On the Calculation of the Solvency Capital Requirement based on Nested Simulations*

Daniel Bauer

Daniela Bergmann†

Andreas Reuss

August 2010

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.

---

*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 second author's doctoral dissertation (cf. Bergmann (2010)). The authors are grateful for helpful comments from participants at the 2009 ARIA meeting, the 2009 CMA Workshop on Insurance Mathematics and Longevity Risk, the 2010 International Congress of Actuaries as well as from seminar participants at Georgia State University, Humboldt University of Berlin, Ulm University, and the University of Duisburg Essen. All remaining errors are ours.
D. Bauer gratefully acknowledges financial support from the Willis Research Network.

†Corresponding author.

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 on a market-consistent valuation of assets and liabilities within a structural *internal model*. However, many life insurers are struggling 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. As a consequence, many companies rely on second-best 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 mathematical framework based on the *Market Consistent Embedded Value* (MCEV) principles issued by the CFO Forum (2008). On the other hand, to provide guidance for the practical implementation, we survey 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 as e.g. measured by the length of a corresponding confidence interval for the SCR can be increased by more than a factor of ten 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 adaptation of the underlying ideas to the insurance setting and 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. We particularly illustrate the relation between these quantities and the concept of MCEV. In Section 3, we introduce the mathematical framework underlying our considera-

[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.

tions and describe the basic nested simulation 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. 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. Finally, Section 7 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, 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 their calculation, the *direct* and the *indirect* approach (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. 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. In particular, as illustrated by our numerical experiments in Section 6, the quality of the resulting estimate can differ significantly. Since the conceptual results of our paper are not affected by the choice of the method and since the indirect method generally presents the practically more accepted approach, we limit our exposition to the indirect method.

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.

[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.

To reduce the arbitrariness in the choice of this model and to ascertain comparability of results across companies, over the last decade, the insurance industry has developed principles for assessing the market-consistent value of a life insurance company's assets and liabilities from the shareholders' perspective. This so-called *Market-Consistent Embedded Value* (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. It is important to note though that the MCEV does not reflect the shareholders' default put option resulting from their limited liability. More precisely, it is assumed that the shareholders would make up any deficit arising in the future with no upper limit on the amount. Consequently, the market-consistent value of insurance liabilities can be derived *indirectly* as the difference between the market value of assets and the MCEV. In particular, the Available Capital (AC) 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.[3]

According to the CFO Forum (2008), 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):

$$\text{MCEV} \quad := \quad \text{ANAV} + \text{PVFP} - \text{CoC}. \tag{1}$$

The ANAV is derived from the (statutory) Net Asset Value (NAV)[4] 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 (2008)). 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[5] 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 (2008)). The derivation of the PVFP is quite challenging since it highly depends on the future development of the financial market, i.e. on the evolution of the yield curve, equity returns, credit spreads etc. Hence, the PVFP needs to be determined based on stochastic models, where, in general, risk-neutral valuation approaches are applied.

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 (2008)).

2.2 The Solvency Capital Requirement

For deriving the SCR, the quantity of interest is the Available Capital at $t = 1$. Assuming that the profit for the first year (denoted by X_1) has not been paid to shareholders yet, it can be described by

$$\text{AC}_1 := \text{MCEV}_1 + X_1. \tag{2}$$

[3]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).

[4]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.

[5]This means that cash flows from future new business are not included in the PVFP.

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.

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

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

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

where $s(0,1)$ is the one-year risk-free rate over $[0,1]$. The SCR is then defined as the α-quantile of L, where the security level α is set equal to 99.5%:[6]

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

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 $\mathrm{AC}_0 - \mathrm{SCR}$ and satisfies the following requirement:

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

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 $t = 0$ and may also include capital for covering losses arising from assets backing positive Excess Capital. In case the Excess Capital is negative, it is implicitly assumed that it is invested in a risk-free asset which can be illustrated by rewriting Equation (4) as follows:

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

Based on this definition of the SCR, the solvency ratio can be calculated as $\mathrm{AC}_0/\mathrm{SCR}$.

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).

[6] 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.[7] 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 profits 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 $(B_t)_{t \in [0,T]}$ (the bank account) with $B_t = \exp\{\int_0^t r_u \, du\}$, where $r_t = r(Y_t)$ is the instantaneous risk-free interest rate at time t. In this market, we take for granted the existence of a risk-neutral probability measure $\mathcal{Q}$ equivalent to $\mathcal{P}$ under which payment streams can be valued via their expected discounted values with respect to the numéraire process $(B_t)_{t \in [0,T]}$.[8]

Based on this market model, we assume that there exists a cash flow projection model of the insurance company, i.e. there exist functionals $f_1, \ldots, f_T$ that derive the future profits at time t from the development of the financial market up to time t, $t = 1, \ldots, T$. This cash flow model reflects legal and regulatory requirements as well as management rules. Hence, we model the future profits from the in-force business as a sequence of random variables $X = (X_1, \ldots, X_T)$ where $X_t = f_t(Y_s, s \in [0,t])$, $t = 1, \ldots, T$. 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. The corresponding cost-of-capital charges as well as other frictional costs could then be considered by an appropriate choice of $\mathcal{Q}$ and f_t, $t = 1, \ldots, T$.

3.2 Calculation of the SCR

According to the risk-neutral valuation formula, we can determine the PVFP at time $t = 0$, V_0, as the expectation of the sum of the discounted future profits X_t, $t = 1, \ldots, T$, under the risk-neutral measure $\mathcal{Q}$:

$$V_0 := \mathbb{E}^{\mathcal{Q}}\left[\sum_{t=1}^{T} \exp\left(-\int_0^t r_u \, du\right) X_t\right] \text{ with } \sigma_0 := \sqrt{\mathrm{Var}^{\mathcal{Q}}\left[\sum_{t=1}^{T} \exp\left(-\int_0^t r_u \, du\right) X_t\right]}.$$

In most cases, V_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 side, or, more

[7]Since insurance contracts are long-term investments, T will usually be in the range of 30-40 years or even longer.

[8]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.

precisely, the shareholders' profits X_t. Thus, in general, we have to rely on numerical methods to estimate V_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 $\mathcal{Q}$. 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.

$$\widetilde{V}_0(K_0) := \frac{1}{K_0}\sum_{k=1}^{K_0}\sum_{t=1}^{T}\exp\left(-\int_0^t r_u^{(k)}\,du\right)X_t^{(k)},$$

where $r_t^{(k)}$ denotes the instantaneous risk-free interest rate at time t in sample path k. By Equation (1) and since the ANAV can be derived from the statutory balance sheet, an estimator for AC_0 is given by $\widetilde{\mathrm{AC}}_0(K_0) = \mathrm{ANAV}_0 + \widetilde{V}_0(K_0)$. The sample version of the standard deviation is denoted by $\widetilde{\sigma}_0(K_0)$.

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 the profit of the first year, X_1, has not been paid to shareholders yet, we need to determine the $\mathcal{P}$-distribution of the $\mathcal{F}_1$-measurable random variable (cf. Equations (1) and (2))

$$\mathrm{AC}_1 := \mathrm{ANAV}_1 + \underbrace{\mathbb{E}^{\mathcal{Q}}\left[\sum_{t=2}^{T}\exp\left(-\int_1^t r_u\,du\right)X_t\,\middle|\,\mathcal{F}_1\right]}_{=:V_1} + X_1.$$

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

$$V_1^{(i)} := \mathbb{E}^{\mathcal{Q}}\left[\underbrace{\sum_{t=2}^{T}\exp\left(-\int_1^t r_u\,du\right)X_t}_{=:PV_1^{(i)}}\,\middle|\,(Y_s^{(i)})_{s\in[0,1]}\right] \text{ with } \sigma_1^{(i)} := \sqrt{\mathrm{Var}^{\mathcal{Q}}\left[PV_1^{(i)}\,\middle|\,(Y_s^{(i)})_{s\in[0,1]}\right]}. \tag{5}$$

Note that the $\sigma_1^{(i)}$ may differ significantly for different scenarios i, i.e. the discounted cash flows $\sum_{t=2}^{T}\exp\left(-\int_1^t r_u\,du\right)X_t$ are usually not identically distributed for different realizations of the state process over the first year.

In addition, realizations for the remaining components of AC_1, X_1 and ANAV_1, can easily be calculated for each of the N first-year paths. Therefore, N realizations of AC_1 are given by

$$\mathrm{AC}_1^{(i)} = \mathrm{ANAV}_1^{(i)} + V_1^{(i)} + X_1^{(i)}.$$

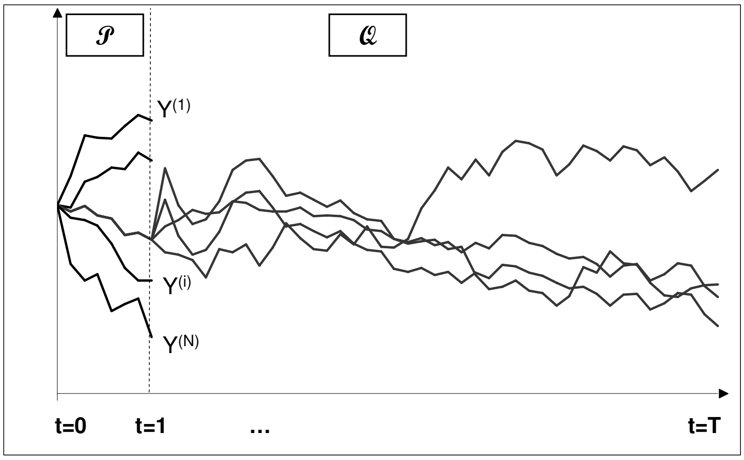

Figure 1: Illustration of the nested simulations approach

Note that these $\mathcal{F}_1$-measurable random variables $\mathrm{AC}_1^{(i)}$, $i = 1, \ldots, N$, 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 $i \in \{1, \ldots, N\}$, 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 $i \in \{1, \ldots, N\}$, by determining the resulting future profits $X_t^{(i,k)}$ ($t = 2, \ldots, T$; $k = 1, \ldots, K_1^{(i)}$) and averaging over all $K_1^{(i)}$ sample paths, we obtain Monte Carlo estimates for $V_1^{(i)}$ via

$$\widetilde{V}_1^{(i)}(K_1^{(i)}) := \frac{1}{K_1^{(i)}} \sum_{k=1}^{K_1^{(i)}} \underbrace{\sum_{t=2}^{T} \exp\left(-\int_1^t r_u^{(i,k)}\, du\right) X_t^{(i,k)}}_{=:PV_1^{(i,k)}}, \quad i \in \{1, \ldots, N\}.$$

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 $PV_1^{(i)}$:

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

Now, we can estimate N realizations of AC_1 by

$$\widetilde{\mathrm{AC}}_1^{(i)}(K_1^{(i)}) := \mathrm{ANAV}_1^{(i)} + \widetilde{V}_1^{(i)}(K_1^{(i)}) + X_1^{(i)}, \quad i = 1, \ldots, N.$$

From Equation (3), it follows that the SCR is the α-quantile of the random variable $L = \mathrm{AC}_0 - \frac{\mathrm{AC}_1}{1+s(0,1)}$. 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 statistic $\left(\tilde{S}_{(1)}, \ldots, \tilde{S}_{(N)}\right)$ and realization $\left(\tilde{s}_{(1)}, \ldots, \tilde{s}_{(N)}\right)$, 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.[9] 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

$$L = \mathrm{AC}_0 - \frac{\mathrm{AC}_1}{1 + s(0,1)} = \mathrm{AC}_0 - \frac{\mathrm{ANAV}_1 + V_1 + X_1}{1 + s(0,1)},$$

but instead we actually consider the distribution of the estimated loss

$$\tilde{L} = \widetilde{\mathrm{AC}}_0(K_0) - \frac{\mathrm{ANAV}_1 + \left(\frac{1}{K_1} \sum_{k=1}^{K_1} \sum_{t=2}^{T} \mathrm{e}^{-\int_1^t r_u^{(k)}\, du} X_t^{(k)} \middle| (Y_s)_{s \in [0,1]} \right) + X_1}{1 + s(0,1)}.$$

In particular, $\tilde{L}$ 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)

$$\tilde{L} \to L \quad \text{a.s. as } K_0, K_1 \to \infty.$$

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:[10]

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

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

[10] In what follows, probabilities and expectations are calculated under a *simulation measure*. More specifically, while the structure of the probability space is modified by the interim change of measure, our simulation procedure implies a new probability measure, which for simplicity is also denoted by $\mathcal{P}$.

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

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

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

$$Z^{K_1} = \frac{\mathrm{ANAV}_1 + \left(\frac{1}{K_1}\sum_{k=1}^{K_1}\sum_{t=2}^{T} \mathrm{e}^{-\int_1^t r_u^{(k)}\,du} X_t^{(k)} \Big| (Y_s)_{s\in[0,1]}\right) + X_1}{1+s(0,1)} - \frac{\mathrm{ANAV}_1 + V_1 + X_1}{1+s(0,1)}$$

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 L and $\tilde{Z}^{K_1} := Z^{K_1} \cdot \sqrt{K_1}$.

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

$$\mathbb{E}\left[\frac{\tilde{s}_{(m)}}{1+s(0,1)}\right] - \frac{s_\alpha}{1+s(0,1)} = \frac{\theta_\alpha}{K_1 \cdot f(\mathrm{SCR})} + o_{K_1}(1/K_1) + O_N(1/N) + o_{K_1}(1)\, O_N(1/N),$$

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

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

$$\theta_\alpha = -\frac{1}{2}\frac{\partial}{\partial u}\left[f(u)\mathbb{E}\left[\mathrm{Var}(\tilde{Z}^{K_1}|(Y_s)_{s\in[0,1]})|L=u\right]\right]\Big|_{u=\mathrm{SCR}}$$

$$= -\frac{1}{2}\int_{-\infty}^{\infty} z^2 \frac{\partial}{\partial u} g_{K_1}(u,z)\,dz\Big|_{u=\mathrm{SCR}}.$$

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, $\frac{\partial}{\partial u} g_{K_1}(u,z)\big|_{u=\mathrm{SCR}}$ 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 $\alpha = 99.5\%$.

To optimize our estimate, we would like to choose K_0, K_1, and N such that the MSE is as small as possible. Disregarding lower order terms, this yields the following optimization problem in K_0, K_1, and N:

$$\frac{\sigma_0^2}{K_0} + \frac{\theta_\alpha^2}{K_1^2 \cdot f^2(\mathrm{SCR})} + \frac{\alpha(1-\alpha)}{(N+2)f^2(\mathrm{SCR})} \to \min$$

subject to the budget restriction $K_0 + N \cdot K_1 = \Gamma$.[11] Using Lagrangian multipliers, we obtain that for any choice of Γ,

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

[11] We disregard the cost for the generation of the N sample paths in the first period, since this effort is small compared to the effort for the nested simulations. Furthermore, we do not consider the fact that the sample paths for the estimation of AC_0 are one period longer than those for the estimation of AC_1 since T is usually relatively large.

i.e. given any choice of K_1 and given θ_α, we may choose an optimal N and K_0.

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.

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 $L^{(i)}$, $1 \leq i \leq N$, 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

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

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

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

$$\Rightarrow \quad \mathcal{P}(L_{(\underline{\psi})} \leq \mathrm{SCR} < L_{(\overline{\psi})}) = \sum_{i=\underline{\psi}}^{\overline{\psi}-1} \binom{N}{i} \alpha^i (1-\alpha)^{N-i}, \quad \underline{\psi}, \overline{\psi} \in \mathbb{N}, \tag{9}$$

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

$$\mathcal{P}(L_{(\underline{\psi})} \leq \text{SCR} < L_{(\overline{\psi})}) = \sum_{i=\underline{\psi}}^{\overline{\psi}-1} \binom{N}{i} \alpha^i \, (1-\alpha)^{N-i} \geq 1 - \alpha_{\text{out}}, \tag{10}$$

and to set $LB := L_{(\underline{\psi})}$ and $UB := L_{(\overline{\psi})}$. 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

$$\lim_{K_1^{(i)}\to\infty} \mathcal{P}\left([LB;UB] \subseteq \left[\widehat{LB};\widehat{UB}\right]\right) \geq 1-\alpha_{\text{in}} \Rightarrow \lim_{K_1^{(i)}\to\infty} \mathcal{P}\left(\text{SCR} \in \left[\widehat{LB};\widehat{UB}\right]\right) \geq 1-\alpha_{\text{out}}-\alpha_{\text{in}}. \tag{11}$$

Hence, $\left[\widehat{LB};\widehat{UB}\right]$ is a 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

$$z_{\text{AC}_0}(K_0) := t_{K_0-1,1-\frac{\alpha_{\text{AC}_0}}{2}} \frac{\widetilde{\sigma}_0(K_0)}{\sqrt{K_0}} \quad \text{and} \quad z_{\text{AC}_1}^{(i)}(K_1^{(i)}, N) := t_{K_1^{(i)}-1,1-\frac{\epsilon}{2}} \frac{\widetilde{\sigma}_1^{(i)}(K_1^{(i)})}{(1+s(0,1)) \cdot \sqrt{K_1^{(i)}}},$$

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

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

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

$$\begin{aligned}
& \mathcal{P}\left(\left(L^{(1)}, \ldots, L^{(N)}\right) \in C\right) \\
\geq \; & \mathcal{P}\left(\widetilde{\text{AC}}_0 - z_{\text{AC}_0}(K_0) \leq \text{AC}_0 \leq \widetilde{\text{AC}}_0 + z_{\text{AC}_0}(K_0)\right) \\
& \prod_{i=1}^{N} \mathcal{P}\left(\widetilde{\text{AC}}_1^{(i)} - z_{\text{AC}_1}^{(i)}(K_1^{(i)}, N) \cdot (1+s(0,1)) \leq \text{AC}_1^{(i)} \leq \widetilde{\text{AC}}_1^{(i)} + z_{\text{AC}_1}^{(i)}(K_1^{(i)}, N) \cdot (1+s(0,1))\right) \\
= \; & (1-\alpha_{\text{AC}_0}) \cdot \prod_{i=1}^{N} (1-\epsilon) = 1 - \underbrace{(\alpha_{\text{AC}_0} + \alpha_{\text{AC}_1} - \alpha_{\text{AC}_0} \cdot \alpha_{\text{AC}_1})}_{=:\alpha_{\text{in}}},
\end{aligned} \tag{12}$$

i.e. C is a confidence region for $\left(L^{(1)},\ldots,L^{(N)}\right)$ 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, simply set

$$\widehat{LB} := \inf_{M\in C} M_{(\underline{\psi})} \quad \text{and} \quad \widehat{UB} := \sup_{M\in C} M_{(\overline{\psi})}, \tag{13}$$

where $M := (M^{(1)},\ldots,M^{(N)})$ is an element in the confidence region C and $M_{(.)}$ is the order statistic of $M^{(1)},\ldots,M^{(N)}$. The following Proposition summarizes the foregoing:

Proposition 4.1. *$\widehat{LB}$ is the $\underline{\psi}^{th}$ order statistic of $\tilde{L}^{(i)}(K_1^{(i)})-z_{AC_0}(K_0)-z_{AC_1}^{(i)}(K_1^{(i)},N)$, $\widehat{UB}$ is the $\overline{\psi}^{th}$ order statistic of $\tilde{L}^{(i)}(K_1^{(i)})+z_{AC_0}(K_0)+z_{AC_1}^{(i)}(K_1^{(i)},N)$, $1\leq i\leq N$, and the confidence interval $\left[\widehat{LB};\widehat{UB}\right]$ for the SCR has an asymptotic confidence level of $(1-\alpha_{out}-\alpha_{in})$.*

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 confidence level. More specifically, on the one hand, the outer confidence level $\mathcal{P}(L_{(\underline{\psi})}\leq \text{SCR} < L_{(\overline{\psi})})$ may be strictly greater than $(1-\alpha_{\text{out}})$ due to the discreteness of the binomial distribution. On the other hand, the inequalities in (11) and (12) 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 $\Gamma = K_0 + K_1\cdot N$. 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 $\widehat{LB} = \tilde{L}^{(i_{LB})}(K_1) - z_{\text{AC}_0}(K_0) - z_{\text{AC}_1}^{(i_{LB})}(K_1,N)$, and let i_{UB} be the index such that $\widehat{UB} = \tilde{L}^{(i_{UB})}(K_1) + z_{\text{AC}_0}(K_0) + z_{\text{AC}_1}^{(i_{UB})}(K_1,N)$. Then the length of the confidence interval is given by

$$\widehat{UB}-\widehat{LB} \quad = \quad \tilde{L}^{(i_{UB})}(K_1) - \tilde{L}^{(i_{LB})}(K_1) + 2\cdot z_{\text{AC}_0}(K_0) + z_{\text{AC}_1}^{(i_{UB})}(K_1,N) + z_{\text{AC}_1}^{(i_{LB})}(K_1,N).$$

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

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

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

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

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 $\Gamma = K_0 + K_1 \cdot N$. While it cannot be solved in closed form, from the first order condition with respect to K_1, we obtain

$$K_1 = \frac{\Gamma}{N + \zeta_1^{\frac{2}{3}} \cdot \zeta_2^{-\frac{2}{3}} \cdot N^{\frac{2}{3}}}, \tag{14}$$

where

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

Hence, for fixed Γ and N the optimal K_1 is given by (14) and since $K_0 = \Gamma - N \cdot K_1$ 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 $\tilde{L}^{(i)}(K_{1,1}) = \widetilde{\mathrm{AC}}_0(K_0) - \frac{\widetilde{\mathrm{AC}}_1^{(i)}(K_{1,1})}{1+s(0,1)}$. 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

$$I \;:=\; \left\{ i : \sum_{j\neq i} \mathbb{1}_{\left\{ \tilde{L}^{(i)}(K_{1,1}) < \tilde{L}^{(j)}(K_{1,1}) - t_{f^{(i,j)},1-\delta} \sqrt{\frac{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2 + \left(\tilde{\sigma}_1^{(j)}(K_{1,1})\right)^2}{(1+s(0,1))^2 K_{1,1}}} \right\}} < N_1 - \underline{\psi} + 1 \right\} \tag{15}$$

where $\delta := \frac{\alpha_{\mathrm{screen}}}{(N_1-\underline{\psi}+1)(\underline{\psi}-1)}$, $\underline{\psi}$ is defined by Equation (10), and $t_{f^{(i,j)},1-\delta}$ is the $(1-\delta)$-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 δ 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.[12] The number of "surviving" scenarios is denoted by $N_2 = |I|$.

In order to limit the number of necessary comparisons, we further use a pre-screening procedure before we start the screening process.[13] Specifically, let $\pi_1(\cdot)$ be a permutation of the indices such that $\tilde{L}^{(\pi_1(i))}$ is non-decreasing in i and define

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

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

Then we pre-screen out all scenarios with

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

[12] 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.

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

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 for this proposition can be found in the Appendix.

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

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

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 $\tilde{L}^{(i)}(K_{1,2}^{(i)})$ and $\tilde{\sigma}_1^{(i)}(K_{1,2}^{(i)})$, 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.

$$K_{1,2}^{(i)} := \left\lfloor \frac{(\Gamma - N_1 \cdot K_{1,1} - K_0)\left(\widetilde{\sigma}_1^{(i)}(K_{1,1})\right)^2}{\sum_{j\in I}\left(\widetilde{\sigma}_1^{(j)}(K_{1,1})\right)^2} \right\rfloor. \tag{16}$$

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

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

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

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 $\tilde{L}^{(i)}(K_{1,2}^{(i)}) - z_{\mathrm{AC}_0}(K_0) - z_{\mathrm{AC}_1}^{(i)}(K_{1,2}^{(i)}, N_2)$ and the $(\overline{\psi} - (N_1 - N_2))^{th}$ order statistic of $\tilde{L}^{(i)}(K_{1,2}^{(i)}) + z_{\mathrm{AC}_0}(K_0) + z_{\mathrm{AC}_1}^{(i)}(K_{1,2}^{(i)}, N_2)$, $i \in I$, 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 $(1 - \alpha_{out} - \alpha_{in})$ as $K_0 \to \infty$, $K_{1,1} \to \infty$, and $K_{1,2}^{(i)} \to \infty$, where*

$$\alpha_{in} \quad := \quad 1 - (1 - \alpha_{screen})(1 - \alpha_{AC_0})(1 - \alpha_{AC_1}). \tag{17}$$

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

In addition to the confidence interval, we may also compute a point estimate $\widetilde{\text{SCR}}^{\text{screen}}$ for the SCR, which is given by the $(m-(N_1-N_2))^{\text{th}}$ order statistic of $\tilde{L}^{(i)}(K_{1,2}^{(i)})$, $i \in I$, 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 $\tilde{L}^{(i)}(K_{1,2}^{(i)})$ for those real-world scenarios that were screened out, they would have been smaller than the $(m-(N_1-N_2))^{\text{th}}$ order statistic of $\tilde{L}^{(i)}(K_{1,2}^{(i)})$, $i \in I$. Under this assumption, the $(m-(N_1-N_2))^{\text{th}}$ order statistic of $\tilde{L}^{(i)}(K_{1,2}^{(i)})$, $i \in I$, coincides with the m^{th} order statistic of $\tilde{L}^{(i)}(K_{1,2}^{(i)})$, $1 \leq i \leq N_1$, 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 considerably 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 are smaller than those in the surviving scenarios is problematic because we may have screening mistakes. More specifically, it is possible that we have screened out a scenario where $\tilde{L}^{(i)}(K_{1,2}^{(i)})$ is greater than the $(m-(N_1-N_2))^{\text{th}}$ order statistic of $\tilde{L}^{(i)}(K_{1,2}^{(i)})$, $i \in I$. 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 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. (2008) to our problem. Here, the authors use multiple stages of screening to estimate the expected shortfall. However, they note that the "procedure does not provide confidence intervals nor guarantees a minimum probability of correctly identifying the tail."

5.2 Efficient Use of Screening Procedures

Obviously, for a fixed computational budget, the efficiency of the screening procedure described in the previous subsection strongly 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 Γ. 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 follows (cf. Equation (17)):

$$\alpha_{\mathrm{AC}_1} = 1 - \frac{(1-\alpha_{\mathrm{in}})}{(1-\alpha_{\mathrm{screen}})\cdot(1-\alpha_{\mathrm{AC}_0})}. \tag{18}$$

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

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

Our target 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.[14] 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 (18). The resulting confidence interval is denoted by $[\widehat{LB}_{\mathrm{pilot}}; \widehat{UB}_{\mathrm{pilot}}]$ with corresponding indices $i_{LB,\mathrm{pilot}}$ and $i_{UB,\mathrm{pilot}}$, respectively. Subsequently, we apply the screening procedure to the results from the first run of the pilot simulation.

Similar to Lan et al. (2007b), 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 survives screening in the pilot simulation.

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

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

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

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

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

where $\epsilon := 1 - (1 - \alpha_{\mathrm{AC}_1})^{\frac{1}{\hat{N}_2}}$ with $\hat{N}_2 := \frac{\tilde{N}_2 \cdot N_1}{\tilde{N}_1}$ being the estimated number of survivors. $\hat{K}_{1,2} := \frac{(\Gamma - N_1 \cdot K_{1,1} - K_0) \cdot \tilde{N}_1}{\tilde{N}_2 \cdot N_1}$ is the estimated number of inner simulations in the second run.[15] Then, the task is to minimize this length which may be carried out numerically.

If we allocate the remaining budget for the second run proportional 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.

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

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:

$$\hat{K}_{1,2}^{(i_{LB})} := (\Gamma - N_1 \cdot K_{1,1} - K_0) \cdot \frac{\tilde{N}_1 \cdot \left(\widetilde{\sigma}_1^{(i_{LB},\text{pilot})}(K_{1,1})\right)^2}{N_1 \cdot \sum\limits_{i \in I_{\text{pilot}}} \left(\widetilde{\sigma}_1^{(i)}(K_{1,1})\right)^2}.$$

$\hat{K}_{1,2}^{(i_{UB})}$ 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

Assets	**Liabilities**
A_t	L_t
	R_t
A_t	A_t

Table 1: Simplified balance sheet

[15] In case $\hat{K}_{1,2}$ is not an integer, we use the largest integer that is smaller than $\hat{K}_{1,2}$.

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.

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. In particular, the shareholders' funds are available to cover potential losses, i.e. they are exposed to risk. Thus, as compensation for the adopted risk, we assume that dividends d_t may be paid out to shareholders each period. 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 $\mathrm{ROI}_T := R_T - \exp\left(\int_0^T r_u\, du\right) R_0$ as their (time value-adjusted) return on investment at the end of the projection period (also referred to as "maturity") T.

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.[16] 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" in Kling et al. (2007)). However, as was pointed out in Section 2.1, the MCEV should not reflect the shareholders' put option, i.e. the MCEV should be calculated under the supposition that shareholders cover any deficit. In accordance with this hypothesis, we assume that the company obtains an additional contribution c_t from its shareholders in case of such a shortfall.

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

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

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

$$\begin{aligned} d_t &= (1-\delta) y \left(A_t^- - A_{t-1}^+\right) \mathbb{1}_{\left\{\delta y \left(A_t^- - A_{t-1}^+\right) > g L_{t-1}\right\}} \\ &\quad + \left[y \left(A_t^- - A_{t-1}^+\right) - g L_{t-1}\right] \mathbb{1}_{\left\{\delta y \left(A_t^- - A_{t-1}^+\right) \leq g L_{t-1} \leq y \left(A_t^- - A_{t-1}^+\right)\right\}}. \end{aligned}$$

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

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 ignore unrealized gains and losses on assets as well as other adjustments, we have $\text{ANAV}_0 = \text{NAV}_0 = R_0$. Therefore, the Available Capital at time $t = 0$ can be described as follows:

$$\begin{aligned}
\text{AC}_0 &= \text{ANAV}_0 + V_0 \\
&= R_0 + \mathbb{E}^{\mathcal{Q}}\left[\sum_{t=1}^{T} \exp\left(-\int_0^t r_u\,du\right)(d_t - c_t) + \exp\left(-\int_0^T r_u\,du\right)\text{ROI}_T\right] \\
&= R_0 + \mathbb{E}^{\mathcal{Q}}\left[\sum_{t=1}^{T} \exp\left(-\int_0^t r_u\,du\right)(d_t - c_t) + \exp\left(-\int_0^T r_u\,du\right)R_T - R_0\right] \\
&= \mathbb{E}^{\mathcal{Q}}\left[\sum_{t=1}^{T} \exp\left(-\int_0^t r_u\,du\right)X_t\right]
\end{aligned}$$

where

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

Similarly, we obtain

$$\text{AC}_1 = \text{ANAV}_1 + V_1 + X_1 = \mathbb{E}^{\mathcal{Q}}\left[\sum_{t=1}^{T} \exp\left(-\int_1^t r_u\,du\right)X_t \middle| \mathcal{F}_1\right].$$

So far, we described the Available Capital based on cash flows from the shareholders' point of view. But as already mentioned in Section 2.1, we can also express AC_0 and AC_1 based on cash flows from the policyholders' perspective, i.e. within this framework we have

$$\text{AC}_0 = A_0 - \mathbb{E}^{\mathcal{Q}}\left[\exp\left(-\int_0^T r_u\,du\right)L_T\right].$$

and

$$\text{AC}_1 = A_1^+ - \mathbb{E}^{\mathcal{Q}}\left[\exp\left(-\int_1^T r_u\,du\right)L_T \middle| \mathcal{F}_1\right] + X_1.$$

The corresponding estimators for the Available Capital based on policyholder cash flows can be derived in analogy to the estimator for the shareholder cash flows (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 policyholder cash flows. As we will see in Section 6.2, the quality of the two different estimation approaches differs considerably.

6.1.3 Asset Model

For the evolution of the financial market, similarly to Zaglauer and Bauer (2009), 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

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

respectively, where $\rho \in [-1,1]$ describes their correlation, $\mu \in \mathbb{R}$, $\sigma_A, \kappa, \xi, \sigma_r > 0$, and W and Z are two independent Brownian motions under the real-world measure $\mathcal{P}$. 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

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

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 $\mathcal{Q}$:

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

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

$$\begin{aligned} A_t^- &= A_{t-1}^+ \exp\left(\int_{t-1}^t r_s ds - \frac{\sigma_A^2}{2} + \rho\, \sigma_A\, (\tilde{W}_t - \tilde{W}_{t-1}) + \sqrt{1-\rho^2}\sigma_A(\tilde{Z}_t - \tilde{Z}_{t-1})\right), \\ r_t &= e^{-\kappa} r_0 + \tilde{\xi}\left(1 - e^{-\kappa}\right) + \int_{t-1}^t \sigma_r e^{-\kappa(t-s)}\, d\tilde{W}_s, \end{aligned}$$

and

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

which can be conveniently used in Monte Carlo algorithms (cf. Zaglauer and Bauer (2009)).

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 portfolio are calibrated to an index consisting of 80% REXP[17] and 20% DAX.[18] 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

[17] The REXP is a total return index of the German bond market.

[18] The DAX is a total return index of the German stock market.

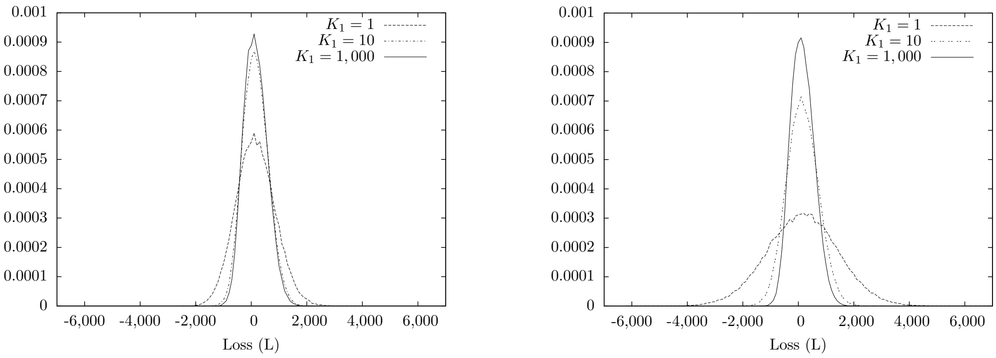

Figure 2: Empirical density function for different choices of K_1 for the estimator based on the policyholder cash flows (left) and the shareholder cash flows (right); $N = 100,000$, $K_0 = 250,000$

$\sigma_r = 0.6\%$. 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 introduce 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. This bias mainly depends on the choice of the estimator and the number of inner simulations. Hence, in order to develop an idea for its magnitude, we analyze the results for the estimator based on cash flows from the policyholders' and from the shareholders' perspective (see Section 6.1.2), and choose 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 functions of the loss L for both estimators and different choices of K_1 are plotted. As expected, for both estimators 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 can also be noticed in Table 2, where the estimated SCR for different choices of K_1 is displayed. Moreover, we observe that the distribution given by the estimator based on shareholder cash flows is more dispersed than that given by the estimator for the policyholder cash flows for the same K_1.

Since the bias mainly depends on the variance of $\widetilde{\mathrm{AC}}_1^{(i)}(K_1^{(i)})$, $1 \leq i \leq N$, this indicates that the former estimator has a higher variance and, thus, we need more inner simulations to obtain reliable results. Further analyses show that in our setting, the estimator based on cash flows from the policyholders' perspective is superior to that based on shareholder cash flows in most cases. Therefore, if not mentioned otherwise, we will rely on the estimator based on cash flows from the policyholders' perspective for the remainder of the paper.

K_1	policyholder cash flows		shareholder cash flows	
	$\widetilde{\mathrm{SCR}}$	$\widetilde{\mathrm{AC}}_0/\widetilde{\mathrm{SCR}}$	$\widetilde{\mathrm{SCR}}$	$\widetilde{\mathrm{AC}}_0/\widetilde{\mathrm{SCR}}$
1	1,994.0	94%	3,432.5	55%
5	1,404.7	134%	1,874.6	100%
10	1,332.7	141%	1,606.5	117%
100	1,261.2	149%	1,279.1	147%
1,000	1,246.3	151%	1,254.6	149%

Table 2: Estimated SCR and estimated solvency ratio for different choices of K_1; $K_0 = 250,000$, $N = 100,000$

The above results show that a proper allocation of 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. Based on these scenarios, we calculate the empirical variances $\left(\tilde{\sigma}_1^{(i)}(\tilde{K}_1)\right)^2$ for each real-world scenario i, $i = 1, \ldots, \tilde{N}$ and estimate the expected conditional variance via a regression analysis. More specifically, we assume

$$\mathbb{E}^{\mathbb{Q}}\left[\mathrm{Var}\left(\tilde{Z}^{K_1}|Y_1, D_1\right)|L\right] \approx \beta_0 + \beta_1 L + \beta_2 L^2$$

and estimate β_0, β_1, and β_2 from our results. Sensitivity analyses show that the optimal choice of K_0, K_1, and N is rather insensitive to different choices of the regression function. 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. The resulting estimate for θ_α is given by $\hat{\theta}_\alpha \approx 0.027$. 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 $\widetilde{\mathrm{SCR}} = 1,249.7$ and a solvency ratio of 150%. At first sight, it might be surprising that K_0 should be chosen that large compared to the two other parameters. However, reducing the variance of $\widetilde{\mathrm{AC}}_0(K_0)$ 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 reasonable to allocate a rather large budget to 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 $\frac{\bar{\theta}_\alpha}{K_1 \cdot \bar{f}(\widetilde{\mathrm{SCR}})}$, where $\bar{\theta}_\alpha$ and $\bar{f}$ denote the average of the estimates resulting from the 150 estimation procedures as explained above. The MSE is then estimated by the sum of the empirical variance and the squared estimated bias. This allows us to correct the mean by the estimated bias. Figure 3 and Table 3 show our results.

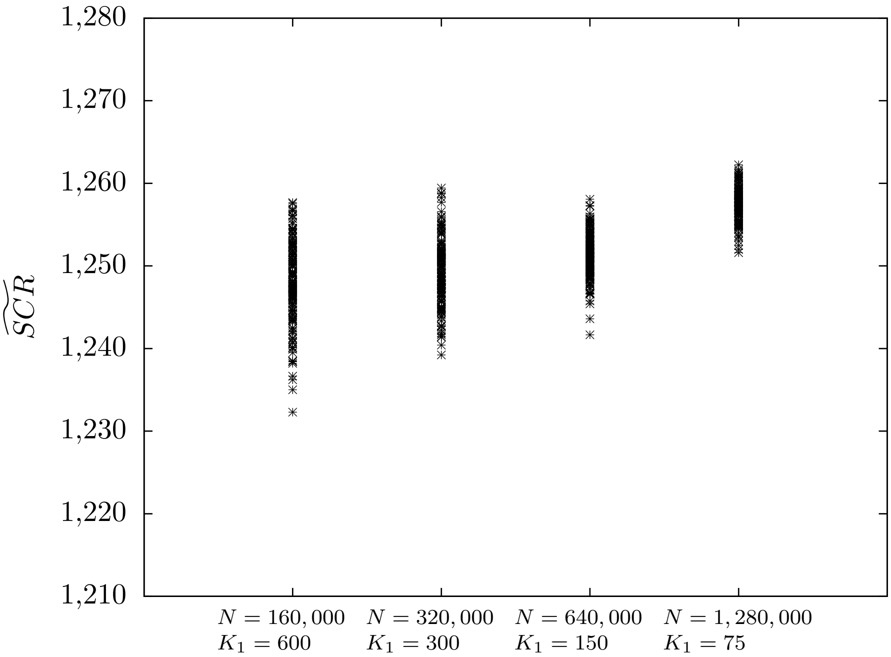

Figure 3: 150 runs for different choices of N and K_1; $K_0 = 1,500,000$

N	K_1	Mean ($\widetilde{\mathrm{SCR}}$)	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
1,280,000	75	1,257.4	4.2	11.4	133.1	1,246.1

Table 3: Results for different choices of N and K_1; 150 runs, $K_0 = 1,500,000$

As expected, the mean of the estimated SCRs increases as K_1 decreases due to the increased bias. In contrast to this, 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 remarkable 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.[19]

6.2.2 Confidence Intervals

Having analyzed the point estimator of the nested simulations approach, 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 (10)) 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}.[20] However, based on some sensitivity analyses we find 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 $\widetilde{\text{AC}}_1^{(i)}(K_1)$, $i = 1, \ldots, N$, $\alpha_{\text{in}} = 8\%$ 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 $\underline{\psi}^{\text{th}}$ and $\overline{\psi}^{\text{th}}$ order statistics of the estimates losses are given by $\tilde{L}_{\underline{\psi}} = 1,241.6$ and $\tilde{L}_{\overline{\psi}} = 1,259.4$. Thus, a very large part of the confidence interval can be attributed to the uncertainty arising from the inner simulation and the estimation of AC_0. Therefore, it may be conductive to reconsider the choice of K_0, K_1, and N for fixed α_{out} and α_{AC_0} and using the optimization approach presented in Section 4.2.

For the sake of simplicity, we use the same pilot simulation as in Section 6.2.1[21] leading to 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 from the previous section, 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 ration 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, we also compute the length of the confidence interval for other numbers of real-world scenarios N, where for each

[19] 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.

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

[21] However, we found that already pilot simulations with about $N = 10,000$ yields suitable estimates.

N we calculate the approximately optimal choice of K_0 and K_1. In Figure 4, we show 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 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. Furthermore, we need to keep in mind that the results in Figure 4 are based on a limited number of N and for every choice of N, only one run was performed. 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. This demonstrates again that the parameters need to be chosen carefully in order to obtain accurate results.

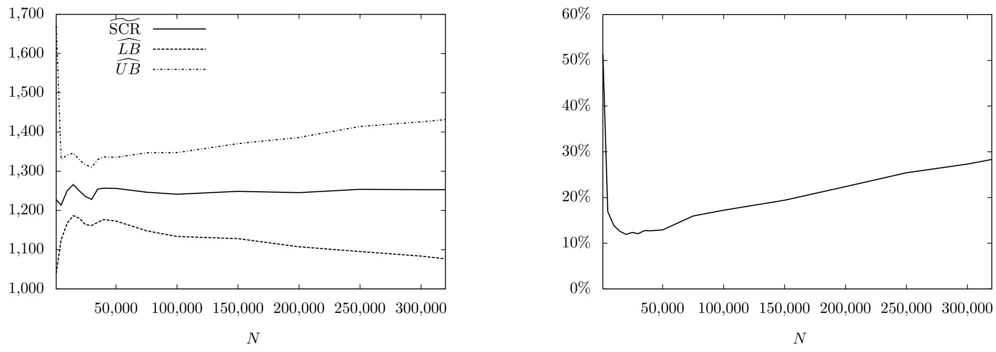

Figure 4: 90% Confidence intervals for different N (left), length of confidence interval as percentage of $\widetilde{\text{SCR}}$ (right); $\Gamma = 97,500,000$

6.2.3 Screening Procedures

In the previous subsection, we used a rather large computational budget of $\Gamma = 97,500,000$ to determine a confidence interval with length corresponding to 12% of $\widetilde{\text{SCR}}$. 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 shorter confidence interval with the same computational budget or to derive a confidence interval of the same length based on a lower computational budget.

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 from Section 6.2.1, i.e. we fix $N_1 = 320,000$, $K_0 = 1,500,000$, and we use a total budget of $\Gamma = 97,500,000$. As before, we choose $\alpha_{\text{in}} = 8\%$ and $\alpha_{\text{AC}_0} = 0.1\%$. Furthermore, we set $\alpha_{\text{screen}} = 4\%$.[22] The remaining budget in the second run is allocated equally to all surviving scenarios. To obtain a 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 $\widetilde{\text{SCR}}^{\text{screen}} = 1,247.8$.[23]

[22]Again, 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 $\alpha_{\text{screen}} = 4\%$ such that the error due to screening is similar

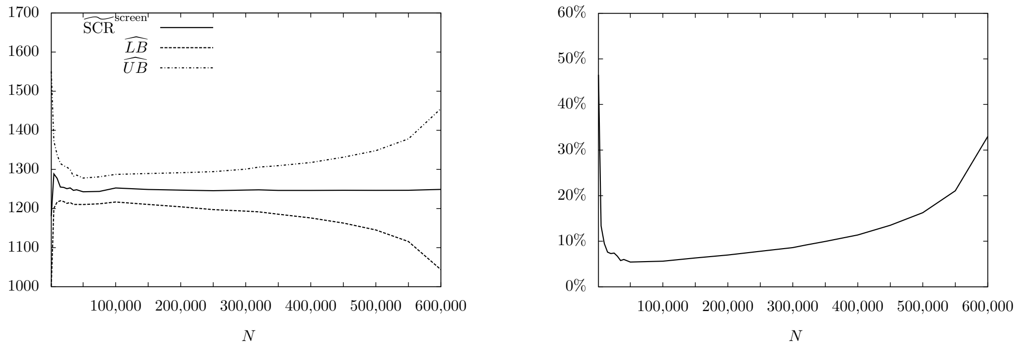

Figure 5: 90%-confidence intervals with screening for different N_1 (left), length of confidence interval as percentage of $\widetilde{\text{SCR}}$ (right); $K_0 = 1,500,000$, $K_{1,1} = 150$, $\Gamma = 97,500,000$

Next, we optimize N_1 for given $K_0 = 1,500,000$ and $K_{1,1} = 150$ to further reduce the length of the confidence interval. Based 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 following confidence interval: [1,212.2;1,281.0]. Thus, we have a length of 68.8 which corresponds to 6% of $\widetilde{\text{SCR}}^{\text{screen}} = 1,243.8$. To show that this choice is 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%.

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

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 sample contract, we found that there is hardly any difference between the two methods.

to the error arising from the estimation of $\text{AC}_1^{(i)}$.

[23]Here, $\widetilde{\text{SCR}}^{\text{screen}}$ is the point estimate resulting from the screening procedure as described at the end of Section 5.

policyholder cash flows					
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%

shareholder cash flows					
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	1,275.3	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 4: Comparison of estimated SCR and estimated solvency ratio with and without antithetic variates

6.2.4 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 is an array of different alternatives available. We refer to Glasserman (2004) for more details and to Bergmann (2010) 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. Table 4 shows the analog of Table 2 when relying on antithetic variates, where 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. We notice two effects: On the one hand, the use of antithetic variates clearly improves the estimate significantly indicating considerable gains in the efficiency of the estimation; on the other hand, it now seems that the estimator based on shareholder cash flows is superior. This is in contrast to the analysis without variance reduction where the estimator based on policyholder cash flows generally performs better.

These findings are also illustrated by Table 5, 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 for the estimator based on policyholder (PH) cash flows by about 70%. The effect for the estimator based on shareholder (SH) cash flows is even more pronounced. Here, the MSE is reduced by almost 90%. Consequently, with antithetic variates, only a budget of 15,760,000 (PH) and 2,352,000 (SH) is necessary in order to obtain results of a similar accuracy 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

	N	K_1, K_1^{AV}	K_0	Mean ($\widetilde{\text{SCR}}$, $\widetilde{\text{SCR}}^{\text{AV}}$)	Emp. Var	Est. Bias	Est. MSE	Corrected Mean
PH with AV	1,070,000	45	600,000	1,247.7	4.6	1.7	7.4	1,246.0
PH with AV	310,000	25	130,000	1,248.8	13.8	3.0	23.1	1,245.7
PH w/o AV	320,000	300	1,500,000	1,249.3	15.8	2.9	24.0	1,246.4
SH with AV	1,375,000	35	625,000	1,247.4	5.2	1.5	7.4	1,246.0
SH with AV	115,000	10	26,000	1,250.9	46.0	5.3	73.7	1,245.7
SH w/o AV	105,000	920	900,000	1,250.4	48.7	4.4	68.4	1,246.0

Table 5: Comparison of nested simulations approach with and without antithetic variates for different parameters

$\Gamma = 97,500,000$ as in Section 6.2.3, our pilot simulation suggests that for $K_0^{\text{AV}} = 750,000$ and $K_{1,1}^{\text{AV}} = 75$, $N_1 = 200,000$ is approximately optimal for the estimator based on policyholder cash flows. The resulting confidence interval is given by $[1,222.7; 1,257.0]$ which corresponds to about 3% of $\widetilde{\text{SCR}}^{\text{screen}}$.

Considering our "first" confidence interval from Section 6.2.2 with a length of approximately 28% of $\widetilde{\text{SCR}}$, this efficiency gain by relying on more advanced techniques is remarkable. However, it is necessary to note that although our example company model is quite simple and we rely on 97,500,000 scenarios, the length of the confidence interval is *still* about 3% of the SCR corresponding to approximately 30 bps of the balance sheet total.

7 Conclusion

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

A first 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.

Clearly, the practical usefulness of the nested simulations estimator depends on its accuracy, which may be assessed via the length of a confidence interval. However, it turns out that these intervals 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 the length of the confidence interval may be decreased by more than a factor of ten. However, within our application, although the example company is quite simple and although we rely an a relatively large number of scenario, the confidence interval may still be too wide to be practicable.

It is arguable that this last result can be interpreted as an indication of a general impracticability of the nested simulations approach in the present context. While there are many parallels between estimating the risk of a portfolio of financial derivatives and determining the capital requirement for an insurance portfolio, there is at least one important difference: In a portfolio of financial derivatives, the single instruments can be valuated independently and hence, the pricing errors diversify away when the portfolio is large (see Gordy and Juneja (2010)). This is generally not the case for an insurance portfolio. Due to management rules being applied at company level (e.g. strategic asset allocation and profit participation), the cash flows of different insurance contracts usually depend on each other. Therefore, we need to simulate the whole portfolio simultaneously based on the same stochastic scenarios and valuation errors in the inner simulation will not diversify away when the portfolio is large.

Nevertheless, we are convinced that in the long run, advanced numerical approaches as presented here should allow for a computationally feasible and sufficiently accurate assessment of an insurer's solvency position.

References

CFO Forum, 2008. Market Consisten Embedded Value Principles. `http://www.cfoforum.nl/pdf/mcev_principles_and_guidance.pdf`.

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

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

Bergmann, D. (2010) *Nested Simulations in Life Insurance*. Dissertation, Ulm University.

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

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

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`.

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

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.

Gordy, M.B., Juneja, D. (2010) Nested simulations in portfolio risk measurement. *Management Science*, forthcoming.

Kling, A., Richter, A., 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., 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., Staum, J. (2007a) *Two-level simulation for risk management*. Proceedings of the 2007 INFORMS Simulation Society Research Workshop.

Lan, H., Nelson, B.L., 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.), Proceeedings of the 2007 Winter Simulation Conference, IEEE Press, pp. 949-957.

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

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. (2008) An Efficient Simulation Prodecure for Point Estimation of Expected Shortfall. Working Paper, Northwestern University, Evanston, Illinois.

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

Ronkainen, V., Koskinen, L., 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.

Zaglauer, K., Bauer, D. (2009) Risk-neutral valuation of participating life insurance contracts in a stochastic interest rate environment. *Insurance: Mathematics and Economics* **43**, 29-40.

Appendix

Proof of Proposition 5.1:
Let $i \in I$ which implies

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

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

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

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

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

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 approximately holds for large samples. Since we are looking to prove an asymptotic result, we adopt it without much loss of generality. We denote by $\mathcal{P}(.|(Y_1, D_1)^{(1),\ldots,(N_1)})$ the probability measure conditional on the event that $(Y_1^{(1)}, D_1^{(1)}), \ldots, (Y_1^{(N_1)}, D_1^{(N_1)})$ 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

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

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

$$
\begin{aligned}
& \mathcal{P}(\gamma \subseteq I |\, (Y_1, D_1)^{(1),\ldots,(N_1)}) \\
\geq\ & \mathcal{P}(\forall i \in \gamma, j \notin \gamma, B_{ij} = 0 |\, (Y_1, D_1)^{(1),\ldots,(N_1)}) \\
\geq\ & 1 - \sum_{i \in \gamma} \sum_{j \notin \gamma} \mathcal{P}(B_{ij} = 1 |\, (Y_1, D_1)^{(1),\ldots,(N_1)}) \\
=\ & 1 - \sum_{i \in \gamma} \sum_{j \notin \gamma} \mathcal{P}\left(\frac{\left(\tilde{L}^{(j)}(K_{1,1}) - \tilde{L}^{(i)}(K_{1,1})\right)(1 + s(0,1))\sqrt{K_{1,1}}}{\sqrt{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2 + \left(\tilde{\sigma}_1^{(j)}(K_{1,1})\right)^2}} > t_{f^{(i,j)},1-\delta} \,\middle|\, (Y_1, D_1)^{(1),\ldots,(N_1)} \right) \\
=\ & 1 - \sum_{i \in \gamma} \sum_{j \notin \gamma} \mathcal{P}\left(\frac{\left(\widetilde{\mathrm{AC}}_1^{(i)}(K_{1,1}) - \widetilde{\mathrm{AC}}_1^{(j)}(K_{1,1})\right)\sqrt{K_{1,1}}}{\sqrt{\left(\tilde{\sigma}_1^{(i)}(K_{1,1})\right)^2 + \left(\tilde{\sigma}_1^{(j)}(K_{1,1})\right)^2}} > t_{f^{(i,j)},1-\delta} \,\middle|\, (Y_1, D_1)^{(1),\ldots,(N_1)} \right) \\
\geq\ & 1 - \left(N_1 - \underline{\psi} + 1\right) \cdot \left(\underline{\psi} - 1\right) \cdot \frac{\alpha_{\text{screen}}}{(N_1 - \underline{\psi} + 1)(\underline{\psi} - 1)} \\
=\ & 1 - \alpha_{\text{screen}}.
\end{aligned}
$$

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

(b) *Inner Simulation*

$$
\begin{aligned}
& \mathcal{P}\left(\left\{[LB, UB] \subseteq [\widehat{LB}, \widehat{UB}]\right\} \cap \{\gamma \subseteq I\} \,\middle|\, (Y_1, D_1)^{(1),\ldots,(N_1)}\right) \\
=\ & \mathcal{P}\left(\{\gamma \subseteq I\} |\, (Y_1, D_1)^{(1),\ldots,(N_1)}\right) \times \\
& \mathcal{P}\left(\left\{[LB, UB] \subseteq [\widehat{LB}, \widehat{UB}]\right\} \,\middle|\, \{\gamma \subseteq I\}, (Y_1, D_1)^{(1),\ldots,(N_1)}\right) \\
\geq\ & \mathcal{P}\left(\{\gamma \subseteq I\} |\, (Y_1, D_1)^{(1),\ldots,(N_1)}\right) \times P\left(L^{(i)} \in C_i, \forall i \in I \,\middle|\, \{\gamma \subseteq I\}, (Y_1, D_1)^{(1),\ldots,(N_1)}\right) \\
\geq\ & \mathcal{P}\left(\{\gamma \subseteq I\} |\, (Y_1, D_1)^{(1),\ldots,(N_1)}\right) \times P\left(\widetilde{\mathrm{AC}}_0 - z_{\mathrm{AC}_0} \leq \mathrm{AC}_0 \leq \widetilde{\mathrm{AC}}_0 + z_{\mathrm{AC}_0}\right) \times \\
& \prod_{i \in I} \mathcal{P}\left(\widetilde{\mathrm{AC}}_1^{(i)} - z_{\mathrm{AC}_1}^{(i)} \cdot (1 + s(0,1)) \leq \mathrm{AC}_1^{(i)} \leq \widetilde{\mathrm{AC}}_1^{(i)} + z_{\mathrm{AC}_1}^{(i)}(1 + s(0,1)) \,\middle|\, (Y_1, D_1)^{(1)\ldots(N_1)}\right) \\
=\ & (1 - \alpha_{\text{screen}})\,(1 - \alpha_{\mathrm{AC}_0}) \prod_{i \in I} (1 - \epsilon) \\
=\ & (1 - \alpha_{\text{screen}})\,(1 - \alpha_{\mathrm{AC}_0})\,(1 - \alpha_{\mathrm{AC}_1}).
\end{aligned}
$$

Hence, we obtain

$$
\begin{aligned}
& \mathcal{P}\left(\left\{[LB, UB] \subseteq [\widehat{LB}, \widehat{UB}]\right\} \cap \{\gamma \subseteq I\}\right) \\
=\ & E\left[\mathcal{P}\left(\left\{[LB, UB] \subseteq [\widehat{LB}, \widehat{UB}]\right\} \cap \{\gamma \subseteq I\} \,\middle|\, (Y_1, D_1)^{(1)\ldots(N_1)}\right)\right] \\
\geq\ & (1 - \alpha_{\text{screen}})\,(1 - \alpha_{\mathrm{AC}_0})\,(1 - \alpha_{\mathrm{AC}_1}).
\end{aligned}
$$

(c) *Total confidence level*

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

Daniel Bauer
Department of Risk Management and Insurance. Georgia State University
35 Broad Street. Atlanta, GA 30303. USA
phone: +1-404-413-7490, fax: +1-404-413-7499
dbauer@gsu.edu

Daniela Bergmann
Institute of Insurance. Ulm University
Helmholtzstrasse 18. 89069 Ulm. Germany
Phone: +49-(731)-50-31259. Fax: +49-(731)-50-31239
daniela.bergmann@uni-ulm.de

Andreas Reuss
Institute for Finance and Actuarial Sciences (ifa)
Helmholtzstrasse 22. 89081 Ulm. Germany
Phone: +49-(731)-50-31251. Fax: +49-(731)-50-31239
a.reuss@ifa-ulm.de