SAFE-SIDE SCENARIOS FOR FINANCIAL AND BIOMETRICAL RISK

BY

MARCUS C. CHRISTIANSEN AND MOGENS STEFFENSEN

ABSTRACT

Premium settlement and calculation of reserves and capital requirements are typically based on worst- or just bad-case assumptions on interest rates, mortality rates, and other transition rates between states defined according to the insurance benefits. If interest and transition rates are chosen independently from each other, the worst choice, i.e. the combination of interest rates and transition rates that maximizes the reserve, can be found by dynamic programming. Here, we generalize this idea by choosing the interest and transition rates from a set that allows for mutual dependence. In general, finding the worst case is much more complicated in this situation, but we characterize a set of relatively tractable problems and present a series of examples from this set. Our approach with mutual dependence is relevant e.g. for internal models in Solvency II.

KEYWORDS

Worst-case scenarios, interest and transition rates, dependence, first-order basis, capital requirement.

1. INTRODUCTION

Safe-side calculations play two important roles in the mathematics of life insurance. First, traditional with-profit life insurance products are based on a settlement of the payments in accordance with a safe-side equivalence principle. Here, the term "safe-side" reflects that the resulting premiums should be high enough to cover the benefits in essentially all realistic economic–demographic scenarios. Second, when evaluating the future payments for accounting and solvency purposes, safe-side calculations also play a role depending on the accounting and solvency regime. Before notions such as market-consistency and risk-based capital entered into the vocabulary of the actuary, managing a life insurance portfolio was easier. The safe-side assumptions used for calculation of the premiums at the initiation of the contract were typically also used throughout the term of the contract in order to calculate (meant-to-be) safe-side reserves for the

Astin Bulletin 43(3), 323-357. doi: 10.1017/asb.2013.16

balance sheet. These book values were then essentially multiplied by a constant in order to calculate the buffer capital required for a business to be solvent. Thus, safe-side calculations for premium calculation and reserving were indistinguishable.

However, during the last decades, the market, the economists, academia and forces inside the actuarial community have shaken these grounds by pointing at deficiencies. How should the financial market evaluate and the management steer the insurance business in a real-economic environment with artificial (book) quantities from the balance sheets as the only input? How should a required capital calculation essentially based on the expectation principle ever be able to measure the real creditworthiness of a risky business? These questions called for a sophistication of the safe-side calculation in connection with premium settlement, accounting and solvency, respectively. Safe-side calculation in connection with premium settlement is still an important issue, but isolated to the initiation of the contract as being a part of the contract design. The role of safe-side calculations in accounting is uncertain. It is exactly the very hiding of reserves resulting from safe-side calculations which market-consistent valuation is supposed to prevent. However, safe-side calculations may still appear in accounting since (1) the premium calculation basis may play a direct role in future payments, e.g. in the calculation of surrender values, and (2) (somewhat) safe-side calculations have still to be used to produce risk margins on top of the best estimates.

Safe-side and bad-case scenarios are commonly used in the actuarial practice with the same meaning and are understood as prudential or conservative scenarios to be adopted when choosing the first-order basis for pricing and (traditional) reserving. Conversely, stress scenarios and worst-case scenarios do not in the actuarial practice refer to pricing or reserving but to capital allocation for solvency purposes, resilience testing, etc. Actually, this is where the use of scenarios appears to have a strong future ahead, cf. Genest *et al.* (2009), Goovaerts *et al.* (2011) and McNeil and Smith (2012). The so-called Solvency II Standard Formula, which will be the future reference for the calculation of solvency capital requirements (SCRs), is methodically based on stress scenarios. Its concept emanates from the Basel accords, where stress tests are applied for market and credit risk capital requirements. The regulatory frameworks of both Solvency II and Basel II require to base the risk measurement on the value-at-risk approach. Studer (1997, 1999) showed that there is a strong relationship between the value-at-risk and the stress scenario concept, proving that certain worst-case scenarios yield conservative bounds for the value-at-risk. The strong link between stress testing and the risk measure value-at-risk was also emphasized by Berkowitz (2000). Whether one needs a bad-case or worst-case scenario depends of course on the purpose of the scenario. Although we do have solvency purposes in mind, our results are not limited to such an application. And since, from a mathematical point of view, the concepts are qualitatively equivalent, we do not carefully distinguish between the scenario terminology “safe-side”, “bad-case”, “worst-case” or “stress”. All in all, safe-side calculations still play important roles that are delicately integrated across the calculation of premium, reserves and capital requirements. A great advantage of safe-side, worst-case and stress scenarios is that they are easy to understand also for non-mathematicians. While their derivation, which includes solving an optimization problem, can be tricky, the result can be easily interpreted.

One of the first systematic, probabilistic modeling frameworks of stress testing was given by Studer (1997, 1999), who defined stress scenarios as scenarios that maximize the loss of a market portfolio with respect to some trust region. By assuming that the risk factors have a multivariate normal distribution, his trust regions have an ellipsoidal shape.

Closely related to stress testing is the so-called worst-case risk measurement, where the worst case of some risk measure is calculated if only partial information on the distribution of the future loss is available, see for example Laeven (2009) and Kaas *et al.* (2009). In a recent paper, Goovaerts *et al.* (2011) study a general form of this problem, maximizing an integral with respect to a set of measures expressing incomplete probabilistic information. With the objective function being an integral, they solve a linear problem. In contrast, our paper deals with a nonlinear objective function.

Another concept for the definition of stress scenarios uses the probabilistic concept of half-space trimming, see for example Kupiec (1998), Aragons *et al.* (2001) and Alexander and Sheedy (2008). McNeil and Smith (2012) extend that concept to quite general multivariate stress scenarios. They establish theoretical properties when the risk factors can be assumed to have a linear effect on the net assets of an insurer. Interestingly, the stress scenarios of McNeil and Smith (2012) directly correspond to a value-at-risk. In our paper, the future loss is a nonlinear mapping of the risk factors, and we obtain upper bounds on the value-at-risk instead.

Christiansen and Denuit (2010) calculate worst-case mortality rates for classical life insurance policies, but interactions with interest rates or lapse rates are not taken into account. Li and Szimayer (2011) calculate worst-case mortality rates for unit-linked life insurance contracts. An advantage of their model is that the worst-case mortality rate may depend on past capital gains. A similar model is used in Li and Szimayer (2010) for the study of surrender guarantees in unit-linked life insurances. The surrender rate takes the role of the mortality and may vary between an upper and a lower bound, while the mortality rate itself is fixed. The worst-case scenarios of Li and Szimayer (2010, 2011) are adapted stochastic process, whereas we aim to find deterministic worst-case scenarios, which are completely known at time zero. In practice, the deterministic approach is typically used for first-order calculation and it appears, e.g., in the Standard Formula of Solvency II. Christiansen (2010, 2011) calculates deterministic worst-case scenarios for multistate policies with several transition rates. The focus is on biometrical parameters only and interest rates are fixed. We want to solve simultaneously for worst-case biometrical parameters and interest rates. In Christiansen (2010, 2011), the biometrical intensities fluctuate independently without interactions. However, it may be natural to impose some dependence, first, mutually between the intensities and, second, maybe even between the intensities and the interest rates. What we mean by this dependence is made clear in some motivating examples following the introduction.

2. Motivating examples

In order to ground the ideas, we consider two small examples of a three-state model for a life insurance contract. They serve to illustrate the "usual" approach, how this approach can be improved by realistic dependence modeling, and the challenges arising from such an enhancement.

In the first example, we illustrate what we mean by dependence between different intensities. We consider a disability model consisting of three states, $a =$ "active", $i =$ "disabled" and $d =$ "dead". We let μ_{ad} and μ_{id} be the transition rates from a and i to d, respectively. Correspondingly, we let b_{ad} and b_{id} be the sums paid out upon these transitions. In the classical Thiele differential equations characterizing the statewise reserves V_a and V_i for the states a and i, respectively, appear the so-called death risk premium rates paid out of the reserves, namely

$$\begin{aligned} &\mu_{ad}(t)\,(b_{ad}(t) - V_a(t))\,, \\ &\mu_{id}(t)\,(b_{id}(t) - V_i(t))\,, \end{aligned} \tag{2.1}$$

respectively. The death risk premium rate is a product of the mortality rate and the loss upon death, the so-called sum at (death) risk. In order to find some "bad-case" (i.e. stressful or "expensive") intensities, one usually looks at these two risk premium rates separately. If, for example, $b_{ad}(t) > V_a(t)$, which means that the insurance company experiences a loss upon death, then death is stressful and a "high" mortality intensity μ_{ad} is a "bad-case". We conclude that, from a set of possible/admissible intensities $M = [\underline{\mu}_{ad}, \overline{\mu}_{ad}]$, we should choose $\overline{\mu}_{ad}$ as the "worst-case". Here and in the following, we use intuitive overlines and underlines in order to mark upper and lower bounds. For the death intensity in state "disabled", a similar argument leads to the conclusion that if $b_{id}(t) < V_i(t)$, then a "low" mortality intensity is a "bad case" and from a set $M = [\underline{\mu}_{id}, \overline{\mu}_{id}]$ we should choose $\underline{\mu}_{id}$ as "worst-case". The conditions $b_{ad}(t) > V_a(t)$ and $b_{id}(t) < V_i(t)$ are recognized in the usual sum at risk sign check for a "safe-side basis", namely that, for example, a safe-side intensity $\mu^*_{ad}(t)$ should be chosen such that

$$\mathrm{sign}(\mu^*_{ad}(t) - \mu_{ad}(t)) = \mathrm{sign}(b_{ad}(t) - V_a(t)), \tag{2.2}$$

see Norberg (1999, formula (3.8)). In (2.2), we see that the spread between the intensities depends on the reserve. In turn, the reserve depends on the intensities. Thus, the intensities and the reserves should in principle be calculated simultaneously. A complete study of this is found in Christiansen (2010).

The conclusions above correspond to choosing a bad combination of μ_{ad} and μ_{id} in a rectangular set $M = [\underline{\mu}_{ad}, \overline{\mu}_{ad}] \times [\underline{\mu}_{id}, \overline{\mu}_{id}]$. But what happens if the tuple (μ_{ad}, μ_{id}) is to be chosen from a non-rectangular set? One could, for example, rule out extreme "opposite" combinations like the above scenario $(\overline{\mu}_{ad}, \underline{\mu}_{id})$, since this may be an unrealistic combination from a socio-biological point of view. Natural shapes of M (at least to think of) could be an ellipsoid or simply a line. A line could come from a constraint like, e.g., $\mu_{id} \equiv \mu_{ad} + d$ for some fixed addend d. What is then a "bad" μ_{ad}? If $b_{ad}(t) > V_a(t)$ and $b_{id}(t) > V_i(t)$, it is clear from the arguments of above that $\left(\overline{\mu}_{ad}, \overline{\mu}_{ad} + d\right)$ is the worst choice in M. But what is the worst choice if $b_{ad}(t) > V_a(t)$ and $b_{id}(t) < V_i(t)$?

In the second example, we illustrate what we mean by dependence between intensities and interest rates. We consider a surrender model consisting of three states, a = "alive and in force", s = "surrender" and d = "dead". In the differential equation characterizing the statewise reserve for the state a appears an interest rate term minus the surrender risk premium rate, namely

$$\varphi\, V_a(t) - \mu_{as}(t)\,(b_{as}(t) - V_a(t))\,.$$

Here, φ is the interest rate. In order to find a "bad" basis, one usually looks at the two terms separately. If $V_a(t) > 0$, which is usually the case, a low interest rate is "bad", and in the set of possible/admissible interest rates $M_\varphi = [\underline{\varphi}, \overline{\varphi}]$ we should choose $\underline{\varphi}$. Similar to the disability model in the first example, a "bad case" μ_{as} depends on the sign of $b_{as}(t) - V_a(t)$, and from the set $M_{as} = [\underline{\mu}_{as}, \overline{\mu}_{as}]$ we choose $\underline{\mu}_{as}$ if $b_{as}(t) < V_a(t)$ and $\overline{\mu}_{as}$ if $b_{as}(t) > V_a(t)$. To have a surrender sum larger than the reserve may sound unrealistic. However, note that the surrender value may be related — or even equal — to a different (and, possibly, larger) reserve, based on a different technical basis, than the "bad-case" reserve $V_a(t)$ we wish to calculate.

However, these conclusions hold only if (φ, μ_{as}) is chosen from the rectangular set $M = [\underline{\varphi}, \overline{\varphi}] \times [\underline{\mu}_{as}, \overline{\mu}_{as}]$. What happens if this set is not rectangular? An interest rate-dependent surrender intensity is not unrealistic and has been suggested in several connections, e.g., De Giovanni (2010). The usual perception is that the surrender intensity increases with the interest rate, see De Giovanni (2010). This means that there are some combinations of the interest rate and the surrender intensity in the rectangular set which should be ruled out. Again, we can imagine an M formed as an ellipsoid or simply a line. What is now the worst-case combination of φ and μ_{as} in M?

The principal idea of this paper is to answer questions such as the ones posed in these two motivating examples.

3. Modeling and Valuation of Life Insurance Policies

Consider an insurance policy that is driven by a jump process $X(t)$, $t \geq 0$, with finite state space $\mathcal{S}$, transition space $J \subset \{(j, k) \in \mathcal{S}^2 \mid j \neq k\}$, and deterministic starting value. We write x for the age of the policyholder at time zero (the beginning of the contract period) and ω_x for the residual limiting lifetime for the policyholder at time zero. The cash-flows of the contract are described by the following functions:

1. The lump sum $b_{jk}(t)$ is payable upon a transition from j to k at time t. We assume that the functions b_{jk}, $(j,k) \in J$, have bounded variation on $[0, \omega_x]$.
2. The function $B_j(t)$ gives the accumulated annuity benefits minus accumulated premiums for sojourns in j during $[0, t]$. We assume that the B_j, $j \in \mathcal{S}$, have bounded variation on $[0, \omega_x]$ and are right-continuous.

We write $v(s,t) = v(s,t;\phi)$ for the stochastic present value at time s of a unit payable at time $t > s$. Following the notion of Norberg (1999), we assume that the sample paths are absolutely continuous such that we have a representation of the form

$$v(s,t;\phi) = e^{-\int_s^t \phi(u)\,\mathrm{d}u},$$

with ϕ being some stochastic interest intensity. As Norberg (1999) points out, apart from the absolute continuity of the paths, no particular specification of the distribution of $v(s,t;\phi)$ is needed. The possible distributions are countless. In particular, for any given (risky) asset portfolio combined with any fixed finite time grid, there exists a distribution of ϕ, such that the finite dimensional distributions of v and the inverse portfolio return coincide on that grid. Therefore, ϕ can be thought of as a fairly general representation of portfolio return. Furthermore, let there be a matrix-valued stochastic process $\nu(t)$, $t \geq 0$, with diagonal values $\nu_{jj}(t) = -\sum_{k:k\neq j} \nu_{jk}(t)$ and such that, conditional on $\nu = \mu \in L_1^{\mathcal{S}\times\mathcal{S}}([0,\omega_x])$, the jump process X is Markovian with transition intensity matrix μ. That means that the conditional distribution $\mathcal{L}(X \mid \nu = \mu)$ of X is completely determined by a so-called transition probability matrix

$$p(s,t;\mu) \overset{a.s.}{=} \Big(P(X_t = k \mid X_s = j,\ \nu = \mu)\Big)_{(j,k)\in\mathcal{S}^2}, \qquad 0 \leq s \leq t \leq \omega_x,$$

that is the unique solution of the Kolmogorov forward equation

$$\frac{\mathrm{d}}{\mathrm{d}t} p(s,t;\mu) = p(s,t;\mu)\,\mu(t), \quad s \leq t,$$

with starting values $p(s,s) = \mathbb{I}$ for all s. The interest intensity ϕ and the process ν form a stochastic valuation basis (ϕ, ν), which we here think of as the "true" basis that is not known a priori. As the elements of ν are completely determined by $\nu_J(t) := (\nu_{jk}(t))_{j\neq k}$ and $\nu_{jj} = -\sum_{k:k\neq j} \nu_{jk}$, from now on we will also write (ϕ, ν_J) for the stochastic valuation basis. The aim of the present paper is to identify a deterministic valuation basis $(\varphi, \mu_J) \in L_1^{1+|J|}([0,\omega_x])$ that is conservative in some sense.

Let $B(s; X, \phi)$ be the present value at time s of future benefits, net of future premiums. Since future payments between insurer and policyholder depend on ν

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

only via the state process X, the present value B does not have ν as an argument. For a fixed process of future net cash-flows, the prospective reserve at time s in state i conditional on $(\phi, \nu_J) = (\varphi, \mu_J)$ is defined by

$$V_i(s; \varphi, \mu) = \mathbb{E}\big(B(s; X, \phi) \,\big|\, (X_s, \phi, \nu) = (i, \varphi, \mu)\big).$$

Since we know that conditional on $\nu = \mu$ the jump process X is a Markovian process with transition intensity matrix μ, we almost surely have

$$\begin{aligned} V_i(s; \varphi, \mu) = &\sum_{j\in\mathcal{S}} \int_{(s,\omega_x]} v(s, t; \varphi)\, p_{ij}(s, t; \mu)\, \mathrm{d}B_j(t) \\ &+ \sum_{(j,k)\in J} \int_s^{\omega_x} v(s, t; \varphi)\, b_{jk}(t)\, p_{ij}(s, t; \mu)\, \mu_{jk}(t)\, \mathrm{d}t. \end{aligned} \tag{3.1}$$

The family of prospective reserves $V_i(s; \varphi, \mu), i \in \mathcal{S}, s \in [0, \omega_x]$, can be obtained as the unique solution of Thiele's integral equation system

$$\begin{aligned} V_i(s; \varphi, \mu) = B_i(\omega_x) - B_i(s) - &\int_s^{\omega_x} V_i(t; \varphi, \mu)\, \varphi(t)\, \mathrm{d}t \\ + \sum_{j:j\neq i} &\int_s^{\omega_x} R_{ij}(t; \varphi, \mu)\, \mu_{ij}(t)\, \mathrm{d}t, \end{aligned} \tag{3.2}$$

with starting values $V_i(\omega_x; \varphi, \mu) = 0$ for all $i \in \mathcal{S}$, where

$$R_{ij}(t; \varphi, \mu) := b_{ij}(t) + V_j(t; \varphi, \mu) - V_i(t; \varphi, \mu)$$

is the so-called sum at risk associated with a possible transition from state i to state j at time t. In the following, we will write $V_i(t)$, $R_{ij}(t)$, $p_{ij}(s, t)$, and $v(s, t)$ instead of $V_i(t; \varphi, \mu)$, $R_{ij}(t; \varphi, \mu)$, $p_{ij}(s, t; \mu)$, and $v(s, t; \varphi)$ when it is clear from the context which valuation basis (φ, μ) we are referring to.

Proposition 3.1. *Let (φ, μ) and $(\widetilde{\varphi}, \widetilde{\mu})$ be two valuation bases and let $V_i(s)$ and $\widetilde{V}_i(s)$ be the corresponding prospective reserves. For all $i \in \mathcal{S}$ and $s \geq 0$, we have*

$$\begin{aligned} \widetilde{V}_i(s) - V_i(s) = \int_s^{\omega_x} \Bigg(&\sum_{(j,k)\in J} v(s, t)\, p_{ij}(s, t)\, \widetilde{R}_{jk}(t)\, \big(\widetilde{\mu}_{jk}(t) - \mu_{jk}(t)\big) \\ &- \sum_{j\in\mathcal{S}} v(s, t)\, p_{ij}(s, t)\, \widetilde{V}_j(t)\, \big(\widetilde{\varphi}(t) - \varphi(t)\big) \Bigg) \mathrm{d}t. \end{aligned} \tag{3.3}$$

Proof. By applying Thiele's integral equation system (3.2), we get an integral equation system for the differences $\widetilde{V}_i(s) - V_i(s)$,

$$\widetilde{V}_i(s) - V_i(s) = C_i(\omega_x) - C_i(s) - \int_s^{\omega_x} \big(\widetilde{V}_i(t) - V_i(t)\big)\,\varphi(t)\,\mathrm{d}t + \sum_{j:j\neq i} \int_s^{\omega_x} \Big(\big(\widetilde{V}_j(t) - V_j(t)\big) - \big(\widetilde{V}_i(t) - V_i(t)\big)\Big)\,\mu_{ij}(t)\,\mathrm{d}t$$

with starting values $\widetilde{V}_i(\omega_x) - V_i(\omega_x) = 0$, $i \in \mathcal{S}$, where the right-continuous functions C_i are (uniquely) defined by

$$C_i(\omega_x) - C_i(s) = -\int_s^{\omega_x} \widetilde{V}_i(t)\,\big(\widetilde{\varphi}(t) - \varphi(t)\big)\,\mathrm{d}t + \sum_{j:j\neq i} \int_s^{\omega_x} \widetilde{R}_{ij}(t)\,\big(\widetilde{\mu}_{ij}(t) - \mu_{ij}(t)\big)\,\mathrm{d}t.$$

The C_i are known in the actuarial literature as "statewise contributions to surplus". Our new integral equation system for the differences $W_i(s) := \widetilde{V}_i(s) - V_i(s)$,

$$W_i(s) = C_i(\omega_x) - C_i(s) - \int_s^{\omega_x} W_i(t)\,\varphi(t)\,\mathrm{d}t + \sum_{j:j\neq i} \int_s^{\omega_x} \Big(W_j(t) - W_i(t)\Big)\,\mu_{ij}(t)\,\mathrm{d}t,$$

can be interpreted as a Thiele integral equation system for a policy with cumulative annuity benefits C_i and no transition benefits. It has the unique solution

$$W_i(s) = \sum_{j\in\mathcal{S}} \int_{(s,\omega_x]} v(s,t)\,p_{ij}(s,t)\,\mathrm{d}C_j(t).$$

To see that, follow the proof of Theorem 4.8 in Milbrodt and Stracke (1997), but note that $W_i(s)$ here has the integration intervals $(s, \omega_x]$ instead of $[s, \infty)$, and that we have no transition benefits. ■

Another way to write $W_i(s)$ can be found in Christiansen (2008), where a Taylor expansion of the prospective reserve $V_i(s)$ — seen as a mapping of the

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

valuation basis (φ, μ_J) — is developed, leading to

$$W_i(s) = \int_s^{\omega_x} \Bigg(\sum_{(j,k)\in J} v(s,t)\, p_{ij}(s,t)\, R_{jk}(t) \left(\widetilde{\mu}_{jk}(t) - \mu_{jk}(t)\right) - \sum_{j\in\mathcal{S}} v(s,t)\, p_{ij}(s,t)\, V_j(t) \left(\widetilde{\varphi}(t) - \varphi(t)\right)\Bigg) \mathrm{d}t + o(\|(\widetilde{\varphi}, \widetilde{\mu}) - (\varphi, \mu)\|)$$
$$=: \int_s^{\omega_x} \nabla_{(\varphi,\mu_J)}(V_i(s))(t) \cdot ((\widetilde{\varphi}(t), \widetilde{\mu}_J(t)) - (\varphi(t), \mu_J(t)))\, \mathrm{d}t + o(\|(\widetilde{\varphi}, \widetilde{\mu}) - (\varphi, \mu)\|), \tag{3.4}$$

where $\|\cdot\|$ denotes the L_1-norm of the maximum row sum. Christiansen (2008) interprets the function $\nabla_{(\varphi,\mu_J)}(V_i(s))(t) \in \mathbb{R}^{1+|J|}$ as some form of generalized gradient. In this version of W, compared with the representation in Proposition 3.1, the sum at risk and the reserve under the integral are based on (φ, μ_J) rather than $(\widetilde{\varphi}, \widetilde{\mu}_J)$, and this gives rise to the "error term" $o\,(\|(\widetilde{\varphi}, \widetilde{\mu}) - (\varphi, \mu)\|)$.

Another quantity that we will need later on is the reserve at time s, conditionally expected at time $u < s$, namely

$$V_{i,u}(s) = \sum_{j\in\mathcal{S}} p_{ij}(u,s)\, V_j(s).$$

Lemma 3.2. *The reserve $V_{i,u}(s)$ uniquely solves the integral equation system*

$$V_{i,u}(s) = \sum_{j\in\mathcal{S}} \int_{(s,\omega_x]} p_{ij}(u,t)\, \mathrm{d}B_j(t) - \int_s^{\omega_x} V_{i,u}(t)\, \varphi(t)\, \mathrm{d}t + \sum_{(j,k)\in J} \int_s^{\omega_x} b_{jk}(t)\, p_{ij}(u,t)\, \mu_{jk}(t)\, \mathrm{d}t \tag{3.5}$$

with starting values $V_{i,u}(\omega_x) = 0$ for all $i \in \mathcal{S}$ and an arbitrary but fixed u.

Proof. From (3.1) and the Chapman–Kolmogorov equation, we get that

$$V_{i,u}(s) = \int_{(s,\omega_x]} v(s,t) \Bigg[\sum_{j\in\mathcal{S}} p_{ij}(u,t)\, \mathrm{d}B_j(t) \Bigg] + \int_s^{\omega_x} v(s,t) \Bigg[\sum_{(j,k)\in J} b_{jk}(t)\, p_{ij}(u,t)\, \mu_{jk}(t) \Bigg] \mathrm{d}t.$$

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

With defining

$$\mathrm{d}B_1(t) := \sum_{j\in\mathcal{S}} p_{ij}(u,t)\,\mathrm{d}B_j(t) + \sum_{(j,k)\in J} b_{jk}(t)\,p_{ij}(u,t)\,\mu_{jk}(t)\,\mathrm{d}t,$$

we can interpret $s \mapsto V_{i,u}(s)$ as the prospective reserve of a one-state policy with state space $\mathcal{S} = \{1\}$ and sojourn payments $\mathrm{d}B_1(t)$. The corresponding Thiele integral equation (3.2) has just the form (3.5). ■

Note that we always have $V_{i,s}(s) = V_i(s)$.

3.1. Calculations on the safe-side

For the calculation of premiums and reserves, the insurer has to choose some deterministic basis since the "true" basis (ϕ, ν) is not known a priori. In order to set premiums and reserves on the safe-side, it is a common method to choose a first-order valuation basis that represents some worst-case scenario from the perspective of the insurer. In mathematical terms, we are looking for a valuation basis $(\widetilde{\varphi}, \widetilde{\mu})$ with

$$V_i(s;\widetilde{\varphi},\widetilde{\mu}) \geq V_i(s;\phi,\nu) \tag{3.6}$$

for all outcomes of (ϕ, ν). We can either aim to meet that condition for some specified time s and state i, or we can even ask for a valuation basis that satisfies this condition for all times s and all states i. The latter requirement was introduced by Hoem (1988) as the "basis safe-side requirement". The "true" basis may contain some quite extreme situations that appear too safe in one way or another. For example, they may be too expensive in terms of pricing or reserving. Therefore, we propose to work with the weaker

$$P\Big(V_i(s;\widetilde{\varphi},\widetilde{\mu}) \geq V_i(s;\phi,\nu)\Big) \geq 1-\alpha. \tag{3.7}$$

Starting from a set of possible future scenarios M that satisfies $P((\phi,\nu)\in M) \geq 1-\alpha$, we can find an admissible scenario $(\widetilde{\varphi}, \widetilde{\mu})$ for (3.7) by calculating

$$(\widetilde{\varphi},\widetilde{\mu}) = \underset{(\varphi,\mu)\in M}{\operatorname{argmax}}\, V_i(s;\varphi,\mu) \tag{3.8}$$

since

$$P\Big(\underset{(\varphi,\mu)\in M}{\operatorname{argmax}}\, V_i(s;\varphi,\mu) \geq V_i(s;\phi,\nu)\Big) \geq P((\phi,\nu)\in M) \geq 1-\alpha. \tag{3.9}$$

The focus of our paper is to solve the optimization problem (3.8), supposing that M is given. Note that in general it is not a trivial task to find a proper scenario set M. For example, M can represent some expert opinion or can be deduced from a stochastic model.

Solutions of the worst-case problem (3.8) are of general interest because of (3.9). However, it is of special interest for calculating regulatory solvency capital.

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

According to the Solvency II directive of the European Parliament and of the Council (2008), the SCR shall be calculated as the 99.5% value-at-risk of the change in the net value of assets minus liabilities over an $m = 1$ year period. With writing $N(t) = A(t) - L(t)$ for the net value of assets minus liabilities at time t, the SCR at time s has the representation

$$\mathrm{SCR}(s) = \mathrm{VaR}_{0.995}(N(s) - v(s, s+m)\, N(s+m)) .$$

The so-called Standard Formula, as suggested by the regulator, uses the approximation

$$\mathrm{SCR}(s) = \mathrm{VaR}_{0.995}\Big(N(s) - N^{\mathcal{F}_{s+m}}(s)\Big)$$

instead, where $N^{\mathcal{F}_{s+m}}(s)$ is the net value at time s given $\mathcal{F}_{s+m}$, and $\mathcal{F}_{s+m}$ denotes the information about (ϕ, ν) on $[0, s+m]$. If the insurer mainly holds riskless assets, then the difference $A(s) - A^{\mathcal{F}_{s+m}}(s)$ is negligible, and we obtain

$$\mathrm{SCR}(s) = \mathrm{VaR}_{0.995}\Big(- L(s) + L^{\mathcal{F}_{s+m}}(s)\Big) .$$

A different motivation for leaving out $A(s) - A^{\mathcal{F}_{s+m}}(s)$ would be to concentrate on the SCR contribution from the interest rate and insurance risk modules only (partial internal model). According to the so-called mark-to-market approach, which is adopted in Solvency II, and given that our policyholder is at present time s in state i, $L(s)$ has here the form

$$L(s) = V_i(s; \varphi^{BE_s}, \mu^{BE_s}),$$

where $(\varphi^{BE_s}, \mu^{BE_s})$ denotes a market-consistent best estimate basis known at time s. Such a basis combines market-consistent valuation of risks that are priced in the market by calibrating to market prices with a plain expectation with respect to risks that are not priced in the market. If e.g. the bond market specifies a forward rate projection of interest rates, these forward rates would be part of the market-consistent basis, see e.g. Fabozzi (2005, p. 148). If e.g. there exists a market for mortality derivatives, forward mortality rates would be part of the market-consistent basis, see e.g. Cairns *et al.* (2008). For $L^{\mathcal{F}_{s+m}}(s)$, we introduce a basis (ϕ^{s+m}, ν^{s+m}) that combines the true stochastic basis (ϕ, ν) on the time interval $[0, s+m)$ and the $\mathcal{F}_{s+m}$-measurable market-consistent best estimate basis $(\varphi^{BE_{s+m}}, \mu^{BE_{s+m}})$ on the time interval $[s+m, \infty)$. Then,

$$\begin{aligned} L^{\mathcal{F}_{s+m}}(s) &= V_i(s; \phi^{s+m}, \nu^{s+m}) \\ &= \mathbb{E}\Big(V_i(s; \phi, \nu) - v(s, s+m; \phi)\, V_{X(s+m)}(s+m; \phi, \nu) \\ &\qquad + v(s, s+m; \phi)\, V_{X(s+m)}(s+m; \varphi^{BE_{s+m}}, \mu^{BE_{s+m}}) \Big| \mathcal{F}_{s+m}\Big). \end{aligned}$$

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

In the second and third terms, we replace the stochastic basis (ϕ, ν) underlying $V_i(s; \phi, \nu)$ by the basis $(\varphi^{BE_{s+m}}, \mu^{BE_{s+m}})$ from time $s+m$ and onward for valuation at time s of the payments from time $s+m$ and onward. The expectation conditional on information about (ϕ, ν) until $s+m$ is then a representation of $V_i(s; \phi^{s+m}, \nu^{s+m})$. Hence, the SCR at time s has the formula

$$\mathrm{SCR}(s) = \mathrm{VaR}_{0.995}\Big(- V_i(s; \varphi^{BE_s}, \mu^{BE_s}) + V_i(s; \phi^{s+m}, \nu^{s+m})\Big) \tag{3.10}$$

and concerns the uncertainty of (ϕ^{s+m}, ν^{s+m}) given $\mathcal{F}_s$. In (3.10), it is of course a delicate issue how to determine $(\varphi^{BE_s}, \mu^{BE_s})$ and (ϕ^{s+m}, ν^{s+m}) given $\mathcal{F}_s$, see Börger (2010). However, this is not central for what follows and is therefore beyond the scope of this paper.

If we set $m = \infty$, formula (3.10) can be interpreted as the liability runoff approach. Advantages and disadvantages of different choices of m can be — and have been — discussed. Although it appears that for $m = 1$ and $m = \infty$ we have substantially different problems in (3.10), it is crucial to realize that this is actually a matter of setting the distribution under which the value-at-risk is taken. Below, specifications of the distribution of (ϕ^{s+m}, ν^{s+m}) correspond to confidence sets for (ϕ^{s+m}, ν^{s+m}) which then become crucially important.

The problems (3.10) and (3.7) are very similar, for $m = \infty$ they are in principle equivalent. However, both problems are in general difficult to solve. Instead, one can look for an upper bound in terms of quantiles of (ϕ^{s+m}, ν^{s+m}). For example, if we assume that M is a given confidence set with $P((\phi^{s+m}, \nu^{s+m}) \in M) \geq 0.995$, then we can obtain the following bound for the $\mathrm{SCR}(s)$:

$$\begin{aligned}
\mathrm{SCR}(s) &= \mathrm{VaR}_{0.995}\Big(- V_i(s; \varphi^{BE_s}, \mu^{BE_s}) + V_i(s; \phi^{s+m}, \nu^{s+m})\Big) \\
&= \inf\Big\{c \in \mathbb{R} \Big| P\Big(V_i(s; \varphi^{BE_s}, \mu^{BE_s}) - V_i(s; \phi^{s+m}, \nu^{s+m}) \geq -c\Big) \geq 0.995\Big\} \\
&\leq - \inf_{(\varphi,\mu)\in M} \Big\{ V_i(s; \varphi^{BE_s}, \mu^{BE_s}) - V_i(s; \varphi, \mu)\Big\} \\
&= \sup_{(\varphi,\mu)\in M} \big\{ V_i(s; \varphi, \mu) - V_i(s; \varphi^{BE_s}, \mu^{BE_s})\big\},
\end{aligned} \tag{3.11}$$

where the inequality is based on the fact that

$$\widetilde{c} = - \inf_{(\varphi,\mu)\in M} \Big\{ V_i(s; \varphi^{BE_s}, \mu^{BE_s}) - V_i(s; \varphi, \mu)\Big\}$$

is an admissible constant in the sense that

$$P\Big(V_i(s; \varphi^{BE_s}, \mu^{BE_s}) - V_i(s; \phi^{s+m}, \nu^{s+m}) \geq -\widetilde{c}\Big) \geq P\big((\phi^{s+m}, \nu^{s+m}) \in M\big) \geq 0.995.$$

This correspondence of the value-at-risk to a worst-case problem was already shown by Studer (1997, 1999), who denotes M as the trust region. The supremum in (3.11) corresponds to problem (3.8) and is the kind of worst-case problem that we study in this paper. A similar idea to identify an upper bound for the SCR can be found in Section 4 of Bauer *et al.* (2009). There, the SCR of a monotonically increasing functional of a finite-dimensional random vector is bounded by taking the minimum over all rectangular confidence sets for the argument. In our case, the functional has an infinite dimensional domain, the supremum is taken over an arbitrarily shaped yet fixed confidence set, and, most important, we avoid the monotonicity requirement. The Solvency II Standard Formula also uses the idea in (3.9) by approximating the value-at-risk by a stress scenario that can be interpreted as a worst-case scenario, see for example Börger (2010).

Finally, we wish to place a comment on the connection between the choice of M and the time consistency of the worst-case basis already at this point. Our results in the next section contain invariance conditions for the set M under which the worst-case basis is indeed time consistent. We provide examples for which the invariance condition holds. However, we do not discuss necessary properties of the set M under which the invariance conditions are fulfilled. In other words, we do not impose any upfront restrictions on M that guarantee the worst-case basis to be time consistent. This would be a different route to take but it is not clear how such restrictions should be formed.

4. Calculation of worst-case scenarios

Let the vector $(\varphi(t), \mu_J(t)) := (\varphi(t), (\mu_{jk}(t))_{(j,k)\in J})$ be an element of a set $M \subset L_1^{1+|J|}([0, \omega_x])$ of integrable intensity vectors. We think of M as a set of future financial and biometrical scenarios that may occur with some probability, possibly one. (Recall that the diagonal elements of μ are determined by μ_J via $\mu_{jj} = -\sum_{k:k\neq j} \mu_{jk}$.) We write $M(t)$, $t \in [0, \omega_x]$, for the t-slices of M, which describe the admissible values that the different scenarios may take just at time t. We are interested in the supremum

$$\sup_{(\varphi,\mu_J)\in M} V_{i_1}(s_1; \varphi, \mu_J) \tag{4.1}$$

for arbitrary but fixed $i_1 \in \mathcal{S}$, $s_1 \in [0, \omega_x]$. If there exists an element $(\widetilde{\varphi}, \widetilde{\mu}_J) \in M$ such that $V_{i_1}(s_1; \widetilde{\varphi}, \widetilde{\mu}_J)$ equals the supremum, we call $(\widetilde{\varphi}, \widetilde{\mu}_J)$ a *worst-case scenario for* $V_{i_1}(s_1)$ *with respect to* M. For general sets M, the calculation of the supremum in (4.1) can be very difficult. We give analytical solutions for a number of special cases for which dynamic programming applies.

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

Proposition 4.1 (Verification Lemma). *Let $(\varphi, \mu_J) \in M$ and $(\widetilde{\varphi}, \widetilde{\mu}_J) \in M$ be fixed scenarios that solve*

$$\begin{aligned}\widetilde{V}_i(s) &= B_i(\omega_x) - B_i(s) - \int_s^{\omega_x} \widetilde{V}_i(t)\,\widetilde{\varphi}(t)\,\mathrm{d}t \\ &\quad + \sum_{j:j\neq i} \int_s^{\omega_x} \widetilde{R}_{ij}(t)\,\widetilde{\mu}_{ij}(t)\,\mathrm{d}t, \quad i \in \mathcal{S},\ s \geq s_1, \\ (\widetilde{\varphi}(t), \widetilde{\mu}_J(t)) &= \underset{(f,m_J)\in M(t)}{\operatorname{argmax}} \Bigg\{ \sum_{(j,k)\in J} v(s_1,t)\,p_{i_1 j}(s_1,t)\,\widetilde{R}_{jk}(t)\,m_{jk} \\ &\quad - \sum_{j\in\mathcal{S}} v(s_1,t)\,p_{i_1 j}(s_1,t)\,\widetilde{V}_j(t)\,f \Bigg\}, \quad t \geq s_1, \end{aligned} \tag{4.2}$$

with initial values $V_i(\omega_x) = 0$ for all $i \in \mathcal{S}$. Further assume that for each $t > s_1$ the argmax *is constant with respect to all factors $v(s_1,t)\,p_{i_1 j}(s_1,t) \in \{v(s_1,t;\varphi)\,p_{i_1 j}(s_1,t;\mu_J) \mid (\varphi,\mu_J) \in M\}$. Then we have*

$$\widetilde{V}_{i_1}(s_1) = V_{i_1}(s_1; \widetilde{\varphi}, \widetilde{\mu}_J) = \sup_{(\varphi,\mu_J)\in M} V_{i_1}(s_1; \varphi, \mu_J),$$

and, thus, $(\widetilde{\varphi}, \widetilde{\mu}_J)$ is a worst-case scenario for $V_{i_1}(s_1)$ with respect to M.

The crucial assumption in this proposition is that the argmax property of $(\widetilde{\varphi}, \widetilde{\mu}_J)$ holds with respect to all alternative scenarios $(\varphi, \mu_J) \in M$ independently of $v(s_1,t)\,p_{i_1 j}(s_1,t)$. This is a rather strong condition, but we will see that there are a number of examples of practical interest where this condition is satisfied.

Proof. Under the assumptions of the proposition, we have at any time $t > s_1$ that

$$\begin{aligned} 0 \leq \max_{(f,m_J)\in M(t)} \Bigg\{ &\sum_{(j,k)\in J} v(s_1,t)\,p_{i_1 j}(s_1,t)\,\widetilde{R}_{jk}(t)\,\big(m_{jk} - \mu_{jk}(t)\big) \\ &- \sum_{j\in\mathcal{S}} v(s_1,t)\,p_{i_1 j}(s_1,t)\,\widetilde{V}_j(t)\,\big(f - \varphi(t)\big) \Bigg\} \end{aligned}$$

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

for all $(\varphi, \mu_J) \in M$. Applying (3.3), we then get

$$
\begin{aligned}
&V_{i_1}(s_1; \widetilde{\varphi}, \widetilde{\mu}_J) - V_{i_1}(s_1; \varphi, \mu_J) \\
&= \int_{s_1}^{\omega_x} \Bigg(\sum_{(j,k)\in J} v(s_1,t)\, p_{i_1 j}(s_1,t)\, \widetilde{R}_{jk}(t) \big(\widetilde{\mu}_{jk}(t) - \mu_{jk}(t)\big) \\
&\quad - \sum_{j\in\mathcal{S}} v(s_1,t)\, p_{i_1 j}(s_1,t)\, \widetilde{V}_j(t) \big(\widetilde{\varphi}(t) - \varphi(t)\big) \Bigg) \mathrm{d}t \\
&= \int_{s_1}^{\omega_x} \max_{(f,m_J)\in M(t)} \Bigg\{ \sum_{(j,k)\in J} v(s_1,t)\, p_{i_1 j}(s_1,t)\, \widetilde{R}_{jk}(t) \big(m_{jk} - \mu_{jk}(t)\big) \\
&\quad - \sum_{j\in\mathcal{S}} v(s_1,t)\, p_{i_1 j}(s_1,t)\, \widetilde{V}_j(t) \big(f - \varphi(t)\big) \Bigg\} \mathrm{d}t \\
&\geq 0
\end{aligned}
$$

for all $(\varphi, \mu_J) \in M$. ■

The question is now if a scenario $(\widetilde{\varphi}, \widetilde{\mu}_J) \in M$ with the properties of Proposition 4.1 really exists. We will give an existence result, but note that we assume stronger conditions than in Proposition 4.1.

Proposition 4.2 (Existence). *Let $M \subset L_1^{1+|J|}([0,\omega_x])$ be a set of intensity vectors (φ, μ_J) where the slices $M(t)$ are compact subsets of $\mathbb{R}^{1+|J|}$ and the function*

$$t \mapsto \sup_{(\varphi(t),\mu_J(t))\in M(t)} \|(\varphi(t), \mu_J(t))\| \tag{4.3}$$

is integrable. Further assume that for each time t the quantity

$$\operatorname*{argmax}_{(f,m_J)\in M(t)} \Bigg\{ \sum_{(j,k)\in J} v(s,t)\, p_{ij}(s,t)\, R_{jk}(t)\, m_{jk} - \sum_{j\in\mathcal{S}} v(s,t)\, p_{ij}(s,t)\, V_j(t)\, f \Bigg\} \tag{4.4}$$

is constant with respect to all times $s \in [0,t)$, all states $i \in \mathcal{S}$, and all factors $v(s,t)\, p_{ij}(s,t) \in \{v(s,t;\varphi)\, p_{ij}(s,t;\mu_J) \mid (\varphi,\mu_J) \in M\}$. Then, the sequence of valuation bases $(\varphi^{(n)}, \mu_J^{(n)})_{n\geq 0}$ defined by an arbitrary but fixed starting point $(\varphi^{(0)}, \mu_J^{(0)}) \in M$, an arbitrary but fixed initial state $X_0 = i_0 \in \mathcal{S}$, and the recursion

$$
\begin{aligned}
(\varphi^{(n+1)}(t), \mu_J^{(n+1)}(t)) \in \operatorname*{argmax}_{(f,m_J)\in M(t)} \Bigg\{ &\sum_{(j,k)\in J} v^{(n)}(0,t)\, p_{i_0 j}^{(n)}(0,t)\, R_{jk}^{(n)}(t)\, m_{jk} \\
&- \sum_{j\in\mathcal{S}} v^{(n)}(0,t)\, p_{i_0 j}^{(n)}(0,t)\, V_j^{(n)}(t)\, f \Bigg\}
\end{aligned}
\tag{4.5}
$$

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

creates a series of prospective reserves that converge to a limit $\lim_{n\to\infty} V_i(t;\varphi^{(n)},\mu_J^{(n)}) =: V_i^*(t)$ *for all* $t \in [0,\omega_x]$ *and* $i \in \mathcal{S}$. *With defining* $R_{jk}^*(t) := b_{ij}(t) + V_j^*(t) - V_i^*(t)$, *the intensity vectors*

$$(\widetilde{\varphi}(t), \widetilde{\mu}_J(t)) = \underset{(f,m_J)\in M(t)}{\operatorname{argmax}} \left\{ \sum_{(j,k)\in J} v(s_1,t)\, p_{i_1 j}(s_1,t)\, R_{jk}^*(t)\, m_{jk} - \sum_{j\in\mathcal{S}} v(s_1,t)\, p_{i_1 j}(s_1,t)\, V_j^*(t)\, f \right\}$$

satisfy the assumptions of Proposition 4.1 *for all* $s_1 \geq 0$ *and all* $i_1 \in \mathcal{S}$.

In Proposition 4.1, we assumed that the argmax in (4.2) does not depend on the factors $v(s,t)\, p_{ij}(s,t)$ for $s = s_1$ and $i = i_1$. In Proposition 4.2, we extended that assumption for the argmax in (4.4) to all $s \geq 0$ and all $i \in \mathcal{S}$. In effect, we get the maximality of $V_i(s;\widetilde{\varphi},\widetilde{\mu}_J)$ not only at $s = s_1$ and $i = i_1$ but for all $s \geq 0$ and all $i \in \mathcal{S}$. That means that the worst-case scenario is time consistent in the sense that it does not change when we look at the same policy at a later point in time. Under the weaker assumptions of Proposition 4.1 that is not necessarily the case.

Proof. Because of the independence of (4.4) from the factors $v(s,t)\, p_{ij}(s,t)$ for all $s \geq 0$ and $i \in \mathcal{S}$, we can substitute (4.5) with

$$(\varphi^{(n+1)}(t), \mu_J^{(n+1)}(t)) \in \underset{(f,m_J)\in M(t)}{\operatorname{argmax}} \left\{ \sum_{(j,k)\in J} v^{(n+1)}(0,t)\, p_{i_0 j}^{(n+1)}(0,t)\, R_{jk}^{(n)}(t)\, m_{jk} - \sum_{j\in\mathcal{S}} v^{(n+1)}(0,t)\, p_{i_0 j}^{(n+1)}(0,t)\, V_j^{(n)}(t)\, f \right\}.$$

Multiplying both sides of (3.3) with -1 and identifying (φ,μ_J) and $(\widetilde{\varphi},\widetilde{\mu}_J)$ with $(\varphi^{(n+1)},\mu_J^{(n+1)})$ and $(\varphi^{(n)},\mu_J^{(n)})$, respectively, we obtain

$$\begin{aligned} & V_i(s;\varphi^{(n+1)},\mu_J^{(n+1)}) - V_i(s;\varphi^{(n)},\mu_J^{(n)}) \\ &= \int_s^{\omega_x} \Bigg(\sum_{(j,k)\in J} v^{(n+1)}(s,t)\, p_{ij}^{(n+1)}(s,t)\, R_{jk}^{(n)}(t) \left(\mu_{jk}^{(n+1)}(t) - \mu_{jk}^{(n)}(t)\right) \\ &\quad - \sum_{j\in\mathcal{S}} v^{(n+1)}(s,t)\, p_{ij}^{(n+1)}(s,t)\, V_j^{(n)}(t) \left(\varphi^{(n+1)}(t) - \varphi^{(n)}(t)\right) \Bigg) \mathrm{d}t \end{aligned}$$

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

$$= \int_s^{\omega_x} \max_{(f,m_J)\in M(t)} \Bigg\{ \sum_{(j,k)\in J} v^{(n+1)}(s,t)\, p_{ij}^{(n+1)}(s,t)\, R_{jk}^{(n)}(t) \left(m_{jk} - \mu_{jk}^{(n)}(t)\right)$$

$$- \sum_{j\in\mathcal{S}} v^{(n+1)}(s,t)\, p_{ij}^{(n+1)}(s,t)\, V_j^{(n)}(t) \left(f - \varphi^{(n)}(t)\right) \Bigg\} \mathrm{d}t$$

$$\geq 0.$$

That means that the sequences $V_i(t; \varphi^{(n)}, \mu_J^{(n)})$ are monotonically increasing. Using the representation formula (3.1), we can show that $V_i(t; \varphi, \mu)$ is uniformly bounded for all $i \in \mathcal{S}$, $t \in [0, \omega_x]$, and $(\varphi, \mu) \in M$, since the payments functions B_j and b_{jk} have finite variation, the discounting factor is bounded because of the integrability of (4.3), and all transition intensities are bounded by (4.3). Thus, the limits $\lim_{n\to\infty} V_i(t; \varphi^{(n)}, \mu_J^{(n)})$ exist pointwise on $[0, \omega_x]$ for all $i \in \mathcal{S}$, and also the limits of the sequences $R_{ij}(t; \varphi^{(n)}, \mu_J^{(n)}) = b_{ij}(s) + V_j(t; \varphi^{(n)}, \mu_J^{(n)}) - V_i(t; \varphi^{(n)}, \mu_J^{(n)})$ exist pointwise on $[0, \omega_x]$. Multiplying both sides of (3.3) with -1, we get that

$$V_i(s; \widetilde{\varphi}, \widetilde{\mu}_J) - V_i^*(s)$$

$$= V_i(s; \widetilde{\varphi}, \widetilde{\mu}_J) - \lim_{n\to\infty} V_i(s; \varphi^{(n)}, \mu_J^{(n)})$$

$$= \lim_{n\to\infty} \int_s^{\omega_x} \Bigg(\sum_{(j,k)\in J} \widetilde{v}(s,t)\, \widetilde{p}_{ij}(s,t)\, R_{jk}^{(n)}(t) \left(\widetilde{\mu}_{jk}(t) - \mu_{jk}^{(n)}(t)\right)$$

$$- \sum_{j\in\mathcal{S}} \widetilde{v}(s,t)\, \widetilde{p}_{ij}(s,t)\, V_j^{(n)}(t) \left(\widetilde{\varphi}(t) - \varphi^{(n)}(t)\right) \Bigg) \mathrm{d}t$$

$$= \lim_{n\to\infty} \int_s^{\omega_x} \Bigg(\sum_{(j,k)\in J} \widetilde{v}(s,t)\, \widetilde{p}_{ij}(s,t)\, R_{jk}^{*}(t) \left(\widetilde{\mu}_{jk}(t) - \mu_{jk}^{(n)}(t)\right)$$

$$- \sum_{j\in\mathcal{S}} \widetilde{v}(s,t)\, \widetilde{p}_{ij}(s,t)\, V_j^{*}(t) \left(\widetilde{\varphi}(t) - \varphi^{(n)}(t)\right) \Bigg) \mathrm{d}t$$

$$+ \lim_{n\to\infty} \int_s^{\omega_x} \Bigg(\sum_{(j,k)\in J} \widetilde{v}(s,t)\, \widetilde{p}_{ij}(s,t)\, (R_{jk}^{(n)}(t) - R_{jk}^{*}(t)) \left(\widetilde{\mu}_{jk}(t) - \mu_{jk}^{(n)}(t)\right)$$

$$- \sum_{j\in\mathcal{S}} \widetilde{v}(s,t)\, \widetilde{p}_{ij}(s,t)\, (V_j^{(n)}(t) - V_j^{*}(t)) \left(\widetilde{\varphi}(t) - \varphi^{(n)}(t)\right) \Bigg) \mathrm{d}t.$$

The integrand of the first integral is never negative because of the definition of $(\widetilde{\varphi}, \widetilde{\mu}_J)$. The second integral converges to zero because its integrand converges pointwise to zero and is uniformly bounded (see the arguments of above).

Hence, we obtain $V_i(s; \widetilde{\varphi}, \widetilde{\mu}_J) - V_i^*(s) \geq 0$. Analogously, we can show that

$$
\begin{aligned}
&V_i(s; \widetilde{\varphi}, \widetilde{\mu}_J) - V_i^*(s) \\
&\quad = \lim_{n\to\infty} \int_s^{\omega_x} \Bigg(\sum_{(j,k)\in J} \widetilde{v}(s,t)\, \widetilde{p}_{ij}(s,t)\, R_{jk}^{(n-1)}(t) \left(\widetilde{\mu}_{jk}(t) - \mu_{jk}^{(n)}(t)\right) \\
&\qquad - \sum_{j\in\mathcal{S}} \widetilde{v}(s,t)\, \widetilde{p}_{ij}(s,t)\, V_j^{(n-1)}(t) \left(\widetilde{\varphi}(t) - \varphi^{(n)}(t)\right) \Bigg) \mathrm{d}t \\
&\qquad + \lim_{n\to\infty} \int_s^{\omega_x} \Bigg(\sum_{(j,k)\in J} \widetilde{v}(s,t)\, \widetilde{p}_{ij}(s,t) \left(R_{jk}^{(n)}(t) - R_{jk}^{(n-1)}(t)\right)\left(\widetilde{\mu}_{jk}(t) - \mu_{jk}^{(n)}(t)\right) \\
&\qquad - \sum_{j\in\mathcal{S}} \widetilde{v}(s,t)\, \widetilde{p}_{ij}(s,t)\, (V_j^{(n)}(t) - V_j^{(n-1)}(t)) \left(\widetilde{\varphi}(t) - \varphi^{(n)}(t)\right) \Bigg) \mathrm{d}t.
\end{aligned}
$$

The integrand of the first integral is never positive because of the definition of $(\varphi^{(n)}, \mu_J^{(n)})$. The second integral converges to zero. Hence, we have $V_i(s; \widetilde{\varphi}, \widetilde{\mu}_J) - V_i^*(s) \leq 0$, and thus $V_i(s; \widetilde{\varphi}, \widetilde{\mu}_J) = V_i^*(s)$ for all states i and times s. That means that $(\widetilde{\varphi}, \widetilde{\mu}_J)$ satisfies the assumptions of Proposition 4.1 for each $s_1 \geq 0$ and $i_1 \in \mathcal{S}$. ■

We now turn to discuss the construction of worst-case scenarios for (4.1).

(a) First we look at the situation where the assumptions of Proposition 4.2 hold. The proof of Proposition 4.2 has a constructive form and does not only show the existence of a solution for (4.2) but also gives an iteration method for the numerical calculation of a solution.

Algorithm

(1) Choose a starting scenario $(\varphi^{(0)}, \mu_J^{(0)}) \in M$.

(2) Calculate a new scenario by using the iteration

$$
\begin{aligned}
(\varphi^{(n+1)}(t), \mu_J^{(n+1)}(t)) \in \underset{(f,m_J)\in M(t)}{\operatorname{argmax}} \Bigg\{ &\sum_{(j,k)\in J} v^{(n)}(0,t)\, p_{i_0 j}^{(n)}(0,t)\, R_{jk}^{(n)}(t)\, m_{jk} \\
&- \sum_{j\in\mathcal{S}} v^{(n)}(0,t)\, p_{i_0 j}^{(n)}(0,t)\, V_j^{(n)}(t)\, f \Bigg\}
\end{aligned}
\tag{4.6}
$$

for an arbitrary but fixed initial state $X_0 = i_0 \in \mathcal{S}$.

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

(3) Repeat step 2 until $\left| V_{i_1}(s_1; \varphi^{(n+1)}, \mu_J^{(n+1)}) - V_{i_1}(s_1; \varphi^{(n)}, \mu_J^{(n)}) \right|$ is below some error tolerance.

If the assumptions of Proposition 4.2 hold, we know that the algorithm always converges to a worst-case scenario $(\widetilde{\varphi}, \widetilde{\mu}_J) \in M$.

(b) Under the weaker assumptions of Proposition 4.1, we still have that the right-hand sides of both equations in (4.2) only depend on the future and not on the past. By replacing t in the second equation of (4.2) by s, which is possible according to the assumptions of Proposition 4.1, we can see (4.2) as a terminal value problem that can be solved backward starting from time $s = \omega_x$ and going backward to $s = s_1$. In other words, we use a dynamic programming approach. If we find a solution, it must be maximal because of Proposition 4.1. In practice, we typically do not find an analytical solution but we can then use standard numerical techniques for ordinary differential/integral equation systems.

The assumption of Proposition 4.1 that the argmax in (4.2) is constant with respect to the factors $v(s_1, t)\, p_{i_1 j}(s_1, t)$ means that we do not need to know the past before time t when calculating $(\widetilde{\varphi}(t), \widetilde{\mu}_J(t))$. This is the minimal assumption that we need in order to allow for dynamic programming leading to a deterministic solution.

(c) If we are in a situation where even the weaker assumptions of Proposition 4.1 do not hold, we can neither use dynamic programming as in (b), nor do we know if the algorithm in (a) really converges to a worst-case. In the optimization literature, a popular approach for dealing with difficult optimization problems is the gradient ascent method. However, this method yields a local maximum but not necessarily a global maximum. Nevertheless, in what follows, we briefly outline how a gradient ascent method looks here.

By applying the first-order Taylor expansion (3.4),

$$V_i(s; \widetilde{\varphi}, \widetilde{\mu}_J) = V_i(s; \varphi, \mu_J) + \int_s^{\omega_x} \nabla_{(\varphi,\mu_J)}(V_i(s))(t) \cdot \big((\widetilde{\varphi}(t), \widetilde{\mu}_J(t)) - (\varphi(t), \mu_J(t))\big)\, \mathrm{d}t + o(\|(\widetilde{\varphi}, \widetilde{\mu}) - (\varphi, \mu)\|),$$

we can locally increase $V_i(s; \varphi, \mu_J)$ by making a small step from (φ, μ_J) to $(\widetilde{\varphi}, \widetilde{\mu}_J) = (\varphi, \mu_J) + \varepsilon\, \nabla_{(\varphi,\mu_J)}(V_i(s))$ in the direction of the generalized gradient $\nabla_{(\varphi,\mu_J)}(V_i(s))$. The step size factor ε has to be sufficiently small. Iterating this idea leads us to the following gradient ascent method.

Algorithm

(1) Choose a Taylor center $(\varphi^{(0)}, \mu_J^{(0)}) \in M$.

(2) Calculate a new Taylor center by using the iteration

$$(\varphi^{(n+1)}(t), \mu_J^{(n+1)}(t)) = (\varphi^{(n)}(t), \mu_J^{(n)}(t)) + \varepsilon_n\, \nabla_{(\varphi^{(n)}, \mu_J^{(n)})}(V_{i_1}(s_1))(t)$$

for some small step size factor $\varepsilon_n > 0$ such that $(\varphi^{(n+1)}, \mu_J^{(n+1)}) \in M$.

(3) Repeat step 2 until $\left| V_{i_1}(s_1; \varphi^{(n+1)}, \mu_J^{(n+1)}) - V_{i_1}(s_1; \varphi^{(n)}, \mu_J^{(n)}) \right|$ is below some error tolerance.

If our series converges, the limit is a local maximum but not necessarily a global maximum. Therefore, this algorithm must be applied carefully.

5. Examples

We give some examples where the assumptions of Proposition 4.2 hold and some examples where at least the weaker assumptions of Proposition 4.1 are satisfied. That means that we can use both the recursion algorithm and the dynamic programming method. Apart from Example 5.3, where both methods are demonstrated, we will always use the latter method, since it has the advantage that we have to numerically solve only one differential equation system. In all numerical examples, the premiums are chosen as best estimate net premiums.

Example 5.1 (No dependence, interest rate fixed). Let the confidence band M have a fixed interest rate φ and rectangular slices $M(t) = \{\varphi(t)\} \times \left(\bigtimes_{(j,k)\in J} M_{jk}(t) \right)$ with compact sets $M_{jk}(t) \subset \mathbb{R}$,

$$M = \left\{ (\varphi, \mu_J) \middle| \varphi \text{ fixed}, \mu_{jk}(t) \in M_{jk}(t) \text{ for compact } M_{jk}(t) \subset \mathbb{R} \right\}.$$

Then, the argmax in (4.2) can be rewritten as

$$\begin{aligned} \widetilde{\varphi}(t) &= \varphi(t), \\ \widetilde{\mu}_{jk}(t) &= \underset{m_{jk} \in M_{jk}(t)}{\operatorname{argmax}} \left\{ \widetilde{R}_{jk}(t)\, m_{jk} \right\}, \quad (j,k) \in J, \end{aligned}$$

and for any time t the choice of $(\varphi(t), \mu_J(t))$ does not depend on the past. The special case with $M_{jk}(t) = [\underline{\mu}_{jk}(t), \overline{\mu}_{jk}(t)]$ for $\underline{\mu}_{jk}, \overline{\mu}_{jk} \in L_1([0, \omega_x])$ can be already found in Christiansen (2010), where also two numerical examples are given.

Example 5.2 (Exit intensities from the same states are dependent, interest rate fixed). We now expand the model setting of Example 5.1 to only partially rectangular slices $M(t) = \{\varphi(t)\} \times \left(\bigtimes_{j\in\mathcal{S}} M_j(t) \right)$ with compact sets $M_j(t) \subset \mathbb{R}^{|\mathcal{S}|-1}$,

$$\begin{aligned} M = \Big\{ (\varphi, \mu_J) \Big| \varphi \text{ fixed}, \mu_j(t) &= (\mu_{jk}(t))_{k:k\neq j} \\ &\in M_j(t) \text{ for compact } M_j(t) \subset \mathbb{R}^{|\mathcal{S}|-1} \Big\}. \end{aligned}$$

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

The argmax in (4.2) can be rewritten as

$$\widetilde{\varphi}(t) = \varphi(t),$$
$$\widetilde{\mu}_j(t) = \underset{m_j \in M_j(t)}{\operatorname{argmax}} \Big\{ \sum_{k:k\neq j} \widetilde{R}_{jk}(t)\, m_{jk} \Big\}, \quad j \in \mathcal{S},$$

and for any time t we still have independency of the choice of $(\varphi(t), \mu_J(t))$ from the past.

Numerical example. Consider a 30-year-old male who contracts a critical illness insurance that pays a lump sum of 10 000 in case of death or a lump sum of 7000 in case of disability, whichever occurs first. The contract terminates after the first claim occurs but at the latest at age 60. A constant premium of 47.5626 is paid yearly in advance. The relevant transitions are ad = "active to dead" and ai = "active to invalid/disabled". For the construction of $M(t) = \{\varphi(t)\} \times M_a(t)$, where $M_a(t) \subset \mathbb{R}^2$ is the set of admissible values for $(\mu_{ad}(t), \mu_{ai}(t))$, we use the model of Christiansen *et al.* (2012). Let $\mu_{ad}(t)$ and $\mu_{ai}(t)$ be stochastic processes that are constant in between years,

$$\mu_{ad}(x+t) = \mu_{ad}(x+[t]), \quad \mu_{ai}(x+t) = \mu_{ai}(x+[t]),$$

with a representation of the form

$$\ln(\mu_{ad}(x+k)) = \alpha_{ad}(x+k) + \beta_{ad}(x+k)\,\gamma_{ad}(k),$$
$$\ln(\mu_{ai}(x+k)) = \alpha_{ai}(x+k) + \beta_{ai}(x+k)\,\gamma_{ai}(k),$$

for all integer values k, where $\alpha_{ad}(x+k)$, $\alpha_{ai}(x+k)$, $\beta_{ad}(x+k)$, $\beta_{ai}(x+k)$ are deterministic and the differences $(\gamma_{ad}(k) - \gamma_{ad}(k-1), \gamma_{ai}(k) - \gamma_{ai}(k-1))^T$ are independent and identically normal distributed random vectors with mean $(-0.037238, -0.072526)^T$ and variance–covariance matrix

$$\Sigma = \begin{pmatrix} 0.001392 & 0.000770 \\ 0.000770 & 0.008808 \end{pmatrix}.$$

Note that the model of Christiansen *et al.* (2012) shows an outlier at age 50 and that we adopt their model as it is without making corrections. One can easily show that also $(\ln(\mu_{ad}(x+k)), \ln(\mu_{ai}(x+k)))^T$ is a normal distributed random vector with some mean m_k and some covariance matrix Σ_k. Now we define the confidence area $M_a(t)$ as the set where the probability density of $(\ln(\mu_{ad}(t)), \ln(\mu_{ai}(t)))^T$ is at or above a certain level, that is

$$M_a(t) = \left\{ (x_1, x_2)^T \in \mathbb{R}^2 \,\middle|\, \left(\begin{pmatrix} \ln(x_1) \\ \ln(x_2) \end{pmatrix} - m_{[t]}\right)^T \Sigma_{[t]}^{-1} \left(\begin{pmatrix} \ln(x_1) \\ \ln(x_2) \end{pmatrix} - m_{[t]}\right) \le r^2 \right\} \tag{5.1}$$

for some constant $r^2 > 0$. (Note that for a two-dimensional normal random vector X with mean m and covariance matrix Σ, we always have that

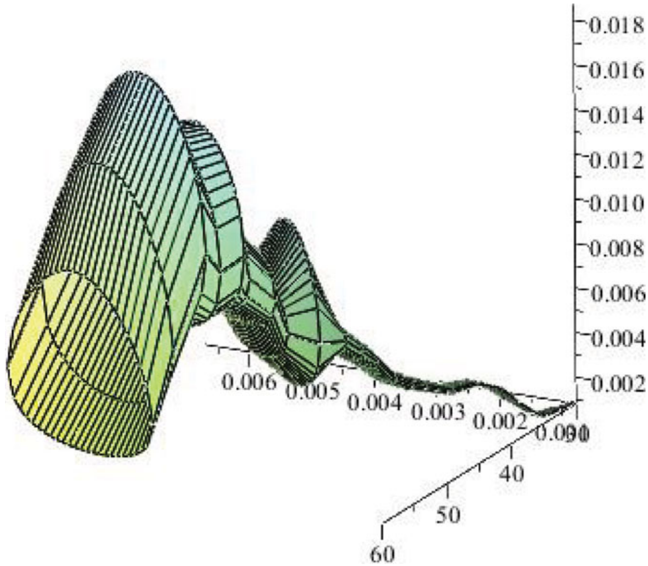

FIGURE 1: Confidence band M_a of Example 5.2 for age span 30–60 (front axis) for mortality intensity μ_{ad} (left axis) and disability intensity μ_{ai} (vertical axis).

$(X-m)^T\Sigma^{-1}(X-m)$ has a chi-square distribution with 2 degrees of freedom.) For any fixed t, the calculation of

$$(\widetilde{\mu}_{ad}(t), \widetilde{\mu}_{ai}(t)) = \underset{(m_{ad}, m_{ai}) \in M_a(t)}{\operatorname{argmax}} \left\{ \widetilde{R}_{ad}(t)\, m_{ad} + \widetilde{R}_{ai}(t)\, m_{ai} \right\}$$

is a linear optimization problem and there is always a maximal solution on the boundary $\partial M_a(t)$ of $M_a(t)$. ($\partial M_a(t)$ is the subset of $M_a(t)$ where we have equality in (5.1).) If $M_a(t)$ is strictly convex, we can find the argmax as explained in the Appendix. Figure 1 shows the confidence band $t \mapsto M_a(t)$ for age span 30–60 and $r^2 = 3$. Figures 3 and 4 show the worst-case valuation basis $(\widetilde{\mu}_{ad}, \widetilde{\mu}_{ai}) \in M$. By comparing the worst-case with the best estimate, the latter being given by $\Sigma = 0$, we see that the worst-case scenario increases the prospective reserve in state active at contract time zero from 0 to 130.96, see Figure 2. A crucial observation from Figure 3 is that, taking dependence into account, the "optimal" transition rates are not necessarily to be found on the univariate boundaries.

Example 5.3 (Multiple causes of decrement model with dependence on interest rates). Assume that $\mathcal{S} = \{a, d_1, ..., d_m\}$, that only the transitions $(a, d_1), ..., (a, d_m)$ are possible, and that annuity benefits and premiums are only payable in state a = "active". Then for arbitrarily shaped confidence bands of the form

$$M = \Big\{(\varphi, \mu_J)\Big|(\varphi(t), \mu_J(t)) \in M(t) \text{ for compact } M(t) \subset \mathbb{R}^{1+|J|},$$
$$\mu_{jk}(t) = 0 \text{ for all } (j,k) \neq (a, d_1), ..., (a, d_m)\Big\},$$

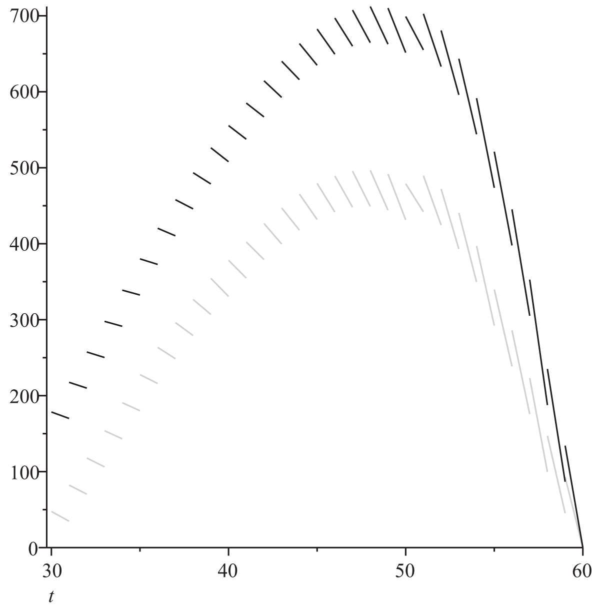

FIGURE 2: Best estimate (lower gray curve) and worst-case (upper black curve) prospective reserve for Example 5.2 in state active at ages 30–60.

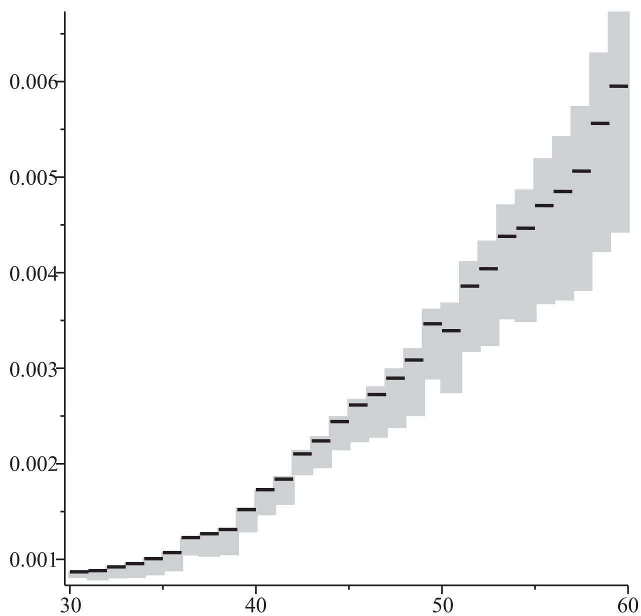

FIGURE 3: Worst-case mortality intensity $\widetilde{\mu}_{ad}$ (black) of Example 5.2 and maximal and minimal possible mortality intensities (gray).

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

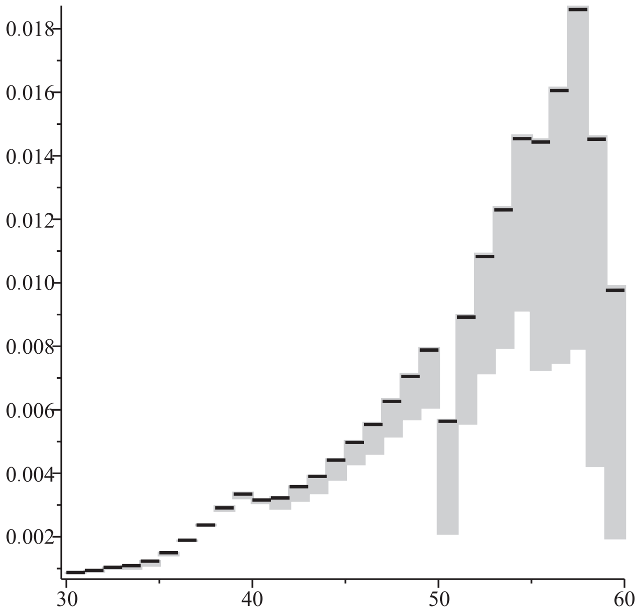

FIGURE 4: Worst-case disability intensity $\widetilde{\mu}_{ai}$ (black) of Example 5.2 and maximal and minimal possible disability intensities (gray).

the argmax in (4.2) can be transformed to

$$(\widetilde{\varphi}(t), \widetilde{\mu}_J(t)) = \underset{(f, m_J) \in M(t)}{\operatorname{argmax}} \left\{ \sum_{j \in \{d_1, \ldots, d_m\}} \widetilde{R}_{aj}(t)\, m_{aj} - \widetilde{V}_a(t)\, f \right\}.$$

Hence, for any time t, the choice of $(\varphi(t), \mu_J(t))$ is independent from the past. We point out that this example also comprises the surrender model of Section 2.

Numerical example 1. Consider a temporary life insurance that starts at age 35 and that pays a lump sum of 10 000 in case of death of the policyholder before termination of the contract at age 65. A constant premium of 28.3435 is paid yearly in advance. For the construction of $M(t) \subset \mathbb{R}^2$ (the set of admissible values for $(\varphi(t), \mu_{ad}(t))$), we use the following model. Let $\varphi(t)$ and $\mu_{ad}(t)$ be stochastic processes that are constant in between years,

$$\varphi(t) = \varphi([t]), \quad \mu_{ad}(x+t) = \mu_{ad}(x+[t]),$$

where the latter has a representation of the form

$$\ln(\mu_{ad}(x+k)) = \alpha(x+k) + \beta(x+k)\,\gamma(k)$$

for all integer values k. Let $(\Delta\varphi(k), \Delta\gamma(k)) = (\varphi(k) - \varphi(k-1), \gamma(k) - \gamma(k-1))^T$ be independent and identically normal distributed random vectors with mean $(0, \theta)$ and covariance matrix Σ. We choose $\sqrt{\mathrm{Var}(\Delta\varphi(k))} = 0.002$ and use for $\alpha(x+k)$, $\beta(x+k)$, $\gamma(0)$, θ, and $\sqrt{\mathrm{Var}(\Delta\gamma(k))}$ the values of Christiansen and Denuit (2010). If we add some correlation factor $corr(\Delta\varphi(k), \Delta\gamma(k)) = \rho$, then the covariance matrix Σ is uniquely determined. One can easily show that $(\varphi(k), \ln(\mu_{ad}(x+k)))^T$ is also a normal distributed random vector with mean $m_k = (\ln(1.035), \alpha(x+k) + \beta(x+k)(\gamma(0) + k\theta))^T$ and covariance matrix

$$\Sigma_k = \begin{pmatrix} (k+1)\sigma^2_{\Delta\varphi(k)} & (k+1)\sigma_{\Delta\varphi(k)}\beta(x+k)\sigma_{\Delta\gamma(k)}\rho \\ (k+1)\sigma_{\Delta\varphi(k)}\beta(x+k)\sigma_{\Delta\gamma(k)}\rho & \beta(x+k)^2(k+1)\sigma^2_{\Delta\gamma(k)} \end{pmatrix}.$$

Now we define confidence areas $M(t)$ for $(\varphi(t), \mu_{ad}(t))^T$ as sets where the probability density of $(\varphi(t), \ln(\mu_{ad}(t)))^T$ is at or above a certain level, that is

$$M(t) = \left\{(x_1, x_2) \in \mathbb{R}^2 \middle| (\begin{pmatrix} x_1 \\ \ln(x_2) \end{pmatrix} - m_{[t]})^T \Sigma^{-1}_{[t]} (\begin{pmatrix} x_1 \\ \ln(x_2) \end{pmatrix} - m_{[t]}) \leq r^2 \right\} \quad (5.2)$$

for some constant $r^2 > 0$. (Recall that for a two-dimensional normal random vector X with mean m and covariance matrix Σ, we always have that $(X-m)^T\Sigma^{-1}(X-m)$ has a chi-square distribution with 2 degrees of freedom.) For any fixed t, the calculation of

$$(\widetilde{\varphi}(t), \widetilde{\mu}_{ad}(t)) = \underset{(f, m_{ad}) \in M_t}{\operatorname{argmax}} \left\{ \widetilde{R}_{ad}(t)\, m_{ad} - \widetilde{V}_a(t)\, f \right\}$$

is a linear optimization problem, and there is always an optimal solution at the boundary $\partial M(t)$ of $M(t)$. ($\partial M(t)$ is the subset of $M(t)$ where we have equality in (5.2).) If $M(t)$ is strictly convex, we can find the argmax as explained in the Appendix. Figure 5 shows the confidence band M for age span 35–60, $\rho = 0.25$, and $r^2 = 3$. The reason that we exemplify our method with $\rho = 0.25$ is that the Fifth Quantitative Impact Study (2010) of Solvency II also uses a correlation assumption of 0.25 when aggregating market risk and life underwriting risk, although it is not clear if this coefficient can be projected to our model. Figures 7 and 8 show the worst-case valuation basis $(\widetilde{\varphi}, \widetilde{\mu}_{ad}) \in M$, and Figure 6 shows the corresponding prospective reserves in state active. The results were obtained by using the dynamic programming method. If we use the recursion algorithm (4.6) with the center of the confidence band as starting scenario, we get a good approximation already after two steps. The corresponding prospective reserves in state active at the beginning of the contract time are

starting scenario	first iter. step	second iter. step	limit / worst-case/ dyn. progr.
0.0199	79.5890	81.7640	81.7642.

In contrast to the case with rectangular confidence bands, where the worst-case interest intensity is always equal to the lower bound, we here have a worst-case interest intensity that is significantly above the minimal values. Toward the end of the contract period, we are even above the mean. The lesson to learn is

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

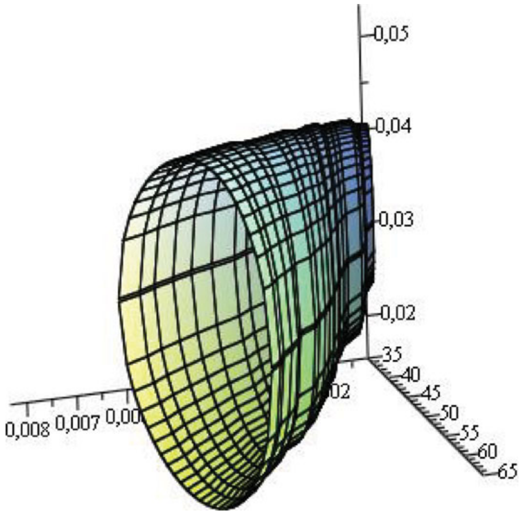

FIGURE 5: Confidence band M of Example 5.3 for ages 35–65 (right axis) for interest intensity φ (vertical axis) and mortality intensity μ_{ad} (left axis).

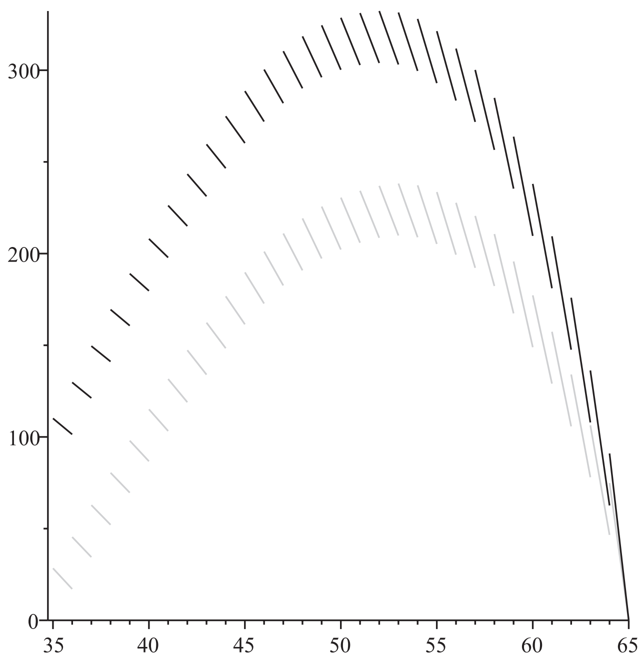

FIGURE 6: Best estimate (lower gray curve) and worst-case (upper black curve) prospective reserve of Example 5.3 in state active at ages 35–65

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

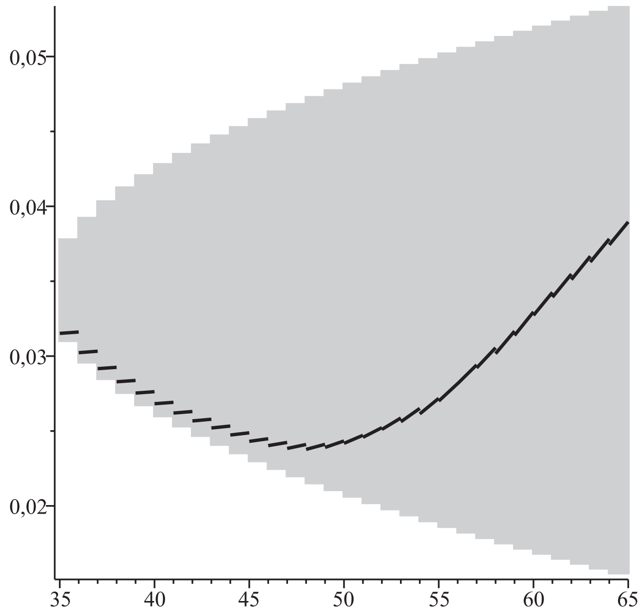

FIGURE 7: Worst-case interest intensity $\widetilde{\varphi}$ (black) of Example 5.3 and maximal and minimal possible interest intensities (gray).

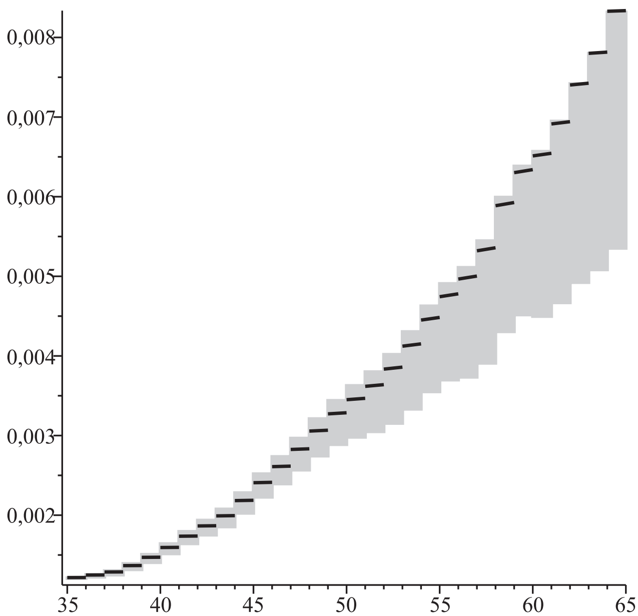

FIGURE 8: Worst-case mortality intensity $\widetilde{\mu}_{ad}$ (black) of Example 5.3 and maximal and minimal possible mortality intensities (gray).

that the optimization problem does not have a so-called corner solution due to the mutual dependence between interest rates and mortality.

Numerical example 2. Consider a unit-linked endowment insurance with surrender guarantees and a state space of $\mathcal{S} = \{a, d, s\}$. The contract starts at age 35, and the policyholder gets a lump sum of

$$b_{ad}(t) = 10\,000 \max\{a(t)\, e^{\int_0^t \varphi^g(u)\,\mathrm{d}u}, e^{\int_0^t \varphi(u)\,\mathrm{d}u}\}, \quad 0 < t \leq 30,$$

in case of death before age 65 or a survival payment of

$$\Delta B_a(30) = 10\,000 \max\{a(30)\, e^{\int_0^{30} \varphi^g(u)\,\mathrm{d}u}, e^{\int_0^{30} \varphi(u)\,\mathrm{d}u}\}$$

at age 65. The function φ is the interest rate of the investment (a priori unknown and reduced by a fee, if any), φ^g is a deterministic minimal interest rate, and $a(t)$ is the fraction of the premium guaranteed to yield the minimal interest rate φ^g. Let $a(t) = \min\{1 - 0.015\,t, 0.85\}$ and $\varphi^g = \ln(1.035)$. If the contract is terminated early, a surrender value of

$$b_{as}(t) = 10\,000\,(1 - c(t))\, e^{\int_0^t \varphi^g(u)\,\mathrm{d}u}$$

is paid, where $c(t)$ is the surrender penalty, which is here set to $c(t) = \max\{0.01\,(5 - t), 0\}$. For μ_{ad} we choose the best estimate of $\ln(\mu_{ad}(x + k))$ according to the previous example. In this example, for all scenarios $(\varphi, \mu_{ad}, \mu_{as})$ and $(\varphi', \mu_{ad}, \mu_{as})$ with $\varphi(t) \geq \varphi'(t)$ for all t, we have that

$$V_a(0; \varphi, \mu_{ad}, \mu_{as}) \leq V_a(0; \varphi', \mu_{ad}, \mu_{as}) \tag{5.3}$$

since the discounted payments $v(0, t)\, b_{ad}(t) = \max\{v(0, t)\, a(t)\, e^{\int_0^t \varphi^g(u)\,\mathrm{d}u}, 1\}$, $v(0, 30)\, \Delta B_a(30) = \max\{v(0, 30)\, a(30)\, e^{\int_0^{30} \varphi^g(u)\,\mathrm{d}u}, 1\}$, and $v(0, t)\, b_{as}(t)$ are non-increasing in φ. A better than expected performance of the investment does not increase the prospective reserve here because the increased benefits can be fully paid from the excess return. At first, let

$$M(t) := [\ln(1.015), \ln(1.100)] \times \{\mu_{ad}(t)\} \times [0.005, 0.100].$$

Because of (5.3), for the worst-case calculation we can replace $M(t)$ by $M(t) = \{\ln(1.015)\} \times \{\mu_{ad}(t)\} \times [0.005, 0.100]$, which implies that $a(t) \exp\{\int_0^t \varphi^g(u)\,\mathrm{d}u\} \geq \exp\{\int_0^t \varphi(u)\,\mathrm{d}u\}$ for all t so that $b_{ad}(t)$ and $\Delta B_a(30)$ do not depend on φ anymore. By applying the dynamic programming method, we get that

$$\varphi(t) = \ln(1.015)\,, \quad \mu_{as}(t) = 0.005\,\mathbf{1}_{[35,58.892)}(t) + 0.1\,\mathbf{1}_{[58.892,65]}(t)$$

is the worst-case scenario. The corresponding prospective reserve at the beginning of the contract in state active is 15 320.69. In a second step, we assume that the surrender intensity depends on the interest rate in such a way that in case of a low interest rate the surrender activity is low as well. We model that by cutting

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

off from $M(t)$ a neighborhood of the edge $(\ln(1.015), \mu_{ad}(t), 0.100)$, namely the triangle

$$D(t) := conv\{(\ln(1.015), \mu_{ad}(t), 0.005), (\ln(1.030), \mu_{ad}(t), 0.100), (\ln(1.015), \mu_{ad}(t), 0.100)\}.$$

We are looking for a worst-case scenario in $M'(t) := M(t) \setminus D(t)$. Because of (5.3), for the worst-case calculation we can replace $M'(t)$ by the line

$$conv\{(\ln(1.015), \mu_{ad}(t), 0.005), (\ln(1.030), \mu_{ad}(t), 0.100)\},$$

where *conv* means the convex hull. Analogous to the above, in the reduced set $b_{ad}(t)$ and $\Delta B_a(30)$ do not depend on φ. By applying the dynamic programming method, we get that

$$\varphi(t) = \ln(1.015)\, \mathbf{1}_{[35,64.109)}(t) + \ln(1.030)\, \mathbf{1}_{[64.109,65]}(t),$$
$$\mu_{as}(t) = 0.005\, \mathbf{1}_{[35,64.109)}(t) + 0.1\, \mathbf{1}_{[64.109,65]}(t)$$

is the worst-case scenario with respect to M'. The corresponding prospective reserve at the beginning of the contract in state active is 14 785.71. Compared with the worst-case with respect to M that is a difference of −534.98.

Example 5.4 (Fixed transition intensities, unknown interest rates). If all transition intensities are fixed and only the interest intensity may fluctuate,

$$M = \left\{(\varphi, \mu_J) \middle| \varphi(t) \in M_\varphi(t) \text{ for compact } M_\varphi(t) \subset \mathbb{R}, \mu_J \text{ fixed}\right\},$$

then the argmax in (4.2) can be transformed to

$$\widetilde{\varphi}(t) = \operatorname*{argmax}_{f \in M_\varphi(t)} \left\{ -\sum_{j \in \mathcal{S}} p_{i_1 j}(s_1, t)\, \widetilde{V}_j(t)\, f \right\},$$
$$\widetilde{\mu}_J(t) = \mu_J(t).$$

The factors $p_{i_1 j}(s_1, t)$ are fixed since μ_J is fixed, and we may apply Proposition 4.1. However, the stronger assumption (4.4) of Proposition 4.2 is not met, because the argmax can vary with respect to i_1 and s_1. In order to obtain an existence result, we use the following trick. Instead of maximizing the prospective reserve $V_{i_1}(s_1)$, we use the property $V_{i_1}(s_1) = V_{i_1,s_1}(s_1)$ and maximize

$$V_{i_1,s_1}(t) := \sum_{j \in \mathcal{S}} p_{i_1 j}(s_1, t)\, V_j(t)$$

by using the Bellman equation

$$\widetilde{V}_{i_1,s_1}(s) = \sum_{j\in\mathcal{S}} \int_{(s,\omega_x]} p_{i_1 j}(s_1,t)\,\mathrm{d}B_j(t) - \int_s^{\omega_x} \widetilde{V}_{i_1,s_1}(t)\,\widetilde{\varphi}(t)\,\mathrm{d}t + \sum_{(j,k)\in J} \int_s^{\omega_x} p_{i_1 j}(s_1,t)\,b_{jk}(t)\,\mu_{jk}(t)\,\mathrm{d}t, \tag{5.4}$$

$$\widetilde{\varphi}(t) = \underset{f\in M_\varphi(t)}{\operatorname{argmax}} \Big\{ - \widetilde{V}_{i_1,s_1}(t)\, f \Big\},$$

which can be derived from (3.5) similar to (4.2). Now we follow the ideas of Proposition 4.2 and obtain the existence of a solution $\widetilde{V}_{i_1,s_1}(t) = V_{i_1,s_1}(t;\widetilde{\varphi},\widetilde{\mu})$ of (5.4) on $[s_1,\omega_x]$ that satisfies $V_{i_1,s_1}(t;\widetilde{\varphi},\widetilde{\mu}) \geq V_{i_1,s_1}(t;\varphi,\mu)$ for all scenarios $(\varphi,\mu_J)\in M$ and any $t \geq s_1$.

Typically, insurance contracts are designed in such a way that all prospective reserves are non-negative. Then, V_{i_1,s_1} is also non-negative and we get that the minimal interest intensity is a worst-case scenario.

Example 5.5 (Disability model with dependent death intensities, dependence with interest rates). Assume that $\mathcal{S} = \{a =$ "active", $i =$ "invalid/disabled", $d =$ "dead"$\}$, that there is no recovery, and that $\mu_{ai}(t)$ and the difference $d(t) = \mu_{id}(t) - \mu_{ad}(t)$ are arbitrary but fixed. Then, for any confidence band of the form

$$M = \Big\{(\varphi,\mu_J)\Big|(\varphi(t),\mu_J(t)) \in M(t) \text{ for compact } M(t) \subset \mathbb{R}^{1+|J|}, \quad \mu_{ia}(t) = 0\,,\, \mu_{id}(t) = \mu_{ad}(t) + d(t),\, \mu_{ai}(t) \text{ fixed}\Big\},$$

the argmax in (4.2) can be transformed to

$$(\widetilde{\varphi}(t),\widetilde{\mu}_J(t)) = \underset{(f,m_J)\in M(t)}{\operatorname{argmax}} \bigg\{ \widetilde{R}_{ad}(t)\,m_{ad} + \frac{p_{ai}(s_1,t)}{p_{aa}(s_1,t)}\widetilde{R}_{id}(t)\,m_{id} - \Big(\widetilde{V}_a(t) + \frac{p_{ai}(s_1,t)}{p_{aa}(s_1,t)}\widetilde{V}_i(t)\Big)\, f \bigg\}.$$

Since we have

$$\frac{p_{ai}(s_1,t)}{p_{aa}(s_1,t)} = \int_{s_1}^t \mu_{ai}(u)\, e^{\int_u^t (\mu_{ai}(v)-d(v))\,\mathrm{d}v}\mathrm{d}u$$

and μ_{ai} and d are assumed to be fixed, we have that $p_{ai}(s_1,t)/p_{aa}(s_1,t)$ is independent from the choice of $(\varphi,\mu_J)\in M$, and we may apply Proposition 4.1. (The assumptions of Proposition 4.2 are not necessarily satisfied.)

Numerical example. In addition to the temporary life insurance of Example 5.3, we add some disability benefits. In the case of disability, the policyholder is freed from the premium and gets a yearly disability annuity of 1000 in arrears.

https://doi.org/10.1017/asb.2013.16 Published online by Cambridge University Press

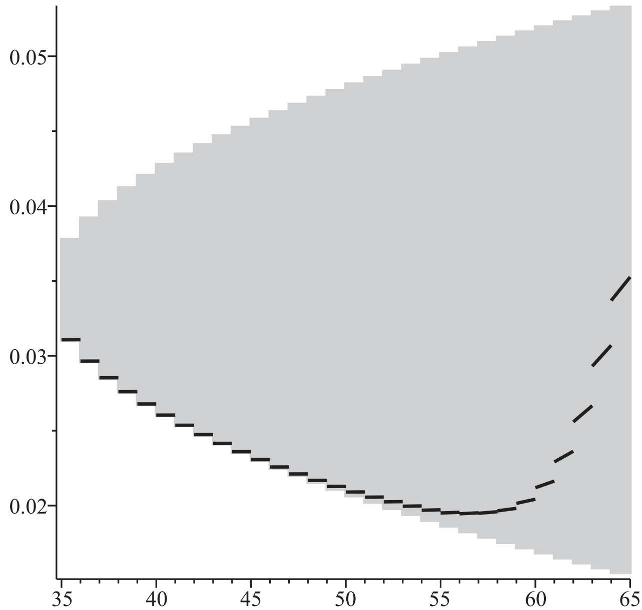

FIGURE 9: Worst-case interest intensity $\widetilde{\varphi}$ (black) of Example 5.5 and maximal and minimal possible interest intensities (gray).

The disability intensity is given by

$$\mu_{ai}(t) = 0.0004 + 0.000003467368\, e^{0.1381551\, t}.$$

Let the difference between mortality in state disabled and in state active be $d(t) = 0.0005$. The constant yearly premium is 65.4756 and was calculated by using the principle of equivalence on the basis of the best estimate, the latter being defined by $\Sigma = 0$. For $(\varphi(t), \mu_{ad}(t))$ we take the confidence band from Example 5.3. Figures 9 and 10 show the worst-case valuation basis $(\widetilde{\varphi}, \widetilde{\mu}_{ad})$. The worst-case scenario $(\widetilde{\varphi}, \widetilde{\mu}_{ad})$ increases the prospective reserve in state active at contract time zero from 0 to 164.77. Note from Figures 9 and 10 that, at certain ages, both the interest rate and mortality intensity are found away from their boundaries.

6. FINAL REMARKS

This paper explains how one can find deterministic worst-case scenarios when the interest and transition rates are mutually dependent. Such worst-case scenarios are useful for the identification and assessment of risks in life insurance, in particular for the calculation of first-order bases and SCRs. In the special case of Solvency II, our calculations deal with interest rate and insurance risk only, which can be motivated by either assuming (approximately) risk-free in-

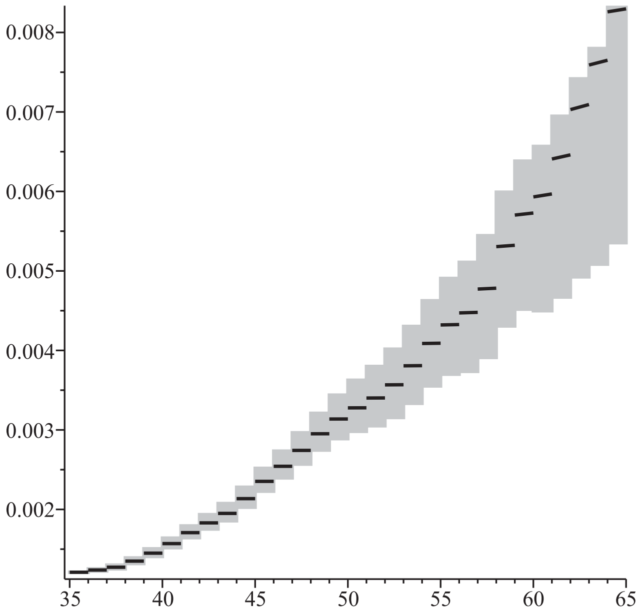

FIGURE 10: Worst-case mortality intensity $\widetilde{\mu}_{ad}$ (black) of Example 5.5 and maximal and minimal possible mortality intensities (gray).

vestments or considering a partial contribution to the SCR. However, our results are not limited to applications inspired from Solvency II since we impose hardly any restrictions on the distribution of future financial and demographical developments. This makes our results robust to model risk.

There are several interesting open questions in the continuation of this work. These topics for future research include:

- Concerning the Solvency II application, one should work out a full model that includes asset risk.
- One of our motivating examples included a dependence between surrender and interest rates. Accounting for and stressing of policyholder behavior is a crucial part of Solvency II, and we foresee that this will challenge companies, in particular in the case of internal modeling. We studied one example with dependence between surrender and interest rates, but a closer look at this kind of dependency is called for.
- One realizes from our figures that, when modeling the dependence between basis elements, we no longer get corner solutions in general. However, it is clearly valuable to better understand under what condition the corner solutions turn out to be the worst-case after all, at least for some marginal elements. Further quantitative and qualitative studies are needed for such a clarification.
- In the present paper, we do not impose time-consistency restrictions upfront in the sense that the worst-case scenario shall not change with progressing

time. However, under certain invariance assumptions in Proposition 4.2, we indeed have time consistency. This is not necessarily the case under the weaker assumptions in Proposition 4.1. It is an important part of future research to further discuss time consistency.

ACKNOWLEDGEMENTS

The authors thank two anonymous referees for comments that have been very helpful for revising a previous version of the present work. Mogens Steffensen acknowledges work supported by the Danish Advanced Technology Foundation (Hoejteknologifonden) (017-2010-3).

REFERENCES

[1] ALEXANDER, C. and SHEEDY, E. (2008) Developing a stress testing framework based on market risk models. *Journal of Banking and Finance*, **32**, 2220–2236.
[2] ARAGONS, J., BLANCO, C. and Dowd, K. (2001) Incorporating stress tests into market risk modelling. *Derivatives Quarterly*, **7**, 44–49.
[3] BAUER, D., BERGMANN, D. and Reuss, A. (2010) *Solvency II and Nested Simulations – A Least-Squares Monte Carlo Approach*. In *Proceedings of the 2010 ICA Congress*.
[4] BERKOWITZ, J. (2000) A coherent framework for stress testing. *Journal of Risk*, **2**, 1–11.
[5] 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*, 225–259.
[6] CAIRNS, A.J.G., BLAKE, D. and Dowd, K. (2008) Modelling and management of mortality risk: A review. *Scandinavian Actuarial Journal*, **2008**(2–3), 79–113.
[7] 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**, 680–690.
[8] CHRISTIANSEN, M.C. (2010) Biometric worst-case scenarios for multi-state life insurance policies. *Insurance: Mathematics and Economics*, **47**, 190–197.
[9] CHRISTIANSEN, M.C. (2011) Making use of netting effects when composing life insurance contracts. *European Actuarial Journal*, **1**(Suppl. 1), 47–60.
[10] CHRISTIANSEN, M.C. and Denuit, M.M. (2010) First-order mortality rates and safe-side actuarial calculations in life insurance. *ASTIN Bulletin*, **40**(2), 587–614.
[11] CHRISTIANSEN, M.C., DENUIT, M.M. and Lazar, D. (2012) The Solvency II square-root formula for systematic biometric risk. *Insurance: Mathematics and Economics*, **50**, 257–265.
[12] DE GIOVANNI, D. (2010) Lapse rate modeling: A rational expectation approach. *Scandinavian Actuarial Journal*, **2010**(1), 56–67.
[13] European Commission (2008) Directive of the European Parliament and of the Council on the taking-up and pursuit of the business of insurance and reinsurance. (Solvency II). COM(2008) 119.
[14] FABOZZI, F.J. (2005) *The Handbook of Fixed Income Securities*, 7th ed. New York: McGraw Hill.
[15] European Commission (2010) Fifth Quantitative Impact Study: Technical Specifications.
[16] GENEST, C., GERBER, H.U., GOOVAERTS, M.J. and Laeven, R.J.A. (2009) Editorial to the special issue on modeling and measurement of multivariate risk in insurance and finance. *Insurance: Mathematics and Economics*, **44**, 143–145.
[17] GOOVAERTS, M.J., KAAS, R. and Laeven, R.J.A. (2011) Worst case risk measurement: Back to the future? *Insurance: Mathematics and Economics*, **49**, 380–392.
[18] HOEM, J.M. (1988) The versatility of the Markov chain as a tool in the mathematics of life insurance. In *Transactions of the 23rd International Congress of Actuaries,* Vol. 3, pp. 171–202. Helsinki, Finland: ICA.

[19] KAAS, R., LAEVEN, R. and Nelsen, R. (2009) Worst VaR scenarios with given marginals and measures of association. *Insurance: Mathematics and Economics*, **44**, 146–158.
[20] KUPIEC, P. (1998) Stress testing in a value at risk framework. *Journal of Derivatives*, **6**, 7–24.
[21] LAEVEN, R. (2009) Worst VaR scenarios. *Insurance: Mathematics and Economics*, **44**, 159–163.
[22] LI, J. and Szimayer, A. (2010) *The Effect of Policyholders Rationality on Unit-Linked Life Insurance Contracts with Surrender Guarantees* (December 15, 2010). Available at SSRN: http://ssrn.com/abstract=1725769 or doi:10.2139/ssrn.1725769
[23] LI, J. and Szimayer, A. (2011) The uncertain force of mortality framework: Pricing unit-linked life insurance contracts. *Insurance: Mathematics and Economics*, **49**, 471–486.
[24] MCNEIL, A.J. and Smith, A.D. (2012) Multivariate stress scenarios and solvency. *Insurance: Mathematics and Economics*, **50**, 299–308.
[25] MILBRODT, H. and Stracke, A. (1997) Markov models and Thiele's integral equations for the prospective reserve. *Insurance: Mathematics and Economics*, **19**, 187–235.
[26] NORBERG, R. (1999) A theory of bonus in life insurance. *Finance and Stochastics*, **3**, 373–390.
[27] STUDER, G. (1997) *Maximum Loss for Measurement of Market Risk*, PhD thesis, ETH Zurich.
[28] STUDER, G. (1999) Market risk computation for nonlinear portfolios. *Journal of Risk*, **1**, 33–53.

MARCUS C. CHRISTIANSEN (CORRESPONDING AUTHOR)
Institut für Versicherungswissenschaften, Universität Ulm,
D-89069 Ulm, Germany
E-mail: marcus.christiansen@uni-ulm.de

MOGENS STEFFENSEN
Department of Mathematical Sciences, University of Copenhagen,
DK-2100 Copenhagen, Denmark

APPENDIX

LOCAL OPTIMIZATION PROBLEM OF EXAMPLES 5.2 AND 5.3

Here, we explain how to calculate the argmax in Example 5.3. The argmax in Example 5.2 can be obtained analogously.

Assume that the set $M(t)$ is strictly convex. Consider the linear optimization problem

$$\min_{(x,y)^T \in M(t)} ax + by = \min_{(x,y)^T \in \partial M(t)} ax + by.$$

With writing $s_{ij} = (\Sigma_t^{-1})_{ij}$ for the ijth entry of the symmetric matrix Σ_t^{-1}, one can show that

$$\partial M(t) = \Big\{ (x, y)^T | x = \mathbb{E}\varphi(t) + \frac{s_{12}}{s_{11}} (\mathbb{E} \ln(\mu_{ad}(t)) - \ln(y)) \pm \frac{1}{s_{11}} \sqrt{s_{11} r^2 - \det(\Sigma_t^{-1})(\mathbb{E} \ln(\mu_{ad}(t)) - \ln(y))^2} \Big\}. \tag{6.1}$$

The point $(x, y)^T \in \partial M(t)$ is the minimal point if and only if its normal vector shows into the same direction as $(-a, -b)^T$. The tangent vector of $(x, y)^T \in \partial M(t)$ has the direction

$$\begin{pmatrix} \mathrm{d}x/\mathrm{d}y \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{s_{12}}{s_{11}}\frac{1}{y} \pm \frac{1}{s_{11}y}\frac{\det(\Sigma_t^{-1})(\ln(\mu_{ad}(t)) - \ln(y))}{\sqrt{s_{11}r^2 - \det(\Sigma_t^{-1})(\ln(\mu_{ad}(t)) - \ln(y))^2}} \\ 1 \end{pmatrix},$$

and for a minimal point $(x, y)^T \in \partial M(t)$ the corresponding tangent vector is orthogonal to $(-a, -b)^T$, that is

$$(-a, -b)\begin{pmatrix} \mathrm{d}x/\mathrm{d}y \\ 1 \end{pmatrix} = 0.$$

This equation, which has only one unknown variable, namely y, can be solved with standard numerical methods, and one of the two solutions is the worst-case mortality intensity at t. Plugging the right solution into the equation in (6.1) yields the worst-case interest intensity.