

*A Note on the Myers and Read Capital  
Allocation Formula*

Stephen J. Mildenhall, FCAS, ASA, MAAA

# A Note on the Myers and Read Capital Allocation Formula

Stephen J. Mildenhall

July 8, 2003

## **Abstract**

The Myers and Read capital allocation formula is an important new actuarial result. In this paper, we give an overview of the Myers and Read result, explain its significance to actuaries, and provide a simple proof. Then we explain the assumption the allocation formula makes on the underlying families of loss distributions as expected losses by-line vary. We show that this assumption does not hold when insurers grow by writing more risks from a discrete group of insureds—as is typically the case.

Next, we discuss whether the inhomogeneity in a realistic portfolio of property casualty risks is material. We show how to decompose the relevant partial derivatives into homogeneous and inhomogeneous parts and examine the behaviour of each. We then apply the theory to some realistic examples. These clearly show that the lack of homogeneity is material. This failure will severely limit the practical application of the Myers and Read allocation formula.

## 1 Introduction

In an important paper for actuaries, Myers and Read (2001) showed how to allocate the expected policy holder deficit in a multi-line insurance company uniquely to each line. Their work can also be used to allocate surplus to each line. Previous work on the allocation problem, including Phillips et al. (1998) and Merton and Perold (2001), had concluded that such an allocation could be inappropriate and misleading. The Myers and Read result is, therefore, potentially a significant breakthrough, with obvious importance to actuaries.

Myers and Read repeatedly stress their result is independent of the distribution of losses by line and of any correlations between lines that may exist. They say their “proof requires no assumptions about the joint probability distributions of line-by-line losses and returns on the firm’s portfolio of assets.” However, while their result makes no assumptions about the *static* distribution of losses with fixed expected loss by line, their derivation does make an important assumption about how the *dynamic* distribution of losses changes shape with changing expected losses by line. This paper will explain the significance of the latter assumption. We will show it is a necessary and sufficient condition for the Myers and Read result to hold. Most importantly, we will show that the assumption does not hold when insurers grow through the assumption of risk from discrete insureds—as is typically the case.

For the convenience of readers not familiar with Myers and Read’s work, we begin with an overview. Consider a simple insurance company which writes two lines of business. The losses from each line are represented by a random variables

$X_1$  and  $X_2$ , with means  $x_1$  and  $x_2$ . Since the company can choose to write more or less of each line, we assume that the families  $X_1(x_1)$  and  $X_2(x_2)$ , with varying means  $x_1$  and  $x_2$ , are specified. For example, losses from line 1 may be normally distributed with mean  $x_1$  and standard deviation 1000 and for line 2 be normally distributed with mean  $x_2$  and coefficient of variation  $\nu$ . Assume the company has capital  $k$  and total assets  $x_1 + x_2 + k$ . Also assume that interest rates are zero. (Myers and Read show how to convert from deterministic investment income to stochastic income. We focus on deterministic income and set it equal to zero for simplicity. Nothing of substance is lost in doing so.) Let

$$I(x_1, x_2, k) = \Pr(X_1 + X_2 > x_1 + x_2 + k)$$

be the probability of insolvency. Finally, assume that the company holds its probability of insolvency constant, by adjusting writings of each line and the amount of capital held. Let  $K(x_1, x_2)$  satisfy

$$I(x_1, x_2, K(x_1, x_2)) = \text{constant}.$$

Then, under certain assumptions on the families  $X_1(x_1)$  and  $X_2(x_2)$  for varying  $x_1, x_2$ , but under *no assumptions on the distributions of losses given fixed  $x_1$  and  $x_2$*  we can prove

$$x_1 \frac{\partial K}{\partial x_1} + x_2 \frac{\partial K}{\partial x_2} = K. \quad (1)$$

This is obviously a very useful result: it tells the company that it should allocate capital at the rate  $\partial K/\partial x_1$  to line 1 and  $\partial K/\partial x_2$  to line 2, and that if it does so the total capital allocation will add up to actual capital! We prove Equation (1) in

Corollary 2, below. It is very similar to the actual Myers and Read result, which we prove in Corollary 1.

The main result of the paper, Proposition 1, states the assumptions on the families  $X_i(x_i)$  required for Equation (1) to hold. We show that in most real-world situations these assumptions will, unfortunately, fail to hold. We also give a straight-forward proof of the Myers and Read “adds-up” result and we prove two related extensions. Finally we give several examples to illustrate the results.

The necessary distributional assumption highlights the difference between a continuous “representative insurer” approach, where each insurer assumes a share of a total market risk, and a discrete approach, where insurers assume risk from distinct and discrete individual insureds. The Myers and Read result requires a continuous view as we show in Proposition 1. Examples 4.4 and 4.5 show the result is not true in a discrete environment. Butsic (1999) used the representative insurer argument in his application of Myers and Read.

The rest of the paper is laid out as follows. In the next section we prove two technical lemmas. Section 3 states and proves the main Proposition. Section 4 gives several examples using the main result. Section 5 examines how the Myers and Read formula fails when losses are inhomogeneous and shows that in realistic examples the failure will be material.

## 2 Two Technical Lemmas

**Lemma 1** *Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a differentiable function of  $n$  variables. Then*

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + \cdots + x_n \frac{\partial f}{\partial x_n} = 0 \quad (2)$$

*if and only if  $f$  is constant along rays from the origin.*

Note: If  $f$  is constant on lines through the origin then  $f$  is called *homogeneous*. The lemma only requires  $f$  be constant along rays from the origin; along a line  $f$  can change as the line passes through the origin. The function  $x \mapsto x/|x|$  is a good example of what can occur: it changes value from  $+1$  to  $-1$  at zero. If  $f$  is constant along rays from the origin, then in half spaces through the origin  $f$  can be expressed as a function of  $x_i/x_j$ ,  $i = 1, \dots, n$  when  $x_j \neq 0$ , for each  $j$ . In our applications of this lemma, the domain of  $f$  will be the positive quadrant, so there is no difference between lines through the origin and rays from the origin in the domain. I would like to thank Christopher Monsour for pointing this out.

**Proof** Sufficiency: if  $f$  is constant along rays through the origin, then by the note we can assume locally that  $f(x_1, \dots, x_n) = \tilde{f}(x_1/x_n, \dots, x_{n-1}/x_n)$  for some function  $\tilde{f}$  of  $n - 1$  variables. An easy calculation shows

$$\begin{aligned} x_1 \frac{\partial f}{\partial x_1} + \dots + x_n \frac{\partial f}{\partial x_n} &= \frac{x_1}{x_n} \tilde{f}_1 + \dots + \frac{x_{n-1}}{x_n} \tilde{f}_{n-1} - x_n \left( \frac{x_1}{x_n^2} \tilde{f}_1 + \dots + \frac{x_{n-1}}{x_n^2} \tilde{f}_{n-1} \right) \\ &= 0, \end{aligned}$$

where  $\tilde{f}_i = \partial \tilde{f}(x_1, \dots, x_{n-1}) / \partial x_i$ .

Necessity: Let  $\mathbf{v} = (x_1, \dots, x_n)$  be a differentiable curve, so  $\mathbf{v} = \mathbf{v}(t) : \mathbb{R} \rightarrow \mathbb{R}^n$ , with  $d\mathbf{v}/dt = \mathbf{v}$ . This means  $\mathbf{v}$  is equal to its own tangent vector for each  $t$ . By separating variables it is easy to see that  $\mathbf{v}$  is a line through the origin. (It has the form  $e^t(k_1, \dots, k_n)$  for constants of integration  $k_i$ .) Then, by the chain-rule

$$\begin{aligned} \frac{d}{dt} f(\mathbf{v}(t)) &= x_1 \frac{\partial f}{\partial x_1} + \dots + x_n \frac{\partial f}{\partial x_n} \\ &= 0, \end{aligned}$$

by assumption, so the directional derivative of  $f$  along each half of any such line  $v$  is constant, i.e.  $f$  is constant along rays from the origin, as required. Since  $v$  never reaches the origin, we cannot assert that  $f$  is constant along lines through the origin.  $\square$

**Lemma 2** Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a differentiable function of  $n$  variables. Then,

$$x_1 \frac{\partial f}{\partial x_1} + \cdots + x_n \frac{\partial f}{\partial x_n} = f \quad (3)$$

on a half-space where  $x_1 > 0$  (resp.  $x_1 < 0$ ) if and only if there exists a differentiable function  $\tilde{f}$  so that  $f(x_1, \dots, x_n) = x_1 \tilde{f}(x_2/x_1, \dots, x_n/x_1)$  on that half space, and similarly for  $x_2, \dots, x_n$ .

**Proof** If  $f(x_1, \dots, x_n) = x_1 \tilde{f}(x_2/x_1, \dots, x_n/x_1)$  then, using subscripts on  $\tilde{f}$  to denote partial derivatives,

$$\begin{aligned} x_1 \frac{\partial f}{\partial x_1} + \cdots + x_n \frac{\partial f}{\partial x_n} &= \left( x_1 \tilde{f} - \sum_{j=2}^n x_j \tilde{f}_{j-1} \right) + \sum_{j=2}^n x_j \tilde{f}_{j-1} \\ &= f. \end{aligned}$$

The first sum comes from the partial derivative with respect to  $x_1$  and the second sum comes from all the remaining partials.

On the other hand, suppose  $f$  satisfies Equation (3) and let  $\tilde{f}(t, s_2, \dots, s_n) = f(t, s_2 t, \dots, s_n t)/t$  where  $t > 0$  (resp.  $t < 0$ ). We must show  $\tilde{f}$  is independent of  $t$ . Differentiating

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{f(t, s_2 t, \dots, s_n t)}{t} \right) &= -\frac{1}{t^2} f + \frac{1}{t} \left( \frac{\partial f}{\partial x_1} + \sum_{j=2}^n s_j \frac{\partial f}{\partial x_j} \right) \\ &= 0 \end{aligned}$$

and the result follows.  $\square$

### 3 Statement and Proof of Main Result

Before stating the proposition we need to define some more notation. We are modeling a multi-line insurance company. Losses from each line are modeled by random variables  $X_i$ ,  $i = 1, \dots, n$ , where  $X_i$  has mean  $x_i$  and distribution function  $F_i$ . We often regard  $x_i$  as a variable (but not a random variable), so each  $X_i$  is really a family of distributions indexed by  $x_i$ . Where necessary we emphasize this by writing  $X_i(x_i)$ . Changes in  $x_i$  correspond to increasing or decreasing volume in line  $i$ , since  $x_i$  is the *a priori* expected loss.

Assume that the company holds total assets equal to  $x_1 + \dots + x_n + k$ , so in a very simplistic sense,  $k$  is the capital or surplus of the company.

Next, define the probability of insolvency function and the expected policyholder deficit function for a single line  $i$  as

$$I_i(x_i, k) = \Pr(X_i > x_i + k) = 1 - F_i(x_i + k) \quad (4)$$

and

$$D_i(x_i, k) = \int_{x_i+k}^{\infty} t - (x_i + k) dF_i(t). \quad (5)$$

In both of these equations  $x_i$  is performing double duty: it is the mean of  $X_i$  and in  $x_i + k$  it determines where  $F_i$  is evaluated. To emphasize this we could write

$$I_i(x_i, k) = 1 - F_i(x_i + k; x_i). \quad (6)$$

Finally, let  $X = X_1 + \dots + X_n$  be the total losses with distribution function  $F$ . Define insolvency and deficit functions for the whole company as

$$I(x_1, \dots, x_n, k) = \Pr\left(\sum X_i > \sum x_i + k\right) = 1 - F(x_1 + \dots + x_n + k) \quad (7)$$



and

$$D(x_1, \dots, x_n, k) = \int_{\sum t_i > \sum x_i + k} \dots \int t_1 + \dots + t_n - (x_1 + \dots + x_n + k) dF(t_1, \dots, t_n). \quad (8)$$

The following definition is key:

**Definition 1** *A family of random variables  $X(x)$  with  $E(X(x)) \propto x$  is called homogeneous if there exists a single random variable  $U$  so that  $X(x)/x$  has the same distribution as  $U$  for all  $x$ .*

Homogeneity is Myers and Read's only distributional assumption, and it means that losses come from a representative insurer. The requirement that  $U$  is independent of  $x$  is important—after all, any random variable can be written as  $X = E(X)(X/E(X))$ ! An exponential variable  $X$  with mean  $x$  is a homogeneous family, since  $X = xU$  where  $U$  has an exponential distribution with mean 1. However, a normal variable with mean  $x$  and standard deviation 1 is not homogeneous.

In order to compute expressions like  $\partial I / \partial x$  we need to know how the family  $X(x)$  changes shape with changes in  $x$ . We need to work with  $X(x + \epsilon)$  as well as  $X(x)$  because

$$\begin{aligned} \frac{\partial I}{\partial x} &= -\frac{d}{dx} F(x+k; x) \\ &= -\lim_{\epsilon \rightarrow 0} \frac{F(x+k+\epsilon; x) - F(x+k; x)}{\epsilon} \\ &\quad - \lim_{\epsilon \rightarrow 0} \frac{F(x+k; x+\epsilon) - F(x+k; x)}{\epsilon}. \end{aligned}$$

The partial derivative has a *static* part, where the mean of the underlying variable does not change, and a *dynamic* part, where the point of evaluation is fixed but the mean changes. This shows computing partial derivatives such as  $\partial I/\partial x$  is inextricably linked to families of random variables.

With this notation we can now state our main result.

**Proposition 1** *The following are equivalent.*

1. For each  $i = 1, \dots, n$ ,  $X_i(x_i)$  is a homogeneous family of random variables.

2. For each  $i = 1, \dots, n$

$$x_i \frac{\partial I_i}{\partial x_i} + k \frac{\partial I_i}{\partial k} = 0. \quad (9)$$

3. For each  $i = 1, \dots, n$

$$x_i \frac{\partial D_i}{\partial x_i} + k \frac{\partial D_i}{\partial k} = D. \quad (10)$$

4. We have equality

$$x_1 \frac{\partial I}{\partial x_1} + \dots + x_n \frac{\partial I}{\partial x_n} + k \frac{\partial I}{\partial k} = 0. \quad (11)$$

5. We have equality

$$x_1 \frac{\partial D}{\partial x_1} + \dots + x_n \frac{\partial D}{\partial x_n} + k \frac{\partial D}{\partial k} = D. \quad (12)$$

The proposition says that each of the five statements holds if and only if all the other four hold. Put another way, if one of the five fails to hold then the other four will also fail. This means that we can construct simple one line examples and can use items 2 and 3 generalize to the multi-line case. This simplifies the mathematics of the examples.

**Proof** We shall prove (4) implies (2) implies (1) implies (4), and then (5) implies (3) implies (1) implies (5), which is enough to show all the statements are equivalent.

(4) implies (2): Set  $x_j = 0$  for  $j \neq i$  in Equation (11) to get Equation (9). This can also be seen geometrically using Lemma 1 which says  $I$  is constant along rays from the origin. Therefore  $I_i$ , which is a restriction of  $I$ , is also constant along such rays.

(2) implies (1): Lemma 1 applied to  $I_i$  shows there exists a function  $\tilde{I}_i$  so that

$$I_i(x_i, k) = \tilde{I}_i(k/x_i).$$

Let  $U_i = X_i/x_i$ , then  $\Pr(U_i > u) = \tilde{I}_i(u - 1)$  is independent of  $x_i$  as required.

(1) implies (4): Assumption (1) implies that  $I$  is constant along rays from the origin, so the result follows from Lemma 1.

(5) implies (3): Set  $x_j = 0$  for  $j \neq i$  in Equation (12) to get Equation (10).

(3) implies (1): Let  $U_i = X_i/x_i$ . We have to show  $\Pr(U_i > u)$  is independent of  $x_i$ . Let  $x^+ = \max(x, 0)$ . Then, notice that

$$\frac{\partial D}{\partial k} = \frac{\partial}{\partial k} E[(\sum x_i U_i - (\sum x_i + k))^+] \quad (13)$$

$$= E[\frac{\partial}{\partial k} (\sum x_i U_i - (\sum x_i + k))^+] \quad (14)$$

$$= E[-\mathbf{1}_{\{\sum x_i U_i > \sum x_i + k\}}] \quad (15)$$

$$= -\Pr(\sum x_i U_i > \sum x_i + k) \quad (16)$$

is minus the probability of default. Next, use Lemma 2 to define  $\tilde{D}_i$  so that  $D_i(x_i, k) = x_i \tilde{D}_i(k/x_i)$ . Therefore

$$\frac{\partial D_i}{\partial k} = \tilde{D}'_i(k/x_i)$$

and so

$$\Pr(U_i > u) = -\tilde{D}'_i(u - 1)$$

is independent of  $x_i$  as required.

(1) implies (5): Assumption (1) shows we can write  $D$  as

$$D(x_1, \dots, x_n, k) = k \tilde{D}(x_1/k, \dots, x_n/k)$$

so the result follows from Lemma 2.  $\square$

The results in Proposition 1 are clearly similar to Myers and Read's results but they are not exactly the same. We shall now explain how to derive their exact result and prove some other similar results. For simplicity we shall assume  $n = 2$  and work with just  $x_1$  and  $x_2$  in the rest of the paper.

Myers and Read's "adds-up" result (their Equation A1-3) involves computing the marginal increase in surplus required to hold the default value constant, given a marginal increase in a particular line. We have been taking a slightly different approach: if we hold the surplus and default value constant, what decrease is needed in line 2 to offset an increase in line 1? However, it is easy to reconcile the two approaches. To do this, let  $\kappa_1$  and  $\kappa_2$  be the marginal surplus requirements for each line. Note that  $\kappa_1$  and  $\kappa_2$  are ratios whereas  $k$  is a dollar amount. Myers and

Read then use a capital amount  $k = \kappa_1 x_1 + \kappa_2 x_2$  and define the default value  $D_M$  (to distinguish from our  $D$ ) as

$$D_M(x_1, x_2) := D(x_1, x_2, \kappa_1 x_1 + \kappa_2 x_2). \quad (17)$$

Myers and Read use the following notation in their Appendix 1. They write  $\tilde{L}_a = L_a \tilde{R}_a$ , where  $L_a$  corresponds to our  $x_1$ ,  $\tilde{R}_a$  to  $U_1$  and  $\tilde{L}_a$  to  $X_1$ . Thus  $\tilde{L}_a = L_a \tilde{R}_a$  translates into our  $X_1 = x_1 U_1$ , i.e. the homogeneity assumption. The value  $L_a$  is the expected value of  $\tilde{L}_a$  at time 0. We are ignoring the time value of money here by assuming an interest rate of zero. Myers and Read also work with a fixed interest rate and then integrate over all possible rates—an extra level of sophistication that need not concern us.

We can now prove their result.

**Corollary 1** (*Myers and Read*) *Assume losses  $X_i$  form a homogeneous family for each  $i$ . Then default values “add-up” in that*

$$x_1 \frac{\partial D_M}{\partial x_1} + x_2 \frac{\partial D_M}{\partial x_2} = D_M. \quad (18)$$

**Proof** Computing using the chain-rule and then applying Proposition 1 item 5 in Equation (21) gives:

$$x_1 \frac{\partial D_M}{\partial x_1} + x_2 \frac{\partial D_M}{\partial x_2} = x_1 \left( \frac{\partial D}{\partial x_1} + \kappa_1 \frac{\partial D}{\partial k} \right) + x_2 \left( \frac{\partial D}{\partial x_2} + \kappa_2 \frac{\partial D}{\partial k} \right) \quad (19)$$

$$= x_1 \frac{\partial D}{\partial x_1} + x_2 \frac{\partial D}{\partial x_2} + (\kappa_1 x_1 + \kappa_2 x_2) \frac{\partial D}{\partial k} \quad (20)$$

$$= D(x_1, x_2, \kappa_1 x_1 + \kappa_2 x_2) \quad (21)$$

$$= D_M(x_1, x_2) \quad (22)$$

as required.  $\square$

**Simple Proof** Here is the simple, self-contained proof we promised in the introduction. Dividing through by  $x_1$  in the definition of  $D$ , Equation (8), it is clear that  $D_M(x_1, x_2) = x_1 \tilde{D}_M(x_2/x_1)$  for some function  $\tilde{D}_M$ . Thus

$$\begin{aligned} x_1 \frac{\partial D_M}{\partial x_1} + x_2 \frac{\partial D_M}{\partial x_2} &= x_1 \left( \tilde{D}_M - \frac{x_2}{x_1} \frac{\partial \tilde{D}_M}{\partial x_1} \right) + x_2 \frac{\partial \tilde{D}_M}{\partial x_2} \\ &= D_M \end{aligned}$$

which completes the proof.  $\square$

We now prove two more Myers and Read-like results which follow easily from Proposition 1. Using the implicit function theorem, Burkill and Burkill (1980), there is a function  $K(x_1, x_2)$  so that  $I(x_1, x_2, K(x_1, x_2)) = c$  is a constant.

**Corollary 2** *Assume losses  $X_i$  form a homogeneous family for each  $i$ . Then surplus values defined by constant probability of default “add-up” in that*

$$x_1 \frac{\partial K}{\partial x_1} + x_2 \frac{\partial K}{\partial x_2} = K. \quad (23)$$

**Proof** Proposition 1 implies

$$x_1 \frac{\partial I}{\partial x_1} + x_2 \frac{\partial I}{\partial x_2} + k \frac{\partial I}{\partial k} = 0. \quad (24)$$

By the implicit function theorem

$$\frac{\partial K}{\partial x_1} = - \frac{\partial I}{\partial x_1} / \frac{\partial I}{\partial k} \quad (25)$$

and similarly for  $x_2$ . Rearranging Equation (24) and substituting Equation (25) gives

$$x_1 \frac{\partial K}{\partial x_1} + x_2 \frac{\partial K}{\partial x_2} = K, \quad (26)$$

so surplus values “add-up” just as Myers and Read’s default values add-up.  $\square$

Next, use the implicit function theorem to define a function  $L(x_1, x_2)$  so that  $D(x_1, x_2, L(x_1, x_2)) = c$ .

**Corollary 3** *Assume losses  $X_i$  form a homogeneous family for each  $i$ . Then surplus values defined by constant expected policy holder deficit satisfy*

$$x_1 \frac{\partial L}{\partial x_1} + x_2 \frac{\partial L}{\partial x_2} = L + T \quad (27)$$

where  $T = TVaR(x_1 + x_2 + L(x_1 + x_2))$  is the tail-value at risk beyond  $x_1 + x_2 + L(x_1 + x_2)$ .

**Proof** Using the implicit function theorem again, and dividing Proposition 1 item 5 by  $-\partial D/\partial k$ , we get

$$x_1 \frac{\partial L}{\partial x_1} + x_2 \frac{\partial L}{\partial x_2} = L - D / \frac{\partial D}{\partial k}. \quad (28)$$

Thus, by Equation (16)

$$x_1 \frac{\partial L}{\partial x_1} + x_2 \frac{\partial L}{\partial x_2} = L + T$$

where  $T$  is the tail-value at risk.  $\square$

## 4 Examples

By Proposition 1, we can give one-dimensional examples and know they will extend to the multivariate situation as expected. We make use of this simplification in several of the examples below.

## 4.1 Examples of Homogeneity

Homogeneous families can be made from a wide variety of continuous distributions. For example, varying the scale parameter  $\theta$  and holding all other parameters constant for any of the distributions listed in Appendix A of Klugman, Panjer and Willmot (1998) which have a scale parameter  $\theta$ , will produce a homogeneous family. This includes suitable parameterizations of the transformed beta, Burr, generalized Pareto, Pareto, transformed gamma, gamma, Weibull, exponential, and inverse Gaussian. By Proposition 1, sums of selected from such families will also be homogeneous. Also, trivially, if  $X$  is any distribution with mean 1 then  $xX$  is a homogeneous family as  $x$  varies.

For example if  $X$  has an exponential distribution with mean  $x$ , so  $\Pr(X > t) = \exp(-t/x)$ , then  $X = xU$  where  $U$  has an exponential distribution with mean 1. This follows since

$$\Pr(X > t) = \exp(-t/x) = \Pr(U > t/x).$$

Here

$$I(x, k) = \Pr(X > x + k) = \exp(-k/x)/e$$

which clearly satisfies item 2 of Proposition 1.



## 4.2 Simple Example where Homogeneity Fails

It is easy to construct examples where the homogeneity assumption fails. All members of a homogeneous family have the same coefficient of variation, therefore a family with a non-constant coefficient of variation will not be homogeneous. For example, let  $X$  be normally distributed with mean  $x$  and constant standard deviation 1. Then  $X$  is not homogeneous. By definition  $I(x, k) = 1 - \Phi(k)$  so

$$x \frac{\partial I}{\partial x} + k \frac{\partial I}{\partial k} = -k\phi(k) \neq 0,$$

where  $\Phi$  and  $\phi$  are the distribution and density for the standard normal.

If the reader is skeptical about using only one variable, he or she will find it easy to construct multivariate distribution examples using normal variables. For example, consider what Corollary 2 says when  $X_1$  is distributed  $N(x_1, 1)$  and  $X_2$  is distributed  $N(x_2, 1)$ .  $X_1 + X_2$  is distributed  $N(x_1 + x_2, \sqrt{2})$ , so

$$I(x_1, x_2, k) = 1 - \Phi(k/\sqrt{2}). \quad (29)$$

Thus  $\partial K/\partial x_i = (\partial I/\partial x_i)/(\partial I/\partial k) = 0$  for  $i = 1, 2$ . Corollary 2 then reads  $K = 0$ , which is absurd! This shows the importance of the homogeneity assumption for the results derived from Proposition 1, including Myers and Read's allocation formula. This example can also be generalized to the case where  $X_1$  and  $X_2$  are correlated.

### 4.3 Homogeneity Fails with Constant Coefficient of Variation

It is less simple, but still possible, to construct examples where the coefficient of variation is a constant function of the mean, but which nevertheless fail to satisfy the homogeneity assumption.

For example let  $X(x)$  be distributed as a gamma random variable with parameters  $\alpha = 4x^2$ ,  $\theta = 1/2$  shifted by  $x(1 - 2x)$ . Here we are using the Klugman, Panjer, Willmot parameterization so  $f(t; \alpha, \theta) = (t/\theta)^\alpha e^{-t/\theta} / t\Gamma(\alpha)$ . It is easy to check  $X(x)$  has mean  $x$ , constant coefficient of variation 1 and skewness  $1/x$ , since the skewness of a gamma  $\alpha, \theta$  is  $2/\sqrt{\alpha}$ .  $I$  is given by the incomplete gamma function,  $I(x, k) = \Gamma(4x^2, 4x^2 + 2k)$ , which does not satisfy the assumptions of Lemma 1, so  $X(x)$  is not homogeneous. The reason is clear: the family  $X(x)$  changes shape with  $x$  and so cannot be homogeneous.

Taking this a step further, it is possible to construct a family all of whose higher cumulants (coefficient of variation, skewness, kurtosis, etc.) are independent of the mean, just as they would be for a homogeneous family, but which nevertheless fails to be homogeneous. To do this, let  $U$  be a lognormal random variable with  $\ln(U)$  distributed as a standard normal. Let  $V$  be a random variable density function  $f_V(x) = f_U(x)(1 + \sin(2\pi \log(x)))$ , where  $f_U$  is the density of  $U$ . Then  $U$  and  $V$  have the same moments—see Feller (1971), Chapter VII.3. This type of trick is possible because the moments of a lognormal grow too quickly to ensure it is determined by its moments—see also Billingsley (1986) Section 30. Let  $X(x)$  be a mixture of  $xU$  and  $xV$  with weights  $p(x) = x/(x + 1)$  and  $1 - p(x)$ . Then

$$I(x, k) := p(x)\Pr(U > 1 + k/x) + (1 - p(x))\Pr(V > 1 + k/x)$$

is not a function of  $k/x$  so the result follows from Lemma 1 and Proposition 1. Alternatively, writing  $I_U(x, k) = \Pr(xU > x + k)$  and similarly for  $V$  one can compute directly

$$x \frac{\partial I}{\partial x} + k \frac{\partial I}{\partial k} = xp'(x)(I_U(x, k) - I_V(x, k)) \neq 0$$

since  $xU$  and  $xV$  are homogeneous,  $xp'(x) > 0$  by construction, and  $I_U - I_V \neq 0$ . Thus  $X(x)$  is not a homogeneous family.

#### 4.4 Aggregate Distributions are Not Homogeneous

Example 4.1 shows a large number of continuous variables satisfy the homogeneity assumption. For our purposes, however, there is a very important class which does not: aggregate loss distributions.

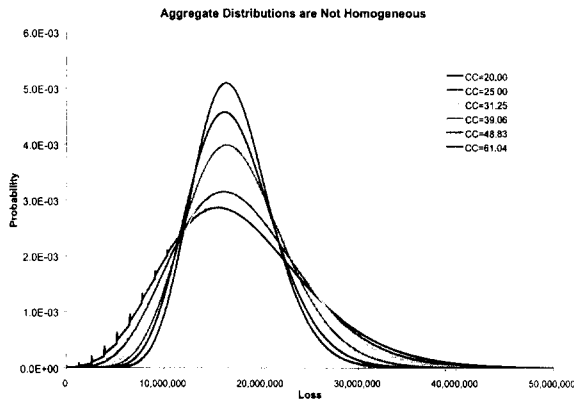
Let  $A = X_1 + \dots + X_N$  where the  $X_i$  are independent, identically distributed severities and  $N$  is a frequency distribution with mean  $n$ . Increasing expected losses in this model involves increasing  $n$ . Suppose  $N$  has contagion  $c$ , so, as suggested by Heckman and Meyers (1983),  $\text{Var}(N) = n(1 + cn)$ . Then

$$\text{CV}(A)^2 = \frac{\text{CV}(X)^2}{n} + \frac{1}{n} + c$$

is clearly not independent of  $n$ . Thus  $A$  does not satisfy the homogeneity assumption. Just as for Example 4.3, the aggregate loss distribution changes shape as  $n$  increases. This is illustrated in the figure below, which shows six aggregate loss distributions with the same severity distribution but different claim counts, indicated by “CC=20” for  $n = 20$ , and so forth. The individual densities have

been scaled so that if the family were homogeneous then all the densities would be identical and only one line would appear in the plot.

If aggregate distributions can be approximated by various of the parametric distributions of Example 4.1, and if those distributions are homogeneous, does the result of this Example really matter? The answer is emphatically “yes”. This example shows that in the real world, where insurers grow by adding discrete insureds, the “adds-up” results do not hold. The way the aggregate distribution changes shape forces parameters other than the scale parameter to change as the mean increases, and thus homogeneity is lost.



#### 4.5 Compound Poisson Distributions are not Homogeneous

This example proves that various aggregate distributions can never be homogeneous families.

**Proposition 2** *Let  $A$  be a compound Poisson aggregate distribution*

$$A = X_1 + \cdots + X_N \quad (30)$$

*where  $N$  has a Poisson distribution with mean  $n$  and the  $X_i$  are independent and identically distributed. Then  $A(n)$  is not a homogeneous family.*

**Proof** The moment generating function of  $N$  is  $M_N(t) = \exp(n(e^t - 1))$ . The moment generating function of  $A$  is therefore  $M_A(t) = \exp(n(M_X(t) - 1))$  where  $M_X$  is the moment generating function for severity  $X$ . If  $A$  is homogeneous with  $A$  distributed as  $nU$  for some fixed  $U$ , then  $M_A(t) = M_U(nt)$ . Thus

$$n(M_X(t) - 1) = \log(M_U(nt)).$$

Differentiating with respect to  $n$  shows

$$(M_X(t) - 1) = tM'_U(nt)/M_U(nt).$$

Therefore  $M'_U(t)/M_U(t)$  must be a constant, since the left hand side is independent of  $n$ . Hence  $M_U(t) = \exp(ct)$  for a constant  $c$ , and so  $U = c$  is a degenerate distribution. But this is impossible unless  $N$  is constant or  $X_i \equiv 0$ .  $\square$

**Corollary 4** *An aggregate distribution with frequency component  $N$  which is a mixture of Poisson distributions cannot be a homogeneous family.*

**Proof** Condition on the mixing parameter and apply Proposition 2.  $\square$

For example, the Corollary applies to negative binomial and Poisson-inverse Gaussian frequency distributions.

## 5 Is Inhomogeneity Material?

In this section we will show that the inhomogeneity inherent in a typical portfolio of property casualty risks is sufficiently large to invalidate the Myers and Read allocation formula. By Proposition 1, we can discuss inhomogeneity in the context of one random variable, rather than two or more, which simplifies the mathematics.

Let  $X(x)$  be a smooth family of random variables with  $E(X(x)) = x$ . Let  $F(t, x) = \Pr(X(x) < t)$  be the distribution function of  $X(x)$  and  $f(t, x) = \partial F / \partial t$  be its density.

Recall that  $X(x)$  is homogeneous (with respect to the mean) if there exists a random variable  $U$  so that

$$X(x) = xU \tag{31}$$

for all  $x$ . In this case, let  $F_U$  and  $f_U$  be the distribution and density functions of  $U$ .

Recall also that the expected default, with capital ratio  $\kappa$ , is defined as

$$D(x) = \int_{x(1+\kappa)}^{\infty} (t - x(1 + \kappa))f(t, x)dt. \tag{32}$$

Note that  $x(1 + \kappa)$  represents total assets:  $x$  from the loss and  $x\kappa$  from allocated capital. In a more sophisticated model we could consider profit in the premium; here we simply assume this is subsumed into the constant  $\kappa$ .

By Proposition 1 we know

$$x \frac{\partial D}{\partial x} = D \tag{33}$$

if and only if  $X(x)$  is a homogeneous family, which is then equivalent to the Myers-Read add-up result.

## 5.1 Heuristics

A homogeneous family offers no diversification benefit as the mean increases. Property casualty insurance is based on diversification, and the inhomogeneity inherent in a portfolio of insurance risks means that the relative riskiness of the portfolio decreases as expected losses increase. Since a lower risk portfolio has a lower expected default, one would expect that

$$x \frac{\partial D}{\partial x} < D \quad (34)$$

for an inhomogeneous insurance portfolio.

Meyers (2003) introduces the heterogeneity multiplier, which is a constant  $\lambda$  defined so that

$$\lambda x \frac{\partial D}{\partial x} = D. \quad (35)$$

He shows that  $\lambda$  is typically greater than 1 (as expected). In further unpublished work, Meyers uses empirical data to estimate that  $\lambda$  is in the range 1.5 to 2.5, depending on the size of the company. This suggests that inhomogeneity is material.

If  $X$  is homogeneous then, for all  $\kappa > 0$ ,

$$x \frac{\partial D}{\partial x} = D \geq 0. \quad (36)$$

However, intuitively, one would expect that for a large enough capital ratio  $\kappa$  it should be possible for the extra capital associated with writing more business to more than offset the extra risk. This would imply that

$$x \frac{\partial D}{\partial x} < 0 \quad (37)$$

should be possible for sufficiently large  $\kappa$ . This is another difference between homogeneous and inhomogeneous families.

## 5.2 Theory

In order to assess the impact of inhomogeneity, we will break the derivative  $\partial D/\partial x$  into two pieces using a homogeneous approximation to the family  $X(x)$ .

For a fixed  $x$ , define a new homogeneous family  $Y(y)$  by

$$Y(y) = \frac{y}{x} X(x). \quad (38)$$

Let  $G(t, y)$  and  $g(t, y)$  be the distribution and density functions of  $Y$ . Note that  $Y(x) = X(x)$  and that

$$g(t, y) = f(tx/y, x)x/y. \quad (39)$$

If  $X$  is already homogeneous, then clearly  $Y(y) = X(y)$  for all  $y$ . Finally define  $E_x(y)$  to be the expected default value of  $Y$ ,

$$E_x(y) = \int_{y(1+\kappa)}^{\infty} (t - y(1+\kappa))g(t, y)dt. \quad (40)$$

The subscript  $x$  on  $E$  highlights the point  $x$  at which we have chosen to “homogenize”  $X$ . By definition  $E_x(x) = D(x)$ .

We can now compute

$$\frac{\partial D}{\partial x} = \lim_{\epsilon \rightarrow 0} \frac{D(x + \epsilon) - D(x)}{\epsilon} \quad (41)$$

$$= \lim_{\epsilon \rightarrow 0} \frac{D(x + \epsilon) - E_x(x + \epsilon) + E_x(x + \epsilon) - D(x)}{\epsilon} \quad (42)$$

$$= \lim_{\epsilon \rightarrow 0} \frac{D(x + \epsilon) - E_x(x + \epsilon)}{\epsilon} + \lim_{\epsilon \rightarrow 0} \frac{E_x(x + \epsilon) - D(x)}{\epsilon} \quad (43)$$

$$= I(x) + \lim_{\epsilon \rightarrow 0} \frac{E_x(x + \epsilon) - E_x(x)}{\epsilon} \quad (44)$$

$$= I(x) + \frac{D(x)}{x} \quad (45)$$



where  $I(x)$  is defined by the first limit, and we have used the fact that  $Y$  is homogeneous to replace  $\partial E_x/\partial x$  with  $E_x(x)/x = D(x)/x$ .

We can now prove the main lemma of this section.

**Lemma 3** *With the above notation*

$$I(x) = \int_{x(1+\kappa)}^{\infty} (s - x(1 + \kappa)) \left( \frac{\partial f}{\partial x} + \frac{s}{x} \frac{\partial f}{\partial t} + \frac{f(s, x)}{x} \right) ds. \quad (46)$$

**Proof** Substitute  $s = tx/(x + \epsilon)$  in the limit defining  $I$ , swap the limit and integral (Lebesgue's dominated convergence theorem), and use the fact that the limit of a product (quotient) is the product (quotient) of the limits to get

$$I(x) = \int_{x(1+\kappa)}^{\infty} (s - x(1 + \kappa)) \lim_{\epsilon \rightarrow 0} \left( \frac{(x + \epsilon)f(s + s\epsilon/x, x + \epsilon) - xf(s, x)}{(x + \epsilon)\epsilon} \right) ds. \quad (47)$$

Now add and subtract a term  $(x + \epsilon)f(s + s\epsilon/x, x)$  in the limit, re-arrange and cancel. The result follows.  $\square$

We will call  $I(x)$  the inhomogeneous derivative of  $D$  with respect to  $x$ . We will use the standard notation  $f_1 = \partial f/\partial t$  and  $f_2 = \partial f/\partial x$ .

If  $X$  is homogeneous then

$$f(t, x) = f_U(t/x)/x \quad (48)$$

$$f_1(t, x) = f'_U(t/x)/x^2 \quad (49)$$

$$f_2(t, x) = -tf'_U/x^2 - f_U/x^2 \quad (50)$$

and so

$$f_2 + \frac{s}{x} f_1 + \frac{1}{x} f = 0. \quad (51)$$

Thus if  $X$  is homogeneous  $I(x) = 0$  as expected.

The Lemma shows that

$$xI(x) = \int_{x(1+\kappa)}^{\infty} (s - x(1 + \kappa))(xf_2 + sf_1)ds + D(x), \quad (52)$$

and so

$$x \frac{\partial D}{\partial x} = \int_{x(1+\kappa)}^{\infty} (s - x(1 + \kappa))(xf_2 + sf_1)ds + 2D(x). \quad (53)$$

When  $X$  is homogeneous, the integral in  $I$  exactly cancels out the extra  $D$  term. In the tail of the distribution, we expect  $f_1 < 0$ , because the density will eventually be decreasing with  $t$ , and  $f_2 > 0$  because for a given  $t$  the density  $f(t, x)$  will increase as the mean  $x$  increases. The exact balance of these two terms depends on the degree of inhomogeneity.

### 5.3 Examples of Inhomogeneity

At this point we have developed enough general theory. For a realistic insurance portfolio we expect  $x\partial D/\partial x < D$ , and possibly that  $x\partial D/\partial x < 0$ . In order to test the magnitude of these effects we will use the following model.

Let  $A(x)$  be an aggregate loss distribution with expected losses  $x$ , severity component  $S(l)$  and frequency component  $N$ , so

$$A(x) = S(l)_1 + \cdots + S(l)_N. \quad (54)$$

We assume that  $S(l) = \min(S, l)$  results from applying a limit  $l$  to a fixed unlimited severity  $S$ . In the example  $S$  is chosen to be reasonably close to ISO's Premises and Operations B curve. The frequency distribution  $N$  is negative binomial with claim count  $n = x/E(S(l))$  and contagion  $c$ , so  $\text{Var}(N) = n(1 + cn)$ . In

Table 1: Total Derivative  $\partial D/\partial x$

$x l$	1,000,000	5,000,000	10,000,000	100,000,000
10,000,000	0.28%	-0.42%	-0.37%	3.74%
50,000,000	0.39%	0.30%	0.13%	-0.40%
100,000,000	0.39%	0.37%	0.31%	-0.02%
1,000,000,000	0.39%	0.39%	0.39%	0.39%

the tables below  $c = 0.15$ , corresponding to an asymptotic coefficient of variation of  $A(x)$  of 38.7%. The capital ratio  $\kappa = 1$ , so the expected loss to surplus ratio is 1 to 1.

In order to compute the necessary derivatives, we will approximate  $A(x)$  with a shifted lognormal distribution, using the method of moments to match the mean, variance, and skewness. For large portfolios, the shifted lognormal is a very good approximation to the true aggregate distribution. This can be seen by comparing the result of using FFTs to compute the true aggregate with the shifted lognormal approximation. Figure 1 shows that the approximation is quite spectacularly good, particularly in the relevant range beyond  $2x$ . Regardless of whether you believe this is a good approximation or not, the approximation has qualitatively the correct shape and behaviour as  $x$  changes.

Let  $X(x)$  be the shifted lognormal approximation to  $A(x)$ . If  $X(x)$  has parameters  $\tau$ ,  $\mu$ , and  $\sigma$ , so  $\ln(Y - \tau)$  is distributed  $N(\mu, \sigma)$ , then the homogeneous approximation  $Y(y)$  to  $X(x)$  has parameters  $y/x\tau$ ,  $\ln(y/x) + \mu$ , and  $\sigma$ . Therefore we can compute  $I$  and  $D$  explicitly.

In each table, expected loss amounts  $x$  are shown vertically and different limits  $l$  are shown across the columns. Patterns in Table 1 are hard to see directly, and are

Table 2: Homogeneous Derivative = Expected Default Ratio =  $D/x$

$x \setminus l$	1,000,000	5,000,000	10,000,000	100,000,000
10,000,000	0.97%	4.21%	8.12%	19.27%
50,000,000	0.49%	0.92%	1.49%	3.99%
100,000,000	0.44%	0.63%	0.87%	1.90%
1,000,000,000	0.40%	0.41%	0.43%	0.49%

Table 3: Inhomogeneous Derivative  $I(x) = \partial D / \partial x - D/x$

$x \setminus l$	1,000,000	5,000,000	10,000,000	100,000,000
10,000,000	-0.69%	-4.64%	-8.49%	-15.53%
50,000,000	-0.10%	-0.62%	-1.36%	-4.39%
100,000,000	-0.05%	-0.26%	-0.56%	-1.92%
1,000,000,000	0.00%	-0.02%	-0.04%	-0.11%

Table 4: Heterogeneity Multiplier

$x \setminus l$	1,000,000	5,000,000	10,000,000	100,000,000
10,000,000	3.45	-9.96	-21.66	5.15
50,000,000	1.26	3.10	11.72	-9.97
100,000,000	1.12	1.72	2.80	-91.90
1,000,000,000	1.01	1.05	1.10	1.28

best understood as the sum of the homogeneous and inhomogeneous derivatives in Tables 2 and 3. In Table 2 we see that the homogeneous derivative (derivative of the homogeneous approximation to  $X$ ) increases with the limit  $l$  and decreases with expected losses  $x$ . This makes sense: increasing the limit increases the riskiness of the portfolio and hence  $D$ . Increasing expected losses yields a diversification benefit and decreases  $D$ .

Table 3 shows that the inhomogeneous derivative increases with  $x$ , eventually tending to zero. This reflects the fact that the aggregate becomes very nearly homogeneous for large  $x$ ,  $x \gg l$ . As  $l$  increases  $I$  decreases, reflecting the fact that the underlying distribution  $A$  is becoming more and more inhomogeneous.

Table 4 shows the heterogeneity multiplier, or ratio of the homogeneous derivative to the total derivative. In a reasonable range of  $x$  between 10M and 100M and smaller limits, this is of the same order of magnitude as in Meyers' study.

In practical applications, where the adds-up formula would be used in the context of allocating surplus between business units or lines of business, expected losses would be in the 10M to 100M range with limits of 1M to 10M. The Tables show that in such a range the lack of homogeneity in an insurance portfolio is material, and would mean the adds-up result would fail to hold by a substantial amount.

## 6 Conclusions

In this paper we have explained the importance of the homogeneity assumption in the derivation of Myers and Read's "adds-up" result. Proposition 1 shows the assumption is necessary as well as sufficient. We have used Proposition 1 to prove two other results in a similar vein, including one involving tail value at risk. Importantly, for practical applications, we have shown that most common families of aggregate distributions will never satisfy the homogeneity assumption. We have given several realistic examples to support the general theory. We conclude that, in a real-world situation, where insurers grow by adding individual risks from discrete insureds, the "adds-up" result will not hold.

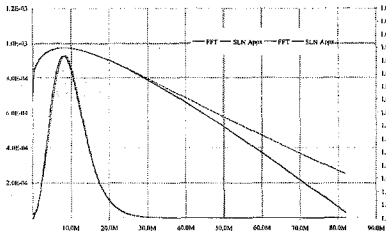
## References

- Billingsley, Patrick. 1986. **Probability and Measure** Second Edition, J. Wiley and Sons
- Burkill, J. C. and H. Burkill. 1980. **A Second Course in Mathematical Analysis**, Cambridge University Press.
- Butsic, Robert P. 1999. "*Capital Allocation for Property-Liability Insurers: A Catastrophe Reinsurance Application*", Casualty Actuarial Society Spring Forum, pp 1–70
- Feller, William. 1971. **An Introduction to Probability Theory and Its Applications** volume II, Second Edition, J. Wiley and Sons.

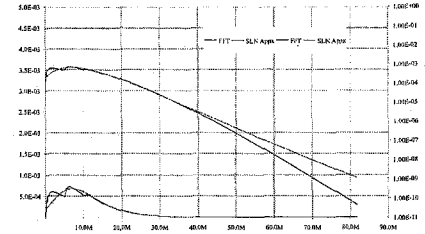
- Heckman, Philip and Glenn G. Meyers. 1983. "*The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions*" Proceedings of the Casualty Actuarial Society, Vol. LXXX pp 22–61.
- Klugman, Stuart A., Harry H. Panjer and Gordon E. Willmot. 1998. **Loss Models From Data to Decisions** J. Wiley and Sons
- Merton, Robert C. and Andre Perold. 2001. "*Theory of Risk Capital in Financial Firms*" pp 438–454 in *The New Corporate Finance, Where Theory Meets Practice*, Edited by Donald H. Chew, Jr., McGraw-Hill.
- Meyers, Glenn. 2003. "*The Economics of Capital Allocation*" Bowles Symposium, Insurance Services Office, Inc. <http://www.casact.org/research/aria/meyers.pdf>
- Meyers, Glenn. 1991. "*The Competitive Market Equilibrium Risk Load Formula for Increased Limits Ratemaking*" Proceedings of the Casualty Actuarial Society, Vol. LXXVIII pp 163–200.
- Myers, Stewart C. and James A. Read Jr. 2001. "*Capital Allocation for Insurance Companies*" *Journal of Risk and Insurance*, Vol 68 no. 4, pp 545–580.
- Phillips, Richard D., J. David Cummins and Franklin Allen. 1998. "*Financial Pricing of Insurance in the Multiple-Line Insurance Company*" *Journal of Risk and Insurance*, Vol 65 no. 4, pp 597–636.

Figure 1  
 Comparison of Shifted Lognormal Approximation to Aggregate Distribution for Variety of Means  $\mu$  and Limits  $l$

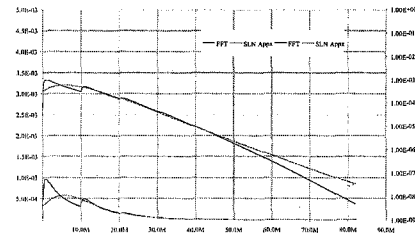
Expected Losses  $\mu=10,000,000$  and Layer  $l=1,000,000$



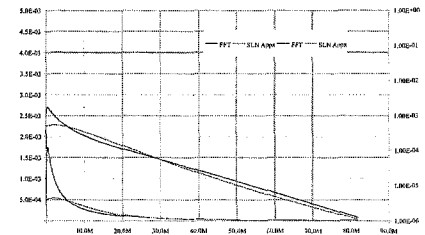
Expected Losses  $\mu=10,000,000$  and Layer  $l=5,000,000$



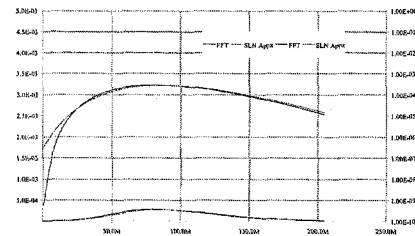
Expected Losses  $\mu=10,000,000$  and Layer  $l=10,000,000$



Expected Losses  $\mu=10,000,000$  and Layer  $l=100,000,000$



Expected Losses  $\mu=100,000,000$  and Layer  $l=1,000,000$



Expected Losses  $\mu=100,000,000$  and Layer  $l=100,000,000$

