ASTIN, AFIR/ERM and IACA Colloquia
Innovation & Invention
23-27 August 2015 | Amora Hotel Jamison Sydney

Actuaries Institute

On the Calculation of the Solvency Capital Requirement Based on Nested Simulations

Prepared by Daniel Bauer, Andreas Reuss and Daniela Singer

Presented to the Actuaries Institute
ASTIN, AFIR/ERM and IACA Colloquia
23-27 August 2015
Sydney

Institute of Actuaries of Australia
ABN 69 000 423 656
Level 2, 50 Carrington Street, Sydney NSW Australia 2000
t +61 (0) 2 9233 3466 **f** +61 (0) 2 9233 3446
e actuaries@actuaries.asn.au **w** www.actuaries.asn.au

ON THE CALCULATION OF THE SOLVENCY CAPITAL REQUIREMENT BASED ON NESTED SIMULATIONS*

BY

DANIEL BAUER, ANDREAS REUSS AND DANIELA SINGER

ABSTRACT

Within the European Union, risk-based funding requirements for insurance companies are currently being revised as part of the Solvency II project. However, many life insurers struggle with the implementation, which to a large extent appears to be due to a lack of know-how regarding both, stochastic modeling and efficient techniques for the numerical implementation.

The current paper addresses these problems by providing a mathematical framework for the derivation of the required risk capital and by reviewing different alternatives for the numerical implementation based on nested simulations. In particular, we seek to provide guidance for practitioners by illustrating and comparing the different techniques based on numerical experiments.

KEYWORDS

Solvency II, Value-at-Risk, nested simulations, screening procedures.

1. INTRODUCTION

Within the European Union, risk-based funding requirements for insurance companies are currently being revised as part of the Solvency II project. One key aspect of the new regulatory framework is the determination of the required risk capital for a one-year time horizon, i.e. the amount of capital the company must hold against unforeseen losses during the following year. In particular, the regulation allows for a company-specific calculation based

* Parts of this paper are taken from an earlier paper called "Solvency II and Nested Simulations — a Least-Squares Monte Carlo Approach" and from the third author's doctoral dissertation (cf. Bergmann (2011)). The authors are grateful for helpful comments from an anonymous referee and seminar participants at the 2009 ARIA meeting, the 2009 CMA Workshop on Insurance Mathematics and Longevity Risk, the 2010 International Congress of Actuaries, Georgia State University, Humboldt University of Berlin, Ulm University, and the University of Duisburg-Essen. All remaining errors are ours.

on a market-consistent valuation of assets and liabilities within a structural *internal model*. However, many life insurers struggle with the implementation, which to a large extent appears to be due to a lack of know-how regarding both, the construction of the underlying model and efficient techniques for implementing the necessary calculations (see e.g. CEIOPS Internal Model Expert Group (2009), p. 23-25, CRO Forum (2009), p. 30, and SOA (2008), Section 2.5). As a consequence, many companies rely on approximations within the so-called *standard model*, which is usually not able to accurately reflect an insurer's risk situation and may lead to deficient outcomes (see e.g. Liebwein (2006), Pfeifer and Strassburger (2008), Ronkainen et al. (2007), or Sandström (2007)).

The current paper addresses these problems. More specifically, our objectives are twofold: On the one hand, we seek to shed light on the proper calculation of the *Solvency Capital Requirement* (SCR) by presenting a concise mathematical framework. On the other hand, to provide guidance for the practical implementation, we survey, extend, and adapt different advanced techniques for the calculation of the SCR based on nested simulations. For instance, we address the optimal allocation of computational resources within the simulation, the construction of confidence intervals for the SCR, the application of variance reduction techniques, and the use of screening procedures to increase the efficiency of the simulation approach. The drawbacks and advantages of the different approaches and techniques are illustrated based on detailed numerical experiments using the model for a participating term-fix contract introduced in Bauer et al. (2006).[1] In particular, we demonstrate that the efficiency of the computation can be increased considerably when relying on a suitable simulation design.

Several of the presented numerical techniques were originally proposed in the context of nested simulations for portfolio risk measurement, and our contribution in this direction lies in the extension and adaptation of the underlying ideas to suit the insurance setting as well as their integration. In particular, we draw on results from Gordy and Juneja (2010), who analyze how to allocate a fixed computational budget to the inner and the outer simulation step within a nested simulation in order to minimize the mean square error when measuring the risk of a derivative portfolio. Furthermore, for the derivation of confidence intervals for the SCR with and without screening procedures, we follow ideas from Lan et al. (2007a,b, 2010), where similar problems were studied.

The remainder of the paper is structured as follows. Section 2 provides background information on the Solvency II requirements and gives precise definitions of the quantities of interest. In Section 3, we introduce the mathematical framework underlying our considerations using a *direct* valuation of the insurer's

[1] As pointed out by Kling et al. (2007), under the assumption that cash flows resulting from premiums roughly compensate for death and surrender benefits, the evolution of a term-fix contract can be considered as an approximation for the evolution of an entire life insurance company offering participating contracts.

liabilities and describe the basic nested simulations approach for estimating the SCR. Aside from presenting the (point) estimation procedure, we address the determination of an optimal allocation of a fixed computational budget and briefly explain how a Jackknife procedure can be used to reduce the inherent bias. In Section 4, we derive confidence intervals for the resulting point estimator. The subsequent Section 5 describes methods to increase the efficiency of the estimation by means of screening procedures. In Section 6, we illustrate the different methods based on detailed numerical experiments. In Section 7, we discuss an alternative estimator for the SCR based on an *indirect* valuation of the insurer's liabilities. Finally, Section 8 summarizes our findings and concludes.

2. The Solvency II Capital Requirement

The quantitative assessment of the solvency position of a life insurer can be split into two components, namely the derivation of the *Available Capital* (AC) at the current point in time ($t = 0$), and the derivation of the *Solvency Capital Requirement* (SCR) based on the Available Capital at the measurement time horizon (one year for Solvency II, $t = 1$).

2.1. Available Capital

The Available Capital, which is also called "own funds" under Solvency II, corresponds to the amount of available financial resources that can serve as a buffer against risks and absorb financial losses. It is derived from a market-consistent valuation approach as the difference between the market value of assets and the market value of liabilities. The market-consistent valuation of assets is usually quite straightforward for the typical investment portfolio of an insurance company since market values are either readily available (mark-to-market, level 1) or can be derived from standard models with market-observable inputs (level 2). This is usually not the case for the liabilities of a life insurance company, and there are two different basic approaches for the calculation of the value, the *direct* and the *indirect* method (cf. Girard (2002)).

As suggested by its name, the *direct method* prescribes a direct valuation of the cash flows associated with an insurance liability, e.g. by determining their expected discounted value under some risk-neutral or risk-adjusted probability measure.[2] In contrast, within the *indirect method*, the valuation is taken out from the shareholders' perspective by considering the free cash flows generated by the insurance business. The market value of liabilities is then equal to the difference between the market value of assets backing liabilities and the market

[2] To keep our focus and without loss of generality, we do not address methods to account for non-financial (non-hedgeable) risks in the current paper, but refer to Babbel et al. (2002), Klumpes and Morgan (2008), and references therein for this discussion.

value of the cash flows from the shareholders' perspective. While of course the quantity to be estimated is — or at least should be — the same for both procedures (see Girard (2002)), the two methods may well yield different estimators for the AC and, hence, for the SCR. Since the conceptual results of our paper are not affected by the choice of the method and since the direct method is more in line with the QIS 5 Technical Specifications for Solvency II (see European Commission (2010)), we limit our technical exposition to the direct method. A more detailed discussion of the indirect method for the calculation of life insurance liabilities based on the *Market Consistent Embedded Value* (MCEV) principles issued by the CFO Forum (2009) is deferred to Section 7. In particular, our numerical experiments illustrate that the quality of the resulting estimates can differ significantly.

In either case, due to the relatively complex financial structure of life insurance liabilities containing embedded options and guarantees, this calculation usually cannot be done in closed form. Therefore, insurance companies usually follow a mark-to-model approach that relies on Monte Carlo simulations. Once the market value of liabilities has been determined, the Available Capital at $t = 0$ can be derived as the difference between the market value of assets (MVA_0) and the market value of liabilities (MVL_0) at $t = 0$, i.e. we have

$$\mathrm{AC}_0 := \mathrm{MVA}_0 - \mathrm{MVL}_0. \tag{1}$$

2.2. The Solvency Capital Requirement

For deriving the SCR, the quantity of interest is the Available Capital at $t = 1$, which can be described by

$$\mathrm{AC}_1 := \mathrm{MVA}_1 - \mathrm{MVL}_1. \tag{2}$$

Then intuitively, an insurance company is considered to be solvent under Solvency II if its AC at $t = 1$ as seen from $t = 0$ is positive with a probability of at least 99.5%, i.e.

$$\mathcal{P}(\mathrm{AC}_1 \geq 0 \,|\, \mathrm{AC}_0 = x) \overset{!}{\geq} 99.5\%.$$

The SCR would then be defined as the smallest amount x satisfying this condition. This is an implicit definition of the SCR ensuring that if the Available Capital at $t = 0$ is greater or equal to the Solvency Capital Requirement, then the probability of the Available Capital at $t = 1$ being positive is at least 99.5%.

However, for practical applications, one usually relies on a simpler but approximately equivalent notion of the SCR, which avoids the implicit nature of the definition given above. For this purpose, we define the one-year loss function evaluated at $t = 0$ as

$$\Delta := \mathrm{AC}_0 - \frac{\mathrm{AC}_1}{1 + s(0,1)},$$

where $s(0,1)$ is the one-year risk-free rate over $[0,1]$, i.e. $s(0,1) := \frac{1}{P(0,1)} - 1$ with $P(0,1)$ the price of a one-year zero coupon bond at time zero. The SCR is then defined as the α-quantile of Δ, where the security level α is set equal to 99.5%[3]:

$$\text{SCR} := \text{argmin}_x \left\{ \mathcal{P}\left(\text{AC}_0 - \frac{\text{AC}_1}{1 + s(0,1)} > x \right) \le 1 - \alpha \right\}. \tag{3}$$

The probability that the loss over one year exceeds the SCR is less or equal to $1 - \alpha$, i.e. we need to calculate a one-year Value-at-Risk (VaR). The *Excess Capital* at $t = 0$, on the other hand, is defined as $\text{AC}_0 - \text{SCR}$ and satisfies the following requirement:

$$\mathcal{P}\left(\frac{\text{AC}_1}{1 + s(0,1)} \ge \text{AC}_0 - \text{SCR} \right) \ge \alpha; \tag{4}$$

thus, the probability (evaluated at $t = 0$) that the Available Capital at $t = 1$ is greater or equal to the Excess Capital is at least α (e.g. 99.5%).

Note that under this definition, the SCR depends on the actual amount of capital held at $t = 0$. In particular, since all assets are included in the calculations, the risk arising from assets backing positive Excess Capital is also reflected in the SCR. In case the Excess Capital is negative, it is implicitly assumed that it is invested in the one-year default-free bond, which can be illustrated by rewriting Equation (4) as follows:

$$\mathcal{P}\left(\text{AC}_1 + (\text{SCR} - \text{AC}_0) \cdot (1 + s(0,1)) \ge 0 \right) \ge \alpha.$$

Based on this definition of the SCR, the solvency ratio can be calculated as AC_0 / SCR.

In the *standard model*, the SCR in Equation (3) is approximated via the so-called *square-root formula* based on a modular approach. However, this formula is usually not able to accurately reflect the insurer's risk situation and may lead to deficient outcomes (see e.g. Pfeifer and Strassburger (2008) and Sandström (2007)). Therefore, in what follows, we describe how to determine the probability distribution of the loss function based on nested simulations in an *internal model*, which enables us to derive the SCR directly as defined in Equation (3).

[3] These simplifications are analogous to the definition used for the Swiss Solvency Test (Federal Office of Private Insurance (2006)).

3. Nested Simulations Approach

3.1. Mathematical Framework

We assume that investors can trade continuously in a frictionless financial market, and we let T be the maturity of the longest-term policy in the life insurer's portfolio. Since insurance contracts are long-term investments, T will usually be in the range of 30-40 years or even longer. Let $(\Omega, \mathcal{F}, \mathcal{P}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]})$ be a complete filtered probability space on which all relevant quantities exist, where Ω denotes the space of all possible states of the financial market and $\mathcal{P}$ is the physical probability measure. $\mathcal{F}_t$ represents all information about the financial market up to time t, and the filtration $\mathbb{F}$ is assumed to satisfy the usual conditions.

The uncertainty with respect to the insurance company's future liabilities from the in-force business[4] arises from the uncertain development of a number of influencing factors, such as equity returns, interest rates, or credit spreads. We introduce the d-dimensional, sufficiently regular Markov process $Y = (Y_t)_{t\in[0,T]} = (Y_{t,1}, \ldots, Y_{t,d})_{t\in[0,T]}$, the so-called *state process*, to model the uncertainty of the financial market, i.e. all risky assets in the market can be expressed in terms of Y. Furthermore, we suppose the existence of a locally risk-free asset $(B_0(t))_{t\in[0,T]}$ (the bank account) with $B_0(t) = \exp\{\int_0^t r_u\,du\}$, where $r_t = r(Y_t)$ is the instantaneous risk-free interest rate at time t. To ease notation, we define $B_s(t) := \frac{B_0(t)}{B_0(s)} = \exp(\int_s^t r_u\,du)$. In this market, we take for granted the existence of a risk-neutral probability measure Q equivalent to $\mathcal{P}$ under which payment streams can be valued via their expected discounted values with respect to the numéraire process $(B_0(t))_{t\in[0,T]}$.[5]

Based on this market model, we assume that there exists a cash flow projection model of the insurance company, i.e. there exist functionals $h_1, \ldots, h_T$ that derive the insurer's liability cash flows at time t from the development of the financial market up to time t, $t = 1, \ldots, T$. In particular, these cash flows include policyholder cash flows (benefits paid minus premiums earned), expenses (both internal and external), and tax payments, and the model reflects legal and regulatory requirements, policyholder behavior, as well as management rules. Hence, we model the future liability cash flows from the in-force business as a sequence of random variables $X = (X_1, \ldots, X_T)$ where $X_t = h_t(Y_s, s\in[0,t])$, $t = 1, \ldots, T$.

In order to keep our presentation concise, as pointed out above, we abstract by limiting our focus to market risk, i.e. non-hedgeable risks as well as the corresponding cost-of-capital charges are ignored (cf. Footnote 2). However, non-financial risk factors such as a mortality index could also be incorporated in the state process (see Zhu and Bauer (2011)). The corresponding cost-of-capital

4 This means that cash flows from future new business are not included in the calculation.

5 Under some mild technical conditions, this assumption is equivalent to the absence of arbitrage in the financial market. See e.g. Bingham and Kiesel (2004) for more details.

charges as well as other frictional costs could then be considered by an appropriate choice of Q and h_t, $t = 1, \ldots, T$.

3.2. Calculation of the SCR

According to the risk-neutral valuation formula, we can determine the market value of liabilities at time $t = 0$ as the expectation of the sum of the discounted liability cash flows X_t, $t = 1, \ldots, T$, under the risk-neutral measure Q:

$$\text{MVL}_0 := \mathbb{E}^Q\left[\sum_{t=1}^{T} \frac{X_t}{B_0(t)}\right] \text{ with } \sigma_0 := \sqrt{\text{Var}^Q\left[\sum_{t=1}^{T} \frac{X_t}{B_0(t)}\right]}.$$

In most cases, MVL_0 cannot be computed analytically due to the complexity of the interaction between the development of the financial market variables Y_t and the liability cash flows X_t. Thus, in general, we have to rely on numerical methods to estimate MVL_0.

A common approach is to use Monte Carlo simulations, i.e. independent sample paths $(Y_t^{(k)})_{t\in[0,T]}$, $k = 1, \ldots, K_0$, of the underlying state process Y generated under the risk-neutral measure Q. Based on these different scenarios for the financial market, we first derive the resulting cash flows $X_t^{(k)}$ $(t = 1, \ldots, T;$ $k = 1, \ldots, K_0)$ using the cash flow projection model. Then, we discount the cash flows with the appropriate discount factor, and average over all K_0 sample paths, i.e.

$$\widetilde{\text{MVL}}_0(K_0) := \frac{1}{K_0}\sum_{k=1}^{K_0}\sum_{t=1}^{Y} \frac{X_t^{(k)}}{B_0^{(k)}(t)},$$

where $B_s^{(k)}(t) := \exp\left(\int_s^t r_u^{(k)} du\right)$ and $r_u^{(k)}$ denotes the instantaneous risk-free interest rate at time u in sample path k. By Equation (1) and since the market value of assets is usually readily available, an estimator for AC_0 is given by $\widetilde{\text{AC}}_0(K_0) = \text{MVA}_0 - \widetilde{\text{MVL}}_0(K_0)$. The sample version of the standard deviation is denoted by $\tilde{\sigma}_0(K_0)$.

For the calculation of the Solvency Capital Requirement, in addition to the Available Capital at $t = 0$, we need to assess the (physical) distribution of the Available Capital at $t = 1$. Assuming that corresponding policyholder cash flows at time $t = 1$ have already been settled but that shareholder cash flows have not been realized yet, we need to determine the $\mathcal{P}$-distribution of the $\mathcal{F}_1$-measurable random variable (cf. Equations (1) and (2))

$$\text{AC}_1 := \text{MVA}_1 - \underbrace{\mathbb{E}^Q\left[\sum_{t=2}^{T} \frac{X_t}{B_1(t)}\middle|\mathcal{F}_1\right]}_{=:\ \text{MVL}_1}.$$

We may now estimate the distribution of AC_1 via the corresponding empirical distribution function: Given $N \in \mathbb{N}$ independent and identically distributed (i.i.d.) sample paths $(Y_s^{(i)})_{s\in[0,1]}$, $i = 1,\dots,N$, for the development of the financial market over the first year under the real-world measure $\mathcal{P}$, the market value of liabilities at $t = 1$ conditional on the state of the financial market in scenario i can be described by

$$\mathrm{MVL}_1^{(i)} := \mathbb{E}^Q\left[\underbrace{\sum_{t=2}^{T}\frac{X_t}{B_1(t)}}_{=:PV_1^{(i)}}\middle|(Y_s^{(i)})_{s\in[0,1]}\right] \text{ with } \sigma_1^{(i)} := \sqrt{\mathrm{Var}^Q[PV_1^{(i)}|(Y_s^{(i)})_{s\in[0,1]}]}. \quad (5)$$

Note that the $\sigma_1^{(i)}$ may differ significantly for different scenarios i, i.e. the discounted cash flows $\sum_{t=2}^{T}\frac{X_t}{B_1(t)}$ are usually not identically distributed for different realizations of the state process over the first year.

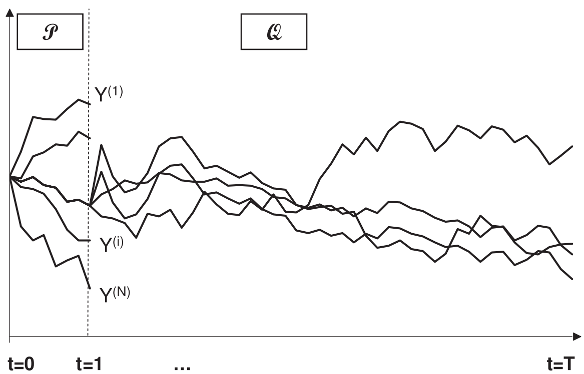

FIGURE 1: Illustration of the nested simulations approach.

In addition, realizations for the market value of assets can easily be calculated for each of the N first-year paths. Therefore, N realizations of AC_1 are given by

$$\mathrm{AC}_1^{(i)} := \mathrm{MVA}_1^{(i)} - \mathrm{MVL}_1^{(i)}.$$

Note that these $\mathcal{F}_1$-measurable random variables $AC_1^{(i)}$, $i = 1,\dots,N$, are independent and identically distributed as Monte Carlo realizations and thus may be used for the construction of an empirical distribution function.

However, just as at time zero, the valuation problem (5) generally cannot be solved analytically, and, again, we may rely on Monte Carlo simulations. As illustrated in Figure 1, based on the first-year path of the state process $(Y_s^{(i)})_{s\in[0,1]}$ in scenario $i\in\{1,\dots,N\}$, we simulate $K_1^{(i)}\in\mathbb{N}$ risk-neutral scenarios and denote them by $(Y_s^{(i,k)})_{s\in(1,T]}$. Then, for each first-year path $i\in\{1, \dots, N\}$, by determining the resulting future liability cash flows $X_t^{(i,k)}(t = 2, \dots, T;$

$k = 1, \ldots, K_1^{(i)}$) and averaging over all $K_1^{(i)}$ sample paths, we obtain Monte Carlo estimates for $\mathrm{MVL}_1^{(i)}$ via

$$\widetilde{\mathrm{MVL}}_1^{(i)}(K_1^{(i)}) := \frac{1}{K_1^{(i)}} \sum_{k=1}^{K_1^{(i)}} \underbrace{\sum_{t=2}^{T} \frac{X_t^{(i,k)}}{B_1^{(i,k)}(t)}}_{=: PV_1^{(i,k)}}, \quad i \in \{1, \ldots, N\}.$$

The number of simulations in the i^{th} real-world scenario may depend on i since for different standard deviations $\sigma_1^{(i)}$, a different number of simulations may be necessary to obtain acceptable results. We obtain the following sample standard deviation for $PV_1^{(i)}$:

$$\tilde{\sigma}_1^{(i)}(K_1^{(i)}) := \sqrt{\frac{1}{K_1^{(i)} - 1} \sum_{k=1}^{K_1^{(i)}} \left(PV_1^{(i,k)} - \widetilde{\mathrm{MVL}}_1^{(i)}(K_1^{(i)}) \right)^2}.$$

Now, we can estimate N realizations of AC_1 by

$$\widetilde{\mathrm{AC}}_1^{(i)}(K_1^{(i)}) := \mathrm{MVA}_1^{(i)} - \widetilde{\mathrm{MVL}}_1^{(i)}(K_1^{(i)}), \quad i = 1, \ldots, N.$$

From Equation (3), it follows that the SCR is the α-quantile of the random variable $\Delta = \mathrm{AC}_0 - \frac{\mathrm{AC}_1}{1+s(0,1)}$. Since AC_0 is approximated by the unbiased estimator $\widetilde{\mathrm{AC}}_0(K_0)$ and $s(0,1)$ is known at $t = 0$, the only remaining random component is AC_1 and the task is to estimate the α-quantile of $-\mathrm{AC}_1$.

Based on the N estimated realizations of the random variable $S = -\mathrm{AC}_1$ with corresponding order statistics $(\tilde{S}_{(1)}, \ldots, \tilde{S}_{(N)})$ and realization $(\tilde{s}_{(1)}, \ldots, \tilde{s}_{(N)})$, a simple approach for estimating the α-quantile s_α is to rely on the corresponding empirical quantile, i.e.

$$\tilde{s}_\alpha = \tilde{s}_{(m)},$$

where $m = \lfloor N \cdot \alpha + 0.5 \rfloor$. The SCR can then be estimated as

$$\widetilde{\mathrm{SCR}} = \widetilde{\mathrm{AC}}_0(K_0) + \frac{\tilde{s}_{(m)}}{1 + s(0,1)}. \tag{6}$$

Alternatively, extreme value theory could be applied to derive a robust estimate of the quantile based on the given observations; see e.g. Embrechts et al. (1997) for details.

3.3. Quality of the Resulting Estimator and Choice of K_0, K_1, and N

Within our estimation process, we have three sources of error: (1) We estimate the Available Capital at $t = 0$ with the help of (only) K_0 sample paths; (2) we only use N real-world scenarios to estimate the distribution function; and (3) the Available Capital at $t = 1$ is estimated with the help of (only) K_1 sample paths in every scenario.[6] As a consequence, Equation (6) does not necessarily present an (unbiased) estimate for the quantile of the distribution function of the "true" $\mathcal{F}_1$-measurable loss

$$\Delta = \mathrm{AC}_0 - \frac{\mathrm{AC}_1}{1+s(0,1)} = \mathrm{AC}_0 - \frac{\mathrm{MVA}_1 - \mathrm{MVL}_1}{1+s(0,1)},$$

but instead we actually consider the distribution of the estimated loss

$$\tilde{\Delta} = \widetilde{\mathrm{AC}}_0(K_0) - \frac{\mathrm{MVA}_1 - \left(\frac{1}{K_1}\sum_{k=1}^{K_1}\sum_{t=2}^{T}\frac{X_t^{(k)}}{B_1^{(k)}(t)}\,\middle|\,(Y_s)_{s\in[0,1]}\right)}{1+s(0,1)}.$$

In particular, $\tilde{\Delta}$ is not $\mathcal{F}_1$-measurable due to the random sampling error resulting from the estimation of AC_0 and the inner simulation.

Obviously, by the law of large numbers (LLN)

$$\tilde{\Delta} \to \Delta \;a.s.\; \text{ as } K_0, K_1 \to \infty.$$

Nevertheless, we base our estimation of the SCR on distorted samples. To analyze the influence of this inaccuracy on our actual estimate $\widetilde{\mathrm{SCR}}$, we follow Gordy and Juneja (2010) and decompose the mean square error (MSE) into the variance of our estimator and a bias[7]:

$$\mathrm{MSE} = \mathbb{E}\left[(\widetilde{\mathrm{SCR}} - \mathrm{SCR})^2\right] = \mathrm{Var}(\widetilde{\mathrm{SCR}}) + \Bigg[\underbrace{\mathbb{E}(\widetilde{\mathrm{SCR}}) - \mathrm{SCR}}_{\text{bias}}\Bigg]^2. \qquad (7)$$

Since $\widetilde{\mathrm{AC}}_0(K_0)$ is an unbiased estimator of AC_0 and since it is independent of $\tilde{s}_{(m)}$, Equation (7) simplifies to

$$\mathrm{MSE} = \mathrm{Var}(\widetilde{\mathrm{AC}}_0(K_0)) + \mathrm{Var}\left(\frac{\tilde{s}_{(m)}}{1+s(0,1)}\right) + \left[\mathbb{E}\left(\frac{\tilde{s}_{(m)}}{1+s(0,1)}\right) - \frac{s_\alpha}{1+s(0,1)}\right]^2. \qquad (8)$$

Obviously, $\mathrm{Var}(\widetilde{\mathrm{AC}}_0(K_0)) = \frac{\sigma_0^2}{K_0}$, and we will now focus on the second and third term in (8). Again following Gordy and Juneja (2010), let

[6] For the sake of simplicity, for the remainder of this section we let $K_1^{(i)} = K_1$ for all $i \in \{1, \dots, N\}$.

[7] In what follows, probabilities and expectations are calculated under the so-called *process distribution*. More specifically, since the joint distributions of the random process governing our problem are given by the simulation procedure, Kolmogorov's construction yields a probability measure, the so-called process distribution, which for simplicity is also denoted by $\mathcal{P}$ (see Gray (2009) for details).

$$Z^{K_1} = \frac{\mathrm{MVA}_1 - \left(\frac{1}{K_1} \sum_{k=1}^{K_1} \sum_{t=2}^{T} \frac{X_t^{(k)}}{B_1^{(k)}(t)} \,\middle|\, (Y_s)_{s \in [0,1]} \right)}{1 + s(0,1)} - \frac{\mathrm{MVA}_1 - \mathrm{MVL}_1}{1 + s(0,1)}$$

denote the difference between the estimated loss and its "true" value under the assumption that $\widetilde{\mathrm{AC}}_0(K_0)$ is exact. Furthermore, define $g_{K_1}(\cdot,\cdot)$ to be the joint distribution function of Δ and $\tilde{Z}^{K_1} := Z^{K_1} \cdot \sqrt{K_1}$.

Then, with Proposition 2 from Gordy and Juneja (2010), under some regularity conditions, we obtain

$$\mathbb{E}\left[\frac{\tilde{s}_{(m)}}{1 + s(0,1)} \right] - \frac{s_\alpha}{1 + s(0,1)} \tag{9}$$

$$= \frac{\theta_\alpha}{K_1 \cdot f(\mathrm{SCR})} + o_{K_1}(1/K_1) + O_N(1/N) + o_{K_1}(1)\, O_N(1/N),$$

and $$\mathrm{Var}\left(\frac{\tilde{s}_{(m)}}{1 + s(0,1)} \right) = \frac{\alpha(1-\alpha)}{(N+2) f^2(\mathrm{SCR})} + O_N(1/N^2) + o_{K_1}(1)\, O_N(1/N), \tag{10}$$

where $f(\cdot)$ denotes the density function of Δ and

$$\begin{aligned} \theta_\alpha &= -\frac{1}{2} \frac{\partial}{\partial u} \left[f(u)\, \mathbb{E}\left[\mathrm{Var}(\tilde{Z}^{K_1} \,|\, (Y_s)_{s \in [0,1]}) \,|\, \Delta = u \right] \right] \Big|_{u = \mathrm{SCR}} \\ &= -\frac{1}{2} \int_{-\infty}^{\infty} z^2 \frac{\partial}{\partial u} g_{K_1}(u,z)\, dz \Big|_{u = \mathrm{SCR}}. \end{aligned} \tag{11}$$

The sign of θ_α — and, hence, the direction of the bias — will eventually be determined by the sign of $\frac{\partial}{\partial u} g_{K_1}(u,z)$. Since the SCR is located in the right-hand tail of the distribution and since $\frac{g_{K_1}(u,z)}{\int_{-\infty}^{\infty} g_{K_1}(l,z)\,dl}$ is a (conditional) density function, $\frac{\partial}{\partial u} g_{K_1}(u,z)|_{u=\mathrm{SCR}}$ will in general be negative. Thus, we expect to overestimate the SCR, i.e. the probability that the company is solvent is on average slightly higher than $\alpha = 99.5\%$.

Equations (8), (9), and (10) can now be readily applied to devise an optimal allocation of computational resources in the sense that the optimal triplet (K_0, N, K_1) minimizes the mean square error for a given computational budget:

Proposition 3.1. *Consider the computational budget constraint*

$$\Gamma = c \cdot K_0 + N \cdot K_1. \tag{12}$$

Then for an optimal triplet that minimizes the mean square error (7) subject to a budget constraint of the form (12), we have asymptotically for large K_1:

$$N \approx \frac{\alpha(1-\alpha) \cdot K_1^2}{2\theta_\alpha^2}, \text{ and}$$

$$K_0 \approx \frac{\sigma_0 \cdot f(\mathrm{SCR}) \sqrt{\alpha(1-\alpha)}}{2\theta_\alpha^2 \sqrt{c}} K_1^2 \sqrt{K_1} \approx \frac{\sigma_0 \cdot K_1 \cdot f(\mathrm{SCR})}{\theta_\alpha} \sqrt{\frac{N \cdot K_1}{2c}}.$$

A proof is provided in the Appendix. Proposition 3.1 extends a corresponding result in Gordy and Juneja (2010) by additionally considering the influence of the simulation error within the computation of AC_0 on the overall error. Its significance in our setting is that given the number of inner simulations K_1, it provides guidance on how to (asymptotically) optimally choose the number of outer simulations and time-zero simulations in order to minimize the mean square error. Here, typically one would choose $c \equiv 1$ since the sample paths for the estimation of AC_0 are only one period longer than those for the estimation of AC_1 and since T usually is relatively large. Similarly, the budget constraint disregards the cost for the generation of the N sample paths in the first period since this effort is small relative to the effort for the nested simulations when T is large.

In practical applications, f, σ_0, and θ_α are unknown but may be estimated in a pilot simulation with only a small number of sample paths. However, the estimation of θ_α generally will be quite inaccurate for large α because it is necessary to estimate a derivative in the very tail of the distribution.

3.4. Bias Reduction via the Jackknife

As a means to reduce the bias within the estimation of large loss probabilities based on nested simulations, Gordy and Juneja (2010) introduce a Jackknife procedure to correct the output from the inner simulation step for the inherent bias. While the method does not find immediate application for the estimation of the Value-at-Risk since the inner-simulation output does not enter the estimate of the SCR linearly, we can follow similar ideas to obtain a bias-corrected estimate. The primary difference is the interpretation of "a sample" as the combination of K_0 fixed time-zero simulations, N fixed outer simulations, and one inner simulation for each of the given N outer realizations. Our estimation of the SCR is then based on K_1 independent and identically distributed samples with coinciding time-zero and outer, but varying inner realizations, say $\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_{K_1}$.

Taking this point of view, we can then divide these samples into I non-overlapping sections $\{\hat{y}_k, k = \frac{(i-1)\cdot K_1}{I} + 1, \frac{(i-1)\cdot K_1}{I} + 2, \ldots, \frac{i\cdot K_1}{I}\}$, $i = 1, 2, \ldots, I$, of size $\frac{K_1}{I}$, where we assume $\frac{K_1}{I}$ is integer-valued. Denoting the estimate of the SCR based on all samples except for those from section i by $\widehat{\mathrm{SCR}}_{-i}$, the Jackknife estimate

$$\begin{aligned}\widehat{\mathrm{SCR}}^{+} &= I \cdot \widehat{\mathrm{SCR}} - (I-1)\cdot \frac{1}{I}\cdot \sum_{i=1}^{I} \widehat{\mathrm{SCR}}_{-i} \\ &= \widehat{\mathrm{SCR}} - (I-1)\cdot \frac{1}{I}\cdot \sum_{i=1}^{I} (\widehat{\mathrm{SCR}}_{-i} - \widehat{\mathrm{SCR}})\end{aligned}$$

then has the property that it eliminates the $O_{K_1}(1/K_1)$ bias term from the bias decomposition (9) (see e.g. Miller (1974)).

With only higher order terms left in the expression for the bias, Gordy and Juneja (2010) conclude that within an optimal allocation, the number of inner steps can even be further reduced relative to the outer simulations for the jackknife estimator.

4. Confidence Interval for the SCR

The practical usefulness of the estimator for the SCR from the previous section clearly depends on its accuracy, which may be described by a confidence interval. This section not only describes how to derive a confidence interval for the SCR based on the ideas by Lan et al. (2007a), but also addresses the allocation of the computational budget to obtain results as accurate as possible.

4.1. Derivation of a Confidence Interval for the SCR

When constructing a confidence interval for the SCR, we have to take into account the same three sources of uncertainty as described in the beginning of Section 3.3. To derive confidence intervals for estimates based on nested simulations, Lan et al. (2007a) propose a two step procedure: First, derive a confidence interval under the assumption that no inner simulations are necessary; then consider the uncertainty arising from the estimation in the inner simulation. However, they do not consider any uncertainty at $t = 0$ which — in our setup — comes into play due to the estimation of AC_0. Thus, in what follows, we extend their approach to derive a confidence interval for the SCR.

If the losses $\Delta^{(i)}$, $1 \leq i \leq N$, are known explicitly, the estimation error is solely due to the fact that the SCR is estimated via the empirical distribution function rather than the "true distribution". We are then looking to determine a lower bound LB as well as an upper bound UB such that

$$\mathcal{P}(\text{SCR} \in [LB; UB]) \geq 1 - \alpha_{\text{out}},$$

where α_{out} is the error resulting from the outer simulation. The derivation of such a confidence interval for the SCR is straightforward since $\sum_{i=1}^{N} \mathbb{1}_{\{\Delta^{(i)} \leq \text{SCR}\}}$ is Binomially distributed with parameters N and $\alpha = \mathcal{P}(\Delta \leq \text{SCR})$ (see e.g. Glasserman (2004), p. 491). More specifically, we have for $n \in \mathbb{N}$

$$\sum_{i=1}^{n-1} \binom{N}{i} \alpha^i (1-\alpha)^{N-i} = \mathcal{P}\left(\sum_{i=1}^{N} \mathbb{1}_{\{\Delta^{(i)} \leq \text{SCR}\}} < n\right) = \mathcal{P}(\Delta_{(n)} > \text{SCR})$$

$$\Rightarrow \mathcal{P}(\Delta_{(\underline{\psi})} \leq \text{SCR} < \Delta_{(\overline{\psi})}) = \sum_{i=\underline{\psi}}^{\overline{\psi}-1} \binom{N}{i} \alpha^i (1-\alpha)^{N-i}, \quad \underline{\psi}, \overline{\psi} \in \mathbb{N}, \tag{13}$$

where $\Delta_{(n)}$ denotes the n^{th} order statistic of the losses $(\Delta^{(i)})_{i=1}^{N}$. Therefore, in order to determine a $(1 - \alpha_{\text{out}})$-confidence interval for the SCR, it suffices to determine $\underline{\psi}, \overline{\psi} \in \mathbb{N}$ such that

$$\mathcal{P}(\Delta_{(\underline{\psi})} \leq \text{SCR} < \Delta_{(\overline{\psi})}) = \sum_{i=\underline{\psi}}^{\overline{\psi}-1} \binom{N}{i} \alpha^i (1-\alpha)^{N-i} \geq 1 - \alpha_{\text{out}}, \tag{14}$$

and to set $LB := \Delta_{(\underline{\psi})}$ and $UB := \Delta_{(\overline{\psi})}$. Clearly, the choice of $\underline{\psi}$ and $\overline{\psi}$ is not unique and the specification depends on the modeler's objective, for example the question of whether one- or two-sided confidence intervals are more appropriate for the application in view. In what follows, we assume that $\underline{\psi}$ and $\overline{\psi}$ are chosen at the beginning of the procedure, and that they remain fixed subsequently.

Within most applications, there exist no closed-form solution for the losses, i.e. they have to be estimated numerically. Therefore, we are looking for bounds $\widehat{LB}$ and $\widehat{UB}$ that can be derived from our nested simulations such that

$$\begin{aligned} &\lim_{\min\{K_1^{(1)},\ldots,K_1^{(N)}\}\to\infty} \mathcal{P}([LB;UB] \subseteq [\widehat{LB};\widehat{UB}]) \geq 1-\alpha_{\text{in}} \\ \Rightarrow &\lim_{\min\{K_1^{(1)},\ldots,K_1^{(N)}\}\to\infty} \mathcal{P}(\text{SCR} \in [\widehat{LB};\widehat{UB}]) \geq 1-\alpha_{\text{out}}-\alpha_{\text{in}}. \end{aligned} \tag{15}$$

Hence, $[\widehat{LB};\widehat{UB}]$ is an asymptotic confidence interval for the SCR.

In order to determine $\widehat{LB}$ and $\widehat{UB}$, we first observe that when determining the loss in the i^{th} real-world scenario, we have two sources of error: the estimation of AC_0 and the estimation of $\text{AC}_1^{(i)}$. Let α_{AC_0} be the error due to the estimation of AC_0 and α_{AC_1} be the error due to the estimation of AC_1 in all real-world scenarios. To simplify notation, we define

$$\begin{aligned} z_{\text{AC}_0}(K_0) &:= t_{K_0-1,1-\frac{\alpha_{\text{AC}_0}}{2}} \frac{\tilde{\sigma}_0(K_0)}{\sqrt{K_0}} \text{ and} \\ z_{\text{AC}_1}^{(i)}(K_1^{(i)},N) &:= t_{K_1^{(i)}-1,1-\frac{\varepsilon}{2}} \frac{\tilde{\sigma}_1^{(i)}(K_1^{(i)})}{(1+s(0,1))\cdot\sqrt{K_1^{(i)}}}, \end{aligned}$$

where $t_{k,\alpha}$ is the α quantile of the t-distribution with k degrees of freedom and $\varepsilon := 1-(1-\alpha_{\text{AC}_1})^{\frac{1}{N}}$. Moreover, we let

$$C := \bigotimes_{i=1}^{N} \left[\widetilde{\Delta}^{(i)}(K_1^{(i)}) - z_{\text{AC}_0}(K_0) - z_{\text{AC}_1}^{(i)}(K_1^{(i)},N); \widetilde{\Delta}^{(i)}(K_1^{(i)}) + z_{\text{AC}_0}(K_0) + z_{\text{AC}_1}^{(i)}(K_1^{(i)},N)\right]$$

where $\otimes$ denotes the cartesian product. If $PV_0^{(k)}$ and $PV_1^{(i,k)}$ are Normally distributed, we directly obtain

$$\begin{aligned} &\mathcal{P}((\Delta^{(1)},\ldots,\Delta^{(N)}) \in C) \\ \geq\ &\mathcal{P}\big(\widetilde{\text{AC}}_0 - z_{\text{AC}_0}(K_0) \leq \text{AC}_0 \leq \widetilde{\text{AC}}_0 + z_{\text{AC}_0}(K_0)\big) \\ &\prod_{i=1}^{N} \mathcal{P}\Big(\widetilde{\text{AC}}_1^{(i)} - z_{\text{AC}_1}^{(i)}(K_1^{(i)},N)\cdot(1+s(0,1)) \leq \text{AC}_1^{(i)} \leq \widetilde{\text{AC}}_1^{(i)} \\ &\qquad + z_{\text{AC}_1}^{(i)}(K_1^{(i)},N)\cdot(1+s(0,1))\Big) \\ =\ &(1-\alpha_{\text{AC}_0})\cdot\textstyle\prod_{i=1}^{N}(1-\varepsilon) = 1-\underbrace{(\alpha_{\text{AC}_0}+\alpha_{\text{AC}_1}-\alpha_{\text{AC}_0}\cdot\alpha_{\text{AC}_1})}_{=:\,\alpha_{\text{in}}}, \end{aligned} \tag{16}$$

i.e. C is a confidence region for $(\Delta^{(1)}, \ldots, \Delta^{(N)})$ with level $(1 - \alpha_{\text{in}})$. While generally, $PV_0^{(k)}$ and $PV_1^{(i,k)}$ will not be Normal, the confidence interval is still asymptotically valid by the central limit theorem (CLT). In order to combine the two confidence intervals for the inner and the outer simulation, let

$$\widehat{LB} := \arg\min_x \Big\{ \sum_i \mathbb{1}_{\{\widetilde{\Delta}^{(i)}(K_1^{(i)}) - z_{\text{AC}_0}(K_0) - z_{\text{AC}_1}^{(i)}(K_1^{(i)}, N) \le x\}} \ge \underline{\psi} \Big\} \text{ and} \tag{17}$$

$$\widehat{UB} := \arg\min_x \Big\{ \sum_i \mathbb{1}_{\{\widetilde{\Delta}^{(i)}(K_1^{(i)}) + z_{\text{AC}_0}(K_0) + z_{\text{AC}_1}^{(i)}(K_1^{(i)}, N) \le x\}} \ge \overline{\psi} \Big\}, \tag{18}$$

i.e. the $\underline{\psi}^{th}$ and $\overline{\psi}^{th}$ order statistic of $\widetilde{\Delta}^{(i)}(K_1^{(i)}) - z_{\text{AC}_0}(K_0) - z_{\text{AC}_1}^{(i)}(K_1^{(i)}, N)$ and $\widetilde{\Delta}^{(i)}(K_1^{(i)}) + z_{\text{AC}_0}(K_0) + z_{\text{AC}_1}^{(i)}(K_1^{(i)}, N)$, respectively. Then

$$\begin{aligned} \mathcal{P}\big([\Delta_{(\underline{\psi})}, \Delta_{(\overline{\psi})}] \subseteq [\widehat{LB}, \widehat{UB}]\big) &= \mathcal{P}\big(\widehat{LB} \le \Delta_{(\underline{\psi})}, \Delta_{(\overline{\psi})} \le \widehat{UB}\big) \\ &\ge \mathcal{P}\big((\Delta^{(1)}, \ldots, \Delta^{(N)}) \in C\big) \\ &\ge 1 - \alpha_{\text{in}} \end{aligned} \tag{19}$$

for $\min\{K_1^{(i)}, \ldots, K_1^{(N)}\} \to \infty$ by (16), so that (15) yields that $[\widehat{LB}, \widehat{UB}]$ is an asymptotic confidence interval for the SCR with level $(1 - \alpha_{\text{out}} - \alpha_{\text{in}})$. The following proposition summarizes the foregoing:

Proposition 4.1. *The confidence interval* $[\widehat{LB}, \widehat{UB}]$ *for the SCR, where* $\widehat{LB}$ *and* $\widehat{UB}$ *are defined in (17) and (18), has an asymptotic confidence level of* $(1 - \alpha_{\text{out}} - \alpha_{\text{in}})$.

It is necessary to note, however, that this confidence interval will in general be very conservative since there are several steps where we underestimate the "true" confidence level. More specifically, on the one hand, the outer confidence level $\mathcal{P}(\Delta_{(\underline{\psi})} \le \text{SCR} < \Delta_{(\overline{\psi})})$ may be strictly greater than $(1 - \alpha_{\text{out}})$ due to the discreteness of the binomial distribution. On the other hand, the inequalities in (15), (16), and (19) will generally not be tight. Hence, our actual confidence level in many cases will be considerably higher than $(1 - \alpha_{\text{out}} - \alpha_{\text{in}})$.

4.2. Choice of Parameters

Clearly, the length of the confidence interval depends on the choice of the parameters, and our aim is to find the shortest confidence interval for the SCR given a fixed computational budget $\Gamma = c \cdot K_0 + K_1 \cdot N$. For the sake of simplicity, we fix α_{out}, α_{in}, and α_{AC_0} although they could easily be included in the optimization process.

Let i_{LB} be the index such that $\widehat{LB} = \widetilde{\Delta}^{(i_{LB})}(K_1) - z_{\text{AC}_0}(K_0) - z_{\text{AC}_1}^{(i_{LB})}(K_1, N)$, and let i_{UB} be the index such that $\widehat{UB} = \widetilde{\Delta}^{(i_{UB})}(K_1) + z_{\text{AC}_0}(K_0) + z_{\text{AC}_1}^{(i_{UB})}(K_1, N)$. Then the length of the confidence interval is given by

$$\widehat{UB} - \widehat{LB} = \widetilde{\Delta}^{(i_{UB})}(K_1) - \widetilde{\Delta}^{(i_{LB})}(K_1) + 2 \cdot z_{\mathrm{AC}_0}(K_0) + z_{\mathrm{AC}_1}^{(i_{UB})}(K_1, N) + z_{\mathrm{AC}_1}^{(i_{LB})}(K_1, N).$$

In order to obtain an estimate for this length based on a pilot simulation, we fix $\tilde{K}_0$ sample paths for the estimation of AC_0, $\tilde{N}$ real-world scenarios, and $\tilde{K}_1$ inner simulations. We derive the corresponding confidence interval as described in the first part of this section and denote the lower and upper limit by $\widehat{LB}_{\text{pilot}}$ and $\widehat{UB}_{\text{pilot}}$, respectively, where $i_{LB,\text{pilot}}$ and $i_{UB,\text{pilot}}$ denote the corresponding indices.

For our approximation of the length of the confidence interval, similarly to Lan et al. (2007b), we make the following assumptions:

1. Sample standard deviations can be approximated by the pilot simulation.
2. K_0 and K_1 are sufficiently large so that the quantiles of the t-distribution can be approximated by those of the standard Normal distribution.
3. The (approximate) length of the outer confidence interval for N real-world scenarios can be derived from the pilot simulation by

$$\widetilde{\Delta}^{(i_{UB})}(K_1) - \widetilde{\Delta}^{(i_{LB})}(K_1) \approx \sqrt{\frac{\tilde{N}}{N}} \left(\widetilde{\Delta}^{(i_{UB,\text{pilot}})}(\tilde{K}_1) - \widetilde{\Delta}^{(i_{LB,\text{pilot}})}(\tilde{K}_1)\right).$$

Assumption 3 can be motivated by assuming that the $\Delta^{(i)}$ are i.i.d. Normally distributed with known variance; in this case the length of a confidence interval for the SCR based on N_1 samples could be derived as $\sqrt{\frac{\tilde{N}_1}{N_1}}$ times the length of the confidence interval based on $\tilde{N}_1$ samples. While of course in general the samples will not be i.i.d. Normal and the variance will be unknown, Assumption 3 may still deliver a reasonable approximation and, thus, provide some guidance on how to choose the simulation parameters.

Based on these assumptions, the length of the confidence interval can be approximated by

$$\begin{aligned}\widehat{UB} - \widehat{LB} \approx& \sqrt{\frac{\tilde{N}}{N}} (\widetilde{\Delta}^{(i_{UB,\text{pilot}})}(\tilde{K}_1) - \widetilde{\Delta}^{(i_{LB,\text{pilot}})}(\tilde{K}_1)) + 2 \cdot z_{1-\frac{\alpha_{\mathrm{AC}_0}}{2}} \frac{\tilde{\sigma}_0(\tilde{K}_0)}{\sqrt{K_0}} \\ &+ z_{1-\frac{\varepsilon}{2}} \frac{\tilde{\sigma}_1^{(i_{UB,\text{pilot}})}(\tilde{K}_1)}{(1+s(0,1))\sqrt{K_1}} + z_{1-\frac{\varepsilon}{2}} \frac{\tilde{\sigma}_1^{(i_{LB,\text{pilot}})}(\tilde{K}_1)}{(1+s(0,1))\sqrt{K_1}},\end{aligned}$$

where z_α denotes the α-quantile of the standard Normal distribution, and the optimization problem is to minimize this length subject to the budget restriction $\Gamma = c \cdot K_0 + K_1 \cdot N$. While it cannot be solved in closed form, from the first order condition with respect to K_1 we obtain

$$K_1 = \frac{\Gamma}{N + c^{\frac{1}{3}} \cdot \zeta_1^{\frac{2}{3}} \cdot \zeta_2^{-\frac{2}{3}} \cdot N^{\frac{2}{3}}}, \tag{20}$$

where

$$\zeta_1 := z_{1-\frac{\alpha_{\mathrm{AC}_0}}{2}} \tilde{\sigma}_0(\tilde{K}_0) \quad \text{and} \quad \zeta_2 := z_{1-\frac{\varepsilon}{2}} \cdot \frac{\tilde{\sigma}_1^{(i_{UB},\mathrm{pilot})}(\tilde{K}_1) + \tilde{\sigma}_1^{(i_{LB},\mathrm{pilot})}(\tilde{K}_1)}{2 \cdot (1+s(0,1))}.$$

Hence, for fixed Γ and N, the optimal K_1 is given by (20) and since $K_0 = \frac{\Gamma - N \cdot K_1}{c}$ the dimension of our optimization problem is reduced to one. Then, numerical methods can be applied to solve the univariate problem for the optimal N.

5. Screening Procedures

As pointed out in the previous section, the confidence interval for the SCR may be relatively wide due to several inequalities in its derivation. Screening procedure present a way to increase the efficiency of the simulation approach.

5.1. Confidence Intervals with Screening

The basic idea behind this method is splitting up the estimation process into two parts: Based on a first run of nested simulations, we "screen" out those scenarios that are not likely to belong to the tail of the distribution. Afterwards, we discard all inner simulations of the first run (this is referred to as "restarting") and generate new inner simulations for those scenarios that survived the screening process. The objective is to screen out as many scenarios as possible, so that we can perform many more inner simulations per real-world scenario in the second run, and, this way, obtain more reliable results. However, when using screening procedures, we have an additional source of error in our computations because we potentially screen out scenarios belonging to the tail.

We follow Lan et al. (2010), who describe a screening procedure for expected shortfall based on nested simulations. Given N_1 real-world scenarios, we simulate a certain number $K_{1,1}$ of inner sample paths for each scenario. The estimated loss in real-world scenario i is denoted by $\tilde{\Delta}^{(i)}(K_{1,1}) = \widetilde{\mathrm{AC}}_0(K_0) - \frac{\widetilde{\mathrm{AC}}_1^{(i)}(K_{1,1})}{1+s(0,1)}$. Based on this first run of inner simulations, we would now like to screen out all scenarios with a "small" loss, i.e. which do not belong to the tail of the $\alpha \cdot N_1$ largest losses. In doing so, we define an "error probability" α_{screen} and keep all scenarios in the set

$$I := \left\{ i : \sum_{j \neq i} \mathbb{1}_{\left\{ \tilde{\Delta}^{(i)}(K_{1,1}) < \tilde{\Delta}^{(j)}(K_{1,1}) - t_{f^{(i,j)},1-\varrho} \sqrt{\frac{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2 + \left(\tilde{\sigma}_1^{(j)}(K_{1,1})\right)^2}{(1+s(0,1))^2 K_{1,1}}} \right\}} < N_1 - \underline{\psi} + 1 \right\} \qquad (21)$$

where $\varrho := \frac{\alpha_{\mathrm{screen}}}{(N_1 - \underline{\psi} + 1)(\underline{\psi} - 1)}$, $\underline{\psi}$ is defined by Equation (14), and $t_{f^{(i,j)},1-\varrho}$ is the $(1-\varrho)$-quantile of the t-distribution with $f^{(i,j)}$ degrees of freedom. Here,

$$f^{(i,j)} := \left\lfloor (K_{1,1}-1)\left(1+\frac{2}{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})/\tilde{\sigma}_1^{(j)}(K_{1,1})\right)^2+\left(\tilde{\sigma}_1^{(j)}(K_{1,1})/\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2}\right)\right\rfloor,$$

which is a consequence of the Welch-Satterthwaite equation. The specific choice of ϱ is required for the proof of the confidence level in Proposition 5.2. Thus, we screen out all scenarios where we can find at least $N_1 - \underline{\psi} + 1$ other realizations yielding a higher loss with a certain predetermined probability. The number of "surviving" scenarios is denoted by $N_2 = |I|$.

Of course, one may also consider screening out scenarios in which the losses are too large, i.e. where we can find at most $N_1 - \overline{\psi} - 1$ other scenarios where the loss is higher with a predetermined probability. However, since we estimate a quantile in the far right tail of the distribution, there will only be very few scenarios that can be screened out in this way. Hence, in most cases this procedure will not be very efficient and thus, it will not be worth the additional computational effort.

In order to limit the number of necessary comparisons, we further use a pre-screening procedure before we start the screening process.[8] Specifically, let $\pi_1(\cdot)$ be a permutation of the indices such that $\tilde{\Delta}^{(\pi_1(i))}$ is non-decreasing in i, and define

$$\tilde{\sigma}_{\max}(K_{1,1}) := \max_{j\in\{\underline{\psi},\dots,N_1\}}\left\{\tilde{\sigma}_1^{(\pi_1(j))}(K_{1,1})\right\} \text{ and}$$

$$t_{\max,1-\varrho} := \max_{i\in\{1,\dots,N_1\}}\left\{\max_{j\in\{\underline{\psi},\dots,N_1\}}\left\{t_{f^{(\pi_1(i),\pi_1(j))},1-\varrho}\right\}\right\}.$$

Then we pre-screen out all scenarios with

$$\tilde{\Delta}^{(i)}(K_{1,1}) < \tilde{\Delta}^{(\pi_1(\underline{\psi}))}(K_{1,1}) - t_{\max,1-\varrho}\sqrt{\frac{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2+(\tilde{\sigma}_{\max}(K_{1,1}))^2}{(1+s(0,1))^2K_{1,1}}},$$

i.e. we pre-screen based on a stricter test using the maximal quantile and the maximal variance in the tail. The great advantage of pre-screening is that actually many scenarios can be screened out by only one comparison, which saves a lot of computational time. Those scenarios that survive pre-screening are screened afterwards. The following proposition shows that screening with and without pre-screening leads to the same result. A proof can be found in the Appendix.

[8] Pre-screening is suggested by Lan et al. (2010) but is not included in their convergence proofs.

Proposition 5.1. *Let $\tilde{I}$ denote the set of scenarios that survive pre-screening, i.e.*

$$\tilde{I} = \left\{ i : \tilde{\Delta}^{(i)}(K_{1,1}) \geq \tilde{\Delta}^{(\pi_1(\underline{\psi}))}(K_{1,1}) - t_{\max,1-\varrho}\sqrt{\frac{(\tilde{\sigma}_1^{(i)}(K_{1,1}))^2 + (\tilde{\sigma}_{\max}(K_{1,1}))^2}{(1+s(0,1))^2 K_{1,1}}} \right\}.$$

Then $I \subseteq \tilde{I}$. Thus, the pre-screening procedure does not screen out scenarios that would survive screening.

Having screened out the irrelevant scenarios, we discard all inner simulations and generate $K_{1,2}^{(i)}$ new inner simulations for each $i \in I$. The corresponding loss estimates and standard deviations are denoted by $\tilde{\Delta}^{(i)}(K_{1,2}^{(i)})$ and $\tilde{\sigma}^{(i)}(K_{1,2}^{(i)})$, respectively, $i = 1, \ldots, N_2$.

We use two different approaches to determine $K_{1,2}^{(i)}$. In the first approach, we allocate the remaining computational budget equally to all scenarios, i.e. $K_{1,2}^{(i)} = K_{1,2}$; within the second allocation, we divide the budget proportional to the variance in the remaining scenarios, i.e.

$$K_{1,2}^{(i)} := \left\lfloor \frac{\left(\Gamma - N_1 \cdot K_{1,1} - c \cdot K_0\right)\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2}{\sum_{j \in I}\left(\tilde{\sigma}_1^{(j)}(K_{1,1})\right)^2} \right\rfloor. \tag{22}$$

To derive a confidence interval, we proceed just like in the previous section. More precisely, we define

$$z_{\mathrm{AC}_0}(K_0) := t_{K_0-1,1-\frac{\alpha_{\mathrm{AC}_0}}{2}} \frac{\tilde{\sigma}_0(K_0)}{\sqrt{K_0}}, \quad \text{and}$$

$$z_{\mathrm{AC}_1}^{(i)}(K_{1,2}^{(i)}, N_2) := t_{K_{1,2}^{(i)}-1,1-\frac{\varepsilon}{2}} \frac{\tilde{\sigma}_1^{(i)}(K_{1,2}^{(i)})}{(1+s(0,1))\sqrt{K_{1,2}^{(i)}}}, \; \varepsilon := 1 - (1-\alpha_{\mathrm{AC}_1})^{\frac{1}{N_2}},$$

where, as before, α_{AC_0} denotes the error resulting from the estimation of AC_0 and α_{AC_1} denotes the error resulting from the estimation of the $\mathrm{AC}_1^{(i)}$, $i \in I$. Now choose $\widehat{LB}$ and $\widehat{UB}$ as the $(\underline{\psi} - (N_1 - N_2))^{th}$ order statistic of $\tilde{\Delta}^{(i)}(K_{1,2}^{(i)}) - z_{\mathrm{AC}_0}(K_0) - z_{\mathrm{AC}_1}^{(i)}(K_{1,2}^{(i)}, N_2)$ and the $(\overline{\psi} - (N_1 - N_2))^{th}$ order statistic of $\tilde{\Delta}^{(i)}(K_{1,2}^{(i)}) + z_{\mathrm{AC}_0}(K_0) + z_{\mathrm{AC}_1}^{(i)}(K_{1,2}^{(i)}, N_2)$, $i \in I$, respectively. Then, we have the following result:

Proposition 5.2. $[\widehat{LB}, \widehat{UB}]$ *is an asymptotically valid confidence interval for the SCR with confidence level* $(1 - \alpha_{\mathrm{out}} - \alpha_{\mathrm{in}})$ *as $K_0 \to \infty$, $K_{1,1} \to \infty$, and $K_{1,2}^{(i)} \to \infty$, where*

$$\alpha_{\mathrm{in}} := 1 - \left(1 - \alpha_{\mathrm{screen}}\right)\left(1 - \alpha_{\mathrm{AC}_0}\right)\left(1 - \alpha_{\mathrm{AC}_1}\right). \tag{23}$$

A proof of the proposition is provided in the Appendix. Note that this confidence interval again will generally be very conservative due to the many inequalities used in the derivation.

Aside from a confidence interval, we may also compute a point estimate $\widehat{\mathrm{SCR}}^{\mathrm{screen}}$ for the SCR, which is given by the $(m-(N_1-N_2))^{th}$ order statistic of $\tilde{\Delta}^{(i)}(K_{1,2}^{(i)})$, $i \in I$, where $m = \lfloor N_1 \cdot \alpha + 0.5 \rfloor$. Clearly, this estimate is based on the assumption that if we had also computed the losses $\tilde{\Delta}^{(i)}(K_{1,2}^{(i)})$ for those real-world scenarios that were screened out, they would have been smaller than the $(m-(N_1-N_2))^{th}$ order statistic of $\tilde{\Delta}^{(i)}(K_{1,2}^{(i)})$, $i \in I$. Under this assumption, the $(m-(N_1-N_2))^{th}$ order statistic of $\tilde{\Delta}^{(i)}(K_{1,2}^{(i)})$, $i \in I$, coincides with the m^{th} order statistic of $\tilde{\Delta}^{(i)}(K_{1,2}^{(i)})$, $1 \leq i \leq N_1$, i.e. this estimate for the SCR is the same as the point estimator from the basic nested simulations approach with N_1 real-world scenarios and $K_{1,2}^{(i)}$ inner simulations. Hence, if $K_{1,2}^{(i)} > K_1^{(i)}$, where $K_1^{(i)}$ denotes the number of inner simulations in the basic nested simulations approach with N_1 real-world scenarios and the same computational budget Γ, the point estimate resulting from the screening procedure will be more precise than the point estimator from the basic nested simulations approach because of the higher number of inner simulations. However, in general the assumption that all estimated losses in those scenarios that have been screened out would be smaller is problematic because we may have screening mistakes. More specifically, it is possible that we have screened out a scenario where $\tilde{\Delta}^{(i)}(K_{1,2}^{(i)})$ is greater than the $(m-(N_1-N_2))^{th}$ order statistic of $\tilde{\Delta}^{(i)}(K_{1,2}^{(i)})$, $i \in I$. Hence, screening introduces an additional type of bias in our point estimate. This bias will be negative, since we may have replaced one of the tail scenarios by a scenario with a smaller loss, i.e. it will generally lead to an underestimation of the SCR. Note, however, that we have a positive bias originating from the uncertainty associated with the inner simulation (cf. Section 3.3), so that the two biases may potentially offset each other.

If we only aim for a good point estimator for the SCR, we may further adapt the approach from Liu et al. (2010) to our problem. Here, the authors use multiple stages of screening to estimate the expected shortfall. However, they note that their "procedure does not provide confidence intervals nor guarantees a minimum probability of correctly identifying the tail." Moreover, an alternative screening algorithm and a heuristic approach to optimize the budget allocation is presented by Broadie et al. (2011). Here, the authors seek to sequentially allocate the computational budget within the inner simulations such that the marginal impact of one additional inner scenario is maximal. While this algorithm could be adapted to our setting, we leave the further exploration for future research.

5.2. Efficient Use of Screening Procedures

For a fixed computational budget, the efficiency of the screening procedure described in the previous subsection obviously depends on the choice of K_0, $K_{1,1}$, and N_1. If we allocate too much of our budget to the screening procedure, there is only a small budget left for the second run. However, choosing the budget for the screening procedure "too small" results in a high number of survivors and thus, the remaining budget for the second run has to be divided between "too many" scenarios. In this section, we describe a procedure how to choose N_1 approximately optimal to minimize the length of the confidence interval for fixed $K_{1,1}$ and K_0, and a given computational budget $\Gamma = cK_0 + N_1 K_{1,1} + N_2 K_{1,2}$. The approach again uses the basic ideas from the adaptive procedure in Lan et al. (2007b).

We first consider the case where the remaining budget is allocated equally to all survivors in the second run. Furthermore, we fix α_{out}, α_{in}, α_{AC_0}, and α_{screen}. α_{AC_1} can then be derived from these values as (cf. Equation (23)):

$$\alpha_{\text{AC}_1} = 1 - \frac{(1-\alpha_{\text{in}})}{(1-\alpha_{\text{screen}}) \cdot (1-\alpha_{\text{AC}_0})}. \tag{24}$$

Akin to the optimization approach for confidence intervals without screening (cf. Section 4.2), let i_{LB} be the index such that $\widehat{LB} = \tilde{\Delta}^{(i_{LB})}(K_{1,2}) - z_{\text{AC}_0}(K_0) - z_{\text{AC}_1}^{(i_{LB})}(K_{1,2}, N_2)$, and i_{UB} be the index such that $\widehat{UB} = \tilde{\Delta}^{(i_{UB})}(K_{1,2}) + z_{\text{AC}_0}(K_0) + z_{\text{AC}_1}^{(i_{UB})}(K_{1,2}, N_2)$. Then the length of the confidence interval is given by

$$\begin{aligned}\widehat{UB} - \widehat{LB} = {} & \tilde{\Delta}^{(i_{UB})}(K_{1,2}) - \tilde{\Delta}^{(i_{LB})}(K_{1,2}) + 2 \cdot z_{\text{AC}_0}(K_0) + z_{\text{AC}_1}^{(i_{UB})}(K_{1,2}, N_2) \\ & + z_{\text{AC}_1}^{(i_{LB})}(K_{1,2}, N_2).\end{aligned}$$

Our objective is now to predict this length for different choices of N_1 based on a pilot simulation with $\tilde{N}_1$ real-world scenarios, $K_{1,1}$ inner simulations, and K_0 sample paths for the estimation of AC_0.[9] Within the pilot simulation, we perform the first run and compute the resulting confidence interval as described in Section 4.1, the only difference being that we use α_{AC_1} from Equation (24). The resulting confidence interval is denoted by $[\widehat{LB}_{\text{pilot}}; \widehat{UB}_{\text{pilot}}]$ with corresponding indices $i_{LB,\text{pilot}}$ and $i_{UB,\text{pilot}}$, respectively. Subsequently, we apply the screening procedure to the results from the first run of the pilot simulation.

Similar to Lan et al. (2007b) and Section 4.2, we make the following assumptions:

1. For fixed K_0 and $K_{1,1}$, the fraction of scenarios that survive screening does not depend on the number of real-world scenarios N_1, i.e.

$$\frac{\tilde{N}_2}{\tilde{N}_1} \approx \frac{N_2}{N_1},$$

where $\tilde{N}_2$ is the number of scenarios that survive screening in the pilot simulation.

[9] Note that once N_1, $K_{1,1}$, and K_0 are specified, the number of survivors N_2 and the number of inner simulations in the second run $K_{1,2}$ result from the screening procedure and the budget constraint.

2. The sample standard deviations can be approximated by the pilot simulation.
3. The length of the outer confidence interval for N_1 real-world scenarios can be approximated from the length for $\tilde{N}_1$ scenarios by

$$\tilde{\Delta}^{(i_{UB})}(K_{1,2}) - \tilde{\Delta}^{(i_{LB})}(K_{1,2}) \approx \sqrt{\frac{\tilde{N}_1}{N_1}}\left(\tilde{\Delta}^{(i_{UB,\text{pilot}})}(K_{1,1}) - \tilde{\Delta}^{(i_{LB,\text{pilot}})}(K_{1,1})\right).$$

Based on these assumptions, the length of the confidence interval can be approximated by

$$\begin{aligned}\widehat{UB} - \widehat{LB} \approx & \sqrt{\frac{\tilde{N}_1}{N_1}}\left(\tilde{\Delta}^{(i_{UB,\text{pilot}})}(K_{1,1}) - \tilde{\Delta}^{(i_{LB,\text{pilot}})}(K_{1,1})\right) + 2 \cdot z_{\text{AC}_0}(K_0) \\ & + t_{\widehat{K}_{1,2}-1,1-\frac{\hat{\varepsilon}}{2}} \frac{\widetilde{\sigma_1}^{(i_{LB,\text{pilot}})}(K_{1,1})}{(1+s(0,1))\sqrt{\widehat{K}_{1,2}}} + t_{\widehat{K}_{1,2}-1,1-\frac{\hat{\varepsilon}}{2}} \frac{\widetilde{\sigma_1}^{(i_{UB,\text{pilot}})}(K_{1,1})}{(1+s(0,1))\sqrt{\widehat{K}_{1,2}}},\end{aligned}$$

where $\varepsilon := 1 - (1 - \alpha_{\text{AC}_1})^{\frac{1}{\widehat{N}_2}}$ with $\widehat{N}_2 := \frac{\tilde{N}_2 \cdot N_1}{\tilde{N}_1}$ being the estimated number of survivors. $\widehat{K}_{1,2} := \left\lfloor \frac{(\Gamma - N_1 \cdot K_{1,1} - c \cdot K_0) \cdot \tilde{N}_1}{\tilde{N}_2 \cdot N_1} \right\rfloor$ is the estimated number of inner simulations in the second run. Then, the task is to minimize this length which may be carried out numerically.

If we allocate the remaining budget for the second run proportionally to the variance in the first run, we need to add one more assumption (cf. Lan et al. (2007b)):

(iv) The average variance in a scenario that survives screening does not depend on the original number N_1 of real-world scenarios, i.e.

$$\frac{\sum_{i \in I}\left(\tilde{\sigma}_1^{(i)}(K_{1,2})\right)^2}{N_2} \approx \frac{\sum_{i \in I_{\text{pilot}}}\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2}{\tilde{N}_2},$$

where I_{pilot} denotes the set of scenarios that survives screening in the pilot simulation.

Then we obtain the following expression for the number of inner simulations in the second run:

$$\widehat{K}_{1,2}^{(i_{LB})} := \left\lfloor \left(\Gamma - N_1 \cdot K_{1,1} - c \cdot K_0\right) \cdot \frac{\tilde{N}_1 \cdot \left(\tilde{\sigma}_1^{(i_{LB,\text{pilot}})}(K_{1,1})\right)^2}{N_1 \cdot \sum_{i \in I_{\text{pilot}}}\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2} \right\rfloor.$$

$\widehat{K}_{1,2}^{(i_{UB})}$ is derived analogously. Subsequently, we proceed as in the case of a constant allocation.

6. Application

6.1. Asset and Liability Model

As an example framework for our considerations, we use the model for a single participating term-fix contract introduced in Bauer et al. (2006).

6.1.1. *General Setup*

A simplified balance sheet is employed to represent the insurance company's financial situation (see Table 1). Here A_t denotes the market value of the insurer's asset portfolio, L_t is the policyholder's statutory account balance, and $R_t = A_t - L_t$ are the free funds (also referred to as "reserve") at time t.

TABLE 1

Simplified balance sheet.

Assets	Liabilities
A_t	L_t R_t
A_t	A_t

Disregarding debt financing, the total assets A_0 at time zero derive from two components, the policyholder's account balance ("liabilities") and the shareholders' capital contribution ("equity"). Ignoring charges as well as unrealized gains or losses, these components are equal to the single up-front premium L_0 and the reserve at time zero, R_0, respectively.

For the bonus distribution scheme, i.e. for modeling the evolution of the liabilities, we rely on the so-called MUST-case from Bauer et al. (2006). This distribution mechanism describes what insurers are obligated to pass on to policyholders according to German regulatory and legal requirements: On the one hand, companies are obligated to guarantee a minimum rate of interest g on the policyholder's account; on the other hand, according to the regulation about minimum premium refunds in German life insurance, a minimum participation rate δ of the earnings on book values has to be credited to the policyholder's account.[10] Since earnings on book values usually do not coincide with earnings on market values due to accounting rules, we assume that earnings on book values amount to a portion y of the latter.

In case the asset returns are so poor that crediting the guaranteed rate g to the policyholder's account will result in a negative reserve R_t, the insurer will default due to the shareholders' limited liability (cf. the notion of a "shortfall"

[10] These earnings reflect the investment income on all assets, including the assets backing shareholders' equity R_t; this reduces the shareholders' return on investment.

in Kling et al. (2007)). However, under Solvency II the market value of liabilities should be calculated under the supposition that shareholders cover any deficit. In accordance with this requirement, we assume that the company obtains an additional contribution c_t from its shareholders in case of such a shortfall. To compensate them for the adopted risk, dividends d_t may be paid to the shareholders each period.

Therefore, the earnings on market values equal to $A_t^- - A_{t-1}^+$, where A_t^- and $A_t^+ = A_t^- - d_t + c_t$ describe the market value of the asset portfolio immediately before and after the dividend payments d_t and capital contributions c_t at time $t \in \mathbb{N}$, respectively. Moreover, we have

$$L_t = (1+g)L_{t-1} + [\delta y(A_t^- - A_{t-1}^+) - gL_{t-1}]^+, \quad t = 1, \ldots, T.$$

Assuming that the remaining part of earnings on book values is paid out as dividends, we obtain

$$\begin{aligned} d_t &= (1-\delta)y(A_t^- - A_{t-1}^+)\,\mathbb{1}_{\{\delta y(A_t^- - A_{t-1}^+) > gL_{t-1}\}} \\ &\quad + [y(A_t^- - A_{t-1}^+) - gL_{t-1}]\,\mathbb{1}_{\{\delta y(A_t^- - A_{t-1}^+) \le gL_{t-1} \le y(A_t^- - A_{t-1}^+)\}}. \end{aligned}$$

Obviously, dividend payments equal zero whenever a capital contribution is required. Therefore, the capital contribution at time t can be described as

$$c_t = \max\{L_t - A_t^-, 0\}.$$

For more details on the contract model we refer to Bauer et al. (2006).

6.1.2. *Relevant Quantities*

Since we have a lump sum premium payment and since we did not model any expenses or tax payments, the only liability cash flow for the contract model specified above is the benefit L_T paid to the policyholder at time $t = T$. Therefore, the Available Capital at time $t = 0$ can be described as follows:

$$\mathrm{AC}_0 = A_0 - \mathrm{MVL}_0 = A_0 - \mathbb{E}^Q\left[\frac{L_T}{B_0(T)}\right].$$

Analogously, we can determine the Available Capital at $t = 1$ as

$$\mathrm{AC}_1 = A_1^- - \mathbb{E}^Q\left[\frac{L_T}{B_1(T)}\middle|\mathcal{F}_1\right].$$

6.1.3. *Asset Model*

For the evolution of the financial market, similarly to Zaglauer and Bauer (2008), we assume a generalized Black-Scholes model with stochastic interest rates (Vasicek model). The asset process and the short rate process evolve according to the stochastic differential equations

$$dA_t = \mu A_t dt + \rho\sigma_A A_t\, dW_t + \sqrt{1-\rho^2}\,\sigma_A A_t dZ_t,\ A_0 > 0, \quad \text{and}$$
$$dr_t = \kappa(\xi - r_t)\,dt + \sigma_r\, dW_t,\ r_0 > 0,$$

respectively, where $\rho \in [-1,1]$ describes their correlation, $\mu \in \mathbb{R}$, σ_A, κ, ξ, $\sigma_r > 0$, and W and Z are two independent Brownian motions under the real-world measure P. Hence, the market value of the assets at $t = 1$ can be expressed as

$$A_1^- = A_0 \exp\left(\mu - \frac{\sigma_A^2}{2} + \rho\sigma_A W_1 + \sqrt{1-\rho^2}\,\sigma_A Z_1\right),$$

and for the short rate process, we have

$$r_1 = e^{-\kappa} r_0 + \xi(1-e^{-\kappa}) + \int_0^1 \sigma_r e^{-\kappa(t-s)} dW_s.$$

Moreover, we assume that the market price of interest rate risk is constant and denote it by λ. Then, we obtain the following dynamics under the risk-neutral measure Q:

$$dA_t = r_t A_t dt + \rho\sigma_A A_t d\tilde{W}_t + \sqrt{1-\rho^2}\,\sigma_A A_t d\tilde{Z}_t,$$
$$dr_t = \kappa(\tilde{\xi} - r_t)dt + \sigma_r d\tilde{W}_t,$$

where $\tilde{\xi} = \xi - \frac{\lambda\sigma_r}{\kappa}$, and $\tilde{W}$ and $\tilde{Z}$ are two independent Brownian motions under Q. Hence, under Q, we have

$$A_t^- = A_{t-1}^+ \exp\left(\int_{t-1}^t r_s\, ds - \frac{\sigma_A^2}{2} + \rho\sigma_A(\tilde{W}_t - \tilde{W}_{t-1}) + \sqrt{1-\rho^2}\,\sigma_A(\tilde{Z}_t - \tilde{Z}_{t-1})\right),$$
$$r_t = e^{-\kappa} r_{t-1} + \tilde{\xi}(1-e^{-\kappa}) + \int_{t-1}^t \sigma_r e^{-\kappa(t-s)} d\tilde{W}_s,$$

and

$$\int_{t-1}^t r_s\, ds = \frac{r_{t-1} - \tilde{\xi}}{\kappa}(1-e^{-\kappa}) + \tilde{\xi} + \frac{\sigma_r}{\kappa}\int_{t-1}^t (1 - e^{-\kappa(t-s)})d\tilde{W}_s,\ t = 2, \ldots, T,$$

which can be conveniently used in Monte Carlo algorithms since $\log\left(\frac{A_t^-}{A_{t-1}^+}\right)$, $r_t | r_{t-1}$, and $\int_{t-1}^t r_s\, ds \,|\, r_{t-1}$ follow a joint Normal distribution (cf. Zaglauer and Bauer (2008)).

We estimate the parameters for our asset model from German data from June 1998 to June 2008 using a Kalman filter. The parameters for the asset

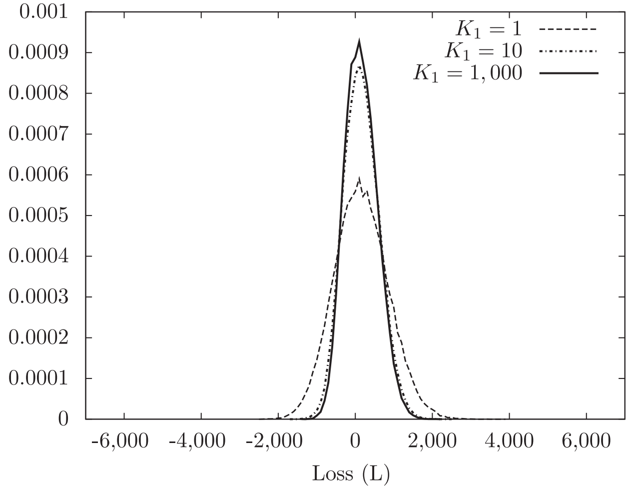

FIGURE 2: Empirical density function for different choices of K_1; N = 100,000, K_0 = 250,000.

portfolio are calibrated to an index consisting of 80% REXP (a total return index of the German bond market) and 20% DAX (a total return index of the German stock market). For the estimation of the short rate process, we use interest rates for government bonds with maturities of 3 months, 1, 3, 5, and 10 years. We obtain the following results: The drift of the asset process is $\mu = 4.25\%$, and its volatility is $\sigma_A = 4.28\%$. For the short rate process we have $\kappa = 14.49\%$, $\xi = 3.64\%$, and $\sigma_r = 0.6\%$. The initial value of the short rate is $r_0 = 4.19\%$. The estimated correlation is $\rho = -0.0597$ and the market price of risk is $\lambda = -0.5061$.

For the insurance contract, similarly to Bauer et al. (2006), we assume a guaranteed minimum interest rate of $g = 3.5\%$, a minimum participation rate of $\delta = 90\%$, an initial premium of L_0 = 10,000, and a maturity of $T = 10$. Moreover, we assume that $y = 50\%$ of earnings on market values are declared as earnings on book values and that the initial reserve quota equals $x_0 = R_0 / L_0 = 10\%$, i.e. $R_0 = x_0 \cdot L_0 = 1{,}000$.

6.2. Results

In Sections 3 to 5, we introduced different methods on how to estimate the SCR and corresponding confidence intervals. In what follows, we implement them in the setup described in Section 6.1. In particular, we focus on the optimal parameter choice for the different methods in view.

6.2.1. *Nested Simulations Approach*

As indicated in Section 3.3, the estimation of the SCR using nested simulations is biased. In order to develop an idea for the magnitude of this bias, we analyze the results for different numbers of inner simulations. We fix $K_0 = 250{,}000$ sample paths for the estimation of AC_0, $N = 100{,}000$ realizations for the simulation over the first year, and choose $K_1^{(i)} = K_1\ \forall 1 \leq i \leq N$.

In Figure 2, the empirical density function of the loss Δ is plotted for different choices of K_1. As expected, the distribution is more dispersed for small K_1, which has a tremendous impact on our problem of estimating a quantile in the tail: We significantly overestimate the SCR for small choices of K_1. This is further documented in Table 2, where the estimated SCR for different choices of K_1 is displayed.

TABLE 2

ESTIMATED SCR AND ESTIMATED SOLVENCY RATIO FOR DIFFERENT CHOICES OF K_1; $K_0 = 250{,}000$, $N = 100{,}000$

K_1	$\widetilde{\mathrm{SCR}}$	$\widetilde{\mathrm{AC}}_0 / \widetilde{\mathrm{SCR}}$
1	1,994.0	94%
5	1,404.7	134%
10	1,332.7	141%
100	1,261.2	149%
1,000	1,246.3	151%

The above results show that a proper allocation of numerical resources, i.e. a careful choice of K_0, K_1, and N, is inevitable in order to obtain accurate results. In order to find (approximately) optimal combinations of K_0, K_1, and N, we estimate the unknown quantities σ_0, f, and θ_α from a pilot simulation with $\tilde{K}_0 = 250{,}000$ sample paths for the estimation of AC_0, $\tilde{N} = 100{,}000$ real-world scenarios, and $\tilde{K}_1 = 200$ inner simulations[11]. Based on these scenarios, we calculate the empirical variances $(\tilde{\sigma}_1^{(i)}(\tilde{K}_1))^2$ for each real-world scenario i, $i = 1, \ldots, \tilde{N}$, and estimate the expected conditional variance via a regression analysis, where we assume

$$\mathbb{E}^Q\left[\mathrm{Var}\left(\tilde{Z}^{K_1} \mid (Y_s)_{s\in[0,1]}\right) \mid \Delta\right] \approx \beta_0 + \beta_1 \Delta + \beta_2 \Delta^2. \tag{25}$$

We obtain $\beta_0 \approx 307{,}280$ (131.68), $\beta_1 \approx -21.186$ (0.29457), $\beta_2 \approx 0.012066$ (0.00045348) for the parameters with the corresponding standard errors in parentheses. R^2 is approximately 5%. However, note that the low coefficient of determination can be attributed to us regressing Δ directly on the $(\tilde{\sigma}_1^{(i)}(\tilde{K}_1))^2$, rather than on $\mathbb{E}^Q\left[\mathrm{Var}\left(\tilde{Z}^{K_1} \mid (Y_s)_{s\in[0,1]}\right) \mid \Delta\right]$, i.e. we have an additional noise term on the left hand side of Equation (25). If we instead partition the first

[11] In practical applications, it may be necessary to determine the SCR on a regular basis. In this case, it might be possible to reduce the computational effort by using the previous period's simulation as the pilot simulation if both the portfolio composition and the capital market parameters remain fairly stable.

year realizations to buckets that yield a similar value of Δ and then calculate the average variances for the regression, the regression function changes only slightly whereas the R^2 increases considerably, although of course this approach comes with an additional source of error due to the choice of the partition. Thus, the low R^2 results from our regression approach rather than a problem with the specification. In particular, sensitivity analyses with respect to the regression function show that the optimal choice of K_0, K_1, and N is rather insensitive when additionally including higher order terms.

In a second step, we derive the empirical density function and approximate its derivative by the average of left- and right-sided finite differences. In this case, sensitivity analyses indicate that the obtained results are not very exact due to the rather small number of observations in the tail. Nevertheless, our estimates provide a rough idea of the optimal ratio. Based on these preliminary calculations and Equation (11), we obtain an estimate for θ_α which is given by $\theta_\alpha \approx 0.027$. σ_0 is approximated by the empirical standard deviation.

In order to obtain an accurate estimate of the 99.5% quantile based on the empirical distribution function, we choose a relatively large number of inner simulations, namely $K_1 = 300$. Then, we find that a choice of approximately $N = 320{,}000$ and $K_0 = 1{,}500{,}000$ is optimal, which results in a total budget of $\Gamma = 97{,}500{,}000$ simulations. In this setting, we obtain $\widetilde{\text{SCR}} = 1{,}249.7$ and a solvency ratio of 150%. At first sight, it might be surprising that K_0 should be chosen relatively large compared to the two other parameters. However, reducing the variance of $\widetilde{\text{AC}}_0(K_0)$ is relatively "cheap" compared to reducing the variance of $\frac{\tilde{s}(m)}{1+s(0,1)}$ because whenever we increase N, we automatically have to perform K_1 inner simulations for every additional real-world scenario. Therefore, it is rather intuitive to allocate a rather large budget to the estimation of $\widetilde{\text{AC}}_0(K_0)$.

To demonstrate that, given a total budget of $\Gamma = 97{,}500{,}000$, this choice is roughly adequate, we estimate the SCR 150 times for fixed K_0 and different combinations of N and K_1, where each combination corresponds to a total budget of 97,500,000 simulations. We estimate the bias by $\frac{\bar{\theta}_\alpha}{K_1 \cdot \bar{f}(\widetilde{\text{SCR}})}$, where $\bar{\theta}_\alpha$ and $\bar{f}$ denote the averages of the estimates resulting from the 150 estimation procedures as explained above. This allows us to correct the mean by the estimated

TABLE 3

RESULTS FOR DIFFERENT CHOICES OF N AND K_1; 150 RUNS, $K_0 = 1{,}500{,}000$.

N	K_1	Mean ($\widetilde{\text{SCR}}$)	Empirical Variance	Estimated Bias	Estimated MSE	Corrected Mean
160,000	600	1,247.7	24.6	1.4	26.6	1,246.3
320,000	**300**	**1,249.3**	**15.8**	**2.9**	**24.0**	**1,246.4**
640,000	150	1,251.3	7.9	5.7	40.6	1,245.6
1,500,00	64	1,259.5	3.2	13.2	178.0	1,246.3

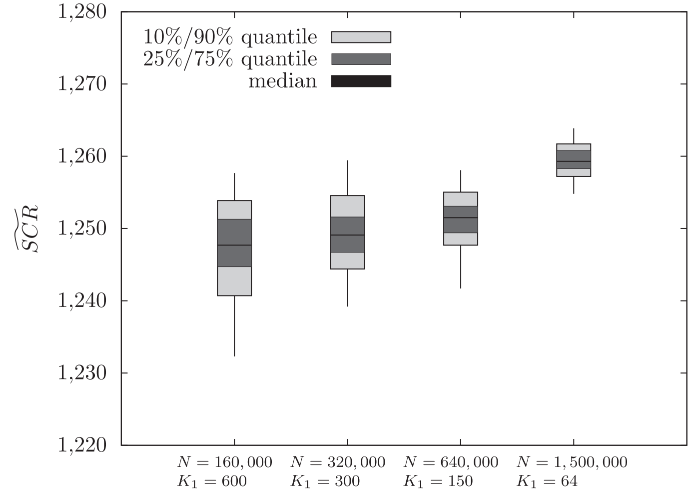

FIGURE 3: 150 runs for different choices of N and K_1; K_0 = 1,500,000.

bias. The MSE is then estimated by the sum of the empirical variance and the squared estimated bias. Figure 3 and Table 3 show our results.

As expected, the mean of the estimated SCRs increases as K_1 decreases due to the increased bias. In contrast, the empirical variance obviously decreases as N increases. Furthermore, we find that our choice of N and K_1 yields the smallest estimated MSE from the combinations given in Table 3. Therefore, our choice appears reasonable within our framework. Moreover, it is worth pointing out that if we correct the means in Table 3 by the corresponding estimated bias, the difference between the results for the different combinations is almost negligible.

Therefore, we will use N = 320,000 and K_1 = 300 in the remaining part of this paper if not stated otherwise, and we refer to this parameter combination as the *base case*. With this parameter combination, it takes about 16 minutes to carry out one run with our C++ implementation[12].

6.2.2. *Jackknife Approach*

In Section 3.4, we introduced a Jackknife procedure to eliminate the bias resulting from the finite number of inner simulations. For our implementation of this approach, we rely on I = 2 non-overlapping sections as advocated by Gordy and Juneja (2010) since — according to them — beyond computational advantages, this choice eliminates nearly all the bias at little cost to variance.

[12] The simulations were carried out on a Windows machine with Intel Core 2 Duo CPU T7500, 2.20GHz, and 2048 MB RAM. Of course, the computational time depends on our particular implementation; optimizations of the code may be possible.

TABLE 4

RESULTS FOR DIFFERENT CHOICES OF N AND K_1; JACKKNIFE ESTIMATOR, 150 RUNS, K_0 = 1,500,000.

		Jackknife Estimator		Nested Simulations Estimator		
N	K_1	Mean ($\widetilde{SCR}^+$)	Empirical Variance	Mean ($\widetilde{SCR}$)	Empirical Variance	Corrected Mean
160,000	600	1,246.3	30.0	1,247.7	24.6	1,246.3
320,000	300	1,246.5	19.2	1,249.3	15.8	1,246.4
640,000	150	1,245.6	9.3	1,251.3	7.9	1,245.6
1,500,000	64	1,246.4	4.9	1,259.5	3.2	1,246.3

Table 4 displays our results for the Jackknife estimator relative to those for the simple nested simulations estimator. Here, we rely on the same parameter combinations as in Table 3.

As might be expected, our results indicate that the Jackknife estimator eliminates most of the bias at the cost of a higher variance. Comparing the ensuing mean square errors, for relatively small choices of K_1, the Jackknife estimator appears far superior to the simple nested simulations estimator due to the reduction in bias, whereas for relatively large choices of K_1, the simple estimator dominates due to the reduction in variance. This documents the conclusion from Gordy and Juneja (2010) that within an optimal allocation, the number of inner steps can even be further reduced for the jackknife estimator as already noted in Section 3.4. However, to determine such an (asymptotically) optimal computational allocation for the Jackknife estimator in analogy to Section 3.3 for the basic estimator, it would be necessary to estimate even more complex terms based on the pilot simulations (see Gordy and Juneja (2010) for details).

Within our application, the MSE for the Jackknife estimator is reduced even for the optimal parameter combination for the simple estimator with $K_1 = 300$. However, of course we may correct the result from the simple estimator for the estimated bias, which may change the precedence. Furthermore, aside from the additional computational complexity, the direction of the increased sample error within the Jackknife estimation is unknown whereas we generally know the direction of the bias for the basic estimator. Nonetheless, our results unambiguously support the application of the Jackknife estimator when there is no reliable bias estimate available for the simple nested simulations estimator and the number of inner simulations is small.

6.2.3. *Confidence Intervals*

Having analyzed the basic point estimator, we now proceed with the derivation of confidence intervals for the SCR as described in Section 4. Within our numerical experiments, we aim for a total confidence level of 90%. In a first step, we determine confidence intervals for the base case from the previous subsection. We derive a two-sided confidence interval and choose the indices $\underline{\psi}$ and $\overline{\psi}$ (cf. Equation (14)) such that they are symmetric around $m = \lfloor \alpha \cdot N_1 + 0.5 \rfloor$, which corresponds to the order statistic of the estimated SCR.

Of course, our results depend on the choice of the error levels α_{out} and α_{AC_0}[13]. However, based on sensitivity analyses we found that the influence of this choice on the length of the confidence interval is not very pronounced. Since in our base case the uncertainty arising from the inner simulation dominates the uncertainty arising from the outer simulation and since the estimation error for $\widetilde{\text{AC}}_0(K_0)$ is significantly smaller than that for $\widetilde{\text{AC}}_1^{(i)}(K_1)$, $i = 1, \ldots, N$, $\alpha_{in} = 8\%$ and $\alpha_{\text{AC}_0} = 0.1\%$ seem to be reasonable choices.

In this case, we obtain a confidence interval of [1,073.4; 1,427.6]. Hence, we have a length of 354.2, which corresponds to about 28% of the point estimate $\widetilde{\text{SCR}}$. However, when analyzing the result in more detail, we find that the $\underline{\psi}^{th}$ and $\overline{\psi}^{th}$ order statistics of the estimated losses are given by $\tilde{\Delta}_{\underline{\psi}} = 1{,}241.6$ and $\tilde{\Delta}_{\overline{\psi}} = 1{,}259.4$. Thus, a large portion of the length of the confidence interval has to be attributed to the uncertainty arising from the inner simulation as well as to the lack of precision in its derivation. More specifically, as already pointed out at the end of Section 4.1, the inequalities (14), (16), and (19) generally will not be very tight, so that the inner confidence level will be very conservative. In particular, note that for large N, the allowed error level in all inner simulations $\varepsilon = 1 - (1 - \alpha_{\text{AC}_1})^{\frac{1}{N}}$ is extremely small. Hence, large choices of N are heavily penalized although it is not strictly necessary for the validity of the confidence interval for all $\Delta^{(i)}$ to be in the confidence region with probability $(1 - \alpha_{\text{in}})$, but only for $\Delta_{\underline{\psi}}$ and $\Delta_{\overline{\psi}}$.

This is also reflected in the "optimal" parameter choice described in Section 4.2: For the sake of simplicity, we use the same pilot simulation as in Section 6.2.1, although we found that already pilot simulations with about $N = 10{,}000$ yield suitable estimates. We obtain the following approximately optimal parameters: $N \approx 20{,}000$, $K_1 \approx 4{,}732$, $K_0 \approx 2{,}860{,}000$. Thus, as expected, in comparison to the base case, the number of inner simulations and the number of sample paths for the estimation of AC_0 increase whereas the number of real-world scenarios decreases. Based on these parameters, we obtain a confidence interval of [1,179.9; 1,329.2]. This translates into a solvency ratio between 141% and 159%. The length of the confidence interval is given by 149.3 which corresponds to approximately 12% of $\widetilde{\text{SCR}}$.

To demonstrate that this choice of parameters is roughly adequate for minimizing the length of the confidence interval, we also compute the length of the confidence interval for other numbers of real-world scenarios N, where for each N we calculate the approximately optimal choice of K_0 and K_1. Figure 4 shows our results. The shortest confidence interval with a length of 148.7 is obtained for $N = 30{,}000$. Nevertheless, we find that our choice resulting from

[13] Note that α_{in} and α_{AC_1} are defined by α_{out} and α_{AC_0} and the total confidence level of 90%.

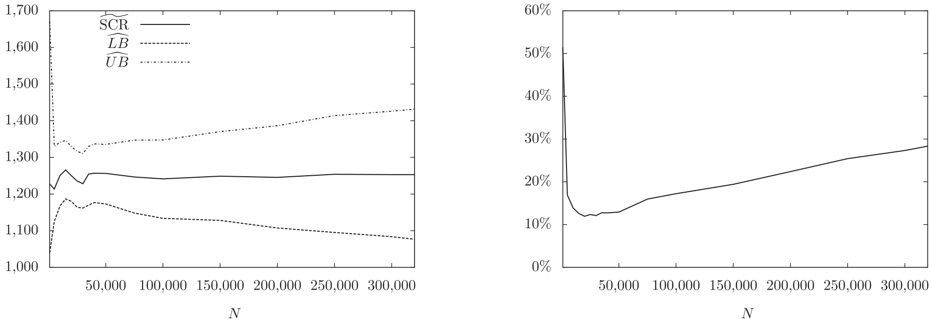

FIGURE 4: 90% Confidence intervals for different N (left), length of confidence interval as percentage of $\widetilde{\text{SCR}}$ (right); $\Gamma = 97{,}500{,}000$.

the optimization algorithm is roughly optimal in the sense that it lies within a range of N where the length of the confidence interval is close to minimal. Compared to our base case with $N = 320{,}000$, the length of the confidence interval with parameters as derived by the optimization approach is decreased to less than 50% of the original length. However, it is important to keep in mind that "optimality" in this case corresponds to the specifics — particularly the rather crude construction — of the confidence interval, whereas optimality with respect to the MSE as for our base case parameters directly pertains the nested simulations estimator. Hence, if we are interested in an accurate point estimate, the MSE certainly appears to be the more satisfactory optimality criterion.

6.2.4. *Screening Procedures*

In the previous subsections, we used a rather large computational budget of $\Gamma = 97{,}500{,}000$ for our calculations. However, within practical applications, due to the complexity of the projection models, it is in general impossible to determine the SCR based on so many sample paths. Therefore, we now apply the screening procedure described in Section 5. This method enables us to either obtain a higher accuracy with the same computational budget or to derive a similar accuracy based on a lower computational budget.

We start by analyzing the influence of screening on our point estimator. In Table 5, we show our results for the base case from Section 6.2.1 ($N = 320{,}000$, $K_0 = 1{,}500{,}000$), $\alpha_{\text{screen}} = 4\%$, and different choices of $K_{1,1}$, i.e. the number of inner simulations within the screening procedure, where we perform 150 runs for each choice of $K_{1,1}$. We find that for all choices, screening improves the results considerably. More precisely, comparing the results to those for the basic estimator from Table 3, we find that screening eliminates most of the bias.

Of course, when comparing the results, we need to keep in mind that the computation takes longer when screening is applied due to the increased number of operations. However, in practical applications, the effort for the projection

TABLE 5

EFFICIENCY OF (PRE-)SCREENING FOR DIFFERENT CHOICES OF $K_{1,1}$;
150 RUNS, $N = 320{,}000$, $K_0 = 1{,}500{,}000$.

$K_{1,1}$	Mean ($\widehat{\text{SCR}}^{\text{screen}}$)	Empirical Variance	Avg. Pre-screened out	Avg. Total screened out	Avg. # inner simulations 2nd run
50	1,247.0	15.5	64%	80%	1,270
100	1,246.6	15.4	89%	92%	2,552
150	1,246.7	15.1	94%	95%	2,984
200	1,246.7	15.6	95%	96%	2,616
250	1,246.8	15.7	96%	97%	1,583

of the insurer's assets and liabilities will in general be the primary source of the numerical complexity such that the additional effort for screening will be negligible. Moreover, our analyses indicate that pre-screening is very efficient. We find that in most cases more than 90% of the real-world scenarios are pre-screened out, whereas the subsequent screening procedure only eliminates a few additional percentage points. Only for the relatively small choice of $K_{1,1} = 50$ we see a significant efficiency gain by screening in comparison to pre-screening.

Hence, a large part of the total number of scenarios that are screened out is already eliminated by pre-screening, which saves much computational time because pre-screening is much faster than screening. In particular, this implies that for practical applications, it may be possible to solely rely on pre-screening.

Since the number of inner simulations for determining the quantile ($K_{1,2}$) now increases relative to the basic estimator, choosing a higher number of outer simulations N than within our base case may further improve the efficiency of the estimate as measured by the mean square error. For instance, if we choose $N = 480{,}000$ and $K_{1,1} = 100$, 150 runs yield a mean $\widehat{\text{SCR}}^{\text{screen}}$ of 1,247.2 with an empirical variance of 10.9. The average number of inner simulations in the second run is 1,228. Hence, the reduction in variance outweighs the increase in bias relative to the results in Table 5 for $N = 320{,}000$. Therefore, with respect to a suitable choice of the simulation parameters, our results suggest that an increase of the number of outer simulations and a "medium" choice of $K_{1,1}$ relative to the optimal budget for the basic estimator are advisable.

Next, we analyze the influence of screening on the length of our confidence interval. As before, we aim for two-sided 90%-confidence intervals for the SCR. In a first step, we analyze the results of the screening procedure for our base case ($N = 320{,}000$, $K_0 = 1{,}500{,}000$, $\Gamma = 97{,}500{,}000$). As before, we choose $\alpha_{\text{in}} = 8\%$, $\alpha_{\text{AC}_0} = 0.1\%$, and $\alpha_{\text{screen}} = 4\%$. Numerical experiments show that at least in our case different choices of α_{screen} do not have a significant impact on the results. Thus, we choose α_{screen} such that the error due to screening is similar to the error arising from the estimation of $\text{AC}_1^{(i)}$. The remaining budget in the second run is allocated equally to all surviving scenarios. To obtain a

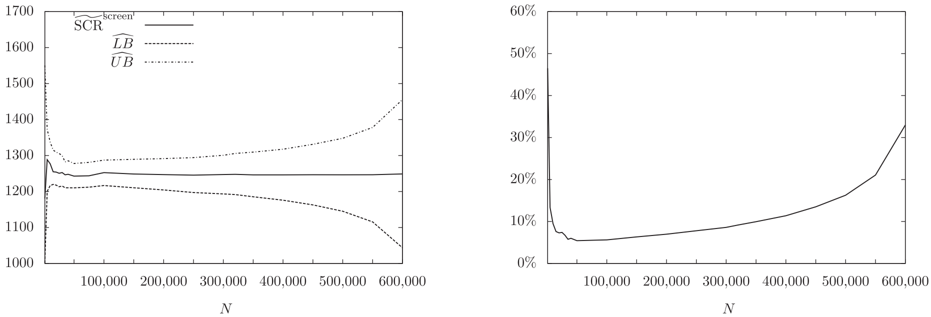

FIGURE 5: 90%-confidence intervals with screening for different N_1 (left), length of confidence interval as percentage of $\widehat{\text{SCR}}$ (right); K_0 = 1,500,000, $K_{1,1}$ = 150, Γ = 97,500,000.

first estimate, we set $K_{1,1} = 150$, i.e. we use half of the maximal number of inner simulations for the first run. In this case, the resulting confidence interval is given by [1,191.5; 1,305.9] and the length corresponds to 9% of the point estimate $\widehat{\text{SCR}}^{\text{screen}}$ = 1,247.8.

In order to optimize the length of the confidence interval as a function of N_1 for given K_0 = 1,500,000 and $K_{1,1}$ = 150, we rely on a pilot simulation with $\tilde{N}_1$ = 10,000 samples. We find that $N_1 \approx 75{,}000$ is optimal. In this case, we obtain the confidence interval [1,212.2; 1,281.0]. Thus, we have a length of 68.8, which corresponds to 6% of the point estimate $\widehat{\text{SCR}}^{\text{screen}}$ = 1,243.8. To show that this choice is approximately adequate, we also derive confidence intervals for other choices of N_1, fixed K_0 = 1,500,000, and $K_{1,1}$ = 150. Figure 5 shows our results. Again, we find that our optimizations approach provides parameters such that the resulting length of the confidence interval is close to minimal. Furthermore, in comparison to the case without screening, the length of the confidence interval is reduced by approximately 50%.

Again, we find that pre-screening is very efficient. More precisely, in all analyzed cases for $K_{1,1}$ = 150, at least 92% of the real-world scenarios are pre-screened out. The subsequent screening procedure eliminates no more than 2 additional percentage points. With respect to the choice of $K_{1,1}$, our sensitivity analyses show that the impact is not very pronounced, i.e. unless $K_{1,1}$ is chosen "too small," we can find an appropriate N_1 such that the confidence interval is close to minimal. Furthermore, we carried out some numerical experiments for an allocation proportional to the variance in the first run. However, at least for our example contract, we found that there is hardly any difference between the two methods.

6.3. Variance Reduction Techniques

Variance reduction techniques present means to further increase the efficiency of our calculations. As an example, we consider the use of antithetic variates although there exists a wide array of different alternatives. We refer to Glasserman

(2004) for more details on variance reduction techniques in general and to Bergmann (2011) for the use of control variates in our context.

The basic idea behind antithetic variates (AV) is to reduce the variance by introducing a negative dependence between pairs of realizations when estimating expected values. In the present context, this means instead of using independent sample paths within the inner simulation step and within the estimation of AC_0, we employ samples of pairs of paths generated based on perfectly negatively correlated Normal random variables.

More precisely, let Z be the vector of standard Normally distributed random variables that is used to derive one sample path of the evolution of our capital market for the estimation of AC_0. Then, a second sample path of the evolution of the capital market can be derived by using $-Z$. The resulting liability cash flows and the resulting money market account are denoted by $\tilde{X}_t$ and $\tilde{B}_0(t)$, $t = 1, \ldots, T$, respectively. Then, the antithetic variates estimator for the market value of liabilities based on K_0^{AV} independent pairs of sample paths is given by

$$\widetilde{\mathrm{MVL}}_0^{\mathrm{AV}}(K_0^{\mathrm{AV}}) := \frac{1}{K_0^{\mathrm{AV}}} \sum_{k=1}^{K_0^{\mathrm{AV}}} \frac{1}{2} \left(\sum_{t=1}^{T} \frac{X_t^{(k)}}{B_0^{(k)}(t)} + \sum_{t=1}^{T} \frac{\tilde{X}_t^{(k)}}{\tilde{B}_0^{(k)}(t)} \right).$$

It can then be shown that

$$\frac{\mathrm{Var}\left(\widetilde{\mathrm{MVL}}_0^{\mathrm{AV}}(K_0^{\mathrm{AV}})\right)}{\mathrm{Var}\left(\widetilde{\mathrm{MVL}}_0(2 \cdot K_0^{\mathrm{AV}})\right)} = 1 + \rho^{\mathrm{AV}},$$

where ρ^{AV} is the correlation of the sum of the discounted liability cash flows in the two antithetic sample paths. Thus, in case the negative correlation of the inputs leads to a negative correlation of the outputs, the use of antithetic variates may significantly reduce the variance of the estimator. Within our calculations, we do not only use antithetic variates for AC_0, but also within each inner simulation.

Table 6 shows our results with and without antithetic variates. We use $K_0^{\mathrm{AV}} = \frac{K_0}{2} = 125{,}000$ and $K_1^{\mathrm{AV}} = \frac{K_1}{2}$ pairs of sample paths in our comparison.

TABLE 6

COMPARISON OF ESTIMATED SCR AND ESTIMATED SOLVENCY RATIO WITH AND WITHOUT ANTITHETIC VARIATES.

K_1	K_1^{AV}	$\widetilde{\mathrm{SCR}}^{\mathrm{AV}}$	$\widetilde{\mathrm{AC}}_0^{\mathrm{AV}} / \widetilde{\mathrm{SCR}}^{\mathrm{AV}}$	$\widetilde{\mathrm{SCR}}$	$\widetilde{\mathrm{AC}}_0 / \widetilde{\mathrm{SCR}}$
4	2	1,286.3	146%	1,436.5	131%
10	5	1,261.7	149%	1,332.7	141%
100	50	1,253.1	150%	1,261.2	149%
1,000	500	1,253.5	150%	1,246.3	151%

TABLE 7

COMPARISON OF NESTED SIMULATIONS APPROACH WITH AND WITHOUT ANTITHETIC VARIATES FOR DIFFERENT PARAMETERS.

	N	K_1, K_1^{AV}	K_0	Mean ($\widehat{SCR}$, $\widehat{SCR}^{AV}$)	Emp. Var	Est. Bias	Est. MSE	Corrected Mean
with AV	1,070,000	45	600,000	1,247.7	4.4	1.7	7.2	1,246.0
with AV	310,000	25	130,000	1,248.8	13.8	3.0	23.1	1,245.7
w/o AV	320,000	300	1,500,000	1,249.3	15.8	2.9	24.0	1,246.4

We notice that the use of antithetic variates clearly improves the estimate significantly indicating considerable gains in the efficiency of the estimation. These findings are further illustrated by Table 7, where different optimal parameter combinations in the sense of Section 3.3 are displayed. We observe that for a fixed computational budget of $\Gamma = 97{,}500{,}000$, the use of antithetic variates reduces the MSE by about 70%. In particular, with antithetic variates, only a budget of 15,760,000 is necessary in order to obtain results of a similar accuracy as for the basic estimator as measured by the MSE.

When applying antithetic variates to the derivation of confidence intervals based on the screening procedure as described in Section 5, and using a computational budget of $\Gamma = 97{,}500{,}000$ as in Section 6.2.4, our pilot simulation suggests that for $K_0^{AV} = 750{,}000$ and $K_{1,1}^{AV} = 75$, $N_1 = 200{,}000$ is approximately optimal in order to minimize its length. The resulting confidence interval is given by [1,222.7; 1,257.0], which corresponds to about 3% of $\widehat{SCR}^{\text{screen}}$.

Considering our "first" confidence interval from Section 6.2.3 with a length of approximately 28% of $\widehat{SCR}$, this efficiency gain by relying on more advanced techniques is remarkable, especially when considering that the confidence interval is very conservative (cf. Section 6.2.3). While of course it is necessary to point out that these efficiency gains — particularly the glaring improvements due to antithetic variates — are closely bound to our example application, our results at the very least should serve as evidence that a suitable simulation design may yield considerable efficiency gains.

7. ESTIMATOR BASED ON THE INDIRECT METHOD

As mentioned in Section 2.1, the market-consistent value of insurance liabilities and thus the Available Capital can be derived based on two approaches — the *direct* and the *indirect* approach. While our technical considerations apply to both methods, so far we have focussed on the direct approach in our exposition. Hence, in what follows, we briefly introduce the indirect method based on the MCEV principles issued by the CFO Forum (2009), adapt it to our example application, and repeat some of the calculations from the previous section. In particular, we show that the quality of the two estimators can differ significantly.

7.1. Available Capital

Within the indirect method (IDM), the Available Capital is derived from the free cash flows generated by the insurance business. Thus, the derivation of the Available Capital is very closely related to the concept of Market-Consistent Embedded Value (MCEV). More precisely, the MCEV corresponds to the present value of shareholders' interest in the earnings distributable from assets backing the life insurance business, after allowance for the aggregate risks in the life insurance portfolio. Here, it is important to note that the MCEV shall not reflect the shareholders' default put option resulting from their limited liability. Consequently, the Available Capital derived under Solvency II principles is usually very similar to the MCEV, so that for the purpose of this paper, we assume that the two quantities coincide[14].

According to the CFO Forum (2009), the MCEV is defined as the sum of the Adjusted Net Asset Value (ANAV) and the Present Value of Future Profits (PVFP) less a Cost-of-Capital charge (CoC):

$$\mathrm{MCEV} := \mathrm{ANAV} + \mathrm{PVFP} - \mathrm{CoC}. \qquad (26)$$

The ANAV is derived from the (statutory) Net Asset Value (NAV)[15] and includes adjustments for intangible assets, unrealized gains and losses on assets etc. It consists of two parts, the free surplus and required capital (cf. Principles 4 and 5 in CFO Forum (2009)). In most cases, the ANAV can be calculated from statutory balance sheet figures and the market value of assets; hence, the calculation does not require simulations. The PVFP corresponds to the present value of post-taxation shareholder cash flows from the in-force business and the assets backing the associated (statutory) liabilities. In particular, it also includes the time value of financial options and guarantees (cf. Principles 6 and 7 in CFO Forum (2009)), so that its calculation presents the primary computational challenge. The CoC is the sum of the frictional cost of required capital and the cost of residual non-hedgeable risks (cf. Principles 8 and 9 in CFO Forum (2009)).

Consequently, we have

$$\mathrm{AC}_0^{\mathrm{IDM}} := \mathrm{MCEV}_0.$$

Assuming that the profit for the first year (denoted by X_1^{IDM}) has not been paid to shareholders yet, the Available Capital at $t = 1$ can be described by

$$\mathrm{AC}_1^{\mathrm{IDM}} := \mathrm{MCEV}_1 + X_1^{\mathrm{IDM}},$$

[14] More specifically, there exist slight differences between the MCEV cost-of-capital and the risk margin under Solvency II, and in the eligibility of certain capital components (e.g. subordinated loans).

[15] For an insurance company, the NAV is defined as the value of its assets less the value of its liabilities based on the statutory balance sheet, and therefore roughly coincides with the statutory equity capital.

where MCEV_1 and X_1^{IDM} denote the MCEV and the shareholder cash flow at time 1.

7.2. Application

Since within our example application introduced in Section 6.1 we ignore unrealized gains and losses on assets as well as other adjustments, here we have $\text{ANAV}_0 = \text{NAV}_0 = R_0$, where R_t denotes the free funds in the statutory balance sheet at time t (see Section 6.1.1 for details). As described above, for the PVFP we have to consider all cash in- and outflows paid to or by the shareholder, respectively. Obviously, this includes all dividend payments d_t and capital injections c_t. Moreover, shareholders may benefit from a favorable evolution of the company in that the market value of their capital contribution increases. More specifically, they may realize $\text{ROI}_T := R_T - B_0(T) \cdot R_0$ as their (time value adjusted) return on investment at the end of the projection period (also referred to as "maturity") T. Thus, disregarding any capital charges, the Available Capital at time $t = 0$ can be described as follows:

$$\begin{aligned}\text{AC}_0^{\text{IDM}} &= R_0 + \mathbb{E}^Q\left[\sum_{t=1}^{T} \frac{(d_t - c_t)}{B_0(t)} + \frac{\text{ROI}_T}{B_0(T)}\right] \\ &= R_0 + \mathbb{E}^Q\left[\sum_{t=1}^{T} \frac{(d_t - c_t)}{B_0(t)} + \frac{R_T}{B_0(T)} - R_0\right] = \mathbb{E}^Q\left[\sum_{t=1}^{T} \frac{X_t^{\text{IDM}}}{B_0(t)}\right],\end{aligned}$$

where

$$X_t^{\text{IDM}} = \begin{cases} d_t - c_t & \text{, if } t \in \{1, \ldots, T-1\} \\ d_T - c_T + R_T & \text{, if } t = T \end{cases}.$$

Similarly, we obtain

$$\text{AC}_1^{\text{IDM}} = \mathbb{E}^Q\left[\sum_{t=2}^{T} \frac{X_t^{\text{IDM}}}{B_1(t)} \middle| \mathcal{F}_1\right] + X_1^{\text{IDM}}.$$

The corresponding nested simulations estimator can be derived in analogy to the estimator based on the direct method (cf. Section 3.2). Furthermore, the optimization procedure described in Section 3.3 as well as the considerations from Section 4 and 5 can easily be adapted for the estimators based on the indirect method.

7.3. Results

Similarly to the estimator derived via the direct method (DM), we analyze the bias of our estimator by calculating the SCR for a fixed number of sample paths

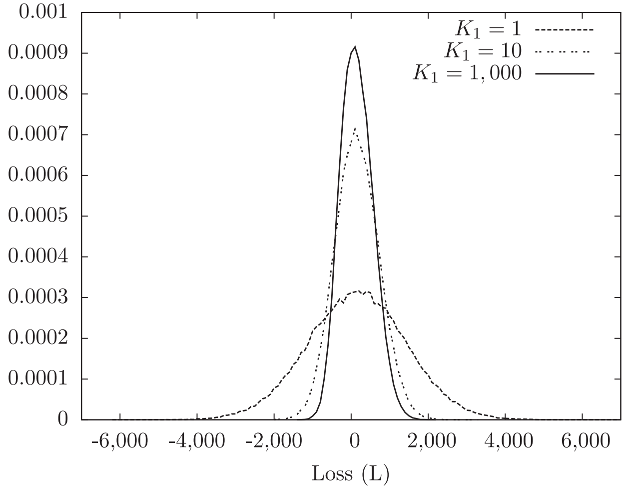

FIGURE 6: Empirical density function for different choices of K_1 for the estimator based on the indirect method; $N = 100{,}000$, $K_0 = 250{,}000$.

for the estimation of AC_0 ($K_0 = 250{,}000$), a fixed number of outer scenarios ($N = 100{,}000$), and varying numbers of inner simulations K_1. Our results are displayed in Figure 6 and Table 8, where for comparison purposes we also include the corresponding results from Section 6.2. We observe that similar to the estimator derived via the direct method, the distributions are more dispersed for smaller K_1 resulting in a higher estimated SCR. However, when comparing the density functions for the two estimators for the same K_1 in Figures 2 and 6, we find that the density function corresponding to the estimator based on the indirect method is even more dispersed.

In line with this result, further analyses show that in our setting, the estimator based on the direct method is superior in most cases. In particular, this highlights that for practical applications, the choice of the estimator for the

TABLE 8

ESTIMATED SCR AND ESTIMATED SOLVENCY RATIO FOR DIFFERENT CHOICES OF K_1; $K_0 = 250{,}000$, $N = 100{,}000$.

K_1	indirect method		direct method	
	$\widetilde{SCR}$	$\widetilde{AC_0} / \widetilde{SCR}$	$\widetilde{SCR}$	$\widetilde{AC_0} / \widetilde{SCR}$
1	3,432.5	55%	1,994.0	94%
5	1,874.6	100%	1,404.7	134%
10	1,606.5	117%	1,332.7	141%
100	1,279.1	147%	1,261.2	149%
1,000	1,254.6	149%	1,246.3	151%

Available Capital may play an important role for the reliability of the estimate, although of course the order of the comparison depends on the application in view. The superiority of the estimate based on the direct method in our setting may be an artifact of the sheer number of stochastic quantities included in the estimation process. More specifically, while for the direct calculation of the Available Capital it is sufficient to simulate L_T (cf. Section 6.1.2), for the indirect method all quantities R_T, c_t, d_t, $1 \leq t \leq T$, enter the expected value calculation.

Similarly as for the direct method, we also carry out some numerical experiments with the use of antithetic variates based on the IDM estimator. As before, we use $K_0^{AV} = \frac{K_0}{2} = 125{,}000$ and $K_1^{AV} = \frac{K_1}{2}$ pairs of sample paths in our comparison. Our results are displayed in Table 9. We find that similar to the direct method, the use of antithetic variates clearly improves the estimate significantly.

TABLE 9

COMPARISON OF ESTIMATED SCR AND ESTIMATED SOLVENCY RATIO WITH AND WITHOUT ANTITHETIC VARIATES (IDM).

K_1	K_1^{AV}	$\widetilde{SCR}^{AV}$	$\widetilde{AC}_0^{AV} / \widetilde{SCR}^{AV}$	$\widetilde{SCR}$	$\widetilde{AC}_0 / \widetilde{SCR}$
4	2	1,275.3	147%	2,024.3	93%
10	5	1,258.7	149%	1,606.5	117%
100	50	1,251.4	150%	1,279.1	147%
1,000	500	1,252.6	150%	1,254.6	149%

Table 10 shows the analog of Table 7 for the indirect method. We observe that for a fixed computational budget of $\Gamma = 97{,}500{,}000$, the reduction in the MSE due to the use of antithetic variates is even more pronounced for the indirect method. Here, the MSE is reduced by almost 90%. In particular, with antithetic variates, only a budget of 2,352,000 is necessary in order to obtain results of a similar accuracy — in terms of MSE — as for the basic estimator. Furthermore, we find that the estimator based on the indirect method is

TABLE 10

COMPARISON OF NESTED SIMULATIONS APPROACH WITH AND WITHOUT ANTITHETIC VARIATES FOR DIFFERENT PARAMETERS (IDM).

	N	K_1, K_1^{AV}	K_0	Mean ($\widetilde{SCR}$, $\widetilde{SCR}^{AV}$)	Emp. Var	Est. Bias	Est. MSE	Corrected Mean
with AV	1,375,000	35	625,000	1,247.5	4.8	1.5	7.0	1,246.0
with AV	115,000	10	26,000	1,250.9	46.0	5.3	73.7	1,245.7
w/o AV	105,000	920	900,000	1,250.4	48.7	4.4	68.4	1,246.0

slightly superior. This is in contrast to the analysis without variance reduction, where the estimator based on the direct method generally performs better.

8. Conclusion

In this paper, we provide a detailed discussion of how to determine the Solvency Capital Requirement within the Solvency II framework based on nested simulations. In particular, we extend and adapt several advanced techniques from the literature on portfolio risk measurement to suit the insurance setting, and we illustrate their potential for application in this context based on numerical experiments.

A first practically important finding is that the allocation of the computational budget significantly affects the results. More precisely, a small number of inner simulations may yield a severe overestimation of the capital requirement due to a bias in the estimation, whereas an increased empirical variance may render the results useless if the number of outer simulations is small. A pilot simulation based on a small number of outer scenarios can be used to determine an approximately optimal allocation.

In order to analyze the reliability of the estimates, we discuss the construction of confidence intervals for the SCR. However, it turns out that these confidence intervals are very conservative so that they are very wide even if computational resources are suitably allocated. In order to increase the efficiency, aside from conventional variance reduction techniques, so-called *screening procedures* can be applied, which *screen* out scenarios that are not likely to belong to the tail of the distribution. These screening procedures — particularly when combined with conventional variance reduction techniques — are able to increase the efficiency tremendously: Our experiments show that within our example application, the length of the confidence interval may be decreased by more than a factor of ten, so that even these very conservative intervals may become practicable.

Of course, the complexity of real-world asset-liability models by far exceeds our example setup so that our quantitative results have limited practical relevance, although of course the computational resources are not on a par either. However, the qualitative insight that a proper allocation of computational resources, screening, and bias/variance reduction techniques can tremendously increase the efficiency of the simulation should also pertain to practical applications. Thus, all in all, our results provide some positive evidence that a nested simulations approach — when properly designed — may be able to provide viable estimates for the SCR with controlled error bounds.

Acknowledgements

D. Bauer gratefully acknowledges financial support from the Willis Research Network.

References

CEIOPS Internal Model Expert Group (2009) Stock-taking report on the use of Internal Models in Insurance. http://www.gcactuaries.org/documents/ceiops_stocktaking_IM_120109.pdf.

CFO Forum (2009) Market Consistent Embedded Value Principles. http://www.cfoforum.nl/downloads/MCEV_Principles_and_Guidance_October_2009.pdf.

CRO Forum (2009) Internal models benchmarking study — Summary results. http://www.croforum.org/assets/files/publications/crofintmodelbms30jan09.pdf.

Babbel, D.F., Gold, J. and Merrill, C.B. (2002) Fair Value of Liabilities: The Financial Economics Perspective. *North American Actuarial Journal* **6**, 12-27.

Bauer, D., Bergmann, D. and Kiesel, R. (2010) On the Risk-Neutral Valuation of Life Insurance Contracts with Numerical Methods in View. *ASTIN Bulletin* **40**, 31-61.

Bauer, D., Kiesel, R., Kling, A. and Russ, J. (2006) Risk-neutral valuation of participating life insurance contracts. *Insurance: Mathematics and Economics* **39**, 171-183.

Bergmann, D. (2011) *Nested Simulations in Life Insurance*. Dissertation. ifa Verlag, Ulm.

Bingham, N. and Kiesel, R. (2004) *Risk-Neutral Valuation (Second Edition)*. Springer, Berlin.

Broadie, M., Du, Y. and Moallemi, C.C. (2011) Efficient Risk Estimation via Nested Sequential Simulation. *Management Science* **57**, 1172-1194.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.

European Commission (2010) QIS5 Technical Specifications. http://ec.europa.eu/internal_market/insurance/docs/solvency/qis5/201007/technical_specifications_en.pdf.

Federal Office of Private Insurance (2006) Technical document on the Swiss Solvency Test. http://www.bvp.admin.ch/themen/00506/00552/index.html?lang=en.

Girard, L. (2002) An Approach to Fair Valuation of Insurance Liabilities using the Firm's Cost of Capital. *North American Actuarial Journal* **6**, 18-41.

Glasserman, P. (2004) *Monte Carlo Methods in Financial Engineering*. Springer, New York.

Gordy, M.B. and Juneja, D. (2010) Nested simulations in portfolio risk measurement. *Management Science* **56**, 1833-1848.

Gray, R.M. (2009) *Probability, Random Processes, and Ergodic Properties*. Springer, 2nd edition, New York.

Kling, A., Richter, A. and Russ., J. (2007) The interaction of guarantees, surplus distribution, and asset allocation in with-profit life insurance policies. *Insurance: Mathematics and Economics* **40**, 164-178.

Klumpes, P.J.M. and Morgan, K. (2008) *Solvency II versus IFRS: Cost of Capital Implications for Insurance Firms*. Proceedings of the 2008 ASTIN Colloquium.

Lan, H., Nelson, B.L. and Staum, J. (2007a) *Two-level simulation for risk management*. Proceedings of the 2007 INFORMS Simulation Society Research Workshop.

Lan, H., Nelson, B.L. and Staum, J. (2007b) *A confidence interval for tail conditional expectation via two-level simulation*. In: Henderson, S.G., Biller, B., Hsieh, M.-H., Shortle, J., Tew, J.D., Barton, R.R. (Eds.), Proceedings of the 2007 Winter Simulation Conference, IEEE Press, pp. 949-957.

Lan, H., Nelson, B.L. and Staum, J. (2010) A confidence interval procedure for expected shortfall risk measurement via two-level simulation. *Operations Research* **58**, 1481-1490.

Liebwein, P. (2006) Risk Models for Capital Adequacy: Applications in the Context of Solvency II and Beyond. *The Geneva Papers on Risk and Insurance: Issues and Practice* **31**, 528-550.

Liu, M., Nelson, B.L., Staum, J. (2010) *An Efficient Simulation Procedure for Point Estimation of Expected Shortfall*. In: Johansson, B., Jain, S., Montoya-Torres, J., Hugan, J., Yücesan, E. (Eds.), Proceedings of the 2010 Winter Simulation Conference, IEEE Press, pp. 2821-2831.

Miller, R.G. (1974) The jackknife — a review. *Biometrica* **61**, 1-15.

Pfeifer, D. and Strassburger, D. (2008) Solvency II: stability problems with the SCR aggregation formula. *Scandinavian Actuarial Journal* **2008(1)**, 61-77.

Ronkainen, V., Koskinen, L. and Bergland, R. (2007) Topical modelling issues in Solvency II. *Scandinavian Actuarial Journal* **2007(2)**, 135-146.

Sandström, A. (2007) Solvency II: Calibration for skewness. *Scandinavian Actuarial Journal* **2007(2)**, 126-134.

Society of Actuaries (SOA) (2008) Economic Capital for Life Insurance Companies. http://www.soa.org/files/pdf/research-ec-report.pdf.
ZAGLAUER, K. and BAUER, D. (2008) Risk-neutral valuation of participating life insurance contracts in a stochastic interest rate environment. *Insurance: Mathematics and Economics* **43**, 29-40.
ZHU, N. and BAUER, D. (2011) Applications of Forward Mortality Factor Models in Life Insurance Practice. *Geneva Papers on Risk and Insurance: Issues and Practice* **36**, 567-594.

APPENDIX

Proof of Proposition 3.1:

From Equations (8), (9), and (10) together with the budget constraint, we obtain the following optimization problem in K_0, K_1, and N:

$$\begin{cases} \frac{\sigma_0^2}{K_0} + \frac{\theta_0^2}{K_1^2 + f^2(\text{SCR})} + \frac{\alpha(1-\alpha)}{(N+2)+f^2(\text{SCR})} + O_N(1/N^2) + o_{K_1}(1/K_1^2) + o_{K_1}(1)\, O_N(1/N) \to \min \\ \text{s.t.} \quad c \cdot K_0 + N \cdot K_1 = \Gamma. \end{cases}$$

Let χ denote the Lagrange multiplier for the constraint. Then the first order condition for K_0 yields

$$\chi = \frac{\sigma_0^2}{K_0^2 \cdot c}.$$

Thus, from the first order condition with respect to K_1 we have

$$\frac{1}{K_0^2} = \frac{2\theta_\alpha^2 \cdot c}{\sigma_0^2 \cdot f^2(\text{SCR})} \cdot \frac{1}{N \cdot K_1^3} + o_{K_1}(1/K_1^3)\, O_N(1/N) + o_{K_1}(1/K_1)\, O_N(1/N^2),$$

and from the first order condition with respect to N we get

$$\frac{1}{(N+2)^2} = \frac{\sigma_0^2 \cdot f^2(\text{SCR})}{\alpha(1-\alpha) \cdot c} \cdot K_1 \cdot \frac{1}{K_0^2} + O_N(1/N^3) + o_{K_1}(1)\, O_N(1/N^2)$$

$$\Rightarrow \frac{1}{(N+2)^2} = \frac{2\theta_\alpha^2}{\alpha(1-\alpha)} \cdot \frac{1}{N \cdot K_1^2} + o_{K_1}(1/K_1^2)\, O_N(1/N) + O_N(1/N^3) + o_{K_1}(1)\, O_N(1/N^2).$$

Therefore,

$$\frac{N}{(N+2)^2} + O_N(1/N^2) + o_{K_1}(1)\, O_N(1/N) = \frac{2\theta_\alpha^2}{\alpha(1-\alpha)} \cdot \frac{1}{K_1^2} + o_{K_1}(1/K_1^2),$$

i.e. for N as a function of K_1

$$\frac{K_1^2}{N_{K_1} + 2} \to \frac{2\theta_\alpha^2}{\alpha(1-\alpha)}, \quad K_1 \to \infty,$$

or

$$N_{K_1} \approx \frac{\alpha(1-\alpha)}{2\theta_\alpha^2} \cdot K_1^2$$

asymptotically. This, on the other hand, implies

$$\frac{1}{K_0^2} = \frac{2\theta_\alpha^2 \cdot c}{\sigma_0^2 \cdot f^2(\mathrm{SCR})} \cdot \frac{K_1^2}{N_{K_1}} \cdot \frac{1}{K_1^5} + o_{K_1}(1/K_1^5)$$
$$\Rightarrow \frac{K_1^2\sqrt{K_1}}{K_{0,K_1}} \to \frac{2\theta_\alpha^2\sqrt{c}}{\sigma_0 \cdot f(\mathrm{SCR})\sqrt{\alpha(1-\alpha)}}, \; K_1 \to \infty,$$

for K_0 as a function of K_1, i.e.

$$K_0 \approx \frac{\sigma_0 \cdot f(\mathrm{SCR})\sqrt{\alpha(1-\alpha)}}{2\theta_\alpha^2\sqrt{c}} K_1^2\sqrt{K_1} \approx \frac{\sigma_0 \cdot K_1 \cdot f(\mathrm{SCR})}{\theta_\alpha} \sqrt{\frac{N \cdot K_1}{2c}}$$

asymptotically.

Proof of Proposition 5.1:

Let $i \in I$ which implies

$$\sum_{j \neq i} \mathbb{1}\left\{\tilde{\Delta}^{(i)}(K_{1,1}) < \tilde{\Delta}^{(j)}(K_{1,1}) - t_{f^{(i,j)},1-\varrho}\sqrt{\frac{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2 + \left(\tilde{\sigma}_1^{(j)}(K_{1,1})\right)^2}{(1+s(0,1))^2 K_{1,1}}}\right\} < N_1 - \underline{\psi} + 1.$$

Now assume that $i \notin \tilde{I}$. Then we have

$$\tilde{\Delta}^{(i)}(K_{1,1}) < \tilde{\Delta}^{(\pi_1(\underline{\psi}))}(K_{1,1}) - t_{\max,1-\varrho}\sqrt{\frac{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2 + \left(\tilde{\sigma}_{\max}(K_{1,1})\right)^2}{(1+s(0,1))^2 K_{1,1}}},$$
$$< \tilde{\Delta}^{(\pi_1(j))}(K_{1,1}) - t_{f^{(i,\pi_1(j))},1-\varrho}\sqrt{\frac{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2 + \left(\tilde{\sigma}_1^{(\pi_1(j))}(K_{1,1})\right)^2}{(1+s(0,1))^2 K_{1,1}}}$$

for $j = \underline{\psi}, \ldots, N_1$. Hence,

$$\sum_{j \neq i} \mathbb{1}\left\{\tilde{\Delta}^{(i)}(K_{1,1}) < \tilde{\Delta}^{(j)}(K_{1,1}) - t_{f^{(i,j)},1-\varrho}\sqrt{\frac{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2 + \left(\tilde{\sigma}_1^{(j)}(K_{1,1})\right)^2}{(1+s(0,1))^2 K_{1,1}}}\right\} \geq N_1 - \underline{\psi} + 1,$$

which is a contradiction.

Proof of Proposition 5.2:

Following Lan et al. (2010), we assume that $PV_0^{(k)}$ and $PV_1^{(i,k)}$ are Normally distributed. While this assumption may not be suitable for small samples, the CLT ascertains that it asymptotically holds for large samples. We denote by $\mathcal{P}(.|(Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]})$ the probability measure conditional on the event that $(Y_s^{(1)},\ldots,Y_s^{(N_1)})_{s\in[0,1]}$ are the simulated real-world scenarios in the first step.

(a) *Screening*
Let γ denote the set of the "true" $N_1 - \underline{\psi} + 1$ tail scenarios. In a first step, we show that the probability of correct screening, i.e. the probability of $\gamma \subseteq I$, is greater or equal to $1 - \alpha_{\text{screen}}$, where we follow the proof for correct screening in Lan et al. (2010).
Let

$$B_{ij} := \mathbb{1}\left\{\tilde{\Delta}^{(i)}(K_{1,1}) < \tilde{\Delta}^{(j)}(K_{1,1}) - t_{f^{(i,j)},1-\varrho}\sqrt{\frac{(\tilde{\sigma}_1^{(i)}(K_{1,1}))^2 + (\tilde{\sigma}_1^{(j)}(K_{1,1}))^2}{(1+s(0,1))^2 K_{1,1}}}\right\}.$$

Then, we have

$$\sum_{i=1}^{N_1} B_{ij} \begin{cases} < N_1 - \underline{\psi} + 1 & \text{, if } i \in I \\ \geq N_1 - \underline{\psi} + 1 & \text{, if } i \notin I. \end{cases}$$

Therefore, we obtain

$$\begin{aligned}
&\mathcal{P}(\gamma \subseteq I \,|\, (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}) \\
&\geq \mathcal{P}(\forall i \in \gamma,\ j \notin \gamma,\ B_{ij} = 0 \,|\, (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}) \\
&\geq 1 - \sum_{i\in\gamma}\sum_{j\notin\gamma}\mathcal{P}(B_{ij} = 1 \,|\, (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}) \\
&= 1 - \sum_{i\in\gamma}\sum_{j\notin\gamma}\mathcal{P}\left(\frac{\left(\tilde{\Delta}^{(j)}(K_{1,1}) - \tilde{\Delta}^{(i)}(K_{1,1})\right)(1+s(0,1))\sqrt{K_{1,1}}}{\sqrt{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2 + \left(\tilde{\sigma}_1^{(j)}(K_{1,1})\right)^2}} > t_{f^{(i,j)},1-\varrho} \,\middle|\, (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}\right) \\
&= 1 - \sum_{i\in\gamma}\sum_{j\notin\gamma}\mathcal{P}\left(\frac{\left(\widetilde{\mathrm{AC}}_1^{(i)}(K_{1,1}) - \widetilde{\mathrm{AC}}_1^{(j)}(K_{1,1})\right)\sqrt{K_{1,1}}}{\sqrt{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2 + \left(\tilde{\sigma}_1^{(j)}(K_{1,1})\right)^2}} > t_{f^{(i,j)},1-\varrho} \,\middle|\, (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}\right) \\
&\geq 1 - (N_1 - \underline{\psi} + 1)\cdot(\underline{\psi} - 1)\cdot\frac{\alpha_{\text{screen}}}{(N_1 - \underline{\psi} + 1)(\underline{\psi} - 1)} \\
&= 1 - \alpha_{\text{screen}}.
\end{aligned}$$

The first equality is a simple consequence of the t-test, where the degrees of freedom are calculated by the Welch-Satterthwaite equation.

(b) *Inner Simulation*

$$\begin{aligned}
&\mathcal{P}\Big(\{[LB,UB] \subseteq [\widehat{LB},\widehat{UB}]\} \cap \{\gamma \subseteq I\} \big| (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}\Big) \\
&= \mathcal{P}\Big(\{\gamma \subseteq I\} \big| (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}\Big) \times \\
&\quad \mathcal{P}\Big(\{[LB,UB] \subseteq [\widehat{LB},\widehat{UB}]\} \big| \{\gamma \subseteq I\},\ (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}\Big) \\
&\geq \mathcal{P}\Big(\{\gamma \subseteq I\} \big| (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}\Big) \times P\Big(\Delta^{(i)} \in C_i, \forall i \in I \big| \{\gamma \subseteq I\}, (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}\Big) \\
&\geq \mathcal{P}\Big(\{\gamma \subseteq I\} \big| (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}\Big) \times P\Big(\widetilde{\mathrm{AC}}_0 - z_{\mathrm{AC}_0} \leq \mathrm{AC}_0 \leq \widetilde{\mathrm{AC}}_0 + z_{\mathrm{AC}_0}\Big) \times \\
&\quad \prod_{i\in I} \mathcal{P}\Big(\widetilde{\mathrm{AC}}_1^{(i)} - z_{\mathrm{AC}_1}^{(i)} \cdot (1+s(0,1)) \leq \mathrm{AC}_1^{(i)} \leq \widetilde{\mathrm{AC}}_1^{(i)} + z_{\mathrm{AC}_1}^{(i)} (1+s(0,1)) \Big| (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}\Big) \\
&= (1-\alpha_{\mathrm{screen}})\big(1-\alpha_{\mathrm{AC}_0}\big) \prod_{i\in I} (1-\varepsilon) \\
&= (1-\alpha_{\mathrm{screen}})\big(1-\alpha_{\mathrm{AC}_0}\big)\big(1-\alpha_{\mathrm{AC}_1}\big).
\end{aligned}$$

Hence, we obtain

$$\begin{aligned}
&\mathcal{P}(\{[LB,UB] \subseteq [\widehat{LB},\widehat{UB}]\} \cap \{\gamma \subseteq I\}) \\
&= E\Big[\mathcal{P}\Big(\{[LB,UB] \subseteq [\widehat{LB},\widehat{UB}]\} \cap \{\gamma \subseteq I\} \big| (Y_s^{(1),\ldots,(N_1)})_{s\in[0,1]}\Big)\Big] \\
&\geq (1-\alpha_{\mathrm{screen}})\big(1-\alpha_{\mathrm{AC}_0}\big)\big(1-\alpha_{\mathrm{AC}_1}\big).
\end{aligned}$$

(c) *Total confidence level*

$$\begin{aligned}
&\mathcal{P}(\mathrm{SCR} \in [\widehat{LB},\widehat{UB}]) \\
&\geq \mathcal{P}(\{\gamma \subseteq I\} \cap \{\mathrm{SCR} \in [LB,UB]\} \cap \{[LB,UB] \subseteq [\widehat{LB},\widehat{UB}]\}) \\
&\geq 1 - \mathcal{P}(\{\mathrm{SCR} \notin [LB,UB]\}) - \mathcal{P}\Big((\{[LB,UB] \subseteq [\widehat{LB},\widehat{UB}]\} \cap \{\gamma \subseteq I\})^C\Big) \\
&= 1 - \alpha_{\mathrm{out}} - \Big(1 - (1-\alpha_{\mathrm{screen}})(1-\alpha_{\mathrm{AC}_0})\big(1-\alpha_{\mathrm{AC}_1}\big)\Big) \\
&= 1 - \alpha_{\mathrm{out}} - \alpha_{in}.
\end{aligned}$$

DANIEL BAUER
Department of Risk Management and Insurance, Georgia State University
35 Broad Street. Atlanta, GA 30303. USA
Tel.: +1-404-413-7490
Fax: +1-404-413-7499
E-Mail: dbauer@gsu.edu

ANDREAS REUSS
Institute for Finance and Actuarial Sciences (ifa)
Helmholtzstrasse 22. 89081 Ulm. Germany
Tel.: +49-(731)-50-31251
Fax: +49-(731)-50-31239
E-Mail: a.reuss@ifa-ulm.de

DANIELA SINGER
Institute for Finance and Actuarial Sciences (ifa)
Helmholtzstrasse 22. 89081 Ulm. Germany
Tel.: +49-(731)-50-31259
Fax: +49-(731)-50-31239
E-Mail: d.singer@ifa-ulm.de