; + \; (1-\tilde{p})\; p\big(1-p+Nnp\big)\big]
\end{aligned}
$$

and then we can deduce the variance from this result, using also (11):

$$
\begin{aligned}
var_2(Z) &= \boldsymbol{E}(Z^2) - \boldsymbol{E}^2(Z) \\
&= Nn\big[\tilde{p}\; q\big(1-q+Nnq\big) \; + \; (1-\tilde{p})\; p\big(1-p+Nnp\big)\big] - N^2 n^2 \Big(\tilde{p}\; q \; + \; (1-\tilde{p})\; p\Big)^2 \\
&= \tilde{p}Nnq\big(1-q+Nnq-\tilde{p}Nnq\big) \; + \\
&\quad (1-\tilde{p})\; Nnp\big(1-p+Nnp-Nn(1-\tilde{p})p\big) - 2N^2 n^2\; \tilde{p}(1-\tilde{p})\; pq \\
&= Nn\big[q(1-q)\tilde{p} + p(1-p)(1-\tilde{p}) + Nn(q-p)^2 \tilde{p}(1-\tilde{p})\big]
\end{aligned}
$$

B Case 3 – Computation of the first two moments of Z

B.1 Computation of the expectation of Z

Using (13), we can write

$$
\begin{aligned}
\boldsymbol{E}(Z) &= \sum_{i=1}^{n} k \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \; (1-\tilde{p})^{n-j} \; \boldsymbol{P}\big[Z_q^{(j)} + Z_p^{(n-j)} = k\big] \\
&= \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \; (1-\tilde{p})^{n-j} \sum_{i=1}^{n} k \; \boldsymbol{P}\big[Z_q^{(j)} + Z_p^{(n-j)} = k\big] \\
&= \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \; (1-\tilde{p})^{n-j} \; \boldsymbol{E}\big[Z_q^{(j)} + Z_p^{(n-j)}\big] \\
&= \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \; (1-\tilde{p})^{n-j} \; \Big(\boldsymbol{E}\big[Z_q^{(j)}\big] + \boldsymbol{E}\big[Z_p^{(n-j)}\big]\Big)
\end{aligned}
$$

$$
\begin{aligned}
&= N \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \ (1-\tilde{p})^{n-j} \left(j(q-p)+np\right) \\
&= N \left((q-p) \sum_{j=0}^{n} \binom{n}{j} j \ \tilde{p}^j \ (1-\tilde{p})^{n-j} \ + \ np \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \ (1-\tilde{p})^{n-j} \right) \\
&= N\left((q-p)n\boldsymbol{E}(U)+np\right) \ = \ N\left((q-p)n\tilde{p}+np\right)
\end{aligned}
$$

from which we obtain, for one contract,

$$
\frac{1}{N}\boldsymbol{E}(L) = \frac{l}{N}\boldsymbol{E}(Z) = nl\left(\tilde{p}(q-p)+p\right) = nN\left(\tilde{p}\ q\ +(1-\tilde{p})\ p\right) \tag{17}
$$

B.2 Computation of the variance of Z

Using (13), then the independence of $Z_q^{(j)}$ and $Z_p^{(n-j)}$ in the third equation, we have

$$
\begin{aligned}
\boldsymbol{E}(Z^2) \ &= \ \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \ (1-\tilde{p})^{n-j} \left(\sum_{i=1}^{n} k^2 \ \boldsymbol{P}[Z_q^{(j)}+Z_p^{(n-j)}=k] - \left(\sum_{i=1}^{n} k \ \boldsymbol{P}[Z_q^{(j)}+Z_p^{(n-j)}=k] \right)^2 \right) \\
&\qquad + \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \ (1-\tilde{p})^{n-j} \left(\sum_{i=1}^{n} k \ \boldsymbol{P}[Z_q^{(j)}+Z_p^{(n-j)}=k] \right)^2 \Big) \\
&= \ \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \ (1-\tilde{p})^{n-j} var\left(Z_q^{(j)}+Z_p^{(n-j)}\right) + \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \ (1-\tilde{p})^{n-j} \left(\boldsymbol{E}[Z_q^{(j)}]+\boldsymbol{E}[Z_p^{(n-j)}]\right)^2 \\
&= \ \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \ (1-\tilde{p})^{n-j} \left(jNq(1-q)+(n-j)Np(1-p)\right) \\
&\qquad + \sum_{j=0}^{n} \binom{n}{j} \tilde{p}^j \ (1-\tilde{p})^{n-j} \left(jNq+(n-j)Np\right)^2
\end{aligned}
$$

$$
\begin{aligned}
\text{So}\quad \boldsymbol{E}(Z^2) &= N\sum_{j=0}^{n}\binom{n}{j}\tilde{p}^j\ (1-\tilde{p})^{n-j}\Big(j[q(1-q)-p(1-p)]+np(1-p)\Big)\\
&\quad +N^2\sum_{j=0}^{n}\binom{n}{j}\tilde{p}^j\ (1-\tilde{p})^{n-j}\Big(j(q-p)+np\Big)^2\\
&= Nn\Big(q(1-q)\tilde{p}+p(1-p)(1-\tilde{p})\Big)+\ N^2n^2p^2+2N^2n^2(q-p)p\tilde{p}\\
&\quad +N^2(q-p)^2\sum_{j=0}^{n}\binom{n}{j}\tilde{p}^j\ (1-\tilde{p})^{n-j}j^2\\
&= Nn\Big(q(1-q)\tilde{p}+p(1-p)(1-\tilde{p})\Big)+\ N^2n^2p^2+2N^2n^2(q-p)p\tilde{p}\\
&\quad +N^2(q-p)^2\boldsymbol{E}\Big[\big(\sum_{i=1}^{n}U_i\big)^2\Big]\\
&= Nn\Big(q(1-q)\tilde{p}+p(1-p)(1-\tilde{p})\Big)+\ N^2n^2p^2+2N^2n^2(q-p)p\tilde{p}\\
&\quad +N^2(q-p)^2\Big(var\big(\sum_{i=1}^{n}U_i\big)+\Big(\boldsymbol{E}\big[\sum_{i=1}^{n}U_i\big]\Big)^2\Big)
\end{aligned}
$$

where $(U_i,\ i=1,...,n)$, denotes the n-sample with parent rv U. Then,

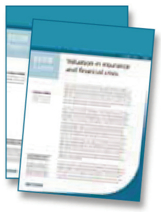
$$
\begin{aligned}
\boldsymbol{E}(Z^2) &= Nn\Big(q(1-q)\tilde{p}+p(1-p)(1-\tilde{p})\Big)+\ N^2n^2p^2\\
&\quad +2N^2n^2(q-p)p\tilde{p}+N^2(q-p)^2\Big(n\tilde{p}(1-\tilde{p})+n^2\tilde{p}^2\Big)\\
&= Nn\Big(q(1-q)\tilde{p}+p(1-p)(1-\tilde{p})\Big)+\ N^2n^2\Big(p(1-\tilde{p})+q\tilde{p}\Big)^2+N^2n(q-p)^2\tilde{p}(1-\tilde{p})
\end{aligned}
$$

from which we deduce the variance $var_3(Z)$ of Z

$$
\begin{aligned}
var_3(Z) &= \boldsymbol{E}[Z^2]-\big(\boldsymbol{E}[Z]\big)^2\\
&= Nn\Big(q(1-q)\tilde{p}+p(1-p)(1-\tilde{p})\Big)+\ N^2n^2\Big(p(1-\tilde{p})+q\tilde{p}\Big)^2+N^2n(q-p)^2\tilde{p}(1-\tilde{p})\\
&\quad -N^2n^2\big(\tilde{p}\ q\ +(1-\tilde{p})\ p\big)^2\\
&= Nn\left[q(1-q)\tilde{p}+p(1-p)(1-\tilde{p})+N(q-p)^2\tilde{p}(1-\tilde{p})\right]
\end{aligned}
$$

which is different from the variance obtained in Case 2.

Former SCOR Papers



SCOR Paper N°5 - July 2009
Securitization, Insurance and Reinsurance

SCOR Paper N°6 - January 2010
Adapting the solvency regulations to times of crisis, accepting the riskiness of the situation

SCOR Paper N°7 - January 2010
The Influence of Risk Measures and Tail Dependencies on Capital Allocation

SCOR Paper N°8 - March 2010
Principle-based Solvency: A Comparison between Solvency II and the Swiss Solvency Test

SCOR Paper N°9 - December 2010
Aging: a Global Phenomenon

SCOR Paper N°10 - December 2010
PrObEx: a New Method for the Calibration of Copula Parameters from Prior Information, Observations and Expert Opinions

SCOR Paper N°11 - December 2010
Development Risk

SCOR Paper N°12 - December 2010
Study of the impact of inflation and GDP growth on property-liability and life insurance premiums over the last 30 years: case of the G7 countries

SCOR Paper N°13 - March 2011
Preparing for Solvency II: Points of debate in the Standard Formula

SCOR Paper N°14 - April 2011
Modelling operational risk in the insurance industry

SCOR Paper N°15 - May 2011
Why do the French not purchase more long-term care cover?

SCOR Paper N°16 - Juin 2011
A new method for modeling dependence via extended common shock type model

SCOR Paper N°17 - October 2011
EU regulation of greenhouse gas emissions: what solutions can insurance companies offer industry?

SCOR Paper N°18 - March 2012
Microscopic longMarket Value Margin: Practical calculations under the Solvency II Cost of Capital approache-vity modeling

SCOR Paper N°19 - May 2012
Microscopic longevity modeling and its practical applications

SCOR Paper N°20 - June 2012
How Long Will We Live? A Demographer's Reflexions on Longevity

SCOR Paper N°21 - February 2013
Microscopic longevity modeling and its practical applications

SCOR Paper N°22 - April 2013
Non-Life Insurance Market Growth in China: can we predict it?

SCOR
Global Risk Center

SCOR Papers, edited by SCOR, are one of the tool supporting the SCOR Global Risk Center.
The SCOR Global Risk Center gathers and analyses the most interesting resources about risks. It operates as a dual resource center, based both on data and resources produced by SCOR itself, and on all other resources available selected specifically by SCOR. Publications will be available in English, French and/or German.
SCOR Global Risk Center is available at www.scorglobalriskcenter.com or on SCOR's website – www.scor.com.