

Minimizing Ruin Probability Under Dependencies for Insurance Pricing

R.L. Gudmundarson ^{*1}, M. Guerra ^{†2}, and A. B. de Moura ^{‡2}

¹Edinburgh Business School, Heriot-Watt University

²ISEG-School of Economics and Management, Universidade de Lisboa; REM - Research in Economics and Mathematics, CEMAPRE

Abstract

In this work the ruin probability of the Lundberg risk process is used as a criterion for determining the optimal security loading of premia in the presence of price-sensitive demand for insurance. Both single and aggregated claim processes are considered and the independent and the dependent cases are analyzed. For the single-risk case, we show that the optimal loading does not depend on the initial reserve. In the multiple risk case we account for arbitrary dependency structures between different risks and for dependencies between the probabilities of a client acquiring policies for different risks. In this case, the optimal loadings depend on the initial reserve. In all cases the loadings minimizing the ruin probability do not coincide with the loadings maximizing the expected profit

1 Introduction

Insurance is based on the idea that society asks for protection against unforeseeable events which may cause serious financial damage. Insurance companies offer a financial protection against these events. The general idea is to build a community where everybody contributes a certain amount and those who are exposed to the damage receive financial reimbursement [35].

When (non-life) insurers set premium prices they usually start by finding the so-called pure premium, which is the expected value of the total claims that will occur in one time unit. However, when pricing insurance policies, insurers must take into account the risk associated with the policy as well as additional costs (e.g. operational cost, capital cost, etc.). Therefore, a so-called security loading is added to cover the risk and additional costs. The security loading is often calculated using some premium calculation principle, and the insurance premium is obtained once the security loading has been determined and added to the pure premium. The main concerns are usually whether the loading is an appropriate measure of the risk and which premium principle to choose. The higher the loading the higher the premium and consequently, the underwriting risk will be lower. However, if the premium price is too high then the exposure will be too low due to competition, and the operational cost of the insurer will engulf the premium income resulting in financial instability. Therefore, insurers usually require sophisticated premium calculations in order to secure stability.

Collective risk models are fundamental in actuarial science to model the aggregate claim amount of a line of business in an insurance company. The collective risk model has two random components, the number of claims and the severity of claims, and is usually modelled with a compound process [15, Chapter 3]. The classical Lundberg risk process has been studied extensively and there exist many variations, for example including reinsurance or investments [14]. It assumes that premia come in a continuous stream while claims happen at discrete times according to a Poisson distribution.

Another common assumption is that the risk can be divided into groups of homogeneous risks such that the pure premia and security loadings can be estimated separately for each risk group. The pure premia of these individual groups are usually modelled with generalized linear models (GLM). GLM's have been applied extensively in actuarial work and a good overview is provided in [26, 36]. Traditional risk theory has usually assumed independence between risks due to its convenience, but it is generally not very realistic. Claims in an insurer's risk portfolio are correlated as they are subject to the same event causes [7]. Completely homogeneous risk groups are extremely rare and dependence among risks has become a flourishing topic in actuarial literature

*E-mail address: rlg2000@hw.ac.uk

†E-mail address: mguerra@iseg.ulisboa.pt

‡E-mail address: amoura@iseg.ulisboa.pt

[2]. Dependence has mostly been measured through linear correlation coefficients [4]. The popularity of linear coefficient is mainly due to the ease with which dependencies can be parameterized, in terms of correlation matrices. Most random variables, however, are not jointly elliptically distributed and it could be very misleading to use linear coefficients [30]. This motivated the use of concordance measures. Two random variables are concordant when large values of one go with large values of the other [25]. The Lundberg risk model is a Lévy jump process, [8] which means that the dependency of two claim processes is best explained through their Lévy measure [32]. This study will not go into details about Lévy processes, but both [8] and [27] provide a very good introduction, and [2, 5, 28, 34, 1] are examples of applications of Lévy copulas to risk processes. For example, van Veesen [34] showed how Lévy copulas can be used in operational modelling and discussed how dependence is implied by the Lévy copula. In this work we consider bivariate claim processes, but the presented theory can be straightforwardly extended to multiple claim processes.

Ruin probability is a classical measure of risk and has been extensively studied [15, 14, 18, 33]. Although there is no absolute meaning to the probability of ruin, it still measures the stability of insurance companies. A high ruin probability indicates instability, and risk mitigation techniques should be used, like reinsurance or raising premia [15]. Most non-life insurance products have a term of one year and therefore it can be argued that the one year ruin probability should be used. The one year ruin probability is the probability that the capital of an insurance company will hit zero within one year. However, the appropriateness of risk measures defined over fixed time horizons can be questioned, since ruin in a given time span can be minimized by increasing the probability of ruin in the aftermath of that period. Lundberg concluded that the actual assumptions behind the classical collective risk model are in fact less restrictive when time-invariant quantities like the infinite time ruin probability are considered [33]. Therefore, we focus on the infinite time ruin probability in this paper.

In this work, the optimal loadings based on two strategies are derived, and compared. One strategy maximizes the profit and the other minimizes the ruin probability. We show that the two loading strategies give different results. Furthermore, we show how the optimal loading with respect to the ruin probability can be found and compare it to the one obtained when the expected profit is maximized. We consider dependencies and illustrate how Lévy copulas can be used to model claim process dependencies and how dependencies can affect the riskiness of the insurance portfolio. We take this idea further and consider dependency between the acquisition of insurance for different risks by policyholders. This is a realistic assumption as policyholders usually buy multiple insurance products from the same insurance company. We also take into account the fact that the market risk process and the company's risk process are not the same, and how the company's risk process depends on its exposure to the market. This is, to our knowledge, the first analysis of the interplay of the ruin probability, the dependency structure of claim, and the dependency structure of acquisition of insurance. We demonstrate that even if there is a strong dependency between insurance products within the market, small insurance companies have less dependency and therefore less risk than bigger insurance companies, provided the dependency between acquisition of insurance for different risks is not too strong.

The paper is organized as follows: Section 2 contains some background material about ruin probabilities in the Lundberg process and aggregation of compound Poisson processes. Section 3 deals with the single-risk case. We characterize the optimal loading and compare it with the loading maximizing the expected profit. Section 4 handles the multiple risks case. We show how the dependency structure existing in the market (i.e. the general population) translates into the risk exposure of the company through its market shares on different risks and the likelihood that clients acquire insurance for more than one risk. Section 5 contains a numerical illustration. A numerical scheme to compute the ruin probabilities is given in the appendix.

2 Preliminaries

2.1 Claim and Surplus Processes

The Lundberg risk model describes the evolution of the capital of an insurance company and assumes that exposure is constant in time, losses follow a compound Poisson process, and premia arrive at a fixed continuous rate:

$$X_t = u + ct - \sum_{i=0}^{N_t} Y_i = u + ct - S_t, \quad Y_0 := 0,$$

where u is the initial surplus, c is the risk premium rate, N_t is a time homogeneous Poisson process with intensity parameter λ , and Y_i are i.i.d. random variables representing the severity of claim i , $i = 0, \dots, N_t$. Here it is assumed that Y_i are positive. In the following sections, Y denotes an arbitrary random variable with the same distribution as any Y_i . The severity distribution is denoted as $F(x)$ and the severity survival distribution as $\bar{F}(x)$. S_t is a compound Poisson process and thus X_t is a stochastic process (sometimes called the surplus process) representing the insurance wealth at time t . X_t increases because of earned premia and decreases when

claims occur. When the capital of an insurance company hits zero, the insurance company is said to be ruined. Formally, the ruin probability is defined as follows.

Definition 2.1 (Probability of Ruin). *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ be a filtered probability space and $X = (X_t)_{t \in [0, \infty[}$ a surplus process which is adapted and Markov with respect to the filtration. The state space is $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. If X is time homogeneous, the infinite time ruin probability is the function $V : \mathbb{R} \mapsto [0, 1]$ such that*

$$V(x) = \mathbb{P}(\exists s \in [0, +\infty[: X_s \leq 0 \mid X_0 = x), \quad x \in \mathbb{R}.$$

Sometimes it is useful to use the survival (non-ruin) probability, defined as $\bar{V}(x) = 1 - V(x)$. The ruin probability can be calculated using the following integro-differential equation [11].

Proposition 2.1. *Assume that X_t is defined as above and the premium rate satisfies $c > \lambda \mathbb{E}[Y]$. If $V \in C^1(]0, \infty[)$, then the probability of ruin with infinite time horizon satisfies the following equation:*

$$0 = c \frac{d}{dx} V(x) + \lambda \left(\int_0^x V(x-y) dF(y) - V(x) + 1 - F(x) \right), \quad x > 0, \quad (2.1)$$

with the following boundary condition:

$$\begin{cases} V(x) = 1 & x \leq 0, \\ \lim_{x \rightarrow 0^+} V(x) = \frac{\lambda}{c} \mathbb{E}[Y]. \end{cases}$$

Furthermore, the probability of non-ruin satisfies the following equation:

$$\bar{V}(x) - \bar{V}(\epsilon) = \frac{\lambda}{c} \int_{\epsilon}^x \bar{V}(x-y) \bar{F}(y) dy \quad (2.2)$$

for $0 < \epsilon \leq x < +\infty$ with the following boundary condition:

$$\begin{cases} \bar{V}(x) = 0 & x \leq 0, \\ \lim_{x \rightarrow 0^+} \bar{V}(x) = 1 - \frac{\lambda}{c} \mathbb{E}[Y]. \end{cases}$$

A numerical scheme solving equation (2.1) can be found in Appendix A.

2.2 Accounting for Claim Dependencies

Consider the surplus process $\mathbf{X} = (X_t^{(1)}, \dots, X_t^{(n)})$ where

$$\begin{aligned} X_t^{(1)} &= u^{(1)} + c^{(1)}t - \sum_{i=0}^{N_t^{(1)}} Y_i^{(1)} \\ &\vdots \\ X_t^{(n)} &= u^{(n)} + c^{(n)}t - \sum_{i=0}^{N_t^{(n)}} Y_i^{(n)} \end{aligned} \quad (2.3)$$

If these processes are independent, it is relatively easy to combine them into a single process using the aggregation property of compound Poisson processes as described in Wütrich [35]. The aggregation property allows the combination of multiple surplus processes into a single risk process as follows:

$$X_t = \sum_{j=1}^n u^{(j)} + \sum_{j=1}^n c^{(j)}t - \sum_{i=0}^{N_t} Y_i,$$

where N_t is a Poisson r.v. with $\lambda = \lambda_1 + \dots + \lambda_n$ and Y_i are i.i.d. random variables, which follow the severity distribution $F(x) = \sum_{j=1}^n \frac{\lambda_j}{\lambda} F_j(x)$. This aggregation property allows us to use the integro-differential equation (2.1) to calculate the ruin of multiple surplus processes.

If the risks are not independent, then we can use the fact that compound Poisson processes are characterized by their Lévy measure to decompose the claim process into independent processes to which the aggregation property can be applied. In particular, for $n = 2$ risks, we obtain the decomposition:

$$X_t = X_t^{(1)} + X_t^{(2)} = u + ct - S_t^{1\perp} - S_t^{2\perp} - S_t^{\parallel}.$$

where $S^{1\perp}$ and $S^{2\perp}$ are compound Poisson processes accounting for events concerning only risk 1 and risk 2, respectively. S^{\parallel} is a compound Poisson process accounting for events concerning both risks simultaneously. Furthermore, $S^{1\perp}$, $S^{2\perp}$ and S^{\parallel} are mutually independent.

In this section, we briefly explain how this can be achieved. Further details can be found in [32]. We will use the following definitions:

Definition 2.2. *The tail integral of a Lévy measure ν on $[0, \infty]^2$ is given by a function $U : [0, \infty]^2 \mapsto [0, \infty]$*

$$\begin{aligned} U(x_1, x_2) &= 0 \quad \text{if } x_1 = \infty \quad \text{or} \quad x_2 = \infty, \\ U(x_1, x_2) &= \nu([x_1, \infty[\times]x_2, \infty[) \quad \text{for } (x_1, x_2) \in]0, \infty[^2, \\ U(0, 0) &= \infty. \end{aligned} \tag{2.4}$$

Definition 2.3 (Lévy Copula for Processes with Positive Jumps). *A two-dimensional Lévy copula for Lévy processes with positive jumps, or for short, a positive Lévy copula, is a 2-increasing grounded function $\mathcal{C} : [0, \infty]^2 \rightarrow [0, \infty]$ with uniform margins, that is, $\mathcal{C}(x, \infty) = \mathcal{C}(\infty, x) = x$.*

Similarly to Sklar's theorem for ordinary copulas [25], it has been shown that the dependency structure of $(X_t^{(1)}, X_t^{(2)})$ can be characterized by a Lévy copula \mathcal{C} such that $\mathcal{C}(U_1(x_1), U_2(x_2))$ where U_1 and U_2 are the marginal tail integrals for $X_t^{(1)}$ and $X_t^{(2)}$. If U_1 and U_2 are absolutely continuous, this Lévy copula is unique, otherwise it is unique on $\text{Range}(U_1) \times \text{Range}(U_2)$, the product of ranges of one-dimensional tail integrals, [8, Theorem 5.4]

Consider a two dimensional claim process:

$$S_t = (S_t^{(1)}, S_t^{(2)}) = \sum_{i=0}^{N_t} (Y_i^{(1)}, Y_i^{(2)}), \tag{2.5}$$

where N_t is a Poisson process with intensity λ and $Y_i = (Y_i^{(1)}, Y_i^{(2)})$, $i \in \mathbb{N}$ are independent random variables with common joint distribution F_Y . The components of S , $S^{(1)}$ and $S^{(2)}$, are one-dimensional compound Poisson processes with intensities λ_1 and λ_2 and severity distributions $F_{Y^{(1)}}$ and $F_{Y^{(2)}}$, respectively. We wish to obtain a decomposition:

$$(S_t^{(1)}, S_t^{(2)}) = \sum_{i=0}^{N_t^{1\perp}} (Y_i^{(1\perp)}, 0) + \sum_{i=0}^{N_t^{2\perp}} (0, Y_i^{(2\perp)}) + \sum_{i=0}^{N_t^{\parallel}} (Y_i^{(1\parallel)}, Y_i^{(2\parallel)}), \tag{2.6}$$

where $\sum_{i=0}^{N_t^{1\perp}} Y_i^{(1\perp)}$, $\sum_{i=0}^{N_t^{2\perp}} Y_i^{(2\perp)}$ and $\sum_{i=0}^{N_t^{\parallel}} (Y_i^{(1\parallel)}, Y_i^{(2\parallel)})$ are independent compound Poisson processes with intensities λ_1^{\perp} , λ_2^{\perp} , λ^{\parallel} and severity distributions $F_{Y^{1\perp}}$, $F_{Y^{2\perp}}$, $F_{Y^{\parallel}}$, respectively. In the above setting, we consider

$$F_Y(0, 0) = F_{Y^{\parallel}}(0, 0) = 0, \quad F_{Y^{(1)}}(0) = F_{Y^{1\perp}}(0) = F_{Y^{(2)}}(0) = F_{Y^{2\perp}}(0) \tag{2.7}$$

A compound Poisson process, S , is a Lévy process with Lévy measure $\nu(dx) = \lambda dF(x)$, with tail integral

$$U(x_1, x_2) = \begin{cases} \lambda \mathbb{P}(Y^{(1)} \geq x_1, Y^{(2)} \geq x_2) & \text{if } x_1 > 0 \text{ or } x_2 > 0 \\ +\infty & \text{if } x_1 = x_2 = 0. \end{cases}$$

The components $S^{(1)}$ and $S^{(2)}$ are independent if and only if $U(x_1, x_2) = 0$ for every $(x_1, x_2) \in]0, +\infty[$, i.e., if and only if $\lim_{x_1 \rightarrow 0^+, x_2 \rightarrow 0^+} U(x_1, x_2) = 0$.

The Lévy measure of the processes $S^{(i)}$, $i = 1, 2$, have tail integrals

$$\begin{aligned} U_1(x_1) &= \lambda_1 \mathbb{P}(Y^{(1)} \geq x_1) = U(x_1, 0) \\ U_2(x_2) &= \lambda_2 \mathbb{P}(Y^{(2)} \geq x_2) = U(0, x_2) \end{aligned}$$

Taking equation (2.7) into account, one obtains

$$\begin{aligned} \lambda_i &= \lim_{x_i \rightarrow 0^+} U_i(x_i), \quad i = 1, 2 \\ \lambda &= \lim_{x_1, x_2 \rightarrow 0^+} (U_1(x_1) + U_2(x_2) - U(x_1, x_2)) = \lambda_1 + \lambda_2 - \lambda^{\parallel} \\ \lambda^{\parallel} &= \lim_{x_1, x_2 \rightarrow 0^+} U(x_1, x_2) \\ \lambda_i^{\perp} &= \lambda_i - \lambda^{\parallel}, \quad i = 1, 2. \end{aligned}$$

The severity distributions $F_{Y^{1\perp}}$, $F_{Y^{2\perp}}$, and $F_{Y^{\parallel}}$ can be recovered from the tail integrals:

$$\begin{aligned}\mathbb{P}(Y^{1\perp} \geq x_1) &= \frac{1}{\lambda_1^\perp} \lim_{x_2 \rightarrow 0^+} (U_1(x_1) - U(x_1, x_2)) \\ \mathbb{P}(Y^{2\perp} \geq x_2) &= \frac{1}{\lambda_2^\perp} \lim_{x_1 \rightarrow 0^+} (U_2(x_2) - U(x_1, x_2)) \\ \mathbb{P}(Y^{1\perp} \geq x_1, Y^{2\perp} \geq x_2) &= \frac{1}{\lambda^\parallel} U(x_1, x_2).\end{aligned}$$

If the dependency between $S^{(1)}$ and $S^{(2)}$ is characterized by a Lévy copula, \mathcal{C} , i.e. $U(x_1, x_2) = \mathcal{C}(U_1(x_1), U_2(x_2))$ for $(x_1, x_2) \in [0, +\infty]^2$, then the relations above can be written using the Lévy copula and one-dimensional tail integrals:

$$\begin{aligned}\lambda^\parallel &= \lim_{u_1 \rightarrow \lambda_1^\perp, u_2 \rightarrow \lambda_2^\perp} \mathcal{C}(u_1, u_2) \\ \mathbb{P}(Y^{1\perp} \geq x_1) &= \frac{1}{\lambda_1^\perp} \lim_{u_2 \rightarrow \lambda_2^\perp} (U_1(x_1) - \mathcal{C}(U_1(x_1), u_2)) \\ \mathbb{P}(Y^{2\perp} \geq x_2) &= \frac{1}{\lambda_2^\perp} \lim_{u_1 \rightarrow \lambda_1^\perp} (U_2(x_2) - \mathcal{C}(u_1, U_2(x_2))) \\ \mathbb{P}(Y^{1\perp} \geq x_1, Y^{2\perp} \geq x_2) &= \frac{1}{\lambda^\parallel} \mathcal{C}(U_1(x_1), U_2(x_2)).\end{aligned}$$

Using the above methodology, the surplus process can be represented as:

$$X_t = u + ct - \sum_{i=0}^{N_t^{1\perp}} Y_i^{1\perp} - \sum_{i=0}^{N_t^{2\perp}} Y_i^{2\perp} - \sum_{i=0}^{N_t^\parallel} (Y_i^{1\parallel} + Y_i^{2\parallel}) = u + ct - \sum_{i=1}^{N_t^*} Y_i^*,$$

where $u = u_1 + u_2$, $c = c_1 + c_2$, N^* is a Poisson process with intensity $\lambda = \lambda_1^\perp + \lambda_2^\perp + \lambda^\parallel$ and Y_i^* are i.i.d. random variables with distribution:

$$F^* = \frac{\lambda_1^\perp}{\lambda} F_{Y^{1\perp}} + \frac{\lambda_2^\perp}{\lambda} F_{Y^{2\perp}} + \frac{\lambda^\parallel}{\lambda} F_{Y^{1\parallel} + Y^{2\parallel}},$$

where

$$F_{Y^{1\parallel} + Y^{2\parallel}}(x) = \int_{x_1 + x_2 \leq x} dF_{Y^\parallel}(x_1, x_2).$$

3 The Optimal Loading for a Single Risk

An insurer can control the volume of its business through the premium loading θ . A reasonable assumption is that the higher the loading, the smaller the number of contracts in its portfolio, which means that the claim intensity (or business volume) will decrease. Therefore, both the claim intensity $\mathbb{E}^\theta[N_1]$, and the premium rate $c(\theta)$, will depend on θ . It is reasonable to assume that $\mathbb{E}^\infty[N_1] = 0$, as abnormal premium rates will not attract customers [13]. To capture these concepts let $\mathbb{E}^\theta[N_1] = \lambda p(\theta)$. Here λ is the average number of claims per unit of time for the whole market, and $p(\theta)$ is the probability that a potential claim is filed as an actual claim to the particular insurer under consideration. In other words, $p(\theta)$ reflects the demand or the market share sensitivity to the loading parameter θ . $p(\theta)$ can be interpreted as a probability that a customer buys an insurance product. For example, we may assume that demand of insurance contracts is described by a logit glm model as in Hardin and Tabari [12]. Thus, $p(\theta)$, will be :

$$p(\theta) = \frac{1}{1 + e^{\beta_0 + \beta_1 \theta}}, \quad (3.1)$$

where β_0 and β_1 are determined from the glm and θ is the loading parameter. β_1 will be a positive number so $p \rightarrow 0$ when $\theta \rightarrow \infty$ and $p \rightarrow 1$ when $\theta \rightarrow -\infty$. Assuming that the company has some fixed costs, independent of the risk exposure, denoted by $r > 0$, the expression for the net premium income becomes:

$$c(\theta) = (1 + \theta) \mathbb{E}^\theta[N_1] \mathbb{E}[Y] - r.$$

The following proposition characterizes the behaviour of the solution of equation (2.1) with respect to the loading θ .

Proposition 3.1. *If $V(x, \theta)$ satisfies equation (2.1) then $V(x, \theta)$ is strictly increasing with respect to the parameter $\alpha = \frac{\mathbb{E}^\theta[N_1]}{c(\theta)}$.*

Proof. It is possible to integrate Equation (2.1) on the interval $]0, x]$ to obtain:

$$V(x, \theta) = \frac{\mathbb{E}^\theta[N_1]}{c(\theta)} \left(\mathbb{E}[Y] + \int_0^x \left(V(z, \theta) - \int_0^z V(z-y) dF(y) + F(z) - 1 \right) dz \right). \quad (3.2)$$

To prove the proposition, we will study equations of the general form:

$$u(x) = \alpha \left(g(x) + \int_0^x \left(u(z) - \int_0^z u(z-y) dF(y) \right) dz \right). \quad (3.3)$$

We introduce the operator Ψ , acting on measurable locally bounded functions $h : [0, +\infty[\mapsto \mathbb{R}$, as:

$$(\Psi h)(x) = \int_0^x \left(h(z) - \int_0^z h(z-y) dF(y) \right) dz, \quad x \geq 0 \quad (3.4)$$

Notice that the transformation $h \mapsto \Psi h$ is linear and for every h , $\Psi h : [0, +\infty[\mapsto \mathbb{R}$ is continuous, hence measurable and locally bounded. Thus, powers of the operator Ψ are defined in the usual way.

$$\Psi^0 h = h, \quad \Psi^n h = \Psi(\Psi^{n-1} h), \quad n \in \mathbb{N}.$$

Let, $\|h\|_{[0,x]} = \sup_{z \in [0,x]} |h(z)|$. Then:

$$|(\Psi h)(x)| \leq \int_0^x \left(|h(z)| + \int_0^z |h(z-y)| dF(y) \right) dz \leq 2x \|h\|_{[0,x]}.$$

If the inequality

$$\|(\Psi^n h)\|_{[0,x]} \leq \frac{2^n x^n}{n!} \|h\|_{[0,x]} \quad (3.5)$$

holds, for some $n \in \mathbb{N}$, then

$$\begin{aligned} |(\Psi^{n+1} h)(x)| &\leq \int_0^x \left(|(\Psi^n h)(z)| + \int_0^z |(\Psi^n h)(z-y)| dF(y) \right) dz \\ &\leq \int_0^x 2 \frac{2^n z^n}{n!} \|h\|_{[0,x]} dz = \frac{2^{n+1} x^{n+1}}{(n+1)!} \|h\|_{[0,x]}. \end{aligned}$$

Thus, by induction, (3.5) holds for every $n \in \mathbb{N}$. Therefore, for every $x \in [0, \infty[$, fixed, there is some $n \in \mathbb{N}$ such that Ψ^n is a contraction in the space of measurable and bounded functions $h : [0, x] \mapsto \mathbb{R}$. It follows from the contraction principle that equation (3.2) has one unique solution. Further, $\lim_{n \rightarrow \infty} (\alpha^n \Psi^n) h = 0$, uniformly in $[0, x]$ for any given h and any fixed $x \in [0, +\infty[$.

Let $u_{\alpha,g}$ be the solution of equation (3.3) for given g and α . Then,

$$\begin{aligned} u_{\alpha,g} &= \alpha(g + \Psi u_{\alpha,g}) = \alpha g + \alpha \Psi(\alpha(g + \Psi u_{\alpha,g})) = \alpha g + \alpha^2 \Psi g + \alpha^2 \Psi u_{\alpha,g} \\ &= \alpha g + \alpha^2 \Psi g + \dots + \alpha^{n+1} \Psi^n g + \alpha^{n+2} \Psi^{n+1} u_{\alpha,g}. \end{aligned}$$

Since $\lim_{n \rightarrow \infty} \alpha^n \Psi^n u_{\alpha,g}(x) = 0$, this shows that $u_{\alpha,g}$ admits the series representation:

$$u_{\alpha,g} = \sum_{n=0}^{\infty} \alpha^{n+1} \Psi^n g,$$

which converges uniformly with respect to α on compact intervals. Thus, we can differentiate term by term and

obtain

$$\begin{aligned}
\frac{d}{d\alpha} u_{\alpha,g}(x) &= \sum_{n=0}^{\infty} (n+1) \alpha^n (\Psi^n g)(x) \\
&= \sum_{n=0}^{\infty} \alpha^n (\Psi^n g)(x) + \sum_{n=1}^{\infty} n \alpha^n (\Psi^n g)(x) \\
&= \frac{1}{\alpha} u_{\alpha,g} + \sum_{n=1}^{\infty} \alpha^n (\Psi^n g)(x) + \sum_{n=2}^{\infty} (n-1) \alpha^n (\Psi^n g)(x) \\
&= \frac{1}{\alpha} u_{\alpha,g} + \sum_{n=0}^{\infty} \alpha^{n+1} (\Psi^{n+1} g)(x) + \sum_{n=1}^{\infty} n \alpha^{n+1} (\Psi^{n+1} g)(x) \\
&= \frac{1}{\alpha} u_{\alpha,g} + (\Psi u_{\alpha,g})(x) + \sum_{n=1}^{\infty} \alpha^{n+1} (\Psi^{n+1} g)(x) + \sum_{n=2}^{\infty} (n-1) \alpha^{n+1} (\Psi^{n+1} g)(x) \\
&= \frac{1}{\alpha} u_{\alpha,g} + (\Psi u_{\alpha,g})(x) + \sum_{n=0}^{\infty} \alpha^{n+2} (\Psi^{n+2} g)(x) + \sum_{n=1}^{\infty} n \alpha^{n+2} (\Psi^{n+2} g)(x) \\
&= \frac{1}{\alpha} u_{\alpha,g} + (\Psi u_{\alpha,g})(x) + (\alpha \Psi^2 u_{\alpha,g})(x) + \dots + (\alpha^{k-1} \Psi^k u_{\alpha,g})(x) + \sum_{n=1}^{\infty} n \alpha^{n+k} (\Psi^{n+k} g)(x) \\
&= \sum_{n=0}^{\infty} \alpha^{n-1} (\Psi^n u_{\alpha,g})(x) = \frac{1}{\alpha^2} u_{\alpha, u_{\alpha,g}}.
\end{aligned}$$

For any $h : [0, x] \mapsto \mathbb{R}$ locally absolutely continuous function:

$$\begin{aligned}
(\Psi h)(x) &= \int_0^x \left(h(z) - \int_0^z h(z-y) dF(y) \right) dz \\
&= \int_0^x \left(h(z) - [h(z-y)F(y)]_{y=0}^{y=z} - \int_0^z h'(z-y)F(y) dy \right) dz \\
&= \int_0^x (h(z) - h(0)F(z)) dz - \int_0^x \int_y^x h'(z-y)F(y) dz dy \\
&= \int_0^x (h(z) - h(0)F(z)) dz - \int_0^x (h(x-y) - h(0))F(y) dy \\
&= \int_0^x h(z) dz - \int_0^x h(x-y)F(y) dy = \int_0^x h(z)(1 + F(x-z)) dz.
\end{aligned}$$

Thus, $h > 0$ implies $(\Psi h) > 0$, which implies $(\Psi^n h) > 0, \forall n \in \mathbb{N}$, and therefore $u_{\alpha,h} > 0$ for any $\alpha > 0$. This argument shows that $\frac{d}{d\alpha} V = \frac{1}{\alpha^2} u_{\alpha,V} > 0$ as $V > 0$. Therefore V is strictly increasing with α . \square

According to Proposition 3.1, in order to find θ minimizing the probability of ruin, it is sufficient to find θ minimizing $\frac{\mathbb{E}^\theta[N_1]}{c(\theta)}$. For example, using the logit demand model (3.1), the optimal loading is found with direct differentiation of α and is given by:

$$\theta_{ruin} = \frac{1}{\beta_1} \left(\ln \left(\frac{\lambda \mathbb{E}[Y]}{r \beta_1} \right) - \beta_0 \right). \quad (3.6)$$

However, the loading that maximizes the expected profit is:

$$\theta_{profit} = \arg \max_{\theta} \mathbb{E}^\theta[X_1 | X_0 = x] = \arg \max_{\theta} \{ \theta \mathbb{E}^\theta[N_1] \mathbb{E}[Y] - r \},$$

which is, in the case of logit demand (3.1), the unique solution of:

$$1 + e^{\beta_0 + \beta_1 \theta} - \beta_1 \theta e^{\beta_0 + \beta_1 \theta} = 0. \quad (3.7)$$

Thus, in general, θ_{ruin} does not coincide with θ_{profit} .

4 The Multiple Risk Case

In this section, we explore how dependencies between risks available in an insurance market translate into risk exposure for a company through its market shares on the different risks. It turns out that this mechanism is non

trivial when the risks are dependent. For the sake of simplicity, we assume that the company offers insurance for two risks in a market constituted by identical individuals, all of them exposed to both risks. Using the notation in equations (2.5) and (2.6) to denote the market claim process, $S_t = (S_t^{(1)}, S_t^{(2)})$ is the vector of the total (accumulated) amount of claims of each risk that occurred in the market, up to time t . The marginal distributions of $S^{(1)}$ and $S^{(2)}$ are characterized by claim intensities λ_1 and λ_2 and the severity distributions $F_{Y^{(1)}}$, $F_{Y^{(2)}}$ and their dependency structure is characterized by a parameter $\lambda^\parallel \in [0, \min(\lambda_1, \lambda_2)]$ and a joint distribution $F_{(Y^{1\parallel}, Y^{2\parallel})}$, as explained in Section 2.

4.1 Risk Exposure as a Function of Market Shares

To extend the demand model outlined in Section 3 to a market with multiple risks where the acquisition of insurance for different risks may not be independent, we propose the following interpretation for the function p .

Let (θ_1, θ_2) be the loadings charged by the company for each risk. We assume that every individual in the market (a potential client) is provided with a vector of bid prices (b_1, b_2) . The client acquires the insurance for risk i if $b_i \geq \theta_i$ (for convenience, we consider prices net of the pure premium). The distribution of the price vectors in the market is modelled by a random vector $B = (B_1, B_2)$. Thus, $p_i(\theta) = p_i(\theta_i) = \mathbb{P}(B_i \geq \theta_i)$ is the company's market share for the insurance of risk i at equilibrium, given the loadings $\theta = (\theta_1, \theta_2)$. Let $p^{(1,0)}$ be the proportion of individuals in the market holding a policy for risk 1 and no policy for risk 2. Similarly, $p^{(0,1)}(\theta)$ and $p^{(1,1)}(\theta)$ denote the proportion of individuals holding a policy only for risk 2 and for both risks, respectively. If the acquisition of policies for different risks is independent, then:

$$p^{(1,1)}(\theta) = p_1(\theta_1)p_2(\theta_2), \quad p^{(1,0)}(\theta) = p_1(\theta_1)(1 - p_2(\theta_2)), \quad p^{(0,1)}(\theta) = p_2(\theta_2)(1 - p_1(\theta_1)). \quad (4.1)$$

Dependency between the acquisition of different risks can be introduced by considering dependent bid prices $B = (B_1, B_2)$. In particular, if the joint distribution of B is characterized by an ordinary copula $C : [0, 1]^2 \mapsto [0, 1]$, then, according to Sklar's theorem $F_B(\theta_1, \theta_2) = C(F_{B_1}(\theta_1), F_{B_2}(\theta_2))$ [25]. This gives:

$$\begin{aligned} p^{(1,0)} &= F_{B_2}(\theta_2^-) - C(F_{B_1}(\theta_1^-), F_{B_2}(\theta_2^-)), \\ p^{(0,1)} &= F_{B_1}(\theta_1^-) - C(F_{B_1}(\theta_1^-), F_{B_2}(\theta_2^-)), \\ p^{(1,1)} &= 1 - F_{B_1}(\theta_1^-) - F_{B_2}(\theta_2^-) + C(F_{B_1}(\theta_1^-), F_{B_2}(\theta_2^-)). \end{aligned}$$

Under this model, the company's surplus process is:

$$\tilde{X}_t = u^{(1)} + u^{(2)} + (c^{(1)}(\theta_1) + c^{(2)}(\theta_2))t - \sum_{i=0}^{\tilde{N}_t^{1\perp}} \tilde{Y}_i^{1\perp} - \sum_{i=0}^{\tilde{N}_t^{2\perp}} \tilde{Y}_i^{2\perp} - \sum_{i=0}^{\tilde{N}_t^\parallel} (\tilde{Y}_i^{1\parallel} + \tilde{Y}_i^{2\parallel}), \quad (4.2)$$

where $\tilde{N}_t^{1\perp}$, $\tilde{N}_t^{2\perp}$, and \tilde{N}_t^\parallel count the number of claims received by the company concerning only risk 1, only risk 2, and both risks, respectively. Their intensities are, respectively,

$$\begin{aligned} \tilde{\lambda}_1^\perp &= p^{(1,0)}(\theta)(\lambda_1^\perp + \lambda^\parallel) + p^{(1,1)}(\theta)\lambda_1^\perp = p_1(\theta_1)\lambda_1^\perp + p^{(1,0)}(\theta)\lambda^\parallel, \\ \tilde{\lambda}_2^\perp &= p_2(\theta_2)\lambda_2^\perp + p^{(0,1)}(\theta)\lambda^\parallel, \\ \tilde{\lambda}^\parallel &= p^{(1,1)}(\theta)\lambda^\parallel. \end{aligned}$$

The distribution of the single risk claim amounts $\tilde{Y}^{1\perp}$ (resp., $\tilde{Y}^{2\perp}$) is a mixture of the distributions $Y^{1\perp}$ and $Y^{1\parallel}$ (resp., $Y^{2\perp}$ and $Y^{2\parallel}$):

$$\begin{aligned} F_{\tilde{Y}^{1\perp}} &= \frac{p_1\lambda_1^\perp}{p_1\lambda_1^\perp + p^{(1,0)}\lambda^\parallel} F_{Y^{1\perp}} + \frac{p^{(1,0)}\lambda^\parallel}{p_1\lambda_1^\perp + p^{(1,0)}\lambda^\parallel} F_{Y^{1\parallel}} \\ F_{\tilde{Y}^{2\perp}} &= \frac{p_2\lambda_2^\perp}{p_2\lambda_2^\perp + p^{(0,1)}\lambda^\parallel} F_{Y^{2\perp}} + \frac{p^{(0,1)}\lambda^\parallel}{p_2\lambda_2^\perp + p^{(0,1)}\lambda^\parallel} F_{Y^{2\parallel}} \end{aligned}$$

This is because some customers insure risk 1, but not risk 2 and vice-versa. Therefore, the aggregate process for the insurer is

$$\tilde{X}_t = u^{(1)} + u^{(2)} + (c^{(1)}(\theta_1) + c^{(2)}(\theta_2))t - \sum_{i=0}^{\tilde{N}_t} \tilde{Y}_i, \quad (4.3)$$

where \tilde{N}_t is a Poisson process with intensity

$$\tilde{\lambda} = p_1\lambda_1^\perp + p_2\lambda_2^\perp + (p^{(1,0)} + p^{(0,1)} + p^{(1,1)})\lambda^\parallel = p_1\lambda_1 + p_2\lambda_2 - p^{(1,1)}\lambda^\parallel, \quad (4.4)$$

and $\tilde{Y}_i, i \in \mathbb{N}$ are i.i.d random variables with distribution

$$\begin{aligned} F_{\tilde{Y}} &= \frac{p_1 \lambda_1^\perp}{\tilde{\lambda}} F_{Y_{1\perp}} + \frac{p_2 \lambda_2^\perp}{\tilde{\lambda}} F_{Y_{2\perp}} + \frac{p^{(1,0)} \lambda^\parallel}{\tilde{\lambda}} F_{Y_{1\parallel}} + \frac{p^{(0,1)} \lambda^\parallel}{\tilde{\lambda}} F_{Y_{2\parallel}} + \frac{p^{(1,1)} \lambda^\parallel}{\tilde{\lambda}} F_{Y_{1\parallel} + Y_{2\parallel}} \\ &= \frac{1}{p_1 \lambda_1 + p_2 \lambda_2 - p^{(1,1)} \lambda^\parallel} \left(p_1 \lambda_1 F_{Y^1} + p_2 \lambda_2 F_{Y^2} + p^{(1,1)} \lambda^\parallel \left(F_{Y_{1\parallel} + Y_{2\parallel}} - F_{Y_{1\parallel}} - F_{Y_{2\parallel}} \right) \right). \end{aligned} \quad (4.5)$$

Thus, if the risks in the market are independent (i.e. if $\lambda^\parallel = 0$), then the risk in the company's portfolio is just a sum of the risks $S^{(1)}$ and $S^{(2)}$, weighted by the respective market shares, p_1 and p_2 , irrespective of any dependency between sales of policies for different risks. However, if the risks in the market are dependent ($\lambda^\parallel \neq 0$), then the company's risk is not, in general, a weighted sum of $S^{(1)}$ and $S^{(2)}$. Further, this effect persists even in the case where sales of different policies are independent (i.e., $p^{(1,1)} = p_1 p_2$). On the other hand, equalities (4.4) and (4.5) show that in the (unlikely) situation where clients always buy insurance for only one risk, the risk exposure of the insurer is accurately computed using only the marginal distributions of each risk (i.e. assuming that the risks are independent). This is due to the static nature of our model. For example, it does not take into account the possibility of external factors changing the frequency of claim events in both risks simultaneously.

4.2 The Impact of Dependencies on Ruin Probability

From the discussion above and Proposition 3.1, it follows that the ruin probability of a company with market shares $(p_1, p_2, p^{(1,1)})$ solves the equation

$$\frac{dV(x)}{dx} = \frac{\tilde{\lambda}}{c^{(1)} + c^{(2)}} \left(V(x) - \int_0^x V(x-y) dF_{\tilde{Y}}(y) + F_{\tilde{Y}}(x) - 1 \right), \quad (4.6)$$

$$V(0^+) = \frac{\tilde{\lambda}}{c^{(1)} + c^{(2)}} \mathbb{E}[\tilde{Y}], \quad (4.7)$$

with $\tilde{\lambda}$ and $F_{\tilde{Y}}$ given by equations (4.4) and (4.5).

Since estimating the dependency structure may pose substantial difficulties, we may wish to have an a-priori bound for the error introduced by neglecting dependencies, that is, by substituting the probability $V_{ind}(x)$ for $V(x)$, where $V_{ind}(x)$ solves the equation.

$$\frac{dV(x)}{dx} = \frac{\hat{\lambda}}{c^{(1)} + c^{(2)}} \left(V(x) - \int_0^x V(x-y) dF_{\hat{Y}}(y) + F_{\hat{Y}}(x) - 1 \right), \quad (4.8)$$

where $\hat{\lambda} = \lambda_1 p_1 + \lambda_2 p_2$ and $F_{\hat{Y}}(x) = \frac{\lambda_1 p_1 F_{Y^{(1)}} + \lambda_2 p_2 F_{Y^{(2)}}}{\hat{\lambda}}$. Notice that $\hat{\lambda} \mathbb{E}[\hat{Y}] = \tilde{\lambda} \mathbb{E}[\tilde{Y}]$ and therefore the boundary condition for (4.8) is again (4.7).

The discussion in Subsection 4.1 shows that the difference $V(x) - V_{ind}(x)$ is expected to be small when $p^{(1,1)}$ is small compared to $p_1 + p_2$. The following proposition gives a precise meaning for this statement.

Proposition 4.1. *With the notation above:*

$$|V(x) - V_{ind}(x)| \leq p^{(1,1)} \lambda^\parallel \frac{e^{\frac{2\tilde{\lambda}x}{c^{(1)} + c^{(2)}}} - 1}{\tilde{\lambda}}$$

for every amount of initial reserve $x \geq 0$.

Proof. From equalities (4.6), (4.7) and (4.8), straightforward computations yield:

$$\begin{aligned} V(x) - V_{ind}(x) &= \frac{p^{(1,1)} \lambda^\parallel}{c^{(1)} + c^{(1)}} \left(\int_0^x V_{ind}(z) - \int_0^z V_{ind}(z-y) dF_{Y_{1\parallel} + Y_{2\parallel}}(y) + F_{Y_{1\parallel} + Y_{2\parallel}}(z) - 1 dz - \right. \\ &\quad \left. \int_0^x V_{ind}(z) - \int_0^z V_{ind}(z-y) dF_{Y_{1\parallel}}(y) + F_{Y_{1\parallel}}(z) - 1 dz - \right. \\ &\quad \left. \int_0^x V_{ind}(z) - \int_0^z V_{ind}(z-y) dF_{Y_{2\parallel}}(y) + F_{Y_{2\parallel}}(z) - 1 dz - \right) \\ &\quad + \frac{\tilde{\lambda}}{c^{(1)} + c^{(2)}} \int_0^x (V - V_{ind})(z) - \int_0^z (V - V_{ind})(z-y) dF_{\tilde{Y}}(y) dz \end{aligned} \quad (4.9)$$

It can be checked that for every distribution function $G : [0, +\infty[\mapsto [0, 1]$,

$$-x \leq \int_0^x V_{ind}(z) - \int_0^z V_{ind}(z-y) dG(y) + G(z) - 1 dz \leq 0$$

Therefore, (4.9) implies:

$$\max_{y \in [0, x]} |V(x) - V_{ind}(x)| \leq \frac{p^{(1,1)} \lambda \parallel}{c^{(1)} + c^{(2)}} 2x + \frac{\tilde{\lambda}}{c^{(1)} + c^{(2)}} \int_0^x 2 \max_{y \in [0, z]} |V(x) - V_{ind}(y)| dz$$

Thus, the result follows by Grönwall's inequality [10]. \square

4.3 The Impact of Dependencies on Small Companies

Now, we proceed with the argument above to explore how dependencies affect companies of different size. We measure the size of the company by its expected total value of claims, $\tilde{\lambda} \mathbb{E}[\tilde{Y}]$ and, to make comparisons meaningful, we consider that the total revenue is proportional to the company's size, i.e.

$$c^{(1)} + c^{(2)} = (1 + \theta) \tilde{\lambda} \mathbb{E}[\tilde{Y}], \quad \text{with } \theta > 0 \text{ constant.}$$

Similarly, we consider the initial reserve to be proportional to size, i.e.:

$$x = x_0 \tilde{\lambda} \mathbb{E}[\tilde{Y}], \quad \text{with } x_0 > 0 \text{ constant.}$$

Notice that, due to equations (4.6), (4.7) and (4.8), the effect of dependencies must be bounded in the sense that

$$|V(x_0 \tilde{\lambda} \mathbb{E}[\tilde{Y}]) - V_{ind}(x_0 \tilde{\lambda} \mathbb{E}[\tilde{Y}])| \leq K_1 x_0 \tilde{\lambda} \mathbb{E}[\tilde{Y}] \leq K_2 (p_1 + p_2),$$

for some constants $K_1, K_2 < +\infty$. However, we can use Proposition 4.1 to obtain a better estimate:

$$|V(x_0 \tilde{\lambda} \mathbb{E}[\tilde{Y}]) - V_{ind}(x_0 \tilde{\lambda} \mathbb{E}[\tilde{Y}])| \leq p^{(1,1)} \lambda \parallel \frac{e^{\frac{2\tilde{\lambda} x_0}{1+\theta}} - 1}{\tilde{\lambda}}. \quad (4.10)$$

Notice that the right-hand side of (4.10) has the same asymptotic behaviour as

$$p^{(1,1)} \lambda \parallel \frac{x_0}{1 + \theta}, \quad \text{when } p_1 + p_2 \rightarrow 0.$$

Further, if the sales of policies for different risks to the same individual are independent, then $p^{(1,1)} = p_1 p_2$ goes to zero faster than $\tilde{\lambda} \mathbb{E}[\tilde{Y}] = p_1 \lambda_1 \mathbb{E}[Y^{(1)}] + p_2 \lambda_2 \mathbb{E}[Y^{(2)}]$, when $p_1 + p_2 \rightarrow 0$. Thus, a small company selling policies for different risks independently is relatively immune to the effects of dependencies between the risks, contrary to a large company (it is obvious that a monopolistic company is fully exposed to the dependencies between risks). This immunity to risk's dependencies may persist even when sales of policies for different risks are not independent, provided the dependency in sales is sufficiently mild. For example, $\lim_{p_1 + p_2 \rightarrow 0} \frac{p^{(1,1)}}{p_1 + p_2} = 0$ if the dependency between sales is modelled by a Clayton or a Frank copula in (4.1). However, small companies are not specially protected from risk dependencies if the dependency between sales is modelled by a Pareto or a Gumbel copula.

4.4 Optimal Loadings and Market Shares

Since the right-hand sides of equalities (4.6) and (4.7) depend on the loadings through both $\frac{\mathbb{E}[\tilde{N}_1]}{c^{(1)} + c^{(2)}}$ and $F_{\tilde{Y}}$, Proposition 3.1 can not be generalized to models with multiple risks. However, it is possible to provide optimality conditions for the loadings $\theta = (\theta_1, \theta_2)$ minimizing the ruin probability.

To do this, we extend the notation introduced in the proof of Proposition 3.1. For any distribution function $G : [0, +\infty[\mapsto [0, 1]$, we consider the compounding operator of type (3.4)

$$(\Psi_G h)(x) = \int_0^x \left(h(z) - \int_0^z h(z-y) d\theta(y) \right) dz, \quad x \geq 0.$$

Thus, the 2-risk version of equation (3.2) can be written as

$$V_\theta(x) = \frac{\tilde{\lambda}_\theta}{c(\theta)} \left(\int_x^\infty 1 - F_\theta(z) dz + (\Psi_{F_\theta} V_\theta)(x) \right), \quad (4.11)$$

where

$$F_\theta = \frac{\lambda_1^\perp}{\lambda_\theta} p_1(\theta) F_{Y_{1\perp}} + \frac{\lambda_2^\perp}{\lambda_\theta} p_2(\theta) F_{Y_{2\perp}} + \frac{\lambda^\parallel}{\lambda_\theta} p^{(1,0)}(\theta) F_{Y_{1\parallel}} + \frac{\lambda^\parallel}{\lambda_\theta} p^{(0,1)}(\theta) F_{Y_{2\parallel}} + \frac{\lambda^\parallel}{\lambda_\theta} p^{(1,1)}(\theta) F_{Y_{1\parallel} + Y_{2\parallel}}.$$

Since $\frac{\lambda_1^\perp}{\lambda_\theta} p_1(\theta) + \frac{\lambda_2^\perp}{\lambda_\theta} p_2(\theta) + \frac{\lambda^\parallel}{\lambda_\theta} p^{(1,0)}(\theta) + \frac{\lambda^\parallel}{\lambda_\theta} p^{(0,1)}(\theta) + \frac{\lambda^\parallel}{\lambda_\theta} p^{(1,1)}(\theta) = 1$, (4.11) becomes

$$\begin{aligned} V_\theta(x) &= \frac{\lambda_1^\perp p_1(\theta)}{c(\theta)} \left(\int_x^\infty 1 - F_{Y_{1\perp}}(z) dz + (\Psi_{F_{Y_{1\perp}}} V_\theta)(x) \right) \\ &+ \frac{\lambda_2^\perp p_2(\theta)}{c(\theta)} \left(\int_x^\infty 1 - F_{Y_{2\perp}}(z) dz + (\Psi_{F_{Y_{2\perp}}} V_\theta)(x) \right) \\ &+ \frac{\lambda^\parallel p^{(1,0)}(\theta)}{c(\theta)} \left(\int_x^\infty 1 - F_{Y_{1\parallel}}(z) dz + (\Psi_{F_{Y_{1\parallel}}} V_\theta)(x) \right) \\ &+ \frac{\lambda^\parallel p^{(0,1)}(\theta)}{c(\theta)} \left(\int_x^\infty 1 - F_{Y_{2\parallel}}(z) dz + (\Psi_{F_{Y_{2\parallel}}} V_\theta)(x) \right) \\ &+ \frac{\lambda^\parallel p^{(1,1)}(\theta)}{c(\theta)} \left(\int_x^\infty 1 - F_{Y_{1\parallel} + Y_{2\parallel}}(z) dz + (\Psi_{F_{Y_{1\parallel} + Y_{2\parallel}}} V_\theta)(x) \right). \end{aligned}$$

We write this in abbreviated form:

$$V_\theta(x) = \langle \alpha(\theta), \Gamma(x) \rangle + \langle \alpha(\theta), \Psi \rangle V_\theta(x),$$

where $\alpha(\theta)$ is the vector

$$\alpha(\theta) = \frac{1}{c(\theta)} \left(\lambda_1^\perp p_1(\theta), \lambda_2^\perp p_2(\theta), \lambda^\parallel p^{(1,0)}, \lambda^\parallel p^{(0,1)}, \lambda^\parallel p^{(1,1)} \right),$$

$\Gamma(x)$ is the vector function

$$\Gamma(x) = \left(\int_x^\infty 1 - F_{Y_{1\perp}}(z) dz, \int_x^\infty 1 - F_{Y_{2\perp}}(z) dz, \int_x^\infty 1 - F_{Y_{1\parallel}}(z) dz, \int_x^\infty 1 - F_{Y_{2\parallel}}(z) dz, \int_x^\infty 1 - F_{Y_{1\parallel} + Y_{2\parallel}}(z) dz \right),$$

Ψ is the vector of operators

$$\Psi = \left(\Psi_{F_{Y_{1\perp}}}, \Psi_{F_{Y_{2\perp}}}, \Psi_{F_{Y_{1\parallel}}}, \Psi_{F_{Y_{2\parallel}}}, \Psi_{F_{Y_{1\parallel} + Y_{2\parallel}}} \right),$$

and $\langle \cdot, \cdot \rangle$ is the usual inner product in \mathbb{R}^5 .

Using the argument in the proof of Proposition 3.1, we see that V_θ admits the series representation

$$V_\theta(x) = \sum_{n=0}^{\infty} \left(\langle \alpha(\theta), \Psi \rangle^n \langle \alpha(\theta), \Gamma \rangle \right)(x).$$

Similarly, any vector $\gamma \in \mathbb{R}^5$ and any bounded measurable function $g : [0, +\infty[\mapsto \mathbb{R}$ define one unique function

$$u_{\gamma, g}(x) = \sum_{n=0}^{\infty} \left(\langle \gamma, \Psi \rangle^n g \right)(x).$$

This function is analytic with respect to γ , with partial derivatives

$$\frac{\partial u_{\gamma, g}}{\partial \gamma_i} = \sum_{n=0}^{\infty} \langle \gamma, \Psi \rangle^n (\Psi_i u_{\gamma, g}) = u_{\gamma, \Psi_i u_{\gamma, g}}, \quad i = 1, \dots, 5.$$

Taking into account the chain rule for derivatives, this proves the following proposition.

Proposition 4.2. *If $\theta \mapsto \alpha(\theta)$ is differentiable, then $\theta \mapsto V_\theta(x)$ is differentiable for every $x \geq 0$ and*

$$\frac{\partial}{\partial \theta_i} V_\theta(x) = \sum_{j=1}^5 u_{\alpha(\theta), (\Gamma_j + \Psi_j V_\theta)} \frac{\partial \alpha_j(\theta)}{\partial \theta_i}, \quad i = 1, 2.$$

By Proposition 4.2, the optimal loadings satisfy the equation

$$\sum_{j=1}^5 u_{\alpha(\theta), (\Gamma_j + \Psi_j V_\theta)} \frac{\partial \alpha_j(\theta)}{\partial \theta_i} = 0, \quad i = 1, 2 \quad (4.12)$$

Contrary to the single-risk case, the odds of finding explicit solutions for this equation seem very low, even in simple cases. However, (4.12) can be numerically solved by Newton's algorithm, the second-order partial derivatives being

$$\frac{\partial^2}{\partial \theta_i \partial \theta_j} V_\theta(x) = \sum_{k=1}^5 u_{\alpha(\theta), (\Gamma_k + \Psi_k V_\theta)} \frac{\partial^2 \alpha_k(\theta)}{\partial \theta_i \partial \theta_j} + \sum_{k=1}^5 \sum_{l=1}^5 u_{\alpha(\theta), \Psi_k} u_{\alpha(\theta), (\Gamma_l + \Psi_l V_\theta)} \frac{\partial \alpha_k(\theta)}{\partial \theta_i} \frac{\partial \alpha_l(\theta)}{\partial \theta_j}.$$

Notice that the expected profit is

$$c^{(1)}(\theta) + c^{(2)}(\theta) - \tilde{\lambda} \mathbb{E}[\tilde{Y}] = \theta_1 p_1(\theta_1) \lambda_1 \mathbb{E}[Y^{(1)}] + \theta_2 p_2(\theta_2) \lambda_2 \mathbb{E}[Y^{(2)}].$$

Thus, it depends only on the marginal distribution of the claim processes $S^{(1)}$, $S^{(2)}$, being independent of the dependency structure. It follows that the loadings minimizing the joint profit coincide with the loadings minimizing the profit on each risk, separately. That is, a pricing strategy that completely focus on expected profit completely fails to take both dependencies between risks and dependencies between sales of policies into account.

5 Numerical Results

Throughout this section, $Y_i^{(i)}$ are assumed to be i.i.d gamma distributed random variables with shape parameter, $a^{(i)}$, and scale parameters, $k^{(i)}$, which means that the mean is, $\mathbb{E}[Y^{(i)}] = a^{(i)} k^{(i)}$, for $i = 1, 2$. In the following numerical analysis let $a^{(1)} = a^{(2)} = 2$, $k^{(1)} = k^{(2)} = 500$, $\lambda^{(1)} = \lambda^{(2)} = 800$, $\beta_0^{(1)} = \beta_0^{(2)} = -0.5$, $\beta_1^{(2)} = 4$ and $\beta_1^{(1)} = 4.5$. That is, the difference stems from surplus process 2 being more sensitive to the loading via the parameter $\beta_1^{(2)}$. $r^{(i)}$ is taken to be 20% of the pure premium if the exposure was 40%, that is $r^{(i)} = 0.4 * 0.2 k^{(i)} a^{(i)} N^{(i)}$. The operational cost is therefore 8% of the expected total amount of claims in the market. The Clayton Lévy copula is considered for positive dependence and the parameter is set to $\omega = 1$. Finally, let θ_{ruin}^* and θ_{profit}^* denote the optimal loading when the ruin probability and expected profit criterion is used, respectively. The programming language R was used for every calculation.

5.1 Single Surplus Process

The surplus processes are first considered separately. The ruin probability and the expected profit is plotted as a function of θ for the two processes in Figures 1 and 2. θ_{ruin}^* was found by minimizing α .

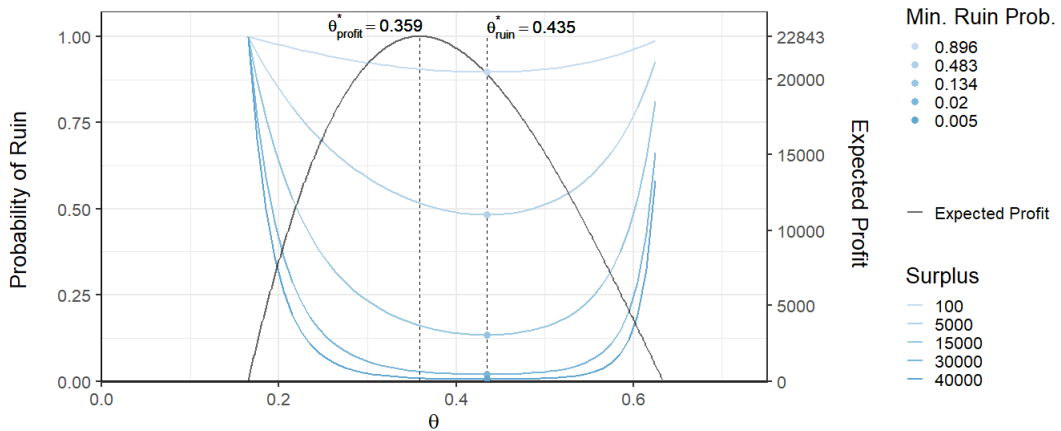


Figure 1: Surplus process 1. The blue lines show the ruin probability as a function of θ for a given surplus x . The black line shows the expected profit per time unit as a function of θ . The blue dots show the minimum ruin probability for each surplus. θ_{profit}^* and θ_{ruin}^* denote the optimal security loading parameter for the expected profit and for the probability of ruin, respectively.

From Figure 1 it can be seen that the optimal security loading parameter for the ruin probability is, $\theta_{ruin}^* = 0.435$, while the θ that maximizes the expected profit is lower, $\theta_{profit}^* = 0.359$. Moreover, in this example, the

maximum expected profit is 22.843 units and is given at θ_{profit}^* . The expected profit taken at the point θ_{ruin}^* is lower, close to 20.000 units.

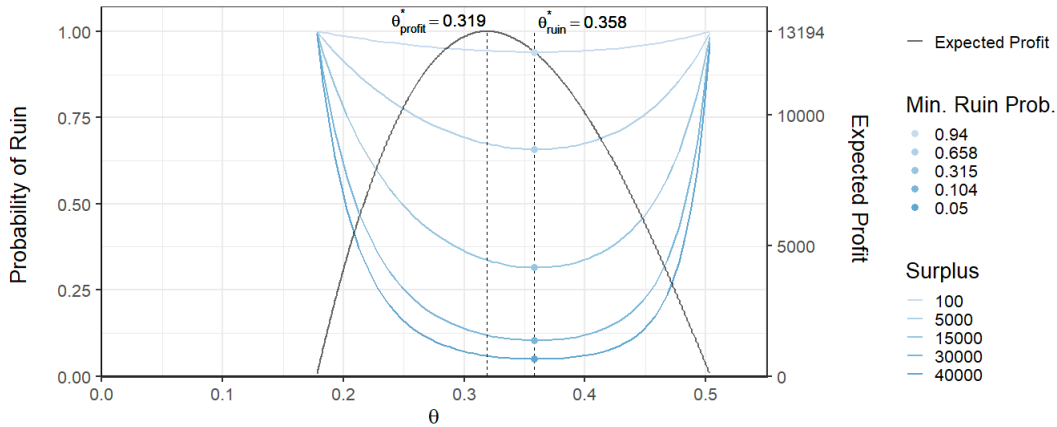


Figure 2: Surplus process 2. The blue lines show the ruin probability as a function of θ for a given surplus x . The black line shows the expected profit per time unit as a function of θ . The blue dots show the minimum ruin probability for each surplus. θ_{profit}^* and θ_{ruin}^* denote the optimal security loading parameter for the expected profit and for the probability of ruin, respectively.

From Figure 2 it can be seen that the optimal security loading parameter for the ruin probability is $\theta_{ruin}^* = 0.358$, while the θ that maximizes the expected profit is again lower or $\theta_{profit}^* = 0.319$.

Obviously, for both processes, the ruin probability decreases with increasing surplus. Moreover, it can be seen that surplus process X_2 has higher probability of ruin than surplus process X_1 for the same amount of surplus. The sensitivity of the demand curve affects the ruin probability and θ_{ruin}^* greatly. The more sensitive to the exposure the demand curve is, the closer the θ_{profit}^* and θ_{ruin}^* are. This more sensitive curve also has higher probability of ruin for a given surplus, which indicates that more competitive insurance products are riskier. These effects can be seen if the two Figures (1 and 2) are compared. Conversely, if the demand curve is not sensitive to the price, then the gap between θ_{profit}^* and θ_{ruin}^* can become quite large. Additionally, it can be seen from the curve at surplus = 100 that the ruin probability for θ_{profit}^* and θ_{ruin}^* are similar but as the surplus grows the values start to differ and once the surplus is great enough the two values θ_{profit}^* and θ_{ruin}^* result in similar ruin probabilities again. This means that if the insurance firm has high enough surplus then they can choose arbitrary θ without risking the chance of ruin. If the surplus is great enough then the value of θ does not matter as much. However, having too much reserves can be bad for insurance companies as it can be seen as a negative leverage. The bowl shape of the blue curves in the two Figures (1 and 2) is because of the interplay between the fixed cost and the demand curve.

θ_{ruin}^* should give the minimum ruin probability at all surplus values. This can be tested by graphing multiple ruin probability curves and compare it with the one obtained by θ_{ruin}^* . Figure 3 shows that θ_{ruin}^* gives the minimum ruin probability indeed.

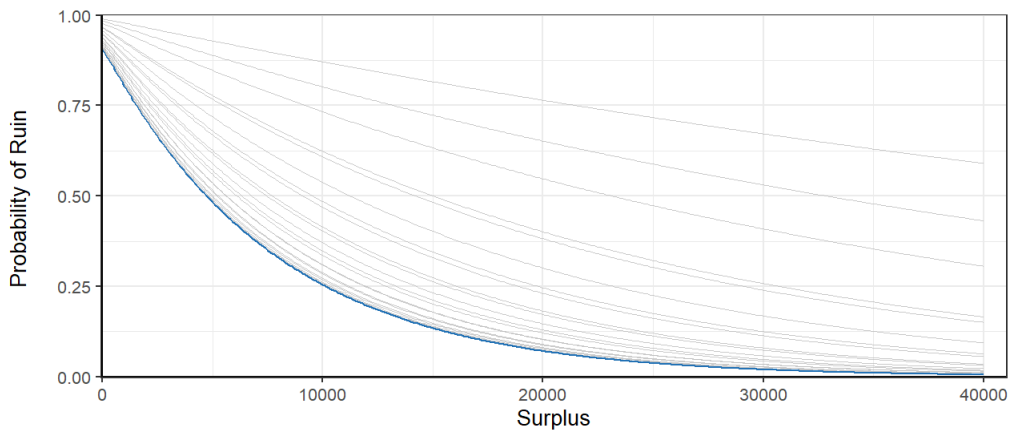


Figure 3: The figure compares the optimal value function of surplus process, X_1 (blue line) to other cost functions (grey lines). The blue line is achieved by setting $\theta = \theta_{ruin}^*$. All the cost functions lie above the optimal value function, as is expected.

5.2 Two Aggregated Surplus Processes with Common Loading

Next, the two surplus processes, X_1 and X_2 are aggregated, both when the claims are independent and dependent. The acquisition is independent in this subsection.

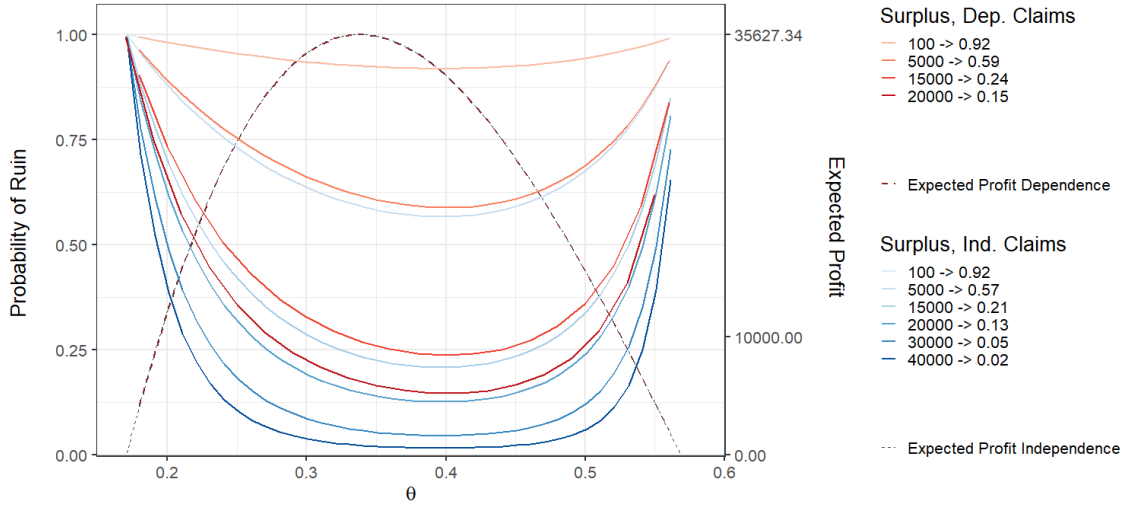


Figure 4: Ruin probability when X_1 and X_2 are aggregated as a function of the security loading parameter, θ , both when they are independent and dependent via Clayton Lévy copula with $\omega = 0.5$. The blue curves show the ruin probability when the two processes are independent for different values of the surplus and the red curves show the same for the dependent case. The curves have similar shapes, but the ruin probability is higher in the case of dependence, for the same surplus. The values in the legend show the minimum ruin probability for a given surplus (surplus \rightarrow probability).

Figure 4 shows the ruin probability of the aggregated surplus process as a function of the security loading parameter, θ , both when they are independent and dependent via Clayton Lévy copula. The red curves represent dependence while the blue curves represent independence.

Firstly, it can be seen that the expected profit is the same for dependence and independence and from the figure, $\theta_{profit}^* \approx 0.34$. The reason is that the claim mean and the claim frequency is almost the same (numerically) for dependence and independence.

Secondly, the dependent case has a higher probability of ruin than the independent case for the same amount of surplus. However, the ruin probability is almost the same for small surplus values as can be seen from the figure. Interestingly, the optimal loading for dependence and independence seem to be the same and numerically the values are $\theta_{ruin,dep}^* = 0.4 = \theta_{ruin,indp}^*$. The surplus value does not change the optimal loading θ^* , as expected. The reason why the ruin probability difference between the dependent and independent cases is relatively small is because of the probability $p^{(0,1)}(\theta)$. The fact that the insurance company does not always have the both claims $Y^{1||}$ and $Y^{2||}$ when a common jump occurs reduces the risk.

Finally, the difference of the two ruin probability curves (red and blue) for a given surplus seems to be increasing with increasing surplus, meaning that the ruin probability in the independent case decreases more rapidly with increasing surplus than for the dependent case. Therefore, it is clear that the positive dependent case is riskier.

Note that $\theta_{ruin}^* \approx 0.4$, which is very close to the weighted average of the optimal loading parameter of the isolated surplus processes where the weight is the exposure ratio of each surplus process, that is

$$\theta_{weighted} = \frac{0.435 \frac{1}{1+\exp(-0.6+4*0.4)} + 0.358 \frac{1}{1+\exp(-0.6+4.5*0.4)}}{\frac{1}{1+\exp(-0.6+4*0.4)} + \frac{1}{1+\exp(-0.6+4.5*0.4)}} \approx 0.4,$$

which strongly indicates that the optimal value, θ_{ruin}^* , is simply the weighted average.

5.2.1 Two Aggregated Surplus Processes with Separate Loadings

It is more realistic to consider θ as a vector so that the loading parameter can be different for each surplus process separately, to spread the total premium over the policies in an optimal way. The two surplus processes, X_1 and X_2 , are aggregated as before and the constants are the same, but let $\theta = (\theta^{(1)}, \theta^{(2)})$.

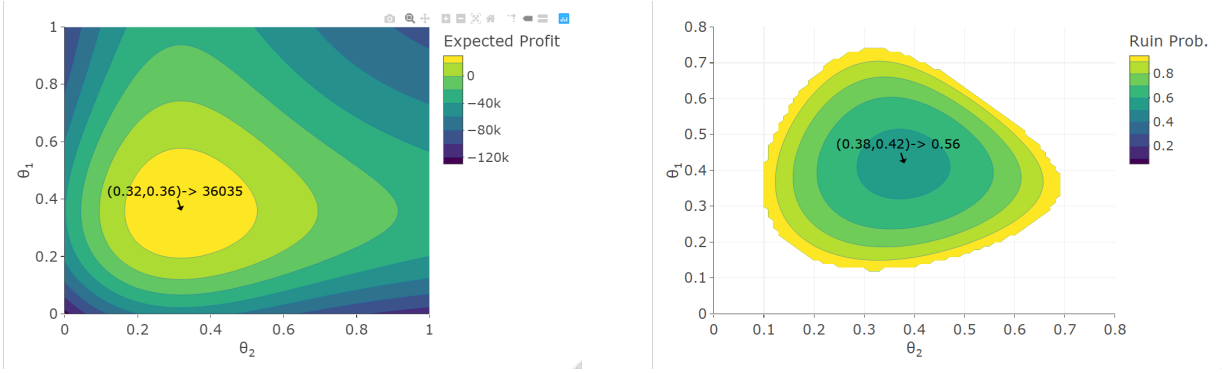


Figure 5: Expected profit (left) and the ruin probability (right) when X_1 and X_2 are aggregated, as a function of the security loading parameters, $\theta^{(1)}$ and $\theta^{(2)}$. The processes are assumed to be independent and the surplus is fixed at $x = 5000$. The parenthesis in the right figure shows the optimal values of $\theta^{(1)}$ and $\theta^{(2)}$ with the ruin probability as a criterion. The arrow indicates which values $\theta^{(1)}$ and $\theta^{(2)}$ are mapped into, thus showing the minimum ruin probability. The parenthesis in the left figure shows the same for the expected profit. The shape of the contour plot is due to the fact that the θ grid considered is sparser for values that give high ruin probability.

Figure 5 shows the expected profit (left) and the ruin probability (right), when X_1 and X_2 are assumed to be independent and aggregated, as a function of the security loading parameters, $\theta^{(1)}$ and $\theta^{(2)}$. The surplus is fixed at $x = 5000$ and the optimal values are shown. It should be noted that many surplus values were tested and they all gave the same value for $\theta_{ruin}^{*(1)}$, $\theta_{ruin}^{*(2)}$, $\theta_{profit}^{*(1)}$, and $\theta_{profit}^{*(2)}$ as shown, only the ruin probability level changed. Note that the optimal loading parameters for the expected profit are the same as those for the individual surplus processes. However, the optimal loading parameters for the ruin probability change when compared to the individual one (compare it with Figures 1 and 2). When compared to the optimal loading parameter for the individual surplus process, $\theta^{(1)}$ decreases from 0.435 to 0.42 and $\theta^{(2)}$ increases from 0.358 to 0.38. Therefore, the optimal security loading parameter decision is to decrease the loading parameter of the less sensitive surplus process while increasing the loading parameter of the more sensitive surplus process. Additionally, when compared to Figure 4, the minimum ruin probability for one shared loading is 0.57 while the ruin for two loadings is 0.56, showing only a marginal difference. When the same is done for other surplus values a similar difference is found. The expected profit is marginally higher.

Lastly, consider the case when the surplus processes are assumed to be dependent via Lévy Clayton copula and the loadings can be different for each surplus process separately. Figure 6 shows the ruin probability when X_1 and X_2 are aggregated as a function of the security loading parameters, $\theta^{(1)}$ and $\theta^{(2)}$. The shape of the contour plots is due to the fact that the θ grid considered is sparser for values that give high ruin probability. The surplus is fixed at $x = 5000$.

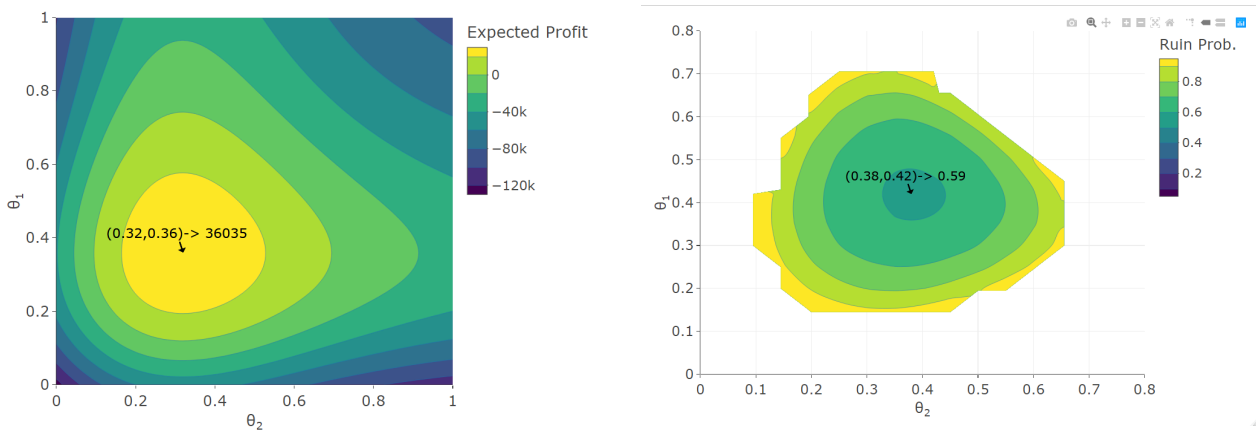


Figure 6: Expected profit (left) and the ruin probability (right) when X_1 and X_2 are aggregated, as a function of the security loading parameters, $\theta^{(1)}$ and $\theta^{(2)}$. The processes are assumed to be dependent via Clayton Lévy copula and the surplus is fixed at $x = 5000$. The parenthesis in the right figure shows the optimal values of $\theta^{(1)}$ and $\theta^{(2)}$ with the ruin probability as a criterion. The arrow shows which values $\theta^{(1)}$ and $\theta^{(2)}$ are mapped into, thus showing the minimum ruin probability. The parenthesis in the left figure shows the same except for the expected profit. The shape of the contour plot is due to the fact that the θ grid considered is sparser for values that give high ruin probability.

It can be seen that the optimal loadings $\theta^{(1)}$ and $\theta^{(2)}$ are the same as the ones in the case of independence and the minimum ruin probability is higher (compared to the case in Figure 5). Both the values and the optimal

loadings of the expected profit are the same as the independent case. Again, the optimal security loading parameter decision is to decrease the loading parameter of the less sensitive surplus process while increasing the loading parameter of the more sensitive surplus process. The difference between the ruin probability in Figure 6 vs Figure 5 is only 0.03 but in this case the surplus is low compared to the expected profit. If the surplus would be increased to ≈ 20.000 the difference would become greater. The difference would then decrease again if the surplus were increased to ≈ 40.000 .

Additionally, when compared to Figure 4, the minimum ruin probability for one common loading is 0.59, which is the same as the ruin probability for separate loading selections, therefore the difference is only marginal.

5.3 Dependent Claims and Dependent Acquisition

It is time to look at the case when we have dependent claims and dependent acquisition. Note that the case when we have independent claims and dependent acquisition is the same as the total independence case. We will look both at the case when the acquisition is modelled with a Gumbel and Clayton dependency structure. To compare these two structures we use Kendall's tau. The following equations relate the copula parameters, $\omega_{clayton}$ and ω_{gumbel} to kendell's tau, τ .

$$\omega_{clayton} = \frac{2\tau}{1-\tau}, \quad \omega_{gumbel} = \frac{1}{1-\tau}.$$

We know that the expected profit is the same as before for all values of τ . Therefore, we analyze the ruin probability.

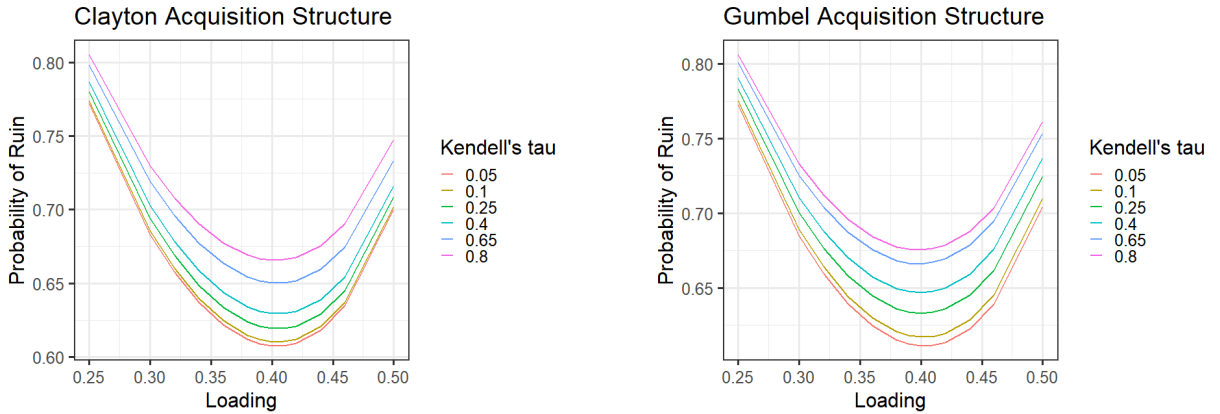


Figure 7: The ruin probability when the acquisition is modelled with a Clayton (left) and a Gumbel (right) dependency structure. The surplus is constant. In both cases, the ruin probability is higher for higher kendell's tau. The Gumbel case is more riskier. The surplus is fixed at 5000 units.

In Figure 7 we can see the ruin probability for different dependency values when the surplus is fixed at 5000 units. We can see that the ruin probability is higher for more dependent acquisition, as we expected. Also, we can see that the Gumbel acquisition model gives higher ruin probabilities than the Clayton model for the same Kendall's tau value. They give the same optimal loading parameter. It can also be seen that when kendell's tau is 0.05 (close to 0) the ruin probability is close to the case of independent acquisition, as expected. The optimal loading parameter is the same for all dependency levels.

A Numerical scheme for equation 2.2, using linear approximation

Consider the process

$$X_t = u + ct - \sum_{i=0}^{N_t} Y_i$$

where Y_i are iid continuous random variables and N_t is a $Poisson(\lambda t)$. To approximate equation 2.2, take a grid of points $\epsilon = x_0 < x_1 < \dots < x_n$, $x_i \in \mathbb{R}$, $\forall i \in \mathbb{N}$, $\epsilon > 0$, with equal interval lengths, $h = x_i - x_{i-1}$. A linear approximation is used to approximate $\bar{V}(x)$

$$\bar{V}(z) = \bar{V}(x_{j-1}) + \frac{\bar{V}(x_j) - \bar{V}(x_{j-1})}{h}(z - x_{j-1}), \quad z \in [x_{j-1}, x_j], j \leq i$$

where $\frac{\bar{V}(x_j) - \bar{V}(x_{j-1})}{h}$ is an approximation of the derivative, $\bar{V}'(x_{j-1})$, using the so-called forward difference. Let \bar{V}_i denote the approximation of $\bar{V}(x_i)$. Let $\bar{S}(x) = \int_0^x \bar{F}(y)dy$ and $\bar{\bar{S}}(x) = \int_0^x \bar{\bar{S}}(y)dy$. For each $x_i, i > 0$ solve the following equation

$$\bar{V}(x_i) - \bar{V}(0^+) = \frac{\lambda}{c} \int_0^x \bar{V}(x_i - y) \bar{F}(y) dy.$$

The goal is to develop a recursive method from x_0 as the value of \bar{V}_0 is known.

if $i = 0$

Set $\bar{V}_0 = 1 - \frac{\lambda}{c} \mathbb{E}[Y]$

if $i = 1$

Calculate

$$\begin{aligned} \bar{V}_1 &= \bar{V}_0 + \frac{\lambda}{c} \int_0^{x_1} ((\bar{V}_0 + \frac{\bar{V}_1 - \bar{V}_0}{h})(x_1 - y - x_0)) \bar{F}(y) dy \\ &= \bar{V}_0 + \frac{\lambda}{c} (\bar{V}_0 (\bar{S}(x_1) - \bar{S}(x_0)) + \\ &\quad \frac{\bar{V}_1 - \bar{V}_0}{h} ([x_1 - y - x_0]_{x_0}^{x_1} + \bar{\bar{S}}(x_1) - \bar{\bar{S}}(x_0))) \\ &= \bar{V}_0 + \frac{\lambda}{c} (a_{1,1} + \frac{\bar{V}_1 - \bar{V}_0}{h} a_{2,1}) \\ &\Leftrightarrow \\ (1 - \frac{\lambda a_{2,1}}{ch}) \bar{V}_1 &= \bar{V}_0 + \frac{\lambda}{c} \bar{V}_0 (a_{1,1} - \frac{a_{2,1}}{h}) \end{aligned}$$

if $i > 1$

Calculate

$$\begin{aligned} \bar{V}(x_i) - \bar{V}(0) &= \frac{\lambda}{c} \left(\sum_{j=2}^i \int_{x_{j-1}}^{x_j} \bar{V}(x_i - y) S(y) dy + \int_{x_{i-1}}^{x_i} \bar{V}(x_i - y) S(y) dy \right) \\ &= \frac{\lambda}{c} \left(\sum_{j=2}^i (\bar{V}_{i-j} (\bar{S}(x_j) - \bar{S}(x_{j-1}))) + \frac{\bar{V}_{i-j+1} - \bar{V}_{i-j}}{h} ([x_i - y - x_{i-j}]_{x_{j-1}}^{x_j} + \bar{\bar{S}}(x_j) - \bar{\bar{S}}(x_{j-1}))) + \right. \\ &\quad \left. \bar{V}_{i-j} (\bar{S}(x_i) - \bar{S}(x_{i-1})) + \frac{\bar{V}_i - \bar{V}_{i-j}}{h} ([x_i - y - x_{i-j}]_{x_{i-1}}^{x_i} + \bar{\bar{S}}(x_i) - \bar{\bar{S}}(x_{i-1}))) \right) \\ &= \frac{\lambda}{c} \left(\sum_{j=2}^i (\bar{V}_{i-j} (\bar{S}(x_j) - \bar{S}(x_{j-1}))) + \frac{\bar{V}_{i-j+1} - \bar{V}_{i-j}}{h} ([x_i - y - x_{i-j}]_{x_{j-1}}^{x_j} + \bar{\bar{S}}(x_j) - \bar{\bar{S}}(x_{j-1}))) + \right. \\ &\quad \left. \bar{V}_{i-j} a_{1,i} + \frac{\bar{V}_i - \bar{V}_{i-j}}{h} a_{2,i} \right) \end{aligned}$$

\Leftrightarrow

$$\begin{aligned} (1 - \frac{\lambda a_{2,i}}{ch}) \bar{V}_i &= \bar{V}_0 + \\ &\frac{\lambda}{c} \left(\sum_{j=2}^i (\bar{V}_{i-j} (\bar{S}(x_j) - \bar{S}(x_{j-1}))) + \frac{\bar{V}_{i-j+1} - \bar{V}_{i-j}}{h} ([x_i - y - x_{i-j}]_{x_{j-1}}^{x_j} + \bar{\bar{S}}(x_j) - \bar{\bar{S}}(x_{j-1}))) + \right. \\ &\quad \left. \bar{V}_{i-j} a_{1,i} - \frac{\bar{V}_{i-j}}{h} a_{2,i} \right) \end{aligned}$$

: Estimation of $\bar{V}(x)$

Let the symbol $\vec{}$ denote a vector.

Initialize

\vec{x} for some x_0, \dots, x_n

\vec{V} with length equal to the length of \vec{x}

$N_x \leftarrow$ length of \vec{x}

loop to estimate each value in \vec{V}

for i in $0, \dots, (N_x - 1)$ **do**

if $i = 0$ **then**

$\bar{V}_0 \leftarrow 1 - \frac{\lambda}{c} \mathbb{E}[Y]$

else if $i = 1$ **then**

$\bar{V}_1 \leftarrow$ take case $i = 1$ from above and isolate \bar{V}_1

else

$\bar{V}_i \leftarrow$ take case $i > 1$ from above and isolate \bar{V}_i

end

end

Return \vec{V}, \vec{x}

Acknowledgements

The second and third authors acknowledge financial support from FCT – Fundação para a Ciência e Tecnologia (Portugal), national funding, through research grant UIDB/05069/2020.

References

- [1] Benjamin Avanzi, Luke C. Cassar, and Bernard Wong. Modelling dependence in insurance claims processes with lévy copulas. *ASTIN Bulletin*, 41(2):575–609, 2011.
- [2] Frank Barning. Counting processes and copulas: Applications in insurance. Master’s thesis, Université du Québec à Montréal, 2018.
- [3] Fred Espen Benth, Valery A Kholodnyi, and Peter Laurence. Quantitative energy finance. In *Modelling, pricing, and hedging in energy and commodity markets*. Springer, 2014.
- [4] Stephen Britt and Albert Napoli. Linear correlation as a measure of dependency. In *XVth General Insurance Seminar, Institute of Actuaries of Australia*, 2005.
- [5] Nicole Bäuerle and Anja Blatter. Optimal control and dependence modeling of insurance portfolios with lévy dynamics. *Insurance: Mathematics and Economics*, 48(3):398–405, 2011.
- [6] Hans Bühlmann. Premium calculation from top down. *Astin Bulletin - ASTIN BULL*, 15:89–101, 11 1985.
- [7] Antonella Campana, Paola Ferretti, et al. Distortion risk measures and discrete risks. Technical report, 2005.
- [8] R Cont and P Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, 2004.
- [9] Michel Denuit, Xavier Maréchal, Sandra Pitrebois, and Jean-François Walhin. *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems*. John Wiley & Sons, 2007.
- [10] Sever S Dragomir. *Some Gronwall type inequalities and applications*. Nova Science, 2003.
- [11] Jan Grandell. *Aspects of risk theory*. Springer Science & Business Media, 2012.
- [12] Patrik Hardin and Sam Tabari. Modelling non-life insurance policyholder price sensitivity: A statistical analysis performed with logistic regression, 2017.
- [13] Christian Hipp. Stochastic control with application in insurance. In *Stochastic methods in finance*, pages 127–164. Springer, 2004.
- [14] Christian Hipp and Michael Plum. Optimal investment for insurers. *Insurance: Mathematics and Economics*, 27(2):215–228, 2000.

- [15] Rob Kaas, Marc Goovaerts, Jan Dhaene, and Michel Denuit. *Modern actuarial risk theory: using R*, volume 128. Springer Science & Business Media, 2008.
- [16] Yehuda Kahane. The theory of insurance risk premiums—a re-examination in the light of recent developments in capital market theory. *ASTIN Bulletin: The Journal of the IAA*, 10(2):223–239, 1979.
- [17] Hilbert J Kappen. An introduction to stochastic control theory, path integrals and reinforcement learning. In *AIP conference proceedings*, volume 887, pages 149–181. American Institute of Physics, 2007.
- [18] Christian Kasumo, Juma Kasozi, and Dmitry Kuznetsov. On minimizing the ultimate ruin probability of an insurer by reinsurance. *Journal of Applied Mathematics*, 2018, 2018.
- [19] Paul C Kettler. Lévy-copula-driven financial processes. *Preprint series. Pure mathematics <http://urn.nb.no/URN:NBN:no-8076>*, 2006.
- [20] Doron Kliger and Benny Levikson. Pricing insurance contracts—an economic viewpoint. *Insurance: Mathematics and Economics*, 22(3):243–249, 1998.
- [21] Stuart A Klugman, Harry H Panjer, and Gordon E Willmot. *Loss models: from data to decisions*, volume 715. John Wiley & Sons, 2012.
- [22] Elena Krashennnikova, Javier García, Roberto Maestre, and Fernando Fernández. Reinforcement learning for pricing strategy optimization in the insurance industry. *Engineering Applications of Artificial Intelligence*, 80:8–19, 2019.
- [23] Hanson Li and Ruodu Wang. Pelve: Probability equivalent level of var and es. *Available at SSRN*, 2019.
- [24] Ronnie Loeffen, Irmina Czarna, Zbigniew Palmowski, et al. Parisian ruin probability for spectrally negative lévy processes. *Bernoulli*, 19(2):599–609, 2013.
- [25] Roger B Nelsen. *An introduction to copulas*. Springer Science & Business Media, 2007.
- [26] Esbjörn Ohlsson and Björn Johansson. *Non-life insurance pricing with generalized linear models*, volume 2. Springer, 2010.
- [27] Antonis Papapantoleon. An introduction to lévy processes with applications in finance, 2008.
- [28] Ken-Iti Sato. *Lévy processes and infinitely divisible distributions*. Cambridge university press, 1999.
- [29] Timothy Sauer. *Numerical Analysis*. Pearson, 2013.
- [30] Daniel Yannik Straumann. Correlation and dependency in risk management: properties and pitfalls. In *Risk Management: Value at Risk and Beyond*. Cambridge University Press, 2001.
- [31] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.
- [32] Peter Tankov. Lévy copulas: review of recent results. In *The fascination of probability, statistics and their applications*, pages 127–151. Springer, 2016.
- [33] Julien Trufin, Hansjoerg Albrecher, and Michel M Denuit. Properties of a risk measure derived from ruin theory. *The Geneva Risk and Insurance Review*, 36(2):174–188, 2011.
- [34] J. L. van Velsen. Parameter estimation of a lévy copula of a discretely observed bivariate compound poisson process with an application to operational risk modelling, 2012.
- [35] Mario V Wuthrich. Non-life insurance: mathematics & statistics. *Available at SSRN 2319328*, 2019.
- [36] Ji Yao. Generalized linear models for non-life pricing-overlooked facts and implications. *Institute and Faculty of Actuaries*, 2013.
- [37] Yuqing Zhang and Neil Walton. Adaptive pricing in insurance: Generalized linear models and gaussian process regression approaches. *arXiv preprint [arXiv:1907.05381](https://arxiv.org/abs/1907.05381)*, 2019.