



A first look back: model performance under Solvency II

Ralf Korn^{1,2} · Gerhard Stahl³

Received: 21 February 2023 / Revised: 5 August 2023 / Accepted: 17 November 2023 /
Published online: 19 December 2023
© The Author(s) 2023

Abstract

We consider an empirical backtesting for the Solvency Capital Required (SCR) under Solvency II. Based on empirical facts that the Basic own Funds (BoF) can be assumed to evolve log-normally and have a much lower volatility than the corresponding equity for our test data, we make a proposal based on Earnings at Risk (EaR) that can be used to reduce the biases from overshooting SCR estimates in a prudential way.

Keywords Solvency II · Backtesting · Capital requirement

1 Introduction

With the current Solvency II review actually in process, it is a good time to have a sharp look at the practical performance of the internal models and the standard formula. We do this in the form of back testing with available data of 28 quarterly values of own funds and the SCR.

In doing so, we restrict ourselves to the presentation of our analysis for a large German insurer, but have performed additional analyses for further companies including one that uses the standard formula. The results are similar and totally in line with our conclusions. What we, however, like to point out is that such an analysis has to be done for every company separately, as aggregation via a pooling of the values of different companies leads to a mixture of distributions. In particular in the case of log-normal distributions such a mixture would typically be no longer log-normal.

✉ Ralf Korn
korn@mathematik.uni-kl.de
Gerhard Stahl
gerhard.stahl@HDI.de

¹ Fachbereich Mathematik, RPTU Kaiserslautern–Landau, 67653 Kaiserslautern, Germany

² Fraunhofer Institute ITWM, Fraunhofer Platz 1, 67663 Kaiserslautern, Germany

³ University Ulm and HDI, HDI Platz 1, 30659 Hannover, Germany

To be able to understand the reasons for the surprising facts and figures collected in the third part of this contribution, the specific nature of the collected data - their *silent features* - are explained in detail in the next section.

Exemplified analyses in the facts and figures section reveal the good model quality of internal models of insurance companies, the usefulness of normal distribution approximations and the much higher stability of the time series of the basic own funds than that of the corresponding share prices of the companies.

2 Task, methodology and the data set

Solvency II improved regulatory approaches for insurers significantly by requiring to model a whole balance sheet. The own funds are defined as the saldo of the SII balance sheet. This valuation approach treats both assets and liabilities in a market-consistent manner. Both options under Solvency II, internal models and the standard formula, define regulatory capital requirements yielding a capital cushion for unfavourable changes of the own funds over a one year time horizon. More precisely, this cushion should exceed the 99.5%-value of the loss function of the Basic Own Funds (BoF_t resp. ΔBoF_t) over one year, the so-called Solvency Capital Required (SCR for short).

The period from 2016 (Solvency II came into force) to the third quarter of 2022 yields 28 data vectors of computed BoF- and SCR-values. Our figures of particular interest are the quarterly increases of the BoF, i.e. $\Delta BoF_{t+1/4} = BoF_{t+1/4} - BoF_t$ (with $t = i/4$, $i = 1, \dots, 27$) and in particular their quotient by the suitably scaled SCR-value at $t + 1/4$

$$\frac{\Delta BoF_{t+1/4}}{\frac{SCR_{t+1/4}}{z_\alpha}} \quad (1)$$

where z_α is the 99.5%-quantile of the standard normal distribution. Our first task in the next section will be a backtesting via a normal QQ-plot. By this, we compare the empirical distribution function of the quotient with the quantiles of the standard normal distribution. If we observe a linear relation, this will support the assumption of having normally distributed data.

Methodological aspects and earnings-at-risk. As an alternative to the bottom-up calculation of the SCR by means of a model we consider a top-down approach based on the earnings-at-risk (EaR) concept used in financial markets (see e.g. Matten [5] and CorporateMetrics, [3]). Consider a portfolio X of financial instruments with a time series of market values, $V_t(X)$, at time t . From the log-returns

$$R_{V,t} = \ln V_t - \ln V_{t-1} \quad (2)$$

one can estimate the portfolio volatility σ_X in the usual manner. This allows for a simple form of the EaR at time t given as (see e.g. [5])

$$EaR_t = q_\alpha \cdot \sigma_X \cdot V_t \quad (3)$$

for some suitable quantile q_α . Using the log-returns of BoF_t

$$R_{BoF,t} = \ln BoF_t - \ln BoF_{t-1}, \tag{4}$$

we can calculate the corresponding EaR in a similar way as in Eq. (3). This value may then be used as a performance yardstick of the used risk model, either an internal model or a standard formula.

Remark 1 (Silent features of the data set) We will highlight some silent features of the data set and shed some light on the assumptions of the overall framework.

- (a) *Data quality*: All SCR figures were approved by both, regulatory authorities and internal validation units from the insurance undertaking as well. The solvency balance sheet has to be approved by external auditors, with the consequence that *BoF* figures are of high quality too.
- (b) *Distinctive features of BoF and SCR*: *BoF* is a feedback variable which reacts to changes of the market and the economic environment in general. Thus, appropriate actions, including corrective ones, may be triggered and influence future *BoF* figures. The *SCR*, however, acts as a feedforward variable, i.e. it serves to absorb disturbances before they affect the system. For this, model changes need to be approved by regulators before implementation. Hence, compared with *BoF*, management actions are mirrored in the *SCR* with a greater time lag. The *SCR* serves merely as a strategic steering variable e.g. in the process of capital allocation.
- (c) *Stakeholders of the data set*: The various stakeholders – in our case such as shareholders represented by the share price, board members, bondholders represented by rating agencies, policy holders represented by regulatory and supervisory agencies (among others) and the government also represented by regulators controlling systemic risks – have in general different interests. Hence, using unbiased figures calculating the *SCR* is the only strategy to satisfy all stakeholders simultaneously.

3 Facts and figures

The following analysis is realized in the spirit of Tukey’s Exploratory Data Analysis, see [8] and is in line with similar analysis given e.g. in Jaschke et al. [2] and Stahl et al. [7].

We will in particular use:

exploratory tools such as QQ-plots to see if our distributional assumptions can be justified and to judge, whether the risk is overestimated or underestimated,

a deviance decomposition of the mean squared error, more precisely, if I is an ideal forecast and F is the realization of our prediction method, then the deviance is a re-scaled version of $\mathbb{E}((F - I)^2)$ and is decomposed into bias, scale difference and imprecision as follows (see Van Belle [9] and note that $\sigma_I = 1$):

$$\frac{\mathbb{E}(F - I)^2}{\sigma_F} = \frac{(\mu_F - 0)^2}{2\sigma_F} + \frac{(\sigma_F - 1)^2}{2\sigma_F} + (1 - \rho), \tag{5}$$

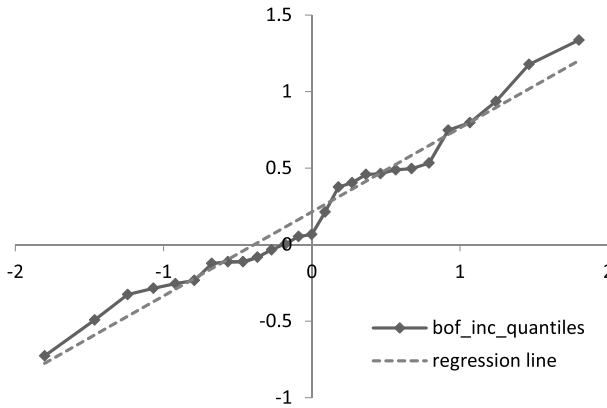


Fig. 1 Normal QQ-plot of the scaled BoF changes of I^1 including the regression line

multiple correlations for judging the possible linear relationship between the forecast distribution and $N(0, 1)$,

confidence intervals for volatility estimates and mean estimates.

We next analyze the data of a large German insurer that we denote by I^1 (for more examples – including a company using the standard formula – we refer to [4]). In Fig. 1, we present the QQ-plot (against $N(0,1)$) of the quartely BoF increments suitably scaled by the SCR as given in Eq. (1). The regression line for those values and the very high multiple correlation value (R^2) of 0.986 between the quantiles of the scaled quarterly BoF increments and the standard normal quantiles indicate that the assumption of a linear relation is highly plausible. I.e. we can model the scaled BoF increments as being normally distributed.

The prediction quality of this normal model is underlined by the low value of the imprecision of 0.14 in the deviance decomposition in

$$\frac{\mathbb{E}(F - I)^2}{\sigma_F} = 0.425 = 0.045 + 0.24 + 0.141, \tag{6}$$

which in particular implies a ρ -value of $\rho = 0.859$.

As the value of the SCR is typically much larger than that of the quarterly change in the own funds, one can in first order keep it constant in the above considered quotient and hope for the log-BoF increments to be normally distributed. This heuristic argumentation is strongly underlined by the corresponding QQ-plot in Fig. 2 where we have an R^2 of 0.985.

The volatilities σ_{BoF} and σ_{share} of the BoF and of the share price of I^1 are estimated as

$$\hat{\sigma}_{BoF, I^1} = 0.075, \quad \hat{\sigma}_{share, I^1} = 0.234. \tag{7}$$

This is exactly what we expected due to our arguments in the preceding sections. The BoF which is based on deep firm-specific, internal information and analysis should

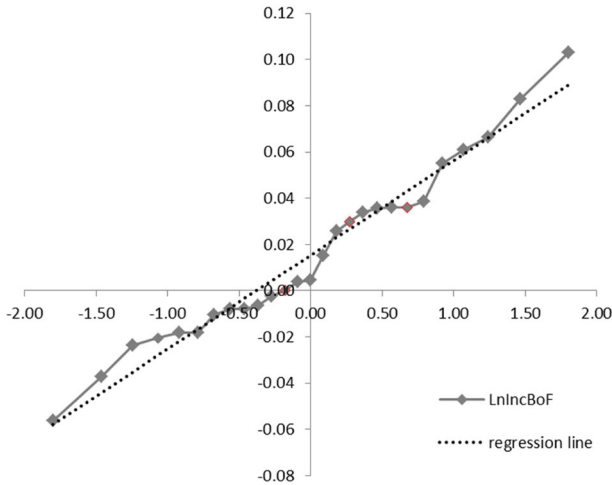


Fig. 2 Normal QQ-plot of BoF log-increases of I^1 including the regression line

mainly equal the idiosyncratic risk, while the share price also contains the systemic risk and might also be subject to market inconsistencies and thus should have a higher volatility. Note, however, that they differ by a factor of three!

Using the log-normality of the BoF increments makes the following approach a tempting one:

- (1) Estimate the annualized mean and volatility of the log BoF as (assuming independence of the log-increments) by a 95%-confidence interval to obtain

$$\mu_{lb} = 0.0622 \pm 0.0285, \quad \sigma_{lb} = 0.0754 \pm 0.0279.$$

- (2) Use the log-normal model representation of the BoF increments given by

$$\Delta(BoF)_t = BoF_{t+1} - BoF_t = BoF_t \left(e^{\mu_{lb} + \sigma_{lb} Y} - 1 \right) \tag{8}$$

for a standard normally distributed random variable Y .

- (3) Plug in the estimated values of the mean and volatility above and replace Y in Eq. (8) by its 0.005-quantile $q_{0.005} = -2.58$. This yields an annual BoF-loss of 12.42% which in the case of $BoF_{2021Q3} = 26.158$ yields $-\Delta(BoF)_{2021Q3}(0.005) = 3.242$. Compare this to the $SCR_{2021Q3} = 10.095$ that I^1 held at the relevant time, the third quarter of 2021.

Note that to obtain the SCR_{2021Q3} -value with the help of Eq. (8), the 0.005-quantile has to be replaced by approx. -7.225 which is much smaller than -4.09 , the 0.00002-quantile, which represents the probability of a 1 in 50000-years event. This dramatic increase of the mean time to the next return of a loss event that is of the level of the SCR shows the prudence inherited of the internal model.

If instead of the mean and volatility of the BoF, we are using the mean and volatility of the share, $\mu_{I^1} = 0.036$, $\sigma_{I^1} = 0.2344$ in the model formula for the annual BoF-increment above, we obtain an annual BoF-increment on the 0.005-quantile level of 44.3%. This yields the corresponding $-\Delta(BoF)_{2021Q3}(0.005) = 11.347$ which indeed is pretty close to the $SCR_{2021Q3} = 10.095$ of I^1 at the third quarter of 2022. It is a coincidence that using the SCR calculated from the annual increment of the BoF with the – here non-relevant – volatility of the share price is so close to the calculated SCR by I^1 while the much more relevant volatility of the BoF leads to a significantly smaller SCR.

Remark 2 (Only one example?) While we presented only one example above, we performed more analyses for other insurers, among them also companies that are using the standard formula for the SCR. Indeed, our results have been similar to what we reported for I^1 . A deep case study for a huge range of companies is beyond the scope of our work.

4 Conclusions

Given the limited amount of data (in both terms of length of our time series and of the number of companies we examined), we are fully aware of the fact that we cannot make claims that have a global validity. However, we want to raise awareness for some aspects and questions that should be discussed in detail, and we encourage individuals and institutions to examine them further, on individual and on large scale:

Forecasting bias. The low volatility of carefully managed BoF are an indicator that companies know how to steer them with a minimum risk. As a consequence, our data analysis shows a significant forecasting bias of internal models (and the standard formula) that leads to an overestimation of risk by a factor of three. The possible capital inefficiencies caused by this overestimation risks require careful considerations, both from a micro and macro economic perspective as well. Given the huge systemic risk we are currently exposed to, the bias might be considered as welcomed. However, the level is worth to be discussed. Like in medicine, it is a question of the dose.

Model adequacy. The adequacy of the model, i.e. the relation between the model and the reality, may be questioned by the observed forecasting. The observed high correlations imply that affine transformations can be applied in order to achieve a nearly perfect prediction quality measured by, say, a deviance.

Modeling uncertainty, modeling process, including validation. An important lesson from our analysis relates to the perception of model uncertainty. Typically, model uncertainty is seen as a downside risk. This cannot be claimed for the data at hand. The bias in the model output (i.e. in SCR),

- which is probably *unintended* in terms of size (in particular, the return periods are way above 200 years and very prudent estimates show return periods beyond 2000 years),

- and *uncontrolled* w.r.t. the sources and hence diffuse w.r.t. preferences of stakeholders (opening the door for unintended arbitrages)

is in general not fully understood w.r.t. its sources and reasons. As far as the sources of overestimation by the SCR are not fully understood, the validation of the adequacy of an internal model or the standard formula is incomplete.

This might have far-reaching implications for the determination of the overall capital needs within the ORSA process. Furthermore, the approach considered in Stahl [6], which requires additional capital charge for model uncertainties in the light of the OCC paper [1], should be reconsidered and re-evaluated. This might also be interesting for the level of economic capital and regulatory capital required by rating agencies and regulators.

Regulatory issues. In the light of the precision of approaches based on EaR, the cost-benefit relation of the current regulatory regime has to be reconsidered. This is especially true for the standard formula. Furthermore, the analysis sheds some light on the importance of the regulatory invention of the SII balance sheet and its saldo, the *BoF* compared to the *SCR*. Regarding the dynamics but also w.r.t. the content and feedback, the *BoF* seems to be the winner of the day. Assumed that a decade ago, the *BoF* were already in place for a number of years, the shape and calibration of the SII framework would be another one. With respect to the on-going SII review process the augmentation of additional capital requirements, e.g. for climate risks or other systemic risk is not indicated.

Possible changes in paradigms. The comparison of the standard formula or internal models with reduced form approaches based on ΔBoF_t (such as the log-normal-based EaR approach) shed some light on silent assumptions of Solvency II. At first sight, the precision of the reduced form model suggests to build the standard formula on such an approach. Of course, this was impossible a decade ago. Secondly, the overarching requirement of market consistency is questioned, given the amount of overestimation, both w.r.t. the standard formula and internal models as well. Thirdly, the inefficient capital requirement might question the use test for internal models.

Reporting. Under Solvency II a series of external and internal reporting requirements have to be fulfilled. In the light of our insights, the provided backtesting analysis might be prominently incorporated. This is true for the SFCR, ORSA and validation reports. With regard to this, the bias which is to be expected should influence the judgements of the model output significantly.

Capital adequacy. A prudent approach in order to overcome unintended overestimation biases could combine risk figures based on volatilities of ΔBoF_t and on the conventional calculation of SCR_t . We give an adhoc suggestion of the form:

$$\max \left\{ \omega_1 SCR_{EaR} + \omega_2 SCR, \beta \times SCR \right\}, \quad (9)$$

with ω_i non-negative weights, i.e. $\omega_1 + \omega_2 = 1$, and $SCRE_{aR}$ denotes a multiple of $\sigma_{BoF}BoF$ given as

$$SCRE_{aR} = \alpha \times \sigma_{BoF} \times BoF. \quad (10)$$

Here, $\beta \in [0.5, 1]$ caps the influence of the correction and we suggest reasonable figures to be

$$\omega_i = 0.5, \alpha = 3, \beta = 0.75, \quad (11)$$

where an in-depth reasoning for (9) and (11) is beyond the scope of this contribution.

Overall judgement. In terms of economic efficiency, the framework of Solvency II is far from optimal. However, the insights gained from its risk modeling, e.g. determination of sensitivities etc. are highly relevant for ERM and as we have seen, the risk ranking is not deteriorated by the actual implementation. Hence, the application of Solvency II is overall far from useless from an ERM perspective. However, from a cost-benefit perspective, our results raise questions.

Perhaps our analysis motivates EIOPA to perform a benchmarking study w.r.t. forecast quality of Solvency II models (internal models and applicants of the standard formula as well) in analogy to the one which EIOPA undertakes for economic scenario generators. Such an analysis might also give further insights w.r.t. the adequacy of regulatory capital, the level playing field, the level of regulatory comfort and insights to systemic risks across Europe.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Comptroller of the Currency (2011) Supervisory guidance on model risk management. OCC Bulletin 2011-12. <https://www.occ.treas.gov/news-issuances/bulletins/2011/bulletin-2011-12a.pdf>
2. Jaschke S, Stahl G, Stehle R (2007) Value-at-risk forecasts under scrutiny—the German experience. *Quant Finance* 7(6):621–636
3. Kim J, Malz AM, Mina J (1999) Longrun: technical document. RiskMetrics Group, New York
4. Korn R, Gerhard S (2023) Insights from backtesting under Solvency II – overdose or implicit change of paradigm? Technical report, discussion paper
5. Matten C (1996) Managing bank capital: capital allocation and performance measurement. Wiley, Chichester
6. Stahl G (2016) Model uncertainty in a holistic perspective. In: Kallsen J, Papapantoleon A (eds) *Advanced modelling in mathematical finance*. Springer, Berlin, pp 189–215

7. Stahl G, Wehn CS, Zapp A (2005) Backtesting within the trading book. *J Risk* 8(2):1
8. Tukey JW (1962) The future of data analysis. *Ann Math Stat* 33(1):1–67
9. Van Belle G (2011) *Statistical rules of thumb*, vol 699. Wiley, Chichester

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.