Aarhus University
Department of Mathematics

PHD DISSERTATION

Decision and evaluation in non-life insurance mathematics

Julie Thøgersen

October, 2019

Julie Thøgersen
Department of Mathematics
Aarhus University
Ny Munkegade 118
8000 Aarhus C
Denmark

Submission:
This thesis was submitted to the Graduate School of Science and Technology, Aarhus University, on October 8, 2019.

Assessment Committee:
Associate Professor Andreas Basse-O'Connor (chairman)
Aarhus University

Professor Filip Lindskog
Stockholm University

Professor Stéphane Loisel
University of Lyon

Supervisors:
Professor Mark Podolskij
Aarhus University

Professor Emeritus Søren Asmussen
Aarhus University

Professor Bent Jesper Christensen
Aarhus University

Summary

This dissertation takes on an applied probability approach to miscellaneous topics of non-life insurance mathematics. It is based on five self-contained papers, Paper A–E.

Widely speaking, insurance provides protection from a financial loss. The insurance market has a supply and a demand side. We exclusively consider the case where it is an individual holding the risk, and an insurance company offering to take on (some of) the risk in exchange for a premium.

First, in Paper A we consider a Markov renewal equation in a heavy-tailed setting. The main objective is to find the asymptotic properties of this Markov renewal equation, and as a step therefor, the asymptotics of random sums with different degrees of heavy-tailedness are also studied. Markov renewal processes and heavy tailed distributions are both tools used for explaining phenomena in observed insurance data. The results in the paper are, however, also relevant to other fields of research than actuarial science.

Next, we present studies aimed for the supply-side of insurance. In Paper B we analyse, for the insurance company, the problem of finding the optimal premium as a function of the fixed amount deductible level. The reserve of the company is modelled by a diffusion approximation and the optimality criterion is either the ruin probability or, if ruin certain, the expected time to ruin. In Paper C and Paper D, we extend this idea to a competitive setting with two suppliers of insurance. In Paper C we analyse the case where market frictions are present and in Paper D the case of product differentiation. In both cases, we model the competition by a stochastic differential game, where the largest company (in terms of initial capital) tries to push their smaller competitor even further away, while the small company tries to pull closer. In the market friction case, we consider Bertrand game and find closed-from Nash equilibrium premiums, and in the product differentiation case, we consider a Stackelberg game and find closed-form Stackelberg equilibrium premiums.

And finally, we finish with Paper E contributing to the theory of demand for insurance. Here we question the simplicity of the product design in today's insurance market, where the standard and often only deductible structure available is the fixed amount deductible. For different pricing mechanisms, we study the welfare loss of an individual when being restricted to this trivial product design compared to being offered a completely flexible one.

Resumé

Denne afhandling benytter anvendt sandsynlighedsteori til at modellere diverse problemer i skadesforsikringsmatematik. Den er baseret på fem selvstændige artikler, ArtikelA–E.

Overordnet sagt tilbyder forsikring beskyttelse mod et økonomisk tab. Forsikringsmarkedet har en udbuds- og efterspørgselsside. Vi betragter udelukkende det tilfælde, hvor det er et individ, som bærer en risiko, og et forsikringsselskab, som tilbyder at påtage sig (noget af) risikoen i gengæld for en præmie.

Vi starter med Artikel A, hvor vi betragter en Markov-fornyelsesligning i en opstilling med tunge haler. Hovedmålet er at finde de asymptotiske egenskaber for denne Markov-fornyelsesligning, og som et skridt på vejen dertil studeres også asymptotikken af stokastiske summer med forskellige grader af tunge haler. Markov-fornyelsesprocesser og fordelinger med tunge haler er begge værktøjer, der bruges til at forklare fænomener observeret i forsikrings-data. Resultaterne i artiklen er dog også relevante for andre forskningsområder end aktuarvidenskab.

Dernæst præsenterer vi artikler, der er rettet mod udbudssiden af forsikring. I Artikel B analyserer vi for forsikringsselskabet, hvordan man kan finde den optimale præmie som en funktion af selvrisikoen. Virksomhedens reserve er modelleret som en diffusionsapproksimation, og optimalitetskriteriet er enten ruin sandsynligheden eller den forventede tid til ruin, hvis ruin er uundgåelig. I Artikel C og Artikel D udvider vi denne idé til en konkurrencemæssig konstruktion, hvor to selskaber udbyder forsikring. I Artikel C analyserer vi tilfældet med markedsfriktion, og i Artikel D tilfældet med produktdifferentiering. I begge tilfælde modellerer vi konkurrencen ved et stokastisk differential spil, hvor den største virksomhed ihht. startkapital forsøger at skubbe deres mindre konkurrent endnu længere væk, imens den mindste virksomhed forsøger at hive sig nærmere. I tilfældet med markedsfriktioner betragter vi et Bertrand spil og finder lukkede løsninger på Nash-ligevægtspræmier, og i tilfældet med produktdifferentiering betragter vi et Stackelberg spil og finder lukkede løsninger på Stackelberg-ligevægtspræmier.

Og endeligt afslutter vi med Artikel E, som bidrager til teorien om efterspørgsel efter forsikring. Her sætter vi spørgsmålstegn ved, hvorvidt produktdesignet på nutidens forsikringsmarked er for enkelt, hvor standardproduktet, som ofte er det eneste udbudt, indeholder en fast selvrisiko. For forskellige prisfastsættelsesmekanismer undersøger vi et individs velfærdstab, når det begrænses til dette trivielle produktdesign sammenlignet med at blive tilbudt et fuldstændigt fleksibelt produktdesign.

Preface

The dissertation is built on five papers, Paper A-E, presented in four chapters with the following structure

CHAPTER 1

- **Paper A**: Markov dependence in renewal equations and random sums with heavy tails. By Søren Asmussen and Julie Thøgersen. Published in Stochastic Models, 2017.

CHAPTER 2

- **Paper B**: Optimal premium as a function of the deductible: customer analysis and portfolio characteristics. By Julie Thøgersen. Published in Risks, 2016.

CHAPTER 3

- **Paper C**: Nash equilibrium premium strategies for push–pull competition in a frictional non-life insurance market. By Søren Asmussen, Bent Jesper Christensen and Julie Thøgersen. Published in Insurance: Mathematics and Economics, 2019.
- **Paper D**: Stackelberg equilibrium premium strategies for push-pull competition in a non-life insurance market with product differentiation. By Søren Asmussen, Bent Jesper Christensen and Julie Thøgersen. Published in Risks, 2019.

CHAPTER 4

- **Paper E**: Personal non-life insurance decisions and the welfare loss from flat deductibles. By Mogens Steffensen and Julie Thøgersen. Published in ASTIN Bulletin, 2019.

Each chapter has its own introduction due to the diversity of the papers. The introductions present some preliminaries to the papers. A state-of-the-art is provided in the papers, and will only supplemented in the introductions. The five papers are self-contained and are presented as such. The presentation of the papers here aligns with their published versions with the exception of minor adjustments. Notational discrepancies among the chapters might occur. Each introduction and each paper is ended by a reference list.

Acknowledgements

During my time as a PhD student, many people have shared my joys, passions and victories, and many have helped me through the frustrations and confusion that occurred at times. I am forever grateful to everyone who has accompanied, inspired and supported me on this journey.

First and foremost, I owe my profound gratitude to my supervisor Søren Asmussen, who trusted me with the opportunity to pursue a PhD. Experiencing his high level of expertise and vast knowledge first-hand has been a great gift. I would like to thank him for the patience he has had with me, and for showing me direction on this journey when I felt I had none. I am also truly grateful for the wisdom and support provided throughout by my supervisor Bent Jesper Christensen. I especially want to thank him for all of his enlightening inputs and his positive mindset. Also, thanks to Mark Podolskij for being the supervisor who always checks in to make sure everything was A-OK, and for bringing good vibes wherever he goes.

At University of Copenhagen, where I have spent a good amount of time, I would like to thank Mogens Steffensen for his hospitality, and for the meetings which I always left encouraged. It has been truly inspiring to work together with him, and for that I am very grateful. I further visited University of Liverpool. Here I would like to thank the IFAM-team for a lovely time, and especially Corina Constantinescu, Apostolos Papaioannou and Lewis Ramsden for the collaborations which I hope will turn out fruitful in the future.

I would like to thank my fellow PhD-students at Aarhus University, in particular Victor, Mikkel, Thorbjørn, Mads, Patrick and Claudio, for creating a pleasant atmosphere both on and off campus, and in the same context, the Cake Club at University of Copenhagen for the weekly "meetings". Also, thanks to the people I have met on conferences for the good times and for giving me renewed energy to go back to my desk and continue the research.

My heartfelt thanks goes to my family, in particular my mom, siblings, dad and step-mom, for their unconditional love and support. A special thanks to my mom, for always believing in me and for cheering me on whenever I needed it.

I am deeply grateful to Niklas, my significant other, for continuously inspiring me to do better. Above all, I thank him for being my best friend and for sharing the wonderful things in life with me. This leads on to the final thanks devoted to our beloved son, Valdemar, whose antics and infectious laughter can make any worries disappear.

Contents

Chapter 1

Markov renewal equations and heavy tails

The first paper fills out a gap in the literature by finding the asymptotics of a Markov renewal equation in the case where the measures involved are heavy-tailed and have spectral radius smaller than one.

In the preliminaries, we will start by introducing a simple concept and increase the level of abstraction little by little. Many of the concepts here has application to multiple areas of study, especially queuing theory. However, we restrict the focus solely to the connection to risk theory.

During a fixed period of time, we can split the uncertainty of the risk composed by a portfolio into two: the random number of claims and the random sizes of the claims reported. The aggregate expenses of claims during that period of time is then a random sum, as introduced in Section 1.1.1. By generalising the claim number to a continuous-time counting process, the random sum evolves over time. The most common counting processes are the Poisson process and the renewal process, which we define in Section 1.1.2. The Poisson process in particular gives rise to the compound Poisson process, which is the standard process to model the development of aggregate claim expenses of a portfolio, as we see in Section 2.1.4. As considered in Wang and Yuen (2005), dependence can occur between different classes of insurance. A car accident is an example of a source of event that may cause claims in multiple classes, e.g. a damaged car in one class and personal injury in another. The aggregate claims are then modelled by a double-sum, similar to the one studied in Theorem A.3 in Paper A.

Renewal processes leads on to renewal theory in Section 1.1.3, including the renewal equation, a strong instrument that has a wide range of application. For example, in Section 2.1.4 we see that the ruin probability, the most common measure of risk in actuarial science, also satisfies a renewal equation.

Another concept with a wide range of application is Markov processes, which we consider in Section 1.1.4. In non-life insurance, Markov processes may be used for Markov-modulation, where the risk process (fx the claim frequency and/or the claim sizes) is influenced by an external environment Markov process. A popular example is automobile insurance, where the risk process obviously depends on the weather states, e.g., {icy, foggy, rainy, other}. References along this type of work are; Asmussen (1989), Asmussen et al. (1995), Lu and Li (2005), and Zhu and Yang (2008). Speaking of automobile insurance, Markov chains are also a common tool for modelling the evolution of the classification of a policyholder in a bonus-malus system (a merit system where the premium depends on the driver's claim history). For a thorough review of such systems, we refer to Denuit et al. (2007).

In Section 1.1.5 we consider the next extension, namely Markov renewal processes and semi-Markov processes, introduced simultaneously by Levy (1954) and Smith (1955). These processes can be considered as generalisations of both renewal processes and Markov processes. Correspondingly, in extension to the Markov-modulated models above, Reinhard (1984) and Janssen (1977) consider a risk process influenced by an external environment semi-Markov process. In this section we also consider the Markov renewal equation, the associated generalisation to the renewal equation previously mentioned. Markov renewal equations are the main concern in Paper A in a heavy-tailed setting.

In insurance, heavy-tailed claims relative to light-tailed claims, constitutes a higher risk and is thus considered more "dangerous". They tend to have large values with many outliers. In actuarial sciences, subexponential is nearly considered to be a synonym for heavy-tailedness. The class of subexponential distributions was first introduced by Chistyakov (1964), followed independently by Chover et al. (1973). Veraverbeke (1977) and Embrechts and Veraverbeke (1982) merged subexponential distributions with risk theory by considering risk models with subexponential claim sizes. One of their main results is stated in Section 2.1.4. A small overview of the theory of heavy-tails appears in Section 1.1.6. The definitions here are similar to the ones in Paper A and, as in the paper, we refer to Foss et al. (2013) for an extensive treatment of the subject.

1.1 Prelimiaries

First we make a short note on notation. For a distribution F on $\mathbb{R}^+$, let $F(x)$ be the cumulative distribution function and $\overline{F}(x) = 1 - F(x)$ the tail function. Let F and G be distributions and let $X \sim F$ and $Y \sim G$ be independent random variables. $F * G$ then denotes the convolution of F and G,

$$F * G(x) = \int_0^x G(x-y)F(dy) = \int_0^x F(x-y)G(dy)$$

which is the distribution of the sum $X + Y$. The n-th fold convolution is defined recursively as

$$F^{*n}(x) = \int_0^x F^{*(n-1)}(x-y)F(dy) \quad \text{where} \quad F^{*0}(x) = \begin{cases} 0 \text{ if } x < 0, \\ 1 \text{ if } x \geq 0. \end{cases}$$

Moreover, $\mathbb{N} = \{1, 2, \ldots\}$ denotes the natural numbers, whereas $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

1.1.1 Random sums

Let $(X_i)_{i\in\mathbb{N}}$ be a collection of random variables, and N be a counting random variable (taking values in $\mathbb{N}_0$) independent of the X_i's. Define the random sum

$$S = X_1 + \cdots + X_N = \sum_{i=1}^{N} X_i \tag{1.1}$$

using the convention $S = 0$ on $\{N = 0\}$. In most cases, the collection $(X_i)_{i\in\mathbb{N}}$ is assumed to be independent and identically distributed (i.i.d) with common distribution F on $\mathbb{R}^+$. This assumption is also made here. (1.1) then has mean and variance

$$\mathbb{E}[S] = \mathbb{E}[N]\mathbb{E}[X_1] \text{ and } \mathbb{V}\mathrm{ar}[S] = \mathbb{E}[N]\mathbb{V}\mathrm{ar}[S] + \mathbb{E}[S]^2\mathbb{V}\mathrm{ar}[N],$$

respectively. Conditioned on $N = n$, S has the distribution F^{*n}. The unconditional compound distribution of S can then be found using the law of total probability,

$$F_S(s) = \sum_{n=1}^{\infty} F^{*n}(s)\mathbb{P}(N = n)$$

In insurance, N could be the number of claims in a given time interval and $X_1, X_2, \ldots$ the claim sizes. S is then the total amount of claims during that time interval.

1.1.2 Counting and compound processes

In previous section, the random counting variable N was considered for a fixed time point. This can be generalised to a counting process $(N_t)_{t\geq 0}$ that evolves in time. The general definition of a counting process is briefly stated below.

Definition 1.1 (Counting process). *$(N_t)_{t\geq 0}$ is a counting process if (i) N(0)=0 a.s., (ii) $N_t \in \mathbb{N}_0$ for all $t \geq 0$, and (iii) it is increasing, i.e. $N_t \leq N_s$ for $0 \leq t < s$.*

Definitions here are similar to the ones in Mikosch (2009). In the insurance example, the random variable N_t is the number of claims occurred by time t. The claim arrives at times $(T_i)_{i\in\mathbb{N}_0}$ that satisfies $0 = T_0 \leq T_1 \leq \ldots$. At the i'th arrival time T_i, a claim of random size X_i occurs. As for the random sum in Section 1.1.1, the sequence of claim sizes $(X_i)_{i\in\mathbb{N}}$ are assumed to be i.i.d and independent of the arrival times $(T_i)_{i\in\mathbb{N}_0}$. The claim number process is then defined as the counting process $N_t = \#\{i \geq 1 : T_i \leq t\}$ for $t \geq 0$. Let $W_i = T_i - T_{i-1}$ for $i \geq 1$ be the inter-arrival times. The arrival times can conversely be written as the sum $T_i = W_1 + \cdots + W_i$.

The total claim amount process, i.e. the generalization of the random sum in the previous section, is then the compound process

$$A_t = X_1 + \cdots + X_{N_t} = \sum_{i=1}^{N_t} X_i \tag{1.2}$$

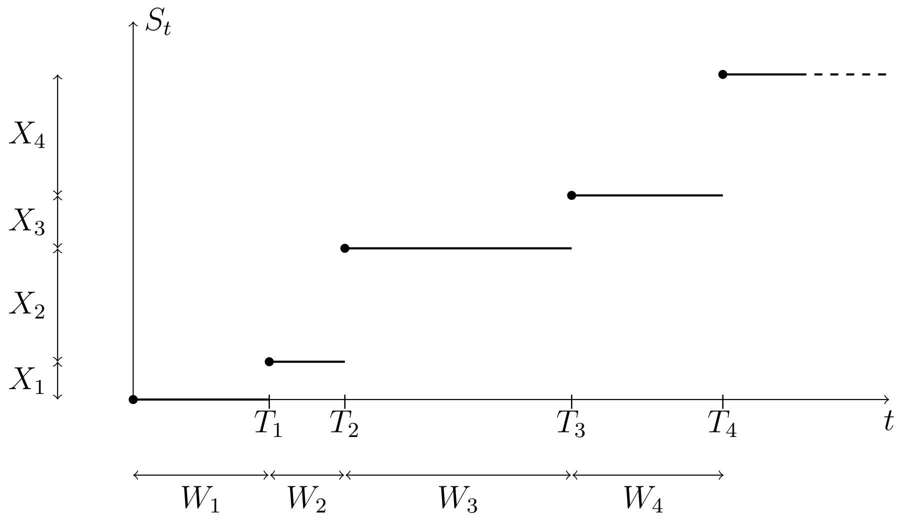


Figure 1.1

A sample path of $(S_t)_{t\geq 0}$ is presented in Figure 1.1.

The most commonly used counting process to model the number of claims is the homogeneous Poisson process. We will not discuss the inhomogeneous Poisson process (having a parameter that is non-constant), so when we refer to Poisson processes the homogeneity is implied.

Definition 1.2 (Poisson process). *A Poisson process $(N_t)_{t\geq 0}$ with intensity λ is defined by the following properties*

(i) *It begins at zero, i.e. $N(0) = 0$ a.s.*

(ii) *It has stationary and independent increments. , hence for any $n \geq 1$ and time points $t_1 < t_2 < \ldots < t_n$, the increments $N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}}$ are mutually independent and $N_{t_i} - N_{t_{i-1}} \sim N_{t_i - t_{i-1}}$.*

(iii) *It has cádlág sample paths, i.e. the sample paths are right continuous and left limits exists.*

(iv) *The number of arrivals in any time interval of length T has an Poisson(λT) distribution, hence $N_{t+T} - N_t \sim$ Poisson(λT) for all $t \geq 0$.*

The terminology above (cádlág, stationary and independent/stationary increments, ect.) stems from Levy processes. However, it is a well-known fact, that the Poisson process is a special case of renewal counting processes, defined as:

Definition 1.3 (Renewal process). *Let $(W_i)_{i\in\mathbb{N}}$ be an i.i.d. sequence of non-negative random variables. Then*

$$T_0, \quad T_n = W_1 + \cdots + W_n \text{ for } n \geq 1 \tag{1.3}$$

is a renewal process. The process $N_t = \#\{i \geq 1 : T_n \leq t\}$ for $t \geq 0$ is the associated renewal counting process.

This leads to the following alternative and perhaps more intuitive definition of Poisson processes.

Definition 1.4 (Poisson process as renewal counting process). *Let W_i in Definition 1.3 be exponentially distributed with parameter λ, then $N_t = \#\{i \geq 1 : T_n \leq t\}$ is a Poisson process with intensity λ.*

The exponential distribution is especially known for its memoryless property, which allows for a Markov interpretation. In fact, it can be shown that the Poisson process is the only renewal process that satisfies the Markov property in Definition 1.7.

When $(N_t)_{t\geq 0}$ is a Poisson process, the random sum (1.2) becomes a compound Poisson process. Compound Poisson processes is inevitable to mention, when it comes to actuarial science. It the most common process to model outgoing claims in the collective model, see 2.1.4.

Definition 1.5 (Compound Poisson process). *Let $(N_t)_{t\geq 0}$ be a Poisson process and $(X_i)_{i\in\mathbb{N}}$ a sequence of i.i.d. random variables independent of $(N_t)_{t\geq 0}$, then $(A_t)_{t\geq 0}$ defined by*

$$A_t = \sum_{i=1}^{N_t} X_i,$$

is a compound Poisson process.

Let $m_Y(x) = \mathbb{E}[\exp(xY)]$ be the moment generating function of the random variable Y. The compound Poisson process then has the following mean, variance and moment generating function,

$$\mathbb{E}[A_t] = \lambda t \mathbb{E}[X_i], \quad \mathbb{V}\mathrm{ar}[A_t] = \lambda t \mathbb{E}[X_i^2]$$
$$m_{A_t}(x) = \mathbb{E}[\exp(xA_t)] = \exp(\lambda(m_{X_i}(x) - 1)).$$

1.1.3 Renewal equations

Of special interest in renewal theory is renewal equation,

$$Z(x) = z(x) + \int_0^x Z(x-y)dF(y) = z(x) + Z * F(x), \tag{1.4}$$

where z is a known non-negative function, Z an unknown non-negative function, and F a measure on $[0, \infty)$. The equation also appears in Paper A as (1.3).

The corresponding renewal function, defined by

$$U(t) = \sum_{i=0}^{\infty} F^{*(n)}(t), \tag{1.5}$$

is a key component to the solution of (1.4). The renewal function satisfies itself the renewal equation

$$U(t) = F(t) + \int_0^t U(t-y)dF(y).$$

Consider a renewal counting process $(N_t)_{t\geq 0}$ with i.i.d. inter-arrivals $(W_i)_{i\in\mathbb{N}}$ having common distribution F_W. The renewal function

$$U(t) = \sum_{i=0}^{\infty} F_W^{*(n)}(t) = \mathbb{E}[N_t] + 1$$

then describes the average behaviour of the renewal counting process. In relation to insurance, the relevance is quite obvious, as it represents the expected number of claims up till time t.

Under weak regularity conditions, see Resnick (2002), the renewal equation (1.4) has the unique solution

$$Z(x) = \int_0^x z(x-y)U(\mathrm{d}y) = z * U(x).$$

In application we are often interested in the asymptotic behaviour as $x \to \infty$. Paper A gives a summary of the asymptotic results, which depend on the mass of F.

1.1.4 Markov chains and processes

In order to distinguish, we use the terminology that a Markov chain is the discrete time version denoted by $(\xi(n))_{n\in\mathbb{N}}$, whereas a Markov process is the continuous-time version written $(\xi_t)_{t\geq 0}$. Note that this notation used for differentiating between discrete and continuous-time is not used consistently for other types of processes.

We define a Markov chain with a finite state space $\mathcal{E} \in \mathbb{R}^d$ (or countable in which case $d = \infty$) as follows.

Definition 1.6 (Markov chain). $(\xi(n))_{n\in\mathbb{N}_0}$ *is a Markov chain with state space* $\mathcal{E}$ *if it satisfies the Markov property*

$$\mathbb{P}(\xi(n) = k_n \mid \xi(1) = k_1, \ldots, \xi(n-1) = k_{n-1}) = \mathbb{P}(\xi(n) = k_n \mid \xi(n-1) = k_{n-1})$$

if $\mathbb{P}\big(\xi(1) = k_1, \ldots, \xi(n-1) = k_{n-1}\big) > 0$, *for any* $n \in \mathbb{N}_0$ *and all* $k_1, \ldots, k_n \in \mathcal{E}$.

In order to complete the definition of a Markov chain properly, we furthermore need two components. Firstly, the transition matrix $Q = (q_{ij})_{i,j\in\mathcal{E}} \in \mathbb{R}^{d\times d}$ defined by $q_{ij} = \mathbb{P}(\xi(n) = i \mid \xi(n-1) = j)$ in the homogeneous case, where the law of the evolution of the system is independent of time. Secondly, an initial probability distribution on the state space, namely $q_i = \mathbb{P}(\xi(0) = i)$ for $i \in \mathcal{E}$.

A Markov process $(\xi_t)_{\{t\geq 0\}}$ is defined similarly by the continuous time version of the Markov property.

Definition 1.7 (Markov process). $(\xi_t)_{\{t\geq 0\}}$ *is a Markov process with state space* $\mathcal{E}$ *if it satisfies the Markov property*

$$\mathbb{P}(\xi_{t_n} = k_n \mid \xi_{t_1} = k_1, \ldots, \xi_{t_{n-1}} = k_{n-1}) = \mathbb{P}(\xi_{t_n} = k_n \mid \xi_{t_{n-1}} = k_{n-1}) \tag{1.6}$$

if $\mathbb{P}(\xi_{t_1} = k_1, \ldots, \xi_{t_{n-1}} = k_{n-1}) > 0$, *for any* $0 \leq t_1 < \cdots < t_n$ *and all* $k_1, \ldots, k_n \in \mathcal{E}$.

A continuous-time Markov process can be perceived as a discrete-time Markov chain combined with a time scale where the waiting times in-between jumps are independent and exponentially distributed.

Indeed, more specifically, define the jump times of the Markov process recursively

$$T_0 = 0, \quad T_i = \inf\{t \geq T_{i-1} \mid \xi_t \neq \xi_{T_{i-1}}\},$$

using the convention $\inf \emptyset = \infty$. The embedded process $(\xi(i))_{i\in\mathbb{N}_0}$ defined by $\xi(i) = \xi_{T_i}$ is then a Markov chain. Further, the waiting times $W_i = T_i - T_{i-1}$ are independent and exponentially distributed with parameter only depending on the current state $\xi(i-1)$.

1.1.5 Markov renewal processes and Markov renewal equations

As the name suggests, Markov renewal processes combine renewal processes and Markov chains. We continue this discussion after the definitions are in place. Definitions and statements in this section are to be found in Çinlar (1969).

Let $(X_t)_{t\geq 0}$ be a stochastic process with a discrete state space $\mathcal{E}$. Let T_i denote the i'th jumping time with $T_0 = 0$, and $W_i = T_i - T_{i-1}$ the interarrival time. Define the sequence $(\xi(n))_{n\in\mathbb{N}_0}$ defined by the successive states visited, i.e. $\xi(n) = X_{T_n}$. A realisation of the system is depicted in Figure 1.2.

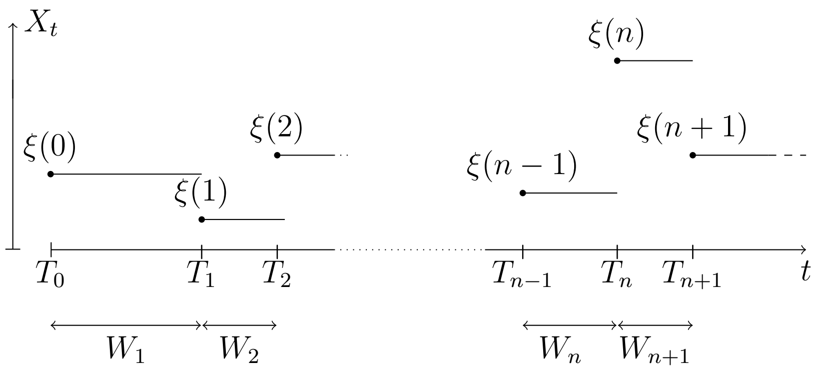


Figure 1.2: Depiction of a Markov renewal process.

Definition 1.8 (Markov renewal process). *The pair $(\xi(n), T_n)_{n\in\mathbb{N}_0}$ is called a Markov renewal process if*

$$\begin{aligned}\mathbb{P}(W_n \leq t, \xi(n) = j \mid (\xi(0), T_0), (\xi(1), T_1), \ldots, (\xi(n-1) = i, T_{n-1}))\\ = \mathbb{P}(W_n \leq t, \xi(n) = j \mid \xi(n-1) = i).\end{aligned}$$

In extension hereof, it is natural to define semi-Markov processes.

Definition 1.9 (semi-Markov process). *If $(\xi(n), T_n)_{n\in\mathbb{N}_0}$ is a Markov renewal process, then $(X_t)_{t\geq 0}$ is a semi-Markov process.*

The two following implications of Definition 1.8 clearly shows how Markov renewal processes and semi-Markov processes are generalisations of Markov processes and renewal processes.

$(\xi(n))_{n\in\mathbb{N}_0}$ is a Markov chain. Hence, the construction of semi-Markov processes are similar to that of Markov processes with the exception of the characterisation of the inter-arrival times. Recall that for the Markov process the inter-arrival times $(W_i)_{i\in\mathbb{N}}$ are independent with an exponential distribution. However, for the semi-Markov process, the inter-arrival times $(W_i)_{i\in\mathbb{N}}$ are merely conditionally independent given $(\xi(n))_{n\in\mathbb{N}_0}$. So where the Markov process satisfies the Markov property at any given time, the semi-Markov process only satisfies it at jump instants. In other words, the Markov process is a special case of semi-Markov processes, where $\mathbb{P}(W_n \le t, \xi(n) = j \mid \xi(n-1) = i) = q_{ij}(1 - \exp(-\lambda_i t))$ for some $\lambda_i > 0$.

On the other hand, the successive entrance times into state $i \in \mathcal{E}$, namely $\{T_n : \xi(n) = i\}$, constitute a renewal process. Hence, a Markov renewal process may also interpreted as a system of renewal processes progressing simultaneously such that the states of the successive renewals form a Markov chain.

A corresponding generalisation of the renewal equation is the Markov renewal equation in (1.1) of Paper A, namely

$$Z_i(x) = z_i(x) + \sum_{j\in\mathcal{E}} \int_0^x Z_j(x-y) F_{ij}(\mathrm{d}y),$$

where $\mathcal{E}$ is a finite index set, $(Z_i)_{i\in\mathcal{E}}$ a set of unknown non-negative functions, $(z_i)_{i\in\mathcal{E}}$ a set of non-negative known functions, and $(F_{ij})_{i,j\in\mathcal{E}}$ a set of measures on $[0,\infty)$.

One of the main concern in Paper A is to find the asymptotics of the Markov renewal equation in a heavy tailed setting, i.e. where the measures $(F_{ij})_{i,j\in\mathcal{E}}$ are related to a locally subexponential distribution, a property defined in the subsequent section.

1.1.6 Heavy-tailed distributions

In general, a distribution F is said to be heavy-tailed if F does not possess any positive exponential moments, that is $\int_{-\infty}^{\infty} \exp(\nu x) F(\mathrm{d}x) = \infty$ for all $\nu > 0$. If, on the contrary, a ν exists such that $\int_{-\infty}^{\infty} \exp(\nu x) F(\mathrm{d}x) < \infty$, then F is instead said to be light-tailed. In practice, most commonly used heavy-tailed distributions belong to the subexponential class defined below. Examples are; the log-normal distribution, Pareto distributions, Weibull distributions having shape parameter less than one.

1.1.6.1 Subexponential distributions

First we define long-tailed distributions.

Definition 1.10 (Long-tailed distribution)**.** *A distribution F is long-tailed if it has right-unbounded support, i.e. $\overline{F}(x) > 0$ for all x, and*

$$\frac{\overline{F}(x+y)}{\overline{F}(x)} \to 1 \qquad \text{as } x \to \infty \text{ for any fixed } y > 0. \tag{1.7}$$

Any long-tailed F is especially heavy-tailed. Next, we define subexponential distributions.

Definition 1.11 (Subexponential distribution)**.** *F is subexponential if it is long-tailed and*

$$\frac{\overline{F^{*2}}(x)}{\overline{F}(x)} \to 2 \qquad \textit{as } x \to \infty. \tag{1.8}$$

Let X_1 and X_2 be i.i.d. with common subexponential distribution F. It follows from the inclusion-exclusion formula that

$$\begin{aligned}
&\mathbb{P}(\max\{X_1, X_2\} > x) = \mathbb{P}(\{X_1 > x\} \cup \{X_2 > x\}) \\
&\quad = \mathbb{P}(X_1 > x) + \mathbb{P}(X_2 > x) - \mathbb{P}(X_1 > x, X_2 > x) \\
&\quad = 2\overline{F}(x) - \overline{F}(x)^2 \sim 2\overline{F}(x) \sim \overline{F^{*2}}(x).
\end{aligned}$$

So the condition (1.8) of F being subexponential can be interpreted as the probability of the set $\{X_1 + X_2 > x\}$ being asymptotically equal to the probability of the subset $\{\max\{X_1, X_2\} > x\}$. This is also known as the principle of a single big jump.

If $F \in \mathcal{S}$, then (1.8) can be generalised to the n-dimensional case

$$\frac{\overline{F^{*n}}(x)}{\overline{F}(x)} \sim n \qquad \text{for all } n \in \mathbb{N}.$$

The interpretation is the same as in the two-dimensional case. The only significant way the sum of independent F-distributed random variables can exceed some large threshold x is if the maximum of the random variables exceeds that x.

1.1.6.2 Subexponential densities

We now proceed to subexponential densities. Following the same order as for subexponential distributions, long-tailed properties are considered first.

Definition 1.12 (Long-tailed density)**.** *A density f is long-tailed if $f(x) > 0$ for all sufficiently large x and*

$$\frac{f(x+y)}{f(x)} \to 1 \qquad \textit{as } x \to \infty \textit{ for any fixed } y > 0.$$

Let F and G be distributions having densities f and g, respectively. The convolution $F * G$ then has density $f * g$ given by

$$f * g(x) = \int_0^x f(x-y)g(y)dy$$

Definition 1.13. *A density f is subexponential if it is long-tailed and*

$$\frac{f^{*2}(x)}{f(x)} \to 2 \qquad \textit{as } x \to \infty.$$

Let $\mathcal{S}_{ac}$ denote the class of subexponential densities. Note that $\mathcal{S}_{ac}$ is a subclass of $\mathcal{S}$. Indeed, if the distribution F has a subexponential density f, then

$$\overline{F}^{*2}(x) = \int_x^\infty f^{*2}(y)dy \sim 2\int_x^\infty f(y)dy = 2\overline{F}(x).$$

Since subexponentiality is a tail property, it is sufficient for a distribution F to only have a density $f(x)$ for large x.

1.1.6.3 Local subexponential distributions

Finally, we want to define of local long-tailedness and local subexponentiality of a distribution. We conclude by elaborating how local subexponentiality can be viewed as an intermediate property of being subexponential and having a subexponential density.

For a fixed and finite $T > 0$ define $\Delta = (0, T]$ and let

$$x + \Delta \equiv \{x + y \mid y \in \Delta\} = (x, x + T].$$

It here becomes more apparent what is meant by local, namely that the distribution is only considered on a fixed length interval $(x, x + T]$ when the location x of the interval tends to infinity. The long-tailed property can be defined locally as follows.

Definition 1.14 (Local long-tailed distribution). *A distribution F is Δ-long-tailed if $F(x + \Delta) > 0$ for all sufficiently large x and*

$$\frac{F(x + y + \Delta)}{F(x + \Delta)} \to 1 \qquad \text{as } x \to \infty \text{ for any fixed } y > 0.$$

Similarly we can define local subexponentiality.

Definition 1.15 (Local subexponential distribution). *F is Δ-subexponential if it is Δ-long-tailed and*

$$\frac{F^{*2}(x + \Delta)}{F(x + \Delta)} \to 2 \qquad \text{as } x \to \infty.$$

Let $\mathcal{S}_\Delta$ denote the class of Δ-subexponential distributions. Then for any finite T, $\mathcal{S}_\Delta$ is a subclass of $\mathcal{S}$, and if we were to allow $T = \infty$ the two classes coincide. It appears that Δ-subexponential distributions possess many similar properties as the ordinary subexponential distributions. A main and crucial difference is that the tail function $\overline{F}(x)$ is monotone whereas $F(x + \Delta)$ may not be.

A distribution F with a subexponential density f is Δ-subexponential for any $T > 0$ since

$$F^{*2}(x + \Delta) = \int_x^{x+T} f^{*2}(y)dy \sim 2\int_x^{x+T} f(y)dy = 2F(x + \Delta) \qquad \text{as } x \to \infty,$$

i.e $\mathcal{S}_{ac}$ is a subclass of $\mathcal{S}_\Delta$.

References

Asmussen, S. (1989). "Risk theory in a Markovian environment". *Scandinavian Actuarial Journal* 1989.2, pp. 69–100.

Asmussen, S., Frey, A., Rolski, T., and Schmidt, V. (1995). "Does Markov-modulation increase the risk?" *ASTIN Bulletin* 25.1, pp. 49–66.

Chistyakov, V. P. (1964). "Theorem on sums of independent random positive variables and its applications to branching processes". *Theory of Probability & Its Applications* 9.4, pp. 640–648.

Chover, J., Ney, P., and Wainger, S. (1973). "Functions of probability measures". *Journal d'Analyse mathématique* 26, pp. 255–302.

Çinlar, E. (1969). "Markov renewal theory". *Advances in Applied Probability* 1.2, pp. 123–187.

Denuit, M., Maréchal, X., Pitrebois, S., and Walhin, J.-F. (2007). *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems*. John Wiley & Sons.

Embrechts, P. and Veraverbeke, N. (1982). "Estimates for the probability of ruin with special emphasis on the possibility of large claims". *Insurance: Mathematics and Economics* 1.1, pp. 55–72.

Foss, S., Korshunov, D., and Zachary, S. (2013). *An introduction to heavy-tailed and subexponential distributions*. 2nd. Springer-Verlag New York.

Janssen, J. (1977). "The semi-Markov model in risk theory". *Advances in Operations Research*, pp. 613–621.

Levy, P. (1954). "Processus semi-markoviens". *Proceedings of the International Congress on Mathematics*. Vol. Amsterdam. 3, pp. 416–426.

Lu, Y. and Li, S. (2005). "On the probability of ruin in a Markov-modulated risk model". *Insurance: Mathematics and Economics* 37.3, pp. 522–532.

Mikosch, T. (2009). *Non-life insurance mathematics: An introduction with the Poisson process*. 2nd. Springer-Verlag Berlin Heidelberg.

Reinhard, J.-M. (1984). "On a class of semi-Markov risk models obtained as classical risk models in a Markovian environment". *ASTIN Bulletin* 14.1, pp. 23–43.

Resnick, S. I. (2002). *Adventures in stochastic processes*. Birkhäuser Basel.

Smith, W. L. (1955). "Regenerative stochastic processes". *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 232.1188, pp. 6–31.

Veraverbeke, N. (1977). "Asymptotic behaviour of Wiener-Hopf factors of a random walk". *Stochastic Processes and their Applications* 5.1, pp. 27–37.

Wang, G. and Yuen, K. C. (2005). "On a correlated aggregate claims model with thinning-dependence structure". *Insurance: Mathematics and Economics* 36.3, pp. 456–468.

Zhu, J. and Yang, H. (2008). "Ruin probabilities of a dual Markov-modulated risk model". *Communications in Statistics – Theory and Methods* 37.20, pp. 3298–3307.

Paper A

Markov dependence in renewal equations and random sums with heavy tails

Søren Asmussen and Julie Thøgersen

Abstract. The Markov renewal equation

$$Z_i(x) = z_i(x) + \sum_{j\in\mathcal{E}} \int_0^x Z_j(x-y)F_{ij}(\mathrm{d}y),\ i \in \mathcal{E},$$

is considered in the subcritical case where the matrix of total masses of the F_{ij} has spectral radius strictly less than one, and the asymptotics of the $Z_i(x)$ is found in the heavy-tailed case involving a local subexponential assumption on the F_{ij}. Three cases occur according to the balance between the $z_i(x)$ and the tails of the F_{ij}. A crucial step in the analysis is obtaining multivariate and local versions of a lemma due to Kesten on domination of subexponential tails. These also lead to various results on tail asymptotics of sums of a random number of heavy-tailed random variables in models which are more general than in the literature.

Keywords: Kesten's bound; key renewal theorem; locally subexponential; Perron-Frobenius theory; terminating Markov chains

Publication details: Stochastic models, 2017, vol. 33(4), pp. 617–632. DOI: 10.1080/15326349.2017.1375860

A.1 Introduction

The occurrence of heavy tails has been argued repeatedly in a variety of application areas covering insurance and finance (Embrechts et al. (1997)), telecommunications and internet traffic (Adler et al. (1998)), optics (Barakat (1976)), cell proliferation (Cao (2015)) and many more; a broad overview is in Resnick (2007). Correspondingly, performance analysis of models for such situations has triggered a vast literature on probabilistic features of heavy tails.

The present paper deals with two particular problems in this area. The first is asymptotics of renewal-type equations of the form

$$Z_i(x) = z_i(x) + \sum_{j\in\mathcal{E}} \int_0^x Z_j(x-y)F_{ij}(\mathrm{d}y),\ i \in \mathcal{E}, \tag{1.1}$$

where $\mathcal{E}$ is a finite index set, $(Z_i)_{i\in\mathcal{E}}$ a set of unknown functions defined on $[0,\infty)$, $(z_i)_{i\in\mathcal{E}}$ a set of non-negative known functions, and $(F_{ij})_{i,j\in\mathcal{E}}$ a set of non-negative heavy-tailed measures on $[0,\infty)$. The second is the tail behaviour of a random sum

$$S = \sum_{i=1}^{N} X_i, \tag{1.2}$$

where $N \in \mathbb{N}$ is a light-tailed random variable and the X_i are non-negative heavy-tailed random variables, such that certain types of dependence to be specified later may occur.

The simple renewal equation

$$Z(x) = z(x) + \int_0^x Z(x-y)F(\mathrm{d}y) \tag{1.3}$$

is a classical structure in applied probability and occurs, for example, in branching processes (Jagers (1975)), ruin problems (Feller (1971)) and ergodicity problems for possibly non-Markovian processes (Asmussen (2003, p. VI.1)). The emphasis is usually on asymptotic properties of $Z(x)$ as $x \to \infty$ where the simplest situation is existence of a limit[1] when F is a probability measure, that is, when $\|F\| = 1$ where $\|F\|$ is the total mass of F. However, in the branching process example $\|F\|$ is the expected number of children of an individual so obviously the case $\|F\| \neq 1$ is of interest. One then typically has an exponential order $\mathrm{e}^{\gamma x}$ of $Z(x)$. Here $\gamma > 0$ when $\|F\| > 1$ and $\gamma < 0$ when $\|F\| < 1$ and F is light-tailed (this last situation also occurs in Cramér-Lundberg asymptotics for ruin probabilities and queues, see Asmussen (2003, p. V.7)). For $\|F\| < 1$ and F heavy-tailed, the order depends on a delicate balance between the tails of F and z and in fact the results are more recent, Asmussen et al. (2003) (see also Wang et al. (2014) for a closely related result).

The system (1.1) goes under the name of the *Markov renewal equation*, see Asmussen (2003); this terminology stems from the case of

$$\boldsymbol{P} = \big(\|F_{ij}\|\big)_{i,j\in\mathcal{E}}$$

[1]This and the following statements require regularity conditions which we omit; see the citations given.

being stochastic, that is, the transition matrix of a Markov chain. In branching processes, it occurs when individuals have several types, and in insurance and finance, it relates to regime-switching; another relevant example comes from computer reliability problems, Asmussen et al. (2016). The known asymptotic results on the $Z_i(x)$ depend crucially on the spectral radius $\rho = \mathrm{spr}(\boldsymbol{P})$: if $\rho = 1$, in particular if $\boldsymbol{P}$ is stochastic, a limit exists whereas otherwise, similar to the simple case, the order is exponential, $\mathrm{e}^{\gamma x}$, where $\gamma > 0$ when $\rho > 1$ and $\gamma < 0$ when $\rho < 1$ and the F_{ij} are light-tailed. The gap is the case $\rho < 1$ and heavy tails. The main contribution of the paper in the setting of (1.1) is to fill this gap. The result is stated as Theorem A.2 below. Three cases occur depending on whether the tail of the z_i or the F_{ij} dominate, or if they are of same order. The conditions involve the non-standard concept of local subexponentiality.

One main example of random sums like (1.2) is total claim distributions in insurance: N is the number of claims in a given period and $X_1, X_2, \dots$ are the claim sizes that classically are taken i.i.d. and independent of N which is assumed light-tailed (e.g. the negative binomial distribution is popular because of its interpretation as a gamma mixture of Poissons). In credit risk, N could be the number of credit defaults in a portolio and $X_1, X_2, \dots$ the losses given default. With light-tailed X_i, $\mathbb{P}(S > x)$ decays roughly exponential, with heavy tails the asymptotic form is $\mathbb{E}N \cdot \mathbb{P}(X_1 > x)$. Our main result in that direction, Theorem A.3 below, is an extension to sums of the form $\sum_{k=1}^{d} \sum_{1}^{N_k} X_{ij}$ where $N_1, \dots, N_d$ are dependent and the distribution of X_{ij} depends on i. For example in the insurance setting, $(N_1, \dots, N_d)$ could be conditionally independent given $(\tau_1, \dots, \tau_d)$ with Poisson rate λ_k for N_k, where τ_k is the time spent by some environmental process in state k in the period $[0, T]$ (dependence occurs because $\tau_1 + \cdots + \tau_d = T$). The main step in the proof of Theorem A.3 is a version of a classical lemma due to Kesten on tail domination of subexponential sums; for the proof of Theorem A.2 we also need a local version of this.

The paper is organised as follows. Section A.2 starts with some necessary background, in particular on local subexponentiality, and proceeds to state the two main results of the paper, Theorems A.2 and A.3 referred to above. The rest of the paper is then proofs supplemented with miscellaneous results of some independent interest. Sections A.3–A.5 give the random sum part in various settings. This together with a number of additional steps then allows to conclude the proof of Theorem A.2 in Section A.6. It proceeds by first assuming $\boldsymbol{P}$ to be substochastic, thereby allowing Markov chain interpretations, and finally link the general case to this by a Perron-Frobenius type transformation.

A.2 Preliminaries and statement of main results

For a distribution F on $\mathbb{R}^+$, let $F(x) = F(0, x]$ be the cumulative distribution function (c.d.f.) and $\overline{F}(x) = 1 - F(x) = F(x, \infty)$ the tail. If F and G are distributions and

X and Y are independent random variables with distribution F and G, respectively, then $F * G$ denotes the convolution of F and G,

$$F * G(x) = \int_0^x G(x-y)F(dy) = \int_0^x F(x-y)G(dy),$$

which is the distribution of the sum $X + Y$.

We next briefly mention the most standard definitions and relations in the heavy-tailed area. For a more detailed and thorough treatment, see Embrechts et al. (1997), Foss et al. (2013). A distribution F is said to be heavy-tailed if F does not possess any positive exponential moments, that is, $\int_{-\infty}^{\infty} \exp(\lambda x)F(\mathrm{d}x) = \infty$ for all $\lambda > 0$. It is long-tailed, written $F \in \mathcal{L}$, if it has unbounded support and $\overline{F}(x+y)/\overline{F}(x) \to 1$ as $x \to \infty$ for any fixed $y \in \mathbb{R}$. Any $F \in \mathcal{L}$ is especially heavy-tailed. F is said to be subexponential, written $F \in \mathcal{S}$, if $\overline{F^{*n}}(x)/\overline{F}(x) \to n$ for all n (actually, it is sufficient that this holds for $n = 2$). The intuition is that the only significant way the sum of independent F-distributed random variables can exceed some large threshold x is if the maximum of the random variables exceeds that x. This is also known as the principle of a single big jump. Similarly, a density f is long-tailed if $f(x) > 0$ for all sufficiently large x and $f(x+t)/f(x) \to 1$ for any fixed t. If F and G have densities f and g, respectively, the convolution $F * G$ then has density $f * g$ given by $f * g(x) = \int_0^x f(x-y)g(y)\mathrm{d}y$. The density f of F is said to be subexponential, written $F \in \mathcal{S}_{ac}$, if f is long-tailed and $f^{*2}(x) = \int_0^x f(x-y)f(y)dy \sim 2f(x)$ as $x \to \infty$. Since subexponentiality is a tail property it is sufficient for a distribution F to only have a density $f(x)$ for sufficiently large x. Note that $\mathcal{S}_{ac}$ is a subclass of $\mathcal{S}$.

We proceed to the less standard concepts of local long-tailedness and subexponentiality, as introduced in Asmussen et al. (2003). The local property can be viewed as intermediate between being long-tailed (subexponential) and having a long-tailed (subexponential) density. First, we need to introduce some notation. For a fixed $T > 0$ define $\Delta = \Delta_T = (0, T]$ and let $x + \Delta = \{x + y \mid y \in \Delta\} = (x, x+T]$.

Definition A.1. *A distribution F is said to be Δ-long-tailed, written $F \in \mathcal{L}_\Delta$, if $F(x+\Delta) > 0$ for all sufficiently large x and $F(x+y+\Delta)/F(x+\Delta) \to 1$ as $x \to \infty$ for any fixed y. It is is Δ-subexponential, written $F \in \mathcal{S}_\Delta$, if $F \in \mathcal{L}_\Delta$ and $F^{*2}(x+\Delta)/F(x+\Delta) \to 2$.*

Notice that if we allow $T = \infty$ the class $\mathcal{L}_\Delta$ corresponds to the ordinary longtailed distributions $\mathcal{L}$; if $T < \infty$, then $\mathcal{L}_\Delta \subset \mathcal{L}$. Similarly, $\mathcal{S}_\Delta \subset \mathcal{S}$ for any finite T, and if we allow $T = \infty$ the two classes coincide. It appears that Δ-subexponential distributions possess many similar properties as the ordinary subexponential distributions. A main and crucial difference is that the tail function $\overline{F}(x)$ is monotone whereas $F(x+\Delta)$ may not be. Also, it it worth noticing that a distribution F with a subexponential density is Δ-subexponential for any $T > 0$.

We are now ready to state our main results. Recall that $\boldsymbol{P}$ is the matrix with ij'th element $p_{ij} = \|F_{ij}\|$, also $\boldsymbol{I}$ denotes the identity matrix.

Theorem A.2. *Consider the Markov renewal equation (1.1). Assume that $\boldsymbol{P}$ is irreducible with $\mathrm{spr}(\boldsymbol{P}) < 1$ and that $F_{ij}(x+\Delta) \sim p_{ij}c_{ij}G(x+\Delta)$ where $c_{ij} > 0$ and*

G is a Δ-subexponential distribution function for all $T > 0$. Define $g(x) = G(x, x+1)$ and $I_j = \int_0^\infty z_j(y)\,\mathrm{d}y$. Let d_{ij} be the ij'th element of the matrix $(\boldsymbol{I}-\boldsymbol{P})^{-1}\boldsymbol{M}(\boldsymbol{I}-\boldsymbol{P})^{-1}$ with $\boldsymbol{M} = (p_{k\ell}c_{k\ell})_{k,\ell\in\mathcal{E}}$, and let k_{ij} be the ij'th element of $\sum_0^\infty \boldsymbol{P}^n = (\boldsymbol{I}-\boldsymbol{P})^{-1}$. Then three cases occur:

(i) *Assume that z_j is directly Riemann integrable and $z_j(x)/g(x) \to 0$ for all $j \in \mathcal{E}$. Then*

$$Z_i(x) \sim \sum_{j\in\mathcal{E}} I_j d_{ij} g(x).$$

(ii) *Assume that z_j is directly Riemann integrable and $z_j(x)/g(x) \to a_j$ where $a_j > 0$ for at least one $j \in \mathcal{E}$. Then*

$$Z_i(x) \sim \sum_{j\in\mathcal{E}} (I_j d_{ij} + k_{ij}a_j)\, g(x).$$

(iii) *Assume that $z_j(y)/I_j$ has a subexponential density for all $j \in \mathcal{E}$ and that $z_j(x)/g(x) \to \infty$ for all j. Then*

$$Z_i(x) \sim \sum_{j\in\mathcal{E}} k_{ij} z_j(x).$$

Theorem A.3. *Let $N_1, \dots, N_d$ be non-negative integer-valued r.v.'s and define*

$$S = \sum_{i=1}^{d}\sum_{j=1}^{N_i} X_{ij},$$

where X_{ij} are mutually independent and independent of $N_1, \dots, N_d$ with distribution F_i. Assume $\overline{F}_i(x) \sim c_i\overline{F}(x)$ for some $F \in \mathcal{S}$ and some $c_1, \dots, c_d \geq 0$, where $c_i = 0$ should be understood as $\overline{F}_i(x) = o(\overline{F}(x))$. If there for all $i = 1, \dots, d$ are $\varepsilon_i > 0$ such that $\mathbb{E}[(1+\varepsilon_i)^{N_i}] < \infty$, then

$$\mathbb{P}(S > x) \sim \big(c_1\mathbb{E}[N_1] + \cdots + c_d\mathbb{E}[N_d]\big)\, \overline{F}(x).$$

Note that $N_1, \dots, N_d$ are not assumed independent. A local version of Theorem A.3 is given below as Theorem A.7.

A.3 Random sums of subexponentials

A classical bound due to Kesten (Athreya and Ney (1972) , Embrechts et al. (1997)) states that for a subexponential distribution F_0 and $\varepsilon > 0$ there exists a $D_0 = D_0(\varepsilon) > 0$ such that

$$\overline{F_0^{*n}}(x) \leq D_0(1+\varepsilon)^n \overline{F_0}(x) \qquad \text{for any } x \geq 0 \text{ and } n \in \mathbb{N}. \tag{1.4}$$

We give here a version involving the convolution of the tails of multiple subexponential distribution functions. It can be deduced from Foss and Richards (2010), but we include the proof since it is much shorter in our set-up.

Proposition A.4. *Let $F \in \mathcal{S}$, and let $F_1, \ldots, F_d$ be distributions with $\overline{F_i}(x) \leq \tilde{c}_i \overline{F}(x)$ with $\tilde{c}_i \geq 0$ for $i = 1, \ldots, n$. Then for every $\varepsilon > 0$ there exists a $D = D(\varepsilon)$ such that the following inequality holds:*

$$\overline{F_1^{*n_1} * \cdots * F_d^{*n_d}}(x) \leq D(1+\varepsilon)^n \overline{F}(x) \qquad \text{for all } x \geq 0 \text{ and } n_1, \ldots, n_d \in \mathbb{N} \tag{1.5}$$

where $n = n_1 + \cdots + n_d$. Furthermore if $\overline{F_i}(x) \sim c_i \overline{F}(x)$ with $c_i \geq 0$ for $i = 1, \ldots, n$ then

$$\overline{F_1^{*n_1} * \cdots * F_d^{*n_d}}(x) \sim (c_1 n_1 + \cdots + c_d n_d)\, \overline{F}(x). \tag{1.6}$$

Proof. Define the distribution function F_0 by having tail

$$\overline{F_0}(x) = \min\{k\overline{F}(x), 1\}$$

with $k = \max_{i=1,\ldots,n} \tilde{c}_i$. Note that taking the minimum (or multiplying by a positive constant) will preserve properties as being cadlag and non-increasing. The minimum will also ensure having values between $[0,1]$. So F_0 is indeed a distribution. Since subexponentiality is a tail property and the class of subexponential distributions is closed under tail equivalency, we must furthermore have $F_0 \in \mathcal{S}$. From the classical Kesten bound it then follows that for every $\varepsilon > 0$ there is a $D_0 = D_0(\varepsilon)$ such that (1.4) holds. By construction $\overline{F_i}(x) \leq \overline{F_0}(x)$, which implies that convolutions must have the corresponding stochastic ordering,

$$\overline{F_1^{*n_1} * \cdots * F_d^{*n_d}}(x) \leq \overline{F_0^{*n}}(x).$$

Combining with (1.4) we now only need to show that there exists a $D = D(\varepsilon)$ such that $D_0\overline{F_0}(x) \leq D\overline{F}(x)$. Let x_0 be chosen to satisfy that $\overline{F_0}(x) = k\overline{F}(x) < 1$ for all $x > x_0$ and $\overline{F_0}(x) = 1$ for all $x \leq x_0$. This is justified since the tail function $\overline{F}(x)$ is non-increasing, which also implies that $\overline{F}(x) \geq \overline{F}(x_0)$ for all $x \leq x_0$. In view of this, we may take $D = D_0 \max\{k, 1/\overline{F}(x_0)\}$.

The asympotics in (1.6) easily follows. To this end, just notice that $\overline{F_i^{*n_i}}(x) \sim n_i c_i \overline{F}(x)$ for all i by standard subexponential theory and proceed by induction, using that $\overline{G_1 * G_2}(x) \sim (\gamma_1 + \gamma_2)\overline{F}(x)$ if $\overline{G_i}(x) \sim \gamma_i \overline{F}(x)$ for $i = 1, 2$ and $F \in \mathcal{S}$. □

Kesten's bound is commonly used as majorant in dominated convergence to find the asymptotics of a randomly stopped sum of i.i.d. subexponential random variables. Proposition A.4 can be used correspondingly for the multidimensional case which we use to give the proof of Theorem A.3.

Proof. The law of total probability yields

$$\begin{aligned}
\frac{\mathbb{P}(S > x)}{\overline{F}(x)} &= \sum_{n_1, n_2, \ldots, n_d = 1}^{\infty} \mathbb{P}(N_1 = n_1, \ldots, N_d = n_d) \frac{\overline{F^{*n_1} * \cdots * F^{*n_d}}(x)}{\overline{F}(x)} \\
&\to \sum_{n_1, n_2, \ldots, n_d = 1}^{\infty} \mathbb{P}(N_1 = n_1, \ldots, N_d = n_d)(c_1 n_1 + \cdots + c_d n_d) \\
&= c_1 \mathbb{E}[N_1] + \cdots + c_d \mathbb{E}[N_d]
\end{aligned}$$

using Proposition A.4 and dominated convergence; this is justified since Hölder's inequality implies that $\mathbb{E}[(1+\varepsilon)^{N_1+\cdots+N_d}] < \infty$ if we take, say, $1+\varepsilon = \min\{1+\varepsilon_1, \ldots, 1+\varepsilon_d\}^{1/d}$. □

A.4 Random sums of local subexponentials

The objective of the following Lemma is to obtain an upper bound for the convolution of local subexponential distribution functions. This is needed to expand the local version of Kesten's bound stated as Proposition 4 in Asmussen et al. (2003) to include the convolution of several local subexponential distribution functions.

Lemma A.5. *Let $H \in \mathcal{S}_\Delta$ for some Δ. Assume that G_1 and G_2 are distributions satisfying*

$$G_i(x+\Delta) \le b_i H(x+\Delta) \qquad \text{for all } x \ge x_0$$

with $b_i > 1$ for $i = 1, 2$ and some sufficiently large x_0. Then there is a constant A independent of G_1 and G_2, such that

$$G_1 * G_2(x+\Delta) \le A b_1 b_2 H(x+\Delta) \qquad \text{for all } x \ge x_0. \tag{1.7}$$

Proof. Define

$$a = \sup_{u,v \ge x_0 \,:\, |u-v| \le 1} \frac{H(u+\Delta)}{H(v+\Delta)}$$

and notice that a is finite for sufficiently large x_0 since H is Δ-subexponential, and therefore especially Δ-longtailed. The convolution of interest can be split into two parts,

$$\begin{aligned} G_1 * G_2(x+\Delta) &= \int_0^x G_2(x-y+\Delta) G_1(\mathrm{d}y) + \int_x^{x+T} G_2(0, x+T-y] G_1(\mathrm{d}y) \\ &\equiv P_1(x) + P_2(x). \end{aligned}$$

Consider the first term. Let k be the smallest integer such that $x/k \le 1$ and $1/k \le T$. Partition the interval $(0, x]$ into k disjoint equally sized parts. For $x \ge x_0$ we can write the first term as a sum and assess it as follows,

$$\begin{aligned} P_1(x) &= \sum_{i=0}^{k-1} \int_{xi/k}^{x(i+1)/k} G_2(x-y+\Delta) G_1(\mathrm{d}y) \\ &\le b_2 \sum_{i=0}^{k-1} \int_{xi/k}^{x(i+1)/k} H(x-y+\Delta) G_1(\mathrm{d}y) \\ &\le a b_2 \sum_{i=0}^{k-1} \int_{xi/k}^{x(i+1)/k} H(x - xi/k + \Delta) G_1(\mathrm{d}y) \end{aligned}$$

$$\le ab_2 \sum_{i=0}^{k-1} H(x - xi/k + \Delta) G_1(xi/k, x(i+1)/k)$$
$$\le ab_1 b_2 \sum_{i=0}^{k-1} H(x - xi/k + \Delta) H(xi/k + \Delta)$$

where the summands can be evaluated backwards

$$\begin{aligned}\int_{xi/k}^{x(i+1)/k} H(x-y+\Delta)H(\mathrm{d}y) &\ge \frac{1}{a}\int_{xi/k}^{x(i+1)/k} H(x - ix/k + \Delta)H(\mathrm{d}y)\\ &= \frac{1}{a} H(x - ix/k + \Delta)H(xi/k + \Delta).\end{aligned}$$

Inserting this upper bound for the summands yields the inequality

$$\begin{aligned}P_1(x) &\le a^2 b_1 b_2 \sum_{i=0}^{k-1} \int_{xi/k}^{x(i+1)/k} H(x-y+\Delta)H(\mathrm{d}y)\\ &= a^2 b_1 b_2 \int_0^x H(x-y+\Delta)H(\mathrm{d}y)\\ &\le a^2 b_1 b_2 H^{*2}(x+\Delta).\end{aligned}$$

Since H is Δ-subexponential there must be a finite δ such that

$$H^{*2}(x+\Delta) \le (2+\delta)H(x+\Delta) \qquad \text{for all } x \ge x_0$$

if x_0 is sufficiently large. This provides the final inequality for the first term

$$P_1(x) \le (2+\delta)a^2 b_1 b_2 H(x+\Delta).$$

Now, consider the second term. It follows directly that

$$P_2(x) \le \int_x^{x+T} G_1(\mathrm{d}y) \le b_1 H(x+\Delta) \le b_1 b_2 H(x+\Delta).$$

Altogether we now have

$$G_1 * G_2(x+\Delta) \le ((2+\delta)a^2 + 1) b_1 b_2 H(x+\Delta).$$

This concludes the proof with $A = (2+\delta)a^2 + 1$. □

We can use this to obtain a local version of Proposition A.4.

Proposition A.6. *Let $F \in \mathcal{S}_\Delta$ for some Δ, and let $F_1, \dots, F_d$ be distributions with $F_i(x+\Delta) \sim c_i F(x+\Delta)$ with $c_i \ge 0$ for $i = 1, \dots, n$. Then for every $\varepsilon > 0$ there exists a $D = D(\varepsilon)$ and a $x_0 = x_0(\varepsilon)$ such that the following inequality holds*

$$F_1^{*n_1} * \cdots * F_d^{*n_d}(x+\Delta) \le D(1+\varepsilon)^n F(x+\Delta) \tag{1.8}$$

for all $x \ge x_0$ and $n_1, \dots, n_d \in \mathbb{N}$ where $n = n_1 + \cdots + n_d$. Furthermore

$$F_1^{*n_1} * \cdots * F_d^{*n_d}(x+\Delta) \sim (c_1 n_1 + \cdots + c_d n_d) F(x+\Delta).$$

Proof. Let $\varepsilon > 0$ be given. From Proposition 4 in Asmussen et al. (2003) it follows that for $i = 1, \ldots, d$ there is a $V_i = V_i(\varepsilon)$ such that

$$F_i^{*n_i}(x + \Delta) \leq V_i(1 + \varepsilon)^{n_i} F(x + \Delta).$$

(1.8) can then be proven by induction using Lemma A.5. For $d = 2$ it follows directly from the Lemma that there is an A such that

$$F_1^{*n_1}(x + \Delta) * F_2^{*n_2}(x + \Delta) \leq AV_1V_2(1 + \varepsilon)^{n_1+n_2} F(x + \Delta).$$

Letting $D = AV_1V_2$ we have the desired inequality. Assume that (1.8) holds for d distribution functions with constant D. Now consider the case with $d+1$ distribution functions, then we have

$$\begin{aligned} F_1^{*n_1} * \cdots * F_{d+1}^{*n_{d+1}}(x + \Delta) &= (F_1^{*n_1} * \cdots * F_1^{*n_d}) * F_{d+1}^{*n_{d+1}}(x + \Delta) \\ &\leq ADV_{d+1}(1 + \varepsilon)^{n_1+\cdots+n_d}(1 + \varepsilon)^{n_{d+1}} F(x + \Delta) \\ &= ADV_{d+1}(1 + \varepsilon)^{n_1+\cdots+n_{d+1}} F(x + \Delta). \end{aligned}$$

From Corollary 2 in Asmussen et al. (2003), it follows that $F_i^{*n_i}(x+\Delta) \sim n_i c_i F(x+\Delta)$ for $i = 1, \ldots, d$, and using Proposition 3 in Asmussen et al. (2003) one can further deduce that

$$F_1^{*n_1} * \cdots * F_d^{*n_d}(x + \Delta) \sim (n_1c_1 + \cdots + n_dc_d) F(x + \Delta). \qquad \square$$

Theorem A.7. *In addition to the assumptions of Theorem A.3 assume that $F \in \mathcal{S}_\Delta$ for some Δ and $F_i(x + \Delta) \sim c_i F(x + \Delta)$ for $c_1, \ldots, c_d \geq 0$. Then*

$$\mathbb{P}(S \in x + \Delta) \sim (c_1\mathbb{E}[N_1] + \cdots + c_d\mathbb{E}[N_d]) F(x + \Delta).$$

Proof. Use dominated convergence justified by Proposition A.6. $\square$

A.5 Random sums with subexponential densities

Proposition 8 in Asmussen et al. (2003) shows a density version of Kesten's bound. We accordingly now seek to obtain a version of Theorem A.3 involving densities instead. To do this, we need an upper bound for convolutions, as for the local subexponential case.

Lemma A.8. *Let $F \in \mathcal{S}_{ac}$ have density f. Assume that f_1 and f_2 are densities satisfying*

$$f_i(x) \leq b_i f(x) \qquad \textit{for all } x \geq x_0.$$

with $b_i > 1$ for $i = 1, 2$ and some sufficiently large x_0. Then there is a constant A independent of f_1 and f_2, such that

$$f_1 * f_2(x) \leq Ab_1b_2 f(x) \qquad \textit{for all } x \geq x_0.$$

Proof. Since f is longtailed then $f(x) > 0$ for all $x \geq x_0$ for x_0 sufficiently large, hence

$$a \equiv \sup_{y \in (0, x-x_0)} \frac{f(x-y)}{f(x)}$$

is finite. Consider the partition

$$\begin{aligned} f_1 * f_2(x) &= \int_0^x f_1(x-y) f_2(y) \mathrm{d}y \\ &= \int_0^{x-x_0} f_1(x-y) f_2(y) \mathrm{d}y + \int_{x-x_0}^{x_0} f_1(x-y) f_2(y) \mathrm{d}y + \int_{x_0}^{x} f_1(x-y) f_2(\mathrm{d}y) \\ &\equiv I_1 + I_2 + I_3. \end{aligned}$$

Now each term is assessed individually, starting with the first for $x \geq x_0$,

$$I_1 \leq b_1 \int_0^{x-x_0} f(x-y) F_2(\mathrm{d}y) \leq a b_1 \int_0^{x-x_0} f(x) F_2(\mathrm{d}y) \leq a b_1 b_2 f(x).$$

Analogously for the third term,

$$I_3 = \int_0^{x-x_0} f_2(x-y) f_1(y) \mathrm{d}y \leq a b_1 b_2 f(x).$$

Only the second term is left to be evaluated. Now let $x \geq 2x_0$, then it follows that

$$I_2 = b_1 b_2 \int_{x-x_0}^{x_0} f(x-y) f(y) \mathrm{d}y \leq b_1 b_2 \int_0^x f(x-y) f(y) \mathrm{d}y = b_1 b_2 f^{*2}(x).$$

Since f is a subexponential density, if x_0 is sufficiently large there must be a $\delta > 0$ such that

$$f^{*2}(x) \geq (2+\delta) f(x) \qquad \text{for all } x \geq x_0.$$

Hence,

$$I_2 \leq b_1 b_2 (2+\delta) f(x).$$

To conclude the proof, let $A = 2a + 2 + \delta$. □

Theorem A.9. *Let $F \in \mathcal{S}_{ac}$ have density f, and let $f_1, \ldots, f_d$ be densities with $f_i(x) \sim c_i f(x)$ with $c_i \geq 0$ for $i = 1, \ldots, n$. Then for any $\varepsilon > 0$ there is a $D = D(\varepsilon)$ and a $x_0 = x_0(\varepsilon)$ such that the following inequality holds*

$$f_1^{*n_1} * \cdots * f_{n_d}^{*n_d}(x) \leq D(1+\varepsilon)^n f(x)$$

for all $x \geq x_0$ and $n_1, \ldots, n_d \in \mathbb{N}$ where $n = n_1 + \cdots + n_d$. Furthermore

$$f_1^{*n_1} * \cdots * f_{n_d}^{*n_d}(x) \sim (n_1 c_1 + \cdots + n_d c_d) f(x).$$

Proof. Analogous to Proposition A.6. □

A.6 Proof of Theorem A.2

Recall that $\boldsymbol{P} = (p_{ij})_{i,j\in\mathcal{E}}$ is the matrix with entries defined by $p_{ij} = \|F_{ij}\|$, and $\rho = \operatorname{spr}(\boldsymbol{P}) < 1$ was defined as the Perron-Frobenius root of $\boldsymbol{P}$. Let $\boldsymbol{v}$ the corresponding right eigenvector. Then $\boldsymbol{P}^n\boldsymbol{v} = \rho^n\boldsymbol{v}$, which by strict positivity of $\boldsymbol{v}$ implies that the ij'th element $p_{ij}^{(n)}$ of $\boldsymbol{P}^n$ will decay at rate ρ^n.

We will first treat the case where in addition to $\rho < 1$ the matrix $\boldsymbol{P}$ is substochastic, which means that $\sum_j p_{ij} \le 1$, with strict inequality for at least one i. We can then introduce an irreducible absorbing Markov chain $(\xi_n)_{n\in\mathbb{N}}$ with state space $\mathcal{E}\cup\{\dagger\}$ where $\dagger$ is the so-called coffin state. Let $\boldsymbol{Q} = (q_{ij})_{i,j\in\mathcal{E}\cup\{\dagger\}}$ be the transition matrix of $(\xi_n)_{n\in\mathbb{N}}$ with $q_{ij} = p_{ij}$ for $i,j\in\mathcal{E}$. Since the state $\dagger$ is absorbing, it must be that $q_{\dagger i} = 0$ and $q_{\dagger\dagger} = 1$. This leaves $q_{i\dagger} = 1 - \sum_{j\in\mathcal{E}} p_{ij}$ with at least one index i such that $q_{i\dagger} > 0$ (due to $\boldsymbol{P}$ being substochastic). The time until absorbtion will then be $N = \inf\{n \ge 0 : \xi_n = \dagger\}$.

Let T_n denote the waiting time between jump n and $n+1$. Then $T_0, T_1, \ldots$ are conditionally independent given $\mathcal{F} = \sigma(\xi_0, \xi_1, \ldots)$ with conditional distribution function satisfying

$$G_{ij}(t) = \mathbb{P}(T_n \le t \mid \mathcal{F}) = \mathbb{P}(T_n \le t \mid \xi_n, \xi_{n+1})$$

on $\{\xi_n = i, \xi_{n+1} = j\}$, and the semi-Markov kernel $\boldsymbol{F} = (F_{ij})_{i,j\in\mathcal{E}}$ is defined by having elements

$$F_{ij}(t) = \mathbb{P}(\xi_{n+1} = j, T_n \le t \mid \xi_n = i) = q_{ij}G_{ij}(t).$$

Henceforth we will only consider $i,j \ne \dagger$. Thus, $F_{ij}(t) = p_{ij}G_{ij}(t)$.

The solution of the Markov renewal equation (1.1) is given by Proposition 4.4 in Asmussen (2003) as

$$Z_i(x) = \sum_{j\in\mathcal{E}} Z_{ij}(x) \text{ where } Z_{ij}(x) = \int_0^x z_j(x-y)U_{ij}(dy) \tag{1.9}$$

with Markov renewal kernel U_{ij} being the expected number of returns to state $j\in\mathcal{E}$ before time t given that the Markov chain starts in state $i\in\mathcal{E}$. That is,

$$U_{ij}(t) = \sum_{n=0}^{\infty}(\boldsymbol{F}^{*n})_{ij}(t) = \sum_{n=0}^{\infty}\mathbb{P}_i(\xi_n = j, S_{n-1} \le t),$$

where $S_{n-1} = T_0 + \cdots + T_{n-1}$. First, it is necessary to consider the local asymptotics of U_{ij}.

Lemma A.10. *Under the assumptions of Theorem A.2,*

$$U_{ij}(t+\Delta) \sim d_{ij}G(t+\Delta). \tag{1.10}$$

Proof. Notice that the nth convolution of the semi-Markov kernel also can be defined locally

$$(\boldsymbol{F}^{*n})_{ij}(t+\Delta) = \mathbb{P}_i(\xi_n = j, S_{n-1} \in t+\Delta).$$

Rewriting the right-hand side using conditional expectations yields

$$\begin{aligned}\mathbb{P}_i(\xi_n = j, S_{n-1} \in t+\Delta) &= \mathbb{P}_i(\xi_n = j, T_0 + \cdots + T_{n-1} \in t+\Delta)\\ &= p_{ij}^{(n)}\mathbb{P}(T_0 + \cdots + T_{n-1} \in t+\Delta \mid \xi_0 = i, \xi_n = j)\\ &= p_{ij}^{(n)}\mathbb{E}\Big[\mathbb{E}\big[\mathbb{P}(T_0 + \cdots + T_{n-1} \in t+\Delta \mid \xi_0 = i, \xi_n = j) \mid \xi_1, \dots, \xi_{n-1}\big]\Big]\\ &= p_{ij}^{(n)}\mathbb{E}\big[G_{ij}^{(n)}(t+\Delta)\big],\end{aligned}$$

where $G_{ij}^{(n)}$ denotes the distribution of the sum $T_0 + \cdots + T_{n-1}$ conditioned on the Markov states $\xi_0 = i, \xi_1, \dots, \xi_{n-1}, \xi_n = j$.

Let $N_{k\ell}^{(n)}$ the the random variable counting the number of jumps from state $k \in \mathcal{E}$ to state $\ell \in \mathcal{E}$ before the nth jump, that is

$$N_{k\ell}^{(n)} = \sum_{m=0}^{n-1} \mathbb{1}_{\{\xi_m = k, \xi_{m+1} = \ell\}}.$$

Correspondingly, let $N_{k\ell|ij}^{(n)}$ be the random variable representing the number of jumps from $k \in \mathcal{E}$ to $\ell \in \mathcal{E}$ before jump n given that $\xi_0 = i$ and $\xi_n = j$, i.e. it is distributed as $N_{k\ell}^{(n)}$ conditioned on $\xi_0 = i, \xi_n = j$. This has expected value

$$\begin{aligned}\mathbb{E}\big[N_{k\ell|ij}^{(n)}\big] &= \mathbb{E}\big[N_{k\ell}^{(n)} \mid \xi_0 = i, \xi_n = j\big] = \frac{\mathbb{E}\left[\sum_{m=0}^{n-1} \mathbb{1}_{\{\xi_0 = i, \xi_m = k, \xi_{m+1} = \ell, \xi_n = j\}}\right]}{\mathbb{P}(\xi_0 = i, \xi_n = j)}\\ &= \frac{\sum_{m=0}^{n-1} \mathbb{P}(\xi_0 = i, \xi_m = k, \xi_{m+1} = \ell, \xi_n = j)}{\mathbb{P}(\xi_0 = i, \xi_n = j)} = \frac{\sum_{m=0}^{n-1} p_{ik}^{(m)} p_{k\ell} p_{\ell j}^{(n-m-1)}}{p_{ij}^{(n)}}\\ &= \frac{\sum_{m=0}^{n-1} (\boldsymbol{P}^m)_{ik} p_{k\ell} (\boldsymbol{P}^{n-m-1})_{\ell j}}{(\boldsymbol{P}^n)_{ij}}.\end{aligned}$$

Recall that $G_{ij}^{(n)}$ is the distribution of the sum of the random variables $T_0, T_1, \dots, T_{n-1}$ conditioned on the Markov chain until the nth jump knowing that it starts in state i and ends in state j. Also, the distribution of T_m merely depends on ξ_m and ξ_{m+1}. Therefore, $G_{ij}^{(n)}$ can be interpreted as the random convolution

$$G_{ij}^{(n)}(t+\Delta) = \mathop{\ast}_{k,\ell \in \mathcal{E}} G_{k\ell}^{*N_{k\ell|ij}^{(n)}}(t+\Delta).$$

Applying Theorem A.7 the local asymptotics of $G_{ij}^{(n)}$ can then be specified as

$$\mathbb{E}\big[\mathop{\ast}_{k,\ell \in \mathcal{E}} G_{k\ell}^{*N_{k\ell|ij}^{(n)}}(t+\Delta)\big] \sim \Big(\sum_{k,\ell \in \mathcal{E}} \mathbb{E}\big[N_{k\ell|ij}^{(n)}\big] c_{k\ell}\Big) G(t+\Delta).$$

This asymptotic specification transfers to the semi-Markov kernel as follows

$$(\boldsymbol{F}^{*n})_{ij}(t+\Delta) \sim \sum_{k,\ell \in \mathcal{E}} \sum_{m=0}^{n-1} (\boldsymbol{P}^m)_{ik} p_{k\ell} (\boldsymbol{P}^{n-m-1})_{\ell j} c_{k\ell} G(t+\Delta)$$

and on to the Markov renewal kernel

$$\begin{aligned}U_{ij}(t+\Delta) &\sim \sum_{n=0}^{\infty}\sum_{k,\ell\in\mathcal{E}}\sum_{m=0}^{n-1}(\boldsymbol{P}^m)_{ik}p_{k\ell}(\boldsymbol{P}^{n-m-1})_{\ell j}c_{k\ell}G(t+\Delta)\\ &= \sum_{k,\ell\in\mathcal{E}}\sum_{m=0}^{\infty}\sum_{n=m+1}^{\infty}(\boldsymbol{P}^m)_{ik}p_{k\ell}(\boldsymbol{P}^{n-m-1})_{\ell j}c_{k\ell}G(t+\Delta).\end{aligned}$$

Since $\boldsymbol{P}$ has spectral radius strictly less than 1, the infinite series above converges with limits

$$\sum_{n=m+1}^{\infty}(\boldsymbol{P}^{n-m-1})_{\ell j} = \sum_{k=0}^{\infty}(\boldsymbol{P}^k)_{\ell j} = (\boldsymbol{I}-\boldsymbol{P})^{-1}_{\ell j}\,, \qquad \sum_{m=0}^{\infty}(\boldsymbol{P}^m)_{ik} = (\boldsymbol{I}-\boldsymbol{P})^{-1}_{ik}.$$

Thus,

$$U_{ij}(t+\Delta) \sim \Big(\sum_{k,\ell\in\mathcal{E}}(\boldsymbol{I}-\boldsymbol{P})^{-1}_{ik}p_{k\ell}(\boldsymbol{I}-\boldsymbol{P})^{-1}_{\ell j}c_{k\ell}\Big)G(t+\Delta)\,,$$

concluding the proof. □

For the proof of Theorem A.2, it suffices in view of (1.9) to find the asymptotics of the summands $Z_{ij}(x)=\int_0^x z_j(x-y)U_{ij}(dy)$ in the solution of the Markov renewal equation. As an introductory remark, notice that $G\in\mathcal{S}_\Delta$ for all $T>0$ has the implications

$$G(x,x+1/n] \sim \frac{g(x)}{n} \text{ for all } n \quad \text{and} \quad \frac{g(x+y)}{g(x)} \to 1 \text{ for all } |y|<y_0<\infty$$

for some appropriate y_0. For a suitable $A<x/2$ decompose $Z_{ij}(x)$ into three parts, namely

$$\begin{aligned}Z_{ij}(x) &= \left(\int_0^A + \int_A^{x-A} + \int_{x-A}^x\right) z_j(x-y)U_{ij}(\mathrm{d}y)\\ &\equiv J_1(x,A)+J_2(x,A)+J_3(x,A).\end{aligned}$$

We now consider and evaluate the three parts separately. In case i) of the Theorem we have

$$\begin{aligned}J_1(x,A) &= \int_0^A z_j(x-y)U_{ij}(\mathrm{d}y)\\ &= g(x)\int_0^A \frac{z_j(x-y)}{g(x-y)}\frac{g(x-y)}{g(x)}U_{ij}(\mathrm{d}y) = o(g(x))\end{aligned}$$

when $x\to\infty$. Also

$$\begin{aligned}J_2(x,A) &= \int_A^{x-A} z_j(x-y)U_{ij}(\mathrm{d}y) = \int_A^{x-A}\frac{z_j(x-y)}{g(x-y)}g(x-y)U_{ij}(\mathrm{d}y)\\ &= o(1)\int_A^{x-A} g(x-y)U_{ij}(\mathrm{d}y) = o(g(x))\int_A^{x-A}\frac{g(x-y)}{g(x)}U_{ij}(\mathrm{d}y).\end{aligned}$$

From this we can conclude

$$\lim_{A\to\infty}\lim_{x\to\infty}\frac{J_2(x,A)}{g(x)} = \lim_{A\to\infty}\lim_{x\to\infty}\frac{o(g(x))}{g(x)}(U(x-A)-U(A)) = 0.$$

Finally, consider a finite partition of the interval $(x-A, x]$ into n equally sized intervals. A is now assumed to be an integer. Furthermore, let

$$\bar{z}_n(x) = \sup_{|y-x|\leq \frac{1}{n}} z(y).$$

Then we can bound J_3 from above by the upper Riemann sum

$$\begin{aligned}
J_3(x,A) &= \int_{x-A}^{x} z_j(x-y)U_{ij}(\mathrm{d}y) \leq \sum_{k=0}^{An-1} \bar{z}_n\left(\frac{k}{n}\right) U_{ij}\left(x-\frac{k+1}{n}, x-\frac{k}{n}\right] \\
&\sim d_{ij}\sum_{k=0}^{An-1} \bar{z}_n\left(\frac{k}{n}\right) G\left(x-\frac{k+1}{n}, x-\frac{k}{n}\right] \\
&\sim g(x)\frac{d_{ij}}{n}\sum_{k=0}^{An-1} \bar{z}_n\left(\frac{k}{n}\right) \sim g(x)\frac{d_{ij}}{n}\sum_{k=0}^{\infty} \bar{z}_n\left(\frac{k}{n}\right)
\end{aligned}$$

as $x\to\infty$ and $A\to\infty$. Since z is assumed to be directly Riemann integrable, we have

$$\frac{1}{n}\sum_{k=0}^{\infty}\bar{z}_n\left(\frac{k}{n}\right) \to I_j$$

as n tends to infinity. So we have now obtained

$$\limsup_{A\to\infty}\limsup_{x\to\infty}\frac{J_3(x,A)}{g(x)} \leq d_{ij}I_j.$$

Using the same approach with $\liminf$, a similar bound from below is obtained using lower Riemann sums. This finishes case i).

Now consider case ii). The first of the decomposed parts has the asymptotics

$$\begin{aligned}
J_1(x,A) &= \int_0^A z_j(x-y)U_{ij}(\mathrm{d}y) \sim a_j g(x)\int_0^A \frac{g(x-y)}{g(x)}U_{ij}(\mathrm{d}y) \\
&\sim a_j g(x) U_{ij}(A)
\end{aligned}$$

which leads to

$$\lim_{A\to\infty}\lim_{x\to\infty}\frac{J_1(x,A)}{g(x)} = \lim_{A\to\infty}\lim_{x\to\infty} a_j U_{ij}(A) = k_{ij}a_j.$$

The second part in case ii) can also be evaluated as follows:

$$\begin{aligned}
J_2(x,A) &= \int_A^{x-A} z_j(x-y)U_{ij}(\mathrm{d}y) = \int_A^{x-A}\frac{z_j(x-y)}{g(x-y)}g(x-y)U_{ij}(\mathrm{d}y) \\
&= O(1)\int_A^{x-A} g(x-y)U_{ij}(\mathrm{d}y).
\end{aligned}$$

Notice the similarity to J_2 in case i). Corresponding calculations show that $J_2(x, A) = o(g(x))$. J_3 is also similar to case i), which concludes case ii).

In case iii) we are no longer able to consider the decomposition we have so far. Instead, let K_j denote the probability measure with density $z_j(x)/I_j$. Recall that $K_j \in \mathcal{S}_\Delta$. Now let $\Delta = (0, 1]$ and consider a new decomposition

$$\begin{aligned}\int_0^x z_j(x-y)U_{ij}(\mathrm{d}y) &= \int_0^{x-A} z_j(x-y)U_{ij}(\mathrm{d}y) + \int_{x-A}^x z_j(x-y)U_{ij}(\mathrm{d}y) \\ &\equiv I_1(x, A) + I_2(x, A),\end{aligned}$$

where the second term satisfies

$$I_2(x, A) \leq A \sup_{y \leq A}|z_j(y)|\; U_{ij}(x - A, x] = o(z_j(x)).$$

Letting A tend to infinity in a slower rate than x (recall that we have chosen A to be less than $x/2$) will preserve this inequaility. Now consider a corresponding decomposition of the convolution $K_j * U_{ij}$.

$$\begin{aligned}(K_j * U_{ij})(x + \Delta) &= \int_0^x K_j(x - y + \Delta)U_{ij}(\mathrm{d}y) \\ &= \int_0^{x-A} K_j(x - y + \Delta)U_{ij}(\mathrm{d}y) + \int_{x-A}^x K_j(x - y + \Delta)U_{ij}(\mathrm{d}y) \\ &\equiv I_1'(x, A) + I_2'(x, A).\end{aligned}$$

As for I_2 we can evaluate $I_2'(x, A) = o(z_j(x))$. Since $z_j \in \mathcal{L}$ we have $z_j(x) \sim I_j \cdot K_j(x + \Delta)$ for $\Delta = (0, 1]$ and therefore $I_1(x, A) \sim I_j \cdot I_1'(x, A)$. This yields

$$\int_0^x z_j(x-y)U_{ij}(\mathrm{d}y) \sim I_j \cdot (K_j * U_{ij})(X + \Delta).$$

Due to the assumption $z_j(x)/g(x) \to \infty$ the use of Proposition 3 in Asmussen et al. (2003) with $c_1 = 1$ and $c_2 = 0$ gives

$$I_j \cdot (K_j * U_{ij})(x + \Delta) \sim k_{ij} I_j K(x + \Delta) \sim k_{ij} z_j(x),$$

which concludes the proof of Theorem A.2 in the substochastic case.

If instead of $\boldsymbol{P}$ being substochastic it merely satisfies $\mathrm{spr}(\boldsymbol{P}) < 1$, we define the measure $\tilde{F}_{ij}(dx) = F_{ij}(dx)v_i/v_j$ for $i, j \in \mathcal{E}$ and let $\tilde{\boldsymbol{P}}$ be the matrix with elements $\|\tilde{F}_{ij}\| = \|F_{ij}\|v_i/v_j$. $\tilde{\boldsymbol{P}}$ will then be a substochastic matrix with row sums λ and therefore it must also have spectral radius less than one. Letting $\tilde{Z}_i(x) = v_i Z_i(x)$ and $\tilde{z}_i(x) = v_i z_i(x)$ another renewal equation occurs

$$\tilde{Z}_i(x) = \tilde{z}_i(x) + \sum_{j \in \mathcal{E}} \int_0^x \tilde{Z}_j(x-y)\tilde{F}_{ij}(dy)\,,$$

which has the same properties as analysed previously in the substochastic case with Markov kernel $\tilde{U}_{ij} = U_{ij}v_i/v_j$. What we have already shown can therefore be applied with coefficients $\tilde{a}_j, \tilde{d}_{ij}, \tilde{k}_{ij}$ and $\tilde{I}_j$. The assumption of the asymptotic properties of G_{ij} stated in Theorem A.2 implies that $\tilde{G}_{ij}(t+\Delta) = v_i G_{ij}(t+\Delta)/v_j \sim \tilde{c}_{ij} G(t+\Delta)$, where $\tilde{c}_{ij} = c_{ij} v_i/v_j$ and same relation transfers to $\tilde{d}_{ij} = d_{ij} v_i/v_j$. Correspondingly for $\tilde{k}_{ij} = \tilde{U}_{ij}(0,\infty) = U_{ij}(0,\infty) v_i/v_j = k_{ij} v_i/v_j$. Last to mention is $\tilde{I}_j = \int_0^\infty \tilde{z}_j(y)dy = v_j \int_0^\infty z_j(y)dy = v_j I_j$ and $\tilde{a}_j = v_j a_j$. Inserting these into the asympotics of $\tilde{Z}_i$ gives the same result as Theorem A.2. For example in case ii), the substochatic result gives $\tilde{Z}_i(x) \sim \sum_{j\in\mathcal{E}} (\tilde{I}_j \tilde{d}_{ij} + \tilde{a}_j \tilde{k}_{ij}) g(x)$ which translates to

$$\begin{aligned} Z_i(x) = \frac{1}{v_i}\tilde{Z}_i(x) &\sim \frac{1}{v_i}\sum_{j\in\mathcal{E}}\Big(v_j I_j \frac{v_i}{v_j} d_{ij} + v_j a_j \frac{v_i}{v_j} k_{ij}\Big) g(x) \\ &= \sum_{j\in\mathcal{E}}\Big(I_j d_{ij} + a_j k_{ij}\Big) g(x). \end{aligned}$$

References

Adler, R., Feldman, R., and Taqqu, M. (1998). *A practical guide to heavy tails: Statistical techniques and applications*. Birkhäuser Boston.

Asmussen, S. (2003). *Applied probability and queues*. 2nd. Vol. 51. Springer-Verlag New York.

Asmussen, S., Foss, S., and Korshunov, D. (2003). “Asymptotics for sums of random variables with local subexponential behaviour”. *Journal of Theoretical Probability* 16.2, pp. 489–518.

Asmussen, S., Lipsky, L., and Thompson, S. (2016). “Markov renewal methods in restart problems in complex systems”. *The Fascination of Probability, Statistics and their Applications*. Springer, Cham, pp. 501–527.

Athreya, K. B. and Ney, P. E. (1972). *Branching processes*. Springer, Berlin, Heidelberg.

Barakat, R. (1976). “Sums of independent lognormally distributed random variables”. *Journal of the Optical Society of America* 66.3, pp. 211–216.

Cao, X. (2015). “Inference for age-dependent branching process and their applications”. PhD thesis. George Mason University.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Vol. 33. Springer-Verlag Berlin Heidelberg.

Feller, W. (1971). *An Introduction to probability theory and its applications*. 2nd. Vol. 2. John Wiley & Sons.

Foss, S., Korshunov, D., and Zachary, S. (2013). *An introduction to heavy-tailed and subexponential distributions*. 2nd. Springer-Verlag New York.

Foss, S. and Richards, A. (2010). “On sums of conditionally independent subexponential random variables”. *Mathematics of Operations Research* 35.1, pp. 102–119.

Jagers, P. (1975). *Branching processes with biological applications*. John Wiley & Sons.

Resnick, S. I. (2007). *Heavy-tail phenomena: probabilistic and statistical modeling.* Springer-Verlag New York.

Wang, K., Chen, Y., and Tan, Z. (2014). "Asymptotics for the solutions to defective renewal equations". *Abstract and Applied Analysis* 2014.2, pp. 1–5.

Chapter 2

Optimal premium selection as function of the deductible

As in any other line of business, choosing the right price for a product in supply is extremely important for creating and maintaining a healthy company. In insurance, if the premium for an insurance policy is too low, it can lead to a loss per policy on average and eventually putting the solvency for the company at risk. On the contrary, if the premium is too high, the company risks its market share as they might loose customers to the competitors. Much thought power has therefore been invested into finding methods of premium calculation. In Paper B, we provide a way of finding the premium optimally with respect to minimising probability of ruin for insurance contracts including a deductible. This involves a three-stage analysis: First, we consider the customer's problem of choosing whether or not to insure. Next, we find the portfolio characteristics, and finally, we solve the company's minimization problem. It is an extension of the work in Asmussen et al. (2013). In preparatory to the paper, we want to consider an overview of the literature on the topic of premium selection. As the number of papers are numerous, only a fraction is mentioned here. A thorough review of premium calculation and their properties is provided in Laeven and Goovaerts (2008).

Early risk theory, commonly referred to as the classical or individual theory of risk, was mainly based on assessing risks individually. A portfolio was then simply aggregated as the sum of individual risks using techniques such as the central limit theorem and law of large numbers. Bohlmann (1909) reviews these methods. A fundamental development in the approach to risk theory on portfolio level was due to work of Lundberg (1909, 1919). The work was advanced mathematically, both the arguments, the techniques, and the notation. It included continuous-time stochastic processes, which not yet was rigorously defined at the time. Especially thanks to Cramér (1930, 1955), it was developed into a coherent theory, known as the collective theory of risk. The idea was to remove the focus from the risk of individuals, and zoom out to the balance of the insurance company at a collective level. The standard collective model by Cramér-Lundberg is considered in Section 2.1.4. As we see there, the model includes an initial reserve, premium payments arriving at a constant rate and claim payments modelled by compound Poisson process. A particular quantity of interest in relation to the reserve is the probability of ruin, which describes how exposed the insurance company is to insolvency. Ruin probability is the most considered measure of risk in actuarial science and has as a matter of consequence been studied extensively. In Section 2.1.4 we further state some of the classical results regarding the ruin probability. As seen here, analysing the ruin probability is rather complex for the Cramér-Lundberg model. Only in a few cases an explicit expression can be found. This motivates the use of diffusion approximations, known for their readily available passage probabilities.

Several extensions to the Cramér-Lundberg model has been proposed since. In Andersen (1957), the assumption of a Poisson counting process is relaxed to a renewal counting process as defined in Section 1.1.3. Hence, inter-arrival times are still assumed to be i.i.d., but generalised to have an arbitrary distribution rather than an exponential. In real-life data, however, seasonal effects and trends (in both claim frequency and sizes) are often observed, which is in disagreement with the i.i.d assumption of the inter-arrival times. A solution to this could be to incorporate an external environment process similar to the references mentioned in the introductory discussion of Chapter A. De Finetti (1957) criticised the fact that in the Cramer-Lundberg model, the reserve is expected to grow towards infinity when premiums are loaded (that is, higher than the expected claim expenses). As a response to this, the author argued that there must be an upper limit to the amount of reserve an insurance company is allowed to accumulate, and if that limit is reached, the excess is paid out as dividends to shareholders. Ruin is then certain, and therefore the expected present value of dividends paid until ruin is introduced as the performance measure. To the list of extensions we can add; investments, capital injections to fulfil capital requirements (e.g. Solvency II), reinsurance, etc.

An indispensable input to modern actuarial science is the theory of stochastic control, which provides methods of employing dynamic decision-making into insurance models. Its relevance was clearly stated by Borch (1967): “The theory of control processes seems to be ‘tailor-made’ for the problems which actuaries have struggled to formulate for more than a century”. The first paper on stochastic control was by Bellman (1958), the pioneer of dynamic programming (optimisation techniques for sequential decision-making problems). The theory was further developed by, inter alia, Florentin (1961) and Kushner (1962), who provided the first formal derivations of the Hamilton-Jacobi-Bellman equation for diffusions. The general theory of stochastic control is technically advanced. In Section 2.1.5 we make a short note on stochastic control in extension to the diffusion approximation of the reserve. We intend only to touch the surface of stochastic control, and therefore various regularity conditions are omitted and notation is simplified to fit in. Azcue and Muler (2014) and the monograph by Schmidli (2008) give proper treatments of stochastic control applied to insurance and provide comprehensive literature reviews.

In Section 2.1.2 we mention some standard premium principles. A premium principle is often based on the distributional properties of the risk. The most rudimentary premium principle is the net premium principle, which is simply the mean of the risk, unloaded. However, in order to avoid ruin, a loading is necessary.

This leads to the expected value premium principle, the variance premium principle, and the standard deviation premium principle. In insurance contracts the pure risk itself is often altered by incorporating a deductible. In Section 2.1.3 we define and discuss the purpose of deductibles. Insurance contracts between an individual and an insurance company often only comes with a so-called fixed amount deductible. This type of deductible structure is also the one considered in Paper B, C and D. For a nuanced discussion on deductible structures, we refer to Paper E.

Another direction is to consider utility functions. The concept utility was formalised by von Neumann and Morgenstern (1947). They show that an agent adhering to certain axioms of rationality has a utility function, and developed expected utility techniques to handle quantifiable risk. A complex decision between random outcomes then reduces to a comparison of real numbers. The construction of utility functions has lead to the theory of risk attitudes and risk premiums, as discussed in Friedman and Savage (1948) and Pratt (1964). Section 2.1.1 explains the relation between utilities, risk attitudes and risk premiums. Utility is also the cornerstone of the zero utility premium principle in Section 2.1.2, including the exponential premium principle. These premium principles are discussed in Gerber (1974) in relation to the property of additivity.

Combining utility functions with general equilibrium theory, Bühlmann (1980) showed that taking general market conditions into account rather than letting a risk be characterised by its own properties, it yields an economic premium principle, and in a special case of independence the economic premium principle corresponds to a change of measure studied by Esscher (1932), and is therefore referred to as the Esscher premium principle. Staying in the genre of measure changes, in financial theory inspiration to an additional way of pricing an insurable risk can be found based on an arbitrage-free argument leading to the existence of a risk neutral pricing measure. However, this is a discussion saved for Chapter 4.

In a different direction to the references mentioned so far, we have Asmussen (1999), where an adapted premium rule is considered. Here the claim payments are modelled by a compound Poisson process as in the Cramér-Lundberg model, but the premium rate is exclusively based on past claim statistics. In this setting, the ruin probability is studied.

We note that in Paper B, the individual is referred to as a customer, all though she has the option not to buy the insurance product offered.

2.1 Preliminaries

2.1.1 Utility and risk attitudes

A utility function $u(x)$ measures the satisfaction the individual gets of consuming or owning x. It is a numerical representation of an individual's preferences. It should be increasing as holding/consuming $x_1 > x_2$ must yields $u(x_1) \geq u(x_2)$.

Let X be a non-negative random variable representing a risk of the individual. An individual with wealth x can have one of the three attitudes towards risk

- *Risk averse*: The risk averse individual has a concave utility function, and by Jensen's inequality
$$\mathbb{E}[u(x-X)] \leq u(x-\mathbb{E}[X]).$$
Hence, the risk adverse individual prefers a predictable payoff at the cost of a possibly lower expected payoff. Examples of concave utility functions used in the literature are: exponential utility where $u(x) = -\exp(-ax))/a$ for $a > 0$, quadratic utility where $u(x) = x + ax^2$ for $a > 0$, and logarithmic utility where $u(x) = \log(x)$.

- *Risk neutral*: The risk neutral individual has a linear utility function,
$$\mathbb{E}[u(x-X)] = u(x-\mathbb{E}[X]).$$

- *Risk loving*: The risk loving individual has a convex utility function, and also by Jensen's inequality
$$\mathbb{E}[u(x-X)] \geq u(x-\mathbb{E}[X]).$$

From the above it is clear that only the risk averse individual would be willing to pay a premium larger than $\mathbb{E}[X]$ in order to avoid the risk X. If the individual has no insurance, and experiences a loss of size X, then she would have expected utility $\mathbb{E}[u(x-X)]$. On the other hand, if she is insured (with no deductible), she pays a premium p and is not exposed to the uncertainty of a loss and therefore has utility $u(x-p)$. The maximum premium the individual is willing to pay for insurance, denoted $\pi(X)$ and called the indifference premium, is chosen by

$$\mathbb{E}[u(x-X)] = u(x-\pi(X)), \tag{2.1}$$

i.e., where she is indifferent between insuring and not insuring. This creates an interval $[0, \pi(X)]$ of premium that the individual would be willing to accept.

2.1.2 Premium principles

A premium principle is a calculation rule for assigning a premium to a risk. Mathematically, if $\mathfrak{P}$ is the premium principle, then the premium charged by the insurance company for assuming the risk X is

$$p = \mathfrak{P}(X).$$

As we saw in the previous section, a fundamental assumption for the existence of an insurance market is that individuals are risk averse, and are therefore willing to pay a premium $p \geq \mathbb{E}[X]$ in order to avoid the risk X. If $p = \mathbb{E}[X]$ we say that the premium is actuarially fair, and if else $p > \mathbb{E}[X]$, we say that the premium includes a risk loading.

Next, we list a few standard example of premium principles, where θ is always positive.

- *The net premium principle*: $\mathfrak{P}(X) = \mathbb{E}[X]$. This is the actuarially fair unloaded premium.
- *Expected value premium principle*: $\mathfrak{P}(X) = (1+\theta)\mathbb{E}[X]$, where θ is the so-called safety loading. This premium principle has a risk loading proportional to the mean.
- *Variance premium principle*: $\mathfrak{P}(X) = \mathbb{E}[X] + \theta\mathbb{V}\mathrm{ar}[X]$. This premium principle has a risk loading proportional to the variance.
- *Standard deviation premium principle*: $\mathfrak{P}(X) = \mathbb{E}[X] + \theta\sqrt{\mathbb{V}\mathrm{ar}[X]}$. This premium principle has a risk loading proportional to the standard deviation.
- *Esscher's premium principle*: $\mathfrak{P}(X) = \mathbb{E}[X\exp(\theta X)]/\mathbb{E}[\exp(\theta X)]$.

In terms of utility functions, the two following premium principles are standard.

- *Equivalent utility principle*: Let $v(x)$ be the concave and increasing utility function of the insurer. The premium is chosen to satisfy
$$v(x) = \mathbb{E}[v(x + \mathfrak{P}_v(X) - X)] \tag{2.2}$$
where x is the initial wealth of the insurer.
- *Exponential premium principle*: If the utility function is exponential, $v(x) = -\exp(-ax)/a$, then by the equivalent utility principle
$$\mathfrak{P}(X) = \frac{1}{a}\log\big(\mathbb{E}[\exp(aX)]\big) \tag{2.3}$$

The equivalent utility principle corresponds to the analysis of (2.1) from the insurer's point of view. It is the minimum premium that the insurer is willing to accept for assuming the risk. Hence, all premiums larger than $\mathfrak{P}_v(X)$ would be accepted by the insurer. So in order for both the individual and the insurer to agree on a contract, the premium must be in the interval $[\pi(X), \mathfrak{P}_v(X)]$.

2.1.3 Deductible

A deductible holds the insured accountable for a part of the loss. Hence, if the insured experiences a loss X and has deductible structure h in her contract, then she must pay $h(X)$ herself, and the remaining $X - h(X)$ is covered by the insurance company, where h is a non-decreasing function. Adding a deductible to a contract serves the purpose of

- *Loss prevention*: The insured is responsible for covering a part of the loss, and can only claim the excess of the deductible to the insurance company. This motivates the insured to prevent a loss from occurring.
- *Loss reduction*: As the deductible structure should be non-decreasing, the insured has incentive to reduce the size of the loss.

- *Reduces administration costs*: A deductible can rule out small losses from being claimed. Small losses happen frequently, and administration costs of processing them might exceed the claimed amount.
- *Premium reduction*: Deductibles decreases the risks transferred to the insurance company, and as compensation the premium should be reduced, which can be an important aspect for the insured, especially for individuals that are less risk averse, and might prefer a high deductible in favour of a small premium.

The fixed amount deductible, considered in the paper of the present chapter, is defined by $h(X) = \min(X, K)$ for some $K \in \mathbb{R}^+$.

2.1.4 Reserve models and ruin probabilities

Reducing to the essentials, a collective model of an insurance company's reserve should include the following components:

- initial reserve r_0, which is the start-up capital of the insurance company.
- income from premium collection by a rate $c_t > 0$ at time t.
- the expenses of claims modelled by a random sum, $\sum_{i=1}^{N_t} X_i$ at time t, where $(N_t)_{t\geq 0}$ is a counting process representing the number of claims and $(X_i)_{i\in\mathbb{N}}$ are the non-negative claims sizes commonly assumed to be independent of $(N_t)_{t\geq 0}$.

Administrative and general costs, such as wages, rent, ect. are usually neglected. The reserve at time t is then

$$R_t = r_0 + \int_0^t c_t \mathrm{d}t - \sum_{i=1}^{N_t} X_i. \tag{2.4}$$

The associated claim surplus process is

$$S_t = \sum_{i=1}^{N_t} X_i - \int_0^t c_t \mathrm{d}t.$$

The Cramér-Lundberg model is the standard collective model. It assumes that the premium rate is constant, i.e. $c = c_t$, that $(N_t)_{t\geq 0}$ is a Poisson process as in Definition 1.2 with intensity λ, and that the claim sizes $(X_i)_{i\in\mathbb{N}}$ are i.i.d. and independent of $(N_t)_{t\geq 0}$, where we denote the common distribution by F with i'th moment $\overline{x_i}$. Further, in accordance with previous notation, let $(T_i)_{i\in\mathbb{N}_0}$ be the claim arrival times, and $(W_i)_{i\in\mathbb{N}}$ the inter-arrival times which, due to the Poisson assumption, are i.i.d. with an exponential distribution. This is also the model mentioned in Paper B.

A common concern in risk theory is the ruin probability,

$$\psi(r_0) = \mathbb{P}(\tau_{r_0} < \infty)$$

where $\tau_{r_0} = \inf\{t \geq 0 : R_t < 0 \mid R_0 = r_0\}$ is the time of ruin, i.e. the first time that the reserve becomes negative. The probability of survival is then $\phi(r_0) = 1 - \psi(r_0)$. Some is also concerned with the finite time ruin probability $\psi(r_0, T) = \mathbb{P}(\tau_{r_0} < T)$, see, e.g., De Vylder and Goovaerts (1988) and Lefévre and Loisel (2008). However, we exclusively consider the infinite-time ruin here. Further, we assume that the net profit condition, $c > \overline{x_1}\lambda$, is satisfied. Otherwise, ruin happens with probability one. Intuitively, the net profit condition says that on average, we should expect the ingoing premium to at least cover the outgoing claims. This is equivalent with having a positive safety loading, i.e. $\theta = c - \lambda\,\overline{x_1} > 0$. Otherwise, the reserve would drift negatively, eventually causing ruin with certainty.

The Pollaczek-Khinchin formula represents the ruin probability as the tail of a compound geometric distribution,

$$\psi(r_0) = \Big(1 - \frac{\overline{x_1}\lambda}{c}\Big) \sum_{n=0}^{\infty} \Big(\frac{\overline{x_1}\lambda}{c}\Big)^n F_{X,I}^{*(n)}(r_0), \tag{2.5}$$

where $F_{X,I}$ is the integrated tail distribution

$$F_{X,I}(r_0) = \frac{1}{\overline{x_1}} \int_0^{r_0} \overline{F}(u)\mathrm{d}u.$$

From this formula, it follows that the ruin probability is available explicitly in the case of exponentially distributed claim sizes, i.e. $F(x) = 1 - \exp(-bx)$ if $x > 0$ with $b > 0$. The ruin probability then simplifies to

$$\psi(r_0) = \frac{1}{1+\theta} \exp\Big(-\frac{b\theta}{1+\theta} r_0\Big).$$

Another case where it is possible to get an exact formula is when $r_0 = 0$, then $\psi(0) = (1+\theta)^{-1}$.

Implicitly stated above, the cases where the ruin probability can be found explicitly are few. Therefore the focus in the literature has instead been redirected at creating upper bounds for the ruin probability and examining the asymptotic behaviour. We next outline some of the standard results within this regard of ruin theory. For further interest, we refer to Asmussen and Albrecher (2010). We distinguish according to the heavy-tailedness of the claim size distribution in the sense of Section 1.1.6.

Recall from Section 1.1.6 that claims are said to be light-tailed distributed if there exists a $\nu > 0$ such that $m_X(\nu) < \infty$. In this case, the moment generating function of the claim surplus process,

$$m_{S_t}(x) = \mathbb{E}[\exp(xS_t)] = \exp\big(t(\lambda(m_X(x) - 1) - cx)\big),$$

is also finite on $(0, \nu)$. Applying the logarithm yields the cumulant generating function

$$\kappa_{S_t}(x) = \log(m_{S_t}(x)) = t\big(\lambda(m_X(x) - 1) - cx\big). \tag{2.6}$$

Note that $\kappa_{S_t}(0) = 0$ and has negative slope in this point as

$$\kappa'_{S_t}(0) = t(\lambda\,\overline{x_1} - c) < 0 \qquad \text{for } t \geq 0$$

due to the net profit condition. $\kappa_{S_t}(x)$ is furthermore a convex function since

$$\kappa''_{S_t}(x) = t\lambda\mathbb{E}[X^2\exp(xX)] > 0 \qquad \text{for } t \geq 0.$$

From the established properties it follows that if there exists a $\gamma > 0$ such that $\kappa_{S_t}(\gamma) = 0$, or equivalently, such that

$$m_X(\gamma) = \frac{c\gamma}{\lambda} + 1, \tag{2.7}$$

then it is unique. This γ is known as the adjustment coefficient (or the Lundberg exponent). It features in two essential results in ruin theory.

The first is Lundberg's inequality from Lundberg (1909) providing an upper bound on the ruin probability.

Theorem 2.1 (Lundberg's inequality)**.** *Assuming that the adjustment coefficient γ exists, then*

$$\psi(r_0) \leq \exp(-\gamma r_0).$$

The second gives the asymptotics of the ruin probability and is due to Cramér (1930).

Theorem 2.2 (Cramer's ruin bound)**.** *Assume that the F has a density and that the adjustment coefficient γ exists, then there is a constant C such that*

$$\psi(r_0) \sim C\exp(-\gamma r_0) \qquad \textit{as } r_0 \to \infty.$$

In the proof of *Cramer's ruin bound*, it is useful to write the ruin probability as a defective renewal equation,

$$\psi(r_0) = q\overline{F}_{X,I}(r_0) + \int_0^{r_0} \psi(r_0 - x)\mathrm{d}(qF_{X,I}(x))$$

where $q = \lambda\,\overline{x_1}/c < 1$. We have previously seen this type of integral equation in Section 1.1.3. Expressing the ruin probability as a renewal equation makes it possibly to use renewal theory techniques to find the asymptotic properties.

For heavy-tailed claims, $m_X(\nu) = \infty$ for all $\nu > 0$. An adjustment coefficient can therefore not exist, and the analysis above becomes inapplicable. Instead, considering the subexponential subclass of heavy-tailed distributions, Embrechts and Veraverbeke (1982) show that the ruin probability asymptotically behaves as follows.

Theorem 2.3. *Assume that $\overline{x_1} < \infty$, that the claim size distribution has a density, and that the integrated claim size distribution $F_{X,I}$ is subexponential. Then the ruin probability satisfies*

$$\psi(r_0) \sim \frac{\lambda\,\overline{x_1}}{c - \lambda\,\overline{x_1}}\overline{F}_{X,I}(r_0) \qquad \textit{as } r_0 \to \infty.$$

It now becomes apparent how heavy-tailed claims endangers the solvency of the insurance company. Theorem 2.2 states that for light-tailed claims the ruin probability decays exponentially asymptotically, whereas in Theorem 2.3 we see that for heavy-tailed claims, the ruin probability is of the same order as the integrated tail of the claims. Hence, in the heavy-tailed case ruin can be spontaneously caused by one large claim, corresponding to the principle of a single big jump discussed in Section 1.1.6

The fact that first passage times are more easily calculated for diffusion processes has lead to the idea of approximating the reserve by a diffusion. Iglehart (1969) shows how a sequence of risk reserve processes converges weakly to the Brownian motion with drift,

$$dR_t^{(d)} = \mu \mathrm{d}t + \sigma \mathrm{d}W_t, \qquad R_0^{(d)} = r_0 \tag{2.8}$$

where $\mu = c - \lambda\,\overline{x_1}$ and $\sigma^2 = \lambda\,\overline{x_2}$. The argument is based on compressing the time scale, such that claims become many and small, meanwhile maintaining the first two moment of the original reserve process. Iglehart (1969) further shows that for a positive drift (corresponding to the net profit condition being satisfied), the ruin probability of the sequence of risk reserves then converges to the ruin probability of the diffusion approximation given by

$$\mathbb{P}(\tau_{r_0}^{(d)} < \infty) = \exp\Big(-2\frac{r_0\mu}{\sigma^2}\Big).$$

where $\tau_{r_0}^{(d)} = \inf\{t \geq 0 : R_t^{(d)} < 0 \mid R_0^{(d)} = r_0\}$ is the time of ruin for the diffusion approximation.

We use this diffusion approximation to model the insurance company's reserve in Paper B, where the simple nature of the associated ruin probability is convenient in the sense that a closed-form solution is obtainable for the optimal single-customer premium.

2.1.5 A short note on stochastic control theory

Until now, we have only considered the bare case of the reserve model, where no control is directly implemented. In reality, how the reserve develops depends on a number of decisions made by the insurance company. We turn to the theory of stochastic control to give a suitable formulation.

Assume that the reserve develops according the stochastic process $(R_t^u)_{t\geq 0}$ with dynamics governed by a control process $u = (u_t)_{t\geq 0}$ as follows

$$\mathrm{d}R_t^u = \mu(R_t^u, u_t)\mathrm{d}t + \sigma(R_t^u, u_t)\mathrm{d}W_t, \qquad R_0^u = r_0. \tag{2.9}$$

The control process u must take value in some admissible set U. The minimum restrictions on admissibility are that the control has to be adapted to the filtration generated by $(W_t)_{t\geq 0}$, since it must only be based on events up to present time, and that the dynamics in (2.9) admits a strong solution.

A value $V^u(t,x)$ is associated depending on the control process u and the current value $x = R_t^u$ of the reserve. The objective of interest is then the value function evaluated in the optimal control, $V(t,x) = \sup_u V^u(t,x)$. Obvious questions to ask are; does there exist an optimal control, and if there does, how do we find it? Further, what is the value function?

Letting τ be an exit time, and say that the insurance company wants to maximise

$$V^u(t,x) = \mathbb{E}\Big[\int_t^\tau \mathrm{e}^{-\zeta(s-t)} \upsilon(R_s^u, u_s)\mathrm{d}s + \mathrm{e}^{-\zeta(\tau-t)}\mathcal{K}(\tau, R_\tau^u) \;\Big|\; R_t^u = x, t < \tau\Big]$$

where $\upsilon : \mathbb{R} \times U \to \mathbb{R}$ is a continuous reward rate, $\mathcal{K} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ a final reward, and ζ a discounting parameter. So the insurance company wants to find the value function $V(t,x)$ and the optimal control process u^* such that

$$V(t,x) = \sup_{u \in U} V^u(t,x) = V^{u^*}(t,x)$$

There are different types of controls depending on how it reacts to the reserve process, e.g.

- *Open loop controls*: The control acts independently of the reserve process.
- *Closed loop (or feedback) controls*: The control process is adapted to the filtration generated by $(R_t^u)_{t \geq 0}$.
- *Markovian controls*: A subset of the closed loop controls, where the control only depends on the current level of the reserve and not on the past. With such controls, the diffusion (2.9) becomes a Markov process, which explains the name.

In many problems, the optimal control is on a Markovian form, i.e. $u_t^* = u^*(R_t^*)$, where $(R_t^*)_{t\geq 0}$ is the optimally controlled reserve. This is also the only type of control considered here.

Paper B considers a setting where the insurance company offers an insurance product with a premium p and fixed amount deductible K. It is stated in the paper how the drift and variance of the reserve diffusion process are controlled by p, which then is chosen to minimise the ruin probability, or, equivalently, maximise survival probability. Compared to the general formulation above, this corresponds to letting $\tau^u = \tau^u(r_0) = \inf\{t \geq 0 : R_t^u < 0 \mid R_0^u = r_0\}$ be the time of ruin of the controlled process, $\upsilon(R_s^u, u_s) = 0$ for any control $u_s \in U$ at any time $s > 0$, $\zeta = 0$, and $\mathcal{K}(\tau^u, R_{\tau^u}^u) = \mathbb{1}_{\{\tau^u = \infty\}}$, such that

$$\begin{aligned} V^u(t,x) &= \mathbb{E}[\mathbb{1}_{\{\tau^u=\infty\}} \mid R_t^u = x, t < \tau] = \mathbb{P}(\tau^u(x) = \infty) \\ &= 1 - \mathbb{P}(\tau^u(x) < \infty) = 1 - \psi(x) = \phi(x). \end{aligned}$$

Due to the use of diffusion approximation, we see in the ending discussion of the previous section that the analysis is significantly simplified, which is also apparent in Paper B, where we find that the optimal premium strategy is constant.

However, in cases where the value function is not directly available, one continues by using dynamic programming techniques. Applying these tools, one is able to show that the value function can be characterised by the so-called Hamilton-Jacobi-Bellman (HJB) equation, a partial differential equation describing the local behaviour of the value function,

$$\sup_{u\in U}\{\mathcal{A}V(x) - \delta V(x) + v(x,u)\} = 0,$$

where

$$\mathcal{A}f(x) = \lim_{t\downarrow 0}\frac{1}{t}\mathbb{E}[f(R_t^u) - f(x) \mid R_0^u = x] = \mu(x,u)f'(x) + \frac{1}{2}\sigma^2 f''(x)$$

is the infinitesimal generator of the process $(R_t^u)_{\{t\geq 0\}}$ with dynamics (2.9).

In this way, the global optimisation problem is reduced to a continuum of local ones. Two important notes to this extend are: i) a solution to the HJB-equation is not necessarily the value function in consideration. One needs to prove a so-called verification theorem to ensure. ii) the HJB-equation obviously requires that the value function is sufficiently smooth, which might not be the case. Then one needs to turn to viscosity solutions. That is, however, well beyond the scope here.

References

Andersen, E. S. (1957). "On the collective theory of risk in case of contagion between claims". *Bulletin of the Institute of Mathematics and its Applications* 12.2, pp. 275–279.

Asmussen, S. (1999). "On the ruin problem for some adapted premium rules". *Probabilistic Analysis of Rare Events*, pp. 3–15.

Asmussen, S. and Albrecher, H. (2010). *Ruin probabilities*. 2nd. World Scientific.

Asmussen, S., Christensen, B. J., and Taksar, M. (2013). "Portfolio size as function of the premium: modelling and optimization". *Stochastics* 85.4, pp. 575–588.

Azcue, P. and Muler, N. (2014). *Stochastic optimization in insurance: a dynamic programming approach*. Springer-Verlag New York.

Bellman, R. (1958). "Dynamic programming and stochastic control processes". *Information and control* 1.3, pp. 228–239.

Bohlmann, G. (1909). "Die Theorie des mittleren Risikos in der Lebensversicherung". *Transactions of the 6th International Congress of Actuaries* 1, pp. 593–683.

Borch, K. (1967). "The theory of risk". *Journal of the Royal Statistical Society. Series B (Methodological)* 29.3, pp. 432–467.

Bühlmann, H. (1980). "An economic premium principle". *ASTIN Bulletin* 11.1, pp. 52–60.

Cramér, H. (1930). *On the mathematical theory of risk*. Försäkringsaktiebolaget Skandia 1855-1930.

— (1955). "Collective risk theory". *Jubilee Volume of Försäkringsaktiebolaget Skandia*, pp. 1–92.

De Finetti, B. (1957). "Su un'impostazione alternativa della teoria collettiva del rischio". *Transactions of the XVth International Congress of Actuaries* 2.1, pp. 433–443.

De Vylder, F. and Goovaerts, M. J. (1988). "Recursive calculation of finite-time ruin probabilities". *Insurance: Mathematics and Economics* 7.1, pp. 1–7.

Embrechts, P. and Veraverbeke, N. (1982). "Estimates for the probability of ruin with special emphasis on the possibility of large claims". *Insurance: Mathematics and Economics* 1.1, pp. 55–72.

Esscher, F. (1932). "On the probability function in the collective theory of risk". *Scandinavian Actuarial Journal* 1932.3, pp. 175–195.

Florentin, J. J. (1961). "Optimal control of continuous time, Markov, stochastic systems". *Journal of Electronics and Control* 10.6, pp. 473–488.

Friedman, M. and Savage, L. J. (1948). "The utility analysis of choices involving risk". *Journal of Political Economy* 56.4, pp. 279–304.

Gerber, H. U. (1974). "On additive premium calculation principles". *ASTIN Bulletin* 7.3, pp. 215–222.

Iglehart, D. L. (1969). "Diffusion approximations in collective risk theory". *Journal of Applied Probability* 6.2, pp. 285–292.

Kushner, H. (1962). "Optimal stochastic control". *IRE Transactions on Automatic Control* 7.5, pp. 120–122.

Laeven, R. and Goovaerts, M. J. (2008). "Premium calculation and insurance pricing". *Encyclopedia of Quantitative Risk Analysis and Assessment*. Vol. 3. John Wiley & Sons, pp. 1302–1314.

Lefévre, C. and Loisel, S. (2008). "On finite-time ruin probabilities for classical risk models". *Scandinavian Actuarial Journal* 2008.1, pp. 41–60.

Lundberg, F. (1909). "Zur Theorie der Rückversicherung". *Transactions of the 6th International Congress of Actuaries* 1, pp. 877–955.

— (1919). "Teorin för riskmassor". *Försäkringsinpektionen* Stockholm.

Pratt, J. W. (1964). "Risk aversion in the small and in the large". *Econometrica* 32.1, pp. 122–136.

Schmidli, H. (2008). *Stochastic control in insurance*. Springer-Verlag London.

von Neumann, J. and Morgenstern, O. (1947). *Theory of games and economic behavior*. 2nd. Princeton University Press.

Paper B

Optimal premium as a function of the deductible: Customer analysis and portfolio characteristics

JULIE THØGERSEN

ABSTRACT. An insurance company offers an insurance contract (p, K), consisting of a premium p and a deductible K. In this paper, we consider the problem of choosing the premium optimally as a function of the deductible. The insurance company is facing a market of N customers, each characterised by their personal claim frequency, α, and risk aversion, β. When a customer is offered an insurance contract, she will, based on these characteristics, choose whether or not to insure. The decision process of the customer is analysed in detail. Since the customer characteristics are unknown to the company, it models them as i.i.d. random variables; $A_1, \dots, A_N$ for the claim frequencies and $B_1, \dots, B_N$ for the risk aversions. Depending on the distributions of A_i and B_i, expressions for the portfolio size $n(p; K) \in [0, N]$ and average claim frequency $\alpha(p; K)$ in the portfolio are obtained. Knowing these, the company can choose the premium optimally, mainly by minimising the ruin probability.

KEYWORDS: microeconomic insurance; customer characteristics; portfolio size; average claim frequency; ruin theory

Publication details: Risks, 2016, vol. 4(4), 42. DOI: 10.3390/risks4040042

B.1 Introduction

An insurance company has several instruments for stochastic control at its disposal. Much studied are dividends, reinsurance and investment; see, e.g., Schmidli (2008). The present paper concentrates on premiums and deductibles. These are also obvious instruments, but have been somewhat less studied in the literature as such.

The standard model for the risk reserve R_t at time t is the Cramer–Lundberg process:

$$R_t = r_0 + ct - A_t,$$

where r_0 is the initial reserve, c is the gross premium rate and $(A_t)_{t\geq 0}$ is a compound Poisson process with parameters λ and F. More explicitly, $A_t = \sum_{i=1}^{N_t} Z_i$ where $(N_t)_{t\geq 0}$ is a Poisson process with parameter λ counting the number of claims until time t, and the Z_i's represent the (positive) claim sizes assumed to be i.i.d. and independent of $(N_t)_{t\geq 0}$, with common distribution F on $(0,\infty)$. Let $\overline{z_1} = \mathbb{E}[Z_i]$ and $\overline{z_2} = \mathbb{E}[Z_i^2]$. The Cramer–Lundberg process can, due to the arguments in Iglehart (1969), be approximated by the diffusion process:

$$\mathrm{d}R_t^{(d)} = \mu \mathrm{d}t + \sigma \mathrm{d}W_t, \quad R_0^{(d)} = r_0 \tag{2.1}$$

with $\mu = c - \lambda\,\overline{z_1}$ and $\sigma^2 = \lambda\,\overline{z_2}$. To put it briefly, this diffusion process is the Brownian motion with drift that matches the mean and variance of the Cramer–Lundberg process at any given point in time.

We will describe an insurance contract by a premium p and a deductible K. The purpose, besides loss prevention and retention, of adding a deductible to a contract is to avoid administrating the numerous number of small claims. The deductible is therefore chosen to serve this purpose, and the premium will then be considered as a function of the chosen deductible. We will consider the methods of how to find the premium optimally in a given market. For simplicity, we assume that all (potential) customers are offered the same insurance contract (p, K). We will also neglect the market effects by assuming that only one company supplies insurance.

A possible extension is to allow the insurance company to offer different deductibles and let the premium in the contract be regulated accordingly. A simple way of doing so is to say that the entire market can be divided into disjoint sub-markets, for example; one where there is a customer demand for a low deductible K_l, one for a medium deductible K_m and one for a high deductible K_h. Therefore, the offered values of the deductibles will be $\{K_l, K_m, K_h\}$, where there is a separate market for each. It will then be possible to apply the same approach as considered here to each of the sub-markets.

There are several types of deductibles. We will consider the classical fixed amount deductible where the claims are truncated, so the loss for the insurance company in relation to a claim Z_i can be described by the random variable:

$$X_i^{(K)} = (Z_i - K)^+ = \begin{cases} 0, & \text{if } Z_i < K, \\ Z_i - K, & \text{if } Z_i \geq K. \end{cases}$$

The risk reserve must therefore be modified:

$$R_t = r_0 + ct - A_t^{(K)},$$

where $A_t^{(K)}$ is the compound Poisson process with losses $X_i^{(K)}$. Let $\overline{x_n^{(K)}}$ denote the n'th moment of $X_i^{(K)}$. Notice that $\overline{z_1} = \overline{x_1^{(0)}}$ and $\overline{z_2} = \overline{x_2^{(0)}}$, since the claims are assumed to be positive. The results here thus contain the no-deductible case.

Assume that the insurance company is facing a market consisting of N potential customers. Let $n(p;K) \in [0, N]$ denote the number of customers the insurance company attracts on the market when offering the contract (p, K). Obviously, increasing the premium should lead to a loss of customers; therefore, it must be that $\frac{\partial n(p;K)}{\partial p} \leq 0$. Furthermore, the average customer claim frequency in the portfolio will be denoted as $\alpha(p;K)$. Raising the premium will make it less attractive to insure for customers having low claim frequencies, and so, the average claim frequency of the portfolio increases, i.e., $\frac{\partial \alpha(p;K)}{\partial p} \geq 0$. This is commonly known as adverse selection. The gross premium c and the aggregate claim frequency λ will then depend on the premium, p, and the chosen deductible, K, as follows:

$$c(p;K) = n(p;K)p \quad \text{and} \quad \lambda(p;K) = n(p;K)\alpha(p;K).$$

The drift and variance of the diffusion process (2.1) are modified accordingly:

$$\begin{aligned}\mu(p;K) &= n(p;K)p - \lambda(p;K)\overline{x_1^{(K)}} - L = n(p;K)\big(p - \alpha(p;K)\overline{x_1^{(K)}}\big) - L,\\ \sigma^2(p;K) &= \lambda(p;K)\overline{x_2^{(K)}} = n(p;K)\alpha(p;K)\overline{x_2^{(K)}}.\end{aligned}$$

In order to avoid trivialities, when minimising the ruin probability at a later stage, a fixed liability payment rate L is introduced in the drift. Otherwise, the insurance company can choose p to be infinitely large such that no customer will insure, and the reserve will remain constantly at r_0, yielding a ruin probability of zero. The term liability is used in a broad sense. It covers any type of costs the insurance company might have. Let $\varphi(r_0)$ denote the ruin probability as a function of the initial reserve when the reserve is modelled by the diffusion process. For more background on ruin theory, we refer to Asmussen and Albrecher (2010). In this reference, it is also stated that when the drift $\mu(p;K)$ in the diffusion approximation is positive, then $\varphi(r_0) = \exp(-r_0\mu(p;K)/\sigma^2(p;K))$, and when the drift is negative, then ruin will be certain, i.e., $\varphi(r_0) = 1$ for any initial reserve r_0.

For a given deductible, changing the premium will have a double-sided effect on the drift (and profit) of the insurance company. Raising the premium will increase the earnings per customer, but will also reduce the size of the portfolio and increase the average claim rate due to adverse selection, and vice versa for decreasing the premium. In order to say which effect is dominating, a specification of the portfolio characteristics is needed.

First of all, we need to gain insight into the decision process of a customer. In this context, we introduce a risk aversion parameter β and motivate a method for incorporating risk aversion. See Rees and Wambach (2008) for a microeconomic perspective of insurance.

In a naive setting, the customers would have the same claim rate, implying the average claim frequency to be constant, $\alpha(p;K) = \alpha$, and $n(p;K)$ could be chosen on some ad hoc form. In a more realistic setting, a market of potential customers is non-homogeneous in the sense that they have different characteristics, namely different α's and β's. A customer knows her own claim rate, but the insurer does not possess this information about the customers. The claim rates are therefore modelled as i.i.d. random variables $A_1, \ldots, A_N$ over the portfolio. Likewise for the risk aversion parameter, which will be considered as the outcomes of the i.i.d. random variables $B_1, \ldots, B_N$. This results in some less ad hoc forms for the functions $\alpha(p;K)$ and $n(p;K)$ to characterise the portfolio. Once the portfolio characteristics are known, the insurance company can use these to choose the premium optimally. The main optimization problem considered is minimising the ruin probability, but this only makes sense in case of a positive drift in the diffusion (2.1). In the case of a negative drift ruin is certain, so the premium will be chosen to maximise the time to ruin.

The ideas here are very similar to the ones considered in Asmussen et al., 2013. The present paper takes however a different approach to risk aversion and, as said, incorporates deductibles. The work in Burnecki et al., 2004 also finds the price of insurance as a function of the deductible for different types of deductibles, though not exploiting the aspects of risk aversion. The work in Højgaard, 2002 also controls the gross premium indirectly by controlling the safety loading. All in all, the contributions of this paper are three-fold; analyzing the customer's behavior in Section B.2, finding portfolio characteristics in Section B.3 and choosing the optimal premium for the insurance company in Section B.4. For examples and illustrations, see Section B.5.

B.2 Customer's problem

A potential customer has to make a decision on whether to insure or not, given that she is offered an insurance contract (p, K). If the customer chooses to insure, she must pay a premium p at constant rate, but the customer can then report claims and get the amount above the threshold K covered. More specifically, if the customer experience a loss Z_i, she will only report it to the insurance company if $Z_i \geq K$ in which case she has to pay K herself; otherwise, if $Z_i < K$, there is no purpose of reporting it, and she will have to cover the entire loss Z_i. On the other hand, if the customer chooses not to insure, she no longer has to pay p continuously. She will instead have to cover all of the uninsured losses by herself.

For the moment, risk aversion is ignored, and the decision is made solely by comparing the present values of the wealth generated by the two options. Later risk aversion will be incorporated by pricing the excess uncertainty when not insuring using the variance premium principle. Let V_I denote the present value of insuring and V_{NI} of not insuring.

The customer is assumed to have a subjective discount rate d and access to a risk-free asset with interest rate r in which all her wealth is assumed to be invested. Furthermore, it is assumed that $d > r$ to ensure finite asset valuation. The customer is characterised by initial wealth w_0 and claim frequency α. The customer is furthermore assumed to have infinite life length. Remark B.2 comments on this assumption. The problem of the customer will then be identical every period, and the decision will therefore not change over time.

We start by finding the present value of not insuring. As said, if the customer chooses not to insure, she will have to cover every loss herself. Therefore, her wealth will develop according to:

$$\mathrm{d}w_t = rw_t\mathrm{d}t - \mathrm{d}A_t^\alpha, \tag{2.2}$$

where $(A_t^\alpha)_{t\geq 0}$ is a compound Poisson process with parameters α and F, representing the total loss until time t for the potential customer when not insuring. Hence, $A_t^\alpha = \sum_{i=1}^{N_t^\alpha} Z_i$ where $(N_t^\alpha)_{t\geq 0}$ is a Poisson process with parameter α assumed to be independent of the Z_i's. Let $(T_i)_{i\in\mathbb{N}}$ denote the arrival times of the Poisson process $(N_t^\alpha)_{t\geq 0}$. Notice that (2.2) is an Ornstein–Uhlenbeck process driven by a Levy process (namely, the compound Poisson process), and the solution is therefore explicitly known as:

$$w_t = \exp(rt)\Big(w_0 - \int_0^t \exp(-rs)\mathrm{d}A_s^\alpha\Big) = \exp(rt)\Big(w_0 - \sum_{i=1}^{N_t^\alpha} \exp(-rT_i)Z_i\Big).$$

The present value of the wealth when not insuring is evaluated as:

$$V_{NI} = \mathbb{E}\Big[\int_0^\infty \exp(-dt)\mathrm{d}w_t\Big].$$

Calculations given in Appendix B.I show that this can be reduced to:

$$V_{NI} = \frac{rw_0}{d-r} - \frac{\overline{z_1}\alpha}{d-r}.$$

On the other hand, if the customer chooses to insure, she will have to pay a premium continuously and cover the parts of the claims below the deductible. Her wealth will then have the following dynamics,

$$\mathrm{d}w_t = (rw_t - p)\mathrm{d}t - \mathrm{d}A_t^{\alpha,K}, \tag{2.3}$$

where $(A_t^{\alpha,K})_{t\geq 0}$ is the compound Poisson process representing the total loss associated with claims until time t for the customer when insuring, that is:

$$A_t^{\alpha,K} = \sum_{i=1}^{N_t^\alpha} \min\{Z_i, K\} = \sum_{i=1}^{N_t^\alpha} \left(K\mathbb{1}_{\{Z_i\geq K\}} + Z_i\mathbb{1}_{\{Z_i<K\}}\right).$$

Once again, we are looking at an Ornstein–Uhlenbeck process driven by a Levy process. This can be solved in the same way as seen previously,

$$\begin{aligned} w_t &= \exp(rt)\Big(w_0 - \int_0^t \exp(-rs)p\mathrm{d}t - \int_0^t \exp(-rs)\mathrm{d}A_s^{\alpha,K}\Big) \\ &= \exp(rt)\left(w_0 - \frac{p}{r}\right) + \frac{p}{r} - \exp(rt)\sum_{i=1}^{N_t^\alpha} \exp(-rT_i)(K\mathbb{1}_{\{Z_i\geq K\}} + Z_i\mathbb{1}_{\{Z_i<K\}}). \end{aligned}$$

The present value of the wealth when insuring will therefore be:

$$V_I = \mathbb{E}\Big[\int_0^\infty \exp(-dt)\mathrm{d}w_t\Big] = \frac{rw_0 - p}{d-r} - \frac{\alpha}{d-r}\mathbb{E}[K\mathbb{1}_{\{Z_i \geq K\}} + Z_i\mathbb{1}_{\{Z_i<K\}}].$$

Since the approach and calculations are very similar to the ones used when finding V_{NI}, the details are skipped. We can now state the following.

Corollary B.1. *Disregarding risk aversion, the customer will insure if and only if $V_I \geq V_{NI}$, which is equivalent to:*

$$p \leq \alpha\mathbb{E}[(Z_i - K)\mathbb{1}_{\{Z_i \geq K\}}] = \alpha\mathbb{E}[X_1^{(K)}] = \alpha\overline{x_1^{(K)}}$$

Hence, a risk-neutral customer will insure if the net premium (taking deductibles into account) exceeds the premium. She has no incentive to pay a loading in order to avoid the risk.

Remark B.2. The customer is assumed to have an infinite life length. This is obviously not realistic, but it is convenient for the analysis and not uncommon in other areas of non-life insurance. For example, it the basis for the standard use of the stationary distribution in bonus-malus systems; see the discussion in Asmussen (2014).

It would be more realistic with a random life length τ. In this case, the discount factor $\exp(-dt)$ has to be replaced by $\mathbb{P}(\tau > t)\exp(-dt)$ in the above formulas for the present values, V_{NI} and V_I. This will in general change the analytic expressions. Though note that in the case where τ is exponential with rate, say, ρ, the same expressions appears with d replaced by $d + \rho$. Due to the forgetfulness property of the exponential distribution, the customer is still facing an identical problem every period. This leads to obvious reinterpretations of the analysis of this section.

In order for insurance to make sense, the customer must of course have some degree of risk aversion. We want to find which excess risk the customer is exposed to when not insuring and how to price this risk when risk aversion is essential.

The first step is to notice that V_{NI} also could have been derived in a more intuitive way. Recall that all wealth is assumed to be invested in the risk-free asset. Therefore, wealth itself has the dynamics $\mathrm{d}w_t = rw_t\mathrm{d}t$ of a bank account. When accumulating interest, this has present value:

$$\int_0^\infty \exp(-dt)\mathrm{d}w_t = \int_0^\infty \exp(-(d-r)t)rw_0 dt = \frac{rw_0}{d-r}. \tag{2.4}$$

When a non-insured customer has to pay a loss Z_i at time T_i, she also loses a possible interest rate income. Therefore, the total loss of paying Z_i at time T_i can be calculated using the formula (2.4) with the discounted loss $\exp(-rT_i)Z_i$ as w_0. Hence,

$$V_{NI} = \frac{rw_0}{d-r} - \frac{r}{d-r}\mathbb{E}\Big[\sum_{i=1}^{\infty}\exp(-rT_i)Z_i\Big].$$

Note that due to Campbell–Mecke's formula (see, e.g., van Lieshout, 2000 for further details), this is in fact equal to the expression for V_{NI} found previously. Furthermore, we have the following alternative characterization of V_I,

$$V_I = \frac{rw_0 - p}{d-r} - \frac{r}{d-r}\mathbb{E}\Big[\sum_{i=1}^{\infty} \exp(-rT_i)(K\mathbb{1}_{\{Z_i \geq K\}} + Z_i \mathbb{1}_{\{Z_i < K\}})\Big]$$

which could have been explained with a similar intuitive approach using truncated losses.

The criteria of insuring in Corollary B.1 when disregarding risk aversion can then be expressed as the inequality:

$$p \leq \mathbb{E}\Big[r \sum_{i=1}^{\infty} \exp(-rT_i)(Z_i - K)\mathbb{1}_{\{Z_i \geq K\}}\Big].$$

The additional risk the customer is exposed to when not insuring can therefore be captured by the random variable:

$$S^{(K)} = r\sum_{i=1}^{\infty} \exp(-rT_i)(Z_i - K)\mathbb{1}_{\{Z_i \geq K\}} = r\sum_{i=1}^{\infty} \exp(-rT_i)X_i^{(K)}.$$

The next step is to find the maximum premium the risk averse customer is willing to pay, also called her reservation price for this risk. Inspired by microeconomics, one can let the customer's preferences be represented by a concave utility function $u(\cdot)$. Assume that the only risk the customer is concerned about when pricing is the additional risk $S^{(K)}$ she is carrying when not insuring. The losses less than or equal to the deductible is a risk that the customer is also facing, but cannot be insured against; see Remark B.5 for further details. Inspired by Gerber and Pafum (1998), the maximal premium $P(K)$ the customer is willing to pay for an insurance of a risk $S^{(K)}$ is the solution to the equation:

$$\mathbb{E}[u(w - S^{(K)})] = u(w - P(K)). \tag{2.5}$$

Therefore, the customer's reservation price satisfies that she is indifferent between carrying the risk $S^{(K)}$ and paying the premium $P(K)$.

A possible choice is to let the customer's preferences be represented by an exponential utility function,

$$u_a(x) = \frac{1}{a}(1 - \exp(-ax)),$$

where a is a risk aversion parameter. The premium in (2.5) can then be solved explicitly as:

$$P_{\exp}(K) = \frac{1}{a}\log\big(\mathbb{E}\big[\exp(aS^{(K)})\big]\big). \tag{2.6}$$

This is the same premium as found in Gerber (1974), where it is the insurer, not the insured, pricing a risk. It is commonly known as the exponential premium principle and is mostly used by the insurer in the literature.

The object is to obtain an analytic expression for the premium, and this is extremely complicated (if not impossible without having to make a lot of simplifying assumptions) to get for $P_{\text{exp}}(K)$. This is mainly due to $S^{(K)}$ being a sum of dependent random variables caused by the dependence structure in the T_i's. Thus, the additivity property that Gerber (1974) proves the exponential premium principle possesses cannot be used. To illustrate this complexity, a simple case example is presented in Appendix B.III. In order to obtain a more simple expression, a second order Taylor approximation of the logarithm in (2.6) around $a = 0$ is considered,

$$P_{\text{exp}}(K) \approx \mathbb{E}[S^{(K)}] + \frac{a}{2}\mathbb{V}\text{ar}[S^{(K)}].$$

Assessments about the premium will be based on the right-hand side hereinafter. Let $\beta = a/2$ and introduce the notation:

$$P_{\text{var}}(K) = \mathbb{E}[S^{(K)}] + \beta\mathbb{V}\text{ar}[S^{(K)}]. \tag{2.7}$$

This the well-known variance premium principle. It expands the net premium principle by adding a risk loading that is proportional by a factor $\beta > 0$ to the variance of the risk. Therefore, the customer is now further characterised by the risk aversion coefficient β. Approximating the exponential premium principle by the variance premium principle is a common approach; see, for example, Gerber and Pafum (1998).

It is indeed possible to find an analytic expression for (2.7). Applying Campbell–Mecke's formula, it appears that the expectation term simply is:

$$\mathbb{E}[S^{(K)}] = \alpha\overline{x_1^{(K)}}.$$

The variance term in (2.7) is calculated using the total law of variance in Appendix B.II, giving the result:

$$\mathbb{V}\text{ar}[S^{(K)}] = \frac{r\overline{x_2^{(K)}}\alpha}{2}.$$

The concluding premium is summarised below.

Corollary B.3. *Given a deductible K, the customer is facing the excess risk $S^{(K)} = r\sum_{i=1}^{\infty}\exp(-rT_i)X_i^{(K)}$ when not insuring and is willing to pay the following price for an insurance:*

$$P_{\text{var}}(K) = \alpha\overline{x_1^{(K)}} + \beta\left(\frac{r\overline{x_2^{(K)}}\alpha}{2}\right) = \left(1 + \beta\frac{r\overline{x_2^{(K)}}}{2\overline{x_1^{(K)}}}\right)\alpha\overline{x_1^{(K)}}. \tag{2.8}$$

In Example B.9, the premium (2.8) is calculated explicitly for log-normally-distributed claim sizes.

Remark B.4. The approach leads to a fairly classical premium calculation principle, which in the literature is mostly seen from the insurer's perspective. This motivates the use of other already developed premium calculation principles applied from the

the customer's point of view. An example of such is the standard deviation principle, where the premium depends on the mean and the standard deviation of the risk in a linear structure,

$$P_{\text{std}}(K) = \mathbb{E}[S^{(K)}] + \beta\sqrt{\mathbb{V}\text{ar}[S^{(K)}]}.$$

Since we already have expressions for the mean and the variance, we can write this more explicitly as:

$$P_{\text{std}}(K) = \alpha\overline{x_1^{(K)}} + \beta\sqrt{\left(\frac{r\overline{x_2^{(K)}}\alpha}{2}\right)}.$$

Another could simply be the expected value premium principle with a safety loading depending on the individual's risk aversion:

$$P_{\text{ev}}(K) = (1+\omega(\beta))\mathbb{E}[S] = (1+\omega(\beta))\alpha\overline{x_1^{(K)}}. \tag{2.9}$$

The latter gives a simple expression, even though it still has a nice intuitive interpretation. All evaluations in the following will be based on the variance premium principle. However, do note that a similar approach can be used with other premium calculation principles.

Remark B.5. The risk the customer is facing can be split into two. First of these is the risk the customer cannot insure. This is the (part of the) losses that the customer must pay regardless of insuring or not and is the uncertainty that appears in both V_n and V_i. The monetary value at Time 0 of this risk can be deduced to:

$$S^{\leq K} = r\sum_{i=1}^{\infty}\exp(-rT_i)(Z_i\mathbb{1}_{\{Z_i<K\}} + K\mathbb{1}_{\{Z_i\geq K\}}).$$

Second of these is the additional risk $S^{(K)}$ the customer can buy insurance to cover. The maximum premium that the customer is willing to pay in (2.5) will be altered as follows if she takes $S^{\leq K}$ into consideration,

$$\mathbb{E}[u(w_0 - S^{\leq K} - S^{(K)})] = \mathbb{E}[u(w_0 - S^{\leq K} - P(K))].$$

Using the exponential utility and solving yields:

$$P(K) = \frac{1}{a}\log\left(\frac{\mathbb{E}[\exp(aS^{\leq K} + aS^{(K)})]}{\mathbb{E}[\exp(aS^{\leq K})]}\right).$$

Proceeding to a similar second order Taylor approximation as seen previously,

$$P_{\exp}(K) \approx \mathbb{E}[S^{(K)}] + \frac{a}{2}\mathbb{V}\text{ar}[S^{(K)}] - a\mathbb{C}\text{ov}(S^{\leq K}, S^{(K)}),$$

it appears that in the final pricing formula, the assumption about the customer only caring about pricing the excess risk $S^{(K)}$ regardless of $S^{\leq K}$ corresponds mathematically to assuming that $\mathbb{C}\text{ov}(S^{\leq K}, S^{(K)}) \approx 0$. A similar comment is also made in Gerber and Pafum (1998).

B.3 Portfolio characteristics

A customer is characterised by her claim frequency α and risk aversion β. As previously commented, these characteristics are most likely customer-dependent. The insurance company therefore considers them as random variables being i.i.d. on the market. The claim frequencies are represented by $A_1, \ldots, A_N$, and the risk aversions by $B_1, \ldots, B_N$. First, we will consider the claim frequency as being random and the risk aversion as constant. Next, we will reverse it, by modelling the risk aversion as random and letting the claim frequency be constant. A third, more advanced, possibility is of course to let the customer characteristics (α, β) be represented by random vectors $(A_1, B_1), \ldots, (A_N, B_N)$. This complicates the evaluations considerably and is therefore left open by this paper.

In each case, we derive an expression for the portfolio size and for the average claim frequency in the portfolio. These expressions will become explicit functions when assuming a concrete distribution.

B.3.1 Stochastic claim frequencies

Consider the first case mentioned above, where the risk aversion is constant, and the claim frequency of the customer is unknown to the insurance company and, therefore, modelled by a random variable denoted by A. As said, we want to find the expected size of the portfolio and the average claim rate in it.

From the reservation price in (2.8), it follows that a customer with characteristics (α, β) will insure if the offered insurance contract (p, K) satisfies the inequality:

$$p \leq \alpha \overline{x_1^{(K)}} + \frac{\beta \alpha r}{2} \overline{x_2^{(K)}}.$$

Since the claim frequency is modelled by a random variable, A, to the insurer, this translates to the relation:

$$A \geq \frac{2p}{2\overline{x_1^{(K)}} + \beta r \overline{x_2^{(K)}}}.$$

The expected portfolio size will then be the probability of this event happening multiplied by the size of the market, namely:

$$n(p; K) = \mathbb{P}\left(A \geq \frac{2p}{2\overline{x_1^{(K)}} + \beta r \overline{x_2^{(K)}}} \right) N.$$

This is the mean demand curve as a function of the premium and deductible. In the following, it is assumed that N is large and that there is a continuum of customers, such that the deviation from the actual demand curve is negligible.

The average claim frequency rate in the portfolio is the expected claim frequency given that the customer chooses to insure, i.e.,

$$\alpha(p; K) = \mathbb{E}\left[A \,\middle|\, A \geq \frac{2p}{2\overline{x_1^{(K)}} + \beta r \overline{x_2^{(K)}}} \right].$$

B.3.1.1 Exponentially distributed claim frequencies

When accounting for unobserved heterogeneity in the Bayesian claim experience rating setting, the claim frequency is frequently assumed to be $\Gamma(s, b)$ distributed, for example, as in Pitrebois et al. (2003). In some cases there are empirical evidence of s being close to one. See Bichsel (1964) for more discussion. This motivates the assumption of an exponential distribution with parameter b of the claim frequency A. Offering the contract (p, K), the customer will insure with probability:

$$\mathbb{P}\left(A \geq \frac{2p}{2\overline{x_1^{(K)}} + \beta r \overline{x_2^{(K)}}}\right) = \exp\left(-b\frac{2p}{2\overline{x_1^{(K)}} + \beta r \overline{x_2^{(K)}}}\right).$$

This yields a portfolio size of:

$$n(p, K) = \exp\left(-b\frac{2p}{2\overline{x_1^{(K)}} + \beta r \overline{x_2^{(K)}}}\right)N.$$

The exponential distribution has a memoryless property, which implies $\mathbb{E}[A \,|\, A \geq a] = 1/b + a$. The expected claim frequency of an insured customer will therefore be:

$$\alpha(p, K) = \frac{2p}{2\overline{x_1^{(K)}} + \beta r \overline{x_2^{(K)}}} + \frac{1}{b}.$$

This equation reflects adverse selection very clearly, since $\alpha(p, K)$ is linearly increasing in p.

Remark B.6. If the customer's reservation price was defined by the expected value premium principle (2.9), the portfolio characteristics would be:

$$n(p; K) = \mathbb{P}\left(A \leq \frac{p}{(1 + \omega(\beta))\overline{x_1^{(K)}}}\right)N,$$
$$\alpha(p; K) = \mathbb{E}\left[A \,\middle|\, A \leq \frac{p}{(1 + \omega(\beta))\overline{x_1^{(K)}}}\right].$$

Assuming an exponentially-distributed claim frequency, the characteristics become:

$$n(p; K) = \exp\left(-b\frac{p}{(1 + \omega(\beta))\overline{x_1^{(K)}}}\right)N,$$
$$\alpha(p; K) = \frac{p}{(1 + \omega(\beta))\overline{x_1^{(K)}}} + \frac{1}{b}.$$

This is slightly easier to work with and has the advantage that $\omega(\beta)$ can be chosen on a simple form.

B.3.2 Stochastic risk aversions

Instead of letting a customer being represented by a stochastic claim frequency and constant risk aversion, we now turn it around. Assume that the customer now has a constant claim frequency α. This is indeed relevant to consider. For example, everyone could be equally disposed to disaster caused by nature. Furthermore, assume that the risk aversion is represented by a random variable B. The criterion (2.8) of insuring for a given contract (p, K) can then be stated as:

$$B \geq \frac{2p - 2\alpha \overline{x_1^{(K)}}}{r\overline{x_2^{(K)}}\alpha}.$$

Therefore, the portfolio size will be characterised by:

$$n(p) = \mathbb{P}\left(B \geq \frac{2p - 2\alpha \overline{x_1^{(K)}}}{r\overline{x_2^{(K)}}\alpha}\right) N.$$

B.3.2.1 Gamma distributed risk aversions

Risk aversion is somewhat an abstract concept. It is therefore difficult to suggest a distribution since there is limited literature available. Since the gamma distribution is a popular choice for the claim frequency, we choose to consider the same distribution for the risk aversion. Assuming that the risk aversion has an $\Gamma(l, q)$ distribution, the demand curve will have the form:

$$n(p) = \frac{\Gamma(l, q(2p - 2\alpha \overline{x_1^{(K)}})/(r\overline{x_2^{(K)}}\alpha))}{\Gamma(l)} N.$$

In the case $l \approx 1$, the distribution approximately reduces to an exponential distribution with parameter q, and the portfolio size will then simply be:

$$n(p) = \exp\left(-q\frac{2p - 2\alpha \overline{x_1^{(K)}}}{r\overline{x_2^{(K)}}\alpha}\right) N.$$

B.4 Ruin probability

In this section, we consider the optimization problem of the insurance company. It mainly wants to minimise the ruin probability, but as this only makes sense for a positive drift, the expected time to ruin is considered in the case of a negative drift. The overall aim is to find the optimal premium as a function of the deductible.

When the controlled reserve develops according to a Brownian motion with positive drift, Hipp and Taksar (2010) shows that minimising the ruin probability is equivalent to maximising the following ratio μ/σ between the drift and volatility. In our setting, this means maximising:

$$\frac{\mu(p;K)}{\sigma^2(p;K)} = \frac{p - \alpha(p;K)\overline{x_1^{(K)}}}{\alpha(p;K)\overline{x_2^{(K)}}} - \frac{L}{n(p;K)\alpha(p;K)\overline{x_2^{(K)}}}. \tag{2.10}$$

This also follows from the relation $\varphi(r_0) = \exp(-2r_0\mu(p;K)/\sigma^2(p;K))$. Recall that if the drift is negative, then $\varphi(r_0) = 1$.

As previously implied, if the market consists of risk-neutral customers, then the drift of the diffusion process will be negative, making ruin certain. In order to avoid this, it is assumed in the following that there is a sufficiently large degree of risk aversion among the (potential) customers to satisfy the net profit condition $p > \alpha(p;K)\overline{x_1^{(K)}}$ for some p. Otherwise, there will be no motivation for selling insurance.

Notice that if $p = 0$, the insurance company will offer insurance for free, and so, $\mu(0;K) < 0$ for any given K. Furthermore, due to the liability rate, L, the diffusion will also have a negative drift when the premium becomes very high, since no customer will then be interested in insuring, hence $\lim_{p\to\infty} \mu(p;K) < 0$. Thus, ruin will be certain for both $p = 0$ and $p \to \infty$.

Now, assume for some K that the premium is zero. What will then happen if the premium was raised marginally? The insurance company will have nearly the same amount of customers and, therefore, also the same amount of claims, but the firm will get a small revenue when collecting premiums. Hence, $\mu(p, K)$ will increase. Note, that this is under the assumption that the effect from portfolio size decrease and adverse selection is smaller than the effect of raising the premium. Conversely, what if the premium was so high that no customer would be interested in insuring? Recall that there are N customers on the market and that each customer has a reservation price for insurance (depending on the customer characteristics). Therefore, if the insurance company sets the premium so high that it is above the reservation price of all N customers, then no customer will insure. Though if the company lowers the premium enough for it to become below the reservation price of the most risky/risk adverse potential customers on the market, then under the condition of $p > \alpha(p, K)\overline{x_1^{(K)}}$, the insurance company will obtain some revenue to cover at least some of the liability cost. Hence, $\mu(p, K)$ will also increase when lowering the premium for very large values. This tells us intuitively that if there is an optimal premium, then it is not obtained for $p = 0$, nor in the limit $p \to \infty$, and we thereby avoid trivialities when finding the optimal premium.

Abbreviate notation by $\mu_p(p, K) = \frac{\partial\mu(p;K)}{\partial p}$ and similarly for $\alpha_p(p, K)$ and $n_p(p, K)$. Assume that there is a unique $\tilde{p}$ that maximises the drift, i.e., satisfies the first order conditions:

$$\mu_p(\tilde{p};K) = n_p(\tilde{p};K)\big(\tilde{p} - \alpha(\tilde{p};K)\overline{x_1^{(K)}}\big) + n(\tilde{p};K)\big(1 - \alpha_p(\tilde{p};K)\overline{x_1^{(K)}}\big) = 0. \tag{2.11}$$

This is equivalent to saying that $\tilde{p}$ must satisfy:

$$\frac{-n_p(\tilde{p};K)}{n(\tilde{p};K)} = \frac{1 - \alpha_p(\tilde{p};K)\overline{x_1^{(K)}}}{\tilde{p} - \alpha(\tilde{p};K)\overline{x_1^{(K)}}}.$$

The intuitive interpretation is that the relative marginal change in the demand curve must equal the relative marginal change in average net revenue per customer due to a change in premium. Two cases then arise:

(1) If $\mu(\tilde{p}; K) > 0$, then $\varphi(r_0) < 1$ for all $r_0 > 0$, $K > 0$ and p in some bounded open interval $I \subset (0, \infty)$ containing $\tilde{p}$.

(2) If $\mu(\tilde{p}; K) \leq 0$, then $\varphi(r_0) = 1$ for all $r_0 > 0$ and $p > 0$.

Notice that the drift is positive if the net profit per customer is greater than the liability cost per customer, that is $p - \alpha(p; K) > L/n(p; K)$. This is more strict than the net profit condition.

Since the current framework is very general, so will the results be. In every concrete application, the existence of a unique solution must be verified. In Case (i), an optimization criterion is given and proven in Theorem B.7.

Theorem B.7. *When $\mu(\tilde{p}; K) > 0$, the optimal premium minimising the ruin probability must be a solution to the equation:*

$$1 - p\frac{\alpha_p(p; K)}{\alpha(p; K)} + \frac{L}{n(p; K)}\left(\frac{n_p(p; K)}{n(p; K)} + \frac{\alpha_p(p; K)}{\alpha(p; K)}\right) = 0.$$

Proof. Differentiating (2.10) with respect to the premium,

$$\begin{aligned}
\frac{\partial\left(\frac{\mu(p;K)}{\sigma^2(p;K)}\right)}{\partial p} &= \frac{1 - \alpha_p(p; K)\overline{x_1^{(K)}}}{\alpha(p; K)\overline{x_2^{(K)}}} - \frac{\left(p - \alpha(p; K)\overline{x_1^{(K)}}\right)\alpha_p(p; K)\overline{x_2^{(K)}}}{\left(\alpha(p; K)\overline{x_2^{(K)}}\right)^2} \\
&\quad + \frac{L(n_p(p; K)\alpha(p; K) + n(p; K)\alpha_p(p; K))\overline{x_2^{(K)}}}{\left(n(p; K)\alpha(p; K)\overline{x_2^{(K)}}\right)^2} \\
&= \frac{\alpha(p; K) - \alpha_p(p; K)p}{\alpha^2(p; K)\overline{x_2^{(K)}}} + \frac{L(n_p(p; K)\alpha(p; K) + n(p; K)\alpha_p(p; K))}{n^2(p; K)\alpha^2(p; K)\overline{x_2^{(K)}}}
\end{aligned}$$

yields the first order condition:

$$\frac{\alpha(p; K) - \alpha_p(p; K)p}{\alpha^2(p; K)\overline{x_2^{(K)}}} + \frac{L(n_p(p; K)\alpha(p; K) + n(p; K)\alpha_p(p; K))}{n^2(p; K)\alpha^2(p; K)\overline{x_2^{(K)}}} = 0.$$

Reducing, we obtain the following optimality criterion:

$$1 - \frac{\alpha_p(p; K)}{\alpha(p; K)}p + \frac{L}{n(p; K)}\left(\frac{n_p(p; K)}{n(p; K)} + \frac{\alpha_p(p; K)}{\alpha(p; K)}\right) = 0.$$

□

In Case (ii), it no longer makes sense to minimise ruin probability since it is constantly one. Instead, we suggest to choose the control to maximise the expected time to ruin, $\mathbb{E}[\tau]$, and thereby extend the expected lifetime of the company as much as possible. This is a non-standard objective function. Since the controlled reserve is a Brownian motion with drift, we are able to obtain a very simple result.

Theorem B.8. *When $\mu(\tilde{p}; K) \leq 0$, then $\tilde{p}$ will optimise the expected time to ruin.*

Proof. As said, when $\mu(\tilde{p}; K) \leq 0$, then (2.1) is a Brownian motion with negative drift. This also has the representation:

$$x_t = r_0 + \mu(p; K)t + \sigma(p; K)W_t.$$

Consider the stopping time $\tau = \inf\{t \geq 0 : x_t \leq 0\}$. Since $(x_t)_{t\geq 0}$ is a continuous process, it must be that $x_\tau = 0$. Furthermore, $\{x_t - \mu(p; K)t\}$ is a martingale with mean r_0. This implies:

$$\mathbb{E}[x_\tau - \mu(p; K)\tau] = -\mu(p; K)\mathbb{E}[\tau] = r_0$$

Hence, $\mathbb{E}[\tau] = -r_0/\mu(p; K)$. Therefore, the expected time to ruin is maximised when the drift is maximised. Due to the assumption of a unique $\tilde{p}$, such that the first order condition (2.11) is satisfied, then this must be the optimal choice. □

In Examples B.10 and B.11 the optimal premium is calculated explicitly for the portfolio characteristics in Sections B.3.1.1 and B.3.2.1, respectively.

B.5 Examples and illustrations

Example B.9. Assume that the claim sizes are log-normally distributed, that is $Z_i \sim \log\mathcal{N}(\mu_\ell, \sigma_\ell^2)$. This distribution has tail function $\overline{F}(x) = 1 - \Phi\left((\log(x) - \mu_\ell)/\sigma_\ell\right)$, where Φ denotes the standard normal distribution function. The k'th moment is given by $\mathbb{E}[Z_i^k] = \exp(k\mu_\ell + k^2\sigma_\ell^2/2)$. We seek to find a closed form solution to the premium $P_{\text{var}}(K)$. The challenge obviously is $\overline{x_1^{(K)}}$ and $\overline{x_2^{(K)}}$. Altering these yields:

$$\begin{aligned} \overline{x_1^{(K)}} &= \mathbb{E}[(Z_i - K)\mathbb{1}_{\{Z_i \leq K\}}] = \mathbb{E}[(Z_i - K) \mid Z_i \geq K]\overline{F}(K) \\ &= (\mathbb{E}[Z_i \mid Z_i \geq K] - K)\overline{F}(K), \end{aligned} \tag{2.12}$$

and:

$$\begin{aligned} \overline{x_2^{(K)}} &= \mathbb{E}[((Z_i - K)\mathbb{1}_{\{Z_i \geq K\}})^2] = \mathbb{E}[(Z_i - K)^2\mathbb{1}_{\{Z_i \leq K\}}] \\ &= \mathbb{E}[(Z_i^2 + K^2 - 2Z_iK)\mathbb{1}_{\{Z_i \geq K\}}] \\ &= K^2\overline{F}(K) + \mathbb{E}[Z_i^2 \mid Z_i \geq K]\overline{F}(K) - 2K\mathbb{E}[Z_i \mid Z_i \geq K]\overline{F}(K). \end{aligned} \tag{2.13}$$

In Benckert and Jung (1974), it is shown that the k'th moment of the truncated random variable is:

$$\mathbb{E}[Z_i^k \mid Z_i \geq K] = \mathbb{E}[Z_i^k]\frac{\Phi\left(\frac{\mu_\ell + k\sigma_\ell^2 - \log(K)}{\sigma_\ell}\right)}{\overline{F}(K)}.$$

From this, it follows that in the log-normal case, (2.12) can be written as:

$$\overline{x_1^{(K)}} = \overline{z_1}\Phi\left(\frac{\mu_\ell + \sigma_\ell^2 - \log(K)}{\sigma_\ell}\right) - K\overline{F}(K),$$

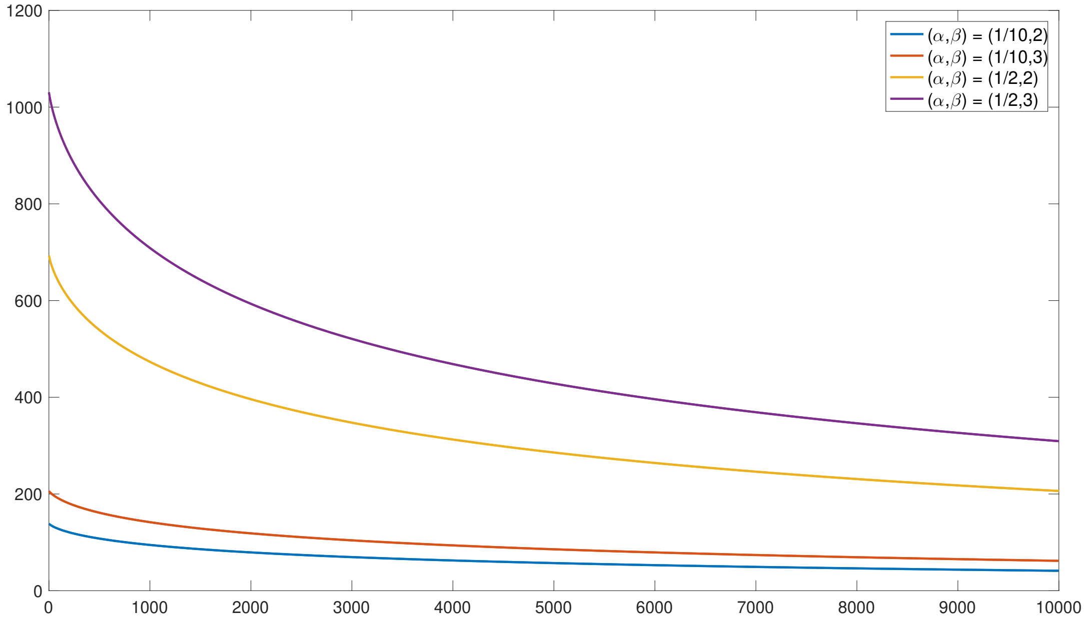


Figure 2.1: The premium (2.14) as function of K for different values of customer characteristics. The interest rate is chosen to be 2%, and the estimates $\hat{\mu_\ell} = 1.6$ and $\hat{\sigma_\ell} = 1.99$ are used.

and (2.13) as:

$$\overline{x_2^{(K)}} = K^2\overline{F}(K) + \overline{z_2}\Phi\left(\frac{\mu_\ell + 2\sigma_\ell^2 - \log(K)}{\sigma_\ell}\right) - 2K\,\overline{z_1}\Phi\left(\frac{\mu_\ell + \sigma_\ell^2 - \log(K)}{\sigma_\ell}\right).$$

Hence, when the claims are log-normally distributed, the maximal premium that a customer with characteristics (α, β) is willing to pay for an insurance contract as a function of the deductible is:

$$\begin{aligned} P_{\text{var}}(K) = \alpha K\overline{F}(K)\left(\frac{K\beta r}{2} - 1\right) + \alpha\,\overline{z_1}\Phi\left(\frac{\mu_\ell + \sigma_\ell^2 - \log(K)}{\sigma_\ell}\right)(1 - \beta r K) \\ + \frac{\beta r\alpha\,\overline{z_2}}{2}\Phi\left(\frac{\mu_\ell + 2\sigma_\ell^2 - \log(K)}{\sigma_\ell}\right). \end{aligned} \tag{2.14}$$

While this is not a straightforward expression, it is computationally easy to evaluate. Some combined data on claims in fire insurance reported 1958–1969 by Swedish fire insurance companies are studied in Benckert and Jung (1974), where the estimates $\hat{\mu_\ell} = 1.6$ and $\hat{\sigma_\ell} = 1.99$ are obtained. These estimates are used in the following. In Figure 2.1, the premium function (2.14) for different combinations of characteristics is illustrated. The function appears to be very sensitive towards changes in characteristics and most so for small deductibles. Notice the considerable change from the combination $(\alpha, \beta) = (1/10, 2)$ to $(\alpha, \beta) = (1/2, 3)$ where the premium gets approximately 7.4-times larger.

Example B.10. An insurance company wants to supply fire insurance. It is entering a market where the customers have unknown, possibly different claim frequencies and constant risk aversions, namely β. The claim frequencies are once again assumed to be independent and identically exponentially distributed with parameter b. The portfolio can thus be characterised as in Section B.3.1.1.

First, the insurer needs to see which region of the premium is profitable to even supply insurance. The criteria $p - \alpha(p,K)\overline{x_1^{(K)}} > 0$ translates into:

$$p > \frac{2}{\beta r b}\frac{\overline{x_1^{(K)}}^2}{\overline{x_2^{(K)}}} + \frac{\overline{x_1^{(K)}}}{b}. \tag{2.15}$$

For a given deductible K, the considered price must exceed this threshold. Otherwise, the insurance company should choose not to supply insurance.

Next, evaluating the drift is of interest. Knowing the portfolio characteristics, it can be written explicitly as:

$$\mu(p;K) = N\exp\left(-b\frac{2p}{2\overline{x_1^{(K)}} + \beta r\overline{x_2^{(K)}}}\right)\left(p - \frac{2p\overline{x_1^{(K)}}}{2\overline{x_1^{(K)}} + \beta r\overline{x_2^{(K)}}} + \frac{\overline{x_1^{(K)}}}{b}\right) - L.$$

Solving the first order criteria (2.11) yields the solution:

$$\tilde{p}(K) = \frac{(2\overline{x_1^{(K)}} + \beta r\overline{x_2^{(K)}})^2}{2\beta b r\overline{x_2^{(K)}}} = \frac{2}{\beta r b}\frac{\overline{x_1^{(K)}}^2}{\overline{x_2^{(K)}}} + \frac{\beta r\overline{x_2^{(K)}}}{2b} + \frac{2\overline{x_1^{(K)}}}{b}.$$

Notice that $\tilde{p}$ obviously satisfies being in the region of (2.15). We now seek to find the conditions under which the drift will be positive. Solving for $\mu(\tilde{p};K) > 0$ yields:

$$\frac{L}{N} < \frac{\beta r\overline{x_2^{(K)}}}{2b}\exp\left(-\frac{2\overline{x_1^{(K)}} + \beta r\overline{x_2^{(K)}}}{\beta r\overline{x_2^{(K)}}}\right).$$

Assuming that this inequality holds, then one must use Theorem B.7 to find the optimal price. The optimality criterion in reduced form is:

$$\frac{N}{L}\frac{2\overline{x_1^{(K)}} + \beta r\overline{x_2^{(K)}}}{2b} = \exp\left(\frac{2bp}{2\overline{x_1^{(K)}} + \beta r\overline{x_2^{(K)}}}\right)\frac{2bp}{2\overline{x_1^{(K)}} + \beta r\overline{x_2^{(K)}}}.$$

This yields the following optimal premium as a function of the deductible,

$$p^*(K) = \frac{2\overline{x_1^{(K)}} + \beta r\overline{x_2^{(K)}}}{2b}W\left(\frac{N}{L}\frac{2\overline{x_1^{(K)}} + \beta r\overline{x_2^{(K)}}}{2b}\right).$$

using the Lambert W function. The Lambert W function is defined as the (multivalued) inverse of the function $w \mapsto \exp(w)w$. For more details, see Corless et al. (1996). In the case $\mu(\tilde{p};K) \leq 0$, it follows from Theorem B.8 that $\tilde{p}(K)$ is the optimal choice. Note that due to the Lambert W function increasing for positive values, then p^* will be preferred to $\tilde{p}$ if $p^*(K) \geq \tilde{p}(K)$. Therefore, for a given deductible, K, it is preferable for the insurance firm to simply choose the maximum of $p^*(K)$ and $\tilde{p}(K)$.

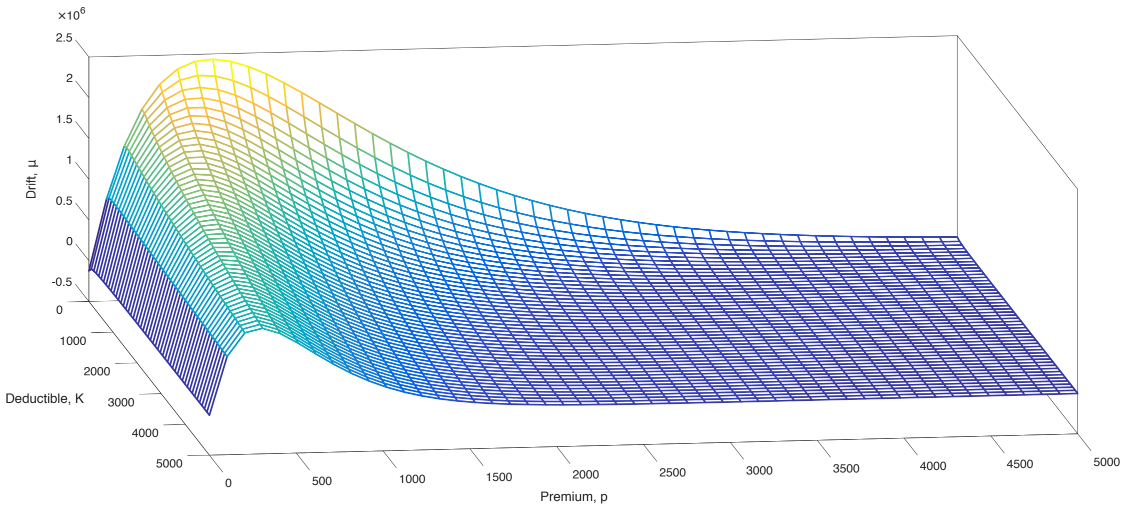


Figure 2.2: A mesh of the drift $\mu(p; K)$ for deductibles and premiums in the range $[0, 5000]$

Assume that the insurance company evaluates that the market consists of $N = 10\,000$ house owners considering buying insurance and that it calculated the liability costs to be $L = 5000$. It also assesses that $b = 3$ and $\beta = 3$. Assume furthermore that the company has some information about the distribution of the claims, and based on this, it believes that the claims are log-normally distributed according to the estimates in Benckert and Jung (1974). It also knows how to reasonably choose a deductible K to serve the purpose described in the introduction. The interest rate applied is 2%.

In Figure 2.2, a mesh of the drift is presented. The preliminary analysis of the drift in Section B.4 is very well illustrated in this. The concavity in the premium is obvious. For all of the considered deductibles in the range $[0, 5000]$, the drift will be positive in $\tilde{p}(K)$.

Next, a contour of the ratio μ/σ^2 is illustrated in Figure 2.3. The concavity in the premium also appears very clearly here. The ratio is at its highest within the region of approximately $K \in [0, 220]$, followed by the regions $K \in (220, 590]$ and then $K \in (590, 1070]$.

In Figure 2.4, the premiums $\tilde{p}(K)$ and $p^*(K)$ as functions of the deductible are plotted.

Notice that the model does not take the cost of processing an increasing number of claims into consideration. The company therefore evaluates that a deductible of $K = 1000$ is suitable. This yields the following values:

$$\tilde{p}(1000) = 474.2, \qquad p^*(1000) = 2458.1,$$

Since $p^*(1000) > \tilde{p}(1000)$, then p^* is the optimal premium.

The optimal premium $\max\{p^*(1000), \tilde{p}(1000)\}$ will of course depend on the parameters chosen, namely the market size N, the liability rate L, the risk aversion β and the exponential parameter b. An illustration of how sensitive the optimal premium are towards changes in these parameters is viewed in Figure 2.5. Figure 2.5a shows a mesh of the optimal premium where N and L take values in a grid of

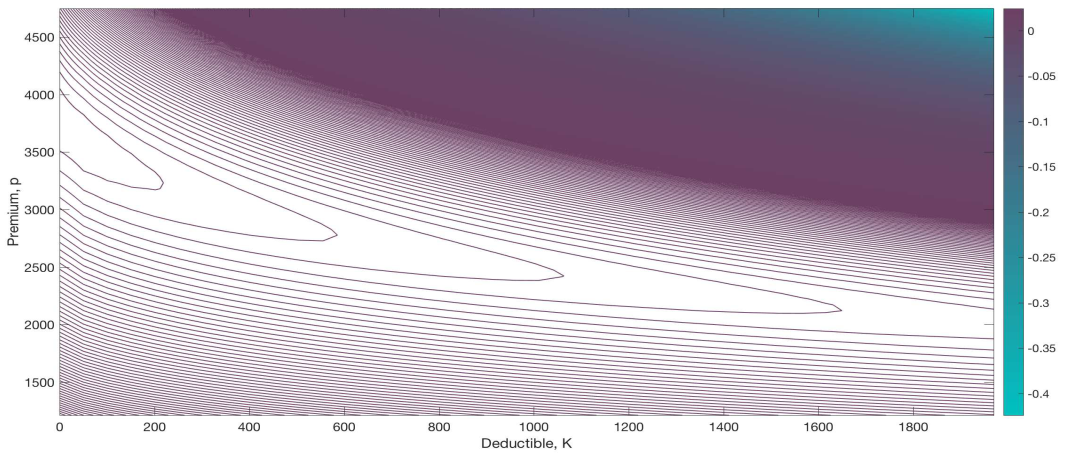


Figure 2.3: A contour of the relation $\mu(p;K)/\sigma(p;K)^2$ for deductibles in the range $[0, 2000]$ and premiums in $[1210, 4750]$.

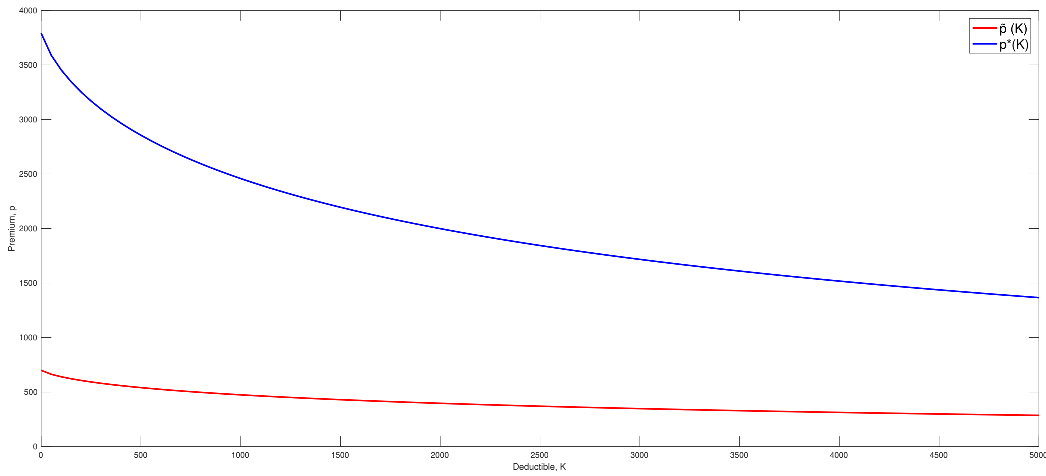
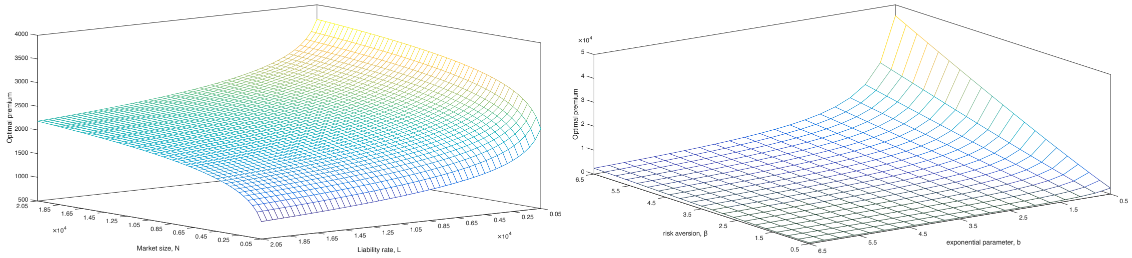


Figure 2.4: The premiums $p^*(K)$ and $\tilde{p}(K)$ as functions of the deductible K.

$[500, 20\,500]$. Here, we can see that the premium is most sensitive towards changes in these parameters for small values. A low liability rate L will yield a high premium. The insurance company does not need many customers when they have small liabilities payments; hence, they can afford to choose a high premium. On the contrary, when N is small, then we see that the premium tends to be low. In Figure 2.5b, a similar mesh of the optimal premium is shown for values of β and b in a grid of $[0.5, 6.5]$. It is observed that when b gets too high or β too low, then the optimal premium will tend towards $\tilde{p}(1000)$. Conversely, when b takes on low values and β gets high, the optimal premium increases considerably. The values $N = 10\,000$, $L = 5000$ $\beta = 3$ and $b = 3$ are chosen such that any of the extremes mentioned above are avoided.

Example B.11. Assume now that the insurance company believes that the claim frequency is constant in the portfolio, but does not possess any information about the customer's risk aversion. The risk aversion is therefore modelled by a random



(a) **(b)**

Figure 2.5: Mesh of the optimal premium $\max\{p^*(1000), \tilde{p}(1000)\}$ for different values of the parameters. If a parameter does not vary, then it has the same value as in previous graphs. (**a**) Mesh of the optimal premium for $L, N \in [500, 20\,500]$; (**b**) mesh of the optimal premium for $\beta, b \in [0.5, 6.5]$.

variable, which is assumed to have an exponential distribution with parameter q. Portfolio characteristics are then as in Section B.3.2.1. The existence criteria of the insurance company will be $p - \alpha\overline{x_1^{(K)}} > 0$, i.e., the premium must simply be larger than the net premium. First, we seek to find the solution of (2.11), namely:

$$\tilde{p}(K) = \frac{r\overline{x_2^{(K)}}\alpha}{2q} + \alpha\overline{x_1^{(K)}}.$$

Next is to find the region for which the drift evaluated in $\tilde{p}$ is positive:

$$\frac{L}{N} < \frac{r\overline{x_2^{(K)}}\alpha}{2q}\exp(-1).$$

In this region:

$$p^*(K) = \frac{r\overline{x_2^{(K)}}\alpha}{2q}\log\left(\frac{N}{L}\frac{r\overline{x_2^{(K)}}\alpha}{2q}\right) + \alpha\overline{x_1^{(K)}}$$

is optimal. Otherwise, $\tilde{p}$ is optimal. Again, due to the logarithm being an increasing function, one can simply state that the insurance company should choose the maximum of $\tilde{p}$ and p^*.

Appendix

B.I Calculations of the present value in Section B.2

The present value of not insuring can be split into three terms:

$$\begin{aligned} V_{NI} &= \mathbb{E}\Big[\int_0^\infty \exp(-dt)\mathrm{d}w_t\Big] = \mathbb{E}\Big[\int_0^\infty \exp(-dt)rw_t\mathrm{d}t - \int_0^\infty \exp(-dt)\mathrm{d}A_t^\alpha\Big] \\ &= \mathbb{E}\Big[\int_0^\infty \exp(-(d-r)t)r\Big(w_0 - \sum_{i=1}^{N_t^\alpha}\exp(-rT_i)Z_i\Big)\mathrm{d}t - \int_0^\infty \exp(-dt)\mathrm{d}A_t^\alpha\Big] \\ &\equiv K_1 - K_2 - K_3. \end{aligned}$$

The three terms are defined and assessed below. The first is easily evaluated as:

$$K_1 = \mathbb{E}\Big[\int_0^\infty \exp(-(d-r)t)rw_0\mathrm{d}t\Big] = \frac{rw_0}{d-r}.$$

The second and third term require a bit more effort. Using Campbell–Mecke's formula yields:

$$\begin{aligned}
K_2 &= \mathbb{E}\Big[\int_0^\infty r\exp(-(d-r)t)\Big(\sum_{i=1}^{N_t^\alpha}\exp(-rT_i)Z_i\Big)\mathrm{d}t\Big] \\
&= \int_0^\infty r\exp(-(d-r)t)\mathbb{E}\Big[\Big(\sum_{i=1}^{N_t^\alpha}\exp(-rT_i)Z_i\Big)\Big]\mathrm{d}t \\
&= \overline{z_1}\alpha r\int_0^\infty \exp(-(d-r)t)\Big(\int_0^t \exp(-rs)\mathrm{d}s\Big)\mathrm{d}t \\
&= \overline{z_1}\alpha\int_0^\infty \exp(-(d-r)t)(1-\exp(-rt))\mathrm{d}t \\
&= \frac{\overline{z_1}\alpha}{d-r} - \frac{\overline{z_1}\alpha}{d}, \\
K_3 &= \mathbb{E}\Big[\int_0^\infty \exp(-dt)\mathrm{d}A_t^\alpha\Big] = \mathbb{E}\Big[\sum_{i=1}^\infty \exp(-dT_i)Z_i\Big] \\
&= \overline{z_1}\alpha\int_0^\infty \exp(-ds)\mathrm{d}s = \frac{\overline{z_1}\alpha}{d}.
\end{aligned}$$

Therefore, to conclude, the present value of not insuring will be:

$$V_{NI} = \frac{rw_0}{d-r} - \frac{\overline{z_1}\alpha}{d-r}.$$

B.II Calculations of the variance in Section B.2

The total law of variance states that:

$$\mathbb{V}\mathrm{ar}[S^{(K)}] = \mathbb{V}\mathrm{ar}\big[\mathbb{E}[S^{(K)} \mid \mathcal{F}]\big] + \mathbb{E}\big[\mathbb{V}\mathrm{ar}[S^{(K)} \mid \mathcal{F}]\big] \tag{B1}$$

where the filtration, $\mathcal{F}$, is the sigma algebra generated by the T_i's. Consider the first term of the variance in (B1). The conditional expectation is easily calculated as:

$$\mathbb{E}[S^{(K)} \mid \mathcal{F}] = \mathbb{E}\Big[r\sum_{i=1}^\infty \exp(-rT_i)X_i^{(K)} \,\Big|\, \mathcal{F}\Big] = r\overline{x_1^{(K)}}\sum_{i=1}^\infty \exp(-rT_i).$$

The square of a sum is used for finding the variance of this expression:

$$\begin{aligned}
&\mathbb{V}\mathrm{ar}\Big[r\overline{x_1^{(K)}}\sum_{i=1}^{\infty}\exp(-rT_i)\Big] \\
&\quad = r^2\overline{x_1^{(K)}}^2\bigg(\mathbb{E}\Big[\Big(\sum_{i=1}^{\infty}\exp(-rT_i)\Big)^2\Big] - \mathbb{E}\Big[\sum_{i=1}^{\infty}\exp(-rT_i)\Big]^2\bigg) \\
&\quad = r^2\overline{x_1^{(K)}}^2\bigg(\mathbb{E}\Big[\sum_{i=1}^{\infty}\exp(-r2T_i) + \sum_{i,j=1\,:\,i\neq j}^{\infty}\exp(-rT_i)\exp(-rT_j)\Big] - \Big(\frac{\alpha}{r}\Big)^2\bigg) \\
&\quad = r^2\overline{x_1^{(K)}}^2\left(\frac{\alpha}{2r} + \Big(\frac{\alpha}{r}\Big)^2 - \Big(\frac{\alpha}{r}\Big)^2\right) = \frac{r\overline{x_1^{(K)}}^2\alpha}{2}.
\end{aligned}$$

The square of a sum is also used for finding the conditional variance in the second term of (B1),

$$\begin{aligned}
&\mathbb{V}\mathrm{ar}[S^{(K)} \mid \mathcal{F}] \\
&\quad = r^2\bigg(\mathbb{E}\Big[\Big(\sum_{i=1}^{\infty}\exp(-rT_i)X_i^{(K)}\Big)^2 \,\Big|\, \mathcal{F}\Big] - \mathbb{E}\Big[\sum_{i=1}^{\infty}\exp(-rT_i)X_i^{(K)} \,\Big|\, \mathcal{F}\Big]^2\bigg) \\
&\quad = r^2\mathbb{E}\Big[\sum_{i=1}^{\infty}\exp(-r2T_i)\,(X_i^{(K)})^2 \\
&\qquad\qquad + \sum_{i,j=1\,:\,i\neq j}^{\infty}\exp(-rT_i)\exp(-rT_j)X_i^{(K)}X_j^{(K)} \,\Big|\, \mathcal{F}\Big] \\
&\qquad - r^2\overline{x_1^{(K)}}^2\Big(\sum_{i=1}^{\infty}\exp(-r2T_i) - \sum_{i,j=1\,:\,i\neq j}^{\infty}\exp(-rT_i)\exp(-rT_j)\Big) \\
&\quad = r^2(\overline{x_2^{(K)}} - \overline{x_1^{(K)}}^2)\sum_{i=1}^{\infty}\exp(-r2T_i).
\end{aligned}$$

Taking the expectation yields:

$$\mathbb{E}\big[\mathbb{V}\mathrm{ar}[S^{(K)} \mid \mathcal{F}]\big] = \mathbb{E}\Big[r^2(\overline{x_2^{(K)}} - \overline{x_1^{(K)}}^2)\sum_{i=1}^{\infty}\exp(-r2T_i)\Big] = \frac{r(\overline{x_2^{(K)}} - \overline{x_1^{(K)}}^2)\alpha}{2}.$$

Concluding, the variance in (B1) will be $\mathbb{V}\mathrm{ar}[S^{(K)}] = (r\overline{x_2^{(K)}}\alpha)/2$.

B.III Illustration of the complexity of the exponential premium in Section B.2

To illustrate the complexity of $P_{\exp}(K)$, consider the simple case where $K = 0$ and the claim size distribution F is degenerate in one, i.e., $Z_i = 1$ almost surely. Let:

$$f(t) = \mathbb{E}\Big[\exp\Big(\beta r\sum_{i=1}^{\infty}\exp(-rT_i)\mathbb{1}_{\{T_i\leq t\}}\Big)\Big].$$

During a short time interval dt, the function $f(t)$ might get an additional term to the sum (corresponding to a new claim) with probability $\alpha \mathrm{d}t$ or $f(t)$ will remain unchanged with probability $1 - \alpha \mathrm{d}t$, i.e.,

$$f(t + \mathrm{d}t) = f(t)(1 - \alpha \mathrm{d}t) + f(t)\exp(\beta \exp(-rt))\alpha \mathrm{d}t.$$

Letting dt tend to zero, an ordinary differential equation appears,

$$f'(t) = \alpha f(t)(\exp(\beta \exp(-rt)) - 1).$$

This can also be expressed as:

$$\begin{aligned}
\log(f(t)) &= \alpha \int_0^t (\exp(\beta \exp(-rs)) - 1)\mathrm{d}s \\
&= \alpha r \int_{-\beta}^{-\beta \exp(-rt)} \frac{1}{u} \exp(-u)\mathrm{d}u - \alpha t \\
&= \alpha r \Big(\int_{-\beta}^{\infty} \frac{1}{u} \exp(-u)\mathrm{d}u - \int_{-\beta \exp(-rt)}^{\infty} \frac{1}{u} \exp(-u)\mathrm{d}u \Big) - \alpha t \\
&= \alpha r (Ei(\beta \exp(-rt)) - Ei(\beta)) - \alpha t,
\end{aligned}$$

where $Ei(.)$ denotes the exponential integral

$$Ei(x) = -\int_{-x}^{\infty} \frac{1}{u} \exp(-u)\mathrm{d}u \qquad \text{for } x > 0.$$

The exponential integral is known not to have a closed form solution. Therefore, even for the simple case of a claim size distribution, an analytic solution to the premium cannot be obtained.

References

Asmussen, S. (2014). “Modeling and performance of bonus-malus systems: stationarity versus age-correction”. *Risks* 2, pp. 49–73.

Asmussen, S. and Albrecher, H. (2010). *Ruin probabilities.* 2nd. World Scientific.

Asmussen, S., Christensen, B. J., and Taksar, M. (2013). “Portfolio size as function of the premium: modelling and optimization”. *Stochastics* 85.4, pp. 575–588.

Benckert, L.-G. and Jung, J. (1974). “Statistical models of claim distributions in fire insurance”. *ASTIN Bulletin* 8.1, pp. 1–25.

Bichsel, F. (1964). “Erfahrungstarifierung in der Motorfahrzeug-Haftpflicht-Versicherung”. *Bulletin of the Swiss Association of Actuaries*, pp. 119–130.

Burnecki, K., Nowicka-Zagrajek, J., and Weron, A. (2004). *Pure risk premiums under deductibles. A quantitative management in actuarial practice*. HSC Research Reports. Hugo Steinhaus Center, Wroclaw University of Technology.

Corless, R. M., Gonnet, G. H., Hare, D. E., Jeffrey, D. J., and Knuth, D. E. (1996). “On the Lambert W function”. *Advances in Computational mathematics* 5.1, pp. 329–359.

Gerber, H. U. (1974). "On additive premium calculation principles". *ASTIN Bulletin* 7.3, pp. 215–222.

Gerber, H. U. and Pafum, G. (1998). "Utility functions: from risk theory to finance". *North American Actuarial Journal* 2.3.

Hipp, C. and Taksar, M. (2010). "Optimal non-proportional reinsurance control". *Insurance: Mathematics and Economics* 47.2, pp. 246–254.

Højgaard, B. (2002). "Optimal dynamic premium control in non-life insurance. Maximizing dividend pay-outs". *Scandinavian Actuarial Journal* 4, pp. 225–245.

Iglehart, D. L. (1969). "Diffusion approximations in collective risk theory". *Journal of Applied Probability* 6.2, pp. 285–292.

Pitrebois, S., Denuit, M., and Walhin, J. (2003). "Fitting the belgian bonus-malus system". *Belgian Actuarial Bulletin* 3.1, pp. 58–62.

Rees, R. and Wambach, A. (2008). *The microeconomics of insurance*. now Publishers Inc.

Schmidli, H. (2008). *Stochastic control in insurance*. Springer-Verlag London.

van Lieshout, M. (2000). *Markov point processes and their applications*. World Scientific.

Chapter 3

Equilibrium premium strategies for push-pull competition

The following papers, Paper C and Paper D, extends the work of Paper B to a setting, where two suppliers of insurance are competing on the same market. If we extend the analysis in the most straightforward and simplest manner, where the insurance companies offer products of same quality and no market frictions are present, then the customer's problem becomes trivial. A potential customer will obviously choose to insure at the company offering the lowest premium. Metaphorically we can consider an individual in demand for an apple. In front of the individual are two similar apples, one cheap and one expensive. The individual will, of course, choose to buy the cheap apple. Same goes for insurance. Therefore, we consider two different directions in the respective papers of the present chapter.

Market frictions are factors that influence the decision-making process. For potential customers examples are; i) search and switching costs preventing customers from making optimal decisions, ii) different accessibility to information (and/or different abilities to process it), iii) personal preferences. In Paper C we consider a competitive insurance environment, where the two suppliers offer the same product, but customers are influenced by market frictions. The quality of an insurance product is measured by the deductible structure and level, as it is the amount of coverage the customer receives. So if two insurance companies offer insurance products of same quality, it means that they include the same deductibles. In Paper D we leave the set-up of market frictions, and instead consider the case of product differentiation, where the two insurance companies both offer a fixed amount deductible, but at different levels. Hence, one of the companies offers a product with a lower deductible, and therefore of better quality, than the other.

The first formal introduction to oligopoly models of strategic interactions between competing companies was by Cournot (1838), where the companies simultaneously/independently choose their outputs in order to maximise their respective profits. The product price is then set by a neutral auctioneer in order to ensure market clearance. The first to review this work was Bertrand (1883) who contended that it is more realistic to let the competing companies choose their product price rather than the produced quantity. Since then Cournot competition has come to represent the overall class of quantity competition and Bertrand competition the class of price competition. The role of commitment in a competitive environment was suggested by von Stackelberg (1934), originally intended as an extension of Cournot (1838) with quantity strategies. In a Stackelberg-type competition there is a leader and a (or several) follower(s). The leading company chooses its strategy first. The following company observes the leader's choice and then plays second. This is the type of competition considered in Paper D, however in a price competition setting, and Paper C deals with Bertrand-type competition.

In order to find the best strategies of the competing insurance companies, we turn to game theory. Game theory puts the strategic interactions between decision-makers into mathematical models. In Section 3.1.1 we go through the basic terminology in game theory. Along with the ground-breaking work on utility theory, von Neumann and Morgenstern (1947) also contributed with pioneering work initiating modern game theory, in particular, in two-person non-cooperative zero-sum games. Work on n-person non-zero-sum games was also presented in this book, however in a cooperative setting, as it was not straightforward how to define an equilibrium when extending to $n > 2$ players. This was followed up by Nash (1951), who established what we today know as the Nash equilibrium intended for non-cooperative, non-zero-sum games with a finite number of players. This type of equilibrium is considered in both Paper C and Paper D. Another equilibrium type considered in Paper D is the Stackelberg equilibrium, obviously relating to the sequential Stackelberg game. Both equilibrium types, Nash and Stackelberg, are defined in Section 3.1.2. For references applying game theory to insurance, we refer to the thorough literature reviews in the papers of the present chapter.

If we add the ingredients competition and game theory to the stochastic control setting in Section 2.1.5, then we end at the stochastic differential game briefly considered in Section 3.1.3. Here we restrict our focus to the case where the insurance companies reserves are modelled by controlled diffusion processes. In general, stochastic differential games has been studied in a number of different settings using various tools. For a treatment of this extremely advanced topic, we refer to Ramachandran and Tsokos (2012).

That some of the researchers contributing within the early stages of stochastic differential games have afterwards proceeded to applying this tool for modelling insurance competition is a clear sign of how relevant stochastic differential games are in actuarial science. Examples are; Bensoussan and Friedman (1977) followed later by the contribution Bensoussan et al. (2014) where insurance companies find their optimal investment-reinsurance strategies in a non-zero-sum stochastic differential game influenced by Markov-modulation, and Elliott (1977), Elliott and Davis (1981) followed by Elliott and Siu (2011a,b) where optimal investment problems are formulated as zero-sum stochastic differential games between the insurer and the market in diverse settings. Among other references dealing with stochastic differential games in insurance are Zeng (2010), Taksar and Zeng (2011), Jin et al. (2013) and Chen et al. (2018) (some of them are also referred to in the papers). All references mentioned here that consider stochastic differential games in insurance do so in order to find optimal reinsurance and/or investment strategies. When it comes to premium competition a more common approach is to consider a small individual insurer vs the market as in Taylor (1986). This work has lead to several extensions, where some are listed in the literature review of Paper C and Paper D. In present papers, we propose a stochastic differential game, where the reserves of two insurance companies are controlled diffusions with dynamics governed by the single customer premiums. The insurance companies then choose their premium strategies in order to optimise (the company with the largest amount of capital maximises and the one with the smallest amount minimises) the probability of the difference in reserve hitting an upper barrier rather than a lower, where we once again benefit from the availability of exit times for diffusion processes.

3.1 Preliminaries

3.1.1 Terminology of game theory

Consider a game in which two players are participating. Each player's preferences are captured by an objective function. However, the object function does not only depend on the decision of the player itself, but also on the strategy of the other.

If the players are able to enter a cooperative environment where binding decisions are done in conjunction, it is called as a cooperative game. However, we will exclusively consider non-cooperative games, where no such agreements are possible. A non-cooperative game is said to be zero-sum if one player's gain is exactly balanced out with another player's loss. Hence, available resources remain constant. Conversely, in a non-zero-sum game the aggregate gains and losses can be different from zero.

If the strategies of the players completely determine the outcome, the game is said to be deterministic. If, on the contrary, a random variable is influencing at least one of the object functions, the game is called stochastic. The game has complete information, if the players, their objective functions, and any underlying probability distribution are common information. If not so, the game is said to have incomplete information.

In a static game, the players only have access to these a priori information and move simultaneously. In a dynamic game, the players move sequentially or repeatedly, where some players are granted access to information about the previous decisions of others. A dynamic game is said to be a differential game if it plays out in continuous time, where a differential equation describing the state trajectory of the game is controlled over time by the decision processes of the players. This type of game is closely related to the theory of stochastic control.

3.1.2 Equilibrium types

Let $x_i \in \mathcal{M}_i$ be the decision variable of player i, where $\mathcal{M}_i$ is the set of possible actions. Let $\mathcal{M} = \mathcal{M}_1 \times \mathcal{M}_2$ ignoring that there might the coupled constraints on the decision variables.

Presume for the moment and without loss of generality that the objective functions of all players are value functions, which they seek to maximise. The value functions are denoted by $V_i^{(x_1,x_2)}$ for $i = 1, 2$. In case a player is a minimiser with a loss function L_i as the objective, it can transformed by $V_i = -L_i$ into a maximisation problem. A Nash equilibrium is then defined as follows.

Definition 3.1 (Nash equilibrium)**.** *$(x_1^*, x_2^*) \in \mathcal{M}$ is a Nash equilibrium if it satisfies*

$$V_1^{(x_1^*,x_2^*)} \geq V_1^{(x_1,x_2^*)} \qquad \textit{for all } x_1 \in \mathcal{M}_1$$

and

$$V_2^{(x_1^*,x_2^*)} \geq V_2^{(x_1^*,x_2)} \qquad \textit{for all } x_2 \in \mathcal{M}_2.$$

Hence, when at a Nash equilibrium neither of the players has the incentive to deviate.

If $V = V_1 = -V_2$ the game is of zero-sum type, where the same objective is maximised by Player 1 and minimised by Player 2. A Nash equilibrium (x_1^*, x_2^*) can then be defined equivalently as the saddle point

$$V^{(x_1,x_2^*)} \leq V^{(x_1^*,x_2^*)} \leq V^{(x_1^*,x_2)} \qquad \text{for all } (x_1, x_2) \in \mathcal{M}. \tag{3.1}$$

This is the type of game and equilibrium considered in Paper C. Note that the sequence of which the players make their decisions is subordinate, since in this type of equilibrium we have

$$V^{(x_1^*,x_2^*)} = \max_{x_1 \in \mathcal{M}_1} \min_{x_2 \in \mathcal{M}_2} V^{(x_1,x_2)} = \min_{x_2 \in \mathcal{M}_2} \max_{x_1 \in \mathcal{M}_1} V^{(x_1,x_2)}.$$

However, in a Stackelberg equilibrium the sequence of the game is indeed relevant. Consider a sequential game where a leader makes the initial play. The follower observes the decision of the leader and then reacts. We define a Stackelberg-equilibrium in a zero-sum game where Player 2 is the (minimising) leader and Player 1 the (maximising) follower as:

Definition 3.2 (Stackelberg equilibrium)**.** *$(x_1^*, x_2^*) \in \mathcal{M}$ is a Stackelberg equilibrium if*

$$V^{(d(x_2^*),x_2^*)} \leq V^{(d(x_2),x_2)} \qquad \textit{for all } x_2 \in \mathcal{M}_2$$

where

$$d(x_2) = \{\varsigma \in \mathcal{M}_1 : V^{(\varsigma,x_2)} \geq V^{(x_1,x_2)} \textit{ for all } x_1 \in \mathcal{M}_1\}.$$

It is solved by backward induction, where the follower finds the optimal response as a function of leader's decision. The leader insert this information about the response function into the value function and solves for the optimal opening play. In Stackelberg equilibrium, the value function is

$$V^{(x_1^*,x_2^*)} = \max_{x_1 \in \mathcal{M}_1} \min_{x_2 \in \mathcal{M}_2} V^{(x_1,x_2)}$$

However, unlike the Nash equilibrium, the maximum and minimum in a Stackelberg equilibrium can not be interchanged. In Paper D a sequential zero-sum game is considered in which we examine the existence of a Stackelberg (and a Nash) equilibrium.

3.1.3 A stochastic differential game

We now intend to extend the formulation of Section 2.1.5 to involve two insurance companies. Insurance Company i has to decide upon a strategy $u_i = (u_{i,t})_{t\geq 0}$ taking values in some admissible set $U_i \subseteq \mathbb{R}$. The reserve is then assumed to be governed by the strategies (u_1, u_2) as follows

$$\begin{aligned} dR_{i,t}^{(u_1,u_2)} &= \mu_i(R_{1,t}^{(u_1,u_2)}, R_{2,t}^{(u_1,u_2)}, u_{1,t}, u_{2,t})\mathrm{d}t \\ &\quad + \sigma_i(R_{1,t}^{(u_1,u_2)}, R_{2,t}^{(u_1,u_2)}, u_{1,t}, u_{2,t})\mathrm{d}W_{i,t} \end{aligned}$$

with $R_{i,0}^{(u_1,u_2)} = r_{i,0}$ for i, and $(W_{1,t})_{t\geq 0}$ and $(W_{2,t})_{t\geq 0}$ are independent Wiener processes.

We here generalise the objective function of Section 2.1.5 by

$$\begin{aligned} V^{(u_1,u_2)}(t, r_1, r_2) = \mathbb{E}\Big[\int_t^\tau & \mathrm{e}^{-\zeta(s-t)} v(R_{1,s}^{(u_1,u_2)}, R_{2,s}^{(u_1,u_2)}, u_{1,s}, u_{2,s})\mathrm{d}s \\ + \mathrm{e}^{-\zeta(\tau-t)}\mathcal{K}(\tau, R_{1,\tau}^{(u_1,u_2)}, R_{2,\tau}^{(u_1,u_2)}) \,\Big|\, & R_{1,t}^{(u_1,u_2)} = r_1, R_{2,t}^{(u_1,u_2)} = r_2, t < \tau\Big], \end{aligned} \tag{3.2}$$

where $v : \mathbb{R} \times \mathbb{R} \times U_1 \times U_2 \to \mathbb{R}$, $\mathcal{K} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, τ is an exit time, and $\zeta > 0$ a discounting parameter. In a zero-sum setting as considered in the previous section, Player 1 chooses its strategy u_1 in order to maximise $V^{(u_1,u_2)}$ and Player 2 chooses its strategy u_2 to minimise it.

This is the type of competition studied in the papers of the present chapter, where the control at an insurance company's disposal is the single customer premium. The exit time is defined as $\tau = \tau(\delta) = \inf\{t > 0 : R_{1,t}^{(u_1,u_2)} - R_{2,t}^{(u_1,u_2)} \notin [\ell_d, \ell_u] \mid \delta = R_{1,0} - R_{2,0} > 0\}$, where the upper and lower value of the interval involved must satisfy that $\delta \in [\ell_d, \ell_u]$. Insurance Company 1 (the larger one based on initial capital) then chooses it's policy premium to maximise the probability that the reserve difference hits the upper barrier before it hits the lower. Insurance Company 1 wants to *push* its competitor even further away. Conversely, Insurance Company 2 (the smaller one) chooses it's policy premium to minimise the same probability, i.e. it wants *pull* closer to Insurance Company 1. The value function is then (3.2) with $v(R_{1,s}^{(u_1,u_2)}, R_{2,s}^{(u_1,u_2)}, u_{1,s}, u_{2,s}) = 0$ for all inputs, $\zeta = 0$, and $\mathcal{K}(\tau, R_{1,\tau}, R_{2,\tau}) = \mathbb{1}_{\{R_{1,\tau}-R_{2,\tau}=\ell_u\}}$.

References

Bensoussan, A. and Friedman, A. (1977). "Nonzero-sum stochastic differential games with stopping times and free boundary problems". *Transactions of the American Mathematical Society* 231.2, pp. 275–327.

Bensoussan, A., Siu, C. C., Yam, S. C. P., and Yang, H. (2014). "A class of non-zero-sum stochastic differential investment and reinsurance games". *Automatica* 50.8, pp. 2025–2037.

Bertrand, J. (1883). "Théorie mathématique de la richesse sociale". *Journal des savants* 67.1883, pp. 499–508.

Chen, S., Yang, H., and Zeng, Y. (2018). “Stochastic differential games between two insurers with generalized mean-variance premium principle”. *ASTIN Bulletin* 48.1, pp. 413–434.

Cournot, A. A. (1838). *Recherches sur les principes mathématiques de la théorie des richesses*. L. Hachette.

Elliott, R. J. and Davis, M. (1981). “Optimal play in a stochastic differential game”. *SIAM Journal on Control and Optimization* 19.4, pp. 543–554.

Elliott, R. J. (1977). “The existence of optimal strategies and saddle points in stochastic differential games”. *Differential Games and Applications (Lecture Notes in Control and Information Sciences)*. Vol. 3. Springer, Berlin, Heidelberg, pp. 123–135.

Elliott, R. J. and Siu, T. K. (2011a). “A BSDE approach to a risk-based optimal investment of an insurer”. *Automatica* 47.2, pp. 253–261.

— (2011b). “A stochastic differential game for optimal investment of an insurer with regime switching”. *Quantitative Finance* 11.3, pp. 365–380.

Jin, Z., Yin, G., and Wu, F. (2013). “Optimal reinsurance strategies in regime-switching jump diffusion models: Stochastic differential game formulation and numerical methods”. *Insurance: Mathematics and Economics* 53.3, pp. 733–746.

Nash, J. (1951). “Non-cooperative games”. *Annals of mathematics* 54.2, pp. 286–295.

Ramachandran, K. M. and Tsokos, C. P. (2012). *Stochastic differential games. Theory and applications*. Vol. 2. Springer Science & Business Media.

Taksar, M. and Zeng, X. (2011). “Optimal non-proportional reinsurance control and stochastic differential games”. *Insurance: Mathematics and Economics* 48.1, pp. 64–71.

Taylor, G. C. (1986). “Underwriting strategy in a competitive insurance environment”. *Insurance: Mathematics and Economics* 5.1, pp. 59–77.

von Neumann, J. and Morgenstern, O. (1947). *Theory of games and economic behavior*. 2nd. Princeton University Press.

von Stackelberg, H. (1934). *Marktform und Gleichgewicht*. J. Springer.

Zeng, X. (2010). “A stochastic differential reinsurance game”. *Journal of Applied Probability* 47.2, pp. 335–349.

Paper C

Nash equilibrium premium strategies for push-pull competition in a frictional non-life insurance market

Søren Asmussen, Bent Jesper Christensen
and Julie Thøgersen

Abstract. Two insurance companies I_1, I_2 with reserves $R_{1,t}, R_2 1, t$ compete for customers, such that in a suitable stochastic differential game the smaller company I_2 with $R_{2,0} < R_{1,0}$ aims at minimising $R_{1,t} - R_{2,t}$ by using the premium p_2 as control and the larger I_1 at maximising by using p_1. The dependence of reserves on premia is derived by modelling the customer's problem explicitly, accounting for market frictions V, reflecting differences in cost of search and switching, information acquisition and processing, or preferences. Assuming V to be random across customers, the optimal simultaneous choice p_1^*, p_2^* of premiums is derived and shown to provide a Nash equilibrium for beta distributed V. The analysis is based on the diffusion approximation to a standard Cramér-Lundberg risk process extended to allow investment in a risk-free asset.

Keywords: stochastic differential game; diffusion approximation; exit problem; market friction; Nash equilibrium; saddle point; beta distribution

Publication details: Insurance: Mathematics and Economics, 2019, vol. 87, pp. 92–100.
DOI: 10.1016/j.insmatheco.2019.02.002

C.1 Introduction

The appropriate choice of premium has received massive attention in insurance, from both academics and practitioners. The main approach in the literature is to base the premium on the expected loss, with an added loading calculated from distributional properties of the risk (the expected value principle, variance principle, utility premium, etc.). An alternative to these premium principles is presented in Asmussen et al. (2013) and Thøgersen (2016), where the individual customer's problem of deciding whether or not to insure at the premium offered is modelled explicitly, allowing a derivation of portfolio size as a function of the premium. The premium is in turn chosen optimally by the insurance company, balancing revenue per customer against portfolio size in order to minimise ruin probability. In the present paper, we extend this premium principle to a situation where insurance companies compete against each other, using premiums as controls.

An early reference on the choice of premium in response to market behavior is Taylor (1986), who observed that individual operators in the Australian insurance market followed the market as the average premium increased and declined. The individual insurer's optimal policy was modelled by adopting a demand function specification (price elasticity) for volume as a function of own premium and market premium and assuming that the market does not react to the policy of the individual insurer. Several extensions have been developed under these basic assumptions. Taylor (1987) considers marginal expense rates. Emms and Haberman (2005) generalise the deterministic discrete-time analysis of Taylor to a stochastic continuous-time model. More recent contributions include Pantelous and Passalidou (2013, 2015, 2017), using stochastic demand functions in discrete time, and Emms (2007) and Emms et al. (2007), adopting stochastic processes for the market average premium and demand conditions in continuous time.

Market reaction to the individual insurer's premium is considered by Emms (2011). More generally, the possibility that the other insurers in the market in fact do react to the policy of the individual insurer, and that the individual insurer takes this into account, leads to game theoretic considerations. Rothschild and Stiglitz (1976) are among the first to apply game theory to the insurance market. They consider a market with two types of customers, low-risk and high-risk. Customers know their own risk, but this information is not available to the insurance companies, and high-risk customers impose an externality on low-risk customers: The latter group is worse off, and the former no better off, than without the existence of the other group. If an equilibrium exists, it is separating, and contracts specify both deductibles and premiums, with low-risk customers willing to accept higher deductibles in exchange for lower premiums. Thus, there is product differentiation.

A number of papers in the more recent literature have applied non-cooperative game theory to non-life insurance markets. Premium controls correspond to Bertrand games, e.g., the one-period games in Polborn (1998) and Dutang et al. (2013) and the insolvency risk model in Rees et al. (1999), in contrast to volume controls as in Cournot games, see, e.g., Powers et al. (1998). Emms (2012) and Boonen et al. (2018) consider differential games in premium controls, based on Taylor (1986) and Taylor

(1987) type demand functions of own and market average premium. Boonen et al. (2018) in addition present a continuous-time extension of a one-period aggregate game of Wu and Pantelous (2017), involving a price elasticity of demand or market power parameter, and the individual insurer's payoff depending on own premium and an aggregate of market premia. The models are deterministic and open-loop Nash equilibria are determined, as opposed to closed-loop or feedback Nash equilibria with policies depending on the current state.

Apart from optimal premium policies, game theory has been used for other purposes in insurance. Borch (1962, 1974), Bühlmann (1980, 1984) and Lemaire (1984, 1991) considered cooperative games of risk transfer between insurer and reinsurer. Chen and Shen (2018) study a Stackelberg stochastic differential game between insurer and reinsurer. Reinsurance strategies have further been considered in zero sum stochastic differential games between insurance companies by Zeng (2010), Taksar and Zeng (2011), and Jin et al. (2013), and the analysis has been extended to non-zero sum games and additional investment controls by Bensoussan et al. (2014), nonlinear risk processes by Meng et al. (2015), ambiguity-aversion by Pun and Wong (2016), and insurance companies with different levels of trust in information by Yan et al. (2017). Finally, besides insurance and reinsurance companies, investment games have been considered, e.g., between portfolio investors by Browne (2000) and Espinosa and Touzi (2015), between money managers by Basak and Makarov (2014), and between defined contribution pension funds by Guan and Liang (2016).

In the present paper, we extend the approach of Asmussen et al. (2013) and Thøgersen (2016) (analyzing the optimal premium of a single company) to a competitive environment. The product in question is a stop-loss non-life insurance contract. We explicitly consider the customer's problem and the resulting stochastic differential game between insurance companies in premium strategies (see Taksar and Zeng (2011) for a thorough theoretical treatment of stochastic differential games in insurance). We focus on the case of two insurance companies I_1 and I_2 offering identical contracts but different premiums p_i, $i = 1, 2$. In this situation without product differentiation, it might be expected that all customers would simply insure at the company offering the lowest premium. However, market frictions may imply that when choosing which insurance company to contact, customers face different costs of search and switching, transportation, or information acquisition, or they simply exhibit differences in preferences. The contribution of this paper is to suggest a model for such market friction and analyse its impact on the decisions of customers as well as companies. This provides an alternative to the approach in the literature on dynamic games in premium controls of assuming a demand structure directly, without reference to the customer's problem or market frictions.

We assume that there is a financial market consisting of a single risk-free asset with dynamics $\mathrm{d}B_t = rB_t\,\mathrm{d}t$, where r is the risk-free interest rate. All excess wealth of customers and reserve of insurers is invested in this asset. There are N customers in the insurance market. We assume that all customers must insure at either I_1 or I_2 and focus the analysis on the choice between the two companies.

The characteristics of an individual customer are unknown to the insurance companies, but their probability distribution known. Based on this distribution, the companies can determine the expected portfolio sizes $n_i(p_1, p_2)$ as functions of the premiums offered. The gross premium rate of I_i is then $c_i(p_1, p_2) = n_i(p_1, p_2)p_i$, and the aggregate claim frequency is $\lambda_i(p_1, p_2) = n_i(p_1, p_2)\alpha$ where α is the rate at which a single customer generates claims.

Let $r_{i,0}$ be the initial reserve of company i. For given premiums (p_1, p_2), the reserve of I_i is governed by the dynamics

$$\mathrm{d}R_{i,t} = (\mu_i(p_1, p_2) + rR_{i,t})\,\mathrm{d}t + \sigma_i(p_1, p_2)\,\mathrm{d}W_{i,t}, \tag{3.1}$$

where $(W_{1,t})_{t\geq 0}$ and $(W_{2,t})_{t\geq 0}$ are independent Wiener processes, and

$$\begin{aligned}\mu_i(p_1, p_2) &= c_i(p_1, p_2) - \lambda_i(p_1, p_2)\mathbb{E}X = n_i(p_1, p_2)\big(p_i - \alpha\,\overline{x_1}\big)\,,\\ \sigma_i^2(p_1, p_2) &= \lambda_i(p_1, p_2)\mathbb{E}X^2 = n_i(p_1, p_2)\alpha\,\overline{x_2}\,.\end{aligned}$$

Here $\overline{x_1} = \mathbb{E}X$, $\overline{x_2} = \mathbb{E}X^2$, with the random variable X representing claim sizes, assumed to be independent and identically distributed. Thus, (3.1) can be considered as a diffusion approximation to the Cramér-Lundberg process extended to the case where the insurance companies have access to investment in a risk-free asset. Such diffusion approximations have been used widely, based on the arguments of Iglehart (1969); see, e.g., Schmidli (2008) and references there.

The aim is to derive value functions for the insurance companies, and determine Nash equilibria in the sense of game theory, that is, considering a Bertrand game with both players simultaneously choosing their premiums and looking for a strategy from which no one would wish to depart. We consider here what we call *push-pull competition*. We assume that the largest company in terms of initial capital selects its premium to try to *push* the small company away, while the small company tries to *pull* closer to the large company.

While we here concentrate on identical contracts, we consider in a companion paper (Asmussen et al. (2019)) product differentiation defined via different deductibles. The object of study is a Stackelberg game with one company being the leader and the other the follower, and we look for the corresponding Stackelberg equilibrium.

The structure of the paper is as follows. In Section C.2 we analyse possible criteria for the customer to choose one company over the other. We proceed to consequences for portfolio sizes in Section C.3. In Section C.4, we use these to find the strategies of I_1 and I_2, and we characterise the corresponding equilibrium. The explicit solution is presented in Section C.5 under what appears to be the most natural parametric assumption on our market friction model, and some numerical examples are given. Section C.6 concludes.

C.2 Customer's problem

Customers may face market frictions when deciding which insurance to buy. It takes time and effort to search for the best company, or to switch companies, which creates sluggishness in the market. Further, customers might have different access to information, and/or different abilities to process it. Search frictions have been studied in economics by Diamond (1982), Mortensen (1982), Mortensen and Pissarides (1994), and others. Brown and Goolsbee (2002) studied the effect of internet search on life insurance premiums in US data. Information frictions have been modelled as differences in the cost of obtaining information, e.g., by Salop and Stiglitz (1977).

We adopt a simple device that can be used to capture various types of market frictions or differences in preferences. Following Hotelling (1929), the market is represented by an interval $[0, 1]$ (we normalise it to unit length), with the insurance companies located at the end points; I_1 at 0 and I_2 at 1. This can be interpreted as a street on which the insurance companies are stores. A given customer is then placed at $v \in [0, 1]$ along this "street" and must pay marginal cost c per unit distance to be transported to either of the companies in order to buy insurance. As the customer has distance $a_1 = v$ to I_1, she must pay $c \cdot a_1$ to buy insurance there. Similarly, the distance to I_2 is $a_2 = 1 - v$, and the customer must pay $c \cdot a_2$ to buy insurance there. This is considered as a one-time cost of frictions, to be paid immediately. It may reflect search and switching costs, costs of information acquisition and processing, or simply different preferences for the two companies and their products. Thus, c is a measure of the degree of market frictions. Specifically, if the costs capture preferences, they can be interpreted in terms of the disutility from buying insurance at a less preferred supplier. This frictional insurance market is illustrated in Figure 1.

Figure 3.1: A frictional insurance market.

The customer has access to the risk-free asset paying interest at rate r. This is the customer's only source of income, and she invests all her wealth in this. The customer is exposed to a risk modelled as a compound Poisson process $\sum_{n=1}^{N_t^\alpha} X_n$, where $(N_t^\alpha)_{t\geq 0}$ is a Poisson process with claim frequency α, and $(X_n)_{n\in\mathbb{N}}$ are the claim sizes, assumed to be i.i.d. and independent of $(N_t^\alpha)_{t\geq 0}$. The customer will then reduce risk by buying insurance. Her calculations are assumed to be using a time horizon $T \leq \infty$ (possibly stochastic, see further Remark C.1 below) and a discount rate d; here one has either $d = r$ or d may be subjective, in which case the traditional transversality assumption in economics is $d > r$, cf. Gordon (1959), accommodating the case $T = \infty$ by allowing present value calculations based on flows while ignoring terminal value at infinity.

We will now present two slightly different approaches which both lead to the customer preferring I_i over I_j if

$$p_i - p_j < \rho c(a_j - a_i) \tag{3.2}$$

for some constant ρ depending on d and the distribution of T. For $\beta > 0$, write

$$\vartheta(\beta) = \mathbb{E} \int_0^T \mathrm{e}^{-\beta t}\, \mathrm{d}t = \frac{1}{\beta}(1 - \mathbb{E}\mathrm{e}^{-\beta T}),$$

and let $\vartheta(0) = \mathbb{E}T$. The first and simplest approach is to calculate the discounted expenses up to T by insuring at company i, resp. j. These are

$$ca_i + \mathbb{E}\int_0^T \mathrm{e}^{-dt} p_i \,\mathrm{d}t = ca_i + \vartheta(d)p_i\,, \quad \text{resp. } ca_j + \vartheta(d)p_j\,.$$

Company I_i will be preferred if the difference is negative, which gives (3.2) with $\rho = 1/\vartheta(d)$.

The second approach is close to Asmussen et al. (2013) and Thøgersen (2016) and assumes $d > r$ if $T = \infty$. Let $w_{i,t}$ be the wealth of the customer at time t when insuring at company i and $w_0 = w_{i,0-}$ her initial wealth. The dynamics are

$$\mathrm{d}w_{i,t} = (rw_{i,t} - p_i)\,\mathrm{d}t, \quad w_{i,0} = w_0 - ca_i\,.$$

The solution of this ODE is

$$w_{i,t} \;=\; \mathrm{e}^{rt}(w_0 - ca_i - p_i/r) + p_i/r\,, \quad \text{giving } \mathrm{d}w_{i,t} \;=\; r\mathrm{e}^{rt}(w_0 - ca_i - p_i/r)\,.$$

The expected total discounted incremental wealth is therefore

$$-ca_i + \mathbb{E}\int_0^T \exp(-dt)\mathrm{d}w_{i,t} = -ca_i + \big[r(w_0 - ca_i) - p_i\big]\vartheta(d-r)\,.$$

If the customer's criterion is to choose the company with the larger value of this quantity, we therefore again arrive at (3.2), this time with $\rho = r + 1/\vartheta(d-r)$.

In summary (recalling that $a_1 = v$ and $a_2 = 1 - v$), the customer will choose I_1 over I_2 if

$$p_1 - p_2 < \rho c(1 - 2v)\,, \tag{3.3}$$

and conversely, I_2 over I_1 if

$$p_2 - p_1 < \rho c(2v - 1)\,. \tag{3.4}$$

The customer is indifferent if equality holds, but this case becomes unimportant in the following since v is taken to be random with a typically continuous distribution.

Remark C.1. The time horizon T could be the time the customer expects to need insurance, or the time in which she expects the companies to keep premiums relatively constant. She could also just take $T = \infty$, in which case ρ reduces to d, the discount rate, in both cases.

We next use relations (3.3) and (3.4) to evaluate the portfolio sizes of the respective companies.

C.3 Portfolio sizes

The location (friction or preference) parameters v of the customers will be considered as random to the firm and denoted by V, assumed to have a continuous distribution. The case (3.3) then corresponds to the event

$$\Omega = \big\{p_1 - p_2 < \rho c(1 - 2V)\big\} \tag{3.5}$$

and (3.4) to the complementary event Ω^{c}. Letting

$$v_0 = \frac{1}{2}\Big(1 - \frac{p_1 - p_2}{\rho c}\Big)\,, \tag{3.6}$$

we have $\Omega = \{V < v_0\}$ and so the expected portfolio sizes are

$$\begin{aligned} n_1(p_1, p_2) &= N\mathbb{P}(\Omega) = N\mathbb{P}(V < v_0)\,, \\ n_2(p_1, p_2) &= N\mathbb{P}(\Omega^{\mathrm{c}}) = N\big(1 - \mathbb{P}(V < v_0)\big)\,. \end{aligned} \tag{3.7}$$

There are two different ways to think of N, the number of customers needing insurance. Either N is fixed and the customers make their choice of company from the start $t = 0$. Or, maybe more naturally, $N = N(t)$ fluctuates randomly. The natural model is then to assume that customers arrive according to a Poisson process, that they need insurance over a period of random length L, that the Ls of different customers are i.i.d., and that the process $N(t)$ is stationary. Such a process is often called a coverage process and coincides with the stationary process of number of customers in the GI/G/∞ queue. With T in Section C.2 being the remaining time the customer needs insurance, T then decreases linearly with time, and so the individual customer will dynamically redo her calculations with the current T (and premiums). However, in a large portfolio the T will simply have the equilibrium distribution of L, i.e., one should take T to have density $\mathbb{P}(L > t)/\mathbb{E}L$.

Remark C.2. The large-portfolio dynamics of a coverage process follows from work of Whitt (1982) and Glynn (1982), giving that $N(t)$ is approximately a certain stationary Gaussian process with characteristics depending on the distribution of T; only the case of T being exponential is simpler, and then $N(t)$ is approximately Ornstein-Uhlenbeck. Fortunately, the dynamics of $N(t)$ will be unimportant since we will see that $N = N(t)$ can just be treated as constant in our game theoretic problem.

Remark C.3. A Bayesian view of letting the customer characteristics (here the frictions V) fluctuate randomly among customers is quite common in insurance modelling. A main example is credibility theory (Bühlmann and Gisler (2006)), another modelling of arrivals of claims as a mixed Poisson process (Grandell (1997)).

Since the market frictions V take values in $[0, 1]$, the obvious parametric choice is a beta distribution, and we treat that case in more detail in Section C.5.

C.4 The strategies of the insurance companies

We now consider the optimization problems of the insurance companies. Inspired by Taksar and Zeng (2011), we analyze the competition among companies using a stochastic differential game.

A control $\pi = (\pi_1, \pi_2)$ is a set $\big(\pi_1(t), \pi_2(t)\big)$ of premium strategies where $\pi_1(t)$, $\pi_2(t)$ denote the premiums set by the companies at time t and are subject to the restrictions

$$\pi_1(t) \geq \underline{p_1}, \quad \pi_2(t) \geq \underline{p_2} \tag{3.8}$$

where $\underline{p_1}, \underline{p_2} \geq -\infty$. If say $\underline{p_1} = -\infty$, this would mean that I_1 could potentially pay customers money to have them insured, in the hope that this loss of customers for I_2 would cause an even bigger deficit for that company. This may seem quite unnatural from a practical point of view, but is on the other hand in line with the set-up of control theory and game theory which is one-eyed by assuming that the decision makers have only one objective in mind. Think for example of dividend optimization, where the optimal strategy in basic examples leads to ruin with probability one, cf. Schmidli (2008). Similar remarks apply to the requirements $\underline{p_1} = \alpha\,\overline{x_1}$, $\underline{p_2} = \alpha\,\overline{x_1}$ of premiums not going below the net premium and $\underline{p_1} = 0$, $\underline{p_2} = 0$ of non-negative premiums. Some illustration of these issues is given below in Remark C.12.

Following Taksar and Zeng (2011), we will only consider Markovian (also called feedback) strategies π, meaning that $\pi_1(t), \pi_2(t)$ only depend on the current value δ of the difference $\Delta^\pi(t) = R^\pi_{1,t} - R^\pi_{2,t}$ between the corresponding controlled reserve processes R^π_1, R^π_2. That is, we can write $\pi_1(t) = p^\pi_1(\Delta^\pi(t))$, $\pi_2(t) = p^\pi_2(\Delta^\pi(t))$ for suitable functions p^π_1, p^π_2. Since the (uncontrolled) reserves have the dynamics of (3.1), this makes $\Delta^\pi(t)$ a diffusion process,

$$\mathrm{d}\Delta^\pi(t) = \mu^\pi(\Delta^\pi(t))\,\mathrm{d}t + \sigma^\pi(\Delta^\pi(t))\,\mathrm{d}W(t)\,, \tag{3.9}$$

where

$$\begin{aligned}\mu^\pi(\delta) &= \mu_1\big(p^\pi_1(\delta), p^\pi_2(\delta)\big) - \mu_2\big(p^\pi_1(\delta), p^\pi_2(\delta)\big) + r\delta\,,\\ \sigma^\pi(\delta)^2 &= \sigma_1\big(p^\pi_1(\delta), p^\pi_2(\delta)\big)^2 + \sigma_2\big(p^\pi_1(\delta), p^\pi_2(\delta)\big)^2\,,\end{aligned}$$

and $W = (W_1 - W_2)/\sqrt{2}$ is again a Wiener process. Without loss of generality, we take $\Delta(0) = \Delta^\pi(0) = r_{1,0} - r_{2,0} > 0$, i.e., I_1 is the large company and I_2 the small. The large company seeks to maximise the reserve difference (to *push* the competitor further away), while the small company seeks to minimise the same (to *pull* closer to the competitor), each taking the current reserve difference as the state variable. The optimality criterion is to consider a fixed interval $[\ell_d, \ell_u]$ with $\ell_d < \Delta(0) < \ell_u$ and let

$$\begin{aligned}\tau(\pi) &= \inf\big\{t > 0 : \Delta^\pi(t) \notin [\ell_d, \ell_u]\big\}\,,\\ V^\pi(\delta) &= \mathbb{P}^\pi\big(\Delta^\pi(\tau(\pi)) = \ell_u \,\big|\, \Delta(0) = \delta\big)\,.\end{aligned}$$

Then the large company I_1 chooses π_1 in order to maximise the probability $V^\pi(\Delta(0))$ to exit at the upper boundary, and the small I_2 chooses π_2 to minimise $V^\pi(\Delta(0))$, or equivalently to maximise the probability $1 - V^\pi(\Delta(0))$ to exit at the lower boundary.

Remark C.4. The feedback assumption implies that this is equivalent to maximising (minimising) $V^\pi(\delta)$ for all $\ell_d < \delta < \ell_u$. The choice of the interval (ℓ_d, ℓ_u) may appear arbitrary, but we will see later that once we have found a Nash equilibrium for a given choice, it will also be Nash equilibrium for any larger interval. This is due to specific features of the problem we consider and certainly not a general principle, but it strongly supports that our optimality criterion is more universal than it looks at a first sight.

Remark C.5. Our formulation of the control problem is basically the same as in Taksar and Zeng (2011) (TZ), although they consider reinsurance strategies whereas we study premium controls and model the customer's problem and market frictions, but it should be remarked that as in TZ, issues associated with ruin are suppressed. An alternative taking ruin into account is to let the control depend on the reserve levels R_1, R_2 rather than just $\Delta = R_1 - R_2$. If τ_1, τ_2 are the ruin times of the two companies and as before τ is the exit time of the reserve difference from $[\ell_d, \ell_u]$, we can then split up into four different outcomes of the game,

$$
\begin{gathered}
F_1 : \Delta(\tau) = \ell_u, \tau < \tau_1 \wedge \tau_2, \quad F_2 : \Delta(\tau) = \ell_d, \tau < \tau_1 \wedge \tau_2, \\
F_3 : \tau_1 < \tau \wedge \tau_2, \quad F_4 : \tau_2 < \tau \wedge \tau_1.
\end{gathered}
$$

Company I_1 would then go for maximising $\mathbb{P}(F_1 \cup F_4)$ and I_2 for maximising $\mathbb{P}(F_2 \cup F_3)$.

The motivation for ignoring ruin (as TZ and we do) could among other possibilities either be that $r_{1,0}, r_{2,0}$ are so large and $\ell_u - \ell_d$ so small that the probabilities of F_3 and F_4 can be ignored, or that investors would reinvest in the case of ruin.

A Nash equilibrium is defined as a strategy pair (π_1^*, π_2^*) satisfying

$$
V^{(\pi_1^*, \pi_2^*)} \geq V^{(\pi_1, \pi_2^*)} \text{ for all } \pi_1 \quad \text{and} \quad V^{(\pi_1^*, \pi_2^*)} \leq V^{(\pi_1^*, \pi_2)} \text{ for all } \pi_2, \tag{3.10}
$$

i.e., neither firm has an incentive to deviate from its strategy unilaterally. An equivalent formulation is that $\pi_1^* = \widehat{\pi}_1(\pi_2^*)$ where $\widehat{\pi}_1(\pi_2)$ is the optimal choice for I_1 in a single-company problem, treating π_2 as fixed, and that similarly $\pi_2^* = \widehat{\pi}_2(\pi_1^*)$ in obvious notation.

We now give some discussion that will allow replacing optimization problems in the space of functions p_1, p_2 by the more elementary problems of pointwise maximization/minimization of the real-valued ratio

$$
\kappa^\pi(\delta) = \frac{\mu^\pi(\delta)}{\sigma^\pi(\delta)^2} \tag{3.11}
$$

between the drift and variance of the reserve difference process in (3.9).

Lemma C.6. *Let $\mu(x), \sigma^2(x)$ be bounded and measurable functions on an interval (ℓ_d, ℓ_u) such that $\inf_{\ell_d < x < \ell_u} \sigma^2(x) > 0$ and let X, W be defined on a suitable probability space such that W is a standard Brownian motion and*

$$
X(t) = \delta + \int_0^t \mu\big(X(s)\big)\,\mathrm{d}s + \int_0^t \sigma\big(X(s)\big)\,\mathrm{d}W(s)\,, \tag{3.12}
$$

for some $\delta \in (\ell_d, \ell_u)$. Define further $\kappa(x) = \mu(x)/\sigma^2(x)$,

$$
s(y) = \exp\Big\{-2\int_{\ell_d}^y \kappa(z)\,\mathrm{d}z\Big\}, \quad S(\delta) = \int_{\ell_d}^\delta s(y)\,\mathrm{d}y\,,
$$

and $\tau = \inf\{t : X(t) \notin (\ell_d, \ell_u)\}$*. Then:*

(i) $\mathbb{P}\big(X(\tau) = \ell_u\big) = S(\delta)/S(\ell_u)$.

(ii) *For a given function* κ *on* $[\ell_d, \ell_u]$ *and a given* $\delta \in [\ell_d, \ell_u]$*, let* $\varphi(\kappa)$ *denote the r.h.s. in* (i)*. Then* $\kappa_0 \le \kappa_1$ *implies* $\varphi(\kappa_0) \le \varphi(\kappa_1)$*.*

Proof. We initially remark that the conditions on $\mu(x), \sigma^2(x)$ ensure the existence of X as a weak solution. (i) This is a standard general formula for diffusions, typically stated under smoothness conditions ensuring the applicability of Itô's formula. For the present version, note first that the absolute continuity and boundedness of $\int_{\ell_d}^y \kappa(z)\,\mathrm{d}z$ ensure that $s(y)$ is of bounded variation, hence of the form $s_1(y) - s_2(y)$ for non-decreasing functions $s_1(y), s_2(y)$ of bounded variation. Letting $S_i(\delta) = \int_{\ell_d}^{\delta} s_i(y)\,\mathrm{d}y$, S_i is then convex (Problem 6.20 p. 213 in Karatzas and Shreve (1998)) and we may apply the Itô-Tanaka formula (Karatzas and Shreve (1998) Section 5.5.B or Rogers and Williams (2000) Section IV.45) to each S_i separately to conclude that

$$S\big(X(t)\big) = S(\delta) + \int_0^t s\big(X(v)\big)\sigma\big(X(v)\big)\,\mathrm{d}W(v)$$

is a local martingale. The boundedness properties of $s(x), \sigma^2(x)$ ensure that we indeed have a proper martingale, and so by optional stopping

$$S(\delta) = \mathbb{E}S\big(X(\tau)\big) = S(\ell_u)\mathbb{P}\big(X(\tau) = \ell_u\big) + S(\ell_d)\big(1 - \mathbb{P}\big(X(\tau) = \ell_u\big)\big)\,.$$

For (ii), define

$$\xi = \kappa_1 - \kappa_0\,, \quad \kappa_t = \kappa_0 + t\xi = (1-t)\kappa_0 + t\kappa_1\,, \quad g(t) = \varphi(\kappa_t)\,.$$

Then

$$g'(t) = \lim_{h \downarrow 0} \frac{\varphi(\kappa_{t+h}) - \varphi(\kappa_t)}{h} = \lim_{h \downarrow 0} \frac{\varphi(\kappa_t + h\xi) - \varphi(\kappa_t)}{h} = \varphi_\xi(\kappa_t)\,,$$

$$\varphi(\kappa_1) - \varphi(\kappa_0) = g(1) - g(0) = \int_0^1 g'(t)\,\mathrm{d}t = \int_0^1 \varphi_\xi(\kappa_t)\,\mathrm{d}t\,,$$

where φ_ξ denotes the directional derivative. Thus (ii) will follow if we can show $\varphi_\xi(\kappa_t) \ge 0$ for all t. Define for a fixed t

$$H(x) = \exp\Big\{-2\int_{\ell_d}^x \kappa_t(y)\,\mathrm{d}y\Big\}\,, \qquad K(x) = \int_{\ell_d}^x \xi(y)\,\mathrm{d}y\,,$$

$$A(\delta) = \int_{\ell_d}^{\delta} H(x)\,\mathrm{d}x\,, \qquad B(\delta) = \int_{\ell_d}^{\delta} H(x)K(x)\,\mathrm{d}x\,.$$

Then $g(t) = \varphi(\kappa_t) = A(\delta)/A(\ell_u)$ and

$$\begin{aligned}\varphi(\kappa_t + h\xi) &= \frac{\displaystyle\int_{\ell_d}^{\delta} \exp\Big\{-2\int_{\ell_d}^{x} \kappa_t(y)\,\mathrm{d}y - 2h\int_{\ell_d}^{x} \xi(y)\,\mathrm{d}y\Big\}\,\mathrm{d}x}{\displaystyle\int_{\ell_d}^{\ell_u} \exp\Big\{-2\int_{\ell_d}^{x} \kappa_t(y)\,\mathrm{d}y - 2h\int_{\ell_d}^{x} \xi(y)\,\mathrm{d}y\Big\}\,\mathrm{d}x}\\ &\approx \frac{\displaystyle\int_{\ell_d}^{\delta} H(x)\big(1-2hK(x)\big)\,\mathrm{d}x}{\displaystyle\int_{\ell_d}^{\ell_u} H(x)\big(1-2hK(x)\big)\,\mathrm{d}x} = \frac{A(\delta)-2hB(\delta)}{A(\ell_u)-2hB(\ell_u)}\\ &= \varphi(\kappa_t)\frac{1-2hB(\delta)/A(\delta)}{1-2hB(\ell_u)/A(\ell_u)}\\ &\approx \varphi(\kappa_t)\big(1-2hB(\delta)/A(\delta)\big)\big(1+2hB(\ell_u)/A(\ell_u)\big)\end{aligned}$$

which gives $\varphi_\xi(\kappa_t) = 2\varphi(\kappa_t)\big[C(\ell_u) - C(\delta)\big]$ where $C(\delta) = B(\delta)/A(\delta)$. Thus it suffices for $\varphi_\xi(\kappa_t) \geq 0$ that C is non-decreasing, which follows since $H \geq 0$ implies

$$C'(\delta) = \frac{A(\delta)B'(\delta) - A'(\delta)B(\delta)}{A(\delta)^2} = \frac{H(\delta)\big[A(\delta)K(\delta) - B(\delta)\big]}{A(\delta)^2} \geq 0,$$

where we used that $\xi \geq 0$ and $\delta \geq x$ entails

$$\begin{aligned}K(\delta) &\geq K(x),\\ A(\delta)K(\delta) &= \int_{\ell_d}^{\delta} H(x)K(\delta)\,\mathrm{d}x \geq \int_{\ell_d}^{\delta} H(x)K(x)\,\mathrm{d}x = B(\delta)\,.\end{aligned}$$

□

Remark C.7. Results of similar type are in Hipp and Taksar (2010) and Taksar and Zeng (2011), but the present version noted in Pestien and Sudderth (1985) (PS) avoids the Hamilton-Jacobi-Bellman equation and certain related smoothness and verification issues. Note that PS use a generalization of Itô's formula given as Theorem 2.10.1 in Krylov (1980) rather than the by now more standard Itô-Tanaka theory we use. Our proof of (ii) is also different. The paper PS seems to have gone relatively unnoticed in the non-life stochastic control literature, but it has recently been exploited in Bäuerle and Bayraktar (2014) for related purposes.

With a slight abuse of notation, define

$$\kappa(p_1, p_2;\, \delta) = \frac{\mu_1(p_1,p_2) - \mu_2(p_1,p_2) + r\delta}{\sigma_1^2(p_1,p_2) + \sigma_2^2(p_1,p_2)}\,, \qquad p_1, p_2 \geq 0,\ \ell_d \leq \delta \leq \ell_u\,. \tag{3.13}$$

To ease notation here and in the following subsections, we use the notation

$$\kappa_i'(p_1,p_2;\,\delta) = \frac{\partial}{\partial p_i}\kappa(p_1,p_2;\,\delta)\,, \quad \kappa_{ij}''(p_1,p_2;\,\delta) = \frac{\partial^2}{\partial p_i \partial p_j}\kappa(p_1,p_2;\,\delta)\,,$$

for partial derivatives, where the number of primes indicates the number of times the function is differentiated, and the subscript specifies with respect to which variable.[1]

Corollary C.8. *The optimal strategy* $\widehat{\pi}_1(\pi_2) = \big(\widehat{p}_1(\delta|\pi_2)\big)_{\ell_d \le \delta \le \ell_u}$ *for* I_1 *at fixed premium strategy* $\pi_2 = \big(p_2(\delta)\big)_{\ell_d \le \delta \le \ell_u}$ *for* I_2 *is given by* $\widehat{p}_1(\delta|\pi_2) = \mathrm{argmax}_{p_1} \kappa(p_1, p_2(\delta);\, \delta)$. *The first and second order conditions are that* $p_1 = \widehat{p}_1(\delta|\pi_2)$ *satisfies*

$$0 = \kappa_1'(p_1, p_2;\, \delta)\,, \quad \textit{resp.} \quad 0 > \kappa_{11}''(p_1, p_2;\, \delta) \tag{3.14}$$

for each $\delta \in (\ell_d, \ell_u)$ *where* $p_2 = p_2(\delta)$. *Similarly, the first and second order conditions for* $\widehat{p}_2(\delta|\pi_1)$ *are that* $p_2 = \widehat{p}_2(\delta|\pi_1)$ *satisfies*

$$0 = \kappa_2'(p_1, p_2;\, \delta)\,, \quad \textit{resp.} \quad 0 < \kappa_{22}''(p_1, p_2;\, \delta), \qquad \textit{where } p_1 = p_1(\delta). \tag{3.15}$$

Proof. The argmax expression is obvious from (ii) in Lemma C.6 and Remark C.4. Determining this argmax then means pointwise maximization, and (3.14)–(3.15) are just the standard calculus characterization of this. □

Combining (3.1) and (3.7), the drift and variance of the reserve process of I_1 can be written as

$$\begin{aligned} \mu_1(p_1, p_2) &= N\mathbb{P}(V < v_0)(p_1 - \alpha\,\overline{x_1}), \\ \sigma_1(p_1, p_2)^2 &= N\mathbb{P}(V < v_0)\alpha\,\overline{x_2}, \end{aligned} \tag{3.16}$$

and for I_2,

$$\begin{aligned} \mu_2(p_1, p_2) &= N\big(1 - \mathbb{P}(V < v_0)\big)(p_2 - \alpha\,\overline{x_1}), \\ \sigma_1(p_1, p_2)^2 &= N\big(1 - \mathbb{P}(V < v_0)\big)\alpha\,\overline{x_2}, \end{aligned} \tag{3.17}$$

where v_0 is given by (3.6). Note in particular that

$$\sigma^\pi(\delta)^2 = \sigma_1(p_1, p_2)^2 + \sigma_2(p_1, p_2)^2 = N\alpha\,\overline{x_2}$$

is independent of premiums, and hence so is the $r\delta/\sigma^\pi(\delta)^2$ term in (3.11). Also, the N in the remaining term $(\mu_1 - \mu_2)/{\sigma^\pi}^2$ cancels, thereby justifying Remark C.2. Altogether, combining with Lemma C.6 (ii), we see that the optimization problem

$$\kappa(p_1^*, p_2^*;\, \delta) = \sup_{p_1}\ \kappa(p_1, p_2^*;\, \delta) = \inf_{p_2}\ \kappa(p_1^*, p_2;\, \delta) \tag{3.18}$$

for a Nash equilibrium takes the form

$$\nu(p_1^*, p_2^*) = \sup_{p_1}\ \nu(p_1, p_2^*) = \inf_{p_2}\ \nu(p_1^*, p_2) \tag{3.19}$$

for the present model, where $\nu(p_1, p_2) = \mu_1(p_1, p_2) - \mu_2(p_1, p_2)$. This makes optimal premiums time-invariant, not dependent on the running reserve difference. If one of the insurance companies chooses to increase its premium (while the competitor keeps its premium constant), it will inevitably lower its portfolio size (and increase the competitor's). Hence, the total effect on ν is non-trivial. Involving Corollary C.8 gives:

[1] The standard notation avoids the primes and considers the subscript as sufficient, but we want to emphasise the differentiation here in order to avoid confusion with other notation in the paper, e.g., $\mu_1(p_1, p_2)$ and $\mu_2(p_1, p_2)$.

Proposition C.9. *The strategy* (p_1^*, p_2^*) *in* (3.19) *with premiums not dependent on the running reserve difference is a Nash equilibrium provided it satisfies* (3.8) *together with the first order conditions*

$$0 = \nu_1'(p_1^*, p_2^*) = \nu_2'(p_1^*, p_2^*) \tag{3.20}$$

and the second order conditions

$$0 > \nu_{11}''(p_1^*, p_2^*), \qquad 0 < \nu_{22}''(p_1^*, p_2^*). \tag{3.21}$$

Proof. Condition (3.18) follows from the definition (3.10) of Nash equilibrium and Lemma C.6. With the denominator of (3.13) not depending on premiums, (3.18) reduces to (3.19). Treating p_2^* as given, I_1 chooses p_1 to satisfy the first and second order conditions for a maximum, respectively given by

$$0 = \nu_1'(p_1, p_2^*), \qquad 0 > \nu_{11}''(p_1, p_2^*),$$

and simultaneously, treating p_1^* as given, I_2 finds p_2^* as the solution to

$$0 = \nu_2'(p_1^*, p_2), \qquad 0 < \nu_{22}''(p_1^*, p_2).$$

Solving the first order conditions yields (3.20). In order for these to be an actual Nash equilibrium, the combined second order conditions yield (3.21). □

Hence, solving for a Nash equilibrium reduces to finding a saddle point of the function $\nu(p_1, p_2)$, with a relative maximum along the first axis and a relative minimum along the second. This is in line with the results of Taksar and Zeng (2011). Specifically, as we consider two-player, zero-sum games, such that the opponents have directly conflicting interests, any Nash equilibrium is a saddle point.

Existence of Nash equilibrium is not guaranteed. The famous existence proof by Nash (1951) for n-player games was specific to the case where each player has a finite number of available actions and possibly chooses a mixed (randomised) strategy over these. For two-player, zero-sum games (as in our case), the corresponding result is the general existence of a saddle point in mixed strategies, which follows from the min-max theorem in matrix theory by von Neumann (1928). In contrast, we consider a non-compact continuum of possible actions (premiums) for each company, and seek pure (non-randomised) strategy equilibria in a stochastic differential game setting.

Remark C.10. The fact that p_1^*, p_2^* do not depend on ℓ_d, ℓ_u substantiates the earlier made claim that the choice of these quantities plays essentially no role at all. Further, the lack of dependence on δ shows that in Nash equilibrium, premiums are frozen, implying that the customer has no incentive to change strategy (switch insurer) as time evolves.

C.5 The beta case

We now consider the main case $V \sim \text{beta}(a, b)$, with c.d.f. $\mathbb{P}(V < v) = B(v; a, b) \cdot B(a, b)^{-1}$ for $v \in (0, 1)$, where

$$B(v; a, b) = \int_0^v t^{a-1}(1-t)^{b-1}\,\mathrm{d}t, \qquad B(a, b) = B(1; a, b)$$

denote the incomplete, resp. the complete Beta function. Further, $\mathbb{P}(V < v) = 1$ for $v > 1$ and $\mathbb{P}(V < v) = 0$ for $v < 0$. Examples of beta distributed market frictions (locations) are illustrated in Figure 3.2 for different values of the parameters a and b. A customer located in $[0, 0.5)$ incurs a lower cost if insuring at I_1 than at I_2, or has a natural preference for the former, while a customer located in $(0.5, 1]$ prefers or minimises cost at I_2. Unless the distribution is symmetric around 0.5, the point of indifference/equal cost, one of the companies has an advantage in terms of customer locations/preferences. For example, from Figure 3.2, if $a = 8$ and $b = 2$, then a large share of customers are located near I_2, i.e., I_2 has an advantage over I_1, and vice versa for the case $a = 3$ and $b = 6$. If $a = 1$ and $b = 1$ then customers are uniformly distributed between the two insurance companies.

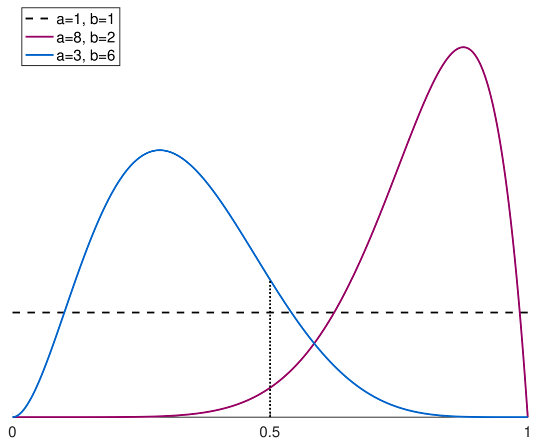


Figure 3.2: Examples of beta market friction densities.

The portfolio sizes (3.7) are now

$$n_1(p_1, p_2) = N\frac{B(v_0; a, b)}{B(a, b)} \quad \text{and} \quad n_2(p_1, p_2) = N\left(1 - \frac{B(v_0; a, b)}{B(a, b)}\right) \tag{3.22}$$

if $v_0 \in (0, 1)$ or, equivalently, if $p_1 - p_2 \in (-\rho c, \rho c)$. If $p_1 - p_2 \geq \rho c$ then the premium charged by I_1 is too high to attract any customers, i.e., $n_1(p_1, p_2) = 0$ and $n_2(p_1, p_2) = N$. Conversely, if $p_1 - p_2 \leq -\rho c$, then I_1 gets the entire market, $n_1(p_1, p_2) = N$ and $n_2(p_1, p_2) = 0$.

Theorem C.11. *Let* $V \sim \mathrm{beta}(a, b)$ *and let* m_β *be the median of* $\mathrm{beta}(a, b)$*. Then a Nash equilibrium exists at*

$$p_1^* = \alpha\,\overline{x_1} + \frac{\rho c}{2}\left(\frac{B(a,b)}{m_\beta^{a-1}(1-m_\beta)^{b-1}} + 1 - 2m_\beta\right),$$

$$p_2^* = \alpha\,\overline{x_1} + \frac{\rho c}{2}\left(\frac{B(a,b)}{m_\beta^{a-1}(1-m_\beta)^{b-1}} - 1 + 2m_\beta\right),$$

provided that p_1^*, p_2^* *are in the feasible region* (3.8) *and that*

$$-4 \le \left(\frac{a-1}{m_\beta} - \frac{b-1}{1-m_\beta}\right)\frac{B(a,b)}{m_\beta^{a-1}(1-m_\beta)^{b-1}} \le 4\,. \tag{3.23}$$

In equilibrium, $n_1(p_1^*, p_2^*) = n_2(p_1^*, p_2^*) = N/2$.

Remark C.12. The verification of conditions (3.8), (3.23) meets the difficulty that there is no explicit analytic expression for the median m_β in terms of a, b (it can be calculated numerically, or for $a, b > 1$ approximated as $m_\beta \approx (a - 1/3)/(a + b - 2/3)$, cf. Kerman (2011)). Aspects of the conditions are illustrated in Figure 3.3 with m_β calculated via Matlab's `median` routine. The range is $0 \le a, b \le 5$ and the conclusion is basically that the conditions are only violated for highly skewed V, that is, if either a or b is close to 0.

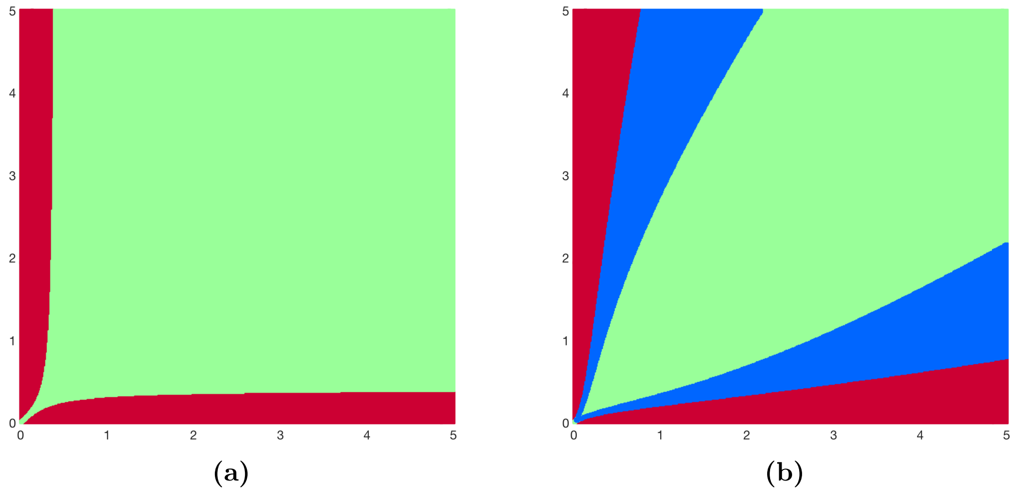


Figure 3.3: (a) Region (3.23) in green. (b) Regions related to (3.8).

In more detail, (3.23) comes out from the proof as the second order condition for a saddlepoint at p_1^*, p_2^*. The region where it holds (fails) is the green (red) region in (a). The picture for condition (3.8) with $\underline{p_1} = \underline{p_2} = 0$ (positive premiums) is somewhat more diverse since it involves not just a, b but also $\alpha\,\overline{x_1}$. More precisely, (3.8) will hold for all values of $\alpha\,\overline{x_1}$ provided

$$\frac{B(a,b)}{m_\beta^{a-1}(1-m_\beta)^{b-1}} \ge |1 - 2m_\beta|. \tag{3.24}$$

This is the green region in (b). Otherwise, (3.8) will fail for sufficiently small values of $\eta = \alpha\,\overline{x_1}/(\rho c/2)$. However, such values appear to be somewhat unrealistic. As examples, we took $\rho = 3\%$, $\eta = 2$ or $\eta = 1/2$. Then for $\eta = 2$, the critical value of the cost c is $\alpha\,\overline{x_1}/\rho$, the expected discounted value of all premiums ever to be paid, and it was found numerically that (3.8) holds everywhere in the considered region $1 \leq a, b \leq 5$. For $\eta = 1/2$, (3.8) continues to hold in the blue region in (b), even if (3.24) fails there, and is only violated in the red region. Alternatively, if condition (3.8) is tightened to $\underline{p_1} = \underline{p_1} = \alpha\,\overline{x_1}$ (premiums never below net level), then this is equivalent to (3.24), i.e., the green region in (b).

Proof. According to Proposition C.9, it suffices to consider the difference in drifts, which using (3.22), (3.16) and (3.17) can be written explicitly as

$$\nu(p_1, p_2) = N\frac{B(v_0; a, b)}{B(a, b)}(p_1 - \alpha\,\overline{x_1}) - N\left(1 - \frac{B(v_0; a, b)}{B(a, b)}\right)(p_2 - \alpha\,\overline{x_1}).$$

The first order conditions are

$$\nu_1'(p_1, p_2) = N\frac{B(v_0; a, b)}{B(a, b)} - \frac{N}{2\rho c}\frac{v_0^{a-1}(1-v_0)^{b-1}}{B(a, b)}(p_1 + p_2 - 2\alpha\,\overline{x_1}) = 0,$$
$$\nu_2'(p_1, p_2) = \frac{N}{2\rho c}\frac{v_0^{a-1}(1-v_0)^{b-1}}{B(a, b)}(p_1 + p_2 - 2\alpha\,\overline{x_1}) - N\left(1 - \frac{B(v_0; a, b)}{B(a, b)}\right) = 0.$$

Combination of these two conditions shows that v_0 must satisfy

$$\frac{B(v_0; a, b)}{B(a, b)} = 1 - \frac{B(v_0; a, b)}{B(a, b)},$$

so v_0 must be the median. Setting (3.6) equal to m_β, it follows that $p_1 = p_2 + \rho c(1 - 2m_\beta)$. Together with the first order conditions, we get the claimed form of p_1^*, p_2^*. The second order partial derivative test in Appendix C.I verifies that we have a Nash equilibrium under the stated parameter restrictions. □

Remark C.13. If $a = b$ then $m_\beta = 1/2$ due to symmetry and premiums coincide, $p_1^* = p_2^*$. If further $a = b = 1$, then the beta distribution reduces to the uniform, and equilibrium premiums are $p_1^* = p_2^* = \alpha\,\overline{x_1} + \rho c/2$.

Remark C.14. Without market frictions, $c = 0$, or if $\rho = 0$, premiums are actuarially fair, $p_i^* = \alpha\,\overline{x_1}$. The higher the cost of frictions or ρ, the greater the departure from actuarially fair premiums.

C.5.1 Numerical example

We present a numerical illustration based on the example with $a = 8, b = 2$, where I_2 has a market advantage. In this case, median market frictions are $m_\beta = 0.82$. Note that for these parameter values, condition (3.23) is satisfied. Say that gross claim sizes Z^* are exponentially distributed with parameter θ but that the contract involves a deductible K. Then

$$\begin{aligned}
\overline{x_1} &= \mathbb{E}[(Z^* - K)^+] \\
&= \mathbb{P}(Z^* \geq K)\mathbb{E}[Z^* - K \mid Z^* \geq K] \\
&= \frac{1}{\theta}\mathrm{e}^{-\theta K}, \\
\overline{x_2} &= \mathbb{E}[(Z^* - K)^2 \mathbb{1}_{\{Z^* \geq K\}}] \\
&= \mathbb{E}[Z^{*2}\mathbb{1}_{\{Z^* \geq K\}}] + K^2\mathbb{E}[\mathbb{1}_{\{Z^* \geq K\}}] - 2K\mathbb{E}[Z^*\mathbb{1}_{\{Z^* \geq K\}}] \\
&= \frac{2}{\theta^2}\mathrm{e}^{-\theta K}.
\end{aligned}$$

Let the market consist of $N = 10\,000$ customers with average claim frequency $\alpha = 0.5$ and marginal cost of frictions $c = 100$, set $\rho = 5\%$, and assume a claim distribution parameter $\theta = 0.01$ (mean claim size 100). Let further $K = 20$, corresponding to 20% of average claim size. This yields $\overline{x_1} = 81.87$ and $\overline{x_2} = 16\,375$. For these values of the parameters, the drift-variance ratio to be optimised, $\kappa(p_1, p_2;\, \delta)$ from (3.13), is illustrated in Figure 3.4 as a function of premiums, p_1 and p_2, for $\delta = 15$. It appears that a saddle point could be present in the graph. The contour diagram in Figure 3.5 zooms in on the middle portion of the mesh in Figure 3.4, and the saddle point becomes readily apparent.

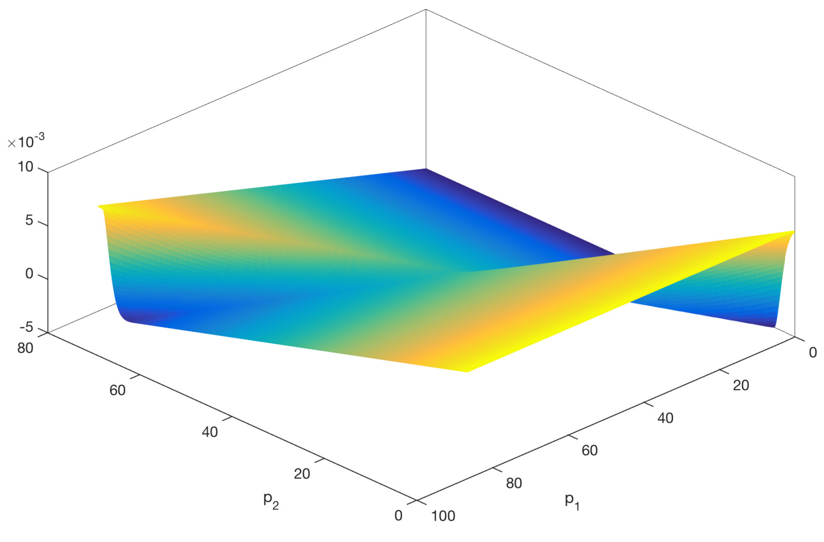


Figure 3.4: Graph of $\kappa(p_1, p_2;\, \delta)$.

Optimal premiums are calculated using Theorem C.11 as

$$p_1^* = 40.11 \quad \text{and} \quad p_2^* = 43.31.$$

Since I_2 has the market advantage, it can charge a higher premium. To see the Nash equilibrium property in Figures 3.4 and 5, note that given $p_1 = 40.11$, I_2 minimises (moves to cooler colors in the figure) by moving to $p_2 = 43.31$, and given $p_2 = 43.31$,

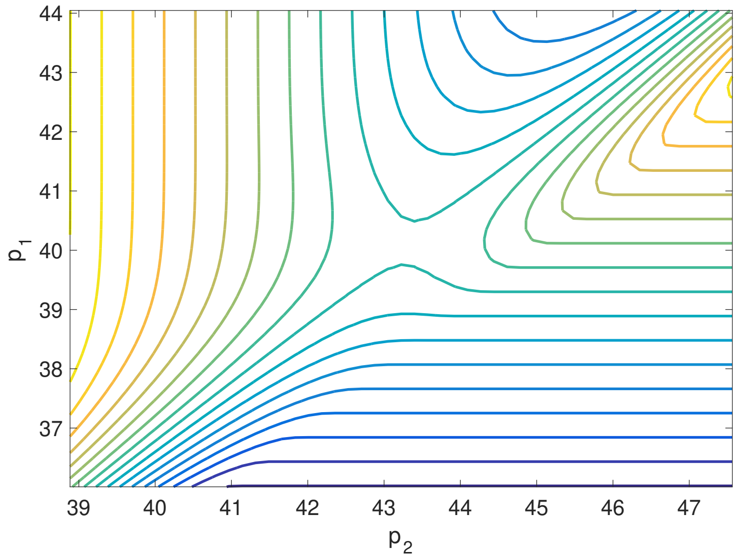


Figure 3.5: Contour diagram of $\kappa(p_1, p_2; \delta)$.

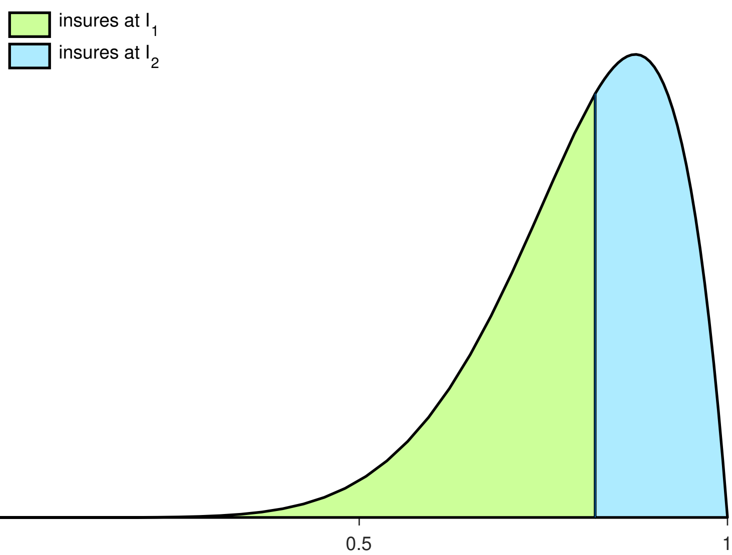


Figure 3.6: Distribution of customers in equilibrium.

I_1 maximises (moves to warmer colors) by moving to $p_1 = 40.11$. From Theorem C.11 it furthermore follows that in equilibrium $n_1(p_1^*, p_2^*) = n_2(p_1^*, p_2^*) = 5000$. The distribution of customers across insurance companies in equilibrium is shown in Figure 3.6, where the separation point from (3.6) is $v_0 = m_\beta = 0.82$.

C.6 Conclusion

We have considered a non-life insurance market in which two insurance companies compete for customers by choice of premium strategies. We model the customer's problem explicitly, paying special attention to market frictions, reflecting differences in cost of search and switching, transportation, information acquisition and processing, or preferences. Each company chooses its strategy to balance revenue per customer against portfolio size, taking into account the strategy of the other company. The reserves of the companies are modelled via the diffusion approximation to a standard Cramer-Lundberg risk process, extended to allow investment in a risk-free asset. The analysis is carried out in continuous time using stochastic differential game techniques. A Nash equilibrium is established for beta distributed market frictions (or customer preferences). A companion paper (Asmussen et al., 2019) of a somewhat different flavour discusses product differentiation via deductibles and Stackelberg games. Future research could consider three or more companies competing for market shares, account explicitly for the risk of ruin, or for the possibility that some potential customers choose not to insure.

Appendix

C.I Second order derivative test in Theorem C.11

Let $x = v_0$ from (3.6) and

$$f(p_1, p_2) = \frac{N}{2\rho c} \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} (p_1 + p_2 - 2\alpha\, \overline{x_1}).$$

Recalling that $\nu(p_1, p_2) = \mu_1(p_1, p_2) - \mu_2(p_1, p_2)$, the first order conditions of Theorem C.11 can then be abbreviated as

$$\nu_1'(p_1, p_2) = N \frac{B(x; a, b)}{B(a,b)} - f(p_1, p_2) = 0,$$

$$\nu_2'(p_1, p_2) = f(p_1, p_2) - N \left(1 - \frac{B(x; a, b)}{B(a,b)}\right) = 0.$$

The second order derivatives are

$$\nu_{11}''(p_1, p_2) = -\frac{N}{2\rho c} \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} - f_1'(p_1, p_2)\,,$$

$$\nu_{22}''(p_1, p_2) = f_2'(p_1, p_2) + \frac{N}{2\rho c} \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}.$$

By symmetry we have the relation

$$f_1'(p_1, p_2) + f_2'(p_1, p_2) = \frac{N}{\rho c} \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}\,,$$

so that

$$\nu_{22}''(p_1, p_2) = \frac{3N}{2\rho c} \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} - f_1'(p_1, p_2).$$

Thus, by the second order conditions from Proposition C.9,

$$\nu''_{11}(p_1, p_2) < 0, \qquad \nu''_{22}(p_1, p_2) > 0,$$

we must have

$$-\frac{N}{2\rho c}\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} < f'_1(p_1, p_2) < \frac{3N}{2\rho c}\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}.$$

Since

$$\begin{aligned} f'_1(p_1, p_2) = {} & \frac{N}{2\rho c}\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} \\ & \cdot \left(1 - \frac{1}{2\rho c}\left(\frac{a-1}{x} - \frac{b-1}{1-x}\right)(p_1 + p_2 - 2\alpha\,\overline{x_1})\right), \end{aligned}$$

the inequalities reduce to

$$-2 \leq -\frac{1}{2\rho c}\left(\frac{a-1}{x} - \frac{b-1}{1-x}\right)(p_1 + p_2 - 2\alpha\,\overline{x_1}) \leq 2.$$

Evaluated at the optimum, we finally get the parameter restrictions (3.23).

References

Asmussen, S., Christensen, B. J., and Taksar, M. (2013). "Portfolio size as function of the premium: modelling and optimization". *Stochastics* 85.4, pp. 575–588.

Asmussen, S., Christensen, B. J., and Thøgersen, J. (2019). "Stackelberg equilibrium premium strategies for push-pull competition in a non-life insurance market with product differentiation". *Risks* 7.2, p. 49.

Basak, S. and Makarov, D. (2014). "Strategic asset allocation in money management". *The Journal of Finance* 69.1, pp. 179–217.

Bäuerle, N. and Bayraktar, E. (2014). "A note on applications of stochastic ordering to control problems in insurance and finance". *Stochastics* 86.2, pp. 330–340.

Bensoussan, A., Siu, C. C., Yam, S. C. P., and Yang, H. (2014). "A class of non-zero-sum stochastic differential investment and reinsurance games". *Automatica* 50.8, pp. 2025–2037.

Boonen, T. J., Pantelous, A. A., and Wu, R. (2018). "Non-cooperative dynamic games for general insurance markets". *Insurance: Mathematics and Economics* 78, pp. 123–135.

Borch, K. H. (1962). "Application of game theory to some problems in automobile insurance". *ASTIN Bulletin* 2.2, pp. 208–221.

— (1974). *The mathematical theory of insurance: an annotated selection of papers on insurance published 1960-1972.* Lexington Books.

Brown, J. R. and Goolsbee, A. (2002). "Does the internet make markets more competitive? Evidence from the life insurance industry". *Journal of Political Economy* 110.3, pp. 481–507.

Browne, S. (2000). "Stochastic differential portfolio games". *Journal of Applied Probability* 37.1, pp. 126–147.

Bühlmann, H. (1980). "An economic premium principle". *ASTIN Bulletin* 11.1, pp. 52–60.

— (1984). "The general economic premium principle". *ASTIN Bulletin* 14.1, pp. 13–21.

Bühlmann, H. and Gisler, A. (2006). *A course in credibility theory and its applications*. Springer.

Chen, L. and Shen, Y. (2018). "On a new paradigm of optimal reinsurance: A stochastic Stackelberg differential game between an insurer and a reinsurer". *ASTIN Bulletin* 48.2, pp. 905–960.

Diamond, P. A. (1982). "Wage determination and efficiency in search equilibrium". *The Review of Economic Studies* 49.2, pp. 217–227.

Dutang, C., Albrecher, H., and Loisel, S. (2013). "Competition among non-life insurers under solvency constraints: A game-theoretic approach". *European Journal of Operational Research* 231.3, pp. 702–711.

Emms, P. (2007). "Dynamic pricing of general insurance in a competitive market". *ASTIN Bulletin* 37.01, pp. 1–34.

— (2011). "Pricing general insurance in a reactive and competitive market". *Journal of Computational and Applied Mathematics* 236.6, pp. 1314–1332.

— (2012). "Equilibrium pricing of general insurance policies". *North American Actuarial Journal* 16.3, pp. 323–349.

Emms, P. and Haberman, S. (2005). "Pricing general insurance using optimal control theory". *ASTIN Bulletin* 35.02, pp. 427–453.

Emms, P., Haberman, S., and Savoulli, I. (2007). "Optimal strategies for pricing general insurance". *Insurance: Mathematics and Economics* 40.1, pp. 15–34.

Espinosa, G.-E. and Touzi, N. (2015). "Optimal investment under relative performance concerns". *Mathematical Finance* 25.2, pp. 221–257.

Glynn, P. W. (1982). "On the Markov property of the GI/G/∞ Gaussian limit". *Advances in Applied Probability* 14.1, pp. 191–194.

Gordon, M. J. (1959). "Dividends, earnings, and stock prices". *The review of economics and statistics* 41, pp. 99–105.

Grandell, J. (1997). *Mixed poisson processes*. Vol. 77. CRC Press.

Guan, G. and Liang, Z. (2016). "A stochastic Nash equilibrium portfolio game between two DC pension funds". *Insurance: Mathematics and Economics* 70, pp. 237–244.

Hipp, C. and Taksar, M. (2010). "Optimal non-proportional reinsurance control". *Insurance: Mathematics and Economics* 47.2, pp. 246–254.

Hotelling, H. (1929). "Stability in competition". *The Economic Journal* 39.153, pp. 41–57.

Iglehart, D. L. (1969). "Diffusion approximations in collective risk theory". *Journal of Applied Probability* 6.2, pp. 285–292.

Jin, Z., Yin, G., and Wu, F. (2013). "Optimal reinsurance strategies in regime-switching jump diffusion models: Stochastic differential game formulation and numerical methods". *Insurance: Mathematics and Economics* 53.3, pp. 733–746.

Karatzas, I. and Shreve, S. (1998). *Brownian motion and stochastic calculus.* 2nd ed. Vol. 113. Graduate Texts in Mathematics. Springer.

Kerman, J. (2011). "A closed-form approximation for the median of the beta distribution". *arXiv: 1111.0433.*

Krylov, N. V. (1980). *Controlled diffusion processes.* Vol. 14. Applications of Mathematics. Springer.

Lemaire, J. (1984). "An application of game theory: cost allocation". *ASTIN Bulletin* 14.1, pp. 61–81.

— (1991). "Cooperative game theory and its insurance applications". *ASTIN Bulletin* 21.1, pp. 17–40.

Meng, H., Li, S., and Jin, Z. (2015). "A reinsurance game between two insurance companies with nonlinear risk processes". *Insurance: Mathematics and Economics* 62, pp. 91–97.

Mortensen, D. T. (1982). "Property rights and efficiency in mating, racing, and related games". *The American Economic Review* 72.5, pp. 968–979.

Mortensen, D. T. and Pissarides, C. A. (1994). "Job creation and job destruction in the theory of unemployment". *The Review of Economic Studies* 61.3, pp. 397–415.

Nash, J. (1951). "Non-cooperative games". *Annals of mathematics* 54.2, pp. 286–295.

Pantelous, A. A. and Passalidou, E. (2013). "Optimal premium pricing policy in a competitive insurance market environment". *Annals of Actuarial Science* 7.2, pp. 175–191.

— (2015). "Optimal premium pricing strategies for competitive general insurance markets". *Applied Mathematics and Computation* 259, pp. 858–874.

— (2017). "Optimal strategies for a non-linear premium-reserve model in a competitive insurance market". *Annals of Actuarial Science* 11.1, pp. 1–19.

Pestien, V. C. and Sudderth, W. D. (1985). "Continuous-time red and black: How to control a diffusion to a goal". *Mathematics of Operations Research* 10.4, pp. 599–611.

Polborn, M. K. (1998). "A model of an oligopoly in an insurance market". *The Geneva Papers on Risk and Insurance Theory* 23.1, pp. 41–48.

Powers, M. R., Shubik, M., and Yao, S. T. (1998). "Insurance market games: Scale effects and public policy". *Journal of Economics* 67.2, pp. 109–134.

Pun, C. S. and Wong, H. Y. (2016). "Robust non-zero-sum stochastic differential reinsurance game". *Insurance: Mathematics and Economics* 68, pp. 169–177.

Rees, R., Gravelle, H., and Wambach, A. (1999). "Regulation of insurance markets". *The Geneva Papers on Risk and Insurance Theory* 24.1, pp. 55–68.

Rogers, L. C. G. and Williams, D. (2000). *Diffusions, Markov processes and martingales: Volume 2, Itô calculus.* Second ed. Cambridge University Press.

Rothschild, M. and Stiglitz, J. (1976). "Equilibrium in competitive insurance markets: An essay on the economics of imperfect information". *The Quarterly Journal of Economics* 90.4, pp. 629–649.

Salop, S. and Stiglitz, J. (1977). "Bargains and ripoffs: A model of monopolistically competitive price dispersion". *The Review of Economic Studies* 44.3, pp. 493–510.

Schmidli, H. (2008). *Stochastic control in insurance.* Springer.

Taksar, M. and Zeng, X. (2011). "Optimal non-proportional reinsurance control and stochastic differential games". *Insurance: Mathematics and Economics* 48.1, pp. 64–71.

Taylor, G. C. (1986). "Underwriting strategy in a competitive insurance environment". *Insurance: Mathematics and Economics* 5.1, pp. 59–77.

Taylor, G. C. (1987). "Expenses and underwriting strategy in competition". *Insurance: mathematics and economics* 6.4, pp. 275–287.

Thøgersen, J. (2016). "Optimal premium as a function of the deductible: customer analysis and portfolio characteristics". *Risks* 4.4, p. 42.

von Neumann, J. (1928). "Zur Theorie der Gesellschaftsspiele". *Mathematische Annalen* 100.1, pp. 295–320.

Whitt, W. (1982). "On the heavy-traffic limit theorem for $GI/G/\infty$ queues". *Advances in Applied Probability* 14.1, pp. 171–190.

Wu, R. and Pantelous, A. A. (2017). "Potential games with aggregation in non-cooperative general insurance markets". *ASTIN Bulletin* 47.1, pp. 269–302.

Yan, M., Peng, F., and Zhang, S. (2017). "A reinsurance and investment game between two insurance companies with the different opinions about some extra information". *Insurance: Mathematics and Economics* 75, pp. 58–70.

Zeng, X. (2010). "A stochastic differential reinsurance game". *Journal of Applied Probability* 47.2, pp. 335–349.

Paper D

Stackelberg equilibrium premium strategies for push-pull competition in a non-life insurance market with product differentiation

Søren Asmussen, Bent Jesper Christensen
and Julie Thøgersen

ABSTRACT. Two insurance companies I_1, I_2 with reserves $R_{1,t}, R_{2,t}$ compete for customers, such that in a suitable differential game the smaller company I_2 with $R_{2,0} < R_{1,0}$ aims at minimising $R_{1,t} - R_{2,t}$ by using the premium p_2 as control and the larger I_1 at maximising by using p_1. Deductibles K_1, K_2 are fixed but may be different. If $K_1 > K_2$ and I_2 is the leader choosing its premium first, conditions for Stackelberg equilibrium are established. For gamma-distributed rates of claim arrivals, explicit equilibrium premiums are obtained, and shown to depend on the running reserve difference. The analysis is based on the diffusion approximation to a standard Cramér-Lundberg risk process extended to allow investment in a risk-free asset.

KEYWORDS: stochastic differential game, product differentiation, adverse selection, Stackelberg equilibrium

Publication details: Risks, 2019, vol. 7(2), 49. DOI: 10.3390/risks7020049

D.1 Introduction

Insurance premiums are typically calculated based on the expected loss, with an added loading depending on distributional properties of the risk (the expected value principle, variance principle, utility premium, etc.). An alternative to these static premium principles is to consider the premium as a dynamic control variable of the insurance company, as suggested in Asmussen et al. (2013) and Thøgersen (2016). In this approach, the individual customer's problem of deciding whether or not to insure at any given premium offered is modelled explicitly, and the premium is chosen optimally by the insurance company, balancing the resulting portfolio size against revenue per customer in order to minimise ruin probability. The analysis is based on the diffusion approximation to a standard Cramér-Lundberg risk process, extended to allow investment in a risk-free asset. In Asmussen et al. (2018), this idea is extended to a situation where insurance companies compete against each other, and Nash equilibria in premium controls of the resulting stochastic differential game are determined under suitable conditions. However, in some cases, no Nash equilibrium exists.

In the present paper, we present a parallel to this analysis dealing with product differentiation, with insurance companies offering different deductibles, and accounting for the possibility of Stackelberg equilibria. Two insurance companies compete against each other such that one company is the leader, choosing its premium first, and the other company is the follower, choosing its premium in response to the leader's. The setting is slightly modified relative to that in Asmussen et al. (2018), in that we do not consider search and switching costs when modelling the customer's choice between insurance products. Our main contributions are, first, to establish the existence of Stackelberg equilibrium under suitable conditions on this strategic game between insurance companies, and to identify the restrictions under which this reduces to the special case of Nash equilibrium. To our best knowledge, this adds at least the following new features to the literature on game theory in insurance: an example of Stackelberg equilibrium in premium controls; a finding of dependence of optimal premiums on reserves; and an occurrence of the phenomenon of adverse selection in a stochastic differential game between insurance companies, i.e., a lower premium charged increases portfolio size but leaves the average customer riskier to the company.

In the literature following Taylor (1986), the individual insurance company is frequently modelled as setting its premium in response to the aggregate insurance market, without explicitly considering the analogous behavior of the other companies constituting this market and the resulting strategic interactions. Examples include Taylor (1987) on marginal expense rates, Emms and Haberman (2005) generalising the deterministic discrete-time analysis of Taylor to a stochastic continuous-time model, Pantelous and Passalidou (2013, 2015) using stochastic demand functions in discrete time, and Emms (2007) and Emms et al. (2007), adopting stochastic processes for the market average premium and demand conditions in continuous time. Pantelous and Passalidou (2017) recently found the optimal premium to depend on the company's reserve in a competitive environment in the sense of this literature, but, again, this is not explicitly a game-theoretic equilibrium in the sense of Nash or Stackelberg, which is where we obtain dependence on reserves.

Game-theoretic aspects arise if the other insurers in the market in fact do react to the policy of the individual insurer, with the latter explicitly taking this into account in setting its policy. Market reaction to the individual insurer's premium is considered by Emms (2011). Explicit games between insurance companies have been studied using non-cooperative game theory, where Cournot games involve volume controls, see, e.g., Powers et al. (1998), whereas premium controls correspond to Bertrand games, e.g., the one-period games in Polborn (1998) and Dutang et al. (2013), who note that one aspect missing in their analysis is adverse selection among policyholders – our analysis includes this. Emms (2012) and Boonen et al. (2018) do consider continuous-time differential games in premium controls, but again based on Taylor (1986) type demand functions of own and market average premium. Boonen et al. (2018) in addition present a continuous-time extension of a one-period aggregate game of Wu and Pantelous (2017), involving a price elasticity of demand or market power parameter, and the individual insurer's payoff depending on own premium and an aggregate of market premiums. The models are deterministic and open-loop Nash equilibria are determined. In contrast, rather than assuming demand functions, we model the customer's choice of where to insure directly and find closed-loop or feedback Nash and Stackelberg equilibria in the resulting continuous-time strategic stochastic differential game between insurance companies. The roles of product differentiation via deductibles, adverse selection, and separating equilibrium in our solution are reminiscent of Rothschild and Stiglitz (1976), one of the first applications of game theory to competition in insurance premiums.

Besides competition in premiums, game theory has found several other applications in insurance, starting with Borch (1962) on risk transfer. Zeng (2010), Taksar and Zeng (2011), and Jin et al. (2013) consider Nash equilibria of stochastic differential games between insurance companies in reinsurance strategies. The analysis has been extended to non-zero sum games and additional investment controls by Bensoussan et al. (2014), nonlinear risk processes by Meng et al. (2015), ambiguity-aversion by Pun and Wong (2016), and insurance companies with different levels of trust in information by Yan et al. (2017). Stackelberg-type equilibria of stochastic differential games have been studied in Lin et al. (2012), where an insurance company selects an investment strategy while the market (or nature) selects a worst-case probability scenario, and in Chen and Shen (2018), where the game is between insurer and reinsurer, but not as here in a game between insurance companies. For some more remote references, see Asmussen et al. (2018). Stackelberg games were introduced by von Stackelberg (1934), and the theory of stochastic differential Stackelberg games is considered by Yong (2002), Bensoussan et al. (2015), and Shi et al. (2016).

Premium competition between insurance companies is likely to arise because the premium charged may affect both portfolio size and revenue per customer. Without market frictions or product differentiation, it might be expected that all customers would simply insure at the company offering the lowest premium. However, this may not be the case in the presence of market frictions. Thus, when choosing which insurance company to contact, customers may face different costs of search and switching, transportation, or information acquisition, or they may simply exhibit differences in preferences. Search frictions have been studied in economics by Diamond (1982), Mortensen (1982), Mortensen and Pissarides (1994), and others. Brown and Goolsbee (2002) studied the effect of internet search on life insurance premiums in US data. Information frictions have been modelled as differences in the cost of obtaining information, e.g., by Salop and Stiglitz (1977). In Asmussen et al. (2018) we study premium competition between insurance companies in the presence of market frictions. In the present paper, we consider instead product differentiation, and for simplicity abstract from market frictions. With the leading example of car insurance in mind, product differentiation may come in several forms. Here, we focus on different deductibles. Other possibilities would be bonus-malus systems, see Denuit et al. (2007), or proportional compensation in deductibles, similar to reinsurance arrangements, see Albrecher et al. (2017).

We consider the case of two insurance companies, referred to as I_1 and I_2. We allow for product differentiation by letting I_i offer an insurance contract with fixed deductible K_i for a premium p_i, $i = 1, 2$. The deductible measures the quality of the insurance product, so the company offering the lower deductible will be able to charge a higher premium.

We assume that there is a financial market consisting of a single risk-free asset with dynamics $\mathrm{d}B_t = rB_t\,\mathrm{d}t$, where r is the risk-free interest rate. All excess wealth of customers and reserve of insurers is invested in this asset. There are N customers in the insurance market. We assume that all customers must insure at either I_1 or I_2 and focus the analysis on the choice between the two companies. This involves several characteristics of both customer and insurance product. We pay special attention to product differentiation and customer risk.

The characteristics of an individual customer are unknown to the insurance companies, but their probability distribution known. Based on this distribution, the companies can determine the expected portfolio sizes $n_i(p_1, p_2)$ and average claim frequencies $\alpha_i(p_1, p_2)$ in their portfolios as functions of the premiums offered. The gross premium rate of I_i is then $c_i(p_1, p_2) = n_i(p_1, p_2)p_i$, and the aggregate claim frequency is $\lambda_i(p_1, p_2) = n_i(p_1, p_2)\alpha_i(p_1, p_2)$.

Let $r_{i,0}$ be the initial reserve of company i. For given premiums (p_1, p_2), the reserve of I_i is governed by the dynamics

$$\mathrm{d}R_{i,t} = (\mu_i(p_1, p_2) + rR_{i,t})\,\mathrm{d}t + \sigma_i(p_1, p_2)\,\mathrm{d}W_{i,t}, \tag{3.1}$$

where $(W_{1,t})_{t\geq 0}$ and $(W_{2,t})_{t\geq 0}$ are independent Wiener processes, and

$$\begin{aligned}
\mu_i(p_1, p_2) &= c_i(p_1, p_2) - \lambda_i(p_1, p_2)\mathbb{E}[(Z - K_i)^+] \\
&= n_i(p_1, p_2)\big(p_i - \alpha_i(p_1, p_2)\mathbb{E}[(Z - K_i)^+]\big)\,, \\
\sigma_i^2(p_1, p_2) &= \lambda_i(p_1, p_2)\mathbb{E}[(Z - Z_i)^{+2}] \\
&= n_i(p_1, p_2)\alpha_i(p_1, p_2)\mathbb{E}[(Z - K_i)^{+2}]\,.
\end{aligned}$$

The random variable Z represents claim sizes, assumed to be independent and identically distributed. Thus, (3.1) can be considered as a diffusion approximation to the Cramér-Lundberg process extended to the case where the insurance companies have access to investment in a risk-free asset. Such diffusion approximations have been used widely, based on the arguments of Iglehart (1969).

The aim is to derive value functions for the insurance companies, and determine game-theoretic equilibria. We consider here what we call push-pull competition. We assume that the largest company in terms of initial capital, I_1, selects its premium to try to *push* the small company away, while the small company tries to *pull* closer to the large company. For $K_1 > K_2$ and I_2 the leader choosing its premium first, we derive conditions for a Stackelberg equilibrium. The stronger feature of a Nash equilibrium may also occur, and we give conditions for that, but our numerical examples indicate that Stackelberg is the more typical case. Subsequently, for completeness, we briefly sketch the solution in the opposite case, $K_1 < K_2$. The claim frequencies of individual customers are considered random to the insurance company, and we obtain explicit solutions for equilibrium premiums in the case of gamma-distributed claim frequencies.

The structure of the paper is as follows. In Section D.2 we analyze the customer's problem. We proceed to portfolio characteristics in Section D.3. In Section D.4, we use the portfolio characteristics to find the strategies of I_1 and I_2. In Section D.5, we obtain explicit solutions in the case of gamma-distributed claim frequencies, and provide numerical examples. Section D.6 concludes. Some calculations and proofs are deferred to Appendix D.I.

D.2 Customer's problem

The customer has access to the risk-free asset paying interest at rate r. This is the customer's only source of income, and she invests all her wealth in this. The customer is exposed to a risk $(A_t^\alpha)_{t\geq 0}$, modelled as a compound Poisson process $A_t^\alpha = \sum_{n=1}^{N_t^\alpha} Z_n$, where $(N_t^\alpha)_{t\geq 0}$ is a Poisson process with claim frequency α, and $(Z_n)_{n\in\mathbb{N}}$ are the claim sizes, assumed to be independent of $(N_t^\alpha)_{t\geq 0}$. The customer will then reduce this risk by buying insurance. If the customer insures at I_i, then she will continuously pay the premium p_i, and in return have the claim sizes reduced to at most K_i. The wealth of the customer $(w_{i,t})_{t\geq 0}$ when insuring at company i thus has dynamics

$$\mathrm{d}w_{i,t} = (rw_{i,t} - p_i)\,\mathrm{d}t - \mathrm{d}A_{i,t}^\alpha, \quad w_{i,0} = w_0,$$

where $(A_{i,t}^\alpha)_{t\geq 0}$ is the compound Poisson process $A_{i,t}^\alpha = \sum_{n=1}^{N_t^\alpha} \min\{Z_n, K_i\}$, and w_0 the customer's initial wealth.

We here use similar evaluation criteria and subsequent arguments as in Thøgersen (2016), which we refer to for a more exhaustive treatment. The first step is to realise that the expected present discounted wealth when insuring at I_i can be evaluated as

$$V_i = \mathbb{E}\left[\int_0^\infty \exp(-dt)\,\mathrm{d}w_{i,t}\right] = \frac{rw_0 - p_i}{d-r} - \frac{\alpha}{d-r}\mathbb{E}[\min\{Z_n, K_i\}],$$

where $d > r$ is a subjective discount rate. If the customer were risk-neutral, she would simply choose the insurance company generating maximum expected present discounted wealth. Thus, she would prefer I_i over I_j if

$$p_i - p_j < -\alpha\big(\mathbb{E}[\min\{Z_n, K_i\}] - \mathbb{E}[\min\{Z_n, K_j\}]\big)\,. \tag{3.2}$$

However, an existence criterion for the insurance industry is that customers are risk averse, and this requires modification of (3.2). If $K_i \neq K_j$, the customer will be facing an excess claim size risk when insuring at the company with the higher deductible. Let this additional (or reduced) risk be denoted $z^e_{i,j} = \mathbb{E}[\min\{Z_n, K_i\} - \min\{Z_n, K_j\}]$ when insuring at I_i rather than I_j. Please note that $z^e_{i,j}$ corresponds to the last factor in (3.2), and is positive if $K_i > K_j$, and vice versa. Let β denote the risk aversion of the customer. By standard arguments of insurance, due to the risk aversion, the customer will be willing to pay a fee to avoid the additional risk. We will take this into account by introducing a personal safety loading $\omega(\beta)$ that the customer is willing to pay to avoid the excess risk present when $K_1 \neq K_2$. This is incorporated in (3.2) by multiplying the excess risk by $(1 + \omega(\beta))$. The more risk averse the customer, the higher the safety loading, i.e., ω is non-negative and increasing in β, with $\omega(0) = 0$. Thus, including risk aversion, the customer will prefer I_1 over I_2 if

$$p_1 - p_2 < -(1 + \omega(\beta))\alpha z^e_{1,2}\,, \tag{3.3}$$

and conversely, I_2 over I_1 if

$$p_2 - p_1 < +(1 + \omega(\beta))\alpha z^e_{1,2}\,. \tag{3.4}$$

In the next section, we use these relations to evaluate the portfolio sizes and average claim frequencies of the respective companies. We remark, however, at this place that in Asmussen et al. (2018) we have presented an in part more sophisticated approach to the customer's problem involving a finite decision horizon with varying interpretations, but for the sake of simplicity, we have not pursued this aspect here.

D.3 Portfolio characteristics

The claim frequencies α of the customers will be considered as random to the firm and denoted by A for a given customer. The case (3.3) then corresponds to the event

$$\Omega = \big\{p_1 - p_2 < -(1 + \omega(\beta))Az^e_{1,2}\big\} \tag{3.5}$$

and (3.4) to the complementary event Ω^{c}.

For I_1 the expected portfolio size $n_1(p_1, p_2)$ and the average claim frequency $\alpha_1(p_1, p_2)$ take the form

$$n_1(p_1, p_2) = N\mathbb{P}(\Omega), \quad \alpha_1(p_1, p_2) = \mathbb{E}[A \mid \Omega],$$

where N is the market size. Vice versa for I_2, where

$$n_2(p_1, p_2) = N\mathbb{P}(\Omega^{\mathrm{c}}), \quad \alpha_2(p_1, p_2) = \mathbb{E}[A \mid \Omega^{\mathrm{c}}].$$

Letting

$$y = \frac{p_2 - p_1}{(1+\omega(\beta))z^e_{1,2}}, \tag{3.6}$$

the probability of (3.5) can for $z^e_{1,2} > 0$ (corresponding to $K_1 > K_2$) be evaluated as

$$\mathbb{P}(\Omega) = \mathbb{P}(A < y),$$

so, the portfolio sizes are

$$n_1(p_1, p_2) = N\mathbb{P}(A < y), \quad n_2(p_1, p_2) = N(1 - \mathbb{P}(A < y)). \tag{3.7}$$

The average claim frequency for I_1 is the conditional expected value of the random claim frequency A given that the customer insures at I_1, i.e.,

$$\alpha_1(p_1, p_2) = \mathbb{E}[A \mid A < y], \tag{3.8}$$

and likewise, for I_2,

$$\alpha_2(p_1, p_2) = \mathbb{E}[A \mid A \geq y], \tag{3.9}$$

if $y > 0$. Otherwise, $\alpha_1(p_1, p_2) = 0$ and $\alpha_2(p_1, p_2) = \mathbb{E}[A]$ if $y < 0$. The criterion $y > 0$ for obtaining information from the customers' choices stems from the assumption $z^e_{1,2} > 0$, which indicates that I_2 offers a better product than I_1, and therefore the premium p_1 should not exceed p_2. Otherwise, every customer would obviously choose to insure at I_2.

In case $z^e_{1,2} < 0$, which means that I_1 offers a better insurance product, i.e., a lower deductible, $K_1 < K_2$, then by symmetry

$$\begin{aligned} n_1(p_1, p_2) &= N\mathbb{P}(A \geq y), \quad \alpha_1(p_1, p_2) = \mathbb{E}[A \mid A \geq y], \\ n_2(p_1, p_2) &= N\mathbb{P}(A < y), \quad \alpha_2(p_1, p_2) = \mathbb{E}[A \mid A < y], \end{aligned}$$

if $y > 0$. Otherwise, if $y < 0$, then I_1 would offer a lower premium for a better product, and would hence win the entire market of customers.

D.4 The strategies of the insurance companies

We now consider the optimization problems of the insurance companies. A control $\pi = (\pi_1, \pi_2)$ is a set $\big(\pi_1(t), \pi_2(t)\big)$ of premium strategies where $\pi_1(t), \pi_2(t)$ denote the premiums set by the companies at time t. As in much of stochastic control theory, we will only consider Markovian (also called feedback) strategies π, meaning that $\pi_1(t), \pi_2(t)$ only depend on the current value δ of the difference $\Delta^\pi(t) = R^\pi_1(t) - R^\pi_2(t)$ between the corresponding controlled reserve processes R^π_1, R^π_2. That is, we can write $\pi_1(t) = p^\pi_1(\Delta^\pi(t))$, $\pi_2(t) = p^\pi_2(\Delta^\pi(t))$ for suitable functions p^π_1, p^π_2. Since the (uncontrolled) reserves have the dynamics (3.1), this makes $\Delta^\pi(t)$ a diffusion process,

$$\mathrm{d}\Delta^\pi(t) = \mu^\pi(\Delta^\pi(t))\,\mathrm{d}t + \sigma^\pi(\Delta^\pi(t))\,\mathrm{d}W(t)\,, \tag{3.10}$$

where

$$\mu^\pi(\delta) = \mu_1\big(p_1^\pi(\delta), p_2^\pi(\delta)\big) - \mu_2\big(p_1^\pi(\delta), p_2^\pi(\delta)\big) + r\delta\,,$$
$$\sigma^\pi(\delta)^2 = \sigma_1\big(p_1^\pi(\delta), p_2^\pi(\delta)\big)^2 + \sigma_2\big(p_1^\pi(\delta), p_2^\pi(\delta)\big)^2\,,$$

and $W = (W_1 - W_2)/\sqrt{2}$ is again a Wiener process. Without loss of generality, we take $\Delta(0) = \Delta^\pi(0) = r_{1,0} - r_{2,0} > 0$, i.e., I_1 is the large company and I_2 the small. The large company seeks to maximize the reserve difference (to *push* the competitor further away), while the small company seeks to minimise the same (to *pull* closer to the competitor), each taking the current reserve difference as the state variable. The optimality criterion is to consider a fixed interval $[\ell_d, \ell_u]$ with $\ell_d < \Delta(0) < \ell_u$ and let

$$\tau(\pi) = \inf\big\{t > 0 :\ \Delta^\pi(t) \notin [\ell_d, \ell_u]\big\}\,, \quad V^\pi(\delta) = \mathbb{P}^\pi\big(\Delta^\pi(\tau(\pi)) = \ell_u \,\big|\, \Delta(0) = \delta\big)\,.$$

Then the large company I_1 chooses π_1 to maximise the probability $\mathrm{V}^\pi(\Delta(0))$ to exit at the upper boundary, and the small I_2 chooses π_2 to minimise $V^\pi(\Delta(0))$, or equivalently to maximise the probability $1 - \mathrm{V}^\pi(\Delta(0))$ to exit at the lower boundary.

Remark D.1. The feedback assumption implies that this is equivalent to maximising (minimising) $V^\pi(\delta)$ for all $\ell_d < \delta < \ell_u$.

Given that deductibles are different, one of the firms offers a product of higher quality (lower deductible) than the other. Therefore, the sequence of the game matters, and so a Stackelberg game is considered, where the companies compete sequentially. The sequence of the game is that at any time t

(1) The insurance company with the better product (i.e., lower deductible) is the leader and thus plays first.

(2) The insurance company with the lower quality product is the follower, and plays second, instantly after observing the leader's choice.

If I_2 (the smallest firm) is the leader and I_1 the follower (i.e., $K_1 > K_2$), then a Stackelberg equilibrium is defined as a strategy pair (π_1^*, π_2^*) satisfying

$$\pi_1^* = \widehat{\pi}_1(\pi_2^*) \quad \text{and} \quad V^{(\pi_1^*,\pi_2^*)} \le V^{(\widehat{\pi}_1(\pi_2),\pi_2)} \qquad \text{for all } \pi_2, \tag{3.11}$$

where $\widehat{\pi}_1(\pi_2) = \arg\sup_{\pi_1} V^{(\pi_1,\pi_2)}$. This case, $K_1 > K_2$, is relevant when the company offering the lower deductible is not able to attract sufficiently many high-risk customers (who need this extra protection) to become the largest company. We briefly discuss the opposite case below, in Remark D.12.

The Stackelberg equilibrium concept involves backward induction. First, the optimal response of the follower is determined as a reaction function. Next, the leader inserts the reaction function of the follower into its optimization problem, and solves for the best first move. As the game evolves in continuous time, the reserve difference changes. At each instant, each firm reconsiders its strategy, taking the running reserve difference as the state variable, and taking into account the future strategies of both companies, as long as the reserve difference remains in $[\ell_d, \ell_u]$. The criteria for a Stackelberg equilibrium are less strict than the ones for the more common Nash equilibrium, defined as a strategy pair (π_1^*, π_2^*) satisfying

$$V^{(\pi_1^*,\pi_2^*)} \geq V^{(\pi_1,\pi_2^*)} \text{ for all } \pi_1 \quad \text{and} \quad V^{(\pi_1^*,\pi_2^*)} \leq V^{(\pi_1^*,\pi_2)} \text{ for all } \pi_2, \tag{3.12}$$

i.e., neither firm has an incentive to deviate from its strategy unilaterally. We later specify the specific (second order) criteria for our solution for both types of equilibrium.

We next quote from Asmussen et al. (2018) some results that will allow replacing optimization problems in the space of functions p_1, p_2 by the more elementary problem of pointwise maximization/minimization of the real-valued ratio

$$\kappa^\pi(\delta) = \frac{\mu^\pi(\delta)}{\sigma^\pi(\delta)^2} \tag{3.13}$$

between the drift and variance of the reserve difference process in (3.10).

Lemma D.2. *Let $\mu(x), \sigma^2(x)$ be bounded and measurable functions on an interval (ℓ_d, ℓ_u) such that $\inf_{\ell_d<x<\ell_u} \sigma^2(x) > 0$ and let X, W be defined on a suitable probability space such that W is a standard Brownian motion and*

$$X(t) = \delta + \int_0^t \mu\big(X(s)\big)\,\mathrm{d}s + \int_0^t \sigma\big(X(s)\big)\,\mathrm{d}W(s) \tag{3.14}$$

for some $\delta \in (\ell_d, \ell_u)$. Define further $\kappa(x) = \mu(x)/\sigma^2(x)$,

$$s(y) = \exp\Big\{-2\int_{\ell_d}^y \kappa(z)\,\mathrm{d}z\Big\}, \quad S(\delta) = \int_{\ell_d}^\delta s(y)\,\mathrm{d}y$$

and $\tau = \inf\big\{t : X(t) \notin (\ell_d, \ell_u)\big\}$. Then:

(i) $\mathbb{P}\big(X(\tau) = \ell_u\big) = S(\delta)/S(\ell_u)$.

(ii) *For a given function κ on $[\ell_d, \ell_u]$ and a given $\delta \in [\ell_d, \ell_u]$, let $\varphi(\kappa)$ denote the r.h.s. in (i). Then $\kappa_0 \leq \kappa_1$ implies $\varphi(\kappa_0) \leq \varphi(\kappa_1)$.*

By slight abuse of notation, define

$$\kappa(p_1, p_2;\, \delta) = \frac{\mu_1(p_1,p_2) - \mu_2(p_1,p_2) + r\delta}{\sigma_1^2(p_1,p_2) + \sigma_2^2(p_1,p_2)}, \quad p_1, p_2 \geq 0,\ \ell_d \leq \delta \leq \ell_u\,. \tag{3.15}$$

To ease notation here and in the following subsections, we use the notation

$$\kappa_i'(p_1,p_2;\,\delta) = \frac{\partial}{\partial p_i}\kappa(p_1,p_2;\,\delta)\,, \quad \kappa_{ij}''(p_1,p_2;\,\delta) = \frac{\partial^2}{\partial p_i \partial p_j}\kappa(p_1,p_2;\,\delta)\,,$$

for partial derivatives, where the number of primes indicates the number of times the function is differentiated, and the subscript specifies with respect to which variable.[1]

[1]The standard notation avoids the primes and considers the subscript as sufficient, but we want to emphasise the differentiation here in order to avoid confusion with other notation in the paper, e.g., $\mu_1(p_1, p_2)$ and $\mu_2(p_1, p_2)$.

Now consider the resulting drift and variance of the reserves in (3.1), focusing on the case $K_1 > K_2$. Writing $\overline{x_{1,i}} = \mathbb{E}[(Z-K_i)^+]$ and $\overline{x_{2,i}} = \mathbb{E}[(Z-K_i)^2 \mathbb{1}_{\{Z \geq K_i\}}]$, it follows from (3.1) and Section D.3 that the drift and variance for the reserve of I_1 can be written as

$$\mu_1(p_1, p_2) = N\mathbb{P}(A < y)\big(p_1 - \mathbb{E}[A \mid A < y]\,\overline{x_{1,1}}\big),$$
$$\sigma_1^2(p_1, p_2) = N\mathbb{P}(A < y)\mathbb{E}[A \mid A < y]\,\overline{x_{2,1}},$$

and for I_2,

$$\mu_2(p_1, p_2) = N\mathbb{P}(A \geq y)\big(p_2 - \mathbb{E}[A \mid A \geq y]\,\overline{x_{1,2}}\big),$$
$$\sigma_2^2(p_1, p_2) = N\mathbb{P}(A \geq y)\mathbb{E}[A \mid A \geq y]\,\overline{x_{2,2}},$$

with y given by (3.6). These expressions show that the denominator $\sigma_1^2(p_1, p_2) + \sigma_2^2(p_1, p_2)$ in (3.15) depends on the controls p_1, p_2 because so does y and $K_1 \neq K_2$ implies $\overline{x_{2,1}} \neq \overline{x_{2,2}}$ (if $K_1 = K_2 = K$ then $\sigma_1^2(p_1, p_2) + \sigma_2^2(p_1, p_2)$ reduces to $N\mathbb{E}[A]\,\mathbb{E}[(Z-K)^+]$). Therefore, we need to optimise over the entire κ function (3.15) and not just the difference in drifts ν as in Asmussen et al. (2018).

From (3.6), by lowering the premium p_1, I_1 (with a high deductible in their product) can increase y and thereby portfolio size $n_1(p_1, p_2)$, for given p_2, but at the expense of simultaneously increasing average claim rate $\alpha_1(p_1, p_2)$, leaving the combined effect on the drift $\mu_1(p_1, p_2)$ in (3.1) of sign that may go either way in general. Thus, there is a tradeoff, reflecting the adverse selection problem, cf. Rothschild and Stiglitz (1976), i.e., lowering the premium brings more but riskier customers. In contrast, by lowering its premium p_2 for given p_1, I_2 (offering the lower deductible) can lower y and thereby simultaneously increase portfolio size $n_2(p_1, p_2)$ and reduce average claim rate $\alpha_2(p_1, p_2)$, but the combined effect on the drift of the reserve difference in (3.10) is nevertheless of ambiguous sign, and further modelling indeed required.

By (i) of Lemma D.2, $V^\pi(\delta)$ takes the form $S(\delta)/S(\ell_u)$, and combination of (ii) of the lemma and Remark D.1 allows characterising a Stackelberg equilibrium with I_2 as the leader and I_1 the follower. It shows that the optimization problem is local: We can just consider maximization or minimization of $\kappa(\cdot,\cdot;\delta)$ separately at each δ. This yields Proposition D.3 below, in which we find the explicit (local) conditions for a Stackelberg equilibrium in (3.11) in terms of the function $\kappa(\cdot,\cdot;\delta)$ from (3.15). For a solution to exist, the maximising company should be facing a (locally, at least) concave problem structure, and the minimising company a convex one. Existence cannot be guaranteed in general, but needs to be verified when considering a specific distribution of A, and hence a specific $\kappa(\cdot,\cdot;\delta)$. For the standard assumption of a gamma-distributed heterogeneity, we see in Section D.5 that an equilibrium does in fact exist and is unique. Although multiple solutions do not occur in this example, they cannot be excluded in general, so that the equilibrium may not be unique. The approach with backward induction should be the same, though giving a set of solutions. As multiple equilibria do not arise in the gamma case, we do not discuss them in more depth, except noting that uniqueness is guaranteed if the (local) concavity and convexity properties exploited in the following proposition extend globally.

For a fixed δ, write $d(p_2) = \widehat{p}_1(\delta|\pi_2)$ for the optimal premium for I_1 given I_2 follows a strategy with premium p_2 at level δ.

Proposition D.3. *In a Stackelberg equilibrium (π_1^*, π_2^*), the optimal set $p_1^* = p_1^*(\delta)$, $p_2^* = p_2^*(\delta)$ of premiums at level δ is a solution to*

$$p_1^* = d(p_2^*), \quad p_2^* = \arg\min_{p_2} \kappa(d(p_2), p_2;\, \delta), \qquad \textit{where } d(p_2) = \arg\sup_{p_1}\ \kappa(p_1, p_2;\, \delta). \tag{3.16}$$

The first order conditions for (p_1^*, p_2^*) are

$$0 = \kappa_1'(p_1^*, p_2^*;\, \delta) = \kappa_2'(p_1^*, p_2^*;\, \delta)\,, \tag{3.17}$$

and the second order conditions are

$$0 > \kappa_{11}''(p_1^*, p_2^*;\, \delta)\,, \tag{3.18}$$
$$0 > \det\big(H(p_1^*, p_2^*;\, \delta)\big), \tag{3.19}$$

where $H(p_1, p_2;\, \delta) = \big(\kappa_{ij}''(p_1, p_2;\, \delta)\big)_{i,j=1,2}$ is the Hessian of $\kappa(\cdot, \cdot;\, \delta)$.

Proof. Condition (3.16) follows from the definition of Stackelberg equilibrium, see (3.11), and the local character of the problem discussed above. Choosing the best p_1 given p_2 means that I_1 takes p_1 as $d(p_2)$, so $d(p_2)$ satisfies

$$0 = \kappa_1'(d(p_2), p_2;\, \delta), \tag{3.20}$$
$$0 > \kappa_{11}''(d(p_2), p_2;\, \delta). \tag{3.21}$$

The problem I_2 is facing is then to minimise $\kappa(d(p_2), p_2;\, \delta)$ so p_2^* is the zero of the function

$$g(p_2) = \kappa_1'(d(p_2), p_2;\, \delta)d'(p_2) + \kappa_2'(d(p_2), p_2;\, \delta).$$

At a Stackelberg equilibrium we have $p_1^* = d(p_2^*)$. We therefore get

$$0 = \kappa_1'(p_1^*, p_2^*;\, \delta)d'(p_2^*) + \kappa_2'(p_1^*, p_2^*;\, \delta), \tag{3.22}$$

and the second order condition $g'(p_2^*) > 0$ means

$$\begin{aligned} 0 < {} & \big[\kappa_{11}''(p_1^*, p_2^*; \delta)d'(p_2^*) + \kappa_{12}''(p_1^*, p_2^*;\, \delta)\big]d'(p_2^*) \\ & + \kappa_1'(p_1^*, p_2^*; \delta)d''(p_2^*) + \kappa_{21}''(p_1^*, p_2^*;\, \delta)d'(p_2^*) + \kappa_{22}''(p_1^*, p_2^*;\, \delta). \end{aligned} \tag{3.23}$$

Now (3.20) implies that the first term in (3.22) vanishes, and using (3.20) again, we arrive at (3.17). Furthermore, differentiating (3.20) gives

$$0 = \kappa_{11}''(d(p_2), p_2;\, \delta)d'(p_2) + \kappa_{12}''(d(p_2), p_2;\, \delta), \tag{3.24}$$

and thus $[\,\cdot\,]d'(p_2^*)$ in (3.23) vanishes. So does the second term, by (3.20), and hence (3.23) reduces to

$$\begin{aligned} 0 &< \kappa''_{21}(p_1^*, p_2^*;\, \delta) d'(p_2^*) + \kappa''_{22}(p_1^*, p_2^*;\, \delta) \\ &= -\kappa''_{21}(p_1^*, p_2^*;\, \delta) \frac{\kappa''_{12}(p_1^*, p_2^*;\, \delta)}{\kappa''_{11}(p_1^*, p_2^*;\, \delta)} + \kappa''_{22}(p_1^*, p_2^*;\, \delta) = \frac{\det\big(H(p_1^*, p_2^*;\, \delta)\big)}{\kappa''_{11}(p_1^*, p_2^*;\, \delta)}, \end{aligned}$$

where the first equality follows from (3.24). Combination with (3.21) produces (3.19). ☐

Corollary D.4. *If, in addition to* (3.18)*, the premiums in* (3.16) *satisfy*

$$0 < \kappa''_{22}(p_1^*, p_2^*\,;\, \delta), \tag{3.25}$$

then (p_1^*, p_2^*) *furthermore meets the conditions of a Nash equilibrium.*

Proof. Follows from Asmussen et al. (2018). ☐

It is clear from Proposition D.3 and Corollary D.4 that the Stackelberg equilibrium concept is more general than Nash equilibrium. In particular, by (3.18) and (3.25), the diagonal entries of the relevant Hessian are of opposite sign in Nash equilibrium, so (3.19) is automatic. Furthermore, since (3.18) and (3.19) are the general conditions for a (local) saddlepoint of $\kappa(p_1, p_2;\, \delta)$, any saddlepoint of this function gives rise to a (local) Stackelberg equilibrium. Geometrically, such a saddlepoint need not be parallel to the axes corresponding to the controls (premiums). In case the cross-partial $\kappa''_{12}(p_1^*, p_2^*;\, \delta) = 0$ (equivalently, the policy of I_1 does not depend on that of I_2 at the optimum), then the saddlepoint is parallel to the axes and, indeed, gives rise to a (local) Nash equilibrium. Again, the conditions are only necessary, whereas sufficient conditions would involve global concavity/convexity.

Heuristically, because the premium controls of the companies are equally powerful and act in opposite directions, they should split customers evenly. This is formalised in the next proposition.

Proposition D.5. *In Stackelberg equilibrium,* I_1 *and* I_2 *share the market equally, i.e.,* $n_1(p_1^*, p_2^*) = n_2(p_1^*, p_2^*) = N/2$.

Proof. Suppressing δ for notational convenience, let $\kappa_n(p_1, p_2)$ and $\kappa_d(p_1, p_2)$ denote the numerator and denominator, respectively, of $\kappa(p_1, p_2\,;\delta)$ in (3.15). Simple calculations show that the partial derivatives satisfy the relations

$$\begin{aligned} \kappa'_{n,2}(p_1, p_2) &= -\kappa'_{n,1}(p_1, p_2) + N(\mathbb{P}(A < y) - \mathbb{P}(A \geq y))\,, \\ \kappa'_{d,2}(p_1, p_2) &= -\kappa'_{d,1}\,, \end{aligned} \tag{3.26}$$

with y from (3.6). Following Proposition D.3, we find the first order condition for I_1,

$$\kappa'_1(p_1, p_2;\, \delta) = \frac{1}{\kappa_d(p_1, p_2)} \kappa'_{n,1}(p_1, p_2) - \frac{\kappa_n(p_1, p_2)}{\kappa_d(p_1, p_2)^2} \kappa'_{d,1}(p_1, p_2) = 0,$$

which can be reduced to

$$\kappa'_{n,1}(p_1,p_2) - \frac{\kappa_n(p_1,p_2)}{\kappa_d(p_1,p_2)}\kappa'_{d,1}(p_1,p_2) = 0. \tag{3.27}$$

From this equation, the optimal response function $d(p_2)$ is deduced. Similarly, the first order condition for I_2 can be reduced to

$$\kappa'_{n,2}(d(p_2),p_2) - \frac{\kappa_n(d(p_2),p_2)}{\kappa_d(d(p_2),p_2)}\kappa'_{d,2}(d(p_2),p_2) = 0.$$

Using the relation (3.26) between the partial derivatives yields

$$-\kappa'_{n,1}(d(p_2),p_2) + N(\mathbb{P}(A<y) - \mathbb{P}(A \geq y)) + \frac{\kappa_n(d(p_2),p_2)}{\kappa_d(d(p_2),p_2)}\kappa'_{d,1}(d(p_2),p_2) = 0,$$

which in combination with (3.27) yields $\mathbb{P}(A<y) = \mathbb{P}(A \geq y)$. Hence, in Stackelberg equilibrium y should be the median of A, and from (3.7), $n_1(p_1^*,p_2^*) = n_2(p_1^*,p_2^*) = N/2$. □

D.5 Gamma-distributed claim frequencies

For modelling purposes, we assume that the claim frequencies are distributed according to $A \sim \text{gamma}(a,b)$, with c.d.f. $\mathbb{P}(A<y) = \gamma(b, y/a)/\Gamma(b)$ where

$$\Gamma(b) = \int_0^\infty t^{b-1}\exp(-t)\,\mathrm{d}t\,,$$

$$\gamma(b,z) = \int_0^z t^{b-1}\exp(-t)\,\mathrm{d}t, \qquad \Gamma(b,z) = \Gamma(b) - \gamma(b,z)$$

are the Gamma function, and the lower resp. upper incomplete Gamma function. The gamma distribution is standard for modelling unobserved heterogeneity in a Poissonian setting (in insurance, a classical case is credibility theory, see Bühlmann and Gisler (2006); in general Bayesian modelling, the gamma has the role of a conjugate prior greatly facilitating calculations, see Robert (2007)). However, the outline calculations can easily be paralleled for other distributions, though the amount of analytic details may be considerable.

The portfolio characteristics (3.7)–(3.9) can then be written explicitly as

$$\begin{aligned} n_1(p_1,p_2) &= N\frac{\gamma(b,y/a)}{\Gamma(b)}, \qquad & \alpha_1(p_1,p_2) &= \frac{a\gamma(b+1,y/a)}{\gamma(b,y/a)}, \\ n_2(p_1,p_2) &= N\frac{\Gamma(b,y/a)}{\Gamma(b)}, \qquad & \alpha_2(p_1,p_2) &= \frac{a\Gamma(b+1,y/a)}{\Gamma(b,y/a)}, \end{aligned} \tag{3.28}$$

if $z^e_{1,2} > 0$ and $y > 0$ which, as explained in Section D.3, is equivalent to $K_1 > K_2$ and $p_1 < p_2$.

Theorem D.6. *Assume $K_1 > K_2$ that $A \sim \mathrm{gamma}(a,b)$, and let m_Γ denote the median of* $\mathrm{gamma}(a,b)$. *Then a Stackelberg equilibrium exists at*

$$\begin{aligned} p_1^* &= d(p_2^*) = a\Big(\frac{p_2^*}{a} - (1+\omega(\beta))z_{1,2}^e\frac{m_\Gamma}{a}\Big), \\ p_2^* &= \frac{a}{2}\Big(\frac{m_\Gamma}{a}\Big)\Big(\frac{1}{2}\mathrm{e}^{m_\Gamma/a}\Big(\frac{m_\Gamma}{a}\Big)^{-b}\Gamma(b)(1+\omega(\beta))z_{1,2}^e \\ &\quad + (1+\omega(\beta))z_{1,2}^e + (\overline{x_{1,1}} + \overline{x_{1,2}}) - \tilde\kappa(\overline{x_{2,2}} - \overline{x_{2,1}})\Big) \end{aligned} \tag{3.29}$$

where

$$\begin{aligned} \tilde\kappa &= \kappa(p_1^*, p_2^*;\, \delta) \\ &= \frac{1}{\frac{1}{2}b\Gamma(b)(\overline{x_{2,1}} + \overline{x_{2,2}}) + \mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^b(\overline{x_{2,2}} - \overline{x_{2,1}})} \\ &\quad \cdot\Big(\mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^b(\overline{x_{1,1}} + \overline{x_{1,2}}) + r\delta\Gamma(b)/(Na) \\ &\quad + \frac{1}{2}\Gamma(b)(b(\overline{x_{1,2}} - \overline{x_{1,1}}) - (m_\Gamma/a)(1+\omega(\beta))z_{1,2}^e)\Big), \end{aligned} \tag{3.30}$$

provided

$$p_1^* \ge 0, \quad p_2^* \ge 0, \tag{3.31}$$

and

$$\mathrm{D}(a,b,K_1,K_2,r,\delta,\omega(\beta)) < 0 \tag{3.32}$$

with

$$\begin{aligned} \mathrm{D}(a,b,K_1,K_2,r,\delta,\omega(\beta)) &= \kappa(p_1^*,p_2^*;\,\delta)(\overline{x_{2,2}} - \overline{x_{2,1}}) - 2(1+\omega(\beta))z_{1,2}^e \\ - (\overline{x_{1,1}} + \overline{x_{1,2}}) &- \tfrac{1}{2}\mathrm{e}^{m_\Gamma/a}(m_\Gamma/a)^{-b}\Gamma(b)(1+\omega(\beta))z_{1,2}^e(m_\Gamma/a - b + 1). \end{aligned} \tag{3.33}$$

Remark D.7. As we discussed in Asmussen et al. (2018), there are arguments that motivate to remove condition (3.31) of non-negative premiums or to tighten it to premiums never below net levels $\alpha_i(p_1,p_2)\,\overline{x_{1,i}}$. However, since $\alpha_i(p_1,p_2)$ now depends on y from (3.6) and hence on premiums unlike in the Nash equilibrium occurring there, this route leads to an implicit condition and is not pursued further here. See; however, the discussion following (3.36) below.

Proof. Using the portfolio characteristics (3.28) and the notation from the proof of Proposition D.5, we can write the numerator of the criterion to be optimised (3.15) as

$$\begin{aligned} \kappa_n(p_1,p_2) &= \mu_1(p_1,p_2) - \mu_2(p_1,p_2) + r\delta \\ &= N\frac{\gamma(b,y/a)\,p_1 - a\gamma(b+1,y/a)\,\overline{x_{1,1}} - \Gamma(b,y/a)\,p_2 + a\Gamma(b+1,y/a)\,\overline{x_{1,2}}}{\Gamma(b)} + r\delta \\ &= N\frac{\gamma(b,y/a)}{\Gamma(b)}(p_1 - ab\,\overline{x_{1,1}}) - N\frac{\Gamma(b,y/a)}{\Gamma(b)}(p_2 - ab\,\overline{x_{1,2}}) \\ &\qquad + N\frac{a(y/a)^b\mathrm{e}^{-y/a}}{\Gamma(b)}(\overline{x_{1,1}} + \overline{x_{1,2}}) + r\delta\,, \end{aligned}$$

using the relations $\gamma(b+1,z) = b\gamma(b,z) - z^b \mathrm{e}^{-z}$ and $\Gamma(b+1,z) = b\Gamma(b,z) + z^b \mathrm{e}^{-z}$. Similarly, for the denominator,

$$\begin{aligned}
\kappa_d(p_1,p_2) &= \sigma_1^2(p_1,p_2) + \sigma_2^2(p_1,p_2) \\
&= \frac{N}{\Gamma(b)}\big(a\gamma\left(b+1, y/a\right)\,\overline{x_{2,1}} + a\Gamma\left(b+1, y/a\right)\,\overline{x_{2,2}}\big) \\
&= \frac{N}{\Gamma(b)}\big(ab\gamma(b, y/a)\,\overline{x_{2,1}} + ab\Gamma(b, y/a)\,\overline{x_{2,2}} \\
&\qquad + a(y/a)^b \mathrm{e}^{-y/a}\big(\,\overline{x_{2,2}} - \,\overline{x_{2,1}}\big)\big).
\end{aligned}$$

The derivatives of the incomplete Gamma functions are

$$\frac{\partial\gamma\left(b,z\right)}{\partial z} = z^{b-1}\mathrm{e}^{-z} = -\frac{\partial\Gamma\left(b,z\right)}{\partial z},$$

and by the definition (3.6) of y, we have $\frac{\partial y}{\partial p_2} = -\frac{\partial y}{\partial p_1} = 1/((1+\omega(\beta))z^e_{1,2})$. Hence, κ_n and κ_d have partial derivatives

$$\begin{aligned}
\kappa'_{n,1}(p_1,p_2) &= \frac{N}{\Gamma(b)}\Bigg(\gamma\left(b, y/a\right) + \frac{\overline{x_{1,1}} + \,\overline{x_{1,2}}}{(1+\omega(\beta))z^e_{1,2}}\mathrm{e}^{-y/a}\left(y/a\right)^b \\
&\qquad - \frac{p_1+p_2}{a(1+\omega(\beta))z^e_{1,2}}\left(y/a\right)^{b-1}\mathrm{e}^{-y/a}\Bigg), \\
\kappa'_{n,2}(p_1,p_2) &= \frac{N}{\Gamma(b)}\Bigg(-\Gamma\left(b, y/a\right) - \frac{\overline{x_{1,1}} + \,\overline{x_{1,2}}}{(1+\omega(\beta))z^e_{1,2}}\mathrm{e}^{-y/a}\left(y/a\right)^b \\
&\qquad + \frac{p_1+p_2}{a(1+\omega(\beta))z^e_{1,2}}\left(y/a\right)^{b-1}\mathrm{e}^{-y/a}\Bigg) \\
&= \frac{N}{\Gamma(b)}(\gamma\left(b, y/a\right) - \Gamma\left(b, y/a\right)) - \kappa'_{n,1}(p_1,p_2), \\
\kappa'_{d,1}(p_1,p_2) &= \frac{N}{\Gamma(b)}\left(\frac{\overline{x_{2,2}} - \,\overline{x_{2,1}}}{(1+\omega(\beta))z^e_{1,2}}\mathrm{e}^{-y/a}\left(y/a\right)^b\right), \\
\kappa'_{d,2}(p_1,p_2) &= -\frac{N}{\Gamma(b)}\left(\frac{\overline{x_{2,2}} - \,\overline{x_{2,1}}}{(1+\omega(\beta))z^e_{1,2}}\mathrm{e}^{-y/a}\left(y/a\right)^b\right) = -\kappa'_{d,1}\left(p_1,p_2\right),
\end{aligned}$$

confirming (3.26) in this case. By Proposition D.5, y must be the median of the gamma distribution, namely, the value m_Γ that solves

$$\gamma(b, m_\Gamma/a)/\Gamma(b) = \Gamma(b, m_\Gamma/a)/\Gamma(b) = 1/2.$$

From the definition (3.6) of y it then follows that p_2^* is chosen to satisfy

$$m_\Gamma = \frac{p_2^* - d(p_2^*)}{(1+\omega(\beta))z^e_{1,2}}.$$

Thus, when evaluated at p_2^*, the optimal response by I_1 is

$$d(p_2^*) = p_2^* - (1+\omega(\beta))z^e_{1,2}m_\Gamma.$$

We may now evaluate the expressions

$$
\begin{aligned}
\kappa(d(p_2^*), p_2^*;\, \delta) &= \frac{1}{\frac{1}{2} b\Gamma(b)(\overline{x_{2,1}} + \overline{x_{2,2}}) + \mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^b(\overline{x_{2,2}} - \overline{x_{2,1}})} \\
&\quad \cdot \Big(\mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^b(\overline{x_{1,1}} + \overline{x_{1,2}}) + r\delta\Gamma(b)/(Na) \\
&\qquad + \tfrac{1}{2}\Gamma(b)\big(b(\overline{x_{1,2}} - \overline{x_{1,1}}) - (m_\Gamma/a)(1+\omega(\beta))z_{1,2}^e\big)\Big) \\
&= \tilde{\kappa}, \\
\kappa'_{n,1}(d(p_2^*), p_2^*) &= \frac{N}{\Gamma(b)}\Big(\frac{-2p_2^*}{a(1+\omega(\beta))z_{1,2}^e}(m_\Gamma/a)^{b-1}\mathrm{e}^{-m_\Gamma/a} + \tfrac{1}{2}\Gamma(b) \\
&\quad + \frac{\overline{x_{1,1}} + \overline{x_{1,2}}}{(1+\omega(\beta))z_{1,2}^e}\mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^b + (m_\Gamma/a)^b\mathrm{e}^{-m_\Gamma/a}\Big), \\
\kappa'_{d,1}(d(p_2^*), p_2^*) &= \frac{N}{\Gamma(b)}\Big(\frac{\overline{x_{2,2}} - \overline{x_{2,1}}}{(1+\omega(\beta))z_{1,2}^e}\mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^b\Big).
\end{aligned}
$$

Substitute these into (3.27) and solve for p_2^* to

$$
\begin{aligned}
p_2^* = \frac{a}{2}\Big(\frac{m_\Gamma}{a}\Big)\Big(&\tfrac{1}{2}\mathrm{e}^{m_\Gamma/a}\Big(\frac{m_\Gamma}{a}\Big)^{-b}\Gamma(b)(1+\omega(\beta))z_{1,2}^e \\
&+ (1+\omega(\beta))z_{1,2}^e + (\overline{x_{1,1}} + \overline{x_{1,2}}) - \tilde{\kappa}(\overline{x_{2,2}} - \overline{x_{2,1}})\Big).
\end{aligned}
$$

The second order conditions, (3.18) and (3.19), are verified in Appendix D.I such that the first order conditions yield the types of optima desired.

We have without loss of generality assumed that $y \geq 0$, i.e., by (3.6), I_2 charges the highest premium. This is reasonable because it offers the best product ($K_1 > K_2$). Indeed, there cannot be an equilibrium in the region $y < 0$, since here, the criterion to be optimised would be $\kappa(p_1, p_2;\, \delta) = (r\delta - N(p_2 - \alpha\,\overline{x_{1,2}}))/(N\alpha\,\overline{x_{2,2}})$, which is decreasing in p_2. Since I_2 seeks to minimise, it would increase p_2 until again $y > 0$. □

Corollary D.8. *If, in addition to (3.32), the premiums in (3.29) satisfy*

$$
\mathrm{D}(a, b, K_1, K_2, r, \delta, \omega(\beta)) > -4(1+\omega(\beta))z_{1,2}^e \tag{3.34}
$$

then (p_1^*, p_2^*) furthermore meets the conditions of a Nash equilibrium.

Proof. Follows from Corollary D.4 and calculations in the Appendix. □

Remark D.9. The median is not analytically available, but can be solved for numerically. Banneheka and Ekanayake (2009) argue that the median for $b \geq 1$ can be approximated as $m_\Gamma \approx ab(3b - 0.8)/(3b + 0.2)$. Further to this, note that by scaling properties of the gamma distribution, m_Γ/a is the median of a gamma$(1, b)$ distribution. Evidently, equilibrium premiums scale in proportion to a.

Remark D.10. If $b = 1$, the gamma distribution reduces to the exponential with parameter $1/a$ and median $m_e = a\log(2)$. In this case, the expressions for equilibrium premiums simplify to

$$\begin{aligned}
p_1^* &= d(p_2^*) = a(p_2^*/a - (1+\omega(\beta))z_{1,2}^e m_e/a),\\
p_2^* &= \frac{a}{2}\Big(\frac{m_e}{a}\Big)\Big((m_e/a)^{-1}(1+\omega(\beta))z_{1,2}^e + (1+\omega(\beta))z_{1,2}^e.\\
&\qquad + (\overline{x_{1,1}} + \overline{x_{1,2}}) - \tilde{\kappa}(\overline{x_{2,2}} - \overline{x_{2,1}})\Big)
\end{aligned}$$

where

$$\tilde{\kappa} = \frac{(m_e/a)(\overline{x_{1,1}} + \overline{x_{1,2}}) + 2r\delta/(aN) + \overline{x_{1,2}} - \overline{x_{1,1}} - (m_e/a)(1+\omega(\beta))z_{1,2}^e}{\overline{x_{2,1}} + \overline{x_{2,2}} + (m_e/a)(\overline{x_{2,2}} - \overline{x_{2,1}})}.$$

Remark D.11. Without product differentiation, $K_1 = K_2$, premiums coincide, $p_1^* = p_2^*$. With product differentiation, the difference between equilibrium premiums is increasing in excess risk $z_{1,2}^e$ and safety loading $\omega(\beta)$.

Remark D.12. In case $K_1 < K_2$, i.e., the large firm I_1 offers the highest-quality insurance product (lowest deductible), then for a gamma-distributed claim frequency, the criterion to be optimised is by symmetry instead

$$\begin{aligned}
\kappa(p_1, p_2;\, \delta) = {}& \frac{1}{a\Gamma\left(b+1, y/a\right)\,\overline{x_{2,1}} + a\gamma\left(b+1, y/a\right)\,\overline{x_{1,2}}}\\
&\cdot\Big(\Gamma\left(b, y/a\right)p_1 - a\Gamma\left(b+1, y/a\right)\,\overline{x_{1,1}} + r\delta\Gamma(b)/N.\\
&\quad - \gamma\left(b, y/a\right)p_2 + a\gamma\left(b+1, y/a\right)\,\overline{x_{1,2}}\Big).
\end{aligned}$$

In this case, I_1 will be the leader of the Stackelberg game, and I_2 the follower. Recall here that because $K_1 < K_2$ we have $z_{1,2}^e < 0$. The same approach as in the proof of Theorem D.6 then yields the equilibrium

$$\begin{aligned}
p_2^* &= d(p_1^*) = a(p_1^*/a + (1+\omega(\beta))z_{1,2}^e(m_\Gamma/a)),\\
p_1^* &= \frac{a}{2}\Big(\frac{m_\Gamma}{a}\Big)\Big(-\tfrac{1}{2}\mathrm{e}^{m_\Gamma/a}(m_\Gamma/a)^{-b}\Gamma(b)(1+\omega(\beta))z_{1,2}^e\\
&\quad - (1+\omega(\beta))z_{1,2}^e + (\overline{x_{1,1}} + \overline{x_{1,2}}) - \tilde{\kappa}(\overline{x_{2,2}} - \overline{x_{2,1}})\Big)
\end{aligned}$$

where

$$\begin{aligned}
\tilde{\kappa} = {}& \frac{1}{\tfrac{1}{2}b\Gamma(b)(\overline{x_{2,1}} + \overline{x_{2,2}}) - \mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^b(\overline{x_{2,2}} - \overline{x_{2,1}})}\\
&\cdot\Big(-\mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^b(\overline{x_{1,1}} + \overline{x_{1,2}}) + r\delta\Gamma(b)/(aN)\\
&\quad + \tfrac{1}{2}\Gamma(b)\big(b(\overline{x_{1,2}} - \overline{x_{1,1}}) - (m_\Gamma/a)(1+\omega(\beta))z_{1,2}^e\big)\Big).
\end{aligned}$$

We have again that $n_i(p_1^*, p_2^*) = N/2$, $i = 1, 2$ (this follows as in the proof of Proposition D.5 and as in that case does not depend on the assumption of gamma-distributed heterogeneity). The case $K_1 < K_2$ is relevant if the company offering best protection (lowest deductible) and therefore charging highest premiums is able to more than cover the extra cost associated with the high-risk customers willing to pay such higher premiums, and thus become the largest company.

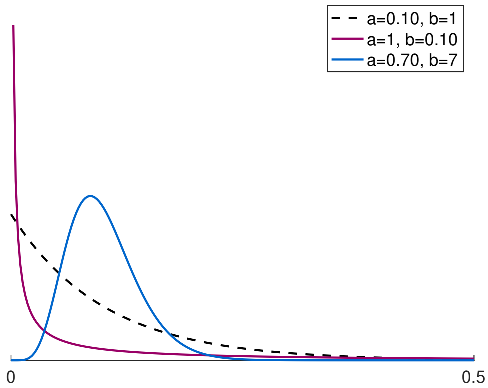


Figure 3.1: The gamma distribution of the claim frequencies for different a and b.

Returning to Theorem D.6 and the discussion of how the Stackelberg equilibrium evolves over time, note that due to interest rates, the strategy of the leader (here, I_2, with strategy p_2^*) changes in an affine fashion with δ. Although δ indicates the difference in initial reserves, the companies may reoptimise at any point in time. The game is repeated every instant, and each new equilibrium in the feedback version of the game takes the same form, with premiums set as in Theorem D.6, and δ the running difference in reserves. As functions of δ, the Stackelberg equilibrium premiums, $p_2^* = p_2^*(\delta)$ and $p_1^* = p_1^*(\delta)$, remain time-invariant. In game-theoretic terms, the equilibrium is time-consistent. Furthermore, the portfolio characteristics actually remain constant through time. The reason is that the difference between premiums, $p_2^* - p_1^* = (1 + \omega(\beta))z_{1,2}^e m_\Gamma$, clearly is constant over time, not dependent on the reserve difference δ, and by Section D.3, portfolio sizes and average claim frequencies for the companies only depend on the difference in premiums.

D.5.1 Numerical illustration

We have aimed for examples with parameters that are somehow realistic in car insurance, taking the time unit as a year and the monetary unit as one €. For gamma-distributed unobserved heterogeneity there are some studies (see Bichsel (1964)) with b very close to 1, so for the sake of illustration, we take $b = 1$. Furthermore, an average claim frequency of order 0.05–0.10 is common in Western countries, so we took $a = 0.1$. A gamma(0.1,1)-distribution has median $m_\Gamma \approx 0.0693$. Examples of gamma-distributed claim frequencies across customers are illustrated in Figure 3.1 for different values of the parameters a and b. The combinations of parameters are chosen to maintain an average claim frequency of 0.1. Assuming for simplicity that the claim sizes are exponentially distributed with parameter θ, then we additionally have that

$$\overline{x_{1,i}} = \frac{1}{\theta}\mathrm{e}^{-\theta K_i}, \qquad \overline{x_{2,i}} = \frac{2}{\theta^2}\mathrm{e}^{-\theta K_i},$$

for $i = 1, 2$, together with excess risk

$$z_{1,2}^e = \mathbb{E}[\min\{Z, K_1\} - \min\{Z, K_2\}] = \frac{1}{\theta}(\mathrm{e}^{-\theta K_2} - \mathrm{e}^{-\theta K_1}).$$

Aiming for an average claim size of 5000 €, we choose $\theta = 1/5000$. We consider a deductible for I_1 of 15% of the average claim size, that is $K_1 = 750$. Similarly for I_2 with 10% of the claim size giving $K_2 = 500$. Note in particularly that $K_1 > K_2$. For these parameter values, we get

$$\overline{x_{1,1}} = 4303.54, \qquad \overline{x_{2,1}} = 43\,035\,398.82,$$
$$\overline{x_{1,2}} = 4524.19, \qquad \overline{x_{2,2}} = 45\,241\,870.90, \qquad z^e_{1,2} = 220.65.$$

Assume further that there are $N = 1\,000\,000$ customers with identical personal safety loadings of $\omega(\beta) = 0.4$ and that the risk-free interest rate is $r = 3\%$. To get an indicator of the level of the reserves, we find a starting point, R, based on a 95% Value at Risk (VaR) principle. As N is rather large, the distribution of the sum $\sum_{i=1}^{N/2}(Z_i - 5000)$ can be approximated by the normal distribution $\mathrm{N}(0, (N/2)/\theta^2)$. Solving for the R that satisfies

$$\mathbb{P}\left(\sum_{i=1}^{N/2}(Z_i - 5000) > R\right) = 0.05,$$

using the inverse of the $\mathrm{N}(0,\, (N/2)/\theta^2)$ cdf, yields $R = 5\,815\,435.77$. Next, I_1 is assumed to have a reserve somewhat more than R, and I_2 somewhat less. More specifically, we let

$$r_{1,0} = (1+\gamma)R \;\text{ and }\; r_{2,0} = (1-\gamma)R, \tag{3.35}$$

which leads to an initial reserve difference of $\delta = 2\gamma R$. Choosing e.g., $\gamma = 0.2$ we get a difference of $\delta = 2\,326\,174.31$. Since the analytic results do not depend on the bounds on the reserve, ℓ_u and ℓ_d, their particular values do not matter, and we just need that the interval $[\ell_d, \ell_u]$ contains the chosen δ. Given this value, the graph of the criterion to be optimised, $\kappa(p_1, p_2;\, \delta)$, appears in Figure 3.2, and the corresponding contour diagram in Figure 3.3.

Recall that we here consider the case where I_2 offers the better product ($K_1 > K_2$) and chooses its premium p_2 first. Given this, I_1 maximises by seeking toward the ridge that appears diagonally when choosing p_1. The market leader, I_2, takes this response function of I_1 into account, and minimises $\kappa(p_1, p_2;\, \delta)$ along the ridge, by choice of p_2. The optimum provides the Stackelberg equilibrium, at the saddle point. However, in this case the saddle is located diagonally, not parallel to the axes, and there is no Nash equilibrium. In particular, given p_1, I_2 would benefit from increasing p_2, moving away from the ridge (toward cooler colors in the figures). While this precludes Nash equilibrium, the analysis demonstrates that it is possible to obtain an equilibrium in finite premiums by having I_2 commit to some p_2 at the given δ, then letting I_1 respond, i.e., a Stackelberg equilibrium. This is also verified by the value

$$\mathrm{D}(a, b, K_1, K_2, r, \delta, \omega(\beta)) = -9603.91,$$

which tells us that condition (3.32) is satisfied, whereas (3.34) is not, as $-4(1 + \omega(\beta)z^e_{1,2} = -1235.62$, i.e., greater than $\mathrm{D}(\cdot)$ in this case.

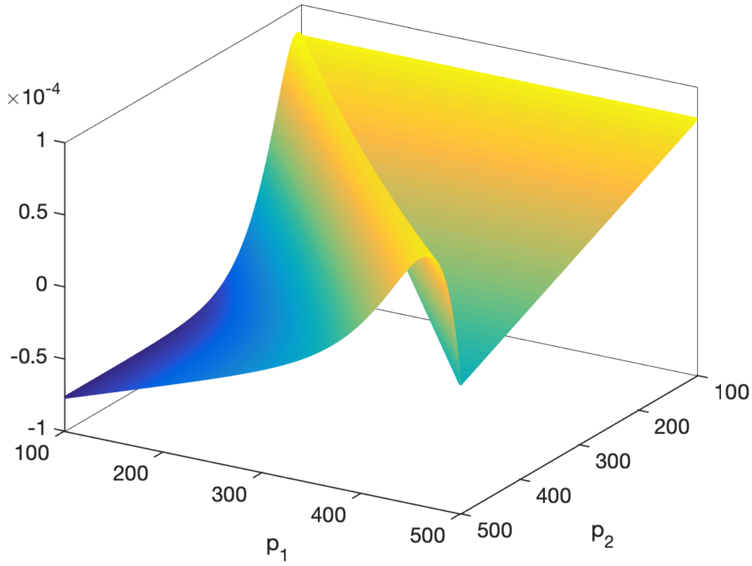


Figure 3.2: Graph of $\kappa(p_1, p_2; \delta)$.

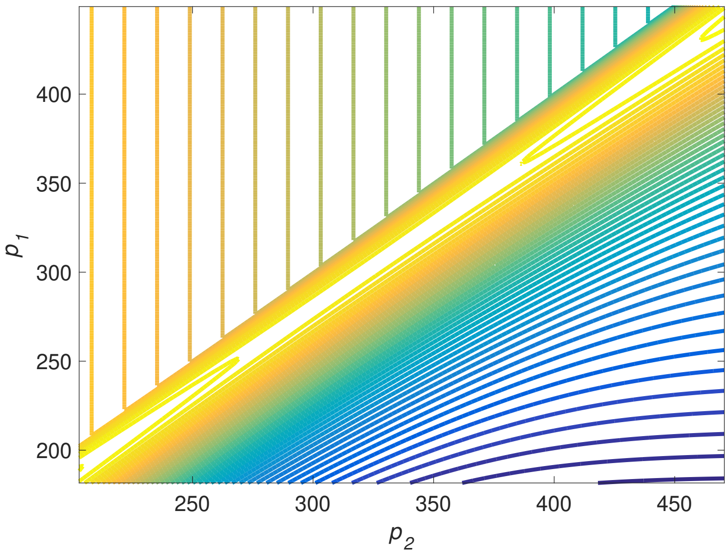


Figure 3.3: Contour diagram of $\kappa(p_1, p_2; \delta)$.

From Theorem D.6, we compute the Stackelberg equilibrium premiums

$$p_1^* = 305.5 \qquad \text{and} \qquad p_2^* = 326.0 \tag{3.36}$$

at the current reserve difference $\delta = 2\,326\,174.31$. These are to be compared with the net premiums

$$\alpha_1(p_1^*, p_2^*)\,\overline{x_{1,1}} = 0.0307 \cdot 4303.54 = 132.1,$$
$$\alpha_2(p_1^*, p_2^*)\,\overline{x_{1,2}} = 0.1693 \cdot 4524.19 = 766.0,$$

so that pursuing solely the competition aspects would lead to a likely loss for I_2 at the current reserve difference. This is not necessarily a paradox since the perspective of control and game theory is to focus solely on a one-eyed goal. Larger δ means I_2 is lagging more behind the large firm I_1, and this gives I_2 greater incentive to compete for customers by lowering its premium, with I_1 responding by letting premiums move in lockstep. Thus, in equilibrium, I_2 always receives a higher premium than I_1, reflecting the higher quality product (lower deductible). This type of product is attractive to “bad” customers, that is, customers with high claim frequency, as seen in Figure 3.4. These customers are expected to experience more losses than “good” customers, and are therefore willing to pay extra for better coverage, yielding a separating equilibrium, with customers’ choices revealing their type, as in Rothschild and Stiglitz (1976). Still, I_2 may remain the smallest company, due to the higher risk of its customers.

In Figures 3.5 and 3.6, exhibiting aspects of D, we take a closer graphical look at the second order criteria. Starting with Figure 3.5, we plot D as a function of δ. All other parameters remain the same as above. Values for δ for which $\mathrm{D} < -4(1+\omega(\beta)z_{1,2}^e$ are plotted in green to indicate that the equilibrium is of Stackelberg type. Values that yield $-4(1+\omega(\beta)z_{1,2}^e < \mathrm{D} < 0$, and hence equilibrium of Nash-type, are plotted in blue. Finally, the values plotted with red give $\mathrm{D} > 0$, which tells us that there is no equilibrium. Here we see that D is indeed a linearly increasing function of δ, as it should be according to (3.30) and (3.33). Hence, for small δ-values we get a Stackelberg equilibrium (green). For a small spectrum in the middle we get a Nash equilibrium (blue), and, finally, for large values of δ there is no equilibrium. The same color codes are used in Figure 3.6, which shows the color plateaus of D, and not the actual values, as depending on the deductibles, K_1 and K_2. As we restrict the analysis to the case where $K_1 > K_2$, it is only the lower triangular part that is illustrated. For simultaneously large values (above 5×10^4) of K_1 and K_2, there appears an area (red) where there is no equilibrium. However, 5×10^4 is ten times the average claim size of 5000 and obviously an unrealistically large value of the deductibles. For K_1 and K_2 being close, i.e., along the diagonal, there is then an equilibrium of Stackelberg type (green area) for smaller values. Moving away from the diagonal, the equilibrium type will change from Stackelberg to Nash (blue area). However, in the most realistic region of deductibles K_1, K_2 being below the mean claim size $5000 = 0.5 \times 10^4$ it is always Stackelberg.

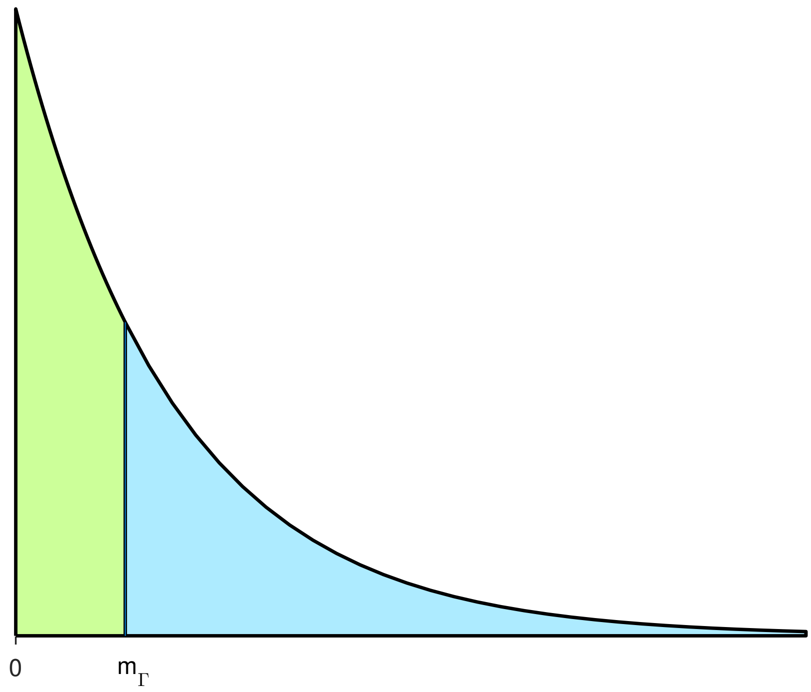


Figure 3.4: Distribution of customers in equilibrium, where customers with claim frequencies in the green (blue) area insure at I_1 (I_2).

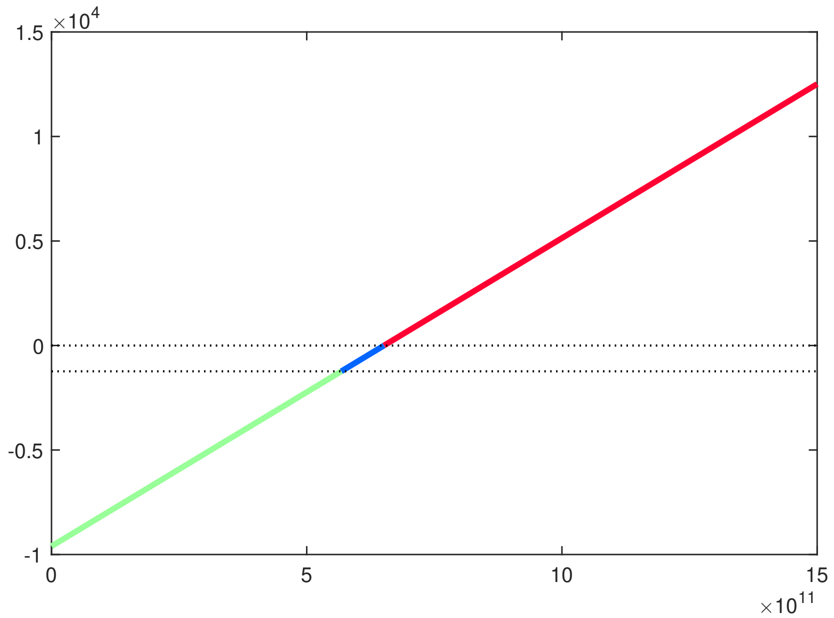


Figure 3.5: D as a function of δ. Green indicates Stackelberg equilibrium, blue indicates Nash equilibrium, and red indicates neither.

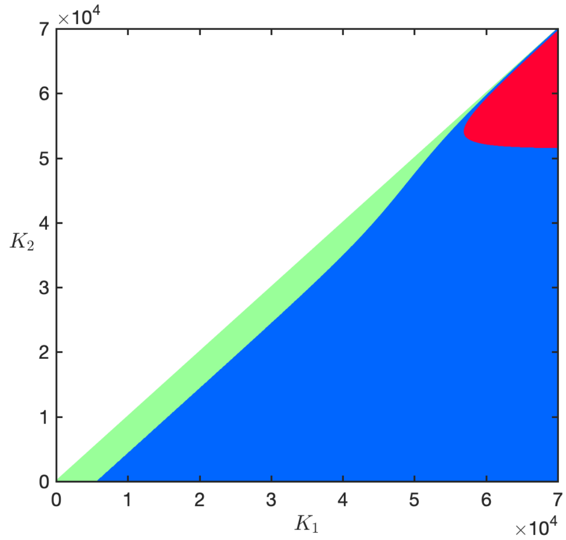


Figure 3.6: D as a function of K_1 and K_2 for $K_1 > K_2$. Green indicates Stackelberg equilibrium, blue indicates Nash equilibrium, and red indicates neither.

D.6 Conclusions

We have considered a non-life insurance market in which two insurance companies compete for customers by choice of premium strategies. Each company chooses its strategy to balance revenue against portfolio size, taking into account the strategy of the other company. We pay special attention to product differentiation and customer risk, while abstracting for simplicity from market frictions. For product differentiation, we focus on different deductibles, noting that alternatives would include bonus-malus systems, and proportional compensation in deductibles. The analysis is carried out in continuous time using stochastic differential game techniques. Adverse selection implies that a change in premium alters the risk composition of the portfolio. With claim arrival rates following a gamma distribution across customers, Stackelberg equilibrium premiums are derived. Conditions under which a Nash equilibrium exists are also established, but our numerical examples indicate that Stackelberg is the more typical case. Equilibrium premiums depend in an affine fashion on the running difference between the reserves of the companies, each modelled using the diffusion approximation to a standard Cramér-Lundberg risk process, extended to allow investment in a risk-free asset. Numerical illustrations of both types of equilibrium are provided.

Overall, the managerial implications are that insurance companies should consider the premium as an active means to control portfolio size and revenue per customer in competition with other companies, as opposed to merely pooling individual risks and setting the premium based on conventional principles. Future research could consider three or more companies competing for market shares, to account explicitly for the risk of ruin or the possibility that some potential customers choose not to insure, or to pursue the more sophisticated ideas of Asmussen et al. (2018) on the customer's problem.

Appendix

D.I Second Order Derivative Tests for Theorem D.6

Since we only need to consider a fixed δ, we write for notational convenience $\kappa(p_1, p_2)$ instead of $\kappa(p_1, p_2; \delta)$. Please note that the first order conditions, in the present case (3.17), can be written as

$$\kappa_i'(p_1, p_2) = \frac{1}{\kappa_d(p_1, p_2)} \left(\kappa_{n,i}'(p_1, p_2) - \kappa(p_1, p_2; \delta)\kappa_{d,i}'(p_1, p_2)\right) = 0 \text{ for } i = 1, 2,$$

and consider the second order partial derivatives

$$\begin{aligned}
\kappa_{ii}''(p_1, p_2) = \frac{1}{\kappa_d(p_1, p_2)} \Big(&\kappa_{n,ii}''(p_1, p_2) - \kappa(p_1, p_2; \delta)\kappa_{d,ii}''(p_1, p_2) \\
&- \frac{1}{\kappa_d(p_1, p_2)}\kappa_{d,i}'(p_1, p_2)\left(\kappa_{n,i}'(p_1, p_2) - \kappa(p_1, p_2; \delta)\kappa_{d,i}'(p_1, p_2)\right)\Big) \\
&- \frac{1}{\kappa_d(p_1, p_2)^2}\kappa_{d,i}'(p_1, p_2)\left(\kappa_{n,i}'(p_1, p_2) - \kappa(p_1, p_2; \delta)\kappa_{d,i}'(p_1, p_2)\right), \\
\kappa_{ij}''(p_1, p_2) = \frac{1}{\kappa_d(p_1, p_2)} \Big(&\kappa_{n,ij}''(p_1, p_2) - \kappa(p_1, p_2; \delta)\kappa_{d,ij}''(p_1, p_2) \\
&- \frac{1}{\kappa_d(p_1, p_2)}\kappa_{d,i}'(p_1, p_2)\left(\kappa_{n,j}'(p_1, p_2) - \kappa(p_1, p_2; \delta)\kappa_{d,j}'(p_1, p_2)\right)\Big) \\
&- \frac{1}{\kappa_d(p_1, p_2)^2}\kappa_{d,j}'(p_1, p_2)\left(\kappa_{n,i}'(p_1, p_2) - \kappa(p_1, p_2; \delta)\kappa_{d,i}'(p_1, p_2)\right).
\end{aligned}$$

In optimum the critical point (p_1^*, p_2^*) must satisfy the first order condition (3.27), which reduces the second order partial derivatives to

$$\begin{aligned}
\kappa_{ii}''(p_1^*, p_2^*) &= \frac{1}{\kappa_d(p_1^*, p_2^*)}\Big(\kappa_{n,ii}''(p_1^*, p_2^*) - \tilde{\kappa}\kappa_{d,ii}''(p_1^*, p_2^*)\Big), \\
\kappa_{ij}''(p_1^*, p_2^*) &= \frac{1}{\kappa_d(p_1^*, p_2^*)}\Big(\kappa_{n,ij}''(p_1^*, p_2^*) - \tilde{\kappa}\kappa_{d,ij}''(p_1^*, p_2^*)\Big).
\end{aligned}$$

From the links between the first order derivatives in the proof of Theorem D.6,

$$\kappa_{22}''(p_1^*, p_2^*) = \frac{1}{\kappa_d(p_1^*, p_2^*)}\Big(\kappa_{n,22}''(p_1^*, p_2^*) - \tilde{\kappa}\kappa_{d,22}''(p_1^*, p_2^*)\Big) \tag{3.37}$$

$$= \frac{1}{\kappa_d(p_1^*, p_2^*)}\Big(\kappa_{n,11}''(p_1^*, p_2^*) - 2f_1'(p_1^*, p_2^*) - \tilde{\kappa}\kappa_{d,11}''(p_1^*, p_2^*)\Big), \tag{3.38}$$

$$\kappa_{12}''(p_1^*, p_2^*) = \frac{1}{\kappa_d(p_1^*, p_2^*)}\Big(\kappa_{n,12}''(p_1^*, p_2^*) - \tilde{\kappa}\kappa_{d,12}''(p_1^*, p_2^*)\Big) \tag{3.39}$$

$$= \frac{1}{\kappa_d(p_1^*, p_2^*)}\Big(- \kappa_{n,11}''(p_1^*, p_2^*) + f_1'(p_1^*, p_2^*) + \tilde{\kappa}\kappa_{d,11}''(p_1^*, p_2^*)\Big), \tag{3.40}$$

where $f(p_1, p_2) = \gamma(b, y/a) - \Gamma(b, y/a)$. The second order derivative test on the Hessian in (3.19),

$$\kappa''_{11}(p_1^*, p_2^*)\kappa''_{22}(p_1^*, p_2^*) - \kappa''_{12}(p_1^*, p_2^*)^2 = -\frac{1}{\kappa_d(p_1^*, p_2^*)^2} f'_1(p_1^*, p_2^*)^2 < 0,$$

then confirms a saddle point, provided we can show the condition (3.18). For this, we need to be more specific and find the actual second order derivatives and evaluate them in equilibrium. Differentiating $\kappa'_{n,1}(p_1, p_2)$ and $\kappa'_{d,1}(p_1, p_2)$ with respect to p_1 yields

$$\begin{aligned}
\kappa''_{n,11}(p_1, p_2) &= \frac{\overline{x_{1,1}} + \overline{x_{1,2}}}{a\big((1+\omega(\beta))z^e_{1,2}\big)^2} \mathrm{e}^{-y/a}(y/a)^{b-1}(y/a - b) \\
&\quad - \frac{2}{a(1+\omega(\beta))z^e_{1,2}} \mathrm{e}^{-y/a}(y/a)^{b-1} \\
&\quad - \frac{p_1 + p_2}{(a(1+\omega(\beta))z^e_{1,2})^2} \mathrm{e}^{-y/a}(y/a)^{b-2}(y/a - b + 1), \\
\kappa''_{d,11}(p_1, p_2) &= \frac{\overline{x_{2,2}} - \overline{x_{2,1}}}{a((1+\omega(\beta))z^e_{1,2})^2} \mathrm{e}^{-y/a}(y/a)^{b-1}(y/a - b).
\end{aligned}$$

Evaluating at the equilibrium premiums,

$$\begin{aligned}
\kappa''_{n,11}(p_1^*, p_2^*) &= \frac{\overline{x_{1,1}} + \overline{x_{1,2}}}{a((1+\omega(\beta))z^e_{1,2})^2} \mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^{b-1}(m_\Gamma/a - b) \\
&\quad - \frac{2}{a(1+\omega(\beta))z^e_{1,2}} \mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^{b-1} \\
&\quad - \frac{2p_2^* - (1+\omega(\beta))z^e_{1,2}m_\Gamma}{(a(1+\omega(\beta))z^e_{1,2})^2} \mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^{b-2}(m_\Gamma/a - b + 1), \\
\kappa''_{d,11}(p_1, p_2) &= \frac{\overline{x_{2,2}} - \overline{x_{2,1}}}{a((1+\omega(\beta))z^e_{1,2})^2} \mathrm{e}^{-m_\Gamma/a}(m_\Gamma/a)^{b-1}(m_\Gamma/a - b).
\end{aligned}$$

Multiplying by the positive constant $\kappa_d(p_1^*, p_2^*)a((1+\omega(\beta))z^e_{1,2})^2 \exp(\frac{m_\Gamma}{a})/(\frac{m_\Gamma}{a})^{b-1}$, the criterion can be written in reduced form explicitly as

$$\begin{aligned}
&-2(1+\omega(\beta))z^e_{1,2} - (2p_2^*/a - (1+\omega(\beta))z^e_{1,2}m_\Gamma/a)\,(m_\Gamma/a)^{-1}\,(m_\Gamma/a - b + 1) \\
&\quad + (\overline{x_{1,1}} + \overline{x_{1,2}})(m_\Gamma/a - b) - \tilde{\kappa}(\overline{x_{2,2}} - \overline{x_{2,1}})(m_\Gamma/a - b) < 0.
\end{aligned}$$

Inserting the optimal premium,

$$\begin{aligned}
p_2^* = \tfrac{a}{2}\Big(&\tfrac{1}{2}\mathrm{e}^{m_\Gamma/a}(m_\Gamma/a)^{1-b}\Gamma(b)(1+\omega(\beta))z^e_{1,2} + (1+\omega(\beta))z^e_{1,2}(m_\Gamma/a) \\
&+(m_\Gamma/a)(\overline{x_{1,1}} + \overline{x_{1,2}}) - (m_\Gamma/a)\tilde{\kappa}(\overline{x_{2,2}} - \overline{x_{2,1}})\Big),
\end{aligned}$$

we can reduce the condition to

$$\begin{aligned}
&\tilde{\kappa}(\overline{x_{2,2}} - \overline{x_{2,1}}) - (\overline{x_{1,1}} + \overline{x_{1,2}}) \\
&\quad - 2(1+\omega(\beta))z^e_{1,2} - \tfrac{1}{2}\mathrm{e}^{m_\Gamma/a}(m_\Gamma/a)^{-b}\Gamma(b)(1+\omega(\beta))z^e_{1,2}(m_\Gamma/a - b + 1) < 0,
\end{aligned}$$

which is the same as (3.32).

The condition (3.25) for a Nash equilibrium can also be found more explicitly by using the link in (3.38) between the second order derivatives. The condition can be rewritten as

$$\frac{1}{\kappa_d(p_1^*, p_2^*)}\big(\kappa''_{n,11}(p_1^*, p_2^*) - 2f_1'(p_1^*, p_2^*) - \tilde{\kappa}\kappa''_{d,11}(p_1^*, p_2^*)\big) > 0,$$

which, using the same approach as above, can be written as

$$\begin{aligned}
&2\frac{a((1+\omega(\beta))z_{1,2}^e)^2}{(m_\Gamma/a)^{b-1}\exp(-m_\Gamma/a)} f_1'(p_1^*, p_2^*) \\
&\quad < \tilde{\kappa}(\,\overline{x_{2,2}} - \overline{x_{2,1}}) - 2(1+\omega(\beta))z_{1,2}^e \\
&\qquad - \tfrac{1}{2}\mathrm{e}^{m_\Gamma/a}(m_\Gamma/a)^{-b}\Gamma(b)(1+\omega(\beta))z_{1,2}^e(m_\Gamma/a - b + 1) - (\,\overline{x_{1,1}} + \overline{x_{1,2}}),
\end{aligned}$$

where

$$\begin{aligned}
&\frac{a((1+\omega(\beta))z_{1,2}^e)^2}{(m_\Gamma/a)^{b-1}\exp(-m_\Gamma/a)} f_1'(p_1^*, p_2^*) \\
&\quad = -2\frac{a((1+\omega(\beta))z_{1,2}^e)^2}{(m_\Gamma/a)^{b-1}\exp(-m_\Gamma/a)} \frac{(m_\Gamma/a)^{b-1}\exp(-m_\Gamma/a)}{a(1+\omega(\beta)z_{1,2}^e} \\
&\quad = -2(1+\omega(\beta))z_{1,2}^e,
\end{aligned}$$

which combined yields (3.34).

References

Albrecher, H., Teugels, J. L., and Beirlant, J. (2017). *Reinsurance: actuarial and statistical aspects*. John Wiley & Sons.

Asmussen, S., Christensen, B. J., and Taksar, M. (2013). “Portfolio size as function of the premium: modelling and optimization”. *Stochastics* 85.4, pp. 575–588.

Asmussen, S., Christensen, B. J., and Thøgersen, J. (2018). “Equilibrium premium strategies for push-pull competition in a non-life insurance market”. *Pending minor revision in Insurance: Mathematics and Economics.*

Banneheka, B. and Ekanayake, G. (2009). “New point estimator for the median of gamma distribution”. *Vidyodaya Journal of Science* 14, pp. 95–103.

Bensoussan, A., Chen, S., and Sethi, S. P. (2015). “The maximum principle for global solutions of stochastic Stackelberg differential games”. *SIAM Journal on Control and Optimization* 53.4, pp. 1956–1981.

Bensoussan, A., Siu, C. C., Yam, S. C. P., and Yang, H. (2014). “A class of non-zero-sum stochastic differential investment and reinsurance games”. *Automatica* 50.8, pp. 2025–2037.

Bichsel, F. (1964). “Erfahrungstarifierung in der Motorfahrzeug-Haftpflicht-Versicherung”. *Bulletin of the Swiss Association of Actuaries*, pp. 119–130.

Boonen, T. J., Pantelous, A. A., and Wu, R. (2018). “Non-cooperative dynamic games for general insurance markets”. *Insurance: Mathematics and Economics* 78, pp. 123–135.

Borch, K. H. (1962). “Application of game theory to some problems in automobile insurance”. *ASTIN Bulletin* 2.2, pp. 208–221.

Brown, J. R. and Goolsbee, A. (2002). “Does the internet make markets more competitive? Evidence from the life insurance industry”. *Journal of Political Economy* 110.3, pp. 481–507.

Bühlmann, H. and Gisler, A. (2006). *A course in credibility theory and its applications*. Springer.

Chen, L. and Shen, Y. (2018). “On a new paradigm of optimal reinsurance: A stochastic Stackelberg differential game between an insurer and a reinsurer”. *ASTIN Bulletin* 48.2, pp. 905–960.

Denuit, M., Maréchal, X., Pitrebois, S., and Walhin, J.-F. (2007). *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems*. John Wiley & Sons.

Diamond, P. A. (1982). “Wage determination and efficiency in search equilibrium”. *The Review of Economic Studies* 49.2, pp. 217–227.

Dutang, C., Albrecher, H., and Loisel, S. (2013). “Competition among non-life insurers under solvency constraints: A game-theoretic approach”. *European Journal of Operational Research* 231.3, pp. 702–711.

Emms, P. (2007). “Dynamic pricing of general insurance in a competitive market”. *ASTIN Bulletin* 37.01, pp. 1–34.

— (2011). “Pricing general insurance in a reactive and competitive market”. *Journal of Computational and Applied Mathematics* 236.6, pp. 1314–1332.

— (2012). “Equilibrium pricing of general insurance policies”. *North American Actuarial Journal* 16.3, pp. 323–349.

Emms, P. and Haberman, S. (2005). “Pricing general insurance using optimal control theory”. *ASTIN Bulletin* 35.02, pp. 427–453.

Emms, P., Haberman, S., and Savoulli, I. (2007). “Optimal strategies for pricing general insurance”. *Insurance: Mathematics and Economics* 40.1, pp. 15–34.

Iglehart, D. L. (1969). “Diffusion approximations in collective risk theory”. *Journal of Applied Probability* 6.2, pp. 285–292.

Jin, Z., Yin, G., and Wu, F. (2013). “Optimal reinsurance strategies in regime-switching jump diffusion models: Stochastic differential game formulation and numerical methods”. *Insurance: Mathematics and Economics* 53.3, pp. 733–746.

Lin, X., Zhang, C., and Siu, T. K. (2012). “Stochastic differential portfolio games for an insurer in a jump-diffusion risk process”. *Mathematical Methods of Operations Research* 75.1, pp. 83–100.

Meng, H., Li, S., and Jin, Z. (2015). “A reinsurance game between two insurance companies with nonlinear risk processes”. *Insurance: Mathematics and Economics* 62, pp. 91–97.

Mortensen, D. T. (1982). “Property rights and efficiency in mating, racing, and related games”. *The American Economic Review* 72.5, pp. 968–979.

Mortensen, D. T. and Pissarides, C. A. (1994). "Job creation and job destruction in the theory of unemployment". *The Review of Economic Studies* 61.3, pp. 397–415.

Pantelous, A. A. and Passalidou, E. (2013). "Optimal premium pricing policy in a competitive insurance market environment". *Annals of Actuarial Science* 7.2, pp. 175–191.

— (2015). "Optimal premium pricing strategies for competitive general insurance markets". *Applied Mathematics and Computation* 259, pp. 858–874.

— (2017). "Optimal strategies for a non-linear premium-reserve model in a competitive insurance market". *Annals of Actuarial Science* 11.1, pp. 1–19.

Polborn, M. K. (1998). "A model of an oligopoly in an insurance market". *The Geneva Papers on Risk and Insurance Theory* 23.1, pp. 41–48.

Powers, M. R., Shubik, M., and Yao, S. T. (1998). "Insurance market games: Scale effects and public policy". *Journal of Economics* 67.2, pp. 109–134.

Pun, C. S. and Wong, H. Y. (2016). "Robust non-zero-sum stochastic differential reinsurance game". *Insurance: Mathematics and Economics* 68, pp. 169–177.

Robert, C. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation.* Springer.

Rothschild, M. and Stiglitz, J. (1976). "Equilibrium in competitive insurance markets: An essay on the economics of imperfect information". *The Quarterly Journal of Economics* 90.4, pp. 629–649.

Salop, S. and Stiglitz, J. (1977). "Bargains and ripoffs: A model of monopolistically competitive price dispersion". *The Review of Economic Studies* 44.3, pp. 493–510.

Shi, J., Wang, G., and Xiong, J. (2016). "Leader–follower stochastic differential game with asymmetric information and applications". *Automatica* 63, pp. 60–73.

Taksar, M. and Zeng, X. (2011). "Optimal non-proportional reinsurance control and stochastic differential games". *Insurance: Mathematics and Economics* 48.1, pp. 64–71.

Taylor, G. C. (1986). "Underwriting strategy in a competitive insurance environment". *Insurance: Mathematics and Economics* 5.1, pp. 59–77.

Taylor, G. C. (1987). "Expenses and underwriting strategy in competition". *Insurance: mathematics and economics* 6.4, pp. 275–287.

Thøgersen, J. (2016). "Optimal premium as a function of the deductible: customer analysis and portfolio characteristics". *Risks* 4.4, p. 42.

von Stackelberg, H. (1934). *Marktform und Gleichgewicht.* J. Springer.

Wu, R. and Pantelous, A. A. (2017). "Potential games with aggregation in non-cooperative general insurance markets". *ASTIN Bulletin* 47.1, pp. 269–302.

Yan, M., Peng, F., and Zhang, S. (2017). "A reinsurance and investment game between two insurance companies with the different opinions about some extra information". *Insurance: Mathematics and Economics* 75, pp. 58–70.

Yong, J. (2002). "A leader-follower stochastic linear quadratic differential game". *SIAM Journal on Control and Optimization* 41.4, pp. 1015–1041.

Zeng, X. (2010). "A stochastic differential reinsurance game". *Journal of Applied Probability* 47.2, pp. 335–349.

Chapter 4

Personal non-life insurance decisions

Within theory of demand for insurance, a popular result by Arrow (1963) (reprinted in Arrow (1971)) states that if the premium is chosen by the expected premium principle, a fixed amount deductible structure is optimal for an individual seeking to optimise expected utility. This conclusion was afterwards supported by different studies, e.g. Gollier and Schlesinger (1996) and Karni (1992), in which the authors show that the optimality of the fixed amount deductible is robust to more general decision models than the expected utility model. However, pricing by the expected premium principle is not as questioned.

As we saw in the introduction of Chapter 2, a lot of effort have been invested into developing more advanced pricing mechanisms. Much less effort seems to have been invested into the product structure for individuals. The fixed amount deductible is still the dominating product structure in non-life insurance both in the literature as well as in practice. In Paper E, we see how this mismatch between the development of premium methods and deductible structures affects the individual in terms of welfare loss.

In the preceding papers, we have seen a way to model the customer's problem. Paper E takes a different approach by introducting consumption. In Section 4.1.1 we explain the difference and how it affects the analysis. Both consumption and amount of insurance are then controls at the individual's disposal. Among studies also dealing with consumption-insurance optimisation problems are Somerville (2004), Briys (1986) and Moffet (1977).

In Paper E, we use a martingale approach to premium calculation motivated by Delbaen and Haezendonck (1989). The inspiration stems from financial pricing where the absence of arbitrage leads to the existence of a risk neutral pricing measure. Section 4.1.2 goes through the justifying arguments for applying such a measure for insurance pricing. For a thorough comparison between classical actuarial pricing and financial pricing, see Embrechts (2000), and for more insight into the financial terms and mathematical details of arbitrage, completeness, change of measure, ect., we refer to Björk (2009). In Paper E, we mention how this way of pricing relates to some of the standard premium principle mentioned in Section 2.1.2.

4.1 Preliminaries

4.1.1 The individual's wealth and objective function

In the three previous papers, as a part of the analysis, the so-called customer's problem is considered by evaluating the present value $\mathbb{E}[\int_0^\infty \exp(-dt)\mathrm{d}w_t]$, where $(w_t)_{t\geq 0}$ is the wealth process only affected by the insurance decision and d is a subjective discount rate. In the present paper we consider an individual who, in addition, consumes. The individual then seeks to find the optimal consumption-insurance strategies by evaluating the expected discounted utility of consumption in (4.4).

As previously, the individual's loss process is assumed to be a compound Poisson process, which (unlike for the insurance company) cannot be argued to be approximated by a diffusion process as losses are few and large relative to the level of wealth and thereby creating significant fluctuations. Hence, the wealth process takes the form of 4.2. Compared to the formulation of the stochastic control problem in Section 2.1.5, this difference in dynamics eliminates the diffusion term of the infinitesimal operator and adds a term related to the jumps of the compound Poisson process, as we see in (4.3). This of course reflects upon the HJB-equation (4.5).

4.1.2 Pricing by change of measure

Let $(A_t)_{t\in[0,T]}$ be a compound Poisson process with parameters (λ, F) as in Definition 1.5. It constitutes a risk in terms of a total claim amount up to a terminal time $T > 0$. Furthermore, let $(\mathcal{F}_t)_{t\geq 0}$ be the filtration generated by $(A_t)_{t\in[0,T]}$.

Assume that the insurance company (or in this chapter: the individual) at any time $t \in [0, T]$ can sell the remaining risk for a predictable premium p_t. The underlying price process $(Y_t)_{t\in[0,T]}$, representing the company's liabilities, then consists of two parts, i) the claims A_t up to time t, and ii) the premium p_t for the remaining risk of $A_T - A_t$, i.e.

$$Y_t = A_t + p_t.$$

The possibility of buying and selling a risk instantly assumes a liquid market, which should rule out arbitrage. If arbitrage opportunities exists, then it is possible to create a portfolio that profits from mispricing without taking on any risk. In line with financial theory, on an arbitrage-free market there exists a risk-neutral pricing measure $\mathbb{Q}$ such that the underlying price process $(Y_t)_{t\in[0,T]}$ is a martingale under $\mathbb{Q}$.

The attention is restricted to premiums on the form

$$p_t = p(T - t),$$

where p is a premium density. In this case it can be shown that the compound Poisson structure of $(A_t)_{t\in[0,T]}$ is maintained under $\mathbb{Q}$, but with altered characteristics $(\lambda^{\mathbb{Q}}, F^{\mathbb{Q}})$. A feasible premium density is then

$$p^{\mathbb{Q}} = \mathbb{E}^{\mathbb{Q}}[A_1] = \mathbb{E}^{\mathbb{Q}}[N_1]\mathbb{E}^{\mathbb{Q}}[X_1] > \mathbb{E}[N_1]\mathbb{E}[X_1],$$

which splits the effect of the measure change into a claim frequency part and a claim size distribution part of the underlying compound Poisson process. The idea is, in general, that by changing the measure, more weight is given to less favourable outcomes.

Delbaen and Haezendonck (1989) links a risk-neutral pricing measure $\mathbb{Q}(\beta)$ to (something we here call) a non-decreasing pricing measure function $\beta : \mathbb{R}^+ \to \mathbb{R}$ by the Radon-Nikodym derivative (4.14). Then for three different choice of the pricing measure function, we compare two cases: one where the individual is restricted to a fixed amount deductible structure and one where the individual freely can choose the deductible structure. The comparison is based on the welfare of the individual.

References

Arrow, K. (1963). "Uncertainty and the Welfare Economics of Medical Care". *American Economic Review* 53.5, pp. 941–969.

— (1971). *Essays in the theory of risk bearing*. Markham Publishing Co.

Björk, T. (2009). *Arbitrage theory in continuous time*. 3rd. Oxford University Press.

Briys, E. (1986). "Insurance and consumption: The continuous time case". *The Journal of Risk and Insurance* 53.4, pp. 718–723.

Delbaen, F. and Haezendonck, J. (1989). "A martingale approach to premium calculation principles in an arbitrage free market". *Insurance: Mathematics and Economics* 8.4, pp. 269–277.

Embrechts, P. (2000). "Actuarial versus financial pricing of insurance". *The Journal of Risk Finance* 1.4, pp. 17–26.

Gollier, C. and Schlesinger, H. (1996). "Arrow's theorem on the optimality of deductibles: a stochastic dominance approach". *Economic Theory* 7.2, pp. 359–363.

Karni, E. (1992). "Optimal insurance: A nonexpected utility analysis". *Contributions to Insurance Economics*. Vol. 13. Springer, Dordrecht, pp. 217–238.

Moffet, D. (1977). "Optimal deductible and consumption theory". *Journal of Risk and Insurance* 44.4, pp. 669–682.

Somerville, R. A. (2004). "Insurance, consumption, and saving: A dynamic analysis in continuous time". *American Economic Review* 94.4, pp. 1130–1140.

Paper E

Personal Non-Life Insurance Decisions and the Welfare Loss from Flat Deductibles

MOGENS STEFFENSEN AND JULIE THØGERSEN

ABSTRACT. We view the retail non-life insurance decision from the perspective of the insured. We formalise different consumption-insurance problems depending on the flexibility of the insurance contract. For exponential utility and power utility we find the optimal flexible insurance decision or insurance contract. For exponential utility we also find the optimal position in standard contracts that are less flexible and therefore, for certain non-linear pricing rules, lead to a welfare loss for the individual insuree compared to the optimal flexible insurance decision. For the exponential loss distribution, we quantify a significant welfare loss. This calls for product development in the retail insurance business.

KEYWORDS: exponential utility, HJB equation, insurance pricing, product design, compound Poisson loss process

Publication details: ASTIN Bulletin, 2018, vol. 49(1), pp. 85–116. DOI: 10.1017/asb.2018.40

E.1 Introduction

The standard marketed retail insurance contract has a fixed amount deductible where the deductible is independent of the size of the loss. This simple product structure is sufficient but only for the simple pricing mechanism where the loading to the expectation is a plain factor to the expectation, also called the expectation principle. Both theoretically and practically other pricing principles are more relevant but the main focus has still been on the fixed deductible. We study how this mismatch of development between pricing and product design affects individuals in the insurance market. We therefore formulate and solve the personal consumption-insurance problem with respect to a non-life risk modelled by a compound Poisson process with the objective to maximise utility of consumption. Two types of insurance products are considered: one with an optimal flexible, claim-dependent deductible and one with an optimal constant fixed amount deductible (the standard product). These are compared by measuring the welfare loss of an individual, which we define as the monetary compensation the individual requires in addition to the standard product in order to be indifferent between that and the optimal flexible product. For the expected value premium principle, henceforth referred to as the linear pricing principle due to its linear relation to the expected value of the risk, the constant deductible is known to be optimal and there is no welfare loss from being offered a standard constant deductible contract. For a certain formalization of the variance pricing principle, we find a log-power deductible to be optimal, and for a certain formalization of the Esscher pricing principle, we find a linear deductible to be optimal. In both of these cases a welfare loss arises if the individual is offered a standard constant deductible contract.

Stochastic control theory has been applied intensively to decision problems in insurance over the last decades. In life and pension insurance the applications are in two separate directions: The Asset-Liability Management decisions of a pension fund and an individual's financial consumption-investment-insurance decisions. In some formulations the two directions have much in common. In non-life insurance, most applications are to the decision making of the insurance company. Here the focus has been on the decisions concerning reinsurance, investments, premium collection, and dividends paid to the owners. A standard objective is the expected accumulated present value of future dividend payouts until ruin. There has been less focus on the non-life insurance decisions made by the individual over her life-cycle in the sense of personal financial consumption-investment-insurance decisions with respect to non-life risks. Our study is a contribution in the latter direction. To formulate an individual's non-life risk decision, we need to think carefully about *how the wealth process is influenced* by non-life risk, *what can be controlled* and how, and *what is the objective function*. In the next three paragraphs we address these three ingredients of the control problem one by one in order to make our standing point clear.

The compound Poisson process is well-established as a benchmark for modelling a portfolio of non-life risks. This is also called a collective risk model. We choose the same model for an individual's non-life risks. This is consistent in the sense that if an individual risk process follows a compound Poisson process, then an aggregation of those in a portfolio is also a compound Poisson process. If further the individual risk processes are homogeneous across individuals, then the collective follows a compound Poisson process with the same claim amount distribution but with a claim intensity corresponding to the individual claim intensity times the number of individuals.

The individual can typically control the non-life risk by the choice of a deductible in her insurance contract only. In practice, most often she can choose among different levels of a constant deductible. Choosing a deductible in an individual's risk process then corresponds to choosing the deductible in an Excess-of-Loss reinsurance program for an insurance company with a collective risk process. However, in the reinsurance program the insurance company may have other decisions to make, e.g. concerning the proportion covered if only proportional reinsurance is bought on top of the deductible. Typically, an individual does not have such a decision to make. We solve different problems regarding this (lack of) flexibility in the insurance products offered to individuals. Both problems where a constant deductible is chosen (the realistic case) and problems where the deductible is a general function of the loss are solved. Thereby we are able to quantify the welfare loss that arises from giving the individual the choice of a constant deductible only. It is a classical result in non-life insurance that the Excess-of-Loss insurance contract is optimal, so therefore at first glance, there should be no welfare loss. But this classical result is only obtained for linear pricing rules. If pricing is non-linear, then the result does not hold anymore and there is, indeed, a welfare loss to detect. We solve the problem for various pricing rules, including non-linear pricing rules.

The objective function in personal financial decision making is often taken to be aggregate utility of consumption financed by wealth, capital gains, and perhaps labor income. This is in contrast to the classical dividend optimization problem for a non-life company where dividends are not bent by a concave utility function before aggregation. On the other hand, dividends run out upon ruin of the insurance company, which in itself forms an indirect aversion towards risk. There exist works where dividends are measured by their utility, but the mainline research counts in dividends linearly in the objective. Apart from the appearance in the objective function with or without a utility function, the consumption of the individual influences individual wealth in the same way as dividends influence the insurance company's risk process. We solve optimal insurance coverage for a general deductible in both cases of exponential and power utility. However, we are able to characterise the solution to the case where the deductible has to be constant in the case of exponential utility only. Therefore, our focus is on this case and the explicit quantification of the welfare loss arising from suboptimal insurance contracts (constant deductible) under non-linear pricing is carried out for exponential utility only. In this type of problem it is possible for the individual to change strategy more often than the time horizon of the objective. It is therefore natural to consider it as a dynamic optimisation problem due to the long-term objective combined with the short-term decisions. This is in contrast to problems where the time horizons for the objective and the decisions are aligned, in which case a one-period (long or short) model suffices. Only with dynamic optimization, future optionality and impact on objectives are correctly counted in and the dependence of state processes time and wealth are revealed. In the versions we consider, time and/or wealth dependence disappear from the optimal controls, but it is important to note that this independence is an endogenous feature of the solution and not exogenously assumed. Generalizations to other cases with dependence should naturally also be based on a dynamic perspective.

The literature on ruin probability minimization and linear dividend optimization until ruin via optimal reinsurance and investment is exhausting. We choose to mention only Schmidli (2002) as well as Schmidli (2008) for an overview. Instead we concentrate here on the somewhat smaller amount of literature where dividends are measured by their utility because the mathematical issues there are more closely related to ours. Hubalek and Schachermayer (2004) considered the problem of optimising power utility of dividend payments until ruin where the risk process is modelled by a Brownian motion with drift. Their insurance risk process itself is not influenced by the control as it is the case if one optimises over reinsurance decisions. Grandits et al. (2007) also considered the Brownian risk process but optimised over exponential utility of dividends. However, instead of taking exponential utility of dividend rates they measure exponential utility of aggregate dividends until ruin. Thonhauser and Albrecher (2011) considered the problem of optimising dividend payouts measured by power utility but with transaction costs related to payouts. Their insurance risk process is not influenced by the control. Common for all these problems is that the dividend payout scheme is the only decision process.

It has been argued that the dividend decision process of the insurance company should not be optimised with respect to a utility function as there is not one distinct individual whose utility function can appropriately represent the corporate decision process, see e.g. the survey paper Avanzi et al. (2016). Only from an individual's point of view, the optimization of utility of consumption financed by a wealth process influenced by non-life risk, partly mitigated by the purchase of insurance contracts, appears to be natural. And as such, it could also form the basis for designing insurance contracts. Namely, the optimal insurance contract is the one where adjustments according to characteristics of the individuals (age, wealth, ect.) is an integrated part of the product design, and must be a service provided by the insurance company according to some agreement between the company and the policy holders, in order to make the decision process as simple as possible for them. The worth of considering these issues are quantified by the welfare loss from offering standardised contracts instead of optimal ones. These are the ideas pursued in this paper.

The result that the fixed deductible contract is optimal for linear pricing is found in e.g. Arrow (1971). A series of papers work within this setup of linear pricing. Cummins and Mahul (2004) and Zhou et al. (2010), for example, implement an upper limitation on the insurance coverage of the fixed deductible contract. Golubin (2016) considers instead joint decisions to be made by both the insurance company and the individual. Aase (2017) considers a different approach by arguing how the presence of costs in insurance impacts the design of (Pareto) optimal insurance contracts. This is a certain form of non-linear pricing that also creates a demand for more general deductible structures.

Finally we relate our work to other works where the individual seeks to maximise utility of consumption or wealth from investment and/or non-life insurance decisions.

Yang and Zhang (2005) consider the investment problems in a jump-diffusion model for insurance risk but they control neither consumption nor insurance risk. Moore and Young (2006) study optimal consumption, investment, and insurance under a diffusive financial market and compound Poisson modelled insurance risk. Compared to that, Perera (2010) generalises both the financial market and the insurance risk model to a general Lévy framework whereas Zou and Cadenillas (2014) generalise to regime shifts in both market and insurance coefficients. Zhang and Siu (2009) control investment and insurance under model uncertainty. Other publications in the area typically generalise the financial market in which an investment decision is made or the preferences of the individual. All the references of this paragraph work with more general financial markets than we do, since we simply earn capital gains from deterministic interest and have no investment decision to make. However, for all the references of this paragraph where insurance is controlled, the pricing mechanism of the insurance market is linear. This makes it optimal, as is shown and used in the references, to buy an insurance product with, in general, a wealth dependent but more importantly in this relation a loss size independent deductible.

The investment decision in financial markets combined with linear insurance pricing in the references mentioned in the prior paragraph marks a clear difference compared to the scope of our work. We concentrate fully on the insurance market, look for the optimal insurance position under non-linear pricing, and quantify the financial sacrifice of being offered a flat deductible only. One may argue that since the non-flat deductible is not offered in the retail market, the approach taken by the references is “correct” and our approach is “useless”. However, both conclusions are false. It is appealing to think that since retail marketed contracts have flat deductibles, the pricing rule in this market is linear and all other decisions on consumption and investment should be made on that basis. This thinking is based on a blind belief in the market’s ability to develop optimal products. If, conversely, the market currently contains suboptimal products only, a couple of important questions arise. Based on the true pricing rules, whatever they are, what is then the optimal insurance decision and how are the consumption and investment decisions altered, respectively, compared to the case of linear pricing? Can the optimal insurance decision inspire to product development with a generalised deductible that actually does represent the market’s development of optimal products? And what is the value created to the individuals following from such development of optimal products? These are the type of questions we address in this exposition. So, the motivation is not to repair the decisions made by individuals but rather to repair the market she faces, or at least to start a discussion about whether and why there is something to repair.

In order to start out with explicit and tractable calculations in this direction of study we do make simplifying assumptions on financial markets and preferences as well as we skip considering the investment decision as an integrated problem. This, however, does not harm the principal discussion we start, the qualitative results that we obtain, or the illustrative power of our quantitative results. Whether our analysis suffers more or less from our simplifying assumptions about financial markets and preferences than the analysis in the references above suffer from simplifying assumptions about linear insurance pricing is unknown. But we conjecture that realistic modelling and controlling of financial risk is of second order importance compared to realistic insurance pricing when studying optimal control of insurance risk and optimal design of insurance contracts.

Among the more restrictive assumptions we make, we highlight already here a few. We deal with a marginal problem and not an equilibrium problem. This means that the insurance company does not change the pricing rule (of course, it changes the price itself) depending on decision made by the individual. The alternative, namely to construct a game, would be much more difficult. Further, we consider specific pricing rules. The rules we consider are, however, well-known and generally accepted for their relevance. In the numerical section, the coefficients within the pricing rules are chosen to be realistic, but are not calibrated to any data or price observation, though. Finally, we take the calculations all the way to the end for the case of exponentially distributed losses. This is clearly very restrictive, but this is just to reach fully explicit results in this particular case. We exploit these results in our numerical examples. The assumptions limit our quantitative conclusions to the cases considered. However, they do not limit the outreach of the qualitative discussion about sub-optimality of (realistic) flat deductibles under (realistic) non-linear pricing which is the very motivation for this exposition.

The outline of the paper is as follows. Section 2 explains how insurance affects the individual and her wealth. Section 3 creates a general view of the optimization problem of the individual. Section 4 introduces in details the mechanics of pricing by changing measure. Section 5 contains the explicit expressions needed to find the welfare loss for the two insurance products in consideration, namely the one with a fully flexible coverage and the one with a fixed amount deductible. Section 6 makes a numerical comparison by illustrating the welfare loss.

E.2 Claims process and insurance contracts

We consider an individual endowed with the initial wealth w. The individual consumes at rate $(c_t)_{t\geq 0}$ and all excess wealth is invested into a risk-free asset with interest rate r. The individual is exposed to a risk that can be modelled by a compound Poisson process $(A_t)_{t\geq 0}$ with parameters (λ, F), i.e.

$$A_t = \sum_{i=1}^{N_t} Z_i,$$

where $(N_t)_{t\geq 0}$ is a Poisson process with parameter λ counting the number of losses until time t and the Z_i's represent the (positive) loss sizes assumed to be i.i.d. and independent of $(N_t)_{t\geq 0}$ with distribution F on $(0, \infty)$. Note that we speak of losses rather than claims as the analysis is performed on an individual level. The losses Z_i are actually the expenses of the individual connected with 'insurable but not yet insured' event number i. Thus, it has not become a claim from a policy holder upon an insurance company yet. Before purchasing insurance the wealth of the individual, denoted by $(w_t)_{t\geq 0}$, develops in accordance with

$$\mathrm{d}w_t = (rw_t - c_t)\,\mathrm{d}t - \mathrm{d}A_t.$$

This wealth process of the individual is modelled similarly to a classical surplus process in risk theory with deterministic capital gains and absolutely continuous dividend payments in terms of consumption. It is conventional to think of this as the surplus process for a portfolio of insurances within an insurance company. Here we interpret the process as the wealth process of an individual with event risk modelled by the compound Poisson process.

The individual can reduce and manage the risk of $(A_t)_{t\geq 0}$ by purchasing insurance. An insurance product is described by a non-decreasing function $g_v : \mathbb{R}^+ \to \mathbb{R}^+$ parameterised by a control v. The function g_v is applied to individual losses in the sense that for a claim of size z, $g_v(z)$ is the deductible that the individual pays herself. Thus, when choosing the insurance contract g_v, the expense of the individual is reduced to $\min\{z, g_v(z)\}$ in relation to a loss z. When the loss $z > g_v(z)$, the loss is reported to the insurance company and the excess $z - g_v(z)$ is claimed and covered.

A standard insurance contract has a fixed amount deductible characterised by g being constant, i.e. $g_K(z) = K$, where $K \in \mathbb{R}^+$ is the deductible level that also parameterises the insurance decision. The individual who we then speak of as a policyholder, reports a claim if the occurred loss z exceeds the fixed amount deductible K. So the policyholder covers the loss $\min\{z, K\}$ by herself, and the rest of the loss, namely $(z - K)^+$, is covered by the insurance company. Our idea is to leave such a restricted class of strategies. The reader may think of a deductible which is more generally dependent on z and, possibly, t and w_t, such that a variable amount is covered by the insured depending on the size of the claim and, possibly, the time (age) and wealth of the policy holder.

Insurance is used by the policyholder as a tool to reduce her risk exposure $(A_t)_{t\geq 0}$ by reducing the size of the losses. Followed by the above-mentioned arguments, when employing insurance characterised by g_v, the policy holder's reduced risk can be represented by the compound Poisson process $(C_{v,t})_{t\geq 0}$ with jump rate λ and claim sizes $(\min\{Z_i, g_v(Z_i)\})_{i=1,2,\dots}$, i.e. $C_{v,t} = \sum_{i=1}^{N_t} \min\{Z_i, g_v(Z_i)\}$. The compound Poisson process $(A_{v,t})_{t\geq 0}$ with $A_{v,t} = A_t - C_{v,t} = \sum_{i=1}^{N_t}(Z_i - g_v(Z_i))^+$ then represents the risk transferred to the insurance company for which the policy holder must pay a premium.

Throughout we adopt the idea of evaluating the premium of an insurance contract by a change of measure. Pricing by a change of measure is mostly considered to be a financial notion (also known as risk neutral pricing). Embrechts (2000) provides a treatment of the link between financial pricing and actuarial pricing. We restrict the focus to the equivalent measures $\mathbb{Q}$ such that the accumulated claim process $(A_{v,t})_{t\geq 0}$ remains a compound Poisson process under $\mathbb{Q}$, but where the characteristics are altered to $(\lambda^{\mathbb{Q}}, F^{\mathbb{Q}})$. We present here how the accumulated premium until time t then can be reduced to a premium density,

$$\mathbb{E}^{\mathbb{Q}}[A_{v,1}] = \mathbb{E}^{\mathbb{Q}}\Big[\sum_{i=1}^{N_1}(Z_i - g_v(Z_i))^+\Big] = \lambda^{\mathbb{Q}}\,\mathbb{E}^{\mathbb{Q}}\left[(Z - g_v(Z))^+\right], \tag{4.1}$$

where Z is an independent copy of $Z_1, Z_2, \ldots$. Since this change of measure is a rather technical concept, we devote Section E.4 to elaborate on the needed ingredients. For now, what is important for us is that we can decompose the measure transformation into its jump part, changing the intensity from λ to $\lambda^{\mathbb{Q}}$, and claim size part, changing the distribution of Z from F to $F^{\mathbb{Q}}$, due to the compound Poisson properties being maintained under $\mathbb{Q}$. The individual therefore must pay a premium rate continuously that depends on the compound Poisson characteristics $(\lambda^{\mathbb{Q}}, F^{\mathbb{Q}})$ under $\mathbb{Q}$, if she wants to buy insurance. In the view of actuarial pricing, this approach is quite general and contains, as special examples, certain versions of the expectation premium principle, a variance premium principle, and an Esscher premium principle, as will be seen in Section E.4.

Purchasing insurance then has the following effect on the dynamics of the individual's wealth process

$$\begin{aligned} \mathrm{d}w_t &= (rw_t - c_t)\,\mathrm{d}t - \mathrm{d}\left(\mathbb{E}^{\mathbb{Q}}[A_{v,t}]\right) - \mathrm{d}C_{v,t} \\ &= \left(rw_t - c_t - \lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}\left[(Z - g_v(Z))^+\right]\right)\mathrm{d}t - \mathrm{d}C_{v,t}. \end{aligned} \tag{4.2}$$

With these dynamics, the infinitesimal operator of a function $f(w)$ is

$$\begin{aligned} \mathcal{A}f = \left(rw - c_t - \lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}\left[(Z - g_v(Z))^+\right]\right)\frac{\partial f}{\partial w}(w) \\ + \lambda\mathbb{E}[f(w - \min\{Z, g_v(Z)\}) - f(w)]. \end{aligned} \tag{4.3}$$

E.3 Optimization problem of the individual

At time t the individual chooses the consumption rate and the insurance strategy in terms of g_v in order to optimise her expected discounted utility of consumption,

$$V(t,w) = \sup_{c,v}\mathbb{E}_{t,w}\left[\int_t^\infty \exp\bigl(-\rho(s-t)\bigr)u(c_s)\mathrm{d}s\right],$$

where $\mathbb{E}_{t,w}$ denotes conditional expectation given that $w_t = w$. The utility function u measures utility from the consumption rate c and ρ is a subjective utility discount factor. We do not impose any control constraints on consumption. A natural (and common) restriction would be $c \geq 0$, but as we prioritise finding a tractable solution, we choose to look past this. The insurance control v, on the other hand, must satisfy that the deductible strategy g_v is non-decreasing and non-negative (as it was defined). We only put a lower limit on the deductible strategy as we avoid any issues with upper limitations by using truncation and minimum in the dynamics of the wealth (4.2). Hence, if a deductible strategy exceeds the actual loss, then this automatically corresponds to having no insurance.

Due to the time-homogeneity of all ingredients of the state process, i.e. the coefficients of $(w_t)_{t\geq 0}$, and the objective, i.e. (ρ, u), the value function is a function of wealth only and we can write

$$V(w) = \sup_{c,v}\mathbb{E}_w\left[\int_0^\infty \exp(-\rho t)u(c_t)\mathrm{d}t\right]. \tag{4.4}$$

The rate ρ is here called the utility discount rate. A different interpretation is the mortality rate of an individual optimising her utility of consumption until death. Then, if an expectation is taken both with respect to time of death and insurance risk, the expectation with respect to time of death gives a survival probability until time t of $\exp(-\rho t)$ and expectation with respect to both risks then leads to (4.4). It is of course a non-realistic restriction, in that interpretation, to work with an age-independent mortality rate. However, for now we work out the details for the time-homogeneous case and it is beyond the scope of this presentation to handle the time-inhomogeneous case. The Hamilton-Jacobi-Bellman (HJB) equation characterising the value function is given by

$$\sup_{c,v}\{\mathcal{A}V + u(c)\} = \rho V(w).$$

Using the infinitesimal operator in (4.3) the HJB equation can be written more explicitly as

$$\begin{aligned}\sup_{c,v}\Big\{&-\rho V(w) + (rw - c - \lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}[(Z - g_v(Z))^+])V_w(w)\\ &+ \lambda\mathbb{E}[V(w - \min\{g_v(Z), Z\}) - V(w)] + u(c)\Big\} = 0,\end{aligned} \tag{4.5}$$

where $V_w(w) = \frac{\partial V}{\partial w}(w)$.

For the special case of an exponential utility function, we can immediately learn something about the structure of V. Since this case plays a crucial role in our considerations we present at this point these principal observations.

Proposition E.1. *Assume that the utility function is of the form*

$$u(c) = \frac{-1}{a}\exp(-ac). \tag{4.6}$$

for $a > 0$. Then, for a sufficiently regular function $g_v(z)$ (in both z and the parameter v), the value function (4.4) can be written by

$$V(w) = \frac{-1}{\alpha}\exp(-raw). \tag{4.7}$$

The optimal consumption c^* is affine in wealth

$$c^* = rw - \frac{1}{a}\log\left(\frac{ra}{\alpha}\right), \tag{4.8}$$

and the optimal insurance control v^* solves

$$\lambda\frac{\partial}{\partial v}\Big(\mathbb{E}[\exp(ra\cdot\min\{Z, g_v(Z)\})]\Big) = ra\lambda^{\mathbb{Q}}\frac{\partial}{\partial v}\left(\mathbb{E}^{\mathbb{Q}}[(Z - g_v(Z))^+]\right), \tag{4.9}$$

and is thus independent of wealth. The parameter α of the value function is determined by the relation

$$\begin{aligned}\alpha = ra\exp\Big(\frac{1}{r}\big(\rho - \lambda\big(\mathbb{E}[\exp(ra\cdot\min\{Z, g_{v^*}(Z)\})] - 1\big)\big)\\ - a\lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}[(Z - g_{v^*}(Z))^+]\Big).\end{aligned} \tag{4.10}$$

Proof. Consider the first order conditions of (4.5) with respect to consumption

$$-V_w(w) + \exp(-ac) = 0,$$

and with respect to deductible

$$-\frac{\partial}{\partial v}\Big(\lambda^{\mathbb{Q}}\mathbb{E}^{\mathbb{Q}}[(Z - g_v(Z))^+]\Big)V_w(w) + \lambda\mathbb{E}[V(w - \min\{Z, g_v(Z)\})]\Bigg) = 0.$$

Note that in the case where the parameter v is multidimensional, the derivative would be replaced by the gradient, which would yield the same number of first order conditions as dimensions of v.

Conjecture that the solution to the HJB equation is of type $V(w) = \frac{-1}{\alpha}\exp(-raw)$. The first order condition with respect to the consumption is then

$$-\frac{ra}{\alpha}\exp(-ra \cdot w) + \exp(-ac) = 0.$$

Reducing this leads to the optimal consumption

$$c^* = rw - \frac{1}{a}\log\left(\frac{ra}{\alpha}\right),$$

Correspondingly for the first order condition with regards to the insurance control,

$$\begin{aligned} -\lambda^{\mathbb{Q}}\frac{\partial}{\partial v}\Big(\mathbb{E}^{\mathbb{Q}}[(Z - g_v(Z))^+]\Big)\frac{ra}{\alpha}\exp(-ra \cdot w) \\ + \frac{\lambda}{\alpha}\frac{\partial}{\partial v}\Big(\mathbb{E}[\exp(-ra \cdot (w - \min\{Z, g_v(Z)\}))]\Big) = 0 \end{aligned} \tag{4.11}$$

and the optimal deductible strategy, v^*, must therefore be a solution to

$$\lambda\frac{\partial}{\partial v}\Big(\mathbb{E}[\exp(ra \cdot \min\{Z, g_v(Z)\})]\Big) = ra\lambda^{\mathbb{Q}}\frac{\partial}{\partial v}\Big(\mathbb{E}^{\mathbb{Q}}[(Z - g_v(Z))^+]\Big).$$

Although not solved explicitly, it is clear that the first order condition for the deductible strategy does not depend on wealth. Regularity of $g_v(z)$ is assumed in the proposition to ensure existence and uniqueness of (4.9). Exact sufficient and necessary conditions are not studied here.

When substituting these optimal values back into the HJB equation (4.5), the supremum will be obtained and we can solve for (4.10), verifying that the initial guess for the structure of the value function was correct, since α does not depend on wealth. □

From (4.8) we learn that the individual stops consuming, $c^* \to 0$, when the wealth approaches some lower level, $w \to \log(ra/\alpha)/ra$. Only the insurable losses will therefore cause the wealth to decrease below this level yielding negative consumption. In relation to this, it is natural to point out that the wealth is allowed to become negative also, but the controlled wealth grows (on average) linearly in time since the optimal insurance control is independent of wealth and the optimal consumption is linear in wealth with a factor equal to the risk free interest rate. Hence, over-accumulation of wealth or long-term bankruptcy is not an issue here. It can not be verified in general that the insurance control chosen by (4.9) satisfies that g_{v^*} is non-decreasing and non-negative, so this must be taken into consideration for the specific structures of g_v and pricing measures $\mathbb{Q}$.

We are going to compare the performance of the optimal decisions (c^*, v^*) with alternative decisions, in particular with suboptimal choices of v. For this purpose we characterise the solution to the problem (4.4) where supremum is taken over c only, i.e.

$$V(w) = \sup_c \mathbb{E}_w \left[\int_0^\infty \exp(-\rho t) u(c_t) \mathrm{d}t \right], \tag{4.12}$$

such that the value function is characterised by the HJB equation

$$\sup_c \{\mathcal{A}V + u(c)\} = \rho V(w).$$

Going through the relevant steps in the proof of Proposition E.1, one can immediately see that the structure of the solution in the exponential utility case is preserved. The value function has again the structure (4.7) with a different α determined by (4.10) with v^* replaced by a given suboptimal deductible. It should be mentioned that any disentanglement of the optimization over c and v is of course true only if the deductible does not depend on wealth. Dependence on wealth, in case of other utility functions, introduces interdependence between consumption and insurance decisions.

We can now compare a suboptimal insurance decision with the optimal one by comparing the two value functions arising from optimising over (c, v) and from optimising over c for a given suboptimal v, respectively. Let us denote by $V^{(1)}$ the value function arising from optimising over (c, v) giving rise to the coefficient α_1 and denote by $V^{(2)}$ the value function arising from optimising over c for a given v giving rise to the coefficient α_2. It is clear that $V^{(1)} \geq V^{(2)}$ and $\alpha_1 \geq \alpha_2$, since the optimal insurance decision beats the suboptimal one. In order to compare, it is standard to transform the difference into monetary units by calculating the so-called certainty equivalent of the loss of utility from implementing the suboptimal insurance decision. This is the solution L to the utility indifference equality

$$V^{(1)}(w - L) = V^{(2)}(w).$$

The idea of utility indifference is a generally accepted way to measure sub-optimality. One may first think of calculating $V^{(1)}(w) - V^{(2)}(w)$ as a measure of sub-optimality. However, the nominal value of the value function does not have a meaning on its own and therefore this difference contains no other information than which one is preferred to the other due to its sign. But how should one then measure e.g. significance of the difference? This is exactly what the utility indifference equality does since when the difference between the value functions is 0, then this nominal value has a clear economic interpretation of the indifference. Thus, the equality translates the difference between value functions, which is non-informative, into an informative quantity in monetary units. It answers the question, what is the loss in monetary units from having access to only a suboptimal control compared to being offered the optimal control?

It is immediately seen that for the exponential utility function, the solution is

$$L = \frac{1}{ra} \log\left(\frac{\alpha_1}{\alpha_2}\right). \tag{4.13}$$

We are ultimately interested in determining this L for different suboptimal choices of v, for different pricing measures $\mathbb{Q}$, and for different distributions of Z in order to get a general view on the loss of not having access to the optimal insurance contract.

E.4 Pricing by change of measure

As briefly explained in connection to (4.1), we want to price insurance by a change of measure. In a risk averse setting, the change of measure has the advantage that one can give more weight to bad outcomes and hence assign a higher price to larger risks. A simple example on how this affects the individual's attitude towards insurance coverage is presented in Appendix E.I.

Let the physical measure be denoted as $\mathbb{P}$. Recall that under this measure the characteristics of the compound Poisson risk process $(A_t)_{t\geq 0}$ can be summarised by (λ, F). Let $\beta : \mathbb{R}^+ \to \mathbb{R}$ be a Borel measurable mapping, henceforth referred to as the pricing measure function, satisfying $\mathbb{E}[\exp(\beta(Z))] < \infty$. Delbaen and Haezendonck (1989) define the Radon-Nikodym derivative

$$M_t^\beta = \exp\Big(\sum_{i=1}^{N_t} \beta(Z_i) - \lambda t \mathbb{E}[\exp(\beta(Z)) - 1]\Big). \tag{4.14}$$

Let $(\mathcal{F}_t)_{t\geq 0}$ denote the filtration generated by the compound Poisson process $(A_t)_{t\geq 0}$. Delbaen and Haezendonck (1989) argues that the measure $\mathbb{Q}(\beta)$ defined by $M_t^\beta = \mathbb{E}[\,\mathrm{d}\mathbb{Q}(\beta)/\mathrm{d}\mathbb{P} \mid \mathcal{F}_t]$ satisfies that

- $\mathbb{Q}(\beta)$ and $\mathbb{P}$ are progressively equivalent, i.e. have the same null sets.
- $(A_t)_{t\geq 0}$ is a $\mathbb{Q}(\beta)$-compound Poisson process with characteristics $(\lambda^{\mathbb{Q}(\beta)}, F^{\mathbb{Q}(\beta)})$ given by

$$\lambda^{\mathbb{Q}(\beta)} = \lambda \mathbb{E}[\exp(\beta(Z))], \quad F^{\mathbb{Q}(\beta)}(\mathrm{d}z) = \frac{\exp(\beta(z))}{\mathbb{E}[\exp(\beta(Z))]} F(\mathrm{d}z). \tag{4.15}$$

  So $\mathbb{E}[\exp(\beta(Z))]$ can be considered as a penalty for claim frequency risk and $\exp(\beta(z))/\mathbb{E}[\exp(\beta(Z))]$ as a penalty for claim size risk.

A premium rate of the risk $(A_t)_{t\geq 0}$ can then be defined as

$$\mathbb{E}^{\mathbb{Q}(\beta)}[A_1] = \mathbb{E}^{\mathbb{Q}(\beta)}\Big[\sum_{i=1}^{N_1} Z_i\Big] = \lambda \mathbb{E}[\exp(\beta(Z))Z], \tag{4.16}$$

where we take the expected value under the pricing measure $\mathbb{Q}(\beta)$ which maintains the compound Poisson structure, but changes the claim frequency and claim size distribution (cf. the second property above).

In Delbaen and Haezendonck (1989) it is argued how certain choices of $\beta(\cdot)$ correspond to certain well-known premium principles. Three choices of the pricing measure function are considered. We recapitulate briefly the findings here.

(1) Constant: $\beta(z) = \delta$. The premium rate (4.16) then reduces to

$$\mathbb{E}^{\mathbb{Q}(\beta)}[A_1] = \lambda \exp(\delta)\mathbb{E}[Z].$$

Only claim frequency risk is penalised, whereas claim size risk is not priced. Since the pricing measure function is constant, the price is linear in the expectation under $\mathbb{P}$, and we are in the case of the expected value premium principle. This corresponds to the case exemplified also at the end of Section E.I with

$$\lambda^{\mathbb{Q}} = \lambda \exp(\delta).$$

(2) Log-linear: $\beta(z) = \log(\theta z + \delta)$. If $\delta = 1 - \theta\mathbb{E}[Z] > 0$ implicating $\mathbb{E}[\exp(\beta(Z))] = 1$, then it appears from (4.16) that

$$\mathbb{E}^{\mathbb{Q}(\beta)}[A_1] = \lambda(\mathbb{E}[Z] + \theta\mathbb{V}[Z]).$$

This is the reversed situation compared to (1), since here the claim frequency risk is not priced whereas claim size risk is priced using the variance premium principle. If $\delta = 1$, the premium principle corresponds to the variance principle used on the total claim amount since

$$\mathbb{E}^{\mathbb{Q}(\beta)}[A_t] = \lambda t\mathbb{E}[(\theta Z^2 + Z)] = \mathbb{E}[A_t] + \theta\mathbb{V}[A_t].$$

(3) Linear: $\beta(z) = \theta z + \delta$. If $\delta = -\log(\mathbb{E}[\exp(\theta Z)])$ it once again implicates that $\mathbb{E}[\exp(\beta(Z))] = 1$. From (4.16) it follows that

$$\mathbb{E}^{\mathbb{Q}(\beta)}[A_1] = \lambda\frac{\mathbb{E}[Z\exp(\theta Z)]}{\mathbb{E}[\exp(\theta Z)]}.$$

As in (2) claim frequency risk has no penalty, and the claim sizes are priced according to the Esscher premium principle.

In Delbaen and Haezendonck (1989), the variable Z_i is the claim on the insurance company from insurance event number i. Recall that here, Z_i is the true loss whereas only a part of this is claimed on the insurance company. The individual does not necessarily buy insurance protection for the entire underlying risk $(A_t)_{t\geq 0}$. Instead, insurance splits the losses, and thus the risk, in two parts, one covered by the policy holder $(C_{v,t})_{t\geq 0}$ and one by the insurance company $(A_{v,t})_{t\geq 0}$, when the insurance product is characterised by g_v. If the same approach as in Delbaen and Haezendonck

(1989) was directly followed, then the premium of the partial risk $(A_{v,t})_{t\geq 0}$ would have to be altered as follows,

$$\mathbb{E}^{\mathbb{Q}(\beta_v)}[A_{v,1}] = \mathbb{E}^{\mathbb{Q}(\beta_v)}\Big[\sum_{i=1}^{N_1}(Z_i - g_v(Z_i))^+\Big] = \lambda\mathbb{E}[\exp(\beta_v(Z))(Z - g_v(Z))^+],$$

where $\beta_v(z) = \beta((z - g_v(z))^+)$. Note that the insurance control v then appears as an endogenous part of the pricing measure function. This considerably complicates the first order condition (4.9) for the insurance control in the policy holder's optimization problem. The result would be a kind of equilibrium insurance strategy. We wish to work within a marginal approach where we separate the pricing measure function from the control problem by making it independent of the insurance control v. We use the change of measure defined by (4.14) applied to the full claim to price the partial risk covered by the insurance company. The following proposition verifies that this is still a meaningful change of measure.

Proposition E.2. *Let compound Poisson characteristics of $(A_{v,t})_{t\geq 0}$ be summarised by (λ, F_v). Under $\mathbb{Q}(\beta)$ the process $(A_{v,t})_{t\geq 0}$ is still a compound Poisson process, but with altered characteristics $(\lambda^{\mathbb{Q}(\beta)}, F_v^{\mathbb{Q}(\beta)})$ where*

$$\lambda^{\mathbb{Q}(\beta)} = \lambda\mathbb{E}[\exp(\beta(Z))], \quad F_v^{\mathbb{Q}(\beta)}(dz) = \frac{\exp(\beta(z))}{\mathbb{E}[\exp(\beta(Z))]}F_v(dz). \tag{4.17}$$

The premium rate is

$$\mathbb{E}^{\mathbb{Q}(\beta)}\left[A_{v,1}\right] = \lambda\mathbb{E}[\exp(\beta(Z))(Z - g_v(Z))^+]. \tag{4.18}$$

The proof appears in Appendxxxxix E.II. The intuition is that even though claim sizes have been transformed by h_v, it is the change of characteristics of the underlying risk $(A_t)_{t\geq 0}$ that is determined by $\mathbb{Q}(\beta)$. The same choices (1)-(3) of β are considered in the following sections, but with different restrictions on the parameters. In general, we are interested in parameters δ and θ such that β is non-decreasing and positive, hence we consider the pricing measure functions

(1) $\beta_1(z) = \delta_1$, where $\delta_1 \in \mathbb{R}^+$.

(2) $\beta_2(z) = \log(\theta_2 z + \delta_2)$, where $(\theta_2, \delta_2) \in \mathbb{R}^+ \times [1, \infty)$.

(3) $\beta_3(z) = \theta_3 z + \delta_3$, where $(\theta_3, \delta_3) \in \mathbb{R}^+ \times \mathbb{R}^+$.

Further restrictions on the parameter values might be imposed in subsequent analysis when necessary. Note that the link from (4.16) to various premium principles does not hold for (4.18) though, since the change of measure no longer relates to the claim size but the entire loss.

E.5 Insurance products

The results in Appendix E.I show that a constant deductible is not optimal in the case of non-linear pricing, i.e. pricing where $\lambda_i^{\mathbb{Q}}/\lambda_i$ is not constant in i. This means that there is a welfare loss connected with having access to only constant deductibles in the market. The welfare loss depends, of course, on the extent and shape of the non-linearity. In section E.4 we have introduced a family of non-linear pricing rules, and we are now going to measure the welfare loss produced by these pricing rules. The welfare loss is measured by comparing the optimal insurance design corresponding to its full flexibility in the structure of the deductible, with the suboptimal constant deductible insurance contract corresponding to no flexibility in the structure of the deductible. To compare these, we use the loss (4.13), which also can be written as

$$\begin{aligned} L = \frac{1}{ra}\Big(& a\lambda(\mathbb{E}[\exp(\beta(Z))(Z - g_v^{(2)}(Z))^+] - \mathbb{E}[\exp(\beta(Z))(Z - g_v^{(1)}(Z))^+]) \\ & + \frac{\lambda}{r}(\mathbb{E}[\exp(ra\min\{Z, g_v^{(2)}(Z)\})] - \mathbb{E}[\exp(ra\min\{Z, g_v^{(1)}(Z)\})])\Big). \end{aligned}$$

Hence for two insurance products (full and no flexibility) and for each pricing measure function $\beta(\cdot)$ we need to calculate the terms

$$\begin{aligned} p_\beta^{(i)} &= \mathbb{E}[\exp(\beta(Z))(Z - g_v^{(i)}(Z))^+] \\ &= \mathbb{P}(Z > g_v^{(i)}(Z))\mathbb{E}[\exp(\beta(Z))(Z - g_v^{(i)}(Z)) \mid Z > g_v^{(i)}(Z)], \\ q_\beta^{(i)} &= \mathbb{E}[\exp(ra\min\{Z, g_v^{(i)}(Z)\})] \\ &= \mathbb{P}(Z \le g_v^{(i)}(Z))\mathbb{E}[\exp(raZ) \mid Z \le g_v^{(i)}(Z)] \\ &\quad + \mathbb{P}(Z > g_v^{(i)}(Z))\mathbb{E}[\exp(rag_v^{(i)}(Z)) \mid Z > g_v^{(i)}(Z)]. \end{aligned} \tag{4.19}$$

for $i = 1, 2$. Recall that $i = 1$ corresponds to the completely flexible insurance product, where the deductible strategy is unparameterised, i.e. $g_v^{(1)}(z) = g(z)$, and $i = 2$ corresponds to the fixed amount deductible insurance product, hence $g_v^{(2)}(z) = K$.

The aim of this section is not to obtain a closed form expression of welfare loss L, but to find the ingredients (4.19) for the different β-functions of interest. The welfare loss is then illustrated by an numerical example in the next section. The relations in (4.19) depend indeed on the distribution of the claim sizes. In order to obtain tractable expressions, we assume that the claim sizes are exponentially distributed with parameter η, i.e. $F(z) = 1 - \exp(-\eta z)$. Only the results will be stated here, for calculations of this section see appendix.

E.5.1 Completely flexible

Assume that the individual can choose a deductible strategy freely as a function of the loss. She is then facing the problem of solving

$$\begin{aligned} \sup_{c, g(\cdot)} \{ & -\rho V(w) + (rw - c - \lambda\mathbb{E}[\exp(\beta(Z))(Z - g(Z))^+])V_w(w) \\ & + \lambda(\mathbb{E}[V(w - \min\{Z, g(Z)\})] - V(w)) + u(c)\} = 0. \end{aligned} \tag{4.20}$$

The first order condition with respect to $g(\cdot)$ is

$$\lambda\mathbb{E}[\exp(\beta(Z))\mathbb{1}_{Z\geq g(Z)}]V_w(w) = \lambda\mathbb{E}[V_w(w-g(Z))\mathbb{1}_{Z\geq g(Z)}].$$

Recall that the value function has the form $V(w) = \exp(-raw)/\alpha$ for an appropriate α, hence the optimal unrestricted insurance strategy, $g^*(\cdot)$, must satisfy

$$\mathbb{E}[\exp(\beta(Z))\mathbb{1}_{Z\geq g(Z)}] = \mathbb{E}[\exp(ra\cdot g(Z))\mathbb{1}_{Z\geq g(Z)}].$$

These moments are matched when $g^*(z) = \beta(z)/(ra)$. We consider the three choices of the pricing measure function β introduced in the end of Section 5. The index of β is reflected in the index of the optimal insurance strategy g^*. Remember that g is unparameterised in this section, so subscripts of the optimal strategy g^* refer to the pricing measure function and not the parameterization of g.

(1) For a constant pricing measure function $\beta_1(z) = \delta_1$, the optimal insurance strategy for the individual is the fixed amount deductible, $g_1^*(z) = \delta_1/(ra)$. This was anticipated due to the preliminary analysis of Section E.I. Let $A_1 = \{Z < \delta_1/(ra)\}$. The two insurance dependent terms (4.19) of L can then be written as

$$\begin{aligned}
p^{(2)}_{\beta_1} &= \mathbb{E}[\exp(ra\cdot\min\{Z, g_1^*(Z)\})] = \mathbb{E}[\exp(ra\cdot\min\{Z,\delta_1/(ra)\})]\\
&= \mathbb{P}(A_1)\mathbb{E}[\exp(raZ)\mid A_1] + \mathbb{P}(A_1^c)\exp(\delta_1),\\
q^{(2)}_{\beta_1} &= \mathbb{E}[\exp(\beta_1(Z))(Z-g_1^*(Z))^+] = \mathbb{E}[\exp(\delta_1)(Z-\delta_1/(ra))^+]\\
&= \mathbb{P}(A_1^c)\exp(\delta_1)\mathbb{E}[Z-\delta_1/(ra)\mid A_1^c].
\end{aligned}$$

Assuming that Z is exponentially distributed with parameter η, then we can write these more explicitly as

$$\begin{aligned}
p^{(2)}_{\beta_1} &= \frac{ra}{ra-\eta}\exp\left(\left(\frac{ra-\eta}{ra}\right)\delta_1\right) - \frac{\eta}{ra-\eta},\\
q^{(2)}_{\beta_1} &= \frac{1}{\eta}\exp\left(\left(\frac{ra-\eta}{ra}\right)\delta_1\right).
\end{aligned}$$

(2) If the pricing measure function is log-linear, $\beta_2(z) = \log(\theta_2 z + \delta_2)$, then the optimal insurance strategy is a logarithmic-power deductible, $g_2^*(z) = \log(\theta_2 z+\delta_2)/(ra)$. Let $A_2 = \{Z < \log(\theta_2 Z+\delta_2)/(ra)\}$. The terms (4.19) are then defined by

$$\begin{aligned}
p^{(2)}_{\beta_2} &= \mathbb{E}[\exp(ra\cdot\min\{Z, g_2^*(Z)\})]\\
&= \mathbb{E}[\exp(ra\cdot\min\{Z,\log(\theta_2 Z+\delta_2)/(ra)\})]\\
&= \mathbb{P}(A_2)\mathbb{E}[\exp(raZ)\mid A_2] + \mathbb{P}(A_2^c)\mathbb{E}[\exp(\theta Z+\delta_2)\mid A_2^c],\\
q^{(2)}_{\beta_2} &= \mathbb{E}[\exp(\beta_2(Z))(Z-g_2^*(Z))^+]\\
&= \mathbb{E}[(\theta_2 Z+\delta_2)(Z-\log(\theta_2 Z+\delta_2)/(ra))^+]\\
&= \mathbb{P}(A_2)\mathbb{E}[(\theta_2 Z+\delta_2)(Z-\log(\theta_2 Z+\delta_2)/(ra))\mid A_2^c].
\end{aligned}$$

The challenge here is to evaluate the event A_2. Calculations in Appendix E.IV show that

$$A_2 = \{Z < Q\} \qquad \text{where } Q = -\frac{1}{ra}\mathcal{W}_{-1}\left(-\frac{ra}{\theta_2}\exp\left(-\delta_2\frac{ra}{\theta_2}\right)\right) - \frac{\delta_2}{\theta_2}.$$

using the lower branch $\mathcal{W}_{-1}$ of the Lambert $\mathcal{W}$ function defined as the inverse of the function $w \mapsto w\exp(w)$.

For an exponential distribution of the claim sizes, we then have

$$\mathbb{P}(A_2) = \mathbb{P}(Z < Q) = 1 - \exp(-\eta Q),$$

which yields

$$\begin{aligned}
p^{(2)}_{\beta_2} &= \frac{\eta}{ra-\eta}(\exp((ra-\eta)Q)-1) + \left(\theta_2 Q + \delta_2 + \frac{\theta_2}{\eta}\right)\exp(-\eta Q),\\
q^{(2)}_{\beta_2} &= \exp(-\eta Q)\Bigg(\theta_2 Q^2 + \left(\delta_2 + 2\frac{\theta_2}{\eta}\right)Q - \frac{1}{\eta}(\theta_2-\delta_2) + 2\frac{\theta_2}{\eta^2}\\
&\qquad - \frac{1}{ra}\frac{\theta_2}{\eta}\exp\left(\frac{\eta}{\theta_2}(\theta_2 Q+\delta_2)\right)E_1\left(\frac{\eta}{\theta_2}(\theta_2 Q+\delta_2)\right)\\
&\qquad - \frac{1}{ra}\left(\theta_2 Q + \delta_2 + \frac{\theta_2}{\eta}\right)\log(\theta_2 Q+\delta_2)\Bigg),
\end{aligned}$$

where E_1 denotes the exponential integral $E_1(x) = \int_x^\infty \exp(-t)/t\,\mathrm{d}t$.

(3) Employing a linear pricing measure function $\beta_3(z) = \theta_3 z + \delta_3$, then the individual optimally chooses a proportional insurance strategy $g_3^*(z) = (\theta_3 z + \delta_3)/(ra)$. Letting $A_3 = \{Z < (\theta_3 Z + \delta_3)/(ra)\} = \{Z < \delta_3/(ra-\theta_3)\}$, then we have

$$\begin{aligned}
p^{(2)}_{\beta_3} &= \mathbb{E}[\exp(ra\cdot\min\{Z, g_3^*(Z)\})]\\
&= \mathbb{E}[\exp(ra\cdot\min\{Z, (\theta Z+\delta)/(ra)\})]\\
&= \mathbb{P}(A_3)\mathbb{E}[\exp(raZ) \mid A_3] + \mathbb{P}(A_3^c)\mathbb{E}[\exp(\theta_3 Z+\delta_3) \mid A_3^c],\\
q^{(2)}_{\beta_3} &= \mathbb{E}[\exp(\beta_3(Z))(Z - g_3^*(Z))^+]\\
&= \mathbb{E}[\exp(\theta_3 Z+\delta_3)(Z - (\theta_3 Z+\delta_3)/(ra))^+]\\
&= \mathbb{P}(A_3^c)\mathbb{E}[\exp(\theta_3 Z+\delta_3)(Z - (\theta_3 Z+\delta_3)/(ra)) \mid A_3^c].
\end{aligned}$$

For an exponential distribution we can calculate these explicitly as

$$\begin{aligned}
p^{(2)}_{\beta_3} &= \left(1 - \exp\left(-(\eta - ra)\frac{\delta_3}{ra-\theta_3}\right)\right)\left(\frac{\eta}{\eta - ra}\right)\\
&\quad + \frac{\eta}{\eta-\theta_3}\exp\left(-(\eta-ra)\frac{\delta_3}{ra-\theta_3}\right),\\
q^{(2)}_{\beta_3} &= \exp\left(-(\eta-ra)\frac{\delta_3}{ra-\theta_3}\right)\frac{(ra-\theta_3)\eta}{ra(\eta-\theta_3)^2},
\end{aligned}$$

where it is assumed that $\theta_3 < \eta$ to ensure finite (left-truncated) exponential moments, and $\theta_3 < ra$ in order for A_3^c not to be a null set.

E.5.2 One level fixed amount deductible

Consider the fixed amount deductible where $g_K(z) = K$ is constant. The HJB equation is then

$$\begin{aligned}\sup_{c,K}\{&-\rho V(w) + (rw - c - \lambda \mathbb{E}[\exp(\beta(Z))(Z-K)^+])V_w(w)\\ &+ \lambda(\mathbb{E}[V(w - \min\{Z,K\})] - V(w)) + u(c)\} = 0.\end{aligned}$$

The first order condition for the fixed amount deductible level follows from using Leibniz integral rule on (4.9) yielding

$$\mathbb{E}[\exp(\beta(Z))\mathbb{1}_{\{Z\geq K\}}] = \exp(raK)\mathbb{E}[\mathbb{1}_{\{Z\geq K\}}].$$

Once again, we take a closer look at the three choices of the pricing measure function. Note that since it is the same insurance product in question, the second characterizations in (4.19), namely

$$\begin{aligned}q_\beta^{(1)} &= \mathbb{E}[\exp(ra\min\{Z,K\})]\\ &= \mathbb{P}(Z \leq K)\mathbb{E}[\exp(raZ) \mid Z \leq K] + \mathbb{P}(Z > K)\exp(raK),\end{aligned}$$

is the same function of the fixed deductible in every case of the pricing measure function. For an exponential distribution it can be calculated as in case (1) for the flexible insurance product,

$$q_\beta^{(1)} = \frac{ra}{ra-\eta}\exp((ra-\eta)K) - \frac{\eta}{ra-\eta}.$$

So the focus in the following is on characterising

$$\begin{aligned}p_\beta^{(1)} &= \mathbb{E}[\exp(\beta(Z))(Z-K)^+]\\ &= \mathbb{P}(Z > K)\mathbb{E}[\exp(\beta(Z))(Z-K) \mid Z > K].\end{aligned} \tag{4.21}$$

(1) For $\beta_1(z) = \delta_1$, the optimal deductible is obviously $K_1^* = \delta_1/(ra)$. Notice that this is the same structure as (1) in previous subsection, so $p_{\beta_1}^{(1)} = p_\beta^{(2)}$ and $q_\beta^{(1)} = q_\beta^{(2)}$ thus follow.

(2) Let $\beta_2(z) = \log(\theta_2 z + \delta_2)$. The optimal fixed deductible, K_2^*, must then satisfy

$$\mathbb{E}[(\theta_2 Z + \delta_2)\mathbb{1}_{\{Z\geq K_2\}}] = \exp(raK_2)\mathbb{E}[\mathbb{1}_{\{Z\geq K_2\}}].$$

For the exponential distribution, we can write this criteria more explicitly as

$$\theta_2\left(K_2 + \frac{1}{\eta}\right) + \delta_2 = \exp(raK_2),$$

which can be solved using the Lambert W function,

$$K_2^* = \frac{-1}{ra}W_{-1}\left(-\frac{ra}{\theta_2}\exp\left(-ra\left(\frac{\delta_2}{\theta_2}+\frac{1}{\eta}\right)\right)\right) - \frac{1}{\eta} - \frac{\delta_2}{\theta_2}.$$

Here we use that $-ra\exp(-ra(\delta_2/\theta_2+1/\eta))/\theta_2 \in [-\exp(-1), 0)$ according similar arguments as in Appendix E.IV. (4.21) can for an exponential distribution be calculated as

$$\begin{aligned} p^{(1)}_{\beta_2} &= \mathbb{E}[\exp(\beta_2(Z))(Z-K_2^*)^+] = \mathbb{E}[(\theta_2 Z+\delta_2)(Z-K_2^*)^+] \\ &= \left(\frac{\theta_2 K_2^*+\delta_2}{\eta}+\frac{2\theta_2}{\eta^2}\right)\exp(-\eta K_2^*). \end{aligned}$$

(3) Let $\beta_3(z) = \theta_3 z + \delta_3$. Then the optimal fixed amount deductible level, K_3^*, satisfies

$$\mathbb{E}[\exp(\theta_3 Z+\delta_3)\mathbb{1}_{\{Z\geq K_3\}}] = \exp(raK_3)\mathbb{E}[\mathbb{1}_{\{Z\geq K_3\}}].$$

For exponential loss distribution, the optimal deductible can be solved explicitly as

$$K_3^* = \frac{1}{ra-\theta_3}\left(\log\left(\frac{\eta}{\eta-\theta_3}\right)+\delta_3\right),$$

which exists and is positive for $ra > \theta_3$ and $\eta > \theta_3$. Furthermore, (4.21) can be expressed as

$$\begin{aligned} p^{(1)}_{\beta_3} &= \mathbb{E}[\exp(\beta_3(Z))(Z-K_3^*)^+] = \mathbb{E}[\exp(\theta_3 Z+\delta_3)(Z-K_3^*)^+] \\ &= \frac{\eta}{(\eta-\theta_3)^2}\exp(\delta_3)\exp(-(\eta-\theta_3)K_3^*). \end{aligned}$$

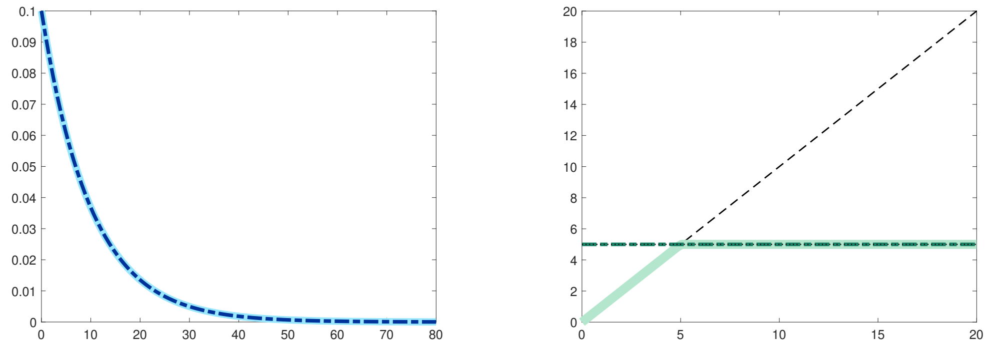
E.6 Numerical illustration

The results of the previous section are now collected and the analysis concluded by a numerical illustration. To do so values of the parameters must be chosen. We here consider an individual with a utility parameter $a = 15$, subjective discount factor $\rho = 10\%$, a claim frequency $\lambda = 0.01$, and losses are assumed to be exponentially distributed with parameter $\eta = 0.1$. The net premium for full insurance of the non-life risk of this individual is thus $\xi = \lambda/\eta = 0.1$. Suppose also that the risk-free interest rate is $r = 5\%$.

E.6.1 The impact of the pricing measure function

We start off by visualising the impact of the pricing measure function. Firstly, we plot the loss density under the pricing measure determined by (4.17), and secondly, we illustrate the optimal insurance strategy of the individual. Let f denote the density of the exponential distribution with parameter η, and $f^{\mathbb{Q}}$ the density under the pricing measure $\mathbb{Q}$.

When the insurance company sets its premium according to a constant pricing measure function, $\beta_1(z) = \delta_1$, it does not charge for claim size risk, and therefore the density of the claims remains unchanged, i.e. $f(z) = f^{\mathbb{Q}}(z)$. The a priori density and the pricing density is presented in Figure 4.1a. Instead, the claim frequency used for pricing increases to $\lambda^{\mathbb{Q}} = \lambda \exp(\delta_1)$. For $\delta_1 = 3.75$ the optimal fixed deductible level is then $K = 5$, which is 50% of the average loss of $1/\eta = 10$. Note that for a constant pricing measure function, a fixed amount deductible is optimal for the individual, and the welfare loss in (4.13) is therefore zero. The optimal deductible strategy is depicted in Figure 4.1b.

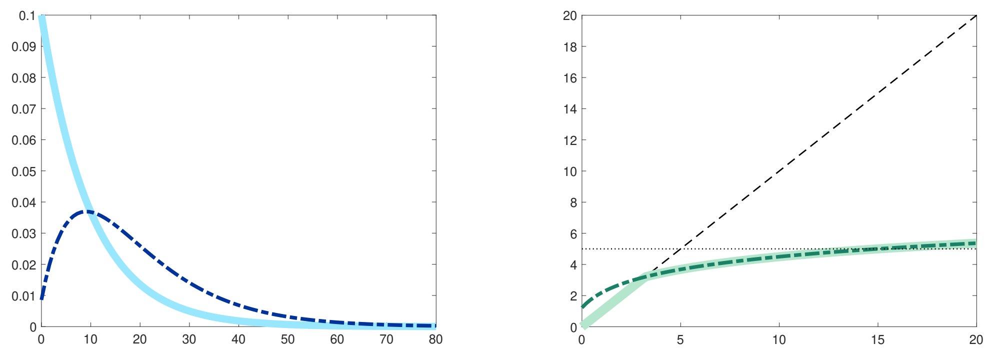


(a) The densities $f(z)$ (in solid) and $f^{\mathbb{Q}}(z)$ (in dash-dotted).

(b) The functions $g_1(z)$ (in dash-dotted) and $\min\{z, g_1(z)\}$ (in solid).

Figure 4.1: For a constant pricing measure function

For a log-linear pricing measure function, $\beta_2(z) = \log(\theta_2 z + \delta_2)$, similar plots are displayed in Figure 4.2a and 4.2b. The pricing parameters are calibrated to satisfy that the individual would optimally choose a fixed amount deductible level of $K_2^* = 5$ if restricted to do so. Choosing $\delta_2 = 2.5$, then this calibration leads to $\theta_2 = 2.6681$. The loss density under the pricing measure, namely $f^{\mathbb{Q}}(z) = \eta \exp(-\eta z)(\theta_2 z + \delta_2)/(\theta_2/\eta + \delta_2)$, has quite a different nature than the a priori exponential distribution, though still staying within the exponential family of distributions. We refer to Figure 4.2a. Recall from (4.17) that claim frequency risk is penalised by $\lambda^{\mathbb{Q}} = \lambda(\theta_2/\eta + \delta_2)$. In Figure 4.2b the optimal flexible deductible



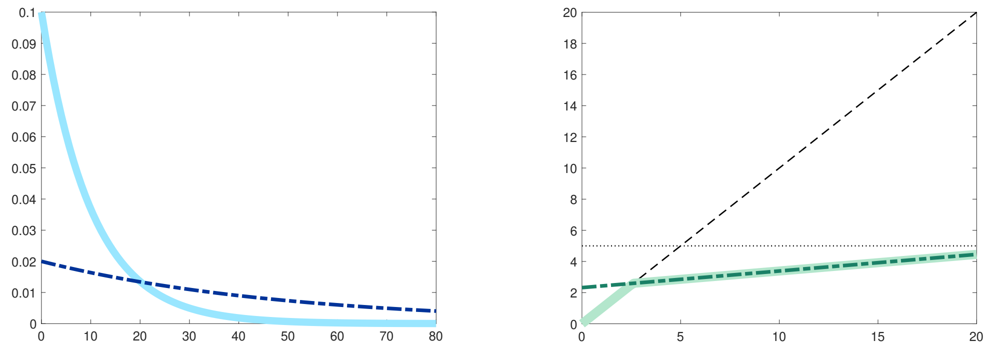
(a) The densities $f(z)$ (in solid) and $f^{\mathbb{Q}}(z)$ (in dash-dotted).

(b) The functions $g_2^*(z)$ (in dash-dotted) and $\min\{z, g_2^*(z)\}$ (in solid).

Figure 4.2: For a log-linear pricing measure function

strategy, namely $g_2^*(z) = \log(\theta_2 z + \delta_2)/(ra)$, as a function of the loss is illustrated. The optimal fixed amount deductible level for these pricing parameters also appears in the graph. We observe that the flexible insurance product yields a deductible level below the standard product for small claim sizes, whereas it slowly grows above for larger claim sizes. In this specific case, the welfare loss of the individuals being restricted to the standard product with a fixed amount deductible rather than the flexible product is $L = 1.2116$, when the pricing measure function is log-linear. As mentioned previously, for this individual the net premium for full insurance is $\xi = 0.1$, so relative to this, the welfare loss is $L/\xi = 12.116$.

Corresponding plots for a linear pricing measure function, $\beta_3(z) = \theta_3 z + \delta_3$, can be seen in Figure 4.3a and 4.3b. Again, the parameters $\theta_3 = 0.25$ and $\delta_3 = 1.2472$ are chosen to such that the optimal fixed deductible level is $K_3^* = 5$. The density under the pricing measure is still exponential, $f^{\mathbb{Q}}(z) = (\theta_3 - \eta)\exp(-(\eta - \theta_3)z)$, but tilted to have a heavier tail, this is apparent in Figure 4.3a. Claim frequency is penalised by $\lambda^{\mathbb{Q}} = \lambda\eta\exp(\delta_3)/(\eta - \theta_3)$.



(a) The densities $f(z)$ (in solid) and $f^{\mathbb{Q}}(z)$ (in dash-dotted).

(b) The functions $g_3^*(z)$ (in dash-dotted) and $\min\{z, g_3^*(z)\}$ (in solid).

Figure 4.3: For a linear pricing measure function

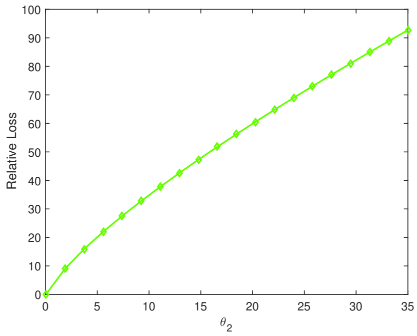
In Figure 4.3b the optimal design of the deductible, $g_3^*(z) = (\theta_3 z + \delta_3)/(ra)$, is illustrated if the individual could choose freely. Compared to being limited to a fixed amount deductible strategy, the individual would optimally choose a linearly growing strategy, which exceeds K_3^* at the point $(raK_3^* - \delta_3)/\theta_3$ (not visible on the graph). The individual's welfare loss of being restricted to the simple product rather than the flexible, when the pricing measure function is linear, is $L = 16.6683$ for this choice of parameters. Relative to the cost of full insurance, the loss is then $L/\xi = 166.683$ times larger.

E.6.2 The welfare loss

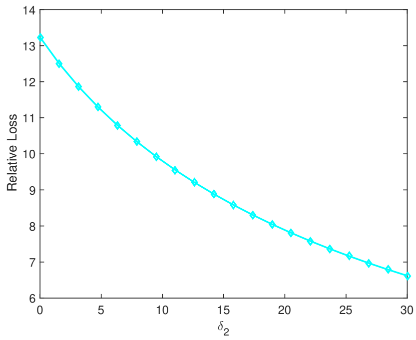
The welfare loss is obviously dependent on the values of the parameters. To illustrate the sensitivity towards changes in the values of these parameters as clear as possible, graphs of the relative loss L/ξ are displayed in Figure 4.4 for a log-linear pricing measure function, and in Figure 4.5 for a linear pricing measure function. Each figure has four subfigures where one of the parameters vary, while the remaining are kept fixed. Figure 4.4a and 4.5a show the relative loss as a function of θ, 4.4b and 4.5b as a function of δ, 4.4c and 4.5c as a function of the absolute risk aversion a and, finally, 4.4d and 4.5d as a function of the loss parameter η. In the latter, we remark that also the net premium for full insurance varies in η. Note that the optimal fixed deductible changes as well, when varying these parameters.

For the log-linear pricing measure function, $\beta_2(z) = \log(\theta_2 z + \delta_2)$, we observe that the relative loss increases when θ_2 increases, see Figure 4.4a. This makes good sense intuitively since $\partial g_2^*(z)/\partial z = 1/(ra(z + \delta_2/\theta_2))$, hence a larger value of θ_2 yields a steeper slope of the optimal flexible deductible strategy, which then will deviate more from the fixed amount deductible strategy. Notice that δ_2 has an inverse impact on $\partial g_2^*(z)/\partial z$, and we can therefore conclude the converse for δ_2. The parameter δ_2 also controls the intersection with the vertical axis and by raising it, a larger part of the function $g_2^*(z)$ with a steep slope will be above the identity line. Hence, as it appears in Figure 4.4b, the relative loss decreases in δ_2. Next, recall that a is the parameter of the exponential utility, and is thus a measure of the absolute risk aversion. An individual with risk neutral preferences (a close to zero) would not pay for insurance as she does not care about the risk. So for this type of individual it does not matter which product is supplied as long as ‘no insurance’ is a possibility. On the other hand, an individual with a very high degree of risk aversion (a large) would prefer to insure fully, and once again, the product structure becomes subordinate as long as a ‘full insurance’ (i.e. a zero deductible) can be chosen. Hence, the difference in product design is the most important for individuals with non-extreme preference, as it appears from the non-monotonicity of Figure 4.4c. At last, in Figure 4.4d the relative welfare loss is decreasing in the loss parameter η as expected. When η increases the tail of the loss distribution gets lighter, and claims will on average get smaller. Hence, the difference between the optimal flexible deductible strategy and the fixed amount deductible for large claim sizes affects the welfare loss less.

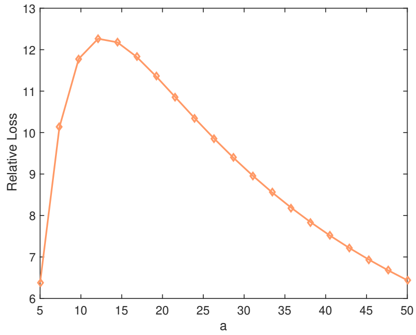
The arguments for the case with a linear pricing measure function, $\beta_3(z) = \theta_3 z + \delta_3$, are similar. Since $\partial g_3^*(z)/\partial z = \theta_3/(ra)$, the parameter θ_3 controls the slope of the optimal flexible deductible strategy. A higher value of θ_3 yields a higher slope, so the distance to the fixed deductible will then be larger and the welfare loss bigger, the graph in Figure 4.5a is therefore increasing. If δ_3 increases then the flexible deductible strategy will exceed the fixed deductible for lower values of the losses, and the individual is therefore forced to buy more insurance for large values of the claim if being restricted to a fixed deductible, leading to a larger welfare loss, which explains the decreasing shape of Figure 4.5b. For the risk aversion coefficient, the effect from being able to choose a slope on the coverage function dominates. If the individual is tending towards risk neutrality (a small), then the product design allows her to choose a high slope giving her a smaller (if not zero) insurance coverage. In contrast, for the risk averse individual (a large) that seeks a high insurance coverage, the best she can obtain in terms of slope is the fixed amount deductible (that is, zero slope), in which case the difference between the flexible product and the product with a fixed amount deductible diminishes, which explains the monotonicity in Figure 4.5c. The



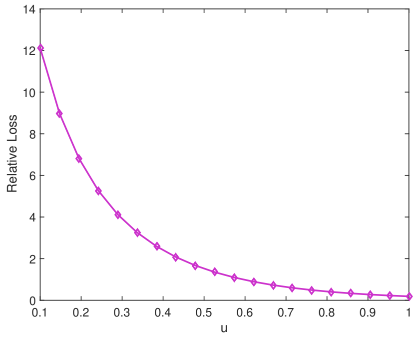
(a) Loss as function of θ_2.



(b) Loss as function of δ_2.

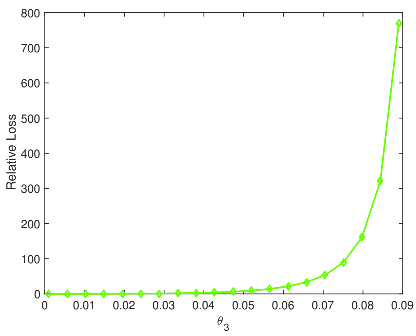


(c) Loss as function of a.

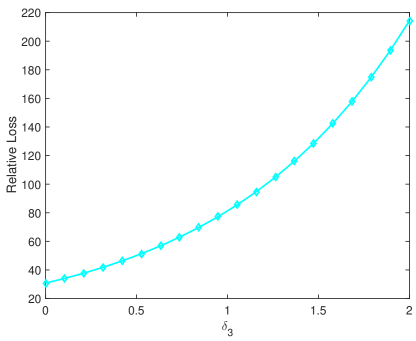


(d) Loss as function of u.

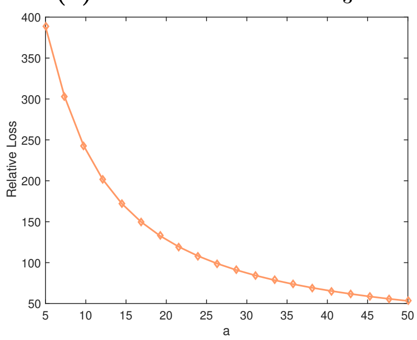
Figure 4.4: Loss function for β_2



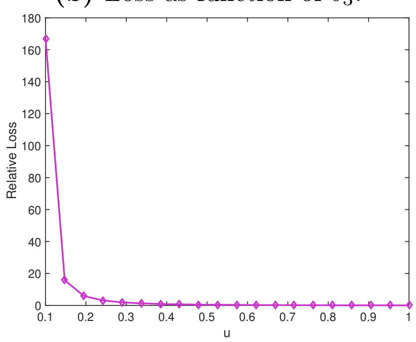
(a) Loss as function of θ_3.



(b) Loss as function of δ_3.



(c) Loss as function of a.



(d) Loss as function of u.

Figure 4.5: Loss function for β_3

sensitivity to the loss parameter η is similar to the second case of pricing measure function, except that here the deviation from the flexible deductible structure to the fixed amount deductible is larger for big losses, which means that the relative loss is more sensitive towards the heaviness of the loss, as we see in Figure 4.5d.

E.7 A digression on power utility

So far we have concentrated almost exclusively on the exponential utility function. Under exponential preferences of the policyholder we could calculate explicitly both the optimal flexible deductible and the best deductible within the suboptimal class of constant deductibles. We benefited from explicit solutions, even to the suboptimal insurance position, along with the wealth-independent insurance decisions in order to obtain an explicit and wealth independent quantification of the welfare loss from being offered a suboptimal product design.

In this section we briefly touch upon the case of power utility. Power utility is a more standard formalization of individual preferences within the area of personal financial decision making. However, in order to find a solution to the optimization problem (4.20), it is necessary to assume that a deductible can exceed the loss, which corresponds to removing the truncation and minimum in (4.20). If the deductible exceeds the loss, it is interpreted as if the individual is actually betting against having a loss of that size. So if the individual finds insurance to be too expensive, then the she will not only choose not to insure, but will actually try to turn it to her advantage that the pricing is too high.

Unfortunately, we therefore do not find explicit solutions to the problem of choosing a fixed deductible, since this standard contract involves the truncation and the minimum. This prevents explicit quantification of the welfare loss from suboptimal contracts in spite of the fact that we can actually find an optimal flexible contract. Although we cannot present the best choice among the standard marketed contract, we choose here to present briefly the optimal flexible one. This serves at the same time as yet another illustration of the HJB machinery exploited in the previous sections as well as a motivating starting point for further studies in the direction.

The result (4.1) is developed under wealth-homogeneous assumptions on the insurance control as the compound Poisson structure is essential and it is therefore necessary to have i.i.d claim sizes $(Z_i - g_v(Z_i))_{i\in\mathbb{N}}$. This works successfully for the exponential utility due to its desirable analytical properties. For a power-utility, the optimal deductible strategy depends on wealth which prevents a compound Poisson structure, but we deliberately use the result of (4.1), conditional on the pre-claim wealth, without further notice. This should be taken as a premium principle (where the jump intensity and the claim size risk are punished separately) rather than a legitimate measure change. The premium at time t is then a function of the time t wealth. Another approach would be to discretise the claim sizes, introduce a pricing rate for every level of the claim sizes, and let the individual choose a distinct deductible for each of these levels. The issue with the measure change is then avoided and similar results are produced, but we will not comment further on this.

Proposition E.3. *Assume that the utility function is of the form*

$$u(c) = \frac{1}{1-\gamma} c^{1-\gamma}. \tag{4.22}$$

for $\gamma > 0, \gamma \neq 1$. Then the value function (4.4) can be written by

$$V(w) = \alpha(w-\kappa)^{1-\gamma}. \tag{4.23}$$

The optimal consumption c^* is affine in wealth

$$c^*(w) = (\alpha(1-\gamma))^{-1/\gamma}(w-\kappa), \tag{4.24}$$

and the optimal insurance control is given by

$$g^*(w,z) = \left(1 - \exp(\beta(z))^{-1/\gamma}\right)(w-\kappa).$$

The parameters α and κ of the value function are determined by

$$\begin{aligned}
\alpha &= \frac{1}{(1-\gamma)\gamma^{-\gamma}} \Bigg(\rho - r(1-\gamma) - \lambda \mathbb{E}\left[1 - \exp(\beta(Z))^{-1/\gamma}\right](1-\gamma) \\
&\qquad\qquad - \lambda\left(\mathbb{E}\left[\exp(\beta(Z))^{(\gamma-1)/\gamma}\right] - 1\right) \Bigg)^{-\gamma}, \\
\kappa &= \frac{\lambda}{r} \mathbb{E}[\exp(\beta(Z))Z].
\end{aligned}$$

Proof. We use the same approach as seen previously. We make the conjecture that the value function is of the type $V(w) = \alpha(w-\kappa)^{1-\gamma}$. First order conditions leads to the optimal choices c^* and $g^*(z)$, and if we insert these back into the HJB equation and solve, we obtain α and κ. □

Note that unlike the exponential utility case this insurance control depends on wealth. When the initial wealth approaches the present value of the cost of the claims, i.e. when $w \to \kappa$, the individual stops consuming, $c^*(w) \to 0$, and insures fully, $g^*(w,z) \to 0$ for any fixed z. The reason why it is so important to protect the value of full coverage is that an individual with power preference must avoid negative consumption and wealth almost surely.

Although we cannot quantify the welfare loss from a suboptimal coverage, we know that there is one, and the result of Proposition E.3 represents an idea for product development in the non-life business. The optimal deductible is affine in the wealth. For $w \gg \kappa$, the optimal deductible is essentially linear in wealth. This could be incorporated in the insurance product or, at least, in product advice given to policy holders.

The proportionality of g^* in w, namely $1 - \exp(\beta(z))^{-1/\gamma}$, has a simple structure which can also be the starting point for further product development. For the linear pricing principle, this is a constant. We note that for the linear pricing rule the optimal deductible is not a constant but a constant fraction of wealth (minus the typically relative small value of full coverage). For our variance principle the fraction becomes a linear-power function of the loss and for our Esscher principle the fraction becomes an exponential-power function. These are quite simple structures that can easily be incorporated in indemnity tables.

The limiting cases for the risk aversion are obvious. As $\gamma \to \infty$, $g^* \to 0$ for all claims and all sizes of wealth. If the individual is extremely risk averse, she avoids risk at any price and demands full protection. As $\gamma \to 0$, $g \to w - \kappa$ for all claim sizes. If the individual is not really risk averse, she can keep the risk at any price but only up to the point where a claim threatens her wealth in order to avoid, above all, negative wealth and consumption.

Appendix

E.I The impact of pricing on the individual's decision

In this section we solve the optimization of the policy holder for a particularly simple pricing rule. It serves as a motivating example for the more abstract calculations in Section E.4. We are going to show very clearly and explicitly with this example that for a specific pricing rule, the size of the claim is indeed, in general, relevant for calculating the optimal deductible. It is natural to speak of the pricing rule studied in this section as being based on a change of measure which is piecewise constant in the size of the loss. Then the optimal deductible indeed also becomes piecewise constant in general.

Let $0 = \ell_1 < \ell_2 < \ldots < \ell_{n-1} < \ell_n = \infty$ be given, such that $[\ell_1, \ell_2), \ldots, [\ell_{n-1}, \ell_n)$ is a finite partition of $\mathbb{R}^+$. When a loss of size z occurs, it falls into one of the regions $[\ell_i, \ell_{i+1})$ for $i = 1, \ldots, n-1$. Assume that the individual can choose a distinct deductible for each region, i.e. a deductible K_i when z takes value in $[\ell_i, \ell_{i+1})$. In this case the insurance product is characterised by

$$g_{(K_1,\ldots,K_n)}(z) = \sum_{i=1}^{n-1} K_i \mathbb{1}_{\{z \in [\ell_i, \ell_{i+1})\}}.$$

The theory of space-decomposition allows us to split, correspondingly, the compound Poisson process $(A_t)_{t \geq 0}$ describing the losses of the individual (without insurance) into n compound Poisson processes with jump rate $\lambda_i = \lambda \mathbb{P}(Z \in [\ell_i, \ell_{i+1}))$ and jump sizes in $[\ell_i, \ell_{i+1})$ for $i = 1, \ldots, n-1$, respectively. Assume that the insurance company prices each of these risks individually with pricing rates $\lambda_1^Q, \ldots, \lambda_n^Q$, where typically $\lambda_{i+1}^Q / \lambda_{i+1} > \lambda_i^Q / \lambda_i$ (large claims constitute larger risks and should be charged accordingly). We can now solve the optimization problem of the policyholder for a given partition of the pricing rule.

Proposition E.4. *Assume that the insurance company applies a piecewise constant pricing rule as described above. Then the value function is of type* $V(w) =$

$-\exp(-raw)/\alpha$ *and the optimal controls are*

$$c^* = rw - \frac{1}{a}\log\Big(\frac{ra}{\alpha}\Big) \quad \textit{and} \quad K_i^* = \frac{1}{ra}\log\left(\frac{\lambda_i^Q}{\lambda_i}\right) \qquad \textit{for } i = 1, \dots, n.$$

Proof. The HJB equation is in this case

$$\begin{aligned}\sup\Big\{ -\rho V(w) + \Big(rw - c - \sum_{i=1}^{n-1} \lambda_i^{\mathbb{Q}} \mathbb{E}[(Z-K_i)^+ \mid Z \in [\ell_i, \ell_{i+1})]\Big) V_w(w) \\ + \sum_{i=1}^{n-1} \lambda_i \mathbb{E}[V(w - \min\{Z, K_i\}) - V(w) \mid Z \in [\ell_i, \ell_{i+1})] + u(c)\Big\} = 0.\end{aligned}$$

Since the nature of the problem is similar to the previous section, we again guess that the value function is of type $V(w) = -\exp(-raw)/\alpha$. The first order condition with respect to consumption then leads to the same structure as seen before,

$$c^* = rw - \frac{1}{a}\log\left(\frac{ra}{\alpha}\right),$$

The Leibniz integral rule is used to find the first order condition with respect to the i'th deductible level,

$$\lambda_i^{\mathbb{Q}} \mathbb{E}[\mathbb{1}_{\{Z_i > K \mid Z_i \in [\ell_i, \ell_{i+1})\}}] - \lambda_i \mathbb{E}[\mathbb{1}_{\{Z_i > K \mid Z_i \in [\ell_i, \ell_{i+1})\}}] \exp(raK_i) = 0.$$

Rearranging and isolating yields

$$K_i^* = \frac{1}{ra}\log\left(\frac{\lambda_i^{\mathbb{Q}}}{\lambda_i}\right) \text{ for } i = 1, \dots, n. \tag{4.25}$$

Inserting the optimal controls in the HJB we obtain the supremum, and we can then solve for

$$\begin{aligned}\alpha = \frac{1}{ra}\exp\Big\{\frac{1}{r}\Big(\rho - \sum_{i=1}^{n} \lambda_i \big(\mathbb{E}[\exp(ra\min\{Z, K_i^*\}) \mid Z \in [\ell_i, \ell_{i+1})] - 1\big)\Big) \\ - a\sum_{i=1}^{n} \lambda_i^{\mathbb{Q}} \mathbb{E}[(Z - K_i^*)^+ \mid Z \in [\ell_i, \ell_{i+1})] - 1\Big\}.\end{aligned}$$

□

Proposition E.4 shows in a simple way how the price of the insurance coverage affects the optimal extent of coverage. The more expensive the insurance is, measured by the pricing ratio $\lambda_i^{\mathbb{Q}}/\lambda_i$, the larger a deductible is optimal for the individual. For the exponential utility case the part is as simple as the logarithm of the pricing ratio times a constant which contains the level of risk aversion.

A special case arises if there is only one 'piece' and piecewise constant really means constant. This corresponds to the expected value pricing principle since the value of a contract is proportional to its expectation with a constant of proportionality equal to $\lambda^{\mathbb{Q}}/\lambda$. As previously argued, if the pricing is based on the expected value premium principle, then a fixed amount deductible is optimal. Proposition E.4 repeats this result for the special case of exponential utility and determines the constant deductible level to be

$$K^* = \frac{1}{ra}\log\left(\frac{\lambda^{\mathbb{Q}}}{\lambda}\right).$$

The n piecewise constant deductible case is a special case of the measure transformation with pricing measure function

$$\beta(z) = \sum_{i=1}^{n} \delta_i \mathbb{1}_{z\in[\ell_i,\ell_{i+1})} \qquad \text{where } \delta_i = \log\left(\frac{\lambda_i^Q}{\lambda_i}\right) \text{ for all } i.$$

E.II Proof of Proposition E.2

Proof. To abbreviate the notation, let $f_\beta(t) = \exp(-\lambda t\mathbb{E}[\exp(\beta(Z))-1])$ and $h_v(z) = (z-g_v(z))^+$. Now simply consider the characteristic function

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}(\beta)}\Big[\exp\Big(\mathrm{i}\, s\cdot\sum_{i=1}^{N_t} h_v(Z_i)\Big)\Big]\\
&\quad= \mathbb{E}\Big[\exp\Big(\sum_{i=1}^{N_t}\beta(Z_i) - \lambda t\mathbb{E}[\exp(\beta(Z))-1]\Big)\exp\Big(\mathrm{i}\, s\cdot\sum_{i=1}^{N_t} h(Z_i)\Big)\Big]\\
&\quad= f_\beta(t)\mathbb{E}\Big[\prod_{i=1}^{N_t}\frac{\exp(\beta(Z_i))\exp(\mathrm{i}\, s\cdot h_v(Z_i))}{\mathbb{E}[\exp(\beta(Z))]}\mathbb{E}[\exp(\beta(Z))]\Big]\\
&\quad= f_\beta(t)\ \mathbb{E}\Big[\mathbb{E}\Big[\prod_{i=1}^{N_t}\frac{\exp(\beta(Z_i))\exp(\mathrm{i}\, s\cdot h_v(Z_i))}{\mathbb{E}[\exp(\beta(Z))]}\mathbb{E}[\exp(\beta(Z))]\ \Big|\ N_t\Big]\Big]\\
&\quad= f_\beta(t)\ \mathbb{E}\Big[\mathbb{E}[\exp(\beta(Z))]^{N_t}\mathbb{E}\Big[\frac{\exp(\beta(Z))\exp(\mathrm{i}\, s\cdot h_v(Z))}{\mathbb{E}[\exp(\beta(Z))]}\Big]^{N_t}\Big]\\
&\quad= f_\beta(t)\ \exp\Big(\lambda t\Big(\mathbb{E}[\exp(\beta(Z))]\mathbb{E}\Big[\frac{\exp(\beta(Z))}{\mathbb{E}[\exp(\beta(Z))]}\exp(\mathrm{i}\, s\cdot h_v(Z_i))\Big]-1\Big)\Big)\\
&\quad= \exp\Big(\lambda t\mathbb{E}[\exp(\beta(Z))]\Big(\mathbb{E}\Big[\frac{\exp(\beta(Z))}{\mathbb{E}[\exp(\beta(Z))]}\exp(\mathrm{i}\, s\cdot h_v(Z_i))\Big]-1\Big)\Big).
\end{aligned}$$

This is the characteristic function of a compound Poisson process with characteristics (4.17). From this, the expected value (4.18) follows directly. □

E.III Calculations for $\beta_1(z) = \delta_1$

$$
\begin{aligned}
&\mathbb{E}[\exp(ra \cdot \min\{Z, \delta_1/(ra)\})] \\
&\quad = \mathbb{P}(A_1)\mathbb{E}[\exp(raZ) \mid A_1] + \mathbb{P}(A_1^c)\mathbb{E}[\exp(\delta_1) \mid A_1^c] \\
&\quad = \mathbb{E}[\exp(raZ)\mathbb{1}_{\{Z<\delta_1/(ra)\}}] + \mathbb{E}[\exp(\delta_1)\mathbb{1}_{\{Z\geq\delta_1/(ra)\}}] \\
&\quad = \int_0^{\delta_1/(ra)} \eta \exp((ra-\eta)z)dz + \exp\left(\frac{ra-\eta}{ra}\delta_1\right) \\
&\quad = \frac{ra}{ra-\eta}\exp\left(\frac{ra-\eta}{ra}\delta_1\right) - \frac{\eta}{ra-\eta}. \\
&\mathbb{E}[\exp(\beta(Z))(Z-g(Z))^+] = \mathbb{P}(A_1^c)\exp(\delta_1)\mathbb{E}[(Z-\delta_1/(ra)) \mid A_1^c] \\
&\quad = \mathbb{E}[\exp(\delta_1)\,(Z-\delta_1/(ra))\,\mathbb{1}_{\{Z\geq\delta_1/(ra)\}}] \\
&\quad = \exp(\delta_1)\int_{\delta/(ra)}^{\infty} \eta\exp(-\eta z)(z-\delta_1/(ra))\mathrm{d}z \\
&\quad = \frac{1}{\eta}\exp\left(\frac{ra-\eta}{ra}\delta_1\right).
\end{aligned}
$$

E.IV Calculations for $\beta_2(z) = \log(\theta_2 z + \delta_2)$

Let $\tilde{Z} = \theta_2 Z + \delta_2$ and $\hat{Z} = -ra\tilde{Z}/\theta_2$. The event $A_2 = \{Z < \log(\theta_2 Z + \delta_2)/(ra)\}$ can then be rewritten as

$$
\begin{aligned}
A_2 &= \{Z < \log(\theta_2 Z + \delta_2)/(ra)\} = \{\exp(raZ) < \theta_2 Z + \delta_2\} \\
&= \{\exp(ra\tilde{Z}/\theta_2)\exp(-\delta_2 ra/\theta_2) < \tilde{Z}\} \\
&= \{\hat{Z}\exp(\hat{Z}) < -ra\exp(-\delta_2 ra/\theta_2)/\theta_2\} \\
&= \{\hat{Z} > \mathcal{W}_{-1}(-ra\exp(-\delta_2 ra/\theta_2)/\theta_2)\},
\end{aligned}
$$

where we use that $-ra\exp(-\delta_2 ra/\theta_2)/\theta_2 \in [-\exp(-1), 0)$, which is necessary in order for $\mathcal{W}_{-1}$ to be defined. The upper boundary is trivial, whereas the lower is a bit less so. First, we recognise that $-x\exp(-\delta_2 x) > -x\exp(-x)$ for any values of x and $\delta_2 \geq 1$. Next, we see that $-x\exp(-x)$ obtains its minimum for $x = 1$, hence $-x\exp(-x) > -\exp(-1)$.

Translated back to Z using substitution, we can now conclude that

$$
A_2 = \{Z < Q\} \qquad \text{where } Q = -\frac{1}{ra}\mathcal{W}_{-1}\left(-\frac{ra}{\theta_2}\exp\left(-\delta_2\frac{ra}{\theta_2}\right)\right) - \frac{\delta_2}{\theta_2}.
$$

We therefore get

$$
\begin{aligned}
&\mathbb{E}[\exp(ra\cdot\min\{Z,\log(\theta_2 Z+\delta_2)/(ra)\})]\\
&\quad=\mathbb{P}(A_2)\mathbb{E}[\exp(raZ)\mid A_2]+\mathbb{P}(A_2^c)\mathbb{E}[\theta_2 Z+\delta_2\mid A_2^c]\\
&\quad=\int_{A_2}\eta\exp((ra-\eta)z)dz+\int_{A_2^c}\eta(\theta_2 z+\delta_2)\exp(-\eta z)dz\\
&\quad=\int_0^Q\eta\exp((ra-\eta)z)dz+\int_Q^\infty\eta(\theta_2 z+\delta_2)\exp(-\eta z)dz\\
&\quad=\frac{\eta}{ra-\eta}(\exp((ra-\eta)Q)-1)+\left(\theta_2 Q+\delta_2+\frac{\theta_2}{\eta}\right)\exp(-\eta Q).\\
&\mathbb{E}[(\theta Z+\delta_2)(Z-\log(\theta_2 Z+\delta_2)/(ra))^+]\\
&\quad=\mathbb{P}(A_2^c)\mathbb{E}[(\theta_2 Z+\delta_2)(Z-\log(\theta_2 Z+\delta_2)/(ra))\mid A_2^c]\\
&\quad=\int_{A_2^c}\eta(\theta_2 z+\delta_2)(z-\log(\theta_2 z+\delta_2)/(ra))\exp(-\eta z)dz\\
&\quad=\int_Q^\infty\eta(\theta_2 z+\delta_2)(z-\log(\theta_2 z+\delta_2)/(ra))\exp(-\eta z)dz\\
&\quad=\exp(-\eta Q)\Bigg(\theta_2 Q^2+\left(\delta_2+2\frac{\theta_2}{\eta}\right)Q-\frac{1}{\eta}\left(\frac{\theta_2}{ra}-\delta_2\right)+2\frac{\theta_2}{\eta^2}\\
&\qquad\qquad-\frac{1}{ra}\frac{\theta_2}{\eta}\exp\left(\frac{\eta}{\theta_2}(\theta_2 Q+\delta_2)\right)E_1\left(\frac{\eta}{\theta_2}(\theta_2 Q+\delta_2)\right)\\
&\qquad\qquad-\frac{1}{ra}\left(\theta_2 Q+\delta_2+\frac{\theta_2}{\eta}\right)\log(\theta_2 Q+\delta_2)\Bigg).
\end{aligned}
$$

E.V Calculations for $\beta_3(z)=\theta_3 z+\delta_3$

$$
\begin{aligned}
&\mathbb{E}[\exp(ra\cdot\min\{Z,(\theta_3 Z+\delta_3)/(ra)\})]\\
&\quad=\int_{A_3}\eta\exp(raz)\exp(-\eta z)\mathrm{d}z+\int_{A_3^c}\eta\exp(\theta_3 z+\delta_3)\exp(-\eta z)\mathrm{d}z\\
&\quad=\int_0^{\delta_3/(ra-\theta_3)}\eta\exp(raz)\exp(-\eta z)\mathrm{d}z+\int_{\delta_3/(ra-\theta_3)}^\infty\eta\exp(\theta_3 z+\delta_3)\exp(-\eta z)\mathrm{d}z\\
&\quad=\frac{\eta}{\eta-ra}\left(1-\exp\left(-(\eta-ra)\frac{\delta_3}{ra-\theta_3}\right)\right)+\frac{\eta}{\eta-\theta_3}\exp\left(-(\eta-ra)\frac{\delta_3}{ra-\theta_3}\right).
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}[\exp(\theta_3 Z+\delta_3)(Z-(\theta_3 Z+\delta_3)/(ra))^+]\\
&\quad=\int_{A_3^c}\exp(\theta_3 z+\delta_3)(z-(\theta_3 z+\delta_3)/(ra))\mathrm{d}z\\
&\quad=\int_{\delta_3/(ra-\theta_3)}^\infty\exp(\theta_3 z+\delta_3)(z-(\theta_3 z+\delta_3)/(ra))\mathrm{d}z\\
&\quad=\exp\left(-(\eta-ra)\frac{\delta_3}{ra-\theta_3}\right)\frac{(ra-\theta_3)\eta}{ra(\eta-\theta_3)^2}.
\end{aligned}
$$

References

Aase, K. K. (2017). "Optimal Insurance Policies in the Presence of Costs". *Risks* 5.3.

Arrow, K. J. (1971). "The theory of risk aversion". *Essays in the Theory of Risk Bearing*, pp. 90–109.

Avanzi, B., Tu, V., and Wong, B. (2016). "A note on realistic dividends in actuarial surplus models". *Risks* 4.4, p. 37.

Cummins, J. D. and Mahul, O. (2004). "The demand for insurance with an upper limit on coverage". *The Journal of Risk and Insurance* 71.2, pp. 253–264.

Delbaen, F. and Haezendonck, J. (1989). "A martingale approach to premium calculation principles in an arbitrage free market". *Insurance: Mathematics and Economics* 8.4, pp. 269–277.

Embrechts, P. (2000). "Actuarial versus financial pricing of insurance". *The Journal of Risk Finance* 1.4, pp. 17–26.

Golubin, A. (2016). "Optimal insurance and reinsurance policies chosen jointly in the individual risk model". *Scandinavian Actuarial Journal* 2016.3, pp. 181–197.

Grandits, P., Hubalek, F., Schachermayer, W., and Žigo, M. (2007). "Optimal expected exponential utility of dividend payments in a Brownian risk model". *Scandinavian Actuarial Journal* 2007.2, pp. 73–107.

Hubalek, F. and Schachermayer, W. (2004). "Optimizing expected utility of dividend payments for a Brownian risk process and a peculiar nonlinear ODE". *Insurance: Mathematics and Economics* 34.2, pp. 193–225.

Moore, K. S. and Young, V. R. (2006). "Optimal insurance in a continuous-time model". *Insurance: Mathematics and Economics* 39.1, pp. 47–68.

Perera, R. S. (2010). "Optimal consumption, investment and insurance with insurable risk for an investor in a Lévy market". *Insurance: Mathematics and Economics* 46.3, pp. 479–484.

Schmidli, H. (2002). "On minimizing the ruin probability by investment and reinsurance". *The Annals of Applied Probability* 12.3, pp. 890–907.

— (2008). *Stochastic control in insurance*. Springer-Verlag London.

Thonhauser, S. and Albrecher, H. (2011). "Optimal dividend strategies for a compound Poisson process under transaction costs and power utility". *Stochastic Models* 27.1, pp. 120–140.

Yang, H. and Zhang, L. (2005). "Optimal investment for insurer with jump-diffusion risk process". *Insurance: Mathematics and Economics* 37.3, pp. 615–634.

Zhang, X. and Siu, T. K. (2009). "Optimal investment and reinsurance of an insurer with model uncertainty". *Insurance: Mathematics and Economics* 45.1, pp. 81–88.

Zhou, C., Wu, W., and Wu, C. (2010). "Optimal insurance in the presence of insurer's loss limit". *Insurance: Mathematics and Economics* 46.2, pp. 300–307.

Zou, B. and Cadenillas, A. (2014). "Explicit solutions of optimal consumption, investment and insurance problems with regime switching". *Insurance: Mathematics and Economics* 58, pp. 159–167.