Stress scenario generation for solvency and risk management

Marcus Christian Christiansen*[1], Lars Frederik Brandt Henriksen†[2], Kristian Juul Schomacker‡[2] and Mogens Steffensen§[3]

[1]University of Ulm
[2]Edlund A/S
[3]University of Copenhagen

Abstract
We derive worst-case scenarios in a life insurance model in the case where the interest rate and the various transition intensities are mutually dependent. Examples of this dependence are that a) surrender intensities and interest rates are high at the same time, b) mortality intensities of a policyholder as active and disabled, respectively, are low at the same time, and c) mortality intensities of the policyholders in a portfolio are low at the same time. The set from which the worst-case scenario is taken reflects the dependence structure and allows us to relate the worst-case scenario-based reserve, qualitatively, to a Value-at-Risk-based calculation of solvency capital requirements. This brings out perspectives for our results in relation to qualifying the standard formula of Solvency II or using a scenario-based approach in internal models. Our results are powerful for various applications and the techniques are non-standard in control theory, exactly because our worst-case scenario is deterministic and not adapted to the stochastic development of the portfolio. The formalistic results are exemplified in a series of numerical studies.

Keywords: Life insurance, worst-case scenario, deterministic control, Solvency II, multi-state Markov chain.

*marcus.christiansen@uni-ulm.de
†lars.henriksen@edlund.dk, Bjerregårds Sidevej 4, 2500 Valby, Denmark
‡kristian.schomacker@edlund.dk
§mogens@math.ku.dk

1 Introduction

From a specific set of scenarios we find the scenario that leads to the worst-case, i.e. largest, value of future obligations. The idea of such worst-case scenarios has a wide range of applications in life insurance pricing, management, and regulation. They include settlement of premiums and surrender values as well as calculation of risk margins and solvency capital requirements (SCR). In particular we draw the attention to the scenario-based standard formula in Solvency II, see Steffen (2008) for an overview of the Solvency II project. So far it has been difficult to say something qualitatively about the relation between a standard scenario and the original Value-at-Risk-based SCR. We generate scenarios such that SCRs based on these scenarios are proven sufficient under the Value-at-Risk measure within a given model. As input to our calculation, we take a set of interest and transition rates such that the probability of realizing interest and transition rates within that set is bounded. For such a set we calculate the scenario that maximizes the reserve. We do not generally address the difficult but interesting question of detecting such a set although we exemplify possible sets in some examples.

Our results are applicable to an inhomogeneous portfolio of contracts. Then the worst-case scenario generated is worst-case for a whole portfolio with different policyholder ages and contracts. This makes the approach particularly useful in the discussion about portfolio SCRs. The results can qualify this discussion in two dimensions. First, for a regulator who wants to develop a stress scenario-based SCR we provide a scenario corresponding to bounds on shortfall probabilities. This qualified standard formula should be derived for a stylized market-realistic portfolio. Second, an insurance company can replace the standard SCR formula by a so-called (partial) internal model-based calculation and still exploit advantages working with scenarios. Our results show how such "internal stress scenarios" can be derived.

In relation to recent academic literature on worst-case scenario generation, we emphasize that our stress scenarios are deterministic in the sense that e.g. the portfolio worst-case mortality intensity at a future time point is calculated today and will not depend on the survivors of the portfolio at that future time point. This means that we are, briefly speaking, finding optimal deterministic processes maximizing an expectation in a stochastic environment. This is a non-standard exercise that contains methodological and computational challenges. The upside is that, once they are overcome, the resulting scenario can be better understood, communicated, implemented and extrapolated for usage in other (similar) portfolios. The idea to look for deterministic worst-case intensities is in sharp contrast to e.g. Li and Szimayer (2011, 2014) who, in a different framework, also study worst-case intensities. They, however, use more standard stochastic control techniques to derive intensities that are adapted to the development of the contract under study. This development amounts, in their cases, to the development of asset prices but, in a more general setting, it could be the number of, or even the names of, survivors in a portfolio. Such adapted scenarios may be useful for other applications but not necessarily for solvency issues. Thus, the conceptual innovation in relation to Li and Szimayer (2011, 2014) is that our worst case is deterministic and its derivation therefore draws on other techniques.

The present work extends the results of Christiansen and Steffensen (2013) in the following way: Like us, they search for optimal deterministic scenarios and obtain simple formulas for these but in a quite restricted class of models. The class is defined endogenously by requiring that certain argmax operations over transition intensities are constant with respect to the transition probabilities they generate, see Christiansen and Steffensen (2013, Proposition 4.1 and 4.2). The work in this paper is very much inspired from the structure of problems and solutions in that article, but we succeed in finding the worst-case scenario also outside their restrictive assumptions. This allows for studying much more realistic and important cases like a hierarchical disability model or a model for a portfolio of heterogeneous contracts hit by the same worst case. Thus, we develop a powerful tool for various applications while sticking to the natural but challenging idea of Christiansen and Steffensen (2013) that the worst-case should be deterministic. Thus, the innovation in relation to Christiansen and Steffensen (2013) is that we find, by means different than theirs, the worst case for interesting products and portfolios that they rule out.

In order to show how the studies in this paper can be applied to solvency calculations we choose to give the reader, already here in the introduction, a glance of the formalistic argument. Details can be found in Christiansen and Steffensen (2013). For (real, unknown) interest and transition rates (ϕ, μ), we want to find deterministic interest and transition rates $(\widetilde{\phi}, \widetilde{\mu})$ such that for the liabilities L, it holds that

$$P\left(L(t,\widetilde{\phi},\widetilde{\mu}) \geq L(t,\phi,\mu)\right) \geq 1-\alpha, \tag{1.1}$$

where $\alpha \in [0,1)$. That is, we want to find a deterministic calculation basis such that the liabilities calculated with this basis with a certain probability are larger than the liabilities calculated with the real (stochastic) basis. This can be obtained by choosing $(\widetilde{\phi}, \widetilde{\mu}) = \text{argmax}_{(\phi,\mu)\in M} L(t,\phi,\mu)$ for a set M such that $P((\phi,\mu) \in M) \geq 1-\alpha$. We do not pay any attention to how the set M is formed, except for in a few numerical examples. The object of study in this paper is, given a set M, to calculate the argmax, $(\widetilde{\phi}, \widetilde{\mu})$.

As shown in Christiansen and Steffensen (2013), we can use this to obtain an upper bound for the SCR given by

$$\sup_{(\phi,\mu)\in M} \{L(t,\phi,\mu)\} - L\left(t,\phi^{\mathrm{BE}},\mu^{\mathrm{BE}}\right),$$

where ϕ^{BE} and μ^{BE} are best estimates for the interest rate and transition intensities, respectively.

The quality of an SCR based on a standard stress scenario compared to a Value-at-Risk-based calculation has been intensively discussed in the literature. Doff (2008) analyses, critically, the Solvency II proposal and the shortcomings of the standard stress model, which is also discussed by Devineau and Loisel (2009). Specific attention has been given to longevity risk and Olivieri and Pitacco (2008) study the pitfalls of approaching longevity risk by means of stress scenarios. A comprehensive numerical study that compares the stress scenarios to the Value-at-Risk calculation can be found in Börger (2010). To the authors' knowledge, all comparative studies are quantitative in the sense that the various principles for SCR calculations are numerically related to each other. If the stress-based SCR is significantly smaller than the Value-at-Risk based SCR, the whole idea of inducing financial stability from the standard formula may be criticised for being an optical illusion. We distinguish ourselves from this quantitative discussion by searching for a stress, such that the SCR derived from this stress scenario is at least as large as the SCR that can be derived from a Value-at-Risk approach. Thus, rather than criticising the idea of stress scenarios, we admit its advantages and seek to qualify the discussion about what the stress scenarios should look like. Our study is general enough to help answer this question no matter if the unit is a contract or a portfolio.

The paper is organized as follows: In Section 2 we introduce the insurance market and get a representation of the probability weighted reserve. In Section 3 we obtain worst-case scenarios and reserves for a single policy and describe numerical methods which are needed for the numerical calculations of these quantities. In Section 4, we extend the theory to cover a portfolio of policyholders, and finally we present some numerical calculations for both single policies and portfolios in Section 5.

2 Modelling and valuation of an insurance policy

Let T be a fixed finite time horizon and $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{0\leq t\leq T}, P)$ a filtered probability space with filtration satisfying the usual conditions of right-continuity and completeness. Let X be a pure jump process defined on this probability space with finite state space $\mathcal{S}$ which we assume consists of n states. The process $(X(t))_{t\in[0,T]}$ represents the state of a policyholder. We do not assume a deterministic starting value for X but only a starting distribution, which we denote π. This enables us to easily calculate the worst-case reserve for a homogeneous portfolio of insurance contracts, where each policyholder can be in different initial states like "Active" and "Disabled". This is a simple special case of the more general theory for portfolios presented in Section 4. We denote by $\mathcal{J}$ the transition space of X defined by $\mathcal{J} = \left\{(j,k) \in \mathcal{S}^2 | j \neq k\right\}$.

Following the lines of Norberg (1999), we assume that the interest rate intensity ϕ is piecewise continuous and that $\int_0^T \phi(s)ds$ is finite. We define a discounting function v by the following forward equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}v(t) \quad = \quad -v(t)\phi(t), \quad v(t_0) = 1, \tag{2.1}$$

where t_0 is not necessarily the initiation time of the contract. One can intuitively think of t_0 as 0 but we introduce the notation t_0 in order to be able to accurately formulate the verification lemma. The solution to (2.1) is given by

$$v(t) = e^{-\int_{t_0}^t \phi(\tau)d\tau},$$

and the discounting factor for the time span from s to t is given by

$$v(s,t) = \frac{v(t)}{v(s)} = e^{-\int_s^t \phi(\tau)d\tau}.$$

We let the $n \times n$ transition matrix for X be denoted by p, which is a function with two arguments. That is, for $t_0 \leq s \leq t \leq T$ we have that

$$p(s,t) = \left(P(X(t) = k | X(s) = j)\right)_{(j,k)\in\mathcal{S}^2}.$$

By μ we denote the corresponding $n \times n$ intensity matrix. It is well-known that p is determined by Kolmogorov's forward equation:

$$\frac{\partial}{\partial t}p(s,t) = p(s,t)\mu(t), \quad p(s,s) = \mathbf{I}_n,$$

where $\mathbf{I}_n$ is the n-dimensional identity matrix.

With a little abuse of notation we denote by p (a function of one variable only) the marginal distribution of the random pattern of states:

$$p(t) = \begin{pmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_n(t) \end{pmatrix},$$

where $p_j(t) = P(X(t) = j)$. The forward equation for p is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} p(t) = \mu^{\mathrm{tr}}(t) p(t), \quad p(t_0) = \pi, \tag{2.2}$$

where π is a given initial distribution over the n states and $\mu^{\mathrm{tr}}(t)$ is the transpose of $\mu(t)$. Equation (2.2) follows by Kolmogorov's forward equation by using that $p_j = \sum_{i \in \mathcal{S}} \pi_i p_{ij}$, which gives us that the dynamics of p_j are equal to p_i times the columns of μ. This is exactly equal to $\mu^{\mathrm{tr}}(t)p(t)$.

We consider an insurance contract with the following type of payments:

1. $b_i(t)$ is the rate of payments in state i at time t.
2. $b_{ij}(t)$ is a lump sum payment payable upon transition from state i to state j at time t.

We denote by $B_i(t)$ the accumulated payments in state i up to time t and by $B(t)$ we denote the present value at time t of future payments of the contract. We assume that all the functions b_{ij} and B_i have bounded variation on $[0, T]$ and that the functions b_i and b_{ij} are C^1 on $[0, T]$. The latter assumption can easily be relaxed to piecewise C^1 but this will result in cumbersome notation, since we need to deal with extra boundary conditions.

To keep the notation simple, we assume that no lump sum payments are paid out while sojourning in a state. That is, for all $t \in [0, T]$ we have that

$$\Delta B_i(t) = B_i(t) - B_i(t-) = 0. \tag{2.3}$$

However, the results of the paper can easily be extended to the case, where assumption (2.3) does not hold.

We are now able to define the *statewise* prospective reserves. The statewise prospective reserve in state $j \in \mathcal{S}$ is given by

$$\begin{aligned} V_j(t) &= \mathrm{E}\left[B(t) | X(t) = j\right] \\ &= \sum_{k \in \mathcal{S}} \int_t^T v(t, u) p_{jk}(t, u) \left(b_k(u) + \sum_{l \in \mathcal{S} : l \neq k} \mu_{kl}(u) b_{kl}(u) \right) du. \end{aligned}$$

The standard Thiele's (backward) differential equation for the reserve is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} V_j(t) = -b_j(t) + \phi(t) V_j(t) - \sum_{k \in \mathcal{S} : k \neq j} \left(b_{jk}(t) + V_k(t) - V_j(t)\right) \mu_{jk}(t), \quad V_j(T) = 0, \tag{2.4}$$

where the boundary condition follows because of assumption (2.3). Note that the integral form of (2.4) is known as Thiele's integral equation of type II. With defining a mapping $W : [t_0, T] \times (0, \infty) \times [0, 1]^{|\mathcal{S}|} \to \mathbb{R}$ by

$$W(t, v, p) := v \sum_{j \in \mathcal{S}} p_j V_j(t), \tag{2.5}$$

the expected present value at time t_0 of the future payments between time t and time T equals

$$\begin{aligned} W(t, v(t), p(t)) &= \sum_{j \in \mathcal{S}} \pi_j \mathrm{E}\left[v(t)B(t) | X(t_0) = j\right] \\ &= v(t) \sum_{j \in \mathcal{S}} p_j(t) V_j(t) \\ &= \sum_{k \in \mathcal{S}} \int_t^T p_k(u) v(u) \left(b_k(u) + \sum_{l \in \mathcal{S}: l \neq k} b_{kl}(u) \mu_{kl}(u) \right) du. \end{aligned} \tag{2.6}$$

Here, the last equality follows from the Chapman-Kolmogorov equation and by collecting the discounting terms. The functions p and v in (2.6) follow from (2.1) and (2.2) and the initial values $v(t)$ and $p(t)$.

3 Calculation of the worst-case reserve

In this section, we derive the worst-case scenario for the probability weighted reserve, $\pi^{\mathrm{tr}} V(t_0)$, which is related to W by (2.5). First, we establish a verification lemma and give a heuristic argument for the main ingredients. Second, we show existence of a worst-case scenario. Third, we translate the verification result for the probability weighted reserve W into a corresponding result for the statewise reserves V_j in a corollary. Finally, we outline two numerical methods for calculation of the worst-case reserve.

3.1 Verification lemma

We note that $\pi^{\mathrm{tr}} V(t_0) = W(t_0, v(t_0), p(t_0)) = W(t_0, 1, \pi)$. This is useful since dynamic programming applies to W and not to $\pi^{\mathrm{tr}} V$. This allows us to attack our optimization problem by maximizing W for all future time points, whereas $V_j, j \in \mathcal{S}$ is not in itself maximized for $t > t_0$.

In the following, we search for the worst-case reserve (the optimal value function) with respect to a set M, where $M \subset L_1^{1+|\mathcal{J}|}([t_0, T])$ is a set of integrable interest rate and transition intensity paths. A set M belongs to $L_1^{1+|\mathcal{J}|}([t_0, T])$ if $\sum_{i=1}^{1+|\mathcal{J}|} \int_{t_0}^T |f_i(s)| ds < \infty$ for all $f_i(t)$ in $M_i(t)$. The "slices" $M(t)$ of M,

$$M(t) = \{(\phi(t), \mu(t)) | (\phi, \mu) \in M\},$$

describe the parameter space at time t. We start by establishing a classical verification lemma.

Proposition 3.1. *(Verification lemma) Let $\bar{W}$ be a solution to the partial differential* *equation*

$$
\begin{aligned}
0 = \frac{\partial}{\partial t}\bar{W}(t,v,p) - \bar{\phi}(t,v,p)v\frac{\partial}{\partial v}\bar{W}(t,v,p) + \sum_{k\in\mathcal{S}} p_k \Bigg(vb_k(t) \\
+ \sum_{l\in\mathcal{S}:l\neq k} \bar{\mu}_{kl}(t,v,p)\left(vb_{kl}(t) + \frac{\partial}{\partial p_l}\bar{W}(t,v,p) - \frac{\partial}{\partial p_k}\bar{W}(t,v,p)\right)\Bigg), \quad \bar{W}(T,v,p) = 0,
\end{aligned}
$$

$$
\begin{aligned}
\big(\bar{\phi}(t,v,p), \bar{\mu}(t,v,p)\big) = \underset{(f,m)\in M(t)}{\operatorname{argmax}} \Bigg\{ - fv\frac{\partial}{\partial v}\bar{W}(t,v,p) + \sum_{k\in\mathcal{S}} p_k \Bigg(vb_k(t) \\
+ \sum_{l\in\mathcal{S}:l\neq k} m_{kl}\left(vb_{kl}(t) + \frac{\partial}{\partial p_l}\bar{W}(t,v,p) - \frac{\partial}{\partial p_k}\bar{W}(t,v,p)\right)\Bigg)\Bigg\}.
\end{aligned} \tag{3.1}
$$

Furthermore, let the ordinary differential equation system

$$
\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}t}\bar{v}(t) + \bar{\phi}(t,\bar{v}(t),\bar{p}(t))\bar{v}(t), \\
0 &= \frac{\mathrm{d}}{\mathrm{d}t}\bar{p}_j(t) - \sum_{k\in\mathcal{S}:k\neq j} \left(\bar{\mu}_{kj}(t,\bar{v}(t),\bar{p}(t))\bar{p}_k(t) - \bar{\mu}_{jk}(t,\bar{v}(t),\bar{p}(t))\bar{p}_j(t)\right), \quad j\in\mathcal{S},
\end{aligned} \tag{3.2}
$$

have a unique solution on $[s,T]$ *for any initial condition* $(\bar{v}(s),\bar{p}(s)) = (v,p)$ *at any time* $s\in[t_0,T]$ *and any pair* $(v,p)\in(0,\infty)\times[0,1]^{|\mathcal{S}|}$*. Then*

$$
\bar{W}(s,v,p) = \sup_{(\phi,\mu)\in M} W(s,v,p;\phi,\mu) = \sup_{(\phi,\mu)\in M} v\sum_{j\in\mathcal{S}} p_j V_j(s;\phi,\mu) \tag{3.3}
$$

for all triples (s,v,p)*.*

Proof. See Bertsekas (2005, Section 3.2). □

Note that the notation "$;\phi,\mu$" in $W(s,v,p;\phi,\mu)$ and $V_j(s;\phi,\mu)$ emphasises that W and V_j depends on the entire processes ϕ and μ.

In the following, we heuristically derive the differential equations in Proposition 3.2. We start by combining two different differential equations for W for a given (ϕ,μ). First, the derivative of W with respect to t, using (2.6), is given by

$$
\frac{\mathrm{d}}{\mathrm{d}t}W(t,v(t),p(t)) = -\sum_{k\in\mathcal{S}} p_k(t)v(t)\left(b_k(t) + \sum_{l\in\mathcal{S}:l\neq k} b_{kl}(t)\mu_{kl}(t)\right). \tag{3.4}
$$

Second, we can also consider W as a function of three variables with dynamics given by

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}W(t,v(t),p(t)) =& \frac{\partial}{\partial t}W(t,v(t),p(t)) + \frac{\partial}{\partial v}W(t,v(t),p(t))\left(\frac{\mathrm{d}}{\mathrm{d}t}v(t)\right) \\
&+ \nabla_p W(t,v(t),p(t))\left(\frac{\mathrm{d}}{\mathrm{d}t}p(t)\right),
\end{aligned} \tag{3.5}
$$

where

$$\nabla_p W = \left(\frac{\partial}{\partial p_1} W, \frac{\partial}{\partial p_2} W, \cdots, \frac{\partial}{\partial p_n} W \right).$$

By inserting the differential equations for $v(t)$ and $p(t)$ (given by (2.1) and (2.2), respectively) into (3.5), combining with (3.4) and rearranging the terms, we obtain the following differential equation

$$\begin{aligned} 0 =& \frac{\partial}{\partial t} W(t, v(t), p(t)) - \phi(t) v(t) \frac{\partial}{\partial v} W(t, v(t), p(t)) \\ &+ \sum_{k \in \mathcal{S}} p_k(t) \left(\sum_{l \in \mathcal{S}: l \neq k} \mu_{kl}(t) \left(v(t) b_{kl}(t) + \frac{\partial}{\partial p_l} W(t, v(t), p(t)) - \frac{\partial}{\partial p_k} W(t, v(t), p(t)) \right) \right. \\ &\left. \qquad + v(t) b_k(t) \right), \end{aligned} \tag{3.6}$$

with terminal condition given by $W(T, v, p) = 0$. Choosing (ϕ, μ) such that the time-derivative is as small as possible at each time point and also using this (ϕ, μ) in the differential equations for v and p give us the equations that characterize the worst-case reserve. We obtain the differential equations for $\bar{v}$ and $\bar{p}$ by inserting the worst-case interest rate into (2.1) and the worst-case intensities into a coordinate-wise version of (2.2). Hereby, we have heuristically derived the system of differential equations in Proposition 3.2.

3.2 Existence

We now turn to the question of existence of a worst-case reserve.

Proposition 3.2. *(Existence of a worst-case scenario) Let M be a compact subset of $L_1^{1+|\mathcal{J}|}([t_0, T])$ which contains only nonnegative interest rates ϕ and nonnegative transition intensities μ_{jk}, $(j, k) \in \mathcal{J}$. Then for each $(t, v, p) \in [t_0, T] \times (0, \infty) \times [0, 1]^{|\mathcal{S}|}$ there exists a maximizing argument $(\bar{\phi}, \bar{\mu}) \in M$ for which $W(t, v, p; \bar{\phi}, \bar{\mu}) = \bar{W}(t, v, p)$.*

Proof. Since ϕ and μ_{jk}, $(j, k) \in \mathcal{J}$ are nonnegative, we necessarily have $|v(t)| \leq 1$ and $|p_j(t)| \leq 1$ for all t and j. Therefore, the reserves $V_j(t)$ are uniformly bounded by

$$|V_j(t)| \leq \sum_{k \in \mathcal{S}} \int_t^T \Big(|b_k(s)| + \sum_{k \in \mathcal{S}: l \neq k} |b_{kl}(s)| \mu_{kl}(s) \Big) ds < C$$

for a finite constant C (which is independent of j, t and μ) since the functions $b_k(t)$ and $b_{kl}(t)$ are bounded and M is a compact subspace of $L_1^{1+|\mathcal{J}|}([t_0, T])$. Consequently, on the set M the mapping $W(t, v, p; \phi, \mu)$ has the uniform upper bound given by

$$W(t, v, p; \phi, \mu) \leq v \sum_{j \in \mathcal{S}} p_j C.$$

Furthermore, Christiansen (2008, Theorem 4.4) showed that the reserves $V_j(t)$ are Fréchet differentiable with respect to the cumulative intensities $t \mapsto \int_{t_0}^t \phi(u) du$ and $t \mapsto \int_{t_0}^t \mu(u) du$ in the total variation norm. Since the operator that maps the intensities to the cumulative intensities is continuous and since the L_1-norm of the intensities equals the total variation norm of the cumulative intensities, we can conclude that $V_j(t;\phi,\mu)$ is continuous with respect to (ϕ,μ). Because of the linear representation (2.5), the mapping $W(t,v,p;\phi,\mu)$ is continuous in (ϕ,μ), as well. From the boundedness and continuity of $W(t,v,p;\cdot,\cdot)$ on M and the compactness and completeness of M we can finally conclude that there exists a maximizing argument in M. □

Proposition 3.2 gives existence of a worst-case reserve but we do not say anything about uniqueness and existence of a solution to the system of differential equations given by (3.2). What type of solution we can possibly expect crucially depends on the set M. We do not dig further into this question in this exposition.

In some examples later on we will construct sets M by specifying the t-slices, and the following lemma will help that we obtain sets that are compact in $L_1^{1+|\mathcal{J}|}$.

Lemma 3.3. *Let B,S be subsets of $L_1^{1+|\mathcal{J}|}([t_0,T])$ where B is a bounded and closed set and*

$$S := \Big\{ f \in L_1^{1+|\mathcal{J}|}([t_0,T]) : \text{ there exists a version of } f \text{ with variation norm not greater than } C \Big\} \tag{3.7}$$

for some $C<\infty$. Then the closure of $B\cap S$ is a compact subset of B.

Proof. Because of Tychonoff's theorem, it suffices to show the lemma just for the space $L_1([t_0,T])$. We let $h\in I\!R$. Using that all elements in S have a version that has finite variation, we can, with the help of Fubini's theorem, show that

$$\begin{aligned}\sup_{f\in B\cap S}\int_{t_0}^{T}|f(t+h)\mathbf{1}_{t+h\le T}-f(t)|dt &\le \sup_{f\in B\cap S}\int_{t_0}^{T}\int_{[t,t+h]}d|f|(s)dt\\ &\le \sup_{f\in B\cap S}\int_{[t_0,T+h]}\int_{s-h}^{s}dt\,d|f|(s)\\ &\le \sup_{f\in B\cap S} Ch,\end{aligned}$$

where $f(t):=0$ for t outside of $[t_0,T]$ and where $|f|:=f_+ + f_-$ for minimal non-decreasing functions f_+,f_- with $f=f_+-f_-$. Since Ch converges to zero for $h\to 0$ (uniformly in f), and since B is bounded, from the Kolmogorov-Riesz theorem we can conclude that $B\cap S$ is pre-compact. Hence, the closure of $B\cap S$ is compact. Since B is closed, the closure of $B\cap S$ must be a subset of B. □

In the following, we approach the differential equations by numerical methods. The existence of a worst-case reserve given by Proposition 3.2 together with convergence in various numerical calculations indicates that we do approximate a worst-case reserve.

3.3 Results for statewise reserves

We have throughout the section worked with the probability weighted reserve W. In order to prepare for the numerical calculations, we reformulate the verification theorem in the following corollary in terms of the statewise reserves V_j. The result follows directly from Proposition 3.1 and (2.6).

Corollary 3.4. *Let the assumptions of Proposition 3.1 be fulfilled, and let the ordinary differential equation system*

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\bar{V}_k(t) &= -\, b_k(t) + \bar{V}_k(t)\bar{\phi}(t) - \sum_{l\in\mathcal{S}:l\neq k} \big(b_{kl}(t) + \bar{V}_l(t) - \bar{V}_k(t)\big)\, \bar{\mu}_{kl}(t), \quad \bar{V}_k(T) = 0,\\
\frac{\mathrm{d}}{\mathrm{d}t}\bar{v}(t) &= -\, \bar{v}(t)\bar{\phi}(t), \quad \bar{v}(t_0) = 1,\\
\frac{\mathrm{d}}{\mathrm{d}t}\bar{p}(t) &= -\, \bar{\mu}^{tr}(t)\bar{p}(t), \quad \bar{p}(t_0) = \pi,\\
\big(\bar{\phi}(t), \bar{\mu}(t)\big) &= \underset{(f,m)\in M(t)}{\operatorname{argmax}} \Big\{ -f \sum_{k\in\mathcal{S}} \bar{p}_k(t)\bar{V}_k(t)\\
&\qquad + \sum_{k\in\mathcal{S}} \bar{p}_k(t) \sum_{l\in\mathcal{S}:l\neq k} m_{kl}\,\big(b_{kl}(t) + \bar{V}_l(t) - \bar{V}_k(t)\big) \Big\}
\end{aligned}
\tag{3.8}
$$

have a unique solution $\bar{V} = \big(\bar{V}_1, \ldots, \bar{V}_n\big)^{tr}$*. Then*

$$
\bar{v}(t)\bar{p}^{tr}(t)\bar{V}(t) = \sup_{(\phi,\mu)\in M} \bar{v}(t) \sum_{j\in\mathcal{S}} \bar{p}_j(t) V_j(t;\phi,\mu), \tag{3.9}
$$

in particular

$$
\pi^{tr}\bar{V}(t_0) = \sup_{(\phi,\mu)\in M} \sum_{j\in\mathcal{S}} \pi_j V_j(t_0;\phi,\mu). \tag{3.10}
$$

By solving the system (3.8) we are also able to calculate $\bar{W}$ and $\bar{V}$. Note, we have implicitly assumed that the interest rate and the intensities are not allowed to depend on the current state of the Markov chain. We do not know $\bar{p}(T)$ and $\bar{v}(T)$ but if we make a guess of the values $\bar{p}(T)$ and $\bar{v}(T)$ (this is a guess of dimension $(n+1)$ in total) and the guess is correct, we get that $\bar{p}(t_0) = \pi$ and $\bar{v}(t_0) = 1$. The problem is that we do not know the boundary conditions for all the differential equations at the same time point. This implies that we cannot use standard iterative, numerical methods to solve the differential equations. We now outline two methods, which can be used to obtain results numerically.

Shooting method: One way to overcome problems with boundary conditions at different time points is to apply the shooting method, see e.g. Orava and Lautala (1976). Here, we need to guess terminal conditions for $\bar{v}$ and $\bar{p}$ and then "shoot" until we hit the "right" starting values for $\bar{v}$ and $\bar{p}$. The shooting method is a method aiming at updating these guesses in order to obtain the right starting values as fast as possible. The system of ordinary differential equations given by (3.8) consists in total of $2n+1$ equations; $n+1$ forward equations and n backward equations. The standard way to choose whether to guess for the missing starting or terminal values is to make a guess of the lowest possible dimension. In the present case, this approach implies that we should guess the starting values for V_j and solve the equation system forward. Note however, that the differences between the two choices for the present case are minimal. Assuming that we make initial guesses $x_{T,0}$ and $y_{T,0}$ for the terminal values of $\bar{p}(T)$ and $\bar{v}(T)$ we obtain results $\bar{p}^{x_{T,0}}(t_0)$ and $\bar{v}^{y_{T,0}}(t_0)$, which we hope are close to π and 1, respectively. One should of course choose $x_{T,0}$ in the set $[0,1]^n$ and $y_{T,0}$ in the set $[0,1]$ since they are transition probabilities and a discount factor. Now the aim is to find roots for the function f given by

$$
f(x,y) = \begin{pmatrix} \pi \\ 1 \end{pmatrix} - \begin{pmatrix} \bar{p}^{x}(t_0) \\ \bar{v}^{y}(t_0) \end{pmatrix}.
$$

We can find these roots by choosing some starting values and apply a standard algorithm like Newton's Method to update the guesses. Hereby, we obtain a series of terminal conditions $(x_{T,i}, y_{T,i})$, $i = 0, 1, \ldots$. We stop the algorithm, when each entry of $f(x_{T,i}, y_{T,i})$ is below a given tolerance level ϵ.

Fixed point equation method: Another numerical method that can be used to solve the system of differential equations is the "fixed point equation method", see e.g. Bailey et al. (1968). They study problems a bit different from ours but the iteration idea is the same. They denote the method "Picard iteration" instead of "fixed-point equation method". We note that this method is the one that has been used to obtain the numerical results in Section 5. The method is an iterative algorithm, which aims at solving the equation system (3.8) within a given tolerance level. The approach is to apply the following algorithm:

1. Choose a reasonable starting interest rate and transition intensities ($\bar{\phi}^0$ and $\bar{\mu}^0$).
2. Solve the equations for $\bar{v}$ and $\bar{p}$ (forwards) using $\bar{\phi}^0$ and $\bar{\mu}^0$ and denote the solutions $\bar{v}^0$ and $\bar{p}^0$.
3. Solve the system of equations for $\bar{V}_j$ and $\left(\bar{\phi}, \bar{\mu}\right)$ (backwards) using the values obtained in the former steps. We denote the solutions $\bar{V}_j^0$ and $\left(\bar{\phi}^1, \bar{\mu}^1\right)$
4. Repeat step 2 and 3 (and increase the numbers of the superscripts accordingly) $\bar{i}$ times, where $\bar{i}$ is defined by

$$\bar{i} = \operatorname*{argmin}_{i \in I\!N} \sup_{t \in [t_0, T]} \max \Bigg\{ \max_{j \in \mathcal{S}} \left\{ \bar{V}_j^i(t) - \bar{V}_j^{i-1}(t) \right\}, \max_{j \in \mathcal{S}} \left\{ \bar{p}_j^i(t) - \bar{p}_j^{i-1}(t) \right\}, \bar{v}^i(t) - \bar{v}^{i-1}(t),$$
$$\bar{\phi}^{i+1}(t) - \bar{\phi}^i(t), \max_{(j,k) \in \mathcal{J}} \left\{ \bar{\mu}_{jk}^{i+1}(t) - \bar{\mu}_{jk}^i(t) \right\} \Bigg\} < \epsilon,$$

where ϵ is a given tolerance level. The fixed-point equation method is not necessarily converging, unless we make further model restrictions; in particular assumptions on the set M in (3.8). However, once we have found a sequence $(\bar{V}^i, \bar{p}^i, \bar{v}^i)_{i \in I\!N}$ that converges pointwise to a limit $(\bar{V}^*, \bar{p}^*, \bar{v}^*)$, we can, under mild conditions, conclude that the latter limit is a solution to (3.8): Given that the assumptions of Proposition 3.2 hold and that the argmax in (3.8) is continuous with respect to the parameters $\bar{V}^i(t)$ and $\bar{p}^i(t)$ at $\bar{V}^*(t)$ and $\bar{p}^*(t)$, we can show that $(\bar{V}^*(t), \bar{p}^*(t), \bar{v}^*(t))$ solves the integral version of (3.8). This is shown by applying the dominated convergence theorem and using the constant C in the proof of Proposition 3.2 as a uniform majorant for the functions $\bar{V}_j^i(t)$.

Remark 3.5. *If we assume that the* argmax *in (3.8) does not depend on any of the factors $\bar{p}_k(t)$ for all $k \in \mathcal{S}$ things get numerically simpler. Unfortunately, this only holds for a very limited type of models like a multiple causes of death model, see Christiansen and Steffensen (2013). The numerical simplification is that the calculation of $\bar{p}$ and $\bar{\mu}$ decouples from the calculation of $\bar{V}_k$. That is, one can first solve the ODEs for $\bar{V}_k$ with* $\bar{v}(t)\bar{p}(t) = 1$ *(backwards) to obtain values for $(\bar{\mu}, \bar{\phi})$, secondly use these results to calculate $(\bar{p}, \bar{v})$ (forwards), and lastly use the results of $(\bar{\mu}, \bar{\phi})$ and $(\bar{p}, \bar{v})$ to calculate $\bar{V}_k$ (backwards). In this case the use of the shooting method is* not *necessary. If the* argmax *in (3.8) is dependent on but constant in the factors $\bar{p}_k(t)$ one can do exactly the same as before by basing the first values $\bar{v}$ and $\bar{p}$ on some arbitrary values of μ and ϕ in M. These simplifications are for a fixed starting state covered by Christiansen and Steffensen (2013).*

3.4 Some notes about the set M

One crucial ingredient of the calculations in the present paper is the set M, over which we maximize the probability weighted reserve. To make the calculations useful for a company they need to know certain probabilistic properties of the set, e.g. the confidence level. One non-statistical way of obtaining a set is to use an expert opinion. However, it is hard to deduce any properties from sets based on an expert opinion.

An alternative and more attractive method is to deduce a set with a certain coverage from a stochastic model by either analytical or numerical methods. One could ask for the advantages of this approach compared to just simulating reserves and finding confidence intervals numerically. The answer to this is twofold. First, you gain some insight about the nature of your risk by calculating the sets. Second, for many life insurance companies this method will be computationally faster and for some easier to implement. The speedup relative to directly simulating the reserves occurs because you only need to solve the system of differential equations characterizing the reserve once, after having found the set. The alternative is to simulate scenarios and solve differential equations over and over again to get reliable results. This is particularly important when making calculations for complex products modeled in Markov chains with many states in which case the differential equations take a long time to solve. An additional advantage comes from the possibility of using the same calculated sets for different products (for the same policyholder) and for different policyholders, who have similar characteristics. We do not think that it is a big drawback to use a parametric approach, since many companies already deal with estimation in such models when e.g. forecasting the future mortality.

4 Worst-case calculations for portfolios

In this section we consider an *inhomogeneous* portfolio, whereby allowing for nonidentical distributions of the jump processes modelling the policyholders. We generally assume that the policyholders are *stochastically independent* and that the force of interest is the same for all contracts. Given this independence assumption and assuming that there exists an intensity matrix for each policyholder, we can conclude that two or more policyholders change state at the same point in time with probability zero.

If the same set M is used for all policyholders in the portfolio but no interactions between the worst cases of different contracts are taken into account, we can easily obtain the worst-case reserve for the portfolio by calculating the worst-case reserve for each policyholder separately and then summing over all the reserves. In practice, we think the mortalities across the population are closely connected (dependent), so this is the case we want to study. Such a study is not doable within the theory of Christiansen and Steffensen (2013), because the requirements of that paper are not fulfilled in this case. This is because dependence between policyholders in a portfolio implies that the argmax in Christiansen and Steffensen (2013, Proposition 4.1) is not constant with respect to all the discounted transition probabilities; see also Remark 3.5. The simple approach outlined above leads to a rough upper bound only, for the worst-case in a portfolio. We illustrate this with a numerical example in Section 5.

There are a few simple cases of inhomogeneous portfolios, where we can confine ourselves to the theory of Section 3. One of the cases, which was shortly mentioned in the beginning of Section 3, is where the policyholders are governed by the same intensities and have the same insurance product but have different starting distributions. This case is covered because it is equivalent to calculation of the worst-case scenario for a single policy with a specific starting distribution. Another simple case is where the policyholders of the portfolios are governed by the same intensities and starting distributions but have different insurance products. In this case, the worst-case reserve of the portfolio can be found as the worst-case reserve of a single policy, where the policyholder has the sum of the products of all the policyholders in the portfolio. However, in general we need some extended results to cover portfolios.

4.1 A representation for the worst-case reserve

The aim of this subsection is to find a simple representation for the worst-case reserve for a portfolio of policyholders. To do so, we start by introducing some additional notation for the setting with multiple policyholders. For the terms where the notation from Section 3 is applicable, we do not repeat these.

- L: Number of policyholders in the portfolio.
- $(\mathbf{X}(t))_{t\in[0,T]} = (X_1(t), \ldots, X_L(t))_{t\in[0,T]}$: Representation of the state of the policyholders of the portfolio.
- $\boldsymbol{\mathcal{S}} = \{(\mathcal{S}_1, \ldots, \mathcal{S}_L) | \mathcal{S}_i \subset \mathcal{S}\}$: State space for all policies.
- $\boldsymbol{\mathcal{J}} = \Big\{(\mathbf{j}, \mathbf{k}) \in \boldsymbol{\mathcal{S}} \times \boldsymbol{\mathcal{S}} | \, \mathbf{j} = (j_1, \ldots, j_l, \ldots, j_L), \mathbf{k} = (j_1, \ldots, \tilde{j}_l, \ldots, j_L), j_l \neq \tilde{j}_l, \; l \in \{1, ..., L\}\Big\}$:
  Transition space for all policies. The condition in the above formula means, that the vectors $\mathbf{j}$ and $\mathbf{k}$ differ in exactly one coordinate.
- $V_i^l(t)$: Statewise reserve at time t for policyholder l in state i.
- $b_{jk}^l(t)$ and $b_j^l(t)$: Lump sum and continuous payments at time t for policyholder l.
- $\mu_{jk}^l(t)$: Transition intensities at time t for policyholder l.
- $p_{jk}^l(s, t)$: Transition probabilities between time t and s for policyholder l.

In the following, we show that the "big model" collapses to something simpler. By "big model" we mean that the portfolio is modeled as a single policy, such that we can use the theory of Section 3. In the "big model", we have $|\mathcal{S}_1| \cdot \ldots \cdot |\mathcal{S}_L|$ states that cover all the different combinations of policyholders and states. By defining

$$b_{\mathbf{j}}(t) = \sum_{l=1}^{L} b_{j_l}^l(t) \text{ and } b_{\mathbf{jk}}(t) = b_{j_l \tilde{j}_l}^l(t)$$

for $\mathbf{j} = (j_1, \ldots, j_L)$ and $\mathbf{k} = (j_1, \ldots, \tilde{j}_l, \ldots, j_L)$ (recall that two or more jumps at the same time occur with probability zero), we get that the present portfolio value $\mathbf{B}(t)$ equals the sum of the single present values,

$$
\begin{aligned}
\mathbf{B}(t) =& \sum_{\mathbf{j}\in\boldsymbol{\mathcal{S}}} \int_t^T v(t,s)\mathbf{1}_{\{\mathbf{X}(s)=\mathbf{j}\}} b_{\mathbf{j}}(s)ds + \sum_{(\mathbf{j},\mathbf{k})\in\boldsymbol{\mathcal{J}}} \int_t^T v(t,s) b_{\mathbf{jk}}(s)dN_{\mathbf{jk}}(s) \\
=& \sum_{\mathbf{j}\in\boldsymbol{\mathcal{S}}} \int_t^T v(t,s)\mathbf{1}_{\{\mathbf{X}(s)=\mathbf{j}\}} \sum_{l=1}^{L} b^l_{j_l}(s)ds + \sum_{(\mathbf{j},\mathbf{k})\in\boldsymbol{\mathcal{J}}} \int_t^T v(t,s) b^l_{j_l\tilde{j}_l}(s)dN^l_{j_l\tilde{j}_l}(s) \\
=& \sum_{l=1}^{L} \sum_{j_1\in\mathcal{S}_1,\dots,j_L\in\mathcal{S}_L} \int_t^T v(t,s) \prod_{k=1}^{L} \mathbf{1}_{\{X_k(s)=j_k\}} b^l_{j_l}(s)ds \\
&+ \sum_{l=1}^{L} \sum_{(j,k)\in\mathcal{J}_l} \int_t^T v(t,s) b^l_{jk}(s)dN^l_{jk}(s) \\
=& \sum_{l=1}^{L} \left(\sum_{j\in\mathcal{S}_l} \int_t^T v(t,s)\mathbf{1}_{\{X_l(s)=j\}} b^l_j(s)ds + \sum_{(j,k)\in\mathcal{J}_l} \int_t^T v(t,s) b^l_{jk}(s)dN^l_{jk}(s) \right) \\
=& \sum_{l=1}^{L} B^l(t),
\end{aligned}
\tag{4.1}
$$

where B^l is the present value of future payments for policyholder l, and N^l_{jk} counts the number of jumps from state j to state k for policyholder l.

We have that the reserve for each policyholder follows a differential equation of the type (2.4) and that the transition probabilities, p^l_j, follow the Kolmogorov equations given by (2.2). We now want to find the worst-case scenario for the portfolio. For this purpose, we define a mapping $\mathbf{W} : [t_0, T] \times (0, \infty) \times [0, 1]^{|\boldsymbol{\mathcal{S}}|} \to \mathbb{R}$ by

$$
\mathbf{W}(t, v, \mathbf{p}) \quad = \quad v \sum_{\mathbf{j}\in\boldsymbol{\mathcal{S}}} \mathbf{p}_{\mathbf{j}} \mathrm{E}\left[\mathbf{B}(t)|\mathbf{X}(t) = \mathbf{j}\right].
$$

That is, we want to find the worst-case scenario for

$$
\begin{aligned}
\mathbf{W}(t, v(t), \mathbf{p}(t)) \quad &= \quad v(t) \sum_{\mathbf{j}\in\boldsymbol{\mathcal{S}}} \pi_{\mathbf{j}} \mathrm{E}\left[\mathbf{B}(t)|\mathbf{X}(t_0) = \mathbf{j}\right] \\
&= \quad v(t) \sum_{\mathbf{j}\in\boldsymbol{\mathcal{S}}} \mathbf{p}_{\mathbf{j}}(t) \mathrm{E}\left[\mathbf{B}(t)|\mathbf{X}(t) = \mathbf{j}\right],
\end{aligned}
$$

where $\mathbf{p}(t) = (P(\mathbf{X}(t) = \mathbf{j}))_{\mathbf{j}\in\boldsymbol{\mathcal{S}}}$ and $\pi_{\mathbf{j}} = \pi^1_{j_1} \cdot \ldots \cdot \pi^L_{j_L}$ because of the independence assumption. Using again the stochastic independence of the policyholders and the additive structure of (4.1), we can show that

$$
\begin{aligned}
\mathbf{W}(t, v(t), p(t)) \quad &= \quad v(t) \sum_{l=1}^{L} \sum_{j\in\mathcal{S}_l} \pi^l_j \mathrm{E}\left[B^l(t)|X_l(t_0) = j\right] \\
&= \quad \sum_{l=1}^{L} \sum_{j\in\mathcal{S}_l} \int_t^T p^l_j(u)v(u) \left(b^l_j(u) + \sum_{k\in\mathcal{S}_l:k\neq j} b^l_{jk}(u)\mu^l_{jk}(u) \right) du.
\end{aligned}
$$

We aim to find

$$
\sup_{(\phi,\mu^1,\dots,\mu^L)\in M} \mathbf{W}(t_0, v(t_0), \mathbf{p}(t_0); \phi, \mu^1, \dots, \mu^L), \tag{4.2}
$$

where $M \subset L_1^{1+|\mathcal{J}|}([t_0, T])$ is a set of integrable interest rate and transition intensity paths. The starting point t_0 is the same for all policyholders of the portfolio and should therefore be thought of as a calendar time point rather than an age in case of a heterogeneous portfolio. Considering this portfolio as a single contract on the state space $\boldsymbol{\mathcal{S}}$, one can apply the theory of Section 3 and obtain the desired results. Because of the huge state space $\boldsymbol{\mathcal{S}}$, the computational workload seems enormous. However, we can simplify the formulas in (3.8).

Proposition 4.1. *For the "big model" the solutions of the differential equation system given by (3.8) are equivalent to the solutions of*

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\bar{V}_j^l(t) =& -b_j^l(t) + \bar{V}_j^l(t)\bar{\phi}(t) - \sum_{k\in\mathcal{S}_l:k\neq j}\left(b_{jk}^l(t) + \bar{V}_k^l(t) - \bar{V}_j^l(t)\right)\bar{\mu}_{jk}^l(t),\\
&\bar{V}_j^l(T) = 0, \quad l = 1,\dots,L,\\
\frac{\mathrm{d}}{\mathrm{d}t}\bar{v}(t) =& -\bar{v}(t)\bar{\phi}(t), \quad \bar{\phi}(t_0) = 1,\\
\frac{\mathrm{d}}{\mathrm{d}t}\bar{p}^l(t) =& -\left(\bar{\mu}^l(t)\right)^{tr}\bar{p}^l(t), \quad \bar{p}(t_0) = \pi^l, \quad l = 1,\dots,L,\\
(\bar{\phi}(t), \bar{\mu}^1(t),\dots,\bar{\mu}^L(t)) =& \underset{(f,m^1,\dots,m^L)\in M(t)}{\operatorname{argmax}}\Bigg\{ - f\sum_{l=1}^{L}\sum_{j_l\in\mathcal{S}_l}\bar{p}_{j_l}^l(t)\bar{V}_{j_l}^l(t) + \sum_{l=1}^{L}\sum_{(j_l,\tilde{j}_l)\in\mathcal{J}_l}\bar{p}_{j_l}^l(t)m_{j_l\tilde{j}_l}^l\\
&\times\left(b_{j_l\tilde{j}_l}^l(t) + \bar{V}_{\tilde{j}_l}^l(t) - \bar{V}_{j_l}^l(t)\right)\Bigg\}.
\end{aligned} \tag{4.3}
$$

Proof. We start by considering the argmax given in (3.2) for the entire portfolio. The interior of the argmax is given by

$$
\begin{aligned}
&- fv\frac{\partial}{\partial v}\bar{\mathbf{W}} + \sum_{\mathbf{j}\in\boldsymbol{\mathcal{S}}}\mathbf{p_j}\left(vb_{\mathbf{j}} + \sum_{\mathbf{k}\in\boldsymbol{\mathcal{S}}:\mathbf{k}\neq\mathbf{j}} m_{\mathbf{jk}}\left(vb_{\mathbf{jk}} + \nabla_{\mathbf{p_k}}\bar{\mathbf{W}} - \nabla_{\mathbf{p_j}}\bar{\mathbf{W}}\right)\right)\\
=& - fv\frac{\partial}{\partial v}\bar{\mathbf{W}} + \sum_{\mathbf{j}\in\boldsymbol{\mathcal{S}}}\mathbf{p_j}vb_{\mathbf{j}} + \sum_{(\mathbf{j},\mathbf{k})\in\boldsymbol{\mathcal{J}}}\mathbf{p_j}m_{\mathbf{jk}}\left(vb_{\mathbf{jk}} + \nabla_{\mathbf{p_k}}\bar{\mathbf{W}} - \nabla_{\mathbf{p_j}}\bar{\mathbf{W}}\right).
\end{aligned} \tag{4.4}
$$

In (4.4), $\mathbf{j}$ and $\mathbf{k}$ are vectors of dimension L. We use that the lives of the policyholders are independent conditional on the intensities and assume that $\mathbf{j} = (j_1,\dots,j_l,\dots j_L)^{\mathrm{tr}}$ and $\mathbf{k} = (j_1,\dots,\tilde{j}_l,\dots,j_L)^{\mathrm{tr}}$ with $j_l \neq \tilde{j}_l$. This means that we obtain:

$$
\begin{aligned}
(4.4) =& - fv\sum_{l=1}^{L}\sum_{j_l\in\mathcal{S}_l}p_{j_l}^l V_{j_l}^l + v\sum_{l=1}^{L}\sum_{j_l\in\mathcal{S}_l}p_{j_l}^l b_{j_l}^l\\
&+\sum_{l=1}^{L}\sum_{(j_l,\tilde{j}_l)\in\mathcal{J}_l}p_{j_l}^l m_{j_l\tilde{j}_l}v\left(b_{j_l\tilde{j}_l}^l + V_{j_1}^1 + \dots + V_{\tilde{j}_l}^l + \dots + V_{j_L}^L - \sum_{i=1}^{L}V_{j_i}^i\right)\\
=& - fv\sum_{l=1}^{L}\sum_{j_l\in\mathcal{S}_l}p_{j_l}^l V_{j_l}^l + \sum_{l=1}^{L}\sum_{j_l\in\mathcal{S}_l}vp_{j_l}^l b_{j_l}^l + \sum_{l=1}^{L}\sum_{(j_l,\tilde{j}_l)\in\mathcal{J}_l}p_{j_l}^l m_{j_l\tilde{j}_l}^l v\left(b_{j_l\tilde{j}_l}^l + V_{\tilde{j}_l}^l - V_{j_l}^l\right).
\end{aligned} \tag{4.5}
$$

We see that (4.5) splits into components relating to each of the policyholders and that the argmax does not depend on the second of the three terms. That is, we have proved the proposition. □

In the following two subsections, we show examples of worst-case scenarios for specific models of a portfolio.

4.2 Example 1: Solvency II

In this subsection we illustrate how the general theory above is used to generate a mortality stress scenario for a portfolio of contracts. This is a way of constructing a stress scenario as for instance the mortality stress in Solvency II. Depending on whether the portfolio is a specific portfolio or a stylised market portfolio the generated scenario is internally based or input to a standard calculation, respectively.

We consider a simple two-state life-death model and assume that there is no payments in the state "dead". We are maximizing the reserve with respect to compact sets of the form

$$M = \overline{S \cap B},$$

where S is given by (3.7) and B is defined via its slices

$$B(t) = \Big\{ \left(\phi(t), \mu^1(t), \dots, \mu^L(t)\right) \in \mathbb{R}_+^{L+1} \Big| \; \phi(t) \in \Phi(t), \mu^1(t) = \hat{\mu}^1(t)\alpha(t), \dots, \mu^L(t) = \hat{\mu}^L(t)\alpha(t) \Big\}.$$

We assume that $\alpha(t) \in [\alpha_l(t), \alpha_h(t)]$ and $\Phi(t) = [\phi_l(t), \phi_h(t)]$ and that $\alpha_l, \alpha_h, \phi_l, \phi_h$ and $\hat{\mu}^i$ are bounded functions. Note that M is compact in $L_1^{1+|J|}$ by Lemma 3.3. Assuming that the L policyholders have different ages $x_1, \dots, x_L$ at the current time, a natural choice would be to model $\hat{\mu}^i$ of the form

$$\hat{\mu}^i(t) = \mu^{\text{be}}(x_i + t)\Lambda(x_i, t),$$

where μ^{be} is a best estimate mortality intensity and Λ is a longevity factor meaning that Λ is decreasing in time. In this setup the interest rate is independent of the mortality intensities and the mortality intensities are linearly dependent.

The argmax in (4.3) becomes

$$\begin{aligned}
&(\bar{\phi}(t), \bar{\mu}^1(t), \dots, \bar{\mu}^L(t)) \\
&= \underset{(f, m^1, \dots, m^L) \in M(t)}{\operatorname{argmax}} \left\{ -f \sum_{l=1}^{L} \bar{p}_a^l(t)\bar{V}_a^l(t) + \sum_{l=1}^{L} \bar{p}_a^l(t) m^l \left(b_{ad}^l(t) - \bar{V}_a^l(t)\right) \right\}.
\end{aligned}$$

Because of linearity, this argmax can be found by calculating

$$\underset{(f,\alpha) \in (\Phi(t) \times [\alpha_l(t), \alpha_h(t)])}{\operatorname{argmax}} \left\{ -f \sum_{l=1}^{L} \bar{p}_a^l(t)\bar{V}_a^l(t) + \alpha \sum_{l=1}^{L} \bar{p}_a^l(t)\hat{\mu}^l(t) \left(b_{ad}^l(t) - \bar{V}_a^l(t)\right) \right\}$$

and multiplying $\hat{\mu}^1, \dots, \hat{\mu}^L$ with α.

We can also consider a portfolio of disability contracts, see Figure 1. That is, we study contracts which can be modeled within a three state Markov chain where recovery from

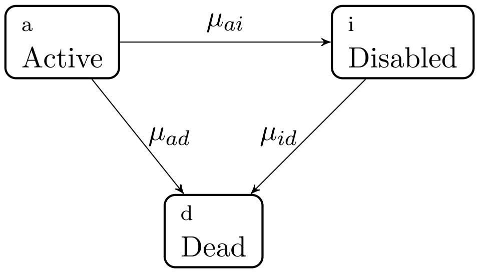

Figure 1: Disability model without recovery.

"disabled" to "active" is not possible. We assume no payments in the state "dead" and that μ_{ai} is fixed in the sense that it is a deterministic function which we are not maximizing the reserve with respect to. We want to find the worst-case reserve of this portfolio with respect to ϕ, μ_{ad} and μ_{id} with respect to the sets of the form

$$M = \overline{S \cap B},$$

where S is given by (3.7) and B is defined via its slices

$$\begin{aligned} B(t) = \Big\{ \, &\big(\phi(t), \mu_{ad}^1(t), \mu_{id}^1(t), \dots, \mu_{ad}^L(t), \mu_{id}^L(t)\big) \in I\!R_+^{2L+1} \Big| \, \phi(t) \in \Phi(t), \\ &\mu_{ad}^1(t) = \hat{\mu}_{ad}^1(t)\alpha(t), \mu_{id}^1(t) = \hat{\mu}_{id}^1(t)\beta(t), \dots, \\ &\mu_{ad}^L(t) = \hat{\mu}_{ad}^L(t)\alpha(t), \mu_{id}^L(t) = \hat{\mu}_{id}^L(t)\beta(t)\Big\}. \end{aligned}$$

We assume that $\Phi(t)$ is defined as above and that $\alpha(t) \in [\alpha_l(t), \alpha_h(t)]$ and $\beta(t) \in [\beta_l(t), \beta_h(t)]$ for bounded functions $\alpha_l, \alpha_h, \beta_l, \beta_h, \hat{\mu}_{ad}^i$ and $\hat{\mu}_{id}^i$. Note that M is compact in $L_1^{1+|J|}$ by Lemma 3.3.

Because of linearity, we can obtain $(\bar{\phi}(t), \bar{\mu}^1(t), \dots, \bar{\mu}^L(t))$ by calculating

$$\begin{aligned} \underset{(f,\alpha,\beta)\in(\Phi(t)\times[\alpha_l(t),\alpha_h(t)]\times[\beta_l(t),\beta_h(t)])}{\operatorname{argmax}} \Bigg\{ &- f \sum_{l=1}^{L} \Big(\bar{p}_a^l(t)\bar{V}_a^l(t) + \bar{p}_i^l(t)\bar{V}_i^l(t)\Big) \\ &+ \alpha \sum_{l=1}^{L} \bar{p}_a^l(t)\hat{\mu}_{ad}^l(t)\left(b_{ad}^l(t) - \bar{V}_a^l(t)\right) \\ &+ \beta \sum_{l=1}^{L} \bar{p}_i^l(t)\hat{\mu}_{id}^l(t)\left(b_{id}^l(t) - \bar{V}_i^l(t)\right) \Bigg\} \end{aligned}$$

and multiplying $\hat{\mu}_{ad}^1, \dots, \hat{\mu}_{ad}^L$ with α and $\hat{\mu}_{id}^1, \dots, \hat{\mu}_{id}^L$ with β.

In the above calculations the death and disability intensities are independent. Another possibility could be to make μ_{ad} and μ_{id} dependent. This is exactly what is described in Section 4.3. In the case $\alpha_l = \alpha_h$, we optimize over a singleton with respect to μ_{ad} for each time point and the argmax becomes trivial.

4.3 Example 2: Dependent version of the Solvency II example

This example is an extension of the result in Section 4.2 where we include sets of the form given by the case "Dependence" in Figure 2. Again, we assume no payments in the state "dead". We are maximizing the reserve with respect to sets of the form

$$M = \overline{S \cap B},$$

where S is given by (3.7) and B is defined via its slices

$$\begin{aligned} B(t) = \Big\{ & \left(\phi(t), \mu_{ad}^1(t), \mu_{id}^1(t), \ldots, \mu_{ad}^L(t), \mu_{id}^L(t)\right) \in \mathbb{R}_+^{2L+1} \Big| \\ & \phi(t) \in \Phi(t), \mu_{ad}^1(t) = \hat{\mu}_{ad}^1(t)\alpha(t), \mu_{id}^1(t) = \hat{\mu}_{id}^1(t)\beta(t), \ldots, \mu_{ad}^L(t) = \hat{\mu}_{ad}^L(t)\alpha(t), \\ & \mu_{id}^L(t) = \hat{\mu}_{id}^L(t)\beta(t), (\alpha(t), \beta(t)) \in \check{B}(t) \Big\}. \end{aligned}$$

We assume that $\Phi(t)$ is given as in Section 4.2 and $\check{B}(t)$ is given on the linear programming form

$$\check{B}(t) = \left\{ (x, y) \in \mathbb{R}_+^2 \,\middle|\, \begin{pmatrix} -\frac{3(\beta_h(t)-\beta_l(t))}{\alpha_h(t)-\alpha_l(t)} & 1 \\ \frac{3(\beta_h(t)-\beta_l(t))}{\alpha_h(t)-\alpha_l(t)} & -1 \\ -\frac{\beta_h(t)-\beta_l(t)}{3(\alpha_h(t)-\alpha_l(t))} & 1 \\ \frac{\beta_h(t)-\beta_l(t)}{3(\alpha_h(t)-\alpha_l(t))} & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} -\frac{3(\beta_h(t)-\beta_l(t))}{\alpha_h(t)-\alpha_l(t)}\alpha_l(t) + \beta_l(t) \\ \frac{3(\beta_h(t)-\beta_l(t))}{\alpha_h(t)-\alpha_l(t)}\alpha_h(t) - \beta_h(t) \\ -\frac{\beta_h(t)-\beta_l(t)}{3(\alpha_h(t)-\alpha_l(t))}\alpha_l(t) + \beta_l(t) \\ \frac{\beta_h(t)-\beta_l(t)}{3(\alpha_h(t)-\alpha_l(t))}\alpha_h(t) - \beta_h(t) \end{pmatrix} \right\}.$$

We choose the functions $\alpha_h, \alpha_l, \beta_h, \beta_l$ in such a way that the slices $\check{B}(t)$ are closed and uniformly bounded (in t). Moreover, we assume that the functions $\hat{\mu}_{ad}^i$ and $\hat{\mu}_{id}^i$ are bounded. Note that M is compact in $L_1^{1+|J|}$ by Lemma 3.3. Because of linearity, we can find the argmax $(\bar{\phi}(t), \bar{\mu}^1(t), \ldots, \bar{\mu}^L(t))$ in (4.3) by calculating

$$\begin{aligned} \underset{(f,\alpha,\beta)\in(\Phi(t)\times\check{B}(t))}{\operatorname{argmax}} \Big\{ & - f \sum_{l=1}^{L} \left(\bar{p}_a^l(t)\bar{V}_a^l(t) + \bar{p}_i^l(t)\bar{V}_i^l(t)\right) + \alpha \sum_{l=1}^{L} \bar{p}_a^l(t)\hat{\mu}_{ad}^l(t) \left(b_{ad}^l(t) - \bar{V}_a^l(t)\right) \\ & + \beta \sum_{l=1}^{L} \bar{p}_i^l(t)\hat{\mu}_{id}^l(t) \left(b_{id}^l(t) - \bar{V}_i^l(t)\right) \Big\} \end{aligned} \tag{4.6}$$

and multiplying $\hat{\mu}_{ad}^1, \ldots, \hat{\mu}_{ad}^L$ with α and $\hat{\mu}_{id}^1, \ldots, \hat{\mu}_{id}^L$ with β. To find (4.6), we must at each time point check each of the four extremal points of the set $\check{B}(t)$ combined with the extremal points of $\Phi(t)$.

5 Numerical calculations

We have performed the numerical calculations in this section by applying the "fixed point equation method" described in Section 3. In all our examples we obtained converging fixed-point sequences, and in a neighborhood of the limit the argmax in (3.8), seen as a mapping of reserves and transition probabilities, turned out to be continuous for almost all t. Taking into account the arguments in the lines before Remark 3.5, our numerical results are indeed approximations for the solutions of (3.8). First, we consider numerical calculations for a single a policy. Next, we consider similar calculations for an inhomogeneous portfolio.

5.1 Numerical calculations for a single policy

We consider the example of a simple disability policy described in Figure 1, where the payments are given by disability benefits in the state "Disabled" at a yearly rate $b_i = 1$

and lump sum payments paying out an amount of 3 upon transcription to the state "Dead" from either of the states "Active" or "Disabled". For simplicity, we assume that no premiums are paid.

In the example we consider a person at the age of 35 and contract expiry at the age of 65. We let the short rate be 2% and let both the intensity from "Active" to "Disabled" (which we consider fixed) and the best estimate death intensity be given on a Gompertz-Makeham form. The exact intensity parameters are given in Table 1. We consider the same lower

μ_{ad}^{be}	μ_{ai}
$0.0025 + 10^{5.804-10+0.038\chi}$	$0.00148 + 10^{4.97136-10+0.06\chi}$

Table 1: Best estimate intensities for a policyholder at age χ.

and upper bounds for both the active-death and the disabled-death intensities. The lower bound is given by $U(t) = 0.8\mu_{ad}^{\text{be}}(t)$ whereas the upper bound is given by $L(t) = 1.15\mu_{ad}^{\text{be}}(t)$.

In the following, we find the worst-case scenario for different sets M using numerical methods. The argmax in (3.8) can be either easy or hard to obtain depending on the form of the slices of the set M. If we can formulate the optimization problem as a linear program, which is the case for all the sets presented in Figure 2, we know that we only need to search for the argmax in the extremal points of the sets. That is, for the two cases "Independence" and "Dependence", we only need to evaluate the object function in four points, whereas for the case "Linear dependence", we only need to evaluate the object function in two points. In the case of a linear program, the extremal points are quite obvious. In the more general case of a strictly convex set $M(t)$, Christiansen and Steffensen (2013, Appendix) outlines a way of obtaining the extremal points.

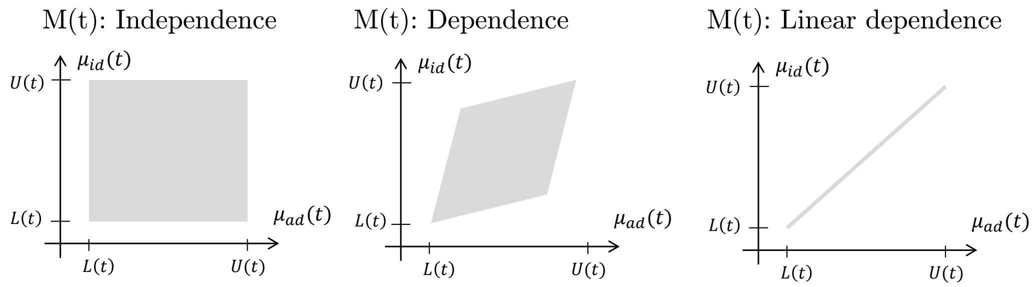

Figure 2: Three different trust regions.

The figures 3-5 show the worst-case bases (conditional on that the current state is "Active") for the three different types of sets depicted in Figure 2. In the case "Dependence", the four extremal points are $\{(L(t), L(t)), (L(t) + 0.25(U(t) - L(t)), L(t) + 0.75(U(t) - L(t))), (U(t), U(t)), (L(t) + 0.75(U(t) - L(t)), L(t) + 0.25(U(t) - L(t)))\}$. In the case of independence, the worst-case scenario is that the intensity μ_{ad} is as high as possible throughout the entire period, since the chances of getting disabled is not that high. On the other hand, the intensity μ_{id} is only high at the very last part of the period of the contract, because there are no more disability benefits after the transition. Note that a bigger relative difference between b_i and $b_{ad} = b_{id}$ would have caused the shift from low to high intensity to happen earlier.

In the case of dependence, the situation is not equally simple. Here, the tradeoff between having a high intensity μ_{ad} and a low intensity μ_{id} results in the worst-case scenario for the middle of the time span of the contract becoming $\mu_{ad}(t) = L(t) + 0.75(U(t) - L(t))$ whereas $\mu_{id}(t) = L(t) + 0.25(U(t) - L(t))$. Note that these are not corner points of the marginals in contrast to the case of independence. The worst-case scenario occurs because the level of μ_{ad} is more important for the size of the reserve than the level of μ_{id}. This is because the level of μ_{id} is a second-order effect in the state "Active", since μ_{id} only matters after transition to the state disabled. However, this second-order effect is so significant that the worst-case scenario is *not* to maximize both μ_{ad} and μ_{id}.

For the linear dependent case we, as in the dependent case, see that the impact of μ_{ad} is more significant than the impact of μ_{id} implying that both are maximized for the entire time span.

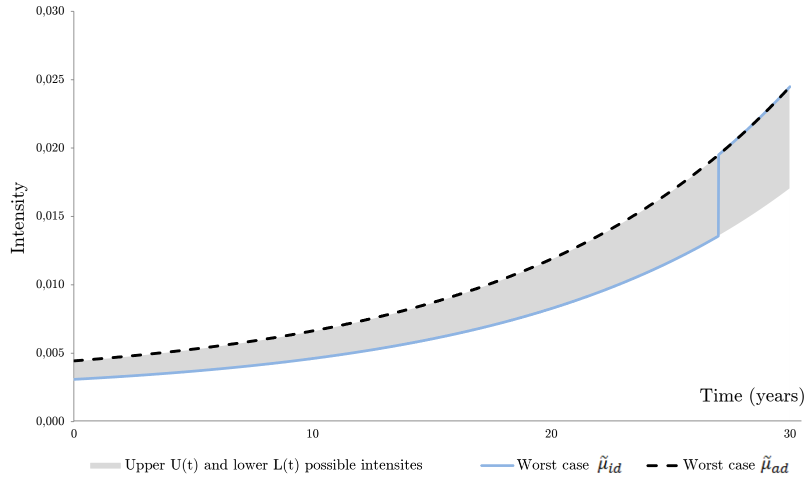

Figure 3: The worst-case death intensities in the case of independence.

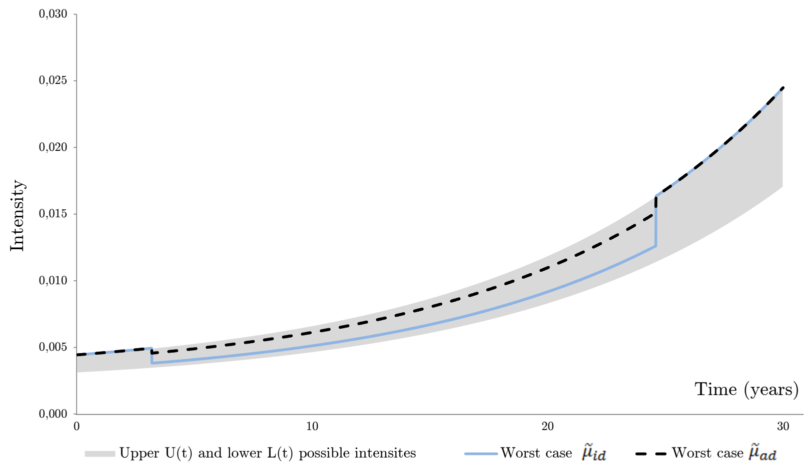

Figure 4: The worst-case death intensities in the case of dependence.

In Figure 6 we see that the convergence to the fixed point is fairly fast: After only four iterations, we have obtained convergence.

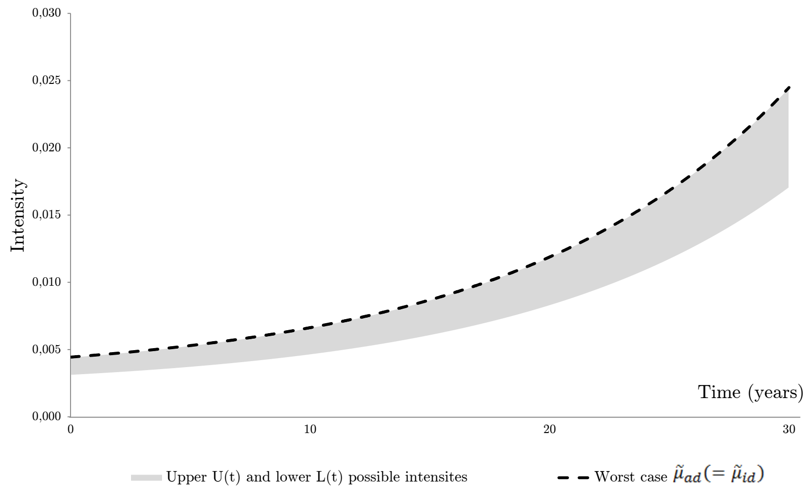

Figure 5: The worst-case death intensities in the case of linear dependence.

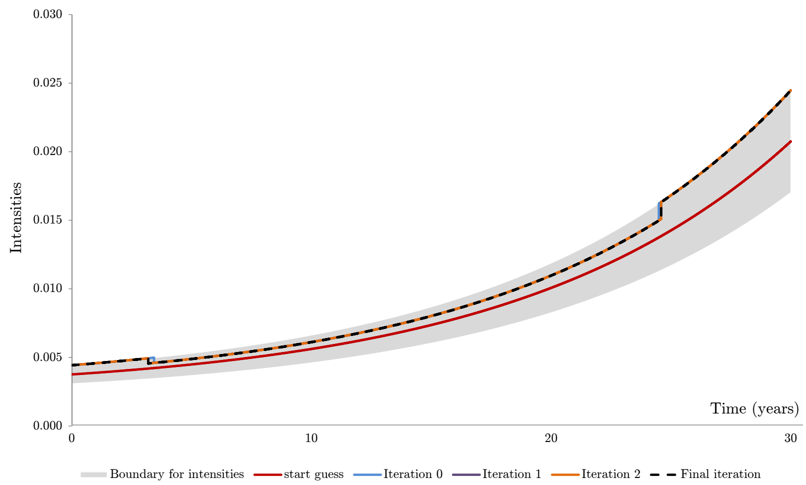

Figure 6: Convergence for the argmax of μ^{ad}.

5.2 Numerical calculations for a portfolio

In this section we illustrate the theory of Section 4 for a representative portfolio consisting of a young, a middle-aged, and a close-to-pension-aged policyholder. We are in the two state life-death model and are maximizing over the type of sets described in Section 4.2. We assume that all three representative policyholders have the same type of contract. That is, they have a term insurance paying 15 at death before retirement (age 67) and a life annuity paying a yearly rate of 1 starting at retirement. Their baseline death intensities are given as μ_{ad}^{be} in Table 1. For the present example the set of interest rates is $\Phi = \{0.02\}$, the lower multiplicative factor is $\alpha_l = 0.8$, and the upper multiplicative factor is $\alpha_h = 1.15$. The values of α_l and α_h are motivated by the mortality and longevity stresses from Solvency II, see EIOPA (2013).

We obtain the worst-case intensities illustrated in Figure 7 on a logarithmic scale. We denote by subscript 30 the youngest policyholder, by subscript 45 the middle-aged policyholder, and by subscript 60 the oldest policyholder. Moreover, we use "I" to indicate that quantities are calculated at an individual level, and "P" to indicate that quantities are calculated on portfolio level. We see from Figure 7 that the worst-case scenario for the oldest person is the same as the worst-case scenario at the portfolio. The scenario is that the intensity is as high as possible for the first seven years (until retirement of the oldest policyholder), and hereafter it is as low as possible. On the other hand, the worst case-scenarios for the two other policyholders are quite different compared to the worst-case scenario for the portfolio. The statewise worst-case reserves for the policyholders corresponding to the intensities in Figure 7 can be found in Figure 8.

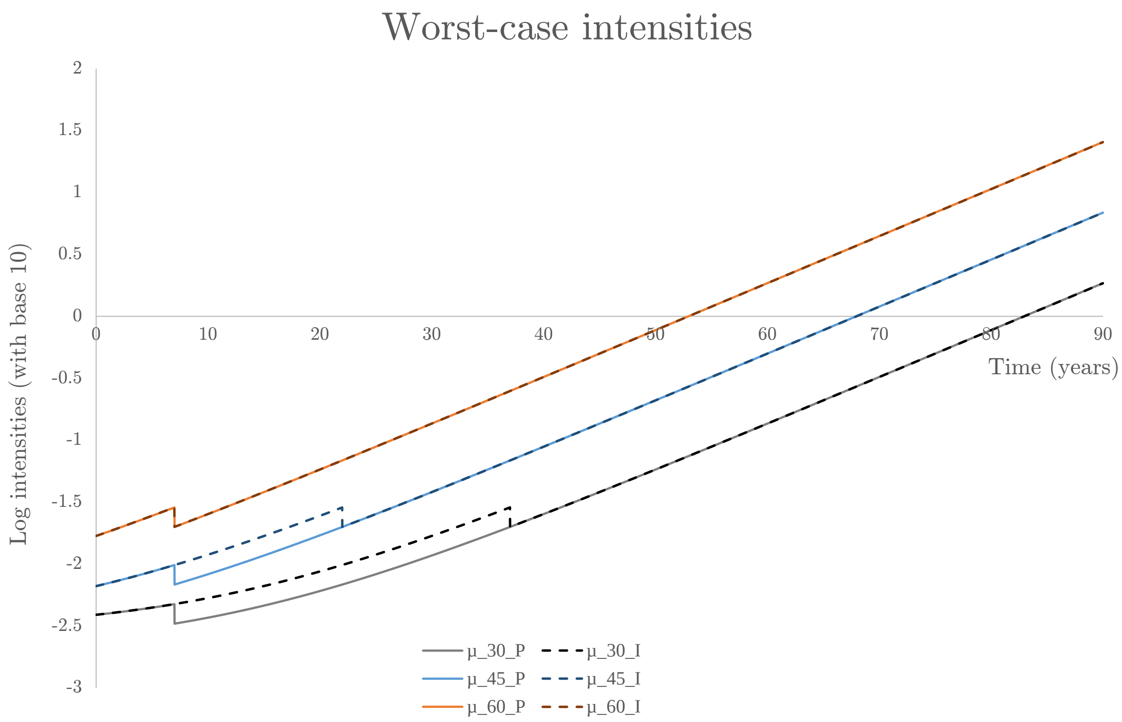

Figure 7: Worst-case intensities for the portfolio and for each individual policyholder.

In Table 2 we compare the reserves for the three policyholders in the portfolio calculated with different bases. The first is the best estimate basis, the second is α_h times the best

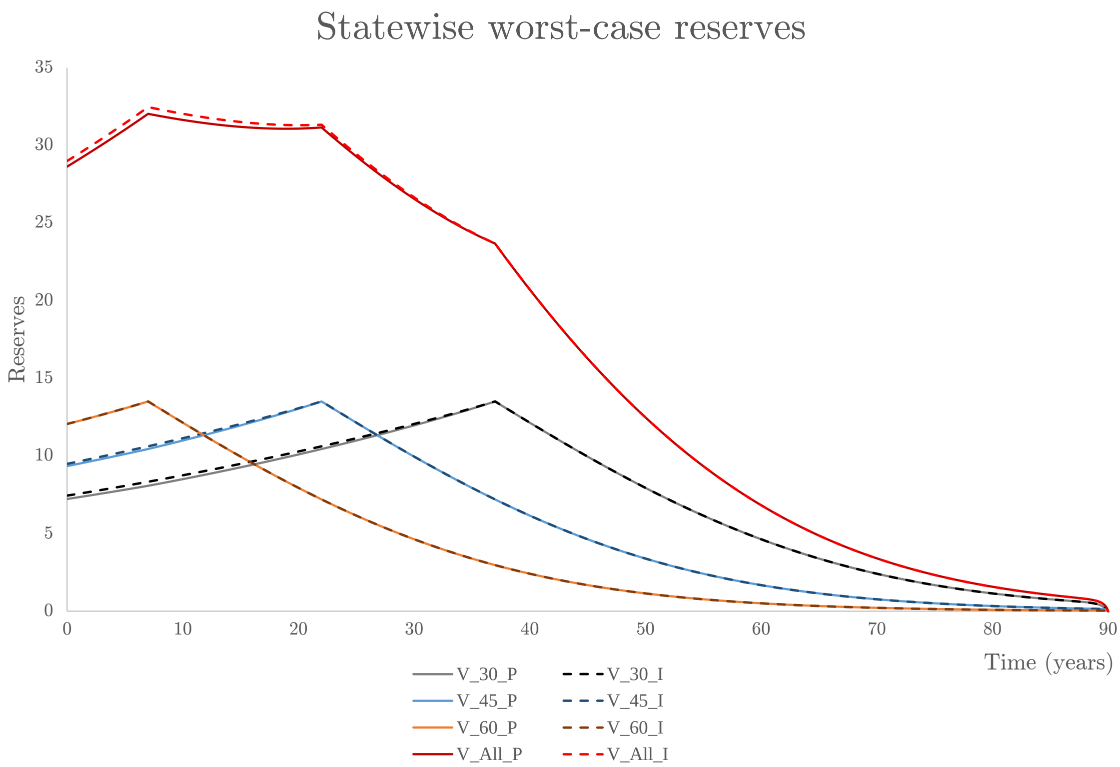

Figure 8: Statewise worst-case reserves calculated on basis of the worst-case scenarios for the individual policyholders and the worst-case scenario for the portfolio.

estimate, the third is α_l times the best estimate, and the fourth and fifth are the worst-case bases for the portfolio and the individual policyholders, respectively.

The numbers for "Solvency II (mortality)" and "Solvency II (longevity)" can be used to calculate the SCR in the "standard model". Assuming that only mortality risk and longevity risk apply to our portfolio, the SCR is defined as

$$\mathrm{SCR} = \sqrt{(\Delta V^{mortality})^2 + (\Delta V^{longevity})^2 - 2 \cdot 0.25 \cdot \Delta V^{mortality} \Delta V^{longevity}}, \tag{5.1}$$

where

$$\Delta V^{mortality} = \sum_{l=1}^{L} \max\left(V^l\left(\mu^{\mathrm{be}} \cdot 1.15\right) - V^l\left(\mu^{\mathrm{be}}\right), 0\right),$$
$$\Delta V^{longevity} = \sum_{l=1}^{L} \max\left(V^l\left(\mu^{\mathrm{be}} \cdot 0.8\right) - V^l\left(\mu^{\mathrm{be}}\right), 0\right).$$

The result of this calculation together with the calculations of the worst-case reserves lead to three different "SCR-like" quantities presented in Table 3. We here see that the SCR for the entire portfolio is significantly smaller for the worst-case scenario for the portfolio ($\approx 6.8\%$ of the reserve) compared to the worst-case scenario for the individual policyholders ($\approx 8.2\%$ of the reserve). However, they are both bigger than the SCR calculated using (5.1) ($\approx 5.9\%$ of the reserve).

	PH 1	PH 2	PH 3	Sum
Best estimate	6.91	8.80	11.09	26.81
Solvency II (mortality)	6.81	8.57	10.60	25.97
Solvency II (longevity)	7.17	9.27	11.97	28.40
Worst-case (PF)	7.23	9.35	12.06	28.64
Worst-case (separate)	7.45	9.49	12.06	29.00

Table 2: Reserves calculated by different methods ($b_{ad} = 15$) for the three policyholders.

Solvency II	Worst-case (PF)	Worst-case (separate)
1.59	1.83	2.19

Table 3: SCR calculated by different methods ($b_{ad} = 15$).

5.2.1 Increasing the term insurance - making more shifts in the worst-case intensities

In the previous example, we saw that there was one shift for the worst-case intensity for the portfolio; the shift was from low to high intensity after seven years. This kind of structure is probably quite normal for a big portfolio. However, this is not necessarily the case. There can be many more shifts, as we illustrate in this section. The only difference compared to the former example is that the term insurance is increased from 15 to 32. This leads to the worst-case intensities in Figure 9, where we have jumps in the worst-case intensities after the retirement of each of the policyholders. We also note, as opposed to the example with $b_{ad} = 15$, that none of the individually worst-case intensities coincide with the worst-case intensity for the portfolio. A table equivalent to Table 3 can be found in Table 5. The results there illustrate that the relative differences between the results of the different calculation methods can be quite big.

A comparison of different types of SCR-like calculations can be found in Table 4. We note that the relative differences of the SCRs are much bigger than in the former example with $b_{ad} = 15$.

	PH 1	PH 2	PH 3	Sum
Best estimate	10.01	11.95	13.08	35.05
Solvency II (mortality)	10.30	12.12	12.86	35.28
Solvency II (longevity)	9.71	11.85	13.58	35.15
Worst-case (PF)	10.28	12.78	13.54	36.60
Worst-case (separate)	10.93	13.04	14.33	38.30

Table 4: Reserves calculated by different methods ($b_{ad} = 32$) for the three policyholders.

The three different calculation methods lead to the three different SCR presented in Table 5.

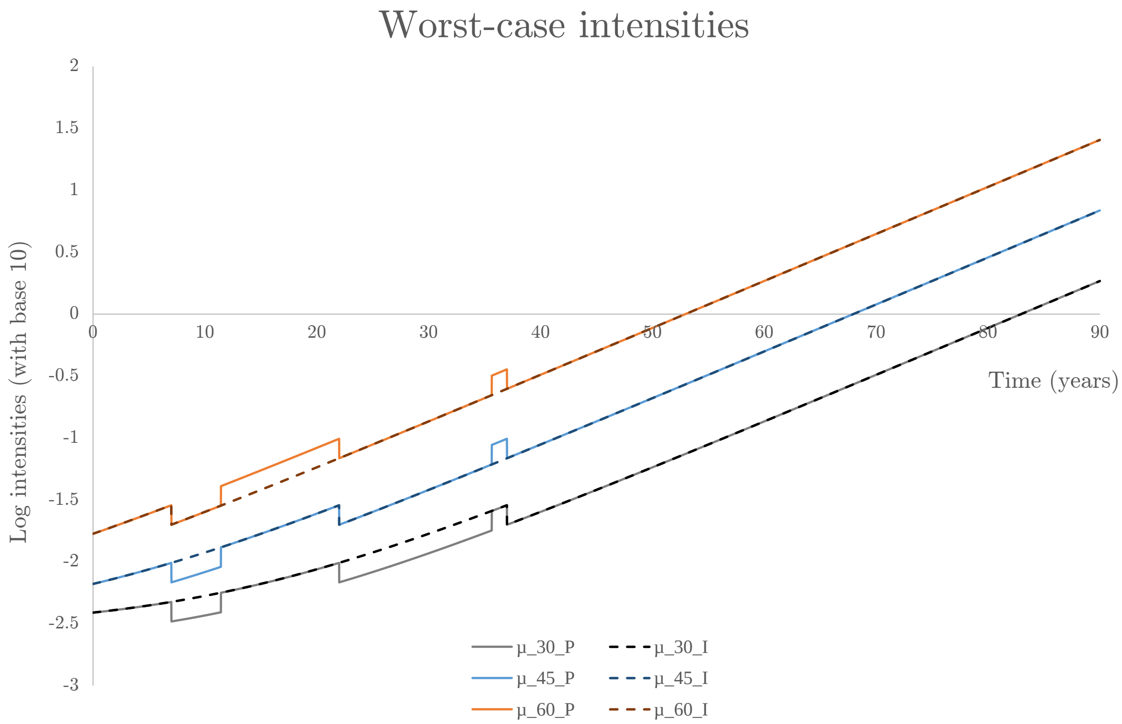

Figure 9: Worst-case intensities for the portfolio and for each individual policyholder ($b_{ad} = 32$).

Solvency II	Worst-case (PF)	Worst-case (separate)
0.59	1.55	3.25

Table 5: SCR calculated by different methods ($b_{ad} = 32$).

5.3 Conclusion

First, what is meant by stressing a portfolio with respect to mortality and longevity by scaling the mortality rate by between 0.8 and 1.15? The standard formula (5.1) represents one interpretation. The negative correlation in (5.1) is a (normal) probabilistic formalization of the idea that experiencing future high mortality rates and future low mortality rates tend not to happen in the same realization. Our calculations do not assume such "restrictions". We actually do find that the worst thing that can happen is that the mortality rate is high in the near future and low in the distant future and our worst-case approach really allows this realization to occur. Our numerical example shows that the standard formula based on a negative correlation of 0.25 leads to a too low capital requirement compared to the one obtained in our calculation. We do not claim that one can draw strong quantitative conclusions from this. What we do claim is that it is urgently important to understand exactly what is calculated and what is not. If one tends to believe that mortality rates actually can be high in the near future and low in the distant future, then calculations based on the standard formula may be dangerous.

Second, what is the intuition behind the high mortality rates in the near future and the low mortality rates in the distant future? This conforms with the basic understanding that in the near future, when policyholders are relatively young and hold positive sums at risk, high mortality rates are undesirable. Conversely, in the distant future, when policyholders are relatively old and hold negative sums at risk, low mortality rates are undesirable. The individual calculations in this section take into account these effects on a policy by policy basis in the sense what defines the near and distant future depends on the age of the individual policyholder. This is the simpler calculation giving a worst-case basis separately for each policy. The portfolio calculations deal with the situation where the same realized mortality rate counts for all policies, thinking of uncertainty in the mortality rate as being at macro-level. Inhomogeneity in the portfolio now reduces the consequences of the worst case and the capital requirement goes down. It is important to understand that this has absolutely nothing to do with diversification but is related to portfolio inhomogeneity exclusively. The inhomogeneity in the portfolio with respect to age and products may be so involved that the worst-case "jumps" up and down before it finds its low when the distant future is finally met, as illustrated in subsection 5.2.1.

Third, what do these solvency calculations have to do with design and pricing, which was mentioned in the introduction? Here, it is important to remember that a policy is an object for safe-side calculation already upon pricing, before the contract is underwritten and goes into the solvency calculation. So the first safe-side calculations are part of the internal pricing and management procedures in contrast to solvency calculations where principles and restrictions are given from outside. Our results illustrate how one can ascertain a given level of prudence by setting the first order pricing basis. Individual and portfolio level calculations allow for setting the first order bases differently for different (groups of) policyholders. A more individual unit for pricing leads to a more prudent pricing basis and, thus, higher surplus contributions from the portfolio. This idea was maybe not relevant in the past due to technological limitations. We have then indirectly illustrated the prudency effects of micro-pricing in life insurance by different tailor-made first order bases used for a portfolio of inhomogeneous policies.

Fourth, what can we conclude in general about scenarios on the basis of our calculations? In contrast to some of the general references mentioned in the introduction, we do not criticize scenarios as a mean of solvency calculations and management for being too uninformative, too inaccurate or too simple. Rather, we push forward scenarios, exactly for being simple to work with and understand. They just have to be chosen such that information and accuracy is not lost in the translation between distributional aspects of intensities and reserves, respectively. We consider general policies and portfolios and find that the worst-case intensities are the ones that maximize the expected sum at risk. Since the intensities occur in the expectation itself, this is a delicate optimization and not just a check of the sign of a given sum at risk. However, once the calculational challenges are overcome, one is left with a stress calculation simple to implement, simple to interpret, and simple to communicate, while actually bounding the insolvency probability, see also the paragraph around (1.1).

Acknowledgments: We are grateful to an anonymous referee for many fruitful comments that greatly improved the paper.

References

Bailey, P. B., Shampine, L. F. and Waltman, P. E. (1968). *Nonlinear two point boundary value problems*, Academic Press, New York.

Bertsekas, D. P. (2005). *Dynamic Programming and Optimal Control*, Vol. 1, 3rd edn, Athena Scientific, Belmont, MA.

Börger, M. (2010). Deterministic shock vs. stochastic value-at-risk - an analysis of the Solvency II standard model approach to longevity risk, *Blätter der DGVFM* **31**(2), 225–259.
URL: *http://dx.doi.org/10.1007/s11857-010-0125-z*

Christiansen, M. C. (2008). A sensitivity analysis concept for life insurance with respect to a valuation basis of infinite dimension, *Insurance: Mathematics and Economics* **42**(2), 680–690.
URL: *http://ideas.repec.org/a/eee/insuma/v42y2008i2p680-690.html*

Christiansen, M. C. and Steffensen, M. (2013). Safe-Side Scenarios for Financial and Biometrical Risk, *ASTIN Bulletin*, 1–35.
URL: *http://www.journals.cambridge.org/abstract_S0515036113000160*

Devineau, L. and Loisel, S. (2009). Risk aggregation in Solvency II: How to converge the approaches of the internal models and those of the standard formula?, *Post-Print hal-00403662*, HAL.
URL: *http://ideas.repec.org/p/hal/journl/hal-00403662.html*

Doff, R. (2008). A Critical Analysis of the Solvency II Proposals, *The Geneva Papers on Risk and Insurance Issues and Practice* **33**, 193–206.

EIOPA (2013). Technical specification on the long term guarantee assessment (part i), *Technical report*, European Insurance and Occupational Pensions Authority.

Li, J. and Szimayer, A. (2011). The uncertain mortality intensity framework: Pricing and hedging unit-linked life insurance contracts, *Insurance: Mathematics and Economics* **49**(3), 471 – 486.

Li, J. and Szimayer, A. (2014). The effect of policyholders' rationality on unit-linked life insurance contracts with surrender guarantees, *Quantitative Finance* **14**(2), 327–342.

URL: *http://www.tandfonline.com/doi/abs/10.1080/14697688.2013.825922*

Norberg, R. (1999). A theory of bonus in life insurance, *Finance and Stochastics* **3**(4), 373–390.
URL: *http://dx.doi.org/10.1007/s007800050067*

Olivieri, A. and Pitacco, E. (2008). Assessing the cost of capital for longevity risk, *Insurance: Mathematics and Economics* **42**(3), 1013–1021.
URL: *http://ideas.repec.org/a/eee/insuma/v42y2008i3p1013-1021.html*

Orava, P. and Lautala, P. (1976). Back-and-forth shooting method for solving two-point boundary-value problems, *Journal of Optimization Theory and Applications* **18**(4), 485–498.
URL: *http://dx.doi.org/10.1007/BF00932657*

Steffen, T. (2008). Solvency II and the Work of CEIOPS, *The Geneva Papers on Risk and Insurance Issues and Practice* **33**, 60–65.