



Special Issue Reprint

---

# Actuarial Mathematics and Risk Management

---

Edited by  
Annamaria Olivieri

[www.mdpi.com/journal/risks](http://www.mdpi.com/journal/risks)



# **Actuarial Mathematics and Risk Management**



# Actuarial Mathematics and Risk Management

Editor

**Annamaria Olivieri**

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin



*Editor*

Annamaria Olivieri  
Department of Economics  
and Management  
University of Parma  
Parma  
Italy

*Editorial Office*

MDPI  
St. Alban-Anlage 66  
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Risks* (ISSN 2227-9091) (available at: [www.mdpi.com/journal/risks/special\\_issues/Actuarial\\_Mathematics\\_Risk](http://www.mdpi.com/journal/risks/special_issues/Actuarial_Mathematics_Risk)).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. <i>Journal Name</i> <b>Year</b> , Volume Number, Page Range.
--

**ISBN 978-3-0365-8391-4 (Hbk)**

**ISBN 978-3-0365-8390-7 (PDF)**

© 2023 by the authors. Articles in this book are Open Access and distributed under the Creative Commons Attribution (CC BY) license, which allows users to download, copy and build upon published articles, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications.

The book as a whole is distributed by MDPI under the terms and conditions of the Creative Commons license CC BY-NC-ND.

# Contents

<b>About the Editor</b> . . . . .	<b>vii</b>
<b>Preface to "Actuarial Mathematics and Risk Management"</b> . . . . .	<b>ix</b>
<b>Annamaria Olivieri</b> Special Issue "Actuarial Mathematics and Risk Management" Reprinted from: <i>Risks</i> <b>2023</b> , <i>11</i> , 134, doi:10.3390/risks11070134 . . . . .	<b>1</b>
<b>Yaser Awad, Shaul K. Bar-Lev and Udi Makov</b> A New Class of Counting Distributions Embedded in the Lee–Carter Model for Mortality Projections: A Bayesian Approach Reprinted from: <i>Risks</i> <b>2022</b> , <i>10</i> , 111, doi:10.3390/risks10060111 . . . . .	<b>3</b>
<b>Anja Breuer and Yves Staudt</b> Equalization Reserves for Reinsurance and Non-Life Undertakings in Switzerland Reprinted from: <i>Risks</i> <b>2022</b> , <i>10</i> , 55, doi:10.3390/risks10030055 . . . . .	<b>21</b>
<b>An Chen, Thai Nguyen and Thorsten Sehner</b> Unit-Linked Tontine: Utility-Based Design, Pricing and Performance † Reprinted from: <i>Risks</i> <b>2022</b> , <i>10</i> , 78, doi:10.3390/risks10040078 . . . . .	<b>63</b>
<b>Silvia Faroni, Olivier Le Courtois and Krzysztof Ostaszewski</b> Equivalent Risk Indicators: VaR, TCE, and Beyond Reprinted from: <i>Risks</i> <b>2022</b> , <i>10</i> , 142, doi:10.3390/risks10080142 . . . . .	<b>91</b>
<b>Anna Jędrzychowska</b> A Bridge Life Insurance for Households—Diagnosis and Motives Reprinted from: <i>Risks</i> <b>2022</b> , <i>10</i> , 81, doi:10.3390/risks10040081 . . . . .	<b>111</b>
<b>Agnieszka Marciniuk and Beata Zmyślona</b> Marriage and Individual Equity Release Contracts with Dread Disease Insurance as a Tool for Managing the Pensioners' Budget Reprinted from: <i>Risks</i> <b>2022</b> , <i>10</i> , 140, doi:10.3390/risks10070140 . . . . .	<b>133</b>
<b>Stefanos C. Orfanos</b> A Comparison of Macaulay Approximations Reprinted from: <i>Risks</i> <b>2022</b> , <i>10</i> , 153, doi:10.3390/risks10080153 . . . . .	<b>149</b>
<b>Silvana M. Pesenti</b> Reverse Sensitivity Analysis for Risk Modelling Reprinted from: <i>Risks</i> <b>2022</b> , <i>10</i> , 141, doi:10.3390/risks10070141 . . . . .	<b>157</b>
<b>Ermanno Pitacco and Daniela Y. Tabakova</b> Special-Rate Life Annuities: Analysis of Portfolio Risk Profiles Reprinted from: <i>Risks</i> <b>2022</b> , <i>10</i> , 65, doi:10.3390/risks10030065 . . . . .	<b>181</b>
<b>Jaap Spreeuw</b> The Copula Derived from the SAHARA Utility Function Reprinted from: <i>Risks</i> <b>2022</b> , <i>10</i> , 133, doi:10.3390/risks10070133 . . . . .	<b>203</b>



# About the Editor

## **Annamaria Olivieri**

Professor of Mathematical Methods for Economics, Actuarial Science and Finance, Department of Economics and Management, University of Parma (Italy).

Main research areas: valuation of the life insurance business, risk management for life insurance and pension funds, longevity risk, solvency for life portfolios, design of post-retirement benefits, multistate models for the insurances of the person, and pricing of life and health insurance products.





# Preface to "Actuarial Mathematics and Risk Management"

This is a reprint of a Special Issue of *Risks*, co-edited with prof. Ermanno Pitacco.

Prof. Pitacco left us, unexpectedly and sadly, before this Special Issue was closed. His contribution to it has been important. Prof. Pitacco was an outstanding scholar. His contribution to the development of actuarial mathematics has been sizable and will not be forgotten.

**Annamaria Olivieri**

*Editor*



# Special Issue “Actuarial Mathematics and Risk Management”

Annamaria Olivieri 

Department of Economics and Management, University of Parma, Via J.F. Kennedy 6, 43125 Parma, Italy; annamaria.olivieri@unipr.it

Among the most important implementations of the principles of enterprise risk management (ERM), the risk management process (RMP) involves various quantitative phases, usually encompassed under the label of quantitative risk management (QRM).

The RMP starts with defining the objectives (of an organization or a line of business) and then proceeding through the phases of risk identification, risk assessment, impact assessment, analysis of actions, choice of actions, and monitoring.

The whole RMP can benefit from the adoption of appropriate quantitative tools. In particular, the risk and impact assessments necessarily involve either stochastic evaluations (frequently implemented via Monte Carlo simulation procedures) or deterministic evaluations, such as sensitivity analysis and stress testing. The costs and efficiency of the alternative actions can be better understood in a quantitative framework. Statistical procedures are required for the monitoring phase, when observations must be elaborated and merged with initial assumptions, yielding updated input for a new cycle of the RMP.

Actuarial mathematics principles and tools can provide substantial support when implementing QRM phases, in particular when facing new risks or risks with changing features. Examples are provided by the development of products suitable for protecting individuals or organizations from emerging risks, the assessment of insurance product and portfolio risk profiles, the modeling of new risks or the revised modeling of traditional risks, and the study of effective risk measures. This background suggests that many areas of modeling and managing risks can benefit from novel research, aiming at both methodological and application innovation, in the insurance (life and non-life) context as well as in other economic sectors.

This Special Issue contributes in this regard with ten high-quality research papers addressing the following specific topics:

1. The design of post-retirement benefits (Chen et al. 2022; Pitacco and Tabakova 2022);
2. Designs of life and health insurance policies against new risks (Jędrzykowska 2022; Marciniuk and Zmysłona 2022);
3. Advancements in mortality modeling (Awad et al. 2022; Spreuw 2022);
4. Advancements in risk measures (Faroni et al. 2022) and risk models (Pesenti 2022);
5. Reserving disclosure tools (Breuer and Staudt 2022);
6. Innovative approximation formulae for the mean duration (Orfanos 2022).

In detail, Awad et al. (2022) discuss an extension of the Lee–Carter model. In particular, they propose a generalization of the Poisson log-bilinear Lee–Carter-type model by introducing a new class of families of counting distributions, namely, the ABM class, which belongs to a wider class of natural exponential families. This class is characterized by its variance functions and contains the Poisson and negative binomial distributions as special cases, offering an infinite class of additional counting distributions to be considered within the Lee–Carter framework. The results of a numerical study demonstrate that when fitting mortality data using this new class of distribution, superior results with respect to more traditional assumptions can be obtained in a number of situations.

Breuer and Staudt (2022) focus on equalization reserves, an insurance liability with features of own capital, with particular regard to the Swiss regulation. Although, according



**Citation:** Olivieri, Annamaria. 2023. Special Issue “Actuarial Mathematics and Risk Management”. *Risks* 11: 134. <https://doi.org/10.3390/risks11070134>

Received: 4 July 2023  
Accepted: 8 July 2023  
Published: 20 July 2023



**Copyright:** © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

to the local GAAP, Swiss reinsurers and non-life insurers must report equalization reserves in their statutory accounts, the solvency regulation does not admit them. As a result, the information about the equalization reserve is not fully disclosed. The purpose of the study is to recover that information and investigate the relationship between the equalization reserves and the publicly available technical account items. A generalized additive model (GAM) and a generalized linear model (GLM) were applied; based on publicly available data, the former proved to work better for reinsurers, whereas the latter worked better for nonlife insurers. The authors obtained encouraging results but also identified areas to be further investigated, such as the opportunity to link the equalization reserves to the insurance/reserving risk assessed from capital modeling.

Chen et al. (2022) address tontines as an alternative retirement product to conventional annuities. In particular, the authors introduce unit-linked tontines, which provide payments linked to an underlying financial asset. Two alternative designs are considered, differing with respect to the guarantee provided. First, the price is obtained using the risk-neutral approach; second, the attractiveness of the products is studied for a utility-maximizing individual. The findings of the numerical assessments stress the main difficulties of implementing retirement products.

Faroni et al. (2022) address the equivalence between VaR and TCE. The authors introduce a new risk indicator that extends TCE to consider higher-order risks. The quantiles of this indicator are compared with the quantiles of VaR in a simple Pareto framework and then in a generalized Pareto framework. The equivalence results between the quantiles of high-order TCEs are also examined.

The purpose of the study by Jędrzykowska (2022) is to describe the initial concept of household bridging insurance. After discussing the research gap regarding the insufficient protection of households against destabilization resulting from the lost personal contribution, the authors discuss the possibility of creating a new product, describing the desired features of its benefit structure.

Marciniuk and Zmyślona (2022) discuss products combining an equity release with a critical illness insurance; this is meant as a solution for protecting the living standards of individuals exposed to longevity risk. Two variants of the policy design are introduced: one addressed to couples and one to single individuals. The possible stream of benefits is analyzed for the two variants. The results suggest that the amount of cash flow related to reverse equity and critical illness insurance benefits depends on several factors, such as the spouse's economic status, age, and health condition.

Orfanos (2022) discusses issues related to the net present value of the cash flows exposed to interest rate risk. In particular, a new approximation formula for the Macaulay duration and convexity is described, which involves hyperbolic functions. The specific purpose of the study is to assess the reliability of each approximation formula under different scenarios. The results may be helpful in a number of actuarial implementations.

Pesenti (2022) proposes a reverse sensitivity analysis framework, which is model-free and allows for stresses on the output such as (a) the mean and variance, (b) any distortion risk measure including the value-at-risk and expected shortfall, and (c) expected utility type constraints. This framework is suitable for risk models. In particular, the author discusses a problem where a modeller needs to understand how a model consisting of random input factors, a corresponding random output of interest, and a baseline probability measure changes under a stress on the output's distribution. The findings not only provide a theoretical description of the stressed distribution but also show how to numerically efficiently calculate it.

Pitacco and Tabakova (2022) analyse special-rate life annuities, i.e., life annuity products rated considering the health status of the applicant. Better annuity rates are applied in the presence of poor health conditions, i.e., when an assumption of a shorter lifetime is acceptable. As the portfolio size should increase and as more potential annuitants can be attracted by more favorable annuity rates, a higher degree of heterogeneity of the portfolio follows as a result of including several risk classes. The pooling effect benefits from the

larger size but not from the increased heterogeneity. The purpose of the study is to analyze the impact on the variability of the total portfolio payout of extending the life annuity portfolio by selling special-rate life annuities. Numerical evaluations are performed by adopting deterministic and stochastic approaches, according to diverse assumptions concerning both lifetime distributions and the portfolio structure and size. The authors suggest that extending the annuity business by issuing special-rate annuities without significantly worsening the portfolio risk profile is possible.

Spreeuw (2022) introduces a new Archimedean copula family that is based on a link between Archimedean generators and utility functions. The family can well fit the mortality data of coupled lives. The parameter estimates suggest the possible existence of short-term dependence, i.e., the mortality of bereaved lives increases on bereavement but diminishes later.

All the papers part of this Special Issue underwent a refereeing process subject to the usual high standards of *Risks*. I would like to thank all the authors for their excellent contributions and all the referees for their thorough and timely reviews. I would also like to thank the MDPI Editorial Team for their active and timely support.

I have edited this Special Issue with Prof. Ermanno Pitacco. Our professional collaboration and our friendship have been lengthy and deep. The promotion of this Special Issue together is only one of the many projects that we had shared. Sadly and unexpectedly, Prof. Pitacco left us last September, when this Special Issue was still open. This was not the only project in which he was involved at that time. Although he was already retired, he was still very active in the field. During his whole professional life, he has deeply contributed to the development and dissemination of the actuarial culture, not only in Italy, but all over the world. He is greatly missed not only by his friends and family but also by the international actuarial community. He would have been proud to see how inspiring this Special Issue is.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

- Awad, Yaser, Shaul K. Bar-Lev, and Udi Makov. 2022. A New Class of Counting Distributions Embedded in the Lee–Carter Model for Mortality Projections: A Bayesian Approach. *Risks* 10: 111. [CrossRef]
- Breuer, Anja, and Yves Staudt. 2022. Equalization Reserves for Reinsurance and Non-Life Undertakings in Switzerland. *Risks* 10: 55. [CrossRef]
- Chen, An, Thai Nguyen, and Thorsten Sehner. 2022. Unit-Linked Tontine: Utility-Based Design, Pricing and Performance. *Risks* 10: 78. [CrossRef]
- Faroni, Silvia, Olivier Le Courtois, and Krzysztof Ostaszewski. 2022. Equivalent Risk Indicators: VaR, TCE, and Beyond. *Risks* 10: 142. [CrossRef]
- Jędrzychowska, Anna. 2022. A Bridge Life Insurance for Households—Diagnosis and Motives. *Risks* 10: 81. [CrossRef]
- Marciniuk, Agnieszka, and Beata Zmyślona. 2022. Marriage and Individual Equity Release Contracts with Dread Disease Insurance as a Tool for Managing the Pensioners' Budget. *Risks* 10: 140. [CrossRef]
- Orfanos, Stefanos C. 2022. A Comparison of Macaulay Approximations. *Risks* 10: 153. [CrossRef]
- Pesenti, Silvana M. 2022. Reverse Sensitivity Analysis for Risk Modelling. *Risks* 10: 141. [CrossRef]
- Pitacco, Ermanno, and Daniela Y. Tabakova. 2022. Special-Rate Life Annuities: Analysis of Portfolio Risk Profiles. *Risks* 10: 65. [CrossRef]
- Spreeuw, Jaap. 2022. The Copula Derived from the SAHARA Utility Function. *Risks* 10: 133. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Article

# A New Class of Counting Distributions Embedded in the Lee–Carter Model for Mortality Projections: A Bayesian Approach

Yaser Awad <sup>1</sup>, Shaul K. Bar-Lev <sup>2</sup> and Udi Makov <sup>3,\*</sup><sup>1</sup> Sakhnin Academic College, Galilee St. 100, Sakhnin 30810, Israel; awad\_y@netvision.net.il<sup>2</sup> Faculty of Industrial Engineering and Technology Management, HIT—Holon Institute of Technology, 52 Golomb Street, Holon 5810201, Israel; shaulb@hit.ac.il<sup>3</sup> Actuarial Research Center, University of Haifa, 199 Aba Khoushy Ave. Mount Carmel, Haifa 3498838, Israel

\* Correspondence: udimakov@gmail.com

**Abstract:** The Lee–Carter model, the dominant mortality projection modeling in the literature, was criticized for its homoscedastic error assumption. This was corrected in extensions to the model based on the assumption that the number of deaths follows Poisson or negative binomial distributions. We propose a new class of families of counting distributions, namely, the ABM class, which belongs to a wider class of natural exponential families. This class is characterized by its variance functions and contains the Poisson and the negative binomial distributions as special cases, offering an infinite class of additional counting distributions to be considered. We are guided by the principle that the choice of distribution should be made from a pool of distributions as large as possible. To this end, and following a data mining approach, a training set of historical mortality data of the population could be modeled using the ABM’s rich choice of distributions, and the chosen distribution should be the one that proved to offer superior projection results on a test set of mortality data. As an alternative to parameter estimation via the singular value decomposition used in the classical Lee–Carter model, we adopted Bayesian estimation, harnessing the Markov Chain Monte Carlo methodology. A numerical study demonstrates that when fitting mortality data using this new class of distributions, while traditional distributions may provide desirable projections for some populations, for others, alternative distributions within the ABM class can potentially produce superior results for the entire population or particular age groups, such as the oldest-old.

**Keywords:** Lee–Carter; counting distributions; mortality projections; natural exponential family

**Citation:** Awad, Yaser, Shaul K. Bar-Lev, and Udi Makov. 2022. A New Class of Counting Distributions Embedded in the Lee–Carter Model for Mortality Projections: A Bayesian Approach. *Risks* 10: 111. <https://doi.org/10.3390/risks10060111>

Academic Editors: Ermanno Pitacco and Annamaria Olivieri

Received: 24 March 2022

Accepted: 18 May 2022

Published: 27 May 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The seminal paper by Lee and Carter (1992) (LC) introduced a model which is one of the most well-known and widely applied models for forecasting mortality rates. Within this model, the time series of the log mortality rates,  $\ln m_{xt}$ , of each age is described by an age-specific intercept  $\alpha_x$  plus a common trend  $k_t$  for all age groups multiplied by an age-specific coefficient  $\beta_x$ ,

$$\ln m_{xt} = \alpha_x + \beta_x k_t + \varepsilon_{xt}.$$

The error term  $\varepsilon_{xt}$  is assumed to be distributed with a mean 0 and variance  $\sigma_\varepsilon^2$ , reflecting influences missed by the model. The age and time-specific mortality rate  $m_{xt}$  is calculated as  $(D_{xt}/E_{xt})$ , where  $D_{xt}$  denotes the number of deaths in a population at age  $x$ ,  $x = 1, 2, \dots, P$ , and at time  $t$ ,  $t = 1, 2, \dots, T$ , and  $E_{xt}$  is the exposure to the risk of death. To ensure the identifiability of model parameters, constraints are imposed such that the sum of  $\beta_x$  over age is 1 and the sum of  $k_t$  over time is 0. To forecast mortality rates into the future, a simple random walk with drift is proposed for  $k_t$ :

$$k_t = k_{t-1} + \theta + w_t.$$



The homoscedastic error assumption of the Lee–Carter model was criticized for its limiting impact on predictions (Brouhns et al. 2002; Danesi et al. 2015; Idrizi 2018). This led to the introduction of the Poisson log-bilinear LC-type model (Brouhns et al. 2002), which, in contrast, is intrinsically heteroscedastic, namely:

$$D_{xt} \sim \text{Poisson}(\mu_{xt}), \mu_{xt} = E_{xt}m_{xt}.$$

Here, the number of dead is directly modelled by a Poisson distribution, whose parameter is estimated by maximum likelihood estimation (MLE). This alternative approach gained momentum and alternative discrete distributions were proposed. In particular, a binomial distribution was proposed by Wang and Lu (2005) and a negative binomial distribution was suggested by Delwarde et al. (2007) and Renshaw and Haberman (2008). See Azman and Pathmanathan (2022) for further discussion of these distributions within the GLM framework.

The Lee–Carter model has been widely used for many purposes (Shair et al. 2018), such as forecasting mortality reduction factors, assessing the adequacy of retirement income, population projections and the projection of mortality trends for the oldest-old (older than 80, and in some sources 85). This age group is of considerable interest for policymakers as it is destined to grow as a proportion of the entire population and can outstrip existing infrastructures' capacity (Buettner 2002). This is a fairly recent phenomenon. In Canada, for instance (Legare et al. 2015), the 21st century brought about the most significant gain in life expectancy at age 85 (7.79% for women and 9.93% for men). Clearly, policies need to be devised that can meet people's special needs in what is called the fourth age Baltes and Smith (2003), and accurate mortality projection for this age group is a must. We shall, therefore, focus on the adequacy of potential underlying discrete distribution functions to produce accurate mortality projections using the Lee–Carter model for this age group. This will be discussed in Section 4.

In essence, while the above-cited papers relied on popular and commonly used discrete distributions, one cannot say that one particular distribution is universally superior. Indeed, it is entirely plausible to assume that there is a winning distribution for any given population or even for a specific population at a certain age range. Ideally, one should consider a rich class of family of counting distributions, much richer than the two already suggested, and use the data to pick the most suitable distribution for the population under study. This paper proposes an infinitely countable set of families of counting distributions, where the Poisson, negative binomial and Abel families of distributions are special cases. Our aim is to study this family, incorporate it into the framework of the LC model and use real data to seek the most suitable distribution for mortality projection. While there is little doubt that the distributions discussed above could prove adequate for specific populations or age groups, other distributions within the suggested family could have the upper hand.

The paper is organized as follows. Section 2 presents the new class of counting distributions. Section 3 is devoted to the new class and its Bayesian framework. Section 4 (divided into Section 4.1: Methods and Section 4.2: Results) reports a numerical study in which superior members of this class are chosen for mortality projections of the oldest-old in three populations. Finally, Section 5 offers a discussion.

## 2. A New Class of Counting Distributions on the Set of Nonnegative Integers $\mathbb{N}_0$

The new class of families of counting distributions on the non-negative integers belongs to a wider class of natural exponential families (NEFs), characterized by their variance functions (VFs). In order to comprehend this class we decompose this section into subsections. We first present some preliminaries on NEFs and their associated VFs. We then introduce a class of NEFs having polynomial structure and then suggest the new class of families of counting distributions, named ABM, first introduced by Awad et al. (2016), where the class was defined and its usefulness for mortality projections was preliminary sketched. Furthermore, such a class has been investigated by Bar-Lev and Ridder (2021a, 2021b) from

a classical frequency approach and has shown superiority with respect to various metrics or goodness-of-fit tests for different count datasets (for further details see item 6 in Section 2.3).

### 2.1. NEFs—Some Preliminaries

The following preliminaries are mainly taken from Letac and Mora (1990) and are briefly presented here for completeness.

Let  $\nu$  be a non-Dirac positive Radon measure on  $\mathbb{R}$ , and  $L(\theta) = \int e^{\theta x} \nu(dx)$  its Laplace transform. Assuming that  $\Theta = \text{int}\{\theta \in \mathbb{R} : L(\theta) < \infty\} \neq \emptyset$ , then the NEF generated by  $\nu$  is defined by the probability distributions

$$\mathcal{F} = \left\{ F_\theta : F_\theta(\nu(dx)) = e^{\theta x - \kappa(\theta)} \nu(dx), \theta \in \Theta \right\}, \tag{1}$$

where  $\kappa(\theta) = \log L(\theta)$ , the cumulant transform of  $\nu$ , is strictly convex and real analytic on  $\Theta$ . If  $X_\theta$  represents a r.v. having distribution  $F_\theta$  of the form given in (1) then the expectation and variance of  $X_\theta$  are given, respectively, by  $E(X_\theta) \doteq m = \kappa'(\theta)$  and  $V(X_\theta) = \kappa''(\theta)$  where  $m = \kappa'(\theta)$  is strictly monotone and thus its inverse, say,  $\theta = \psi(m), m \in M = \kappa'(\Theta)$  is well defined. The set  $M$  of all means of (1) is called the mean parameter space of  $\mathcal{F}$ . The variance of  $F_\theta$  can be expressed in terms of  $m$  by  $V(m) = \kappa''(\theta) = \kappa''(\psi(m))$ . The pair  $(V, M)$  is called the VF of  $\mathcal{F}$  and it uniquely determines  $\mathcal{F}$  within the class of NEFs. For example,  $(m, \mathbb{R}^+)$  and  $(m^2, \mathbb{R}^+)$  are, respectively, the VFs of the Poisson and exponential NEFs and are uniquely determined by them.

### 2.2. The Mean Value Parametrization of NEFs

As indicated above, the VF of an NEF  $\mathcal{F}$  uniquely determines  $F$  within the class of NEFs. Let  $(V, M)$  be a given VF of an NEF  $\mathcal{F}$  generated by  $\nu$ . Then, simple calculations show both  $\theta = \psi(m)$  and the cumulant transform  $\kappa(\theta) = \kappa(\psi(m))$  of  $\nu$  can be expressed in terms of  $m$  as:

$$\theta = \psi(m) = \int \frac{dm}{V(m)} + c_1, \quad k(\theta) = k(\psi(m)) = \int \frac{m}{V(m)} dm + c_2, \tag{2}$$

where one needs to determine the constants  $c_1$  and  $c_2$  so that  $F_\theta, \theta \in \Theta$ , constitutes a probability distribution (not an easy task). Accordingly, a mean value parametrization of an NEF  $\mathcal{F}$  generated by a measure  $\nu$  is given by:

$$\mathcal{F} = \{ \exp\{\psi(m)x - k(\psi(m))\} \nu(dx), m \in M \}. \tag{3}$$

Such a representation of  $\mathcal{F}$  is more natural as it is expressed in terms of the mean  $m$  rather than a somewhat artificial parameter  $\theta$ . A comprehensive description of NEFs in terms of their mean value representation is reviewed in Bar-Lev and Kokonendji (2017).

**Remark 1.** *The task of computing the constants  $c_1$  and  $c_2$  is not simple and might be rather cumbersome. However, from a Bayesian perspective, when (3) is used as a prior distribution on  $m$ , then in the calculation of the respective posterior distribution, such constants are cancelled out (as the likelihood function is the only relevant component). As this paper is concerned with a Bayesian framework, one can assume without any loss of generality that  $c_1 = c_2 = 0$ . Henceforth, we indeed assume so.*

### 2.3. Polynomial VFs of Counting NEFs Supported on the Set of Nonnegative Integers $\mathbb{N}_0$

The innovative and breakthrough Proposition 4.4 of Letac and Mora (1990, p. 13) provided conditions under which a given VF  $(V, M)$  is associated with a counting NEF  $F$  supported on the set of non-negative integers  $\mathbb{N}_0$ , i.e., where all members of  $F$  are composed of counting distributions on  $\mathbb{N}_0$ . They provided general examples of two classes of VFs which fulfill the premises of their Proposition 4.4 and thus their associated NEFs'

distributions are supported on the non-negative integers. One of these two classes has the form:

$$V(m) = m \prod_{i=1}^k \left(1 + \frac{m}{p_i}\right), p_i > 0, i = 1, \dots, k, k \in \mathbb{N}_0, M = \mathbb{R}^+, \text{ where } \prod_{i=1}^0 \doteq 1. \quad (4)$$

They proved that such VFs constitute counting NEFs supported on  $\mathbb{N}_0$ , namely, counting distributions with non-negative integer support. Moreover, their Proposition 4.4 enables to compute (at least theoretically and numerically) the corresponding measure  $\nu$  (we skip details as they are irrelevant for our Bayesian framework analysis). Note that the two special cases of (4) with  $k = 0$  and  $k = 1$  correspond, respectively, to the Poisson and negative binomial NEFs. However, the general setting (4) for  $k \geq 3$  does not allow an explicit calculation of  $\theta = \psi(m)$  and  $k(\theta) = k(\psi(m))$  in (2), implying that the mean value parametrization of the corresponding NEFs in the form (3) is not explicitly expressible in terms of  $m$  and thus becomes useless for any practical consideration.

#### 2.4. A New Class of Polynomial VFs—The ABM NEFs

As we already noted, the fact that a given pair  $(V, M)$  is known to be a VF of some NEF does not necessarily enable the construction of the corresponding mean value parameterization (3), as in most cases the integrals for  $\psi(m)$  and  $k(\psi(m))$  in (2) are not explicitly expressible analytically in closed forms, and indeed, this is the situation for the class (4) in its general form. Consequently, one needs to search for subclasses of (4) for which the integrals in (2) can be computed explicitly. One such special subclass takes the above point into consideration. Indeed, by taking in (4) the special case where

$$p_1 = p_2 = \dots = p_k,$$

and denoting

$$p_2 \doteq k \in \mathbb{N}_0,$$

we obtain a subclass of (4) with VFs with the form:

$$(V, M) = m \left(1 + \frac{m}{p_1}\right)^{p_2}, \mathbb{R}^+, p_1 > 0, p_2 \in \mathbb{N}_0. \quad (5)$$

As (5) is a subclass of (4) and (4) satisfies the premises of Proposition 4.4 of Letac and Mora (1990) it follows that the subclass (5) are VFs associated with counting NEFs supported on the non-negative integers.

The subclass of VFs in (5) (hereafter called the ABM class) was first introduced by Awad et al. (2016) who showed that the corresponding  $\psi(m)$  and  $k(\psi(m))$  (calculated from (2)) have, as opposed to the general form in (4), the following closed forms (the exact proof details appear in Bar-Lev and Kokonendji 2017):

$$\theta = \psi(m) = \ln \frac{m}{p_1 + m} + \sum_{i=1}^{p_2-1} \frac{1}{i} \frac{p_1^i}{(p_1 + m)^i} + c_1, \text{ where } \sum_{i=1}^0 = 0,$$

and

$$\kappa(\psi(m)) = -\frac{p_1^{p_2}}{(p_2 - 1)(m + p_1)^{p_2-1}} + c_2.$$

Thus, its mean value parametrization is given by the probability distribution:

$$F(m, \nu(dx)) = \exp \left\{ x \left[ \ln \frac{m}{p_1 + m} + \sum_{i=1}^{p_2-1} \frac{1}{i} \frac{p_1^i}{(p_1 + m)^i} + c_1 \right] + \frac{p_1^{p_2}}{(p_2 - 1)(m + p_1)^{p_2-1}} + c_2 \right\}, m \in \mathbb{R}^+, p_1 > 0, p_2 \in \mathbb{N} \quad (6)$$

where hereafter we denote this probability distribution by  $ABM(p_1, p_2)$ , where  $p_1$  is a positive real number and  $p_2$  is a non-negative integer. (For a classical frequency approach, the constants  $c_1$  and  $c_2$  have been computed by Bar-Lev and Ridder 2021b). However, as noted above, for a Bayesian framework they are cancelled out when computing the posterior distribution and thus can be taken to be  $c_1 = c_2 = 0$  without any loss of generality).

Note that the ABM class of VFs  $\left\{m\left(1 + \frac{m}{p_1}\right)^{p_2}, p_1 > 0\right\}_{p_2 \in \mathbb{N}_0}$ , or alternatively, the corresponding class  $\{ABM(p_1, p_2)\}_{p_2 \in \mathbb{N}_0}$  of NEFs is composed of an infinitely countable set of families of counting NEFs supported on the non-negative integers. As special cases, this class contains the Poisson NEF ( $p_2 = 0$ ), the negative binomial NEF ( $p_2 = 1$ ) and the Abel NEF ( $p_2 = 2$ ), (c.f., Letac and Mora 1990, p. 31; Bar-Lev and Ridder 2019, for applications to car accident claims of a Swedish insurance company dataset).

Summarizing, this ABM NEF has the following features:

1. It is a class of counting distributions supported on the non-negative integers;
2. It is overdispersed as  $V(m)/m > 1$ ;
3. It allows a mean value parameterization in a closed form;
4. It is infinitely divisible, which allows the construction of an exponential dispersion model (EDM) with dispersion parameter space equal to  $\mathbb{R}^+$ . EDMs are used to describe the error distribution in generalized linear models (see Jorgensen 1987, 1997);
5.  $p_1$  is an unknown parameter to be estimated (see next section).  $p_2 \in \mathbb{N}_0$  is a parameter governing the particular model within the ABM class and is considered to be a decision variable (note that different values of  $p_2$  determine different ABM NEFs). Accordingly, for given national datasets (i.e., those of US, Ireland and Ukraine), the goal will be to locate that value of  $p_2$ , which minimizes a respective RMSE (see in the sequel). However, due to the rather cumbersome and intractable structure of the ABM probabilities (or likelihood) in (6) and the fact that the larger the  $p_2$ , the larger the number of elements in the summands appearing in (6), no analytic solution for an optimal  $p_2$  is feasible at all for achieving such a goal. Consequently, only numerical search algorithms are plausible. The search starts with  $p_2 = 0$  (the Poisson NEF),  $p_2 = 1$  (the negative binomial NEF),  $p_2 = 2$  (the Abel NEF) and so on;
6. As already noted, the ABM class  $\{ABM(p_1, p_2)\}_{p_2 \in \mathbb{N}_0}$  is composed of infinitely countable set of families of counting NEFs supported on the non-negative integers and thus can also be used to model real datasets by employing the classical frequency approach (and not only Bayesian). Indeed, the ABM class has been compared in Bar-Lev and Ridder (2021a, 2021b) with other common counting probability models (such as Poisson-inverse Gaussian distribution, new logarithmic distribution, an exponentiated discrete Lindley distribution) for various real count datasets stemming from automobile insurance claims, marketing, biometry, health, and social sciences (none of which is related to mortality projections). Members of the ABM counting class have shown superiority with respect to various metrics for goodness-of-fit tests (chi-squared test, Akaike information criterion (AIC), root-mean-square error (RMSE) and Kullback–Leibler divergence (KL)), and provided a much better fit for each of the datasets considered (more details can be found in Bar-Lev and Ridder 2021b).

### 3. ABM Based LC Model and its Bayesian Framework

As an alternative to parameter estimation via the singular value decomposition used in the classical LC model or the MLE in the cases discussed above, we adopt the Bayesian approach which offers advantages succinctly expressed in Antonio et al. (2015): a. The calibration and forecast steps are combined, which leads to more consistent estimates of the period effects; b. The Bayesian approach provides a natural framework for incorporating parameter uncertainty in mortality forecasts, which is relevant—for example—in the new insurance regulatory framework of Solvency II. The Bayesian approach allows adequate handling of small populations and missing data. Like Czado et al. (2005) and Pedroza (2006), we harness the power of the Markov Chain Monte Carlo (MCMC) methodology to

estimate the model parameters and execute mortality projection. We note that the interest in Bayesian solutions in the context of mortality projections has recently gained momentum (Ellison et al. 2020; Graziani 2020; Hilton et al. 2019; Hunt and Blake 2020; Kogure et al. 2019; Liu et al. 2020; Njenga and Sherris 2020; Wong et al. 2018).

Suppose the number of deaths  $D_{xt}$  in a population at age  $x$  and time  $t$  is distributed as follow:

$$D_{xt} \sim ABM(p_1, p_2)(\mu_{xt}), \mu_{xt} = E_{xt}m_{xt}, m_{xt} = \exp(\alpha_x + \beta_x k_t),$$

where

$$\alpha = (\alpha_{x_{\min}}, \dots, \alpha_{x_{\max}})', \beta = (\beta_{x_{\min}}, \dots, \beta_{x_{\max}})', k = (k_{t_{\min}}, \dots, k_{t_{\max}}).$$

Bayesian estimation of the unknown parameters  $\alpha, \beta, k$  and  $p_1$  are based on the joint posterior distribution function of  $\alpha, \beta, k$  and  $p_1$  given  $(E_{xt}, D_{xt})$ , when  $x = x_{\min}, x_{\min} + 1, \dots, x_{\max}, t = t_{\min}, t_{\min} + 1, \dots, t_{\max}$ . The first step in the Bayesian estimation is to determine the prior probability functions for these parameters.

**The prior distribution for  $k_t$  and  $\theta$**

Let  $k_t = k_{t-1} + \theta + w_t$ , and let  $w_t \sim N(0, \sigma_w^2)$  and hence  $k_t \sim N(\theta, t\sigma_w^2)$ . we assume  $\sigma_w^{-2} \sim \text{gamma}(a_k, b_k)$  and  $\theta \sim N(\theta_0, \sigma_\theta^2)$ . The hyper-parameters  $\theta_0, a_k, b_k$  and  $\sigma_\theta^2$  are arbitrary initial values.

**The prior distribution for  $\beta_x$**

We assume  $\beta_x \sim N(0, \sigma_\beta^2) \forall x$ , where  $\sigma_\beta^{-2} \sim \text{gamma}(a_\beta, b_\beta)$ . The hyper-parameters  $a_\beta, b_\beta$  are arbitrary initial values.

**The prior distribution for  $\alpha_x$**

We suppose that the prior distribution of  $\alpha_x \sim N(\alpha_{0x}, \sigma_\alpha^2) \forall x$ , where  $\sigma_\alpha^{-2} \sim \text{gamma}(a_\alpha, b_\alpha)$ . The hyper-parameters  $\alpha_{0x}, a_\alpha$  and  $b_\alpha$  are arbitrary initial values.

**The prior distribution for  $p_1$**

We let  $p_1 \sim \text{gamma}(a_{p_1}, b_{p_1})$ . The hyper-parameters  $a_{p_1}, b_{p_1}$  are arbitrary initial values.

*MH (Metropolis–Hastings) Algorithm for Estimating the Parameters  $\alpha, \beta, k$  and  $p_1$*

Suppose the  $D_{xt}$  's are independent random variables, which are distributed as (6) and  $g(\Xi)$  is the joint prior distribution of the unknown parameters  $\Xi = (\alpha, \beta, k, p_1)'$ . Then, the posterior distribution of  $\Xi$ , given all available data  $D = \{d_{xt}\}$  and  $p_2$ , can be represented as follows:

$$f(\Xi | D, p_2) \propto \prod_x \prod_t \exp \left[ d_{xt} \left( \ln \frac{E_{xt}m_{xt}}{p_1 + E_{xt}m_{xt}} + \sum_{i=1}^{p_2-1} \frac{1}{i} \frac{p_1^i}{(p_1 + E_{xt}m_{xt})} \right) + \frac{p_1^{p_2}}{(p_2-1)(p_1 + E_{xt}m_{xt})^{p_2-1}} \right] \times g(\Xi).$$

See Appendix A for the marginal posterior distributions of  $\alpha, \beta, k$  and  $p_1$ . We now describe the estimation of  $\alpha, \beta, k$  and  $p_1$  using the MH, conditioned on the data and all other parameters at their respective iterations. The superscript denotes the iteration number of the parameter of interest.

**Estimation of  $k_t$  using the MH algorithm**

Let the marginal posterior distribution of  $k_t$  be  $f(k_t | D, \alpha, \beta, k_{-t}, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1)$ . The estimation of  $k_t$ , is achieved by the following steps, where.

1. Draw  $k_t^*$  from the proposal density function  $N(k_t^{(i)}, \sigma_t^2)$ , such that  $\sigma_t^2$  is assumed known;
2. Calculate the following probability:

$$\Psi(k_t^{(i)}, k_t^*) = \min\left(1, \frac{f(k_t^* | D, \alpha, \beta, k_{-t}^{(i)}, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1)}{f(k_t^{(i)} | D, \alpha, \beta, k_{-t}^{(i)}, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1)}\right),$$

where  $k_{-t} = (k_{t\min}, \dots, k_{t-1}, k_{t+1}, \dots, k_{t\max})'$ ;

3. Draw a value  $u$  from uniform probability function in range  $U(0, 1)$  and decide in accordance with the following formula:

$$\left\{ \begin{array}{l} \text{if } u \leq \Psi(k_t^{(i)}, k_t^*) \text{ then } k_t^{(i+1)} = k_t^* \\ \text{if } u > \Psi(k_t^{(i)}, k_t^*) \text{ then } k_t^{(i+1)} = k_t^{(i)}; \end{array} \right\}.$$

4. Going over all values of  $t$ , we have:

$$k^{(i+1)} = (k_{t\min}^{(i+1)}, \dots, k_t^{(i+1)}, k_{t+1}^{(i)}, \dots, k_{t\max}^{(i)})'$$

5. Transforming  $k^{(i+1)}$  and  $\alpha^{(i)}$  to assure identifiably:

$$k^{(i+1)} - \bar{k} \rightarrow k^{(i+1)}, \alpha^{(i)} + \beta^{(i)}\bar{k} \rightarrow \alpha^{(i)},$$

where

$$\bar{k} = \frac{1}{T} \left( \sum_{j \leq t} k_j^{(i+1)} + \sum_{j > t} k_j^{(i)} \right);$$

6. Repeat steps 1 to 5.

### Estimation of $\beta_x$ using MH algorithm

Let the marginal posterior distribution of  $\beta_x$  be  $f(\beta_x | D, \alpha, \beta_{-x}, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1)$ . The estimation of  $\beta_x$  is achieved by the following steps.

1. Draw  $\beta_x^*$  from the proposal density function  $N(\beta_x^{(i)}, \sigma_\beta^2)$ , such that  $\sigma_\beta^2$  is assumed to be known;
2. Calculate the following probability:

$$\Psi(\beta_x^{(i)}, \beta_x^*) = \min\left(1, \frac{f(\beta_x^* | D, \alpha, \beta_{-x}, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1)}{f(\beta_x^{(i)} | D, \alpha, \beta_{-x}, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1)}\right),$$

where

$$\beta_{-x} = (\beta_{x\min}, \dots, \beta_{x-1}, \beta_{x+1}, \dots, \beta_{x\max})'$$

3. Draw a value  $u$  from uniform probability function in range  $U(0, 1)$  and decide in accordance with the following formula:

$$\left\{ \begin{array}{l} \text{if } u \leq \Psi(\beta_x^{(i)}, \beta_x^*) \text{ then } \beta_x^{(i+1)} = \beta_x^* \\ \text{if } u > \Psi(\beta_x^{(i)}, \beta_x^*) \text{ then } \beta_x^{(i+1)} = \beta_x^{(i)}; \end{array} \right\}.$$

4. Going over all values of  $x$ , we have:

$$\beta^{(i+1)} = (\beta_{x\min}^{(i+1)}, \dots, \beta_x^{(i+1)}, \beta_{x+1}^{(i)}, \dots, \beta_{x\max}^{(i)})'$$

- Transforming  $k^{(i+1)}$  and  $\beta^{(i+1)}$  to assure identifiably:

$$\frac{\beta^{(i+1)}}{\beta_{sum}} \rightarrow \beta^{(i+1)}, k^{(i+1)} \times \beta_{sum} \rightarrow k^{(i+1)},$$

where

$$\beta_{sum} = \left( \sum_{j \leq x} \beta_j^{(i+1)} + \sum_{j > x} \beta_j^{(i)} \right);$$

- Repeat steps 1 to 5.

### Estimation of $\alpha_x$ using MH algorithm

Let the marginal posterior distribution of  $\alpha_x$  be  $f(\alpha_x | D, \alpha_{-x}, \beta, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1)$ . The estimation of  $\alpha_x$ , is achieved by the following steps;

- Draw  $\alpha_x^*$  from the proposal density function  $N(\alpha_x^{(i)}, \sigma_\alpha^2)$ , such that  $\sigma_\alpha^2$  is assumed known;
- Calculate the following probability:

$$\Psi(\alpha_x^{(i)}, \alpha_x^*) = \min \left( 1, \frac{f(\alpha_x^* | D, \alpha_{-x}^{(i)}, \beta, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1)}{f(\alpha_x^{(i)} | D, \alpha_{-x}^{(i)}, \beta, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1)} \right),$$

where

$$\alpha_{-t} = (\alpha_{x_{\min}}, \dots, \alpha_{x-1}, \alpha_{x+1}, \dots, \alpha_{x_{\max}})';$$

- Draw a value  $u$  from uniform probability function in range  $U(0, 1)$  and decide in accordance with the following formula:

$$\left\{ \begin{array}{l} \text{if } u \leq \Psi(\alpha_x^{(i)}, \alpha_x^*) \text{ then } \alpha_x^{(i+1)} = \alpha_x^* \\ \text{if } u > \Psi(\alpha_x^{(i)}, \alpha_x^*) \text{ then } \alpha_x^{(i+1)} = \alpha_x^{(i)} \end{array} \right\}.$$

- Receiving  $\alpha^{(i+1)}$  in  $(i + 1)$ th iteration as follows:

$$\alpha_x^{(i+1)} = (\alpha_{x_{\min}}^{(i+1)}, \dots, \alpha_x^{(i+1)}, \alpha_{x+1}^{(i)}, \dots, \alpha_{x_{\max}}^{(i)});$$

- Repeat steps 1 to 4.

### Estimation of $p_1$ using MH algorithm

Let the marginal posterior distribution of  $p_1$  be  $f(p_1 | D, \alpha, \beta, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2)$ , proportional to the product of the likelihood (6) and the gamma prior distribution of  $p_1$ . The estimation of  $p_1$  is achieved by the following steps;

- Draw  $p_1^*$  from the probability function  $gamma(\alpha_{p_1}, b_{p_1})$ , such that  $\alpha_{p_1}$  and  $b_{p_1}$  are hyperparameters and are assumed known;
- Calculate the following probability:

$$\Psi(p_1^{(i)}, p_1^*) = \min \left( 1, \frac{f(p_1^* | D, \alpha, \beta, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2)}{f(p_1^{(i)} | D, \alpha, \beta, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2)} \right);$$

3. Draw a value  $u$  from uniform probability function in range  $U(0, 1)$  and decide in accordance with the following formula:

$$\left\{ \begin{array}{l} \text{if } u \leq \Psi(p_1^{(i)}, p_1^*) \text{ then } p_1^{(i+1)} = p_1^* \\ \text{if } u > \Psi(p_1^{(i)}, p_1^*) \text{ then } p_1^{(i+1)} = p_1^{(i)}; \end{array} \right\}.$$

4. Then receiving  $p^{(i+1)}$  in  $(i + 1)$ th iteration;
5. Repeat steps 1 to 4.

### Estimation of $\theta, \sigma_\alpha^2, \sigma_\beta^2$ and $\sigma_w^2$ using the Gibbs sampler

The Gibbs sampler can be used for estimating  $\theta, \sigma_\alpha^2, \sigma_\beta^2$  and  $\sigma_w^2$  since the marginal posterior distribution of these parameters can be written explicitly (See: Czado et al. 2005). The following are the marginal posterior sampling distributions of each of these parameters, conditioned on the data and all other parameters at their respective iterations.

1. Sampling  $\theta$ :

The posterior probability function of the parameter  $\theta$ , is presented as follows:

$$f(\theta | D, \alpha, \beta, k, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1) = f(\theta | k, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1) \propto f(k | \theta, \sigma_w^2) f(\theta).$$

The prior probability function of the parameter  $\theta$  is  $N(\theta_0, \sigma_\theta^2)$ , and the hyper-parameters  $\theta_0$  and  $\sigma_\theta^2$  are set by the user, hence the posterior probability function of the parameter is:

$$(\theta | k, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2) \sim N\left(\frac{\theta_0 \sigma_w^2}{T \sigma_\theta^2 + \sigma_w^2}, (T \sigma_\theta^2 + \sigma_w^2)^{-1}\right).$$

2. Sampling  $\sigma_\alpha^2$ :

The posterior probability function of the parameter  $\sigma_\alpha^2$  is presented as follows:

$$f(\sigma_\alpha^2 | D, \alpha, \beta, k, \theta, \sigma_\beta^2, \sigma_w^2, p_1) \propto f(\alpha | \sigma_\alpha^2) f(\sigma_\alpha^2).$$

The prior probability function of the parameter  $\sigma_\alpha^2$  such that  $\sigma_\alpha^{-2} \sim \text{gamma}(a_\alpha, b_\alpha)$ , and the hyper-parameters  $a_\alpha$  and  $b_\alpha$  are set by the user, so the posterior probability function of the parameter is:

$$(\sigma_\alpha^{-2} | D, \alpha, \beta, k, \theta, \sigma_\beta^2, \sigma_w^2, p_1) \sim \text{gamma}\left(a_\alpha + \frac{x_{\max}}{2}, b_\alpha + \frac{1}{2} \sum_{x=x_{\min}}^{x_{\max}} (\alpha_x - \bar{\alpha})^2\right),$$

where

$$\bar{\alpha} = \frac{1}{x_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \alpha_x.$$

3. Sampling  $\sigma_\beta^2$ :

The posterior probability function of the parameter  $\sigma_\beta^2$  is presented as follows:

$$f(\sigma_\beta^2 | D, \alpha, \beta, k, \theta, \sigma_\alpha^2, \sigma_w^2, p_1) \propto f(\beta | \sigma_\beta^2) f(\sigma_\beta^2).$$

The prior probability function of the parameter  $\sigma_\beta^2$  such that  $\sigma_\beta^{-2} \sim \text{gamma}(a_\beta, b_\beta)$ , and the hyper-parameters  $a_\beta$  and  $b_\beta$  are set by the user, so the posterior probability function of the parameter is:

$$(\sigma_\beta^{-2} | D, \alpha, \beta, k, \theta, \sigma_\alpha^2, \sigma_w^2, p_1) \sim \text{gamma}\left(a_\beta + \frac{x_{\max}}{2}, b_\beta + \frac{1}{2} \sum_{x=x_{\min}}^{x_{\max}} (\beta_x)^2\right).$$



#### 4. Sampling $\sigma_w^2$ :

The posterior probability function of the parameter  $\sigma_w^2$ , is presented as follows:

$$f(\sigma_w^2 | D, \alpha, \beta, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, p_1) \propto f(k | \sigma_w^2) f(\sigma_w^2).$$

The prior probability function of the parameter  $\sigma_w^2$  such that  $\sigma_w^{-2} \sim \text{gamma}(a_k, b_k)$ , and the hyper-parameters  $a_k$  and  $b_k$  are set by the user, so the posterior probability function of the parameter is:

$$(\sigma_w^{-2} | D, \alpha, \beta, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, p_1) \sim \text{gamma}\left(a_k + \frac{T}{2}, b_k + \frac{1}{2} \sum_{t=t_{\min}}^{t_{\max}} (k_t - k_{t-1} - \theta)^2\right).$$

## 4. Numerical Experiment

### 4.1. Methods

To test the adequacy of the ABM class, we analyzed mortality data of men in Ireland, Ukraine and the USA, downloaded from the database of Human Mortality Database (<https://www.mortality.org> accessed on 8 March 2013). The data contain the number of dead and the size of the population exposed to risk by age and year; Ireland's and the USA's data are for 1950–2007, and Ukraine's data are for 1959–2009. This analysis aims to examine sixteen models within the ABM class,  $p_2 = 0, \dots, 15$ , with a particular emphasis on forecasting the mortality of the oldest-old (see below an argumentation for the restriction of the values of  $p_2$  to  $\{0, \dots, 15\}$ ). These models also include the Poisson and negative binomial models which feature widely in the literature, for which  $p_1 = 0$  and  $p_2 = 1$ , respectively. Adopting a data mining approach, the models were fitted using training sets and were examined using test sets. The training sets contained data up to 2000 and the test sets, aimed at monitoring the quality of predictions, contained data from 2001 to 2007 for Ireland and the USA, and from 2001 to 2009 for Ukraine. Predictions are carried out with the estimated parameters,  $\ln[m_{x,t+s}] = \hat{\alpha}_x + \hat{\beta} * \hat{k}_{t+s}$ ,  $s = 1, 2, \dots$ , where model performance (using the test sets) was checked using the root of the mean squared errors (RMSE), which was calculated in two ways:

1. Predicting mortality rates ( $\mu$ ) by age. In other words, after model parameters were estimated, mortality rates were predicted for a given age across years. For instance, predicting mortality rates for those age 70 was carried out over the years beyond 2000;
2. Predicting mortality rates ( $\mu$ ) by cohort. In other words, after model parameters were estimated, mortality rates were predicted for a cohort that was at a particular age at the beginning of the test period. For example, predicted mortality rates in 2001–2007 for a cohort aged 70 in 2001.

For every member of the ABM class (controlled by  $p_2$ ), the Markov chains used to obtain posterior distributions/parameter estimates comprised 4000 iterations with the first 1000 considered a burn-in period. Convergence was established using graphical means and a sensitivity analysis ascertained that the choice of arbitrary initial hyper-parameters did not affect the final outcomes. We report the outcomes for the Poisson distribution ( $p_2 = 0$ ) and the negative binomial distribution ( $p_2 = 1$ ). In addition, we examined the ABM members for which  $p_2 \in \{2, \dots, 15\}$ . We limited our reporting to  $p_2 \in \{2, \dots, 15\}$  since, for the data under study, increasing  $p_2$  beyond 15 (our study explored all models up to  $p_2 = 50$ ) did not alter our findings of the optimal  $p_2$  for varying ages and resulted in much larger RMSE than those found up to  $p_2 = 15$ . In practice, we analyzed all 16 models within the  $p_2$  range and, for each, we estimated  $p_1$  as well as all other unknown parameters. Finally, we reported graphically the RMSE for the Poisson, negative binomial and for the ABM member which produced the minimal RMSE for various ages.

4.2. Results

Figure 1a,b shows Ireland’s mortality projections by age and cohort, respectively. A superior model for an age range is the one which produced the smallest RMSE. It is evident that the Poisson performs best for most ages above 70, with the negative binomial lagging behind. However, at a very old age, the Poisson diverges and the best ABM member to be chosen instead is  $ABM(\cdot, p_2 = 3)$ . A similar result is shown for the USA (Figure 2a,b), except that the best ABM for the very old is  $ABM(\cdot, p_2 = 4)$ . A different picture emerges when we focus on Ukraine’s mortality projections by age and cohort (Figure 3a,b, respectively). While the Poisson and negative binomial perform well (with Poisson being better), for ages above 96 (by cohort) or 104 (by age), the negative binomial drifts away, leaving  $ABM(\cdot, p_2 = 10)$  to be the winner of the ABM class. Clearly, the recommended projection policy for Ukraine is to use the Poisson for most ages but to rely on  $ABM(\cdot, p_2 = 10)$  for very old ages. Naturally, other countries with their specific national datasets may yield different ABM models (i.e., different  $p_2$ ’s) for mortality projections.

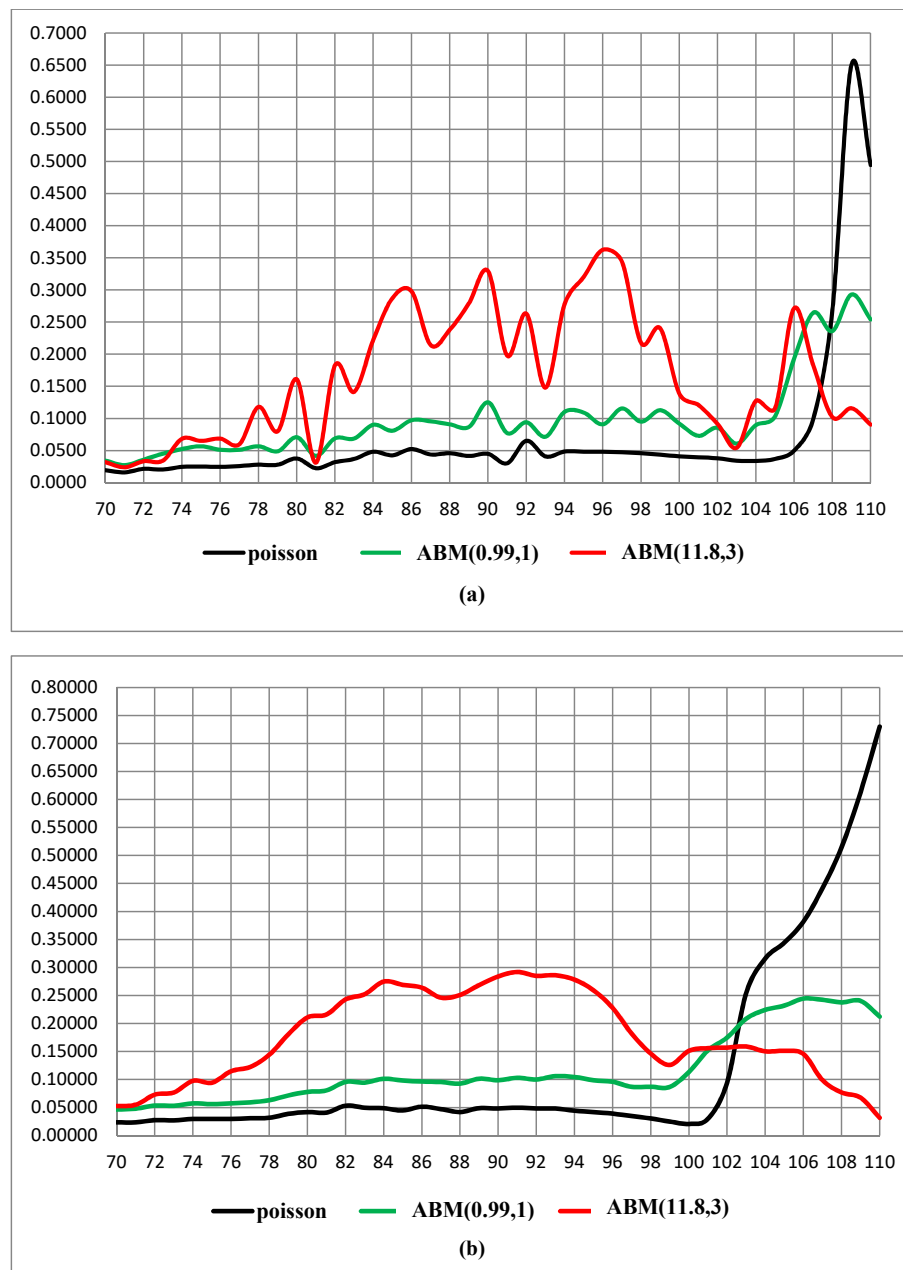


Figure 1. RMSE for projecting mortality rate, Ireland, 2001–2007. (a) by age; (b) by cohort.

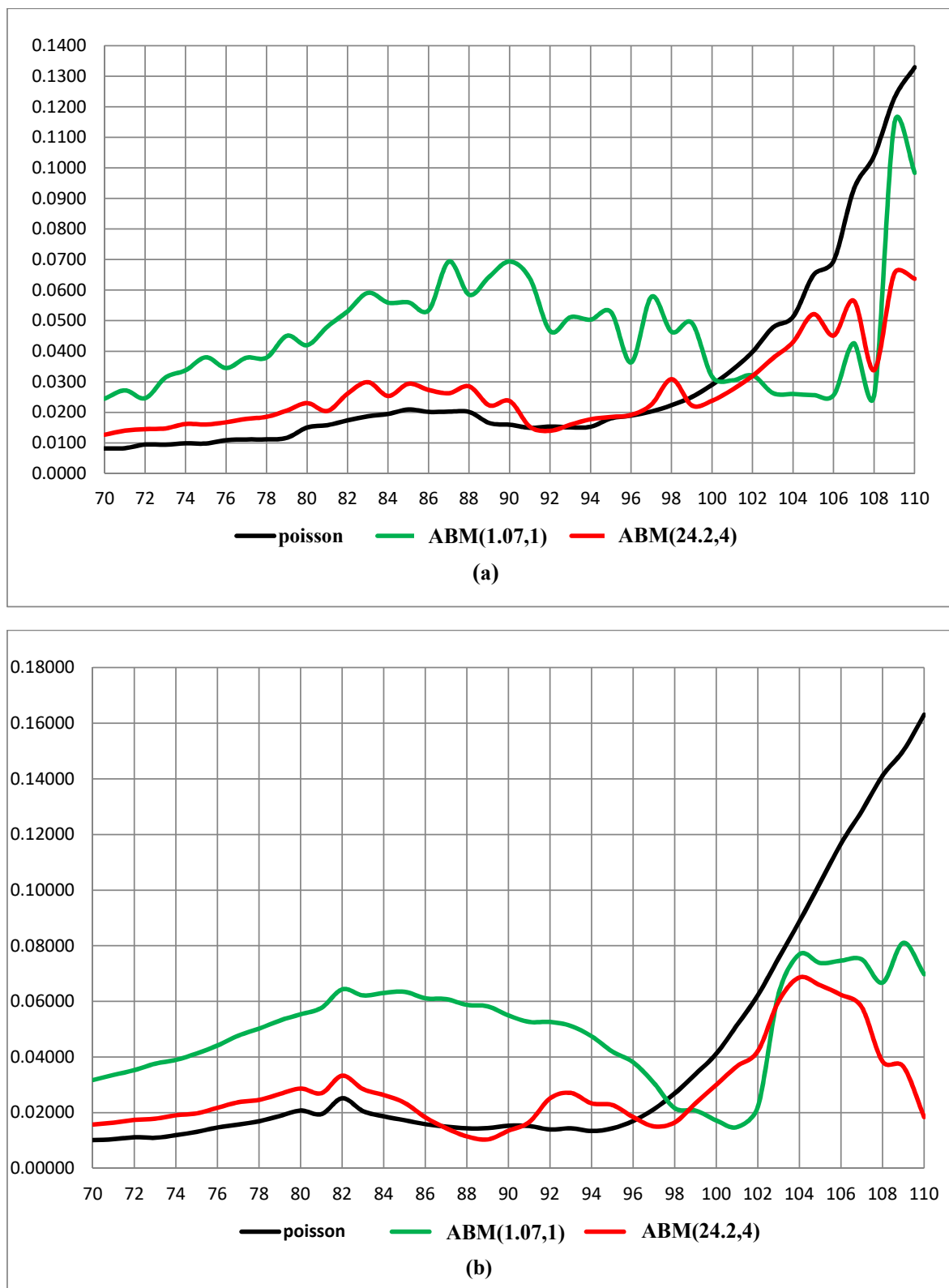


Figure 2. RMSE for projecting mortality rate, USA, 2001–2007. (a) by age; (b) by cohort.

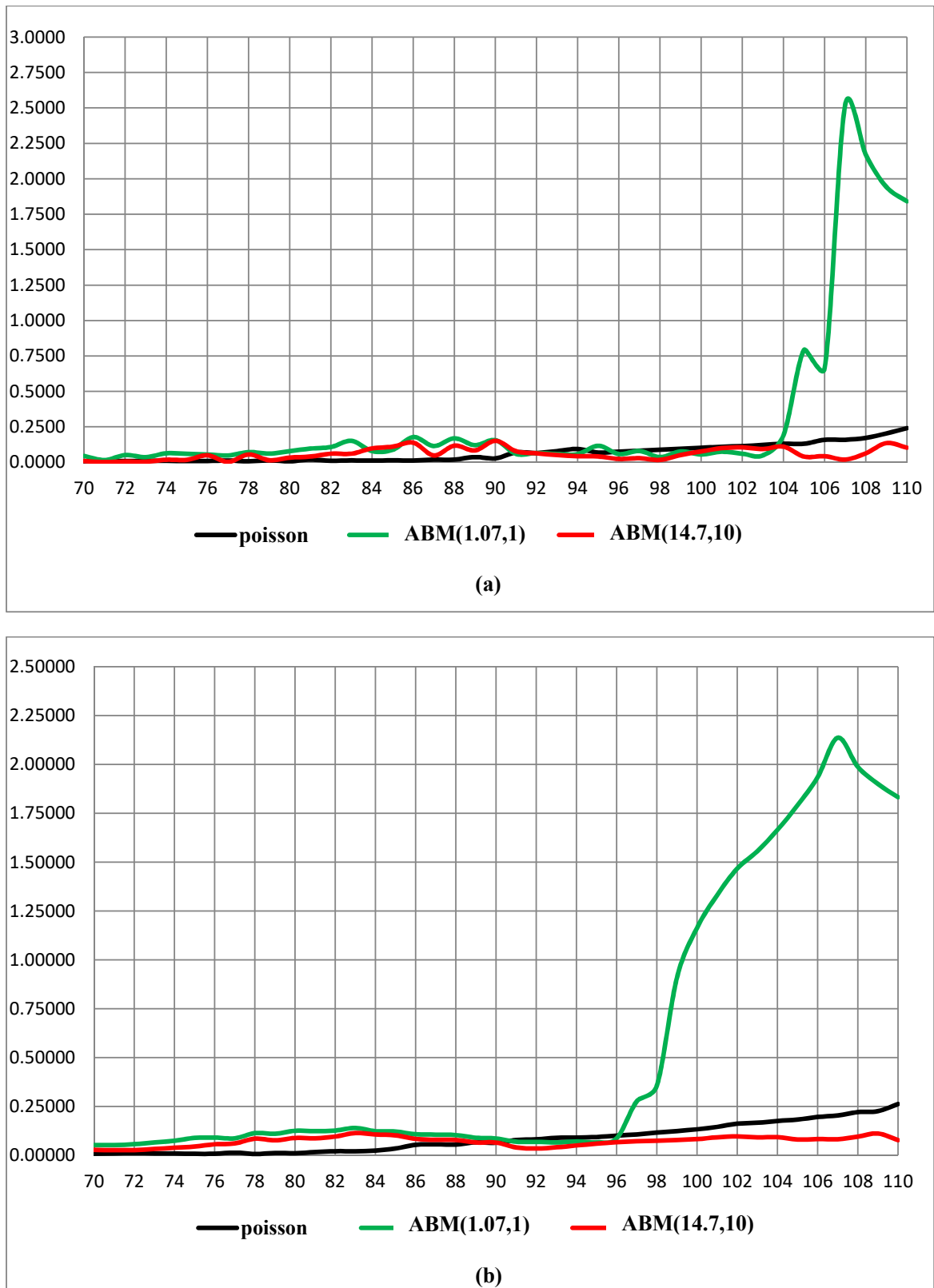


Figure 3. RMSE for projecting mortality rate, Ukraine, 2001–2009. (a) by age; (b) by cohort.

### 5. Discussion

Several extensions to the LC model assume that the number of deaths is distributed Poisson or negative binomial. These distributions have offered adequate mortality projections in several populations reported in the literature. It is not implausible that cases where

these two failed were not reported. Rather than deciding a priori to choose a particular distribution, we aimed to enrich the LC model by allowing a richer pool of candidate distributions. The chosen distribution would be the one providing the best projection for the population and age range under study. To achieve this goal, we proposed a new class of counting distributions on the non-negative integers, the ABM class, which belongs to a wider class of natural exponential families characterized by their variance functions. This class includes the Poisson and negative binomial distributions which are included in an infinitely countable set of additional members. A data mining approach was adopted whereby the model is fitted using a training set and tested using a test set with the RMSE used to pick the winning model. As an alternative to parameter estimation via the singular value decomposition (SVD) used in the classical LC model, we adopted Bayesian estimation, harnessing the Markov Chain Monte Carlo (MCMC) methodology. While we do not suggest that MCMC is superior to SVD (for a comparison of the two, see Ichikawa et al. 2021), we still promote the former since the Bayesian framework frees us from the burden of calculating the normalizing constant of the ABM. This is indeed a great plus, even though running the MCMC requires more computer time for the mid-size databases of the kind used for national mortality projections. We note, however, that the use of MCMC is rather costly and might fail if the dataset is huge. So, perhaps other Bayesian techniques such as Variational Bayes can be more helpful. However, employing such a suggestion is beyond the scope of this paper. We examined ABM models for three countries and established that, for the countries examined, the commonly adopted Poisson distribution is justified except for a very old age for which an alternative member of the ABM class offers better projection. We do not claim that to suggest a superior model. When deciding on an underlying model, one can adopt as an example the Poisson model or the negative binomial model. Rather than adopting a model, we suggest adopting a class of models (the ABM) comprising the Poisson, negative binomial, and numerous other counting distributions. The superiority of this approach lies in the ability to choose a model amongst candidate models. Since no one single model necessarily fits every population and every age group well, the ABM class could allow picking, as an example, the Poisson for members of the population aged under 50, the negative binomial for those aged 50 to 80, and another member of the class to the oldest-old. The suggested criteria for preferring one member of the class over another is the mean squared projection errors (RMSE). This advantage is gained at the cost of additional complexity, which is justified given the financial benefits associated with more accurate modeling. We conclude that it is no longer appropriate to assume a single distribution for the whole process of mortality projection. Instead, for every country and every relevant range of ages, a desirable approach is to pick a member of the ABM class that provides the best mortality projection. In the numerical study reported here, neither the Poisson nor the negative binomial distributions adequately serve the very oldest-old and superior alternatives are within reach in the suggested novel ABM class of distributions.

**Author Contributions:** Conceptualization, Y.A., S.K.B.-L. and U.M.; methodology, Y.A., S.K.B.-L. and U.M.; software, Y.A.; validation, Y.A., S.K.B.-L. and U.M.; formal analysis, Y.A., S.K.B.-L. and U.M.; investigation, Y.A., S.K.B.-L. and U.M.; resources, Y.A., S.K.B.-L. and U.M.; data curation, Y.A. and U.M.; writing—original draft preparation, Y.A., S.K.B.-L. and U.M.; writing—review and editing, Y.A., S.K.B.-L. and U.M.; visualization, Y.A., S.K.B.-L. and U.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Publicly available datasets were analyzed in this study. This data can be found here: the database of Human Mortality Database (<https://www.mortality.org> access on 8 March 2013).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

### Appendix A.1

The marginal posterior probability function of the time index  $k_t$  is

$$f(k_t | D, \alpha, \beta, k_{-t}, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1) \propto f(D_{t_{\min}} | k_{t_{\min}}, \alpha, \beta, p_1) \times f(k_{t_{\min}} | \theta, \sigma_w^2) \times \prod_{j=t_{\min}+1}^{t_{\max}} f(D_j | k_j, \alpha, \beta, p_1) \times f(k_j | k_{j-1}, \theta, \sigma_w^2),$$

where

$$f(D_t | \alpha, \beta, k_t, p_1) = \prod_x \exp \left[ d_{xt} \left( \ln \frac{E_{xt} \mu_{xt}}{p_1 + E_{xt} \mu_{xt}} + \sum_{i=1}^{p_2-1} \frac{1}{i} \frac{p_1^i}{(p_1 + E_{xt} \mu_{xt})} \right) + \frac{p_1^{p_2}}{(p_2-1)(p_1 + E_{xt} \mu_{xt})^{p_2-1}} \right].$$

For the remaining expressions we distinguish between three cases:

1. For  $t = t_{\min}$ , the marginal posterior probability function of the time index  $k_t$  is:

$$f(k_t | D, \alpha, \beta, k_{-t}, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1) \propto f(D_t | \alpha, \beta, k_t, p_1) \times f(k_t | \theta, \sigma_w^2) \times f(k_{t+1} | k_t, \theta, \sigma_w^2),$$

where

$$f(k_t | \theta, \sigma_w^2) = \exp \left( -\frac{1}{2\sigma_w^2} (k_t - \theta)^2 \right) \text{ and}$$

$$f(k_{t+1} | k_t, \theta, \sigma_w^2) = \exp \left( -\frac{1}{2\sigma_w^2} (k_t - k_{t-1} - \theta)^2 \right).$$

2. For  $t_{\min} < t < t_{\max}$ , the marginal posterior probability function of the time index  $k_t$  is:

$$f(k_t | D, \alpha, \beta, k_{-t}, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1) \propto f(D_t | \alpha, \beta, k_t, p_1) \times f(k_t | k_{t-1}, \theta, \sigma_w^2) \times f(k_{t+1} | k_t, \theta, \sigma_w^2),$$

where

$$f(k_t | k_{t-1}, \theta, \sigma_w^2) = \exp \left( -\frac{1}{2\sigma_w^2} (k_t - k_{t-1} - \theta)^2 \right) \text{ and}$$

$$f(k_{t+1} | k_t, \theta, \sigma_w^2) = \exp \left( -\frac{1}{2\sigma_w^2} (k_{t+1} - k_t - \theta)^2 \right).$$

3. For  $t = t_{\max}$ , the marginal posterior probability function of the time index  $k_t$  is:

$$f(k_t | D, \alpha, \beta, k_{-t}, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1) \propto f(D_t | \alpha, \beta, k_t, p_1) \times f(k_t | k_{t-1}, \theta, \sigma_w^2),$$

where

$$f(k_t | k_{t-1}, \theta, \sigma_w^2) = \exp \left( -\frac{1}{2\sigma_w^2} (k_t - k_{t-1} - \theta)^2 \right).$$

### Appendix A.2

The marginal posterior probability function of  $\beta_x$  is

$$f(\beta_x | D, \alpha, \beta_{-x}, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1) \propto \prod_{j=x_{\min}}^{x_{\max}} f(D_j | \alpha, \beta_j, k, p_1) \times f(\beta_j) \propto f(D_x | \alpha, \beta_x, k, p_1) \times f(\beta_x),$$

where

$$f(D_x | \alpha, \beta_x, k, p_1) = \prod_t \exp \left[ d_{xt} \left( \ln \frac{E_{xt} \mu_{xt}}{p_1 + E_{xt} \mu_{xt}} + \sum_{i=1}^{p_2-1} \frac{1}{i} \frac{p_1^i}{(p_1 + E_{xt} \mu_{xt})} \right) + \frac{p_1^{p_2}}{(p_2-1)(p_1 + E_{xt} \mu_{xt})^{p_2-1}} \right],$$

and

$$f(\beta_x) = \exp\left(-\frac{1}{2\sigma_\beta^2}(\beta_x)^2\right).$$

### Appendix A.3

The marginal posterior probability function of  $\alpha_x$  is

$$f(\alpha_x | D, \alpha_{-x}, \beta, k, \theta, \sigma_\alpha^2, \sigma_\beta^2, \sigma_w^2, p_1) \\ \propto \prod_{j=x_{\min}}^{x_{\max}} f(D_j | \alpha_j, \beta, k, p_1) \times f(\alpha_j) \propto f(D_x | \alpha_x, \beta, k, p_1) \times f(\alpha_x),$$

where

$$f(D_x | \alpha_x, \beta, k, p_1) = \\ \prod_t \exp\left[d_{xt} \left( \ln \frac{E_{xt}\mu_{xt}}{p_1 + E_{xt}\mu_{xt}} + \sum_{i=1}^{p_2-1} \frac{1}{i} \frac{p_1^i}{(p_1 + E_{xt}\mu_{xt})} \right) + \frac{p_1^{p_2}}{(p_2-1)(p_1 + E_{xt}\mu_{xt})^{p_2-1}} \right],$$

and

$$f(\beta_x) = \exp\left(-\frac{1}{2\sigma_k^2}(\alpha_x - \alpha_{0x})^2\right).$$

## References

- Antonio, Katrien, Anastasios Bardoutsos, and Wilbert Ouburg. 2015. Bayesian Poisson log-bilinear models for mortality projections with multiple populations. *Eur. Actuar. J.* 5: 245–281. [CrossRef]
- Awad, Yaser, Shaul K. Bar-Lev, and Udi Makov. 2016. *A New Class Counting Distributions Embedded in the Lee–Carter Model for Mortality Projections: A Bayesian Approach*. Technical Report No. 146. Israel: Actuarial Research Center, University of Haifa.
- Azman, Shafiqah, and Dharini Pathmanathan. 2020. The GLM framework of the Lee–Carter model: A multi-country study. *Journal of Applied Statistics* 49: 752–63. [CrossRef]
- Baltes, B. Paul, and Jacqui Smith. 2003. New frontiers in the future of aging: From successful aging of the young old to the dilemmas of the fourth age. *Gerontology* 49: 123–35. [CrossRef] [PubMed]
- Bar-Lev, K. Shaul, and Ad Ridder. 2019. Monte Carlo methods for insurance risk computation. *International Journal of Statistics and Probability* 8: 54–74. [CrossRef]
- Bar-Lev, K. Shaul, and Ad Ridder. 2021a. Exponential dispersion models for overdispersed zero-inflated count data. *Communications in Statistics-Simulation and Computation*. Available online: <https://doi.org/10.1080/03610918.2021.1934020> (accessed on 10 August 2021).
- Bar-Lev, K. Shaul, and Ad Ridder. 2021b. New exponential dispersion models for count data—The ABM and LM classes. *ESAIM: Probability and Statistics* 25: 31–52. [CrossRef]
- Bar-Lev, K. Shaul, and Clestin C. Kokonendji. 2017. On the mean value parametrization of natural exponential families—A revisited review. *Mathematical Methods of Statistics* 26: 159–75. [CrossRef]
- Brouhns, Natacha, Michel Denuit, and Jeroen K. Vermunt. 2002. A Poisson log-bilinear regression approach to the construction of project life-table. *Insurance: Mathematics and Economics* 31: 373–93.
- Buettner, Thomas. 2002. Approaches and experiences in projecting mortality patterns for the oldest-old. *North American Actuarial Journal* 6: 14–29. [CrossRef]
- Czado, Claudia, Antoine Delwarde, and Michel Denuit. 2005. Bayesian Poisson log-bilinear mortality projections. *Insurance: Mathematics and Economics* 36: 260–84. [CrossRef]
- Danesi, L. Ivan, Steven Haberman, and Pietro Millosovich. 2015. Forecasting mortality in subpopulations using Lee–Carter type models: A comparison. *Insurance: Mathematics and Economics* 62: 151–61. [CrossRef]
- Delwarde, Antoine, Michel Denuit, and Christian Partrat. 2007. Negative binomial version of the Lee–Carter model for mortality forecasting. *Applied Stochastic Models in Business and Industry* 23: 385–401. [CrossRef]
- Ellison, Joanne, Erengul Dodd, and Jonathan J. Forster. 2020. Forecasting of cohort fertility under a hierarchical Bayesian approach. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 183: 829–56. [CrossRef]
- Graziani, Rebecca. 2020. Stochastic Population Forecasting: A Bayesian Approach Based on Evaluation by Experts. *Developments in Demographic Forecasting* 49: 21–42.
- Hilton, Jason, Erengul Dodd, Jonathan J. Forster, and Peter W. F. Smith. 2019. Projecting UK mortality by using Bayesian generalized additive models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 68: 29–49. [CrossRef]
- Hunt, Andrew, and David Blake. 2020. A Bayesian Approach to Modeling and Projecting Cohort Effects. *North American Actuarial Journal* 25: S235–S254. [CrossRef]



- Ichikawa, Shota, Hiroyuki Yamamoto, and Takumi Morita. 2021. Comparison of a Bayesian estimation algorithm and singular value decomposition algorithms for 80-detector row CT perfusion in patients with acute ischemic stroke. *La Radiologia Medica* 126: 795–803. [CrossRef]
- Idrizi, Olgerta. 2018. The Heteroscedasticity Impact on Actuarial Science: Lee Carter Error Simulation. *European Journal of Engineering and Formal Sciences* 1: 1–12. [CrossRef]
- Jorgensen, Bent. 1987. Exponential dispersion models (with discussion). *Journal of the Royal Statistical Society, Series B* 49: 127–62.
- Jorgensen, Bent. 1997. *The Theory of Exponential Dispersion Models*. Monographs on Statistics and Probability 76. London: Chapman and Hall.
- Kogure, Atsuyuki, Takahiro Fushimi, and Shinichi Kamiya. 2019. Mortality Forecasts for Long-Term Care Subpopulations with Longevity Risk: A Bayesian Approach. *North American Actuarial Journal* 25: S534–44. [CrossRef]
- Lee, D. Ronald, and Lawrence R. Carter. 1992. Modelling and forecasting the time series of US mortality. *Journal of the American Statistical Association* 87: 659–71.
- Légaré, Jacques, Yann Décarie, Kim Deslandes, and Yves Carrière. 2015. Canada's oldest old: A population group which is fast growing, poorly apprehended and at risk from lack of appropriate services. *Population Change and Lifecourse Strategic Knowledge Cluster Discussion Paper Series/Un Réseau Stratégique de Connaissances Changements de Population et Parcours de vie Document de Travail* 3: 1. Available online: <https://ir.lib.uwo.ca/pclc/vol3/iss2/1> (accessed on 1 April 2021).
- Letac, Gerard, and Marianne Mora. 1990. Natural real exponential families with cubic variance functions. *Annals of Statistics* 18: 1–37. [CrossRef]
- Liu, Zhen, Xiaoqian Sun, Leping Liu, and Yu-Bo Wang. 2020. Bayesian Poisson Log-normal Model with Regularized Time Structure for Mortality Projection of Multi-population. *arXiv* arXiv:2010.04775.
- Njenga, N. Carolyn, and Michael Sherris. 2020. Modeling mortality with a Bayesian vector autoregression. *Insurance: Mathematics and Economics* 94: 40–57. [CrossRef]
- Pedroza, Claudia. 2006. A Bayesian forecasting model: predicting U.S male mortality. *Biostatistics* 7: 530–50. [CrossRef]
- Renshaw, Arthur, and Steven Haberman. 2008. On simulation-based approaches to risk measurement in mortality with specific reference to Poisson Lee–Carter modelling. *Insurance: Mathematics and Economics* 42: 797–816. [CrossRef]
- Shair, N. Syazreen, Muhammad A. S. Rosmizan, Mohammad J. M. S. Ting, and Mohd A. A. Zaini. 2018. Projected Malaysian Lifetable: Evaluations of The Lee–Carter and Poisson Log-Bilinear Models. *International Journal of Modern Trends in Social Sciences* 1: 60–72.
- Wang, Duolao, and Pengjun Lu. 2005. Modelling and forecasting mortality distributions in England and Wales using the Lee–Carter model. *Journal of Applied Statistics* 32: 873–85. [CrossRef]
- Wong, S. T. Jackie, Jonathan J. Forster, and Peter W. F. Smith. 2018. Bayesian mortality forecasting with overdispersion. *Insurance: Mathematics and Economics* 83: 206–21. [CrossRef]





Article

# Equalization Reserves for Reinsurance and Non-Life Undertakings in Switzerland

Anja Breuer <sup>1,\*</sup>  and Yves Staudt <sup>2,3</sup> 

<sup>1</sup> Department of Actuarial Science, Faculty of Business and Economics (HEC Lausanne), University of Lausanne, Extranef, 1015 Lausanne, Switzerland

<sup>2</sup> Institute for Tourism and Leisure, Department Alpine Region Development, University of Applied Sciences of the Grisons, Comercialstrasse 19, 7000 Chur, Switzerland; yves.staudt@fhgr.ch

<sup>3</sup> Center of Data Analysis, Simulation and Visualization, Department Applied Future Technologies, University of Applied Sciences of the Grisons, Ringstrasse 34, 7000 Chur, Switzerland

\* Correspondence: anja.breuer@unil.ch

**Abstract:** Equalization reserves is an insurance liability with features of own capital. By law, Swiss reinsurance and non-life undertakings must hold equalization reserves within their statutory accounts. Regarding Swiss solvency modeling, the equalization reserves are set to zero. Swiss reinsurance and non-life undertakings define the upper limit and the corresponding transfer rule to the equalization reserves; however, this information is not disclosed. The goal of the study is to find a relationship between the equalization reserves and the publicly available technical account items, applying a generalized additive model (GAM). Thereafter, we transform the continuous variables into discrete ones, and we apply a generalized linear model (GLM). The study is based on published data from 1997 to 2018, whereby we restate the implicitly published equalization reserves. For reinsurance undertakings, the GAM model captures the relationship better than the GLM one; for non-life undertakings, the GLM model performs better. For reinsurance undertakings, the equalization reserves depend on the equalization reserves of the previous year, on the calendar year, on the legal form, on the technical result, on the administration and commission costs and on other costs. For non-life undertakings, the equalization reserves depend on the net claims payments, on the equalization reserves of the previous year, on the net change in claims reserves without change in equalization reserves, on the calendar year and on the net earned premium. Furthermore, we look at the need for equalization reserves: do the undertakings accumulate and release the equalization reserves? Further, the impact of taxes on the equalization reserves is looked at. The concept of equalization reserves avoids the misuse of tax optimization. We conclude that the discussion about disclosure of equalization reserves will restart. In addition, the definition of the upper limit of the equalization reserves could be widened by linking the equalization reserves to the insurance/reserving risk from the capital modeling.

**Keywords:** equalization reserves; GAM; GLM; reinsurance; non-life; Switzerland; tax

**Citation:** Breuer, Anja, and Yves Staudt. 2022. Equalization Reserves for Reinsurance and Non-Life Undertakings in Switzerland. *Risks* 10: 55. <https://doi.org/10.3390/risks10030055>

Academic Editors: Ermanno Pitacco and Annamaria Olivieri

Received: 2 February 2022

Accepted: 25 February 2022

Published: 3 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Within the nature of the insurance business, and well-documented, for instance, by Wüthrich and Merz (2008) or by Farny (2011), the determination of claims reserves is related to uncertainty. Models are used to reflect complex circumstances. Parameters are estimated, resulting in parameter risks. Unpredictable loss occurrence is observed and caused, for instance, by random fluctuations. Changes in the claims handling process or legislative amendments bear further unexpected uncertainties. Safety considerations finally join the reason to build up equalization reserves for reinsurance and non-life undertakings, requested by the Swiss Financial Market Supervisory Authority—(FINMA 2008a, rank no. 8; 2011, rank no. 37). For example, a natural catastrophe (CAT) has a low probability of occurrence, and by occurring, CAT causes a high loss. In good financial times, the insurance undertaking put aside the equalization reserve to dampen the future loss burden,

see Dacorogna et al. (2013). These authors explain how even shareholders benefit from the equalization reserves.

### 1.1. Equalization Reserves in Other Countries

Hindley (2017) defines “equalization reserve” in the UK as an amount to smooth the reserves over time, typically considering low frequency/high severity events. Akhurst et al. (1992) point out the philosophical viewpoint of holding either a specific reserve or a surplus solvency fund after tax to finance a future large claim. Furthermore, non-technical elements such as investment and management costs could require a kind of equalization reserve in some countries, see Akhurst et al. (1992). These authors present the calculation and the handling of the equalization reserves for the UK, Denmark, Finland, France, Germany, Italy, The Netherlands, Norway, Spain and Sweden.

### 1.2. Swiss Insurance Supervision

In order to protect creditors, investors and policyholders, FINMA supervises, among others, private insurance companies, complying with the Insurance Supervision Act (ISA)<sup>1</sup> and with the Ordinance on the Supervision of Private Insurance Companies (ISO)<sup>2</sup>. FINMA has regulatory power and regulates using ordinances and circulars as stated in the core tasks, see FINMA (2019b). FINMA’s predecessor was the Federal Office of Private Insurance (FOPI), being responsible until 2008. In Switzerland, reinsurance and non-life undertakings are obliged to hold equalization reserves within their statutory accounts, regardless of whether these entities cover low frequency/high severity events. The handling of the equalization reserves has to be described in the technical part of the business plan of an undertaking. Appendix C of ISO-FINMA (2015) does not explicitly list the equalization reserves to be disclosed.

The reserves of the Swiss capital model (SST) are based on the best estimate principle, as stated by FINMA (2017): the insurance liability includes all future financial expenditures to fulfill those liabilities, excluding the capital costs. As a consequence, the equalization reserves are set to zero. Target capital is the calculated capital to meet the quantitative requirements under the SST, see FINMA (2018a). The target capital has to be covered by own funds. Insurance/reserving risk is part of this target capital. In the case in which the actual case reserves would not be sufficient, the shareholder’s own fund would fill this gap. Due to safety considerations, the SST requests for insurance/reserving risk and the statutory accounts request for equalization reserves. However, within SST, the insurance/reserving risk is covered by own funds, and within the statutory accounts, the equalization reserves are insurance liabilities.

Due to solvency requirements, the focus of the recent actuarial studies lay on market-consistent valuation of the balance sheet. Wüthrich (2016) presents market-consistent actuarial valuation of the insurance reserves. In the literature the classical claims reserving methods are well studied, see Wüthrich and Merz (2008). On the other hand, the calculation or the handling of the statutory required equalization reserves are not explicitly covered. FINMA’s circulars clarify the legal requirements.

### 1.3. Literature Review

The Swiss Association of Actuaries, SAV (2006), publishes guidelines for loss reserves in non-life insurance, focusing on required reserves for claims handling costs. However, the equalization reserves are not mentioned. FINMA’s circulars define the legal scope to handle this kind of reserve for reinsurance and for non-life undertakings. Hindley (2017) discusses the equalization reserves in the UK. Akhurst et al. (1992) present equalization reserves modeling in UK, Denmark, Finland, France, Germany, Italy, The Netherlands, Norway, Spain and Sweden. De Vylder and Goovaerts (1999) present a theoretical evaluation of the equalization reserves. A broad literature regarding Swiss equalization reserves does not exist. This article steps forward to fill the gap.

#### 1.4. Present Research

We start defining “equalization reserves” separately for reinsurance and for non-life undertakings. Thereafter, we look at the handling of the equalization reserves within International Financial Reporting Standards 17 (IFRS 17) and within the capital modeling. As tax consideration impacts the insurance undertakings, we look closer at the mechanism from an actuarial point of view. We do not know how the undertakings calculate the equalization reserve. FOPI and FINMA implicitly disclose the equalization reserves. Therefore, we describe our method to estimate the equalization reserves from the reported statutory accounts and we apply a model to detect the relationship between the equalization reserves and the statutory accounts items.

This paper is structured as follows: in Section 2, we define equalization reserves and we look closer at accounting and capital model aspects. Hereafter, in Section 3, we formalize the model and based on this, we consider taxes. Further, in Section 4, we concretize the formalism on the publicly available data, we present the general approach of the generalized additive model (GAM), the generalized linear model (GLM) and the model results. Finally, in Section 5, we discuss the results and in Section 6 we conclude.

## 2. Definition, Accounting and Capital Model Aspects

In Switzerland, equalization reserves are defined separately for reinsurance and for non-life undertakings. In principle, these definitions are the same for both types of undertakings. However, there are small differences, which we point out in the following. In Section 2.1, we define equalization reserves for reinsurance and for non-life undertakings in Switzerland. Thereafter in Section 2.2, we look at the handling of the equalization reserves within statutory accounting, within IFRS and within the Swiss capital modeling.

### 2.1. Definition

In Switzerland, the ISA distinguishes between legal entities, writing direct business, for instance, non-life undertakings, and legal entities only writing indirect business for reinsurance undertakings. Non-life undertakings control each single risk, by selecting the risk and by complying with the underwriting policy. In general, reinsurance undertakings protect the portfolio of undertakings, whereby the single risks could not be controlled. In consequence, the risk profiles of non-life undertakings and reinsurance undertakings are different and a separate analysis is appropriate. In this section, we define equalization reserves for reinsurance and for non-life undertakings in Switzerland.

#### 2.1.1. Equalization Reserves for Reinsurance Undertakings

In Switzerland, the equalization reserves<sup>3</sup> for reinsurance undertakings are defined as the technical reserve, balancing unfavorable development and fluctuations of incurred claims, see FINMA (2011, rank no. 37). According to FINMA (2011, rank no. 8\*) reinsurance undertaking could keep equalization reserves within the statutory accounts and are requested to specify methods and principles to establish and to dissolve equalization reserves within the business plan in accordance with Art. 4, section 2, letter d of ISA.

#### 2.1.2. Equalization Reserves for Non-Life Undertakings

In Switzerland, the equalization reserves<sup>4</sup> for non-life undertakings are defined as the technical reserves, wholly or partially balancing unfavorable development and fluctuation of incurred claims, see FINMA (2008a, rank no. 8). In particular, the equalization reserves comprise:

- The reserves covering parameter risk, safety considerations and unpredictable loss fluctuation according to FINMA (2008a, rank no. 16);
- Equalization reserves for the credit insurance business, see FINMA (2008a, rank no. 18).

There is no general standard formula for the non-life undertakings to calculate the equalization reserves<sup>5</sup>. Thus, each company determines its own method. According to

Art. 69 of ISO, the equalization reserves are part of the total technical reserves of a non-life insurance undertaking. FINMA (2008a, rank no. 14 ff) requests the non-life undertaking to specify methods and principles to establish and to dissolve equalization reserves within the business plan in accordance with Art. 4, section 2, letter d of ISA.

## 2.2. Accounting and Capital Model Aspects

In this section, we look at the handling of the equalization reserves within statutory accounting, within IFRS and within the Swiss capital modeling.

### 2.2.1. Statutory Accounting

The required equalization reserves are technical reserves reported as a liability in the statutory balance sheet, see ISO-FINMA (2015). For non-life undertakings, according to Art. 71 of ISO, the equalization reserves are part of the amount covered by the tied assets. From the Swiss statutory accounting point of view, the equalization reserves belong to the policyholders. However, the equalization reserves cannot be uniquely assigned to a single customer. In the case of discontinuation of a company, the release of the equalization reserves could strengthen either the case reserves or the own capital, taking into account the tax requirements.

### 2.2.2. IFRS

Regarding IFRS 17, the best estimate of liabilities should be among others “current”, reflecting the most recent existing information, for instance, see Dunne et al. (2017). As a consequence, within IFRS 17, future catastrophe losses occurring outside the contract boundaries are not considered within the insurance liabilities. Nevertheless, within the Swiss statutory accounting, these catastrophe losses could be covered by the equalization reserves.

### 2.2.3. Capital Modeling

The SST assesses the capital solvency situation of an insurance undertaking, see FINMA (2018b). The valuation of the liability as “best estimate” for claims, incurring prior to the reference date of the balance sheet, is the basis for the risk-bearing capital-calculation according to FINMA (2017). As a consequence, equalization reserves are not considered in the SST balance sheet. The European Solvency II framework uses a similar approach to define the insurance liabilities: equalization reserves are excluded from the best estimate of liabilities, see EIOPA (2015). Thus, the handling of the equalization reserves within Solvency II and IFRS 17 is the same.

According to Art. 41 para. 3 of ISO, the market value margin (MVM) corresponds to the capital costs of the risk bearing capital. MVM has to be kept as an insurance liability in the SST balance sheet, see FINMA (2017, rank no. 51) and MVM intends to fulfill the insurance liabilities, see FINMA (2017, rank no. 52). Similar to MVM, Solvency II requires a risk margin, based on the one-year view. In IFRS 17, a margin is requested for risk adjustment, based on a lifetime view, to be taken into account within the technical provisions, see England et al. (2019). However, this discounted capital cost does not have the same purpose as the statutory equalization reserves: MVM’s purpose is to cover capital costs and the equalization reserves’ purpose is to balance unfavorable development and the fluctuation of incurred claims. Similarly, MVM and the equalization reserves intend to provide an additional safety cushion to develop incurred claims.

Within the SST standard-model for non-life undertakings, target capital consists, among others, of the insurance risk comprising reserving risk, normal, large and NatCat claims according to FINMA (2018a). The definitions are as follows:

- Reserve risk: risk out of claims being greater than expected;
- Normal claims: risk out of claims below a threshold—high frequencies/low severities;
- Large claims: risk out of claims above a threshold—low frequencies/high severities;
- Nat Cat: risk out of claims occurred due to natural catastrophe event.

SST's insurance risk and the statutory equalization reserves have some features in common. However, SST's insurance risk is part of the target capital, and the equalization reserves are statutory balance sheet liabilities.

Similar to SST's insurance risk, Solvency II handles reserves risk, for which England et al. (2019) present analytic and simulation-based approaches. In the classical sense, the prediction error to quantify reserve risk consists, in general, of the uncertainty within the estimation of parameters and within the underlying claims generating process. This reserve risk heads in the same direction as the equalization reserves, balancing unfavorable development and fluctuation of incurred claims.

### 3. Preparation for the Model and Tax Considerations

This section starts with basic ideas to model equalization reserves. Thereafter, in Section 3.2, we use the introduced formalism to develop tax considerations linked to these equalization reserves. While FOPI and FINMA implicitly published the equalization reserves as part of the technical reserves, we restate these equalization reserves by using FOPI's and FINMA's published figures. In Section 3.3.1, we present the restatement for reinsurance undertakings, and in Section 3.3.2, for non-life undertakings.

#### 3.1. Basic Idea of the Model

We present a model to study the factors influencing the equalization reserves. We start looking at the modeling in Europe and thereafter, we develop our conceptual approach.

##### 3.1.1. Modelling in Europe

Akhurst et al. (1992) list two features to build up the equalization reserves: the upper limit of the equalization reserves and the transfer rule to build and to release reserves. They present four formulas to define the upper limit:

- Short-term fluctuation, cycles, trends and potential risk cumulation;
- Standard deviation of the claims;
- A coefficient, depending on the line of business, multiplied by the premium income;
- A combination out of premium income, variance and oversea risks.

Within Europe Sandström (2005) promotes further harmonization of the equalization reserves by linking the equalization reserves to the volatility of business. As transfer rules Akhurst et al. (1992) present:

- Recognition of the underwriting result;
- Deviation of the incurred claims ratio from the average level;
- Deviation of the incurred claims ratio from a fixed constant level;
- An asymmetric formula depending on the underwriting result;
- An open and half-open transfer (not only depending on the underwriting results);
- Other specific rules.

##### 3.1.2. Conceptual Approach

In its business plan, each Swiss insurance undertaking defines the upper limit and the corresponding transfer rule to the equalization reserves. Both are unknown to us since they are not disclosed. Nevertheless, we propose to approximate a transfer rule to the equalization reserves with the help of the profit and loss items. We analyze the transfer rule based on the underwriting result: premium income, other income, insurance costs (payments and change in technical reserves) and administration costs. The classical approach to calculate the technical result ( $TRe$ ) of an undertaking is, according to Farny (2011), the sum of the income, premium ( $Pre$ ) and other income ( $OIn$ ), minus the costs, insurance cost ( $ICo$ ) and operation cost ( $OCo$ ):

$$TRe = Pre + OIn - ICo - OCo. \quad (1)$$

Insurance cost ( $OIn$ ) consists of claims payments and changes (“ $\Delta$ ”) in total technical reserves. The equalization reserves ( $ER$ ) are part of the total technical reserves. Therefore, insurance cost is the sum of insurance cost without change in equalization reserves ( $ICo\_wo\Delta ER$ ) plus change in equalization reserves ( $\Delta ER$ ). Accordingly, we transform the insurance costs:

$$ICo = ICo\_wo\Delta ER + \Delta ER. \quad (2)$$

The actual equalization reserves ( $ER$ ) are the sum out of the change in equalization reserves ( $\Delta ER$ ) and the equalization reserves of the prior year ( $ER\_PY$ ):

$$ER = \Delta ER + ER\_PY. \quad (3)$$

Using (2) we transform (1) into:

$$\Delta ER = Pre + OIn - ICo\_wo\Delta ER - OCo - TRe, \quad (4)$$

which yields for (3):

$$ER = Pre + OIn - ICo\_wo\Delta ER - OCo - TRe + ER\_PY. \quad (5)$$

The size of the equalization reserves depends on the profit and loss items and on their previous year’s amount.

In addition to the above, we are interested in the potential influence of the accounting year, of the company age and of the legal form. In Section 4, we use FOPI’s and FINMA’s published net data to analyze the impact of these items.

Equation (5) could take outward reinsurance into account. The equalization reserves have to cover the whole gross business on the one hand. On the other hand, the net business reflects the remaining risk exposure to be covered. An undertaking does not transfer parts of the equalization reserves to the reinsurer and independently, the reinsurer estimates the equalization reserves on its own portfolio. As net data are available, we analyze the numbers net of reinsurance.

### 3.2. Tax Considerations

To quote the words from Benjamin Franklin, “Nothing is certain but death and taxes”. Therefore, we consider tax in connection with equalization reserves.

In Switzerland, the net profits of corporations are taxed and corporations pay tax on their own share capital, see Swiss Confederation (2018b). Thereby, a net loss at the end of the fiscal year can be fiscally offset by future profits for seven years, see Swiss Confederation (2018a). Swiss GAAP FER 40 provides accounting principles for insurance companies, see Foundation for Accounting and Reporting Recommendations (2018): equalization reserves<sup>6</sup> from the insurance techniques are considered as part of the liability “technical reserves”. The company has to define the valuation of the equalization reserves in alignment with Swiss GAAP FER 40 (Foundation for Accounting and Reporting Recommendations 2018); in the business plan of the insurance company, this assessment is fixed and supervised by FINMA.

Change in the equalization reserves impacts the net profit of the fiscal year and herewith the taxable amount. Therefore, we take a closer look at the taxation and the net profit of insurance companies. An insurance undertaking’s net profit consists of a non-technical and technical part. We are only interested in the technical one ( $TRe$ ). The taxation rate ( $Tax$ ) could change over the years. To keep the system simple, we assume a constant taxation rate over time ( $t$ ). Adding the time reference ( $t$ ) to the items we obtain from (1) in combination with (2):

$$\begin{aligned} Tax * TRe_t &= Tax * (Pre_t + OIn_t - ICo\_wo\Delta ER_t - \Delta ER_t - OCo_t) \\ &= Tax * (Pre_t + OIn_t - ICo\_wo\Delta ER_t - OCo_t) - Tax * \Delta ER_t. \end{aligned} \quad (6)$$

The change in equalization reserves reduces the taxable amount in the fiscal year  $t$ . Let us assume  $\hat{T}$  is the year in which a significant loss occurs of height  $L_{\hat{T}}$ , triggering the equalization reserves and assuming  $L_{\hat{T}} > ER_{\hat{T}-1}$ . Thus, the equalization reserves of the prior year are dissolved completely ( $ER_{\hat{T}-1}$ ). The taxable amount of (6) is transformed to:

$$TRe_{\hat{T}} = (Pre_{\hat{T}} + OIn_{\hat{T}} - ICo\_wo\Delta ER_{\hat{T}} - OCo_{\hat{T}}) + ER_{\hat{T}-1} \tag{7}$$

Due to the assumption  $L_{\hat{T}} > ER_{\hat{T}-1}$  we know that  $ICo\_wo\Delta ER_{\hat{T}} > ER_{\hat{T}-1}$ . Furthermore, we assume that the income does not cover the cost. This results in

$$TRe_{\hat{T}} < 0. \tag{8}$$

The release of the equalization reserves reduces the loss of the technical result in the year  $\hat{T}$ . This reduced loss may be offset by future profits over a period of seven years after the year  $\hat{T}$ . Changes in equalization reserves reduce future profits. However, to benefit from loss-offsetting with future profits, a profit has to be reported. This conflict of interest is born within the loss-offsetting with future profits over a period of seven years and the conflict is solved by the written business plan: the rules to establish the equalization reserves define how much the undertaking has to build up the equalization reserves. Table 1 illustrates this circumstance:

**Table 1.** Illustration of the impact of taxes from equalization reserves on the net profit.

Time	$\Delta ER_t$	$ER_t$	Tax Saving in $t$
1	$\Delta ER_1$	$ER_1 = \Delta ER_1$	$Tax * \Delta ER_1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\hat{T} - 1$	$\Delta ER_{\hat{T}-1}$	$ER_{\hat{T}-1} = \sum_1^{\hat{T}-1} \Delta ER_t$	$Tax * \Delta ER_{\hat{T}-1}$
$\hat{T}$	$-ER_{\hat{T}-1}$	0	0
$\hat{T} + 1$	$\Delta ER_{\hat{T}+1}$	$ER_{\hat{T}+1} = \Delta ER_{\hat{T}+1}$	$Tax * (\Delta ER_{\hat{T}+1} + \min[TRe_{\hat{T}+1};  TRe_{\hat{T}} ])$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\hat{T} + 7$	$\Delta ER_{\hat{T}+7}$	$ER_{\hat{T}+7} = \sum_{\hat{T}+1}^{\hat{T}+7} \Delta ER_t$	$Tax * (\Delta ER_{\hat{T}+7} + \min[TRe_{\hat{T}+7}; \max[0;  TRe_{\hat{T}}  - \sum_{\hat{T}+1}^{\hat{T}+6} TRe_t]])$

The Swiss tax regulation counteracts tax-optimization, using the equalization reserves: the scope of the business plan limits the establishment and the release of the equalization reserves. The loss-offsetting with future profit incentivizes an optimization-problem of profit shifting over time. Going concerns assumed, the equalization reserves belong to the insured, and some day, they will be taxed. From the pure actuarial point of view, tax-optimization is not taken into account to define rules to establish or release the equalization reserves, and by defining an upper bound for the equalization reserves. As equalization reserves may be easily misused for tax-optimization, governments looked at different possibilities to handle the valuation of equalization reserves. Akhurst et al. (1992) state that the taxation legislations “vary widely in different countries”. In some countries, the established equalization reserves should be released within a duration of, e.g., ten years.

### 3.3. Restatement of Equalization Reserves

By act of law Art. 25 of ISA, insurance undertakings report to FINMA the financial statements and the supervisory report, and FINMA publishes the annual financial statements<sup>7,8</sup>. FOPI (2007) disclosed the financial years 1997 to 2007 and since 2008 FINMA is in charge, see FINMA (2019a).

FOPI and FINMA have disclosed the figures in a different setup. In Appendices A and B, further details of the main changes and the source of the variables, which we use as explanatory ones, are provided. In Section 4, we describe the variables of



interest. Based on FOPI's and FINMA's items, one database was created in order to analyze the equalization reserves.

Akhurst et al. (1992) remark that disclosure of the equalization reserves is an open question. The equalization reserves reveal the prosperity of a company. To compare among insurance companies, Akhurst et al. (1992) use among others the equalization reserves. Since the reporting year 2019, the Swiss equalization reserves have been officially published. Nevertheless, the prior amounts are included in the published total of technical reserves and could be restated.

In the following, we describe the method used to estimate the equalization reserves for reinsurance and for non-life undertakings, taking into account FOPI's and FINMA's published figures for the years 1997 to 2018. We use FOPI's and FINMA's notation referring to the corresponding documents.

### 3.3.1. Restatement of Equalization Reserves for Reinsurance Undertakings

Since 2008, FINMA (2008b) has published the key metrics for insurers on its website: files reporting the balance sheets and profit and loss items, and files reporting statistical details<sup>9</sup>. FOPI's and FINMA's reported file "AR13A" states the main liability items for reinsurance undertakings: "net claim reserves" (*AR13A\_CR*) including the equalization reserves and "net other reserves" (*AR13A\_oR*) excluding the equalization reserves. In file "AR14K" the reinsurance undertakings report the single component of the technical reserves for the whole business: "net claim reserves" (*AR14K\_CR*) excluding the equalization reserves and "net other reserves" (*AR14K\_oR*) including the equalization reserves. Table 2 provides an overview of the handling of the equalization reserves for reinsurance undertakings.

**Table 2.** Equalization reserves (*ER*) within the reinsurers' files.

Reporting File	Net Claim Reserves	Net Other Reserves
AR13A	<i>ER</i> included	<i>ER</i> excluded
AR14K	<i>ER</i> excluded	<i>ER</i> included

Thus, the equalization reserves (*ER*) should be the difference of "AR13A" and "AR14K" by comparing "net claim reserves" and "net other reserves". These restated equalization reserves should be the same. However, some reinsurance undertakings stated the *ER*, e.g., in file "AR14K" within the "net claim reserves" and not within "net other reserves", resulting in correct the sign.

$$\begin{aligned} ER &= |AR13A\_CR - AR14K\_CR| \\ &= |AR13A\_oR - AR14K\_oR|. \end{aligned} \quad (9)$$

Due to data quality or maybe due to non-disclosure of the equalization reserves, for some undertakings, the calculated amount of the equalization reserves using "net claim reserves" on the one hand and "net other reserves" on the other hand is not the same. Therefore, the maximum of both values is taken for the analysis. *ER* calculated regarding "net claim reserves" and regarding "net other reserves" is not normally distributed. Their Spearman's rank correlation coefficient is 0.99, and both values are monotonically related.

### 3.3.2. Restatement of Equalization Reserves for Non-Life Undertakings

For non-life undertakings, the method, laid out above, cannot be applied: "AS14K" captures only the Swiss business, whereas "AS13A" captures the Swiss and the abroad written business. Nevertheless, FOPI has published the item "general reserves"<sup>10</sup> in file "AS14K", not being allocated to any line of business. We use the general reserves as an approximation of the equalization reserves.

In addition, in file "AS14K", FOPI published the total net technical reserves, including equalization reserves for the Swiss business (*AS14K\_TR*), in file "AS14A" the net premium reserves for the Swiss business (*AS14A\_PR*), in file "AS14B" the net claim reserves excluding

equalization reserves for the Swiss business (*AS14B\_CR*) and in file “AS14D” the net annuities reserves for the Swiss business (*AS14D\_AN*). Table 3 provides an overview of the handling of the equalization reserves for non-life undertakings.

**Table 3.** Equalization reserves (*ER*) within the non-life’s files.

Reporting File	Total Reserves	Premium Reserves	Claim Reserves	Annuities
AS14K	<i>ER</i> included	<i>ER</i> excluded	<i>ER</i> excluded	<i>ER</i> excluded
AS14A				
AS14B				
AS14D				<i>ER</i> excluded

Thus, the equalization reserves (*ER*) could be approximated by calculating the differences of the reported numbers. Nevertheless, we need to keep in mind that herewith “other reserves” would be a part of the equalization reserves. We have:

$$ER = AS14K_{TR} - AS14A_{PR} - AS14B_{CR} - AS14D_{AN}. \quad (10)$$

FINMA has changed file “AS14D”: the sum out of gross annuities, gross other technical reserves and gross equalization reserves are reported. Thus, Equation (10) cannot be evaluated anymore. In consequence, the equalization reserves cannot be restated with the above-listed files from the year 2008 onward. As a consequence, for non-life undertakings we study the calendar years 1997 to 2007.

We observe that several insurance undertakings report the general reserves as zero while reporting the equalization reserves non-zero. Therefore, we use the maximum value for our analysis. *ER* is calculated by using either the reported “general reserves” or restated using Equation (10). *ER* is not normally distributed. The Spearman’s rank correlation coefficient is 0.95.

#### 4. Regression Models for Equalization Reserves

In Switzerland, (re-)insurance undertakings accumulate equalization reserves, as stated in their business plan. Each undertaking defines its own rule to build and to release the equalization reserves. However, this information is not publicly available and unknown to us. The objective is to identify the items of the technical result, which could have influenced the amount of the equalization reserves. In this section, we present two regression models to study the factors impacting the equalization reserves.

In Section 4.1, we present FOPI’s and FINMA’s items as input data for our models. Thereafter, in Section 4.2, we explain the approach to set up the GAM and the GLM model. The models for reinsurance undertaking and the ones for non-life undertakings are provided in Sections 4.3 and 4.4.

##### 4.1. FOPI and FINMA Data as Explanatory Variables

In this section, we describe FOPI’s and FINMA’s items, used as explanatory variables.

Equation (5) is the formula to evaluate *ER*, using profit and loss items. FOPI’s and FINMA’s data offer further granularity in respect of operation and insurance costs. We change the notation to be used within our R modeling; we abbreviate all, except one, possible explanatory variables with three capital letters. In Table 4, an overview of all variables, a short description and the mark, whether we use the item in Equation (5) or whether the variable could be an explanatory one within our model of *ER*, are presented. The operation Cost (*OCo*) is split into “administration cost and commission” (*ACC*) and “other costs” (*OCC*). “Administration cost and commission” are the costs to administer and to conclude the insurance contracts, considering the related costs of the inward and outward reinsurance contracts and the costs out of profit participation. “Other costs” comprise all further underwriting costs, for example, the costs for preventive measures,

whereby detailed components of “other cost” are not available. The insurance cost ( $ICo$ ) is broken down into claims payments ( $NPA$ ), change in claims reserves and change in other technical reserves ( $NOR$ ). The correction of the change in equalization reserves is taken into account within the change of claims’ reserves ( $NCR\_woD\_ER$ ). The items are net of reinsurance. Using the new notation, Equation (5) is transformed to:

$$ER = NEP + OIN - NPA - NCR\_woD\_ER - NOR - ACC - OCC - TRE + ERP. \quad (11)$$

In addition to the profit and loss items, FOPI’s and FINMA’s data offer information about the legal form ( $LEF$ ), the company age ( $COA$ ) and the calendar year ( $YEA$ ), which we consider as a further additional explanatory variable in our model. As branch offices do not report profit and loss files within FOPI’s and FINMA’s reporting, the corresponding cost items are missing. Therefore, we only retain stock companies and mutuals. We analyze 799 data points for reinsurance undertakings, having the legal form “stock company”. For non-life undertakings, we look at 340 data points, having the legal form “stock companies”, and at 73 data points, having the legal form “mutuals”.

**Table 4.** Description of variables in alphabetical order.

Abbreviation	Description	Variable in Equation (5)	Explanatory Variable
$ACC$	Administration and commission cost		✓
$COA$	Company age		✓
$ER$	Equalization reserves, used as response variable in our R model	✓	
$ERP$	Equalization reserves of previous year = $ER\_PY$	✓	✓
$ER\_PY$	Equalization reserves of previous year	✓	
$ICo$	Insurance cost = $NCR\_woD\_ER + NOR + NPA$	✓	
$LEF$	Legal form		✓
$NCR\_woD\_ER$	Net change in claim reserves without change in $ER$		✓
$NEP$	Net earned premium = $Pre$	✓	✓
$NPA$	Net claims payments		✓
$NOR$	Net change in other technical reserves		✓
$OCC$	Other costs		✓
$OCo$	Operation cost = $ACC + OCC$	✓	
$OIN$	Other technical income = $OIn$	✓	✓
$Pre$	Premium = $NEP$	✓	✓
$TRE$	Technical result = $TRe$	✓	✓
$YEA$	Calendar year		✓

For the modeling, the software environment R is used. Profit and loss items are modeled as continuous variables. In our analysis, undertakings can have either the legal form ( $LEF$ ) mutual or stock company (plc). As all reinsurance undertakings have the legal form stock company, we use the explanatory variable, legal form, to capture (reinsurance) captives, 470 data points being available, and professional reinsurer<sup>11</sup>, 329 data points being available. Thus, the categorical variable legal form contains the items “mutual”, “stock company”, “captive” and “professional reinsurer”. Calendar years ( $YEA$ ) are 1997 to 2018. The company age ( $COA$ ) is calculated as the difference out of the calendar and foundation year, being in the range of 0 to 180. To reflect the categorical character of the variables legal form, calendar year and company age, we enter these variables as factors in our model, see

Spector (2011). We look at 12 available explanatory variables: 9 continuous and 3 categorical. In Sections 4.3.1 and 4.4.1, we comment and provide descriptive statistics for the variables. In Appendix B, the source of the variables is presented.

FOPI and FINMA have published the figures mentioned in Table 4. We used the data as reported and checked these figures. Statistical outliers are double checked. As the peculiarity could be explained, we keep these records. Some companies do not follow the instructions, to implicitly disclose the equalization reserves. Therefore, these companies report the equalization reserves as zero. Therefore, we study only these data points, having positive equalization reserves for a given calendar year, resulting in limiting the insurance market to a subset. Small, medium and large undertakings are reported and we keep all these records. We intend to have a picture of all available records and not of a smaller subset.

#### 4.2. Model for the Response Variable, Equalization Reserves

In this section, we explain the approach to setting up the GAM and the GLM model. Our objective is to find a GAM and a GLM model once for the reinsurance and once for non-life undertakings. We acknowledge that reinsurance undertakings could be seen as a special line of business and that in consequence we could consider a joint modeling, see Jeong and Dey (2020); Merz et al. (2012); Shi and Frees (2011). To reflect the two different types of business, we separately consider reinsurance and non-life undertakings. Nevertheless, we use the same approach for both. We apply a popular approach, described in the literature, for example, by Staudt and Wagner (2021). In a first step, we define the GAM model by selecting the explanatory variables, having the highest impact on the equalization reserves. In a second step, we transform the selected continuous variables to discrete ones by defining optimal classifications with the help of *evtree*. In the third step, we apply the GLM model by using the discrete and the transformed variables.

This section is structured, as follows: first, we determine the distribution of the equalization reserves, presented in Section 4.2.1. In Section 4.2.2, we define the GAM model by selecting the explanatory variables. Thereafter in Section 4.2.3, we present how we find a GLM model by applying *evtree* to define classifications for the continuous explanatory variable.

##### 4.2.1. Identification of the Distribution

In this section, we determine the distribution of the equalization reserves. De Vylder and Goovaerts (1999) present a theoretical evaluation of the equalization reserves. However, the literature does not discuss the class of distribution for the equalization reserves, and a recommendation is missing. As mentioned, we apply a GAM model, and one assumption of the GAM is that the response variable has an exponential family, see Hastie and Tibshirani (1990). We identify the type of distribution of the equalization reserves by calculating the AIC value of the exponential, Weibull, Gamma and log-normal distribution, see Akaike's Information Criterion in Parzen et al. (1998)<sup>12</sup>. AIC's formula values the performance of the fitted model by taking into account the log-likelihood of this model and the number of parameters of the model; the smallest AIC defines the best fit, see R Core Team (2021).

##### 4.2.2. Approach to Set Up the GAM Model

Our objective is to set up a GAM model to analyze the relationship between the equalization reserves and its explanatory variables, being discrete and continuous.

##### Selection of Explanatory Variables to Define GAM

Hastie and Tibshirani (1990) explain the generalized linear models (GLM): assumed is that the predictor effects are linear in the parameters; however, the distribution of the responses and the link between the predictors and this distribution could be general. As we assume that the response variable, equalization reserves, has a non-linear dependency from the prior mentioned discrete and continuous explanatory variables, a GLM model

is not suitable, and instead, we use its extension, the generalized additive model (GAM). Hastie and Tibshirani (1990) discuss GAM models and how to replace the linear predictor with an additive one. As Fridley (2010) illustrates, a GAM model lets “the data to ‘speak for themselves’” by using a smoothing function<sup>13</sup>.

Our goal is to link the response variable, equalization reserves, with the explanatory variables, presented in Table 4. We utilize a common approach, described for example by Staudt and Wagner (2021) to obtain the optimal GAM model: the best performance of variously defined models determines the best GAM model. The start is a GAM model having only one explanatory variable: each of the 12 available possible explanatory variables defines its own GAM model and the corresponding AIC value captures its performance. The explanatory variable, having the lowest AIC value, is kept as the best suitable explanatory variable for the next iteration. In the next iteration, each of the remaining 11 possible explanatory variables defines its own GAM model, having two explanatory variables; again, the corresponding AIC value captures the performance. The iteration process is continued until the AIC value would not change and would increase. This approach selects the best explanatory variables. Illustrative for the algorithm in iteration  $i$ , having the smoothing function  $s(\cdot)$ , the best explanatory  $x_i$  out of the set  $X$  of all explanatory variables is defined by

Illustrative<sup>14</sup> for iteration  $i$ :

$$\min_{\text{AIC}(\text{GAM}_{k,k \in \{1, \dots, i\}})} \operatorname{argmin}_{x_i \in X \setminus \{x_1, \dots, x_{i-1}\}} \text{AIC}(\text{GAM}(ER \sim \sum_{j=1}^i s(x_j))) \quad (12)$$

The selection is needed to construct the  $\text{GAM}_{\text{Select}}$ . Note that the order of the explanatory variables is not relevant for the modeling.

#### 4.2.3. Approach to Set Up GLM

Our objective is to set up a GLM model to analyze the relationship between the equalization reserves and its explanatory variables, being discrete and transformed to discrete ones. To transform the continuous variables to discrete ones, we classify these variables by using `evtree`, as presented in the paragraph hereafter.

##### 4.2.3.1. `evtree` to Define Classifications for Explanatory Variable

Having determined the  $\text{GAM}_{\text{Select}}$ , we look for optimal categorization of the significant explanatory variables, using evolutionary trees according to Staudt and Wagner (2021). Grubinger et al. (2014) explain the optimal classification and regression trees in R: a recursive partitioning by choosing splits that maximize the homogeneity at the next step. The response variable of the `evtree` is the fitting effect ( $\text{Fit}_{x_i}$ ) of  $\text{GAM}_{\text{Select}}$  and the predictor variable of the `evtree` is one out of our selection. The split consists out of splitting variables  $v_r^l \in \{\text{observation of } x_l\} =: X_l$  and splitting rules  $s_r^l$  for the internal nodes, while  $r \in \{1, \dots, M_l - 1\}$ ,  $M_l$  denoting the partition of the input space of the predictor variables. Among others, the minimum number of observations ( $b$ ) within an interval  $[v_r^l; v_{r+1}^l]$  can be controlled and the complexity of the splitting rule ( $\alpha$ )<sup>15</sup>. Our goal is to find the best splitting variables  $v_r^l: X_l = \cup_{r \in M_l - 1} [v_r^l; v_{r+1}^l]$  conditional to minimize the AIC and conditional  $b \in B$  and  $\alpha \in A$ ,  $B$  being a set of possible minimum number of nodes and  $A$  a set of complexity parameters. Again, we let the selected data speak for themselves. For each variable of the selection  $x_l$  we determine the optimal buckets size of the tree and regulate the complexity of the tree, whereby optimal is in the sense of minimizing the AIC value of the  $\text{GAM}_{\text{Select}}$  in which we replace  $x_l$  by the new classification. The minimal AIC-value determines the best buckets' size ( $\hat{b}$ ) and the complexity parameter ( $\hat{\alpha}$ ), receiving the best categorization of the analyzed explanatory variables of the selection ( $x_l \in \text{Selection}$ ). Illustration of the R algorithm is as follows:

Illustrative  $\forall x_l \in \text{Selection}$  :

$$\underset{\alpha \in A, b \in B}{\operatorname{argmin}} \operatorname{AIC}(\operatorname{GAM}(ER \sim \sum_{x_j \in \text{Selection} \setminus \{x_l\}} s(x_j) + \bigcup_{r \in M_l - 1} [v_r^l; v_{r+1}^l] \mid \operatorname{evtree}(b, \alpha))) \quad (13)$$

#### 4.2.3.2. Defining GLM

The result of the previous paragraph is the transformation of the continuous variables to discrete ones. Thus, we can apply a GLM model on the previously selected explanatory variables, according to Hastie and Tibshirani (1990).

#### 4.2.3.3. General Remark

We are aware that we base our model on 799 observations for reinsurance undertakings and on 413 observations for non-life undertakings, not reflecting the whole Swiss insurance market. Furthermore, we are aware that the 12 explanatory variables are correlated, which we looked at. Due to the fact that our model can only provide a first indication of the real circumstances, we refrain from modeling the correlations between the explanatory variables.

### 4.3. Model of the Equalization Reserves for Reinsurance Undertakings

To explore the equalization reserves for reinsurance undertakings, we base the analysis on FOPI and on FINMA data. In Section 4.3.1, we look at the data, and thereafter, in Section 4.3.2, we present the model result, using the approach described in Section 4.2.

#### 4.3.1. Data for Reinsurance Undertakings

FOPI and FINMA have published equalization reserves of reinsurance undertakings for the calendar years 1997 to 2020. As FINMA changed the reporting system in the year 2020 and presents the data for the reporting years 2019 onward in a different setup, we focus on the calendar years 1997 to 2018. For our analysis, we filter the data points, having positive equalization reserves, and 799 data points are available for our study. Table 5 captures the summary of the equalization reserves and of the explanatory variables, covering the years 1997 to 2018. The currency is CHF. Minimum (Min.), first quantile (1st Qu.), median, mean, third quantile (3rd Qu.) and maximum (Max.) are presented.

**Table 5.** Variable summary for reinsurance undertakings, observed in the years 1997 and 2018.

Variable	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
ER	1000	533,000	4,615,772	58,007,935	16,883,679	3,531,957,000
ACC	−234,455,888	215,796	736,604	165,830,866	14,841,470	7,165,666,533
COA	0	26	56	62	92	135
ERP	0	25,000	2,128,000	52,440,000	14,570,000	3,532,000,000
NCR_woD_ER	−10,670,000,000	−220,400	1,005,000	72,290,000	10,300,000	8,241,000,000
NEP	−1,260,000,000	2,994,000	7,439,000	572,500,000	81,780,000	20,330,000,000
NOR	−1,338,000,000	0	0	3,208,000	1,769,000	429,200,000
NPA	−5,580,000,000	147,900	3,024,000	329,200,000	39,540,000	20,090,000,000
OCC	−63,649,261	0	0	12,202,336	58,062	1,487,351,655
OIN	−120,478,524	0	43,051	51,501,126	1,553,194	2,502,998,729
TRE	−2,206,000,000	−175,700	462,700	18,820,000	4,974,000	2,088,000,000

“Min.” stands for minimum, “1st Qu.” for first quantile, “3rd Qu.” for third quantile and “Max.” for maximum.

#### Equalization Reserves

The smallest equalization reserves amount to CHF 1000 and the highest to CHF 3.53 billion (Swiss Re 2001)<sup>16</sup>. In the Appendix C.1, Figure A1 shows the relative frequency of the equalization reserves and of the one of the previous year. In total, 82% report the equalization reserves lower than CHF 24 million, and 82% report the equalization reserves of the previous year lower than CHF 20 million.

### Calendar Years and Legal Form

In the year 2007, most observations are noticed; hence, we re-level the variable calendar year to the year 2007. The number of observations in a calendar year increased from 21 in the year 1997 to 62 in the year 2007 and dropped to 34 in the year 2008. In the year 1997, 3% and in the year 2018, 4% of the observations occurred. FOPI published the years 1997 to 2007 and FINMA uses another reporting system for the calendar year 2008 to 2018. FINMA (2009) supervised 28 reinsurance undertakings (2007: 25) and 42 captives (2007: 46) in the calendar year 2008; as a consequence, mergers and acquisitions would not explain the drop from the year 2007 to 2008. We think our counting method, companies having a positive equalization reserves, causes the drop: some reinsurance undertakings would have used the occasion of the new reporting system to change the reporting of the equalization reserves.

All analyzed reinsurance undertakings and reinsurance captives are stock companies; therefore, the variable “legal form” would not be suitable for an explanatory one. Instead, we analyze the kind of reinsurance, captive and professional reinsurer. In the year 1997, 38% were captives, increasing to 78% in the year 2011 and dropping to 56% in the year 2018. As more observations for captives are available, we re-level the explanatory variable to “captive”.

### Net Earned Premium

The smallest net earned premium amount was CHF –1.3 billion and the highest was CHF 20.3 billion (Swiss Re 2008). Europäische Rückversicherung reported in the year 2005 indeed a higher ceded premium (written: CHF 6.0 billion) than the gross one (written: CHF 5.1 billion), resulting in a negative net earned premium (CHF –1.3 billion). In Appendix C.1, Figure A2 displays the relative frequency of the net earned premium and the one of the calendar year. In total, 94% have a net earned premium lower than CHF 2.0 billion and 62% lower than CHF 20 million.

### Insurance Costs

The insurance net claims payments are within the range from CHF –5.6 billion to CHF 20.1 billion (Swiss Re 2009). Europäische Rückversicherung reported in the year 2009 CHF 5.6 billion as payment, having an opposite sign to all other undertakings. In total, 87% have net claims payments lower than CHF 200 million and 70% lower than CHF 20 million. The highest release in claims reserves without change in equalization reserves was CHF 10.7 billion (Swiss Re 2009) and the highest accumulation CHF 8.2 billion (Swiss Re 2016). Furthermore, Swiss Re reported in the year 2017 the highest release in other reserves of CHF 1.3 billion and in the year 2014 the highest accumulation of CHF 429.2 million. In total, 80% report a change in claims reserves without change in equalization reserves in the range from CHF –10.7 billion to CHF 20 million and 62% a change in other reserves in the range from CHF –1.3 billion to CHF 320 million. The relative frequency of the insurance cost of reinsurance undertakings are shown in Appendix C.1 in Figure A3.

### Operational Costs and Other Technical Income

The operational costs consist of administration cost and commission, being in the range from CHF –0.2 to CHF 7.2 billion (Swiss Re 2006), and other cost, having values from CHF –63.6 million (Swiss Re 2013) to CHF 1.5 billion (Swiss Re 2018). Europäische Rückversicherung reported in the year 2008 CHF –0.2 billion as operational costs, having an opposite sign to all other undertakings. This could occur due to ceded costs from outward reinsurance. Indeed, in the year 2013, Swiss Re reported other costs with a negative sign. Other technical income comprises technical interest income, attributed to the technical part, and other technical insurance income. This item is in the range from CHF –120 million to CHF 2.5 billion (Swiss Re 2007). Europäische Rückversicherung reported in the years 2013 CHF 120 million as other technical income, having an opposite sign to all other undertakings. Due to ceded cost out of outward reinsurance this could happen. In

Appendix C.1, Figure A4 presents the relative frequency of the operational costs and of the other technical income. In total, 70% of the administration and commission costs are below CHF 8 million, 73% of the other costs are below CHF 13,000 and 69% of the other insurance income is below CHF 1 million.

#### Technical Result and Company Age

The smallest technical result amounts to a loss of CHF 2.2 billion (Swiss Re 2001) and the highest to CHF 2.1 billion (Swiss Re 2007). A total of 78% report a technical result lower than CHF 6.4 million. In Appendix C.1, Figure A5 shows the relative frequency of the technical result and of the company age. The company age is in the range from founded to 135 years. In our analysis, 88% of the companies have a company age below 30.

#### 4.3.2. Model for the Reinsurance Undertakings

Our objective is to find a GAM and a GLM model for the reinsurance undertakings, as presented in Section 4.2. First, we determine the distribution of the equalization reserves for the reinsurance undertakings. In the second step, we define the GAM model by selecting the explanatory variables. Thereafter, we study the result and apply *evtree* to define classifications for explanatory variables. Finally, we apply the GLM model by using the discrete and the transformed variables and present the main results.

#### Equalization Reserves' Distribution

As mentioned in Section 4.2.1, the literature does not propose a class of distribution for the equalization reserves. Therefore, we let our equalization reserves' observations speak for themselves. We fit some distribution of the exponential family (exponential, Weibull, Gamma and log-normal distribution) on our data and compare the theoretical and the sample quantiles, presented in Appendix D Figure A11. In Figure A11, the q-q plots of all fitted distributions are far off the diagonal. The log-normal distribution is the only suitable one; for the other three distributions, insufficient observation points exist to properly fit the tail. In addition, we calculate the AIC values for these distributions, presented in Table 6. For the reinsurance undertakings, the log-normal distribution best approximates the response variable, equalization reserves.

**Table 6.** AIC values for possible distributions of equalization reserves (reinsurer).

	Exponential	Weibull	Gamma	Log-Normal
AIC	30,166	27,702	27,944	27,685
Number of Parameters	1	2	2	2
Log-likelihood	−15,082	−13,849	−13,970	−13,841

#### Selection of the Explanatory Variables to Define GAM

Our goal is to figure out the relationship between the response variable, equalization reserves, with the explanatory variables, as presented in the Section 4.2.2.

In the previous paragraph, we worked out that equalization reserves could be fitted by a log-normal distribution; thus,  $\log ER$  is normal distributed. According to Wood (2010) we implement in R a GAM model with the family Gaussian and we use the link function identity, ignoring the correlation of the explanatory variables, as mentioned in Section 4.2.3.3. We apply the algorithm, outlined in Equation (12), and obtain the GAM model, as follows:

$$\mathbb{E}(\log(ER)) = f_1(ERP) + \beta_2 \cdot YEA + \beta_3 \cdot LEF + f_4(TRE) + f_5(ACC) + f_6(OCC) + \beta_0. \quad (14)$$

Equalization reserves of the previous year head the selection of the explanatory variables for reinsurance undertakings. The categorical variable calendar year has the second highest impact on the equalization reserves, according to our approach. Thereafter, the legal form (capitves or professional reinsurers) shapes the equalization reserves. The variable



technical result has position four, administration and commission cost position five, and the variable other costs terminate the selection of the explanatory variables, shaping the equalization reserves of the reinsurance undertakings. An additional explanatory variable would not improve the fitting. In Appendix E, Table A3 shows the AIC values for the iterations.

In the first column of Table 7, we list the intercept, the coefficients of the discrete variables, the estimated degree of freedom (edf) of the continuous variables and the significance level of the GAM model. We use the star-notation to mark the statistical significance level of the variables in Table 7: "." for a  $p$ -value below 0.1, "\*" for a  $p$ -value below 0.05, "\*\*\*" for a  $p$ -value below 0.01 and "\*\*\*\*" for a  $p$ -value below 0.001. The smoothing functions of the continuous variables are statistically significant, the intercept and some coefficients of the discrete variables. As Fridley (2010) explains with respect to the GAM model, the predictor variable is separated into sections and fitted by polynomial functions in each section separately. These turning points are measured as edf: the higher the edf, the more complex the smoothing. The continuous variables equalization reserves of the previous year, technical result and the administration and commission costs have an edf-value of 9.0, 7.44 and 8.45, which have the highest complexity to be smoothed. The variable other costs has an edf value of 5.03, which is the lowest complexity to be smoothed. The categorial variable year has 2007 as the base line and the categorial variable legal form has captive as the baseline.

#### GAM Result and `evtree` to Define Classifications for Explanatory Variable

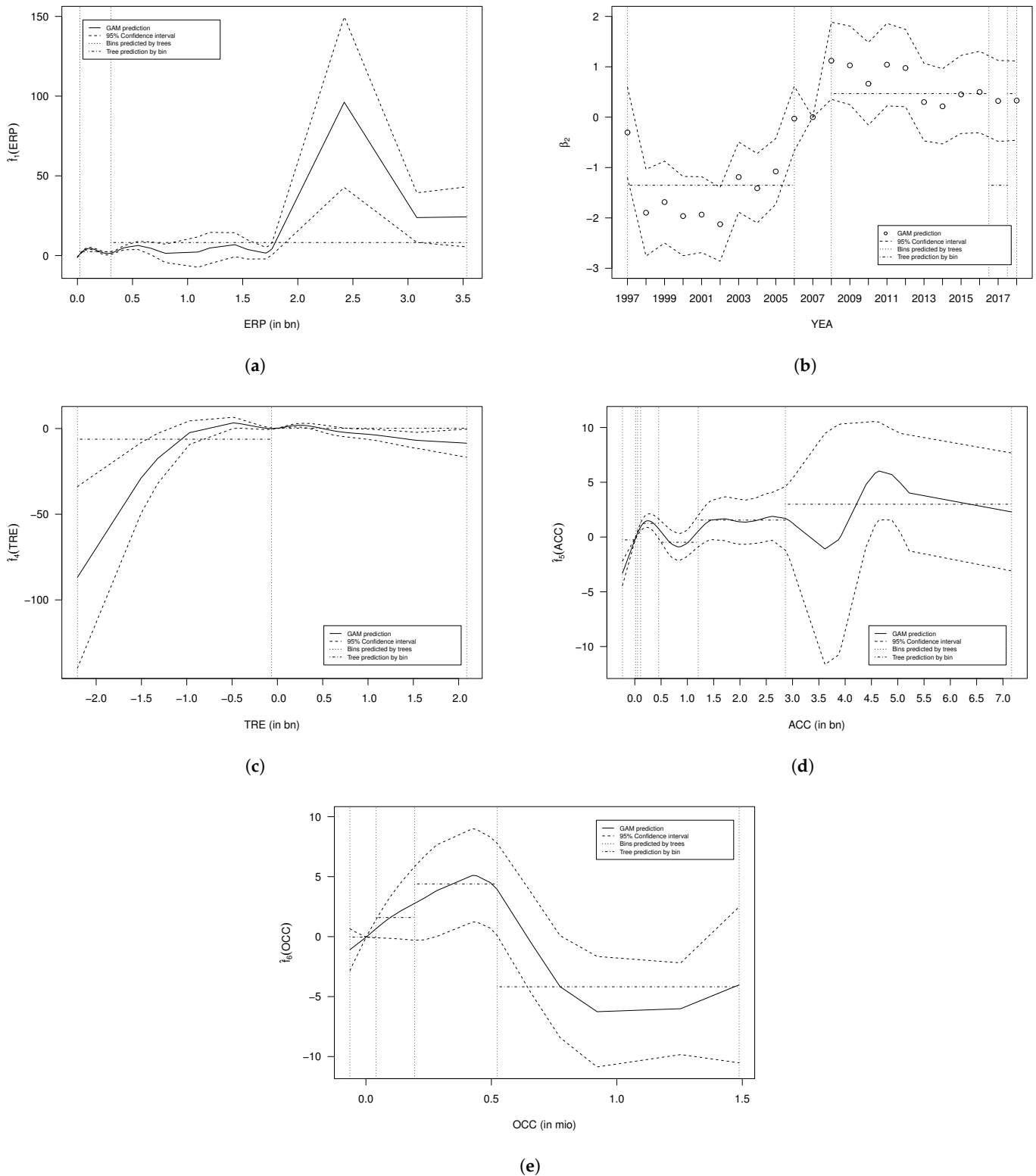
Visualization of the fitted smoothing curves reveals the highest insight into the relationship of the explanatory variables on the equalization reserves. As outlined above in Equation (13), we define the optimal classes of the explanatory variable and in Table A5 the used parameters of the `evtree` are listed. We add this information in Figure 1: for each explanatory variable its GAM prediction on the equalization reserves is shown (solid line), the 95% confidence interval is marked (dashed lines), the boundaries of the bins are displayed by the vertical lines and the predicted mean effect for each class is indicated by the dashed-pointed horizontal lines.

#### Equalization Reserves of the Previous Year

Equalization reserves of the previous year are shown in Figure 1a and we find three classifications. Around 680 companies report the equalization reserves of the previous year lower than CHF 22 million, which has a negative impact on the equalization reserves. The next class is up to CHF 305 million, consisting of around 90 observations, influencing the equalization reserves to increase. The higher the equalization reserves of the previous year, the more equalization reserves are built up. As a take away, we note that the smaller the equalization reserves of the previous year, the more these equalization reserves would be released in the next year. To transform this continuous variable into the discrete one, we define three classes. In the class CHF 0 to 22 million, most observations occur, on which we level this variable.

#### Calendar Year

In Figure 1b, the impact of the discrete variable calendar year is displayed and two classes are found. The calendar year is re-leveled to the year 2007. The years could be split in the range 1997 to 2006, having a negative impact on the equalization reserves, and 2008 to 2018, having a positive impact on the equalization reserves. By adding the categorial variable legal form to the GAM model, the year 2017 moved from the class 2008–2018 to 1997–2006 due to a reduction in the equalization reserves of a reinsurer of around CHF 1 billion. In the years 1997 to 2007, FOPI was in charge of the supervision and since 2008, it was FINMA.



**Figure 1.** Effects on the equalization reserves for reinsurance undertakings. (a) Effect of the equalization reserves of the previous year. (b) Effect of the calendar year. (c) Effect of the technical result. (d) Effect of the accumulation and commission costs. (e) Effect of the other costs.

Table 7. Output of the GAM and GLM for reinsurance undertakings.

GAM			GLM		
Explanatory Variables		Star	Explanatory Variables		Star
(Intercept)	14.77	***	(Intercept)	14.10	***
<b>Categorical:</b>	Coefficient		<b>Categorical:</b>	Coefficient	
	YEA (baseline: 2007)			YEA (baseline: 2007)	
1997	−0.30		1997	−0.99	*
1998	−1.90	***	1998	−2.28	***
1999	−1.69	***	1999	−2.08	***
2000	−1.96	***	2000	−2.32	***
2001	−1.93	***	2001	−2.28	***
2002	−2.13	***	2002	−2.47	***
2003	−1.19	***	2003	−1.55	***
2004	−1.41	***	2004	−1.78	***
2005	−1.08	***	2005	−1.42	***
2006	−0.03		2006	−0.34	
2008	1.12	**	2008	1.15	**
2009	1.03	**	2009	1.18	**
2010	0.66		2010	0.95	*
2011	1.04	*	2011	1.19	**
2012	0.98	*	2012	0.93	*
2013	0.30		2013	0.26	
2014	0.22		2014	0.15	
2015	0.45		2015	0.49	
2016	0.50		2016	0.66	
2017	0.32		2017	0.55	
2018	0.33		2018	0.53	
	LEF (baseline: Captive)			LEF (baseline: Captive)	
Prof. RE	1.03	***	Prof. RE	0.82	***
<b>Continuous:</b>	edf		<b>Categorical:</b>	Coefficient	
s(ERP)	9.00	***		ERP (baseline: 0–22 m.)	
			22–305 m.	2.41	***
			305–3550 m.	4.51	***
s(TRE)	7.44	***		TRE (baseline: −60–2100 m.)	
			−2200–−60 m.	0.53	
s(ACC)	8.45	***		ACC (baseline: −235–15 m.)	
			15–55 m.	1.18	***
			55–115 m.	2.05	***
			115–470 m.	2.09	***
			470–1300 m.	1.75	**
			1300–2900 m.	2.58	***
			2900–7200 m.	2.50	**
s(OCC)	5.03	**		OCC (baseline: −64–40 m.)	
			40–200 m.	0.06	
			200–550 m.	−1.18	
			550–1500 m.	−2.41	*
AIC	3218		AIC	3337	
N	799		N	799	

Star represents significance levels of  $p$ -values: 0 '\*\*\*\*' 0.001 '\*\*\*' 0.01 '\*\*' 0.05 '.' 0.1 '' 1. edf stands for estimated degree of freedom and Prof. RE for professional reinsurance undertakings.

### Legal Form

Reinsurance undertakings are either captives or professional reinsurance undertakings. Baseline is the category captive. As presented in Table 7, the coefficient for the professional reinsurer is 1.03: compared to captives, professional reinsurers tend to build more equalization reserves.

### Technical Result

The effect of the technical result on the equalization reserves is presented in Figure 1c, with two classes. The first classification consists of 20 nodes, and the loss of the undertaking is higher than CHF 60 million. The effect on the equalization reserves is negative. The second class comprises 780 observations, and the impact on the equalization reserves is slightly positive. As a take away, we note that for reinsurance undertakings, even positive technical results have only a slightly positive impact on equalization reserves. To transform this continuous variable to the discrete one, we define two classes. As the most observations occur in the class CHF −60 to 2100 million, we level this variable on this class.

### Administration and Commission Costs

In Figure 1d, the effect of the administration and commission costs on the equalization reserves is presented, with seven classes. The first class consists of 600 companies, with administration and commissions costs up to CHF 15 million and a slightly negative impact on the equalization reserves. In the next class, around 70 companies are collected, with costs up to CHF 55 million and a slightly positive impact on the equalization reserves. The class, with costs between CHF 470 million and CHF 1.3 billion, has a negative impact on the equalization reserves, whereby this class consists of 15 observations. The higher the administration and commission costs, the higher the accumulation of the equalization reserves. To transform this continuous variable to the discrete one, we define seven classes. Most observations occur in the class CHF −235 to 15 million on which we level this variable.

### Other Costs

Figure 1e presents the effect of other costs on the equalization reserve, with four classes. The first class contains 775 observations with other costs up to CHF 40 million. This class has a slightly negative effect on the equalization reserves. In the second class, 12 companies are captured, having other costs up to CHF 200 million; the impact on equalization reserves is positive. Seven observations are in the third class, having other costs up to CHF 550 and having a higher positive effect on equalization reserve than for the first class. The class number four has five observations, having other costs up to CHF 1.5 billion and a negative effect on equalization reserves. To transform this continuous variable to the discrete one, we define four classes. In the class CHF −64 to 40 million, most observations occur, on which we level this variable.

### GLM Model

We apply the GLM model on the initial discrete variables year and legal form and on the initial continuous variables, as transformed and described in the previous paragraph. We mark the transformed variable with “ $c$ ”, and Equation (14) turns to the GLM model:

$$\mathbb{E}(\log(ER)) = \beta_1 \cdot ERP_c + \beta_2 \cdot YEA + \beta_3 \cdot LEF + \beta_4 \cdot TRE_c + \beta_5 \cdot ACC_c + \beta_6 \cdot OCC_c + \beta_0. \quad (15)$$

In the second column of Table 7, we list the intercept, the coefficients of the discrete and of the transformed variables and the significance level of the GLM model. In the GAM model and in the GLM one, the intercept has a statistically significant level “\*\*\*\*” and amounts to 14.77 in the GAM and to 14.10 in the GLM one. The significance level of the initial discrete variables either remains the same or changes to the next significance level. All except one coefficient of the initial continuous variables have a significance level of “\*\*\*\*” or “\*\*\*” in the GLM. However, the coefficient of the technical result is not statistically significant in the GLM model. Furthermore, the GAM model has an AIC value of 3218 and the GLM one an AIC value of 3337, making the GAM a better one, see Table 7.

### Result

For reinsurance undertakings, the GAM model performs better than the GLM one. Based on the GAM model, we find the use of equalization reserves: accumulation and release of the equalization reserves are observed. From the GAM and from the GLM model,

we note that, compared to captives, professional reinsurers tend to build more equalization reserves. Based on the GAM model, we show that reinsurance undertakings, having small equalization reserves of the previous year, would release the equalization reserves. In addition, from the GAM model, we find that for reinsurance undertakings, even positive technical results have only a small positive impact on equalization reserves.

#### 4.4. Model of the Equalization Reserves for Non-Life Undertakings

To explore the equalization reserves for non-life undertakings, we base the analysis on FOPI data. In Section 4.4.1, we look at the data, and thereafter, in Section 4.4.2, we present the model result using the approach described in Section 4.2.

##### 4.4.1. Data for Non-Life Undertakings

FOPI published the equalization reserves for non-life undertakings for the calendar years 1997 to 2007. For our analysis we filter the data points, having positive equalization reserves, and 413 data points are available for our study. Table 8 captures the summary of the equalization reserves and of the explanatory variables, covering the years 1997 to 2007. The currency is CHF. Minimum (Min.), first quantile (1st Qu.), median, mean, third quantile (3rd Qu.) and maximum (Max.) are presented:

**Table 8.** Variable summary of non-life undertakings, observed in the years 1997 and 2007.

Variable	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
ER	1000	3,006,000	13,370,000	50,401,719	64,925,000	375,394,000
ACC	−171,143	4,123,740	12,809,683	209,586,622	159,078,572	6,319,819,394
COA	0	33	74	77	113	181
ERP	1000	323,000	9,499,000	44,157,007	43,442,000	375,394,000
NCR_woD_ER	−314,512,646	127,408	3,165,646	93,267,811	25,688,850	3,984,956,385
NEP	−12,140	13,710,000	45,890,000	749,700,000	519,300,000	22,760,000,000
NOR	−40,069,000	0	0	4,487,718	1,395,000	379,109,740
NPA	0	7,280,000	28,800,000	468,700,000	328,400,000	12,160,000,000
OCC	−27,034	0	229	4,603,862	742,376	327,609,822
OIN	−18,428,688	52,965	2,682,174	74,735,601	22,955,164	2,456,748,842
TRE	−1,192,000,000	−394,700	875,600	40,420,000	7,049,000	2,420,000,000

“Min.” stands for minimum, “1st Qu.” for first quantile, “3rd Qu.” for third quantile and “Max.” for maximum.

#### Equalization Reserves

The smallest equalization reserves amount to CHF 1000 and the highest to CHF 375.4 million (Mobilier 2004). In Appendix C.2, Figure A6 shows the relative frequency of the equalization reserves and of the one of the previous year. In total, 70% report the equalization reserves lower than CHF 40 million, and 71% report the equalization reserves of the previous year lower than CHF 30 million. For the year 1997, information regarding the equalization reserves of the previous year could not be calculated and the analysis will lack this item for the year 1997.

#### Calendar Years and Legal Form

In the year 1998, most observations are made; hence, we re-level the variable calendar year to the year 1998. The number of non-life entities were reduced, on the one hand, due to mergers and acquisition and, on the other hand, due to our selection criteria, analyzing records with positive equalization reserves. In the year 1997, 10% and, in the year 2007, 8% of the observations occurred. The share of mutual companies increased from 15% in the year 1997 to 25% in year 2007.

#### Net Earned Premium

The smallest net earned premium amounted to CHF −12,140 and the highest to CHF 22.8 billion (Zurich 2008). Unifun reported in the years 2004 and 2005 a higher ceded premium than the gross one, resulting in a negative net earned premium. In Appendix C.2,

Figure A7 displays the relative frequency of the net earned premium and the one of the calendar year. In total, 70% have a net earned premium lower than CHF 353.9 million and 47% lower than CHF 40 million.

#### Insurance Costs

The insurance net claims payments are within the range from CHF 0 (Polygon) to CHF 12.2 billion (Zurich 2007). In the year 2003, Baloise reported the highest release in claims reserves without change in equalization reserves of CHF 314.5 million, and in the year 2003, Zurich reported the highest accumulation of CHF 3.98 billion. Furthermore, the highest release in other reserves was CHF 40.1 million (Mobiliar 2005) and the highest accumulation CHF 379.1 million (Zurich 2006). The relative frequency of the insurance cost of non-life undertakings is shown Appendix C.2 in Figure A8. In total, 70% have net claims payments lower than CHF 823.3 million and 43% lower than CHF 20 million. Furthermore, 60% report a change in claims reserves without change in equalization reserves in the range from CHF −15.1 million to CHF 11.7 million.

#### Operational Costs and Other Technical Income

The operational costs consist of administration cost and commission, being in the range from CHF −0.2 million (Appenzeller 1997) to CHF 6.3 billion (Zurich 2007), and other costs, having values from CHF −27,034 (Orion 2006) to CHF 327.6 million (Zurich 2004). Appenzeller received more compensation for the cost from the ceded reinsurance contract than indeed paid, resulting in a negative administration and commission cost. Indeed, Orion reported other costs with a negative sign. Other technical income comprises technical interest income, attributed to non-life insurance, and other technical insurance income; this item is in the range from CHF −18.4 million (Vaudoise 2002) to CHF 2.5 billion (Zurich 1998). Vaudoise reported in the year 2002 a negative technical interest income of CHF 20.2 million compensated by a positive other technical insurance income of CHF 1.8 million. In Appendix C.2, Figure A9 presents the relative frequency of the operational costs and of the other technical income. In total, 72% of the administration cost and commission are below CHF 99.8 million, 73% of the other costs are below CHF 0.6 million and 73% of the other insurance income is below CHF 21.6 million.

#### Technical Result and Company Age

Zurich reported in the year 2002 the highest technical loss of CHF 1.2 billion and in the year 2006 Zurich had the highest technical result of CHF 2.4 billion. In Appendix C.2, Figure A10 shows the relative frequency of the technical result and of the company age. In total, 70% report a technical result lower than CHF 5 million. The company age is in the range from founded to 181 years. In our analysis, 67% of the companies have a company age below 100.

#### 4.4.2. Model of the Non-Life Undertakings

Our objective is to find a GAM and a GLM for the non-life undertakings, as presented in Section 4.2. First, we determine the distribution of the equalization reserves for the non-life undertakings. In a second step, we define the GAM by selecting the explanatory variables. Thereafter, we study the result and apply `evtree` to define classifications for explanatory variables. Finally, we apply the GLM by using the discrete and the transformed variables and present the main results.

#### Equalization Reserves' Distribution

Like the reinsurance procedure, we let our equalization reserves observations speak for themselves. We fit the exponential, Weibull, Gamma and log-normal distribution, being of the exponential distribution family, on our data and compare the theoretical and the sample quantiles, presented in Appendix D Figure A12. Furthermore, we calculate the AIC values for these distributions, presented in Table 9. For the non-life undertakings, the Gamma distribution best approximates the response variable, equalization reserves.

**Table 9.** AIC values for possible distributions of equalization reserves (non-life).

	Exponential	Weibull	Gamma	Log-Normal
AIC	15,478	15,053	15,050	15,135
Number of Parameters	1	2	2	1
Log-likelihood	−7738	−7525	−7523	−7566

#### Selection of Explanatory Variables to Define GAM

Our goal is to figure out the relationship between the response variable, equalization reserves, with the explanatory variables, as presented in Section 4.2.2.

In the previous paragraph, we worked out that equalization reserves could be fitted by a Gamma distribution. According to Wood (2010), we implement in R a GAM model with the family Gamma and we use the link function *log*, ignoring the correlation of the explanatory variables, as mentioned in Section 4.2.3.3. We apply the algorithm, outlined in Equation (12), and obtain the GAM model, as follows:

$$\log \mathbb{E}(ER) = f_1(NPA) + f_2(ERP) + f_3(NCR\_woD\_ER) + \beta_4 \cdot YEA + f_5(NEP) + \beta_0. \quad (16)$$

Net claims payments head the selection of the explanatory variables. The equalization reserves of the previous year have the second-highest impact on the equalization reserves. Thereafter, the net change in claim reserve without change in equalization reserves shapes the equalization reserves. The categorical variable calendar year has position four, and the net earned premium terminates the selection of the explanatory variables, impacting the equalization reserves of the non-life undertakings. An additional explanatory variable would not improve the fitting. In Appendix E, Table A4 presents the AIC values for the iterations.

In the first column of Table 10, we list the intercept, the coefficient of the categorical variables, the edf of the continuous variables and the significance levels of the GAM model. We use the star-notation to mark the statistical significance of the variables in Table 10. The smoothing function of the net claims payments has significance code “\*\*\*”, whereby the smoothing functions of the continuous variables equalization reserves of the previous year, net change in claim reserves without change in equalization reserves and net earned premium have each the significance code “\*\*\*\*”. Thus, all smoothing functions of the continuous variables are statistically significant. In addition, the intercept and the coefficient of the year 1997 are statistically significant. The continuous variables equalization reserves of the previous year and the net change in claim reserves without equalization reserves have an edf value of 8.79 and 7.88, having the highest complexity to be smoothed. Net claims payments and net earned premium have an edf value of 4.94 and 3.41, having the lowest complexity to be smoothed. The categorical variable year has 1998 as the baseline.

#### GAM Result and `evtree` to Define Classifications for Explanatory Variable

Visualization of the fitted smoothing curves reveals the highest insight into the relationship of the explanatory variables on the equalization reserves. As outlined above in Equation (13), we defined the optimal classes of the explanatory variable and in Appendix F in Table A6 the used parameters of the `evtrees` are listed. We added this information in Figure 2: for each explanatory variable of the selection, its GAM prediction on the equalization reserves is shown (solid line), the 95% confidence interval is marked (dashed lines), the boundaries of the bins are displayed by the vertical lines and the predicted mean effect for each class is indicated by the dashed-pointed horizontal lines.

#### Net Claims Payments

Net claims payments are shown in Figure 2a, and we find seven classes. Around 300 claims payments, being lower than CHF 180 million, have a negative impact on the equalization reserves. The next class is up to CHF 470 million claims payments, consisting of 45 observations without any impact on the equalization reserves. The higher the claims

payments of the company the more equalization reserves are built up. As a take away, we note that companies with small claims payments tend to release equalization reserves, and the ones with high claims payments tend to establish more equalization reserves. To transform this continuous variable into the discrete one, we define seven classes. Most observations occur in the class CHF 0–180 million, on which we level this variable.

#### Equalization Reserves of the Previous Year

In Figure 2b, the impact of the equalization reserves of the previous year is displayed and eight classes are marked. The first class has around 130 companies, having equalization reserves of the previous year lower than CHF 2.0 million. Its impact on the equalization reserves is negative; small companies more often release the equalization reserves of the previous year to finance the claims. Each of the remaining seven classes have around 40 nodes. To transform this continuous variable to the discrete one, we define eight classes. In the class CHF 0–2.0 million, most observations occur, on which we level this variable.

#### Net Change in Claim Reserves without Change in Equalization Reserves

Net change in claim reserves without change in equalization reserves is presented in Figure 2c and has eight classes. The first classification consists of 20 nodes and the release in claim reserves is higher than CHF 20 million; the effect on the equalization reserves is negative. The next class has around 250 observations covering the release of claim reserves up to CHF 20 million and the accumulation of claim reserves up to CHF 11 million. The average effect on equalization reserves is negative. Within the third class, the claim reserve would accumulate up to CHF 37 million, whereby the impact on the equalization reserves would still be negative half as much as in the second class. The higher the accumulation of claim reserves of the company, the more equalization reserves are built up. Releasing of claim reserves impacts a release in equalization reserves; an accumulation of claim reserves higher than CHF 35 million would impact an accumulation of equalization reserves. In consequence, a small accumulation of claim reserves up to CHF 35 million would shift the reserves from equalization to claim. To transform this continuous variable to the discrete one, we define eight classes. In the class CHF –20–11 million, most observations occur, on which we level this variable.

#### Calendar Year

In Figure 2d, the effect of the calendar year on the equalization reserves is presented; no evtree could be found. The baseline is the year 1998. As for the year 1997, information regarding the equalization reserves of the previous year is missing, and the effect of the year 1997 on the equalization reserves could not be observed. All other years have a small effect on the equalization reserves.

#### Net Earned Premium

Net earned premium is presented in Figure 2e, having five classes. The first class contains around 290 observations up to CHF 300 million net earned premium. This class has a positive effect, being close to 2, on the equalization reserves. In the second class, around 60 companies are captured up to a net earned premium to CHF 1000 million, and the effect is slightly negative. The higher the net earned premium of the company, the fewer equalization reserves are built up. To transform this continuous variable to the discrete one, we define eight classes. In the class CHF 0–300 million, most observations occur, on which we level this variable.

#### GLM Model

We apply the GLM model on the initial discrete variables year and legal form and on the initial continuous variables as transformed and described in the previous paragraph. We mark the transformed variable with “*c*” and Equation (16) turns to the GLM model:



$$\log E(ER) = \beta_1 \cdot NPA_c + \beta_2 \cdot ERP_c + \beta_3 \cdot NCR\_woD\_ER_c + \beta_4 \cdot YEA + \beta_5 \cdot NEP_c + \beta_0. \quad (17)$$

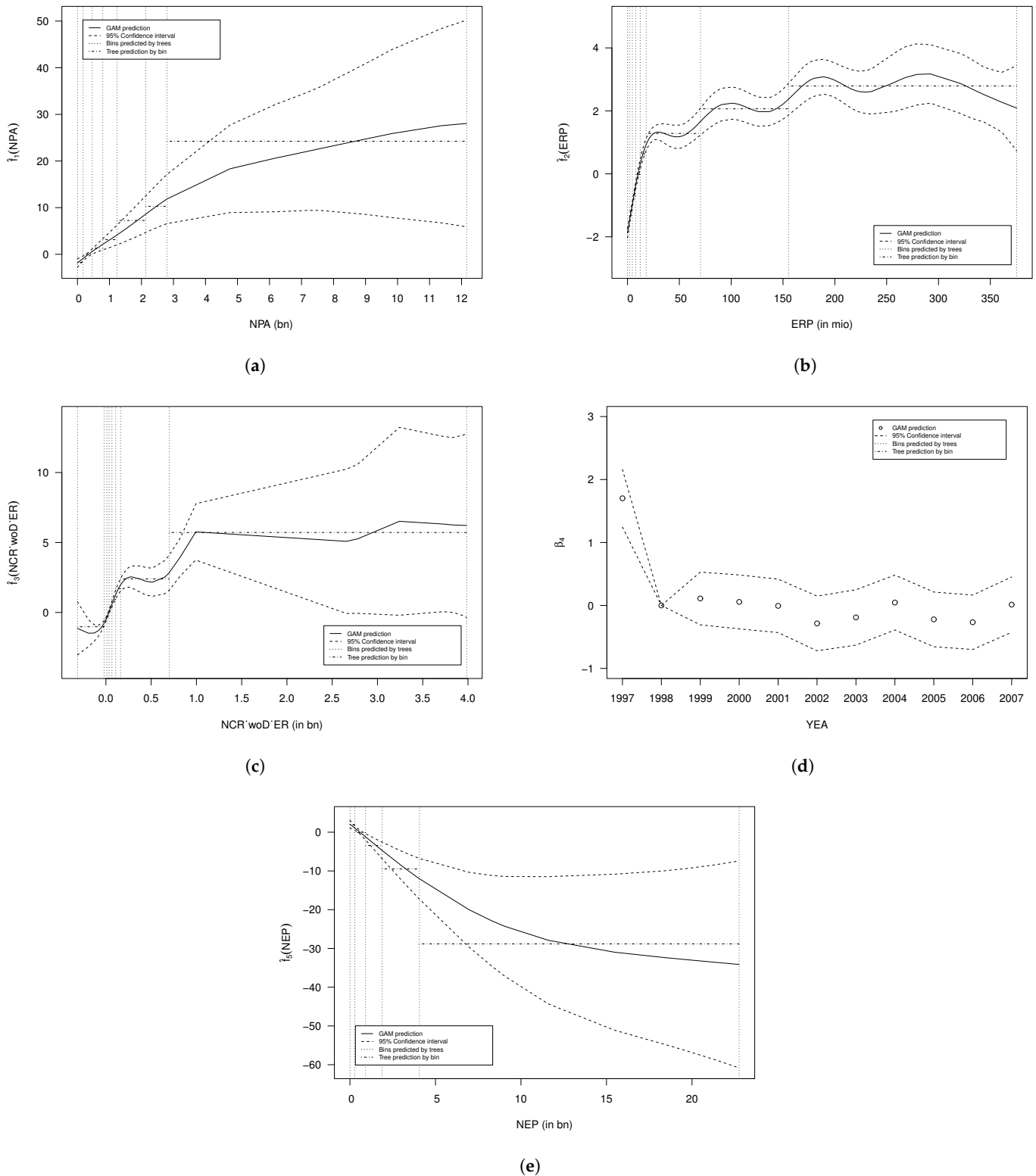
In the second column of Table 10, we list the intercept, the coefficients of the discrete and of the transformed variables and the significance level of the GLM model. In the GAM and GLM models, the intercept of the initial discrete variable year has a statistically significant level “\*\*\*” and amounts to 16.35 for the GAM model and 13.86 for the GLM one. For the discrete variable year in the GLM model, only one coefficient has a statistically significant level “\*” and the other years are not significant. All coefficients of the initial continuous variables  $ERP$  and  $NCR\_woD\_ER_c$  have a significance level of “\*\*\*” or “\*” in the GLM model. Two coefficients of net claims payments have a statistical significant level “\*\*\*” and “\*”; the others are not significant in the GLM model.

In the GLM model for the initial continuous variable net earned premium, three coefficients could not be defined, as net earned premium and net claims payments have an exact linear relationship between them. In Section 4.2.2, we explain how to select the best GAM model by looking at the AIC values. In Appendix E Table A4, we present the AIC values: in the first iteration, the AIC value of the net earned premium is 14,948, while AIC of the net claims payments is 14,944. The net claims payments are the best explanatory variable in the first iteration and the net earned premium is selected in the last iteration. We cannot explain why the last iteration, selecting the net earned premium as the explanatory variable, should improve the GAM model. This is a limitation of our paper.

The AIC values of the GAM and GLM models are reported in Table 7: the GAM model has an AIC value of 14,485 and the GLM one has an AIC value of 14,454, making the GLM a better one.

### Result

For non-life undertakings, the GLM performs better than the GAM model. Our approach to select the best GAM model by looking at AIC values has not detected the strong relationship between net earned premium and net claims payments. Based on the GAM model, we find the use of equalization reserves: accumulation and release of the equalization reserves are observed. From the GAM model, we notice an opposite effect between small-sized and large-sized non-life companies: for small-sized companies, the release effect is observed regarding the insurance expenses, whereby the premium income variable indicates an accumulation one. For large-sized companies, the release effect is observed regarding the premium income variable, whereby the insurance expenses indicate an accumulation effect. Based on the GAM model, we note that non-life undertakings, having small equalization reserves of the previous year, would release the equalization reserves.



**Figure 2.** Effects on the equalization reserves for non-life undertakings. (a) Effect of the net claims payments. (b) Effect of the equalization reserves of the previous year. (c) Effect of the net change in claim reserves without change in ER. (d) Effect of the calendar year. (e) Effect of the net earned premium.

Table 10. Output of the GAM and GLM for non-life undertakings.

GAM			GLM		
Explanatory Variables		Star	Explanatory Variables		Star
(Intercept)	16.35	***	(Intercept)	13.86	***
<b>Categorical:</b>	Coefficient		<b>Categorical:</b>	Coefficient	
	YEA (baseline: 1998)			YEA (baseline: 1998)	
1997	1.70	***	1997	1.81	
1999	0.11		1999	0.01	
2000	0.06		2000	0.04	
2001	−0.01		2001	0.04	
2002	−0.29		2002	−0.45	
2003	−0.19		2003	−0.16	
2004	0.05		2004	0.04	
2005	−0.22		2005	−0.33	
2006	−0.27		2006	−0.47	*
2007	0.01		2007	−0.32	
<b>Continuous:</b>	edf		<b>Categorical:</b>	Coefficient	
s(NPA)	4.94	**		NPA (baseline: 0–180 m.)	
			180–470 m.	−0.40	
			470–800 m.	−0.10	
			800–1500 m.	−0.28	
			1500–2200 m.	−1.37	*
			2200–3000 m.	0.58	
			3000–12,200 m.	−2.06	**
s(ERP)	8.79	***		ERP (baseline: 0–2.0 m.)	
			2.0–4.8 m.	1.57	***
			4.8–8.0 m.	1.90	***
			8.0–12.4 m.	2.36	***
			12.4–18.0 m.	2.78	***
			18.0–71.0 m.	3.38	***
			71.0–160.0 m.	3.97	***
			160.0–380.0 m.	4.42	***
s(NCR_woD_ER)	7.88	***		NCR_woD_ER (baseline: −20–11 m.)	
			−315–−20 m.	0.62	*
			11–37 m.	0.86	***
			37–70 m.	1.06	***
			70–115 m.	1.30	***
			115–170 m.	2.67	***
			170–700 m.	2.76	***
			700–4000 m.	3.91	***
s(NEP)	3.41	***		NEP (baseline: 0–300 m.)	
			300–1000 m.	0.11	
			remaining	na	
AIC	14,485		AIC	14,454	
N	413		N	413	

Star represents significance levels of *p*-values: 0 '\*\*\*\*' 0.001 '\*\*\*' 0.01 '\*\*' 0.05 '.' 0.1 '' 1. edf stands for estimated degree of freedom.

### 5. Discussion

One result of this paper is that the Swiss equalization reserves could be restated out of the publicly available figures for the years 1997 to 2018 for reinsurance undertakings and for the years 1997 to 2007 for the non-life undertakings, although the equalization reserves are not explicitly published. We detect that profit and loss items are identified as explanatory variables for the equalization reserves and that the equalization reserves could not be misused for tax optimization by defining the upper limit and their accumulation and their release. Undertakings need the equalization reserves.

In the literature, the Swiss equalization reserves are not studied, and Swiss undertakings do not publish how their equalization reserves are handled. This paper fills the gap and starts the discussion about the equalization reserves.

Disclosing the equalization reserves reveals the financial cushion of the undertaking and could awaken the covetousness of some greedy short-term investors.

Equalization reserves should smooth the volatility of the case reserves, which should be studied in detail, being a further research topic. Another further research topic could be a study about the determination of the optimal upper limit of the equalization reserves. The basis of this paper is the restated equalization reserves. Thus, the study could be enriched by using the equalization reserves reported by the undertakings. In addition, a future research direction could be the impact of the equalization reserves on the (tied) assets.

## 6. Conclusions

In this paper, we use publicly available data from FOPI and FINMA to restate the equalization reserves for the years 1997 to 2018 for reinsurance and for non-life undertakings in Switzerland and we analyze the relationship between these equalization reserves and the profit and loss items, using a GAM and a GLM model. The analysis is limited to 799 observations for reinsurance undertakings and to 413 observations for non-life undertakings.

For reinsurance undertakings, the equalization reserves depend on the equalization reserves of the previous year, on the calendar year, on the legal form, on the technical result, on the administration and commission costs and on other costs. The GAM model performs better than the GLM one for reinsurance undertakings. For non-life undertakings, the equalization reserves depend on the net claims payments, on the equalization reserves of the previous year, on the net change in claims reserves without change in equalization reserves, on the calendar year and on the net earned premium. The GLM model performs better than the GAM one, for non-life undertakings.

We figure out the use of the equalization reserves for reinsurance and for non-life undertakings: accumulation and release of the equalization reserves are observed. Based on the GAM model, we note that reinsurance and non-life undertakings, having small equalization reserves of the previous year, would release the equalization reserves. From the GAM and from the GLM model, we note that, compared to captives, professional reinsurers tend to build more equalization reserves.

We illustrated that by fixing the accumulation and the release of the equalization reserves within the business plan, no tax advantage could be gained out of the equalization reserves. The equalization reserves belong to the policyholder and to the portfolio.

The discussion about the disclosing of the equalization reserves should be restarted.

After finding a positive consensus, the analysis could be repeated based on more observations and real figures.

Reinsurance and non-life undertakings should reflect whether the definition of the equalization reserves, fixed in the business plan, should be linked to the outcome of the capital modeling for insurance/reserving risk or linked to other profit and loss items or on a mixture of both.

Another future research question could be derived: do the equalization reserves impact the variance of the technical reserves within the statutory accounts for an undertaking? Furthermore, a future research topic could be the missing explanation as to why the last iteration improves the GAM model when defining the best GAM model, whereby the selected last explanatory variable is highly correlated to an already identified one.

The work is limited to publicly available data, whereby the equalization reserves were restated. To improve the GAM and the GLM model, the database should take the reported equalization reserves as a basis. First, this would enlarge the number of observations the analysis is based on. Second, the restated equalization reserve could be replaced by the observation, avoiding unnecessary inaccuracies.

**Author Contributions:** Conceptualization, A.B.; methodology, A.B.; software, A.B. and Y.S.; validation, A.B.; formal analysis, A.B.; investigation, A.B.; data curation, A.B.; writing—original draft preparation, A.B.; writing—review and editing, A.B.; visualization, A.B.; supervision, A.B.; project administration, A.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** In this section, we provide the links to the publicly available raw data. Link to FOPI's data: [finma.ch/FinmaArchiv/bpv/e/dokumentation/00439/01389/index.html?lang=en](https://finma.ch/FinmaArchiv/bpv/e/dokumentation/00439/01389/index.html?lang=en) (accessed on 1 March 2018); Link to FINMA's data: [finma.ch/en/documentation/archiv/insurers-reporting-portal/](https://finma.ch/en/documentation/archiv/insurers-reporting-portal/) (accessed on 1 March 2018).

**Acknowledgments:** The authors are very thankful to Joël Wagner for his helpful comments and suggestions and for organizing the funding of the publication.

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A. Main Changes within FOPI's and FINMA's Reporting

The appendix starts in Appendix A with a short overview of the main changes within FOPI's and FINMA's reporting. In Appendix B the source of the available explanatory variables within the R model is presented: once for FOPI in Table A1 and once for FINMA in Appendix B.2. Thereafter, in Appendix C the relative frequency of all available explanatory variables is presented: once for reinsurance undertakings in Appendix C.1 and once for non-life undertakings in Appendix C.2. In Appendix D the fit of the response variable's distribution function, equalization reserves, is depicted for reinsurance and for non-life undertakings. Then, in Appendix E the AIC values for reinsurance and for non-life undertakings are shown. Finally, in Appendix F the parameters of the *evtree* for reinsurance and for non-life undertakings are listed.

FOPI and FINMA have disclosed the figures in a slightly different setup. The main changes are:

- For life undertakings, FINMA changed the presentation of the number of individual and collective contracts and the sum insured, written directly in Switzerland;
- For non-life undertakings, FINMA changed the presentation of the number of contracts and number of risks, written directly in Switzerland;
- For life and for non-life undertakings, FINMA changed the presentation of the equalization reserves, written directly in Switzerland;
- For life undertakings, FINMA publishes the written periodic premium of the direct business in Switzerland;
- FOPI published the number of workers within the insurance industry;
- FOPI published the profit and loss and the balance sheet information of each single (re-)insurance, undertaking each in one file. Furthermore, FOPI published these gross and net items; FINMA publishes only the net ones;
- For non-life undertakings, FOPI published tied assets information, whereas FINMA publishes tied assets for life and non-life undertakings within one file. Furthermore, the composition of "Sollbetrag" changed over time.

## Appendix B. Source of the Available Explanatory Variables

### Appendix B.1. FOPI's Available Explanatory Variables

**Table A1.** FOPI published the files “Jahresrechnungen” in each calendar year for each undertaking. This is the source of the available explanatory variables for reinsurance and non-life undertakings.

Variable	FOPI's Title and Calculation	Cell
ACC	Aufwendungen für den Versicherungsbetrieb für eigene Rechnung	C32
COA	Calendar Year - Foundation Year (as stated in Zefix)	
LEF	Included in the name of the undertaking (as stated in Zefix)	
NCR_woD_ER	Veränderung der Schadenrückstellungen für eigene Rechnung $NCR\_woD\_ER = C19 + ER - ERP$	C19
NEP	Verdiente Prämien für eigene Rechnung	C9
NOR	Nicht anderweitig auszuweisende Veränderung vt. Nettorückstellungen	C21
NPA	Zahlungen für Versicherungsfälle für eigene Rechnung	C16
	Nicht realisierte Verluste aus Kapitalanlagen für anteilgebundene Lebensversicherungen	C33
OCC	Sonstige vt. Aufwendungen für eigene Rechnung $OCC = C33 + C34$	C34
	Der technischen Rechnung zugeordneter Zinsertrag für eigene Rechnung	C10
OIN	Nicht realisierte Gewinne aus Kapitalanlagen für anteilgebundene Lebensversicherungen	C11
	Sonstige vt. Erträge für eigene Rechnung $OIN = C10 + C11 + C12$	C12
TRE	Ergebnis der vt. Rechnung	C35
YEA	Calendar Year	

“vt.” stands for versicherungstechnisch. Zefix is the Swiss Companies's register.

### Appendix B.2. FINMA's Available Explanatory Variables

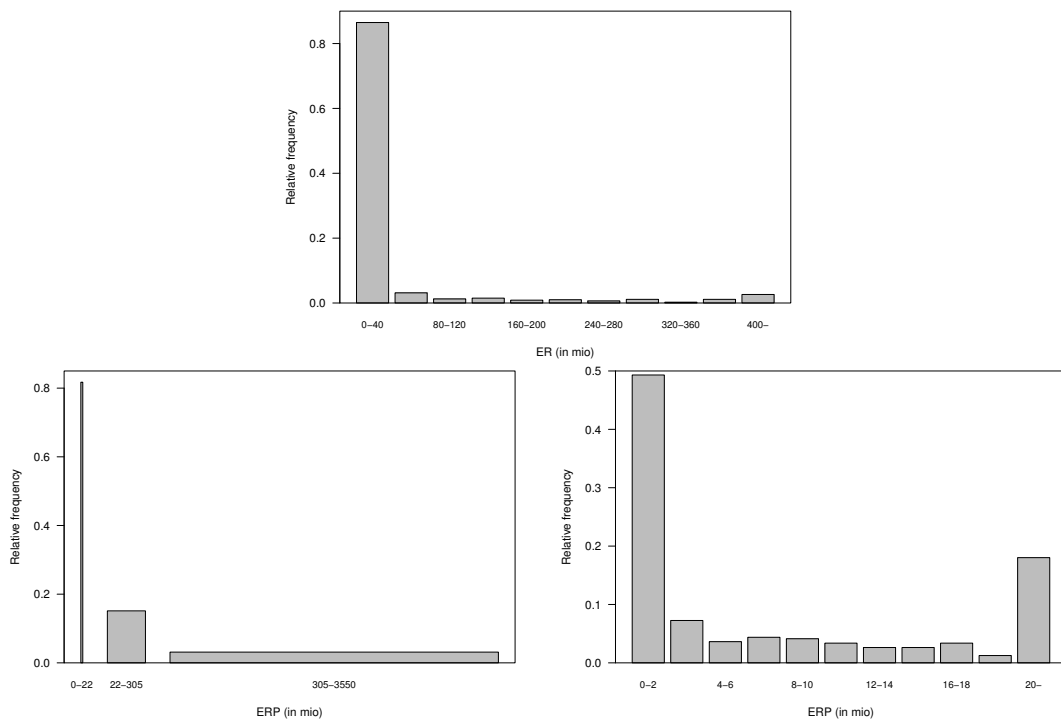
**Table A2.** FINMA published the data within a “cube” for the years 2008 to 2018. Section “Bilanz and Erfolgsrechnung” is the source of the available explanatory variables for reinsurance and non-life undertakings.

Variable	FINMA's Title and Calculation	Section
ACC	Aufwendungen für den Versicherungsbetrieb für eigene Rechnung	Erfolgsrechnung - Aufwand
COA	Calendar Year - Foundation Year (as stated in Zefix)	
LEF	Included in the name of the undertaking (as stated in Zefix)	
NCR_woD_ER	Veränderung der Rückstellungen für Versicherungsfälle/ Schadenrückstellungen (Leben und Schaden) $NCR\_woD\_ER = \Delta \text{claims reserves} + ER - ERP$	Erfolgsrechnung - Aufwand vt. Aufwand
NEP	Versicherungstechnische Erträge Veränderung der übrigen vt. Verbindlichkeiten für eigene Rechnung (Leben) $=: X_1$	Erfolgsrechnung - Ertrag Erfolgsrechnung - Aufwand vt. Aufwand
NOR	Veränderung der übrigen vt. Verbindlichkeiten für eigene Rechnung (Schaden) $=: X_2$ $NOR = X_1 + X_2$	Erfolgsrechnung - Aufwand vt. Aufwand
NPA	Zahlungen für Versicherungsfälle für eigene Rechnung (Leben und Schaden)	Erfolgsrechnung - Aufwand vt. Aufwand
OCC	Sonstige Aufwendungen aus der Versicherungstätigkeit	Erfolgsrechnung - Aufwand
OIN	Sonstige Erträge aus der Versicherungstätigkeit vt. Aufwendungen	Erfolgsrechnung - Ertrag Erfolgsrechnung - Aufwand
TRE	$=: X_3$ $TRE = NEP + OIN + X_3 + OCC + ACC$	
YEA	Calendar Year	

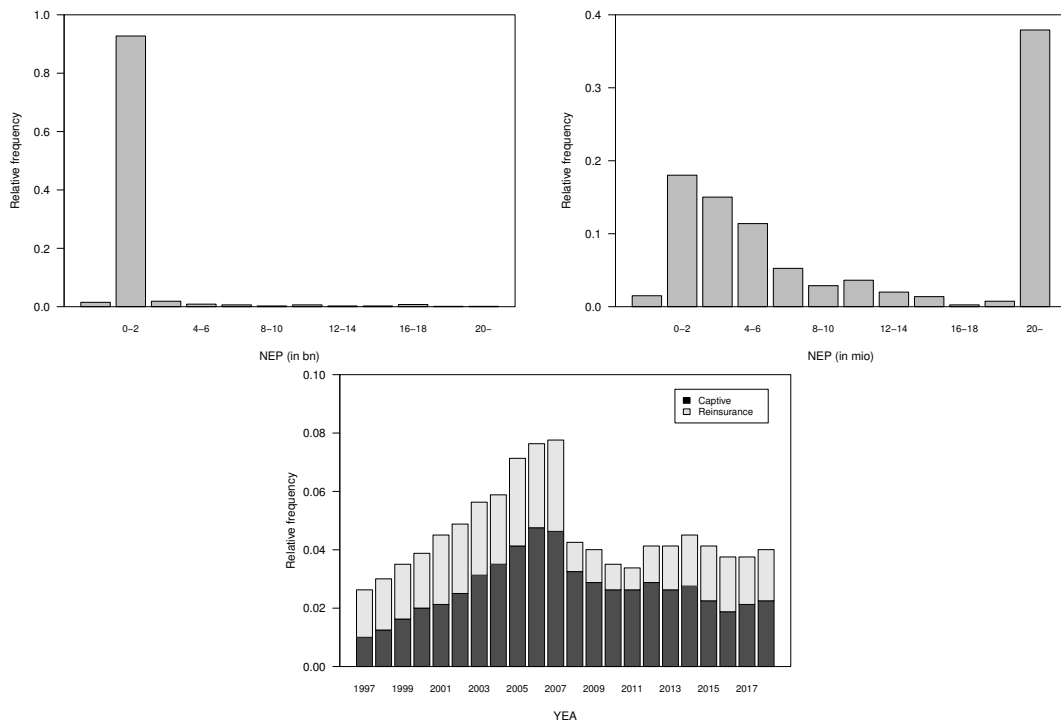
“vt.” stands for versicherungstechnisch. FINMA's cube offers the possibility to download the data as selected; we state the section where the items are reported. Income items have positive and costs negative signs. Zefix is the Swiss Companies's register.

### Appendix C. Relative Frequency

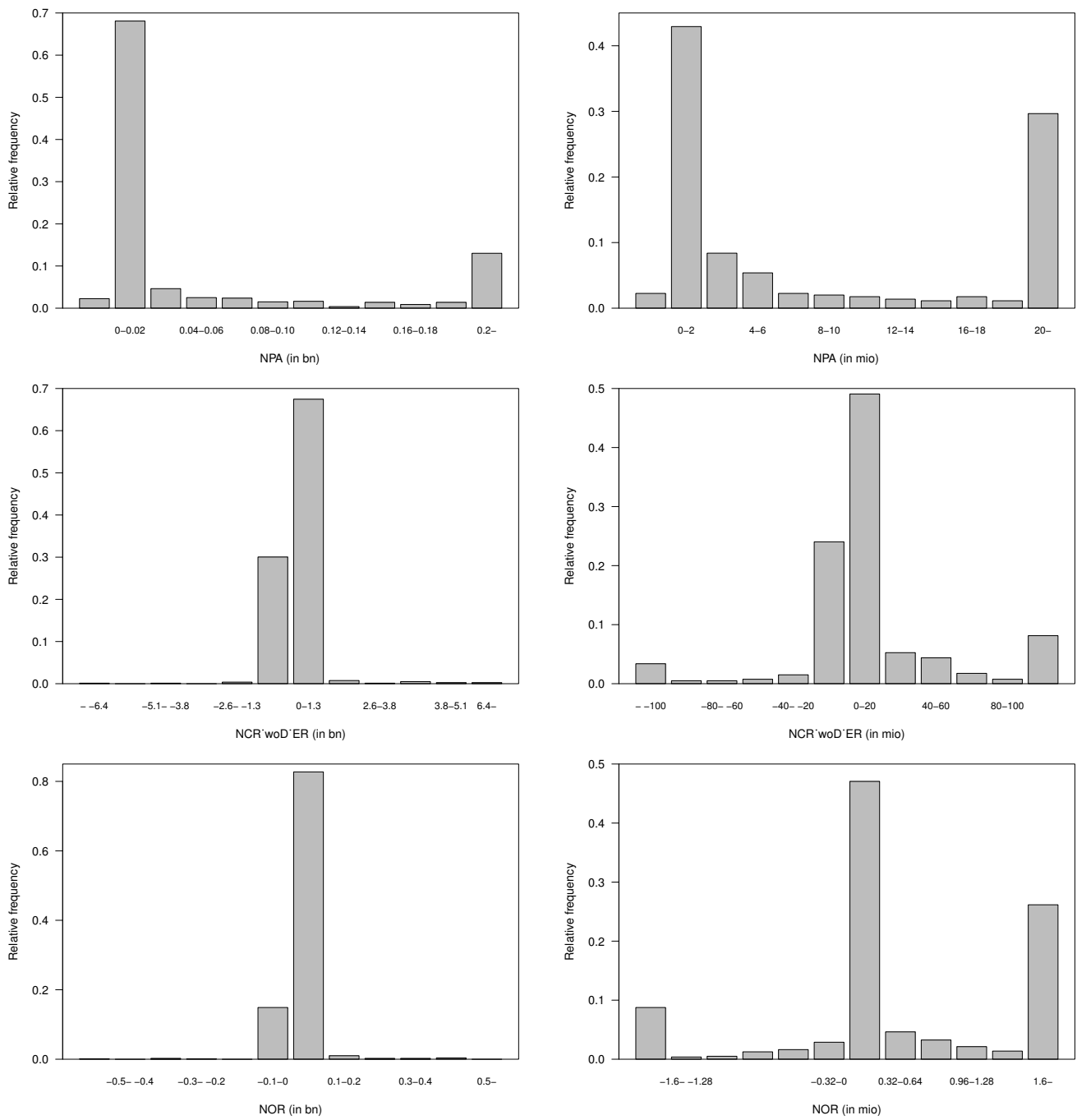
#### Appendix C.1. Relative Frequency—Reinsurance Undertakings



**Figure A1.** Relative frequency of the equalization reserves (*ER*) and of the one of the prior year (*ERP*) for reinsurance undertakings. The figure presents twice the relative frequency of *ERP*: once sliced by the optimal bin out of our model and once sliced by equal bands, whereby the last band includes all remaining higher observations.

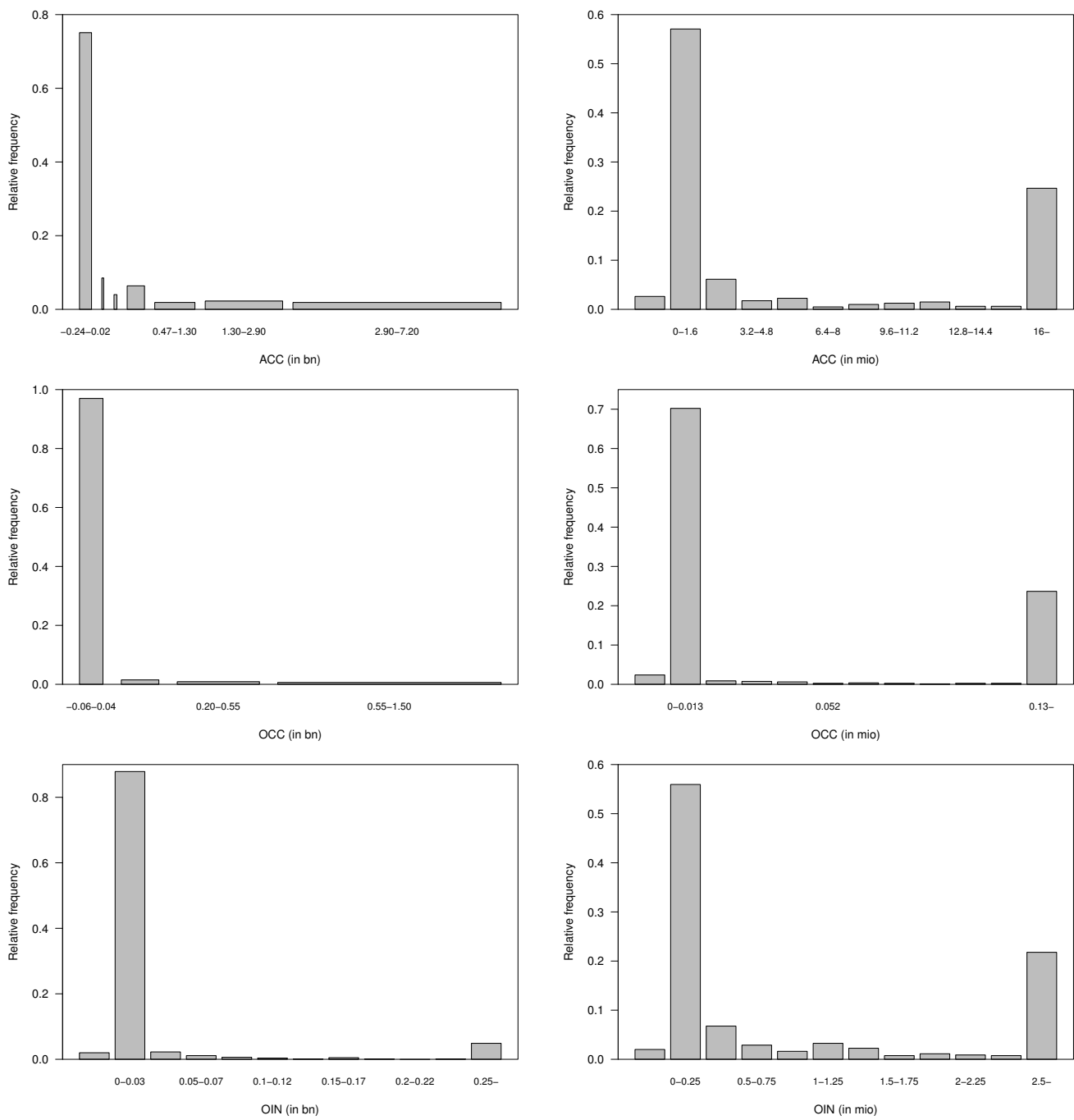


**Figure A2.** Relative frequency of the net earned premium (*NEP*) and of the calendar year (*YEA*) of reinsurance undertakings. The figure presents twice the relative frequency of the *NEP*: once sliced by billion and once sliced by million, whereby the last band includes all remaining higher observations.

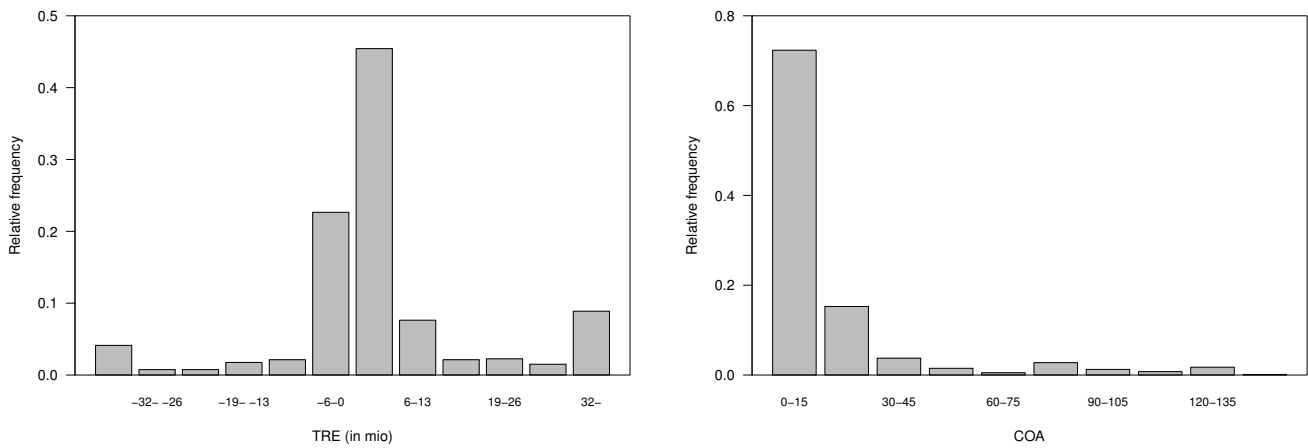


**Figure A3.** Relative frequency of the insurance cost for reinsurance undertakings: net claims payments (*NPA*), net change in claim reserves without change in *ER* (*NCR\_woD\_ER*) and net change in other technical reserves (*NOR*). The figure presents twice the relative frequency of the *NPA*, *NCR\_woD\_ER* and *NOR*: once sliced by billion and once sliced by million, whereby the last band includes all remaining higher observations.



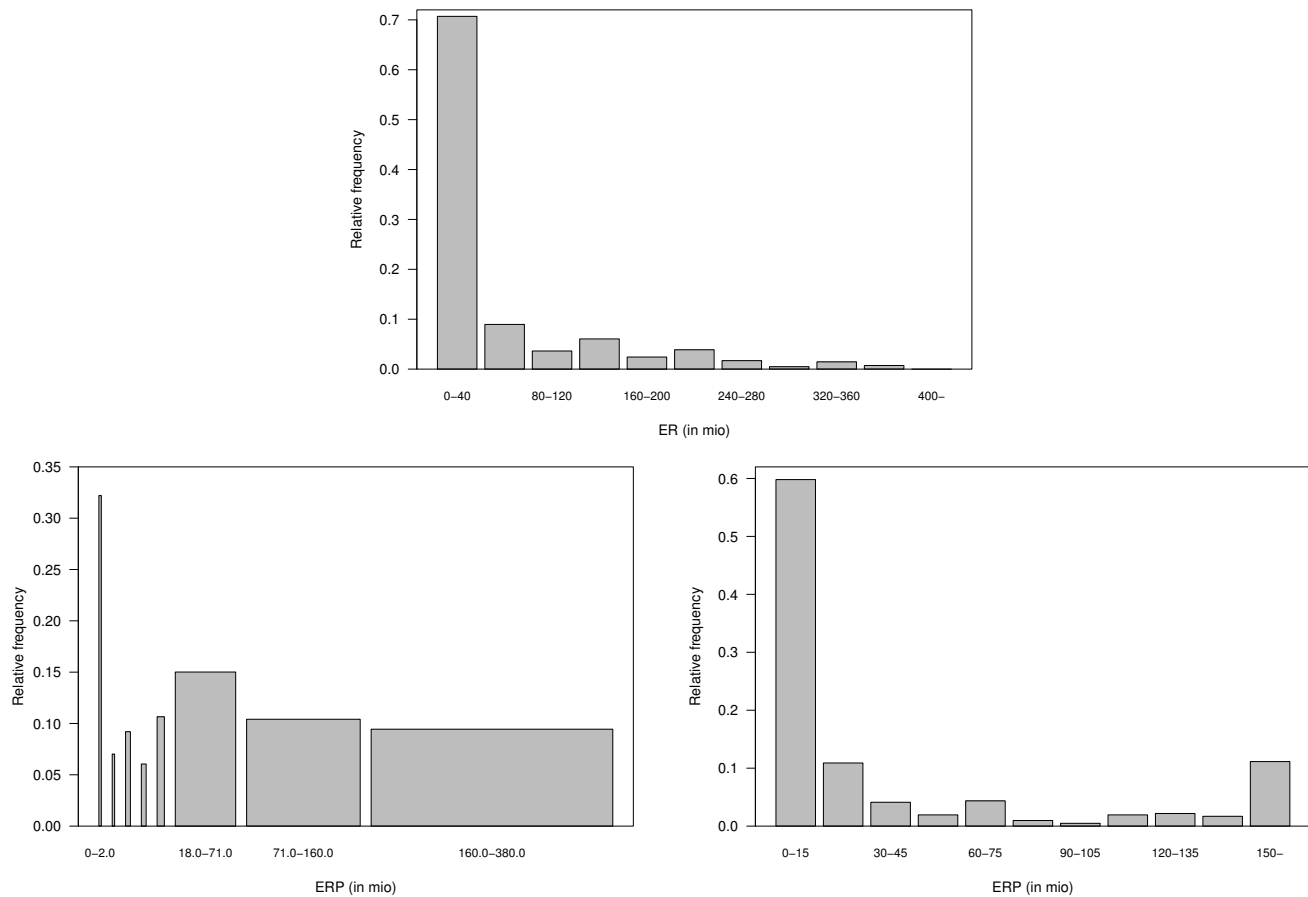


**Figure A4.** Relative frequency of the operation cost, administration cost and commission (*ACC*) and other costs (*OCC*), and of the other income (*OIN*) for reinsurance undertakings. The figure presents twice the relative frequency of *ACC*, *OCC* and *OIN*: *ACC* and *OCC* are once sliced by the optimal bin out of our model and once sliced by equal bands. *OIN* is once sliced by billion and once sliced by million.

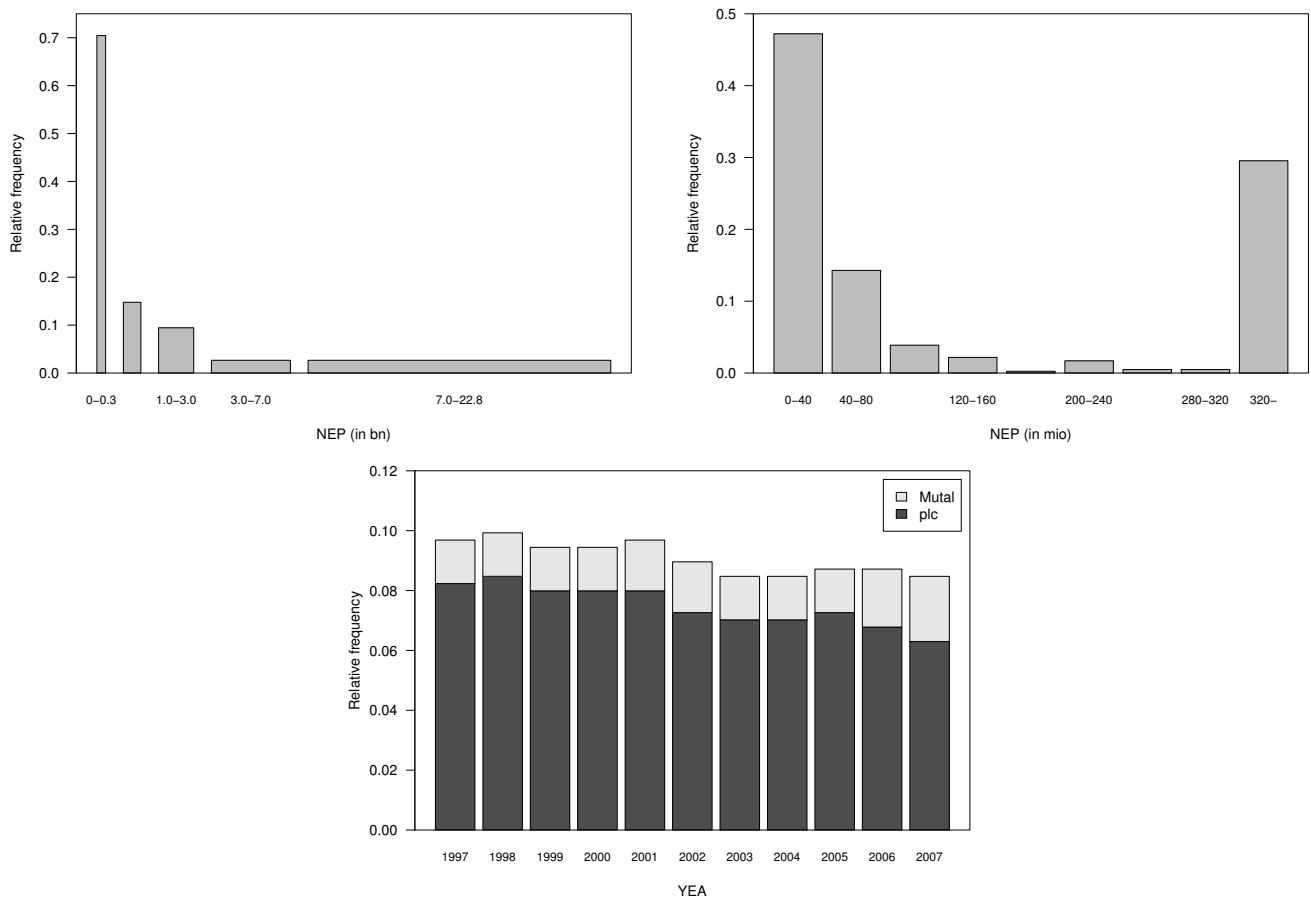


**Figure A5.** Relative frequency of the technical result (*TRE*) and of the company age (*COA*) for reinsurance undertakings.

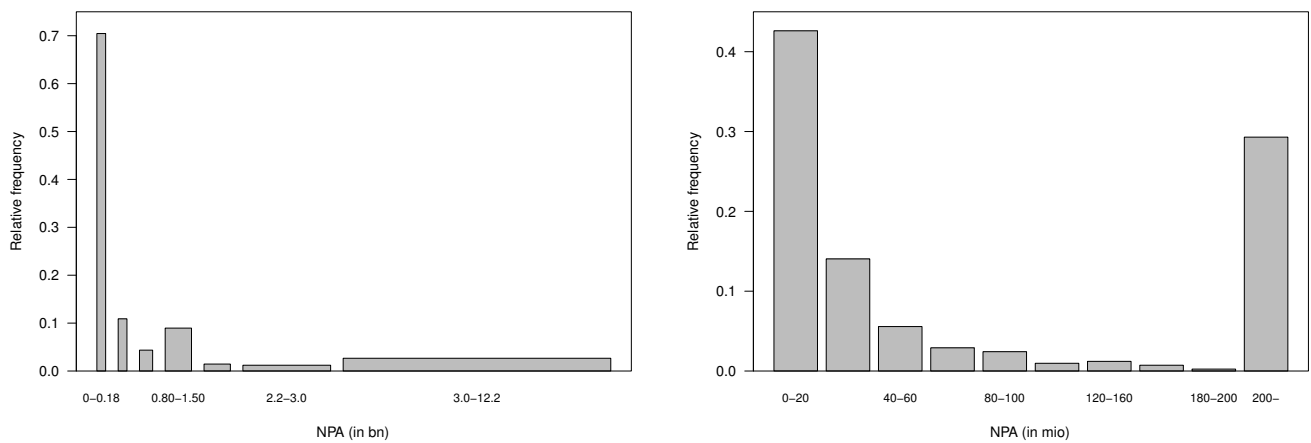
*Appendix C.2. Relative Frequency—Non-Life Undertakings*



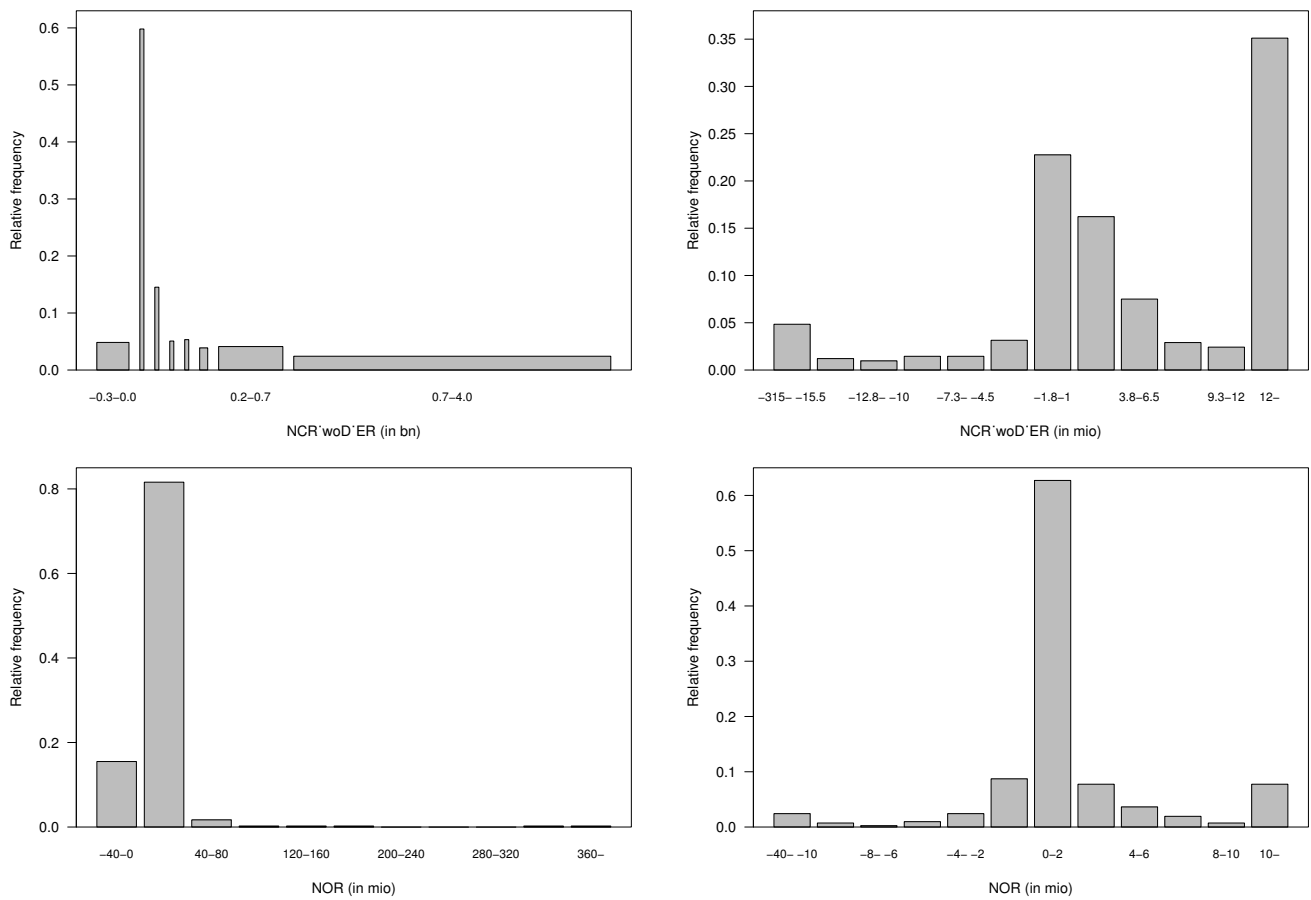
**Figure A6.** Relative frequency of the equalization reserves (*ER*) and of the one of the prior year (*ERP*) for non-life undertakings. The figure presents twice the relative frequency of *ERP*: once sliced by the optimal bin out of our model and once sliced by equal bands, whereby the last band includes all remaining higher observations.



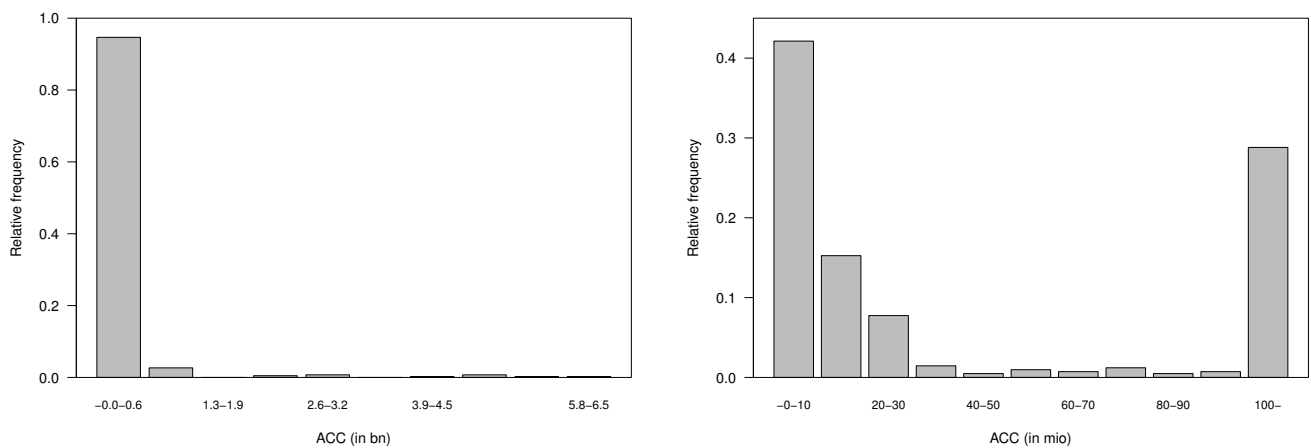
**Figure A7.** Relative frequency of the net earned premium (*NEP*) and of the calendar year (*YEA*) for non-life undertakings. The figure presents twice the relative frequency of the *NEP*: once sliced by the optimal bin out of our model and once sliced by equal bands, whereby the last band includes all remaining higher observations.



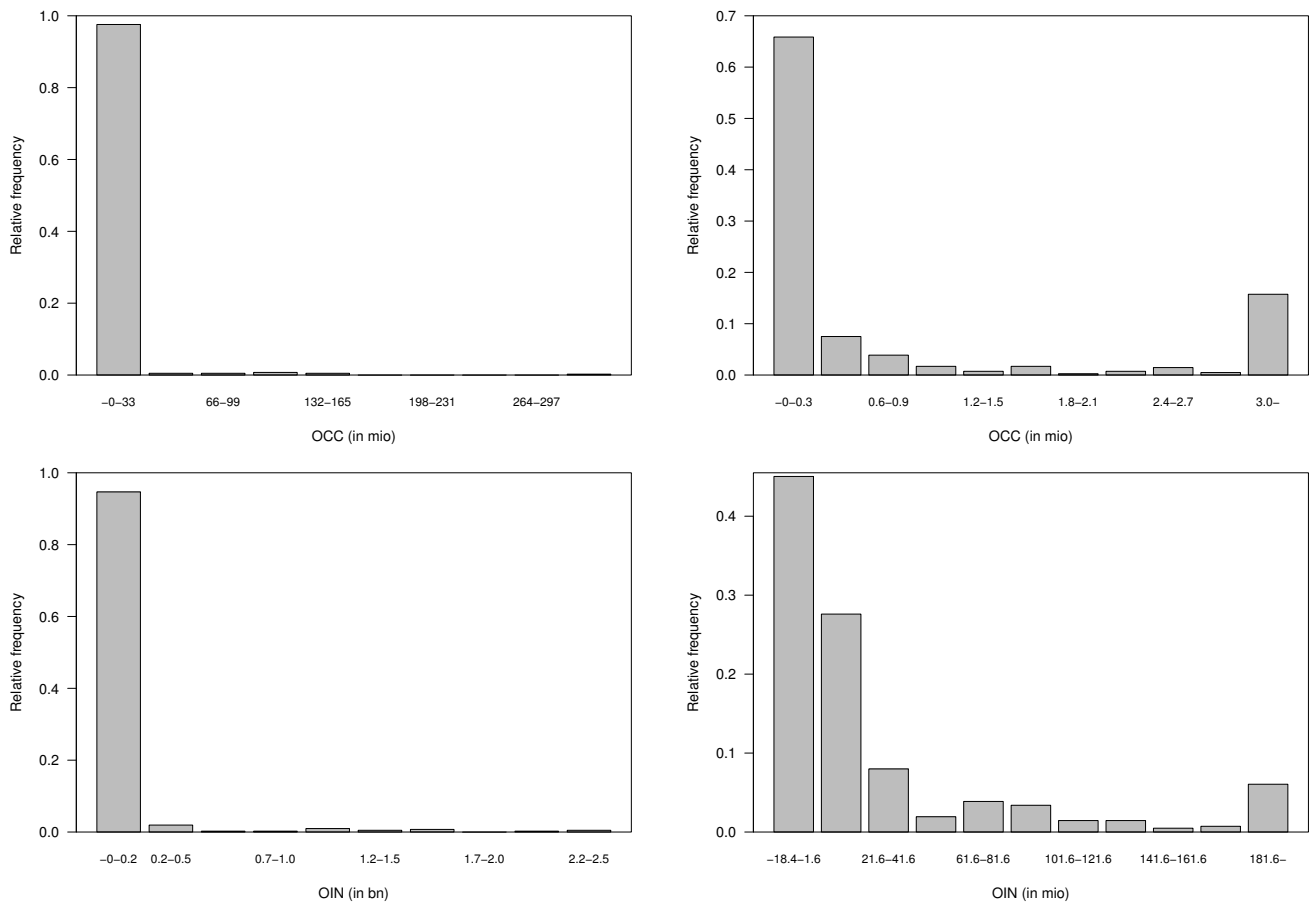
**Figure A8.** *Cont.*



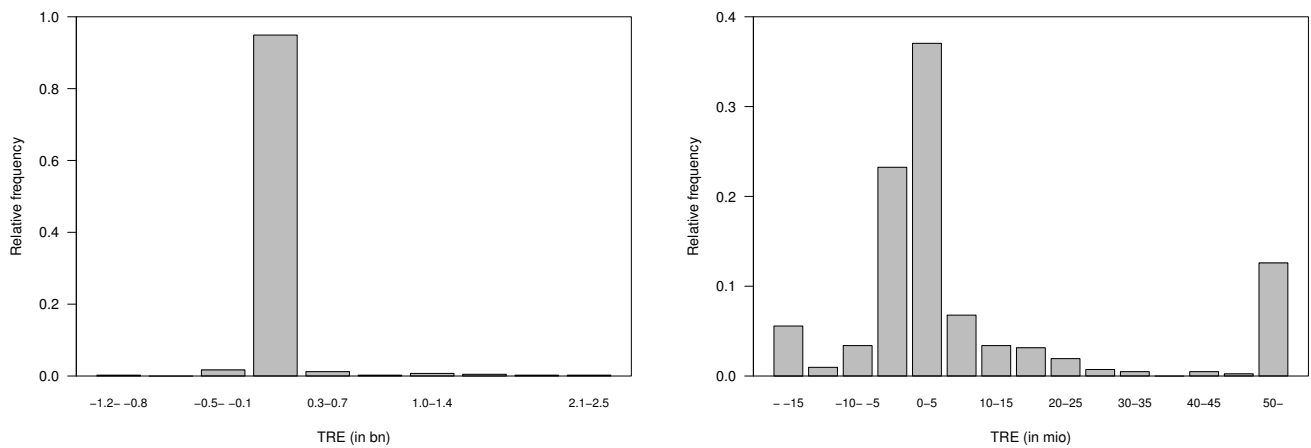
**Figure A8.** Relative frequency of the insurance cost for non-life undertakings: net claims payments (*NPA*), net change in claim reserves without change in *ER* (*NCR\_woD\_ER*) and net change in other technical reserves (*NOR*). The figure presents twice the relative frequency of the *NPA* and *NCR\_woD\_ER*: once sliced by the optimal bin out of our model and once sliced by equal bands, whereby the last band includes all remaining higher observations. *NOR* is sliced in two different intervals.



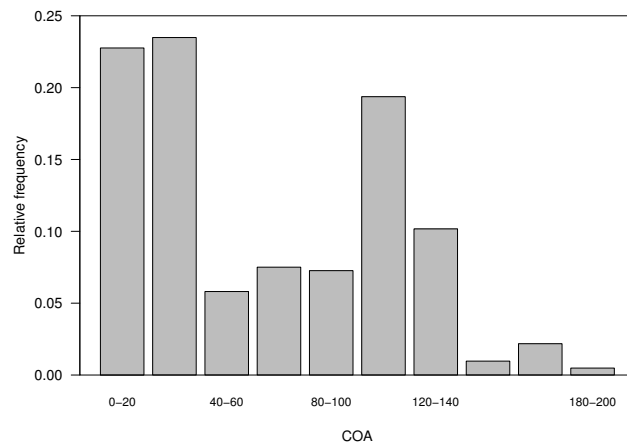
**Figure A9.** *Cont.*



**Figure A9.** Relative frequency of the operation cost, administration cost and commission (*ACC*) and other costs (*OCC*), and of the other income (*OIN*) for non-life undertakings. The figure presents twice the relative frequency of *ACC*, *OCC* and *OIN*: *ACC* and *OCC* are once sliced by billion and once sliced by million. *OIN* is sliced in two different intervals.

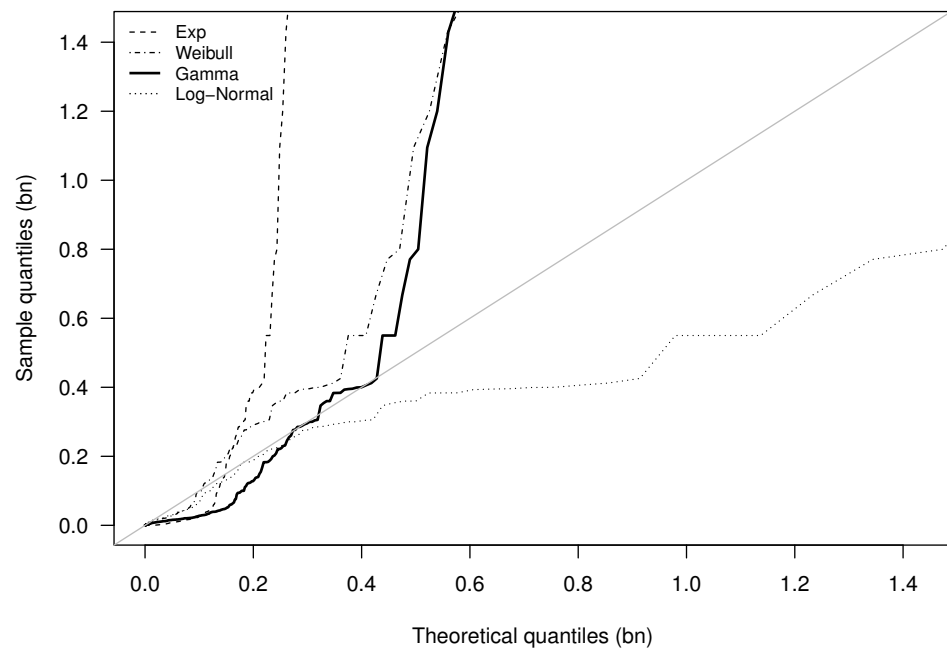


**Figure A10.** Cont.



**Figure A10.** Relative frequency of the technical result (*TRE*) and of the company age (*COA*) of non-life undertakings. The figure presents twice the relative frequency of *TRE*: once sliced by billion and once sliced by million.

**Appendix D. Fit of the Distribution Function of the Equalization Reserves**



**Figure A11.** Fit of distribution of the equalization reserves for reinsurance undertakings.

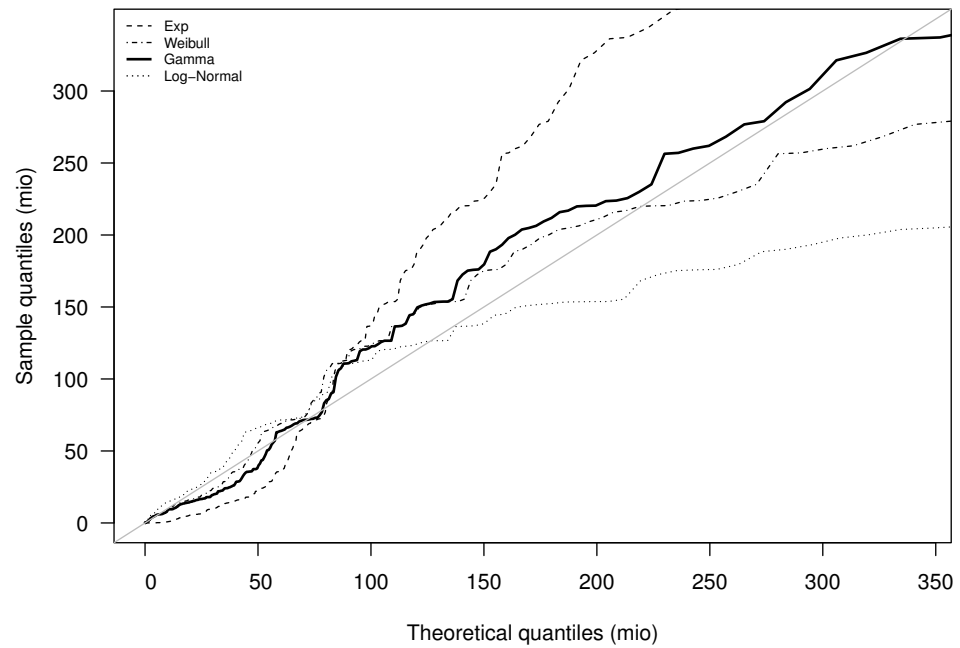


Figure A12. Fit of distribution of the equalization reserves for non-life undertakings.

Appendix E. Model Output—AIC Values

Table A3. AIC values of the explanatory variables for the reinsurance undertakings.

AIC	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5	Iteration 6
ACC	3761	3466	3303	3257	3232	0
COA	3894	3466	3354	3262	3252	3259
ERP	3528	0	0	0	0	0
LEF	3888	3468	3289	0	0	0
NCR_woD_ER	3854	3519	3378	3274	3247	3224
NEP	3742	3466	3312	3257	3241	3233
NOR	3885	3521	3402	3286	3256	3234
NPA	3744	3478	3314	3261	3249	3234
OCC	3885	3518	3401	3283	3248	3218
OIN	3824	3507	3343	3264	3241	3221
TRE	3816	3489	3357	3256	0	0
YEA	3820	3406	0	0	0	0

Table A4. AIC values of the explanatory variables for the non-life undertakings.

AIC	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 5
ACC	14,952	14,895	14,929	14,585	14,493
COA	15,076	14,995	14,677	14,643	14,608
ERP	14,999	14,793	0	0	0
NCR_woD_ER	14,971	14,942	14,591	0	0
NEP	14,948	14,946	14,930	14,572	14,485
NOR	15,020	14,948	14,932	14,583	14,501
NPA	14,944	0	0	0	0
OCC	15,006	14,949	14,992	14,588	14,500
OIN	15,003	14,908	14,866	14,593	14,501
TRE	15,000	14,946	14,926	14,593	14,502
YEA	15,070	14,957	14,674	14,500	0
LEF	15,052	14,828	14,747	14,592	14,502

## Appendix F. Parameters of evtree

Table A5. Parameters of the evtree for reinsurance undertakings.

Explanatory Variables	Reinsurance				
	<i>A</i>	$\hat{\alpha}$	<i>B</i>	$\hat{b}$	# Classes
<i>ERP</i>		2	{5, 10, 15, 20, 25, 30, 35, 50}	30	3
<i>YEA</i>	{1, 1.5, 2, ..., 9.5,	950	{10}	10	2
<i>TRE</i>	10, 15, 20, ..., 95,	950	{10, 20, 30, 50, 100}	20	2
<i>ACC</i>	100, 150, 200, ..., 950}	1	{5, 10, 15, 20, 30}	15	7
<i>OCC</i>		15	{5, 10, 15, 20, 30}	5	4

For all explanatory variable the same set *A* is used.

Table A6. Parameters of the evtree for non-life undertakings.

Explanatory Variables	Non-Life				
	<i>A</i>	$\hat{\alpha}$	<i>B</i>	$\hat{b}$	# Classes
<i>NPA</i>	{1, 1.5, 2, ..., 9.5,	1	{5, 10, 15}	5	7
<i>ERP</i>	10, 15, 20, ..., 95,	4.5	{5, 10, 15, 50, 100}	15	8
<i>NCR_woD_ERP</i>	100, 150, 200, ..., 950}	2.5	{5, 10, 15}	10	8
<i>NEP</i>		7	{5, 10, 15}	10	5

For all explanatory variable the same set *A* is used.

## Notes

- 1 In German: Versicherungsaufsichtsgesetz—VAG, see ISA (2006).
- 2 In German: Aufsichtsverordnung—AVO, see ISO (2006).
- 3 In FINMA (2011) the German expression “Schwankungsrückstellungen” is used.
- 4 For non-life undertakings the German expression “Sicherheits- und Schwankungsrückstellungen” is used to reflect the nature of the reserves “safety” and “fluctuation”. In this paper we simplify this expression by using the name “equalization reserves”; however, another reasonable name would be “fluctuation reserves”.
- 5 For the Credit Insurance line of business a methodology is given; see Swiss Confederation (1993).
- 6 In FER (Foundation for Accounting and Reporting Recommendations 2018) the German words “Schwankungsrückstellungen” and “Sicherheitsrückstellungen” are used; in this paper we refer with “equalization reserve” to both of these reserves.
- 7 With respect to Art. 958 of OR, the financial statements comprise balance sheet, profit and loss accounts and the notes of these accounts; the financial statements “must be signed by the chair-person of the supreme management or administrative body and the person responsible for financial reporting within the undertaking”.
- 8 The statutory accounting is the basis for the calculation, see FINMA (2019c), whereby the Swiss Code of Obligations (OR) regulates the statutory accounting, see Swiss Confederation (1912).
- 9 For the years 1997 to 2018, the structure of the name of the statistical files is, as follows: two letters, two figures, followed by one letter. The first two letters mark the type of insurance such as reinsurance or non-life undertakings. The two figures refer to the line of business, in general. The last letter indicates the reporting item, in general.
- 10 In German: “Pauschalrückstellung”.
- 11 In order to keep the terminology simple, we use the word “reinsurance” designated for both captives and for professional reinsurers. In the subsection for reinsurance, we are more precise.
- 12 Burnham and Anderson (2004) study AIC and explain the relationship between reasonable data and a good model by separating “information” and “noise”. “Here, information relates to the structure of relationships, estimates of model parameters, and components of variance. Noise, then, refers to the residuals: variation left unexplained.” Burnham and Anderson (2004).
- 13 To cite Fridley (2010): “GAMs take each predictor variable in the model and separate it into sections (delimited by ‘knots’), and then fit polynomial functions to each section separately, with the constraint that there are no kinks at the knots (second derivatives of the separate functions are equal at the knots).”
- 14 Simplicity within the notation is made; for  $i = 1$  we define  $X \setminus \{x_1, \dots, x_0\} = X$ . Furthermore, we neglect that categorical variables should not have any smoothing functions.
- 15 Grubinger et al. (2014) explain that the parameter alpha of an evtree “regulates the complexity part of the cost function. Increasing values of alpha encourage decreasing tree sizes.”



- <sup>16</sup> Notation is simplified: e.g., CHF 3.53 billion (Swiss Re 2001) is the shortcut for “the reinsurance undertaking Swiss Re reports in the calendar year 2001 CHF 3.53 billion”.

## References


- Akhurst, Ron, Kathryn Morgan, George Orros, Jonathan Piper, Peter Hinton, Andrew Macnair, Teivo Pentikainen, Joachim Hertig, Charles Levi, Arne Sandström, and et al. 1992. Equalisation Reserves on a European Basis [Conference]. *General Insurance Convention*. Available online: <https://www.actuaries.org.uk/system/files/documents/pdf/0393-0461.pdf> (accessed on 2 October 2019).
- Burnham, Kenneth Paul, and David Anderson. 2004. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research* 33: 261–304.
- Dacorogna, Michel, Hansjörg Albrecher, Michael Moller, and Suzane Sahiti. 2013. Equalization reserves for natural catastrophes and shareholder value: A simulation study. *European Actuarial Journal* 3: 1–21. [CrossRef]
- De Vylder, Florian, and Marc Goovaerts. 1999. Solvency margins and equalization reserves. *Insurance: Mathematics and Economics* 24: 103–15. [CrossRef]
- Dunne, Carla, Donal Lehane, Eimear McCarthy, Angela McNally, Maaz Mushir, Ciara Regan, and David Walsh. 2017. IFRS 17 Insurance Contracts: Impacts for Reserving. 2 August 2017. Deloitte & Touche. Available online: <https://de.scribd.com/document/379850262/Ie-2017-IFRS-17-Impacts-for-Reserving-Deloitte-Ireland> (accessed on 7 October 2019).
- England, Peter, Richard Verrall, and Mario Valentin Wüthrich. 2019. On the lifetime and one-year views of reserve risk, with application to IFRS 17 and Solvency II risk margins. *Insurance: Mathematics and Economics* 85: 74–88. [CrossRef]
- European Insurance and Occupational Pensions Authority (EIOPA). 2015. Guidelines on the Valuation of Technical Provisions. Available online: [https://www.eiopa.europa.eu/sites/default/files/publications/eiopa\\_guidelines/tp\\_final\\_document\\_en.pdf](https://www.eiopa.europa.eu/sites/default/files/publications/eiopa_guidelines/tp_final_document_en.pdf) (accessed on 4 September 2018).
- Farny, Dieter. 2011. *Versicherungsbetriebslehre*. Karlsruhe: Verlag Versicherungswirtschaft.
- Federal Office of Private Insurance. 2007. Facts and Figures [Webside]. Available online: <https://www.finma.ch/FinmaArchiv/bpv/d/dokumentation/00439/01389/index.html?lang=de> (accessed on 1 September 2019).
- Foundation for Accounting and Reporting Recommendations—FER. 2018. DE: Swiss GAAP FER 40 Rechnungslegung für Versicherungsunternehmen; EN: Swiss GAAP FER 40 Consolidated Financial Statements of Insurance Companies. Available online: <https://www.fer.ch/content/uploads/2016/09/Swiss-GAAP-FER-40-final.pdf> (accessed on 10 October 2019).
- Fridley, Jason. 2010. Generalized Additive Models in R [Website]. Available online: <http://plantecology.syr.edu/fridley/bio793/gam.html> (accessed on 25 September 2021).
- Grubinger, Thomas, Achim Zeileis, and Karl-Peter Pfeiffer. 2014. emtree: Evolutionary Learning of Globally Optimal Classification and Regression Trees in R. *Journal of Statistical Software* 61: 1–29. [CrossRef]
- Hastie, Trevor John, and Robert John Tibshirani. 1990. *Generalized Additive Models*. London: Chapman and Hall/CRC.
- Hindley, David. 2017. *Claims Reserving in General Insurance*. Cambridge: Cambridge University Press.
- Jeong, Himchan, and Dipak Dey. 2020. Application of a vine copula for multi-line insurance reserving. *Risks* 8: 111. [CrossRef]
- Merz, Michael, Mario Valentin Wüthrich, and Enkelejd Hashorva. 2012. *Dependence Modelling in Multivariate Claims Run-Off Triangles*. Cambridge: Cambridge University Press.
- Parzen, Emanuel, Kunio Tanabe, and Genshiro Kitagawa. 1998. *Selected Papers of Hirotugu Akaike*. New York: Springer.
- R Core Team. 2021. R: A Language and Environment for Statistical Computing. Available online: <https://www.r-project.org/> (accessed on 25 September 2021).
- Sandström, Arne. 2005. *Solvency: Models, Assessment and Regulation*. London: Chapman & Hall/CRC.
- Schweizer Aktuar Vereinigung—SAV. 2006. Guidelines for Loss Reserves in Non-Life Insurance. Available online: [https://www.actuaries.ch/de/downloads/aid!b4ae4834-66cd-464b-bd27-1497194efc96/id!48/Schadenrueckstellungen\\_NL\\_EN.pdf](https://www.actuaries.ch/de/downloads/aid!b4ae4834-66cd-464b-bd27-1497194efc96/id!48/Schadenrueckstellungen_NL_EN.pdf) (accessed on 4 September 2019)
- Shi, Peng, and Edward Frees. 2011. *Dependent Loss Reserving Using Copulas*. Cambridge: ASTIN Bulletin.
- Spector, Phil. 2011. Stat 133 Class Notes [Website]. Available online: <https://www.stat.berkeley.edu/~s133/all2011.pdf> (accessed on 25 September 2021).
- Staudt, Yves, and Joël Wagner. 2021. Assessing the performance of random forests for modeling claim severity in collision car insurance. *Risks* 9: 53. [CrossRef]
- Swiss Confederation. 1912. DE: *Bundesgesetz Betreffend die Ergänzung des Schweizerischen Zivilgesetzbuches (Fünfter Teil: Obligationenrecht; OR)*; EN: *Federal Act on the Amendment of the Swiss Civil Code Section (Part Five: The Code of Obligations)*. 220. Fedlex. Available online: <https://www.admin.ch/opc/en/classified-compilation/19110009/201704010000/220.pdf> (accessed on 1 September 2019).
- Swiss Confederation—ISA. 2006. DE: *Bundesgesetz Betreffend die Aufsicht über Versicherungsunternehmen (Versicherungsaufsichtsgesetz, VAG)*; EN: *Federal Law on the Supervision of Insurance Undertakings (Insurance Supervision Act, ISA)*. 961.0. Fedlex. Available online: <https://www.fedlex.admin.ch/eli/cc/2005/734/de> (accessed on 1 September 2019)
- Swiss Confederation—ISO. 2006. DE: *Verordnung über die Beaufsichtigung von privaten Versicherungsunternehmen (Aufsichtsverordnung, AVO)*; EN: *Insurance Supervision Ordinance (ISO)*. 961.011. Fedlex. Available online: <https://www.fedlex.admin.ch/eli/cc/2005/735/de> (accessed on 1 September 2019).

- Swiss Confederation. 2018a. Distribution of Earnings and Carryover of Loss [Website]. Available online: <https://www.kmu.admin.ch/kmu/en/home/concrete-know-how/finances/accounting-and-auditing/annual-financial-statements/distribution-earnings-carryover-loss.html> (accessed on 18 October 2019).
- Swiss Confederation. 2018b. Taxation of Share Capital Companies [Website]. Available online: <https://www.kmu.admin.ch/kmu/en/home/concrete-know-how/finances/taxes/taxation-share-capital-companies.html> (accessed on 9 October 2019).
- Swiss Confederation, Europäischen Wirtschaftsgemeinschaft. 1993. *Abkommen Zwischen der Schweizerischen Eidgenossenschaft und der Europäischen Wirtschaftsgemeinschaft Betreffend die Direktversicherung mit Ausnahme der Lebensversicherung*. Fedlex. Available online: [https://www.fedlex.admin.ch/eli/cc/1992/1894\\_1894\\_1894/de](https://www.fedlex.admin.ch/eli/cc/1992/1894_1894_1894/de) (accessed on 3 September 2019).
- Swiss Financial Market Supervisory Authority—FINMA. 2008a. *DE: Rundschreiben 2008/42 Versicherungstechnische Rückstellungen in der Schadenversicherung; EN: Circular 2008/42 Technical Provisions for Non-Life Insurer*. Available online: <https://www.finma.ch/de/~media/finma/dokumente/dokumentencenter/myfinma/rundschreiben/finma-rs-2008-42.pdf?la=de> (accessed on 1 September 2019).
- Swiss Financial Market Supervisory Authority—FINMA. 2008b. FINMA's Key Metrics for Insurers [Website]. Available online: <https://www.finma.ch/en/documentation/finma-publications/kennzahlen-und-statistiken/kennzahlen/kennzahlen-versicherer/> (accessed on 1 September 2021).
- Swiss Financial Market Supervisory Authority—FINMA. 2009. Die Privaten Versicherungsunternehmen in der Schweiz im Jahr 2008. Available online: [https://www.finma.ch/de/~media/finma/dokumente/dokumentencenter/myfinma/finma-publikationen/versicherungsbericht/versicherungsreport-2008-d\\_einleitung.pdf?la=de](https://www.finma.ch/de/~media/finma/dokumente/dokumentencenter/myfinma/finma-publikationen/versicherungsbericht/versicherungsreport-2008-d_einleitung.pdf?la=de) (accessed on 29 September 2021).
- Swiss Financial Market Supervisory Authority—FINMA. 2011. *DE: Rundschreiben 2011/3 Versicherungstechnische Rückstellungen in der Rückversicherung; EN: Circular 2011/3 Technical Provisions for Reinsurer*. Available online: <https://www.finma.ch/de/~media/finma/dokumente/dokumentencenter/myfinma/rundschreiben/finma-rs-2011-03.pdf?la=de> (accessed on 1 September 2019).
- Swiss Financial Market Supervisory Authority—FINMA. 2015. *Versicherungsaufsichtsverordnung-FINMA, AVO-FINMA. 961.011.1*. Fedlex. Available online: <https://fedlex.data.admin.ch/filestore/fedlex.data.admin.ch/eli/cc/2005/736/20151215/de/pdf-a/fedlex-data-admin-ch-eli-cc-2005-736-20151215-de-pdf-a.pdf> (accessed on 3 October 2019).
- Swiss Financial Market Supervisory Authority—FINMA. 2017. *FINMA Circular 2017/3 Swiss Solvency Test (SST)*. Available online: <https://www.finma.ch/en/~media/finma/dokumente/dokumentencenter/myfinma/rundschreiben/finma-rs-2017-03-20201104.pdf?la=en> (accessed on 7 October 2019).
- Swiss Financial Market Supervisory Authority—FINMA. 2018a. SST 2018 Survey. Available online: <https://www.finma.ch/en/~media/finma/dokumente/dokumentencenter/myfinma/2ueberwachung/sst/sst-survey-2018.pdf?la=en> (accessed on 7 October 2019).
- Swiss Financial Market Supervisory Authority—FINMA. 2018b. The Swiss Solvency Test [Factsheet]. Available online: <https://www.finma.ch/~media/finma/dokumente/dokumentencenter/myfinma/faktenblaetter/faktenblatt-schweizer-solvenztest-sst.pdf> (accessed on 7 October 2019).
- Swiss Financial Market Supervisory Authority—FINMA. 2019a. An Overview of the Insurance Sector [Website]. Available online: <https://www.finma.ch/en/documentation/finma-publications/reports/insurance-reports/> (accessed on 1 September 2019).
- Swiss Financial Market Supervisory Authority—FINMA. 2019b. FINMA's Core Tasks [Website]. Available online: <https://www.finma.ch/en/finma/supervisory-objectives/tasks/> (accessed on 12 October 2021).
- Swiss Financial Market Supervisory Authority—FINMA. 2019c. Periodic Collection of Data [Website]. Available online: <https://www.finma.ch/en/supervision/insurers/cross-sectoral-tools/periodic-collection-of-data/> (accessed on 1 September 2019).
- Wood, Simon. 2010. mgcv: GAMs in R [Website]. Available online: <https://www.maths.ed.ac.uk/~swood34/mgcv/> (accessed on 24 September 2021).
- Wüthrich, Mario Valentin. 2016. *Market-Consistent Actuarial Valuation*. Cham: Springer.
- Wüthrich, Mario Valentin, and Michael Merz. 2008. *Stochastic Claims Reserving Methods in Insurance*. Chichester: Wiley Finance.



Article

# Unit-Linked Tontine: Utility-Based Design, Pricing and Performance <sup>†</sup>

An Chen <sup>1</sup>, Thai Nguyen <sup>2,\*</sup> and Thorsten Sehner <sup>1</sup> 

<sup>1</sup> Institute of Insurance Science, University of Ulm, Helmholtzstrasse 20, 89069 Ulm, Germany; an.chen@uni-ulm.de (A.C.); thorsten.sehner@uni-ulm.de (T.S.)

<sup>2</sup> École d'Actuariat, Université Laval, 2425, rue de l'Agriculture, Quebec, QC G1V 0A6, Canada

\* Correspondence: thai.nguyen@act.ulaval.ca

<sup>†</sup> An earlier version of this article is a part of the dissertation of Thorsten Sehner which is published on the Open Access Repository of Ulm University and Ulm University of Applied Sciences.

**Abstract:** Due to the low demand for conventional annuities, alternative retirement products are sought. Quite recently, tontines have been frequently brought up as a promising option in this respect. Inspired by unit-linked life insurance and retirement products, we introduce unit-linked tontines in this article, where the tontine payoffs are directly linked to the development of the underlying financial market. More specifically, we consider two different tontine payoff structures differing in the (non-)inclusion of guaranteed payments. We first price the unit-linked tontines by using the risk-neutral pricing approach. Consequently, we study the attractiveness of these products for a utility-maximizing policyholder and compare them with non-unit-linked tontines. Our numerical analysis sheds light on the design challenges and gives explanations why similar products might not be widely adopted already.

**Keywords:** unit-linked tontine; product design; risk neutral pricing; utility optimization; utility performance

**JEL Classification:** G13; G22

**Citation:** Chen, An, Thai Nguyen, and Thorsten Sehner. 2022.

Unit-Linked Tontine: Utility-Based Design, Pricing and Performance.

*Risks* 10: 78. <https://doi.org/10.3390/risks10040078>

Academic Editors: Ermanno Pitacco and Annamaria Olivieri

Received: 27 January 2022

Accepted: 29 March 2022

Published: 7 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Unit-linked insurance policies belong to the most frequently concluded contracts in the life insurance sector; for example, more than 50% of the UK life (re)insurance gross written premiums were attributed to the index- and unit-linked insurance field in 2019 according to Statista (2020b). Among other attractive features, higher return expectations, flexibility, design possibilities and tax advantages (see, e.g., Schiereck et al. 2020) certainly play a driving role in the attractiveness of these policies. Interesting subject areas related to unit-linked insurance contracts, such as variable annuities, include pricing and valuation from the insurer's or the customers' perspective (see, e.g., Aase and Persson 1994; Ekern and Persson 1996; Gatzert et al. 2011), hedging strategies (see, e.g., Møller 1998), impact of stochastic interest rates (see, e.g., Schrage and Pelsser 2004) or guarantee components (see, e.g., Ledlie et al. 2008). In this paper, inspired by variable annuities, we design and investigate a new type of tontine that is directly linked to the developments in the financial market.

Yet, why is it even reasonable to consider tontines when dealing with old-age provision? From a theoretical point of view, actuarially fairly priced annuities should actually be regarded favorably by rational customers (see, e.g., Peijnenburg et al. 2016; Yaari 1965). However, annuitization rates are rather low in reality (see, e.g., Hu and Scott 2007). This adverse phenomenon known as the annuity puzzle (see, e.g., Ramsay and Oguledo 2018) is hitting conventional annuities. Moreover, due to low interest rate environments and tightening solvency regulations, it is hard to expect that annuitization rates will go up any

time soon. Therefore, alternative retirement products are naturally searched by insurers and customers, which brings up tontines as an option. Due to the backdrop of the demographic change (see, e.g., Margaras 2019), the so-called tontine retirement investment has become more and more important (see, e.g., Milevsky and Salisbury 2015; Sabin 2010). A main characteristic of tontines is that, in contrast to annuities, longevity risk is borne, to a great extent, by the pooled policyholders themselves. Hence, tontines are normally cheaper and, thus, potentially more attractive. Further discussions on practicalities, qualitative regulatory, technological and risk management issues associated with a tontine product can be found in Milevsky et al. (2018); Winter and Planchet (2021).

Let us briefly mention some of the recent literature that has addressed relevant topics related to tontine products. A general and historical view on tontines, as well as their possible applications for retirement income planning, is given in Milevsky (2015). Specific forms for the tontine payout structure are discussed in, e.g., Milevsky and Salisbury (2015). The question regarding how the tontine principle can be used to create tontine pensions for employees is studied in Forman and Sabin (2015). In Gemmo et al. (2020), investment possibilities in both tontines and traditional financial assets are investigated. Fairness issues when considering heterogeneous cohorts are considered in, e.g., Chen et al. (2020); Denuit (2019); Donnelly et al. (2014); Milevsky and Salisbury (2016); Sabin (2010). Bernhardt and Donnelly (2019) study the inclusion of bequest motives in tontine products. Recently, research on reasonable ways to combine tontines and annuities has been more extensively explored, see, e.g., Chen and Rach (2019); Chen et al. (2019, 2020); Weinert and Gründl (2020). However, to the best of our knowledge, the idea to consider a tontine as a unit-linked product has not yet been considered in detail in the literature.

In this article, inspired by unit-linked life and retirement insurance products, we introduce unit-linked tontines (see Sehner (2021)). We analyze the pricing and attractiveness of such products where two concrete unit-linked tontine payoffs are considered. We base our product model on the tontine concept applied in, e.g., Milevsky and Salisbury (2015), where the deterministic payout function is replaced by a stochastic payout process that depends on the developments in the financial market. In the specific setting, one tontine payoff is designed to coincide with the pure value of a portfolio following a certain investment strategy in the financial market, while the other one includes guaranteed payments, such that the policyholders participate in high portfolio values, but are secured in bad market scenarios. We rely on the risk-neutral pricing approach to determine the premiums required to buy the corresponding unit-linked tontines. In order to highlight the potential of our unit-linked tontine variant, we conduct an expected utility analysis that is commonly used in such a context (see, e.g., Mitchell 2002; Yaari 1965). More specifically, we first search for the optimal investment strategy that maximizes the expected utility of the policyholder for a given unit-linked tontine variant. We then numerically compare the maximum expected utilities of the two variants. Our comparison also takes two traditional tontine alternatives without unit-linked payments into account, namely the optimal and the natural traditional tontine.

The main observations and results, which can be drawn from our numerical analysis, are as follows: The unit-linked tontine may perform better than the traditional tontine alternatives if the following circumstances are present: First, the initial number of pooled individuals is either very low or high. Second, the expected return of the tradable risky asset is high or its volatility is low, which leads to a higher market price of risk, working naturally in favor of the unit-linked tontine. Third, the policyholder's risk aversion or subjective discount rate is low. The additional financial risk component in the unit-linked tontine and the steady increase of the expected payment of the unit-linked tontine over time are respectively responsible for this. For our baseline parametrization, the certainty equivalent induced by the variant, whose payout process is defined by the pure portfolio value, is, for instance, about 8% higher than the one belonging to the optimal traditional tontine and about 11% higher than the one belonging to the natural traditional tontine. As the unit-linked tontine can be more successful among customers than the traditional

counterpart, it seems reasonable to further study it. We further observe that, if the pure portfolio value stipulates its payout process, the unit-linked tontine may yield a higher utility level than in the case where it includes guaranteed payments. For our baseline parametrization, the corresponding certainty equivalent is, for instance, about 27% higher. Nevertheless, the latter case might be attractive, especially to customers who consider additional guarantee components important. In particular, its performance approaches that of the superior variant if the expected return of the risky asset decreases or if the volatility of the risky asset or the policyholder's risk aversion increases.

The remainder of this article is organized as follows: Section 2 introduces the model setting including the general nature of the unit-linked tontine product and the underlying financial and mortality risks. In Section 3, we derive the pricing formulas not only for the general payment structure, but also for both concrete variants of the unit-linked tontine. In Section 4, we discuss the solution of the utility optimization problem for our two particular unit-linked payment designs. In Section 5, we conduct the numerical study and present its outputs. Section 6 concludes the article. Some additional mathematical derivations can be found in the Appendices A–D.

## 2. Model Setting

### 2.1. Unit-Linked Tontine Product

In order to model the unit-linked tontine product, we employ the tontine concept presented in, e.g., Milevsky and Salisbury (2015), and modify it according to our purposes. Therefore, the idea behind the tontine type established in Milevsky and Salisbury (2015), to which we also refer as the traditional tontine, is shortly reviewed here first. Initially, i.e., at time 0, the buyer of such a tontine pays a single premium to the providing life insurance company. After the insurer has issued tontines to  $n \in \mathbb{N}$  individuals at time 0, they are grouped together into a pool. For simplicity, it is assumed that these  $n$  individuals, who are also referred to as policyholders or participants, are homogeneous, i.e., they are all of the same age  $x \geq 0$  at time 0 and of the same gender (which implies that they all have the same mortality rate). As time goes by, the insurance company disburses contractually predetermined payments to living participants. Specifically, a living individual holding one of the traditional tontine contracts receives at time  $t \geq 0$ , in the first place, a specific amount of money determined by the so-called tontine payout function denoted by  $d_t$ , which is deterministic and initially stipulated. What is more, contingent on being alive at time  $t$ , there is the possibility that she obtains more than  $d_t$  due to the fact that the theoretical payments to the dead participants, if existent, are distributed among the survivors in the pool. Owing to the homogeneity between the participants, this extra payment is given by  $\frac{(n-N_t)d_t}{N_t}$ , where the random variable  $N_t$  denotes the stochastic number of participants alive at time  $t$ . Overall, we can summarize the total payment that is disbursed to the considered traditional tontine holder at time  $t$ , given that she is alive, in the following expression:

$$\left( \frac{(n - N_t)d_t}{N_t} + d_t \right) \mathbb{1}_{\{\zeta_x > t\}} = \frac{nd_t}{N_t} \mathbb{1}_{\{\zeta_x > t\}}, \quad (1)$$

where the random variable  $\zeta_x$  represents the stochastic remaining lifetime of the individual aged  $x$  at time 0. As there are no death benefits, it is clear that the policyholder's payments proportionally increase if more individuals in the pool pass away. Note that throughout the following sections, we always assume that the payments of the insurer to a tontine holder are continuously disbursed.

When considering the unit-linked tontine product, we focus on payments stemming from the purchase of this tontine that are explicitly linked to the financial market. In this way, the participants directly partake in the developments in the financial market. Our corresponding product model is adopted, to a great extent, from the traditional tontines described above. The only difference is that the deterministic tontine payout function  $d_t$  is replaced by the so-called tontine payout process denoted by the stochastic process  $\Psi_t$ . This process depends on the performance of the financial market and, hence, makes the tontine

a unit-linked product. Apart from that, the role of  $\Psi_t$  stays the same as the one of  $d_t$ . Note that in this article, we study two specified variants for  $\Psi_t$  that are introduced in Section 3.2. On the whole, similar to (1), the total payment being disbursed to a unit-linked tontine holder at time  $t$  and described by the stochastic process  $D_t$  is given by

$$D_t = \frac{n\Psi_t}{N_t} \mathbb{1}_{\{\zeta_x > t\}}. \tag{2}$$

### 2.2. Financial Market and Mortality Risk

For the examination of the unit-linked tontine introduced in Section 2.1, we need to model the financial market. Hereinafter, we always consider the financial market in continuous time that consists of one risky and one risk-free asset. We assume that there are no transaction costs or liquidity risk when trading the assets in the market. Following the well-known Black–Scholes model (see Black and Scholes 1973), the stochastic value of the risky asset at time  $t$ , denoted by  $S_t$ , is described by the following geometric Brownian motion:<sup>1</sup>

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S_0 > 0, \tag{3}$$

where  $W$  is a standard Brownian motion. The dynamics of the risk-free asset is given by

$$dB_t = rB_t dt, \quad B_0 = 1, \tag{4}$$

where  $r$  is the risk-free interest rate. The three parameters  $\mu, \sigma$  and  $r$  are constant over time in our setting and  $\mu > r$  is assumed. Note that possible dividend payments existing in the described financial market are neglected in our framework.

Let  $V_t$  be the value of a portfolio at time  $t$  that is generated by the investments of the insurer in the financial market. We assume that the fraction of the portfolio invested in the risky asset at time  $t$  is described by the *deterministic* trading strategy  $\pi_t \in [0, 1]$ . This means that neither short selling of the risky portfolio nor leverage is allowed. The remaining fraction  $(1 - \pi_t)$  is invested in the risk-free asset. By the self-financing property, the dynamics of  $V_t$  under  $\mathbb{P}$  is given by

$$dV_t = \pi_t \frac{V_t}{S_t} dS_t + (1 - \pi_t) \frac{V_t}{B_t} dB_t = (r + \pi_t(\mu - r))V_t dt + \sigma \pi_t V_t dW_t, \quad V_0 > 0. \tag{5}$$

It can be shown that the explicit solution of the stochastic differential Equation (5) is given by

$$V_t = V_0 e^{rt + (\mu - r) \int_0^t \pi_s ds - \frac{\sigma^2}{2} \int_0^t \pi_s^2 ds + \sigma \int_0^t \pi_s dW_s}. \tag{6}$$

Besides the financial risk, mortality risk is also contained in the unit-linked tontine. It stems from two sources, namely the unsystematic mortality risk and the systematic mortality risk (see, e.g., Dahl et al. 2008). The unsystematic mortality risk arises from the randomness of deaths in the pool with a known mortality law. This risk is diversifiable, i.e., it disperses if the size of the pool grows. In contrast, the systematic mortality risk is not diversifiable, even if the pool size is large, as it results from overarching changes in the underlying mortality intensity. For the traditional tontines (with mortality risk exclusively) and an infinite pool size, all the mortality risk is shared by the policyholders. With a finite pool size, the insurer only has the risk generated by the death time of the last survivor, at which the insurer stops its payment. Additionally, in unit-linked tontines, there is financial market risk. Depending on the risk management strategies the insurer chooses, the insurer might still retain some financial market risk.

To model the mortality risk, we use the following framework: The probability (under  $\mathbb{P}$ ) that the considered individual survives the next  $t$  years from time 0 on, at which she is  $x$  years old, is denoted by  ${}_t p_x \in (0, 1]$ . To include the above-mentioned systematic mortality

risk component, we, similar to, e.g., Lin and Cox (2005), allow for a mortality shock that is represented by a random variable denoted by  $\epsilon$ . We assume that  $\epsilon$  has a density function denoted by  $f_\epsilon$  and that its moment-generating function denoted by  $M_\epsilon$  exists. The shocked survival curve is then given by  ${}_t p_x^{1-\epsilon}$ . We set the range of  $\epsilon$  to  $(-\infty, 1)$ , so that  ${}_t p_x^{1-\epsilon} \in (0, 1]$  is preserved. If no mortality shock is existent, simply let  $\epsilon = 0$  a.s. We remark that the latest insurance solvency regulations require insurers to test their balance sheets against various stress-test scenarios. For instance, in the Canadian solvency regulation, a 10–20% decrease of mortality rates (depending on the type of annuity is assumed for a longevity shock). The U.S. regulation assumes a stress on mortality improvement between 16–40% (depending on the age). This results in lower mortality rates between 0.7–6%. In Solvency II, which is implemented for insurance undertakings in the EU, a longevity shock is defined as a decrease of annual death probabilities by 20%. The simple model we have chosen reflects the spirit of these realistic regulation frameworks.

For the random variable  $N_t$ , which is affected by mortality risk, we can obtain the following distribution under  $\mathbb{P}$  when conditioning on the survival of the considered policyholder and on  $\epsilon$ :

$$(N_t - 1 | \zeta_x > t, \epsilon) \stackrel{\mathbb{P}}{\sim} \text{Bin}(n - 1, {}_t p_x^{1-\epsilon}), \quad (7)$$

where we use the assumption that the lifetimes of the participants are stochastically independent under  $\mathbb{P}$ .

Following the main stream of unit-linked insurance products (e.g., Aase and Persson 1994; Bacinello et al. 2018; Bernhardt and Donnelly 2019; Briys and de Varenne 1994), we suppose that  $W$ , constituting the financial risk, is stochastically independent of  $(\zeta_x, N, \epsilon)$  under  $\mathbb{P}$ . Note that this requirement does usually not pose a restriction as the development of the value of the risky asset and the chances of survival do generally not interact. We remark that the independence assumption of actuarial and financial risk in the real world may be quite reasonable in many situations. Recent research however finds that shocks in stock market wealth might have an impact on mortality. For example, Giulietti et al. (2020) provide evidence that daily fluctuations in the stock market have important effects on fatal car accidents. Schwandt (2018) demonstrates that stock wealth shocks that lead to losses in the wealth of stock-holding retirees affect the health of retirees in the US. In our paper, the independence assumption of these risks allows us to analyze the pricing problem and individual welfare of the unit-linked tontine in a semi-explicit way.

Let  $\mathcal{G} = \{\mathcal{G}_t\}_{t \geq 0}$  be the filtration generated by the Brownian motion  $W$  and denote the natural filtration with respect to  $\zeta_x$ ,  $N$  and  $\epsilon$  by  $\mathbb{H} = \{\mathcal{H}_t\}_{t \geq 0}$ . The resulting progressively enlarged filtration is given by  $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ , whose element  $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t$  contains all relevant information revealed until time  $t$ .

### 3. Pricing

In this section, we aim at pricing the unit-linked tontine product established in Section 2.1, i.e., we determine the single initial premium denoted by  $P_0 > 0$  that needs to be paid by a policyholder to the insurance company. As we employ the standard risk-neutral pricing approach to find  $P_0$ , we have to clarify how a risk-neutral probability measure denoted by  $\mathbb{Q}$  is chosen when mortality risk is also taken into account. First, it is clear that, due to the dependence of  $D$  on the survival of the policyholder and the other participants, the market, in which the unit-linked tontine is traded, is incomplete. Thus, a risk-neutral probability measure is not unique and, hence, there is, in general, also no unique price  $P_0$ . For a concrete choice of  $\mathbb{Q}$ , we assume that the insurer considers the financial risk and the mortality risk separately when determining  $\mathbb{Q}$ , whereby the stochastic independence of these two risk categories is also supposed under  $\mathbb{Q}$ . Further discussions about the independence property between financial and actuarial risks in the  $\mathbb{P}$ - and the  $\mathbb{Q}$ -worlds can be found in, e.g., Dhaene et al. (2013), where the authors investigate



the conditions under which it is possible (or not) to transfer the independence assumption from the physical measure  $\mathbb{P}$  to the risk-neutral pricing measure  $\mathbb{Q}$ .

Regarding the financial risk that is captured by the filtration  $\mathbb{G}$ , we expect the insurer to use the risk-neutral probability measure, which, if we restrict ourselves to  $\mathbb{G}$ , exists and is unique due to the completeness of the financial market described in Section 2.2. Note that the explicit solution for  $V_t$ , which is under  $\mathbb{P}$  given in (6), changes accordingly under  $\mathbb{Q}$  to

$$V_t = V_0 e^{rt - \frac{\sigma^2}{2} \int_0^t \pi_s^2 ds + \sigma \int_0^t \pi_s dW_s^{\mathbb{Q}}}, \tag{8}$$

where  $(W_t^{\mathbb{Q}})_{t \geq 0}$  is a standard Brownian motion under  $\mathbb{Q}$ .

Following Chen and Rach (2019), we assume that the choice of  $\mathbb{Q}$  on  $\mathbb{H}$  for pricing purposes depends on the nature of the overall insurance business of the life insurance company. If a large product range is offered, there may already be some natural hedges between the products and, thus, the insurer would be faced with less mortality risk than in the case in which it solely concentrates on one specific product field. We assume that the insurer only trades tontine products and that, also due to the resulting higher mortality risk exposure, the insurer is prudent when charging premiums, i.e., safety loadings are to be included in some way. If  ${}_t\tilde{p}_x \in (0, 1]$  denotes the survival probability under  $\mathbb{Q}$ , and since tontines belong to the retirement product type, a possibility to reflect the insurer’s pricing prudence is to require that

$${}_t\tilde{p}_x \geq {}_t p_x, \tag{9}$$

and that the mortality shock  $\varepsilon$  follows the same distribution under  $\mathbb{Q}$  as under  $\mathbb{P}$ . Given these requirements, to which we stick in the following, the (shocked) survival curve under  $\mathbb{Q}$  runs at a higher level than the one under  $\mathbb{P}$ , which leads to the inclusion of implicit safety loadings in premiums. If the insurer increases  ${}_t\tilde{p}_x$ , the company is more conservative about pricing. The choice of the magnitude of  ${}_t\tilde{p}_x$  usually depends on the pool size  $n$  since, as already pointed out in Section 2.2, the unsystematic mortality risk becomes less relevant if  $n$  grows. Therefore,  ${}_t\tilde{p}_x$  normally attains a rather low value if the pool size is large. As it is determined that changing the probability measure from  $\mathbb{P}$  to  $\mathbb{Q}$  does not have an impact on the distribution type of the random variable  $N_t$ , we simply replace  $\mathbb{P}$  by  $\mathbb{Q}$  and  ${}_t p_x$  by  ${}_t\tilde{p}_x$  in (7) when specifying the distribution of  $N_t$  under  $\mathbb{Q}$ . The stochastic independence of the remaining lifetimes of the participants is preserved under  $\mathbb{Q}$  accordingly.

Having clarified the risk-neutral probability measure  $\mathbb{Q}$ , we discuss the pricing of the unit-linked tontine in the following. We will start with a general tontine payout process  $\Psi_t$  and then continue by examining specified alternatives for it. We always assume that the rates of convergence of  ${}_t p_x$  and  ${}_t\tilde{p}_x$  towards 0 if  $t$  goes to infinity exceed the rates of convergence or divergence of the other time-dependent quantities in order to guarantee that all improper integrals with respect to  $t$  necessary throughout the subsequent sections exist.<sup>2</sup>

### 3.1. General Payment Structure

First, let  $\Psi_t$  be a general tontine payout process. Next, the single initial premium  $P_0$  can be calculated via the risk-neutral pricing approach as

$$\begin{aligned} P_0 &= E_{\mathbb{Q}} \left[ \int_0^{\infty} e^{-rt} D_t dt \middle| \mathcal{F}_0 \right] = E_{\mathbb{Q}} \left[ \int_0^{\infty} e^{-rt} \frac{n \Psi_t}{N_t} \mathbb{1}_{\{\zeta_x > t\}} dt \right] \\ &= n \int_0^{\infty} e^{-rt} E_{\mathbb{Q}}[\Psi_t] E_{\mathbb{Q}} \left[ \frac{\mathbb{1}_{\{\zeta_x > t\}}}{N_t} \right] dt, \end{aligned} \tag{10}$$

where the stochastic independence between  $W^{\mathbb{Q}}$  and  $(\zeta_x, N, \epsilon)$  is applied in the last step. The latter expected value in (10) is given by

$$E_{\mathbb{Q}} \left[ \frac{\mathbb{1}_{\{\zeta_x > t\}}}{N_t} \right] = \frac{1}{n} I_t, \tag{11}$$

where

$$I_t = \int_{-\infty}^1 \left( 1 - \left( 1 - {}_t\tilde{p}_x^{1-z} \right)^n \right) f_{\epsilon}(z) dz. \tag{12}$$

The detailed derivation of (11) is reported in Appendix A. Consequently, we obtain the following general pricing formula:

$$P_0 = \int_0^{\infty} e^{-rt} I_t E_{\mathbb{Q}}[\Psi_t] dt. \tag{13}$$

### 3.2. Specified Payment Structures

In the following, we consider two specified variants for the tontine payout process  $\Psi_t$ , which can be interesting to examine and may have potential for tontine product design. We determine the single premiums that need to be contributed by the individual if she wants to buy the corresponding unit-linked tontine.

As our focus is on payments with a direct linkage to the financial market, i.e., to the developments of the risky and of the risk-free asset, hereafter, we assume that for payout purposes, the insurer creates a tontine payment account  $\Psi$  whose value can be amounted to the portfolio given in (8). We assess the following cases on how to potentially define  $\Psi_t$ :

- (A) Let us first consider the case where the tontine payout process is equal to the portfolio value  $V$  explicitly given in (8), i.e.

$$\Psi_t = V_t. \tag{14}$$

This means that the tontine payout process at time  $t$  simply complies with a money stock amounting to  $V_t$ . To generate this amount, the insurance company can invest in the risky and the risk-free asset according to the trading strategy applied in the corresponding portfolio. By the choice given in (14), the full potential of the financial market will be passed on to the customers within a tontine framework. By (2), the total tontine payment to the policyholder at time  $t$  in this case is given by  $\frac{nV_t}{N_t} \mathbb{1}_{\{\zeta_x > t\}}$ .

- (B) Second, inspired by participating life insurance policies with guaranteed payments (see, e.g., Briys and de Varenne 1994), we stipulate

$$\Psi_t = G_t + \alpha(V_t - G_t)^+, \tag{15}$$

where  $G_t > 0$  denotes the guaranteed payment at time  $t$  and  $\alpha \in (0, 1]$  is the constant participation rate, and where  $(V_t - G_t)^+ = \max\{V_t - G_t, 0\}$ . Thus, the tontine payout process coincides here with a predetermined payment function represented by  $G_t$  as long as the financial market performs poorly, i.e.,  $V_t$  is low, so that  $V_t \leq G_t$  holds. On the contrary, if the financial market performs well, i.e.,  $V_t$  is high, so that  $V_t > G_t$ , an additional participation in the positive difference  $V_t - G_t$  at the rate  $\alpha$  is included. Employing the choice given in (15) can satisfy customers, who appreciate additional guarantee components smoothing uncertain payout structures. By (2), the total tontine payment to the policyholder at time  $t$  in this case is given by  $\frac{n(G_t + \alpha(V_t - G_t)^+)}{N_t} \mathbb{1}_{\{\zeta_x > t\}}$ .

The tontine pricing in Cases A and B is summarized in the following two propositions:

**Proposition 1** (Case A). *If  $\Psi_t$  is defined as in (14), the single initial premium of the resulting version of the unit-linked tontine product is given by*

$$P_0 = V_0 \int_0^\infty I_t dt. \tag{16}$$

**Proof.** With the aid of the general pricing formula given in (13) and by using

$$E_{\mathbb{Q}}[\Psi_t] = E_{\mathbb{Q}}[V_t] = V_0 e^{rt}$$

due to the fact that the discounted portfolio value process is a  $\mathbb{Q}$ -martingale, we obtain (16).  $\square$

**Proposition 2** (Case B). *If  $\Psi_t$  is defined as in (15), the single initial premium of the resulting version of the unit-linked tontine product is given by*

$$P_0 = \int_0^\infty e^{-rt} I_t \left( G_t + \alpha \left( V_0 e^{rt} \Phi(\tilde{d}_t) - G_t \Phi(\hat{d}_t) \right) \right) dt, \tag{17}$$

where  $\Phi$  is the distribution function of the standard normal distribution and the functions  $\tilde{d}_t$  and  $\hat{d}_t$  are given by

$$\tilde{d}_t = \frac{\ln\left(\frac{V_0}{G_t}\right) + rt + \frac{\sigma^2}{2} \int_0^t \pi_s^2 ds}{\sigma \sqrt{\int_0^t \pi_s^2 ds}} \quad \text{and} \quad \hat{d}_t = \tilde{d}_t - \sigma \sqrt{\int_0^t \pi_s^2 ds}. \tag{18}$$

**Proof.** The proof of Proposition 2 is reported in Appendix B.1.  $\square$

**Remark 1.** *From the insurer’s perspective, managing such unit-linked products would require the insurer to pay transaction costs that are linked to hedging activities against fluctuations of the risky asset in the financial market and of the mortality development. For instance, in Case B, if we ignore the mortality risk, the insurer has to hedge against selling a guaranteed amount plus the call option, which by put-call parity is equivalent to selling the portfolio value plus the put option. In bad market scenarios when the risk asset price goes down, more hedging activities would be needed; hence, it is true that the relative transaction price will be higher if the tail is longer. A thorough analysis that includes transaction costs is interesting and left for future research. In the real-world implementation, these transaction costs do impact the product design. We remark that in the presence of transaction costs, hedging and pricing are no longer valid in the classical Black and Scholes model. In such contexts, Leland’s increasing volatility method, as per Leland (1985), would be helpful for compensating transaction costs and an approximately complete replication can be expected by using the delta strategy calculated from a modified Black–Scholes equation with an appropriate modified volatility. This prescription is based on the idea that the presence of transaction costs implies an extra fee, which is necessary for the option seller in the replication problem, i.e., options become more expensive in the presence of transaction costs.*

#### 4. Utility Optimization

In the following, we conduct a utility maximization analysis to find out which of the two unit-linked tontine variants suggested in Section 3.2 is more preferable to an individual investor. To this end, for a given unit-linked tontine variant, we search for the optimal investment strategy that maximizes the discounted expected utility of the policyholder in this section. We numerically compare the utility optimums of the two variants with each other and with those of the traditional tontine alternatives without unit-linked payments in Section 5.

Subsequently, we always assume that the policyholder’s utility function, denoted by  $u$ , is of constant relative risk aversion:

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}, \tag{19}$$

where  $c > 0$  represents the consumable input and  $\gamma > 0$  adhering to  $\gamma \neq 1$  is the measure of the policyholder’s relative risk aversion. This choice is one of the most frequently used utility functions to capture the preferences of individuals (see, e.g., Levy 1994; Sharpe 2017).<sup>3</sup> In the design problems below, we assume that expectations are not subjective.

#### 4.1. General Payment Structure

For a general tontine payout process  $\Psi_t$ , the objective of the optimization problem, i.e., the discounted expected utility, can be formulated and transformed as follows:

$$E_{\mathbb{P}} \left[ \int_0^{\infty} e^{-\rho t} u \left( \frac{n\Psi_t}{N_t} \right) \mathbb{1}_{\{\zeta_x > t\}} dt \right] = \frac{n^{1-\gamma}}{1-\gamma} \int_0^{\infty} e^{-\rho t} \kappa_t E_{\mathbb{P}} [\Psi_t^{1-\gamma}] dt, \tag{20}$$

where  $\rho$  is the constant subjective discount rate of the individual and

$$\begin{aligned} \kappa_t &= E_{\mathbb{P}} \left[ \frac{\mathbb{1}_{\{\zeta_x > t\}}}{N_t^{1-\gamma}} \right] \\ &= \sum_{k=0}^{n-1} \frac{1}{(k+1)^{1-\gamma}} \binom{n-1}{k} \int_{-\infty}^1 \left( {}_t p_x^{1-z} \right)^{k+1} \left( 1 - {}_t p_x^{1-z} \right)^{n-1-k} f_{\epsilon}(z) dz. \end{aligned} \tag{21}$$

The formulation of the discounted expected utility in (20) arises from translating the formula in (10) into the utility framework, while its transformation results from applying the power utility function given in (19) and similar calculation techniques as before. Since the individual has to provide a single initial premium out of her available initial wealth, denoted by  $v > 0$ , to buy the tontine product, the pricing formula found in Section 3, where the general version is given in (13) and the specified ones in (16) and (17), naturally forms the budget constraint in the optimization problem. The decision variables in the optimization problem are typically appropriate quantities occurring in the tontine payout process  $\Psi_t$ . This means that we eventually search for the optimal specific form of  $\Psi_t$ , which determines the tontine disbursements in such a way that the policyholder is endowed with the highest utility level possible. The general representative maximization problem overall is given by:

#### Problem 1.

$$\begin{aligned} &\max_{(\Psi_t)_{t \geq 0}} \frac{n^{1-\gamma}}{1-\gamma} \int_0^{\infty} e^{-\rho t} \kappa_t E_{\mathbb{P}} [\Psi_t^{1-\gamma}] dt \\ &s.t. \ v = P_0 = \int_0^{\infty} e^{-rt} I_t E_{\mathbb{Q}} [\Psi_t] dt. \end{aligned}$$

Note that, strictly speaking, we shall put  $v \geq P_0$  in the budget constraint. However, as is typically done in this kind of optimization problem, the budget constraint is binding in the optimal solution due to the steadily positive slope of  $u$ , such that we start immediately with equality in the constraint.

#### 4.2. Specified Payment Structures

Now, we consider the particular unit-linked payment designs from Section 3.2 specifying the tontine payout process  $\Psi_t$  in two different ways and modify Problem 1 accordingly. The emerging optimization problems are then, if possible, solved analytically.

Concerning the fractions invested in the risky and the risk-free asset, we henceforth assume that they stay constant over time and are non-negative and bounded from above by 1, i.e.,  $\pi_t = \pi \in [0, 1]$  for all  $t$ . Note that these assumptions do not actually pose a strict restriction: By their invariability, the fractions can also be regarded as the perpetual average percentages which determine the long-term mean composition of the portfolio. By generally forbidding short selling, we account for the fact that bans on short selling (can) indeed exist, as in the case in Europe in March 2020 during the coronavirus pandemic showed (see, e.g., Smith 2020). Applying a constant  $\pi$  simplifies the equations in (5), (6), (8) and (18), accordingly.

**Case A:** Recall that we assume for Case A that  $\Psi_t = V_t$  holds. Therefore, it is reasonable to choose  $\pi$  and  $V_0$  (note that  $V_0$  is not  $v$ , the initial wealth) as the decision variables in the corresponding optimization problem. In other words, we look for the optimal portfolio parameter combination, namely for the fraction invested in the risky asset and the initial investment amount that is supposed to be determined in such a way that the policyholder comes off best.<sup>4</sup> The appropriate maximization problem derived from Problem 1 and (16) is given by

**Problem 2** (Case A-bounded investment strategy).

$$\begin{aligned} & \max_{(\Psi_t)_{t \geq 0}} \frac{n^{1-\gamma}}{1-\gamma} \int_0^\infty e^{-\rho t} \kappa_t E_{\mathbb{P}}[\Psi_t^{1-\gamma}] dt \\ & \text{s.t. } v = P_0 = \int_0^\infty e^{-rt} I_t E_{\mathbb{Q}}[\Psi_t] dt. \end{aligned}$$

The objective of Problem 2 results from employing  $E_{\mathbb{P}}[V_t^{1-\gamma}] = \left( V_0 e^{\left( r + (\mu-r)\pi - \frac{\gamma^2}{2} \pi^2 \right) t} \right)^{1-\gamma}$ .

As it is possible to solve this problem analytically, we summarize the related optimizing quantities in a proposition:

**Proposition 3.** *The optimal values  $\pi^{*A}$  and  $V_0^{*A}$  for  $\pi$  and  $V_0$  solving Problem 2 are given by*

$$\pi^{*A} = \frac{\mu - r}{\gamma \sigma^2} \mathbb{1}_{\{\mu - r \leq \gamma \sigma^2\}} + \mathbb{1}_{\{\mu - r > \gamma \sigma^2\}} \quad \text{and} \quad V_0^{*A} = \frac{v}{\int_0^\infty I_t dt}. \tag{22}$$

**Proof.** The proof of Proposition 3 is reported in Appendix B.2.  $\square$

We observe that the optimal value for the trading strategy in Proposition 3 coincides with Merton’s fraction if  $\mu - r \leq \gamma \sigma^2$  (see Merton 1969).

**Case B:** As we assume for Case B that  $\Psi_t = G_t + \alpha(V_t - G_t)^+$  holds, it is sensible to again choose  $\pi$  and  $V_0$  as the decision variables in the corresponding optimization problem.<sup>5</sup> By means of Problem 1 and (17), the maximization problem for Case B can be formulated as follows:

**Problem 3** (Case B-bounded investment strategy).

$$\begin{aligned} & \max_{(\pi, V_0) \in [0, 1] \times (0, \infty)} \frac{n^{1-\gamma}}{1-\gamma} \int_0^\infty e^{-\rho t} \kappa_t \\ & \quad \cdot G_t^{1-\gamma} \left( \Phi(\bar{d}_t) + \int_0^\infty \left( 1 + \alpha \left( e^{\sigma \pi \sqrt{t} y} - 1 \right) \right)^{1-\gamma} \phi(y + \bar{d}_t) dy \right) dt \\ & \text{s.t. } v = P_0 = \int_0^\infty e^{-rt} I_t \left( G_t + \alpha \left( V_0 e^{rt} \Phi(\tilde{d}_t) - G_t \Phi(\hat{d}_t) \right) \right) dt. \end{aligned}$$

The objective of Problem 3 results from employing similar calculation techniques as before, which, inter alia, leads to

$$\begin{aligned} & E_{\mathbb{P}} \left[ \left( G_t + \alpha(V_t - G_t)^+ \right)^{1-\gamma} \right] \\ &= G_t^{1-\gamma} \Phi(\bar{d}_t) + \int_{\bar{d}_t}^{\infty} \left( G_t + \alpha \left( V_0 e^{rt + (\mu-r)\pi t - \frac{\sigma^2}{2} \pi^2 t + \sigma \pi \sqrt{t} z} - G_t \right) \right)^{1-\gamma} \phi(z) dz \quad (23) \\ &= G_t^{1-\gamma} \left( \Phi(\bar{d}_t) + \int_0^{\infty} \left( 1 + \alpha \left( e^{\sigma \pi \sqrt{t} y} - 1 \right) \right)^{1-\gamma} \phi(y + \bar{d}_t) dy \right), \end{aligned}$$

where  $\phi$  is the density of the standard normal distribution. Further, the substitution  $y = z - \bar{d}_t$  is applied in the third line and the function  $\bar{d}_t$  is given by

$$\bar{d}_t = \frac{\ln\left(\frac{G_t}{V_0}\right) - rt - (\mu - r)\pi t + \frac{\sigma^2}{2} \pi^2 t}{\sigma \pi \sqrt{t}}. \quad (24)$$

If we try to solve this problem by using the method of Lagrange multipliers (see, e.g., Bertsekas 2014), the corresponding Lagrange function  $\mathcal{L}(\pi, V_0, \lambda)$ , where  $\lambda$  is the Lagrange multiplier, is defined as

$$\begin{aligned} \mathcal{L}(\pi, V_0, \lambda) &= \frac{n^{1-\gamma}}{1-\gamma} \int_0^{\infty} e^{-\rho t} \kappa_t G_t^{1-\gamma} \left( \Phi(\bar{d}_t) + \int_0^{\infty} \left( 1 + \alpha \left( e^{\sigma \pi \sqrt{t} y} - 1 \right) \right)^{1-\gamma} \phi(y + \bar{d}_t) dy \right) dt \\ &\quad + \lambda \left( v - \int_0^{\infty} e^{-rt} I_t \left( G_t + \alpha \left( V_0 e^{rt} \Phi(\tilde{d}_t) - G_t \Phi(\hat{d}_t) \right) \right) dt \right). \quad (25) \end{aligned}$$

The first-order condition with respect to  $\pi$  is given as

$$\begin{aligned} \frac{\partial}{\partial \pi} \mathcal{L}(\pi, V_0, \lambda) &= \frac{n^{1-\gamma}}{1-\gamma} \int_0^{\infty} e^{-\rho t} \kappa_t G_t^{1-\gamma} \left( \phi(\bar{d}_t) \frac{1}{\pi} \tilde{d}_t + \int_0^{\infty} \left( 1 + \alpha \left( e^{\sigma \pi \sqrt{t} y} - 1 \right) \right)^{-\gamma} \phi(y + \bar{d}_t) \right. \\ &\quad \cdot \left. \left( (1-\gamma) \alpha e^{\sigma \pi \sqrt{t} y} \sigma \sqrt{t} y - \left( 1 + \alpha \left( e^{\sigma \pi \sqrt{t} y} - 1 \right) \right) \right) (y + \bar{d}_t) \frac{1}{\pi} \tilde{d}_t dy \right) dt \\ &\quad - \lambda \alpha V_0 \sigma \int_0^{\infty} I_t \phi(\tilde{d}_t) \sqrt{t} dt = 0. \quad (26) \end{aligned}$$

The first-order condition with respect to  $V_0$  is given as

$$\begin{aligned} \frac{\partial}{\partial V_0} \mathcal{L}(\pi, V_0, \lambda) &= \frac{n^{1-\gamma}}{(1-\gamma)V_0 \sigma \pi} \int_0^{\infty} e^{-\rho t} \kappa_t G_t^{1-\gamma} \frac{1}{\sqrt{t}} \left( \int_0^{\infty} \left( 1 + \alpha \left( e^{\sigma \pi \sqrt{t} y} - 1 \right) \right)^{1-\gamma} \right. \\ &\quad \cdot \left. \phi(y + \bar{d}_t) (y + \bar{d}_t) dy - \phi(\bar{d}_t) \right) dt - \lambda \alpha \int_0^{\infty} I_t \Phi(\tilde{d}_t) dt = 0, \quad (27) \end{aligned}$$

and the one with respect to  $\lambda$  naturally coincides with the budget constraint:

$$v = \int_0^{\infty} e^{-rt} I_t \left( G_t + \alpha \left( V_0 e^{rt} \Phi(\tilde{d}_t) - G_t \Phi(\hat{d}_t) \right) \right) dt. \quad (28)$$

From (26) and (27), the following equation must hold true:

$$\begin{aligned} & \int_0^{\infty} e^{-\rho t} \kappa_t G_t^{1-\gamma} \left( \phi(\bar{d}_t) \tilde{d}_t + \int_0^{\infty} \left( 1 + \alpha \left( e^{\sigma \pi \sqrt{t} y} - 1 \right) \right)^{-\gamma} \phi(y + \bar{d}_t) \right. \\ & \cdot \left. \left( (1-\gamma) \alpha e^{\sigma \pi \sqrt{t} y} \sigma \pi \sqrt{t} y - \left( 1 + \alpha \left( e^{\sigma \pi \sqrt{t} y} - 1 \right) \right) \right) (y + \bar{d}_t) \tilde{d}_t dy \right) dt \int_0^{\infty} I_t \Phi(\tilde{d}_t) dt \\ &= \int_0^{\infty} e^{-\rho t} \kappa_t G_t^{1-\gamma} \frac{1}{\sqrt{t}} \left( \int_0^{\infty} \left( 1 + \alpha \left( e^{\sigma \pi \sqrt{t} y} - 1 \right) \right)^{1-\gamma} \phi(y + \bar{d}_t) (y + \bar{d}_t) dy - \phi(\bar{d}_t) \right) dt \\ & \cdot \int_0^{\infty} I_t \phi(\tilde{d}_t) \sqrt{t} dt. \quad (29) \end{aligned}$$

For the calculations in (26) and (27), the following identities are applied:  $\frac{\partial \tilde{d}_t}{\partial \pi} = \frac{1}{\pi} \tilde{d}_t$ ,  $\frac{\partial \hat{d}_t}{\partial \pi} = \frac{\partial \tilde{d}_t}{\partial \pi} - \sigma \sqrt{t}$  and

$$V_0 e^{rt} \phi(\tilde{d}_t) - G_t \phi(\hat{d}_t) = 0. \tag{30}$$

The detailed derivation of (30) is reported in Appendix A. The solution of the system of Equations (28) and (29) (when it exists) provides the optimal values for  $\pi$  and  $V_0$  in Case B. However, due to the complexity of this system of equations, we are unable to find explicit formulas for the solution of Problem 3. Therefore, in what follows, we numerically solve Problem 3 to find the optimal values  $\widetilde{\pi}^{*B}$  and  $\widetilde{V}_0^{*B}$ .

### 5. Numerical Analysis

In this section, we aim at discovering distinct characteristics of the introduced unit-linked tontine product by means of numerical studies. For these studies, concrete assumptions about definite numbers for the various appearing parameters and about other modeling implementations need to be made initially. Subsequently, the specific main objective is to compare, in terms of the utility of a policyholder, our two different variants for the unit-linked tontine product established and priced in Section 3.2, and optimized in Section 4.2 within several sensitivity analyses. Additionally, we seek to integrate the traditional tontine with non-unit-linked payments into this comparison. Thereby, we are able to indicate whether the individual, in the analyzed instances, prefers that the tontine payment is linked to the financial market.

#### 5.1. Setup

First, let us set up the overall framework with the different assumptions for our numerical studies. We start with the determination of the modeling of the shocked survival curves  ${}_t p_x^{1-\epsilon}$  and  ${}_t \tilde{p}_x^{1-\epsilon}$ , respectively. We initially specify the survival probabilities  ${}_t p_x$  and  ${}_t \tilde{p}_x$  as

$${}_t p_x = e^{-\int_0^t m_{x+\tau} d\tau} = e^{\frac{x-g_2}{s_1} \left(1 - e^{-\frac{t}{s_1}}\right)} \quad \text{and} \quad {}_t \tilde{p}_x = e^{-\int_0^t \tilde{m}_{x+\tau} d\tau} = e^{\frac{x-\tilde{g}_2}{\tilde{s}_1} \left(1 - e^{-\frac{t}{\tilde{s}_1}}\right)}, \tag{31}$$

where

$$m_{x+\tau} = \frac{1}{g_1} e^{\frac{x+\tau-g_2}{s_1}} \quad \text{and} \quad \tilde{m}_{x+\tau} = \frac{1}{g_1} e^{\frac{x+\tau-\tilde{g}_2}{\tilde{s}_1}}, \tag{32}$$

are the individual's forces of mortality at the age of  $x + \tau$  with  $\tau \geq 0$  following the Gompertz law of mortality (see, e.g., Milevsky and Salisbury 2015) under  $\mathbb{P}$  and  $\mathbb{Q}$ , respectively. We refer to  $g_1 > 0$ ,  $g_2 > 0$  and  $\tilde{g}_2 > 0$  as the first Gompertz parameter describing the dispersion and the second Gompertz parameters describing the modal ages at death. Note that we assume that  $g_1$  remains the same under both probability measures  $\mathbb{P}$  and  $\mathbb{Q}$ , and that  $\tilde{g}_2 \geq g_2$ , so that (9) is fulfilled. For the mortality shock  $\epsilon$ , following Chen et al. (2019), we assume its distribution to be truncated normal:

$$\epsilon \sim \mathcal{N}_{(-\infty, 1)}(\eta_1, \eta_2^2), \tag{33}$$

where  $\eta_1$  and  $\eta_2^2$  are the mean and the variance parameter of the normal distribution truncated on the interval  $(-\infty, 1)$ , respectively. Table 1 summarizes the assumed baseline values for the relevant parameters and their corresponding ranges for  $n$ ,  $\mu$ ,  $\sigma$ ,  $\gamma$  and  $\rho$ , used in our sensitivity analyses.

When choosing the parameter values given in Table 1, we include the following considerations:

- For the choice of the value for  $r$ , we take account of the current low interest rate environments in many European countries. For example, the average risk-free rate on investments in the United Kingdom in the year 2020 equals only 1.1% (see Statista 2020a);
- For the choice of the value for  $\gamma$ , we refer to Thomas (2016); Thomas et al. (2010), who mention an estimate of the average risk aversion of British citizens that amounts to 0.85 when considering the power utility function. In Thomas et al. (2010); Waddington et al. (2013), an average risk aversion  $\gamma \in (0.8, 1)$  is obtained;<sup>6</sup>
- For simplicity, we equate the value for  $\rho$  with the one of the risk-free interest rate. This is a common assumption. However, note that the cases  $r > \rho$  and  $r < \rho$  are also considered when letting  $\rho$  vary in the sensitivity analyses;
- For the choice of the value for  $v$ , we are guided by Royal London (2018). In this report, it is estimated that an average British employee needs to invest £260,000 in her private pension provision to maintain the same standard of living as in her working period during the retirement phase;
- For the choices of the values for  $g_1$  and  $g_2$ , we follow Milevsky (2020), who presents 9.38 and 88.85 for the two Gompertz parameters for British females. For the choice of the value for  $\tilde{g}_2$ , we roughly convert the corresponding applied numbers from Chen and Rach (2019) into our framework, where we take into account that the (implicit) safety loadings included in the premiums, that stem from the usage of the risk neutral probability measure  $\mathbb{Q}$  during pricing can depend on the pool size  $n$ . As described in Section 3, a higher  $n$  implies that less unsystematic mortality risk is incorporated in the tontines and, consequently, lower (implicit) safety loadings can be chosen. We handle this by considering  $\tilde{g}_2$  as a function of  $n$ . By linearly interpolating, we find  $\tilde{g}_2(n) = -0.0062n + 95.08$ . Using this relation guarantees that  ${}_t\tilde{p}_x$ , and thereby also the (implicit) safety loadings, decreases in  $n$ . Note that the condition  $\tilde{g}_2(n) \geq g_2$  is fulfilled in all considered instances, such that  ${}_t\tilde{p}_x(n) \geq {}_t p_x$ .

**Table 1.** Specification of relevant parameters for numerical studies.

Symbol	Description	Value	Range
$n$	Initial number of participants	100	[1, 1000]
$x$	Initial age of the participants	65	–
$\mu$	Drift rate of the risky asset	0.1	(0.01, 0.2]
$\sigma$	Volatility of the risky asset	0.35	(0, 0.7]
$r$	Risk-free interest rate	0.01	–
$\gamma$	Measure of the policyholder’s risk aversion	0.85	(0, 5] \ {1}
$\rho$	Subjective discount rate	0.01	[0, 0.05]
$v$	Available initial wealth	£260,000	–
$g_1$	First Gompertz parameter	9.38	–
$g_2$	Second Gompertz parameter under $\mathbb{P}$	88.85	–
$\tilde{g}_2$	Second Gompertz parameter under $\mathbb{Q}$	94.46	–
$\eta_1$	Mean parameter of the truncated normal distribution	−0.0035	–
$\eta_2^2$	Variance parameter of the truncated normal distribution	0.0814 <sup>2</sup>	–

We also need to introduce a practicable choice in Case B for the guaranteed payment  $G_t$ , which has not been specified so far. Since we aim at taking account of the circumstance that the individual’s attitude towards the guaranteed payment can change if she gets older, we choose

$$G_t = Ge^{\delta t}, \tag{34}$$

where  $G > 0$  is the prescribed constant initial guarantee amount and  $\delta$  the guarantee growth rate. By this stipulation, we can consider different situations, such as the case in which the liquidity needs of the policyholder increase with age, which can be modeled by choosing a positive  $\delta$ . If it is required that  $G_t$  is time-independent, i.e., a constant over time, simply



let  $\delta = 0$ . We choose  $G$  in such a way that the value of the guaranteed payments at time 0 corresponds to a fraction, say  $g \in (0, 1)$ , of the total premium. Relying on (13), which represents the described correspondence if  $P_0 = v$  is multiplied by  $g$  and  $\Psi_t$  is replaced by  $G_t$ , we obtain

$$G = \frac{gv}{\int_0^\infty e^{-(r-\delta)t} I_t dt}. \tag{35}$$

For the three case-related parameters  $\alpha$ ,  $\delta$  and  $g$ , we summarize their assumed baseline values in Table 2. For the sensitivity analyses below, the corresponding ranges of  $\delta$  and  $g$  are also presented in Table 2.

**Table 2.** Specification of relevant parameters related to Case B for numerical studies.

Symbol	Description	Value	Range
$\alpha$	Participation rate	0.9	–
$\delta$	Guarantee growth rate	0.01	[−0.03, 0.05]
$g$	Guaranteed premium fraction	0.75	(0, 1)

The following considerations are taken into account when choosing the parameter values given in Table 2:

- For the choice of the value for  $\alpha$ , we first notice that participation rates between 80% and 100% are commonly practiced in reality (see, e.g., Bacinello et al. 2018). Applying the mean value appears appropriate;
- We choose the value for  $\delta$  to be equal to  $r = 0.01$ . Note that the cases where  $r > \delta$  or  $r < \delta$  are also considered when letting  $\delta$  vary in the sensitivity analyses;
- For the choice of the value for  $g$ , we first notice that the guaranteed premium fraction is often chosen between 60% and 90% in practice, as this can be exemplarily observed for the product “GarantieRente Performance” offered by Gothaer (2021). Again, applying the mean value appears appropriate.

### 5.2. Comparison

The main questions we intend to answer in this numerical analysis are as follows:

- From the individual’s viewpoint, which of the two introduced unit-linked tontine variants is preferred? How does this preference depend on the parameter values?
- From the individual’s viewpoint, how does the introduced unit-linked tontine product perform in comparison to the traditional tontine product with no financial market component? How does this performance ordering depend on the parameter values?

These questions will be answered in Section 5.2.2, where we present our numerical results and sensitivity analyses based on the assumptions made in Section 5.1. In preparation for this, a short overview of the necessary details on the traditional tontine is given and the precise comparison approach is explained in Section 5.2.1.

#### 5.2.1. Traditional Tontine and Comparison Approach

Recall that the traditional tontines established in Milevsky and Salisbury (2015) are introduced in Section 2.1, where its total payment is given in (1). In order to buy a traditional tontine, we assume that the individual also spends her available initial wealth  $v$  to pay the single initial premium charged for it. By replacing the tontine payout process  $\Psi_t$  in (13) by the tontine payout function  $d_t$ , this premium can be calculated via

$$P_0 = \int_0^\infty e^{-rt} I_t d_t dt. \tag{36}$$

We consider two different variants of specific forms of  $d_t$ , one rather theoretical and one rather practical. The first one, which is also examined in, e.g., Chen et al. (2019),

arises directly from the maximization of the discounted expected utility associated with the purchase of the traditional tontine. In the corresponding optimization problem,  $d_t$  is naturally chosen as the decision variable. Details on this problem and its solution that is given by the optimal version  $d_t^*$  for  $d_t$  are reviewed in Appendix C. We refer to the resulting product as the *optimal traditional tontine*. For the second specific form of  $d_t$ , we use one of the so-called natural tontines proposed by Milevsky and Salisbury (2015). This more practicable payout function is given by

$$d_t = E_{\mathbb{Q}} \left[ \mathbb{1}_{\{\zeta_x > t\}} \right] d = {}_t\tilde{p}_x E_{\mathbb{Q}} \left[ e^{-\ln({}_t\tilde{p}_x)\epsilon} \right] d = {}_t\tilde{p}_x M_{\epsilon}(-\ln({}_t\tilde{p}_x)) d, \tag{37}$$

where  $d > 0$  is constant over time and determined by plugging (37) in the budget constraint  $v = P_0$ , where  $P_0$  is given in (36):

$$d^* = \frac{v}{\int_0^{\infty} e^{-rt} I_t {}_t\tilde{p}_x M_{\epsilon}(-\ln({}_t\tilde{p}_x)) dt}. \tag{38}$$

Note that by applying (37), the total payment to the living traditional tontine holder is actually also constant over time if deaths in the pool occur as expected. We refer to the product resulting from (37) and (38) as the *natural traditional tontine*.

For the comparison, we look at the (maximized) discounted expected utilities arising from the optimal findings in Section 4.2 and from above that the individual attains when acquiring the respective tontine product alternatives. They are denoted by  $EU^{*A}$  and  $EU^{*B}$  in case of the two unit-linked tontines from Case A and Case B, respectively, by  $EU^{*OT}$  in case of the optimal traditional tontine and by  $EU^{*NT}$  in case of the natural traditional tontine. For the sake of completeness, an overview of the formulas for the different (maximized) discounted expected utilities is given in Appendix D.<sup>7</sup> The reason why such a direct comparison approach is valid within our framework is that the individual spends the same initial wealth  $v$  for every product variant. Therefore, since the purchase costs for the policyholder are always identical, she rationally prefers the tontine that provides her with the highest utility. To make our comparison results easier to interpret, we do not straightforwardly consider the different (maximized) discounted expected utility levels, but the corresponding certainty equivalents, which are the safe amounts that make the individual indifferent between obtaining them and the optimal uncertain total payments of the tontine products. These certainty equivalents, which are denoted by  $CE^{*j}$  with  $j \in \{A, B, OT, NT\}$  marking the respective product variant, are thus calculated by using the same concept as in (20) and the quantities  $EU^{*j}$  for equating:

$$E_{\mathbb{P}} \left[ \int_0^{\infty} e^{-\rho t} u(CE^{*j}) \mathbb{1}_{\{\zeta_x > t\}} dt \right] = EU^{*j} \\ \Leftrightarrow CE^{*j} = \left( (1 - \gamma) EU^{*j} \left( \int_0^{\infty} e^{-\rho t} \int_{-\infty}^1 {}_t p_x^{1-z} f_{\epsilon}(z) dz dt \right)^{-1} \right)^{\frac{1}{1-\gamma}}. \tag{39}$$

As  $EU^{*j}$  is strictly increasing in  $CE^{*j}$ , comparing the (maximized) discounted expected utilities is equivalent to comparing the certainty equivalents.

### 5.2.2. Numerical Results and Sensitivity Analyses

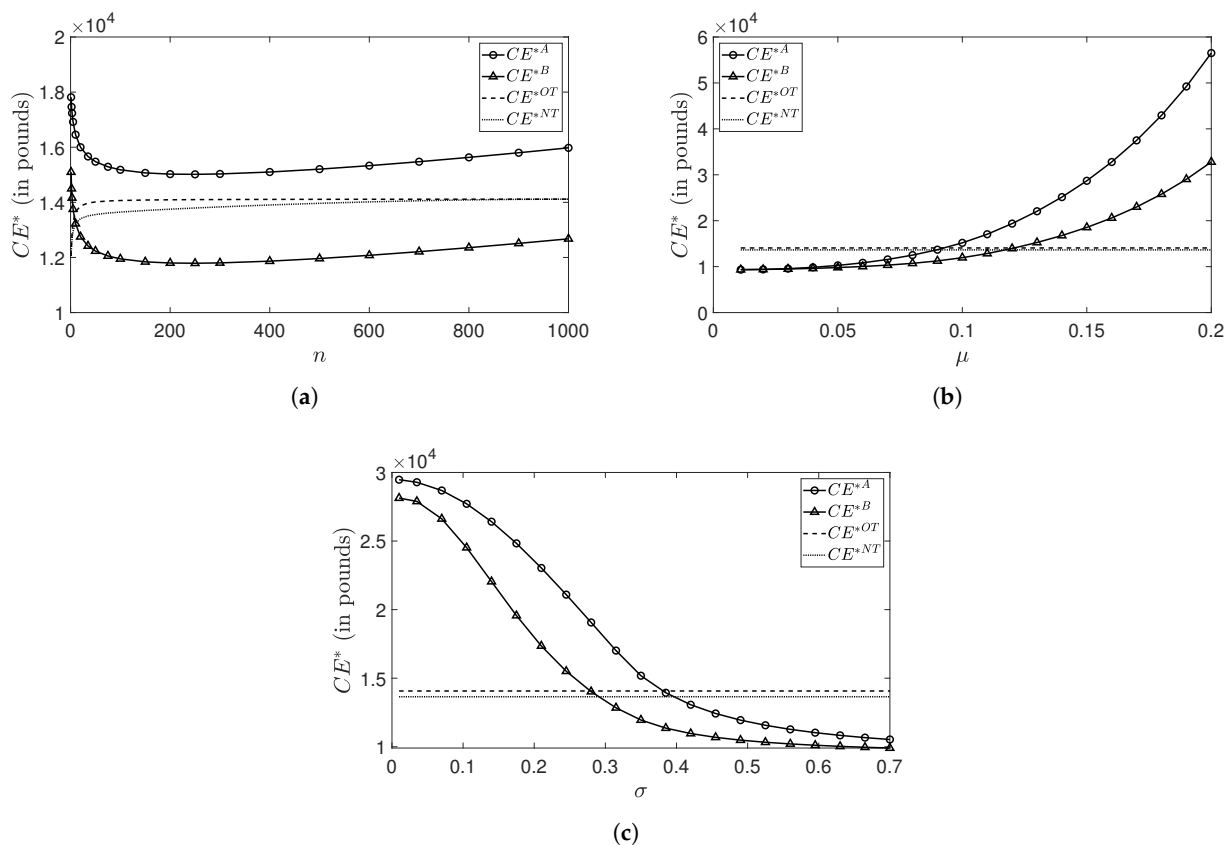
In Table 3, we show the first numerical findings, namely the ones for  $CE^{*j}$ , that emerge from applying the baseline parameter values given in Tables 1 and 2 (for Case B).<sup>8</sup>

**Table 3.** Certainty equivalents of different tontines with baseline parameter values.

$CE^{*A}$	£15,180.83
$CE^{*B}$	£11,948.69
$CE^{*OT}$	£14,066.46
$CE^{*NT}$	£13,647.26

Comparing the certainty equivalents reported in in Table 3 shows that the policyholder is in the best position as long as she holds the unit-linked tontine designed in Case A. When comparing only the unit-linked tontine variants, it is more beneficial for the individual if the tontine payout process does not include an additional guaranteed payment as in Case B, but rather simply complies with the entire portfolio value that arises entirely out of optimally investing in the financial market. The unit-linked tontine variant from Case B actually performs worse than the traditional tontine, where even the more practicable version, the natural traditional tontine, surpasses it by far, i.e.,  $CE^{*NT} \gg CE^{*B}$ . Do the previous observations also hold if certain parameter values change?

In Figures 1 and 2, we show the numerical comparison findings that emerge from applying the parameter ranges given in Table 1. In particular, we present the resulting curves for  $CE^{*j}$  if the parameters  $n, \mu, \sigma, \gamma$  and  $\rho$  vary, respectively.<sup>9</sup>



**Figure 1.** Effects of  $n$  (a),  $\mu$  (b) and  $\sigma$  (c) on certainty equivalents.

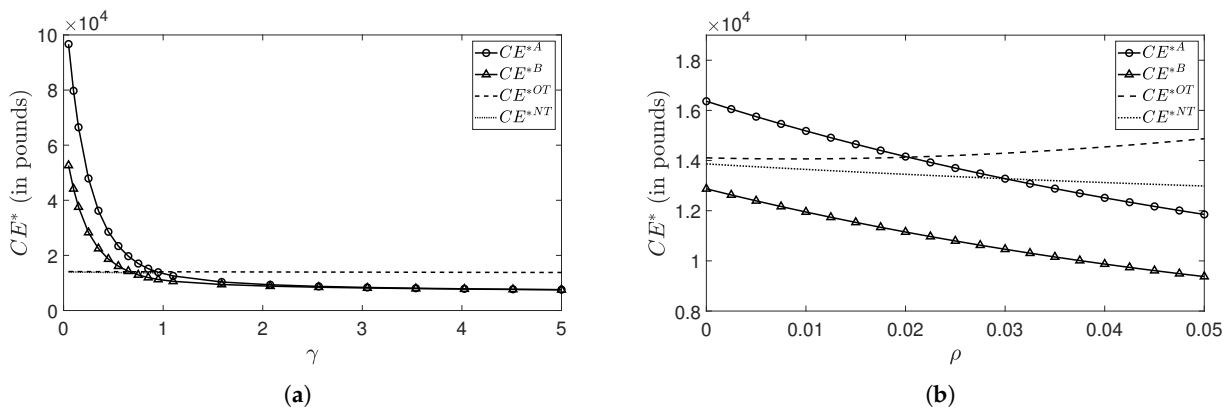


Figure 2. Effects of  $\gamma$  (a) and  $\rho$  (b) on certainty equivalents.

Two main observations can be universally drawn from Figures 1 and 2:

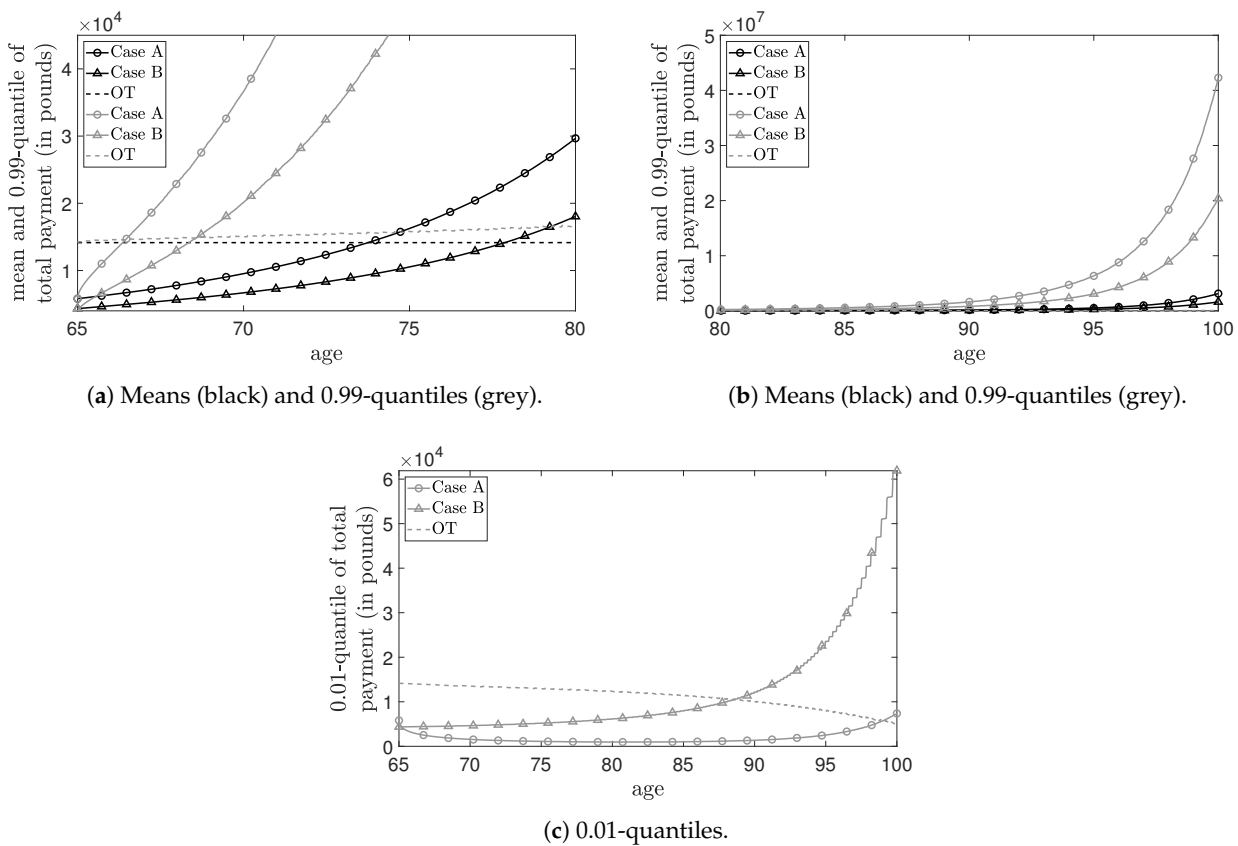
- Overall, we detect in each graph that, like in Table 3, the unit-linked product from Case A provides the policyholder with a higher certainty equivalent than the one from Case B. As such, varying parameter values does not seem to affect the performance order between the two unit-linked tontine alternatives (at least not for the parameters and their ranges under consideration). Nevertheless, the performance of the tontine from Case B more and more approaches that of the one from Case A if  $\mu$  decreases or if  $\sigma$  or  $\gamma$  increases;
- There exist regions in which the unit-linked tontine variants make the individual better off than the traditional tontine variants. This is not very surprising for Case A, as is known. However, it reveals that our Case B can also outperform the traditional tontine in some parameter constellations. This emphasizes the potential attractiveness of this participating tontine, especially to customers who consider additional guarantee components important. We remark that participants preferring guarantees are typically loss averse, see e.g., Berkelaar et al. (2004); Kahneman and Tversky (1979). In particular, the unit-linked tontine performs well if  $n$  is either very low or high, if  $\mu$  is high or if  $\sigma$ ,  $\gamma$  or  $\rho$  is low. On the whole, we conclude that if the traditional tontine product is consulted as a basis for comparison, it is possible that the unit-linked counterpart is more successful among the customers and, thus, it seems reasonable to promote it.

As already pointed out by Chen et al. (2021) (Theorem 5.2), the impact of the pool size on the attractiveness of a tontine is not monotonically increasing. In their context, they compare tontines with annuities and the critical pool size determines the preference ordering between annuities and tontines. After the pool size reaches a certain magnitude, tontines will become, for instance, more attractive than conventional annuities. They observe that this number is rather small for a conventional tontine case. Now, in our unit-linked products, this number seems rather large, shown in Figure 1a to be larger than 200, beyond which the attractiveness of the unit-linked tontine products increases in the pool size.

In order to get a better understanding of the findings derived from Table 3 and of the above-mentioned observations based on Figures 1 and 2, we show in Figure 3 the means and 0.01-/0.99-quantiles under  $\mathbb{P}$  of the optimal total payments for Cases A and B and the traditional tontine with respect to age. For the generation of the graphs, we assume that the considered policyholder is always alive and that the parameters attain their baseline values given in Tables 1 and 2 (for Case B).<sup>10</sup>

By comparing Case A and Case B by means of Figure 3, the effect of the guaranteed payment picked up in Case B becomes clear: In Figure 3c, we notice that the 0.01-quantile curve for Case B is almost always significantly above the one for Case A. This implies

that the inclusion of the guaranteed payment prevents the policyholder in Case B from receiving a very low total payment in bad market scenarios. Yet, at the same time, the guaranteed payment also limits a possible positive development of the total payment in good market scenarios, which is, however, completely exploited by the unit-linked tontine from Case A. This is recognizable by the 0.99-quantile curves displayed in Figure 3a,b. Since the scale of the 0.99-quantiles, especially in Figure 3b, is much larger than the one of the 0.01-quantiles in Figure 3c, the dominance of Case A in good market scenarios clearly outperforms the dominance of Case B in bad market scenarios. Hence, as visible in Figure 3a,b, the mean of the total payment for Case A is consistently higher than the one for Case B. Due to the fact that the power utility function is strictly increasing, we can infer from this finding that  $CE^{*A} > CE^{*B}$  holds, as observed above for the given parameters. Moreover, we, particularly in Figure 3a,c, see that the curves for the traditional tontine, which is represented here by the optimal version, can be above or below the ones for the unit-linked tontine. That is why the policyholder prefers the traditional tontine to the unit-linked tontine in some instances, while in others she does not, as apparent from Figures 1 and 2. The partial dominance of the traditional tontine explicitly shown in Figure 3 suffices to beat the performance of the unit-linked tontine from Case B, but not the one from Case A. This can be observed from Table 3, where all parameters also attain their baseline values.



**Figure 3.** Means and 0.99-quantiles at earlier retirement ages (a) and at more advanced retirement ages (b), and 0.01-quantiles (c) of optimal total payments for Cases A and B and the optimal traditional tontine (OT) depending on age, assuming that the policyholder is always alive and the parameters attain baseline values.

In the following, the impacts of the parameters  $n, \mu, \sigma, \gamma$  and  $\rho$  and, eventually of the varying parameters  $\delta$  and  $g$ , being only related to Case B, will be discussed in detail.

### Sensitivity Analyses Regarding $n$

In Figure 1a, we notice right away the converse behavior of  $CE^{*A}$  and  $CE^{*B}$  with regard to  $CE^{*OT}$  and  $CE^{*NT}$  as long as the initial number  $n$  of participants in the pool ranges within relatively small values. Especially when an extremely small pool takes in a very few new participants, the policyholder’s benefit drops sharply in case of the unit-linked tontine, whereas it rises quickly for the traditional tontine. From around  $n = 250$  on, the courses of the curves belonging to the unit-linked tontine switch to an upward movement, which becomes even steeper than the one for the traditional tontine. In summary, a purchase decision in favor of the unit-linked tontine is wise if the pool size is either very small or large.

In order to get a better understanding of the recorded observations, we let  $n$  vary again and study the resulting optimal values  $V_0^{*A}$  and  $\widetilde{V}_0^{*B}$  for the initial investment amount in Figure 4.

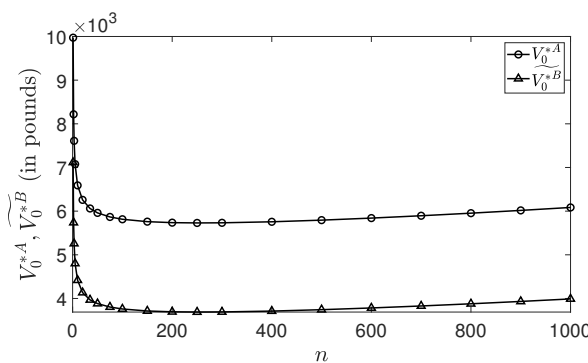


Figure 4. Effect of  $n$  on optimal values for the initial investment amount.

We observe very similar curve shapes for  $V_0^{*A}$  and  $\widetilde{V}_0^{*B}$  in Figure 4 compared to the ones for  $CE^{*A}$  and  $CE^{*B}$  in Figure 1a, namely the strong decline in  $n$  at the beginning, which quickly lessens and, from around  $n = 250$  on, turns into an increase. Consequently, it seems that the behavior of the initial investment amount for a varying pool size causes the performance development of the unit-linked tontine described above. If we exemplarily consider the formula in (22) for  $V_0^{*A}$ , only the initial decrease appears plausible at first glance. However, when we recall that lower (implicit) safety loadings included in the premiums can be chosen if  $n$  grows by reducing  ${}_t\widetilde{p}_x$ , just like we do, it is comprehensible why the decrease can be slowed down and possibly even be reversed at some point. Below, we provide more interpretations to the impact of the pool size  $n$ :

- Unit-linked products can outperform the traditional tontines (both the natural tontine and the optimal tontine), but can also be beaten by the traditional ones. With the chosen parameters, the unit-linked tontine type A outperforms, while the unit-linked tontine type B is beaten by, the traditional ones;
- For the given parameters, we observe that the unit-linked products with  $n = 1$  leads to the highest utility level. It is implied that the unit-linked annuity is most favored. However, let us point out that the result depends substantially on the choice of the parameters;
- The main message is that, depending on the design of the unit-linked tontine products including the pool size, the unit-linked tontine product can be attractive for some individuals. Among all these products, there is no dominance in terms of expected utility. The unit-linked products enriches the variety of the products.

### Sensitivity Analyses Regarding $\mu$ and $\sigma$

If the policyholder chooses the unit-linked tontine, we observe in Figure 1b,c that her utility enhances more and more as long as the drift rate  $\mu$  of the risky asset increases and

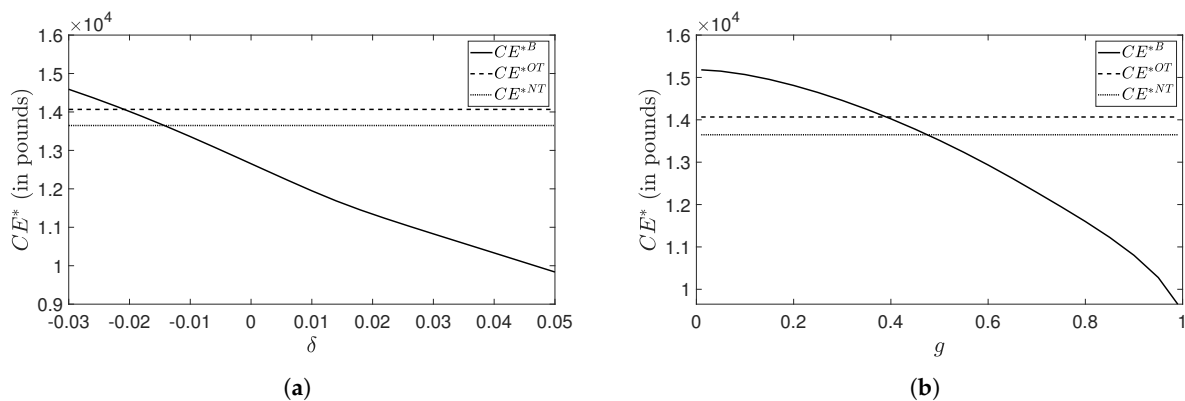
its volatility  $\sigma$  decreases, respectively. This is because the risky asset, in which investments are made within the framework of the unit-linked tontine, is clearly more profitable if its return grows and its risk reduces, as can be seen, for example, from a higher Sharpe ratio  $\frac{\mu-r}{\sigma}$ , which eventually is naturally also more beneficial to the policyholder. As the certainty equivalents associated with the traditional tontine are apparently not affected by a varying  $\mu$  or  $\sigma$  due to its payout's independence of the financial market, there is a certain level at which the performance of the risky asset is so good that the traditional tontine is no longer preferred.

**Sensitivity Analyses Regarding  $\gamma$  and  $\rho$**

In Figure 2a, we find that  $CE^{*j}$  declines for all  $j$  for higher values of  $\gamma$ , which means that each tontine variant gets less interesting for the individual when she becomes more risk-averse. This is because the risk inherent in the tontines is borne, to a great extent, by all participants in the pool, and the payments to the policyholder are, hence, uncertain to some extent. If the policyholder embraces less of this risk, i.e., she is more risk-averse and prefers more stable payments, her personal benefit is, thus, smaller. However, the curves displaying  $CE^{*A}$  and  $CE^{*B}$  exhibit (partly much) steeper slopes than those for  $CE^{*OT}$  and  $CE^{*NT}$  due to the fact that the unit-linked tontine alternatives contain more risk, namely not only the mortality risk but also the financial risk component. Therefore, if the policyholder tolerates more risk, i.e., she is less risk-averse ( $\gamma$  decreases) and prefers riskier payments, the unit-linked tontine is definitely the better choice. In Figure 2b, it can be observed that when the subjective discount rate  $\rho$  grows, the personal utilities induced by buying the examined tontines constantly diminish. The only exception is the optimal traditional tontine that regains some attractiveness for higher values of  $\rho$  in consequence of the specific structure of  $d_t^*$ , which is explicitly given in Appendix C. Since a higher subjective discount rate means that the individual tends to consume more at earlier retirement ages, the decreases of  $CE^{*A}$  and  $CE^{*B}$  in  $\rho$  are explainable by the steady increases of the means of the total unit-linked tontine payments over time, as this is exemplarily illustrated in Figure 3a,b. In these two figures, we also observe that the magnitudes of the two mean curves for the unit-linked tontine variants are a lot greater compared to the traditional tontine. This gives a reason for the steeper slopes of the curves displaying  $CE^{*A}$  and  $CE^{*B}$  in Figure 2b.

**Sensitivity Analyses Regarding  $\delta$  and  $g$**

When considering the choice for  $G_t$  as introduced in (34) and (35) for Case B, we are especially interested in the impact of the guarantee growth rate  $\delta$  and the guaranteed premium fraction  $g$  on the policyholder's tontine product preference. To analyze this, we look at the resulting curves for  $CE^{*B}$ ,  $CE^{*OT}$  and  $CE^{*NT}$  depicted in Figure 5, where the ranges given in Table 2 are applied.



**Figure 5.** Effects of  $\delta$  (a) and  $g$  (b) on certainty equivalents.

Both graphs of Figure 5 demonstrate a similar curve progression for  $CE^{*B}$ . In particular, the resulting certainty equivalents are negatively proportional to  $\delta$  and  $g$ . However, as the payout of the traditional tontine does not depend on  $\delta$  and  $g$ , neither  $CE^{*OT}$  nor  $CE^{*NT}$  changes. As a consequence, it is possible that the policyholder benefits more from the unit-linked tontine designed in Case B than from the traditional tontine if the guaranteed payment is low enough. On the other hand, a high guaranteed component in the unit-linked tontine may adversely affect the performance of the product due to stronger limitations on possible investment gains.

## 6. Conclusions

In the present article, we propose unit-linked tontine products that combine the tontine concept with the idea underlying unit-linked insurance policies, i.e., to tie payouts to the developments in the financial market. We examine a general payment structure of the product and analyze two specified payment structures. The two risk types contained in the unit-linked product are the financial risk stemming from the risky asset existing in the financial market and the mortality risk, for which we actually also incorporate the systematic part in our model. The premium required to buy the unit-linked tontine is determined in a risk-neutral pricing framework. Further, we study the optimal expected utility of an individual purchasing the unit-linked tontine by adjusting the payment structure. In our numerical comparison and sensitivity analyses, we contrast the policyholder's benefits arising out of the two optimized unit-linked tontine variants, as well as the optimal and the natural traditional tontine. In particular, we find that there exist circumstances in which the unit-linked tontine endows the policyholder with a higher utility level than the traditional tontine, emphasizing the potential of the suggested unit-linked tontines. More precisely, under our numerical setting with power utility functions, the unit-linked tontines might be a potential choice for the policyholder when the expected return of the risky asset is high or if the volatility of the risky asset, the policyholder's risk aversion or her subjective discount rate is low. Moreover, we observe that if its payout process is stipulated by the pure financial market portfolio value, the unit-linked tontine consistently makes the policyholder better off than in the case where it includes guaranteed payments. However, its performance approaches more and more that of the superior variant if the expected return of the risky asset decreases or if the volatility of the risky asset or the policyholder's risk aversion increases. Furthermore, when comparing the case with guaranteed payments with the traditional tontine with no financial market component, this case can nevertheless be attractive, especially to customers who consider additional guarantee elements important. Our findings would give reason to further study this new type of product in more realistic settings that take practical aspects into account, for instance, how the provider hedges the mortality and financial market risks related to the unit-linked tontines and what the net loss of the provider is. A thorough analysis of the hedging perspective requires a more dynamic framework and will be left for future research.

**Author Contributions:** Conceptualization, A.C. and T.N.; Methodology, A.C., T.N. and T.S.; Software, T.S.; Writing—original draft, T.S.; Writing—review & editing, A.C., T.N. and T.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** Thai Nguyen acknowledges the support of the Natural Sciences and Engineering Research Council of Canada [RGPIN-2021-02594].

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.



### Appendix A. Detailed Derivations

The equality in (11) holds for all  $t$  as

$$\begin{aligned}
 E_{\mathbb{Q}} \left[ \frac{\mathbb{1}_{\{\zeta_x > t\}}}{N_t} \right] &= E_{\mathbb{Q}} \left[ t \tilde{p}_x^{1-\epsilon} E_{\mathbb{Q}} \left[ \frac{1}{N_t} \mid \zeta_x > t, \epsilon \right] \right] = E_{\mathbb{Q}} \left[ t \tilde{p}_x^{1-\epsilon} E_{\mathbb{Q}} \left[ \frac{1}{N_t - 1 + 1} \mid \zeta_x > t, \epsilon \right] \right] \\
 &= E_{\mathbb{Q}} \left[ t \tilde{p}_x^{1-\epsilon} \sum_{k=0}^{n-1} \frac{1}{k+1} \binom{n-1}{k} (t \tilde{p}_x^{1-\epsilon})^k (1 - t \tilde{p}_x^{1-\epsilon})^{n-1-k} \right] \\
 &= E_{\mathbb{Q}} \left[ \sum_{k=0}^{n-1} \frac{n}{n} \frac{(n-1)!}{(k+1)!(n-1-k)!} (t \tilde{p}_x^{1-\epsilon})^{k+1} (1 - t \tilde{p}_x^{1-\epsilon})^{n-1-k} \right] \\
 &= \frac{1}{n} E_{\mathbb{Q}} \left[ \sum_{k=0}^{n-1} \binom{n}{k+1} (t \tilde{p}_x^{1-\epsilon})^{k+1} (1 - t \tilde{p}_x^{1-\epsilon})^{n-1-k} \right] \\
 &= \frac{1}{n} E_{\mathbb{Q}} \left[ \sum_{k=1}^n \binom{n}{k} (t \tilde{p}_x^{1-\epsilon})^k (1 - t \tilde{p}_x^{1-\epsilon})^{n-k} \right] \\
 &= \frac{1}{n} E_{\mathbb{Q}} \left[ \sum_{k=1}^n \binom{n}{k} (t \tilde{p}_x^{1-\epsilon})^k (1 - t \tilde{p}_x^{1-\epsilon})^{n-k} \right. \\
 &\quad \left. + \binom{n}{0} (t \tilde{p}_x^{1-\epsilon})^0 (1 - t \tilde{p}_x^{1-\epsilon})^n - \binom{n}{0} (t \tilde{p}_x^{1-\epsilon})^0 (1 - t \tilde{p}_x^{1-\epsilon})^n \right] \\
 &= \frac{1}{n} E_{\mathbb{Q}} \left[ \sum_{k=0}^n \binom{n}{k} (t \tilde{p}_x^{1-\epsilon})^k (1 - t \tilde{p}_x^{1-\epsilon})^{n-k} - (1 - t \tilde{p}_x^{1-\epsilon})^n \right] \\
 &= \frac{1}{n} E_{\mathbb{Q}} \left[ (1 - t \tilde{p}_x^{1-\epsilon} + t \tilde{p}_x^{1-\epsilon})^n - (1 - t \tilde{p}_x^{1-\epsilon})^n \right] = \frac{1}{n} E_{\mathbb{Q}} \left[ 1 - (1 - t \tilde{p}_x^{1-\epsilon})^n \right] \\
 &= \frac{1}{n} \int_{-\infty}^1 (1 - (1 - t \tilde{p}_x^{1-z})^n) f_{\epsilon}(z) dz = \frac{1}{n} I_t.
 \end{aligned}$$

The equality in (30) holds for all  $t$  as, with  $\pi$  denoting the ratio of a circle's circumference to its diameter,

$$\begin{aligned}
 V_0 e^{rt} \phi(\tilde{d}_t) - G_t \phi(\hat{d}_t) &= \frac{1}{\sqrt{2\pi}} \left( e^{\ln(V_0) + rt - \frac{\tilde{d}_t^2}{2}} - e^{\ln(G_t) - \frac{\hat{d}_t^2}{2}} \right) \\
 &= \frac{1}{\sqrt{2\pi}} \left( e^{\ln(V_0) + rt - \frac{\tilde{d}_t^2}{2}} - e^{\ln(G_t) - \frac{\tilde{d}_t^2 - 2\tilde{d}_t \sigma \pi \sqrt{t} + \sigma^2 \pi^2 t}{2}} \right) \\
 &= \frac{1}{\sqrt{2\pi}} e^{-\frac{\tilde{d}_t^2}{2}} \underbrace{\left( e^{\ln(V_0) + rt} - e^{\ln(G_t) + \ln\left(\frac{V_0}{G_t}\right) + rt + \frac{\sigma^2}{2} \pi^2 t - \frac{\sigma^2}{2} \pi^2 t} \right)}_{=0} = 0.
 \end{aligned}$$

### Appendix B. Proofs

#### Appendix B.1. Proposition 2

**Proof.** With the aid of the general pricing formula given in (13), we obtain the claim since the equality in (17) holds as

$$E_{\mathbb{Q}}[\Psi_t] = E_{\mathbb{Q}} \left[ G_t + \alpha (V_t - G_t)^+ \right] = G_t + \alpha E_{\mathbb{Q}} \left[ (V_t - G_t)^+ \right],$$

with

$$\begin{aligned}
 E_{\mathbb{Q}}[(V_t - G_t)^+] &= E_{\mathbb{Q}}\left[\left(V_0 e^{rt - \frac{\sigma^2}{2} \int_0^t \pi_s^2 ds + \sigma \sqrt{\int_0^t \pi_s^2 ds} Z - G_t\right) \mathbb{1}_{\{Z > \beta_t\}}\right] \\
 &= V_0 e^{rt} \int_{\beta_t}^{\infty} e^{-\frac{\sigma^2}{2} \int_0^t \pi_s^2 ds + \sigma \sqrt{\int_0^t \pi_s^2 ds} z} \phi(z) dz - G_t \int_{\beta_t}^{\infty} \phi(z) dz \\
 &= V_0 e^{rt} \int_{\beta_t - \sigma \sqrt{\int_0^t \pi_s^2 ds}}^{\infty} \phi(z) dz - G_t \Phi(-\beta_t) \\
 &= V_0 e^{rt} \Phi\left(\sigma \sqrt{\int_0^t \pi_s^2 ds} - \beta_t\right) - G_t \Phi(-\beta_t) = V_0 e^{rt} \Phi(\tilde{d}_t) - G_t \Phi(\hat{d}_t),
 \end{aligned}$$

where

$$\beta_t = \frac{\ln\left(\frac{G_t}{V_0}\right) - rt + \frac{\sigma^2}{2} \int_0^t \pi_s^2 ds}{\sigma \sqrt{\int_0^t \pi_s^2 ds}}.$$

□

Appendix B.2. Proposition 3

**Proof.** As the budget constraint in Problem 2 depends only on  $V_0$  and not on  $\pi$ , the optimal value for  $V_0$  is already completely determined by this constraint, so that we immediately obtain

$$V_0^{*A} = \frac{v}{\int_0^{\infty} I_t dt},$$

which is obviously positive, so that we also stick to the condition that  $V_0 > 0$ . Consequently, the budget constraint is entirely taken care of by  $V_0^{*A}$  and, thus, the determination of the optimal value of the trading strategy  $\pi$  can be done by simply maximizing the objective of Problem 2 with respect to  $\pi$ . To this end, we realize the shape of the objective as a function of  $\pi$  by considering the corresponding derivative:

$$\begin{aligned}
 &\frac{\partial}{\partial \pi} \left( \frac{(nV_0^{*A})^{1-\gamma}}{1-\gamma} \int_0^{\infty} e^{-\rho t} \kappa_t e^{(1-\gamma)\left(r+(\mu-r)\pi - \frac{\gamma\sigma^2}{2}\pi^2\right)t} dt \right) \\
 &= \frac{(nV_0^{*A})^{1-\gamma}}{1-\gamma} \int_0^{\infty} e^{-\rho t} \kappa_t e^{(1-\gamma)\left(r+(\mu-r)\pi - \frac{\gamma\sigma^2}{2}\pi^2\right)t} (1-\gamma)(\mu-r-\gamma\sigma^2\pi) t dt \\
 &= \underbrace{(nV_0^{*A})^{1-\gamma}}_{>0} (\mu-r-\gamma\sigma^2\pi) \underbrace{\int_0^{\infty} e^{-\rho t} \kappa_t e^{(1-\gamma)\left(r+(\mu-r)\pi - \frac{\gamma\sigma^2}{2}\pi^2\right)t} t dt}_{>0}.
 \end{aligned}$$

The identified derivative is positive (negative), i.e., the objective is strictly increasing (decreasing) in  $\pi$ , if

$$\mu - r - \gamma\sigma^2\pi \begin{matrix} (>) \\ (<) \end{matrix} 0 \Leftrightarrow \pi \begin{matrix} (<) \\ (>) \end{matrix} \frac{\mu - r}{\gamma\sigma^2}.$$

Since we also need to adhere to the condition that  $\pi \in [0, 1]$ , it is clear that, as long as  $\mu - r \leq \gamma\sigma^2$ , the optimal value for  $\pi$  is given by  $\pi^{*A} = \frac{\mu - r}{\gamma\sigma^2}$ . Otherwise, if  $\mu - r > \gamma\sigma^2$ , it is  $\pi^{*A} = 1$ . Overall, we find

$$\pi^{*A} = \frac{\mu - r}{\gamma\sigma^2} \mathbb{1}_{\{\mu - r \leq \gamma\sigma^2\}} + \mathbb{1}_{\{\mu - r > \gamma\sigma^2\}}.$$

□

### Appendix C. Review of Optimization Problem for Traditional Tontine

In the style of Problem 1, the maximization problem for the traditional tontine with the decision variable  $d_t$  can be, by using (36) and replacing  $\Psi_t$  by  $d_t$ , formulated as follows:

$$\begin{aligned} \max_{(d_t)_{t \geq 0}} & \frac{n^{1-\gamma}}{1-\gamma} \int_0^\infty e^{-\rho t} \kappa_t d_t^{1-\gamma} dt \\ \text{s.t. } & v = P_0 = \int_0^\infty e^{-rt} I_t d_t dt. \end{aligned}$$

By applying the techniques in Chen et al. (2019), it can be shown that the optimal solution is given by

$$d_t^* = \left( \frac{\lambda^* e^{-rt} I_t}{n^{1-\gamma} e^{-\rho t} \kappa_t} \right)^{-\frac{1}{\gamma}},$$

where the optimal Lagrange multiplier is given by

$$\lambda^* = v^{-\gamma} \left( \int_0^\infty \frac{(e^{-rt} I_t)^{1-\frac{1}{\gamma}}}{(n^{1-\gamma} e^{-\rho t} \kappa_t)^{-\frac{1}{\gamma}}} dt \right)^\gamma.$$

### Appendix D. Overview of Formulas for (Maximized) Discounted Expected Utilities

The formulas for the different (maximized) discounted expected utilities  $EU^{*j}$  with  $j \in \{A, B, OT, NT\}$  that are applied for the comparison are listed in the following overview:

$$\begin{aligned} EU^{*A} &= \frac{(nV_0^{*A})^{1-\gamma}}{1-\gamma} \int_0^\infty e^{-\rho t} \kappa_t e^{(1-\gamma)\left(r + (\mu - r)\pi^{*A} - \frac{\gamma\sigma^2}{2}(\pi^{*A})^2\right)t} dt, \\ EU^{*B} &= \frac{n^{1-\gamma}}{1-\gamma} \int_0^\infty e^{-\rho t} \kappa_t G_t^{1-\gamma} \left( \Phi(\bar{d}_t^*) + \int_0^\infty \left(1 + \alpha \left(e^{\sigma\widetilde{\pi}^B \sqrt{t}y} - 1\right)\right)^{1-\gamma} \phi(y + \bar{d}_t^*) dy \right) dt, \\ EU^{*OT} &= \frac{n^{1-\gamma}}{1-\gamma} \int_0^\infty e^{-\rho t} \kappa_t (d_t^*)^{1-\gamma} dt, \\ EU^{*NT} &= \frac{n^{1-\gamma}}{1-\gamma} \int_0^\infty e^{-\rho t} \kappa_t ({}_t\tilde{p}_x M_e(-\ln({}_t\tilde{p}_x)) d_t^*)^{1-\gamma} dt, \end{aligned}$$

where  $\bar{d}_t^*$  is given as in (24), but with  $\pi$  replaced by  $\widetilde{\pi}^B$  and  $V_0$  replaced by  $\widetilde{V}_0^{*B}$ .

### Notes

<sup>1</sup> For simplicity, we have assumed log-normal risky asset dynamics, which, as well documented, may not be very realistic. It would be interesting to look at the unit-linked tontine design problem in more general settings where the asset volatility is random when fat-tailed returns and volatility clustering are taken into account (see, e.g., Cont and Tankov 2004; Fouque et al. 2000). The continuity assumption of the stock price is relaxed in order to capture sudden and unpredictable market changes (see, e.g., Cont and Tankov 2004). Also, for such long-term investment problems, it would be more realistic to incorporate interest rate fluctuations (see, e.g., Hull and White 1990; Vasicek 1977).

- 2 This assumption is actually fulfilled for our specific choices of  ${}_t p_x$  and  ${}_t \bar{p}_x$  in our numerical analysis, where they are given in (31) in Section 5.1.
- 3 Further properties of and a detailed discussion about the power utility function family can be found in, e.g., Wakker (2008).
- 4 In general, it is true that an optimal value for  $V_0$  can theoretically become arbitrarily large, which would not be feasible in reality. However, due to the budget constraint, this can be prevented and, thus, choosing  $V_0$  as a decision variable is reasonable. For example, the optimal value for  $V_0$  in Case A given in (22) is high only when it is justified, namely if the initial wealth spent by the individual is large or if her survival probability is low, for instance.
- 5 It should be pointed out that  $G_t$  and  $\alpha$  can theoretically also serve as decision variables. A practicable option for  $G_t$  is presented and discussed in Section 5.1.
- 6 We remark that a baseline  $\gamma > 1$  or another type of utility function may lead to different conclusions.
- 7 The discounted expected utility for the natural traditional tontine diverges for too high values of  $\gamma$ , i.e.,  $EU^{*NT}$  goes to minus infinity if  $\gamma \gg 1$ . Consequently, we do not consult the natural traditional tontine if  $\gamma$  attains rather large values.
- 8 If the two required values  $\pi^{*B}$  and  $V_0^{*B}$  for Case B introduced in Section 4.2 cannot be uniquely determined for the numerical outcomes in this section, this is adequately reported in the related paragraphs hereinafter.
- 9 Due to the divergence of the discounted expected utility for the natural traditional tontine as soon as  $\gamma$  gets too large, we show, in the case, where  $\gamma$  varies,  $CE^{*NT}$  only as long as  $\gamma$  ranges within  $(0, 1)$ . Further note that as long as we assess the sensitivity towards a parameter in any analysis in this section, all the remaining parameters attain their baseline values, if not stated otherwise.
- 10 In detail, the applied total payments in Figure 3 are determined, for Case A, by  $\frac{nV_t^*}{N_t}$ , where  $V_t^* = V_0^{*A} e^{rt + (\mu - r)\pi^{*A}t - \frac{\sigma^2}{2}(\pi^{*A})^2 t + \sigma\pi^{*A}W_t}$ , for Case B, by  $\frac{n(G_t + \alpha(V_t^{*+} - G_t)^+)}{N_t}$ , where  $V_t^* = V_0^{*B} e^{rt + (\mu - r)\pi^{*B}t - \frac{\sigma^2}{2}(\pi^{*B})^2 t + \sigma\pi^{*B}W_t}$ , and, for the optimal traditional tontine, by  $\frac{nd_t^*}{N_t}$ . Note that the computation of all depicted quantities is done numerically, where we divide the relevant time line running from  $t = 0$  to  $t = 35$  by a constant discretization step size of 0.025, which means that we overall analyze 1401 points, and simulate each occurring random variable 450,000 times.

## References

- Aase, Knut K., and Svein-Arne Persson. 1994. Pricing of Unit-Linked Life Insurance Policies. *Scandinavian Actuarial Journal* 1994: 26–52. [CrossRef]
- Bacinello, Anna Rita, Pietro Millosovich, and An Chen. 2018. The Impact of Longevity and Investment Risk on a Portfolio of Life Insurance Liabilities. *European Actuarial Journal* 8: 257–90. [CrossRef]
- Berkelaar, Arjan B., Roy Kouwenberg, and Thierry Post. 2004. Optimal Portfolio Choice under Loss Aversion. *Review of Economics and Statistics* 86: 973–87. [CrossRef]
- Bernhardt, Thomas, and Catherine Donnelly. 2019. Modern Tontine with Bequest: Innovation in Pooled Annuity Products. *Insurance: Mathematics and Economics* 86: 168–88. [CrossRef]
- Bertsekas, Dimitri P. 2014. *Constrained Optimization and Lagrange Multiplier Methods*. Cambridge: Academic Press.
- Black, Fischer, and Myron Scholes. 1973. The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy* 81: 637–54. [CrossRef]
- Briys, Eric, and François de Varenne. 1994. Life Insurance in a Contingent Claim Framework: Pricing and Regulatory Implications. *The Geneva Papers on Risk and Insurance Theory* 19: 53–72. [CrossRef]
- Chen, An, and Manuel Rach. 2019. Options on Tontines: An Innovative Way of Combining Tontines and Annuities. *Insurance: Mathematics and Economics* 89: 182–92.
- Chen, An, Peter Hieber, and Jakob Klein. 2019. Tonuity: A Novel Individual-Oriented Retirement Plan. *ASTIN Bulletin* 49: 5–30. [CrossRef]
- Chen, An, Peter Hieber, and Manuel Rach. 2021. Optimal Retirement Products under Subjective Mortality Beliefs. *Insurance: Mathematics and Economics* 101: 55–69. [CrossRef]
- Chen, An, Linyi Qian, and Zhixin Yang. 2020. Tontines with Mixed Cohorts. *Scandinavian Actuarial Journal* 2021: 437–55. [CrossRef]
- Chen, An, Manuel Rach, and Thorsten Sehner. 2020. On the Optimal Combination of Annuities and Tontines. *ASTIN Bulletin* 50: 95–129. [CrossRef]
- Cont, Rama, and Peter Tankov. 2004. *Financial Modelling with Jump Processes*. CRC Mathematics Series. Chapel Hill: Chapman & Hall.
- Dahl, Mikkel, Martin Melchior, and Thomas Møller. 2008. On Systematic Mortality Risk and Risk-Minimization with Survivor Swaps. *Scandinavian Actuarial Journal* 2008: 114–46. [CrossRef]
- Denuit, Michel. 2019. Size-Biased Transform and Conditional Mean Risk Sharing, with Application to P2P Insurance and Tontines. *ASTIN Bulletin* 49: 591–617. [CrossRef]
- Dhaene, Jan, Alexander Kukush, Elisa Luciano, Wim Schoutens, and Ben Stassen. 2013. On the (In-)Dependence between Financial and Actuarial Risks. *Insurance: Mathematics and Economics* 52: 522–31. [CrossRef]
- Donnelly, Catherine, Montserrat Guillén, and Jens Perch Nielsen. 2014. Bringing Cost Transparency to the Life Annuity Market. *Insurance: Mathematics and Economics* 56: 14–27. [CrossRef]

- Ekern, Steinar, and Svein-Arne Persson. 1996. Exotic Unit-Linked Life Insurance Contracts. *The Geneva Papers on Risk and Insurance Theory* 21: 35–63. [CrossRef]
- Forman, Jonathan Barry, and Michael J. Sabin. 2015. Tontine Pensions. *University of Pennsylvania Law Review* 163: 755–831.
- Fouque, Jean-Pierre, George Papanicolaou, and K. Ronnie Sircar. 2000. *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge: Cambridge School Press.
- Gatzert, Nadine, Carin Huber, and Hato Schmeiser. 2011. On the Valuation of Investment Guarantees in Unit-Linked Life Insurance: A Customer Perspective. *The Geneva Papers on Risk and Insurance—Issues and Practice* 36: 3–29. [CrossRef]
- Gemmo, Irina, Ralph Rogalla, and Jan-Hendrik Weinert. 2020. Optimal Portfolio Choice with Tontines under Systematic Longevity Risk. *Annals of Actuarial Science* 14: 302–15. [CrossRef]
- Giulietti, Corrado, Mirco Tonin, and Michael Vlassopoulos. 2020. When the Market Drives You Crazy: Stock Market Returns and Fatal Car Accidents. *Journal of Health Economics* 70: 102245. [CrossRef]
- Gothaer. 2021. Fondsgebundene Rentenversicherung mit Garantien. Gothaer Lebensversicherung AG. Available online: <https://www.gothaer.de/privatkunden/rentenversicherung/fondsgebundene-rentenversicherung/> (accessed on 15 February 2021).
- Hu, Wei-Yin, and Jason S. Scott. 2007. Behavioral Obstacles in the Annuity Market. *Financial Analysts Journal* 63: 71–82. [CrossRef]
- Hull, John, and Alan White. 1990. Pricing Interest-Rate-Derivative Securities. *The Review of Financial Studies* 3: 573–92. [CrossRef]
- Kahneman, Daniel, and Amos Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica: Journal of the Econometric Society* 47: 263–92. [CrossRef]
- Ledlie, M. Collin, Dermot P. Corry, Gary S. Finkelstein, A. J. Ritchie, K. Su, and D. C. E. Wilson. 2008. Variable Annuities. *British Actuarial Journal* 14: 327–89. [CrossRef]
- Leland, Hayne E. 1985. Option Pricing and Replication with Transactions Costs. *The Journal of Finance* 40: 1283–301. [CrossRef]
- Levy, Haim. 1994. Absolute and Relative Risk Aversion: An Experimental Study. *Journal of Risk and Uncertainty* 8: 289–307. [CrossRef]
- Lin, Yijia, and Samuel H. Cox. 2005. Securitization of Mortality Risks in Life Annuities. *Journal of Risk and Insurance* 72: 227–52. [CrossRef]
- Margaras, Vasilis. 2019. Demographic Trends in EU Regions. Available online: <https://ec.europa.eu/futurium/en/system/files/ged/eprs-briefing-633160-demographic-trends-eu-regions-final.pdf> (accessed on 15 December 2021).
- Merton, Robert C. 1969. Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics* 51: 247–57. [CrossRef]
- Milevsky, Moshe A. 2015. *King William's Tontine: Why the Retirement Annuity of the Future Should Resemble its Past*. Cambridge: Cambridge University Press.
- Milevsky, Moshe A. 2020. Calibrating Gompertz in Reverse: What is Your Longevity-Risk-Adjusted Global Age? *Insurance: Mathematics and Economics* 92: 147–61. [CrossRef]
- Milevsky, Moshe A., and Thomas S. Salisbury. 2015. Optimal Retirement Income Tontines. *Insurance: Mathematics and Economics* 64: 91–105. [CrossRef]
- Milevsky, Moshe A., and Thomas S. Salisbury. 2016. Equitable Retirement Income Tontines: Mixing Cohorts without Discriminating. *ASTIN Bulletin* 46: 571–604. [CrossRef]
- Milevsky, Moshe A., Thomas S. Salisbury, Gabriela Gonzalez, and Hanna Jankowski. 2018. Annuities Versus Tontines in the 21st Century: A Canadian Case Study. Available online: <https://www.soa.org/globalassets/assets/files/resources/research-report/2018/annuities-vs-tontines.pdf> (accessed on 15 December 2021).
- Mitchell, Olivia S. 2002. Developments in Decumulation: The Role of Annuity Products in Financing Retirement. In *Ageing, Financial Markets and Monetary Policy*. Edited by Alan J. Auerbach and Heinz Herrmann. Berlin: Springer, chp. 3, pp. 97–125.
- Møller, Thomas. 1998. Risk-Minimizing Hedging Strategies for Unit-Linked Life Insurance Contracts. *ASTIN Bulletin* 28: 17–47. [CrossRef]
- Peijnenburg, Kim, Theo Nijman, and Bas J. M. Werker. 2016. The Annuity Puzzle Remains a Puzzle. *Journal of Economic Dynamics and Control* 70: 18–35. [CrossRef]
- Ramsay, Colin M., and Victor I. Oguledo. 2018. The Annuity Puzzle and an Outline of Its Solution. *North American Actuarial Journal* 22: 623–45. [CrossRef]
- Royal London. 2018. *Will we Ever Summit the Pension Mountain?* Royal London Policy Paper No. 21. London: Royal London Mutual Insurance Society Limited.
- Sabin, Michael J. 2010. Fair Tontine Annuity. Social Science Research Network. Available online: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1579932](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1579932) (accessed on 26 February 2021).
- Schiereck, Dirk, Jochen Ruß, Rolf Tilmes, and Torsten Haupt, eds. 2020. *Ruhestandsplanung—Beratungsansatz für die Zielgruppe 50plus*. Wiesbaden: Springer.
- Schrager, David F., and Antoon A. J. Pelsser. 2004. Pricing Rate of Return Guarantees in Regular Premium Unit Linked Insurance. *Insurance: Mathematics and Economics* 35: 369–98. [CrossRef]
- Schwandt, Hannes. 2018. Wealth Shocks and Health Outcomes: Evidence from Stock Market fluctuations. *American Economic Journal: Applied Economics* 10: 349–77. [CrossRef]
- Sehner, Thorsten. 2021. Life and Long-Term Care Insurance: Valuation and Product Design. Ph.D. dissertation, October 2021, Ulm University of Applied Sciences, Ulm, Germany [CrossRef]

- Sharpe, William Forsyth. 2017. Retirement Income Analysis with Scenario Matrices. Leland Stanford Junior University. Available online: <https://web.stanford.edu/~wfsarpe/RISMAT/RIAbok.pdf> (accessed on 3 March 2021).
- Smith, Elliot. 2020. Short-Selling Bans Sweep Europe in the Hope of Stemming Stock Market Bleeding. CNBC. Available online: <https://www.cnbc.com/2020/03/17/short-selling-bans-sweep-europe-in-hope-of-stemming-stock-market-bleeding.html> (accessed on 13 October 2021).
- Statista. 2020a. Average Risk Free Investment Rate in the United Kingdom (UK) 2015–2020. Statista GmbH. Available online: <https://www.statista.com/statistics/885750/average-risk-free-rate-united-kingdom/> (accessed on 16 December 2020).
- Statista. 2020b. Distribution of Direct Life Insurance and Reinsurance Gross Written Premiums in the United Kingdom (UK) in 2019, by Lines of Business. Statista GmbH. Available online: <https://www.statista.com/statistics/1126649/direct-and-reinsurance-life-premiums-distribution-united-kingdom-uk/> (accessed on 25 February 2021).
- Thomas, Philip J. 2016. Measuring Risk-Aversion: The Challenge. *Measurement* 79: 285–301. [CrossRef]
- Thomas, Philip J., Roger D. Jones, and William J. O. Boyle. 2010. The Limits to Risk Aversion: Part 1. The Point of Indiscriminate Decision. *Process Safety and Environmental Protection* 88: 381–95. [CrossRef]
- Vasicek, Oldrich. 1977. An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics* 5: 177–88. [CrossRef]
- Waddington, Ian, William J. O. Boyle, and James Kearns. 2013. Computing the Limits of Risk Aversion. *Process Safety and Environmental Protection* 91: 92–100. [CrossRef]
- Wakker, Peter P. 2008. Explaining the Characteristics of the Power (CRRA) Utility Family. *Health Economics* 17: 1329–44. [CrossRef] [PubMed]
- Weinert, Jan-Hendrik, and Helmut Gründl. 2020. The Modern Tontine. *European Actuarial Journal* 11: 49–86 [CrossRef]
- Winter, Pascal, and Frédéric Planchet. 2021. Modern Tontines as a Pension Solution: A Practical Overview. Available online: [https://www.researchgate.net/publication/355260987\\_Modern\\_tontines\\_as\\_a\\_pension\\_solution\\_a\\_practical\\_overview](https://www.researchgate.net/publication/355260987_Modern_tontines_as_a_pension_solution_a_practical_overview) (accessed on 15 December 2021).
- Yaari, Menahem E. 1965. Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *The Review of Economic Studies* 32: 137–50. [CrossRef]



# Equivalent Risk Indicators: VaR, TCE, and Beyond

Silvia Faroni <sup>1,2</sup> , Olivier Le Courtois <sup>1,\*</sup> and Krzysztof Ostaszewski <sup>3</sup> 

<sup>1</sup> EMLyon Business School, 23, Avenue Guy de Collongue, CEDEX, 69134 Ecully, France; faroni@em-lyon.com

<sup>2</sup> COACTIS (EA4161), Université de Lyon/Lyon 2, ISH, 14-16 Avenue Berthelot, 69007 Lyon, France

<sup>3</sup> College of Arts and Science, Illinois State University (ISU), Normal, IL 61790-4520, USA; krzysio@ilstu.edu

\* Correspondence: lecourtois@em-lyon.com or olecourt@yahoo.fr

**Abstract:** While a lot of research concentrates on the respective merits of VaR and TCE, which are the two most classic risk indicators used by financial institutions, little has been written on the equivalence between such indicators. Further, TCE, despite its merits, may not be the most accurate indicator to take into account the nature of probability distribution tails. In this paper, we introduce a new risk indicator that extends TCE to take into account higher-order risks. We compare the quantiles of this indicator to the quantiles of VaR in a simple Pareto framework, and then in a generalized Pareto framework. We also examine equivalence results between the quantiles of high-order TCEs.

**Keywords:** VaR; TCE; extended TCE; insurance regulation; risk measurement

## 1. Introduction

In the second half of the twentieth century, developed market economies had undergone an inflationary period after World War II, followed by a disinflationary period that started in the early 1980s accompanied by less regulation and greater reliance on market forces. These developments resulted in improved economic performance, but also increased volatility and market pressure on financial institutions following traditional fixed capital standards. In order to address these developments, regulators have moved towards risk-based capital requirements for financial institutions, while also allowing gradual relaxation of certain rules of governance of financial institutions and replacing them with a principle-based approach.

The risk measures imposed by regulators on insurance companies share a common feature: they are all related to the behavior of tails of the probability distribution of a firm's financial results. The reason is that the regulatory purpose of capital requirements is to make capital available for absorbing losses occurring in extreme events, i.e., events of large financial losses, which could bring about insolvency. Risk-based capital requirements are a natural outgrowth of traditional prudential regulation aiming at preserving the solvency of private financial institutions.

Hence, safety capital should be computed by looking at the extreme risks that can impact financial institutions in general and insurance enterprises in particular. However, regulators differ in their specification of the tail indicator that they recommend. Some regulators impose the use of quantiles of the distribution (value at risk, or VaR), while other regulators impose the use of partial moments (tail conditional expectation or TCE, or an equivalent measure of expected shortfall). Furthermore, regulators also differ in their choice of time horizon. Finally, they also differ in their choice of confidence level (i.e., probability value for the quantile, or for the conditional expectation).

Comité Européen des Assurances and Mercer Oliver Wyman Limited (2005) and CEA (2007) provide a comparison of different regulatory regimes for capital requirements. The most consequential regulation of risk-based capital is the European Union's Solvency II. Solvency II imposes risk-based capital requirements computed using the VaR as a risk measure over a one-year period and with a confidence level of 99.5%. When Solvency II

**Citation:** Faroni, Silvia, Olivier Le Courtois, and Krzysztof Ostaszewski. 2022. Equivalent Risk Indicators: VaR, TCE, and Beyond. *Risks* 10: 142. <https://doi.org/10.3390/risks10080142>

Academic Editors: Ermanno Pitacco and Annamaria Olivieri

Received: 31 May 2022

Accepted: 18 July 2022

Published: 22 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).



was being designed, it was to some degree modeled on the Basel II banking regulation, also using VaR as a risk measure for capital requirement purposes, for market risk. The credit crisis of 2008 in many ways exposed the weaknesses of VaR, and a new system of capital regulation for banking, Basel III (see Basel Committee on Banking Supervision's documents from 2019 and 2022), has been developed and implemented since (see also Gatzert and Wesker (2011) for a comparison of Solvency II and Basel regulations).

Solvency II is not only the main regulatory law for prudential regulation of insurance enterprises in the European Union, but it has become a model for risk-based capital requirements worldwide. The United States is one major exception to this trend, as the U.S. model regulation of insurance firms' capital preceded Solvency II, and it is designed around a formula provided by the regulatory body, the National Association of Insurance Commissioners (NAIC) (see also National Association of Insurance Commissioners (2007) on ORSA perspectives).

The Canadian regulation of insurance capital differs from that of the European Union, and it is based on risk assessment of the firm in the context of certain extreme events. Canada's Office of Supervision of Financial Institutions (OSFI) risk assessment process begins with an evaluation of the inherent risk within each significant activity of an insurer and the quality of risk management applied to mitigate these risks (see Canada Office of Supervision of Financial Institutions 2022). After considering this information, OSFI determines the level of net risk and direction (i.e., whether it is decreasing, stable, or increasing) of the rating for each significant activity. The net risks of the significant activities are combined, by considering their relative importance, to arrive at the overall net risk (ONR) of the insurer. Furthermore, OSFI provides capital requirement guidelines, which must be then included in insurers' own risk and solvency assessment.

In the cases of both the United States and Canada, we see a significant regulatory involvement in the supervision of risk-based capital. All regulators are, of course, involved in this process, but the approach of the European Union and, notably, Switzerland, is more principle-based than rule-driven. The Swiss regulation of risk-based capital for insurers includes a capital standard, and stress-testing of certain extreme scenarios. The capital standard is based on the expected shortfall, or, equivalently, tail conditional expectation (TCE).

In this work, we focus on two regulations: Solvency II and the Swiss Solvency Test (SST). The Swiss Solvency Test, implemented in 2004, preceded Solvency II, but in 2015 the European Union recognized the SST as the first regime to be fully equivalent to Solvency II. Solvency II imposes a capital requirement computed using value at risk (VaR) as a risk measure over a 1-year period and with a confidence level of 99.5%, whereas SST uses TCE with a confidence level of 99% over a 1-year period.

VaR and TCE are the most classic examples of risk measures (see for instance Linsmeier and Pearson 2000; Klugman et al. 2012; Acerbi et al. 2001). A risk measure is a mapping from the random variable representing risk exposure to the set of real numbers. It can be interpreted as the amount of capital required to protect against adverse outcomes of a given risk. The paper by Artzner et al. (1999) introduced the concept of coherence of risk measures and has been very influential in the further development of risk measurement. A coherent risk measure is defined by the following four properties: subadditivity, monotonicity, positive homogeneity, and translation invariance. The properties of risk measures in the context of insurance are discussed by Wang and Zitikis (2021). Acerbi and Tasche (2002) noted that TCE is a coherent risk measure, while VaR is not (see also Society of Actuaries (2000) on TCE).

Rostek (2010) provides an interesting alternative to evaluations of risk, a model of preferences, in which, given beliefs about uncertain outcomes, an individual evaluates an action by a quantile of the induced distribution. Fadina et al. (2021) designed a unified axiomatic framework for risk evaluation principles, which quantifies jointly a loss random variable and a set of plausible probabilities. They called such an evaluation principle a generalized risk measure.

Fuchs et al. (2017) show that a notion of a quantile risk measure is a natural generalization of that of a spectral risk measure and provides another view of the distortion risk measures generated by a distribution function on the unit interval. In this general setting, they prove several results on quantile or spectral risk measures and their domain with special consideration of the expected shortfall. They also present a particularly short proof of the subadditivity of the expected shortfall risk measure. Denuit et al. (2006) provide a comprehensive review or modeling risk in incomplete markets, with emphasis on insurance risks, expanding on and combining in a comprehensive review the existing literature on quantitative risk management.

Li and Wang (2022) noted that the Basel Committee on Banking Supervision proposed the shift from the 99% value at risk (VaR) to the 97.5% expected shortfall (ES) for internal models in market risk assessment (see Basel Committee on Banking Supervision 2019, 2022). Inspired by that development, Li and Wang introduced a new distributional index, the probability equivalence level of VaR and ES (PELVE), which identifies the balancing point for the equivalence between VaR and ES. PELVE has desirable theoretical properties, and it distinguishes empirically heavy-tailed distributions from light-tailed ones.

Barczy et al. (2022) generalized and further developed the PELVE measure and applied this indicator to a high-order TCE, which is distinct from the extended TCE introduced in the present paper. Fiori and Rosazza Gianin (2021) construct another generalization by constructing an indicator that relates monotone risk measures.

Our paper was developed independently from this stream of papers and develops related comparisons of VaR and TCE quantiles, but without introducing an intermediate indicator such as PELVE. A key contribution of our paper is the introduction of a new risk indicator that extends TCE to take into account higher-order risks. We also provide comparisons of this new indicator with more classic risk indicators.

More deeply, the goal of our paper is to understand how regulators choose such or such risk indicators for solvency computations. Thus, our goal is to understand the implicit utility function of insurance regulators and to understand how equivalent risk valuation systems can be put in place, but we do so without introducing any type of utility function. It is, of course, entirely possible that the regulations put in place are not fully a result of a certain intent, but rather of a political process so that what we determine may not be the output of an actual utility function of a specific regulator. However, making the implicit functioning of actual regulations explicit should be a valuable contribution in assuring that regulations function in an effective and efficient manner.

The paper makes use of two probabilistic assumptions, where we assume that claims can be either Pareto or generalized Pareto (GPD) distributed. We make these assumptions for two main reasons: they allow us to very conveniently derive readable results and they are consistent with what is observed for claims with heavily distributed tails. The reader interested in exploring situations where claims could be associated with semi-heavy tails can be referred for instance to Le Courtois (2018) or Le Courtois and Walter (2014).

The paper is organized as follows. In Section 2, we provide a general overview of the two main risk measures, VaR and TCE, and briefly discuss equivalence results that relate to the confidence levels of VaR and TCE. In Section 3, we introduce a new and generalized tail indicator, and we discuss the relation between the quantiles of this indicator and the quantiles of VaR, in a simple Pareto framework, and then in a generalized Pareto framework. Section 4 examines equivalence results between the quantiles of generalized indicators.

## 2. VaR and TCE

In this section, we recall key preparatory elements on VaR and TCE. We first recall classic definitions. Then, we examine the link between VaR and TCE. We conclude with an illustration of this link. These elements will be useful in the next section, where the core contribution of the paper is developed.

### 2.1. Definitions

The most used risk measure is value at risk (VaR), which is the expected worst loss over a given horizon at a given confidence level. Linsmeier and Pearson (2000) define VaR as the loss that is expected to be exceeded with a probability of only  $(1 - \alpha)$  percent during the next holding period  $T$ . The role of regulators is to choose the value of the confidence level  $\alpha$  and of the horizon  $T$ .

From a mathematical viewpoint, suppose that  $F_{X_T}(x)$  represents the distribution function of outcomes over a fixed period of time  $T$  of a portfolio of risks. An adverse outcome is a loss and, in this case, positive values of the random variable  $X_T$  are losses. The VaR of the random variable is the  $\alpha$  percentile of the distribution of  $X_T$ , denoted by

$$\text{VaR}_\alpha(X_T) = F_{X_T}^{-1}(\alpha).$$

According to the report of the National Association of Insurance Commissioners (2007), tail conditional expectation (TCE)—or conditional tail expectation (CTE)—measures the amount of risk within the tail of a distribution of outcomes, expressed as the probability-weighted average of the outcomes beyond a chosen point in the distribution.

In the report produced by the CEA (2007), TCE measures the average losses over the defined threshold (typically set as the VaR at a given confidence level  $\alpha$ ). In other words, TCE is a conditional mean value, given that the loss exceeds the  $(1 - \alpha)$  percentile. It is also often called tail value at risk (TVaR) or expected shortfall (ES). A broader analysis of TCE and its properties can be found in Society of Actuaries (2000). From a mathematical viewpoint, TCE is defined as follows:

$$\text{TCE}_\alpha(X_T) = \mathbb{E}[X_T \mid X_T \geq \text{VaR}_\alpha(X_T)].$$

Furthermore, if the random variable is continuous, we can write:

$$\text{TCE}_\alpha(X_T) = \frac{1}{1 - \alpha} \int_{\alpha}^1 \text{VaR}_u(X_T) du.$$

### 2.2. Relation between VaR and TCE Quantiles

If we compute VaR and TCE using the same quantile, TCE will be always higher than VaR, by construction. However, the quantiles for VaR and TCE are usually chosen to be different by regulators. We study which relation should exist between these two quantiles. Specifically, we examine how it is possible to find  $c$  and  $q$  such that  $\text{VaR}_q = \text{TCE}_c$ , where  $q > c$ . We conduct our analysis when risks follow a generalized Pareto distribution.

Let  $X_T$  be a random variable that follows a generalized Pareto distribution with three parameters: location  $\mu$ , scale  $\sigma$ , and shape  $\zeta$ . The cumulative distribution function of  $X_T$  admits the Jenkinson–von Mises representation, which can be expressed as follows:

$$F(X_T) = 1 - \left(1 + \frac{\zeta(x - \mu)}{\sigma}\right)^{-\frac{1}{\zeta}}, \tag{1}$$

for  $\zeta \neq 0$ .

In this situation, we can show that the VaR can be computed as follows:

$$\text{VaR}_q(X_T) = \mu + \left((1 - q)^{-\zeta} - 1\right) \cdot \frac{\sigma}{\zeta}, \tag{2}$$

while the tail conditional expectation admits the following expression:

$$\text{TCE}_c(X_T) = \mu + \frac{\sigma}{\zeta} \left(\frac{(1 - c)^{-\zeta}}{1 - \zeta} - 1\right).$$

To solve  $\text{VaR}_q(X_T) = \text{TCE}_c(X_T)$ , we first rewrite  $\text{TCE}_c(X_T)$  as a function of  $\text{VaR}_c(X_T)$ :

$$\text{TCE}_c(X_T) = \frac{1}{1 - \zeta} \cdot \text{VaR}_c(X_T) + \frac{\sigma - \zeta\mu}{1 - \zeta}.$$

Thus,  $\text{TCE}_c(X_T) = \text{VaR}_q(X_T)$  is equivalent to

$$\frac{1}{1 - \zeta} \cdot \text{VaR}_c(X_T) + \frac{\sigma - \zeta\mu}{1 - \zeta} - \text{VaR}_q(X_T) = 0. \tag{3}$$

Using the above equality, we can relate the quantities  $c$  and  $q$  as follows.

**Theorem 1.** *In the generalized Pareto framework, the quantile of TCE and the quantile of VaR obey the following relationship when the two risk indicators are equal:*

$$c = 1 - (1 - \zeta)^{-\frac{1}{\zeta}} \cdot (1 - q) \tag{4}$$

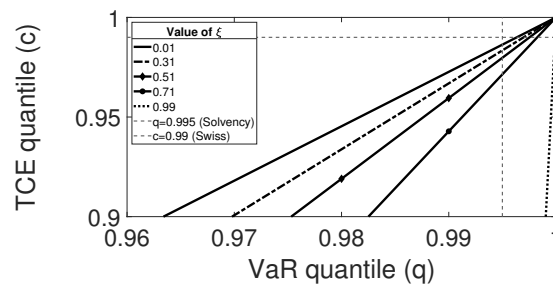
where  $0 < \zeta < 1$ .

**Proof.** See Appendix A.  $\square$

Note that Equation (4) depends on  $\zeta$ , but not on  $\mu$  or  $\sigma$ . Also note that if  $\zeta = \frac{1}{a}$ , the result obtained using the generalized Pareto distribution boils down to an identical result in the subcase of Pareto Type I distributions.

### 2.3. Illustration

Next, we illustrate Theorem 1 in Figure 1, where we plot the TCE quantile as a function of its equivalent VaR quantile. We plot this relation for different values of the market risk parameter  $\zeta$ , where we let  $\zeta$  take values between 0.01 and 0.99. We recall that a higher value of  $\zeta$  is equivalent to an increased presence of extreme risks in the phenomenon under study.



**Figure 1.** TCE quantile as a function of VaR quantile.

We can see that when there is more risk on the market, that is, when  $\zeta \rightarrow 1$ , the VaR quantile has to be a number close to 1, whereas the TCE quantile can take a broad range of values between 0.9 and 1.

This feature implies that in the presence of a lot of extreme risks, there is a large variability in the choice of the TCE quantile, making it a difficult value to choose by regulators. Conversely, in the presence of a lot of extreme risks, the VaR quantile is easy to set, where the regulator only needs to choose a sufficiently high value, as is observed in the case of the Solvency II regulation.

However, this feature is not observed when  $\zeta$  is small, where a lot of admissible values can be taken by both the VaR and TCE quantiles, and where these two quantiles vary linearly.

Figure 1 also illustrates that a lot of the solutions to Theorem 1 are consistent with the Solvency II and the Swiss insurance regulations. However, the figure also illustrates that the two regulations are inconsistent. Indeed, it is practically impossible to find a pair of reasonable values of the VaR and TCE quantiles that is consistent with both regulations.

### 3. A New High-Order TCE Indicator

In this section, we introduce a new generalized TCE indicator, which is a conditional higher-order moment of the probability distribution under study. We compute this indicator when losses follow a Type I Pareto distribution and we derive equivalence relations with value at risk. We also conduct a similar study when losses follow a generalized Pareto distribution.

#### 3.1. Definition

By analogy with higher-order moments, which are key characteristics of probability distributions, we construct a higher-order measure of risk, which is a TCE at order  $m$ . We denote this indicator by  $TCE_c^{(m)}$  and we define it as follows:

$$TCE_c^{(m)}(X_T) = \mathbb{E}[X_T^m \mid X_T \geq VaR_c(X_T)]. \tag{5}$$

As an illustration,  $TCE_c^{(2)}$  is a conditional non-central second-order moment, where the condition is that losses exceed the  $1 - c$  percentile. Further,  $TCE_c^{(1)}$  is the standard tail conditional expectation indicator.

Because

$$\mathbb{E}[X_T^m \mid X_T \geq VaR_c(X_T)] = \frac{\mathbb{E}\left[X_T^m \mathbb{1}_{X_T \geq VaR_c(X_T)}\right]}{Pr(X_T \geq VaR_c(X_T))},$$

we can rewrite the extended TCE indicator as follows:

$$TCE_c^{(m)}(X_T) = \frac{1}{1 - c} \mathbb{E}\left[X_T^m \mathbb{1}_{X_T \geq VaR_c(X_T)}\right],$$

so that

$$TCE_c^{(m)}(X_T) = \frac{1}{1 - c} \int_{VaR_c(X_T)}^{+\infty} x^m dF(x), \tag{6}$$

where  $c = F(VaR_c(X_T))$ .

Let us change variables as follows:  $F(x) = s$ ,  $x = F^{-1}(s) = VaR_s(X_T)$ , and  $ds = dF(x)$ . We readily obtain a third equivalent representation of the extended TCE indicator:

$$TCE_c^{(m)}(X_T) = \frac{1}{1 - c} \int_c^1 (VaR_s(X_T))^m ds. \tag{7}$$

Note that another extended TCE indicator  $\Xi^{(m)}$  can be found in the risk management literature (see for instance Barczy et al. (2022)). This indicator is defined by:

$$\Xi^{(m)} = \frac{m}{1 - c} \int_c^1 \left(\frac{s - c}{1 - c}\right)^{m-1} VaR_s(X_T) ds = \frac{m}{1 - c} \int_c^1 \left(\frac{s - c}{1 - c}\right)^{m-1} F^{-1}(s) ds.$$

If we again change variables as follows:  $F(x) = s$ ,  $x = F^{-1}(s) = VaR_s(X_T)$ , and  $ds = dF(x)$ , we obtain:

$$\Xi^{(m)} = \frac{m}{1 - c} \int_{VaR_c(X_T)}^{+\infty} \left(\frac{F(x) - c}{1 - c}\right)^{m-1} x dF(x).$$

All of these expressions are distinct from Equations (5)–(7) and confirm that  $\Xi^{(m)}$  cannot be interpreted as a conditional higher-order moment, contrary to the indicator examined in this paper.

### 3.2. Pareto Distributed Losses

Let us now assume that losses follow a Type I Pareto distribution, whose probability density function is represented by:

$$f_{X_T}(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, \tag{8}$$

where  $\alpha$  is the shape parameter and  $\theta$  the scale parameter.

We apply the definition in Equation (6) to compute  $TCE_c^{(m)}$  when losses follow a Type I Pareto distribution. We write:

$$TCE_c^{(m)}(X_T) = \frac{1}{1-c} \int_{\text{VaR}_c(X_T)}^{+\infty} x^m \frac{\alpha\theta^\alpha}{x^{\alpha+1}} dx.$$

This expression can be developed in closed form as a function of value at risk, as shown in the next theorem.

**Theorem 2.** *When losses are Pareto distributed, the extended TCE indicator admits the following expression.*

$$TCE_c^{(m)}(X_T) = \frac{\alpha}{\alpha - m} (\text{VaR}_c(X_T))^m \tag{9}$$

when  $\alpha > m$ .

**Proof.** See Appendix A.  $\square$

From this theorem, we deduce that  $\text{VaR}_q(X_T) = TCE_c(X_T)$  is equivalent to

$$\frac{\alpha}{\alpha - m} \cdot (\text{VaR}_c(X_T))^m - \text{VaR}_q(X_T) = 0. \tag{10}$$

This equation allows us to find the relation between the extended TCE and the VaR quantiles. Indeed, we obtain:

**Theorem 3.** *In the Pareto framework, the extended TCE quantile ( $c$ ) and the VaR quantile ( $q$ ) obey the following relationship when the two risk indicators are equal:*

$$c = 1 - \left( \frac{(\alpha - m) \theta^{1-m}}{\alpha} \right)^{-\frac{\alpha}{m}} (1 - q)^{\frac{1}{m}}. \tag{11}$$

**Proof.** See Appendix A.  $\square$

For consistency with the generalized Pareto approach, we replace the shape parameter  $\alpha$  of the classic Pareto distribution with  $\zeta = \frac{1}{\alpha}$ , which can also be interpreted as a shape parameter.

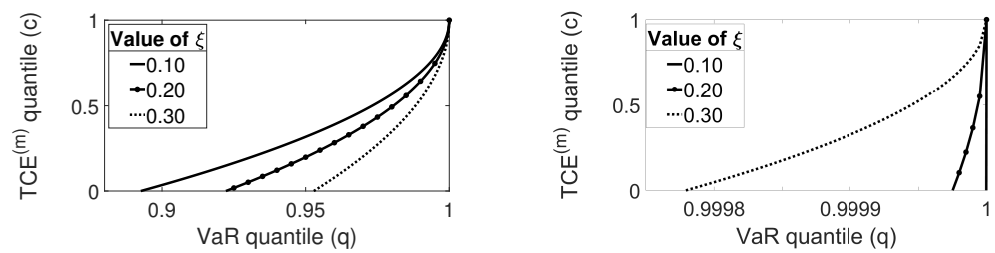
Thus, Equation (11) can be rewritten as follows:

$$c = 1 - \left( (1 - m\zeta) \theta^{1-m} \right)^{-\frac{1}{m\zeta}} (1 - q)^{\frac{1}{m}}, \tag{12}$$

where  $0 < \zeta < \frac{1}{m}$ .

It appears that Equation (12) generalizes the result of Theorem 1. Indeed, Equation (12) reduces to Equation (4) when  $m = 1$ .

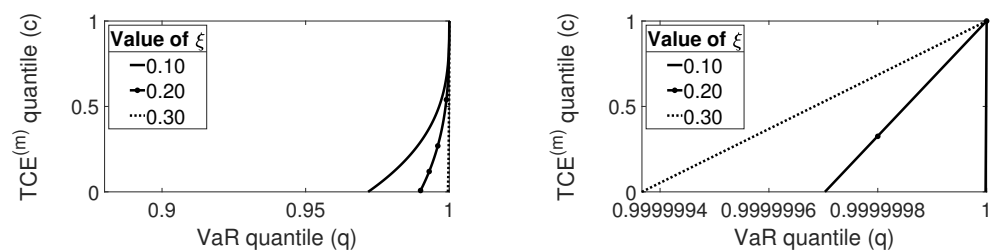
Figure 2 shows the extended TCE quantile as a function of the VaR quantile, for different values of the shape parameter  $\zeta$ . In the left panel of Figure 2, the scale parameter is equal to  $\theta = 1$ , while the scale parameter is equal to  $\theta = 5$  in the right panel of this figure. Both panels are plotted assuming that  $m = 2$ .



**Figure 2.** Extended TCE quantile as a function of VaR quantile. (Left panel):  $m = 2$  and  $\theta = 1$ . (Right panel):  $m = 2$  and  $\theta = 5$ .

As in the case of Figure 1, we see that a higher value of the shape parameter leads to more intricate situations, where the VaR quantile takes a value close to one, while the extended TCE quantile can take a broad range of values. However, Figure 2 also shows that a higher value of the scale parameter leads to even more intricate situations. Thus, when value at risk is not able to distinguish between extreme risk situations, a more sophisticated indicator such as the extended TCE indicator is able to produce such a distinction.

Figure 3 is constructed in a similar way as Figure 2, but now both panels are plotted assuming that  $m = 3$ . By comparing Figures 2 and 3, we see that the curves are pushed to the right for higher values of  $m$ , all other parameters being equal. Thus, varying the value of  $m$  allows us to construct risk indicators that are more or less sensitive to the presence of extreme risks.



**Figure 3.** Extended TCE quantile as a function of VaR quantile. (Left panel):  $m = 3$  and  $\theta = 1$ . (Right panel):  $m = 3$  and  $\theta = 5$ .

### 3.3. GPD Losses

Let us now generalize the previous study to the case where the loss random variable  $X$  follows a generalized Pareto distribution. In accordance with the cdf shown in Equation (1), we rely on the following probability density function:

$$f_{X_T}(x) = \frac{1}{\sigma} \left( 1 + \frac{\xi(x - \mu)}{\sigma} \right)^{\left( -\frac{1}{\xi} - 1 \right)}, \tag{13}$$

which generalizes Equation (8) by introducing a location parameter  $\mu$ . The scale parameter is now denoted by  $\sigma$  and the shape parameter is now  $\xi \neq 0$ . For information on the estimation of tail parameters, see for instance Hill (1975) or Hosking and Wallis (1987).

We obtain a quasi-closed-form formula for the extended TCE indicator in this setting. Indeed, we have:

**Theorem 4.** In the GPD case, the extended TCE indicator can be computed as follows:

$$\begin{aligned}
 TCE_c^{(m)}(X_T) &= \frac{1}{1-c} \cdot \int_{VaR_c(X_T)}^{+\infty} x^m \frac{1}{\sigma} \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{\left(-\frac{1}{\xi}-1\right)} dx \\
 &= \frac{1}{1-c} \left[ -(-1)^{-\left(m-\frac{1}{\xi}+1\right)} \frac{1}{\xi} \frac{\Gamma\left(-m+\frac{1}{\xi}\right)\Gamma(m+1)}{\Gamma\left(1+\frac{1}{\xi}\right)} \left(\frac{\mu\xi-\sigma}{\sigma}\right)^{-\frac{1}{\xi}} \left(\frac{\mu\xi-\sigma}{\xi}\right)^m \right. \\
 &\quad \left. - (1-c) \frac{(VaR_c(X_T))^{m+1}}{(m+1) \cdot (\sigma-\mu\xi)} {}_2F_1\left(1, m-\frac{1}{\xi}+1; m+2; \frac{VaR_c(X_T)\xi}{\mu\xi-\sigma}\right) \right], \tag{14}
 \end{aligned}$$

where  ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$  is the hypergeometric function,  $\Gamma(\cdot)$  is the gamma function,  $0 < \xi < \frac{1}{m}$ , and  $VaR_c(X_T) = \mu + \frac{\sigma}{\xi} \left( (1-c)^{-\xi} - 1 \right)$ .

**Proof.** See Appendix A.  $\square$

To solve  $TCE_c^{(m)}(X_T) = VaR_q(X_T)$ , we can equivalently solve:

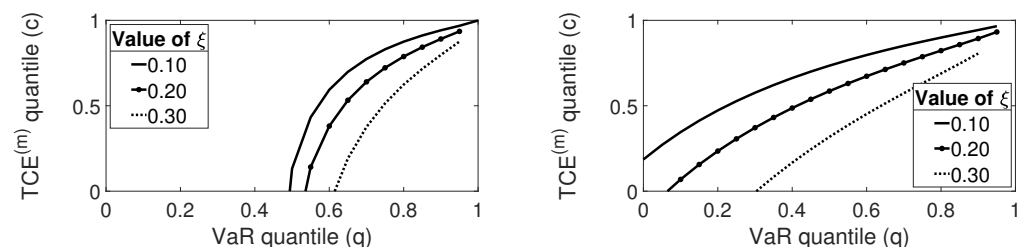
$$\begin{aligned}
 \frac{1}{1-c} \left[ -(-1)^{-\left(m-\frac{1}{\xi}+1\right)} \frac{1}{\xi} \frac{\Gamma\left(-m+\frac{1}{\xi}\right)\Gamma(m+1)}{\Gamma\left(1+\frac{1}{\xi}\right)} \left(\frac{\mu\xi-\sigma}{\sigma}\right)^{-\frac{1}{\xi}} \left(\frac{\mu\xi-\sigma}{\xi}\right)^m \right. \\
 \left. - \frac{(1-c)(VaR_c(X_T))^{m+1}}{(m+1)(\sigma-\mu\xi)} {}_2F_1\left(1, m-\frac{1}{\xi}+1; m+2; \frac{VaR_c(X_T)\xi}{\mu\xi-\sigma}\right) \right] = VaR_q(X_T). \tag{15}
 \end{aligned}$$

To numerically solve Equation (15), the following three conditions must be met:

- $0 < \xi < \frac{1}{m}$ , to ensure the convergence of the integral in Equation (14) and to avoid the appearance of complex numbers in Equation (15).
- $\mu\xi - \sigma \geq 0$ , to avoid the appearance of complex numbers in Equation (15).
- $\left| \frac{VaR_c \xi}{\mu\xi - \sigma} \right| < 1$ , which is a necessary property of the fourth parameter of the hypergeometric function.

Although Equation (15) does not admit closed-form solutions, we numerically solve it to show the relation between the  $TCE^{(m)}$  quantile and the VaR quantile, as a function of the order  $m$  and of the three generalized Pareto distribution parameters  $\xi$ ,  $\mu$ , and  $\sigma$ .

We start by plotting in Figure 4 the relation between the  $TCE^{(m)}$  and the VaR quantiles when  $m = 2$  and  $\sigma = 0.1$ . The left panel of the figure presents the situation where  $\mu = -0.05$ , while the right panel presents the situation where  $\mu = 0.05$ .



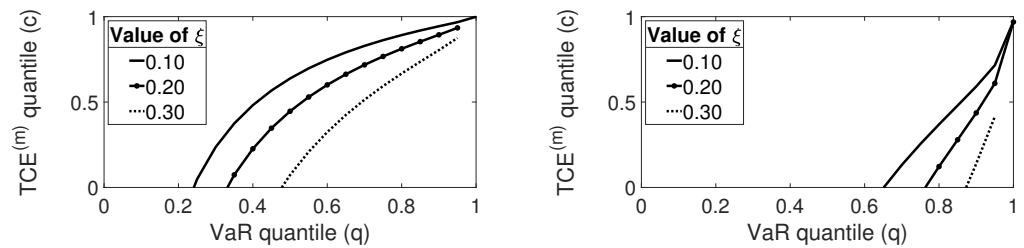
**Figure 4.** Extended TCE quantile as a function of VaR quantile. (Left panel):  $m = 2, \sigma = 0.1$ , and  $\mu = -0.05$ . (Right panel):  $m = 2, \sigma = 0.1$ , and  $\mu = +0.05$ .

From Figure 4, we see that the location parameter  $\mu$  has a sharp impact on quantile dependences. When  $\mu$  is high, the relation between indicator quantiles becomes nearly linear, which is an ideal situation from a risk management viewpoint.

Next, we plot in Figure 5 the relation between the  $TCE^{(m)}$  and the VaR quantiles when  $m = 2$  and  $\mu = 0$ . The left panel of the figure presents the situation where  $\sigma = 0.1$ , while the right panel presents the situation where  $\sigma = 0.4$ . From the comparison of the two

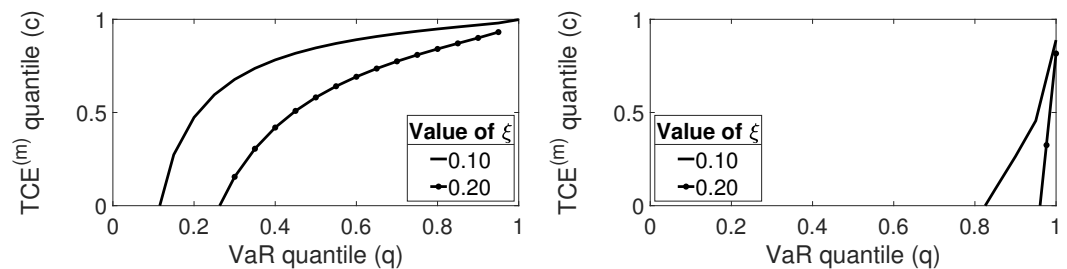


panels of the figure, we see that the ideal situation where the relation between the indicator quantiles is quasi-linear occurs for small values of the scale parameter  $\sigma$ .



**Figure 5.** Extended TCE quantile as a function of VaR quantile. (Left panel):  $m = 2, \sigma = 0.1$ , and  $\mu = 0$ . (Right panel):  $m = 2, \sigma = 0.4$ , and  $\mu = 0$ .

Finally, we plot in Figure 6 the relation between the  $TCE^{(m)}$  and the VaR quantiles when  $m = 3, \mu = 0$ . We set  $\sigma = 0.1$  in the left panel and  $\sigma = 0.4$  in the right panel. Thus, Figure 6 presents the same comparison as Figure 5, but for a larger value of the order parameter  $m$ .



**Figure 6.** Extended TCE quantile as a function of VaR quantile. (Left panel):  $m = 3, \sigma = 0.1$ , and  $\mu = 0$ . (Right panel):  $m = 3, \sigma = 0.4$ , and  $\mu = 0$ .

From the comparison of Figures 5 and 6, we see that higher values of the order parameter  $m$  lead to more intricate situations from a management viewpoint. Thus, the presence of extreme risks in the system being considered, and their taking into account via higher-order conditional moments, makes risk management more complicated in the sense that choosing an indicator quantile becomes a more critical and sensitive decision.

#### 4. Equivalence between High-Order Indicators

In this section, we study the relation between the quantiles  $q^{(m)}$  and  $q^{(n)}$  of distinct extended tail conditional expectation indicators, where each indicator is associated with a different order  $m$  or  $n$ . Namely, we examine the situation where:

$$TCE_{q^{(m)}}^{(m)}(X_T) = TCE_{q^{(n)}}^{(n)}(X_T). \tag{16}$$

We also study the sub-case where TCE is compared with a high-order TCE, that is, we study the quantiles  $q$  and  $q^{(m)}$  that satisfy:

$$TCE_q(X_T) = TCE_{q^{(m)}}^{(m)}(X_T). \tag{17}$$

##### 4.1. Pareto Distributed Losses

If we model losses as random variables that follow a classic type I Pareto distribution, we can use Equation (9) to compute  $TCE_{q^{(m)}}^{(m)}(X_T)$  where  $\alpha > m$ . Thus, we can compare two higher-order TCEs as follows:

$$\frac{\alpha}{\alpha - m} \left( \text{VaR}_{q^{(m)}}(X_T) \right)^m = \frac{\alpha}{\alpha - n} \left( \text{VaR}_{q^{(n)}}(X_T) \right)^n,$$

where

$$\text{VaR}_{q^{(m)}}(X_T) = \theta \left(1 - q^{(m)}\right)^{-\frac{1}{\alpha}}.$$

We obtain:

$$\frac{\alpha}{\alpha - m} \theta^m \left(1 - q^{(m)}\right)^{-\frac{m}{\alpha}} = \frac{\alpha}{\alpha - n} \theta^n \left(1 - q^{(n)}\right)^{-\frac{n}{\alpha}},$$

which leads us to:

$$q^{(n)} = 1 - \left(\frac{\alpha - n}{\alpha - m} \theta^{m-n}\right)^{-\frac{\alpha}{n}} \left(1 - q^{(m)}\right)^{-\frac{m}{n}} \tag{18}$$

where  $\alpha > m, \alpha > n$ , and  $\theta > 0$ .

If we denote  $\xi = \frac{1}{\alpha}$ , we can rewrite Equation (18) as in the following proposition.

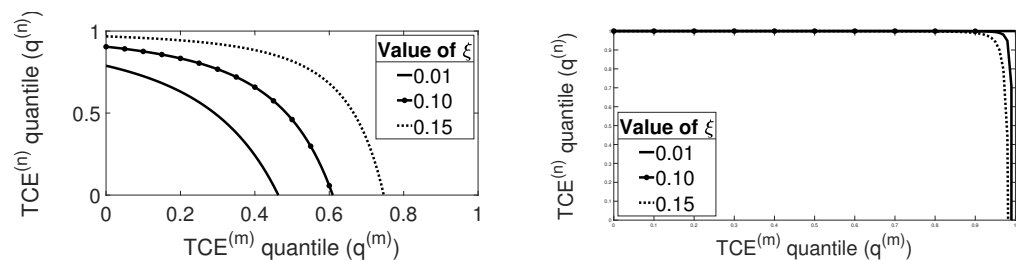
**Proposition 1.** *When losses follow a classic Pareto distribution, the quantiles of high-order TCEs that solve Equation (16) can be related as follows:*

$$q^{(n)} = 1 - \left(\frac{\frac{1}{\xi} - m}{\frac{1}{\xi} - n}\right)^{\frac{1}{\xi n}} \theta^{\frac{n-m}{\xi n}} \left(1 - q^{(m)}\right)^{-\frac{m}{n}}, \tag{19}$$

where  $0 < \xi < \frac{1}{m}, 0 < \xi < \frac{1}{n}$ , and  $\theta > 0$ .

We now illustrate this proposition.

Figure 7 plots the relation between the quantiles  $q^{(m)}$  and  $q^{(n)}$  when  $m = 5$  and  $n = 2$ . The left panel of the figure shows the situation where  $\theta = 1$ , while the right panel of the figure shows the situation where  $\theta = 2$ .



**Figure 7.** Extended TCE quantile at order  $n = 2$  as a function of extended TCE quantile at order  $m = 5$ . (Left panel):  $\theta = 1$ . (Right panel):  $\theta = 2$ .

The left panel of Figure 7 can be interpreted as follows. The relation between the high-order quantiles is countermonotonic contrary to the relation between the TCE quantile and the VaR quantile, for instance. This means that a high value of  $q^{(m)}$  corresponds to a small value of  $q^{(n)}$ , and conversely.

This feature is a consequence of the fact that high-order TCEs concentrate on different parts of probability tails. Thus, the figure shows us that a manager that reduces high-order extreme risks at a given order, say  $m$ , is not simultaneously reducing high-order extreme risks at another order, say  $n$ . The right panel of the figure tells us that this aspect is even more pronounced for higher values of  $\theta$ .

We now come to the specific case where  $n = 1$ , that is to the study of the relation between TCE and a higher-order TCE:

$$\frac{\alpha}{\alpha - 1} (\text{VaR}_q(X_T))^1 = \frac{\alpha}{\alpha - m} \left(\text{VaR}_{q^{(m)}}(X_T)\right)^m.$$

Equation (18) becomes

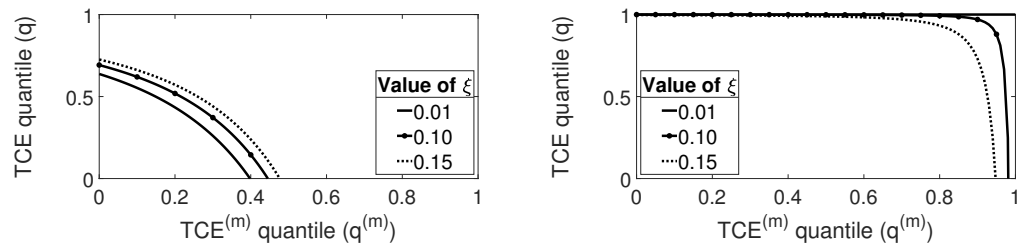
$$q = 1 - \left( \frac{\alpha - m}{(\alpha - 1)\theta^{m-1}} \right)^\alpha (1 - q^{(m)})^{-m},$$

when  $\alpha > m$  and  $\theta > 0$ . Similarly, Equation (19) becomes

$$q = 1 - \left( \frac{\frac{1}{\zeta} - m}{\left(\frac{1}{\zeta} - 1\right)\theta^{m-1}} \right)^{\frac{1}{\zeta}} (1 - q^{(m)})^{-m}$$

when  $0 < \zeta < \frac{1}{m}$  and  $\theta > 0$ .

We show in Figure 8 the relation between the TCE quantile and the high-order TCE quantile when  $m = 2$ . The left panel of the figure shows the situation where  $\theta = 1$ , while the right panel of the figure shows the situation where  $\theta = 2$ .



**Figure 8.** TCE quantile as a function of extended TCE quantile at order 2. (Left panel):  $\theta = 1$ . (Right panel):  $\theta = 2$ .

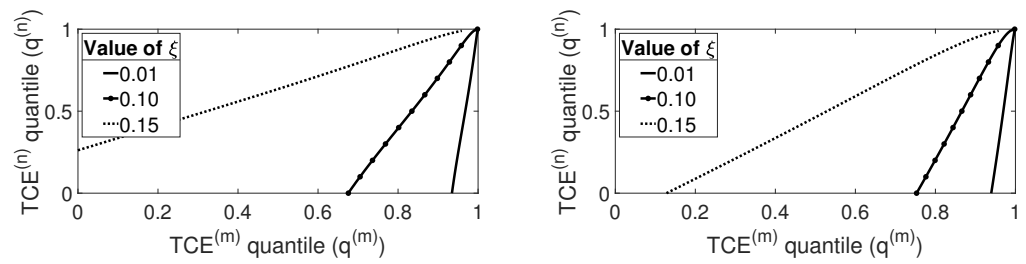
Figure 8 confirms the results of Figure 7. Reducing risks using TCE does not necessarily reduce risks as measured by a high-order TCE, and conversely. Again, this effect is more pronounced for higher values of  $\theta$ .

#### 4.2. GPD Losses

Let us now come to the more general situation where losses are modeled using a generalized Pareto distribution. Our goal is to solve Equation (16) when the extended TCE indicator  $TCE_{q^{(m)}}^{(m)}$  is given by Equation (14). Thus, to derive the relation between  $q^{(m)}$  and  $q^{(n)}$ , we numerically solve:

$$\begin{aligned} & \frac{1}{1 - q^{(m)}} \left[ -(-1)^{-(m - \frac{1}{\zeta} + 1)} \frac{1}{\zeta} \frac{\Gamma(-m + \frac{1}{\zeta})\Gamma(m + 1)}{\Gamma(1 + \frac{1}{\zeta})} \left(\frac{\mu\zeta - \sigma}{\sigma}\right)^{-\frac{1}{\zeta}} \left(\frac{\mu\zeta - \sigma}{\zeta}\right)^m \right. \\ & \left. - (1 - q^{(m)}) \frac{(\text{VaR}_{q^{(m)}}(X_T))^{m+1}}{(m + 1) \cdot (\sigma - \mu\zeta)} {}_2F_1\left(1, m - \frac{1}{\zeta} + 1; m + 2; \frac{\text{VaR}_{q^{(m)}}(X_T) \zeta}{\mu\zeta - \sigma}\right) \right] \\ & = \frac{1}{1 - q^{(n)}} \left[ -(-1)^{-(n - \frac{1}{\zeta} + 1)} \frac{1}{\zeta} \frac{\Gamma(-n + \frac{1}{\zeta})\Gamma(n + 1)}{\Gamma(1 + \frac{1}{\zeta})} \left(\frac{\mu\zeta - \sigma}{\sigma}\right)^{-\frac{1}{\zeta}} \left(\frac{\mu\zeta - \sigma}{\zeta}\right)^n \right. \\ & \left. - (1 - q^{(n)}) \frac{(\text{VaR}_{q^{(n)}}(X_T))^{n+1}}{(n + 1) \cdot (\sigma - \mu\zeta)} {}_2F_1\left(1, n - \frac{1}{\zeta} + 1; n + 2; \frac{\text{VaR}_{q^{(n)}}(X_T) \zeta}{\mu\zeta - \sigma}\right) \right] \end{aligned} \tag{20}$$

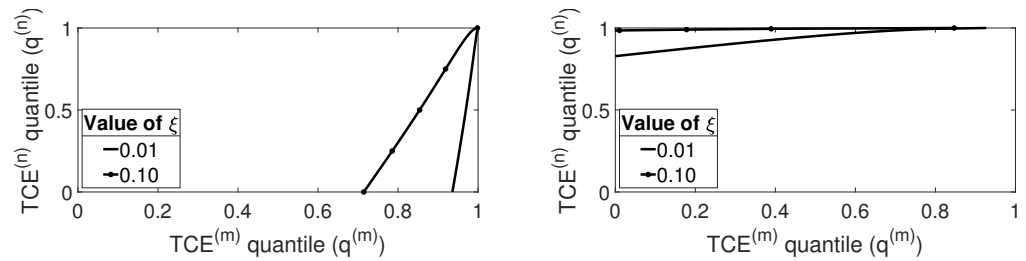
Figure 9 plots the relation between the quantiles  $q^{(m)}$  and  $q^{(n)}$  when  $m = 5, n = 2$ , and  $\sigma = 0.1$ . The left panel of the figure shows the situation where  $\mu = -0.05$ , while the right panel of the figure shows the situation where  $\mu = 0.05$ .



**Figure 9.** Extended TCE quantile at order  $n = 2$  as a function of extended TCE quantile at order  $m = 5$  with  $\sigma = 0.1$ . (**Left panel**):  $\mu = -0.05$ . (**Right panel**):  $\mu = +0.05$ .

From Figure 9, we deduce that the link between the high-order TCE quantiles is linear when  $\sigma = 0.1$ , so this parameter of the GPD distribution is not problematic. By comparing the two panels of the figure, we see that the parameter  $\mu$  has little effect on the curves linking the high-order TCE quantiles.

Figure 10 plots the relation between the quantiles  $q^{(m)}$  and  $q^{(n)}$  when  $m = 5$ ,  $n = 2$ , and  $\mu = 0$ . The left panel of the figure shows the situation where  $\sigma = 0.1$ , while the right panel of the figure shows the situation where  $\sigma = 0.4$ .

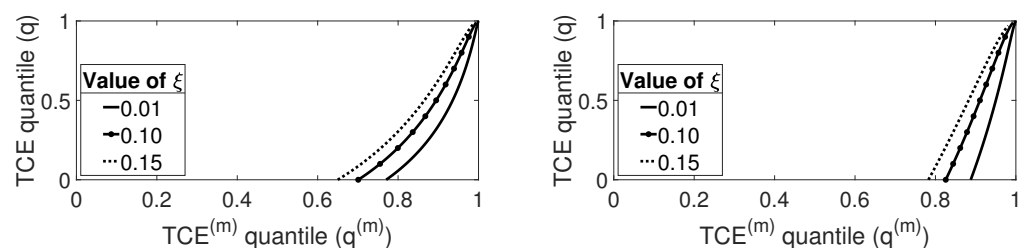


**Figure 10.** Extended TCE quantile at order  $n = 2$  as a function of extended TCE quantile at order  $m = 5$  with  $\mu = 0$ . (**Left panel**):  $\sigma = 0.1$ . (**Right panel**):  $\sigma = 0.4$ .

Figure 10 shows us that high values of  $\sigma$  can yield problematic links between the high-order TCE quantiles, hinting at probability tails that are quantified differently by distinct high-order TCE indicators.

We now come to the specific case where  $n = 1$ , that is, to the study of the relation between TCE and a higher-order TCE. In that case also, the solutions are numerically obtained by solving Equation (20).

Figure 11 plots the relation between the quantiles  $q^{(m)}$  and  $q$  when  $m = 2$  and  $\sigma = 0.1$ . The left panel of the figure shows the situation where  $\mu = -0.05$ , while the right panel of the figure shows the situation where  $\mu = 0.05$ .

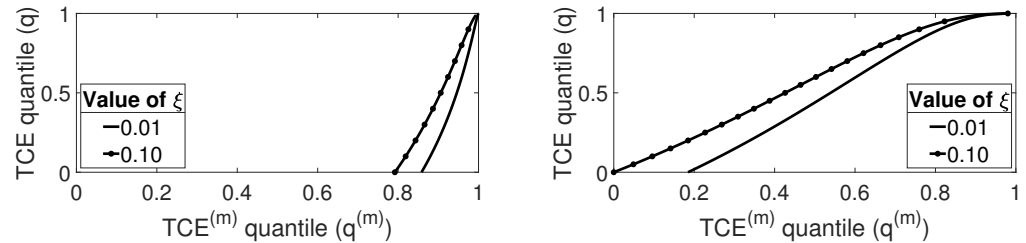


**Figure 11.** TCE quantile as a function of extended TCE quantile at order  $m = 2$  with  $\sigma = 0.1$ . (**Left panel**):  $\mu = -0.05$ . (**Right panel**):  $\mu = +0.05$ .

Figure 11 tells us that the link between the second order TCE quantile and the TCE quantile is close to linear when  $\sigma = 0.1$ , so that, again, this parameter of the GPD distribution is not problematic when it is not set to a high value.

By comparing the two panels of Figure 11, we see, as in Figure 9, that large variations of the parameter  $\mu$  have a quite limited impact on the position of the curves relating a high-order TCE quantile to the TCE quantile.

Figure 12 plots the relation between the quantiles  $q^{(m)}$  and  $q$  when  $m = 2$  and  $\mu = 0$ . The left panel of the figure shows the situation where  $\sigma = 0.1$ , while the right panel of the figure shows the situation where  $\sigma = 0.4$ .



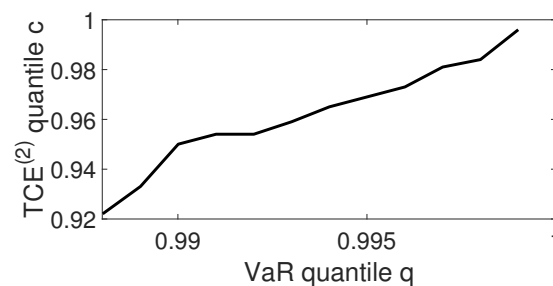
**Figure 12.** TCE quantile as a function of extended TCE quantile at order  $m = 2$  with  $\mu = 0$ . (Left panel):  $\sigma = 0.1$ . (Right panel):  $\sigma = 0.4$ .

Figure 12 confirms the conclusion of Figure 10. Specifically, high values of  $\sigma$  can yield problematic links between a high-order TCE quantile and the TCE quantile.

### 5. Conclusions

We end this paper with a brief illustration based on actual data. We use a fire insurance claim data set, labeled “beaonre” within the R package CASdatasets. This dataset includes 1823 observations of fire insurance claims from the year 1997. We transform this dataset of claim costs into a dataset of reimbursements, from which we can compute VaR, TCE, and  $TCE^{(m)}$ . To compare VaR or TCE with a high-order TCE indicator, we need to adjust the latter quantity in terms of scale. For instance, we may want to solve  $VaR_q(X_T) = (TCE_c^{(m)}(X_T))^{1/m}$ . While exact solutions to this equation do not exist, we can still deduce a relation between VaR and  $TCE^{(m)}$ .

We show in Figure 13 the link between VaR and the high-order TCE indicator (computed with  $m = 2$ ) in the case of fire data. This figure is consistent with the theoretical results shown, for instance, in Figure 6. Thus, Figure 13 confirms, in passing, the relevance of the GPD assumption. Further illustration with actual data is out of the scope of the present paper but could be a matter of an interesting extension.



**Figure 13.** Extended TCE quantile, with  $m = 2$ , as a function of VaR.

To conclude, we introduce in this paper a new risk indicator that is a high-order TCE risk measure. We compare the quantiles of this indicator to the quantiles of VaR in a simple Pareto framework, and then in a generalized Pareto framework. We also examine equivalence results between the quantiles of high-order TCEs. By doing so, we aim at illustrating the interplay between implicit choices of risk measures by regulators and the characteristics of probability distribution tails.

Among the possible theoretical extensions of our paper, one could cite the verification that the high-order indicator that we introduce is indeed a coherent risk measure. While this is out of the scope of the present paper, a separate document is in the process of being written on this aspect. See for instance Krokmal (2007) or Barbosa and Ferreira (2004) for references on the coherence of related indicators. Another possible extension of our paper could consist in examining the relation between high-order TCEs when the probability distribution admits tails that are not modeled using the generalized Pareto distribution, but using, for instance, the semi-heavy tails of infinitely divisible probability distributions. Finally, it could be interesting to examine the stability of high-order TCE indicators (see for instance the discussion in Le Courtois et al. (2020) on the cross-stability of second and fourth-order moments).

**Author Contributions:** All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

### Appendix A.1. Proof of Theorem 1

Our goal is to solve Equation (3):

$$\frac{1}{1-\xi} \text{VaR}_c + \frac{\sigma - \xi\mu}{1-\xi} - \text{VaR}_q = 0.$$

We replace VaR with its expression given in Equation (2):

$$\frac{1}{1-\xi} \left( \left( (1-c)^{-\xi} - 1 \right) \cdot \frac{\sigma}{\xi} + \mu \right) + \frac{\sigma - \xi\mu}{1-\xi} - \left( \left( (1-q)^{-\xi} - 1 \right) \cdot \frac{\sigma}{\xi} + \mu \right) = 0.$$

We have:

$$\frac{1}{1-\xi} \left( \left( (1-c)^{-\xi} - 1 \right) \cdot \frac{\sigma}{\xi} \right) + \frac{\mu}{1-\xi} + \frac{\sigma - \xi\mu}{1-\xi} - \left( \left( (1-q)^{-\xi} - 1 \right) \cdot \frac{\sigma}{\xi} \right) - \mu = 0,$$

or

$$\frac{\sigma}{\xi} \left( \frac{1}{1-\xi} \left( (1-c)^{-\xi} - 1 \right) - \left( (1-q)^{-\xi} - 1 \right) \right) = \frac{\mu - \xi\mu - \mu - \sigma + \xi\mu}{1-\xi},$$

so that

$$\frac{1}{1-\xi} \left( (1-c)^{-\xi} - 1 \right) - \left( (1-q)^{-\xi} - 1 \right) = -\frac{\sigma}{1-\xi} \frac{\xi}{\sigma}.$$

Next, we write:

$$\frac{1}{1-\xi} \cdot (1-c)^{-\xi} - (1-q)^{-\xi} - \frac{\xi}{1-\xi} = -\frac{\xi}{1-\xi}$$

and

$$\frac{1}{1-\xi} \cdot (1-c)^{-\xi} = (1-q)^{-\xi}.$$

Finally, we obtain:

$$1-c = \left( (1-\xi)(1-q)^{-\xi} \right)^{-\frac{1}{\xi}},$$

which can also be reformulated as follows:

$$c = 1 - (1-\xi)^{-\frac{1}{\xi}} \cdot (1-q).$$

Appendix A.2. Proof of Theorem 2

Our goal is to solve

$$\text{TCE}_c^{(m)}(X_T) = \frac{1}{1-c} \cdot \int_{\text{VaR}_c(X_T)}^{+\infty} x^m f_{X_T}(x) dx,$$

where  $f_{X_T}(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}$ .

We have:

$$\begin{aligned} \text{TCE}_c^{(m)}(X_T) &= \frac{1}{1-c} \cdot \int_{\text{VaR}_c(X_T)}^{+\infty} x^m \frac{\alpha\theta^\alpha}{x^{\alpha+1}} dx, \\ &= \frac{1}{1-c} \left[ -\frac{\alpha\theta^\alpha}{\alpha-m} \cdot x^{m-\alpha} \right]_{\text{VaR}_c(X_T)}^{+\infty}. \end{aligned}$$

The quantity  $\lim_{x \rightarrow +\infty} \left( -\frac{\alpha\theta^\alpha}{\alpha-m} \cdot x^{m-\alpha} \right)$  converges to zero only if  $m - \alpha < 0$ . Assuming that  $\alpha > m$ , we obtain:

$$\text{TCE}_c^{(m)}(X_T) = \frac{1}{1-c} \frac{\alpha\theta^\alpha}{\alpha-m} \cdot (\text{VaR}_c(X_T))^{m-\alpha}.$$

Next, we write:

$$\text{TCE}_c^{(m)}(X_T) = \frac{1}{1-c} \frac{\alpha\theta^\alpha}{\alpha-m} (\text{VaR}_c(X_T))^{-\alpha} \cdot (\text{VaR}_c(X_T))^m,$$

and we replace VaR with its expression:

$$\text{VaR}_c(X_T) = \theta(1-c)^{-\frac{1}{\alpha}},$$

to obtain:

$$\text{TCE}_c^{(m)}(X_T) = \frac{1}{1-c} \frac{\alpha\theta^\alpha}{\alpha-m} \left( \theta \cdot (1-c)^{-\frac{1}{\alpha}} \right)^{-\alpha} \cdot (\text{VaR}_c(X_T))^m.$$

Finally, we have:

$$\text{TCE}_c^{(m)}(X_T) = \frac{\alpha}{\alpha-m} \cdot (\text{VaR}_c(X_T))^m,$$

which is our result.

Appendix A.3. Proof of Theorem 3

Our goal is to solve Equation (10) when  $\text{VaR}_c(X_T) = \theta(1-c)^{-\frac{1}{\alpha}}$ , so when  $X_T$  follows a Pareto Type I distribution. Replacing VaR with its expression, we can rewrite Equation (10) as follows:

$$\frac{\alpha}{\alpha-m} \cdot \left( \theta(1-c)^{-\frac{1}{\alpha}} \right)^m - \left( \theta(1-q)^{-\frac{1}{\alpha}} \right) = 0.$$

Next, we write:

$$\frac{\alpha}{\alpha-m} \cdot \theta^m (1-c)^{-\frac{m}{\alpha}} - \theta(1-q)^{-\frac{1}{\alpha}} = 0,$$

so that

$$c = 1 - \left( \frac{\alpha-m}{\alpha\theta^m} \theta (1-q)^{-\frac{1}{\alpha}} \right)^{-\frac{\alpha}{m}},$$

which can readily be rewritten as Equation (11).

Appendix A.4. Proof of Theorem 4

The aim of this appendix is to demonstrate that the integral:

$$\text{TCE}^{(m)}(X_T) = \frac{1}{1-c} \int_{\text{VaR}_c(X_T)}^{+\infty} x^m \frac{1}{\sigma} \left( 1 + \frac{\zeta(x-\mu)}{\sigma} \right)^{-\frac{1}{\zeta}-1} dx$$

can be computed to provide the result in Equation (14).

Using classic results on special functions (see for instance Lebedev (1972)), we rewrite the integral as follows:

$$\begin{aligned} \text{TCE}^{(m)}(X_T) &= \frac{1}{1-c} \left[ \frac{x^{m+1}}{(m+1) \cdot (\sigma - \mu\zeta)} \left( \frac{\sigma + \zeta(x-\mu)}{\sigma} \right)^{-\frac{1}{\zeta}} \right. \\ &\quad \left. \times {}_2F_1 \left( 1, m - \frac{1}{\zeta} + 1; m + 2; \frac{x\zeta}{\mu\zeta - \sigma} \right) \right]_{\text{VaR}_c(X_T)}^{+\infty}. \end{aligned} \tag{A1}$$

To compute the limit when  $x$  tends to infinity of the quantity  $J$  defined by:

$$J = \frac{x^{m+1}}{(m+1) \cdot (\sigma - \mu\zeta)} \left( \frac{\sigma + \zeta(x-\mu)}{\sigma} \right)^{-\frac{1}{\zeta}} {}_2F_1 \left( 1, m - \frac{1}{\zeta} + 1; m + 2; \frac{x\zeta}{\mu\zeta - \sigma} \right),$$

we rewrite the hypergeometric function using a linear transformation:

$$\begin{aligned} &{}_2F_1 \left( 1, m - \frac{1}{\zeta} + 1; m + 2; \frac{x\zeta}{\mu\zeta - \sigma} \right) \\ &= \frac{\Gamma \left( m - \frac{1}{\zeta} \right) \Gamma(m+2)}{\Gamma \left( m - \frac{1}{\zeta} + 1 \right) \Gamma(m+1)} \left( -\frac{x\zeta}{\mu\zeta - \sigma} \right)^{-1} {}_2F_1 \left( 1, -m; -m + \frac{1}{\zeta} + 1; \frac{\mu\zeta - \sigma}{x\zeta} \right) \\ &+ \frac{\Gamma \left( -m + \frac{1}{\zeta} \right) \Gamma(m+2)}{\Gamma(1) \Gamma \left( 1 + \frac{1}{\zeta} \right)} \left( -\frac{x\zeta}{\mu\zeta - \sigma} \right)^{-\left( m - \frac{1}{\zeta} + 1 \right)} {}_2F_1 \left( m - \frac{1}{\zeta} + 1, -\frac{1}{\zeta}; m - \frac{1}{\zeta} + 1; \frac{\mu\zeta - \sigma}{x\zeta} \right). \end{aligned} \tag{A2}$$

Thus,  $J$  can be rewritten as follows:

$$\begin{aligned} J &= K x^m \left( 1 + \frac{\zeta(x-\mu)}{\sigma} \right)^{-\frac{1}{\zeta}} {}_2F_1 \left( 1, -m; -m + \frac{1}{\zeta} + 1; \frac{\mu\zeta - \sigma}{x\zeta} \right) \\ &+ L x^{\frac{1}{\zeta}} \left( 1 + \frac{\zeta(x-\mu)}{\sigma} \right)^{-\frac{1}{\zeta}} {}_2F_1 \left( m - \frac{1}{\zeta} + 1, -\frac{1}{\zeta}; m - \frac{1}{\zeta} + 1; \frac{\mu\zeta - \sigma}{x\zeta} \right), \end{aligned}$$

where  $K$  and  $L$  are functions of the parameters that are independent of  $x$ . Specifically,

$$K = \frac{\Gamma \left( m - \frac{1}{\zeta} \right)}{\Gamma \left( m - \frac{1}{\zeta} + 1 \right) \zeta}$$

and

$$L = \frac{\Gamma \left( -m + \frac{1}{\zeta} \right) \Gamma(m+1)}{\Gamma \left( 1 + \frac{1}{\zeta} \right)} \frac{1}{(\sigma - \mu\zeta)} \left( \frac{\zeta}{\sigma - \mu\zeta} \right)^{-\left( m - \frac{1}{\zeta} + 1 \right)}.$$



Then, we use the fact that

$$\lim_{x \rightarrow +\infty} {}_2F_1\left(1, -m; -m + \frac{1}{\xi} + 1; \frac{\mu\xi - \sigma}{x\xi}\right) = \lim_{x \rightarrow \infty} \left[ \left(\frac{\mu\xi - \sigma}{x\xi}\right)^0 + \frac{-m}{-m + \frac{1}{\xi} + 1} \frac{\mu\xi - \sigma}{x\xi} + \dots \right] = 1$$

and

$$\begin{aligned} \lim_{x \rightarrow +\infty} {}_2F_1\left(m - \frac{1}{\xi} + 1, -\frac{1}{\xi}; m - \frac{1}{\xi} + 1; \frac{\mu\xi - \sigma}{x\xi}\right) &= \\ = \lim_{x \rightarrow \infty} \left[ \left(\frac{\mu\xi - \sigma}{x\xi}\right)^0 + \frac{-\frac{1}{\xi}(m - \frac{1}{\xi} + 1)}{m - \frac{1}{\xi} + 1} \frac{\mu\xi - \sigma}{x\xi} + \dots \right] &= 1, \end{aligned}$$

and also

$$\lim_{x \rightarrow +\infty} x^m \left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-\frac{1}{\xi}} = 0$$

and

$$\lim_{x \rightarrow +\infty} x^{\frac{1}{\xi}} \left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-\frac{1}{\xi}} = \left(\frac{\xi}{\sigma}\right)^{-\frac{1}{\xi}},$$

to show that

$$\lim_{x \rightarrow \infty} J = L \left(\frac{\xi}{\sigma}\right)^{-\frac{1}{\xi}},$$

when  $\xi < \frac{1}{m}$ .

A few elementary operations allow us to write that

$$\begin{aligned} \lim_{x \rightarrow +\infty} \frac{x^{m+1}}{(m+1) \cdot (\sigma - \mu\xi)} \left(\frac{\sigma + \xi(x - \mu)}{\sigma}\right)^{-\frac{1}{\xi}} {}_2F_1\left(1, m - \frac{1}{\xi} + 1; m + 2; \frac{x\xi}{\mu\xi - \sigma}\right) &= \\ = -(-1)^{-(m - \frac{1}{\xi} + 1)} \frac{1}{\xi} \frac{\Gamma(-m + \frac{1}{\xi})\Gamma(m+1)}{\Gamma(1 + \frac{1}{\xi})} \left(\frac{\mu\xi - \sigma}{\sigma}\right)^{-\frac{1}{\xi}} \left(\frac{\mu\xi - \sigma}{\xi}\right)^m &\quad (A3) \end{aligned}$$

when  $\xi < \frac{1}{m}$ .

Finally, we compute the value of the primitive in Equation (A1) when  $x$  is equal to  $\text{VaR}_c(X_T)$ . This quantity is equal to

$$-\frac{(\text{VaR}_c(X_T))^{m+1}}{(m+1) \cdot (\sigma - \mu\xi)} \left(\frac{\sigma + \xi(\text{VaR}_c(X_T) - \mu)}{\sigma}\right)^{-\frac{1}{\xi}} {}_2F_1\left(1, m - \frac{1}{\xi} + 1; m + 2; \frac{\text{VaR}_c(X_T)\xi}{\mu\xi - \sigma}\right). \quad (A4)$$

We input Equations (A3) and (A4) into Equation (A1) and we derive Equation (14), which is our result.

## References

- Acerbi, Carlo, and Dirk Tasche. 2002. On the coherence of expected shortfall. *Journal of Banking & Finance* 26: 1487–503.
- Acerbi, Carlo, Claudio Nardio, and Carlo Sirtori. 2001. Expected shortfall as a tool for financial risk management. *arXiv*, arXiv:cond-mat/0102304.
- Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. Coherent measures of risk. *Mathematical Finance* 9: 203–28. [CrossRef]
- Barbosa, Antonio, and Miguel Ferreira. 2004. Beyond Coherence and Extreme Losses: Root Lower Partial Moment as a Risk Measure. Available online: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=609221](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=609221) (accessed on 15 May 2022).

- Barczy, Matyas, Fanni K. Nedényi, and László Sütő. 2022. Probability equivalent level of value at risk and higher-order expected shortfalls. *arXiv* arXiv:2202.09770.
- Basel Committee on Banking Supervision. 2019. The Market Risk Framework: In Brief. Available online: [https://www.bis.org/bcbs/publ/d457\\_inbrief.pdf](https://www.bis.org/bcbs/publ/d457_inbrief.pdf) (accessed on 15 May 2022).
- Basel Committee on Banking Supervision. 2022. Basel iii: International Regulatory Framework for Banks. Available online: <https://www.bis.org/bcbs/basel3.htm> (accessed on 15 May 2022).
- Canada Office of Supervision of Financial Institutions. 2022. Regulatory Capital and Internal Capital Targets. Available online: [https://www.osfi-bsif.gc.ca/Eng/fi-if/rg-ro/gdn-ort/gl-ld/Pages/a4\\_gd18.aspx](https://www.osfi-bsif.gc.ca/Eng/fi-if/rg-ro/gdn-ort/gl-ld/Pages/a4_gd18.aspx) (accessed on 15 May 2022).
- CEA. 2007. Solvency ii Glossary. Available online: [https://piu.org.pl/public/upload/ibrowser/sol2\\_glossary\\_final\\_160307.pdf](https://piu.org.pl/public/upload/ibrowser/sol2_glossary_final_160307.pdf) (accessed on 15 May 2022).
- Comité Européen des Assurances and Mercer Oliver Wyman Limited. 2005. Solvency Assessment Models Compared: Essential Groundwork for the Solvency ii Project. Available online: [https://www.naic.org/documents/committees\\_smi\\_int\\_solvency\\_eu\\_II-cea.pdf](https://www.naic.org/documents/committees_smi_int_solvency_eu_II-cea.pdf) (accessed on 15 May 2022).
- Denuit, Michel, Jan Dhaene, Marc Goovaerts, and Rob Kaas. 2006. *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. Hoboken: John Wiley & Sons.
- Fadina, Tolulope, Peng Liu, and Ruodu Wang. 2021. One Axiom to Rule Them All: An Axiomatization of Quantiles. Available online: <https://ssrn.com/abstract=3944312> (accessed on 15 May 2022).
- Fiori, Anna Maria, and Emanuela Rosazza Gianin. 2021. Generalized Pelve and Applications to Risk Measures. Available online: <https://ssrn.com/abstract=3949592> (accessed on 15 May 2022).
- Fuchs, Sebastian, Ruben Schlotter, and Klaus D. Schmidt. 2017. A review and some complements on quantile risk measures and their domain. *Risks* 5: 59. [CrossRef]
- Gatzert, Nadine, and Hans Wesker. 2011. A comparative assessment of basel ii/iii and solvency ii. *Geneva Papers on Risk and Insurance: Issues and Practice* 37: 539–70. [CrossRef]
- Hill, Bruce. 1975. A simple general approach to inference about the tail of a distribution. *Annals of Statistics* 3: 1163–74. [CrossRef]
- Hosking, John, and Jeremy Wallis. 1987. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics* 29: 339–49. [CrossRef]
- Klugman, Stuart A., Harry H. Panjer, and Gordon E. Willmot. 2012. *Loss Models: From Data to Decisions*. Hoboken: John Wiley & Sons, vol. 715.
- Krokhmal, Pavlo. 2007. Higher moment coherent risk measures. *Quantitative Finance* 7: 373–87. [CrossRef]
- Lebedev, Nicolai. 1972. *Special Functions and Their Applications*. Mineola: Dover Publications.
- Le Courtois, Olivier. 2018. Some further results on the tempered multistable approach. *Asia-Pacific Financial Markets* 25: 87–109. [CrossRef]
- Le Courtois, Olivier, and Christian Walter. 2014. The computation of risk budgets under the lévy process assumption. *Revue Finance* 35: 203–28. [CrossRef]
- Le Courtois, Olivier, Jacques Lévy-Véhel, and Christian Walter. 2020. Regulation risk. *North American Actuarial Journal* 24: 463–74. [CrossRef]
- Li, Hanson, and Ruodu Wang. 2022. Pelve: Probability equivalent level of var and es (November 25, 2019). *Journal of Econometrics*. [CrossRef]
- Linsmeier, Thomas J., and Neil D. Pearson. 2000. Value at risk. *Financial Analysts Journal* 56: 47–67. [CrossRef]
- National Association of Insurance Commissioners. 2007. *Own Risk and Solvency Assessment (ORSA) Guidance Manual*. Kansas City: NAIC.
- Rostek, Marzena. 2010. Quantile maximization in decision theory. *The Review of Economic Studies* 77: 339–71. [CrossRef]
- Society of Actuaries. 2000. Getting to Know cte. Risk Management Newsletter. Available online: <https://www.soa.org/globalassets/assets/library/newsletters/risk-management-newsletter/2004/july/rm-2004-iss02-ingram-b.pdf> (accessed on 15 May 2022).
- Wang, Ruodu, and Ričardas Zitikis. 2021. An axiomatic foundation for the expected Shortfall. *Management Science* 67: 1413–29. [CrossRef]



Article

# A Bridge Life Insurance for Households—Diagnosis and Motives

Anna Jędrzychowska

Department of Insurance, Faculty of Economics and Finance, Wrocław University of Economics and Business, ul. Komandorska 118-120, 53-345 Wrocław, Poland; anna.jedrzychowska@ue.wroc.pl

**Abstract:** Purpose: The purpose of this article is to describe the initial concept of household bridging insurance. Design/methodology/approach: In the first part of the article, an extensive literature review is made. This is made to show the research gap of insufficient protection of households against destabilization resulting from the lost personal contribution. Data shown in the text present the scale of the loss of lost unpaid work (based on household time budgets). The existing methods of managing this loss, based on social insurance, are also shown. Findings: This paper discusses the possibility of creating a new insurance. Its need is indicated (research gap, the scale of the problem, and insufficient protection by the social insurance system) and a preliminary outline of its structure is indicated (annuities character, dynamic sum insured related to the lifecycle of the household). The article contains the theoretical background of the new product, and introduces further research on the use of multistate models in the construction and calculation of insurance premiums. Originality/value: So far, studies concerning, inter alia, personal damage indicate the lost personal contribution (unpaid work for household members) and even try to evaluate it. However, no private insurance has been proposed to mitigate the destabilization resulting from the death of an adult household member. The article therefore proposes a new life insurance (a separated policy or as an extension option) that would help the household to return to normal operation after the death of one of the household members.

**Citation:** Jędrzychowska, Anna. 2022. A Bridge Life Insurance for Households—Diagnosis and Motives. *Risks* 10: 81. <https://doi.org/10.3390/risks10040081>

Academic Editors: Ermanno Pitacco and Annamaria Olivieri

Received: 23 February 2022

Accepted: 29 March 2022

Published: 8 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** life insurance; personal finance; human capital

## 1. Introduction

The concept of a life-insurance policy for households (HHs) presented herein was developed from this author's research on compensation for personal injury damage. The postulated insurance product was designed in response to the inadequate representation of the demand for financial support for HHs following the demise of a main provider, as demonstrated by multiple studies. This identified gap suggests an ostensible need for a revised life-insurance product addressed specifically to HHs, one that will provide them with adequate financial backing following the demise of a main provider while they struggle to adapt to the new situation. The product is not designed to provide full damage compensation, but rather to furnish the affected HHs with material support to alleviate the resulting income gap, to purchase or arrange for replacement of the missing material and personal contribution, and to cover any urgent needs that may arise in the context of their loss (such as psychological support or temporary work absence on the part of the remaining providers).

Modern societies offer solutions to finance this type of personal loss. In Poland, the social security system is primarily responsible for this range of services. However, the social security compensations and payouts offered are calculated in relation to minimum subsistence levels of income (without any regard for the significance and financial standing of HHs). In addition, many services of this type only cover the formally recognized members of a family—this may present additional barriers to collection of the associated

claims given the rapid cultural and lifestyle changes observed in multiple segments of the general population. In cases involving wrongful death, additional compensation may be sought by families from the guilty party. In this context, the civil law provides for the extension of claim rights to all HH members, regardless of formal family status. Wrongful death claims are typically covered by liability insurance (obligatory for drivers, medical personnel, and employers) or from private assets held by the guilty party. Because it is difficult to perform a satisfactory evaluation of the economic effects of a HH provider's demise, claims often need to be presented before court, which can deter HH members from seeking a compensation.

The above arguments emphasize the need for additional forms of insurance for HHs. The proposed solution should fulfil the following key requirements:

1. Providing coverage for two adult providers of a shared HH, with payout claimed after the demise of one;
2. Payout in the form of monthly annuities for a specified period of time;
3. Payout adjusted to actual needs dictated by the HH's lifecycle.

This paper presents a diagnostic evaluation of the need for the proposed insurance product, along with a justification for this type of product.

## 2. Household Finances—A Literature Survey

For many years, HH finances have been a topic of avid scientific dispute. In its earlier stages, the principal focus was on evaluations of aggregate values of HH income, consumption, and savings (for details, see Frenette 2014). One of the key components in this type of approach is the Keynesian assumption of direct correlations between income and consumption levels (Keynes 1936). The assumption holds that, for HHs, the average propensity to consume falls as income increases, following a decline in marginal propensity to consume. In later years, the Keynesian approach was empirically tested by Kuznets and Goldsmith (Kuznets 1942; Goldsmith 1955), the resulting evidence for which gave rise to the so-called Kuznets Consumption Puzzle. The authors demonstrated that, as disposable income rises, the share of consumption in HHs remains fairly constant over a long period of time. Further investigations were conducted in response to the intense individualization and financialization of HHs. Here we should point out the ongoing discussion in the literature of the introduction of excluded social groups (e.g., Latin Americans, Asians) and children into financial services—Shim et al. (2009, 2010); Xiao et al. (2009, 2011). In addition, cultural and country differences in consumer financial behavior are also examined (Xiao and Fan 2002; Fan and Xiao 2011; Yao et al. 2015). They indicate a continuous increase in the phenomenon of financialization.

A concept of HH lifecycles was developed, formally attributed to Fisher and Harrod, expanded upon by Ando and Modigliani (1957), and widely disputed by many other authors (e.g., North (1994)—limitations of consumption; Thaler and Shefrin (1981)—the economic theory of self-control; Friedman (1957)—the permanent income hypothesis; Duesenberry (1949)—relative income theory). Initially, two stages of the HH lifecycle were distinguished—gainful employment stage and pension stage. This model was later expanded to include the initial stage of education (Ando and Modigliani 1957). Each of the identified stages focuses on specific financial decisions—education mostly involves human capital investment, while the gainful employment stage emphasizes resource accumulation (also indebtedness) and is followed by rentier orientation in the final stage of the HH lifecycle. This concept was contested by observations (Campbell and Mankiw 1989) that HHs formulate their consumption plans based on expected income and that interest rates have no bearing upon their consumption decisions. Relations between consumption and interest rates were also explored by Parker (1999) and Hsieh (2003). However, the concept was heavily contested by White (1978) on the basis that the actual levels of savings of HHs do not fit the above formula, as the second stage of a HH lifecycle should be associated with savings and the third with consumption. In response to understandably critical arguments of this nature voiced against the various HH lifecycle models, the current scholarly focus is

on the concept of a lifecycle framework as a broader construct that includes a multitude of possible empirical models that present the effective inter-stage patterns related to HH allocation of time, effort and resources. There is still discussion in the literature about how HHs use, throughout their lives, financial services (of various kinds) to achieve their goals (e.g., Campbell 2006, p. 1553; Tufano 2009, p. 229; Xiao 2016). Koijen et al. (2016) developed a pair of risk measures, health and mortality delta, for the universe of life and health insurance products. Research by Xiao and Yao (2014), Campbell and Cocco (2003), and Koijen et al. (2009) refer to changes in the demand for credit products along with changes in the life phases of HH. During the lifecycle, along with the investing time, the increase of financial wealth and the decrease of human capital, households will go through the value ladder, from investing growth stocks transition to value (Betermier et al. 2017). Research using Chinese data show mixed evidence. Some research asserts that the lifecycle investment effect is insignificant (Wu and Qi 2007), but other research shows it exists to some degree (Wu et al. 2010). At the same time, it is important to emphasize that no universal pattern can be identified as fitting all HHs (Browning and Crossley 2001), and that analytical evaluations of HH finances should include the following elements: HH approaches to financial decisions and the preferences of individual members, characteristics and specifics of various types of HHs (including the biological structure of HHs and their child-bearing desires), and the particulars of effective phases of educational and vocational development (for details, see: Ellis 1988; Pahl 2005).

A large proportion of the available literature has been produced in response to the observed changes in HH structure. For instance, Pahl (1989) explored the increased trend towards cohabitation and the related asset isolation dilemmas (see also: Vogler 2005). Other topics include the impact of the social security system upon financial decisions made in times of a temporary fall in HH disposable income (Browning and Crossley 2001), and the relationships between the biological structure of HHs and changes in consumption patterns (Browning and Mette 2000). An emphasis is also placed on changing female roles in HHs, such as the gradual transition from a patriarchal to an egalitarian model of role distribution, which is related to the rise in vocational aspirations among women. Consequently, numerous authors (Pahl 1995, 2005; Burgoyne et al. 2006) insist on the proper recognition of several HH models, depending on their specific features and the patterns of financial decision-making adopted.

Another important issue is the scientific approach applied to the topic of financial planning by HHs, as part of general financial management approaches (Kapoor et al. 2001; Brounen et al. 2016). Financial planning should apply a long-term perspective and cover all the categories of assets and liabilities held by HHs (Campbell 2006; Jajuga et al. 2015, pp. 19–26). Therefore, it is necessary to properly recognize all elements of HH revenue and expenditure (including labor income and non-labor revenues: social security benefits, income from capital, loans and credits, endowments etc.). The effective HH revenue structure is a combination of fixed income (such as labor), periodic revenues (temporary social benefits), and incidental income (inheritance). Analyses of income in the context of financial planning should properly identify those HH activities that have a direct impact on the level of revenues. These may include the level of education (Cooper and Zhu 2014; Walker and Zhu 2013) or place of residence (Nelson and Patton 1990; Mullet et al. 1990; Gohmann et al. 1998). Inevitably, HH income is also influenced by external factors that include globalization processes (for their direct impact on national economic growth—United Nations 2012), demographic changes (for their strong impact on labor markets—Kowal et al. 2016), and the operation of social security systems (Kawiński 2016). Labor income, however, is highly susceptible to factors of a random nature. A good illustration of this can be found in the catalogue published in 1944 by the International Labor Organization (International Labour Organization (ILO) 1944), which lists the following risk factors: maternity, sickness, death of breadwinner, old age, unemployment, invalidity, and emergency expenses. This cohort of studies also extends to the context of pension/retirement income (Blau 2008; Laitner and Silverman 2005), redrawing attention

to the already emphasized significance of HH lifecycle stages for financial planning in the strict context of HH income. These revenues are counterbalanced by HH expenditure and liability. These can be divided into fixed expenditure (current consumption needs and living expenses), periodic dues (cost of child upbringing and education, healthcare and medical expenses) and incidental expenses (decision to buy a new car—regardless of whether it is planned or dictated by unforeseen circumstances). The structure and level of expenditures is influenced by HH needs. These may be expressed on a micro-scale (for instance, gender, HH lifecycle phase, HH biological structure, place of residence, health condition of HH members, education, social status, income level), and as part of a macro dimension (e.g., pricing of commodities and services, national economic growth, climate and geographic conditions, culture and social factors, market information) (Żelazna 2002; Xiao and Tao 2021). Analyses of HH finances should also strive to determine levels of individual and shared consumption (Jędrzychowska et al. 2018). The impact of the biological composition of HHs and the HH lifecycle is revisited in a later section of this paper.

The two categories of HH finances identified above—revenue and expenses—should be supplemented by another group of factors, namely HH resources. The importance of this category relates to the use of resources for achieving HH goals, such as the desire to improve living standards. The available resources also help HHs maintain financial liquidity under temporary financial duress or a fall in income. It must be noted in this context that certain types of HH resources are not easily transferable (immovable resources are one such example). Aside from these, HH resources may include financial reserves, consumer goods (furniture, home appliances, vehicles, and apparel) and immaterial resources—mainly in the form of human capital. In general, literary sources associate human capital with the ability to generate revenue (Goldsmith 1983; Baek and DeVaney 2005). The current paper, however, requires a broader perspective on the immaterial aspect of HH resources—one that offers proper recognition of the individual effort of members expressed through labor and service for the benefit of the HH, also referred to as a personal contribution.

Underlining the insufficient recognition by the scientific community of the topic of HH resources (including unpaid work for HH), this author would like to refer to the report of the Munich Center for the Economics of Aging (Hanemann and Johannes 2020). This report (using SHARE data) discusses extensively the survivor pension programs and highlights the risk of poverty while being a widow or widower. However, it also fails to attempt to assess the value of the loss of personal capital and the organizational destabilization that results from a partner's death.

The utility of measurable indices of HH production output has been explored for decades, both in domestic and international studies, albeit solely in the context of national accounts. An emphasis is placed on the fact that personal effort from HH members presents an important contribution to the national GDP. This approach focuses on the macroeconomic scale, disregarding HH finances. Postulates from this segment include the introduction of extended national accounts, developed by a specialized Eurostat task force, to incorporate calculations of the non-marketable production of HHs as part of regular public statistics (Varjonen and Alto 2006).

As observed by Kuznets (1995) and Clark (1958), GDP values are greatly understated if formal calculations fail to recognize income in kind generated for the benefit of HHs by their members. Female scientists Walker and Gauger (1975) were the first to stress that the economic value of the female contribution to HH production is drastically underestimated in conventional statistical analyses, even though approximately two thirds of HH duties are delegated to female members, and contrasts with the observation that HH duties constitute between 60 and 80% of total HH production output value. In her analyses, Folbre (2006) evokes the term “care economy”, with its key component of care work. This type of work expresses both the market and non-market value of services and labor rendered for the benefit of the HH and, as such, should be analyzed in parallel with other forms of HH production as it contributes to the general improvement of welfare. The significance of reliably estimating all the elements of production and the need to include them in

formal social and economic analyses have been stressed, among others, by Stiglitz et al. (2009). In a report produced at the behest of the President of the French Republic and the European Commission, the role of HHs as non-market producers is clearly recognized. This assumes that the apparent lack of market value does not preclude specific elements of the non-marketable production of HHs from being recognized as sources of otherwise tangible value. The above report also observes that, regardless of the volatile character of current economic conditions and social patterns, numerous tedious tasks that may safely be delegated to third-party market providers continue to be performed by HH members. These include, among others, home renovation, childcare and senior care, and meal preparation.

Invaluable input into the ongoing debate on the methods and (most of all) the need for an effective evaluation of the personal contribution of each HH member can be gained from research on compensation for personal damage claims. This segment of research emphasizes that these services constitute “unpaid (but still productive) household work” (Greenwood 1996, p. 89). Tinari deserves a special mention here for his extensive output on the subject (Tinari 1998, 2005, 2011a, 2011b) and for a compendium of the state of research (2016, as editor) on methods and prospects for personal damage compensation (in a broad sense of the term, i.e., not reduced to claims of personal loss). For the purposes of this paper, the most instructive observations are provided in chapter 10 of the above, which is devoted to a valuation of lost HH services and offers a list of prime elements that require proper consideration in this context, namely: categories of lost services, the volume and range of such services, and the estimation of their joint value. Categories invoked in the context of HH services include (without limitation) regular home maintenance, such as cleaning, shopping, and washing (Ward and Krueger 1994, p. 95), but also those of a more social dimension that are related to specific HH roles. These may include such values as sharing pastime activities, offering advice, or even being present (children rest safely in the presence of a parent) and being available to offer help in need (analogous to the services of public officers on paid duty, who are remunerated for time on duty rather than the number of interventions). This subject was elaborated upon by Olson and Rodgers (1999) who drew a distinction between services for the HH and services of an emotional nature. Due to the compound character of this category of services, the valuation task is difficult and rarely employed in practice with respect to claims for compensation. In a survey administered to members of the National Association of Forensic Economics (NAFE), only 11% of respondents said they included a measure of companionship services, and 19% said they included valuation of guidance, counsel and/or advice services (Slesnick et al. 2013, p. 90). Moreover, the forms of such services (and time needed to perform them—this aspect is revisited in a later section) are extremely volatile. Thus, the need to provide childcare arises after childbirth, to be later replaced by needs of another type, such as educational support. Another problematic issue raised in the professional literature is that of estimating the true value of lost services. Does the time spent by a HH member on a specific type of service correspond with that of a paid worker employed for the task? Dulaney et al. (1992) highlight the difference between the time needed to perform a task and the actual amount of time devoted to its performance. The authors also stress that HH members are more likely to perform a task as a team, while paid workers tend to operate single-handedly (but are potentially more time-efficient). These reservations gave rise to the concept of direct production valuation (based on time spent on a task), which is in opposition to the concept of purchasable labor (with standardized task performance times). In contrast to the above, we provide evidence to confirm that valuation based on purchasable labor is decidedly lower (Cushing and Rosenbaum 2012, p. 49). Another important caveat invoked in this context is the observation that time required to perform a task may vary between different stages in the HH lifecycle. These may arise in relation to HH composition structure (with a baby in the house, it takes more time to prepare a meal), but also to the age of the performer, as the labor efficiency of seniors is markedly lower than that of younger persons (Ireland 2011). In addition, studies on compensation claims suggest that identifying the real beneficiaries of certain HH tasks may be problematic



(Martin and Weinstein 2012, section 631; Olson and Rodgers 1999, p. 260). However, this does not constitute a problem in the context of this paper as the postulated individual life-insurance product is designed to provide a fixed value of support for HHs, irrespective of the particularities of personal loss experienced by each of the surviving members. Each approach requires proper knowledge of the time spent by HH members on such tasks and the time required to perform them. This type of information can be obtained, for instance, from the American Time Use Survey (dating back to 2003), and by national statistics offices. Inevitably, such values are mere averages, with substantial disproportions observed for specific types of HHs. Nonetheless, they serve their role as points of reference.

The final issue under scrutiny is the problem of valuation. The professional literature offers two methods: opportunity cost and replacement cost (Ireland 1999; Ireland and Ward 1999). The first method emphasizes the fact that, following the demise of a member responsible for given tasks, the duty of their performance falls upon another member. The latter will, as a result, forfeit their opportunity for gainful employment, at least for the time reserved to perform said duties. The negative consequence of this approach lies in the fact that valuations of this type emphasize the acquisition power of certain professions (thus, an MD delegated to kitchen duty will receive a remuneration ten times that of a security guard). The opposite approach calls for a valuation based on the pricing of service substitutes offered on the market. Thus, cooking tasks are valued by catering wages, house maintenance by those of cleaning operators, and so on. Although the former approach does not entail the need for identification of tasks, but rather the amount of time spent jointly on those tasks expressed in terms of forfeited profit, the latter method requires a list of specific tasks, along with the times assigned to each task and their pricing. Periodic surveys administered to NAFE members (Luthy et al. 2015) indicate that, when working out the value of lost HH services, most respondents prefer to use the “cost of hiring one or more individuals to replace the particular services that were lost” (p. 66). This 50% response matched the response in 2003, leading the authors to conclude that “Clearly, this is one area where there is not necessarily agreement among all forensic economists, but opinions are remarkably stable over time” (p. 82).

Invariably, scientists analyze the role of life insurance in securing HH. For example, Harris and Yelowitz (2018), using the Health and Retirement Study, examine individuals whose spouses died during or soon after his or her peak earning years and find that sizable lump-sum life-insurance payouts do not significantly influence spousal well-being. Satrovic and Muslija (2018) show economic and demographic determinants that are used to predict the demand for life insurance for 150 countries during the period 2005–2010. Claims raised on lost HH service as a form of tangible loss befalling surviving HH members are not hypothetical: many researchers provide evidence of such claims being raised and successfully won before a court of law (c.f. Dulaney et al. 1992, p. 124; Boss 1999, pp. 295–96).

Members of HH can buy many types of insurance. A wide variety of experimental methods used in research about insurance demand has been researched by Jaspersen (2016). In his work, he reviews 95 hypothetical surveys and experimental studies. In a review paper, Harrison and Ng (2019) argue that theories of the demand for insurance products are well developed, but the empirical literature has gaps. A lot of the scientific work concerns private health insurance. The theory and evidence concerning selection in competitive health insurance markets are reviewed by Geruso and Layton (2017). They also discuss the standard policy tools used to address the problems it creates. Based on 45 studies from countries such as the United States, Germany, the Netherlands, and Switzerland, Pendzialek et al. (2016) focused on a systematic review of empirical studies on price elasticity for health insurance. However, Saltzman (2019) estimated demand for health insurance using data from the California and Washington ACA exchanges. Using data from China Cheng and Yu (2019) found changes in demographic conditions associated with the one-child policy. Doiron and Kettlewell (2020) based their research on data from Australia. Panel research of young women showed that women purchase insurance in

preparation for pregnancy but transition out of insurance after they have finished family building. Nayak et al. (2018) studied customer preferences when purchasing a health-insurance policy. They also provided a total view of what customers expect from the health insurance industry and what the industry is prepared to provide. An overview of consumer financial issues in health care in the United States is provided by Sharpe (2016).

Additionally, a lot of work is related to disability insurance. Jimenez-Martín et al. (2019) explored the Spanish market and showed the relationship between economic conditions and disability insurance participation. According to data from the Swedish sickness insurance system, Engström et al. (2017) found that one of the interventions increases the flow to disability benefits by 20%. Armour (2018) presented research for the US Disability Insurance market and exploited a natural experiment in information provision. Researchers Le et al. (2019) also checked the US market and found that spousal coverage is associated with a reduced labor supply of secondary earners. The work on life insurance in HH financial management also includes Li and Trivedi (2016); Corea (2017); Han and Lavetti (2017); Soika (2018); Sloan et al. (2018); Briand and Lesueur (2019); Ali et al. (2019). These works concern the demand for insurance (factors shaping it) and are a discussion with classical models, e.g., the classical model of insurance demand proposed by Rothschild and Stiglitz (1976).

For subsequent deliberations, it is also necessary to emphasize another trend observed in professional literature, one stimulated by the growing significance of services purchased by HHs on the market (Chadeau 1985; Dąbrowska 2010). These are often associated with civilizational changes, such as the need to place more focus on leisure activities, and are typically realized through the delegation of certain other tedious services (both material and immaterial) to third-party providers (p. 249). One of the most potent examples in this regard is the growing servitization of elderly care, accompanied by the accretion of assets assigned for this purpose; for example, in the form of long-term care insurance (LTC) products.

### 3. Investigation of Motivations and Potential Developments for the Postulated Introduction of a Bridge Life-Insurance Product

The bulk of research discussed in the previous chapter focuses on the requirement to provide effective financial management solutions to HHs. Therefore, methods to optimize this process should be sought in the form of new products designed to provide dedicated support. This paper presents such a product: a bridge life-insurance policy formula aimed at providing material support to HHs faced with the loss of an adult member (typically a parent or a spouse).

Let us first examine the motives underpinning the proposed introduction of a bridge life-insurance product for HHs.

#### 3.1. *The Scale of the Early/Premature Death Phenomenon*

First, it is useful to note that sudden death constitutes the most challenging form of loss for HHs, at least in terms of adaptation to change. How pronounced is the statistical probability of an early death in Europe? Eurostat data for 2018 indicate that accidents were responsible for 8.7% of deaths before the age of 65 in the male population (37% of these were traffic accidents), and for 4.4% of deaths before the age of 65 in the female population (38.6% were traffic accidents). These values fall significantly for the 65+ age segment, with accidental deaths dropping to 2.5% of all deaths for males (13.6%—traffic accidents), and to 2.3% for women (8.4%—traffic accidents). Thus, it seems that the incidence of premature death is fairly contained. However, early death and its consequences should not by any means be considered moderate.

Another argument for the provision of material support for HHs faced with the loss of a member is the formal postponement of some claims. For deaths caused by default of a third party (such as traffic accidents, accidents at work, medical error, or homicide), compensation may be sought before a court of law (as already noted, this approach is well

supported in the professional literature), but the payout in such claims is inevitably delayed. Attention then turns to the average span of such claims, namely the time span between the incidence of death and passing of the final judgement. The analytical evaluation of 1875 verdicts passed by Polish civil courts (in the years 2017–2020 inclusive) in matters related to personal loss compensation claims from a surviving family member gave an average time span of 18 years between incidence and verdict, with a median of 2.8 years (this serves to justify the postulated duration of such support to be set at 3 years).

### 3.2. Social Security as a Source for Financing Premature Death Risk in Households

Another question raised in this context can be formulated as follows: do the affected HHs really need such support from a private source? After all, social security coverage should take care of their problem. However, the public social security system has its limitations. First, it does not extend to the entire population: a sizable proportion remains outside the system. In Poland, overall system participation was reported to be 28 million (data for 2021), which corresponds to 73% of the general population (26.6 million registered by the national Social Insurance Institution and 1.4 million on the Agricultural Social Insurance Fund). As a result, a large segment of the public is effectively devoid of support from this source. Another limitation is that aid instruments related to the death of a family member are only offered to HHs with a formal family status. This means that a large segment of the population in HHs formed on cohabitation patterns (fairly frequent in many countries, Poland included) have very limited access to this type of benefit. Informal partners will not receive any support, and the same applies to their offspring if they are not formally recognized by partners/parents before the proper registry office.

Regarding the recent changes observed in the structure of European populations, the following require special mention:

- The steady rise in out-of-wedlock births—Eurostat data suggest that the rate of live births outside marriage in the EU area grew from 25.4% in 2000 to 42.7% in 2019. This trend is evident in all European countries except for Latvia (a 2 percentage points (pp) decline between 2000 and 2019), Estonia and Sweden (a 0.8 pp decline). The highest rates were reported for Portugal (a 34.6 pp decline) and Spain (a 30.7 pp decline);
- A fall in the number of contracted marriages, which also serves as an indirect measure of the rise of informal relationship patterns—Eurostat data show that the crude marriage rate, namely the annual number of marriages per 1000 population, fell by 0.9 pp from 2000 to 2019. Eight member states (Estonia, Latvia, Lithuania, Hungary, Austria, Romania, Slovakia, and Sweden) registered a rise, with the most notable increase occurring in Latvia (by 2.8 pp, up to 6.6%). The most notable falls were registered for Cyprus (by 4.5 pp, down to 8.9%), Portugal (by 3 pp, down to 3.2%), and Denmark (by 1.9 pp, down to 3.5%);
- A rise in the average age of persons entering a contracted marriage and the associated postponement of child-bearing decisions—Eurostat data confirm that Europeans enter marriage at increasingly later stages in their lives. The average age of females entering formal marriage was 30.6 years of age in 2019 (an increase from 26.8 in 2000); for males, it was estimated at 33.3 years of age (compared to 29.6 in 2000). The most pronounced shifts were observed for Portugal and Spain. At the same time, the average age of first-time mothers is on the rise, and is presently calculated at 29.2 (a rise from 27.9 registered in 2010). However, a comparison of the two average values (age at first birth vs. age at marriage) for European women (data for 2019) suggests that in only three EU member states does the registered marriage come before the birth of a first child (in average terms)—Slovakia (ca. 7 months before first childbirth), Croatia (ca. 5 months) and Switzerland (ca. 4 months). The remaining states seem to follow the new trend where first childbirth occurs before the mother enters formal marriage. The most contrasting differences are observed in Sweden (where marriage comes, on average, 4 years and 8 months after the first child is born), France (4 years and 4 months), and Denmark (3 years and 2 months).

These three indicators, as perfect representations of ongoing changes in the structure of European societies, suggest a steady rise in the number of persons (cohabiting partners, non-recognized offspring) that may be effectively excluded from this form of social support (even if the deceased or their child was duly covered by the social security system).

The third important aspect is that both the pool and the amount of social security benefits assigned is simply inadequate. For instance, following the death of an adult breadwinner or parent, survivors may apply for one of the following two benefits (on top of the regular funeral benefit paid to cover last ceremony costs): surviving spouse and divorced spouse, and child-raising supplement for a single parent. The details of these are as follows:

- Surviving spouse and divorced spouse: a survivor pension in amounts decided by the number of authorized recipients and paid as a percentage of the old-age or invalidity pension to which the deceased was or would have been entitled:
  - One person: 85%;
  - Two persons: 90%;
  - Three or more persons: 95%.

This amount is then divided equally among all recipients. The benefit is not means-tested. The survivor pension is paid in monthly instalments. Amounts are calculated from the associated old-age pension or invalidity pension (for existing entitlements) or based on regular calculations of pension received from the capital stored in pension accounts (if the deceased was gainfully employed). The latter scenario is particularly unfavorable to HHs. With the former scenario, HHs are better equipped to deal with a reduction in their income and are eligible to at least 80% of the income received prior to the fact. This may be perceived as an effective reduction of consumption by the value previously assigned to one HH member. The latter scenario follows a more complex path. First, the HH is accustomed to high living standards supported by the previous earnings of the now-deceased person, and the amount of pension will always be much lower than monthly earnings. The present pension replacement rate for Poland is approximately 43% of final earnings. This already indicates that the HH will be entitled to 80–95% of a sum representing a mere 43% of hitherto supplied gains. However, the problem extends far beyond this. The cited value of the pension replacement rate describes benefits received by a person entering their retirement age, namely with the “entire” capital already placed on their pension accounts. If death befalls a person with a short history of employment, their accounts may not provide a sufficient replacement rate. Therefore, the effective payoffs will never reach the 43% return margin. Similar types of benefit products can be found in the social security systems of other European countries;

- Child-raising supplement for a single parent: granted to the parent or guardian of a child or to a full-age student whose parents have died or who is not dependent on them further to a court decision on alimony, in the amount of PLN 193 (EUR 42) per child, up to a maximum of PLN 386 (EUR 85) per family. The amount is increased by PLN 80 (EUR 18) in the case of disabled children, but within the limit of PLN 160 (EUR 35) per family monthly amounts. (source: MISSOC-Mutual Information System on Social Protection n.d.). For comparison purposes, the average salary in Poland in 2020 was PLN 5 167.47 (EUR 1131), representing an average net gain of PLN 3 731.33 (EUR 816). Thus, the volume of such support is marginal. However, many member states do not provide this form of support. These include Bulgaria, Croatia, the Czech Republic, Germany, Greece, Hungary, Latvia, Luxemburg, Slovakia, Spain, and The Netherlands.

Neither of these benefits is available to all families and, as can be seen, their value is extremely low.

The consequences of widowhood have been studied in the literature. The article by Zick and Smith (1986) used the Panel Study of Income Dynamics (PSID) and found that

both widowers and widows had more of poverty in the first five years of widowhood. Research concentrates on the economic consequences for women after the death of their husband—Hurd and Wise (1989) and Sevak et al. (2003). We also find studies containing international comparisons on the financial situation of widows (e.g., Ahn 2005, Bíró 2013). There is a discussion in the community about the survivor pension system—James (2009). However, this research concerns the moment of death “in old age”, and does not cover the risk of death at an early age—the risk of premature death.

### 3.3. The Moment of Death Risk Realization in Relation to HH Lifecycle Phases

Another important aspect arises in relation to phases in HH lifecycle. For current purposes, deliberations are restricted to the target HHs of the postulated product, namely those with two adult members and a child/children. Based on information on HH time budgets (according Eurostat), target groups will represent the following HH lifecycle phase scenarios: a childless HH with two adults aged 45 or younger, a HH with the youngest child aged 6 or younger, a HH with the youngest child aged 7–17, a childless HH with two adults past the age of 45, HH with adults past the age of 65. Inevitably, some HHs will never yield any offspring (or will otherwise have no prospect of such HH development), but the mathematical multistate model for the postulated insurance product will operate on probabilities of transitions between states, and each subsequent phase model will be calculated with suitable adjustments. It is, therefore, useful to present an illustrative description of changes in human capital (i.e., personal contribution) for each of the categories identified above, and separately for male and female survivors:

- Total—all the categories identified in research, also those excepted from the illustrative set of target HHs (type I);
- Person in a couple with youngest child less than 6 years old (type II);
- Person in a couple with youngest child between 7 and 17 years old (type III);
- Person less than 45 years old, in a couple, with no children younger than 18 years old (type IV);
- Single parent with youngest child less than 18 years old (type V) (this is meant to illustrate changes following a death of a parent, as the scenario is not limited to widower HHs).

Categories of activities are also identified, by gender roles, describing the involvement of both sexes in HH duties. Three categories are included:

- Home maintenance: food management except dish washing, dish washing, house-cleaning, household upkeep except housecleaning, laundry, ironing, handicraft and textile production/maintenance, gardening; other pet care, construction and repairs, shopping and services, travel related to shopping and services;
- Childcare, including supervision (without teaching), reading and talking, teaching, reading and communication with a child, transporting a child;
- Other HH duties: HH management, assisting a family member, travel related to other HH purposes.

Table 1 presents an overview of selected HH types and HH activities in average time units based on responses from 15 member states participating in the 2010 edition of a Eurostat survey. The most striking observation is that house maintenance and (obviously enough) childcare activities are largely intensified in HHs with children, and the main burden is generally borne by female members. By comparing couples with no children below the age of 18, couples with babies, and couples with adolescent offspring (7–17), it is clear that, for childless couples, the daily burden borne by women averages 176 min, and 96 min for their male partners. For a couple with a baby, these daily burdens reach 217 min for women and 92 min for men. For couples with an adolescent child, the burden reaches 242 min for women and 105 min for men. More importantly, the above figures describe the age of the youngest progeny, regardless of the actual number of children in the HH. For this reason, data for couples with children should be examined with care. Nonetheless,

the average value of female involvement in HH activities is twice that of males—this is another important aspect that merits attention. Several types of activities—HH upkeep (w/o housecleaning), gardening, other pet care, and construction and services—are more often performed by male HH members. With childcare activities, the burden placed on women is again twice that borne by men. For the third category—other HH duties—values for women and men are more equal, and differences are much less pronounced. It is also important to note that the average annual value of male capital placed in HH duties (Type I) is 815 h (9.3% of a year), compared to 1600 h for women (18.3% of a year). Thus, a female contribution to the benefit of their HH is considerably more pronounced and their loss will, as a result, be felt more acutely. Values of personal contribution reach their peak for HHs with an infant child (aged 6 or younger) and are observed in all types of HHs, amounting to 1089 h (12.4% of a year) for men, and 2427 h (27.7% of a year) for women.

**Table 1.** Average time spent performing activities (minutes) in particular types of households.

	Type of Activities	Type I		Type II		Type III		Type IV		Type V	
		M	F	M	F	M	F	M	F	M	F
Housework	Food management except for dish washing	20	67	17	71	19	76	17	52	31	59
	Dish washing	6	23	5	24	5	26	5	18	9	20
	Cleaning the dwelling	10	38	9	41	9	44	8	29	16	35
	Household upkeep except for cleaning the dwelling	16	14	13	13	16	15	11	11	28	14
	Laundry	1	10	1	12	0	13	1	7	4	10
	Ironing	0	9	0	10	0	13	0	7	0	8
	Handicraft and producing textiles, and other care for textiles	0	5	0	2	0	3	0	2	0	2
	Gardening; other pet care	12	8	6	4	10	7	7	5	6	4
	Construction and repairs	14	2	14	2	16	2	14	3	8	1
	Shopping and services	22	29	18	26	20	30	23	28	25	29
Travel related to shopping and services	11	14	9	12	10	13	10	14	13	14	
	Sum	112	219	92	217	105	242	96	176	140	196
Childcare	Childcare, except for teaching, reading, and talking	7	21	40	112	6	16	1	2	17	48
	Teaching, reading, and talking with a child	7	11	35	49	9	17	0	0	20	28
	Transporting a child	2	4	7	15	4	8	0	0	6	12
	Sum	16	36	82	176	19	41	1	2	43	88
Other household activities	Household management and helping a family member	4	7	3	5	4	7	4	5	5	5
	Travel related to other household purposes	2	1	2	1	2	1	2	1	3	0
	Sum	6	8	5	6	6	8	6	6	8	5
	TOTAL for the day	134	263	179	399	130	291	103	184	191	289

Another important observation derived from the data presented in Table 1 relates to changes in time budgets in HHs with a child in scenarios with both parents present, compared with those with a single parent. For HH maintenance activities, the contribution of a single father is higher than that of a male partner with a spouse (140 min per day for single fathers, and 92 or 105 min for male partners in couples, depending on the age of the youngest child). For women, this is reversed: single mothers spend less time on HH maintenance than those in couples. For the second category—childcare—and given the

fact that a single parent (Type V) is also a parent of an infant and then an adolescent child, weighted averages may be required to represent the average childcare contribution from each member over the entire 17-year span. The resulting annual averages are 30 min per day for a father and 89 min per day for a mother. Therefore, childcare contribution from a single mother will be comparable in value to that of a mother in a couple, while a single father will be required to assign a substantially greater time budget (from 30 to 43 min per day). This demonstrates that shifts in time budget following a death of an adult HH member must be properly recognized, particularly if a female member of a HH dies.

### 3.4. Loss of Household Services—A Latent Loss

The final aspect to merit special attention regarding the motives for devising the postulated insurance product for HHs relates to problems faced by surviving members of a HH following the demise of an adult member, particularly the losses resulting from such death and the extent of operational changes required to help them deal with the new situation. Based on the available catalogues of such losses (e.g., Tinari 2016), the consequences of death in a HH can be divided into three categories: loss of monetary income, loss of HH services (non-monetary income), and elevated needs.

Regarding loss of monetary income, a plethora of methods are used for compensation purposes. The UK is an example of a system based on actuarial life tables. Due to the lack of cohesion in court judgements passed in this context, a special commission was established in the 1970s and 1980s to advise on reforms and updates to the existing actuarial base. This resulted in the publication of the first edition in 1984 of *The Ogden Tables—Actuarial Tables, with Explanatory Notes for Use in Personal Injury and Fatal Accident Cases*. These were designed to support the evaluation of lump-sum compensation for loss related to personal injury—mainly the loss of monetary income and any cost incurred in relation to this (e.g., the cost of care). Tables were produced by a dedicated team under the guidance of Sir Michael Ogden, comprising expert actuaries and law specialists of the government Actuary Department. The tables have since been updated, modified, and improved in response to problems voiced by economists and to resolve issues related to their use in practice. However, despite the frequent modifications, the tables remain fairly limited and imperfect. Polemic criticisms of this method have been voiced by many researchers (e.g., Haberman and Bloomfield 1990; Ritchie 1994). The most recent edition can be accessed by its title: *Actuarial Tables, with Explanatory Notes for Use in Personal Injury and Fatal Accident Cases*, eighth edition updated May 2021 (Latimer-Sayer 2021).

In the United States, lost income is calculated—in principle—from gross wages, and most states insist on discounting the values of lost income, services, and the future cost of medical care from current values. In contrast with the British system, which is based on a unified system of multipliers directly accessible by court officers, the US approach is based on opinions from court-appointed expert economists, actuaries, and other specialists (e.g., psychologists in relation to lost non-material income). This procedure is applied per case and may differ widely depending on local state jurisdiction. A complete evaluation of the US system can be found in Ward (2009), while examples of economic analyses for individual US states are presented in Tinari (2016); Spizman and Tinari (2011); Anderson and Roberts (1989); Bryan and Linke (1988); Lane and Glennon (1985); Gilbert (1994, 1997); Thornton et al. (1997); Rodgers et al. (1996).

Importantly, Ward (2009) also provides a comparative analysis or calculation results obtained from two sources, the Ogden Tables (6th edition) and the VCF fund tables. Conclusions from analyses of factors included in the comparison demonstrate that neither source properly recognizes levels of education. In addition, the Ogden Tables disregard the impact of unemployment and disability in the cycle of vocational development. In consequence—as concluded by Ward—the Ogden Tables tend to undervalue the volume of compensation dues of younger generations and overestimate the dues of elder employees. Finally, Ward emphasizes that while the tables offer good predictability and cohesiveness, the US approach has the benefit of being formed through market competition and expert

economic input in the crucible of the courtroom, and the concepts are subject to debates and public critique.

As demonstrated above, this particular type of loss receives proper research support, with vivid disputes and a wealth of instruments designed for the purpose. Attention now shifts to the problem of the lost immaterial income of HHs, which refers to the loss of human capital defined in terms of HH services and duties and the problem of elevated needs resulting from the demise of a HH member. Practical consequences of these two aspects are presented below, complete with real-life examples of changes observed in the context studied. In place of a standard questionnaire survey, this author decided to formulate conclusions from rulings and substantiations of real-life court cases issued in relation to compensation claims raised by physical persons following the death of a family member. Information presented before the court can safely be considered equivalent to a questionnaire survey or interview. Furthermore, the gravity of proceedings and the dignity of a courtroom may in fact improve their reliability for the purpose of this study. The following are some of the most illustrative excerpts from rulings made by Polish courts of law in the cases studied, by type of consequence produced (case signatures are provided in square brackets):

- Disruption of the course of education or decline of school results (the eldest son—aged 20, a student at the University of Health and Sports Science—was forced to drop the course in order to attend to the needs of the agricultural holding left by the deceased father [case signature I C 828/11]; a second-year student had to take a sabbatical following the demise of both parents [XII C 42/19; when her father died, a daughter dropped her studies for psychological and financial reasons, and moved back to live with her mother [I C 420/15]);
- Disruption of employment (3 months work leave following the death of a cohabiting fiancée [V ACa 849/12]; a widow on 6 months work leave following the demise of her husband [I C 832/12]; limitation of gainful employment resulting from the need to take over house maintenance and childcare duties by the surviving male cohabitant with an infant daughter [I C 304/12]);
- Discontinuation or change of employment (upon the death of her partner and father of unborn child, a six-months-pregnant survivor requires full support—the survivor's mother decides to leave her job to help [I C 777/16]);
- Taking up gainful employment by a person unaccustomed to the task (following the death of the only breadwinner, a widow is forced to work as a sewist [I C 832/12]);
- Discontinuation or limitation of economic activities (a widow is compelled to discontinue a family business [I C 828/11]; a widow drastically limits the scope of her agricultural business and is forced to delegate some services to third-party providers [I ACa 878/12]);
- Changes in career development (following the death of his father, a son is forced to decline a profitable position with a pharmaceutical company to provide care, support, and full attendance for his mother [I C 420/15]);
- Taking over or delegating duties and services after a departed HH member (after his wife died, a widower is incompetent to take over the fiscal duties and other aspects of financial management of his now one-man HH and is compelled to delegate them to his son [I C 896/11]; care duties over an infant child after the death of his mother and prolonged hospitalization of his father are assumed by the brother of the late mother [I ACa 896/13]; a sister helps her brother deal with house management duties when his wife dies, such as washing, cleaning, and cooking, and the verdict calls for limitation of such involvement [III Ca 517/18]; a widow aged 33 is helped by her mother who moves in with her to offer house management and childcare support [I C 120/18]; following the demise of her husband, the scope of his colossal immaterial involvement and strenuous physical labor, such as cleaning, cooking, gardening, car servicing and construction work appliance servicing, was beyond the capabilities of the widow and many tasks had to be purchased on the market, such as mowing, roof



- repairs, etc./[I C 219/14]: when her mother died, her daughter aged 11 had to take over childcare duties for her younger brothers, including babysitting and education [I C 851/12]);
- Change of residence (when her husband died, a widow aged 46 moved with her children to another city [I C 876/13]; a widow with two children moved to live with her father in a cramped space, as her previous residence was no longer affordable [I ACa 78/16]; a widow was forced to vacate parts of residence after the share previously promised as assigned to the deceased husband had been declared by the late husband's parents as the property of the deceased's sister—the verdict supported the duty to vacate [I ACa 807/12]);
  - Auctioning of material assets (after her husband died, a widow was pressed to sell their car [I C 518/12]; after her husband died, a widow could not find solace in living alone and asked her son and his family to move in with her. She was then driven to sell the house and build a new one near her son, which was in accordance with previous arrangements with the late husband, but well below the planned standard [I ACa 922/16]);
  - Change of HH plans and prospects (when a son died, his intentions to invest in a workshop with the prospect of employing his father and similar plans made with reference to the late man's brother were no longer feasible [I ACa 307/12]; a surviving member of a steady cohabiting HH who is engaged with the marriage date already set and applicable expenses covered is forced to surrender all plans, including the forgoing house construction plans [V ACa 849/12]; following the death of his father, an eldest son takes over the agricultural business, while the youngest son aged 15 is forced to revise his plans of becoming a farmer and urged to apply for a culinary school [I ACa 878/12]; following the tragic death of a husband, family plans for house construction are shattered [I C 120/18]; a late father was planning a house development project and brewery in tandem with his son, with loans already taken out in the amount of ca. PLN 2.5 m/EUR 500,000/[I ACa 922/16]);
  - Need for psychological support, including pharmacotherapy (one and a half years after the demise of her sister, a woman is still in need of psychological intervention and therapy [I C 858/12]; for one year, a widow receives anti-psychotic and stress moderation treatment [I ACa 843/12], a mother experiences emotional breakdown and takes medication prescribed for the purpose after death of her son, with a two-year history of particularly acute symptoms [I ACa 845/12]).

As is evident from the above examples, a premature death potentially brings a whole wealth of diverse consequences covering both the vocational and functional aspects of HH operation. In addition, some of the consequences tend to correspond with or reinforce one another, for instance: the extra burden of house management will naturally limit the scope of gainful employment or disrupt the educational path of those affected by a sudden loss.

#### 4. Results—The Concept of a Bridge Life-Insurance Product for Households

Based on conclusions derived from analytical studies of professional literature, it can safely be stated that the financial aspects of HHs continue to constitute a major focus of economic analyses. New instruments, processes, and financial products should be sought to offer adequate support for HHs for the effective realization of HH goals adjusted in line with specific phases and scenarios of the HH lifecycle.

However, given previous deliberations on factors that serve to destabilize HH operation under specific scenarios, it must be noted that the specificity of effects produced by the demise of an adult HH member (and not at all limited to the purely emotional responses associated with such a loss) is not yet adequately reflected in research. Most attention is placed on lost material income. Although a few attempts have been made to explore the context of human capital in HHs, this is chiefly in relation to HH investments in education of their members made with a view to improving future HH revenues. Therefore, following the demise of an adult HH member, the survivors not only face a

loss of material income (related to education) but also the loss of specific non-monetary services and personal contributions to the benefit of the HH which were hitherto provided by the late member. The effective volume of such contributions was presented in an earlier section of this paper. In practice, only two economic domains place a mild emphasis on the valuation of selected aspects related to the personal contribution of HH members, namely macroeconomic studies (in relation to satellite accounts) and economic analyses of law and compensation of personal loss. This paper presents the potential implementation of results produced by the two domains regarding analyses of HH finances. This takes the form of a bridge life-insurance product for HHs as a vehicle allowing for the transfer of the risk of a temporary HH dysfunction following the demise of an adult provider.

Clearly, this purpose may well be served by a regular life-insurance product properly adjusted to account for this type of loss. However, as already noted, dysfunctions resulting from the demise of an adult HH member are mostly of a transient nature and can be addressed by a separate bridge-type product designed to provide temporary assistance to HHs for a period that enables them to adjust to change in their operational patterns. The postulated product would provide temporary financial support for a set period of time (e.g., three years) to help the survivors make necessary replacements and adjustments related to lost immaterial income and/or a portion of the lost monetary income (e.g., before they receive a full compensation verdict or before a child reaches adulthood).

Some elements of the lost immaterial income are easily measurable, typically by referring to market pricing for a similar type of service (in this context, it may be useful to reinforce the already mentioned servitization trends observed among HHs). The broad scope of potential types of immaterial loss borne by HHs has already been outlined, so in this section some of their most obvious “market substitutes” are provided. The list of measurable and readily replaceable services includes the following (with their market substitute presented in brackets):

- Housework (housekeeper);
- Childcare (nursery, babysitter);
- Transport, such as taking children to school, shopping, etc. (cab service);
- Educational support (coach, tutor);
- Psychological support for children (professional therapist);
- provision of HH services associated with the late member’s profession—a car mechanic responsible for service and care of the family car, a dentist provides free service to family members (purchase of the service on the market, but only in person-power, as the necessary purchases are still required; thus, the effective loss for the HH may be estimated at 40% of the market price);
- Access to social services, medical service packages, trade benefits, etc. (purchase of a market product);
- Passing own skills and abilities on to children: swimming, skiing, playing an instrument (a market product of the relevant category).

It is clear that, aside from determination of the catalogue of lost services, a method is needed to assess their proportionate time assignment, including the person-hour load per service required to cover the loss and their relevant market value.

The key properties of the postulated life-insurance product are as follows:

1. Insurance covers two adult members (breadwinners) in a HH, and is effected upon the demise of one such person (a form of “first risk insurance”).

The postulated product is designed for HHs with children or to those likely to have children. In addition, the term breadwinner extends to providers of exclusively immaterial income, such as housework and childcare duties. Originally, the concept envisaged a whole-of-life type of insurance—in this scenario, support would be offered without time limitation, even for HHs with adult and self-reliant progeny, up to the moment of demise of one of the policy holders. A set term solution may be adopted in its stead, to ensure the product does not extend past the moment the offspring reaches adulthood. However, this

author favors the whole-of-life approach, as the destabilizing effects produced by loss of a family member will not be any less acute past that point.

2. Support is paid in monthly instalments for a set period of time.

The product is intended to provide support for a period of time required by a HH to adapt to the new situation. This would involve provision of material resources in volumes adequate to the loss experienced by survivors and expressed by a drop in material income, non-monetary income, and/or elevated needs. In an earlier section, the list of easily measurable elements of such a loss was established—these may serve as a basis for the effective determination of adequate payout volumes.

The value of loss experienced by a HH can be calculated as the sum of values associated with the existing financial streams that offer the prospect of replacing the lost services (in a broad sense of the term, i.e., including the loss of social and emotional relations). As already noted, this type of calculation is akin to those adopted in macroeconomic analyses, particularly in evaluations of HH productive value, satellite accounts, and the social cost of traffic accidents. Moreover, the amount of insurance should not be set at a fixed value, as the passing of an adult provider may realize the risk at any given phase of the HH lifecycle.

3. Payout volumes are adjusted to the phase of the HH lifecycle.

The amount of insurance would be a variable sum, and payout would be calculated in relation to the HH lifecycle phase in which the risk was realized. For instance, the procedure following the death of a mother of two infant children would entail calculating her previous non-monetary contribution for the benefit of both children and her partner—washing, cooking, cleaning, assistance, transport, etc. When such a death occurs two decades later, the HH situation is substantially different, and losses in this category will be decidedly less pronounced. Similar scenarios may be presented in relation to the lost monetary income, with values established for a vocationally active person being substantially different to those for a pensioner. As already established, the product does not entail a lump-sum payoff, but is provided in monthly instalments in amounts adequate to replace the lost services associated with particular stages of a HH lifecycle. It may also involve diminishing return rates, such as 100% of pension payout in the first year, followed by 70% for the second year of insurance coverage and 50% for the third year. This approach would offer a considerable reduction in insurance premiums, and reflect the fact that adjustments in such scenarios progress in a fluid fashion: some problems are solved early after the fact, others require more time and effort.

Inevitably, the postulated product does not offer protection against any destabilizing factor that may apply in the context of a loss, as some losses are simply irreplaceable and certain services cannot be “purchased” on the market. This context calls for realization of the risk of decline in HH situations and quality of life. Consequences in this category include the following:

- Severing of family ties;
- Loss of intimacy and the prospect of family support;
- Change in lifestyle (e.g., less time for leisure and social activities);
- Lost prospects for skill acquisition (e.g., comfort of training in affable conditions);
- The risk of losing social contacts and reduction in social status (particularly in scenarios where the late person was a member of an elite social group—lawyer, member of a medical profession, etc.);
- The risk of decreased self-assessment in orphans;
- Limitation of time assigned for leisure, personal development, regeneration, and respite.

These categories are not included in the volume of compensation, and may be sought by way of compensation for pain and suffering.

In conclusion, with respect to the proposed insurance product, it may be helpful to note that HHs may safely be covered against the risk of decrease in monetary and non-monetary income by whole-of-life products, with insurance amounts adjusted to the practical requirements of HHs. Although such a solution offers greater benefits, its

practical implementation may face the barrier engendered by the general lack of financial knowledge and awareness of practical consequences that may arise following the demise of a breadwinner. The difference in insurance premiums may form another barrier to such development. Therefore, the postulated bridge life-insurance product for HHs should not be regarded as a complete answer to the problem of loss, but as an instrument bridging the apparent gap in support experienced by many HHs following the loss of an adult member; for instance, up to the moment when their liability claims are duly recognized and settled.

## 5. Conclusions

The paper presents a bridge life-insurance product addressed to HHs that is designed to cover the risk of temporary destabilization of HH functions caused by the death of a breadwinner—this term also extends to providers of exclusively immaterial income, such as housework and childcare duties.

It demonstrates that the postulated product may comprise dynamic insurance sums, with values related to phases in HH lifecycle and payout in monthly instalments. Support would be offered for a set period of time, making this an attractive alternative to standard, properly calculated personal life-insurance products.

There are still some issues to be resolved before the concept presented herein takes its final form. The first task is to design an adequate multistate model capable of recognizing value of loss for each phase of the HH lifecycle, complete with a matrix of phase-to-phase transitions. The valuation of losses associated with each phase requires a catalogue of lost services, their range, market substitutes (if any), and the associated price of service (per hour). Conversely, the dynamic construct of the matrix of transitions requires proper recognition of probabilities (both of the demise on an adult HH member and of a child being born (firstborn or otherwise)). In addition, the model should recognize changes in HH lifecycle, such as children coming of age or an adult member taking retirement. The proposed product will also need to be examined in the context of national legislations; for example, to ensure that absence of certain requirements or standards (e.g., the requirement of fixed sum) shall not be questioned. Finally, based on actuarial principles and data produced by the multi-stage model, proper estimations of insurance premiums for such a product should be performed to verify whether the product is affordable to HHs under local economic conditions. These issues will be addressed in further studies.

**Funding:** The article is a part of the project is financed by the Ministry of Science and Higher Education in Poland under the program “Regional Initiative of Excellence” 2019–2022 project number 015/RID/2018/19 total funding amount 10 721 040,00 PLN”.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data from EUROSTAT: Time spent, participation time and participation rate in the main activity by sex and age group ([https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=tus\\_00age&lang=en](https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=tus_00age&lang=en), accessed on 10 March 2022).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Ahn, Namkee. 2005. Economic Consequences of Widowhood in Europe: Cross-country and Gender Differences, European Network of Economic Policy Research Institutes (ENEPRI) Working Paper No. 32. Available online: [www.enepri.org/Publications/WP032.pdf](http://www.enepri.org/Publications/WP032.pdf) (accessed on 1 February 2022).
- Ali, S. Nageeb, Maximilian Mihm, Lucas Siga, and Chloe Tergiman. 2019. Adverse and Advantageous Selection in the Laboratory. *American Economic Review* 111: 2152–78. [CrossRef]
- Anderson, Gary A., and David L. Roberts. 1989. Stability in the present value assessment of lost earnings. *The Journal of Risk and Insurance* 56: 50–66. [CrossRef]
- Ando, Albert K., and Franco Modigliani. 1957. Tests of the life cycle hypothesis of savings comments and suggestions. *Bulletin of the Oxford University Institute of Economics & Statistics* 19: 99–124.

- Armour, Philip. 2018. The role of information in disability insurance application: An analysis of the social security statement phase-in. *American Economic Journal: Economic Policy* 10: 1–41. [CrossRef]
- Baek, Eunyoung, and Sharon A. DeVaney. 2005. Human capital, bequest motives, risk, and the purchase of life insurance. *Journal of Personal Finance, Middletown* 4: 62–84.
- Betermier, Sebastian, Laurente E. Calvet, and Paolo Sodini. 2017. Who are the value and growth investors? *The Journal of Finance* 72: 5–46. [CrossRef]
- Bíró, Aniko. 2013. Adverse effects of widowhood in Europe. *Advances in Life Course Research* 18: 68–82. [CrossRef]
- Blau, David M. 2008. Retirement and Consumption in a Life Cycle Model. *Journal of Labor Economics* 26: 35–71. [CrossRef]
- Boss, Pauline. 1999. A qualitative assessment of loss in the family system. In *Reading 24 in Assessing Family Loss in Wrongful Death Litigation: The Special Roles of Lost Services and Personal Consumption*. Tucson: Lawyers & Judges Publishing Co., pp. 289–301.
- Briand, Steve, and Jean-Yves Lesueur. 2019. Does Health Care Program Participation Favor Advantageous Selection? *Some Econometric Results from Insurance Company Data, No. halshs-02194880*. Available online: Ideas.repec.org/p/hal/journal/halshs-02194886.html (accessed on 1 February 2022).
- Brounen, Dirk, Kees G. Koedijk, and Rachel A. J. Pownall. 2016. Household financial planning and savings behavior. *Journal of International Money and Finance* 69: 95–107. [CrossRef]
- Browning, Martin, and Ejnæs Mette. 2000. *Consumption and Children*. Mimeo: University of Copenhagen.
- Browning, Martin, and Thomas F. Crossley. 2001. The Life-Cycle Model of Consumption and Saving. *Journal of Economic Perspectives* 15: 3–22. [CrossRef]
- Bryan, William, and Charles M. Linke. 1988. The estimation of the age/earnings profiles in wrongful death and injury cases: Comment. *The Journal of Risk and Insurance* 55: 168–73. [CrossRef]
- Burgoyne, Carole B., Victoria Clarke, Janet Reibstein, and Anne Edmunds. 2006. ‘All My Worldly Goods I Share with you’? Managing Money at the Transition to Heterosexual Marriage. *The Sociological Review* 54: 619–37. [CrossRef]
- Campbell, John Y. 2006. Household Finance. *The Journal of Finance* 61: 1553–604. [CrossRef]
- Campbell, John Y., and Gregory Mankiw. 1989. Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence. *NBER Macroeconomic Annual*. 4: 185–216. [CrossRef]
- Campbell, John Y., and Joao F. Cocco. 2003. Household risk management and optimal mortgage choice. *The Quarterly Journal of Economics* 118: 1449–94. [CrossRef]
- Chadeau, Ann. 1985. Measuring Household Activities: Some International Comparison. *Review of Income and Wealth*, 237–253. [CrossRef]
- Cheng, Jiang, and Lu Yu. 2019. Life and health insurance consumption in China: Demographic and environmental risks. *The Geneva Papers on Risk and Insurance* 44: 67–101. [CrossRef]
- Clark, Colin. 1958. The Economics of Housework. *Bulletin of the Oxford Institute of Statistics*, 205–11.
- Cooper, Russell, and Guozong Zhu. 2014. *Household Finance over the Life-Cycle: What does Education Contribute?* Cambridge: National Bureau of Economic Research Working Paper, No. 20684.
- Corea, Francesco. 2017. Big data and insurance: Advantageous selection in European markets. *Data Science Journal* 16: 1–15. [CrossRef]
- Cushing, Matthew J., and David I. Rosenbaum. 2012. Valuing household services: A new look at the replacement cost approach. *Journal of Legal Economics* 19: 37–60.
- Dąbrowska, Anna. 2010. *Serwicyzacja Konsumpcji w Polskich Gospodarstwach Domowych: Uwarunkowania i Tendencje*. Warszawa: Difin.
- Doiron, Denise, and Nathan Kettlewell. 2020. Family formation and the demand for health insurance. *Health Economics* 29: 523–33. [CrossRef] [PubMed]
- Duesenberry, James S. 1949. *Income, Saving and the Theory of Consumer Behavior*. Cambridge: Harvard University Press.
- Dulaney, Ronald A., John H. Fitzgerald, Matthew S. Swenson, and John H. Wicks. 1992. Market valuation of household production. *Journal of Forensic Economics* 5: 115–26. [CrossRef]
- Ellis, Frank. 1988. *Peasant Economics: Farm Households and Agrarian Development*. Cambridge: Cambridge University Press.
- Engstrom, Per, Pathric Hagglund, and Per Johansson. 2017. Early interventions and disability insurance: Experience from a field experiment. *The Economic Journal* 127: 363–92. [CrossRef]
- Fan, Jessie X., and Jing Jang Xiao. 2011. Cross-cultural differences in risk tolerance: A comparison between Chinese and Americans. *Journal of Personal Finance* 5: 54–75.
- Folbre, Nancy. 2006. Measuring Care: Gender, Empowerment and the Care Economy. *Journal of Human Development* 7: 183–99. [CrossRef]
- Frenette, Marc. 2014. An investment of a lifetime?: The long-term labour market premiums associated with a postsecondary education. In *Analytical Studies Branch Research Paper Series 1205–9153, No. 359*. Ottawa: Statistics Canada.
- Friedman, Milton. 1957. *A Theory of the Consumption Function*. Princeton: Princeton University Press.
- Geruso, Michael, and Timothy J. Layton. 2017. Selection in health insurance markets and its policy remedies. *Journal of Economic Perspectives* 31: 23–50. [CrossRef]
- Gilbert, Roy F. 1994. Estimates of earnings growth rates based on earnings profiles. *Journal of Legal Economics* 4: 1–18.
- Gilbert, Roy F. 1997. Long-term and short-term changes in earnings profiles. *Journal of Forensic Economics* 10: 29–49. [CrossRef]
- Gohmann, Stephan F., Myra J. McCrickard, and Frank Slesnick. 1998. Age-earnings profiles estimates: Do they change over time? *Journal of Forensic Economics* 11: 173–90. [CrossRef]

- Goldsmith, Raymond W. 1955. *A Study of Saving in the United States*. New York: Greenwood Press.
- Goldsmith, Art. 1983. Household Life Cycle Protection: Human Capital versus Life Insurance. *The Journal of Risk and Insurance* 50: 473–86. [CrossRef]
- Greenwood, Daphne T. 1996. Estimating hours of lost household production using time-use data: A caution. *Litigation Economics Digest* 2: 89–91.
- Haberman, Steven, and Della Suzanne Freeth Bloomfield. 1990. Work time lost to sickness, unemployment and stoppages: Measurement and application. *Journal of the Institute of Actuaries* 117: 533–78. [CrossRef]
- Han, Tony, and Kurt Lavetti. 2017. Does Part D abet advantageous selection in Medicare Advantage? *Journal of Health Economics* 56: 368–82. [CrossRef] [PubMed]
- Hanemann, Felizia, and Rausch Johannes. 2020. *Poor Survivors?* Muenchen: Economic Consequences of Death of Spouse, Munich Center for the Economics of Aging.
- Harris, Timothy F., and Aaron Yelowitz. 2018. Life insurance holdings and well-being of surviving spouses. *Contemporary Economic Policy* 36: 526–38. [CrossRef]
- Harrison, Glenn. W., and Jia Min Ng. 2019. Behavioral insurance and economic theory: A literature review. *Risk Management and Insurance Review* 22: 133–82. [CrossRef]
- Hsieh, Chang-Tai. 2003. Do Consumers React to Anticipated Income Changes? Evidence from the Alaska Permanent Fund. *American Economic Review* 93: 397–405. [CrossRef]
- Hurd, Michael D., and David A. Wise. 1989. *The Wealth and Poverty of Widows: Assets Before and After the Husband's Death, Economics of Aging*. Edited by David A. Wise. Chicago: University of Chicago Press.
- International Labour Organization (ILO). 1944. R067-Income Security Recommendation. Available online: [www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO::P12100\\_INSTRUMENT\\_ID:312405](http://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO::P12100_INSTRUMENT_ID:312405) (accessed on 10 March 2022).
- Ireland, Thomas R. 1999. Opportunity cost versus replacement cost in a lost service analysis for a wrongful death action. *Reading 6 in Ireland and Depperschmidt* 6: 55–64.
- Ireland, Thomas. R. 2011. Uses of the American Time Use Survey to measure household services: What works and does not work. *Journal of Legal Economics* 18: 61–77.
- Ireland, Thomas. R., and John O. Ward, eds. 1999. Replacement cost valuation of production by homemakers: Conceptual questions and measurement problems. *Reading 12 in Ireland and Depperschmidt*, 131–142.
- Jajuga, Krzysztof, Łukasz Feldman, Radosław Pietrzyk, and Paweł Rokita. 2015. *Integrated Risk Model in Household Life Cycle*. Wrocław: Publishing House of Wrocław University of Economics.
- James, Estelle. 2009. *Rethinking Survivor Benefits*. Washington: World Bank.
- Jaspersen, Johannes. G. 2016. Hypothetical surveys and experimental studies of insurance demand: A review. *Journal of Risk and Insurance* 83: 217–55. [CrossRef]
- Jędrzychowska, Anna, Radosław Pietrzyk, and Rokita Paweł. 2018. Oszacowanie poziomu konsumpcji indywidualnej i wspólnej w gospodarstwach domowych. In *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*. vol. 508, pp. 66–78.
- Jimenez-Martin, Sergi, Arnau Juanmarti Mestres, and Judit Vall Castello. 2019. Great Recession and disability insurance in Spain. *Empirical Economics* 56: 1623–45. [CrossRef]
- Kapoor, Johannes R., Les R. Dlabay, and Robert J. Hughes. 2001. *Personal Finance*. Boston: McGraw Hill/Irwin.
- Kawiński, Marcin. 2016. Systemy zabezpieczenia społecznego—organizacja i rozwój. In *Ubezpieczenia*. Edited by W. Ronka-Chmielowiec. Warszawa: C.H. Beck, pp. 467–81.
- Keynes, John Maynard. 1936. *The General Theory of Employment, Interest and Money*. New York: Harcourt, Brace.
- Koijen, Ralph S., Otto Van Hemert, and Stijn Van Nieuwerburgh. 2009. Mortgage timing. *Journal of Financial Economics* 93: 292–324. [CrossRef]
- Koijen, Ralph S. J., Stijn Van Nieuwerburgh, and Motohiro Yogo. 2016. Health and mortality delta: Assessing the welfare cost of household insurance choice. *The Journal of Finance* 71: 957–1009. [CrossRef]
- Kowal, Paul, Daniel Goodkind, and Wan He. 2016. *An Aging World: 2015, International Population Reports*. Washington, DC: U.S. Government Printing Office. Available online: [Census.gov/library/publications/2016/demo/P95-16-1.html](http://www.census.gov/library/publications/2016/demo/P95-16-1.html) (accessed on 10 March 2022).
- Kuznets, Simon. 1942. *Uses of National Income in Peace and War*. New York: National Bureau of Economic Research.
- Kuznets, Simon. 1995. National Income and Its Composition, 1919–1938, National Bureau in Economic Research, 1994. *What is Household Non-Market Production Worth?* Edited by Ann Chadeau. OECD Economic Studies, No. 18. Available online: <http://www.oecd.org/eo/outlook/34252981.pdf> (accessed on 10 March 2022).
- Laitner J., and D. Silverman. 2005. *Estimating Life-Cycle Parameters from Consumption Behavior at Retirement*. National Bureau of Economic Research NBER Working Paper 11163. [CrossRef]
- Lane, Julia, and Dennis Glennon. 1985. The estimation of age/earnings profiles in wrongful death and injury cases. *Journal of Risk and Insurance* 52: 686–95. [CrossRef]
- Latimer-Sayer, William, ed. 2021. Actuarial Tables with Explanatory Notes for Use in Personal Injury and Fatal Accident Cases Eighth Edition (updated) (updated May 2021), Government Actuary's Department. Available online: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/989906/Ogden\\_Tables\\_8th\\_Edition\\_Updated\\_Final\\_25.5.21.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/989906/Ogden_Tables_8th_Edition_Updated_Final_25.5.21.pdf) (accessed on 10 March 2022).

- Le, Nga, Wim Groot, Sonila M. Tomini, and Florian Tomini. 2019. Effects of health insurance on labour supply: A systematic review. *International Journal of Manpower* 40: 4. [CrossRef]
- Li, Qian, and Pravin K. Trivedi. 2016. Adverse and advantageous selection in the Medicare supplemental market: A Bayesian analysis of prescription drug expenditure. *Health Economics* 25: 192–211. [CrossRef] [PubMed]
- Luthy, Michael R., Michael L. Brookshire, David Rosenbaum, David Schap, and Frank L. Slesnick. 2015. A 2015 survey of forensic economists: Their methods, estimates, and perspectives. *Journal of Forensic Economics* 26: 53–83. [CrossRef]
- Martin, Gerard D., and M. A. Weinstein. 2012. *Determining Economic Damages*. Costa Mesa: James Publishing, Inc.
- MISSOC-Mutual Information System on Social Protection. n.d. Available online: <https://www.missoc.org/missoc-database/comparative-tables/> (accessed on 10 March 2022).
- Mullet, Matthew J., David M. Nelson, and Robert T. Patton. 1990. Alternative measures of earnings growth. *Journal of Forensic Economics* 3: 29–50. [CrossRef]
- Nayak, Bishwajit, Bala Krishnamoorthy, Som Sekhar Bhattacharrya, and Prasanta Pathak. 2018. Customer preferences for health insurance product attributes. *Journal of Services Research* 18: 59–77.
- Nelson, David M., and Robert T. Patton. 1990. Measuring earnings growth in the U.S. *Journal of Legal Economics* 3: 11–22.
- North, Douglass. 1994. Economic Performance Through Time. *The American Economic Review* 84: 359–68.
- Olson, Gerald W., and James D. Rodgers. 1999. The problem of valuing emotional services: An analysis of legal and economic criteria. *Reading 21 in Ireland and Depperschmidt* 21: 253–62.
- Pahl, Jan. 1989. *Money and Marriage*. New York: St. Martin's Press.
- Pahl, Jan. 1995. His money, her money: Recent research on financial organisation in marriage. *Journal of Economic Psychology* 16: 361–76. [CrossRef]
- Pahl, Jan. 2005. Individualisation in Couple Finances: Who Pays for the Children? *Social Policy and Society* 4: 381–91. [CrossRef]
- Parker, Jonathan A. 1999. The Reaction of Household Consumption to Predictable Changes in Payroll Tax Rates. *American Economic Review* 89: 959–73. [CrossRef]
- Pendzialek, Jonas B., Dusan Simic, and Stephanie Stock. 2016. Differences in price elasticities of demand for health 23 insurance: A systematic review. *The European Journal of Health Economics* 17: 5–21. [CrossRef]
- Ritchie, David. 1994. Smith v Manchester awards: How do courts assess loss of capacity on the labour market. *Journal of Personal Injury Litigation* 4: 103–7.
- Rodgers, James D., Michael L. Brookshire, and Robert J. Thornton. 1996. Forecasting earnings using age-earnings profiles and longitudinal data. *Journal of Forensic Economics* 9: 169–210. [CrossRef]
- Rothschild, Michael, and Joseph Stiglitz. 1976. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics* 90: 629–49. [CrossRef]
- Saltzman, Emory. 2019. Demand for health insurance: Evidence from the California and Washington ACA exchanges. *Journal of Health Economics* 63: 197–222. [CrossRef] [PubMed]
- Satrovic, Elma, and Adnan Mushlaja. 2018. Economic and demographic determinants of the demand for life insurance: Multivariate analysis. *Journal of Management and Economics Research* 1: 102–15. [CrossRef]
- Sevak, Purvi, David R. Weir, and Robert J. Willis. 2003. The Economic Consequences of a Husband's Death: Evidence from the HRS and AHEAD. *Social Security Bulletin* 65: 31–45.
- Sharpe, Deanna L. 2016. Chapter 22: Consumer financial issues in health care. In *Handbook of Consumer Finance Research*, 2nd ed. Edited by Jing Jang Xiao. New York: Springer, pp. 267–80.
- Shim, Soyeon, Jing Jang Xiao, Bonie Barber, and Angela Lyons. 2009. Pathway to life success: A conceptual model of financial well-being for young adults. *Journal of Applied Developmental Psychology* 30: 708–23. [CrossRef]
- Shim, Soyeon, Bonie Barber, Noel Card, Jing Jian Xiao, and Joyce Serido. 2010. Financial socialization of first-year college students: The roles of parents, work, and education. *Journal of Youth and Adolescence* 39: 1457–70. [CrossRef]
- Slesnick, Frank L., Michael R. Luthy, and Michael L. Brookshire. 2013. A 2012 survey of forensic economists: Their methods, estimates, and perspectives. *Journal of Forensic Economics* 24: 67–99. [CrossRef]
- Sloan, Frank A., Patricia A. Robinson, and Lindsey M. Eldred. 2018. Advantageous selection, moral hazard, and insurer sorting on risk in the US automobile insurance market. *Journal of Risk and Insurance* 85: 545–75.
- Soika, Sebastian. 2018. Moral hazard and advantageous selection in private disability insurance. *The Geneva Papers on Risk and Insurance-Issues and Practice* 43: 97–125. [CrossRef]
- Spizman, Lawrence, and Frank D. Tinari. 2011. Assessing Economic Damages in Personal Injury and Wrongful Death Litigation: The State of New York. *Journal of Forensic Economics* 22: 75–100. [CrossRef]
- Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi. 2009. Report by the Commission in the Measurement of Economic Performance and Social Progress. Available online: [http://www.stiglitz-sen-fitoussi.fr/documents/rapport\\_anglais.pdf](http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf) (accessed on 10 March 2022).
- Thaler, Richard, and Hersch Shefrin. 1981. An Economic Theory of Self-Control. *Journal of Political Economy* 89: 392–406. [CrossRef]
- Thornton, Robert J., James D. Rodgers, and Michael L. Brookshire. 1997. On the interpretation of age-earnings profiles. *Journal of Labour Research* 18: 351–65. [CrossRef]
- Tinari, Frank D. 1998. Household services: Toward a more comprehensive measure. *Journal of Forensic Economics* 11: 253–65. [CrossRef]

- Tinari, Frank D. 2005. A note on household services: Toward a more comprehensive measurement. *Journal of Forensic Economics* 17: 383–85.
- Tinari, Frank D. 2011a. Comment on “Green v. Bittner and Progeny”. *Forensic Rehabilitation and Economics* 4: 109–12.
- Tinari, Frank D. 2011b. Comment on Smith, Smith & Uhl, “Estimating the value of family household management services: Approaches and markups”. *Forensic Rehabilitation and Economics* 4: 33–36.
- Tinari, Frank D., ed. 2016. *Forensic Economics*. Palgrave Macmillan: Assessing Personal Damages in Civil Litigation.
- Tufano, Peter. 2009. Consumer finance. *Annual Review Financial Economics* 1: 227–47. [CrossRef]
- United Nations. 2012. *Development and Globalization: Facts and Figures 2012*. Switzerland: United Nations Publication.
- Varjonen, Johanna, and Kristiina Alto. 2006. *Household Production and Consumption in Finland 2001*. Helsinki: Household Satellite Account, Tilastokeskus Statistikcentralen Statistics Finland & National Consumer Research Centre, pp. 3–70.
- Vogler, Carolyn. 2005. Cohabiting couples: Rethinking money in the household at the beginning of the twenty first century. *Sociological Review* 53: 1–29. [CrossRef]
- Walker, Kathryn E., and William H. Gauger. 1975. Time and Its Dollar Value in Household Work. *Family Economics Review* 60: 8–13.
- Walker, Ian, and Yu Zhu. 2013. *The Impact of University Degrees on the Lifecycle of Earnings: Some Further Analysis*, Bis Research Paper NO. 112.
- Ward, John O. 2009. Economic damages and tort reform: A comparative analysis of the calculation of economic damages in personal injury and death litigation in the United States and the United Kingdom. *Personal Injury and Wrongful Death Damages Calculations: Transatlantic Dialogue* Edited by John O. Ward and Robert J. Thornton. Contemporary Studies in Economic and Financial Analysis: Emerald Group Publishing Limited, vol. 91.
- Ward, John O., and Kurt V. Krueger. 1994. *Establishing Damages in Catastrophic Injury Litigation*. Tucson: Lawyers & Judges Publishing Co., Inc.
- White, Betsy Buttrill. 1978. Empirical Tests of the Life Cycle Hypothesis. *The American Economic Review* 68: 547–60.
- Wu, Yangru W., and Yaxuan Qi. 2007. Liquidity, lifecycle, and heterogeneity of investment portfolio: Empirical analyses of Chinese investor behavior. *Economic Research* 42: 97–110.
- Wu, W., J. Yi, and J. Zheng. 2010. Investment structures of Chinese households: Empirical analyses based on lifecycle, wealth and housing. *Economic Research* 45: 72–82.
- Xiao, Jing Jian, ed. 2016. *Handbook of Consumer Finance Research*, 2nd ed. New York: Springer.
- Xiao, Jing Jian, and Jessie X. Fan. 2002. A comparison of saving motives of urban Chinese and American workers. *Family and Consumer Sciences Research Journal* 30: 463–95. [CrossRef]
- Xiao, Jing Jian, and Chunsheng Tao. 2021. Consumer finance/household finance: The definition and scope. *China Finance Review International* 11: 1–25. [CrossRef]
- Xiao, Jing Jian, and Rui Yao. 2014. Consumer debt delinquency by family lifecycle categories. *International Journal of Bank Marketing* 32: 43–59. [CrossRef]
- Xiao, Jing Jian, Chuanyi Tang, and Soyeon Shim. 2009. Acting for happiness: Financial behavior and life satisfaction of college students. *Social Indicator Research* 92: 53–68. [CrossRef]
- Xiao, Jing Jian, Chuanyi Tang, Joyce Serido, and Soyeon Shim. 2011. Antecedents and consequences of risky credit behavior among college students: Application and extension of the Theory of Planned Behavior. *Journal of Public Policy and Marketing* 30: 239–45. [CrossRef]
- Yao, Rui, Jing Jang Xiao, and Li Liao. 2015. Effects of age on saving motives of Chinese urban consumers. *Journal of Family and Economic Issues* 36: 224–38. [CrossRef]
- Żelazna, Krystyna. 2002. *Ekonomika Konsumpcji: Elementy Teorii*. Warszawa: SGGW.
- Zick, Cathleen D., and Ken R. Smith. 1986. Immediate and Delayed Effects of Widowhood on Poverty: Patterns from the 1970s. *The Gerontologist* 26: 669–75. [CrossRef]





Article

# Marriage and Individual Equity Release Contracts with Dread Disease Insurance as a Tool for Managing the Pensioners' Budget

Agnieszka Marciniuk \*  and Beata Zmysłona

Department of Statistics, Faculty of Economics and Finance, Wrocław University of Economics and Business, ul. Komandorska 118/120, 53-345 Wrocław, Poland; beata.zmyslona@ue.wroc.pl

\* Correspondence: agnieszka.marciniuk@ue.wroc.pl

**Abstract:** In many countries around the world, demographic and civilization changes have brought about the phenomenon of aging societies. This phenomenon affects the economy, especially the pension and health care systems, causing difficulties in their financing. The implementation of a policy that would effectively manage the problem of the longevity risk is thus required. Using housing resources and private health insurance to improve retirees' living standards may serve this purpose. The instruments we propose comprise two variants of contracts: the first for a marriage, the second for an individual client. We analysed the cash flow in both the cases. The results suggest that the amount of cash flows related to reverse equity and dread disease insurance benefits depends on the spouse's economic status, age, and health conditions. The benefits of the two variants of the contract vary. This paper examines numerous strategies for selecting the type of the contract, taking into consideration the abovementioned factors.

**Keywords:** equity release; reverse annuity contract; critical health insurance; cash flow; financial protection of elderly

**Citation:** Marciniuk, Agnieszka, and Beata Zmysłona. 2022. Marriage and Individual Equity Release Contracts with Dread Disease Insurance as a Tool for Managing the Pensioners' Budget. *Risks* 10: 140. <https://doi.org/10.3390/risks10070140>

Academic Editors: Ermanno Pitacco, Annamaria Olivieri and Montserrat Guillén

Received: 24 February 2022

Accepted: 8 July 2022

Published: 12 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Demographic changes have been caused by many factors. On the one hand, the developments and progress in medicine and the improved living conditions have extended life expectancy. On the other hand, the birth rate has been systematically decreasing. The consequence of these phenomena is the rapid increase in the over-65 population. In 2019, the percentage of people in this age group reached about 20.3% of the total European population (Eurostat 2020b). This group has become the fastest-growing segment of the population. It would not be an exaggeration to say that the world is turning grey.

The declining birth rate and increasing longevity worldwide have contributed to significant changes in the population's makeup and affected many branches of the economy, especially the labour and medical service markets. An increase in the elderly dependency ratio overburdens the work-age population. This indicates that in the future, pensions' funding gap will be a significant social problem. Moreover, the burden on health care triggered by the expenditures for the 65 and over population is becoming a large problem due to deteriorating health conditions, the presence of chronic diseases, and the need to provide long-term care for this social group. Ensuring financial security and health services to such a large elderly population is becoming a big challenge (Chen et al. 2021).

Aging implies changes in financial behaviour and the perception of many aspects of life, as well as the deterioration of health. Older adults' life satisfaction depends on such factors as their pensions, social support, and ability for self-care. Overall health and financial worries are significantly associated with life satisfaction (Borg et al. 2006, 2008). Therefore, it is also a crucial issue to provide financial peace and treatment options for the elderly (Korenman et al. 2021).

Many members of society, including retirees, own apartments or homes that they do not want to sell. On average, 70 % of people own their own properties in European Union (EU) countries (Eurostat 2020a); in Poland, this figure reaches almost 84 %, in Hungary almost 90 %, and in Romania over 95 %. Therefore, financial institutions in different countries have introduced equity release contracts for the retired (Hanewald et al. 2016), which provide an additional income in exchange for surrendering their real estates. Many variants of these contracts are available in European countries (also in Poland). The largest European market for equity release contracts is the United Kingdom (UK) market (Shao et al. 2015), while the largest world markets are the United States of America and Australia (Lee and Shi 2021).

Many studies and debates focus on equity release contracts. It is suggested that the housing equity of elderly homeowners may provide significant financial resources. The authors of the paper by (Mullings and Hamnett 1992) examined the participation of elderly persons in equity release schemes. In the article by (Toussaint and Elsinga 2009), attitudes toward the use of housing equity in several EU countries are examined. Based on 2005 data, researchers have observed that respondents used housing equity in various ways to help plan their finances. Since 2005, more people are aware that a possible method for funding the expenses associated with old age is through the use of equity release contracts (Marciniuk et al. 2020). Retirees are asset-rich and cash-poor; therefore, their housing wealth can be used to supplement their retirement income. Thus, the equity release contract is socially significant (IFoA 2019). The authors of the article (Sharma et al. 2020) point to various schemes for approaching equity release contracts, such as focussing on government subsidies for contracts. They conclude that the focus of government policy on equity release to tackle the challenges of an aging population is misplaced. However, treating equity release as an additional private source of income has many benefits.

In the richer countries of the world, an increase in demand for medical services has been observed. This increase is caused not only by demographic changes, such as the phenomenon of an aging population, but also by the occurrence of civilization diseases and technological progress in medicine. These additional factors result in increased health care expenses; consequently, it is impossible to provide high-level medical services without a mechanism of co-financing by patients. Two basic forms of co-payment can be considered: private payments, i.e., out of pocket expenditures, or private health insurance. Private health insurance can ensure access to additional funding, especially for critical or chronic illnesses. The funds obtained from this type of insurance can be used to improve the quality of life and treatment standards even in places where the state finances medical care. However, not everyone can afford to buy this insurance. Therefore, it has been necessary to introduce financial products combining different types of contracts (e.g., life insurance or equity release) with health insurance. Additional contracts are a source of financing the health insurance premium. The paper by (Webber 1993) presents an example of the financial contract, which meets the healthcare costs of the elderly. The authors emphasize that a major demand for the new products have been created by demographic changes, more widespread homeownership, and other political and cultural changes. A rapidly expanding market for products providing finance and healthcare for the elderly was observed, and its analysis indicates that the aged care insurance is both feasible and welfare-enhancing (Paolucci et al. 2015). Contracts that combine with critical health insurance are described in many papers. An example of a contract with life insurance is examined in the papers by (Dębicka and Zmyślona 2018, 2019). The second example, considering a contract combined with a reverse annuity contract is described in article (Dębicka et al. 2015), which discusses an individual contract and one for a married couple (Zmyślona and Marciniuk 2020). In the two mentioned articles, home equity release products combined with dread disease insurance are promoted as a potential solution to long-term care costs for the elderly.

The aim of this paper is to present two variants of these contracts, the first for both spouses and the second is a combination of two separate contracts for the wife and the husband. In Section 2, we describe the theoretical background of the contracts, and the

multiplied Markov models are applied. Two varieties of contract with their cash flows are presented in Section 3. In Section 4, we analyse and compare the cash flows related to both options. The analysis results suggest the contract version affects the amount of cash flows connected to the reverse annuity contract and the dread disease insurance benefits, depending on the financial needs, health status, and spouses' age. The construction of the presented model is original. It is a continuation of the research conducted by the authors of the current work. Firstly, the individual model described in (Dębicka et al. 2015) was generalized to the marital model. In addition, the model was extended to include health benefit payments in two states (not only at the moment of diagnosis but also when the health condition worsens). Secondly, all analyses are performed based on the authors' own life expectancy tables for lung cancer patients in the critical stage described in (Dębicka and Zmyślona 2016, 2019). Finally, all calculations are performed based on original computer programs in MATLAB.

## 2. Contract Combining Reverse Annuity with Critical Health Insurance

The new product consists of two elements, namely, the reverse annuity contract and the dread disease insurance. One of the equity release forms similar to the home reversion scheme is the reverse annuity contract (Marciniuk et al. 2020). The benefits derived from the new product can help improve the living conditions and the quality of life for the elderly as well as provide additional financial resources in case of a critical illness. The construction of this kind of contract is based on a model that describes an extended lifetime considering the risk of morbidity, severe disease, and life expectancy in the case of the critical stage.

Usually, both spouses are property owners and so the joint further lifetime of the wife and the husband should be considered. The spouses, who want to contract reverse annuity, receive a lifetime annuity in exchange for surrendering their real estate to a company. We consider two kinds of cash flows; the first group is connected with the reverse annuity contract. We distinguish a single premium and annuity benefits. This premium is not actually paid. It depends on the percentage  $\alpha$  of real estate value. In the considered contract, the annuity benefits are paid with the option of selecting the last-surviving status (Marciniuk et al. 2020). This status means that the annuity benefits are paid every year while at least one of the spouses is alive and healthy. The annuity benefits depend on the pensioner's age and his/her further life expectancy and the value of the real estate  $W$ .

The second kind of cash flow concerns critical disease insurance. The survival time in case of falling ill with a severe disease depends on its course. We consider the health insurance premium that is paid every year that the insured party is healthy and alive. This insurance premium is paid as a percentage  $(1 - \beta)$  of the reverse annuity benefits. The health insurance benefits are firstly paid in case of a diagnosis. Moreover, the contract enables a benefit payment in the situation of deterioration of a patient's condition. This is important because the deterioration of health conditions in the case of a critical illness involves the necessity of care in the terminal state. The end-of-life health expenditures are often huge. In many cases, families caring for terminally ill patients require financial support. For this reason, mild and critical health conditions are distinguished. In the critical stage the fatality rates grow rapidly, which decreases the probability of survival year by year. The analysis of the joint cash flows connected with the reverse annuity contract and health insurance requires the introduction of a multistate model, which is built separately for both spouses.

We have distinguished the following steps:

First step: Equity release—in exchange for surrendering their real estate, the clients at the retirement age receive a lifetime annuity (they do not pay a cash premium for that); the level of this theoretical annuity is computed in relation to the value of the house.

Second step: from the theoretical annuity, a portion  $1 - \beta$  is paid to the insurer to finance health insurance; the remaining portion  $\beta$  is paid in cash to the annuitant as a lifetime annuity (level 1).

Third step: in case of a mild stage illness, the annuitant does not have to pay the health premium and therefore receives an annuity (level 2) equal to the initial theoretical annuity; at the moment of a diagnosis, additional benefits are paid through health insurance.

Fourth step: in case of the critical stage of an illness; an additional single benefit financed by health insurance is paid to the client (level 3).

2.1. Multistate Markov Models

The multistate model is described by the state space  $S = \{1, 2, \dots, N\}$  and a set of direct transitions between states  $T$  of the state space, where  $(i, j)$  denotes a direct transition from state  $i$  to  $j$  ( $i \neq j, i, j \in S$ ). All the possible insured risk events up until the end of the contract can be described by this model (Pitacco 2014; Dhaene et al. 2017; Haberman and Pitacco 2018). We consider the Markov model in which the stationary Markov chain is described as a discrete-time process  $\{X(k), k \geq 0\}$ . The model describes the state and its changes for the contract at time  $k = \{0, 1, \dots, n\}$ .

The construction of a contract combining the reverse annuity and dread disease insurance enables the introduction of two models. The first one is related to the reverse annuity contract and the latter with health insurance. The model connected with the reverse annuity contract considers only two states:

1. the insured is alive and healthy;
2. the insured is dead.

The model connected with health insurance considers the course of a dread disease. Thus, the following states are distinguished (Zmyślona and Marciniuk 2020):

1. the insured is alive and healthy;
2. the insured became mildly ill during the last year;
3. the insured has been mildly ill for at least one year;
4. the insured became critically ill during the last year;
5. the insured has been critically ill for at least one year;
6. the insured is dead (D—dead).

We assume that an insured could survive in the critical stages of illness during  $h$  years, and for that reason the fifth stage is extended by the introduction of states  $5^{(1)}, 5^{(2)}, \dots, 5^{(h)}$ . The Markov model has a lack of memory; therefore, the presented model is extended by one additional state denoted by the number 3. This state describes a case when the insured fell mildly ill during the year. The multistate model related to the critical health insurance is presented in Figure 1 (Zmyślona and Marciniuk 2020). The circles present the states, and the arcs describe the direct transitions between the states.

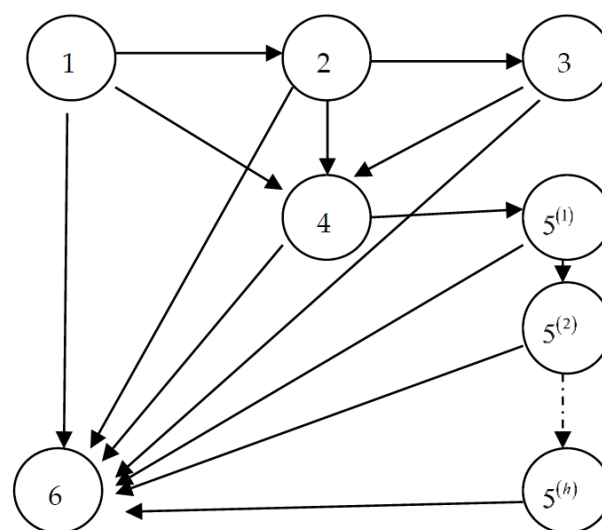


Figure 1. The multistate model for one spouse.

2.2. The Probability Structure of the Model

The probabilistic structure of the multistate model for the considered contract is given by the transition probability matrix  $q_{ij}^{[x]}(k) = P(X(x, k + 1) = j | X(x, k) = i)$ . We denote the elements of matrix  $\mathbf{Q}(k)$  by  $q_{ij}^{[x]}(k)$ . We define the following nonzero transition probabilities matrix for a person initially aged  $x$  (during the time interval  $[x + k, x + k + 1)$ ) connected with the model for the reverse annuity contract

$$\mathbf{Q}^{[x]}(k) = \begin{pmatrix} q_{11} & q_{12} \\ 0 & 1 \end{pmatrix}, \tag{1}$$

where  $q_{11}$  stands for the probability of surviving denoted as  $p_{x+k}$  and  $q_{12}$  represents the probability of death denoted as  $q_{x+k}$ . Considering the sex of the insured requires the introduction of two separate matrixes  $\mathbf{Q}^{[x]}$  and  $\mathbf{Q}^{[y]}$ , where  $x$  denotes the initial age of a woman and  $y$  the initial age of a man.

The transition probabilities matrix connected with the multistate model related to critical health insurance is given by

$$\mathbf{Q}_h^{[x]}(k) = \begin{pmatrix} q_{11} & q_{12} & 0 & q_{14} & 0 & 0 & \dots & 0 & q_{16} \\ 0 & 0 & q_{23} & q_{24} & 0 & 0 & \dots & 0 & q_{26} \\ 0 & 0 & q_{33} & q_{34} & 0 & 0 & \dots & 0 & q_{36} \\ 0 & 0 & 0 & 0 & q_{45^{(1)}} & 0 & \dots & 0 & q_{46} \\ 0 & 0 & 0 & 0 & 0 & q_{5^{(1)5^{(2)}}} & \dots & 0 & q_{5^{(1)6}} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2}$$

The transition probabilities matrixes in the female and male populations are denoted by  $\mathbf{Q}_h^{[x]}$  and  $\mathbf{Q}_h^{[y]}$ , respectively.

The formulas for these elements are obtained based on the methodology of the multistate life table (increment-decrement table). The descriptions of the formulation and estimation processes are described in (Dębicka and Zmyślona 2016, 2019). These elements depend on the following population rates:

- $q_{x+k}$ —the probability of death;
- $\varphi_{x+k}$ —the dread disease mortality rate;
- $\chi_{x+k}$ —the dread disease incidence rate (the morbidity rate);
- $\psi_{x+k}$ —the percentage of patients diagnosed in the critical stage;
- $\zeta_{x+k}$ —the probability of health deterioration to the critical state;
- $d_{x+k}^{(i,j)}$ —the fatality rate in the population of the critically ill.

The above first three values are calculated for the whole population and the remaining three for the patient population.

The formulas for the elements of the matrixes  $\mathbf{Q}_h^{[x]}$  are presented in Table 1 (Dębicka and Zmyślona 2016, 2019). The matrix formulas for the male  $\mathbf{Q}_h^{[y]}$  population are obtained analogously to those for the female population.

**Table 1.** The formulas for the estimators of the transition probabilities for health insurance.

The Number of State	Estimators of Transition Probabilities
1	$q_{11} = 1 - (q_{x+k} - \varphi_{x+k}) - \chi_{x+k}$ , $q_{12} = \chi_{x+k}(1 - \psi_{x+k})$ $q_{13} = \chi_{x+k} \cdot \psi_{x+k}$ , $q_{16} = q_{x+k} - \varphi_{x+k}$
2 and 3	$q_{ij} = 1 - q_{x+k} - \zeta_{x+k}$ , for $i = 2, 3$ and $j = 3$ $q_{ij} = \zeta_{x+k}$ , for $i = 2, 3$ and $j = 4$ $q_{i6} = q_{x+k}$ , for $i = 2, 3$
4	$q_{45^{(1)}} = 1 - d_{x+k}^{(4,5^{(1)})}$ , $q_{46} = d_{x+k}^{(4,5^{(1)})}$
$5^{(1)}, \dots, 5^{(h)}$	$q_{5^{(i)}5^{(j)}} = 1 - d_{x+k}^{(5^{(i)},5^{(j)})}$ , for $i = 1, 2, \dots, h - 1$ and $j = i + 1$ $q_{5^{(l)}6} = d_{x+k}^{(5^{(l)}6)}$ , for $l = 1, 2, \dots, h$ .

### 3. Two Version of Combining Contract

A property is usually owned by both spouses, so the contract should be concluded by both a husband and a wife. The spouses may sign either one marriage contract for which they will receive one annuity benefit, or two individual contracts, in which two annuity benefits are paid independently. In each of the abovementioned cases, health insurance benefits are paid independently for the husband and the wife. Both solutions have advantages and disadvantages and provide a varied number of benefits. The choice of the type of contract will have an impact on the formation and differentiation of the reverse annuity and health insurance benefits.

#### 3.1. Marriage Contract

The spouses sign one contract in which, in exchange for surrendering their rights to the property, they receive an annuity for life. A percentage  $\alpha$  of the real value of the property  $W$  funds the benefit. Then, they pay two separate premiums for private health insurance. The funds designated for paying health insurance form a part of the annuity premium. We define  $\beta \in [0, 1]$  as the reverse annuity parameter. This parameter describes the percentage of the paid annuity. The remaining part  $1 - \beta$  describes the health insurance premium. The impact of this parameter on the benefits is described in (Zmysłona and Marciniuk 2020). The annuity benefit is paid in advance for both spouses at the beginning of the year, if at least one of the spouses is alive.

The cash flows connected with the marriage version of the model are obtained in three steps. In the first step, the annuity benefits are calculated based on the multiple state Markov model, which describes the further lifetime of a couple. It consists of 4 states (1—both spouses are alive, 2—the wife is dead, 3—the husband is dead, and 4—both spouses are dead).

Let  $p_i(k) = P(X(k) = i)$  and  $\mathbf{P}(k) = (p_1(k), p_2(k), p_3(k), p_4(k))^T$  be the probabilities of the remaining process  $\{X(k), k \geq 0\}$  at states  $k$ , for  $k = 1, 2, 3, \dots$ . The matrixes describing the process duration in each state are shown in Table 2 for the wife and the husband.

**Table 2.** The process duration matrixes for a marriage reverse annuity contract.

For a Wife	For a Husband
$\mathbf{H}^X(k) = \begin{pmatrix} p_{x+k} & q_{x+k} & p_{x+k} & q_{x+k} \\ 0 & 1 & 0 & 1 \\ 0 & 0 & p_{x+k} & q_{x+k} \\ 0 & 0 & 0 & 1 \end{pmatrix}$	$\mathbf{H}^Y(k) = \begin{pmatrix} p_{y+k} & p_{y+k} & q_{y+k} & q_{y+k} \\ 0 & p_{y+k} & 0 & q_{y+k} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

The transition probability matrix for the 4-state Markov model (for a marriage) is defined as the following Hadamard product

$$\mathbf{Q}^{X \circ Y}(k) = \mathbf{H}^X(k) \circ \mathbf{H}^Y(k), \tag{3}$$

where  $x$  denotes the age of the wife and  $y$  the age of the husband. An element of matrix (2) equals the following product  $q_{ij}^{(X \circ Y)} = h_{ij}^X \cdot h_{ij}^Y$  for each  $i$  and  $j$ .

The details connected with this model are presented in (Dębicka and Marciniuk 2014). The formulas are determined based on the matrix notations described in (Dębicka 2013).

The course of the process is described by the matrix  $D$ . This matrix is defined as the following formula

$$D = \begin{pmatrix} \mathbf{P}^T(0) \\ \mathbf{P}^T(1) \\ \dots \\ \mathbf{P}^T(n) \end{pmatrix} \in R^{(n+1) \times (N \cdot N)}, \tag{4}$$

where  $\mathbf{P}(0) = (1, 0, 0, \dots, 0)^T \in R^{(N \cdot N)}$  denotes the initial distribution vector,  $\mathbf{P}^T(t) = \mathbf{P}^T(0) \prod_{k=0}^{t-1} \mathbf{Q}^{(X \circ Y)}(t)$ , where  $N$  is the state number and  $n$  is a contract duration.

This course takes into consideration both spouses' survival times.

The annuity benefit is determined by means of the following formula (Dębicka et al. 2022)

$$\ddot{b} = \frac{\alpha W}{\mathbf{M}^T (\mathbf{I} - \mathbf{I}_{n+1} \mathbf{I}_{n+1}^T - \mathbf{I}_1 \mathbf{I}_1^T) \mathbf{D} (\mathbf{S} - \mathbf{J}_4) + \mathbf{1}}, \tag{5}$$

where  $\mathbf{M}^T = (1, v, v^2, \dots, v^n) \in R^{n+1}$  denotes a discounting factor, which relates to the interest rate. The auxiliary vectors are used, where  $\mathbf{I}_t = (0, 0, \dots, 1, 0, \dots, 0)^T \in R^{(n+1) \times 1}$ ,  $\mathbf{S} = (1, 1, \dots, 1) \in R^{(N \cdot N) \times 1}$ ,  $\mathbf{J}_i = (0, 0, \dots, 1, 0, \dots, 0)^T \in R^{(N \cdot N) \times 1}$  and  $\mathbf{I}$  is an identity matrix with  $n+1$  rows and columns. In this case,  $N = 2$ . The formula for the reverse annuity contract benefit in the classical notation is presented in (e.g., Marciniuk et al. 2020).

In the second step, the health insurance benefits are estimated. The health insurance premium is separately payable for the husband and wife. The further lifetime of a couple is described by the model presented in Figure 1. The part of the pension allocated to the payment of health insurance is divided between the spouses proportionally to the  $R$  parameter, which denotes a fraction of the men in the population who fall ill with the disease that is subject to dread disease insurance within a year. This parameter determines the risk of morbidity connected with a critical disease in the population of men. The decomposition of the annuity benefit is represented by the formula  $\ddot{b} = \beta \ddot{b} + R(1 - \beta) \ddot{b} + (1 - R)(1 - \beta) \ddot{b}$ . Thus, the annual health insurance premium is given by the following formulas: for a husband  $p_Y = R(1 - \beta) \ddot{b}$  and for a wife  $p_X = (1 - R)(1 - \beta) \ddot{b}$ . The health premiums are paid separately by the spouses for when they are healthy. Health insurance benefits are paid individually to the spouses depending on their health condition (at the moment of the diagnosis of a critical illness and in the event of deterioration of health), for a wife  $c_X$  and for a husband  $c_Y$ . The formula for calculating the value of the health benefits is given by (compare Appendix A)

$$c = \frac{\mathbf{M}^T \text{Diag}(\mathbf{C}_{out} \mathbf{D}^T) \mathbf{S}}{\mathbf{M}^T \text{Diag}(\mathbf{C}_{in}^{(1)} \mathbf{D}^T) \mathbf{S}}. \tag{6}$$

Matrix  $D$ , describing the probability structure, is calculated on the basis of the matrix of transition probabilities given by (2) for the husband and wife separately. Thus, the number of columns in matrix  $D$  is reduced to  $N$ . The matrix of all premiums paid during the term of the contract is denoted as  $\mathbf{C}_{out} \in R^{(n+1) \times N}$ . The matrix  $\mathbf{C}_{in}^{(1)} \in R^{(n+1) \times N}$  equals 1 in the second and the fourth columns.

In summary, during the duration of the contract, spouses receive the following benefits:

- $\ddot{b} - p_X - p_Y$  (Both spouses are healthy);
- $\ddot{b} - p_Y$  (The husband is healthy; the wife is ill or dead);



- $\ddot{b} - p_X$  (The husband is ill or dead; the wife is healthy);
- $\ddot{b} - p_X + c_Y$  (The husband has become sick or his health has worsened; the wife is healthy);
- $\ddot{b} - p_Y - p_X$  (The husband is healthy; the wife has become sick or her health has worsened);
- $\ddot{b} + c_Y$  (The husband has become sick or his health has worsened; the wife is sick or dead);
- $\ddot{b} + c_X$  (The husband is sick or dead; the wife has become sick or her health has worsened);
- $\ddot{b} + c_X + c_Y$  (The spouses have both become sick or their health has worsened).

Therefore, the spouses receive annuity benefits reduced by the health insurance premium (if they are healthy). In the event of a diagnosis or a deterioration of their health state(s), the spouses receive health insurance benefits individually. From the moment the diagnosis is made, the spouse does not pay the health insurance premium.

### 3.2. Two Individual Contracts

An individual version of the contract implies that a married couple signs two separate contracts. In this case, the premiums and benefits connected with health insurance and the reverse annuity contract are paid separately. The benefits of the reverse annuity contract are financed and determined based on the part of the property value owned by a given spouse (usually 50%) and is equal to  $\pi_{II} = (0.5 \cdot W) \cdot \alpha$ . The annuity benefits are calculated separately for a husband  $\ddot{b}_Y$  and for a wife  $\ddot{b}_X$ . The sum of the annuity benefits in an individual version is higher than the annuity benefit paid under the marriage contract. However, in the individual case, the annuity benefits are paid independently to each of the spouses only until their death. Thus, in the event of the death of one of the spouses, the surviving partner will be deprived of the part of the dead spouse.

Within individual contracts, the spouses may separately determine a part of an annuity benefit, which will be spent on the health insurance premium by setting the value of the reverse annuity parameter  $\beta$ . Therefore, the amounts of the health insurance premium are equal to  $p_Y^{II} = (1 - \beta_Y) \cdot \ddot{b}_Y$  and  $p_X^{II} = (1 - \beta_X) \cdot \ddot{b}_X$  for a husband and a wife, respectively. In the next step, the health insurance benefits are estimated for a husband  $c_Y^{II}$  and a wife  $c_X^{II}$ .

The cash flows related to the individual version of the contract are separately modelled for a husband and a wife. The probability structure describing the further lifetime of a wife is denoted by  $\mathbf{Q}_h^{[x]}$ , whereas for a husband it is denoted by  $\mathbf{Q}_h^{[y]}$  (compare (2)).

The benefits are obtained separately for individual models for a wife and a husband in three steps. In the first step, the annuity benefits are obtained separately on the basis of the further lifetime for a wife and a husband using the following formula (compare Dębicka and Marciniuk 2014)

$$\ddot{b} = \frac{\alpha(0.5W)}{\mathbf{M}^T(\mathbf{I} - \mathbf{I}_{n+1}\mathbf{I}_{n+1}^T)\mathbf{D}\mathbf{J}_1'} \tag{7}$$

which is calculated separately for a husband and a wife. The matrix  $\mathbf{D}$  is obtained on the basis of the transition probability matrix given by (1) separately for the spouses.

In the second step, the health insurance benefits are obtained using the model described in Figure 1. The health insurance benefits are calculated in a similar way as in the case of a marriage version of the contract using the Formula (6).

During two individual contracts, the spouses receive the following benefits:

- $\ddot{b}_X - p_X^{II} + \ddot{b}_Y - p_Y^{II}$  (Both spouses are healthy);
- $\ddot{b}_Y - p_Y^{II} + \ddot{b}_X$  (The husband is healthy; the wife is ill);
- $\ddot{b}_Y - p_Y^{II}$  (The husband is healthy; the wife is dead);
- $\ddot{b}_X - p_X^{II} + \ddot{b}_Y$  (The husband is ill; the wife is healthy);
- $\ddot{b}_X - p_X^{II}$  (The husband is dead; the wife is healthy);

- $\ddot{b}_X - p_X^{II} + \ddot{b}_Y + c_Y^{II}$  (The husband has become sick or his health has worsened; the wife is healthy);
- $\ddot{b}_X + c_X^{II} + \ddot{b}_Y - p_Y^{II}$  (The husband is healthy; the wife has become sick or her health has worsened);
- $\ddot{b}_Y + c_Y^{II} + \ddot{b}_X$  (The husband has become sick or his health has worsened; the wife is sick);
- $\ddot{b}_Y + c_Y^{II}$  (The husband has become sick or his health has worsened; the wife is dead);
- $\ddot{b}_X + c_X^{II} + \ddot{b}_Y$  (The husband is sick; the wife has become sick or her health has worsened);
- $\ddot{b}_X + c_X^{II}$  (The husband is dead; the wife has become sick or her health has worsened);
- $\ddot{b}_X + c_X^{II} + \ddot{b}_Y + c_Y^{II}$  (The spouses have become sick or their health has worsened).

The spouses receive annuity benefits reduced by the health insurance premium (if they are alive and healthy). The death of a spouse deprives the living partner of a part of their pension. Payments of dread disease insurance benefits are made in the event of the diagnosis of a serious illness or deterioration of the health state. From the moment the diagnosis is made, the spouse does not pay the health insurance premium.

#### 4. Results

The retirees' family, economic, and health conditions determine their financial needs. The state of their health of has an important impact on their financial resources. Simultaneously, the elderly often have limited funds to buy an insurance policy against the risk of a dread disease. Therefore, the described contract may be an alternative to obtaining additional pensioner's income and funds for treatment and palliative care in the case of a dread disease.

A sample of numerical examples using two variants of the equity release contract is presented in this section. In the *marriage variant*, it is assumed that a married couple has a property valued at  $W$ . Together, they decide to sign the marriage reverse annuity and spend the percentage of  $1 - \beta$  on withdrawing from both of their dread disease insurance plans. In the *individual variant*, the spouses also have a property valued at the same sum  $W$ . They decide to sign the individual reverse annuity contracts for  $0.5W$  and allocate the percentage of  $1 - \beta$  to buy dread disease insurance from their part of the benefit payment. In this case, they have two separate contracts.

The empirical examples are based on actual data. The examples concern the risk of lung cancer mortality and the morbidity of the Lower Silesia population (one of Poland's voivodeships). The calculations are made separately for male and female groups based on datasets provided by the National Cancer Registry for the Lower Silesia Region (Wojciechowska and Didikowska 2014) and the Lower Silesia Department of the National Health Fund (the public payer in Poland). The available, unpublished data covers the period between 2006 and 2011. The transition probabilities matrixes are obtained from the life expectancy tables described in the (Dębicka and Zmyślona 2016). The maximum survival time in a critical condition was 4 years. A critical stage entails diagnosis with distant metastases or an inoperable life-threatening tumour (Dębicka and Zmyślona 2019).

Let  $x$  denote a woman's age and  $y$  denote a man's age. The percentage of a real value of property  $\alpha = 50$  and the real value of a property is 100,000 euro. This assumption allows for the easy rescaling of the results obtained in case of different property prices. The discounting factor  $v$ , which is calculated from formula  $v = (1 + i)^{-1}$ , is closely connected with the analysed period of the disease and a constant long-term interest rate  $i = 0.0579$ . The interest rate  $i$  was estimated on the basis of actual Polish market data based on the yield to maturity on the fixed interest bonds and Treasury bills from 2008 (Zmyślona and Marciniuk 2020). The percentage of men who fall ill with lung cancer in 2008, denoted by  $R$ , equals 0.691. We consider couples aged 65, 70, 75, 80, and 85, as well as mixed-age couples. We analyse a case when  $\beta = 0.99$ , which means that 1% of the annuity is intended for paying the premium. It is obvious that the levels of the premium may vary (Zmyślona

and Marciniuk 2020). Annual marriage reverse annuity payments, single benefit payments of illness insurance, which could be paid twice, and premiums in euros are calculated using self-developed programs written in MATLAB. Table 3 presents all cash flows in both variants.

**Table 3.** Cash flows in two variants of equity release.

First Variant—Marriage Contract							
$x = y$	Annuity $\ddot{b}$	Benefit for Woman $c_x$	Benefit for Man $c_y$	Premium for Woman $p_x$	Premium for Man $p_y$		
65	4085.4	8417.8	4945.8	12.6	28.2		
70	4628.7	9558.4	5413.1	14.3	32.0		
75	5458.0	11,903.0	6803.5	16.9	37.7		
80	6717.5	15,593.0	10,410.0	20.8	46.4		
85	8612.4	20,667.0	16,076.0	26.6	59.5		
Second Variant—Individual Contracts							
$x = y$	Annuity for Woman $\ddot{b}_x$	Annuity for Man $\ddot{b}_y$	Total Annuity $\ddot{b}$	Benefit for Woman $c_x$	Benefit for Man $c_y$	Premium for Woman $p_x$	Premium for Man $p_y$
65	2287.5	2786.9	5074.4	15,253.0	4882.6	22.9	27.9
70	2657.4	3257.5	5914.9	17,760.0	5513.1	26.6	32.6
75	3252.0	3952.1	7204.1	22,956.0	7129.0	32.5	39.5
80	4195.0	4984.0	9179.0	31,513.0	11,177.0	42.0	49.9
85	5673.3	6487.8	12,161.1	44,058.0	17,526.0	56.7	64.9

In both variants, an increase in all the benefits was observed. This growth is positively correlated with the spouses' age. Women pay lower premiums because the incidence of lung cancer is lower in the female population. Consequently, the critical illness insurance benefit is significantly higher for females than for the male population. It is obvious that in case of another disease the relation may be opposite. The percentage of the morbidity rate under consideration  $R$  determines the distribution of the premium. In the marriage variant, the total benefit payment of the equity release is lower than the same benefit in the individual variant. The differences in the benefit payments increase with the rise in the spouses' age and reach almost 30% for  $x = y = 85$ . The same relationship is observable for illness benefits; however, women pay an almost two-fold higher premium in the second option, thus receiving twice as much payment. This shows that women could pay half the premium for this particular disease, which means  $0.5 \cdot (1 - \beta)$ , to receive the same chronic illness benefit payment. The critical illness insurance payment for men is comparable in both cases. This is due to the fact that the premium is divided according to the morbidity risk in the marriage contract. In both variants, the spouses receive a reverse annuity as a yearly payment, and they can allow themselves to buy critical illness insurance. The annuity income is not significantly reduced because the level of the annual premium is very low. Whereas the health insurance benefit allows for the acquirement of significant financial resources for health care, treatment, and improvement of the quality of life in the case of dread disease.

As shown by the above analysis, it can be assumed that the second variant is more favourable for the spouses. Therefore, let us analyse the relative increases in the number of total benefits payments received by spouses at 65, 75, and 85 years depending on their further life expectancy, which is presented in Figures 2–4.

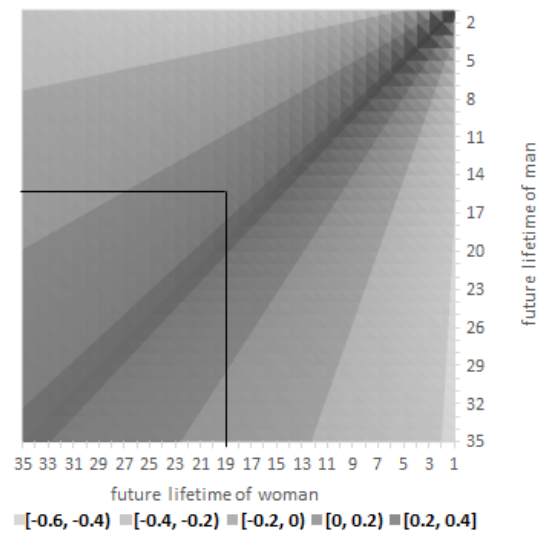


Figure 2. The relative increase in benefits in both variants for spouses aged 65.

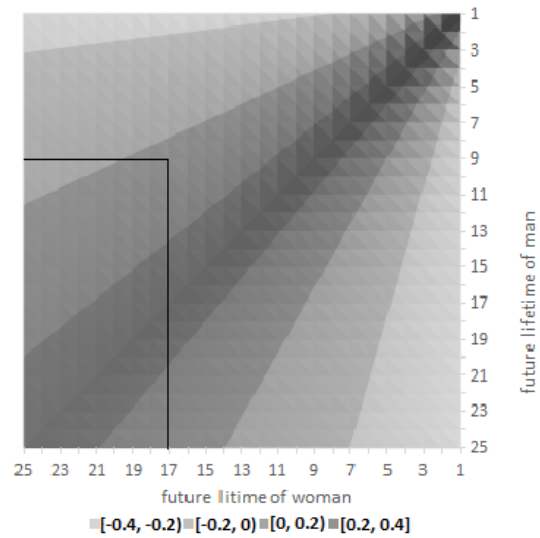


Figure 3. The relative increase in benefits in both variants for spouses aged 75.

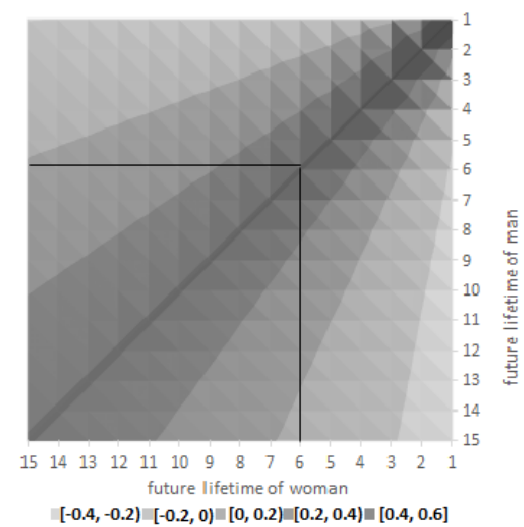


Figure 4. The relative increase in benefits in both variants for spouses aged 85.

The dark colours indicate positive differences, i.e., a higher total benefit in the second variant. Lighter colours represent negative differences, which means a higher total payment in the first option. The graphs also show lines that illustrate the future life expectancy of people at the considered age. The future life expectancy of a 65, 75, and 85-year-old woman is 19.04, 16.54, and 6.05 years, respectively, and for men of the same age it is 14.73, 9.22, and 5.2 years, respectively.

It is visible that the similar age of the spouses' death determines the total benefit of the equity release contract, increasing up to 40%, and results in a higher dread disease insurance payment in the event of sickness. This means that the second variant of the contract is more favourable for spouses. When one of the spouses dies earlier than the other, and the second spouse lives a long time, the first variant is more beneficial. This can especially be observed for women, who become widows more often than men, and receive a higher marital annuity, even by over 40%. However, when the husband additionally suffers from a serious illness, in the considered case of lung cancer, the health insurance benefit is of a similar amount. The lower left rectangle shows situations when the spouses live longer than their future life expectancy. Then, the probability of receiving a higher benefit increases. When spouses are at 65, the higher payments of the second variant of the contract occur in 40.7% of cases, of which 61.3% concerns a situation when spouses live above their future life expectancy. These numbers increase with the rise in the spouses' age. The second variant's higher benefit occurs in 50.9% and 63.1% when spouses are at 75 and 85 years, respectively. Of these situations, 72.3% are placed in the left lower rectangle.

The spouses are not very often at the same age; therefore, Tables 4 and 5 present the amount of the annuity and health insurance benefits in both variants for spouses of various ages. An exemplary five-year age difference between the spouses was assumed.

**Table 4.** Total benefits of equity release for various ages of spouses in two variants of contract.

First Variant—Marriage Contract						Second Variant—Individual Contracts					
	<i>x</i>						<i>X</i>				
<i>y</i>	65	70	75	80	85	<i>y</i>	65	70	75	80	85
65	4085	4394	4717	5006	5226	65	5074	5444	6039	6982	8460
70	4224	4629	5084	5526	5889	70	5545	5915	6510	7453	8931
75	4340	4839	5458	6115	6712	75	6240	6610	7205	8147	9625
80	4428	5008	5788	6718	7647	80	7272	7641	8237	9179	10,657
85	4493	5133	6050	7241	8612	85	8775	9145	9740	10,683	12,161

It is evident that the reverse annuity payments increase alongside the rise in the spouses' age in both cases. However, the benefits are higher in the second variant. The difference in both benefits increases for younger women and decreases for older women as the men's age increases. For example, when  $x = 65$  and  $y = 85$ , the benefit is over 95% higher, and when men become older, it is only 41.2% higher, but it is still huge. It is not difficult to notice that the payment is lower for a younger wife and an older husband than contrariwise. In the case of men, age has a higher impact on the benefit. This relates to the fact that men live shorter lives than women.

The critical illness insurance payment for women is almost always twice as high in the second variant. Wives do not share the premium for this insurance with husbands. This benefit does not depend on the age of the second spouse in the second case. For a younger husband and an older wife, the dread disease insurance payment is higher in the first contract; thus, it is more favourable for men.

**Table 5.** Critical illness insurance payment depending on the age of spouses in two variants of contract.

First Variant—Marriage Contract						Second Variant—Individual Contract
Critical illness insurance payment for man						
	x					different x
y	65	70	75	80	85	
65	4946	5320	5711	6060	6326	4882.6
70	4940	5413	5946	6463	6887	5513.1
75	5410	6032	6804	7622	8367	7129.3
80	6862	7761	8970	10,410	11,850	11,171
85	8387	9581	11,293	13,516	16,076	17,526
Critical illness insurance payment for woman						
	x					different y
y	65	70	75	80	85	
65	8418	9075	10,288	11,619	12,539	15,253
70	8704	9558	11,088	12,828	14,132	17,760
75	8943	9993	11,903	14,194	16,107	22,596
80	9124	10,342	12,624	15,593	18,350	31,513
85	9258	10,600	13,194	16,807	20,667	44,058

## 5. Discussion and Conclusions

The article presents two variants of a contract that is a combination of an equity release and critical illness insurance. This new proposition could protect against the effects of longevity. The contractual cash flow and different scenarios were analysed. It cannot be clearly stated which of these contracts is better. Assuming a pessimistic variant, when both spouses fall ill and die quickly, neither option will be beneficial, although they will receive the health insurance benefit payment. If one spouse dies earlier and the other lives healthily longer than the life expectancy, then the marriage variant is more favourable. On the other hand, assuming an optimistic version, when both spouses live together in good health for a long time, particularly above the life expectancy, the individual variant is definitely a better choice. Of course, when concluding a contract, most people do not consider the risk of death. The knowledge of the genetic burden concerning a critical illness can help to more adequately adjust the client's contract. It was shown that both contracts offer high benefits payments, but for women, who become widows more often, the marriage contract is more beneficial. The calculations show a very significant gender impact on the amount of the benefits. In the considered examples, net premiums and net benefits were presented. All the contracts have additional costs. Combining arrangements into one lowers the costs; hence, the first variant seems to be a cheaper option than buying individual policies.

The introduction of critical illness insurance into the contract has caused two important constraints for the presented model. Firstly, the limited critical illness survival time (a four-year period for lung cancer as described in the studied example) enforced the application of the periodic life tables. Secondly, since an estimation of the relationship between cancer risk for both the spouses seems impossible, we have assumed independence between the spouses' life expectancies. A justification of the possibility of such an assumption of independence is presented in (Zmysłona and Marciniuk 2020).

The studied contracts could be used as part of longevity risk management policies as they provide an instrument of incentive for the retirees to use housing resources to improve their living conditions. In case of a dread disease, additional funds are likely to improve the quality of life during the treatment. The two variants of contracts, marriage and individual, could help pensioners to manage their home budget.

**Author Contributions:** Conceptualization, A.M. and B.Z.; methodology, A.M. and B.Z.; software, A.M.; validation A.M. and B.Z.; formal analysis, A.M. and B.Z.; investigation, A.M. and B.Z.; resources, B.Z.; data curation, A.M. and B.Z.; writing—original draft preparation, A.M. and B.Z.; writing—review and editing, A.M. and B.Z.; visualization, A.M. and B.Z.; supervision, A.M. and B.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** The project was financed by the Ministry of Science and Higher Education in Poland under the programme “Regional Initiative of Excellence” 2019–2022 project number 015/RID/2018/19 total funding amount 10 721 040,00 PLN.

**Data Availability Statement:** Analysis are based on: 1. The Life Tables from Dębicka, J., Zmysłona, B. (2016). Construction of Multi-State Life Tables for Critical Illness Insurance—Influence of Age and Sex on the Incidence of Health Inequalities. *Śląski Przegląd Statystyczny*, 14, 41–63. <https://doi.org/10.15611/sps.2016.14.03>; 2. The National Cancer Registry for the Lower Silesia Region from Wojciechowska, U., and Didkowska, J. (2014). Zachorowania i zgony na nowotwory złośliwe w Polsce. *Krajowy Rejestr Nowotworów. Centrum Onkologii—Instytut im. Marii Skłodowskiej—Curie*. [www.onkologia.org.pl/raporty/](http://www.onkologia.org.pl/raporty/). 3. The Life Table for the Lower Silesia from 2008 (followed from Statistical Offices).

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A

The formula for the value of the health benefits is given by

$$c = \frac{\mathbf{M}^T \text{Diag}(\mathbf{C}_{\text{out}} \mathbf{D}^T) \mathbf{S}}{\mathbf{M}^T \text{Diag}(\mathbf{C}_{\text{in}}^{(1)} \mathbf{D}^T) \mathbf{S}} \tag{A1}$$

**Proof.** Let  $\mathbf{C} = \mathbf{C}_{\text{in}} - \mathbf{C}_{\text{out}}$ , where  $\mathbf{C}_{\text{out}} \in R^{(n+1) \times N}$  denotes the matrix of all premiums paid during the term of the contract and  $\mathbf{C}_{\text{in}} \in R^{(n+1) \times N}$  is the matrix of all the benefits, which are paid at state 2 and 4 in the amount  $c$ . Hence

$$\mathbf{C}_{\text{in}} = \begin{bmatrix} 0 & c & 0 & c & \dots & 0 \\ 0 & c & 0 & c & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & c & 0 & c & \dots & 0 \end{bmatrix} = c \cdot \begin{bmatrix} 0 & 1 & 0 & 1 & \dots & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 1 & 0 & 1 & \dots & 0 \end{bmatrix} = c \cdot \mathbf{C}_{\text{in}}^{(1)}.$$

We use the equivalence principle for the multistate insurance (cf. Dębicka 2013), which has the following form

$$\mathbf{M}^T \text{Diag}(\mathbf{C} \mathbf{D}^T) \mathbf{S} = 0,$$

where  $\mathbf{M}^T = (1, v, v^2, \dots, v^n) \in R^{n+1}$ ,  $\mathbf{S} = (1, 1, \dots, 1) \in R^{(N \cdot N) \times 1}$  and  $\mathbf{D} \in R^{(n+1) \times (N \cdot N)}$  describes the matrix probability (such as in Formula (4)).

Hence

$$\mathbf{M}^T \text{Diag}((\mathbf{C}_{\text{in}} - \mathbf{C}_{\text{out}}) \mathbf{D}^T) \mathbf{S} = 0$$

and

$$\mathbf{M}^T \text{Diag}(\mathbf{C}_{\text{out}} \mathbf{D}^T) \mathbf{S} = \mathbf{M}^T \text{Diag}(\mathbf{C}_{\text{in}} \mathbf{D}^T) \mathbf{S}$$

Due to  $\mathbf{C}_{\text{in}} = c \cdot \mathbf{C}_{\text{in}}^{(1)}$ , we receive

$$\mathbf{M}^T \text{Diag}(\mathbf{C}_{\text{out}} \mathbf{D}^T) \mathbf{S} = c \cdot \mathbf{M}^T \text{Diag}(\mathbf{C}_{\text{in}}^{(1)} \mathbf{D}^T) \mathbf{S}$$

which completes the proof.  $\square$

## References

- Borg, Christel, Ingalill Hallberg, and Kerstin Blomqvist. 2006. Life satisfaction among older people (65+) with reduced self-care capacity: The relationship to social, health and financial aspects. *Journal of Clinical Nursing* 15: 607–18. [CrossRef] [PubMed]
- Borg, Christel, Cecilia Fagerström, Cristian Balducci, Vanessa Burholt, Dieter Ferring, Germain Weber, Clare Wenger, Göran Holst, and Ingalill Hallberg. 2008. Life Satisfaction in 6 European. *Geriatric Nursing* 29: 48–57. [CrossRef] [PubMed]
- Chen, Mengni, Yuanyuan Fu, and Qingsong Chang. 2021. Life Satisfaction among Older Adults in Urban China: Does Gender Interact with Pensions, Social Support and Self-Care Ability? *Ageing and Society* 2021: 1–20. [CrossRef]
- Dębicka, Joanna. 2013. An approach to the study of multistate insurance contracts. *Applied Stochastic Models in Business Industry* 29: 224–40. [CrossRef]
- Dębicka, Joanna, Agnieszka Marciniuk, and Beata Zmyślona. 2015. Combination reverse annuity contract and critical illness insurance. Paper presented at 18th AMSE, Applications of Mathematics and Statistics in Economics, Conference Proceedings, Jindřichův Hradec, Czech Republic, September 2–6; pp. 1–9. [CrossRef]
- Dębicka, Joanna, Agnieszka Marciniuk, and Stanisław Heilpern. 2022. Modelling marital annuity contract in a stochastic economic environment. *Statistika 2. in press*.
- Dębicka, Joanna, and Agnieszka Marciniuk. 2014. Comparison of Reverse Annuity Contract and Reverse Mortgage on the Polish Market. Paper presented at 17th AMSE, Applications of Mathematics and Statistics in Economics, Conference Proceedings, Wrocław, Poland, August 27–31; pp. 55–64. [CrossRef]
- Dębicka, Joanna, and Beata Zmyślona. 2016. Construction of Multi-State Life Tables for Critical Illness Insurance—Influence of Age and Sex on the Incidence of Health Inequalities. *Śląski Przegląd Statystyczny* 14: 41–63. [CrossRef]
- Dębicka, Joanna, and Beata Zmyślona. 2018. A Multiple State Model for Premium Calculation When Several Premium-Paid States Are Involved. *Central European Journal of Economic Modelling and Econometrics* 10: 27–52.
- Dębicka, Joanna, and Beata Zmyślona. 2019. Modelling of Lung Cancer Survival Data for Critical Illness Insurances. *Statistical Methods and Applications* 28: 723–47. [CrossRef]
- Dhaene, Jan, Els Godecharle, Katrien Antonio, Michel Denuit, and Hamza Hanbali. 2017. Lifelong health insurance covers with surrender values: updating mechanism in the presence of medical inflation. *ASTIN Bulletin* 47: 803–36. [CrossRef]
- Eurostat. 2020a. Housing Statistics. Available online: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Housing\\_statistics/pl](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Housing_statistics/pl) (accessed on 13 March 2021).
- Eurostat. 2020b. Population Structure and Ageing. Available online: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population\\_structure\\_and\\_ageing/pl](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_structure_and_ageing/pl) (accessed on 13 March 2021).
- Haberman, Steven, and Ermanno Pitacco. 2018. *Actuarial Models for Disability Insurance. Actuarial Models for Disability Insurance*. London: Routledge. [CrossRef]
- Hanewald, Katja, Thomas Post, and Michael Sherris. 2016. Portfolio Choice in Retirement—What Is The Optimal Home Equity Release Product? *Journal of Risk and Insurance* 83: 421–46. [CrossRef]
- Institute and Faculty of Actuaries (IFoA). 2019. Equity Release Mortgages (ERM). *British Actuarial Journal* 24: e25. [CrossRef]
- Korenman, Sanders, Dahlia Remler, and Rosemary Hyson. 2021. Health insurance and poverty of the older population in the United States: The importance of a health inclusive poverty measure. *Journal of the Economics of Ageing* 18: 100297. [CrossRef]
- Lee, Yung-Tsung, and Tianxiang Shi. 2021. Valuation of Reverse Mortgages with Surrender: A Utility Approach. *J Real Estate Finan Econ. The Journal of Real Estate Finance and Economics*. [CrossRef]
- Marciniuk, Agnieszka, Emília Zimková, Vlastimil Farkašovský, and Colin W. Lawson. 2020. Valuation of Equity Release Contracts in Czech Republic, Republic of Poland and Slovak Republic. *Prague Economic Papers* 29: 505–21. [CrossRef]
- Mullings, Beverley, and Chris Hamnett. 1992. Equity Release Schemes and Equity Extraction by Elderly Households in Britain. *Ageing & Society* 12: 413–42.
- Paolucci, Francesco, Przemysław M. Sowa, Manuel García-Goñi, and Henry Ergas. 2015. Mandatory Aged Care Insurance: A Case for Australia. *Ageing and Society* 35: 231–45. [CrossRef]
- Pitacco, Ermanno. 2014. *Health Insurance. Basic Actuarial Models*. Berlin/Heidelberg: Springer.
- Shao, Adam W., Katja Hanewald, and Michael Sherris. 2015. Reverse Mortgage Pricing and Risk Analysis Allowing for Idiosyncratic House Price Risk and Longevity Risk. *Insurance: Mathematics and Economics* 63: 76–90. [CrossRef]
- Sharma, Tripti, Declan French, and Donald McKillop. 2020. Risk and Equity Release Mortgages in the UK. *The Journal of Real Estate Finance and Economics* 64: 274–97. [CrossRef]
- Toussaint, Janneke, and Marja Elsinga. 2009. Exploring ‘Housing Asset-Based Welfare’. Can the UK Be Held Up as an Example for Europe? *Housing Studies* 24: 669–92. [CrossRef]
- Webber, James. 1993. Financing Healthcare for the Elderly. *Journal of the Staple Inn Actuarial Society* 33: 77–122. [CrossRef]
- Wojciechowska, U., and J. Didikowska. 2014. *Zachorowania i Zgony Na Nowotwory Złośliwe w Polsce. Krajowy Rejestr Nowotworów. Centrum Onkologii—Instytut Im. Marii Skłodowskiej—Curie*. Available online: [www.onkologia.org.pl/raporty2014](http://www.onkologia.org.pl/raporty2014) (accessed on 13 March 2021).
- Zmyślona, Beata, and Agnieszka Marciniuk. 2020. Financial Protection for the Elderly—Contracts Based on Equity Release and Critical Health Insurance. *European Research Studies Journal* XXIII: 867–82. [CrossRef]





Article

# A Comparison of Macaulay Approximations

Stefanos C. Orfanos 

Department of Risk Management & Insurance, J. Mack Robinson College of Business, Georgia State University, Atlanta, GA 30302, USA; sorfanos@gsu.edu; Tel.: +1-404-413-7485

**Abstract:** We discuss several known formulas that use the Macaulay duration and convexity of commonly used cash flow streams to approximate their net present value, and compare them with a new approximation formula that involves hyperbolic functions. Our objective is to assess the reliability of each approximation formula under different scenarios. The results in this note should be of interest to actuarial candidates and educators as well as analysts working in all areas of actuarial practice.

**Keywords:** Macaulay duration; Macaulay convexity; net present value of cash flows

## 1. Introduction

Actuaries and actuarial science students at universities all over the world are familiar with approximation formulas for the present value of cash flow streams using some notion of cash flow duration or convexity. For example, the syllabus of Exam FM of the US-based Society of Actuaries includes the topic of approximations using the Macaulay and modified duration and convexity, while the UK-based Institute and Faculty of Actuaries in its material for exam CM1 mentions approximations derived from a Taylor series expansion.

Beside the academic and pedagogical interest in such approximation formulas, one may also consider the practical value in the management of interest rate risk. Although abundant computing power has enabled firms to implement elaborate immunization strategies that incorporate multi-factor stochastic interest rate models, non-parallel yield curve shifts, and complicated asset and liability characteristics, the restrictions posed by a simplistic valuation model are not unreasonable if rates remain historically low, yield curves stay relatively flat, and we can control the potential errors. Indeed, it may be helpful to know which approximation formula proves to be the most reliable, and to use it as a quick validation tool when time constraints preclude the use of a more sophisticated approach.

Alps (2017) describes a realistic scenario involving an investment actuary and her CEO, where the use of an approximation formula would be warranted or even necessitated. This is especially true in today's world of fast-changing rates, when companies have to react almost instantly to benchmark fund rates and quantitative tightening decisions by the Federal Reserve or other central banks.

In this note, we discuss several known formulas that use the Macaulay duration and convexity of commonly used cash flow streams to approximate their net present value, and compare them with a new approximation formula that involves hyperbolic functions. In addition to annuities, dividend stocks, and bonds, we also consider the cases of negative payments and embedded options to perform a deeper assessment. The notions of effective duration and convexity are defined in the next section and used to price the embedded options. Our objective is to measure the reliability of each approximation formula under different scenarios. As alluded to earlier, we only consider parallel interest rate shocks to flat yield curves.

**Citation:** Orfanos, Stefanos C. 2022.

A Comparison of Macaulay Approximations. *Risks* 10: 153. <https://doi.org/10.3390/risks10080153>

Academic Editors: Ermanno Pitacco and Annamaria Olivieri

Received: 13 June 2022

Accepted: 15 July 2022

Published: 29 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

### 1.1. Literature Review

The idea of using a bond's duration to approximate changes to its price goes back to Macaulay (1938). Some authors credit Fischer and Weil (1971) with the publication of the first duration–convexity approximation formula. Enhancements of that formula by controlling the missing higher-order terms and incorporating passage of time were announced in Jarjir and Rakotondratsimba (2008, 2012), though the resulting formulas contain parameters that are unintuitive and hard to calibrate. A conceptually simpler formula was given in Barber (1995), and independently in Livingston and Zhou (2005) for the modified duration, which was subsequently generalized to a duration–convexity model in Tchuindjo (2008). Further work in Barber and Dandapani (2017) considered negative-yielding bonds, and Johansson (2012) added passage of time. A separate duration–convexity formula appears in Alps (2017) and is applied to an empirical study of basic immunization strategies in Nie et al. (2021), while a very recent paper by Barber (2022) further generalizes a duration–convexity approximation by introducing an additional ‘compounding’ parameter. Finally, traditional approximations have been implemented in statistical analysis packages; see Lee (2021) for R code.

### 1.2. Notation

We denote the net present value of a cash flow stream by  $P$ . The interest rate  $r$  is annualized and continuously compounded (i.e., force of interest).  $\Delta r$  is the change in interest rates from the initial value  $r_0$  to  $r$ . Finally, the annual discount factor  $v$  is by definition equal to  $e^{-r}$ .

Throughout the remainder of this paper and for convenience, assume  $r_0 = 1.6\%$ , which is approximately the yield on the 10-year T-bond at the beginning of this year.

## 2. Materials and Methods

Recall that the *Macaulay duration* of a stream of cash flows  $\{CF_{t_j}\}_{j=1}^n$  being paid at future times  $\{t_j\}_{j=1}^n$  is defined by

$$d = \frac{\sum_{j=1}^n CF_{t_j} v^{t_j} t_j}{\sum_{j=1}^n CF_{t_j} v^{t_j}} = -\frac{dP/dr}{P},$$

while its *Macaulay convexity* is given by

$$c = \frac{\sum_{j=1}^n CF_{t_j} v^{t_j} t_j^2}{\sum_{j=1}^n CF_{t_j} v^{t_j}} = \frac{d^2P/dr^2}{P}.$$

We do not consider the modified duration here because the Macaulay duration has a more intuitive interpretation (being the ‘average’ timing of the cash flows) and tends to result in tighter approximations for non-negative rates.

In the case of bonds with embedded options, it will be necessary to price the value of the option using a simple Black model. Recall that the pricing formula for, say, a European call option with expiration at time  $t$  and strike  $K$  is

$$V = v^t (P\Phi(d_1) - K\Phi(d_2))$$

where  $\Phi$  represents the standard normal c.d.f. and the quantities  $d_{1,2}$  are given by

$$d_{1,2} = \frac{\ln(P/K)}{\sigma\sqrt{t}} \pm \frac{\sigma\sqrt{t}}{2}$$

with  $\sigma$  the bond price volatility. We also need more flexible measures of bond duration and convexity. To that end, define the *effective duration* by means of

$$d_e = -\frac{P(r_0 + \Delta r) - P(r_0 - \Delta r)}{2P_0 \Delta r}$$

and the *effective convexity* as

$$c_e = \frac{P(r_0 + \Delta r) - 2P_0 + P(r_0 - \Delta r)}{P_0 (\Delta r)^2}.$$

### 2.1. Fischer–Weil’s Approximation

This follows immediately from Calculus and the definitions above.

$$\frac{\Delta P}{P_0} \approx -d_0 \Delta r + \frac{c_0}{2} (\Delta r)^2. \quad (1)$$

It is assumed that the Macaulay duration and convexity are computed at rate  $r_0$ , hence the subscripts.

### 2.2. Barber’ 1995 Approximation

Instead of the second-order Taylor polynomial of  $P$ , we consider the first-order Taylor polynomial of  $\ln P$ , thus obtaining

$$\ln P \approx \ln P_0 - d_0 \Delta r,$$

thus

$$P \approx P_0 e^{-d_0 \Delta r}. \quad (2)$$

Unlike the first-order Taylor polynomial in  $P$  that has no convexity, the functional form of Barber’s approximation bequeaths it with a certain degree of positive curvature. This leads to good approximation results whenever  $c_0 \approx d_0^2$  and poor performance for  $c_0 < 0$ .

### 2.3. Tchuindjo’ Approximation

Similar to Barber’s approximation, but involving the second-order Taylor polynomial of  $\ln P$

$$\ln P \approx \ln P_0 - d_0 \Delta r + \frac{c_0 - d_0^2}{2} (\Delta r)^2, \quad (3)$$

from which one solves for  $P$ . The added quadratic term gives better results in cases where  $c_0 - d_0^2$  is non-trivial, but may still introduce large errors whenever  $c_0 < 0$ .

### 2.4. Alps’ Approximation

The approximation formula and its derivation can be found in Alps (2017). The central idea in the derivation of this approximation is to compute a Taylor polynomial for the current value of the cash flow stream at time  $t = d_0$ . This choice results in high accuracy in situations where  $d_0 > 0$ , and less so for  $d_0 < 0$ .

We have rewritten it below in terms of continuously compounded interest rates.

$$P \approx P_0 e^{-d_0 \Delta r} \left( 1 + \frac{c_0 - d_0^2}{2} (e^{\Delta r} - 1)^2 \right). \quad (4)$$

### 2.5. Hyperbolic Approximation

We have not encountered this approximation formula in the literature and we assume its derivation is presented here for the first time. Consider the homogeneous differential equation  $P'' - c_0 P = 0$  that mimics the definition of Macauley convexity given earlier in this note. Its general solution takes the form

$$P = a e^{\sqrt{c_0} r} + b e^{-\sqrt{c_0} r}$$

(do not worry for the time being about the case  $c_0 < 0$ .) Setting  $P(r_0) = P_0$  and  $P'(r_0) = -d_0 P_0$ , which is a reformulation of the definition of Macauley duration, one obtains the approximation

$$P \approx P_0 \left( \frac{1}{2} \left( 1 - \frac{d_0}{\sqrt{c_0}} \right) e^{\sqrt{c_0} \Delta r} + \frac{1}{2} \left( 1 + \frac{d_0}{\sqrt{c_0}} \right) e^{-\sqrt{c_0} \Delta r} \right)$$

which can be rewritten as

$$P \approx P_0 \left( \cosh(\sqrt{c_0} \Delta r) - \frac{d_0}{\sqrt{c_0}} \sinh(\sqrt{c_0} \Delta r) \right). \tag{5}$$

The well-known trig identities

$$\cosh(i\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2} = \cos \theta, \quad \sinh(i\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2} = i \sin \theta$$

can be used in the case  $c_0 < 0$  to obtain

$$P \approx P_0 \left( \cos(\sqrt{|c_0|} \Delta r) - \frac{d_0}{\sqrt{|c_0|}} \sin(\sqrt{|c_0|} \Delta r) \right),$$

which is useful whenever there is a computational issue with imaginary numbers.

In the next section, we demonstrate that the hyperbolic approximation is less prone to errors than other well-known approximations in situations where the duration and/or convexity are negative. Recall that negative convexity cash flow streams can be easily constructed with the addition of negative cash flows to a stream of positive payments, when considering callable bonds, or with mortgage-backed securities due to the prepayment option in conventional residential mortgages.

### 3. Results

Approximation formulas such as Equations (1)–(5) should ideally be intuitive and behave well in special cases.

- (i) The simplest cash flow is cash, which has trivial duration and convexity and is unaffected by interest rate changes. By substituting  $d_0 = c_0 = 0$  or taking the corresponding limit in the case of (5) and using the fact that

$$\lim_{\theta \rightarrow 0} \frac{\sinh \theta}{\theta} = \lim_{\theta \rightarrow 0} \frac{\sin \theta}{\theta} = 1,$$

we obtain  $P = P_0$  as expected.

- (ii) Next, take a zero-coupon bond, for which  $c_0 = d_0^2$ : Except for Fischer–Weil’s approximation, the rest reduce to Barber’s approximation, which is perfectly accurate in this case. On the other hand, the error associated with Fischer–Weil’s approximation increases with the bond duration and it can be as high as 0.56% for a 30-year zero-coupon bond after a 100 bp increase in rates.
- (iii) For a convexity-hedged ( $c_0 \rightarrow 0$ ) portfolio, Fischer–Weil’s and the hyperbolic approximations reduce to the first-order approximation  $\Delta P \approx -d_0 P_0 \Delta r$ . The corresponding results for the other approximation formulas are not as intuitive and their accuracy relative to the above approximation cannot be determined without additional details about the cash flow characteristics.

We supplement the theoretical tests above with some concrete examples.

- (iv) Consider a 10-year annuity-immediate with annual payments of 10. Recall that our assumption is  $r_0 = 1.6\%$  and compute the present value  $P_0 = 10a_{\overline{10}|} = 91.6728$ . Another easy calculation gives the Macauley duration and convexity as  $d_0 = 5.3681$  and  $c_0 = 37.0554$ , respectively.

In Table 1, the exact value of  $P$  is computed using the same formula as for  $P_0$  but at the new continuously compounded rate.

**Table 1.** PV of 10-year annuity with annual payments of 10.

$\Delta r$	Exact	Fischer-Weil	Barber	Tchuindjo	Alps	Hyperbolic
−100 bp	96.7682	96.7637	96.7283	96.7682	96.7669	96.7668
−80 bp	95.7207	95.7184	95.6954	95.7207	95.7199	95.7199
−60 bp	94.6876	94.6866	94.6735	94.6876	94.6871	94.6873
−40 bp	93.6687	93.6684	93.6625	93.6687	93.6685	93.6686
−20 bp	92.6639	92.6638	92.6623	92.6639	92.6638	92.6639
0 bp	91.6728	91.6728	91.6728	91.6728	91.6728	91.6728
20 bp	90.6954	90.6954	90.6939	90.6954	90.6953	90.6954
40 bp	89.7313	89.7316	89.7254	89.7313	89.7311	89.7314
60 bp	88.7804	88.7813	88.7672	88.7804	88.7800	88.7807
80 bp	87.8425	87.8447	87.8193	87.8425	87.8417	87.8432
100 bp	86.9173	86.9216	86.8815	86.9173	86.9162	86.9186

We can observe that Tchuindjo’s approximation outperforms the rest, while Barber’s lags behind for sizable rate changes. This was to be expected, since Barber’s approximation lacks a convexity term and will not do well in cases when  $c_0 - d_0^2$  is non-trivial. On the other hand, Alps’ and the hyperbolic approximations are roughly equally accurate behind Tchuindjo’s.

- (v) Next, add a negative cash flow at time 20. We have chosen  $CF_{20} = -120$  in the example below; the net present value is  $P_0 = 4.5349$  and the Macaulay duration and convexity are  $d_0 = -275.7817$  and  $c_0 = -6,936.8498$ , respectively.

Looking at Table 2 below, it may come as a surprise that the approximations by Tchuindjo and Alps blow up completely. However, we can provide a simple mathematical explanation for the bizarre behavior. Whenever  $c_0 < 0$ , the quadratic term of these two approximations that includes the expression  $c_0 - d_0^2$  has the potential to be extremely influential. As  $\Delta r$  increases, said term can overwhelm the baseline value  $P_0$  and the linear term, resulting in large errors. Barber’s approximation exhibits the opposite weakness: missing a quadratic term implies that the negative convexity is not accounted for at all. In fact, for suitable  $CF_{20}$ , we can obtain  $d_0 = 0$ , in which case Barber’s approximation fails to yield any results.

**Table 2.** NPV of 10-year annuity with annual payments of 10 and a payment of −120 at time 20.

$\Delta r$	Exact	Fischer-Weil	Barber	Tchuindjo	Alps	Hyperbolic
−100 bp	−9.6622	−9.5445	0.2877	0.0045	−0.8684	−8.0590
−80 bp	−6.5366	−6.4769	0.4994	0.0351	−0.7850	−5.7162
−60 bp	−3.5601	−3.5352	0.8669	0.1946	−0.3873	−3.2151
−40 bp	−0.7266	−0.7193	1.5048	0.7747	0.5372	−0.6250
−20 bp	1.9698	1.9707	2.6123	2.2128	2.1924	1.9824
0 bp	4.5349	4.5349	4.5349	4.5349	4.5349	4.5349
20 bp	6.9742	6.9733	7.8725	6.6685	6.6069	6.9619
40 bp	9.2929	9.2859	13.6664	7.0357	4.8785	9.1962
60 bp	11.4960	11.4726	23.7243	5.3262	−10.6004	11.1758
80 bp	13.5885	13.5335	41.1846	2.8930	−64.7471	12.8461
100 bp	15.5748	15.4686	71.4950	1.1275	−215.8388	14.1608

We conclude this example by mentioning that the top-performing approximation is Fischer–Weil’s, while the hyperbolic approximation is second-best.

(vi) Let us now consider a dividend stock, whose theoretical price is computed using Gordon’s dividend discount model

$$P = \frac{D}{r - g}$$

with  $D$  representing next year’s dividend and  $g$  its constant continuously compounded growth rate in perpetuity. A quick calculation gives  $d = (r - g)^{-1}$  and  $c = 2(r - g)^{-2}$ ; for  $g = 0.6\%$  we obtain  $d_0 = 100$  and  $c_0 = 20,000$ . Assume  $D = 1$ .

Some of the results in Table 3 may appear counterintuitive at first sight. Gordon’s model suggests that  $P$  has an inverse relationship to  $r$ ; however, all approximations except for Alps’ and Barber’s eventually produce a divergent estimate for  $P$  as  $\Delta r$  increases. However, this is explained by the fact that we are attempting to trace a hyperbola using quadratic curves. Moreover, all approximations struggle to keep up with  $P$  for large negative values of  $\Delta r$ .

**Table 3.** Price of a dividend stock with  $D = 1$  and  $g = 0.6\%$ .

$\Delta r$	Exact	Fischer-Weil	Barber	Tchuindjo	Alps	Hyperbolic
−100 bp	n/a	300.0000	271.8282	448.1689	403.4619	354.6482
−80 bp	500.0000	244.0000	222.5541	306.4854	291.5285	269.3175
−60 bp	250.0000	196.0000	182.2119	218.1472	213.9771	205.6762
−40 bp	166.6667	156.0000	149.1825	161.6074	160.7412	158.5990
−20 bp	125.0000	124.0000	122.1403	101.8813	101.8813	101.8813
0 bp	100.0000	100.0000	100.0000	100.0000	100.0000	100.0000
20 bp	83.3333	84.0000	81.8731	83.5270	83.4590	83.7590
40 bp	71.4286	76.0000	67.0320	72.6149	72.2257	74.2635
60 bp	62.5000	76.0000	54.8812	65.7047	64.4487	70.7488
80 bp	55.5556	84.0000	44.9329	61.8783	58.8586	72.9319
100 bp	50.0000	100.0000	36.7879	60.6531	54.6026	80.9885

Overall, Alps’ approximation proves to be the most dependable for moderate changes in the interest rates.

(vii) Next, consider a 10-year bond with a coupon rate of  $r_0$  and face value of 100. A quick calculation yields  $d_0 = 9.3151$  and  $c_0 = 90.6932$ .

It turns out that the last three approximation formulas clearly outperform the rest, with Tchuindjo’s having a slight advantage over Alps’ and the hyperbolic approximation, as evidenced from Table 4. The subpar performance of Fischer–Weil on bonds is one of the reasons why this approximation is not widely utilized, despite its robustness in cases such as (v).

It has been shown empirically that although investment-grade bonds fall in price when interest rates rise, that is not necessarily the case with high-yield bonds whose duration can be negative due to default risk; see Melentyev and Yu (2020). For such bonds, care should be exercised when using the approximations by Tchuindjo or Alps.

**Table 4.** PV of 10-year par bond with coupon rate  $r_0$ .

$\Delta r$	Exact	Fischer-Weil	Barber	Tchuindjo	Alps	Hyperbolic
−100 bp	109.7839	109.7686	109.7628	109.7843	109.7836	109.7830
−80 bp	107.7501	107.7423	107.7368	107.7503	107.7499	107.7497
−60 bp	105.7556	105.7523	105.7482	105.7557	105.7555	105.7554
−40 bp	103.7996	103.7986	103.7963	103.7996	103.7995	103.7995
−20 bp	101.8813	101.8812	101.8805	101.8813	101.8813	101.8813
0 bp	100.0000	100.0000	100.0000	100.0000	100.0000	100.0000
20 bp	98.1550	98.1551	98.1542	98.1550	98.1550	98.1550
40 bp	96.3456	96.3465	96.3425	96.3455	96.3454	96.3456
60 bp	94.5710	94.5742	94.5642	94.5709	94.5707	94.5712
80 bp	92.8306	92.8381	92.8188	92.8304	92.8301	92.8310
100 bp	91.1238	91.1383	91.1056	91.1234	91.1229	91.1246

(viii) Finally, assume the bond is callable, with the European call strike set at  $K = 101.0000$  and bond price volatility  $\sigma = 8\%$ . The call is exercised a year ahead of the bond's maturity and has price  $V = 9.9431$ , which is subtracted from the price of a conventional bond to arrive at the callable bond price. Using  $\Delta r = 20$  bp in the calculation of the effective duration and convexity, we obtain  $d_e = 5.3333$  and  $c_e = 50.3993$ . The positive convexity may surprise some readers, but note that the convexity turns negative when the interest rate gets closer to 0 and the bond price approaches the strike.

It is important to observe in Table 5 that none of the approximation formulas can consistently outperform the rest, if our objective is to estimate the full range of prices for such a bond. In effect, we are trying to approximate a function with an inflection point using quadratic curves, and thus significant approximation errors are inevitable.

**Table 5.** NPV of 10-year par bond with coupon rate  $r_0$ , callable for 101 at  $t = 9$ .

$\Delta r$	Exact	Fischer-Weil	Barber	Tchuindjo	Alps	Hyperbolic
−100 bp	95.0594	95.0868	94.9903	95.0946	95.0913	95.0910
−80 bp	94.0367	94.0445	93.9824	94.0485	94.0464	94.0466
−60 bp	93.0197	93.0204	92.9853	93.0220	93.0209	93.0213
−40 bp	92.0148	92.0144	91.9987	92.0149	92.0144	92.0146
−20 bp	91.0265	91.0265	91.0226	91.0266	91.0265	91.0266
0 bp	90.0569	90.0569	90.0569	90.0569	90.0569	90.0569
20 bp	89.1053	89.1053	89.1014	89.1053	89.1052	89.1053
40 bp	88.1692	88.1720	88.1560	88.1715	88.1710	88.1717
60 bp	87.2437	87.2568	87.2207	87.2552	87.2541	87.2559
80 bp	86.3227	86.3597	86.2953	86.3559	86.3540	86.3577
100 bp	85.3991	85.4808	85.3797	85.4735	85.4705	85.4768

The only useful conclusion is that the hyperbolic approximation is never the worst one, since it tends to be “sandwiched” between other approximations.

#### 4. Discussion

We have established through a number of theoretical considerations and concrete examples that the accuracy of various Macaulay approximations can vary widely. Approximations that outperform in one case turn out to be unreliable in another case. The hyperbolic approximation, introduced in this paper, exhibited modest errors in most cases and thus the most reliability among the five approximations studied.

We can envision a variety of uses for the results presented here:

- To perform expeditious interest risk calculations by practitioners;



- As a study note to gain insight into risk management concepts that are tested in the actuarial examinations in the US and Europe;
- As potential areas of student research or as assigned projects that utilize real financial data in actuarial science classes taught by academics.

There is also potential to expand the scope of this study by incorporating non-flat yield curves, key rate durations, passage of time, and more complex financial instruments.

**Funding:** This research received no external funding.

**Acknowledgments:** The author wishes to thank the anonymous referees for their constructive comments. Partial results were presented at the 56th Actuarial Research Conference (see Orfanos 2021) and this work benefitted from the attendees' questions.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

- Alps, Robert. 2017. Using Duration and Convexity to Approximate Change in Present Value. SOA Financial Mathematics Study Note FM-24-17. pp. 1–18. Available: <https://www.soa.org/globalassets/assets/Files/Edu/2017/fm-duration-convexity-present-value.pdf> (accessed on 22 July 2022).
- Barber, Joel R. 1995. A Note on Approximating Bond Price Sensitivity Using Duration and Convexity. *The Journal of Fixed Income* 4: 95–98. [CrossRef]
- Barber, Joel R. 2022. Power Law Bond Price and Yield Approximation. *The Journal of Risk Finance* 23: 14–31. [CrossRef]
- Barber, Joel R., and Krishnan Dandapani. 2017. Interest Rate Risk in a Negative Yielding World. *Frontiers in Finance and Economics* 14: 1–19.
- Fischer, Lawrence, and Roman L. Weil. 1971. Coping with the Risk of Interest-Rate Fluctuations: Returns to Bondholders from Naive and Optimal Strategies. *The Journal of Business* 44: 406–31. [CrossRef]
- Livingston, Miles, and Lei Zhou. 2005. Exponential Duration: A More Accurate Estimate of Interest Rate Risk. *The Journal of Financial Research* 28: 343–61. [CrossRef]
- Jarjir, Souad L., and Yves Rakotondratsimba. 2008. Revisiting the Bond Duration-Convexity Approximation. *SSRN Electronic Journal*. [CrossRef]
- Jarjir, Souad L., and Yves Rakotondratsimba. 2012. Enhancement of the Bond Duration-Convexity Approximation. *International Journal of Economics and Finance* 4: 115–25.
- Johansson, Bo. 2012. A Note on Approximating Bond Returns Allowing for Both Yield Change and Time Passage. Munich Personal RePEc Archive Paper No. 92607. Available online: [https://mpira.ub.uni-muenchen.de/92607/1/MPRA\\_paper\\_92607.pdf](https://mpira.ub.uni-muenchen.de/92607/1/MPRA_paper_92607.pdf) (accessed on 22 July 2022).
- Lee, Sang-Heon. 2021. Percent Change of Bond Price Using Duration and Convexity in R. *R bloggers*. Available online: <https://www.r-bloggers.com/2021/09/percent-change-of-bond-price-using-duration-and-convexity-in-r/> (accessed on 22 July 2022).
- Macaulay, Fblackerick R. 1938. *Some Theoretical Problems Suggested by Movement of Interest Rates, Bonds, Yields, and Stock Prices in the United States Since 1856*. New York: National Bureau of Economic Research, pp. 45–53.
- Melentyev, Oleg, and Eric Yu. 2020. *Bank of America/Merrill Lynch High Yield Strategy*. New York: Bank of America Merrill Lynch, September 4.
- Nie, Junming, Zhihao Wu, Shiyao Wang, and Ye Chen. 2021. Duration Strategy for Bond Investment Based on an Empirical Study. *Advances in Economics, Business and Management Research* 203: 1726–30.
- Orfanos, Stefanos. 2021. A Comparison of Macaulay approximations. Research talk at the 56th Actuarial Research Conference. Available online: <https://csh.depaul.edu/academics/mathematical-sciences/about/arc2021/Documents/Parallel-Session-Abstracts.pdf> (accessed on 22 July 2022).
- Tchuindjo, Leonard. 2008. An Accurate Formula for Bond-Portfolio Stress Testing. *The Journal of Risk Finance* 9: 262–77. [CrossRef]

# Reverse Sensitivity Analysis for Risk Modelling

Silvana M. Pesenti 

Department of Statistical Sciences, University of Toronto, Toronto, ON M5S 3G3, Canada;  
silvana.pesenti@utoronto.ca

**Abstract:** We consider the problem where a modeller conducts sensitivity analysis of a model consisting of random input factors, a corresponding random output of interest, and a baseline probability measure. The modeller seeks to understand how the model (the distribution of the input factors as well as the output) changes under a stress on the output's distribution. Specifically, for a stress on the output random variable, we derive the unique stressed distribution of the output that is closest in the Wasserstein distance to the baseline output's distribution and satisfies the stress. We further derive the stressed model, including the stressed distribution of the inputs, which can be calculated in a numerically efficient way from a set of baseline Monte Carlo samples and which is implemented in the R package SWIM on CRAN. The proposed reverse sensitivity analysis framework is model-free and allows for stresses on the output such as (a) the mean and variance, (b) any distortion risk measure including the Value-at-Risk and Expected-Shortfall, and (c) expected utility type constraints, thus making the reverse sensitivity analysis framework suitable for risk models.

**Keywords:** distortion risk measures; expected utility; Wasserstein distance; robustness and sensitivity analysis; model uncertainty

## 1. Introduction

Sensitivity analysis is indispensable for model building, model interpretation, and model validation, as it provides insight into the relationship between model inputs and outputs. A key tool used for sensitivity analysis are sensitivity measures, that assign to each model input a score, representing an input factor's ability to explain the variability of a model output's summary statistic; see Saltelli et al. (2008) and Borgonovo and Plischke (2016) for an in-depth review. One of the most widely used output summary statistic is the variance, which gives rise to sensitivity measures, e.g., the Sobol indices, that apportion the uncertainty in the output's variance to input factors. In many applications, such as reliability management and financial and insurance risk management, however, the variance is not the output statistic of concern and instead quantile-base measures are used; indicatively, see Asimit et al. (2019); Fissler and Pesenti (2022); Maume-Deschamps and Niang (2018); Tsanakas and Millossovich (2016). Furthermore, typical for financial risk management applications is that model inputs are subject to distributional uncertainty. Probabilistic (or global) sensitivity measures, however, tacitly assume that the model's distributional assumptions are correctly specified; indeed, sensitivity measures based on the difference between conditional (on a model input) and unconditional densities (of the output) are termed "common rationale" Borgonovo et al. (2016). Examples include indices, such as Borgonovo's sensitivity measures Borgonovo (2007), the  $f$ -sensitivity index Rahman (2016), and sensitivity indices based on the Cramér–von Mises distance Gamboa et al. (2018), we also refer to Plischke and Borgonovo (2019) for a detailed overview and to Gamboa et al. (2020) for estimation of these sensitivity measures. Recently, Plischke and Borgonovo (2019) define sensitivity measures that depend only on the copula between input factors, whereas Pesenti et al. (2021) propose a sensitivity measure based on directional derivatives that take dependence between input factors into account. Estimating these sensitivities, however, may render difficult in application where joint observations are

**Citation:** Pesenti, Silvana M. 2022. Reverse Sensitivity Analysis for Risk Modelling. *Risks* 10: 141. <https://doi.org/10.3390/risks10070141>

Academic Editors: Annamaria Olivieri and Ermanno Pitacco

Received: 26 May 2022

Accepted: 7 July 2022

Published: 18 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

scarce, e.g., insurance portfolios, and their interpretation may be limited as dependence structures are commonly specified by expert opinions Denuit et al. (2006).

We consider an alternative sensitivity analysis framework proposed in Pesenti et al. (2019) that (a) considers statistical summaries relevant to risk management, (b) applies to models subject to distributional uncertainty, thus instead of relying on correctly specified distributions from which to calculate sensitivity measures we derive alternative distributions that fulfil a specific probabilistic stress and are “closest” to the baseline distribution; and (c) studies reverse sensitivity measures. Differently to the framework proposed in Pesenti et al. (2019) who use the Kullback–Leibler divergence to quantify the closedness of probability measures, in this work we consider the Wasserstein distance of order two to measure the distance between distribution functions. The Wasserstein distance allows for more flexibility in the choice of stresses including survival probabilities (via quantiles) used in reliability analysis, risk measures employed in finance and insurance, and utility functions relevant for decision under ambiguity.

Central to the reverse sensitivity analysis framework is a *baseline model*, the 3-tuple  $(X, g, \mathbb{P})$ , consisting of random input factors  $X = (X_1, \dots, X_n)$ , an aggregation function  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  mapping input factors to a univariate output  $Y = g(X)$ , and a probability measure  $\mathbb{P}$ . The methodology has been termed *reverse sensitivity analysis* by Pesenti et al. (2019) since it proceeds in a reverse fashion to classical sensitivity analysis where input factors are perturbed and the corresponding altered output is studied. Indeed, in the reverse sensitivity analysis proposed by Pesenti et al. (2019) a stress on the output’s distribution is defined and changes in the input factors are monitored. The quintessence of the sensitivity analysis methodology is, however, not confined to stressing the output’s distribution, it is also applicable to stressing an input factor and observing the changes in the model output and in the other inputs. Throughout the exposition, we focus on the reverse sensitivity analysis that proceeds via the following steps:

- (i) Specify a stress on the baseline distribution of the output;
- (ii) Derive the unique stressed distribution of the output that is closest in the Wasserstein distance and fulfils the stress;
- (iii) The stressed distribution induces a canonical Radon–Nikodym derivative  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ ; a change of measures from the baseline  $\mathbb{P}$  to the stressed probability measure  $\mathbb{Q}^*$ ;
- (iv) Calculate sensitivity measures that reflect an input factors’ change in distribution from the baseline to the stressed model.

Sensitivity testing using divergence measures—in the spirit of the reverse sensitivity methodology—has been studied by Cambou and Filipović (2017) using  $f$ -divergences on a finite probability space; by Pesenti et al. (2019) and Pesenti et al. (2021) using the Kullback–Leibler divergence; and Makam et al. (2021) consider a discrete sample space combined with the  $\chi^2$ -divergence. It is however known that the set of distribution functions with finite  $f$ -divergence, e.g., the Kullback–Leibler and  $\chi^2$  divergence—around a baseline distribution function depends on the baseline’s tail-behaviour, thus the choice of  $f$ -divergence should be chosen dependent on the baseline distribution Kruse et al. (2019). The Wasserstein distance on the contrary, automatically adapts to the baseline distribution function in that the Wasserstein distance penalises dissimilar distributional features such as different tail behaviour Bernard et al. (2020). The Wasserstein distance has enjoyed numerous applications to quantify distributional uncertainty, see, e.g., Blanchet and Murthy (2019) and Bernard et al. (2020) for applications to financial risk management. In the context of uncertainty quantification, Moosmüller et al. (2020) utilise the Wasserstein distance to elicit the (uncertain) aggregation map  $g$  from the distributional knowledge of the inputs and outputs. Fort et al. (2021) utilises the Wasserstein distance to introduce global sensitivity indices for computer codes whose output is a distribution function. In this manuscript we use the Wasserstein distance as it allows for different stresses compared to the Kullback–Leibler divergence. Indeed, the Wasserstein distance allows for stresses on any distortion risk measures, while the Kullback–Leibler divergence only allow for stresses on risk measures which are Value-at-Risk (VaR) and VaR and Expected Shortfall jointly, see Pesenti et al. (2019).

This paper is structured as follows: In Section 2, we state the notation and definitions necessary for the exposition. Section 3 introduces the optimisation problems and we derive the unique stressed distribution function of the output which has minimal Wasserstein distance to the baseline output’s distribution and satisfies a stress. The considered stresses include constraints on risk measures, quantiles, expected utilities, and combinations thereof. In Section 4, we characterise the canonical Radon–Nikodym derivative, induced by the stressed distribution function, and study how input factors’ distributions change when moving from the baseline to the stressed model. An application of the reverse sensitivity analysis is demonstrated on a mixture model in Section 5.

All proofs are delegated to Appendix A.

### 2. Preliminaries

Throughout we work on a measurable space  $(\Omega, \mathcal{A})$  and denote the sets of distribution functions with finite second moment by

$$\mathcal{M} = \left\{ G: \mathbb{R} \rightarrow [0, 1] \mid G \text{ non-decreasing, right-continuous, } \lim_{x \searrow -\infty} G(x) = 0, \right. \\ \left. \lim_{x \nearrow +\infty} G(x) = 1, \text{ and } \int x^2 dG(x) < +\infty \right\},$$

and the corresponding set of square-integrable (left-continuous) quantile functions by

$$\check{\mathcal{M}} = \left\{ \check{G} \in \mathbb{L}^2([0, 1]) \mid \check{G} \text{ non-decreasing \& left-continuous} \right\}.$$

or any distribution function  $G \in \mathcal{M}$ , we denote its corresponding (left-continuous) quantile function by  $\check{G} \in \check{\mathcal{M}}$ , that is  $\check{G}(u) = \inf\{y \in \mathbb{R} \mid G(y) \geq u\}$ ,  $u \in [0, 1]$ , with the convention that  $\inf \emptyset = +\infty$ . We measure the discrepancy between distribution functions on the real line using the Wasserstein distance of order 2, defined as follows.

**Definition 1** (Wasserstein Distance). *The Wasserstein distance (of order 2) between two distribution functions  $F_1$  and  $F_2$  is defined as Villani (2008)*

$$W_2(F_1, F_2) = \inf_{\pi \in \Pi(F_1, F_2)} \left\{ \left( \int_{\mathbb{R}^2} |z_1 - z_2|^2 \pi(dz_1, dz_2) \right)^{\frac{1}{2}} \right\},$$

where  $\Pi(F_1, F_2)$  denotes the set of all bivariate probability measures with marginal distributions  $F_1$  and  $F_2$ , respectively.

The Wasserstein distance is the minimal quadratic cost associated with transporting the distribution  $F_1$  to  $F_2$  using all possible couplings (bivariate distributions) with fixed marginals  $F_1$  and  $F_2$ . The Wasserstein distance admits desirable properties to quantify model uncertainty such as the comparison of distributions with differing support, e.g. with the empirical distribution function. Moreover it is symmetric and forms a metric on the space of probability measures; we refer to Villani (2008) for an overview and properties of the Wasserstein distance. It is well known (Dall’Aglio 1956) that for distributions on the real line, the Wasserstein distance admits the representation

$$W_2(F_1, F_2) = \left( \int_0^1 |\check{F}_1(u) - \check{F}_2(u)|^2 du \right)^{\frac{1}{2}}.$$

### 3. Deriving the Stressed Distribution

Throughout this section we assume that the modeller’s *baseline model* is the 3-tuple  $(\mathbf{X}, g, \mathbb{P})$  consisting of a random vector of input factors  $\mathbf{X} = (X_1, \dots, X_n)$ , an aggregation function  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  mapping input factors to a (for simplicity) univariate output  $Y = g(\mathbf{X})$ ,

and a probability measure  $\mathbb{P}$ . The baseline probability measure  $\mathbb{P}$  reflects the modeller’s (statistical and expert) knowledge of the distribution of  $X$  and we denote the distribution function of the output by  $F(y) = \mathbb{P}(Y \leq y)$ . The modeller then performs reverse sensitivity analysis, that is tries to understand how prespecified stresses/constraints on the output distribution  $F$ , e.g., an increase in jointly its mean and standard deviation or a risk measures such as the Value-at-Risk (VaR) or Expected Shortfall (ES), affects the baseline model, e.g., the joint distribution of the input factors. For this, we first define the notion of a *stressed distribution*. Specifically, for given constraints we call a solution to the optimisation problem

$$\arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to stresses/constraints on } G, \tag{1}$$

a stressed distribution. In problem (1), the baseline distribution  $F$  is fixed and we seek over all alternative distributions  $G \in \mathcal{M}$  the one who satisfies the stress(es) and which has smallest Wasserstein distance to  $F$ . The solution to problem (1)–the stressed distribution–may be interpreted as the most “plausible” distribution function arising under adverse circumstances. Examples of stresses and constraints considered in this work include an increase (decrease), compared to their corresponding values under the reference probability  $\mathbb{P}$ , in e.g., the mean, mean and standard deviation, distortion risk measures, and utility functions, and combinations thereof.

Next, we recall the concept of weighted isotonic projection which is intrinsically connected to the solution of optimisation problem (1); indeed the stressed quantile functions can be uniquely characterised via weighted isotonic projections.

**Definition 2** (Weighted Isotonic Projection Barlow et al. (1972)). *The weighted isotonic projection  $\ell^{\uparrow w}$  of a function  $\ell \in \mathbb{L}^2([0, 1])$  with weight function  $w: [0, 1] \rightarrow [0, +\infty)$ ,  $w \in \mathbb{L}^2([0, 1])$ , is its weighted projection onto the set of non-decreasing and left-continuous functions in  $\mathbb{L}^2([0, 1])$ . That is, the unique function satisfying*

$$\ell^{\uparrow w} = \arg \min_{h \in \mathcal{M}} \int_0^1 (\ell(u) - h(u))^2 w(u) du .$$

When the weight function is constant, i.e.,  $w(x) \equiv c$ ,  $c > 0$ , we write  $\ell^{\uparrow}(\cdot) = \ell^{\uparrow c}(\cdot)$ , as in this case the isotonic projection is indeed independent of  $c$ . The weighted isotonic projection admits not only a graphical interpretation as the non-decreasing function that minimises the weighted  $\mathbb{L}^2$ -distance from  $\ell$  but has also a discrete counterpart: the weighted isotonic regression Barlow et al. (1972). Numerically efficient algorithms for calculating weighted isotonic regressions are available, e.g., the R package `isotone` De Leeuw et al. (2010).

In the next sections, we solve problem (1) for different choices of constraints. Specifically, for risk measures constraints (Section 3.1), integral constraints (Section 3.2), Value-at-Risk constraints (Section 3.3), and expected utility constraint (Section 3.4), and in Section 3.5 we consider ways to smooth stressed distributions. Using these stressed distributions, we derive the stressed probability measures in Section 4 and study how a stress on the output is reflected on the input distribution(s).

### 3.1. Risk Measure Constraints

This section considers stresses on distortion risk measures, that is we derive the unique stressed distribution that satisfies an increase and/or decrease of distortion risk measures while minimising the Wasserstein distance to the baseline distribution  $F$ .

**Definition 3** (Distortion Risk Measures). Let  $\gamma \in \mathbb{L}^2([0, 1])$  be a square-integrable function with  $\gamma: [0, 1] \rightarrow [0, +\infty)$  and  $\int_0^1 \gamma(u) du = 1$ . Then the distortion risk measure  $\rho_\gamma$  with distortion weight function  $\gamma$  is defined as

$$\rho_\gamma(G) = \int_0^1 \check{G}(u)\gamma(u) du \quad \text{for } G \in \mathcal{M}. \tag{2}$$

The above definition of distortion risk measures makes the assumption that positive realisations are undesirable (losses) while negative realisations are desirable (gains). The class of distortion risk measures includes one of the most widely used risk measures in financial risk management, the Expected Shortfall (ES) at level  $\alpha \in [0, 1)$  (also called Tail Value-at-Risk), with  $\gamma(u) = \frac{1}{1-\alpha} \mathbb{1}_{\{u>\alpha\}}$ , see, e.g., Acerbi and Tasche (2002). The often used risk measure Value-at-Risk (VaR), while admitting a representation given in (2), has a corresponding weight function  $\gamma$  that is not square-integrable. We derive the solution to optimisation problem (1) with a VaR constraint in Section 3.3.

**Theorem 1** (Distortion Risk Measures). Let  $r_k \in \mathbb{R}$ ,  $\rho_{\gamma_k}$  be a distortion risk measure with weight function  $\gamma_k$  and assume there exists a distribution function  $\check{G} \in \mathcal{M}$  satisfying  $\rho_{\gamma_k}(\check{G}) = r_k$  for all  $k \in \{1, \dots, d\}$ . Then, the optimisation problem

$$\arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to} \quad \rho_{\gamma_k}(G) = r_k \quad k = 1, \dots, d, \tag{3}$$

has a unique solution given by

$$\check{G}^*(u) = \left( \check{F}(u) + \sum_{k=1}^d \lambda_k \gamma_k(u) \right)^\uparrow, \tag{4}$$

where the Lagrange multipliers  $\lambda_k$  are such that the constraints are fulfilled, that is  $\rho_{\gamma_k}(\check{G}^*) = r_k$  for all  $k = 1, \dots, d$ .

In the above theorem, and also in later results, we assume that there exists a distribution function which satisfies all constraints. This assumption is not restrictive and requires that, particularly, multiple constraints are chosen carefully, e.g., imposing that  $\int_0^1 \check{G}(u) du > \frac{1}{1-\alpha} \int_\alpha^1 \check{G}(u) du$  for  $\alpha \in (0, 1)$ , i.e., the mean being larger than the  $ES_\alpha$ , cannot be fulfilled by any distribution function; thus, a combination of stresses not of interest to a modeller.

We observe that the optimal quantile function is the isotonic projection of a weighted linear combination of the baseline’s quantile function  $\check{F}$  and the distortion weight functions of the risk measures. A prominent group of risk measures is the class of coherent risk measures, that are risk measures fulfilling the properties of monotonicity, positive homogeneity, translation invariance, and sub-additivity; see Artzner et al. (1999) for a discussion and interpretation. It is well-known that a distortion risk measure is coherent, if and only if, its distortion weight function  $\gamma(\cdot)$  is non-decreasing Kusuoka (2001). For the special case of a constraint on a coherent distortion risk measure that results in a larger risk measure compared to the baseline’s, we obtain an analytical solution without the need to calculate an isotonic projection.

**Proposition 1** (Coherent Distortion Risk Measure). If  $\rho_\gamma$  is a coherent distortion risk measure and  $r \geq \rho_\gamma(F)$ , then optimisation problem (3) with  $d = 1$  has a unique solution given by

$$\check{G}^*(u) = \check{F}(u) + \frac{r - \rho_\gamma(F)}{\int_0^1 (\gamma(u))^2 du} \gamma(u).$$

We illustrate the stressed distribution functions for constraints on distortion risk measures in the next example. Specifically, we look at the  $\alpha$ - $\beta$  risk measures which are a parametric family of distortion risk measures.

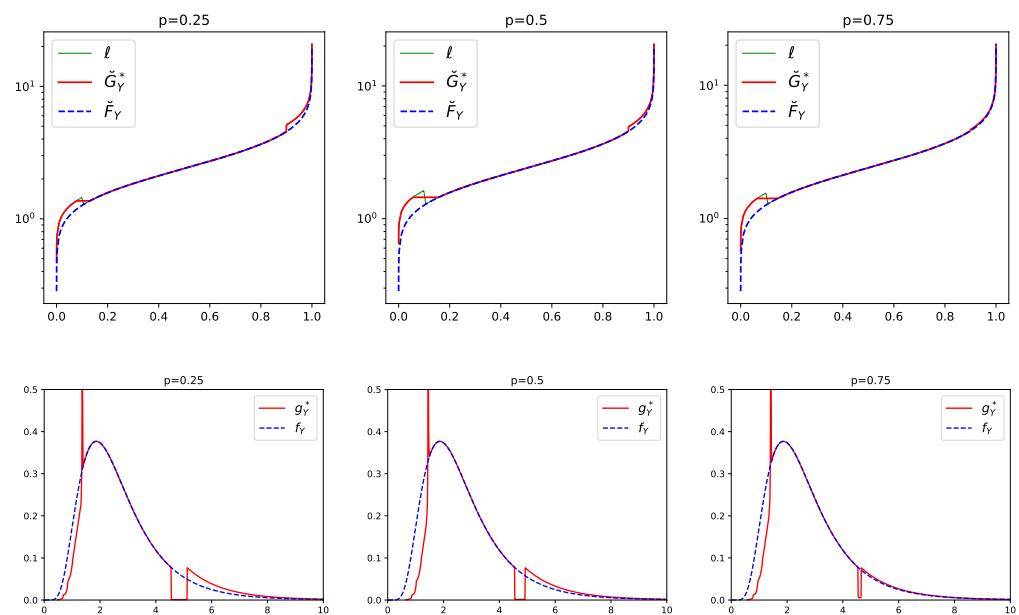
**Example 1** ( $\alpha$ - $\beta$  Risk Measure). The  $\alpha$ - $\beta$  risk measure,  $0 < \beta \leq \alpha < 1$ , is defined by

$$\gamma(u) = \frac{1}{\eta} \left( p \mathbb{1}_{\{u < \beta\}} + (1 - p) \mathbb{1}_{\{u \geq \alpha\}} \right),$$

where  $p \in [0, 1]$  and  $\eta = p\beta + (1 - p)(1 - \alpha)$  is the normalising constant. This parametric family contains several notable risk measures as special cases: for  $p = 0$  we obtain  $ES_\alpha$ , and for  $p = 1$  the conditional lower tail expectation (LTE) at level  $\beta$ .

Moreover, if  $p < \frac{1}{2}$  ( $p > \frac{1}{2}$ ) the  $\alpha$ - $\beta$  risk measure emphasises losses (gains) relative to gains (losses). For  $\alpha = \beta$  and  $p < \frac{1}{2}$ , the risk measure is equivalent to  $\kappa(ES_\alpha[Y] - \lambda \mathbb{E}[Y])$ , where  $\kappa = \frac{(1-2p)(1-\alpha)}{\eta}$  and  $\lambda = \frac{p}{\kappa\eta}$ .

Figure 1 displays the baseline  $\check{F}_Y$  and the stressed  $\check{G}_Y^*$  quantile functions of a random variable  $Y$  under a 10% increase on the  $\alpha$ - $\beta$  risk measure with  $\beta = 0.1$ ,  $\alpha = 0.9$ , and various  $p \in \{0.25, 0.5, 0.75\}$ . The baseline distribution is chosen to be  $F_Y$  is Lognormal( $\mu, \sigma^2$ ) with parameters  $\mu = \frac{7}{8}$  and  $\sigma = 0.5$ . We observe in Figure 1 that the stressed quantile functions  $\check{G}_Y^*$  have, in all three plots, a flat part which straddles  $\beta = 0.1$  and a jump at  $\alpha = 0.9$ . The length of the flat part is increasing with increasing  $p$  while the size of the jump is decreasing with increasing  $p$ . This can also be seen in the stressed densities  $g_Y^*$  which have, for all values of  $p$ , a much heavier right albeit a much lighter left tail than the density of the baseline model. Thus, under this stress, both tails of the baseline distribution are altered.



**Figure 1.** Top panels: Baseline quantile function  $\check{F}_Y$  (blue dashed) compared to the stressed quantile function  $\check{G}_Y^*$  (red solid) for a 10% increase on the  $\alpha$ - $\beta$  risk measure with  $\beta = 0.1$ ,  $\alpha = 0.9$ , and various values of  $p$ . The green line  $l(\cdot)$  is the function, whose isotonic projection equals  $\check{G}_Y^*(\cdot)$ . Bottom panels: corresponding baseline  $f_Y$  and stressed  $g_Y^*$  densities.

### 3.2. Integral Constraints

The next results are generalisations of stresses on distortion risk measures to integral constraints, and include as a special case a stress jointly on the mean, the variance, and distortion risk measures.

**Theorem 2 (Integral).** Let  $h_k, \tilde{h}_l: [0, 1] \rightarrow [0, \infty)$  be square-integrable functions and assume there exists a distribution function  $\tilde{G} \in \mathcal{M}$  satisfying  $\int_0^1 h_k(u)\tilde{G}(u) du \leq c_k$  and  $\int_0^1 \tilde{h}_l(u)(\tilde{G}(u))^2 du \leq \tilde{c}_l$  for all  $k = 1, \dots, d$ , and  $l = 1, \dots, \tilde{d}$ . Then the optimisation problem

$$\arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to} \quad \int_0^1 h_k(u)\tilde{G}(u) du \leq c_k, \quad k = 1, \dots, d,$$

$$\int_0^1 \tilde{h}_l(u)(\tilde{G}(u))^2 du \leq \tilde{c}_l, \quad l = 1, \dots, \tilde{d},$$

has a unique solution given by

$$\check{G}^*(u) = \left( \frac{1}{\tilde{\Lambda}(u)} \left( \check{F}(u) + \sum_{k=1}^d \lambda_k h_k(u) \right) \right)^{\uparrow_{\tilde{\Lambda}}},$$

where  $\tilde{\Lambda}(u) = 1 + \sum_{k=1}^{\tilde{d}} \tilde{\lambda}_k \tilde{h}_k(u)$  and the Lagrange multipliers  $\lambda_1, \dots, \lambda_d$  and  $\tilde{\lambda}_1, \dots, \tilde{\lambda}_d$  are non-negative and such that the constraints are fulfilled.

A combination of the above theorems provides stresses jointly on the mean, the variance, and on multiple distortion risk measures.

**Proposition 2 (Mean, Variance, and Risk Measures).** Let  $m' \in \mathbb{R}, \sigma' > 0, r_k \in \mathbb{R}$ , and distortion risk measures  $\rho_{\gamma_k}, k = 1, \dots, d$ . Assume there exists a distribution function  $\tilde{G} \in \mathcal{M}$  with mean  $m'$ , standard deviation  $\sigma'$ , and which satisfies  $\rho_{\gamma_k}(\tilde{G}) = r_k$ , for all  $k = 1, \dots, d$ . Then the optimisation problem

$$\arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to} \quad \int x dG(x) = m',$$

$$\int (x - m')^2 dG(x) = (\sigma')^2 \quad \text{and}$$

$$\rho_{\gamma_k}(G) = r_k, \quad k = 1, \dots, d,$$

has a unique solution given by

$$\check{G}^*(u) = \left( \frac{1}{1 + \lambda_2} \left( \check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right) \right)^{\uparrow},$$

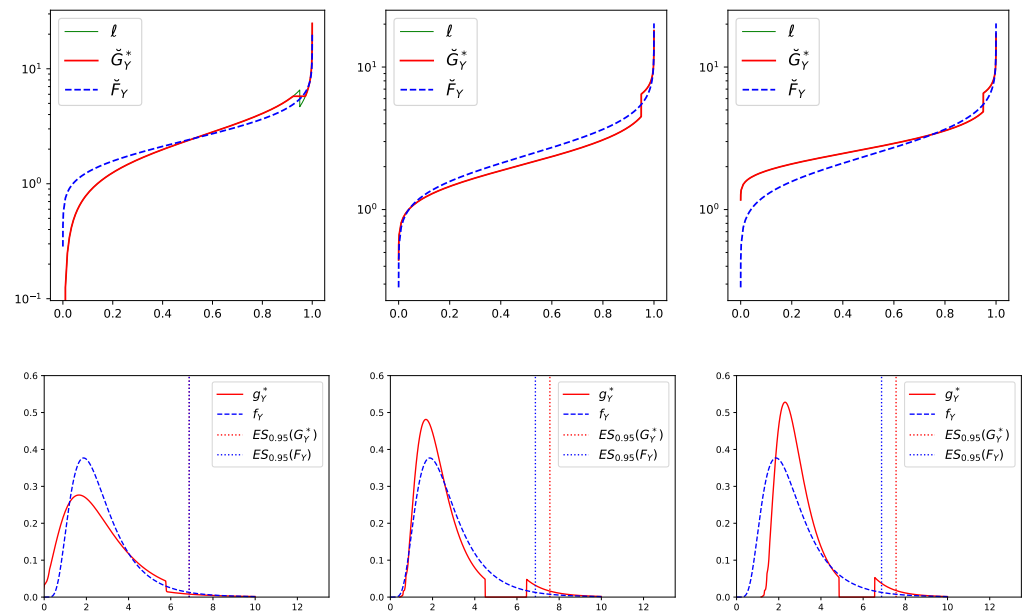
and the Lagrange multipliers  $\lambda_1, \dots, \lambda_{d+2}$  with  $\lambda_2 \neq -1$  are such that the constraints are fulfilled.

**Example 2 (Mean, Variance, and ES).** Here, we illustrate Proposition 2 with the ES risk measure and three different stresses. The top panels of Figure 2 display the baseline quantile function  $\check{F}_Y$  and the stressed quantile function  $\check{G}_Y^*$  of  $Y$ , where the baseline distribution  $F_Y$  of  $Y$  is again Lognormal( $\mu, \sigma^2$ ) with parameters  $\mu = \frac{7}{8}$  and  $\sigma = 0.5$ . The bottom panels display the corresponding baseline and stressed densities. The left panels correspond to a stress, where, under the stressed model, the  $ES_{0.95}$  and the mean are kept fixed at their corresponding values under the baseline model, while the standard deviation is increased by 20%. We observe, both in the quantile and density plot, that the stressed distribution is more spread out indicating a larger variance. Furthermore, at  $y \approx 5.77$  the stressed density  $g_Y^*(y)$  drops to ensure that  $ES_{0.95}(G_Y^*) = ES_{0.95}(F_Y)$ . This drop is due to the fact that a stress composed of a 20% increase in the standard deviation while fixing the mean (i.e., without a constraint on ES) results in an ES that is larger compared to the baseline's. Indeed, under this alternative stress (without a constraint on ES) we obtain that  $ES_{0.95}(G_Y^*) \approx 7.70$  compared to  $ES_{0.95}(F_Y) \approx 6.87$ .

The middle panels correspond to a 10% increase in  $ES_{0.95}$  and a 10% decrease in the mean, while keeping the standard deviation fixed at its value under the baseline model. The density plot clearly indicates a general shift of the stressed density to the left, stemming from the decrease in the



mean, and a single trough which is induced by the increase in ES. The right panels correspond to a 10% increase in  $ES_{0.95}$ , a 10% increase in the mean, and a 10% decrease in the standard deviation. The stressed density still has the trough from the increase in ES; however, the density is less spread out (reduction in the standard deviation) and generally shifted to the right (increase in the mean).



**Figure 2.** **Top:** Baseline quantile function  $\check{F}_Y$  compared to the stressed quantile function  $\check{G}_Y^*$ . **Bottom:** corresponding baseline  $f_Y$  and stressed  $g_Y^*$  densities. **Left:**  $ES_{0.95}$  and the mean being fixed and a 20% increase in the standard deviation. **Middle:** 10% increase in  $ES_{0.95}$ , 10% decrease in the mean, and fixed standard deviation. **Right:** 10% increase in  $ES_{0.95}$ , 10% increase in the mean, and 10% decrease in standard deviation. Note that in the middle and right panel the green lines are equal to the red lines.

### 3.3. Value-at-Risk Constraints

In this section we study stresses on the risk measure Value-at-Risk (VaR). The VaR at level  $\alpha \in (0, 1)$  of a distribution function  $G \in \mathcal{M}$  is defined as its left-continuous quantile function evaluated at  $\alpha$ , that is

$$VaR_\alpha(G) = \check{G}(\alpha).$$

We further define the right-continuous  $VaR^+$ , that is the right-continuous quantile function of  $G \in \mathcal{M}$  evaluated at  $\alpha$ , by

$$VaR_\alpha^+(G) = \check{G}^+(\alpha) = \inf \{y \in \mathbb{R} \mid F(y) > \alpha\}.$$

**Theorem 3 (VaR).** Let  $q \in \mathbb{R}$  and consider the optimisation problem

$$\begin{aligned} \arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to} \quad & (a) \quad VaR_\alpha(G) = q \quad \text{or} \\ & (b) \quad VaR_\alpha^+(G) = q, \end{aligned}$$

and define  $\alpha_F$  such that  $VaR_{\alpha_F}(F) = q$ . Then, the following holds

(i) under constraint (a), if  $q \leq VaR_\alpha(F)$ , then the unique solution is given by

$$\check{G}^*(u) = \check{F}(u) + (q - \check{F}(u)) \mathbb{1}_{\{u \in (\alpha_F, \alpha]\}};$$

if  $q > VaR_\alpha(F)$ , then there does not exist a solution.

(ii) under constraint (b), if  $q \geq \text{VaR}_\alpha^+(F)$ , then the unique solution is given by

$$\check{G}^*(u) = \check{F}(u) + (q - \check{F}(u))\mathbb{1}_{\{u \in (\alpha, \alpha_F]\}};$$

if  $q < \text{VaR}_\alpha^+(F)$ , then there does not exist a solution.

The above theorem states that if the optimal quantile function exists it is either the baseline quantile function  $\check{F}$  or constant equal to  $q$ . Moreover, the stressed quantile function (if it exists) jumps at  $\alpha$  which implies that the existence of a solution hinges on the careful choice of the stress. For a stress on VaR (constraint (a)) for example, a solution exists if and only if the constraint satisfies  $q \leq \text{VaR}_\alpha(F)$ ; a decrease in the  $\text{VaR}_\alpha$  from the baseline to the stressed model. The reason for the non-existence of a solution when stressing VaR upwards is that the unique increasing function that minimises the Wasserstein distance and satisfies the constraint is not left-continuous and thus not a quantile function.

Alternatively to stressing VaR or  $\text{VaR}^+$ , and in particularly in the case when a desired stressed solution does not exist, one may stress instead the distortion risk measure Range-Value-at-Risk (RVaR) Cont et al. (2010). The RVaR at levels  $0 \leq \alpha < \beta \leq 1$  is defined by

$$\text{RVaR}_{\alpha, \beta}(G) = \frac{1}{\beta - \alpha} \int_\alpha^\beta \check{G}(u) du, \quad \text{for } G \in \mathcal{M},$$

and belongs to the class of distortion risk measures. The RVaR attains as limiting cases the VaR and  $\text{VaR}^+$ . Indeed, for any  $G \in \mathcal{M}$  it holds

$$\text{VaR}_\alpha(G) = \lim_{\alpha' \nearrow \alpha} \text{RVaR}_{\alpha', \alpha}(G) \quad \text{and} \quad \text{VaR}_\alpha^+(G) = \lim_{\beta \searrow \alpha} \text{RVaR}_{\alpha, \beta}(G).$$

The solution to stressing RVaR is provided in Theorem 1.

### 3.4. Expected Utility Constraint

This section considers the change from the baseline to the stressed distribution under an increase of an expected utility constraint. In the context of utility maximisation, the next theorem provides a way to construct stressed models with a larger utility compared to the baseline.

**Theorem 4** (Expected Utility and Risk Measures). *Let  $u: \mathbb{R} \rightarrow \mathbb{R}$  be a differentiable concave utility function,  $r_k \in \mathbb{R}$ , and  $\rho_{\gamma_k}$  be distortion risk measures, for  $k = 1, \dots, d$ . Assume there exists a distribution function  $\check{G}$  satisfying  $\int_{\mathbb{R}} u(x) d\check{G}(x) \geq c$  and  $\rho_{\gamma_k}(\check{G}) = r_k$  for all  $k = 1, \dots, d$ . Then the optimisation problem*

$$\arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to} \quad \int_{\mathbb{R}} u(x) dG(x) \geq c \quad \& \quad \rho_{\gamma_k}(G) = r_k, \quad k = 1, \dots, d$$

has a unique solution given by

$$\check{G}^*(u) = \check{v}_{\lambda_1} \left( \left( \check{F}(u) + \sum_{k=1}^d \lambda_{k+1} \gamma_k(v) \right)^\uparrow \right), \tag{5}$$

where  $\check{v}_{\lambda_1}$  is the left-inverse of  $v_{\lambda_1}(x) = x - \lambda_1 u'(x)$ , and  $\lambda_1 \geq 0, (\lambda_2, \dots, \lambda_{d+1}) \in \mathbb{R}^d$  are such that the constraints are fulfilled.

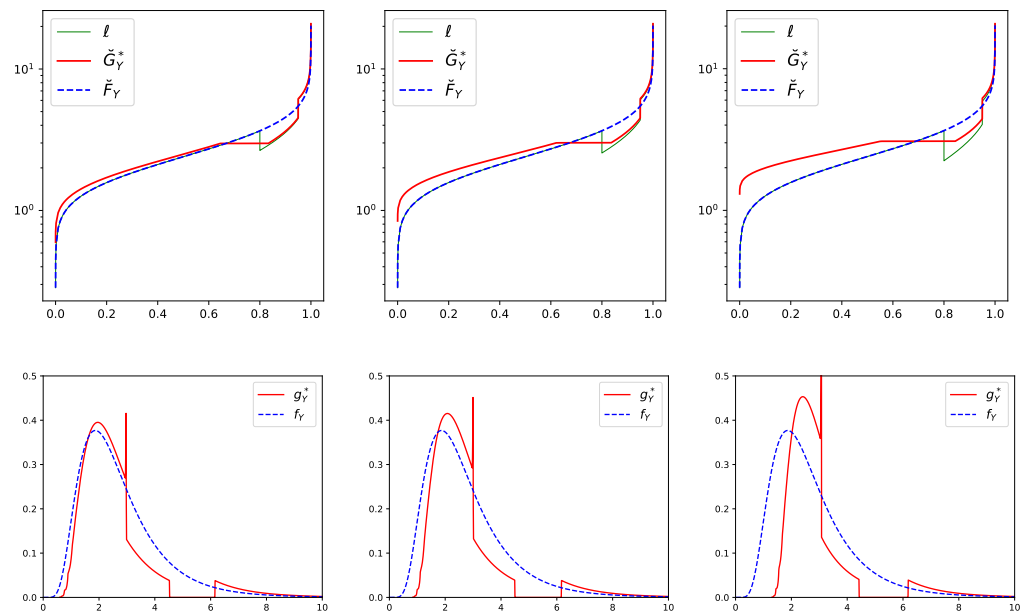
The utility function in Theorem 4 need not be monotone, indeed the theorem applies to any differentiable concave function, without the need of an utility interpretation. Moreover, Theorem 4 also applies to differentiable convex (disutility) functions  $\tilde{u}$  and constraint  $\int_{\mathbb{R}} \tilde{u}(x) dG(x) \leq c$ ; a situation of interest in insurance premium calculations. In this case, the solution is given by (5) with  $u(x) = -\tilde{u}(x)$ .

**Example 3 (HARA Utility & ES).** The Hyperbolic absolute risk aversion (HARA) utility function is defined by

$$u(x) = \frac{1 - \eta}{\eta} \left( \frac{ax}{1 - \eta} + b \right)^\eta,$$

with parameters  $a > 0$ ,  $\frac{ax}{1 - \eta} + b > 0$ , and where  $\eta \leq 1$  guarantees concavity.

We again choose the baseline distribution  $F_Y$  of  $Y$  to be Lognormal( $\mu, \sigma^2$ ) with  $\mu = \frac{7}{8}$  and  $\sigma = 0.5$  and consider utility parameters  $a = 1$ ,  $b = 5$ , and  $\eta = 0.5$ . Figure 3 displays the baseline and the stressed quantile functions  $\check{F}_Y$  and  $\check{G}_Y^*$ , respectively, for a combined stress on the HARA utility and on ES at levels 0.8 and 0.95. Specifically, for all three stresses we decreasing  $ES_{0.8}$  by 10% and increasing  $ES_{0.95}$  by 10% compared to their values under the baseline model. Moreover, the HARA utility is increased by 0%, 1%, and 3%, respectively, corresponding to the panels from the left to the right. The flat part in the stressed quantile function  $\check{G}^*(u)$  around  $u = 0.8$ , visible in all top panels of Figure 3, is induced by the decrease in  $ES_{0.8}$  while the jump at  $u = 0.95$  is due to the increase in  $ES_{0.95}$ . From the left to the right panel in Figure 3, we observe that the larger the stress on the HARA utility, the more the stressed quantile function shifts away from the baseline quantile function  $\check{F}_Y$ .



**Figure 3.** Top panels: Baseline quantile function  $\check{F}_Y$  compared to the stressed quantile function  $\check{G}_Y^*$ , for a 10% decrease in  $ES_{0.8}$ , and a 10% increase in  $ES_{0.95}$ , and, from left to right, a 0%, 1%, and 3% increase in the HARA utility, respectively. The function  $\ell(\cdot)$  (solid green) is the function whose isotonic projection equals  $\check{G}^*(\cdot)$ . Bottom panels: Corresponding baseline  $f_Y$  and stressed  $g_Y^*$  densities.

### 3.5. Smoothing of the Stressed Distribution

We observe that the stressed quantile functions derived in Section 3 generally contain jumps and/or flat parts even if the baseline distribution is absolutely continuous. In situation where this is not desirable, one may consider a smoothed version of the stressed distributions. For this, we recall that the isotonic regression, the discrete counterpart of the weighted isotonic projection, of a function  $\ell$  evaluated at  $u_1, \dots, u_n$  with positive weights  $w_1, \dots, w_n$ , is the solution to

$$\min_{u_1, \dots, u_n} \sum_{i=1}^n (u_i - \ell(u_i))^2 w_i, \quad \text{subject to } u_i \leq u_j, \quad i \leq j. \tag{6}$$

There are numerous efficient algorithms that solve (6) most notably the pool-adjacent-violators (PAV) algorithm Barlow et al. (1972). It is well-known that the solution to the

isotonic regression contains flat parts and jumps. A smoothed isotonic regression algorithm, termed smooth pool-adjacent-violators (SPAV) algorithm, using an  $\mathbb{L}^2$  regularisation was recently proposed by Sysoev and Burdakov (2019). Specifically, they consider

$$\min_{u_1, \dots, u_n} \sum_{i=1}^n (u_i - \ell(u_i))^2 w_i + \sum_{i=1}^n \zeta_i (u_{i+1} - u_i)^2, \quad \text{subject to } u_i \leq u_j, \quad i \leq j,$$

where  $\zeta_i \geq 0, i = 0, \dots, n - 1$ , are prespecified smoothing parameters. Using a probabilistic reasoning, Sysoev and Burdakov (2019) argue that  $\zeta_i$  may be chosen proportional to a (e.g., quadratic) kernel evaluated at  $u_i$  and  $u_{i+1}$ , that is

$$\zeta_i = \zeta K(u_i, u_{i+1}) \quad \text{with} \quad K(u_i, u_{i+1}) = \frac{1}{|u_i - u_{i+1}|^2} \quad \text{and} \quad \zeta \geq 0.$$

The choice of smoothing parameter  $\zeta = 0$  correspond to the original isotonic regression larger values of  $\zeta$  correspond to a greater degree of smoothness of the solution.  $\zeta$  can either be prespecified or estimated using cross-validation, see e.g., Sysoev and Burdakov (2019).

To guarantee that the smoothed quantile function still fulfils the constraint, one may replace in every step of the optimisation for finding the Lagrange parameter the PAV with the SPAV algorithm. Thus, the Lagrange parameter are indeed found such that the constraints are fulfilled.

**Remark 1.** *There are numerous works proposing smooth versions of isotonic regressions. Approaches include kernel smoothers, e.g., Hall and Huang (2001), and spline techniques, e.g., Meyer (2008). These algorithms, however, are computationally heavy in that their computational cost is  $O(n^2)$ , where  $n$  is the number of data points. Furthermore, these algorithm require a careful choice of the kernel or the spline basis which is in contrast to the SPAV. We refer the reader to Sysoev and Burdakov (2019) for a detailed discussion and references to smooth isotonic regression algorithms.*

#### 4. Analysing the Stressed Model

Recall that a modeller is equipped with a baseline model, the 3-tuple  $(\mathbf{X}, g, \mathbb{P})$ , consisting of a set of input factors  $\mathbf{X} = (X_1, \dots, X_n)$ , a univariate output random variable of interest,  $Y = g(\mathbf{X})$ , and a probability measure  $\mathbb{P}$ . For a stress on the output’s baseline distribution  $F_Y$ , we derived in Section 3 the corresponding unique stressed distribution function, denoted here by  $G_Y^*$ . Thus, to fully specify the stressed model we next define a stressed probability measure  $\mathbb{Q}^*$  that is induced by  $G_Y^*$ .

##### 4.1. The Stressed Probability Measures

A stressed distribution  $G_Y^*$  induces a canonical change of measure that allows the modeller to understand how the baseline model including the distributions of the inputs changes under the stress. The Radon–Nikodym (RN) derivative of the baseline to the stressed model is

$$\frac{d\mathbb{Q}^*}{d\mathbb{P}} = \frac{g_Y^*(Y)}{f_Y(Y)},$$

where  $f_Y$  and  $g_Y^*$  denote the densities of the baseline and stressed output distribution, respectively. The RN derivative is well-defined since  $f_Y(Y) > 0, \mathbb{P}$ -a.s. The distribution functions of input factors under the stressed model – the stressed distributions – are then given, e.g., for input  $X_i, i \in \{1, \dots, n\}$ , by

$$\mathbb{Q}^*(X_i \leq x_i) = \mathbb{E} \left[ \mathbb{1}_{\{X_i \leq x_i\}} \frac{d\mathbb{Q}^*}{d\mathbb{P}} \right] = \mathbb{E} \left[ \mathbb{1}_{\{X_i \leq x_i\}} \frac{g_Y^*(Y)}{f_Y(Y)} \right], \quad x_i \in \mathbb{R},$$

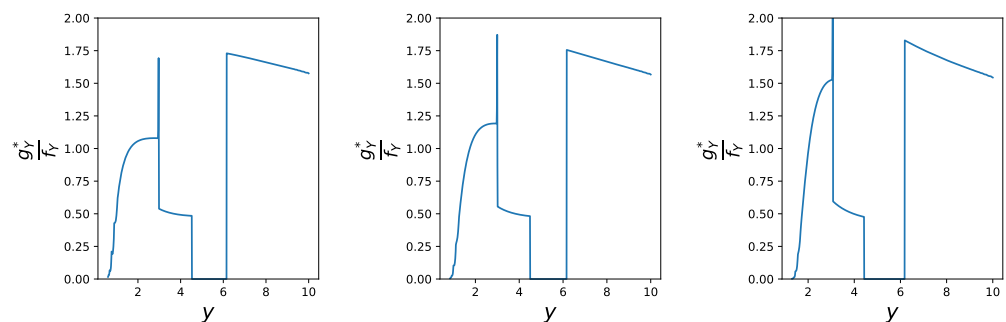
and for multivariate inputs  $\mathbf{X}$  by

$$\mathbb{Q}^*(\mathbf{X} \leq \mathbf{x}) = \mathbb{E} \left[ \mathbb{1}_{\{\mathbf{X} \leq \mathbf{x}\}} \frac{g_Y^*(Y)}{f_Y(Y)} \right], \quad \mathbf{x} \in \mathbb{R}^n,$$

where  $\mathbb{E}[\cdot]$  denotes the expectation under  $\mathbb{P}$ . Note that under the stressed probability measure  $\mathbb{Q}^*$ , the input factors' marginal and joint distributions may be altered.

**Example 4 (HARA Utility & ES continued).** We continue Example 3 and illustrate the RN-densities  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$  for the following three stresses (from the left to the right panel): a 10% decrease in  $ES_{0.8}$  and a 10% increasing  $ES_{0.95}$  for all three stresses, and a 0%, 1%, and 3% increase in the HARA utility, respectively.

We observe in Figure 4, that for all three stresses large realisations of  $Y$  obtain a larger weight under the stressed probability measures  $\mathbb{Q}^*$  compared to the baseline probability  $\mathbb{P}$ . Indeed, for all three stresses it holds that  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}(\omega) > 1$  whenever  $Y(\omega) > 6$  and  $\omega \in \Omega$ . This is in contrast to small realisations of  $Y$  which obtain a weight smaller than 1. The impact of the different levels of stresses of the HARA utility (0%, 1%, and 3%, from the left to the right panel) can be observed in the left tail of  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ ; a larger stress on the utility induces larger weights. The length of the trough of  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$  is increasing from the left panel (approx. (4.53, 6.15)) to the right panel (approx. (4.43, 6.18)), and corresponds in all cases to the constant part in  $G_Y^*$  (see Figure 3, top panels) which is induced by the decrease in  $ES_{0.8}$  under the stressed model.



**Figure 4.** RN-densities for following stresses: a 10% decrease in both  $ES_{0.8}$  and  $ES_{0.95}$ , and an increase in the HARA utility. The change in HARA utility is 0%, 1%, and 3%, respectively, from left to right.

#### 4.2. Reverse Sensitivity Measures

Comparison of the baseline and a stressed model can be conducted via different approaches depending on the modeller's interest. While probabilistic sensitivity measures underlie the assumption of a fixed probability measure and quantify the divergence between the conditional (on a model input) and the unconditional output density Saltelli et al. (2008), the proposed framework compares a baseline and a stressed model, i.e., distributions under different probability measures. Therefore, to quantify the distributional change in input factor  $X_i$  from the baseline  $\mathbb{P}$  to the stressed  $\mathbb{Q}^*$  probability, a sensitivity measure introduced by Pesenti et al. (2019) may be suitable which quantifies the variability of an input factor's distribution from the baseline to the stressed model. A generalisation of the reverse sensitivity measure is stated here.

**Definition 4** (Marginal Reverse Sensitivity Measure Pesenti et al. (2019)). For a function  $s: \mathbb{R} \rightarrow \mathbb{R}$ , the reverse sensitivity measure to input  $X_i$  with respect to a stressed probability measure  $\mathbb{Q}^*$  is defined by

$$S_i^{\mathbb{Q}^*} = \begin{cases} \frac{\mathbb{E}^{\mathbb{Q}^*}[s(X_i)] - \mathbb{E}[s(X_i)]}{\max_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X_i)] - \mathbb{E}[s(X_i)]} & \mathbb{E}^{\mathbb{Q}^*}[s(X_i)] \geq \mathbb{E}[s(X_i)], \\ -\frac{\mathbb{E}^{\mathbb{Q}^*}[s(X_i)] - \mathbb{E}[s(X_i)]}{\min_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X_i)] - \mathbb{E}[s(X_i)]} & \text{otherwise,} \end{cases}$$

where  $\mathcal{Q} = \{\mathbb{Q} \mid \mathbb{Q} \text{ probability measure with } \frac{d\mathbb{Q}}{d\mathbb{P}} \stackrel{\mathbb{P}}{=} \frac{d\mathbb{Q}^*}{d\mathbb{P}}\}$  is the set of all probability measures whose RN-derivative have the same distribution as  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$  under  $\mathbb{P}$ . We adopted the convention that  $\pm \frac{\infty}{\infty} = \pm 1$  and  $\frac{0}{0} = 0$ .

The sensitivity measure is called “reverse”, as the stress is applied to the output random variable  $Y$  and the sensitivity monitors the change in input  $X_i$ . The definition of 4 applies, however, also to stresses on input factors, in which case the RN-density  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$  is a function of the stressed input factor and we refer to Pesenti et al. (2019) for a discussion. Note, that the reverse sensitivity measure can be viewed as a normalised covariance measure between the input  $s(X_i)$  and the Radon Nikodym derivative  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ .

The next proposition provides a collection of properties that the reverse sensitivity measure possesses, we also refer to Pesenti et al. (2019) for a detailed discussion of these properties. For this, we first recall the definition of comonotonic and counter-monotonic random variables.

**Definition 5.** Two random variables  $Y_1$  and  $Y_2$  are comonotonic under  $\mathbb{P}$ , if and only if, there exists a random variable  $W$  and non-decreasing functions  $h_1, h_2: \mathbb{R} \rightarrow \mathbb{R}$ , such that the following equalities hold in distribution under  $\mathbb{P}$

$$Y_1 = h_1(W) \quad \text{and} \quad Y_2 = h_2(W). \tag{7}$$

The random variables  $Y_1$  and  $Y_2$  are counter-monotonic under  $\mathbb{P}$ , if and only if, (7) holds with one of the functions  $h_1(\cdot), h_2(\cdot)$  being non-increasing, and the other non-decreasing.

If two random variables are (counter) comonotonic under one probability measure, then they are also (counter) comonotonic under any other absolutely continuous probability measure, see, e.g., Proposition 2.1 of Cuestaalbertos et al. (1993). Thus, we omit the specification of the probability measure when discussing counter- and comonotonicity.

**Proposition 3** (Properties of Reverse Sensitivity Measure). The reverse sensitivity measure possesses the following properties:

- (i)  $S_i^{\mathbb{Q}^*} \in [-1, 1]$ ;
- (ii)  $S_i^{\mathbb{Q}^*} = 0$  if  $(s(X_i), \frac{d\mathbb{Q}^*}{d\mathbb{P}})$  are independent under  $\mathbb{P}$ ;
- (iii)  $S_i^{\mathbb{Q}^*} = 1$  if and only if  $(s(X_i), \frac{d\mathbb{Q}^*}{d\mathbb{P}})$  are comonotonic;
- (iv)  $S_i^{\mathbb{Q}^*} = -1$  if and only if  $(s(X_i), \frac{d\mathbb{Q}^*}{d\mathbb{P}})$  are counter-comonotonic.

The function  $s(\cdot)$  provides the flexibility to create sensitivity measures that quantify changes in moments, e.g., via  $s(x) = x^k, k \in \mathbb{N}$ , or in the tail of distributions, e.g., via  $s(x) = \mathbb{1}_{\{x > \text{VaR}_\alpha(X_i)\}}$ , for  $\alpha \in (0, 1)$ .

Next, we generalise Definition 4 to a sensitivity measure that accounts for multiple input factors. While  $S_i^{\mathbb{Q}^*}$  measures the change of the distribution of  $X_i$  from the base-line to the stressed model the sensitivity  $S_{i,j}^{\mathbb{Q}^*}$  introduced below, quantifies how the joint distribution of  $(X_i, X_j)$  changes when moving from  $\mathbb{P}$  to  $\mathbb{Q}^*$ .

**Definition 6** (Bivariate Reverse Sensitivity Measure). For a function  $s: \mathbb{R}^2 \rightarrow \mathbb{R}$ , the reverse sensitivity measure to inputs  $(X_i, X_j)$  with respect to a stressed probability measure  $\mathbb{Q}^*$  is defined by

$$\mathcal{S}_{i,j}^{\mathbb{Q}^*} = \begin{cases} \frac{\mathbb{E}^{\mathbb{Q}^*}[s(X_i, X_j)] - \mathbb{E}[s(X_i, X_j)]}{\max_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X_i, X_j)] - \mathbb{E}[s(X_i, X_j)]} & \mathbb{E}^{\mathbb{Q}^*}[s(X_i, X_j)] \geq \mathbb{E}[s(X_i, X_j)], \\ -\frac{\mathbb{E}^{\mathbb{Q}^*}[s(X_i, X_j)] - \mathbb{E}[s(X_i, X_j)]}{\min_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X_i, X_j)] - \mathbb{E}[s(X_i, X_j)]} & \text{otherwise,} \end{cases}$$

where  $\mathcal{Q}$  is given in Definition 4.

The bivariate sensitivity measure satisfies all the properties in Proposition 3 when  $s(X_i)$  is replaced by  $s(X_i, X_j)$ . The bivariate sensitivity  $\mathcal{S}_{i,j}^{\mathbb{Q}^*}$  can also be generalised to  $k$  input factors by choosing a function  $s: \mathbb{R}^k \rightarrow \mathbb{R}$ .

**Remark 2.** Probabilistic sensitivity measures are typically used for importance measurement and take values in  $[0, 1]$ ; with 1 being the most important input factor and 0 being (desirably) independent from the output Borgonovo et al. (2021). This is in contrast to our framework where  $\mathcal{S}_i^{\mathbb{Q}^*}$  lives in  $[-1, 1]$  and e.g., a negative dependence, such as negative quadrant dependence between  $s(X_i)$  and  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$  implies that  $\mathcal{S}_i^{\mathbb{Q}^*} < 0$ , see Pesenti et al. (2019) [Proposition 4.3]. Thus, the proposed sensitivity measure is different in that it allows for negative sensitivities where the sign of  $\mathcal{S}_i^{\mathbb{Q}^*}$  indicates the direction of the distributional change.

### 5. Application to a Spatial Model

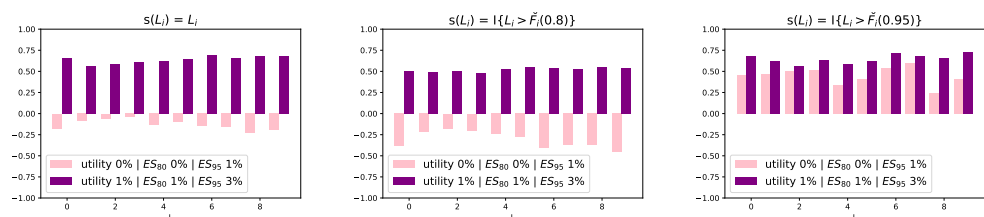
We consider a spatial model for modelling insurance portfolio losses where each individual loss occurs at different locations and the dependence between individual losses is a function of the distance between the locations of the losses. Mathematically, denote the locations of the insurance losses by  $z_1, \dots, z_{10}$ , where  $z_m = (z_m^1, z_m^2)$  are the coordinates of location  $z_m$ ,  $m = 1, \dots, 10$ . The insurance loss at location  $m$ , denoted by  $L_m$ , follows a  $\text{Gamma}(5, \frac{0.2}{m})$  distribution with location parameter 25. Thus, the minimum loss at each location is 25 and locations with larger mean also exhibit larger standard deviations. The losses  $L_1, \dots, L_m$  have, conditionally on  $\Theta = \theta$ , a Gaussian copula with correlation matrix given by  $\rho_{i,j} = \text{Cor}(L_i, L_j) = e^{-\theta \|z_i - z_j\|}$ , where  $\|\cdot\|$  denotes the Euclidean distance. Thus, the further apart the locations  $z_i$  and  $z_j$  are the smaller the correlation between  $L_i$  and  $L_j$ . The parameter  $\Theta$  takes values  $(0, 0.4, 5)$  with probabilities  $(0.05, 0.6, 0.35)$  that represent different regimes. Indeed,  $\Theta = 0$  corresponds to a correlation of 1 between all losses, independently of their location. Larger values of  $\Theta$  correspond to smaller albeit still positive correlation. Thus, regime with  $\Theta = 0$  can be viewed as, e.g., circumstances suitable for natural disasters. We further define the total loss of the insurance company by  $Y = \sum_{m=1}^{10} L_m$ .

We perform two different stresses on the total loss  $Y$  detailed in Table 1. Specifically, we consider as a first stress a 0% change in HARA utility, a 0% change in  $\text{ES}_{0.8}(Y)$ , and a 1% increase in  $\text{ES}_{0.95}(Y)$  from the baseline to the stressed model. The second stress is composed of a 1% increase in HARA utility, a 1% increase in  $\text{ES}_{0.8}(Y)$ , and a 3% increase in  $\text{ES}_{0.95}(Y)$  compared to the baseline model. As the second stress increases all three metrics it may be viewed as a more severe distortion of the baseline model.

**Table 1.** Summary of the stresses applied to the portfolio loss  $Y$  represented in relative increases of the stressed model from the baseline model.

	HARA Utility	ES <sub>0.8</sub> ( $Y$ )	ES <sub>0.95</sub> ( $Y$ )
Stress 1: $Q_1^*$	0%	0%	1%
Stress 2: $Q_2^*$	1%	1%	3%

Next, we calculate reverse sensitivity measures for the losses  $L_1, \dots, L_{10}$  for both stresses  $Q_1^*$  and  $Q_2^*$ . Figure 5 displays the reverse sensitivity measures for functions  $s(x) = x$ ,  $s(x) = \mathbb{1}_{\{x > \check{F}_i(0.8)\}}$ , and  $s(x) = \mathbb{1}_{\{x > \check{F}_i(0.95)\}}$ , from the left to the right panel, and where  $\check{F}_i$ , denotes the  $\mathbb{P}$ -quantile function of  $L_i$ ,  $i = 1, \dots, 10$ .



**Figure 5.** Reverse sensitivity measures with  $s(x) = x$ ,  $s(x) = \mathbb{1}_{\{x > \check{F}_i(0.8)\}}$ , and  $s(x) = \mathbb{1}_{\{x > \check{F}_i(0.95)\}}$  (left to right), for two different stresses on the output  $Y$ . First stress (salmon) is keeping the HARA utility and  $ES(Y)_{0.8}$  fixed and increasing the  $ES(Y)_{0.85}$  by 1%. Second stress (violet) is an increase of 1% in HARA utility, 1%  $ES(Y)_{0.8}$ , and 3% in  $ES(Y)_{0.85}$ .

We observe that for stress 2, the reverse sensitivities to all losses  $L_i$  and all choices of function  $s(\cdot)$  are positive. This contrasts the reverse sensitivities for stress 1. Indeed, for stress 1 the reverse sensitivities with both  $s(x) = x$  and  $s(x) = \mathbb{1}_{\{x > \check{F}_i(0.8)\}}$  are negative, with the former values being smaller indicating a smaller change in the distributions of the  $L_i$ 's. By definition of the reverse sensitivity, the left panel corresponds to the (normalised) difference between the expectation under the stressed and baseline model. The middle and right panels correspond to the (normalised) change in the probability of exceeding  $\check{F}_i(0.8)$  and  $\check{F}_i(0.95)$ , respectively. Thus, as seen in the plots, while the expectations and probabilities of exceeding the 80%  $\mathbb{P}$ -quantile are smaller under the stressed model, the probabilities of exceeding the 95%  $\mathbb{P}$ -quantile are increased substantially. The first stress increases the ES at level 0.95 while simultaneously fixes the utility and ES at level 0.8 to its values under the baseline model. This induces under the stressed probability measure a reduction of the mean and of the probability of exceeding the 80%  $\mathbb{P}$ -quantile while the probability of exceeding the 95%  $\mathbb{P}$ -quantile increases. Thus, the reverse sensitivity measures provide a spectrum of measures to analyse the distributional change of the losses  $L_i$  from the baseline to the stressed model.

Next, for a comparison we calculate the delta sensitivity measure of introduced by Borgonovo (2007). For a probability measure  $\mathbb{Q}$  the delta measure of  $L_i$  is defined by

$$\zeta^{\mathbb{Q}}(L_i) = \frac{1}{2} \int \int \left| f_Y^{\mathbb{Q}}(y) - f_{Y|L_i}^{\mathbb{Q}}(y|z) \right| f_i^{\mathbb{Q}}(z) dy dz,$$

where  $f_Y^{\mathbb{Q}}(\cdot)$  and  $f_i^{\mathbb{Q}}(\cdot)$  are the densities of  $Y$  and  $L_i$  under  $\mathbb{Q}$ , respectively, and where  $f_{Y|L_i}^{\mathbb{Q}}(\cdot|\cdot)$  is the conditional density of the total portfolio loss  $Y$  given  $L_i$  under  $\mathbb{Q}$ .

Table 2 reports the delta measures under the baseline model  $\zeta^{\mathbb{P}}$  and the two stresses, i.e.,  $\zeta^{Q_1^*}$  and  $\zeta^{Q_2^*}$ . We observe that the delta measures are similar for all losses  $L_i$  and do not change significantly under the different probability measures. As the delta sensitivity measure quantifies the importance of input factors under a probability measure, having similar values for  $\zeta^{\mathbb{P}}$ ,  $\zeta^{Q_1^*}$ , and  $\zeta^{Q_2^*}$ , means that the importance ranking of the  $L_i$ 's under different stresses does not change. We also report, in the first two columns of Table 2, the

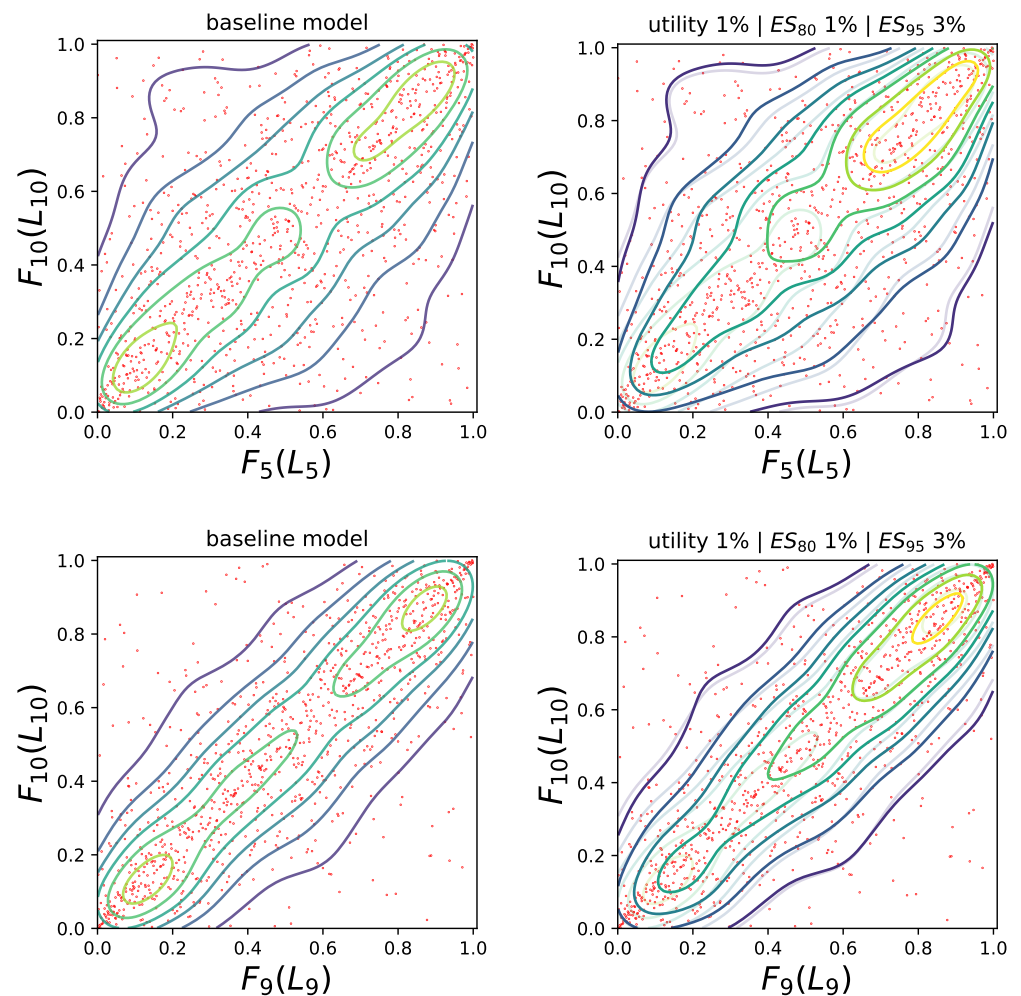


reverse sensitivity measures with  $s(x) = \mathbb{1}_{\{x > \check{F}(0.95)\}}$ . The reverse sensitivity measures provide, in contrast to the delta measure, insight into the change in the distributions of the  $L_i$ 's from  $\mathbb{P}$  and  $\mathbb{Q}^*$ .

**Table 2.** Comparison of different sensitivity measures: First two columns correspond to the reverse sensitivity measures with  $s(x) = \mathbb{1}_{\{x > \check{F}(0.95)\}}$  and stressed models  $\mathbb{Q}_1^*$  and  $\mathbb{Q}_2^*$ , respectively. The last three columns are the delta measure under  $\mathbb{P}$ ,  $\mathbb{Q}_1^*$ , and  $\mathbb{Q}_2^*$ , respectively.

	$\mathcal{S}_i^{\mathbb{Q}_1^*}$	$\mathcal{S}_i^{\mathbb{Q}_2^*}$	$\zeta^{\mathbb{P}}$	$\zeta^{\mathbb{Q}_1^*}$	$\zeta^{\mathbb{Q}_2^*}$
$L_1$	0.45	0.68	0.38	0.38	0.38
$L_2$	0.47	0.62	0.29	0.29	0.29
$L_3$	0.51	0.57	0.30	0.30	0.29
$L_4$	0.52	0.63	0.30	0.30	0.29
$L_5$	0.34	0.58	0.33	0.34	0.33
$L_6$	0.41	0.62	0.34	0.34	0.32
$L_7$	0.54	0.72	0.40	0.40	0.38
$L_8$	0.60	0.69	0.38	0.39	0.39
$L_9$	0.24	0.66	0.40	0.40	0.38
$L_{10}$	0.41	0.73	0.39	0.38	0.37

Alternatively to considering the change in the marginal distributions  $L_i$  from the baseline to the stressed model, we can study the change in the dependence between the losses when moving from the baseline to a stressed model. For this, we consider the bivariate reverse sensitivity measures for the pairs  $(L_5, L_{10})$  and  $(L_9, L_{10})$  for the second stress  $\mathbb{Q}_2^*$ , that is a 1% increase in HARA utility and  $ES_{0.8}$ , and a 3% increase in  $ES_{0.95}$ . Specifically, we look at the function  $s(L_i, L_j) = \mathbb{1}_{\{L_i > \check{F}_i(0.95)\}} \mathbb{1}_{\{L_j > \check{F}_j(0.95)\}}$ , where  $\check{F}_i(\cdot)$  and  $\check{F}_j(\cdot)$  are the  $\mathbb{P}$ -quantile functions of  $L_i$  and  $L_j$  respectively. This bivariate sensitivity measure quantifies the impact a stress has on the probability of joint exceedances with values  $\mathcal{S}_{5,10}^{\mathbb{Q}_2^*} = 0.78$  and  $\mathcal{S}_{9,10}^{\mathbb{Q}_2^*} = 0.81$  indicating that the probabilities of joint exceedances increase more for stress 2. This can also be seen in Figure 6 which shows the bivariate copulae contours of  $(L_5, L_{10})$  (top panels) and  $(L_9, L_{10})$  (bottom panels). The left contour plots correspond to the baseline model  $\mathbb{P}$  whereas the right panels display the contours under the stress model  $\mathbb{Q}_2^*$  (solid lines) together with the baseline contours (reported using partially transparent lines). The red dots are the simulated realisations of the losses  $(L_5, L_{10})$  and  $(L_9, L_{10})$ , respectively (which are the same for the baseline and stressed model). We observe that for both pairs  $(L_5, L_{10})$  and  $(L_9, L_{10})$  the copula under the stressed model admits larger probabilities of joint large events, which is captured by the bivariate reverse sensitivity measure admitting positive values close to 1.



**Figure 6.** Contour plots of the bivariate copulae of  $(L_5, L_{10})$  (**top panels**) and  $(L_9, L_{10})$  (**bottom panels**) under different models. The left contour plots correspond to the baseline model and the right panels to the stress  $\mathbb{Q}_2^*$  (solid lines) with the baseline contours reported using partially transparent lines. Red points are simulated realisations.

## 6. Concluding Remarks

We extend the reverse sensitivity analysis proposed by Pesenti et al. (2019) which proceeds as follows. Equipped with a baseline model which comprises of input and output random variables and a baseline probability measure, one derives a unique stressed model such that the output (or input) under the stressed model satisfies a prespecified stress and is closest to the baseline distribution. While Pesenti et al. (2019) consider the Kullback–Leibler divergence to measure the difference between the baseline and stressed models we utilise Wasserstein distance of order two. Compared to Pesenti et al. (2019) the Wasserstein distance allows for additional and different stresses on the output including the mean and variance, any distortion risk measure including the Value-at-Risk and Expected-Shortfall, and expected utility type constraints, thus making the reverse sensitivity analysis framework suitable for models used in financial and insurance risk management. We further discuss reverse sensitivity measures which quantify the change the inputs’ distribution when moving from the baseline to a stressed model and illustrate our results on a spatial insurance portfolio application. The reverse sensitivity analysis framework (including the results from this work and from Pesenti et al. (2019) are implemented in the R package SWIM which is available on CRAN.

**Funding:** This research was funded by the Connaught Fund, the Canadian Statistical Sciences Institute (CANSSI), and the Natural Sciences and Engineering Research Council of Canada (NSERC) with funding reference numbers DGECR-2020-00333 and RGPIN-2020-04289.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** S.M.P. would like to thank Judy Mao for her help in implementing the numerical examples.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Appendix A. Proofs**

**Proof of Theorem 1.** We solve the optimisation on the set of quantile functions  $\check{\mathcal{M}}$  and define the Lagrangian with Lagrange multipliers  $\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{R}^d$

$$\begin{aligned} L(\check{G}, \lambda) &= \int_0^1 (\check{G}(u) - \check{F}(u))^2 - 2 \sum_{k=1}^d \lambda_k (\check{G}(u) \gamma_k(u) - r_k) \, du \\ &= \int_0^1 \left( \check{G}(u) - \left( \check{F}(u) + \sum_{k=1}^d \lambda_k \gamma_k(u) \right) \right)^2 \\ &\quad - 2 \sum_{k=1}^d \lambda_k (\check{F}(u) \gamma_k(u) - r_k) - \left( \sum_{k=1}^d \lambda_k \gamma_k(u) \right)^2 \, du. \end{aligned}$$

Thus, the optimisation problem (3) is equivalent to first solving, for fixed  $\lambda$ , the optimisation problem

$$\arg \min_{\check{G} \in \check{\mathcal{M}}} L(\check{G}, \lambda) \tag{A1}$$

and then finding  $\lambda$  such that the constraints are fulfilled. For fixed  $\lambda$ , the solution to (A1) is equal to the solution to

$$\arg \min_{\check{G} \in \check{\mathcal{M}}} \int_0^1 \left( \check{G}(u) - \left( \check{F}(u) + \sum_{k=1}^d \lambda_k \gamma_k(u) \right) \right)^2 \, du,$$

which is given by the isotonic projection of  $\check{F}(u) + \sum_{k=1}^d \lambda_k \gamma_k(u)$  onto the set  $\check{\mathcal{M}}$  and the Lagrange multipliers are such that the constraints are satisfied. Existence of the Lagrange multipliers follows since the set  $\mathcal{M}$  is non-empty. Uniqueness follows by convexity of the Wasserstein distance, by convexity of the constraints on the set of quantile functions.  $\square$

**Proof of Proposition 1.** For coherent distortion risk measures the corresponding weight function  $\gamma$  is non-decreasing. Moreover the optimal quantile function is given by Theorem 1 and is of the form  $\check{G}_\lambda(u) = (\check{F}(u) + \lambda \gamma(u))^\uparrow$  for some  $\lambda$  such that  $\check{G}_\lambda$  fulfils the constraint. The choice

$$\lambda^* = \frac{r - \rho_\gamma(F)}{\int_0^1 (\gamma(u))^2 \, du} \geq 0$$

implies that  $\check{G}_{\lambda^*}(u) = \check{F}(u) + \lambda^* \gamma(u)$  is a quantile function of the form (4) that fulfils the constraint. By uniqueness of Theorem 1,  $\check{G}_{\lambda^*}$  is indeed the unique solution.  $\square$

**Proof of Theorem 2.** Since both constraints are convex in  $\check{G}$  the Lagrangian with parameters  $\lambda = (\lambda_1, \dots, \lambda_d)$  and  $\tilde{\lambda} = (\tilde{\lambda}_1, \dots, \tilde{\lambda}_{\bar{d}}) \geq 0$  becomes

$$\begin{aligned} L(\check{G}, \lambda, \tilde{\lambda}) &= \int_0^1 (\check{G}(u) - \check{F}(u))^2 + 2 \sum_{k=1}^d \lambda_k (h_k(u)\check{G}(u) - c_k) du \\ &\quad + \sum_{k=1}^{\bar{d}} \tilde{\lambda}_k (\tilde{h}_k(u)(\check{G}(u))^2 - \tilde{c}_k) du \\ &= \int_0^1 \tilde{\Lambda}(u) \left( \check{G}(u) - \frac{1}{\tilde{\Lambda}(u)} \left( \check{F}(u) - \sum_{k=1}^d \lambda_k h_k(u) \right) \right)^2 \\ &\quad - \frac{1}{\tilde{\Lambda}(u)} \left( \check{F}(u) - \sum_{k=1}^d \lambda_k h_k(u) \right)^2 + (\check{F}(u))^2 - 2 \sum_{k=1}^d \lambda_k c_k - \sum_{k=1}^{\bar{d}} \tilde{\lambda}_k \tilde{c}_k, \end{aligned}$$

where  $\tilde{\Lambda}(u) = 1 + \sum_{k=1}^{\bar{d}} \tilde{\lambda}_k \tilde{h}_k(u)$ . Since  $\tilde{\lambda} \geq 0$  by the KKT-condition, we obtain that  $\tilde{\Lambda}(u) \geq 0$  for all  $u \in [0, 1]$ . Therefore, for fixed  $\lambda, \tilde{\lambda}$ , using an argument similar to the proof of Theorem 1, the solution (as a function of  $\lambda, \tilde{\lambda}$ ) is given by the weighted isotonic projection of  $\frac{1}{\tilde{\Lambda}(u)} (\check{F}(u) - \sum_{k=1}^d \lambda_k h_k(u))$ , with weight function  $\tilde{\Lambda}(\cdot)$ .  $\square$

**Proof of Proposition 2.** The mean and variance constraint can be rewritten as

$$\begin{aligned} m' &= \int x dG(x) = \int_0^1 \check{G}(u) du \quad \text{and} \\ (\sigma')^2 &= \int (x - m')^2 dG(x) = \int_0^1 (\check{G}(u) - m')^2 du. \end{aligned}$$

Thus, the Lagrangian with Lagrange multipliers  $\lambda = (\lambda_1, \dots, \lambda_{k+2})$  is, if  $\lambda_2 \neq -1$ ,

$$\begin{aligned} L(\check{G}, \lambda) &= \int_0^1 (\check{G}(u) - \check{F}(u))^2 du - 2\lambda_1 \left( \int_0^1 \check{G}(u) du - m' \right) \\ &\quad + \lambda_2 \left( \int_0^1 (\check{G}(u) - m')^2 du - (\sigma')^2 \right) \\ &\quad - 2 \sum_{k=1}^d \lambda_{k+2} \left( \int_0^1 \check{G}(u) \gamma_k(u) du - r_k \right) \\ &= (1 + \lambda_2) \int_0^1 \left( \check{G}(u) - \frac{1}{1 + \lambda_2} \left( \check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right) \right)^2 \\ &\quad - \frac{1}{1 + \lambda_2} \left( \check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right)^2 + (\check{F}(u))^2 du \\ &\quad + 2\lambda_1 m' + \lambda_2 \left( (m')^2 - (\sigma')^2 \right) + 2 \sum_{k=1}^d \lambda_{k+2} r_k. \end{aligned}$$

For fixed Lagrange multipliers  $\lambda$  with  $\lambda_2 \neq -1$ , the optimal quantile function is characterised by the isotonic projection and given by (using an analogous argument to the proof of Theorem 1)

$$\begin{aligned}
 \check{G}^*(u) &= \left( \frac{1}{1 + \lambda_2} \left( \check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right) \right)^\uparrow \\
 &= \frac{1}{|1 + \lambda_2|} \left( \operatorname{sgn}(1 + \lambda_2) \left( \check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right) \right)^\uparrow \\
 &= \frac{1}{|1 + \lambda_2|} \check{H}(u),
 \end{aligned}
 \tag{A2}$$

where we define  $\check{H}(u) = \left( \operatorname{sgn}(1 + \lambda_2) \left( \check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right) \right)^\uparrow \in \check{\mathcal{M}}$ , and  $\operatorname{sgn}(\cdot)$  denotes the sign function. Next we show that  $\lambda_2$  cannot be in a neighbourhood of  $-1$ . It holds that for  $\lambda_2 \neq -1$ ,

$$\int_0^1 (\check{G}^*(u))^2 du = \frac{1}{(1 + \lambda_2)^2} \int_0^1 (\check{H}(u))^2 du.
 \tag{A3}$$

Since the rhs of (A3) is increasing for  $|\lambda_2 + 1| \searrow 0$ , there exists a  $\varepsilon_0 > 0$  such that for all  $\varepsilon < \varepsilon_0$  and  $\lambda_2 \in (-1 - \varepsilon, -1 + \varepsilon)$ , it holds that

$$\frac{1}{(1 + \lambda_2)^2} \int_0^1 (\check{H}(u))^2 du > (\sigma')^2 + (m')^2,$$

which is a contradiction to the optimality of  $\check{G}^*$ . Thus,  $\lambda_2$  is indeed bounded away from  $-1$  and the unique solution is given in (A2).  $\square$

**Proof of Theorem 3.** We split this proof into the two cases (i), that is constraint (a) and (ii), i.e., constraint (b).

Case (i): For constraint a), i.e.,  $\operatorname{VaR}_\alpha(G) = q$ , we first assume that  $q \leq \operatorname{VaR}_\alpha(F)$  which implies  $\check{F}(\alpha_F) = q \leq \check{F}(\alpha)$  and thus  $\alpha_F \leq \alpha$ . Therefore,  $\check{G}^*(u) = \check{F}(u) + (q - \check{F}(u)) \mathbb{1}_{\{u \in (\alpha_F, \alpha]\}}$  is a quantile function which satisfies the constraint. Next, we show that  $G^*$  has a smaller Wasserstein distance to  $F$  than any other distribution function satisfying the constraint. For this, let  $\check{H}$  be a quantile function satisfying the constraint and  $\check{H}(u) \neq \check{G}(u)$  on a measurable set of non-zero measure. Then

$$\begin{aligned}
 W_2(H, F) &= \int_0^{\alpha_F} (\check{H}(u) - \check{F}(u))^2 du + \int_{\alpha_F}^\alpha (\check{H}(u) - \check{F}(u))^2 du + \int_\alpha^1 (\check{H}(u) - \check{F}(u))^2 du \\
 &\geq \int_{\alpha_F}^\alpha (\check{H}(u) - \check{F}(u))^2 du.
 \end{aligned}$$

By non-decreasingness of  $\check{H}$  and  $\check{F}$  and by the constraint it holds for all  $u \in [\alpha_F, \alpha]$  that  $\check{H}(u) \leq \check{H}(\alpha) = q = \check{F}(\alpha_F) \leq \check{F}(u)$ . Thus, on the interval  $[\alpha_F, \alpha]$ , we obtain  $(\check{H}(u) - \check{F}(u))^2 \geq (q - \check{F}(u))^2$  and therefore

$$W_2(H, F) \geq \int_{\alpha_F}^\alpha (\check{H}(u) - \check{F}(u))^2 du \geq \int_{\alpha_F}^\alpha (q - \check{F}(u))^2 du = W_2(G^*, F),$$

where at least one inequality is strict since  $\check{H}(u) \neq \check{G}(u)$  on a measurable set of non-zero measure. Uniqueness follows by the strict convexity of the Wasserstein distance and since the constraint is convex on the set of quantile functions.

Second, we assume that  $q > \operatorname{VaR}_\alpha(F)$  and show that there does not exist a solution. Assume by contradiction that  $\check{G}$  is an optimal quantile function satisfying the constraint. By definition of  $\alpha_F$ , we have that  $q = \check{F}(\alpha_F) > \check{F}(\alpha)$  and thus  $\alpha_F \geq \alpha$ . We apply a similar argument to the first part of the proof using non-decreasingness of  $\check{G}$ ,  $\check{G}(\alpha) = q$ , and optimality of  $\check{G}$ , to obtain that  $\check{G}$  is constant equal to  $q$  on  $[\alpha, \alpha_F]$  and equal to  $\check{F}(u)$  for  $u > \alpha_F$ . Specifically, it holds that

$$\check{G}(u) = \check{F}(u) + (q - \check{F}(u))\mathbb{1}_{\{u \in (\alpha, \alpha_F]\}}, \quad \text{for all } u > \alpha.$$

Moreover, since the optimal quantile function minimises the Wasserstein distance to  $F$ , it holds that, for all  $\varepsilon > 0$ ,  $\check{G}$  satisfies

$$\check{G}(u) = \check{F}(u), \quad \text{for all } u \leq \alpha - \varepsilon.$$

Thus, we can define for all  $\varepsilon \in (0, \alpha)$  the family of quantile functions

$$\check{H}_\varepsilon(u) = \check{F}(u) + (q - \check{F}(u))\mathbb{1}_{\{u \in (\alpha - \varepsilon, \alpha_F]\}},$$

which satisfies  $W_2(H_{\varepsilon_1}, F) < W_2(H_{\varepsilon_2}, F)$  for all  $0 \leq \varepsilon_1 < \varepsilon_2$ , and  $\check{H}_\varepsilon(\alpha) = q$  for all  $\varepsilon > 0$ . However,  $\lim_{\varepsilon \searrow 0} \check{H}_\varepsilon(\alpha) = \check{F}(\alpha) < q$  and thus the quantile function  $\lim_{\varepsilon \searrow 0} \check{H}_\varepsilon(u)$  does not fulfil the constraint. Hence, we obtain a contradiction to the optimality of  $\check{G}$ .

Case (ii): First, we assume that  $q \geq \text{VaR}_\alpha^+(F)$  which implies that  $\check{F}(\alpha_F) = q \geq \check{F}^+(\alpha) \geq \check{F}(\alpha)$  and thus  $\alpha_F \geq \alpha$ . Therefore  $\check{G}^*(u) = \check{F}(u) + (q - \check{F}(u))\mathbb{1}_{\{u \in (\alpha, \alpha_F]\}}$  is a quantile function. Moreover,  $\check{G}^*$  satisfies the constraint since by right-continuity of  $\check{G}^*$ , we have that

$$\check{G}^{*+}(\alpha) = \lim_{\varepsilon \searrow 0} \check{G}^{*+}(\alpha + \varepsilon) = q.$$

The proof that  $\check{G}^*$  has the smallest Wasserstein distance to  $F$  compared to any other distribution function satisfying the constraint is analogous to the one in case (i).

For the case when  $q > \text{VaR}_\alpha^+(F)$ , the argument of non-existence of the solution follows using similar arguments as those in case (i).  $\square$

**Proof of Theorem 4.** By concavity of the utility function, the constraint is convex and can be written as  $-\int_0^1 u(\check{G}(v)) dv + c \leq 0$ . Thus, we can define the Lagrangian with  $\lambda_1 \geq 0$  and  $(\lambda_2, \dots, \lambda_{d+1}) \in \mathbb{R}^d$  by

$$\begin{aligned} L(\check{G}, \lambda) &= \frac{1}{2} \int_0^1 (\check{G}(v) - \check{F}(v))^2 - \lambda_1 (u(\check{G}(v)) - c) - \sum_{k=1}^d \lambda_{k+1} (\check{G}(v) \gamma_k(v) - r_k) dv \\ &= \int_0^1 T(\check{G}(v)) - \check{G}(v) \left( \check{F}(v) + \sum_{k=1}^d \lambda_{k+1} \gamma_k(v) \right) \\ &\quad + \frac{1}{2} (\check{F}(v))^2 + \lambda_1 c + \sum_{k=1}^d \lambda_{k+1} r_k dv, \end{aligned}$$

where  $T(x) = \frac{1}{2}x^2 - \lambda_1 u(x)$ . Therefore, for fixed  $\lambda_1, \dots, \lambda_{d+1}$ , we apply Theorem 3.1 by Barlow and Brunk (1972) and obtain the unique optimal quantile function (as a function of  $\lambda_1, \dots, \lambda_{d+1}$ ), that is  $\check{G}^*(v) = \check{\nu}_{\lambda_1} \left( (\check{F}(v) + \sum_{k=1}^d \lambda_{k+1} \gamma_k(v))^\uparrow \right)$ , where  $\check{\nu}_{\lambda_1}$  is the left-inverse of  $\nu_{\lambda_1}(x) = x - \lambda_1 u'(x)$ .

Next, we show that if  $d = 0$ , the utility constraint is binding, that is  $\lambda_1 > 0$ . For this, assume by contradiction that the  $\lambda_1 = 0$ , then the optimal quantile function becomes  $\check{G}^*(u) = \check{\nu}_0(\check{F}(u))$ . Since  $\nu_0(x) = x$ , we obtain that  $\check{G}^*(u) = \check{F}(u)$ .  $\check{F}$ , however, does not fulfil the constraint, which is a contradiction to the optimality of  $\check{G}^*$ .  $\square$

**Proof of Proposition 3.** We prove the properties one-by-one:

(i) We first define for a random variable  $Z$  with  $\mathbb{P}$ -distribution  $F_Z$  the random variable  $U_Z := F_Z(Z)$ . Then,  $U_Z$  and  $Z$  are comonotonic and  $U_Z$  has a uniform distribution under  $\mathbb{P}$ . Next, recall that for any random variables  $Y_1, Y_2$  it holds that Rüschemdorf (1983)

$$\mathbb{E} \left[ Y_1 F_{Y_2}^{-1}(1 - U_{Y_1}) \right] \leq \mathbb{E}[Y_1 Y_2] \leq \mathbb{E} \left[ Y_1 F_{Y_2}^{-1}(U_{Y_1}) \right]. \tag{A4}$$

where  $F_{Y_2}^{-1}(U_{Y_1})$  is the random variable that is comonotonic to  $Y_1$  and has the same  $\mathbb{P}$ -distribution as  $Y_2$ . Similarly,  $F_{Y_2}^{-1}(1 - U_{Y_1})$  is the random variable that is counter-monotonic to  $Y_1$  and has the same  $\mathbb{P}$ -distribution as  $Y_2$ . The left (right) inequality in (A4) become equality if and only if the random variables  $Y_1$  and  $Y_2$  are counter-comonotonic (comonotonic).

Thus, we can rewrite the maximum in the normalising constant of the reverse sensitivity measure as follows

$$\max_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X)] = \max_{Z \stackrel{\mathbb{P}}{=} \frac{d\mathbb{Q}^*}{d\mathbb{P}}} \mathbb{E}[s(X) Z] = \mathbb{E} \left[ s(X) F_{\frac{d\mathbb{Q}^*}{d\mathbb{P}}}^{-1} \left( U_{s(X)} \right) \right]$$

and the minimum in the normalising constant is

$$\min_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X)] = \min_{Z \stackrel{\mathbb{P}}{=} \frac{d\mathbb{Q}^*}{d\mathbb{P}}} \mathbb{E}[s(X) Z] = \mathbb{E} \left[ s(X) F_{\frac{d\mathbb{Q}^*}{d\mathbb{P}}}^{-1} \left( 1 - U_{s(X)} \right) \right].$$

The reverse sensitivity for the case  $\mathbb{E}^{\mathbb{Q}^*}[s(X_i)] \geq \mathbb{E}[s(X_i)]$  then becomes

$$S_i^{\mathbb{Q}^*} = \frac{\mathbb{E}[s(X_i) \frac{d\mathbb{Q}^*}{d\mathbb{P}}] - \mathbb{E}[s(X_i)]}{\mathbb{E} \left[ s(X) F_{\frac{d\mathbb{Q}^*}{d\mathbb{P}}}^{-1} \left( U_{s(X)} \right) \right] - \mathbb{E}[s(X_i)]},$$

which satisfies  $0 \leq S_i^{\mathbb{Q}^*} \leq 1$  using again (A4). For the case  $\mathbb{E}^{\mathbb{Q}^*}[s(X_i)] \leq \mathbb{E}[s(X_i)]$ , it holds that

$$S_i^{\mathbb{Q}^*} = - \frac{\mathbb{E}[s(X_i) \frac{d\mathbb{Q}^*}{d\mathbb{P}}] - \mathbb{E}[s(X_i)]}{\mathbb{E} \left[ s(X) F_{\frac{d\mathbb{Q}^*}{d\mathbb{P}}}^{-1} \left( 1 - U_{s(X)} \right) \right] - \mathbb{E}[s(X_i)]},$$

which satisfies  $-1 \leq S_i^{\mathbb{Q}^*} \leq 0$ .

(ii) Assume that  $s(X_i)$  and  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$  are independent under  $\mathbb{P}$ , then

$$\mathbb{E}[s(X_i) \frac{d\mathbb{Q}^*}{d\mathbb{P}}] = \mathbb{E}[s(X_i)] \mathbb{E} \left[ \frac{d\mathbb{Q}^*}{d\mathbb{P}} \right] = \mathbb{E}[s(X_i)],$$

and the reverse sensitivity measure is indeed zero.

(iii) From property (i) we observe that  $s(X_i)$  and  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$  are comonotonic, if and only if,  $S_i^{\mathbb{Q}^*} = 1$  since in this case the right inequality in Equation (A4) becomes equality.

(iv) From property (i) we observe that  $s(X_i)$  and  $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$  are counter-comonotonic, if and only if, then  $S_i^{\mathbb{Q}^*} = -1$  as in this case left inequality in Equation (A4) becomes equality.

The proof that the joint reverse sensitivity  $S_{i,j}^{\mathbb{Q}^*}$  also fulfils the above properties follows using analogous arguments and replacing  $s(X_i)$  with  $s(X_i, X_j)$ .  $\square$

## References


- Acerbi, Carlo, and Dirk Tasche. 2002. On the coherence of Expected Shortfall. *Journal of Banking & Finance* 26: 1487–503.
- Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. Coherent measures of risk. *Mathematical Finance* 9: 203–28. [CrossRef]
- Asimit, Vali, Liang Peng, Ruodu Wang, and Alex Yu. 2019. An efficient approach to quantile capital allocation and sensitivity analysis. *Mathematical Finance* 29: 1131–56. [CrossRef]
- Barlow, Richard E., and Hugh D. Brunk. 1972. The isotonic regression problem and its dual. *Journal of the American Statistical Association* 67: 140–47. [CrossRef]
- Barlow, Richard E., David J. Bartholomew, John. M. Bremner, and Hugh D. Brunk. 1972. *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression*. Hoboken: Wiley.
- Bernard, Carole, Silvana M. Pesenti, and Steven Vanduffel. 2020. Robust distortion risk measures. *arXiv arXiv:2205.08850*.

- Blanchet, Jose, and Karthyek Murthy. 2019. Quantifying distributional model risk via optimal transport. *Mathematics of Operations Research* 44: 565–600. [CrossRef]
- Borgonovo, Emanuele. 2007. A new uncertainty importance measure. *Reliability Engineering & System Safety* 92: 771–84.
- Borgonovo, Emanuele, and Elmar Plischke. 2016. Sensitivity analysis: A review of recent advances. *European Journal of Operational Research* 48: 869–87. [CrossRef]
- Borgonovo, Emanuele, Gordon B. Hazen, and Elmar Plischke. 2016. A common rationale for global sensitivity measures and their estimation. *Risk Analysis* 36: 1871–95. [CrossRef]
- Borgonovo, Emanuele, Gordon B. Hazen, Victor Richmond R. Jose, and Elmar Plischke. 2021. Probabilistic sensitivity measures as information value. *European Journal of Operational Research* 289: 595–610. [CrossRef]
- Cambou, Mathieu, and Damir Filipović. 2017. Model uncertainty and scenario aggregation. *Mathematical Finance* 27: 534–67. [CrossRef]
- Cont, Rama, Romain Deguest, and Giacomo Scandolo. 2010. Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance* 10: 593–606. [CrossRef]
- Cuestaalbertos, Juan-Alberto, Ludger Rüschendorf, and Araceli Tuerodiaz. 1993. Optimal coupling of multivariate distributions and stochastic processes. *Journal of Multivariate Analysis* 46: 335–61. [CrossRef]
- Dall’Aglia, Giorgio. 1956. Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze* 10: 35–74.
- De Leeuw, Jan, Kurt Hornik, and Patrick Mair. 2010. Isotone optimization in R: Pool-adjacent-violators algorithm (pava) and active set methods. *Journal of Statistical Software* 32: 1–24. [CrossRef]
- Denuit, Michel, Jan Dhaene, Marc Goovaerts, and Rob Kaas. 2006. *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. Hoboken: John Wiley & Sons.
- Fissler, Tobias, and Silvana M. Pesenti. 2022. Sensitivity measures based on scoring functions. *arXiv* arXiv:2203.00460.
- Fort, Jean-Claude, Thierry Klein, and Agnès Lagnoux. 2021. Global sensitivity analysis and Wasserstein spaces. *SIAM/ASA Journal on Uncertainty Quantification* 9: 880–921. [CrossRef]
- Gamboa, Fabrice, Pierre Gremaud, Thierry Klein, and Agnès Lagnoux. 2020. Global sensitivity analysis: A new generation of mighty estimators based on rank statistics. *arXiv* arXiv:2003.01772.
- Gamboa, Fabrice, Thierry Klein, and Agnès Lagnoux. 2018. Sensitivity analysis based on Cramér–von Mises distance. *SIAM/ASA Journal on Uncertainty Quantification* 6: 522–48. [CrossRef]
- Hall, Peter, and Li-Shan Huang. 2001. Nonparametric kernel regression subject to monotonicity constraints. *The Annals of Statistics* 29: 624–47. [CrossRef]
- Kruse, Thomas, Judith C. Schneider, and Nikolaus Schweizer. 2019. The joint impact of f-divergences and reference models on the contents of uncertainty sets. *Operations Research* 67: 428–35. [CrossRef]
- Kusuoka, Shigeo. 2001. On law invariant coherent risk measures. In *Advances in Mathematical Economics*. Berlin and Heidelberg: Springer, pp. 83–95.
- Makam, Vaishno Devi, Pietro Millossovich, and Andreas Tsanakas. 2021. Sensitivity analysis with  $\chi^2$ -divergences. *Insurance: Mathematics and Economics* 100: 372–83. [CrossRef]
- Maume-Deschamps, Véronique, and Ibrahima Niang. 2018. Estimation of quantile oriented sensitivity indices. *Statistics & Probability Letters* 134: 122–27.
- Meyer, Mary C. 2008. Inference using shape-restricted regression splines. *The Annals of Applied Statistics* 2: 1013–33. [CrossRef]
- Moosmüller, Caroline, Felix Dietrich, and Ioannis G. Kevrekidis. 2020. A geometric approach to the transport of discontinuous densities. *SIAM/ASA Journal on Uncertainty Quantification* 8: 1012–35. [CrossRef]
- Pesenti, Silvana M., Alberto Bettini, Pietro Millossovich, and Andreas Tsanakas. 2021. Scenario weights for importance measurement (SWIM)—An R package for sensitivity analysis. *Annals of Actuarial Science* 15: 458–83. [CrossRef]
- Pesenti, Silvana M., Pietro Millossovich, and Andreas Tsanakas. 2019. Reverse sensitivity testing: What does it take to break the model? *European Journal of Operational Research* 274: 654–70. [CrossRef]
- Pesenti, Silvana M., Pietro Millossovich, and Andreas Tsanakas. 2021. Cascade sensitivity measures. *Risk Analysis* 31: 2392–414. [CrossRef]
- Plischke, Elmar, and Emanuele Borgonovo. 2019. Copula theory and probabilistic sensitivity analysis: Is there a connection? *European Journal of Operational Research* 277: 1046–59. [CrossRef]
- Rahman, Sharif. 2016. The f-sensitivity index. *SIAM/ASA Journal on Uncertainty Quantification* 4: 130–62. [CrossRef]
- Rüschendorf, Ludger. 1983. Solution of a statistical optimization problem by rearrangement methods. *Metrika* 30: 55–61. [CrossRef]
- Saltelli, Andrea, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. 2008. *Global Sensitivity Analysis: The Primer*. Hoboken: John Wiley & Sons.
- Sysoev, Oleg, and Oleg Burdakov. 2019. A smoothed monotonic regression via l2 regularization. *Knowledge and Information Systems* 59: 197–218. [CrossRef]
- Tsanakas, Andreas, and Pietro Millossovich. 2016. Sensitivity analysis using risk measures. *Risk Analysis* 36: 30–48. [CrossRef] [PubMed]
- Villani, Cédric. 2008. *Optimal transport: Old and New*. Berlin and Heidelberg: Springer Science & Business Media, vol. 338.





# Special-Rate Life Annuities: Analysis of Portfolio Risk Profiles

Ermanno Pitacco \* and Daniela Y. Tabakova

Insurance and Risk Management, MIB Trieste School of Management, Largo Caduti di Nasiriya 1, 34142 Trieste, Italy; tabakova@mib.edu

\* Correspondence: ermanno.pitacco@deams.units.it

**Abstract:** Special-rate life annuities are life annuity products whose single premium is based on a mortality assumption driven (at least to some extent) by the health status of the applicant. The health status is ascertained via an appropriate underwriting step (which explains the alternative expression “underwritten life annuities”). Better annuity rates are then applied in presence of poor health conditions. The worse the health conditions, the smaller the modal age at death (as well as the expected lifetime), but the higher the variance of the lifetime distribution. The latter aspect is due to significant data scarcity as well as to the mix of possible pathologies leading to each specific rating class. A higher degree of (partially unobservable) heterogeneity inside each sub-portfolio of special-rate annuities follows, and this results in a higher variability of the total portfolio payout. The present research aims at analyzing the impact of extending the life annuity portfolio by selling special-rate life annuities. Numerical evaluations have been performed by adopting a deterministic approach as well as a stochastic one, according to diverse assumptions concerning both lifetime distributions and portfolio structure and size. Our achievements witness the possibility of extending the annuity business without taking huge amounts of risk. Hence, the risk management objective “enhancing the company market share” can be pursued without significant worsening of the annuity portfolio risk profile.

**Keywords:** life annuities; standard annuities; underwritten annuities; enhanced annuities; impaired annuities; preferred risks; substandard lives

**Citation:** Pitacco, Ermanno, and Daniela Y. Tabakova. 2022.

Special-Rate Life Annuities: Analysis of Portfolio Risk Profiles. *Risks* 10: 65. <https://doi.org/10.3390/risks10030065>

Academic Editor: Nadine Gatzert

Received: 5 February 2022

Accepted: 10 March 2022

Published: 13 March 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction and Motivation

Considerable attention is currently being devoted in insurance work (and, in particular, in the actuarial work) to the management of life annuity portfolios and to the annuity product design, because of the growing importance of annuity benefits paid by private pension schemes and individual policies.

In particular, the progressive shift in many countries from defined benefit to defined contribution pension schemes has increased the interest in life annuity products with a guaranteed periodic benefit. Nevertheless, various “weak” features of the (standard) life annuities should be noted, looking at the product from both the annuity provider’s and the customer’s perspective.

However, many features can be improved by moving from traditional products to more complex products, for example, by adding riders (that is, supplementary benefits), or by adopting restrictions on the age intervals covered, or by allowing for individual risk factors; hence, “tailoring” the annuity rates (at least to some extent) to specific features of the customer.

Special-rate life annuities are life annuity products whose single premium is based on a mortality assumption driven by the health status of the applicant. The health status is ascertained via an appropriate underwriting step (which explains the alternative expression “underwritten life annuities”). Better annuity rates are then applied in presence of poor health conditions. The worse the health conditions, the smaller the modal age at death (as well as the expected lifetime), but the higher the variance of the lifetime distribution. The

latter aspect is due to significant data scarcity as well as to the mix of possible pathologies leading to each specific rating class. A higher degree of (partially unobservable) heterogeneity inside each sub-portfolio of special-rate annuities follows, and this results in a higher variability of the total portfolio payout. By selling special-rate annuities, on the one hand, a higher premium income can be expected, and on the other, a higher variability of the portfolio payout must be faced. What about the “balance”? Our achievements witness the possibility of extending the annuity business without taking huge amounts of risk. Hence, the risk management objective “enhancing the company market share” can be pursued without significant worsening of the annuity portfolio risk profile.

Diverse input data might lead to worse risk profiles. An appropriate sensitivity testing can then help in checking risk profile changes. Assuming the extension of the life annuity business as the insurer’s target, the present paper aims at providing a simple technical tool for assessing how and to what extent selling special-rate annuities impacts the portfolio risk profile. The structure of the proposed tool arises from a trade-off between strictly pragmatic approaches (frequently adopted in current actuarial practice) and rigorous mathematical settings (which may result in implementation difficulties, notably because of data scarcity).

The remainder of the paper is organized as follows. A compact literature review is provided in Section 2, while Section 3 describes the main products in the area of special-rate annuities. Biometric assumptions underlying the assessment of portfolio risk profiles are defined and commented in Section 4. In Section 5, portfolio structures are specified in terms of sub-portfolio sizes; then, results of interest, which can express the portfolio risk profile, are defined. Our main achievements are discussed in Sections 6 and 7 where numerical results obtained by adopting a deterministic and a stochastic approach are respectively presented. Finally, Section 8 concludes the paper.

## 2. Literature Review

The main features of life annuity products are discussed in many life insurance and actuarial textbooks. A presentation of basic actuarial models for premium and reserve calculations for life annuities as well as a discussion of possible innovations in life annuity products are provided by Pitacco (2021).

Heterogeneity in mortality and risk classification constitute the natural frameworks in which the basic features of special-rate life annuities can be analyzed. Risk classification in life insurance and life annuities is addressed in many books and papers; a compact review, together with an extensive reference list, is provided by Haberman and Olivieri (2014). The impact of risk classification on the structure of life annuity portfolios is dealt with by Gatzert et al. (2012), Hoermann and Russ (2008) and Olivieri and Pitacco (2016).

The impact of heterogeneity on portfolio results and the consequent capital requirements are analyzed by Denuit and Frostig (2006). More specifically, Denuit and Frostig (2007) focusses on heterogeneity among lifetimes in the context of stochastic mortality according to the Lee-Carter model.

Heterogeneity in mortality is due to both observable and unobservable risk factors. The reader can refer to Pitacco (2019) for a literature review from an actuarial perspective, as well as for a discussion of models, which can be used to represent specific mortality rates accounting for observable risk factors.

Special-rate life annuities are described in various papers and technical reports: see, in particular Ainslie (2000), Drinkwater et al. (2006), Ridsdale (2012) and Rinke (2002). The article by Edwards (2008) is specifically devoted to life annuity rating based on the postcode (that is, a proxy for social class and location of housing).

An interesting analysis of market issues related to special-rate annuities is presented by Gatzert and Klotzki (2016), where barriers on the supply side and the demand side are in particular addressed. Practical aspects of pricing special-rate life annuities are dealt with by Gracie and Makin (2006) and James (2016).

Special-rate annuities in the context of new product development (NPD) processes are addressed in Chapter 9 of Pitacco (2020). In particular, the NPD according to the structure

of the risk management process and the logic of the Stage-Gate<sup>®</sup> process is described and commented on<sup>1</sup>.

Underwriting for special-rate life annuities can be implemented in a number of ways, and several classifications can be conceived. An interesting classification has been proposed by Rinke (2002), and summarized in Chapter 9 of Pitacco (2021).

Statistics regarding extra-mortality by various causes are beyond the scope of this paper. Here, we only cite the contribution by Weinert (2006), which has suggested some baseline choices for the lifetime distributions. For a list of references the reader can refer to Pitacco and Tabakova (2020).

### 3. The Products

The terminology adopted in the technical literature to denote the various types of special-rate annuities is not univocally defined. For example, the term “enhanced annuity” is frequently used in the wider sense of all life annuities where, given the single premium, the annual benefit is of a higher level than the standard, due to some customer’s characteristics (health, lifestyle, etc.). In what follows, we refer to the terminology originally adopted to define special-rate annuities, which are sold in the UK market (see Ridsdale 2012). The same terminology has also been adopted in the book (Pitacco 2021), from which the following descriptions have been taken.

The following special-rate annuities are sold in several markets.

1. Given the single premium amount, a *lifestyle annuity* pays out benefits higher than a standard life annuity because of risk factors (e.g., smoking and drinking habits, marital status, occupation, height and weight, blood pressure and cholesterol levels), which might result, to some extent, in a shorter life expectancy. Specific lifestyle annuities are the following ones.
  - (a) *Smoker life annuities*: if the applicant has smoked at least a given number of cigarettes for a certain number of years, then they are eligible for a smoker annuity.
  - (b) Mortality differences between married and unmarried individuals underpin the use of special rates in pricing the *unmarried lives annuities*. The observed higher mortality rates of unmarried individuals justify a higher annuity rate.
2. The *enhanced life annuity* pays out an income to a person with a reduced life expectancy, in particular because of a personal history of medical conditions. Of course, the “enhancement” in the annuity benefit (compared to a standard-rate life annuity, same premium) comes from the use of a higher mortality assumption.
3. The *impaired life annuity* pays out a higher income than an enhanced life annuity, as a result of medical conditions which significantly shorten the life expectancy of the annuitant (e.g., diabetes, chronic asthma, cancer, etc.).
4. Finally, *care annuities* are aimed at individuals with very serious impairments or individuals who are already in a senescent-disability (or long-term care) state. These annuities are frequently placed in the context of long-term care insurance products, and labeled as providing benefits “in point of need” (see, for example, Pitacco 2014).

Thus, moving from type 1 to type 4 results in progressively higher mortality assumptions, shorter life expectancy, and hence, for a given single premium amount, in higher annuity benefits. Of course, an insurer can decide to offer a more limited set of products.

The applicant’s health status and, notably, the presence of past or current diseases is explicitly considered in the special-rate annuities of types 2, 3 and 4. Various factors concerning the health status can be accounted for, and medical ascertainment is of course required. In particular, the underwriting process for impaired-life annuities and care annuities must result in classifying the applicant as a *substandard risk*, because of ascertainment of significant extra-mortality. For this reason, annuities of types 3 and 4 are sometimes named *substandard life annuities*.

The above list of special-rate annuity types can be completed by the *postcode life annuities*, which constitute an interesting example of “environment-based” rating. The postcode can provide a proxy for social class and location of housing; that is, risk factors that may have a significant impact on the lifestyle and hence on the life expectancy. Then, its use as a rating factor for pricing life annuities can be justified.

#### 4. The Mortality Model

To assess present values (that is, random present values and expected present values) of benefits paid by special-rate annuities, a model quantifying the mortality of annuitants (standard annuitants and special-rate annuitants) must be chosen. The following aspects must be considered in particular:

- Heterogeneity in mortality inside the (total) life annuity portfolio (see Sections 4.1 and 4.2);
- Possible future mortality trends (see Section 4.1);
- The individual age-pattern of mortality (see Section 4.2).

##### 4.1. General Aspects

A higher degree of heterogeneity in mortality affects a life annuity portfolio also including special-rate annuities, with respect to a standard annuity portfolio. As is well known, heterogeneity may be due to both observable and unobservable risk factors, which imply diverse modeling choices. In the case of a life annuity portfolio, the practical problem is: to what extent the underwriting process can detect, for each applicant, the outcomes of the risk factors that entitle the individual to purchase a special-rate annuity? Even a rigorous underwriting process can leave some degree of “residual” heterogeneity inside each special-rate class, because:

- Some risk factors are unobservable;
- Diverse pathologies entitle one to the same annuity rate.

Given that some degree of heterogeneity inside each risk class is unavoidable, the problem is how to model its impact on the individual age-pattern of mortality. We focus on the two following choices.

The modeling of heterogeneity in mortality due to unobservable risk factors has found an elegant and rigorous solution in the concept of (constant) frailty, initially described by Beard (1959), but formally defined by Vaupel et al. (1979). A well known implementation of frailty modeling leads to the so-called Gompertz–Gamma model, which results in one of the Perks laws (see Perks 1932). A number of generalizations of the concept of frailty have been proposed, in particular looking at possible dynamic features of the individual frailty (a survey is presented, for example, by Pitacco 2019).

Frailty modeling to express heterogeneity in mortality has been adopted by various Authors: see, for example, Denuit and Frostig (2007), and, in the context of special-rate annuities, by Olivieri and Pitacco (2016).

Given the practical difficulties in calibrating the frailty model, the paper by Olivieri and Pitacco (2016) specifically aims at assessing, via sensitivity analysis, the impact of diverse assumptions for the frailty parameter values (notably, the Gamma parameters) on portfolio results of interest.

A simpler choice can conversely consist in implicitly expressing the presence of heterogeneity in mortality by directly assuming a higher variance in the individual lifetime distribution, as suggested by statistical data analyzed by various Authors (see, for example, Weinert 2006), that is, the stronger the assumed degree of heterogeneity, the higher the variance. If a mortality law is chosen to express the individual age-pattern of mortality, a sensitivity analysis can be performed also in this setting.

As far as the future mortality trends are concerned, the following aspects can be singled out:

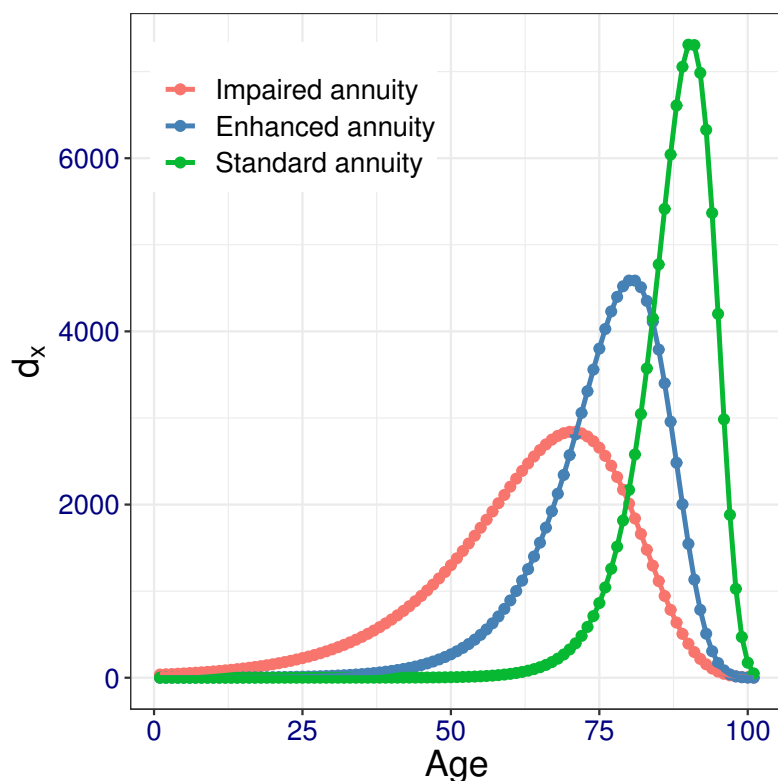
- An assumed future mortality trend can be taken into account by adopting projected life tables or projected mortality laws;
- Representing uncertainty in future mortality trend, which implies systematic risk (that is, the aggregate longevity risk), calls for the use of stochastic mortality models.

In what follows, we express heterogeneity inside each rating class via the variance of the individual lifetime distribution (see Section 4.2). We disregard systematic longevity risk, as heterogeneity mainly affects the idiosyncratic risk in each rating class; that is, the risk of random fluctuations around the expected value.

Assumptions regarding the relations among individual lifetimes must supplement the mortality model. Given the above hypotheses, in what follows we assume that individual lifetimes are independent random variables. While this is, of course, a simplifying assumption, correlation among lifetimes could be assumed, in particular to express uncertainty about future mortality trends in the context of a stochastic mortality model. Conditional independence would replace, in that case, the independence assumption.

#### 4.2. Age Pattern of Mortality

Three (hypothetical) curves showing expected number  $d_x$  of deaths between exact age  $x$  and  $x + 1$  out of a notional cohort of 100,000 individuals are shown in Figure 1. We recall that the expected numbers of deaths are proportional to the values of the probability density function in the time-continuous context defined below.



**Figure 1.** Curves of deaths for different life annuities.

The worse the health conditions, the smaller the modal age at death (as well as the life expectancy), but the higher the variance of the lifetime distribution. The latter aspect is due to the mix of possible pathologies leading to each specific individual classification (and also due to data scarcity). A higher degree of (partially unobservable) heterogeneity in mortality follows, inside each sub-portfolio of special-rate annuities. However, this heterogeneity can be reduced by restricting the range of pathologies that entitle one to a special-rate annuity, then making the relevant sub-portfolio more homogeneous.

It is also worth noting that most of the available mortality statistics refer to specific sets of pathologies (e.g., diabetes), rather than to broad sets of diseases. Splitting a class of special-rate annuities (for example, the impaired life annuities) into pathology-related subclasses can, hence, be an appropriate choice.

To describe the age patterns of mortality in quantitative terms, appropriate life tables can be chosen. However, the use of a mortality law (in particular a “simple” law) significantly eases the implementation of a sensitivity analysis, which can be performed by assigning diverse values to the relevant parameters. In line with the main purpose of life annuities, that is, providing a post-retirement income, adult and old ages have only been addressed in what follows. Then, the Gompertz law has been used to express the age-pattern of mortality. Parameter values have been chosen to represent the features of the curves of deaths, as described in Section 4.1.

In terms of the force of mortality,  $\mu_x$ , the Gompertz law is as follows:

$$\mu_x = B c^x, \quad \text{with } B, c > 0 \tag{1}$$

Instead of referring to the usual parametrization in (1), we refer to the “informative” parametrization (see, for example, Carriere 1992), that is:

$$\mu_x = \frac{1}{D} \exp\left(\frac{x - M}{D}\right), \quad \text{with } M, D > 0 \tag{2}$$

where  $M$  denotes the mode of the Gompertz probability density function and  $D$  a measure of dispersion. Relations with the usual parameters are as follows:

$$c = \exp\left(\frac{1}{D}\right) \tag{3}$$

$$B = \frac{\exp\left(-\frac{M}{D}\right)}{D} \tag{4}$$

Sensitivity analysis can simply be performed by assigning values to the mode parameter  $M$  and the dispersion parameter  $D$  (as described in Sections 6.2 and 7.2).

It can be proved that the elements of the corresponding life table  $\{\ell_x\}$ , with  $\ell_0 = 100,000$ , are given by the following expression:

$$\ell_x = 100,000 \times \exp\left(\exp\left(-\frac{M}{D}\right) - \exp\left(\frac{x - M}{D}\right)\right) \tag{5}$$

From the life table  $\{\ell_x\}$ , all the biometric functions of interest (e.g.,  $d_x$ ,  $q_x$ , etc.) can immediately be derived.

As already noted, the shape of the lifetime distribution can be driven by choosing specific values  $M_k$  and  $D_k$  for the parameters of Equation (2). The (baseline) parameter values shown in Table 1 determine the curves of death plotted in Figure 1. We note that the choice of the parameter values is only aimed at performing a sensitivity analysis and does not reflect real statistical data.

**Table 1.** Parameters of the Gompertz law.

Rating Class	$k$	$M_k$	$D_k$
Standard	1	90	5
Enhanced	2	80	8
Impaired	3	70	13

### 5. The Actuarial Model

After defining the portfolio structures used in the various evaluations, we define the quantities referred to in the deterministic and the stochastic assessments.

### 5.1. Portfolio Structures

We will consider a life annuity portfolio P generally consisting of three sub-portfolios:

- Sub-portfolio SP1 initially consisting of  $n_1$  standard life annuities;
- Sub-portfolio SP2 initially consisting of  $n_2$  enhanced life annuities;
- Sub-portfolio SP3 initially consisting of  $n_3$  impaired life annuities.

Of course, one of the  $n_k$  can be set equal to 0. Let  $n$  denote the size of the portfolio P, that is:

$$n = n_1 + n_2 + n_3$$

Assumptions underlying the actuarial model are as follows:

- The lifetime distribution for annuitants in the sub-portfolio SP $k$  follows the Gompertz law with parameters  $M_k$  and  $D_k$ ,  $k = 1, 2, 3$  (see Section 4.2);
- All the annuitants are age  $x$  at policy issue;
- The individual lifetimes in each sub-portfolio and in the portfolio P are independent random variables;
- The same benefit  $b$  is paid by all the life annuity policies;
- Each sub-portfolio is closed to new entries (and hence consists of a generation of policies).

### 5.2. Actuarial Values

Our ultimate object is to analyze the behavior of various quantities defined as functions of  $n_1, n_2, n_3$ , in particular: expected value, variance and coefficient of variation (risk index) of the portfolio payouts.

To this purpose, we first recall the basic formulae for a life annuity-immediate, with benefit  $b = 1$  paid to an individual age  $x$  at policy issue and assigned to sub-portfolio SP $k$ .

The expected present value (shortly, the actuarial value) of the annual benefits paid to the individual is given (according to the traditional actuarial notation) by:

$$a_x^{(k)} = \sum_{h=1}^{\omega-x} a_{h|} \cdot {}_h|q_x^{(k)} \tag{6}$$

where:

- $\omega$  denotes the maximum attainable age;
- $a_{h|} = \frac{1-(1+i)^{-h}}{i}$  is the present value of an annuity-certain, with  $i$  denoting the interest rate used for discounting;
- ${}_h|q_x^{(k)} = \frac{\ell_{x+h}^{(k)} - \ell_{x+h+1}^{(k)}}{\ell_x^{(k)}}$  is the probability of a person age  $x$  dying between age  $x + h$  and  $x + h + 1$ , according to the biometric model with parameters  $M_k$  and  $D_k$ .

For example, with the parameter values given in Table 1 and  $x = 65$ , Equation (6) yields:

$$\begin{aligned} a_{65}^{(1)} &= 17.29 \\ a_{65}^{(2)} &= 11.00 \\ a_{65}^{(3)} &= 8.20 \end{aligned}$$

The variance of the present value of the annual benefits is given by:

$$\sigma_x^{2(k)} = \sum_{h=1}^{\omega-x} a_{h|}^2 \cdot {}_h|q_x^{(k)} - \left(a_x^{(k)}\right)^2 \tag{7}$$

For example, with the above data, from Equation (7) we obtain:



$$\begin{aligned}\sigma_{65}^{2(1)} &= 16.858 \\ \sigma_{65}^{2(2)} &= 26.436 \\ \sigma_{65}^{2(3)} &= 27.446\end{aligned}$$

We denote with  $\mathbb{E}^{(k)}(n_k)$  and  $\mathbb{V}\text{ar}^{(k)}(n_k)$  the expected value and the variance of the benefit payouts of sub-portfolio SPk. Assuming a benefit  $b = 1$ , we obviously have:

$$\mathbb{E}^{(k)}(n_k) = n_k a_x^{(k)}; \quad k = 1, 2, 3 \tag{8}$$

and, thanks to the assumption of independence among the individual lifetimes:

$$\mathbb{V}\text{ar}^{(k)}(n_k) = n_k \sigma_x^{2(k)}; \quad k = 1, 2, 3 \tag{9}$$

For a generic portfolio P, consisting of  $n = n_1 + n_2 + n_3$  policies, we then find:

$$\mathbb{E}(n_1, n_2, n_3) = \sum_{k=1}^3 \mathbb{E}^{(k)}(n_k) \tag{10}$$

$$\mathbb{V}\text{ar}(n_1, n_2, n_3) = \sum_{k=1}^3 \mathbb{V}\text{ar}^{(k)}(n_k) \tag{11}$$

### 5.3. The Risk Index

The risk index (or coefficient of variation) is a relative risk measure that expresses the variability of a random quantity in terms of standard deviation per unit of expected value. It is frequently adopted in risk theory and risk management to assess the so-called pooling effect, that is, the diversification effect which is achieved by constructing a pool of risks.

For a generic portfolio P, the risk index  $\rho$  is defined as follows:

$$\rho(n_1, n_2, n_3) = \frac{\sqrt{\mathbb{V}\text{ar}(n_1, n_2, n_3)}}{\mathbb{E}(n_1, n_2, n_3)} \tag{12}$$

We note that  $\rho$  is a unit-free risk measure.

### 5.4. Cash Flows

Annual cash flows are, of course, random quantities. For the generic sub-portfolio SPk, the random cash flow, that is the sub-portfolio payout, at time  $t$ ,  $X_k(t)$ , depends on the number  $N_k(t)$  of annuitants alive at that time (out of the initial  $n_k$ ), and is of course given by:

$$X_k(t) = b N_k(t); \quad k = 1, 2, 3 \tag{13}$$

Referring to a generic portfolio P, consisting of three sub-portfolios, the total payout at time  $t$  is then given by:

$$X(t) = \sum_{k=1}^3 X_k(t) \tag{14}$$

## 6. Portfolio Risk Profiles: Deterministic Approach

In this Section, we first assess the impact, in terms of the risk index, of the portfolio structure on the portfolio risk profile. Biometric assumptions are as specified in Section 4.2, with parameter values given in Table 1 (if not otherwise stated).

We then assess the impact, again in terms of the risk index, of diverse biometric assumptions.

### 6.1. Impact of the Portfolio Structure

#### 6.1.1. Cases 1.1

We analyze the impact of the size of the sub-portfolio SP2 of enhanced annuities. Then:

$$\begin{aligned}n_1 &= 10,000 \\n_2 &= 100, 200, \dots, 1000 \\n_3 &= 0\end{aligned}$$

Results are shown in Table 2.

**Table 2.** Cases 1.1—Impact of the portfolio structure on the risk index.

Portfolio	$n_2$	$\rho(10,000, n_2, 0)$
P01	100	0.002378268
P02	200	0.002381505
P03	300	0.002384564
P04	400	0.002387452
P05	500	0.002390176
P06	600	0.002392740
P07	700	0.002395152
P08	800	0.002397415
P09	900	0.002399535
P10	1000	0.002401517

#### 6.1.2. Cases 1.2

We analyze the impact of the size of the sub-portfolio SP3 of impaired annuities. Then:

$$\begin{aligned}n_1 &= 10,000 \\n_2 &= 0 \\n_3 &= 100, 200, \dots, 1000\end{aligned}$$

Results are shown in Table 3.

**Table 3.** Cases 1.2—Impact of the portfolio structure on the risk index.

Portfolio	$n_3$	$\rho(10,000, 0, n_3)$
P01	100	0.002382799
P02	200	0.002390524
P03	300	0.002398028
P04	400	0.002405318
P05	500	0.002412401
P06	600	0.002419283
P07	700	0.002425969
P08	800	0.002432465
P09	900	0.002438776
P10	1000	0.002444908

#### 6.1.3. Cases 1.3

We assume that both enhanced annuities and impaired annuities are sold (together with standard annuities), and analyze the joint impact by assuming that  $n_3 = n_2/2$ . Then:

$$\begin{aligned}n_1 &= 10,000 \\n_2 &= 500, 600, \dots, 1000 \\n_3 &= 250, 300, \dots, 500\end{aligned}$$

Results are shown in Table 4.

**Table 4.** Cases 1.3—Impact of the portfolio structure on the risk index.

Portfolio	$n_2$	$n_3$	$\rho(10,000, n_2, n_3)$
P01	500	250	0.002407197
P02	600	300	0.002412496
P03	700	350	0.002417448
P04	800	400	0.002422070
P05	900	450	0.002426375
P06	1000	500	0.002430381

#### 6.1.4. Cases 1.4

The launch of special-rate annuities might negatively impact on the sale of standard annuities (the so called “cannibalization effect”). To analyze this aspect in terms of portfolio risk profile, we assume that one half of the enhanced annuity sales (sub-portfolio SP2) are “subtracted” from the standard annuity business (sub-portfolio SP1). Then, we consider portfolios with the following sub-portfolio sizes:

$$n_1 = 10,000 - \frac{n_2}{2}$$

$$n_2 = 500, 600, \dots, 1000$$

$$n_3 = \frac{n_2}{2} = 250, 300, \dots, 500$$

Furthermore, it is reasonable to assume that, in case of a cannibalization effect, the mortality in the standard annuity sub-portfolio improves. To represent this aspect, we assume  $M_1 = 91$  (instead of  $M_1 = 90$ ). Results are shown in Table 5.

**Table 5.** Cases 1.4—Impact of the portfolio structure on the risk index.

Portfolio	$n_1$	$n_2$	$n_3$	$\rho(n_1, n_2, n_3)$
P01	9750	500	250	0.002340041
P02	9700	600	300	0.002352541
P03	9650	700	350	0.002364783
P04	9600	800	400	0.002376774
P05	9550	900	450	0.002388521
P06	9500	1000	500	0.002400030

#### 6.1.5. Some Comments

When considering a given set of cases, the size of subportfolios and structure of the total portfolio change, and this of course impacts both the numerator and denominator of the risk index. Hence, the analysis of the risk index values in the various portfolio structures provides interesting information. We note that, in all the sets of cases we have considered, the range of values assumed by the risk index is very narrow. From a mathematical perspective, this is the straight consequence of a higher variability in terms of standard deviation (the numerator of fraction (12)) offset, to a large extent, by a higher expected value (the denominator), and, in practice, a higher volume of premiums. Therefore, the almost constant value of the risk index witnesses this offset. A wider range of values (anyway very limited) can be noted as the effect of the number of impaired annuities: see, for example, the set of cases 1.2, where the increase in the risk index is equal to 2.6%, compared to the set 1.1 where the increase is smaller than 1%.

The presence of special-rate annuities impacts on the standard annuity portfolios, in particular, in terms of the lifetime distribution of standard annuitants. This has been considered in Section 6.1.4 by increasing the modal age at death from 90 to 91, then increasing the expected value of benefits paid by standard annuities. It is worth noting that a (reasonable) impact on standard annuity premiums should follow, and this might, in turn, impact the demand of standard annuities. Further interesting results, regarding the

variability of the annual payouts, can be achieved via stochastic analysis and are presented in Section 7.

6.2. Impact of Lifetime Distributions

Given the uncertainty in biometric assumptions, a sensitivity analysis is appropriate. While keeping unchanged the parameters  $M_k$  (representing the modal age at death), we propose diverse assumptions regarding the dispersion of the lifetime distributions, which might more heavily impact on the portfolio risk profile. Hence, various values of the parameters  $D_k$  are considered.

Figure 2 shows the graphs of the lifetime distribution for enhanced annuities, corresponding to different values of dispersion (parameter  $D$ ) while keeping the same modal value (parameter  $M$ ).

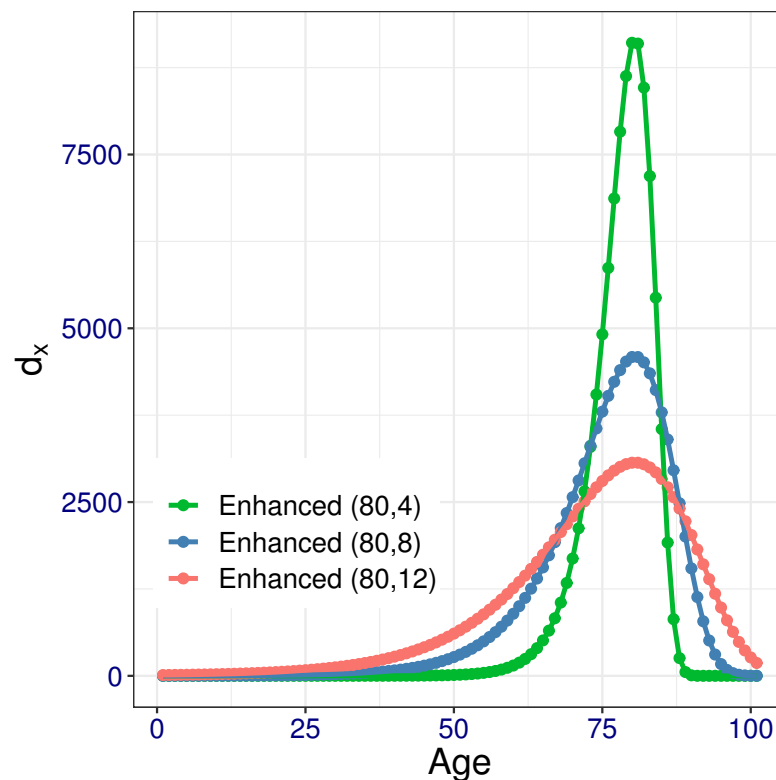


Figure 2. Three assumptions on lifetime dispersion for enhanced annuities.

6.2.1. Cases 2.1

We consider a portfolio only consisting of standard annuities and enhanced annuities, with given sub-portfolio sizes. Hence:

$$\begin{aligned} n_1 &= 10,000 \\ n_2 &= 1000 \\ n_3 &= 0 \end{aligned}$$

We analyze the impact of diverse assumptions on the dispersion of lifetimes in sub-portfolio SP2. Then:

$$D_2 = 4, 5, \dots, 13$$

(while keeping  $D_1 = 5$ ). Results are shown in Table 6.

**Table 6.** Cases 2.1—Impact of the lifetime distribution on the risk index.

Portfolio	$D_2$	$\rho(10,000, 1000, 0)$
P01	4	0.002315649
P02	5	0.002341998
P03	6	0.002364379
P04	7	0.002383918
P05	8	0.002401517
P06	9	0.002417793
P07	10	0.002433146
P08	11	0.002447834
P09	12	0.002462022
P10	13	0.002475816

6.2.2. Cases 2.2

We consider a portfolio only consisting of standard annuities and impaired annuities, with given sub-portfolio sizes. Hence:

$$\begin{aligned}
 n_1 &= 10,000 \\
 n_2 &= 0 \\
 n_3 &= 1\,000
 \end{aligned}$$

We analyze the impact of diverse assumptions on the dispersion of lifetimes in sub-portfolio SP3. Then:

$$D_3 = 11, 12, \dots, 15$$

Results are shown in Table 7.

**Table 7.** Cases 2.2—Impact of the lifetime distribution on the risk index.

Portfolio	$D_3$	$\rho(10,000, 0, 1000)$
P01	11	0.002422885
P02	12	0.002433728
P03	13	0.002444908
P04	14	0.002456358
P05	15	0.002468019

6.2.3. Cases 2.3

We consider a portfolio consisting of standard annuities, enhanced annuities and impaired annuities, with given sub-portfolio sizes. Hence:

$$\begin{aligned}
 n_1 &= 10,000 \\
 n_2 &= 1000 \\
 n_3 &= 500
 \end{aligned}$$

We analyze the joint impact of diverse assumptions on the dispersion of lifetimes in both sub-portfolios SP2 and SP3. To this purpose, we assume:

$$D_2 = D_3 = 4, 5, \dots, 13$$

We note that lower dispersions can be achieved by restricting the range of pathologies, which entitle the purchase of enhanced annuities and impaired annuities. Results are shown in Table 8.

**Table 8.** Cases 2.3—Impact of the lifetime distributions on the risk index.

Portfolio	$D_2 = D_3$	$\rho(10,000, 1000, 500)$
P01	4	0.002360437
P02	5	0.002366549
P03	6	0.002373800
P04	7	0.002382362
P05	8	0.002392205
P06	9	0.002403223
P07	10	0.002415280
P08	11	0.002428234
P09	12	0.002441949
P10	13	0.002456293

#### 6.2.4. Some Comments

Although dispersion in lifetime distributions does affect the risk profile of the annuity portfolio, the sensitivity analysis we have performed witnesses a rather limited impact on the risk index. We note that, of course, the broadest range of risk index values can be found when a portfolio consisting of standard annuities, enhanced annuities and impaired annuities is addressed, and for both the types of special-rate annuities higher values for the dispersion parameter are considered.

### 7. Portfolio Risk Profiles: Stochastic Approach

Deterministic assessments performed in Section 6 only provide values of specific markers, notably the risk index. To obtain better insights into the risk profile of a portfolio, stochastic assessments are required. To this purpose, stochastic (Monte Carlo) simulation procedures are commonly adopted.

As we focus on the biometric features of the various portfolios, simulation of the numbers of survivors, that is  $N_k(t)$ ,  $k = 1, 2, 3$  and  $t = 1, 2, \dots$ , is only needed. Then, via Equations (13) and (14), the simulated outcomes of the payouts and, finally, the relevant (empirical) distributions are obtained. Consistent with the approach adopted in Section 6, we assume that all the assessments are performed at time  $t = 0$ , and hence our information is given by the initial sizes of the sub-portfolios, that is,  $n_1$ ,  $n_2$  and  $n_3$ .

The following Sections 7.1 and 7.2 are organized similarly to Sections 6.1 and 6.2, respectively, but with a reduction in the number of cases analyzed.

Besides “descriptive” results in terms of (empirical) distributions of the annual payouts, the stochastic approach can also yield “operational” results: an example is provided in Section 7.3, where amounts of assets are calculated, which are needed to meet the annual payouts with an assigned probability.

#### 7.1. Impact of the Portfolio Structure

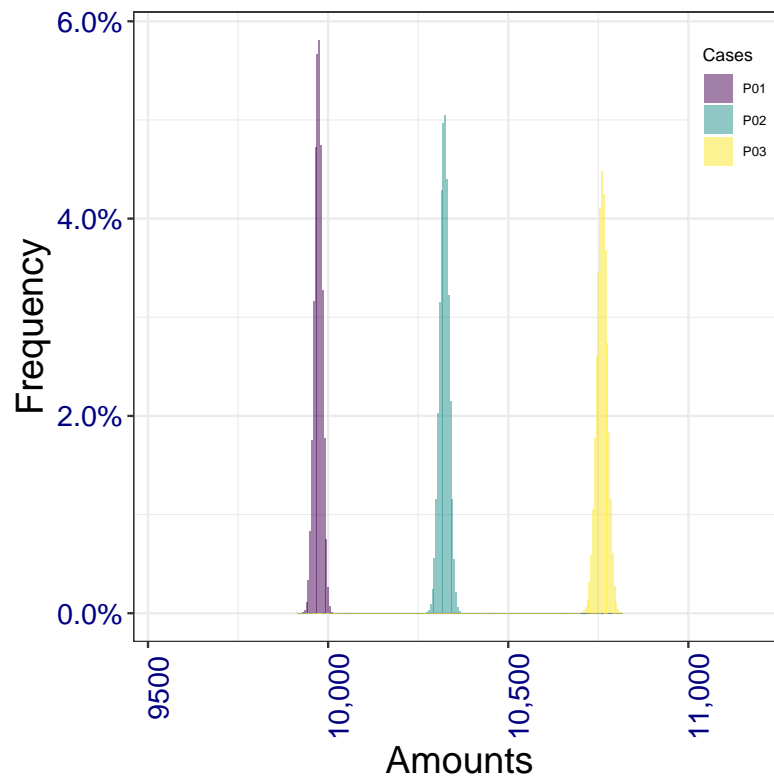
As already noted, we follow the organization in the cases adopted in Section 6.1, although reducing the number of alternatives.

##### 7.1.1. Cases 1.1

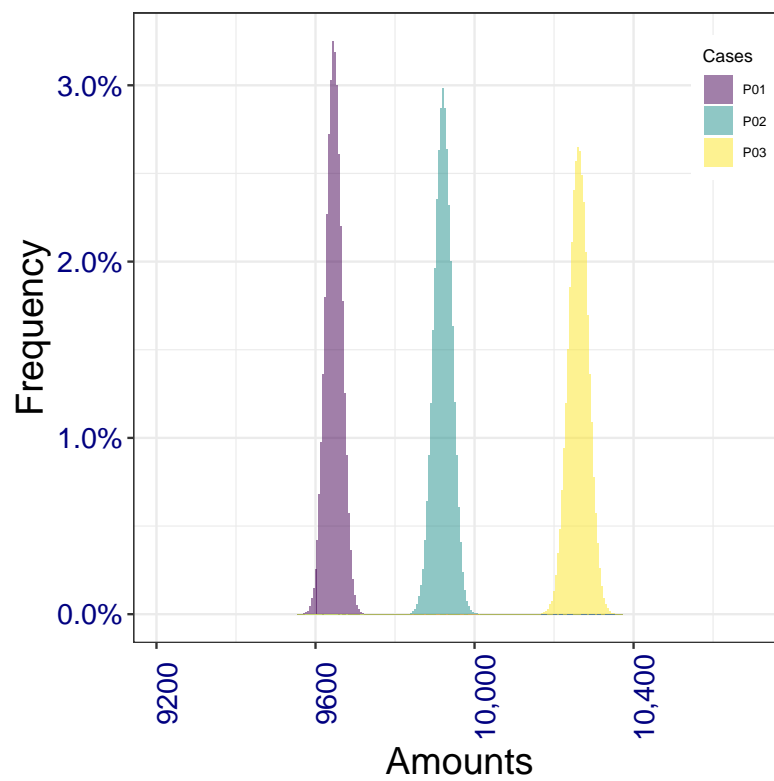
We analyze the impact of the size of the sub-portfolio SP2 of enhanced annuities. Then:

$$\begin{aligned}n_1 &= 10,000 \\n_2 &= 100, 500, 1000 \\n_3 &= 0\end{aligned}$$

Empirical distributions at times 5 and 10 of the portfolio payout, that is, empirical distributions of  $X(5)$  and  $X(10)$ , are sketched in Figures 3 and 4, where the three portfolios are denoted by P01, P02 and P03, respectively.



**Figure 3.** Impact of the number of enhanced annuities: empirical distributions of the annual benefit payout at time  $t = 5$ .



**Figure 4.** Impact of the number of enhanced annuities: empirical distributions of the annual benefit payout at time  $t = 10$ .

## 7.1.2. Cases 1.4

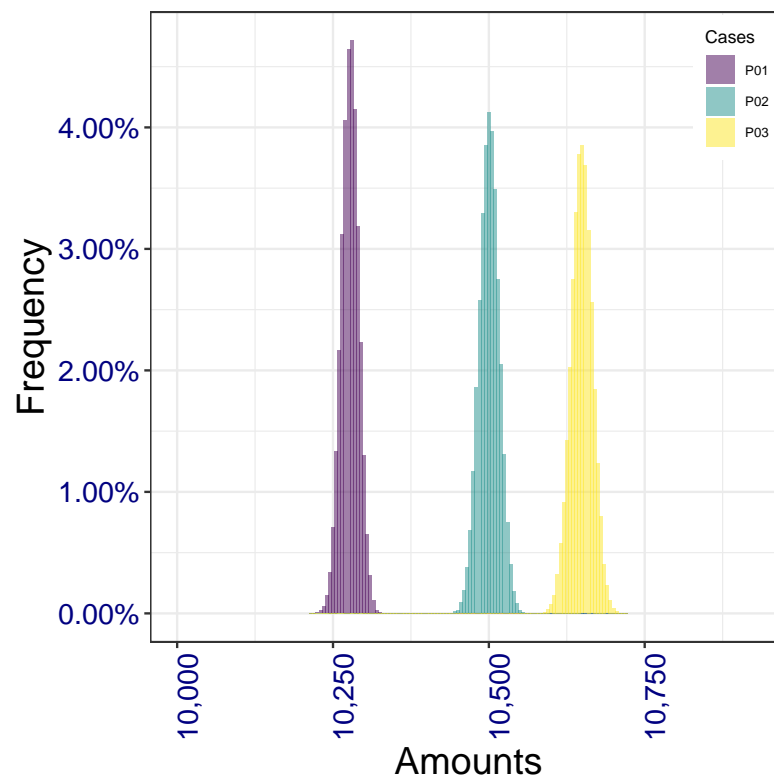
To assess the impact of a possible cannibalization effect, we consider three portfolios with the following sub-portfolio sizes:

$$n_1 = 10,000 - \frac{n_2}{2}$$

$$n_2 = 500, 800, 1000$$

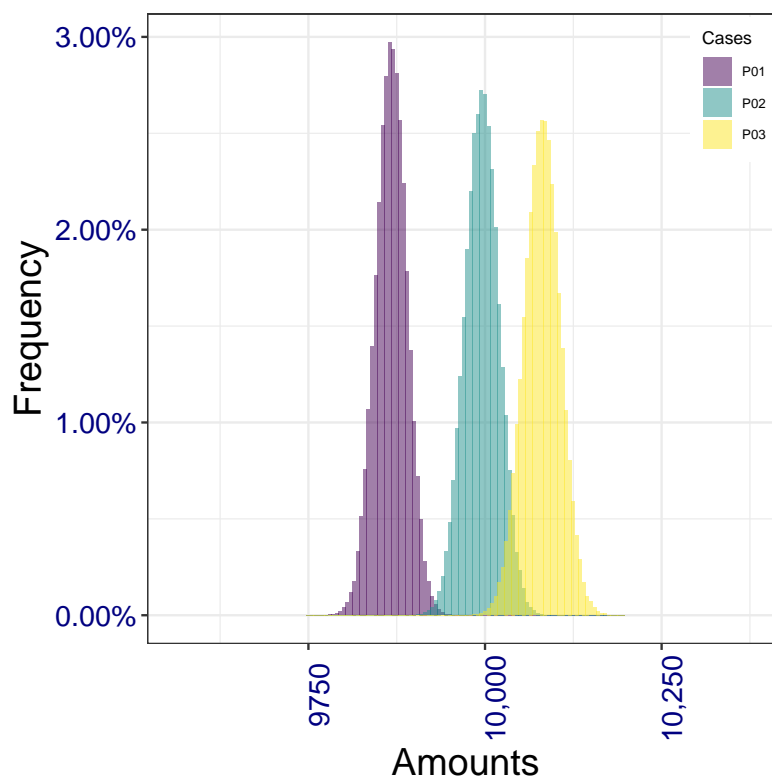
$$n_3 = \frac{n_2}{2} = 250, 400, 500$$

As previously noted, to represent an improvement in mortality in the standard annuity sub-portfolio, we assume  $M_1 = 91$  (instead of  $M_1 = 90$ ). Empirical distributions of the portfolio payout  $X(5)$  and  $X(10)$  are sketched in Figures 5 and 6, respectively, where the three portfolios are denoted by P01, P02 and P03.



**Figure 5.** Impact of cannibalization effect: empirical distributions of the annual benefit payout at time  $t = 5$ .





**Figure 6.** Impact of cannibalization effect: empirical distributions of the annual benefit payout at time  $t = 10$ .

7.1.3. Some Comments

Results are self-explanatory, and in line with the findings in the deterministic setting: the larger the (initial) number of enhanced annuities in Cases 1.1, the higher the dispersion in the annual payouts. The same effect is, of course, witnessed by the distributions of payouts in Cases 1.4, where presence of both enhanced annuities and impaired annuities is assumed.

7.2. Impact of the Lifetime Distribution

To assess the impact of uncertainty in biometric assumptions, we only analyze the Cases 2.1 considered in Section 6.2.

7.2.1. Cases 2.1

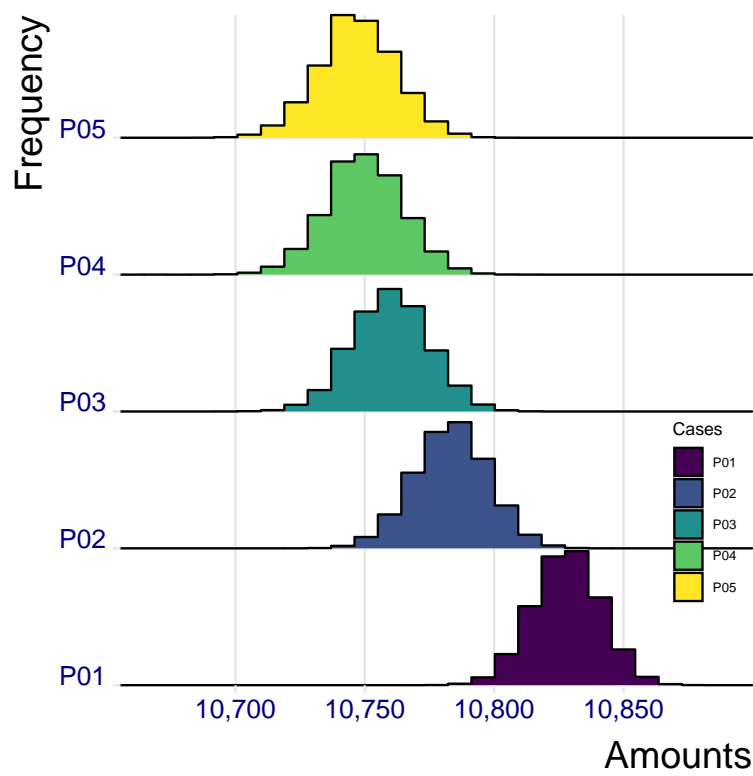
We consider a portfolio only consisting of standard annuities and enhanced annuities, with given sub-portfolio sizes. Hence:

$$\begin{aligned} n_1 &= 10,000 \\ n_2 &= 1000 \\ n_3 &= 0 \end{aligned}$$

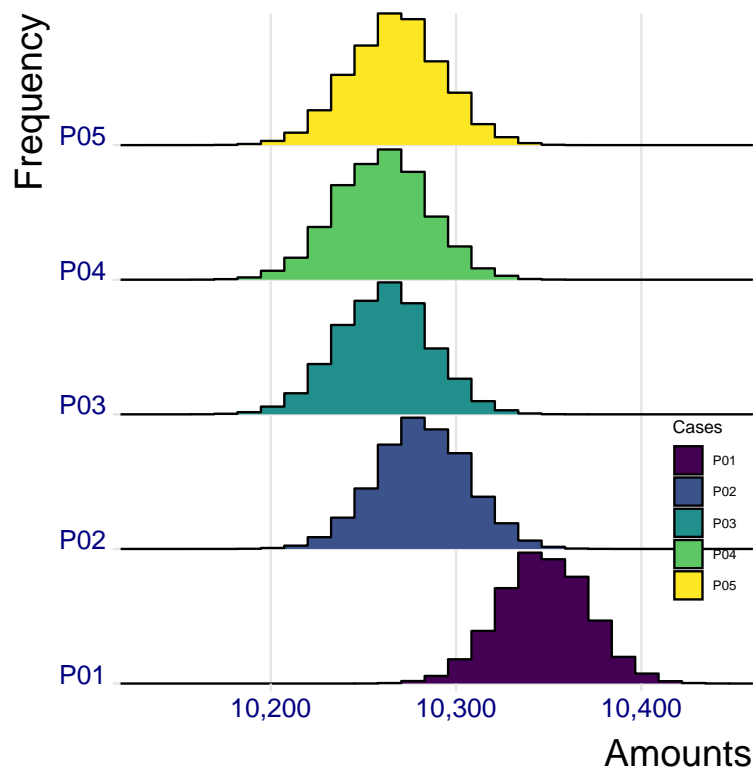
We analyze the impact of diverse assumptions on the dispersion of lifetimes. Then:

$$D_2 = 4, 6, 8, 10, 12$$

(while keeping  $D_1 = 5$ ). Empirical distributions of the portfolio payout  $X(5)$  and  $X(10)$  are sketched in Figures 7 and 8, where the portfolios are respectively denoted by  $P01, \dots, P05$ .



**Figure 7.** Impact of different dispersion parameters: empirical distributions of the total payout at time  $t = 5$ .



**Figure 8.** Impact of different dispersion parameters: empirical distributions of the total payout at time  $t = 10$ .

### 7.2.2. Some Comments

These results are self-evident: a larger dispersion in the lifetime distribution of annuitants with enhanced annuity implies a larger dispersion in the total portfolio payout (compare, in particular, distributions in portfolio P05 and P01. Moreover, for all the portfolios, dispersions increase with time (compare the distributions in Figure 7 to the ones in Figure 8).

### 7.3. Meeting the Annual Payouts

Appropriate resources must be assigned to the portfolio in order to meet the annual payouts with a high probability. Diverse criteria can be adopted to quantify the above resources which, whatever the criterion adopted, will partly be provided by single premiums cashed at policies issued (via decumulation of the portfolio reserve) and partly by shareholders’ capital allocated to the portfolio. In what follows, we focus on the annual total amount of resources needed, disregarding the funding source.

#### 7.3.1. The Percentile Principle

Referring to a generic portfolio and the relevant cash flows, we recall that  $X_1(t)$ ,  $X_2(t)$ ,  $X_3(t)$  denote the random payouts at time  $t$ , related to standard annuities, enhanced annuities and impaired annuities, respectively (see Section 5.4), and:

$$X(t) = X_1(t) + X_2(t) + X_3(t) \tag{15}$$

denotes the portfolio total payout at time  $t$ .

We adopt the percentile principle. Hence, we have to find, for  $t = 1, 2, \dots$ , the amount  $A(t)$  such that:

$$\Pr[X(t) > A(t)] = \varepsilon \tag{16}$$

where  $\varepsilon$  denotes an assigned (small) probability; hence,  $1 - \varepsilon$  can be interpreted as the “adequacy” level.

A more detailed analysis could be performed by separately addressing the risk profile of each sub-portfolio, thus calculating, for  $k = 1, 2, 3$  and  $t = 1, 2, \dots$  the quantities  $A_k(t)$  such that:

$$\Pr[X_k(t) > A_k(t)] = \varepsilon_k \tag{17}$$

However, we only focus on the overall requirement (16), which clearly takes into account the pooling effect.

#### 7.3.2. Numerical Results

We consider four portfolios with the structures defined in Table 9.

**Table 9.** Portfolio structures.

Portfolio	$n_1$	$n_2$	$n_3$
P01	10,000	0	0
P02	10,000	1000	0
P03	10,000	0	500
P04	10,000	1000	500

This way, we can analyze the risk profile of a “traditional” portfolio only consisting of standard annuities (P01), a portfolio including standard annuities and enhanced annuities (P02) or standard annuities and impaired annuities (P03) and finally a portfolio including both types of special-rate annuities (P04).

Values of the parameters  $M_k$  and  $D_k$  for standard life annuities ( $k = 1$ ), enhanced annuities ( $k = 2$ ) and impaired annuities ( $k = 3$ ), respectively, are as specified in Table 1.

Asset requirements (at a given time  $t$ ), in terms of ratios  $\frac{\text{Assets}}{EV}$ , where  $EV$  denotes the expected value of the total payout, are plotted in Figures 9–11 against the probability  $1 - \varepsilon$ .

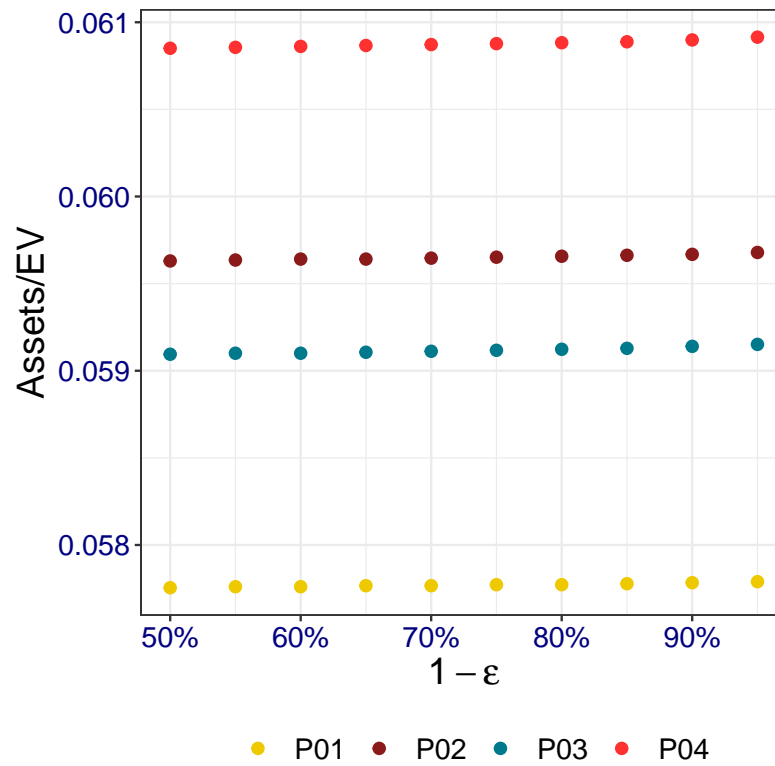


Figure 9. Impact of the portfolio structure on assets requirements at time  $t = 1$ .

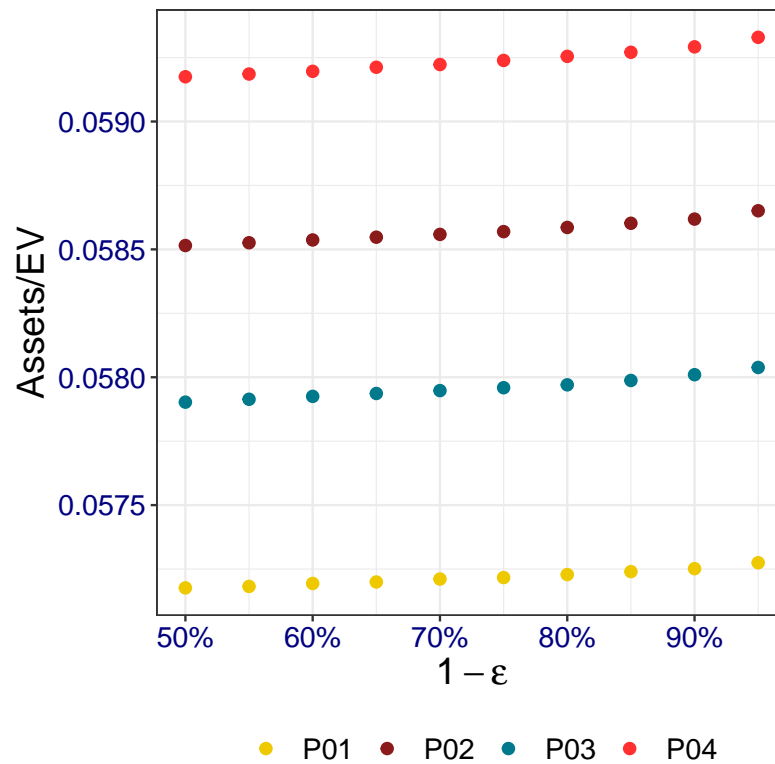
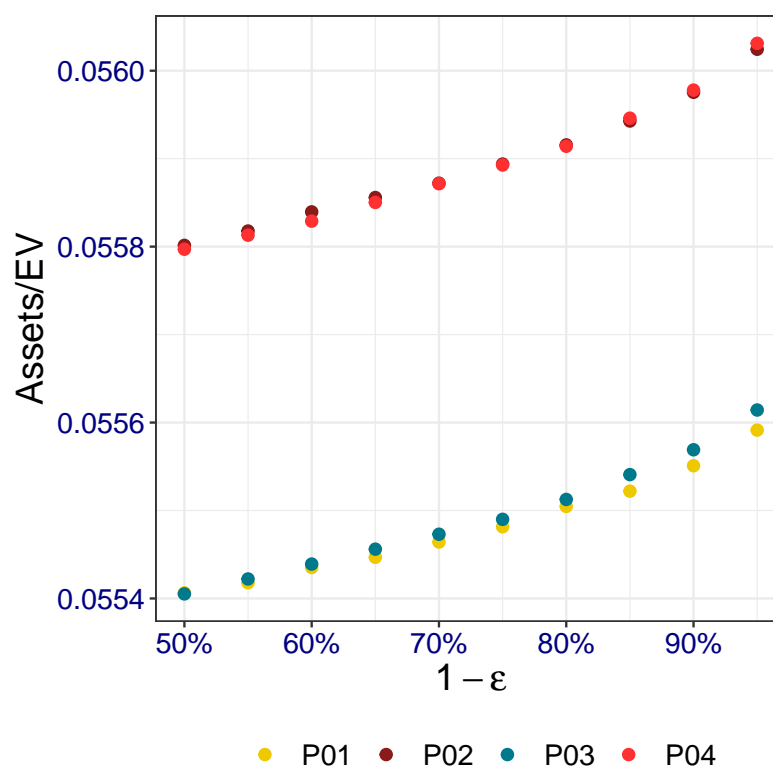


Figure 10. Impact of the portfolio structure on assets requirements at time  $t = 5$ .



**Figure 11.** Impact of the portfolio structure on assets requirements at time  $t = 10$ .

### 7.3.3. Some Comments

Results in terms of assets requirements are also encouraging. We note that the range of values, expressed by the ratio between assets required and expected values of total payout, corresponding to the various portfolio structures are very limited, whatever the adequacy level chosen. A higher sensitivity with respect to the adequacy level  $1 - \epsilon$  can be observed in particular for  $t = 10$ , because of a dispersion of the numbers of survivors and hence of the total payouts, which increases with time.

We note that the portfolio structure of course evolves over time and, notably, the share of standard annuities progressively increases because of longer lifetimes of standard annuitants. Increasing shares of standard life annuities (characterized, according to a reasonable assumption, by a variance lower than that of special-rate annuities) imply a lower riskiness in the (total) annuity portfolio. In terms of the ratio Assets/EV, this effect appears in particular when comparing the requirements (according to each probability  $1 - \epsilon$ ) represented in Figures 9–11.

## 8. Concluding Remarks

By offering special-rate life annuities, on the one hand, a higher premium income can be expected, while on the other hand, a higher variability of the total portfolio payout will follow because of both the larger size and the specific higher variability of payouts related to special-rate annuities.

The analysis, in quantitative terms, of the “balance” between the two aspects (that is, higher risk and higher premium income) has been the aim of this research. A number of numerical evaluations have been performed by adopting both a deterministic approach and a stochastic one as well. Diverse hypotheses on lifetime distributions have been assumed, and various portfolio sizes and structures (in terms of numbers of standard, enhanced and impaired annuities) have been considered.

It is worth noting that, whatever the choice of the parameter values for the Gompertz law, the mortality model is deterministic, in the sense that no uncertainty in future mortality trend is accounted for. In this context, it is reasonable to assume independence among

the individual lifetimes. An interesting extension of our simple model should allow for uncertainty in mortality trends via an appropriate stochastic mortality model. Then, correlation among lifetimes would follow, so that independence assumption would be replaced by conditional independence. Moreover, diverse trends and diverse degrees of uncertainty could be considered for the various special-rate annuities by taking into account possible improvements in medical treatments, surgery, etc.

Special attention should be placed on the “cannibalization” effect. First, a reduction in the size of the standard annuity sub-portfolio may occur, provided that some applicants can be eligible to special-rate annuities. A lower average mortality in the standard sub-portfolio then follows (as noted in Section 6.1.5), leading to a (reasonable) increase in standard annuity premiums. This might in turn impact the demand of standard life annuities (even beyond the reduction mentioned above).

A further aspect that is interesting to investigate is the impact on the insurer’s liabilities (and, notably, on the portfolio risk profile) of an incorrect allocation of individual risks to the various rating classes. Because of the presence of unobservable risk factors, a misspecification of the rating class is always possible. This would result in an unfair annuity rate applied to some individuals and then in an increased (or reduced) probability of loss for the insurer. In this regard, the research task should concern, in particular, the modeling of the incorrect specification of the rating class.

The results we have obtained of course depend on assumptions (notably, regarding both the portfolio structure, the mortality model and the relevant parameters). Nevertheless, the broad range of assumptions regarding both the portfolio structure and the lifetime distributions has allowed us to perform an effective sensitivity analysis, whose interesting achievements witness the possibility of extending the life annuity business without taking huge amounts of risk. Hence, the creation of values for customers (and an increase in the insurer’s market share) can be pursued without a significant worsening of the company’s risk profile.

**Author Contributions:** Conceptualization, E.P. and D.Y.T.; methodology, E.P. and D.Y.T.; software, E.P. and D.Y.T.; validation, E.P. and D.Y.T.; formal analysis, E.P. and D.Y.T.; investigation, E.P. and D.Y.T.; writing—original draft preparation, E.P. and D.Y.T.; writing—review and editing, E.P. and D.Y.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors thank the anonymous referees for their helpful comments and suggestions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Note

<sup>1</sup> Stage-Gate® is a registered trademark of Stage-Gate Inc. ([www.stage-gate.com](http://www.stage-gate.com) (accessed on 10 January 2022)).

## References

- Ainslie, Ross. 2000. *Annuity and Insurance Products for Impaired Lives*. Working Paper. London: Staple Inn Actuarial Society.
- Beard, Robert E. 1959. Note on some mathematical mortality models. In *CIBA Foundation Colloquia on Ageing*. Edited by Gordon Ethelbert Ward Wolstenholme and Maeve O. Connor. Boston: John Wiley & Sons, Ltd., vol. 5, pp. 302–11.
- Carriere, Jacques F. 1992. Parametric models for life tables. *Transaction of Society of Actuaries* 44: 77–99.
- Denuit, Michel, and Esther Frostig. 2006. Heterogeneity and the need for capital in the individual model. *Scandinavian Actuarial Journal* 2006: 42–66. [CrossRef]
- Denuit, Michel, and Esther Frostig. 2007. Association and heterogeneity of insured lifetimes in the Lee–Carter framework. *Scandinavian Actuarial Journal* 2007: 1–19. [CrossRef]
- Drinkwater, Matthew, Joseph E. Montminy, Eric T. Sondergeld, Christopher G. Raham, and Chad R. Runchey. 2006. Substandard Annuities. Technical Report. LIMRA International Inc. and the Society of Actuaries, in Collaboration with Ernst & Young LLP. Available online: <https://www.soa.org/Files/Research/007289-Substandard-annuities-full-rpt-REV-8-21.pdf> (accessed on 10 January 2022).

- Edwards, Matthew. 2008. The Last Post. *The Actuary* 2008: 30–31. Available online: <http://www.theactuary.com/archive/2008/09/> (accessed on 10 January 2022).
- Gatzert, Nadine, and Udo Klotzki. 2016. Enhanced annuities: Drivers of and barriers to supply and demand. *The Geneva Papers on Risk and Insurance Issues and Practice* 41: 53–77. [CrossRef]
- Gatzert, Nadine, Gudrun Schmitt-Hoermann, and Hato Schmeiser. 2012. Optimal risk classification with an application to substandard annuities. *North American Actuarial Journal* 16: 462–86. [CrossRef]
- Gracie, Stewart, and Stephen Makin. 2006. The Price to Pay for Enhanced Annuities. Healthcare Conference 2006. Available online: <https://www.actuaries.org.uk/system/files/documents/pdf/Gracie.pdf> (accessed on 10 January 2022).
- Haberman, Steven, and Annamaria Olivieri. 2014. Risk Classification/Life. In *Wiley StatsRef: Statistics Reference Online*. Hoboken: Wiley.
- Hoermann, Gudrun, and Jochen Russ. 2008. Enhanced annuities and the impact of individual underwriting on an insurer's profit situation. *Insurance: Mathematics & Economics* 43: 150–57.
- James, Mick. 2016. Enhanced annuities: Caring for at-retirement needs. *Reinsurance News* 84: 24–27.
- Olivieri, Annamaria, and Ermanno Pitacco. 2016. Frailty and risk classification for life annuity portfolios. *Risks* 4: 39. Available online: <http://www.mdpi.com/2227-9091/4/4/39> (accessed on 10 January 2022). [CrossRef]
- Perks, Wilfred. 1932. On some experiments in the graduation of mortality statistics. *Journal of the Institute of Actuaries* 63: 12–57. [CrossRef]
- Pitacco, Ermanno. 2014. *Health Insurance. Basic Actuarial Models*. EAA Series. Berlin: Springer.
- Pitacco, Ermanno. 2019. Heterogeneity in mortality: A survey with an actuarial focus. *European Actuarial Journal* 9: 3–30. [CrossRef]
- Pitacco, Ermanno. 2020. *ERM and QRM in Life Insurance. An Actuarial Primer*. Springer Actuarial. Berlin: Springer.
- Pitacco, Ermanno. 2021. *Life Annuities. From Basic Actuarial Models to Innovative Designs*. London: Risk Books.
- Pitacco, Ermanno, and Daniela Y. Tabakova. 2020. Innovation in Life Insurance Products: Special-Rate Annuities. MIB Trieste School of Management, DemoLab Paper 2 /2020. Available online: <https://mib.edu/en/about-mib/demographic-laboratory/research/special-rate-annuities> (accessed on 10 January 2022).
- Ridsdale, Brian. 2012. Annuity Underwriting in the United Kingdom. Note for the International Actuarial Association Mortality Working Group. Available online: [https://www.actuaries.org/CTTEES\\_TFM/Documents/IAA20MWG20Annuity20underwriting20in20the20United20Kingdom20BPR120May202012.pdf](https://www.actuaries.org/CTTEES_TFM/Documents/IAA20MWG20Annuity20underwriting20in20the20United20Kingdom20BPR120May202012.pdf) (accessed on 10 January 2022).
- Rinke, Cord Roland. 2002. The Variability of Life Reflected in Annuity Products. In *Hannover Re's Perspectives—Current Topics of International Life Insurance*. Issue No. 8. London: ERM.
- Vaupel, James W., Kenneth G. Manton, and Eric Stallard. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16: 439–54. [CrossRef] [PubMed]
- Weinert, Ted. 2006. Enhanced annuities on the move. In *Hannover Re's Perspectives—Current Topics of International Life Insurance*. Issue No. 13. London: ERM.

# The Copula Derived from the SAHARA Utility Function

Jaap Spreeuw 

Faculty of Actuarial Science and Insurance, Bayes Business School (Formerly Cass), University of London, 106 Bunhill Row, London EC1Y 8TZ, UK; j.spreeuw@city.ac.uk

**Abstract:** A new Archimedean copula family is presented that was derived from the SAHARA utility function introduced in the economic literature in 2011. Its properties are discussed, and its flexibility and versatility are demonstrated. It is left tail decreasing or right tail increasing, but unlike mainstream Archimedean families, not necessarily stochastically increasing at the same time. It is shown that the family fits very well to a dataset of previously studied coupled lives in the literature.

**Keywords:** copula; Archimedean generator; dependence; coupled lives

## 1. Introduction

Archimedean copulas, which in two dimensions are of the form  $C_\psi(u_1, u_2) = \psi(\psi^{-1}(u_1) + \psi^{-1}(u_2))$  where  $\psi$  is the one-dimensional generator, have become a popular mode of modelling dependence in both finance and insurance. Several ways of constructing copula families are given in Chapter 3 of Joe (2015). The interpretation of the generator as the Williamson transform of a radial random variable has given rise to new Archimedean families; see McNeil and Nešlehová (2009, 2010). Archimedean copulas are a flexible class due to the ease with which new Archimedean copulas with an enriched parameter space can be constructed from existing ones using transformations. For the bivariate case, five of such transformations, namely, left composition, right composition, scaling, exponentiation, and the linear combination of the (inverse) generator, were introduced in the literature by Genest et al. (1998). They were reviewed by Michiels and De Schepper (2012), and in more detail in Michiels and De Schepper (2009), with the focus on the so-called  $\lambda$  function (which is the ratio of the inverse generator to its derivative). For the latter, see also Michiels et al. (2011).

In the literature, several generalized families were constructed that contain the Archimedean class as a special case, e.g., the Archimax family in Capéraà et al. (2000) (which includes extreme value copulas as another special case). The background risk model where one random variable (“systemic risk”) acts multiplicatively on a series of other random variables (“idiosyncratic risks”) is the basis of generalization of Archimedean copulas in several ways, as demonstrated in Côté and Genest (2019) and Marri and Moutanabbir (2022).

Commonly used families typically feature a generator being completely monotonic and thereby the Laplace transform of a mixing random variable. Common examples include Clayton and Frank (as far as dependence is positive), Gumbel–Hougaard, and Joe. This subfamily of Archimedean copulas, also known as shared frailty models, has the advantage of being valid in any dimension. Recent applications regarding shared frailty models involve the aforementioned background risk model; see (Albrecher et al. 2011; Furman et al. 2021; Sarabia et al. 2018).

The Gumbel–Hougaard copula is a good fit to the well-known dataset of loss vs. Allocated Loss Adjustment Expenses (ALAE), which is the object of statistical inference in several publications, starting with Genest et al. (1998). For extensive analysis, consult Joe (2015). Probably there are several other case studies of dependence in insurance where the use of shared frailty models is appropriate.

**Citation:** Spreeuw, Jaap. 2022. The Copula Derived from the SAHARA Utility Function. *Risks* 10: 133. <https://doi.org/10.3390/risks10070133>

Academic Editors: Ermanno Pitacco and Annamaria Olivieri

Received: 28 February 2022

Accepted: 10 June 2022

Published: 28 June 2022

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).



Shared frailty dependency models have the property of being conditionally increasing (CI) and multivariate totally positive of order 2 (MTP<sub>2</sub>), as shown in Müller and Scarsini (2005). In the specific case of two dimensions, this is known as TP<sub>2</sub>, implying stochastically increasing (SI), which in turn implies both left tail decreasing (LTD) and right tail increasing (RTI). If the two involved marginal random variables are remaining lifetimes, LTD/RTI implies that the hazard rate of the one upon hazard (e.g., death or default) of the other goes up. The stronger condition SI implies that the hazard rate of the one given the hazard of the other (e.g., death or default) at time  $t$  decreases as  $t$  goes up. This was noted in Spreeuw (2006). The underlying assumption of SI may work out well in reliability theory. Consider two printers available to office staff. When one fails, the other one, pending the repair of the first, is used more intensively and thereby exposed to greater strain. It is sensible to assume that the longer the first printer is out of order, the higher the failure rate of the remaining machine would be.

For coupled lives, however, it is not so clear-cut. On the one hand, the event of the death of one life usually triggers an elevated mortality of the surviving life, so the assumption of LTD/RTI seems sensible. In addition, such lives are exposed to common risks due to permanently living together and due to a similar background (“birds of a feather flock together”, the so called long-term dependence according to Hougaard (2000)). On the other hand, however, there is also the phenomenon of the event of the one life dying leading to the mortality of the remaining life temporarily going up, which is the so-called broken-heart syndrome (short-term dependence, also attributed to Hougaard (2000)). Similar nonstandard features may apply to other cases in insurance (and finance as well). In short, there is a case for constructing copula families allowing for flexibility in terms of type of dependence, such as LTD/RTI but not necessarily SI.

In this paper, we introduce a new Archimedean copula family that is based on a link between Archimedean generators and utility functions; see Spreeuw (2010) for more details. Unlike mainstream copulas, this family has the property of being LTD/RTI, but not necessarily SI, the latter being clearly indicated by the sign of one of the parameters.

The outline of the paper is as follows. Section 2 gives the basic definitions of Archimedean copula, and the dependence concepts of LTD/RTI and SI. Section 3 introduces the new family and analyzes its basic properties. Section 4 fits the new Archimedean family to the section of censored remaining lifetime data of coupled lives, as in Luciano et al. (2008). Section 5 sets out a conclusion.

## 2. Basic Definitions

Define  $\psi$  as the generator of a 2-dimensional Archimedean copula, being strictly continuous, strictly decreasing, convex, with  $\psi(0) = 1$  and  $\lim_{x \rightarrow \infty} \psi(x) = 0$ . The copula itself is then specified as

$$C_{\psi}(u_1, u_2) = \psi\left(\psi^{-1}(u_1) + \psi^{-1}(u_2)\right), \quad (1)$$

where  $u_1, u_2$  each take values between 0 and 1.

Next are definitions of the tail concepts of left tail decreasing (LTD), right tail increasing (RTI) and stochastically increasing (SI), based on two random variables  $X$  and  $Y$ , and their copula  $C$ . They can all be found in Chapter 5 of Nelsen (2006).

**Definition 1.**  $Y$  is LTD in  $X$  (notation  $LTD(Y|X)$ )  $\Leftrightarrow \Pr[Y \leq y|X \leq x]$  is nonincreasing in  $x$  for all  $y$ . For an exchangeable copula  $C$  (i.e.,  $C(u, v) = C(v, u)$  for  $0 \leq u, v \leq 1$ ) of random variables  $X$  and  $Y$ ,  $LTD(Y|X)$  and  $LTD(X|Y)$  are equivalent.

**Definition 2.**  $Y$  is RTI in  $X$  (notation  $RTI(Y|X)$ )  $\Leftrightarrow \Pr[Y > y|X > x]$  is nondecreasing in  $x$  for all  $y$ . For an exchangeable copula  $C$  (i.e.,  $C(u, v) = C(v, u)$  for  $0 \leq u, v \leq 1$ ) of random variables  $X$  and  $Y$ ,  $RTI(Y|X)$  and  $RTI(X|Y)$  are equivalent.

**Definition 3.**  $Y$  is SI in  $X$  (notation  $SI(Y|X)$ )  $\Leftrightarrow \Pr[Y \leq y|X = x]$  is nonincreasing in  $x$  for all  $y$ .

The following propositions are from Avérous and Dortet-Bernadet (2004). The second one was originally shown in Capéraà and Genest (1993).

**Proposition 1.** If  $C$  is Archimedean with generator  $\psi$ ,  $LTD(Y|X)$  or  $LTD(X|Y)$  if and only if  $\psi$  is logconvex. Likewise, if  $\hat{C}$ , the rotated copula (also known as survival copula) of  $C$  (so  $\hat{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$  for  $0 \leq u, v \leq 1$ ) is Archimedean with generator  $\hat{\psi}$  is  $RTI(Y|X)$  or  $RTI(X|Y)$  if and only if  $\hat{\psi}$  is logconvex.

**Proposition 2.** If  $C$  is Archimedean with differentiable generator  $\psi$ ,  $SI(Y|X)$  or  $SI(X|Y)$  if and only if  $-\psi'$  is logconvex.

### 3. SAHARA Family

The SAHARA copula family is derived from the Symmetric Asymptotic Hyperbolic Absolute Risk Aversion (SAHARA) utility function introduced in Chen et al. (2011). This utility function is specified below.

$$\varphi_{\theta,\epsilon}(s) = \begin{cases} -\frac{1}{(1+1/\theta)^2-1} \left( s - \epsilon + \sqrt{\delta^2 + (s - \epsilon)^2} \right)^{-(1+1/\theta)} & \theta \neq 0 \\ \cdot \left( s - \epsilon + (1 + 1/\theta)\sqrt{\delta^2 + (s - \epsilon)^2} \right) & \\ \frac{1}{2} \ln \left( s - \epsilon + \sqrt{\delta^2 + (s - \epsilon)^2} \right) & \\ + \frac{1}{2} \delta^{-2} (s - \epsilon) \left( \sqrt{\delta^2 + (s - \epsilon)^2} - (s - \epsilon) \right) & \theta = 0. \end{cases}, \quad \theta \in (-\infty, -1) \cup (0, \infty), \delta > 0, \epsilon \in \mathbb{R}. \quad (2)$$

As shown in Spreeuw (2010), a strict Archimedean generator can be obtained from a utility function  $\varphi$  if  $\varphi(\infty) = \lim_{s \rightarrow \infty} \varphi(s) < \infty$ . For SAHARA, this is the case when  $\theta > 0$ . Then, applying the formula  $\psi_{\theta,\epsilon}(s) = \{\varphi_{\theta,\epsilon}(\infty) - \varphi_{\theta,\epsilon}(s)\} \{\varphi_{\theta,\epsilon}(\infty) - \varphi_{\theta,\epsilon}(0)\}^{-1}$  leads to the Archimedean generator

$$\psi_{\theta,\epsilon}(s) = \left( \frac{s - \epsilon + \sqrt{\delta^2 + (s - \epsilon)^2}}{-\epsilon + \sqrt{\delta^2 + \epsilon^2}} \right)^{-(1+1/\theta)} \left( \frac{s - \epsilon + (1 + 1/\theta)\sqrt{\delta^2 + (s - \epsilon)^2}}{-\epsilon + (1 + 1/\theta)\sqrt{\delta^2 + \epsilon^2}} \right). \quad (3)$$

**Remark 1.** This approach of obtaining an Archimedean generator from a utility function is not to be confused with the method of obtaining the inverse of an Archimedean generator from a utility function. For the latter, consult Spreeuw (2014).

The SAHARA utility function was inspired by nonmonotone risk aversion coefficient

$$AR_{\varphi}(s) = \frac{1 + 1/\theta}{\sqrt{(s - \epsilon)^2 + \delta^2}},$$

which, unlike common utility functions, is not monotone in its argument. It is rather increasing for  $s < \epsilon$  attaining a maximum for  $s = \epsilon$ , and decreasing for  $s > \epsilon$ . The SAHARA utility function found applications in both finance and insurance (Bernard et al. 2021; Bernard and Kwak 2016; Brachetta and Schmidli 2020; Chen et al. 2021; Chen and Vellekoop 2017; Li and Ma 2018; Schumacher 2018). As shown in Spreeuw (2010), a risk aversion monotone decreasing (increasing) on the positive real line in general implies that the corresponding Archimedean copula is stochastic increasing (stochastic decreasing) in two dimensions. The former property applies to the vast majority of commonly applied

copula families (including all those whose generator is a completely monotonic function). For  $\epsilon > 0$  the copula is neither stochastic increasing nor stochastic decreasing. To the best of our knowledge, there are hardly any copula families that share this property.

For  $\epsilon < 0$ , the condition imposed on  $\delta$  can be somewhat relaxed to  $\delta \geq 0$ ,  $\delta = 0$  corresponding to the Clayton copula with parameter  $\theta$ . Due to scaling, for some real-valued  $\delta^* > 0$  and  $\epsilon^*$ ,  $(\delta, \epsilon) = (\delta^*, \epsilon^*)$  gives exactly the same Archimedean copula as  $(\delta, \epsilon) = (1, \epsilon^*/\delta^*)$ . To see this, we divide in (3) both numerator and denominator by  $\delta > 0$ . This gives:

$$\psi_{\theta, \epsilon}(s) = \left( \frac{\frac{s}{\delta} - \frac{\epsilon}{\delta} + \sqrt{1 + \left(\frac{s}{\delta} - \frac{\epsilon}{\delta}\right)^2}}{-\frac{\epsilon}{\delta} + \sqrt{1 + \left(\frac{\epsilon}{\delta}\right)^2}} \right)^{-(1+1/\theta)} \left( \frac{\frac{s}{\delta} - \frac{\epsilon}{\delta} + (1 + 1/\theta)\sqrt{1 + \left(\frac{s}{\delta} - \frac{\epsilon}{\delta}\right)^2}}{-\frac{\epsilon}{\delta} + (1 + 1/\theta)\sqrt{1 + \left(\frac{\epsilon}{\delta}\right)^2}} \right).$$

Now,  $\epsilon$  being real-valued implies that  $\epsilon/\delta$  is real-valued as well. In addition, it is well-known that a generator is only defined up to a multiple constant. In other words, for  $\beta > 0$ ,  $\psi_{\theta, \epsilon}(s)$  and  $\psi_{\theta, \epsilon}(\beta s)$  generate the same Archimedean copula.  $\delta^* \rightarrow^+ 0$  is equivalent to  $|\epsilon^*/\delta^*| \rightarrow \infty$ . Hence, without loss of generality, we take  $\delta = 1$  from now on bearing in mind that, for  $\epsilon \rightarrow -\infty$ , the Clayton copula with parameter  $\theta$  is obtained as a limiting case. For  $\epsilon \rightarrow \infty$ , the Clayton copula is again obtained as a limiting case, although now with parameter  $-\frac{\theta}{2\theta+1}$ .

It can be numerically shown that this copula (a) for a fixed  $\theta$  decreases in concordance with increasing  $\epsilon$ ; (b) for negative fixed  $\epsilon$ , it increases in concordance with increasing  $\theta$ ; and (c) for a positive fixed  $\epsilon$ , it first decreases and then increases in concordance in terms of  $\theta$ . For any finite  $\epsilon$ ,  $\theta \downarrow 0$  and  $\theta \rightarrow \infty$  lead to independence and comonotonicity, respectively.

According to Theorem 4.3 of Joe (1997), p. 91, for Archimedean copulas with a strict generator, the population version of Kendall’s tau can be written as:

$$\tau = 1 - 4 \int_{s=0}^{\infty} s \left( \frac{d}{ds} \psi_{\theta, \epsilon}(s) \right)^2 ds.$$

Using software system *Wolfram Mathematica* (see Wolfram Research Inc. (2017)) this gives the expression:

$$\tau = 1 - \frac{(2\theta + 1) \left( (\theta + 2)(3\theta + 2) + 4(\theta + 2)\epsilon^4 + 12(\theta + 1)\epsilon^2 + 2(\theta + 4)\epsilon\sqrt{\epsilon^2 + 1} + 4(\theta + 2)\epsilon^3\sqrt{\epsilon^2 + 1} \right)}{(\theta + 2)(3\theta + 2) \left( \theta + \epsilon \left( \sqrt{\epsilon^2 + 1} + \epsilon \right) + 1 \right)^2},$$

for  $\epsilon = 0$  considerably reducing to  $\tau = \{\theta/(\theta + 1)\}^2$ . Taking the limit for  $\epsilon \rightarrow -\infty$ , keeping  $\theta$  constant, gives  $\tau \rightarrow \theta/(\theta + 2)$ . This is the well-known formula of Kendall’s tau for the Clayton copula, and therefore not very surprising. Taking the limit  $\epsilon \rightarrow \infty$ , keeping  $\theta$  constant, gives  $\tau \rightarrow -\theta/(3\theta + 2)$ . This implies that, for increasing  $\theta$ , the range of values taken by  $\tau$  increases. So, the lowest possible value of  $\tau$  for this family is  $\lim_{\theta \rightarrow \infty} -\theta/(3\theta + 2) = -1/3$ . For fixed nonpositive  $\epsilon$ ,  $\tau$  is monotone increasing in  $\theta$  from 0 to 1. For fixed and finite positive  $\epsilon$ ,  $\tau$  as a function of  $\theta$  first decreases until a certain negative minimum that is greater than  $-1/3$ , and increases afterwards. The greater the value of  $\epsilon$  is, the greater the value of  $\theta$  for which the minimum is reached and the smaller the minimal value. The difference between Kendall’s tau for  $\epsilon = 0$  and  $\epsilon \rightarrow -\infty$  is

$$\frac{\theta}{\theta + 2} - \left( \frac{\theta}{\theta + 1} \right)^2 = \frac{\theta}{(\theta + 2)(\theta + 1)^2},$$

which is zero for  $\theta = 0$ , increasing until  $\theta = (\sqrt{5} - 1)/2 \approx 0.618$  (which for  $\epsilon = 0$  and  $\epsilon \rightarrow -\infty$  gives values of Kendall’s tau of 0.236 and 0.146, respectively), then decreasing for

increasing  $\theta$  and ultimately vanishing. There is no upper tail dependence, while the lower tail dependence coefficient is  $2^{-1/\theta}$ .

Another interesting feature of this family concerns the conditional survival copula, that is, if the copula of the conditional joint survival function  $\bar{H}(\mathbf{x}|\mathbf{y}) = \Pr[X_1 > x_1, X_2 > x_2 | X_1 > y_1, X_2 > y_2]$ . If the joint survival function  $\bar{H}(\mathbf{x})$  has an Archimedean copula with generator  $\psi(s), s \geq 0$ , the conditional joint survival function  $\bar{H}(\mathbf{x}|\mathbf{y})$  also has an Archimedean copula with generator

$$\psi_y(s) = \psi(s + t) / \psi(t), \tag{4}$$

where  $t = \psi^{-1}(\bar{H}(\mathbf{y}))$ . (Usually the conditional copula is rather given in terms of the inverse generator  $\psi_y^{-1}(s) = \psi^{-1}(s\bar{H}(\mathbf{y})) - \psi^{-1}(\bar{H}(\mathbf{y}))$ , see (Charpentier 2003; Spreuw 2006; Sungur 2002)) Applying (4) to the SAHARA family gives

$$\psi_{y,\theta,\epsilon}(s) = \left( \frac{s + t - \epsilon + \sqrt{\delta^2 + (s + t - \epsilon)^2}}{t - \epsilon + \sqrt{\delta^2 + (t - \epsilon)^2}} \right)^{-(1+1/\theta)} \left( \frac{s + t - \epsilon + (1 + 1/\theta)\sqrt{\delta^2 + (s + t - \epsilon)^2}}{t - \epsilon + (1 + 1/\theta)\sqrt{\delta^2 + (t - \epsilon)^2}} \right), \tag{5}$$

so the conditional copula is again SAHARA with parameters  $\theta$  and  $\epsilon - t$ . It follows that dependence increases over time, and the copula converges to Clayton with parameter  $\theta$ . Again, this is unlike most other copula families where the limiting dependence is either none (independence) or perfect positive.

Some scatterplots follow in Figures 1–3, for Kendall’s tau fixed at 0.25 and varying values for  $\epsilon$  and  $\theta$ . As  $\epsilon$  went up, we encountered on the one hand increasingly positive dependence in the bottom left part, and increasingly negative dependence in the top right half. Such families could be considered when data feature strongly positive dependence for small values, and weakly positive, no, or even negative dependence for large values.

The SAHARA copula is clearly flexible and versatile. A drawback is that the inverse of the generator is not available in closed form, like some families introduced in McNeil and Nešlehová (2010), Hua and Joe (2011) and Hua (2015), rendering computations more complicated.

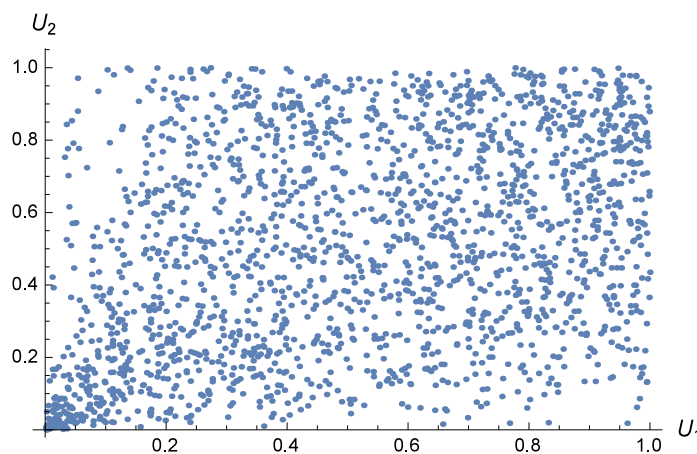


Figure 1. Scatterplot of the SAHARA copula for  $\theta = 1$  and  $\epsilon = 0$ ;  $\tau = 0.25$ .

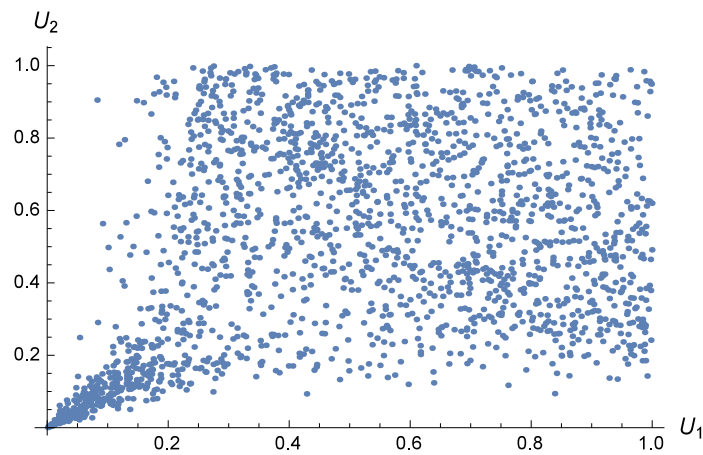


Figure 2. Scatterplot of the SAHARA copula for  $\theta = 4.5464$  and  $\epsilon = 2$ ;  $\tau = 0.25$ .

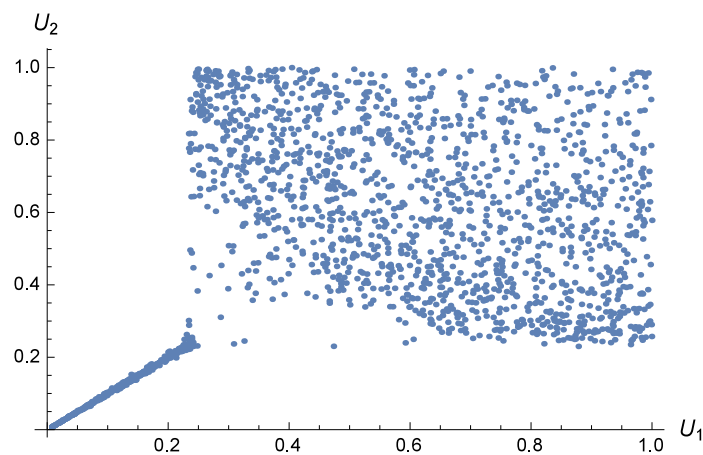


Figure 3. Scatterplot of the SAHARA copula for  $\theta = 69.11$  and  $\epsilon = 10$ ;  $\tau = 0.25$ .

#### 4. Application

For the numerical application in this section, we use the example about modelling dependence of coupled lives in Luciano et al. (2008) and Spreeuw (2014). The two publications used different data, although they both concern samples specified as generations from the same large dataset of annuitants from a Canadian insurer. In this section, copula families are fitted into the data from Luciano et al. (2008) rather than those from Spreeuw (2010). We follow the same procedure of modelling and calibration as that in Luciano et al. (2008) and Spreeuw (2014). Some elaboration on deriving the empirical generator is in order to render this paper self-contained.

The joint survival function of two remaining lifetimes  $T_x^m$  (male, age  $x$  at the start of the observation) and  $T_y^f$  (female, age  $y$  at the start of the observation) is given in terms of a survival copula  $C_{xy}$  as

$$S_{xy}(s, t) = C_{xy}(S_x^m(s), S_y^f(t)).$$

In this setup, the lives are coupled at the time when they are observed (rather than at birth, as in, e.g., Frees et al. (1996)), just like in Carriere (2000). Using a modified version of the procedure by Wang and Wells (2000), the performance of a candidate Archimedean copula is judged on the basis of distance between the empirical Kendall function, denoted by  $\hat{K}_{n(xy)}$ , and the theoretical Kendall function, denoted by  $K_{\psi_{\hat{\mathbf{A}}}^{-1}(xy)}(v)$ , where  $\psi_{\hat{\mathbf{A}}}^{-1}$  is the inverse generator of the copula concerned, with  $\hat{\mathbf{A}}$  being the parameter vector estimate minimizing the distance between  $\hat{K}_{n(xy)}$  and  $K_{\psi_{\hat{\mathbf{A}}}^{-1}(xy)}(v)$ . For single parameter copulas,

$\mathbf{A} = \theta$ , while for families with two parameters,  $\mathbf{A} = \{\theta, \epsilon\}$ . The distance or error is defined under the  $L^2$  norm (so, in the usual quadratic sense). Therefore,

$$error(\psi_{\hat{\mathbf{A}}}^{-1}(xy)) = \int_{\zeta}^1 \left( K_{\psi_{\hat{\mathbf{A}}}^{-1}(xy)}(v) - \hat{K}_{n(xy)}(v) \right)^2 dv,$$

with

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A}} \int_{\zeta}^1 \left( K_{\psi_{\mathbf{A}}^{-1}(xy)}(v) - \hat{K}_{n(xy)}(v) \right)^2 dv.$$

Given that the data were right censored, the lower bound  $\zeta$  was greater than zero. In this example, it is taken to be the smallest value for which  $\hat{K}_{n(xy)}$  is positive:

$$\zeta = \min\{v : \hat{K}_{n(xy)}(v) > 0\}.$$

The empirical Kendall function, denoted by  $\hat{K}_{n(xy)}$ , was derived from Dabrowska’s nonparametric estimator of the joint survival function (see Dabrowska (1988)). Given that the data were right censored, with many observations being doubly censored,  $\hat{K}_{n(xy)}$  is zero between 0 and a certain value  $\zeta_1 > 0$ , at which point it jumps. In this case,  $\zeta_1 = 0.23$ . The pseudo-maximum likelihood (PML) procedure uses as input rescaled Kaplan–Meier estimates of the marginal survival functions in order to accommodate censoring.

Luciano et al. (2008) fit the data to Clayton, Gumbel-Hougaard, Frank, entry 20 of Table 4.2 in Nelsen (2006) (“4.2.20 Nelsen”) and the so-called Special copula. For convenience, the last two are listed below with their generators:

1. 4.2.20 Nelsen:  $\psi_{\theta}(t) = \{\log[e + t]\}^{-\frac{1}{\theta}}$ ,  $\theta > 0$ .
2. Special:  $\psi_{\theta}(t) = \left(\frac{-t + \sqrt{t^2 + 4}}{2}\right)^{\frac{1}{\theta}}$ ,  $\theta > 0$ .

Luciano et al. (2008) concluded that 4.2.20 Nelsen fit the data best. In this section, we compare its performance with that of SAHARA and the best contender of common two-parameter families from Joe (1997, 2015), i.e., BB2. Its generator is:

$$\psi_{\theta}(t) = \left\{ 1 + \frac{\log[1 + t]}{\epsilon} \right\}^{-\frac{1}{\theta}}, \quad \theta, \epsilon > 0.$$

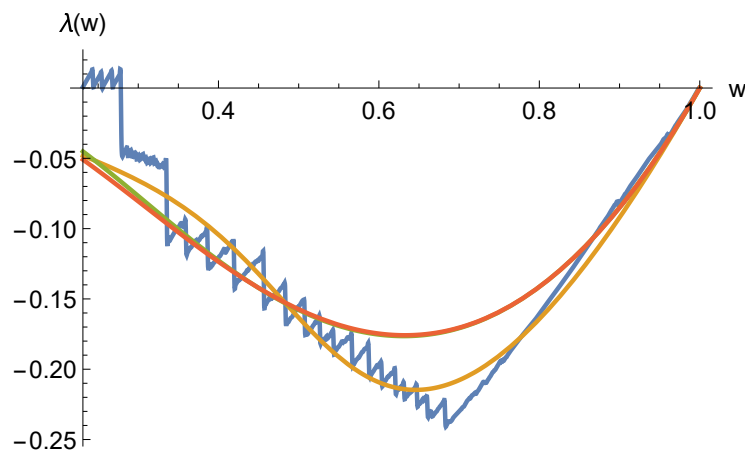
The 4.2.20 Nelsen is a special case of BB2 arising for  $\epsilon = 1$ .

Results are given in Table 1. The positive estimate for  $\epsilon$  indicates the absence of SI.

**Table 1.** Results for several copula families.

Copula	Parameter Estimates	Error $\varphi_{\hat{\theta}}^{[-1]}(xy)$
4.2.20 Nelsen	$\hat{\theta} = 1.005$	0.720
BB2 Joe (1997)	$\hat{\theta} = 1.469; \hat{\epsilon} = 0.383$	0.667
SAHARA	$\hat{\theta} = 0.204; \hat{\epsilon} = 0.914$	0.293

As in Luciano et al. (2008), we performed a graphical comparison between the theoretical and the empirical K functions through transformation  $\lambda(w) = w - K(w)$  for  $\zeta_1 \leq w \leq 1$ . The result can be found in Figure 4. SAHARA achieved a significant improvement to the fit compared to the other families.



**Figure 4.** Graphical comparison between theoretical  $\lambda(w) = w - K(w)$  for SAHARA (yellow), BB2 (green) and 4.2.20 Nelsen (orange) and empirical one (blue).

Now consider the notion of SI in more detail. For two random variables  $X_1$  and  $X_2$ ,  $X_2$  being SI in  $X_1$  is equivalent to  $\Pr[X_2 > x_2 | X_1 = x_1]$  being nondecreasing in  $x_1$  for all  $x_2$ . Related to this is the notion of long-term dependence as introduced in Hougaard (2000). If we define  $\mu^m(t | T_y^f = t_y)$  as the conditional force of mortality of life ( $x$ ) at duration  $t$  given  $T_y^f = t_y < t$  ( $y$  dies at duration  $t_y$ ), then the dependence between  $T_x^m$  and  $T_y^f$  is of the long-term type if  $\mu^m(t | T_y^f = t_y)$  is constant or decreasing as a function of  $t_y$ , while dependence is short-term if  $\mu^m(t | T_y^f = t_y)$  is increasing as a function of  $t_y$ . To understand this, it is important to that, as indicated before, for Archimedean copulas, stochastic increasing (SI) is equivalent to  $-\psi'$  being logconvex. Spreeuw (2006) showed that this property of the generator also applies to long-term dependence, implying that SI and long-term dependence are actually equivalent. On the other hand, however, for the SAHARA family, we have:

$$\frac{\partial \left\{ \ln \left[ -\psi'_{\theta, \epsilon}(s) \right] \right\}}{\partial s} = -\frac{1 + 1/\theta}{\sqrt{(s - \epsilon)^2 + 1}},$$

which is monotone increasing in  $s$  across the board for negative  $\epsilon$ . For positive values of  $\epsilon$ , however, the expression decreases in  $s$  for  $0 \leq s < \epsilon$ , so the SI property does not hold. The positive parameter estimate for  $\epsilon$  suggests that short-term rather than long-term dependence may prevail between the coupled lives. To investigate this further, we analysed the data in the same vein as in Spreeuw and Owadally (2013), who devise an augmented Markov model to allow for short-term dependence for the entire dataset. Results are reported in Table 2.

In this table,  $e$  denotes the time in which an integer number of years that have elapsed since the death of the partner. So, e.g.,  $e = 0$  concerns the lives that were bereaved less than a year ago. For each possible value of  $e$  (noting that each life was observed for 5 years or less), we calculate number of deaths reported, the risk exposure, and the overall mortality rate being the ratio of the values in the second and third column. So, for instance, the risk exposure of lives who lost their partner less than a year ago is 604.87, and there were 69 lives that died within one year after their partner. Now, long-term dependence implies that the mortality rate in the last column should be increasing as a function of  $e$ , but the results in Table 2 show that this is not the case, and that short-term dependence may be present even though the aggregate mortality rate for  $e = 4$  is higher than for  $e$  equal to 1, 2 or 3.

Ideally, one such table should be shown for each gender. However, due to the small number of observed deaths in the dataset, in particular for higher values of  $e$  caused by

heavy censoring, males and females were combined. Results in Table 2 should thus be interpreted as an indication of possible short-term dependence, rather than firm evidence.

**Table 2.** Mortality for all couples, with  $e$  denoting the number of years since partner's death.

	Deaths	Exposure	Mortality
Partner dead			
$e = 0$	69	604.87	0.114075
$e = 1$	17	428.44	0.039679
$e = 2$	9	277.76	0.032403
$e = 3$	4	155.08	0.025590
$e = 4$	3	49.67	0.060395
Partner alive	751	34,631.45	0.021685

## 5. Conclusions

In this paper, we introduced a new Archimedean copula family derived from the SAHARA utility function. With SAHARA utility first increasing to a maximum and subsequently decreasing, the corresponding copula family allows for stochastically increasing (SI) and non-SI at the same time, depending on the sign of one of the parameters. As the numerical application shows, this family could fit the mortality data of coupled lives well. The parameter estimates suggest the possible existence of short-term dependence, i.e., the mortality of bereaved lives increases on bereavement but diminishes later.

**Funding:** This research received no external funding.

**Data Availability Statement:** Anyone who would like to use the data should consult Professor Edward W. (Jed) Frees of University of Wisconsin-Madison, USA, e-mail: jfrees@bus.wisc.edu, or Professor Emiliano Valdez of University of Connecticut, USA, e-mail: emiliano.valdez@uconn.edu.

**Acknowledgments:** We wish to acknowledge the Society of Actuaries, through the courtesy of Professors Edward (Jed) Frees and Emiliano Valdez, for providing the data used in this paper.

**Conflicts of Interest:** The author declares no conflict of interest.

## References

- Albrecher, Hansjörg, Corina Constantinescu, and Stéphane Loisel. 2011. Explicit ruin formulas for models with dependence among risks. *Insurance: Mathematics and Economics* 48: 265–70. [CrossRef]
- Avérous, Jean, and Jean-Luc Dortet-Bernadet. 2004. Dependence for Archimedean copulas and aging properties of their generating functions. *Sankhyā: The Indian Journal of Statistics* 66: 607–20.
- Bernard, Carole, Luca De Gennaro Aquino, and Lucia Levante. 2021. Optimal annuity demand for general expected utility agents. *Insurance: Mathematics and Economics* 101: 70–79. [CrossRef]
- Bernard, Carole, and Minsuk Kwak. 2016. Dynamic preferences for popular investment strategies in pension funds. *Scandinavian Actuarial Journal* 2016: 398–419. [CrossRef]
- Brachetta, Matteo, and Hanspeter Schmidli. 2020. Optimal reinsurance and investment in a diffusion model. *Decisions in Economics and Finance* 43: 341–61. [CrossRef]
- Capéraà, Philippe, Anne-Laure Fougères, and Christian Genest. 2000. Bivariate distributions with given extreme value attractor. *Journal of Multivariate Analysis* 72: 30–49. [CrossRef]
- Capéraà, Philippe, and Christian Genest. 1993. Spearman's  $\rho$  is larger than Kendall's  $\tau$  for positively dependent random variables. *Journal of Nonparametric Statistics* 2: 183–94. [CrossRef]
- Carriere, Jacques F. 2000. Bivariate survival models for coupled lives. *Scandinavian Actuarial Journal* 2000: 17–32. [CrossRef]
- Charpentier, Arthur 2003. Dependence and tail distributions. Paper Presented at Seventh International Conference on Insurance: Mathematics and Economics, Lyon, France, June 25–27.
- Chen, An, Antoon Pelsler, and Michel Vellekoop, 2011. Modeling non-monotone risk aversion using SAHARA utility functions. *Journal of Economic Theory* 146: 2075–92. [CrossRef]
- Chen, An, Thai Nguyen, and Nils Sørensen. 2021. Indifference pricing under SAHARA utility. *Journal of Computational and Applied Mathematics* 388: 113288. [CrossRef]



- Chen, An, and Michel Vellekoop. 2017. Optimal investment and consumption when allowing terminal debt. *European Journal of Operational Research* 258: 385–97. [CrossRef]
- Côté, Marie-Pier, and Christian Genest. 2019. Dependence in a background risk model. *Journal of Multivariate Analysis* 172: 28–46. [CrossRef]
- Dabrowska, Dorota M. 1988. Kaplan-Meier estimate on the plane. *The Annals of Statistics* 16: 1475–89. [CrossRef]
- Frees, Edward W., Jacques Carriere, and Emiliano Valdez. 1996. Annuity valuation with dependent mortality. *Journal of Risk and Insurance* 63: 229–61. [CrossRef]
- Furman, Edward, Yisub Kye, and Jianxi Su. 2021. Multiplicative background risk models: Setting a course for the idiosyncratic risk factors distributed phase-type. *Insurance: Mathematics and Economics* 96: 153–67. [CrossRef]
- Genest, Christian, Kilani Ghoudi, and Louis-Paul Rivest. 1998. Discussion on “Understanding relationships using copulas” by E. Frees and E. Valdez. *North American Actuarial Journal* 2: 143–49. [CrossRef]
- Hougaard, Philip. 2000. *Analysis of Multivariate Survival Data*. New York: Springer, vol. 564.
- Hua, Lei. 2015. Tail negative dependence and its applications for aggregate loss modeling. *Insurance: Mathematics and Economics* 61: 135–45. [CrossRef]
- Hua, Lei, and Harry Joe. 2011. Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis* 102: 1454–71. [CrossRef]
- Joe, Harry. 1997. *Multivariate Models and Dependence Concepts*. Boca Raton: Chapman & Hall/CRC.
- Joe, Harry. 2015. *Dependence Modeling with Copulas*. Monographs on Statistics and Applied Probability 134. Boca Raton, London and New York: CRC Press, Taylor and Francis Group.
- Li, Wenyuan, and Jingtang Ma. 2018. Optimal investment strategies for general utilities under dynamic elasticity of variance models. *Quantitative Finance* 18: 1379–88. [CrossRef]
- Luciano Elisa, Jaap Spreeuw, and Elena Vigna. 2008. Modelling stochastic mortality for dependent lives. *Insurance: Mathematics and Economics* 43: 234–44. [CrossRef]
- Marri, Fouad, and Khouzeima Moutanabbir. 2022. Risk aggregation and capital allocation using a new generalized Archimedean copula. *Insurance: Mathematics and Economics* 102: 75–90. [CrossRef]
- McNeil, Alexander J., and Johanna Nešlehová. 2009. Multivariate Archimedean copulas, d-monotone functions and  $\ell_1$ -norm symmetric distributions. *The Annals of Statistics* 37: 3059–97. [CrossRef]
- McNeil, Alexander J., and Johanna Nešlehová. 2010. From Archimedean to Liouville copulas. *Journal of Multivariate Analysis* 101: 1772–90. [CrossRef]
- Michiels, Frederik, and Ann De Schepper. 2009. *Understanding Copula Transforms: A Review of Dependence Properties*. Working Paper. Antwerp: University of Antwerp.
- Michiels, Frederik, and Ann De Schepper. 2012. How to improve the fit of Archimedean copulas by means of transforms. *Statistical Papers* 53: 345–55. [CrossRef]
- Michiels, Frederik, Inge Koch, and Ann De Schepper. 2011. A new method for the construction of bivariate Archimedean copulas based on the  $\lambda$  function. *Communications in Statistics—Theory and Methods* 40: 2670–79. [CrossRef]
- Müller, Alfred, and Marco Scarsini. 2005. Archimedean copulae and positive dependence. *Journal of Multivariate Analysis* 93: 434–45. [CrossRef]
- Nelsen, Roger B. 2006. *An Introduction to Copulas*, 2nd ed. New York: Springer Science & Business Media.
- Sarabia, José María, Emilio Góómez-Déniz, Faustino Prieto, and Vanesa Jordá. 2018. Aggregation of dependent risks in mixtures of exponential distributions and extensions. *ASTIN Bulletin: The Journal of the IAA* 48: 1079–107. [CrossRef]
- Schumacher, Johannes M. 2018. Linear versus nonlinear allocation rules in risk sharing under financial fairness. *ASTIN Bulletin* 48: 995–1024. [CrossRef]
- Spreeuw, Jaap. 2006. Types of dependence and time-dependent association between two lifetimes in single parameter copula models. *Scandinavian Actuarial Journal* 2006: 286–309. [CrossRef]
- Spreeuw, Jaap. 2010. Relationships between Archimedean copulas and Morgenstern utility functions. In *Copula Theory and Its Applications*. Edited by Piotr Jaworski, Fabrizio Durante, Wolfgang Karl Härdle and Tomasz Rychlik. Lecture Notes in Statistics. Berlin and Heidelberg: Springer, vol. 198. [CrossRef]
- Spreeuw, Jaap. 2014. Archimedean copulas derived from utility functions. *Insurance: Mathematics and Economics* 59: 235–42. [CrossRef]
- Spreeuw, Jaap, and Iqbal Owadally. 2013. Investigating the broken-heart effect: a model for short-term dependence between the remaining lifetimes of joint lives. *Annals of Actuarial Science* 7: 236–57. [CrossRef]
- Sungur, Engin A. 2002. Some results on truncation dependence invariant class of copulas. *Communications in Statistics—Theory and Methods* 31: 1399–422. [CrossRef]
- Wang, Weijing, and Martin T. Wells. 2000. Model selection and semiparametric inference for bivariate failure-time data. *Journal of the American Statistical Association* 95: 62–72. [CrossRef]
- Wolfram Research Inc. 2017. *Mathematica*, version 11.2. Champaign: Wolfram Research Inc.

MDPI  
St. Alban-Anlage 66  
4052 Basel  
Switzerland  
Tel. +41 61 683 77 34  
Fax +41 61 302 89 18  
[www.mdpi.com](http://www.mdpi.com)

*Risks* Editorial Office  
E-mail: [risks@mdpi.com](mailto:risks@mdpi.com)  
[www.mdpi.com/journal/risks](http://www.mdpi.com/journal/risks)







Academic Open  
Access Publishing

[www.mdpi.com](http://www.mdpi.com)

ISBN 978-3-0365-8390-7