EBA STAFF PAPER SERIES

N. 14 – JULY/2021

A UNIVERSAL STRESS SCENARIO APPROACH FOR CAPITALISING NON-MODELLABLE RISK FACTORS UNDER THE FRTB

by Martin Aichele, Marco Giovanni Crotti, and Benedikt Rehle

EBA EUROPEAN BANKING AUTHORITY

ABSTRACT

EU legislators mandated the European Banking Authority to propose a stress scenario methodology for capitalising non-modellable risk factors (NMRF) as foreseen under the Basel Fundamental Review of the Trading Book (FRTB) rules for market risk. In this paper, we present the foundations of such a methodology.

By design, it is universally applicable to all kinds of risk factors to which a bank may be exposed, and it caters for a wide range of data availability by adjusting the stress scenario for the number of returns observed in the calibration period. It captures non-linearities in the portfolio loss profile against changes in the NMRF, while reducing the computational effort and being simple.

To motivate the values set for some parameters in the methodology, we use a set of skewed generalised 't' (SGT) distributions as a generic tool for describing a wide universe of real historical returns from all asset classes.

Finally, we extend the methodology from single risk factors to segments of curves or surfaces as envisaged in the FRTB.

KEYWORDS

Market risk; FRTB; NMRF; capital requirements for non-modellable risk factors; sampling error for the expected shortfall; SGT distributions.

JEL CLASSIFICATION:

C13; C46; G21; G28; G32

1 Introduction

In January 2019, the Basel Commitee on Banking Supervision (BCBS) published the final revised global standard for banks' minimum capital requirements for market risk, also known as the Fundamental Review of the Trading Book (FRTB) [Bas19]. Under those standards, banks have the choice to use a so-called standardised approach based on sensitivities and prescribed "risk weights" or, subject to supervisory approval, to devise their own internal risk measurement model for computing capital requirements.

A key-feature of the FRTB is the classification of risk-factors that a bank has identified for their risk measurement as 'modellable' or 'non-modellable'. The risk-factor eligibility test (RFET) determines whether a risk factor is modellable or not. To be modellable, the number of 'real-price observations' observed over the last year needs to be at least 100 and be free of long streaks for no observations. Real price observations must be a real transaction or committed quote and be representative for the risk factor, i.e. the bank must be able to extract the value of the risk factor from the value of the real-price observations to use the observation for the RFET of that risk factor.

Modellable risk factors are capitalised in an integrated expected shortfall measure allowing for diversification and hedging among all modellable risk factors, with some restrictions for diversification across risk classes (general interest rates, credit spreads, equities, FX, and commodities). The ES measure is calibrated to a period of financial stress, i.e. a period that maximises the capital requirements for the bank's portfolio.

Non–modellable risk factors (NMRF) are capitalised, outside and incremental to the expected shortfall (ES) measure, under the 'stress scenario risk measure' (SSRM) which the FRTB does not specify in detail. Paragraph MAR 33.16 of the FRTB only states that it should be calibrated to be at least as prudent as the expected shortfall used for modelled risks (i.e. a loss calibrated to a 97.5% confidence level over a period of stress). It should be determined for each risk factor or, subject to supervisory approval, for each risk factor bucket (i.e. segments of a risk curve or surface). This implies a computational effort scaling with the number of risk factors or risk factor buckets. The risk factor (bucket) correlations are implicitly prescribed by an aggregation formula for the total capital requirement for non-modellable risk factors.

In May 2019, the European Parliament and Council adopted amendments to the Capital Requirements Regulation [PC13, PC19] implementing the FRTB standards in the European Union. In that context, the legislators mandated the European Banking Authority (EBA) to develop a methodology for determining the SSRM for capitalising NMRF. According to EU regulation an 'extreme scenario of future shock' must be determined for each non-modellable risk factor (bucket) that leads to a loss, which is the SSRM of that NMRF. In this paper, we present the design goals and mathematical foundations of the methodology that we designed to deliver on that mandate. Probably due to the FRTB novelties, and except for [MBP17], we are not aware of any prior work putting forward a methodology for capitalising NMRFs.

Banks could capitalise a non–modellable risk factor by direct historical simulation of losses per risk factor (bucket), which we call the 'direct method': first, the returns observed for the NMRF in a period of financial stress are obtained, then those returns are applied as shocks to the current value of the NMRF so as to obtain a sample of losses. The stress scenario risk measure would be the expected shortfall of the losses at 97.5% confidence level. However, the direct method requires a significant computational effort: for each risk factor (bucket) the losses of the portfolio positions susceptible to this risk factor would need to be calculated for each return in the stress period. This method is thus computationally very demanding for a trading book having potentially thousands of NMRFs and is not suited for risk factors with sparse data. Hence, there is a need for a much more efficient approach, which also has to be extendible to a low number of returns down to no observed returns at all in the stress period.

A recurrent underlying theme of this work is to consider the situation of less than daily returns for non-modellable risk factors, because NMRFs are typically less liquid and even more so in a stress period. However, it should be noted that there are in practice daily data for some non-modellable risk factors (i.e. daily risk factor returns can be computed), while the requirement to find real-price observations, so as to pass the RFET, is not met.

The paper is organised as follows[1]. In Section 2, we outline the methodology by first introducing its design goals and assumptions. We describe the concept of nearest to 10-(business-) day returns rescaled to 10-(business-) days catering for sparse data in the stress period, we put forward the 'asymmetrical sigma method', a method to obtain a robust approximation of the expected shortfall from a small sample of data, and we develop a computationally efficient approximation of the expected shortfall of the losses due to a risk factor by the loss corresponding to the expected shortfall of the risk factor distribution. In Section 3, we discuss corner cases under which the methodology may not lead to appropriate results and present ad-hoc solutions to address them. In Section 4, we show how to compensate for the small sample uncertainty in the estimation of the risk measures to a controlled target confidence level. In Section 5, we first give analytical expressions for Value-at-Risk (VaR) VaR and ES of the skewed-generalised t (SGT) distributions. We then use the SGT distributions as a tool for motivating the values taken by the parameters of our methodology, and in particular for investigating the small sample uncertainty of the ES estimates by means of a quantile approximation and Monte Carlo simulations. In Section 6, we show how the methodology is naturally extended to a bucket of risk factors in a curve or surface. In Section 7, we recall the final aggregated capital requirements for non-modellable risks and in Section 8, we conclude.

The building blocks of the methodology that are laid down in a detailed manner in this paper, namely, (i) a parsimonious way to obtain a time series of 10 days returns for risk factors with non-daily data; (ii) the asymmetrical sigma method ES estimator to robustly estimate ES measures on the basis of a volatility measure for small samples; (iii) the efficient approximation of the ES of losses with the loss at the ES of the risk factors' distribution by means of a non-linearity correction coefficient; (iv) the introduction of a sampling error compensation factor; and (v) the use, for calibration and analyses purposes, of a SGT distribution family broadly matching a large set of historically observed risk factors, are our main contributions. Linking all pieces together to obtain a universally applicable and efficient methodology for the SSRM that works also from a regulatory perspective can be considered our most relevant contribution. Finally, we believe that the analysis and results obtained for SGT distributions may also be useful in contexts other than the one discussed in this paper.

2 The methodology and the goals that drove its design

In this section we first introduce the goals that drove the design of the methodology and a few assumptions for it to be applied. We then present the methodology designed to overcome the drawbacks of a direct calculation of the expected shortfall of the losses while meeting those goals.

2.1 Goals and assumptions

The methodology is designed to meet the following goals:

- **G1**: Be applicable to any kind of risk factor. Typically banks' portfolios are susceptible to a vast range of risk factors that heavily differ in their nature: equity risk factors, interest rate risk factors, parameters used for modelling curves and surfaces. Thus, one of the goals of the methodology is its universal applicability.
- **G2**: Capture a wide range of different cases with respect to the number of observations available for a given risk factor. There may be risk factors for which banks have daily data, while for some others, the data availability is very limited. Thus, the methodology is designed to work with any data availability.
- **G3**: Ensure an adequate level of capitalisation for the non-modellable risk factors. In order to be in line with the FRTB standards, the methodology should lead to a stress scenario risk measure at least as prudent as the expected shortfall calibration used for modellable risk factors.

[1]In this version dated February 28, 2022, the derivation of the uncertainty correction factor uses a more generic approach, and the temporal autocorrelation structure of risk factor returns in our methodology as outlined in Annex B is considered. Additionally, it includes some corrections, while the main results and conclusions remain unchanged.

- **G4**: Be efficient and simple. The methodology should be efficient and use as few loss evaluations as possible without compromising other goals. It should be simple and remove complexity where possible.
- **G5**: Be applicable both at risk-factor level and at bucket level. The FRTB specifies that a stress scenario has to be generated for each NMRF separately. However, the FRTB also provides for the possibility to assess the modellability of risk factors belonging to a curve or a surface via either the own bucketing approach or the regulatory bucketing approach. Where a bank opts to use regulatory buckets, it can generate a stress scenario for all risk factors in the bucket, and determine the capitalisation at bucket level. Thus, the methodology must be designed to be applicable at bucket level as well.
- **G6**: Capture losses accurately. The methodology must accurately capture the characteristics of the loss functions with respect to movements in the risk factors.
- **G7**: Less data, more capital. The methodology should be built to obtain higher capital requirements when less data are available so as to reflect the higher uncertainty that is present when obtaining figures with less data.

When developing the methodology we assumed that:

- **A1**: When applying the methodology, banks have identified the portfolio that is capitalised in accordance with the internal models approach - we refer to it as the 'SSRM portfolio'. Specifically, in accordance with the FRTB, only positions in trading desks meeting the backtesting and P&L attribution requirements can be capitalised by means of the internal models approach.
- **A2**: When applying the methodology, banks have a fixed set of risk factors that passed the risk factor eligibility test and have been classified as modellable and a fixed set of risk factors that did not pass that test and have been classified as non–modellable.
- **A3**: For each of the five risk classes identified in the FRTB (e.g. equity risk), the bank identified a 12-month period of financial stress S – we refer to it as 'stress period' – for which a sufficient amount of data is available for calibrating shocks applicable to the NMRFs in the risk class. Paragraph 33.16 of the FRTB sets out that a common 12-month period of stress across all NMRFs in the same risk class must be used for obtaining the stress scenario risk measures.
- **A4**: For each non–modellable risk factor j, banks can identify the loss that their SSRM portfolio would suffer following a change in the value of the NMRF. Equivalently, banks are able for each shock x to determine the corresponding portfolio's loss $l_j(x)$,

$$l_j(x) = \text{Loss}_{D^*}(r_j^* \oplus x)\,, \tag{1}$$

where D^* denotes the date for which the stress scenario risk measure is calculated - we refer to it as 'reference date', r_j^* the value of the NMRF j at the reference date, and $\oplus$ denotes the application of the shock x to the risk factor value r_j^* in accordance with the risk factor return modelling used (e.g. absolute returns, logarithmic returns, etc.)[2]. The loss function $\text{Loss}_{D^*}(r_j)$ is defined as follows:

$$\text{Loss}_{D^*}(r_j) = PV(r_j^*) - PV(r_j)\,, \tag{2}$$

where $PV(r_j)$ denotes the present value of the portfolio as a function of r_j, the value of the NMRF j.

In the paper, we will flag important cases where a design choice has been made to fulfil one of the listed goals. We will also recall, where relevant, the assumptions made.

[2]The shock x can be either positive or negative. When a shock is positive, the operator $\oplus$ applies an upward shock to the risk factor. Vice versa when x is negative a downward shock is applied.

2.2 The methodology

Having set the goals and laid down some basic assumptions, we are now ready to present the methodology to determine the stress scenario risk measure for a NMRF.

In accordance with paragraph MAR 33.12 of the FRTB, each risk factor in the bank's internal model is mapped to a specific liquidity horizon. Following the mapping, modellable risk factors are capitalised determining an expected shortfall measure calibrated on a 10-day horizon which is then rescaled to reflect the liquidity horizons of the modelled risks. To be consistent with the treatment envisaged for modellable risk factors, we reduce the problem of obtaining a stress scenario risk measure already capturing the risk factor liquidity horizon to the one of obtaining a stress scenario risk measure on a 10-day horizon, $SS_{10\mathrm{d}}$. Only at the final stage, we will rescale $SS_{10\mathrm{d}}$ to reflect the liquidity horizon of the non-modellable risk factor. Hence, the outcome of the methodology should be the identification of a loss under a stress scenario calibrated on a 10-day horizon.

To address the drawbacks resulting from a direct calculation of the expected shortfall of the losses, i.e. via the direct method presented in the introductory section, we propose in our methodology an approach based on the determination of the stress scenario risk measure $SS_{10\mathrm{d}}$ starting from the expected shortfall of the distribution of the 10-days returns observed for the NMRF in the stress period. Consistently with goal G3, such an approach should lead to:

$$SS_{10\mathrm{d}} \approx \mathrm{ES}(l(X), \alpha)\,, \tag{3}$$

having denoted with X the random 10-day return for the NMRF in the stress period, and $1-\alpha = 97.5\%$ being the confidence level of the ES measure.

We summarise the methodology as follows: first, we determine a time series of 10-business-day returns for the NMRF in the stress period and, from that time series, we estimate or approximate the left-tail and right-tail expected shortfall of the return distribution. Subsequently, we investigate the loss profile due to a change in the NMRF within the range identified by the two tail measures. Finally, we link that loss profile with the expected shortfall of the losses, so as to obtain a stress scenario risk measure that attracts a sufficient level of capital as in approximation in Eq. 3. As it will be clearer later, that link will mostly be based on the switch between the loss function operator $l(.)$ and the ES operator, i.e. we will link $\mathrm{ES}(l(X))$ with $l(\mathrm{ES}(X))$.

2.2.1 Building a time series of 10-business-day returns

For a risk factor with daily data, the determination of a time series of rolling 10-business-day returns is trivial as the identification of a 10-days-rolling period is always possible. However, as mentioned in goal G2, NMRF data may be sparse; hence, our methodology generalises the concept of a 10-business-day return to a 'nearest' to 10-business-day return to meet that goal. More specifically, we built up a parsimonious recipe based on the identification of the return on the nearest to 10-business-day window, that is then rescaled to obtain a return reflecting a 10-business-day period using the square-root-of-time rule employed in the FRTB for modellable risk factors.

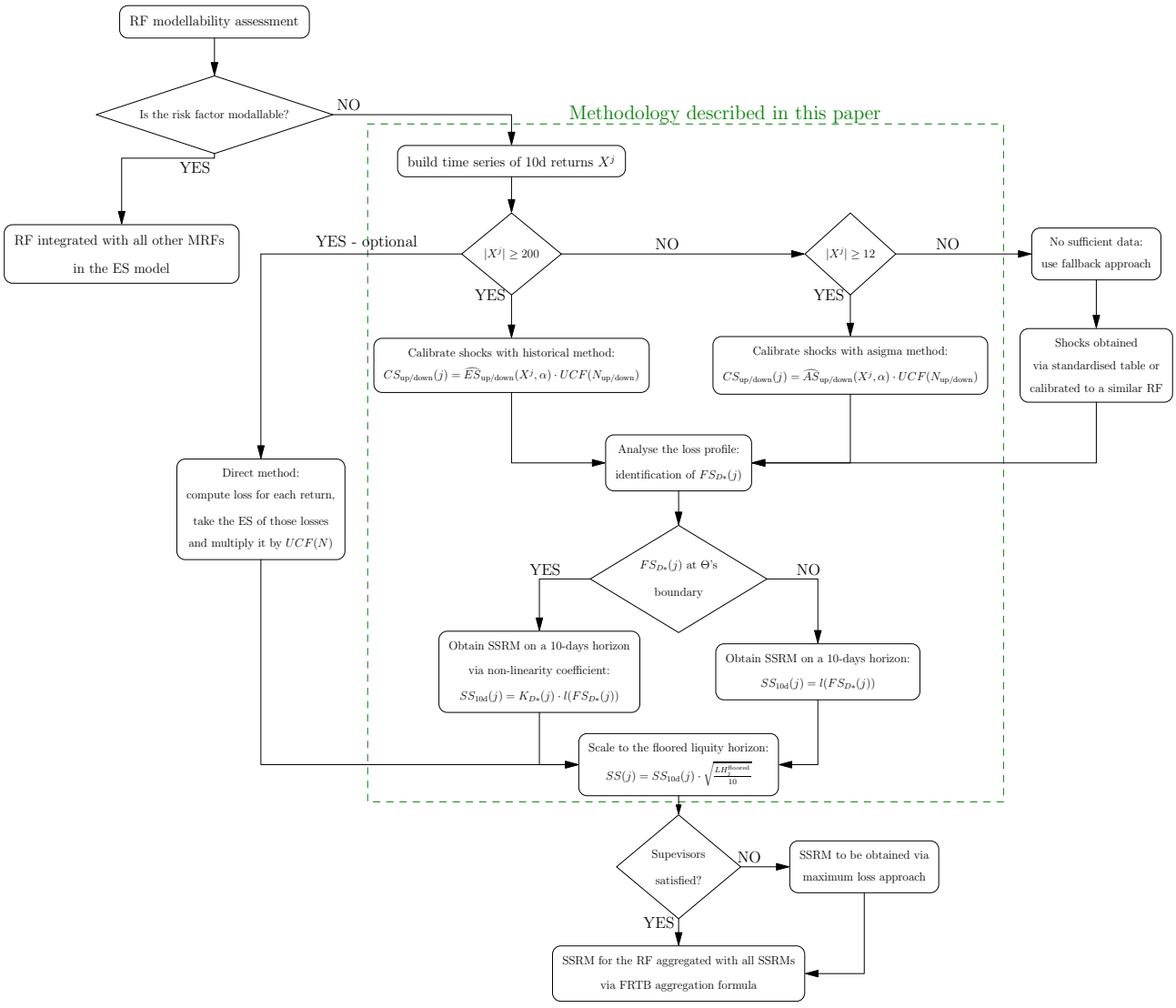
Let $D_1, D_2, \ldots, D_M$ be the ordered dates within the stress period S for which an observation for the NMRF j is present, and let $D_{M+1}, \ldots, D_{M+d-1}, D_{M+d}$ be the ordered dates representing the observation dates in the 20-business-day period following the stress period S. For each date index $t = 1, 2, \ldots, M-1$, we identify a "nearest next to 10 days" date index $t_{\mathrm{nn}}(t)$ by minimising the absolute relative deviation[3]:

$$t_{\mathrm{nn}}(t) = \underset{t'>t,\, t'\in\{2,\ldots,M+d\}}{\operatorname{argmin}} \left(\left| \frac{10\,\mathrm{days}}{D_{t'} - D_t} - 1 \right| \right) . \tag{4.a}$$

Once $t_{\mathrm{nn}}(t)$ is determined, the return $\mathrm{return}^j\left(D_t, D_{t_{\mathrm{nn}}(t)}\right)$ over the period identified by the two dates D_t and $D_{t_{\mathrm{nn}}(t)}$ is computed using a return calculation appropriate for this risk factor (e.g. log, relative, or absolute return). The date index $t \in \{1, \ldots, M-1\}$, indexing the 'starting observation' used to determine a return,

[3] As a rare special case, where an observation at both the 6th and 30th business days forward date $D_{t'}$ is available, with no dates in between (i.e. they both minimise the relative deviation), t' corresponding to the observation that occurred 30 business days after D_t is to be selected as $t_{\mathrm{nn}}(t)$.

Figure 1: This figure shows a flow-chart of the various steps that banks are to perform to capitalise their risk factors (RF) according to the currently proposed EU technical standards. First, banks assess with the risk-factor eligibility test (RFET) whether a RF is modellable or not. Then, the rescaled nearest to 10-day returns are computed. For the non-modellable RF the green box groups the steps of the methodology described in this paper. Depending on the number of returns in the one year stress period, different estimation methods are used to calibrate shocks to the left and right ES(97.5%). To cater for different signs of the sensitivity to the shocks (long or short) the highest loss occurring in the range from down to up shock is taken as the stress scenario risk measure basis value. Corrections for non-linearity, estimation uncertainty and a rescaling to the applicable liquidity horizon are applied as needed. For completeness we also include the 'direct method', i.e. where banks compute the losses when shocking the portfolio by the 10-day returns in the stress period from which the expected shortfall of those losses is used directly as stress scenario risk measure on a 10-day horizon. That method can be used only where the number of returns available ensure a robust ES estimation as in the historical method presented in this paper, [EBA20b] limits its use to cases where $N \geq 200$.



corresponds always to a date in the 1-year stress period, while $t' \in \{2, \ldots, M, M+1, \ldots, M+d\}$ might correspond to an 'ending observation' date in the 20-business-day period following the stress period. We extended the stress period S with 20 business days so as to avoid that too many returns at the end of S are obtained using the observation at D_M - potentially leading the last observation in S to be overly represented.

The return on the period identified by the two dates D_t and $D_{t_{\mathrm{nn}}(t)}$, is finally rescaled by $\sqrt{\frac{10\,\mathrm{days}}{D_{t_{\mathrm{nn}}(t)} - D_t}}$ to get the return X_t^j on a 10-business-day period:

$$X_t^j = \sqrt{\frac{10\,\mathrm{days}}{D_{t_{\mathrm{nn}}(t)} - D_t}} \times \mathrm{return}^j \left(D_t, D_{t_{\mathrm{nn}}(t)}\right) . \tag{4.b}$$

Eq. 4.a for $t_{\mathrm{nn}}(t)$ privileges the rescaling of a longer-than-10-day return to a 10-day return over the rescaling of a shorter-than-10-day return to a 10-day return if compared to a formula that identifies $t_{\mathrm{nn}}(t)$ by minimising $|D_{t'} - D_t - 10\ \mathrm{days}|$ or $\left|\sqrt{\frac{10\,\mathrm{days}}{D_{t'} - D_t}} - 1\right|$. In particular, dates distancing from D_t for more than 10 business days are always preferred to dates distancing for 5 or fewer business days. Favouring longer-dated returns was done considering that paragraph MAR 33.16 of the FRTB sets out that the liquidity horizon of a NMRF must be greater than or equal to 20 days – thus, longer horizons will likely be closer to the final applicable liquidity horizon. In addition, rescaling shorter returns to a longer period brings with itself the possibility of amplifying short term movements; vice versa, rescaling longer returns to a shorter period tends to remove those short term effects.

Other methods for constructing a 10-business-day returns time-series by imputation of data in case of missing observations could be envisaged. However, they might require a time-series modelling (e.g. via a GARCH model approach) and might need a fitting or regression procedure. This would increase complexity, in contrast to goal G4, and different kinds of risk factors might require different approaches, in contrast to goal G1.

As a result, regardless of the number of NMRF observations in the stress period, a time series with $N = M-1$ returns is obtained. We denote with X^j, the set of returns for the NMRF j in the time series obtained as a result of the steps of Eq. 4. From that sample, our objective is to obtain two shocks resembling the right and left-tail expected shortfall of the return distribution, with the idea of linking them with the expected shortfall of the losses calculated on that return distribution (i.e. linking $\mathrm{ES}(X)$ with $\mathrm{ES}(l(X))$.

2.2.2 Calibrating extreme shocks from the sample X^j

As mentioned, we now want to calibrate an upward and downward shock for the NMRF j, respectively $CS_{\mathrm{up}}(j)$ and $CS_{\mathrm{down}}(j)$, from the sample of returns X^j. The two 'calibrated shocks' are estimates of the right-tail and left-tail expected shortfall of the distribution of the 10-business-day returns for the NMRF j in the stress period.

In goal G2, we stated that processing for a wide of range of cases with respect to number of returns available in the stress period should be possible. An α-tail ES estimation from a large sample of N returns of which $\lfloor \alpha N \rfloor$ fall in the α-tail is not posing any problem. Having at least 5 observations in the tail, i.e. $N \geq 200$, leads to a sufficiently robust estimation, as we will show in Section 5.2.2. As a result, for calibrating $CS_{\mathrm{up}}(j)$ and $CS_{\mathrm{down}}(j)$, a historical estimator is used when at least $N^{\mathrm{hist}} = 200$ 10-day returns are available in the sample of returns X^j. $CS_{\mathrm{up}}(j)$ and $CS_{\mathrm{down}}(j)$ as determined by this 'historical method' are:

$$CS_{\mathrm{down}}(j) = \widehat{ES}_{\mathrm{down}}(X^j, \alpha) \cdot UCF(N) \tag{5}$$

and:

$$CS_{\mathrm{up}}(j) = \widehat{ES}_{\mathrm{up}}(X^j, \alpha) \cdot UCF(N)\,, \tag{6}$$

where $\widehat{ES}_{\mathrm{down}}(X^j, \alpha)$ and $\widehat{ES}_{\mathrm{up}}(X^j, \alpha)$ denote respectively the left α-tail and right α-tail historical expected shortfall estimators introduced in Eq. 7 or 8 applied to the sample X^j of returns for the NMRF j, and $\alpha = 2.5\%$. $UCF(N)$ is the uncertainty compensation factor for N returns to be introduced in Section 4 – it implements the principle introduced in G7 that less data should lead to more capital, and reflects the error in estimating the expected shortfall.

The following historical estimator for the expected shortfall for the left tail of the sample distribution is used:

$$\widehat{ES}_{\text{down}}(X^j, \alpha) = \frac{-1}{\alpha N} \left\{ \sum_{i=1}^{\lfloor \alpha N \rfloor} X^j_{(i)} + (\alpha N - \lfloor \alpha N \rfloor) X^j_{(\lfloor \alpha N \rfloor + 1)} \right\}, \tag{7}$$

where $X^j_{(i)}$ is the order statistics of the sample X^j of size N, and $\lfloor \alpha N \rfloor$ denotes the integer part (floor) of the product αN. We adopt a sign convention leading to a positive number for the left (negative) tail of a distribution centred around zero like in [AT02]. This estimator (Eq. 23 in [NZC14]) is slightly different from the simple historical estimator which uses only the $\lfloor \alpha N \rfloor$ worst losses. It is more natural for the expected shortfall being an α-tail mean (cf. Definition 2.6 in [AT02]), accounts for αN not being an integer and is somewhat more stable by incorporating $X^j_{(\lfloor \alpha N \rfloor + 1)}$. The historical estimator for the α ES of the right tail is:

$$\widehat{ES}_{\text{up}}(X^j, \alpha) = \widehat{ES}_{\text{down}}(-X^j, \alpha)\,. \tag{8}$$

The sign convention is also leading to a positive number for the right (positive) tail of a distribution centered around zero. The estimators deliberately do not remove a mean (drift) in the return sample which is kept as a feature of return shocks in a stress period. While such a 'de-meaning' could be appropriate e.g. for FX rates and short liquidity horizons, where a zero-mean assumption might be warranted, we note that the risk factors are varied as stated in goal G1 and could e.g. also be parameters of a yield curve or volatility surface where a zero mean assumption would not apply. Furthermore, such 'de-meaning' would fit more for pricing purposes rather than for the purpose of generating risk measures calibrated to a stress period, i.e. based on historically observed data. That said, typically the means are small compared to the volatility or the ES.

If the number of returns N in the sample X^j is lower than N^{hist}, the robust estimation of the ES becomes more challenging. The key idea in our methodology for this case is to approximate the expected shortfall by rescaling a volatility measure, because the standard deviation can be estimated more robustly for a small sample than the ES with the historical estimator. The HÜRLIMANN bound [Hür02] states $\text{ES}(\alpha) \leq \mu + \sqrt{\frac{1-\alpha}{\alpha}}\sigma$ for any continuous distribution on $\mathbb{R}$ with given mean μ and standard deviation σ and motivates that the ES can be approximated by a (distribution dependent) multiple of σ when the mean is negligible. We will recover this location-scale dependence explicitly for the SGT distribution family in Subsection 5.1.

Instead of rescaling the standard deviation on the sample comprising all returns in the time series, our methodology splits the set of returns at the median into two halves, and on each half, a quantity resembling the mean and standard deviation is estimated. This allows for capturing asymmetry in a distribution which is often found in risk factor returns [EBA20a, EBA20b]. Higher than second moments are deliberately not used so as to make the estimator more robust against outliers and more suitable for small samples. Finally, to ensure that the standard deviations on the two halves are robustly estimated, the determination of $CS_{\text{down}}(j)$ and $CS_{\text{up}}(j)$ via such rescaling is limited to cases where there are at least $N \geq N^{\text{asigma}} = 12$ returns in the sample X^j – we motivate the value set for N^{asigma} in Subsection 5.2.2.

Formally, we first define the sets of returns in each half, $X^j_{\text{down}} := \{X^j \leq \text{med}(X^j)\}$ and $X^j_{\text{up}} := \{X^j > \text{med}(X^j)\}$ with $\text{med}(X^j)$ denoting the median of the sample of returns X^j and $N_{\text{down}} = \lfloor \frac{N}{2} \rfloor + N \bmod 2$, $N_{\text{up}} = \lfloor \frac{N}{2} \rfloor$, or $N_{\text{down,up}} = \frac{N}{2} \pm \frac{N \bmod 2}{2} \approx \frac{N}{2}$, the cardinalities of the two sets.

The calibrated shocks of this 'asymmetrical sigma' method are:

$$CS_{\text{down}}(j) = \widehat{AS}_{\text{down}}(X^j) \cdot UCF(N_{\text{down}}) \tag{9}$$

and

$$CS_{\text{up}}(j) = \widehat{AS}_{\text{up}}(X^j) \cdot UCF(N_{\text{up}})\,, \tag{10}$$

where $UCF(N_{\text{down}})$ and $UCF(N_{\text{up}})$ capture the uncertainty in approximating the expected shortfall measures, and where $\widehat{AS}_{\text{down}}(X^j)$ and $\widehat{AS}_{\text{up}}(X^j)$ are the 'asigma' estimators approximating the left and right tail expected shortfalls from the set X^j:

$$\widehat{AS}_{\text{down}} = -\widehat{\mu}_{X^j_{\text{down}}} + \frac{C^{\text{asigma}}_{\text{ES}}}{\sqrt{N_{\text{down}} - \frac{3}{2}}} \sqrt{\sum_{X^j_i \in X^j_{\text{down}}} \left(X^j_i - \widehat{\mu}_{X^j_{\text{down}}}\right)^2} = -\widehat{\mu}_{X^j_{\text{down}}} + C^{\text{asigma}}_{\text{ES}}\, \widehat{\sigma}_{X^j_{\text{down}}}\,, \tag{11}$$

where $\widehat{\mu}_{X^j_{\text{down}}}$ and $\widehat{\sigma}_{X^j_{\text{down}}}$ denote the mean and standard deviation estimators of the returns below the median. The minus sign in front of $\widehat{\mu}_{X^j_{\text{down}}}$ leads to a positive calibrated shock for the typical case that $\widehat{\mu}_{X^j_{\text{down}}} < 0$ when the median is near zero.

Analogously,

$$\widehat{AS}_{\text{up}} = \widehat{\mu}_{X^j_{\text{up}}} + \frac{C_{\text{ES}}^{\text{asigma}}}{\sqrt{N_{\text{up}} - \frac{3}{2}}} \sqrt{\sum_{X_i^j \in X^j_{\text{down}}} \left(X_i^j - \widehat{\mu}_{X^j_{\text{up}}}\right)^2} = \widehat{\mu}_{X^j_{\text{up}}} + C_{\text{ES}}^{\text{asigma}}\, \widehat{\sigma}_{X^j_{\text{down}}} \,. \tag{12}$$

The two estimators rescale estimates of the standard deviation on the two sets by $C_{\text{ES}}^{\text{asigma}}$. For small sample sizes, taking the square root of the usual sample estimator for the variance as the estimator for the standard deviation leads to a biased estimate due to JENSEN's inequality [Jen06, Ben03] and the convexity of the square root function. We use the approximately unbiased sample estimator for the standard deviation [GT71], keeping only the leading term $N - \frac{3}{2}$. Setting the constant $C_{\text{ES}}^{\text{asigma}}$ so as to ensure an approximation of the ES measure that leads to an adequate level of capitalisation of NMRFs as laid down in goal G3 is key. We set it to 3 for all risk factors - in Section 5, we motivate in detail our choice which was also confirmed by a calibration on a large collection of risk factors' historical data [EBA20a, EBA20b].

We have covered $N \geq 200$ with the historical method and $N \geq 12$ with the asymmetrical sigma method. Below $N^{\text{asigma}} = 12$ returns, there are insufficient data to robustly approximate an ES with reasonable accuracy. In this case, Assumption A3 is considered to not be met in practice, hence, we do not address it in this section. However, we will briefly discuss this case in Section 7 where a fallback method is sketched.

As a result, for any N, the methodology yields a downward shock $CS_{\text{down}}(j)$ and an upward shock $CS_{\text{up}}(j)$. The large upward and downward shock span the shock range $[-CS_{\text{down}}(j), CS_{\text{up}}(j)]$. We now aim to identify the shock from this interval which leads to the highest portfolio loss in order to have a sufficient understanding of the loss profile to infer information about $\text{ES}(l(x))$ – in particular, we are chiefly interested to determine whether the worst loss in the range occurs at its boundaries. This shock will be called the extreme scenario of future shock $FS_{D^*}(j)$.

2.2.3 Analysing the loss profile in the range $[-CS_{\text{down}}(j), CS_{\text{up}}(j)]$

The loss function $l_j(x)$ describes the loss of a whole SSRM portfolio under variation of the single risk factor j. In the most typical situation, $l_j(x)$ is not available in analytical form and no shape properties, e.g. neither monotonicity nor extremal values, are known a priori. Therefore, the only way to understand the loss profile is to explicitly calculate $l_j(x)$ for various shocks x.

A natural and simple approach to finding the highest loss is to cover the range $[-CS_{\text{down}}(j), CS_{\text{up}}(j)]$ with a finite grid of points, evaluate $l_j(x)$ for every vertex and pick the largest value. Clearly, the finer the grid is chosen, the more precisely the maximum can be determined. However, each calculation of $l_j(x)$ amounts to a portfolio revaluation, which is computationally costly given that in that portfolio many instruments dependent on the NMRF j may be present. Moreover, the procedure has to be repeated for each NMRF and computation results cannot be reused over several reference dates as every change in the portfolio composition potentially changes the loss profiles with respect to all NMRFs. Scanning the shock range with a very granular grid would therefore entail a very large overall computational cost, in conflict with goal G4.

As a pragmatic solution, the methodology requires the evaluation of $l_j(x)$ on a grid Θ of only four points, comprising the two boundary and two inner points:

$$\Theta = \underbrace{\{-CS_{\text{down}}(j), CS_{\text{up}}(j)\}}_{\text{boundary points}} \cup \underbrace{\{-0.8\, CS_{\text{down}}(j), 0.8\, CS_{\text{up}}(j)\}}_{\text{inner points}} \,. \tag{13}$$

The extreme scenario of future shock $FS_{D^*}(j)$ is given by the shock that leads to the highest portfolio loss:

$$FS_{D^*}(j) = \operatorname*{argmax}_{x \in \Theta} l_j(x) \,. \tag{14}$$

With such a coarse-grained scanning grid, the methodology cannot be expected to identify the highest loss in the range under all circumstances. Importantly, though, it yields the correct outcome in the most common case. When the portfolio has a directional exposure to the NMRF j and hence $l_j(x)$ is monotonic, the worst loss in the range occurs for one of the boundary points. Concretely, for a net long position in j, the methodology correctly identifies $-CS_{\text{down}}(j)$, for a net short position, it yields $CS_{\text{up}}(j)$. Adding the loss evaluations at the 80% grid points makes the methodology more robust when extreme tail hedges are in place. The choice of including the point $0.8\,CS_{\text{up}}(j)$ was also done with a view to the non-linearity correction introduced in Subsection 2.2.4 and is motivated in Subsection 5.2.3, showing $\text{VaR}(97.5\%) \approx 0.8\,\text{ES}(97.5\%)$ for a wide range of return distributions. In theory, $l_j(x)$ can be an arbitrarily complicated function with maxima well inside the interval, which will not be picked up by the methodology. However, given that portfolios tend to be fairly well hedged against small risk factor fluctuations, focusing exclusively on large returns yields adequate results in most practical cases [EBA20a, EBA20b] – hence, our choice to limit the loss evaluation to those four points only.

Consistently with goal G3, the methodology aims at estimating the expected shortfall of losses. We therefore have to link the loss $l_j(FS_{D^*}(j))$ to $\text{ES}(l_j(X))$. By definition, the expected shortfall is the average of the worst 2.5% losses and hence depends on both the return distribution and the shape of the loss function $l_j(x)$ in a potentially non-trivial manner[4]. As said, we cannot use detailed information on $l_j(x)$ as this would increase the computational footprint. Instead, we rely on further assumptions, dependent on whether the scenario corresponds to a shock from the inside or the boundary of the shock range. For the time being, let us assume that at least one of the four points from Θ actually leads to a loss so that $l_j(FS_{D^*}(j)) > 0$. We distinguish the case where $FS_{D^*}(j)$ is at the boundaries of Θ from the case where it is in the inner point. We will briefly discuss corner cases, e.g. cases where $l_j(FS_{D^*}(j)) \leq 0$, in Section 3.

2.2.4 Obtaining $SS_{10\text{d}}$ when $FS_{D^*}(j)$ is at the boundaries of Θ via a non-linearity correction

We expect in the vast majority of cases the extreme scenario of future shock $FS_{D^*}(j)$ to be at the boundaries of Θ, given that trading portfolios tend to be fairly well hedged against small risk factor fluctuations, and that, typically, for large shocks – like those included in Θ – the losses are monotonic.

Therefore, where $FS_{D^*}(j)$ is at the boundaries of Θ, we assume that the worst 2.5% losses are due to tail events of the return distribution. Concretely, where $CS_{\text{up}}(j)$ leads to the highest loss, we assume that the 2.5% worst losses come from the 2.5% largest returns. Analogously, where the highest loss is associated with $-CS_{\text{down}}(j)$, we assume that the 2.5% worst losses come from the 2.5% smallest (or most negative) returns. As part of a data collection exercise [EBA19], data on portfolio loss functions upon change in single risk factors were collected both for standardised portfolios (the EBA Supervisory benchmarking exercise portfolios [EBA21]) and some banks' real trading portfolios which included also structured products. Those data backed the presumption that banks portfolios are in general directional, i.e. that in the vast majority of cases the extreme shocks will be at the boundaries of Θ, and that the losses are in general monotonic for large shocks. We formally investigate the case where $CS_{\text{up}}(j)$ is the extreme scenario of future shock. The analysis for $-CS_{\text{down}}(j)$ is fully analogous and differs only in signs. In Figure 2, we show the various elements that are discussed in this subsection.

We employ a general return distribution X in this subsection and while the results are general, it will be the sample of 10-business-day obtained from Eq. 4 in Subsection 2.2.1 to which the results are applied.

Let $q_{97.5\%}(l_j(X))$ and $q_{97.5\%}(X)$ denote the 97.5th percentile of the loss and the return distribution, respectively. As stated above, we assume that $\{x|l_j(x) > q_{97.5\%}(l_j(X))\} = \{x|x > q_{97.5\%}(X)\}$. Using the equality of these two sets, the expected shortfall of losses can be calculated as

$$\text{ES}(l_j(X)) = \frac{1}{2.5\%}\mathbb{E}\left(l_j(X) \cdot I_{X>q_{97.5\%}(X)}\right) , \tag{15}$$

where $I_{(\cdot)}$ denotes the indicator function. In other words, the expected shortfall is the average loss over tail events of shocks X applied to risk factor j. To make this explicit, let X_{T} be a new random variable, distributed

[4]Different from our definition, the mathematics literature typically defines expected shortfall as an average of quantiles. However, the two definitions are equivalent for continuous distributions and hence for all practical purposes. Moreover, our definition coincides with the one given in paragraph MAR 10.18 of the FRTB.

like X but conditional on X being larger than $q_{97.5\%}(X)$, i.e. $X_\mathrm{T} := X \cdot I_{X>q_{97.5\%}(X)}$. Note that $\mathbb{E}(X_\mathrm{T}) = \mathrm{ES}(X)$ by definition of the expected shortfall of X. We have:

$$\mathrm{ES}(l_j(X)) = \mathbb{E}(l_j(X_\mathrm{T}))\,. \tag{16}$$

If $l_j(x)$ was a linear function one could exchange it with the integral to find the equality $\mathrm{ES}(l_j(X)) = l_j(\mathrm{ES}(X))$. However, depending on whether $l_j(x)$ is convex or concave, Jensen's inequality for expected values [Jen06, Ben03] applies and the left-hand side can be significantly larger or smaller than the right-hand side, respectively. We capture this non-linearity effect by an analytical approximation.

To this end, we use a second-order Taylor expansion of $l_j(x)$ around $\mathrm{ES}(X)$:

$$\mathbb{E}\left(l_j(X_\mathrm{T})\right) = \mathbb{E}\left(l_j(\mathrm{ES}(X)) + \left.\frac{\partial l_j}{\partial x}\right|_{x=\mathrm{ES}(X)} (X_\mathrm{T} - \mathrm{ES}(X)) + \frac{1}{2}\left.\frac{\partial^2 l_j}{\partial x^2}\right|_{x=\mathrm{ES}(X)} (X_\mathrm{T} - \mathrm{ES}(X))^2\right) \tag{17}$$

and by noting that $\mathbb{E}(X_\mathrm{T}) = \mathrm{ES}(X)$, the linear term drops out from Eq. 17 and we find:

$$\mathbb{E}\left(l_j(X_\mathrm{T})\right) \approx l_j(\mathrm{ES}(X)) + \frac{1}{2}\left.\frac{\partial^2 l_j}{\partial x^2}\right|_{x=\mathrm{ES}(X)} \left(\mathbb{E}\left(X_\mathrm{T}^2\right) - \mathrm{ES}(X)^2\right)\,. \tag{18}$$

After reorganising terms and going back from X_T to X, the approximate expected shortfall of losses can be written as

$$\mathrm{ES}(l_j(X)) \approx l_j(\mathrm{ES}(X)) + \frac{1}{2}\,\Gamma\,(\phi - 1)\,\mathrm{ES}(X)^2 = l_j(\mathrm{ES}(X)) \cdot K\,, \tag{19}$$

where we have introduced the quantities

$$\Gamma := \left.\frac{\partial^2 l_j}{\partial x^2}\right|_{x=\mathrm{ES}(X)}\,, \quad \phi := \frac{\mathbb{E}\left(X_\mathrm{T}^2\right)}{\mathrm{ES}(X)^2} = \frac{\mathbb{E}\left(X^2 \cdot I_{X>q_{97.5\%}(X)}\right)}{\mathbb{E}\left(X \cdot I_{X>q_{97.5\%}(X)}\right)^2} \geq 1\,, \tag{20}$$

remembering that $\mathrm{ES}(X) = \mathbb{E}(X_\mathrm{T})$. The fact that $\phi \geq 1$ can be proved by applying Jensen's inequality with the convex function $f(x) = x^2$. With those quantities we define

$$K := 1 + \frac{1}{2}\frac{\Gamma}{l_j\left(\mathrm{ES}(X)\right)}\,(\phi - 1)\,\mathrm{ES}(X)^2\,. \tag{21}$$

This is our analytical approximation and we call K the 'non-linearity correction'. The formula for K nicely disentangles the two components that drive $\mathrm{ES}(l_j(X))$: Γ, the curvature of the loss function, and ϕ, which is a tail shape measure for how heavy the tails of the return distribution are. Both super-linear loss profiles and heavy-tailed return distributions are common features in financial markets, so arguably such non-linearity corrections can be non-negligible for real-world portfolios.

While the Taylor polynomial is a *local* approximation of $l_j(x)$ around $\mathrm{ES}(X)$, we use it as an approximation of $l_j(x)$ for all $x > q_{97.5\%}(X)$. The quality of the approximation can therefore vary substantially, depending on the global shape of $l_j(x)$. Later in this section we account for this by capping and flooring the non-linearity correction leading to Eq. 26.

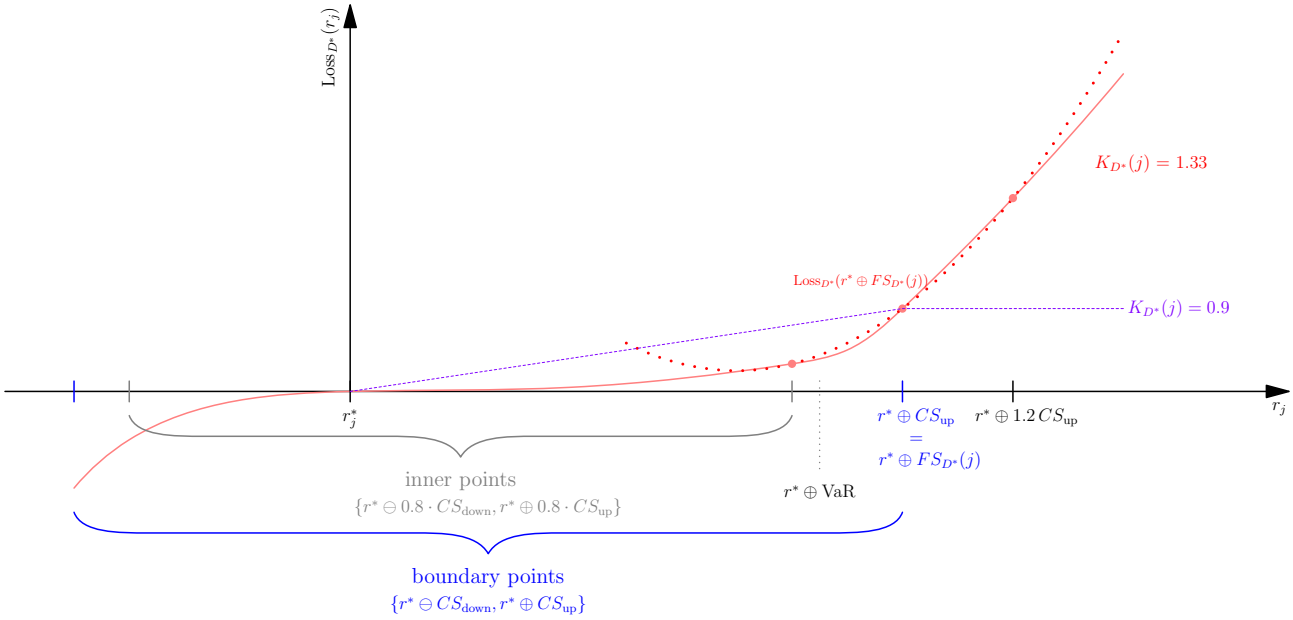
Γ is the second derivative of $l_j(x)$ with respect to the return x around $\mathrm{ES}(X)$. We note that it can be non-zero even for portfolios consisting exclusively of linear products in the risk factor. This is the case when returns are applied non-linearly to a risk factor, i.e. when $r_j^* \oplus x$ is a non-linear function of x. The typical example are log returns for which $l_j(x) = \mathrm{Loss}_{D^*}(r_j^* \oplus x) = \mathrm{Loss}_{D^*}(r_j^* \cdot e^x)$.

We determine the derivative numerically using a second-order finite central differences formula [Far93] with step $h = 0.2\,\mathrm{ES}(X)$:

$$\Gamma \approx \frac{l_j\left(\mathrm{ES}(X) - h\right) - 2\,l_j\left(\mathrm{ES}(X)\right) + l_j\left(\mathrm{ES}(X) + h\right)}{h^2}\,. \tag{22}$$

The relatively large step width parameter h limits the dependence on local peculiarities of $l_j(x)$ around $\mathrm{ES}(X)$. Instead, the derivative based on these three points can rather be thought of as a global indicator of the curvature of $l_j(x)$ in the tail regime for the purpose of integration from $\mathrm{VaR}(X)$ to ∞ (see Figure 2). In Subsection 5.2.3 we motivate the value for h by showing that for a wide range of risk factors, $q_{97.5\%}(X) =$

Figure 2: We show in an example for a risk factor with absolute returns how the worst loss is determined for a loss profile and the computation of the non-linearity correction $K_{D^*}(j)$ (cf. Eq. 26). The risk factor values of the grid Θ for which the loss must be evaluated are indicated on the x-axis. With the continuous red line, we depict the loss profile. In this example, the worst loss in the grid Θ occurs at CS_{up}, which is an estimate of the expected shortfall for upward shocks, $\text{ES}(X)$. Hence, CS_{up} is the extreme scenario of future shock $FS_{D^*}(j)$. The three points to calculate the non-linearity coefficient $K_{D^*}(j)$ are identified with a red circle on the loss profile. The parabola passing through the three points in the return space is the same for in the risk factor space for absolute returns and represented with a dotted red line. We obtained a numerical value for the non-linearity correction $K_{D^*}(j) = 1.33$ for the represented loss function assuming $\phi^{\text{asigma}} = 1.04$. Assuming $\phi^{\text{asigma}} = 1.04$, we also show with a purple dashed line a loss profile that would lead to reaching the floor 0.9 for $K_{D^*}(j)$. The dotted grey vertical line indicates $r^* \oplus \text{VaR}(X) = r^* \oplus q_{97.5\%}(X)$, which is the start of the integration to infinity for calculating the ES of losses (cf. Eq. 15).



$\text{VaR}(X, 97.5\%) \approx 0.8\,\text{ES}(X, 97.5\%)$, so that $h = 0.2$ means to take the inner point of the Γ computation approximately as the lower bound of the integration for the non-linearity correction.

An alternative interpretation is the following: rather than replacing the loss profile by the tangential Taylor parabola, we effectively replace $l_j(x)$ by the single parabola that goes through the three points defining the derivative. To see this, note that linear terms drop out in the calculation above so that one can, without changing the result, tilt the Taylor parabola around $\text{ES}(X)$ until it hits all three points.

The quantity ϕ characterises the heaviness of the tails of the return distribution, i.e. the dispersion of the returns in the tail around the ES value. Being a tail property, a sufficiently large sample of returns is required to estimate ϕ. Where we have recorded at least $N \geq N^{\text{hist}} = 200$ returns, we determine ϕ using the statistical estimator $\widehat{ES}_{\text{up}}(X^j, \alpha)$ introduced in Eq. 8 combined with Eq. 20. We define the following estimator:

$$\widehat{\phi}_{\text{up}}(X^j, \alpha) = \frac{\frac{1}{\alpha N}\left\{\sum_{i=1} \lfloor \alpha N \rfloor (-X^j)^2_{(i)} + (\alpha N - \lfloor \alpha N \rfloor)\,(-X^j)^2_{(\lfloor \alpha N \rfloor + 1)}\right\}}{\widehat{ES}^2_{\text{up}}(X^j, \alpha)}\,. \tag{23}$$

where, as in Eq. 7, the index (i) is used to indicate the order statistics. If less than $N^{\text{hist}} = 200$ returns are available, we do not rely on an estimator - hence, where the asigma method is used to determine the calibrated shocks, we use a flat value of $\phi^{\text{asigma}} = 1.04$. In Section 5.2.1 we motivate this constant which corresponds to the assumption that the return distribution has moderately heavy tails and was also confirmed by historical data [EBA20a, EBA20b].

We now combine these components to arrive at a concrete formula for the non-linearity correction coefficient $K_{D^*}(j)$. Where the scenario of future shock is associated with $-CS_{\text{down}}(j)$, all occurrences of the right-tail expected shortfall $\text{ES}(X)$ have to be replaced by its left-tail counterpart. Moreover, the calculations in ϕ must refer to the left rather than the right tail of the distribution of X. Overall, we obtain:

$$\tilde{K}_{D^*}(j) = 1 + \frac{25}{2} \frac{l_j(0.8\, FS_{D^*}(j)) - 2\, l_j(FS_{D^*}(j)) + l_j(1.2\, FS_{D^*}(j))}{l_j(FS_{D^*}(j))} \left(\widehat{\phi} - 1\right) , \tag{24}$$

where we set $\text{ES}(X)$ equal to $FS_{D^*}(j)$ in Eqs. 20 and 22 given that the latter is the estimation or approximation of the former, and where:

$$\widehat{\phi} = \begin{cases} \widehat{\phi}_{\text{up}} & \text{if} \quad N \geq N^{\text{hist}} \text{ and } FS_{D^*}(j) = CS_{\text{up}}(j); \\ \widehat{\phi}_{\text{down}} = \widehat{\phi}_{\text{up}}(-X^j, \alpha) & \text{if} \quad N \geq N^{\text{hist}} \text{ and } FS_{D^*}(j) = -CS_{\text{down}}(j); \\ \phi^{\text{asigma}} = 1.04 & \text{if} \quad 12 \leq N < N^{\text{hist}}. \end{cases} \tag{25}$$

Finally, we apply a cap and a floor,

$$K_{D^*}(j) = \max\left(K_{\min}, \min\left(\tilde{K}_{D^*}(j), K_{\max}\right)\right) , \tag{26}$$

where $K_{\min} = 0.9$ and $K_{\max} = 5$. The floor at $K_{\min}$ limits unintended consequences of our approximation approach when the second derivative Γ at $FS_{D^*}(j)$ is negative. Typically, this is indicative of a flattening loss profile for large returns, which justifies $K_{D^*}(j) < 1$. However, recall that in this case, our methodology replaces the loss function $l_j(x)$ with a downward opening parabola. That is, the approximate loss profile does not only flatten, but actually bends and eventually turns negative (i.e. turns from losses into gains) for very large returns, which is highly unrealistic. It can be shown that under optimistic assumptions a maximum benefit of $K_{D^*}(j) = 0.9$ can be expected; this can be seen by considering a hypothetical loss profile that grows linearly from zero, reaches the value $l_j(\text{ES}(X))$ at $x = \text{ES}(X)$ and then stays constant for all larger returns, which serves as our optimistic benchmark (cf. Figure 2). Γ can be calculated explicitly, and using $\phi = \phi^{\text{asigma}} = 1.04$ we find that $K_{D^*}(j) = 0.9$ in this case. Therefore, where the formula for $K_{D^*}(j)$ yields values smaller than 0.9, this is considered a relic of the approximation and the floor applies. The large cap at $K_{\max}$ corresponds to a very steep loss profile and is merely a precaution to prevent implausibly large results that could be triggered e.g. by numerical issues in approximating the second order derivative.

We also note that for $0.8\, FS_{D^*}(j)$, we have already evaluated the loss function when searching the highest loss in Θ. Hence, in line with goal G4, only one additional loss function evaluation at $1.2\, FS_{D^*}(j)$ is necessary to calculate $\tilde{K}_{D^*}(j)$.

This marks the end of our discussion and we can finally calculate the 10-day SSRM for the case where the scenario of future shock is one of the boundary points of the shock range:

$$SS_{10\text{d}}(j) = K_{D^*}(j) \cdot l(FS_{D^*}(j)) \quad \text{if} \quad FS_{D^*}(j) \in \{-CS_{\text{down}}(j), CS_{\text{up}}(j)\} . \tag{27}$$

As part of the data collection exercise [EBA19], the stress scenario risk measures resulting from the application of our methodology were provided. In a few cases, banks even reported the stress scenario risk measure by directly computing the expected shortfall of the losses, i.e. via the direct method. For risk factors with more than 200 returns in the stress period, results showed that our methodology ensures a level of capitalisation comparable to the one resulting from the direct method. For example, one bank participating in the data collection exercise which made its results public [Int20], obtained a total stress scenario risk measure by applying the historical method (with a grid for Θ also having 20%, 40% and 60% of the boundaries) which was only 1% higher than when calculating the expected shortfall of the losses with the direct method, for both linear and structured interest rate strategies.

2.2.5 Obtaining $SS_{10\text{d}}$ when $FS_{D^*}(j)$ is at the inner points of Θ

Where the extreme scenario of future shock corresponds to one of the inner points, we are left with no clear guess how $l_j(x)$ might possibly look, apart from the indication that the worst losses presumably do not occur

for the most extreme returns of the NMRF. Possible refinements of our strategy could require the evaluation of $l_j(x)$ at further grid points at this stage to get a clearer picture, but to keep the methodology light and simple in line with the goals, our approach refrains from such extensions. This choice has been made also considering that, as mentioned, losses are in general directional in the risk factors - hence, $FS_{D^*}(j)$ is rarely expected to be in the inner points of Θ. Instead, we make the blunt assumption that the scenario of future shock at the inner points of Θ is representative for the worst 2.5% losses and set the 10-day SSRM to the corresponding loss,

$$SS_{10\mathrm{d}}(j) = l_j(FS_{D^*}(j)) \quad \text{if} \quad FS_{D^*}(j) \in \{-0.8\, CS_{\mathrm{down}}(j), 0.8\, CS_{\mathrm{up}}(j)\}\ , \tag{28}$$

accepting that this could underestimate, or overestimate, $\mathrm{ES}(l_j(X))$ slightly.

3 Corner cases and maximum loss approach

As mentioned before, the methodology is expected to work reasonably well in the vast majority of cases. However, special cases could also occur. For example, a very rare case would occur when all four points from the grid Θ lead to zero or negative losses, i.e. actual gains. In this rare situation, we set the 10-day SSRM for the risk factor j to zero:

$$SS_{10\mathrm{d}}(j) = 0 \quad \text{if} \quad l_j(FS_{D*}(j)) \leq 0\,. \tag{29}$$

In addition, when putting forward our methodology, we introduced some assumptions. Assumption A1 and A2, i.e. the identification of the SSRM portfolio and of the NMRFs via the RFET, are prerequisites envisaged in the FRTB. Unlike A1 and A2, A3 and A4 are assumptions that may not be met in practice; for example, there could be shocks for which the pricers of a bank are not able to calculate the corresponding loss, and there could be NMRFs with less than 12 observations in the stress period. [EBA20b] covers those cases by means of additional provisions making our methodology even more complete. To address cases where the pricers cannot determine the loss corresponding to a shock, [EBA20b] set out that banks are to use sensitivity based pricing methods. However, a sensitivity based loss can be computed only in relation to those instruments for which the pricers do not work - this to reflect cases where for the same shock, the pricers of some instruments provide the loss results, while for others they do not, as that shock would lead to arbitrage conditions in the context of those instruments (e.g. because only the NMRF is shocked, and all other risk factors are held constant). For NMRFs with less than 12 returns in the stress period, [EBA20b] envisages a fallback approach which relies on our methodology. Depending on the nature of the risk factor, CS_{up} and CS_{down} under such a fallback approach are either based on the FRTB standardised approach pre-defined shocks or on the shocks calibrated on the returns of a 'similar risk factor' for which at least 12 returns are available. Once those shocks are identified, our methodology is applied in the same way as for risk factors for which more than 12 returns in the stress period are available. Considering that the assumption A4 addresses a problem that may also occur in the ES for modellable risk factors, and that the fallback approach leads back to our methodology, we decided this paper need not be focused on those ad-hoc solutions.

More generally, to address rare cases where supervisors are not satisfied, paragraph 33.16(3) of the FRTB foresees the possibility for supervisors to require a bank to set the stress scenario risk measure to the maximum loss that can occur due to a NMRF. This may happen e.g. when the worst loss for a change in an NMRF occurs as a result of a small change to its current value, far below the 80% of the calibrated shock, or when the SSRM portfolio composition is such that the loss profile for a NMRF cannot be duly captured by the methodology as it could be for very exotic options. Whenever there is a need to, the supervisor could revert to the maximum loss. However, the maximum possible loss could be infinite and a sensible finite replacement is needed. To this end [EBA20b] stipulates that where the maximum loss is infinite, banks are to identify a VaR(99.95) of the losses that may occur due to the NMRF on the 10-business-day horizon rescaled to the NMRF's liquidity horizon. In Table 2 in Section 5, we show how the ES(97.5%) compares to the VaR(99.95%) for a set of SGT distributions.

4 Uncertainty compensation factor

In order to achieve goals G2 and G7, our methodology incorporates in Eqs. 5, 6, 9, and 10, an 'uncertainty compensation factor' $UCF(N_{\text{eff}})$ which increases when the number of 'effective' returns N_{eff} decreases – where $N_{\text{eff}} = N$ in the historical method and $N_{\text{eff}} = N_{\text{down,up}} = \frac{N}{2} \pm \frac{N \bmod 2}{2} \approx \frac{N}{2}$ under the asymmetrical sigma method. We keep denoting with N the number of returns in X^j.

In this section, we derive an approximative formula for the uncertainty compensation factor compensating the estimation error in the calibrated shocks $CS_{\text{down}}(j)$ and $CS_{\text{up}}(j)$, while it also implicitly covers the generally lower market observability of non-modellable risk factors and approximations in the methodology (such as those introduced to derive the non-linearity coefficient $K_{D^*}(j)$).

We analyse the two drivers of the estimation error to understand the dependence on N_{eff}, return distribution and method, while in the end using the same simple approximation for the uncertainty compensation factor $UCF(N_{\text{eff}})$ for all methods for the calibrated shocks and for all distributions of returns as an acceptable simplification and retaining only the N_{eff}-dependence.

We obtained in our methodology a sample of 10-business-day returns X^j from Eq. 4 in Subsection 2.2.1. By construction, these returns are typically based on overlapping observation windows and hence the calibrated shocks for a risk factor j are typically estimated from autocorrelated returns X^j, with lower autocorrelation for fewer and higher autocorrelation for more (e.g. daily) risk factor observations, see Annex B. While we formulate our results for a return parent distribution X in this subsection, it will be the sample of those 10-business-day returns X^j for NMRF j to which the results are applied and for which figures are shown.

4.1 Uncertainty compensation factor and quantile convergence

This subsection first defines the uncertainty compensation factor formally and relates it to the quantile of the sampling distribution of the calibrated shocks. The estimation error is the difference of an estimator for a sample metric $\widehat{M}$ and the true but unknown estimand value M^*. In our methodology, $\widehat{M}$ is an estimator for the calibrated shocks which are ES estimates given the 10-business day returns, i.e. $\widehat{M} = \widehat{ES}_{\text{down,up}}$ in the historical method (indicated by superscripts "hist") or $\widehat{M} = \widehat{AS}_{\text{down,up}}$ in the asigma method (indicated by superscripts "asigma"), and M^* is the true ES for the true (but generally unknown) 10-day return parent distribution X.

We assume in the following that the estimator converges in probablity to a finite value if the sample size N gets infinitely large, $\lim_{N\to\infty} \widehat{M}(N,X) \overset{\text{prob}}{=} M^{\infty}(X)$. $\widehat{M}(N,X)$ is called a (weakly) consistent estimator, if the infinite sample size value converges in probability to the true value, $\lim_{N\to\infty} \widehat{M}(N,X) \overset{\text{prob}}{=} M^*(X)$, i.e. $M^{\infty}(X) = M^*(X)$. The historical ES estimators are consistent. On the contrary, the asigma estimator is in general not consistent[5]. $\widehat{M}(N,X)$ is a mean (median) unbiased estimator of $M^*(X)$, if its expectation for a certain N is the true value, i.e. $\mathbb{E}\left(\widehat{M}(N,X)\right) = M^*(X)$ (med $\left(\widehat{M}(N,X)\right) = M^*(X)$). Both the historical and asigma estimators are somewhat biased towards zero for finite N (cf. Section B.4).

In our methodology we only have return time-series for one year (cf. Subsection 2.2.1) and N is thus bounded by the number of business days in a year. When we use results for $N \to \infty$ it is therefore in a hypothetical sense of an infinite time-series of daily observations.

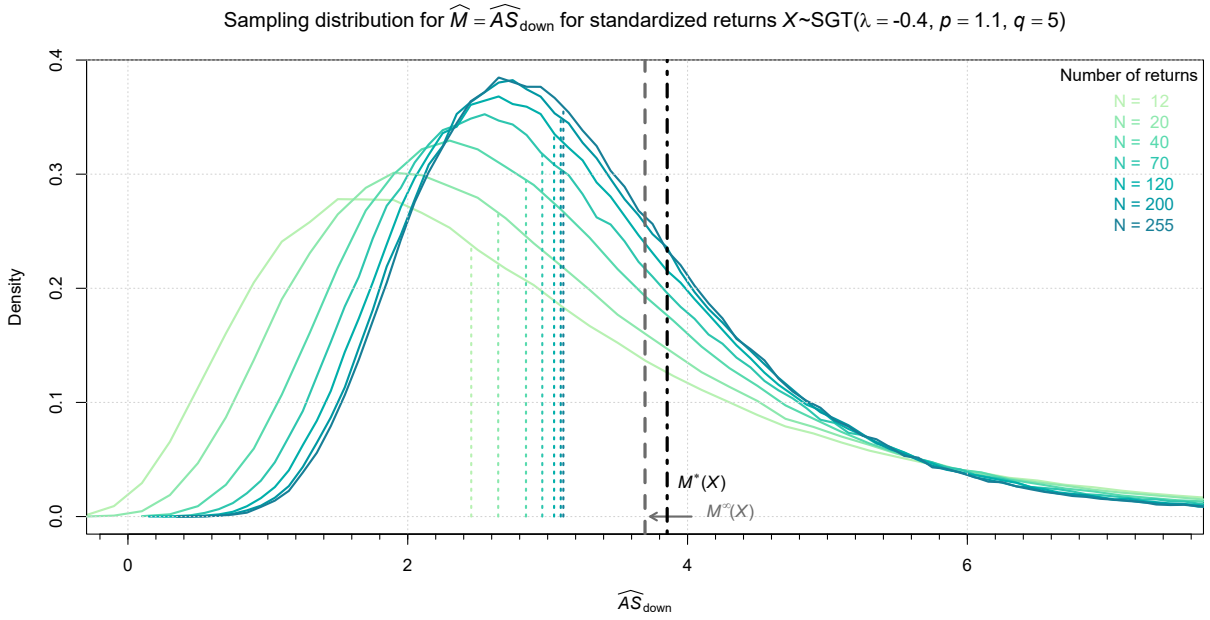
We start with defining the general uncertainty compensation factor $U_{\alpha^{\text{UC}}}\left(\widehat{M}(N,X)\right)$, which ensures that when multiplying the sampling metric $\widehat{M}(N,X)$ calculated from N return realisations of a parent distribution X with it, a certain target confidence level $CL^{\text{UC}} = 1 - \alpha^{\text{UC}}$ for not underestimating the true value $M^*(X)$ is achieved. For the asigma method, $N = N_{\text{down,up}}$ can be as low as six and $U_{\alpha^{\text{UC}}}$ becomes substantial. For ease of notation, the dependence of $\widehat{M}$ on N and X is not always written out.

We set:

$$U_{\alpha^{\text{UC}}}\left(\widehat{M}(N,X)\right) := \frac{M^*(X)}{q_{\alpha^{\text{UC}}}\left(\widehat{M}(N,X)\right)}, \tag{30}$$

[5]Because the constant $C_{\text{ES}}^{\text{asigma}} = 3$ is not exact for all X, as elaborated in Subsection 4.2 and Eq. 39.

Figure 3: Sampling distributions for the asigma method, i.e. using the estimator $\widehat{M} = \widehat{AS}_{\text{down}}$ for different sample sizes N of overlapping returns X drawn from an SGT distribution (cf. Annex B). The (slow) convergence to the normal distribution is visible for the largest N. The black dot-dashed black vertical line is the true value for the ES, M^*, the short-dashed grey vertical line is the large N limit M^∞. The estimator exhibits a bias for all N due to the choice of $C_{\text{ES}}^{\text{asigma}} = 3$. The vertical dotted lines indicate the median values $\text{med}\left(\widehat{M}\right)$ of the sampling distributions for different N, converging to M^∞.



which ensures

$$P\left(U_{\alpha^{\text{UC}}}\,\widehat{M} \le M^*\right) = P\left(\widehat{M} \le q_{\alpha^{\text{UC}}}\left(\widehat{M}\right)\right) = \alpha^{\text{UC}}\,, \tag{31}$$

the last equation being the definition of the quantile for a continuous probability distribution.

$U_{\alpha^{\text{UC}}}\left(\widehat{M}(N,X)\right)$ depends on the tail probability α^{UC} and via the quantile $q_{\alpha^{\text{UC}}}\left(\widehat{M}\right)$ on the estimator $\widehat{M}(N,X)$ used and thus, on the size of the sample N for the estimation and the properties of X. In our methodology we target confidence levels $CL^{\text{UC}} = 1-\alpha^{\text{UC}} = 50\%$ that $\widehat{M}$ does not underestimate the true value, in order to achieve the accuracy goal G6. This means our target is median accuracy, because $q_{\alpha^{\text{UC}}=50\%}(N)$ is the median. While $\alpha^{\text{UC}} = 50\%$ is our focus, we first develop some equations for the uncertainty compensation factor for any tail probility α^{UC} and specialize to $\alpha^{\text{UC}} = 50\%$ when necessary to obtain concrete numerical values.

As we are interested in the N-dependence of $U_{\alpha^{\text{UC}}}\left(\widehat{M}(N,X)\right)$, we need to look at the sampling distributions of the calibrated shocks in more detail and investigate the quantile convergence, i.e. the N dependence of $q_{\alpha^{\text{UC}}}\left(\widehat{M}\right)$ and in particular the convergence of the median (dotted vertical lines in Figure 3).

To get a first impression of the sampling distributions, we first turn to sampling distributions in the i.i.d. case: for continuous distributions with finite variance, the sample distribution of $\widehat{ES}_{\text{down,up}}$ from i.i.d. samples converges in $\sqrt{N}$ to the normal distribution in the large sample limit [MH05, BJPZ08]. Note that because the order statistics $X_{(i)}$ in the tail are highly dependent, this convergence does not follow immediately from the central limit theorem.

We are not aware of comparable results for the asigma estimator $\widehat{AS}_{\text{down,up}}$, while our simulations with i.i.d. returns show a similar convergence to the normal distribution. Its first term $\widehat{\mu}_{X^j_{\text{down,up}}}$ is the ES at the 50% tail probability, converging to a normal distribution. For the second term $\widehat{\sigma}_{X^j_{\text{down}}}$ of the asigma estimator we could argue that the returns below resp. above the median are identical and only weakly dependent when N gets larger. For i.i.d. normal returns $\widehat{\sigma}_{X^j_{\text{down}}}$ would follow a χ-distribution which in turn converges to a normal

distribution. This is a non-rigorous reasoning for why $\widehat{AS}_{\text{down,up}}$ should show a similar convergence behavior as $\widehat{ES}_{\text{down,up}}$.

Our setting is however not i.i.d., because returns are autocorrelated and for return samples of size $12 \leq N \leq 255$ the large N limit is not reached in any case. We present some numerical results from simulations of overlapping rescaled nearest to 10-day returns X from a family of SGT distributions as described in Annex B in the following.

Figure 3 illustrates the sample distributions $\widehat{M}(N, X)$ for some N in the asigma method, i.e. $\widehat{M} = \widehat{AS}_{\text{down}}$, when the returns X follow a typical skewed and moderately fat-tailed SGT distribution (cf. Section 5). One can see the slow convergence to a GAUSSian distribution. $U_{50\%}\left(\widehat{M}\right) = \frac{M^*}{q_{50\%}\left(\widehat{M}\right)}$ can be read off for different N as the true value M^* for the ES (dot-dashed) divided by the median values $q_{50\%}\left(\widehat{M}\right)$ (dotted). Qualitatively, for all $\widehat{M}(N, X)$, i.e. the asigma (Figure 3) and the historical method (not shown) and for both downward and upward shocks, the sample distributions look similar and have positive skewness (are left leaning). One can see that the sampling uncertainty increases as N decreases. Yet, even for $N = 12$, the asigma method still delivers acceptable results, using only six values below or above the median!

Motivated by the i.i.d. convergence of the ES and Figure 3 we make the additional assumption that the estimators' variance converges to zero for infinite sample sizes (which is sufficient for convergence in probability, but not necessary). This assumption is also supported by the quantile convergences shown in Figure 4. Now, for finite mean and variance (both depending on the sampling distribution $\widehat{M}(N, X)$) the difference between a quantile $q_{\alpha^{\text{UC}}}\left(\widehat{M}\right)$ and mean is bounded by the standard deviation times a factor depending only on the tail probability according to Theorem 3.1 of [BB04]:

$$\mu\left(\widehat{M}\right) - \sigma\left(\widehat{M}\right)\sqrt{\frac{1-\alpha^{\text{UC}}}{\alpha^{\text{UC}}}} \leq q_{\alpha^{\text{UC}}}\left(\widehat{M}\right) \leq \mu\left(\widehat{M}\right) + \sigma\left(\widehat{M}\right)\sqrt{\frac{\alpha^{\text{UC}}}{1-\alpha^{\text{UC}}}}\,. \tag{32}$$

This shows that under this convergence assumption all quantiles for non-degenerate probabilities converge to the same limit value,

$$\lim_{N\to\infty} q_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right) = M^{\infty}(X),\ \forall \alpha^{\text{UC}} \in (0, 1)\,. \tag{33}$$

We split $U_{\alpha^{\text{UC}}}$ in two factors, the first depending on the ratio of the true value M^* and the $N \to \infty$ limit M^{∞} not depending on N nor α^{UC}, and the second depending on the convergence of the quantile in N to the limit:

$$\begin{aligned} U_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right) &= \frac{M^*(X)}{q_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right)} = \frac{M^*(X)}{M^{\infty}(X)}\,\frac{M^{\infty}(X)}{q_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right)} \\ &= \left(1 + \frac{M^*(X) - M^{\infty}(X)}{M^{\infty}(X)}\right)\left(1 + \frac{M^{\infty}(X) - q_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right)}{q_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right)}\right)\,. \end{aligned} \tag{34}$$

Introducing the relative deviation for the limit values,

$$\Delta^{\infty}\left(M^{\infty}(X)\right) := \frac{M^*(X) - M^{\infty}(X)}{M^{\infty}(X)}\,, \tag{35}$$

and the relative deviation from convergence in N,

$$\Delta^{\text{conv}}_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right) := \frac{M^{\infty}(X) - q_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right)}{q_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right)}\,, \tag{36}$$

we can write:

$$U_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right) = \left(1 + \Delta^{\infty}\left(M^{\infty}(X)\right)\right)\left(1 + \Delta^{\text{conv}}_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right)\right)\,, \tag{37}$$

where only $\Delta^{\text{conv}}_{\alpha^{\text{UC}}}$ depends on the number of returns in the sample, N. The uncertainty compensation factor can thus be written as unity in the absense of any estimation error plus the sum of three terms,

$$U_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right) = 1 + \Delta^{\infty}\left(M^{\infty}(X)\right) + \Delta^{\text{conv}}_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right) + \Delta^{\infty}\left(M^{\infty}(X)\right)\Delta^{\text{conv}}_{\alpha^{\text{UC}}}\left(\widehat{M}(N, X)\right)\,. \tag{38}$$

The first, $\Delta^{\infty}\left(M^{\infty}(X)\right)$, measures the relative deviation of the true value to the estimator in the large N limit, the second, $\Delta^{\text{conv}}_{\alpha^{\text{UC}}}\left(\widehat{M}(N,X)\right)$, measure the sampling error due to the quantile converging in N, and the third is the product of the two former. Equation 38 is valid for any estimator $\widehat{M}$ with vanishing variance and any confidence level.

4.2 Approximation of the uncertainty compensation factor

We now use the properties of our estimators in the historical and asigma method obtain a simpler approximation. The estimators of the historical method, $\widehat{M} = \widehat{ES}_{\text{down,up}}$ is consistent and the median converges to the true value, thus $\Delta^{\infty,\text{hist}}(X) = 0$. In the asigma method, the constant $C^{\text{asigma}}_{\text{ES}} = 3$ is not exact, but only a reasonable approximation for a wide range of SGT distributions as explained in Subsection 5.2.1 and visualized in Figure 5. The relative deviation the true ES from the ES obtained in the asigma method per Eq. 47 for the SGT distributions X analysed (cf. Section 5) is

$$-0.12 \lesssim \Delta^{\infty,\text{asigma}}(X) \lesssim 0.075\,. \tag{39}$$

Focussing now on the median, $\alpha^{\text{UC}} = 50\%$, we note that the convergence of the median to the large N limit is from below (cf. Figure 3), so that $\Delta^{\text{conv}}_{50\%} > 0\,\forall N$. Simulations of N overlapping rescaled nearest to 10-day returns X from a family of SGT distributions as described in Annex B yields:

$$0.06 \lesssim \Delta^{\text{conv,hist}}_{50\%}(N) \lesssim 0.3,\ 255 \geq N \geq N^{\text{hist}} = 200\,, \tag{40}$$

$$0.03 \lesssim \Delta^{\text{conv,asigma}}_{50\%}(N) \lesssim 0.7,\ 255 \geq N \geq N^{\text{asigma}} = 12\,. \tag{41}$$

The product terms of Eq. 38 in the asigma method are quite small,

$$-1\% \lesssim \Delta^{\infty,\text{asigma}}\Delta^{\text{conv,asigma}}_{50\%}(N) \lesssim 5\%, N \geq N^{\text{asigma}} = 12\,, \tag{42}$$

while vanishing for the historical method.

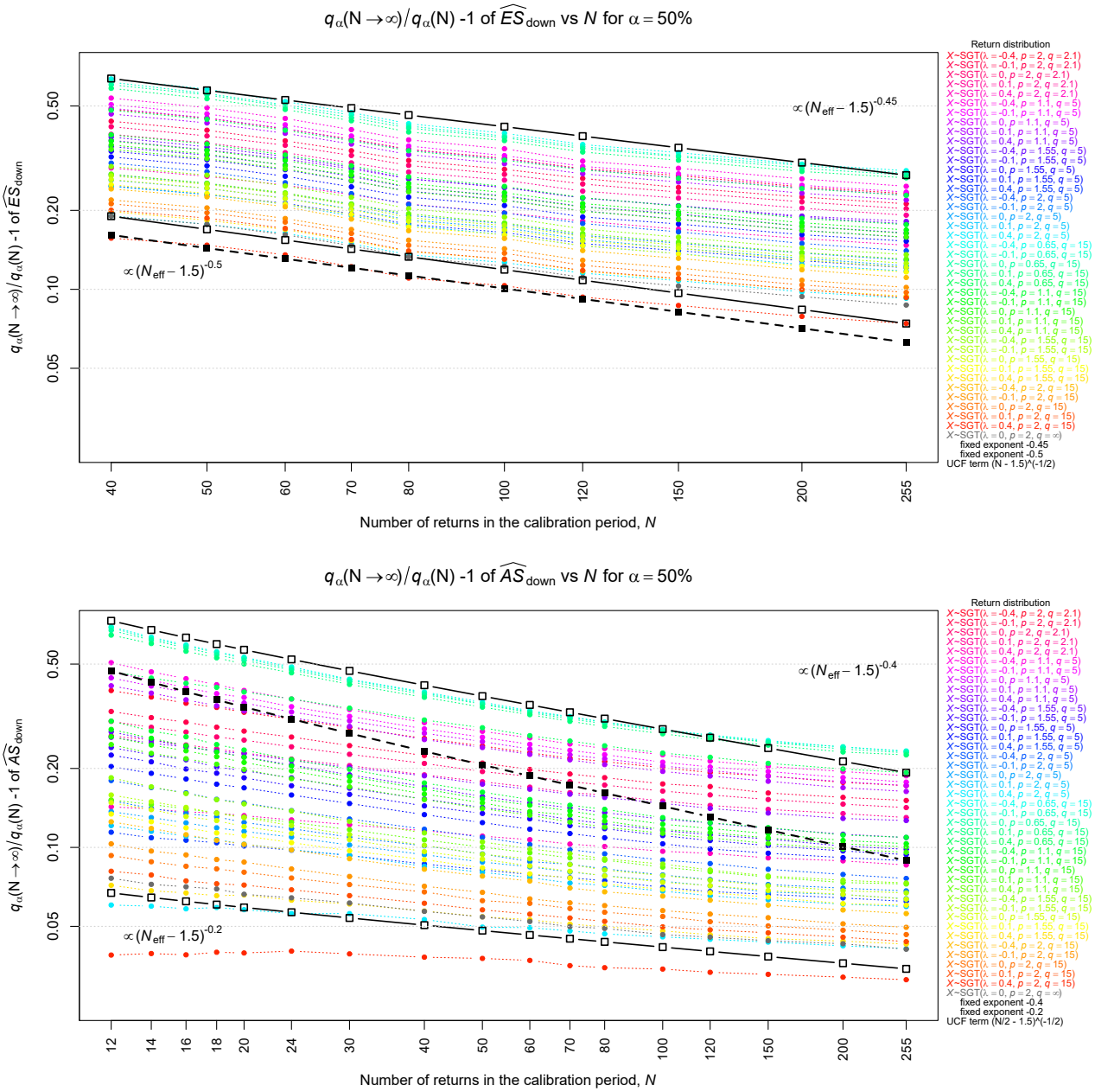
Inspired by the convergence in the i.i.d. case, we make the ansatz for the N-dependence of $\Delta^{\text{conv}}_{50\%}(N)$:

$$\Delta^{\text{conv}}_{50\%}(N) = \frac{C^{\text{UC}}_{50\%,\text{B}}\left(\widehat{M}\right)}{\left(N_{\text{eff}} - \frac{3}{2}\right)^{\xi_{50\%}\left(\widehat{M}\right)}}\,, \tag{43}$$

with a constant $C^{\text{UC}}_{50\%,\text{B}}\left(\widehat{M}\right)$ that depends on the method (i.e. which estimator $\widehat{M}$ is used) and the return distribution, but not on N. N_{eff} is the 'effective' number of returns, with $N_{\text{eff}} = N$ for the historical method and $N_{\text{eff}} = \frac{N}{2} \pm \frac{N \bmod 2}{2}$ for the asigma method. The N-dependence is completely captured by the denominator with an exponent $\xi_{50\%}\left(\widehat{M}\right)$ that depends on the method and the return distribution, which would be $\frac{1}{2}$ for the historical method and i.i.d. returns. Due to the increasingly strong autocorrelation when the number of observations in the stress period increases, the convergence of the ES from overlapping returns X is expected to be somewhat slower than in the i.i.d. case, so that we expect $\xi_{50\%}\left(\widehat{M}\right) \leq \frac{1}{2}$ for the historical method. For the asigma method, we gave some reasoning why a similar behaviour could be expected. Indeed, we found from our simulations of overlapping returns following SGT distributions (cf. Annex B) presented in Figure 4, that Eq. 43 is a good approximation in general. We use $N_{\text{eff}} - \frac{3}{2}$ in the denominator in alignment to the definition of $\widehat{AS}_{\text{down,up}}$ in Eqs. 11 and 12 due to the small number N_{eff}, because the standard deviation is a strong driver of the quantile convergence, cf. Eq. 32. While not in the focus, we note that for other confidence levels the proportionality constant $C^{\text{UC}}_{\alpha^{\text{UC}},\text{B}}\left(\widehat{M}\right)$ increases strongly with increasing confidence level $CL^{\text{UC}} = 1-\alpha^{\text{UC}}$, while the exponents $\xi_{\alpha^{\text{UC}}}\left(\widehat{M}\right)$ vary only slightly.

Turning to the exponents $\xi_{50\%}\left(\widehat{M}\right)$ in Figure 4 with $\Delta^{\text{conv}}_{50\%}\left(\widehat{M}(N,X)\right) \propto \left(N_{\text{eff}} - \frac{3}{2}\right)^{-\xi_{50\%}\left(\widehat{M}\right)}$, we infer the exponents are $\xi^{\text{hist}}_{50\%} \approx 0.5$ for the historical method and $0.2 \leq \xi^{\text{asigma}}_{50\%} \leq 0.4$ for the asigma method. There is variation in $\xi_{50\%}\left(\widehat{M}\right)$ depending on the distribution and the also the method. However, to simplify our

Figure 4: Relative median deviation $\Delta^{\text{conv}}_{50\%}\left(\widehat{M}\right) = \frac{M^{\infty}(X)}{q_{50\%}\left(\widehat{M}(N,X)\right)} - 1$ of the sampling distribution for the estimator $\widehat{M} = \widehat{ES}_{\text{down}}$ used in the historical method (top) and $\widehat{M} = \widehat{AS}_{\text{down}}$ in the asigma method (bottom) versus number of returns N - on both axes logarithmic scales are used. $5 \cdot 10^5$ samples of autocorrelated rescaled nearest to 10-day returns of size N were drawn for each N from various SGT distributions (coloured dots) including a GAUSSian (grey dots). The open squares indicate curves $\propto \left(N_{\text{eff}} - \frac{3}{2}\right)^{-\xi_{50\%}\left(\widehat{M}\right)}$ with exponents $\xi^{\text{hist}}_{50\%} = \{0.45, 0.5\}$ (top) and $\xi^{\text{asigma}}_{50\%} = \{0.4, 0.2\}$ (bottom). In both cases, $\frac{C^{\text{UC}}_{50\%,\text{B}}}{\left(N_{\text{eff}} - \frac{3}{2}\right)^{\xi_{50\%}\left(\widehat{M}\right)}}$ describes the N-dependence reasonably well, with more visible deviations for the asigma method, where $N_{\text{eff}} = N_{\text{down,up}} = \frac{N}{2} \pm \frac{N \bmod 2}{2} \approx \frac{N}{2}$. The closed squares with dashed lines indicate the N-dependent term used in the final expression for the uncertainty compensation factor in Eq. 44, $\left(N_{\text{eff}} - \frac{3}{2}\right)^{-\frac{1}{2}}$, which uses the same prefactor $C^{\text{UC}}_{50\%,\text{B}} = 1$ for both methods. It overall works well for the asigma method, while being somewhat too small for the historical method.



methodology in line with goal G4 we universally use the N_{eff}-dependence $\propto \left(N_{\text{eff}} - \frac{3}{2}\right)^{-\frac{1}{2}}$ for both the historical and asigma method. This choice for the exponent is well matching the exponents found for the historical method and for strongly non-normal returns in the asigma method.

Regarding the constant in the nominator of Eq. 43 we opted to use a single prefactor $C^{\text{UC}}_{50\%,\text{B}} = 1$ for both the historical and the asigma method and regardless of the distribution instead of different prefactors per method and distribution to further simplify our methodology in line with goal G4. Figure 4 shows that this choice leads to values that are somewhat too small for the historical method, while working overall well for a wide range of distributions for the asigma method. We can justify this choice for the historical method by observing that $N \geq N^{\text{hist}} = 200$ and the term $\Delta^{\text{conv}}_{50\%}\left(\widehat{ES}(N,X)\right)$ is thus rather small, so that the additional complexity of a separate constant calibration for the historical method was not considered warranted. As we will see in Subsection 5.2.2, the resulting underestimation is tolerable.

In the exact expression in Eq. 38, we drop the last product term, because it is small as noted in Eq. 42. We combine $1 + \Delta^{\infty}\left(M^{\infty}(X)\right)$ into a single constant C^{UC}_{A}, which in the asigma method neglects the variation for different parent distributions X (cf. Eq. 39) for the sake of simplicity (goal G4). We note that while C^{UC}_{A} not explicitly depends on the quantile probability α^{UC} as it is an infinite N quantity, for a concrete finite N calibration it can implicitly depend on $C^{\text{UC}}_{50\%,\text{B}}$, however. This is discussed for Eq. 49 in Subsection 5.2.2.

Using these simplifications, the uncertainty compensation factor becomes finally:

$$U_{\alpha^{\text{UC}}=50\%}\left(\widehat{M}(N,X)\right) \approx UCF(N_{\text{eff}}) = C^{\text{UC}}_{\text{A}} + \frac{C^{\text{UC}}_{50\%,\text{B}}}{\sqrt{N_{\text{eff}} - \frac{3}{2}}}\,, \tag{44}$$

with $C^{\text{UC}}_{\text{A}} = 0.95$ and $C^{\text{UC}}_{50\%,\text{B}} = 1$. In Subsection 5.2.2, we further motivate the choice of the two constants and show to what extent a confidence level $CL^{\text{UC}} \approx 50\%$ is obtained. Only this simple form $UCF(N_{\text{eff}})$ is used in Eqs. 5, 6, 11, and 12 of our methodology to achieve goals G2 and G7.

5 Motivating the methodology constants using SGT distributions

The methodology and the uncertainty compensation factor depend on some parameters. Those parameters should be set so as to meet our goals. In this section, we discuss and motivate the values at which they were set by investigating return time-series generated from SGT distributions. In Annex A, the definition of the density of the SGT distributions is provided in a parametrisation which is suitable for the R language [R C19]. It also provides explicit expressions for the first four moments, VaR, and ES.

It is a well known stylised fact that return distributions of financial risk factors are often strongly non-normal. A generalisation of the Student-t distribution are the skewed generalised t (SGT) distributions, which describe strongly skewed and heavy tailed financial returns well [The98, HMN10, KM13, AMLSG16, MM17]. While the SGT family is perhaps the most used, there are several other generalisations of the Student-t distribution [LN20].

Studies of SGT distributions for financial risk factors typically use stock market data (e.g. [The98, HMN10, AMLSG16, MM17]), probably because of their good availability. However, goal G1 states that the methodology should be universal. [EBA19, EBA20a, EBA20b] analysed a wide variety of risk factors (almost 50,000 in total) across all risk classes (foreign exchange, interest rates, equity and commodities) used in the regulatory capital market risk models in some of the largest European banks. The analysis in [EBA20a] showed that risk factor returns on the 10-business-day horizon calculated according to our methodology are often strongly skewed and leptokurtic and that the SGT distributions can be used to describe those returns well.

Accordingly, aiming at setting the methodology's parameters, we first provide explicit expressions for the VaR and ES of SGT distributions. Making use of those analytical results, and leveraging on the historical risk factor data analysed in [EBA20a], we motivate the choices made on the values set for $C^{\text{asigma}}_{\text{ES}}$ in Eqs. 11 and 12, ϕ^{asigma} in Eq. 25, C^{UC}_{A} and $C^{\text{UC}}_{50\%,\text{B}}$ in Eq. 44. We also back our choice on the step width h employed to derive the coefficient Γ in Eq. 22, and assess the confidence level prescribed in [EBA20b] for calculating a maximum possible loss when such loss is infinite as described in Subsection 3.

Table 1: Standardised SGT distribution parameters ranges which overall describe the rescaled 10-business-day risk factors returns for all risk factors, and lists of the parameter sets that we analyse in detail if the kurtosis is finite (i.e. $pq > 4$).

Parameter	Lower	Upper	Values analysed
λ (skew)	-0.4	0.4	{ -0.4, -0.1, 0, 0.1, 0.4 }
q (tail thickness)	2.1	∞	{ 2.1, 5, 15, ∞ }
p (peakedness)	0.65	2	{ 0.65, 1.1, 1.55, 2 }

5.1 VaR and ES of SGT distributions

SGT distributions form a location-scale family that encompasses many other distributions e.g. the normal distribution and Student's t distribution. In this paper (see Annex A for more details), we use the notation in the following independent parameters: μ, σ, λ, p, and q. μ and σ are the mean and standard deviation, λ with $-1 < \lambda < 1$ is the skewness parameter, $p > 0$ is the peakedness, and $q > 0$ the tail thickness parameter. In Annex A, we show (Eqs. 70 and 71) that the ES and VaR for SGT distributions can be expressed as:

$$\mathrm{ES}^{\mathrm{SGT}}(\alpha;\mu,\sigma,\lambda,p,q) = -\mu + C_{\mathrm{ES}}^{\mathrm{SGT}}(\alpha;\lambda,p,q)\sigma\,, \tag{45}$$

$$\mathrm{VaR}^{\mathrm{SGT}}(\alpha;\mu,\sigma,\lambda,p,q) = -\mu + C_{\mathrm{VaR}}^{\mathrm{SGT}}(\alpha;\lambda,p,q)\sigma\,. \tag{46}$$

with explicit constants $C_{\mathrm{ES}}^{\mathrm{SGT}}$ and $C_{\mathrm{VaR}}^{\mathrm{SGT}}$ in Eqs. 45 and 46 capturing the terms depending on the shape of the normalised distribution and the tail probability α.

The SGT distributions split at the median have the same location and volatility scale properties and thus, the same ansatz of the ES being a linear function of mean and standard deviation on the support below or above the median *med* is used for the asigma method, while we do not have analytical expressions. We thus write for the asigma method:

$$\mathrm{ES}^{\mathrm{SGT}}(\alpha;\mu,\sigma,\lambda,p,q) = -\mu^{\mathrm{asigma}}_{\mathrm{down/up:}X>/\leq med} + C_{\mathrm{ES}}^{\mathrm{SGT,asigma}}(\alpha;\lambda,p,q)\sigma^{\mathrm{asigma}}_{\mathrm{down/up:}X>/\leq med}\,, \tag{47}$$

We recall that the analogue of the mean $\mu^{\mathrm{asigma}}_{\mathrm{down/up:}X>/\leq med}$ and the analogue of the standard deviation $\sigma^{\mathrm{asigma}}_{\mathrm{down/up:}X>/\leq med}$ are computed on the arguments below or above the median in the asigma method.

When making the simplification that the constant $C_{\mathrm{ES}}^{\mathrm{SGT,asigma}}(\alpha;\lambda,p,q)$ can be set to a single constant $C_{\mathrm{ES}}^{\mathrm{asigma}}$, we get the expression for the asigma method estimator in Eqs. 11 and 12. We show detailed numerical results in Subsection 5.2.1.

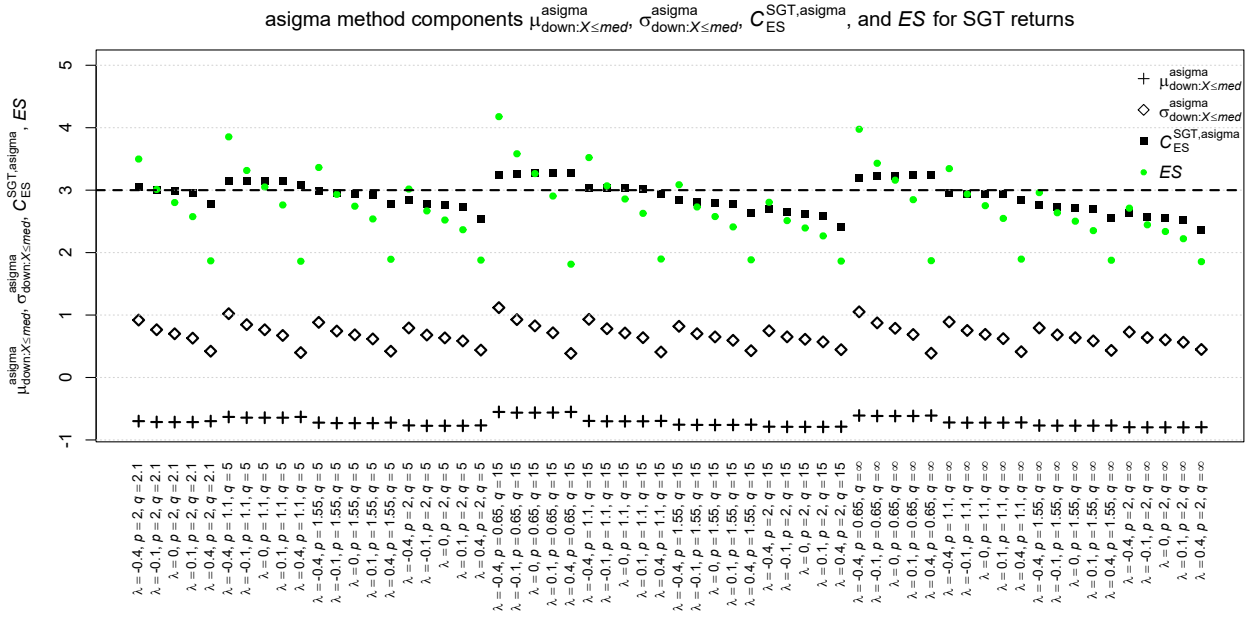
5.2 Motivating the constants of the methodology

As mentioned, the analysis in [EBA20a, EBA20b] showed that SGT distributions describe well the features observed in risk factor returns on the 10-business-day horizon. SGT distributions were fitted to the rescaled nearest to 10-business-day risk factor returns obtained by applying our methodology to the historical data. The SGT parameter ranges which overall describe the 10-business-day risk factor returns well [EBA20a, EBA20b] are stated in Table 1, where we also list the parameter sets we analyse concretely. Only parameter combinations leading to finite first four moments, i. e. $pq > 4$ were considered and without loss of generality the standardised distributions, i. e. $\mu = 0$ and $\sigma = 1$, were used.

In other words, we employ the SGT distributions with the set of parameter values shown in Table 1 as a surrogate of all 10-business-day risk factor returns collected in [EBA19, EBA20a] in all risk categories of Pillar 1 market risk models. Together with the explicit expressions for the moments, ES and VaR, we believe that this is an efficient tool-set to assess questions in risk modelling across all risk categories under the FRTB in general.

The set of SGT distributions are used for the following: (i) the asigma method's constants $C_{\mathrm{ES}}^{\mathrm{SGT,asigma}}$ and ϕ^{asigma} motivating the value chosen in the SSRM methodology (additionally, [EBA20a, EBA20b] checked the

Figure 5: Asigma method components $\mu^{\text{asigma}}_{\text{down}:X\leq med}$ $\sigma^{\text{asigma}}_{\text{down}:X\leq med}$ $C^{\text{SGT,asigma}}_{\text{ES}}$ to obtain the expected shortfall for SGT distributions. $\sigma^{\text{asigma}}_{\text{down}:X\leq med}$ explains a large part of the variation of the ES. Thus, $C^{\text{SGT,asigma}}_{\text{ES}}$ (black squares) shows much less variation than the ES. Finally, $C^{\text{SGT,asigma}}_{\text{ES}} \approx 3$ for many SGT distributions, while being conservative for near normal distributions ($p = 2$, $q = \infty$).



historical values); (ii) the ratio of VaR(α) over ES(α) which motivates 0.8 used for the inner points in Θ (Eq. 13) and the step width $h = 0.2$ (Eq. 22); (iii) simulation analysis of the probability of underestimating the exact ES or asigma method approximation of it when calculating them from a sample of size N in order to determine the constants in the uncertainty compensation factor formula; (iv) ratio of VaR($1 - 99.95\%$) over ES($\alpha = 2.5\%$). The VaR(99.95%) is used as the finite value replacement in case a maximum loss incurring on shocking a risk factor would be infinite theoretically, cf. Subsection 5.2.4.

Table 2 shows the SGT parameter sets of Table 1 with finite first four moments along with the corresponding normalised skewness, excess kurtosis, VaR(α), ES(α), $C^{\text{SGT,asigma}}_{\text{ES}}$, the tail shape parameter $\phi(\alpha)$ (cf. Eq. 20), and $\frac{\text{VaR}(1-99.95\%)}{\text{ES}(\alpha)}$ for the standardised distributions for tail probability $\alpha = 2.5\%$. The value $q = \infty$ was numerically approximated by $q = 10^5$. For standardised SGT distributions $\text{ES}(\alpha) = C^{\text{SGT}}_{\text{ES}}(\alpha; \lambda, p, q)$ (cf. Eq. 70). Only the values for the left tail are reported, because the right tail values are the ones for $-\lambda$. Results were obtained using the analytical expressions of Subsection 5.1 or numerically from large samples of size $N = 10^7$.

We observe in Table 2 that the ES varies significantly depending on λ and to a lesser extent on p and q, it stays well below the HÜRLIMANN bound [Hür02] which implies ES(2.5%) ≤ 6.245 in our case.

5.2.1 Constants $C^{\text{asigma}}_{\text{ES}}$ and ϕ^{asigma} for the asigma method

While the historical method is parameter free except for the loss evaluation grid and the non-linearity correction specification, the constant $C^{\text{asigma}}_{\text{ES}}$ is the main parameter in the asigma method. The key idea for the asigma method was that even for few returns, a mean and standard deviation can still be estimated below and above the median, when an estimate of the ES becomes infeasible. But then those quantities need to be converted according to Eq. 47 to an approximation of the ES, which varies strongly depending on the distributional properties.

Why this works well in practice using a single constant can be inferred from Figure 5, where the components of the calibrated shock according to the asigma method of Eg. 47 for the set of SGT distribution is plotted along with ES(α).

Table 2: SGT parameters, skewness, excess kurtosis, VaR(α), ES(α), $\frac{\text{VaR}(\alpha)}{\text{ES}(\alpha)}$, ES($\alpha$) $= C_{\text{ES}}^{\text{SGT}}$, $C_{\text{ES}}^{\text{SGT,asigma}}$, tail shape parameter ϕ, and $\frac{\text{VaR}(1-99.95\%)}{\text{ES}(\alpha)}$ for standardised SGT distributions ($\mu = 0$, $\sigma = 1$) with $\alpha = 2.5\%$.

λ	p	q	skew.	ex. kurt.	VaR(α)	ES(α)	$\frac{\text{VaR}(\alpha)}{\text{ES}(\alpha)}$	$C_{\text{ES}}^{\text{SGT,asigma}}$	ϕ	$\frac{\text{VaR}(1-99.95\%)}{\text{ES}(\alpha)}$
-0.4	2	2.1	-2.18	58.27	2.34	3.50	0.67	3.05	1.188	2.24
-0.1	2	2.1	-0.61	32.16	2.09	3.01	0.69	3.00	1.165	2.15
0	2	2.1	0	30.00	1.97	2.80	0.70	2.99	1.146	2.10
0.1	2	2.1	0.61	32.16	1.85	2.57	0.72	2.96	1.134	2.05
0.4	2	2.1	2.18	58.27	1.45	1.87	0.77	2.78	1.084	1.83
-0.4	1.1	5	-2.77	27.25	2.49	3.86	0.65	3.16	1.189	2.28
-0.1	1.1	5	-0.83	17.78	2.19	3.31	0.66	3.15	1.175	2.22
0	1.1	5	0	16.89	2.04	3.05	0.67	3.15	1.165	2.20
0.1	1.1	5	0.83	17.78	1.87	2.76	0.68	3.15	1.159	2.16
0.4	1.1	5	2.77	27.25	1.35	1.86	0.73	3.09	1.112	1.98
-0.4	1.55	5	-1.37	5.16	2.42	3.36	0.72	2.99	1.096	1.90
-0.1	1.55	5	-0.39	3.35	2.16	2.93	0.74	2.96	1.083	1.85
0	1.55	5	0	3.20	2.04	2.74	0.75	2.95	1.078	1.82
0.1	1.55	5	0.39	3.35	1.92	2.54	0.76	2.93	1.072	1.79
0.4	1.55	5	1.37	5.16	1.51	1.89	0.80	2.78	1.047	1.64
-0.4	2	5	-0.88	1.78	2.31	3.02	0.77	2.84	1.060	1.71
-0.1	2	5	-0.24	1.06	2.09	2.67	0.78	2.78	1.050	1.65
0	2	5	0	1.00	1.99	2.52	0.79	2.76	1.047	1.63
0.1	2	5	0.24	1.06	1.89	2.37	0.80	2.72	1.042	1.60
0.4	2	5	0.88	1.78	1.58	1.88	0.84	2.54	1.027	1.47
-0.4	0.65	15	-3.79	40.08	2.54	4.18	0.61	3.25	1.229	2.42
-0.1	0.65	15	-1.21	28.24	2.20	3.58	0.62	3.26	1.225	2.40
0	0.65	15	0	26.96	2.03	3.26	0.62	3.27	1.215	2.39
0.1	0.65	15	1.21	28.24	1.83	2.91	0.63	3.27	1.214	2.37
0.4	0.65	15	3.79	40.08	1.22	1.81	0.67	3.27	1.167	2.22
-0.4	1.1	15	-1.57	5.82	2.52	3.52	0.72	3.04	1.088	1.88
-0.1	1.1	15	-0.46	4.02	2.24	3.07	0.73	3.03	1.081	1.84
0	1.1	15	0	3.85	2.10	2.86	0.73	3.03	1.078	1.82
0.1	1.1	15	0.46	4.02	1.95	2.63	0.74	3.02	1.072	1.79
0.4	1.1	15	1.57	5.82	1.48	1.90	0.78	2.94	1.052	1.67
-0.4	1.55	15	-0.95	1.86	2.38	3.09	0.77	2.84	1.051	1.65
-0.1	1.55	15	-0.27	1.20	2.14	2.73	0.78	2.81	1.045	1.61
0	1.55	15	0	1.14	2.04	2.58	0.79	2.80	1.042	1.59
0.1	1.55	15	0.27	1.20	1.93	2.41	0.80	2.78	1.039	1.57
0.4	1.55	15	0.95	1.86	1.57	1.89	0.83	2.63	1.026	1.47
-0.4	2	15	-0.66	0.59	2.25	2.81	0.80	2.70	1.035	1.53
-0.1	2	15	-0.18	0.26	2.05	2.51	0.82	2.65	1.030	1.49
0	2	15	0	0.23	1.97	2.40	0.82	2.62	1.028	1.47
0.1	2	15	0.18	0.26	1.89	2.27	0.83	2.59	1.025	1.45
0.4	2	15	0.66	0.59	1.61	1.86	0.86	2.42	1.016	1.36
-0.4	0.65	∞	-2.51	14.43	2.64	3.97	0.66	3.20	1.134	2.09
-0.1	0.65	∞	-0.78	10.37	2.30	3.43	0.67	3.22	1.129	2.08
0	0.65	∞	0	9.96	2.13	3.16	0.67	3.23	1.128	2.06
0.1	0.65	∞	0.78	10.37	1.94	2.85	0.68	3.24	1.123	2.04
0.4	0.65	∞	2.51	14.43	1.35	1.87	0.72	3.24	1.097	1.93
-0.4	1.1	∞	-1.27	3.40	2.50	3.34	0.75	2.95	1.061	1.72
-0.1	1.1	∞	-0.37	2.37	2.23	2.94	0.76	2.94	1.056	1.69
0	1.1	∞	0	2.28	2.10	2.75	0.76	2.94	1.054	1.67
0.1	1.1	∞	0.37	2.37	1.96	2.55	0.77	2.93	1.051	1.65
0.4	1.1	∞	1.27	3.40	1.53	1.90	0.81	2.85	1.036	1.55
-0.4	1.55	∞	-0.82	1.13	2.34	2.96	0.79	2.76	1.038	1.55
-0.1	1.55	∞	-0.23	0.69	2.12	2.64	0.80	2.73	1.033	1.52
0	1.55	∞	0	0.65	2.02	2.50	0.81	2.72	1.032	1.50
0.1	1.55	∞	0.23	0.69	1.92	2.35	0.82	2.70	1.029	1.48
0.4	1.55	∞	0.82	1.13	1.59	1.88	0.85	2.56	1.020	1.40
-0.4	2	∞	-0.58	0.25	2.22	2.71	0.82	2.63	1.027	1.46
-0.1	2	∞	-0.16	0.02	2.04	2.45	0.83	2.58	1.023	1.42
0	2	∞	0	0.00	1.96	2.34	0.84	2.55	1.021	1.41
0.1	2	∞	0.16	0.02	1.88	2.22	0.85	2.52	1.019	1.39
0.4	2	∞	0.58	0.25	1.62	1.85	0.87	2.36	1.013	1.31

Because only the relevant half of the distribution is considered to cater for asymmetry, $\sigma^{\text{asigma}}_{\text{down:}X\leq med}$ picks up a lot of the variation for different SGT parameters, while the variation of $C^{\text{SGT,asigma}}_{\text{ES}}(\alpha;\lambda,p,q)$ in λ for fixed p,q is small. $\mu^{\text{asigma}}_{\text{down:}X\leq med}$ in turn does not depend much on λ and picks up some variation for different p,q, albeit to a lesser degree. The effect of the variations in $\sigma^{\text{asigma}}_{\text{down:}X\leq med}$ and $\mu^{\text{asigma}}_{\text{down:}X\leq med}$ lead to a more stable value $C^{\text{SGT,asigma}}_{\text{ES}}(\alpha;\lambda,p,q)$, which exhibits much less variation than $\text{ES}(\alpha) = C^{\text{SGT}}_{\text{ES}}(\alpha;\lambda,p,q)$. The values for $C^{\text{SGT,asigma}}_{\text{ES}}(\alpha;\lambda,p,q)$ are indeed around 3 for many SGT distributions. Somewhat lower for near normal distributions ($p=2$, $q=\infty$) and somewhat higher for strongly non-GAUSSian cases.

Considering the relatively modest variation of $C^{\text{SGT,asigma}}_{\text{ES}} \approx 3$, for the sake of universality of goal G1 and simplicity of goal G4 of the methodology, a single calibration constant $C^{\text{asigma}}_{\text{ES}} = 3$ was considered overall appropriate for all risk factors. A calibration based on historical risk factor data confirms this choice [EBA20a].

The tail shape parameter ϕ (cf. Eq. 20) can vary very strongly, mainly driven by excess kurtosis. The non-linearity correction, where ϕ enters, can go in both directions, as the loss function could be convex or concave at the outer points of Θ. Therefore, the constant ϕ^{asigma} approximating ϕ in the asigma method for all risk factors should be targeting a typical value as no conservative choice can be made. The value $\phi^{\text{asigma}} = 1.04$ in the asigma method captures typical moderately non-normal distributions well and has the practical advantage that the term $\frac{\phi^{\text{asigma}}-1}{h^2}$ in the non-linearity correction becomes unity for the loss grid step width $h = 0.2$ (cf. Subsection 2.2.4). To conclude the discussion of constants for the asigma method, we note that our analysis based on the SGT distributions is showing a qualitatively similar picture to the historical return data analysis in [EBA20a, EBA20b]. The SGT based analysis captures very well the ranges of the metrics in which we are interested, while the historical data adds information on the frequency of the occurrence of certain distributional features like skewness. For a particular portfolio of financial instruments, only few risk factors might be driving the losses. Therefore, it is important for reaching goal G1 to ensure that the methodology works well for all sets of risk factors, as opposed to on average only.

5.2.2 Constants for the uncertainty compensation factor

To fulfil goals G2, G4 and G7 for a universally applicable yet simple methodology which increases capital for less data, we simplifed the uncertainty compensation factor by choosing universal constants C^{UC}_{A} and $C^{\text{UC}}_{50\%,\text{B}}$ for both methods and all risk factors (cf. Eq. 44):

$$U_{\alpha^{\text{UC}}=50\%}\left(\widehat{M}(N,X)\right) \approx UCF(N_{\text{eff}}) = C^{\text{UC}}_{\text{A}} + \frac{C^{\text{UC}}_{50\%,\text{B}}}{\sqrt{N_{\text{eff}} - \frac{3}{2}}}\,. \tag{48}$$

recalling $N_{\text{eff}} = N$ for $N \geq N^{\text{hist}} = 200$, and $N_{\text{eff}} = N_{\text{down,up}} = \frac{N}{2} \pm \frac{N \bmod 2}{2}$ for $12 = N^{\text{asigma}} \leq N \leq N^{\text{hist}}$.

To set C^{UC}_{A} we consider the number of business days in a year. The TARGET2 real-time gross settlement system owned and operated by the Eurosystem has typically at least 255 operating days per calendar year, so that $M = 256$ observations and $N = 255$ returns is a reasonable choice for the largest number of returns in a year. In line with the calibration goal G3, the uncertainty compensation factor for a year of daily data was chosen to be close to one,

$$UCF\left(N_{\text{eff}} = 255\right) \approx 1\,, \tag{49}$$

while a bit higher than one to cater for non-sampling related uncertainty stemming from the lower observability of NMRF. Consequently, the constant C^{UC}_{A} needs to be smaller than one, as $\frac{C^{\text{UC}}_{50\%,\text{B}}}{\sqrt{N_{\text{eff}}-\frac{3}{2}}} = 0.0628 > 0$.

We set the constants $C^{\text{UC}}_{\text{A}} = 0.95$ and $C^{\text{UC}}_{50\%,\text{B}} = 1$ in Eqs. 44 and 48 for both estimators $\widehat{M} = \widehat{ES}_{\text{down,up}}$ and $\widehat{M} = \widehat{AS}_{\text{down,up}}$ used for determining the calibrated shocks. For investigating if $UCF(N_{\text{eff}})$ actually provides the desired level of uncertainty compensation, we simulate the probability of underestimation

$$P\left(\widehat{M}\cdot UCF(N_{\text{eff}}) \leq M^* = \text{ES}^{\text{SGT}}(\alpha;\mu=0,\sigma=1,\lambda,p,q)\right) \tag{50}$$

by drawing overlapping 10-day returns from the set of SGT distributions with the parameters of Table 1 which mimic the universe of risk factors, as described in Annex B. Overall, the target was to ensure the probability of

underestimation is $\alpha^{\text{UC}} \approx 50\%$, i.e. the median of the sampled calibrated shocks is close to the true value. If the probability is lower, the uncertainty compensation factor is conservative.

With $C_{\text{A}}^{\text{UC}} = 0.95$ and $C_{50\%,\text{B}}^{\text{UC}} = 1$ for the historical method where $N \geq 200$, $P\left(\widehat{ES} \cdot UCF(N_{\text{eff}}) \leq \text{ES}^{\text{SGT}}\right) \lesssim 70\%$ and for the asigma method where $N \geq 12$, $P\left(\widehat{AS} \cdot UCF(N_{\text{eff}}) \leq \text{ES}^{\text{SGT}}\right) \lesssim 70\%$ for all SGT parameter sets considered. The highest probabilities of underestimation occur for strongly non-GAUSSian distributions, where a higher value for $UCF(N_{\text{eff}})$ would be needed to achieve median unbiasedness. For GAUSSian returns, $P\left(\widehat{ES} \cdot UCF(N_{\text{eff}}) \leq \text{ES}^{\text{SGT}}\right) \lesssim 65\%$ (historical method) and $P\left(\widehat{ES} \cdot UCF(N_{\text{eff}}) \leq \text{ES}^{\text{SGT}}\right) \lesssim 25\%$ (asigma method). Annex 1 of [EBA20b] showed heat-maps of those probabilities of underestimation which were based on simulations using non-overlapping returns. They not change qualitatively when using simulations of overlapping rescaled nearest to 10-day returns from a family of SGT distributions, cf. Section B.4.

Since the asigma method uses a single constant $C_{\text{ES}}^{\text{asigma}} = 3$, the bias term $\Delta_{50\%}^{\infty,\text{asigma}}$ of Eq. 38 which is implicit in C_{A}^{UC} of Eq. 44 cannot be equally matched for all SGT distributions, and $-0.12 \lesssim \Delta^{\infty,\text{asigma}} \lesssim 0.075$ (Eq. 39). Thus, instances where $1 + \Delta^{\infty,\text{asigma}}$ is substantially different from $C_{\text{A}}^{\text{UC}} = 0.95$, exacerbate instances where $\Delta_{50\%}^{\text{conv,asigma}}(N)$ deviates from $\frac{C_{50\%,\text{B}}^{\text{UC}}}{\sqrt{N_{\text{eff}} - \frac{3}{2}}}$ as shown in Figure 4 (lower panel) and hence, stronger variations of the probability of underestimation from the target level occur for the asigma method.

To conclude the discussion, the values $N^{\text{hist}} = 200$ and $N^{\text{asigma}} = 12$ were considered to be appropriate in combination with the uncertainty compensation factor and the chosen constants C_{A}^{UC} and $C_{50\%,\text{B}}^{\text{UC}}$. $N^{\text{hist}} = 200$ was also motivated by the notion that in a year having about 255 business days, 55 business days correspond roughly to the Christmas and year-end period plus the northern hemisphere summer holiday period with reduced trading activity.

5.2.3 Step width h

The non-linearity correction described in Subsection 2.2.4 is applied in order to approximate the ES of the tail losses, which implies an integration from the VaR(97.5%) to infinity. Table 2 indicates that $\frac{\text{VaR}(97.5\%)}{\text{ES}(97.5\%)} \approx 0.8$, so that the step width $h = 0.2$ in Eq. 22 approximates VaR(97.5%) overall well.

5.2.4 Confidence level for the maximum loss

For the analytical ratio $\frac{\text{VaR}(1-99.95\%)}{\text{ES}(\alpha)}$ we can infer from Table 2 that it is about 1.4 for the normal distribution, while growing for non-GAUSSian parameter choices up to 2.4. Thus, the approach laid down in Section 3 setting $SS_{10\text{d}}(j)$ to a VaR measure with a 99.95% confidence level targets a level of conservatism that is set to be 1.4 to 2.4 times higher than the one which would result from the application of the historical calibration of the stress scenario risk measure, ignoring the estimation errors when obtaining the quantities from observed data and the uncertainty compensation factor.

6 The methodology at bucket level

The regulatory bucketing approach in paragraph MAR 31.16(2) of the FRTB introduces a set of regulatory buckets that a bank can use for proving the modellability of risk factors belonging to curves or surfaces. Where banks apply that approach, they are allowed to determine a single stress scenario for all risk factors in the regulatory bucket. The methodology presented at risk factor level is applicable at bucket level with minimal and natural extensions in line with goal G5.

Assumptions A1, A2 and A3 are still valid. Assumption A4 is slightly revised for the regulatory bucket case by assuming that the bank is able to determine the loss that the SSRM portfolio would suffer due to changes in the values of all the M_B risk factors within the bucket:

$$l(\mathbf{x}) = l(x_1, x_2, \ldots, x_{M_B}) = \text{Loss}_{D^*}(r_1^* \oplus x_1, r_2^* \oplus x_2, \ldots, r_{M_B}^* \oplus x_{M_B}), \tag{51}$$

where $\mathbf{x}$ is the vector of shocks $(x_1, x_2, \ldots, x_{M_B})$ applied to the risk factors $r_1, r_2, .., r_{M_B}$ in the bucket B. The application of those shocks leads to the loss $l(\mathbf{x})$.

The steps of the methodology applicable at single risk factor level are adapted as follows to be applicable at bucket level. Analogously to the single risk factor case, a time series of rescaled nearest to 10-business-day returns is determined for each of the M_B risk factors in the bucket B. We denote with N_j the number of returns in the time series for the risk factor j in the bucket B, and with $N_B = \min(N_1, N_2, \ldots, N_{M_B})$. Where $N_B \geq N^{\text{hist}} = 200$, a downward calibrated shock $CS_{\text{down}}(j)$ and an upward calibrated shock $CS_{\text{up}}(j)$ are determined with the historical method separately for each of the M_B risk factors in the bucket B. Where $200 = N^{\text{hist}} > N_B \geq N^{\text{asigma}} = 12$, those shocks are determined with the asymmetrical sigma method.

Our methodology finds the extreme scenario of future shocks, among the shifts following the contour of the calibrated shocks of the bucket's risk factors. We define the β downward contoured shift $\zeta_{\text{down}}(\beta) = (-\beta\, CS_{\text{down}}(1), -\beta\, CS_{\text{down}}(2), \ldots, -\beta\, CS_{\text{down}}(M_B))$ and the β upward contoured shift $\zeta_{\text{up}}(\beta) = (\beta\, CS_{\text{up}}(1), \beta\, CS_{\text{up}}(2), \ldots, \beta\, CS_{\text{up}}(M_B))$ and we obtain a set of scenarios Z by spanning β in the interval $[0, 1]$:

$$Z = \bigcup_{\beta\in[0,1]} \zeta_{\text{down}}(\beta) \;\cup \bigcup_{\beta\in[0,1]} \zeta_{\text{up}}(\beta)\,. \tag{52}$$

We decided to focus our search of the extreme scenario of future shocks among the contoured shifts to reflect that in reality, shifts for risk factors may be larger, for example, in the short end of the bucket along the maturity dimension rather than at its long end. Along the same lines of the reasoning presented for the single risk factor case, it would be desirable to know the behaviour of the loss $l(\zeta)$ for each $\zeta \in Z$; however, mindful of the computational effort, the methodology requires the valuation of the loss only for those scenarios in Z_Θ:

$$Z_\Theta = \bigcup_{\beta\in\{0.8,1\}} \zeta_{\text{down}}(\beta) \;\cup \bigcup_{\beta\in\{0.8,1\}} \zeta_{\text{up}}(\beta)\,. \tag{53}$$

$FS_{D^*}(B)$, the extreme scenario of future shocks for the bucket B, is the scenario ζ in the set Z_Θ leading to the worst loss:

$$FS_{D^*}(B) = \underset{\zeta\in Z_\Theta}{\text{argmax}}\; l(\zeta)\,. \tag{54}$$

Analogously to the determination of the stress scenario risk measure for a single NMRF, for a bucket B:

$$SS_{10\text{d}}(B) = \begin{cases} K_{D^*}(B)\cdot l(FS_{D^*}(B)) & FS_{D^*}(B) \in \{\zeta_{\text{down}}(1); \zeta_{\text{up}}(1)\} \\ l(FS_{D^*}(B)) & FS_{D^*}(B) \in \{\zeta_{\text{down}}(0.8); \zeta_{\text{up}}(0.8)\} \end{cases} \tag{55}$$

where $K_{D^*}(B)$ is the non–linearity coefficient applicable at bucket level analogously to Eq. 26:

$$K_{D^*}(B) = \max\left(K_{\min}, \min\left(\tilde{K}_{D^*}(B), K_{\max}\right)\right)\,, \tag{56}$$

with $K_{\min} = 0.9$, $K_{\max} = 5$ and:

$$\tilde{K}_{D^*}(B) = 1 + \frac{25}{2}\,\frac{l(0.8\, FS_{D^*}(B)) - 2\, l(FS_{D^*}(B)) + l(1.2\, FS_{D^*}(B))}{l(FS_{D^*}(B))}\,\left(\widehat{\phi_B} - 1\right)\,, \tag{57}$$

where for simplicity and robustness the tail-parameter at bucket level $\widehat{\phi_B}$ is the median of the tail parameters $(\widehat{\phi_1}, \widehat{\phi_2}, \ldots, \widehat{\phi_{M_B}})$ (cf. Eq. 25) of the M_B risk factors in bucket B. Also in this case, $SS_{10\text{d}}(B)$ is floored at zero in the rare case that $l(FS_{D^*}(B)) \leq 0$, and the maximum loss approach described in Subsection 5.2.4 can be used by supervisors to address specific cases.

7 Aggregated capital requirements for non-modelled risks

To finally determine the capital charge associated with non-modellable risk factors, $SS_{10\text{d}}(j)$ is rescaled to reflect the NMRF's liquidity horizon LH_j. We obtain the final stress scenario risk measure $SS(j)$ by employing the FRTB square-root-of-time rule used for the liquidity horizon scaling for modellable risk factors:

$$SS(j) = SS_{10\text{d}}(j)\,\sqrt{\frac{LH_j^{\text{floored}}}{10}}\,, \tag{58}$$

where $LH_j^{\text{floored}} = \max(LH_j, 20)$ is the liquidity horizon of the risk factor floored at 20 business days in accordance with paragraph MAR 33.16(1) of the FRTB standards. All risk factors within a bucket have the same liquidity horizon; thus, Eq. 58 is equivalently applicable at bucket level.

The capital requirements corresponding to all non-modellable risks are finally obtained by aggregating the stress scenario risk measure for each risk factor or bucket, as set out in paragraph MAR 33.17 of the FRTB:

$$SES = \sqrt{\sum_{i \in ICSR} SS(i)^2} + \sqrt{\sum_{j \in IER} SS(j)^2} + \sqrt{\left(\rho \sum_{k \in OR} SS(k)\right)^2 + (1-\rho^2) \sum_{k \in OR} SS(k)^2}\,, \tag{59}$$

where $\rho = 0.6$, $ICSR$ and IER are respectively the sets of risk factors reflecting idiosyncratic credit spread risk and idiosyncratic equity risk only, and for which the bank is able to demonstrate that a zero correlation is appropriate. OR is the set of all other risk factors.

8 Summary and conclusions

In this paper, we design a methodology for capitalising the risk of non-modellable risk factors under the FRTB framework. We built it (i) to be applicable to any kind of risk factor by ensuring that its building blocks are not risk-factor dependent; (ii) to capture a wide range of different cases with respect to the number of observations available for a given risk factor by generalising the concept of 10-business-day return and by introducing the asigma method under which an expected shortfall measure is approximated by rescaling a volatility measure of the returns; (iii) to target a level of capitalisation in line with the FRTB standards by setting the constants of the methodology consistently with the level of capital targeted in the standards; (iv) to capture the portfolio losses susceptible to a risk factor accurately while; (v) being efficient by evaluating only a few selected risk factor shocks and capturing non-linear behaviours of losses with only one additional portfolio loss evaluation.

Our methodology reduces the computational effort stemming from a naïve straightforward application of the FRTB standards, i.e. from a direct computation of the expected shortfall of the losses for each non-modellable risk factor, by a factor of around 50 (approximately 250 business days per year compared to 5 selected shocks, including approximating non-linearity in the tails). Furthermore, it provides for a reasonably accurate ES calculation even where few data are available making the methodology universal, not only because it is applicable to any kind of risk factor, but also because it provides results in line with the FRTB standards even when few data are available. While the methodology is applicable in most cases, we outline measures for supervisors to take in cases the methodology may not work well.

As a tool to analyse the statistical properties of risk factors, we use a family of SGT distributions with moments matched to a large set of historical risk factor data [EBA19, EBA20a, EBA20b]. Analysing the values taken by the relevant measures for those SGT distributions, we motivate the values of the methodology's parameters.

We study the sampling error in estimating the expected shortfall with different estimators by simulation of the family of SGT distributions, and design a simple uncertainty compensation factor to capture that uncertainty and reflect it in the capital requirements.

We finally show that the methodology can be naturally extended from the single risk factor level and be applied at the level of a segment of a risk factor curve or surface (the so-called 'regulatory buckets' described in the FRTB).

The methodology with minor differences has been successfully field tested by some of the largest banks in Europe on standardised and real trading portfolios [EBA19]. To our knowledge it is the first universal methodology with approximately controlled accuracy for the capitalisation of non-modellable risk factors publicly described in detail. More results will be available when banks use it to perform capital calculations under the FRTB internal model approach.

A Skewed generalized t distributions

This annex defines the SGT distribution [The98] in the notation used in this paper and provides explitit expressions for value-at-risk and expected shortfall. For SGT distributions there are several parametrisations in literature. For our analyses using the R language [R C19] with the R package "sgt" [Dav15], we use the notation in the following independent parameters: μ, σ, λ, p, and q. μ and σ are the mean and standard deviation, λ with $-1 < \lambda < 1$ is the skewness parameter, $p > 0$ is the peakedness, and $q > 0$ the tail thickness parameter. The SGT distribution family encompasses many other distributions, cf. e.g. [HMN10], the most simple one being the standard normal distribution obtained for $\mu = 0$, $\sigma = 1$, $\lambda = 0$ (no skew), $p = 2$, and $q = \infty$. The same parameter values with a finite q lead to a Student's t distribution with $\nu = 2q$ degrees of freedom.

In this notation, the SGT distribution probability density is

$$f_{\mathrm{SGT}}(x;\mu,\sigma,\lambda,p,q) = \frac{p}{2v\sigma q^{\frac{1}{p}} B(\frac{1}{p},q)\left(\frac{|x-\mu+m|^p}{q(v\sigma)^p(\lambda\,\mathrm{sign}(x-\mu+m)+1)^p}+1\right)^{\frac{1}{p}+q}}\,, \tag{60}$$

and where for $pq > 1$, the helper variables m and v which are not parameters themselves, are defined as

$$m(\sigma,\lambda,p,q) := \frac{2v(\lambda,p,q)\sigma\lambda q^{\frac{1}{p}} B(\frac{2}{p},q-\frac{1}{p})}{B(\frac{1}{p},q)} = \tilde{m}(\lambda,p,q)v(\lambda,p,q)\sigma\,, \tag{61}$$

and

$$v(\lambda,p,q) := q^{-\frac{1}{p}}\left[(3\lambda^2+1)\frac{B(\frac{3}{p},q-\frac{2}{p})}{B(\frac{1}{p},q)} - 4\lambda^2\frac{B(\frac{2}{p},q-\frac{1}{p})^2}{B(\frac{1}{p},q)^2}\right]^{-\frac{1}{2}}. \tag{62}$$

B denotes the beta function. m is a location shift variable, such that we can express the mode as $mode = \mu - m$. With $m(\sigma,\lambda,p,q) = \tilde{m}(\lambda,p,q)v(\lambda,p,q)\sigma$ we separate the dependency on the rescaled standard deviation $v\sigma$, where v is a scaling variable for σ.

We first give the explicit expressions for the first four moments of the SGT distributions [TS16, MM17, The18b] (in different parametrizations) in our chosen parametrisation.

The i^{th} centralized moment $\mu_i := E[(R-E(R))^i]$ is finite for $pq > i$ only. Mean $\mu_1 = \mu$ and variance $\mu_2 = \sigma^2$ are distribution parameters, the third centralized moment is

$$\begin{aligned}\mu_3 = \frac{2q^{\frac{3}{p}}\lambda(v\sigma)^3}{B\left(\frac{1}{p},q\right)^3}\Bigg[&8\lambda^2 B\left(\frac{2}{p},q-\frac{1}{p}\right)^3 - 3(1+3\lambda^2)B\left(\frac{1}{p},q\right)B\left(\frac{2}{p},q-\frac{1}{p}\right)B\left(\frac{3}{p},q-\frac{2}{p}\right)\\ &+2\left(1+\lambda^2\right)B\left(\frac{1}{p},q\right)^2 B\left(\frac{4}{p},q-\frac{3}{p}\right)\Bigg],\end{aligned} \tag{63}$$

and the forth centralized moment is

$$\begin{aligned}\mu_4 = \frac{q^{\frac{4}{p}}(v\sigma)^4}{B\left(\frac{1}{p},q\right)^4}\Bigg[&-48\lambda^4 B\left(\frac{2}{p},q-\frac{1}{p}\right)^4 + 24\lambda^2\left(1+3\lambda^2\right)B\left(\frac{1}{p},q\right)B\left(\frac{2}{p},q-\frac{1}{p}\right)^2 B\left(\frac{3}{p},q-\frac{2}{p}\right)\\ &-32\lambda^2\left(1+\lambda^2\right)B\left(\frac{1}{p},q\right)^2 B\left(\frac{2}{p},q-\frac{1}{p}\right)B\left(\frac{4}{p},q-\frac{3}{p}\right)\\ &+\left(1+10\lambda^2+5\lambda^4\right)B\left(\frac{1}{p},q\right)^3 B\left(\frac{5}{p},q-\frac{4}{p}\right)\Bigg].\end{aligned} \tag{64}$$

The ES and VaR are key quantities in our analysis and can be expressed analytically for the SGT distribution by adding tail probability and shape dependent summands to the mode [The18a](in another parametrization). We give explicit formulae for the ES and VaR in our parametrization considering them useful for other studies.

The correspondence of the variables k_{Th}, n_{Th}, m_{Th}, ϕ_{Th}, and q_{Th} in [The18a] is as follows: $k_{\mathrm{Th}} = p$, $n_{\mathrm{Th}} = pq$, $m_{\mathrm{Th}} = mode$, $\phi_{\mathrm{Th}} = (v\sigma)/p^{\frac{1}{p}}$ and $q_{\mathrm{Th}} = \alpha = 2.5\%$ in our notation.

We define $t^*(\alpha)$ (Eq. (8) of [The18a]) which is controlled by the tail probability α:

$$t^*(\alpha;\lambda) := \frac{2\left|\alpha-\frac{1-\lambda}{2}\right|}{1+\mathrm{sign}\left(\alpha-\frac{1-\lambda}{2}\right)\lambda}\,, \tag{65}$$

Like in the R language, let $\text{pbeta}(x,a,b) = I_x(a,b)$ denote the incomplete beta function (definition 8.17.2 in [DLMF], sometimes also called incomplete beta function ratio [DJ66, DM92]), and $\text{qbeta}(x,a,b) = I_x^{-1}(a,b)$ denote the inverse of the incomplete beta function.

After setting

$$\tilde{W}(\alpha;\lambda,p,q) := \frac{(1-\lambda)^2}{2\alpha}\, q^{\frac{1}{p}}\, \frac{B\left(\frac{2}{p}, q-\frac{1}{p}\right)}{B\left(\frac{1}{p},q\right)} \left\{1 - \text{pbeta}\left[\text{qbeta}\left(t^*(\alpha;\lambda), \frac{1}{p}, q\right), \frac{2}{p}, q-\frac{1}{p}\right]\right\}, \tag{66}$$

Eq. 31 of [The18a] for the expected shortfall (same sign convention as our convention with a positive left tail ES of a distribution centred around zero) becomes

$$\text{ES}^{\text{SGT}}(\alpha) = -mode + \tilde{W}(\alpha;\lambda,p,q)\, v\sigma\,. \tag{67}$$

After dividing by $p^{\frac{1}{p}}$, Eq. 7 of [The18a] becomes:

$$\tilde{w}(\alpha;\lambda,p,q) = \text{sign}\left(\alpha - \frac{1-\lambda}{2}\right) q^{\frac{1}{p}} \left(\frac{\text{qbeta}(t^*(\alpha;\lambda), \frac{1}{p}, q)}{1-\text{qbeta}(t^*(\alpha;\lambda), \frac{1}{p}, q)}\right)^{\frac{1}{p}} \tag{68}$$

and the α-quantile in Eq. 6 of [The18a] is the negative value-at-risk (in our sign convention):

$$\text{VaR}^{\text{SGT}}(\alpha) = -mode - (1+\text{sign}(\tilde{w}(\alpha;\lambda,p,q))\lambda)\, \tilde{w}(\alpha;\lambda,p,q)\, v(\lambda,p,q)\sigma\,. \tag{69}$$

Hence, the ES and VaR are the mode plus the rescaled sigma times a pre-factor depending on the shape of the distribution and the tail probability. Recalling that $mode = \mu - m = \mu - \tilde{m}(\lambda,p,q)v(\lambda,p,q)\sigma$ is a linear function of $v(\lambda,p,q)\sigma$, too, we see that both risk measures are indeed a linear function of the standard deviation, as it should be for a location-scale family of distributions:

$$\text{ES}^{\text{SGT}}(\alpha;\mu,\sigma,\lambda,p,q) = -\mu + \tilde{m}v\sigma + \tilde{W}(\alpha;\lambda,p,q)\, v\sigma = -\mu + C_{\text{ES}}^{\text{SGT}}(\alpha;\lambda,p,q)\sigma\,, \tag{70}$$

and

$$\text{VaR}^{\text{SGT}}(\alpha;\mu,\sigma,\lambda,p,q) = -\mu + \tilde{m}v\sigma - (1+\text{sign}(\tilde{w}(\alpha;\lambda,p,q))\lambda)\, \tilde{w}(\alpha;\lambda,p,q)\, v\sigma = -\mu + C_{\text{VaR}}^{\text{SGT}}(\alpha;\lambda,p,q)\sigma\,. \tag{71}$$

The variables $C_{\text{ES}}^{\text{SGT}}$ and $C_{\text{VaR}}^{\text{SGT}}$ formally capture the terms depending on the shape of the normalized distribution and the tail probability α.

The analytical expressions for ES and VaR were checked against large sample simulated values, which were used in the investigation of the sampling distributions for the uncertainly compensation factor.

B Simulations and properties of the rescaled nearest to 10-business-day returns for a single risk factor

This annex describes the assumptions for the simulation and some properties of the N rescaled nearest to 10-business-day returns X^j obtained from market data at $M = N + 1$ random observation dates in the stress period for a single risk factor j according to our methodology as described by Eq. 4 in Section 2.2.1. To recall, the construction ensures that in case there are no dates with observations exactly 10 business days apart, then the "nearest" date in the sense of smallest relative deviation is used and the return is rescaled to 10 business days using the square root of time scaling. These returns are in general from varying overlapping periods, which introduces temporal autocorrelation in X^j.

B.1 Assumptions and correlations of the rescaled nearest to 10-business-day returns

We made the following assumptions for the simulations:

- The stress period is assumed to have 256 business days as an upper bound for the number of business days in a year as motivated in Section 5.2.2.
- We assume that risk factor observation dates are uniformly distributed throughout the stress period. While this is certainly not exactly true during the course of a given year, it is a reasonable overall approximation for capturing the effect of having less than daily data with random start and ending dates for the individual returns for a risk factor calculated according to our methodology. The relevant stress period can differ for different SSRM portfolios and the market data availability varies for different markets and risk factor categories, thus, a more precise description would need to be portfolio and risk factor specific.
- For the 20 day extension period after the stress period (cf. Section 2.2.1), used for potentially finding a closer to 10-business days ending date for the return calculation for the N starting dates in the stress period, we assume a constant probability $M/256$ to have a market data observation for each day and perform 20 independent Bernoulli trials leading to some additional observations in the 20 day extension period.
- A risk factor j has a fixed number of observation dates M_j in the stress period. Observation dates and returns for risk factor j are independent from all other risk factors k. This means that when simulating a particular number $N = M - 1$ of returns in the stress period for a risk factor, the returns are independent from all other risk factors with the same or other numbers of returns in the stress period.
- A zero mean is assumed for this generic simulation study, in line with the fact that the mean is typically small relative to the standard deviation of risk factors. In the actual estimation of calibrated upward and downward shocks the mean in the stress period is kept (cf. Section 2.2.2).
- The marginal distributions of the rescaled nearest to 10-day returns X^j constructed by Eq. 4 are assumed to follow a family of standardized SGT distributions with distribution parameters as in Table 1. For some analyses, the skewed normal distributions (i.e. $p = 2$, $q = \infty$, and $\lambda \neq 0$) were not analysed, given that results of the SGT distributions with $q = 15$ were quite similar.
- We assume that a risk factor j can be described as a Lévy process where successive incremental changes are random and independent in disjunct time intervals, identical for the same length of a time interval, and having finite variance. This means we assume that the standardized SGT distributed 10-day returns X^j can be written as a sum of daily increments[6]. The only stable distribution with finite variance is

[6] We are not aware of a mathematical proof that SGT distributions are infinitely divisible from which this assumption would follow. One could argue that the SGT distributions could be approximated by generalized hyperbolic distributions, which are infinitely divisible [BNH77].

the normal distribution, which would thus be the only case where the rescaled nearest to 10-day return and the daily increments would have the same distribution up to a standard deviation factor (and up to centering for non-vanishing means).

- For each of the $M-1$ starting dates D_k for calculating returns in the stress period, the corresponding ending date $D_{k_{\mathrm{nn}}}$ for the return minimises the relative deviation of the period from start to ending date (return horizon) to 10 business days according to Eq. 4.a, leading to $N = M-1$ returns, which typically overlap.

  Let $\tilde{X}_k^j$ be the return for risk factor j from date D_k to the date nearest to 10 business days $D_{k_{\mathrm{nn}}}$, spanning a return date period length of $L_k^j = D_{k_{\mathrm{nn}}} - D_k$ business days for all returns $k = 1, \ldots, N$, $N = M-1$. The daily increments Z_i^j have a standard deviation $\frac{1}{\sqrt{10}}$ by construction of $\tilde{X}_k^j$. The rescaled nearest to 10-business-day return based on the return in the date period of length L_k^j according to Eq. 4 can be written as

$$X_k^j = \sqrt{\frac{10}{L_k^j}}\tilde{X}_k^j = \sqrt{\frac{10}{L_k^j}}\sum_{i=D_k}^{D_{k_{\mathrm{nn}}}} Z_i^j \,. \tag{72}$$

  An overlap of return periods leads to temporal autocorrelation and a non-trivial copula for the temporal correlation structure of a risk factor's returns starting at different dates: any two returns with starting dates D_k and $D_l \geq D_k$ have a correlation given by the number of overlapping business days $o_{k,l} = D_{k_{\mathrm{nn}}} - D_l$, i.e. the number of business days which are both in $[D_k, D_{k_{\mathrm{nn}}}]$ and $[D_l, D_{l_{\mathrm{nn}}}]$, divided by the square root of the product of the length of the returns periods,

$$\rho_{k,l}^j = \mathrm{corr}(X_k^j, X_l^j) = \frac{o_{k,l}}{\sqrt{(D_{k_{\mathrm{nn}}} - D_k)(D_{l_{\mathrm{nn}}} - D_l)}} = \frac{D_{k_{\mathrm{nn}}} - D_l}{\sqrt{(D_{k_{\mathrm{nn}}} - D_k)(D_{l_{\mathrm{nn}}} - D_l)}} \,, \tag{73}$$

  because

$$\mathrm{corr}\left(X_k^j, X_l^j\right) = \mathbb{E}(X_k^j \cdot X_l^j) = \mathbb{E}\left(\sqrt{\frac{10}{L_k^j}}\sum_{i=D_k}^{D_{k_{\mathrm{nn}}}} Z_i^j \cdot \sqrt{\frac{10}{L_l^j}}\sum_{i'=D_l}^{D_{l_{\mathrm{nn}}}} Z_{i'}^j\right) = \frac{10}{\sqrt{L_k^j L_l^j}}\sum_{i=D_k}^{D_{k_{\mathrm{nn}}}}\sum_{i'=D_l}^{D_{l_{\mathrm{nn}}}} \mathbb{E}\left(Z_i^j \cdot Z_{i'}^j\right) \,, \tag{74}$$

  from which Eq. 73 follows, because the daily increments are i.i.d. and $\mathbb{E}\left(Z_i^j \cdot Z_{i'}^j\right) = \frac{\delta_{i,i'}}{10}$.

  In case of daily returns, $D_{k_{\mathrm{nn}}} = D_k + 10$ and $D_{k_{\mathrm{nn}}} - D_l = 10 + D_k - D_l$ for $D_l \geq D_k$. For swapped order of dates, $D_k \geq D_l$, the signs are swapped, and thus for any dates order $\rho_{k,l}^j = 1 - \frac{|D_l - D_k|}{10} = 1 - \frac{|l-k|}{10}$, i.e. the autocorrelation is a symmetric Toeplitz-matrix (cf. [Cla21]).

- The copula describing the temporal autocorrelation of the (X_k^j, X_l^j) pairs is assumed to be a multivariate uniform joint distribution function with correlation ρ^j for risk factor j.

B.2 Simulation of the rescaled nearest to 10-business-day returns

To generate the temporally autocorrelated rescaled nearest to 10-day returns for our simulations, the following steps are performed independently for every set of M observation dates leading to $N = M-1$ returns and every set of SGT distribution parameters of Table 1 with finite kurtosis:

- M observation dates in the stress period of 256 days are randomly drawn from a uniform distribution.
- Additional observation dates in the 20 day extension period following the stress period are drawn by 20 Bernoulli trials with probability $M/256$.
- From those observations, the nearest to 10-business-day return periods are determined according to Eq. 4, i.e. for each D_k, $D_{k_{\mathrm{nn}}}$ is determined, and $o_{k,l}$ are determined.
- The autocorrelation matrix ρ^j is computed according to Eq. 73. It is random, except for daily data.

- $u_1^j, \ldots, u_N^j$ are drawn from an N-dimensional multivariate uniform random distribution with correlation matrix ρ^j . Technically, following [Fal99], the R code uses the GAUSSian copula as a close approximation for the multivariate uniform distribution with the same correlation parameters, effectively approximating SPEARMAN's ρ_S with the PEARSON correlation ρ, $\rho_S \approx \frac{6}{\pi}\arcsin\left(\frac{\rho}{2}\right)$.

- The rescaled nearest to 10-day returns timeseries $X^j = X_1^j, \ldots, X_N^j$ are obtained by applying the quantile function of the SGT distribution: $X_k^j = q^{\mathrm{SGT}}(u_k^j), k = 1, \ldots, N$, cf. Eq. 69.

- These steps are repeated at least $5 \cdot 10^4$ times to obtain many independent samples (realisations) for the analyses in this Annex and for the sampling distributions of the calibrated shocks.

  This approach ensures a defined SGT distributed outcome on the 10-day horizon. Other studies worked with defined daily returns (e.g. [SNHG09, DZ16]). Summing non-normal daily returns could lead to more normally distributed longer period returns by the central limit theorem, and the distribution of the sum of daily returns might not be known analytically, except for the GAUSSian or other special cases.

B.3 Some properties of the rescaled nearest to 10-business-day returns

In the case of daily risk factor data ($N = M - 1 = 255$), the return periods are all exactly 10-(business)-days long, $o_{k,l} = 10, \forall k, l$ and the correlation matrix is a TOEPLITZ matrix with the n-th superdiagonal and n-th subdiagonal having the entry $1 - \frac{|n|}{10}$ for $|n| < 10$, and 0 otherwise, where $n = |l - k|$ is the lag between two returns. Such overlapping daily returns have been investigated previously, e.g. in [HB02, SNHG09, Cla21].

In general, the random observation dates lead to returns that span different periods different from 10 days (before rescaling to 10 days), have consequently different overlaps, and thus are not correlated with fixed pairwise correlations, but with randomly distributed correlations. We are not aware that the overlapping return structure resulting from our methodology has been investigated previously, and present some properties in this section.

First, we show the averages of the period length L_k^j (cf. Eq. 72) of the nearest to 10-business-day returns before rescaling in Figure 6. For daily data, all returns span a period of ten days, as it should be. Reducing the number of observations leads to longer return periods, and for 12 returns from 13 observations in the stress period, the average period increases to about 24 business days. As the minimum liquidity horizon for non-modellable risk factors is 20 business days, this result shows that our methodology is suitable for a wide range of numbers of observations in the stress period.

The 20 business days extension period (cf. Section 2.2.1) was meant to reduce over-using the last observations in the stress period, and thereby keeping the returns at the end of the stress period similar to the rest. For the last return X_N^j and second last return X_{N-1}^j (legend "N" and "$N-1$"), only small deviations from the average period length are noticeable, and we can see that the extension works well in this sense.

Second, the correlation structure from Eq. 73 is visualized in Figure 7 for various numbers of returns in the stress period, N, by showing the average autocorrelation of returns X_k^j with lagged $X_{k-n}^j, n = 1, \ldots, 9$, returns. For a small number of risk factor observations in the stress period, i.e. M much smaller than 256, the temporal autocorrelations wane, and conversely, where $M \to 256$ (and $N = M - 1 \to 255$) the regular TOEPLITZ matrix is obtained. As expected, subsequent returns (lag $n = 1$) have the highest autocorrelation, as their date overlap is the largest. It persists even at the lowest investigated number of returns in the stress period. The higher the lag, the lower the correlation. Note that the actual autocorrelation values are randomly distributed around the average for each sample, except for daily data.

B.4 Effects of using rescaled nearest to 10-business-day returns on moment and calibrated shock estimations

The sample estimators for the different quantities in our methodology are not specifically crafted for the overlapping - and thus autocorrelated - rescaled nearest to 10-business-day returns of our methodology, or generally

Figure 6: Average period length of the nearest to 10-business-day returns (before rescaling to 10 days) vs. number of returns in the stress period, N, for simulated data. For daily data, all returns span a period of ten days. The returns span longer times periods as the number of returns in the stress period get smaller, as there are fewer observations and hence the gaps between observations for obtaining the nearest to 10 business days for computing returns get larger. The 20 business days extension at the end of the stress period foreseen in our methodology keeps the period of the last return X_N^j (legend "N") and second last return X_{N-1}^j (legend "$N-1$") close to the average (legend "all"), while small deviations are visible.

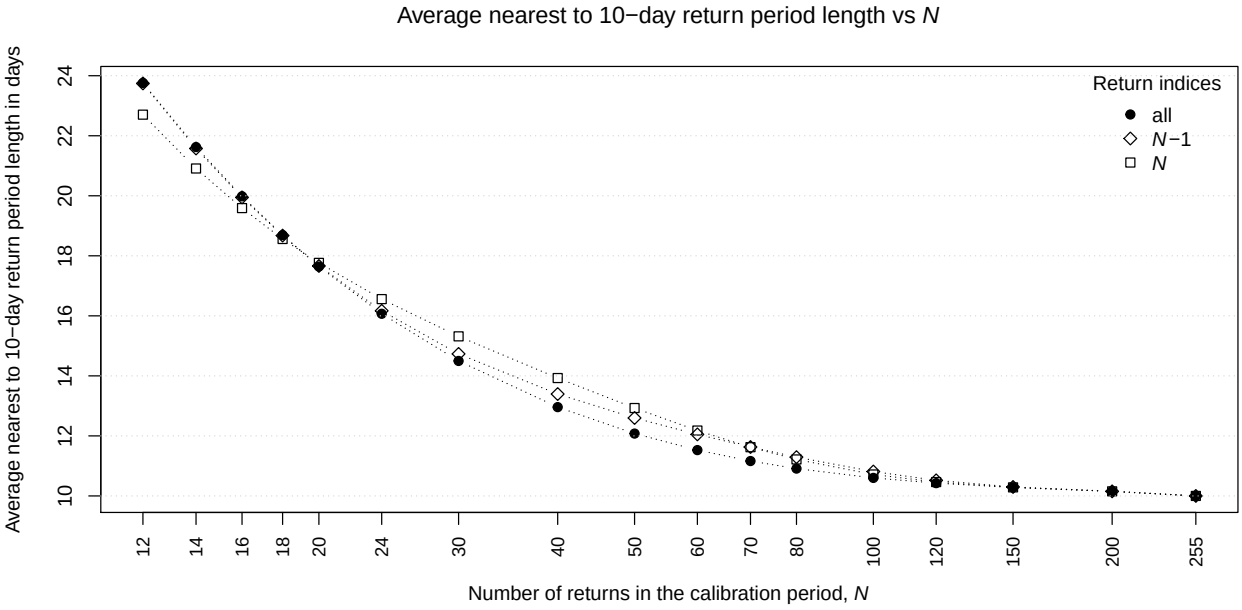
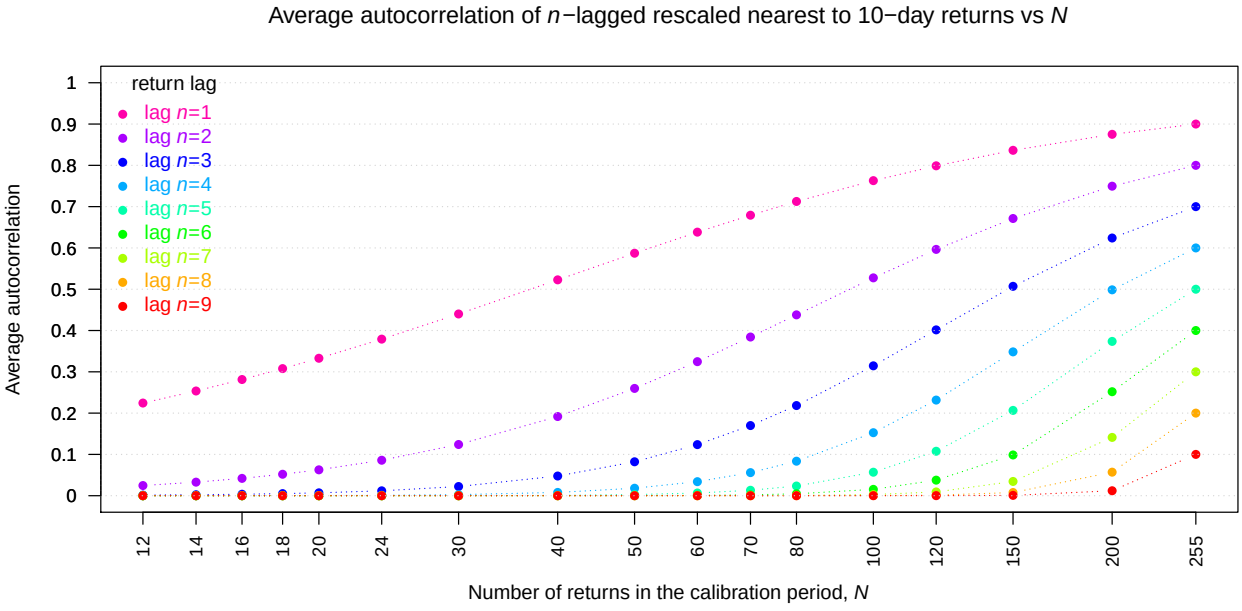


Figure 7: Average temporal autocorrelation of n-lagged rescaled nearest to 10-business-day returns (cf. Eq. 73) vs. number of returns in the stress period, N, for simulated data. The longer the lag and the fewer returns in the stress period, the lower the average correlation. In the simulation, the returns are obtained from randomly generated dates, so that the individual correlations are fluctuating around the average shown. For daily data $N = 255$, the temporal n-lagged correlations are $\frac{1}{n}, 1 \leq n < 10$ and 0 for $n \geq 10$.



speaking, correlated samples. It is known that estimators that are (fairly) unbiased for independent returns (i.e. the usual ones [JG98]) are biased for overlapping returns. We provide some comments on effects observed, as they could be useful for interpreting some results in [EBA20a, EBA20b].

Regarding moment estimators, analytical results are available for the effect of using overlapping returns from daily data for the estimation of variance [BBFK02, SNHG09] and thereby for the standard deviation, as well as for the skewness [TF18]. To our knowledge, there are no analogous analytical results for (excess) kurtosis. Those analytical results would however not be accurate in the general case of our methodology for less than daily data in the risk factor time series, due to the different correlation structure. From our simulations we found that the effect of estimating quantities with the samplte estimators typically used for i.i.d. samples from the rescaled nearest to 10-business-day returns for daily data in a year with $N = 255$ returns leads to a relative mean bias vs. the analytical values for the standard deviation of -2% to -7%, for skewness of -10% to -50% and for the excess kurtosis of -40% to -90%. For skewness and kurtosis the sample estimators G1 and G2 of [JG98] were used. A negative sign means underestimation relative to the analytical value. The ranges correspond to the ranges of values observed for the different SGT distributions analysed. The bias depends strongly on the order of the moment and also on the distribution, reaching the most negative values for the excess kurtosis of strongly non-normal distributions. The bias becomes more negative when fewer observations are available, as expected.

Skewness and excess kurtosis sample estimates from rescaled nearest to 10-business-day returns computed according to our methodology from historical financial risk factor time-series using the typical sample estimators [JG98] were shown in [EBA20a, EBA20b] and compared to the analytical skewness and excess kurtosis of SGT distributions. Using sample estimators which are not compensating the effects of overlapping returns leads to skewness and excess kurtosis estimates in [EBA20a, EBA20b] underestimating the true (but unknown) values.

Regarding VaR and ES, the effect of overlapping returns from daily data on the VaR was investigated numerically in [SNHG09]. [DZ16] uses extreme value theory (EVT) to derive some asymptotic properties of the VaR and ES for overlapping returns from daily data and analysed the mean bias of VaR and ES estimated from those overlapping returns with simulations, which were found to be in the order of some percent for 300 days of 10-day overlapping returns for i.i.d. daily data. Calibrated shocks obtained with the historical method here correspond closely to the ES estimator of [DZ16].

The median bias of calibrated shock estimates in the historical and asigma method estimated from the rescaled nearest to 10-business-day returns X^j vs. the theoretical ES was introduced for the uncertainty compensation factor in Section 4, where the focus was understanding the main drivers and a arriving at a simple generic expression.

For daily data in a year with $N = 255$ returns, the relative median bias (i.e. exact uncertainty compensation factor) in the historical method is -7% to -22%, and in the asigma method is +10% to -24%. A negative sign means underestimation relative to the analytical value. The ranges correspond to the ranges of values observed for the different SGT distributions analysed. The bias depends on the distribution and is most negative for strongly non-normal distributions, like for the moments. Similarly, the biases become more negative when fewer observations are available, too. The sign can be positive where the asigma method yields a higher value than the theoretical ES, which can happen for SGT parameter sets if the constant $C_{\mathrm{ES}}^{\mathrm{asigma}} = 3$ is higher than the exact value for the specific SGT distribution in question would be, i.e. $3 > C_{\mathrm{ES}}^{\mathrm{SGT,asigma}}(\alpha; \lambda, p, q)$ (cf. Eq. 47). The mean biases are somewhat less negative than the median biases and display qualitatively the same dependence on SGT distribution parameters and N.

The median biases for calibrated shocks based on independent 10-day returns are less negative than the biases for rescaled nearest to 10-business-day returns for the same number of returns in the stress period, N, while exhibiting qualitatively the same dependence on SGT distribution parameters and N. This likely explains why the heat-maps of the probability of underestimation of the calibrated shocks in Figures 21 and 22 of [EBA20b] (where independent 10-day returns were used) do not change qualitatively when using rescaled nearest to 10-business-day returns described here.

References

[AMLSG16] Pilar Abad, Sonia Muela, Carmen Lopez, and Miguel Sánchez-Granero. Evaluating the performance of the skewed distributions to forecast value-at-risk in the global financial crisis. *The Journal of Risk*, 18:1 – 28, May 2016.

[AT02] Carlo Acerbi and Dirk Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, Jul 2002.

[Bas19] Basel Committee on Banking Supervision. Minimum capital requirements for market risk. Standard d457, Bank for International Settlements, January 2019. https://www.bis.org/bcbs/publ/d457.pdf.

[BB04] Subhash C. Bagui and Dulal K. Bhaumik. Glimpses of inequalities in probability and statistics. *International Journal of Statistical Sciences*, 3:9–15, 2004.

[BBFK02] Pauline Bod, David Blitz, Philip Hans Franses, and Roy Kluitman. An unbiased variance estimator for overlapping returns. *Applied Financial Economics*, 12(3):155–158, 2002. https://doi.org/10.1080/09603100110090127.

[Ben03] Joseph Bennish. A proof of Jensen's inequality. *Missouri J. Math. Sci.*, 15(1):33–35, 02 2003.

[BJPZ08] Vytaras Brazauskas, Bruce Jones, Madan Puri, and Ricardas Zitikis. Estimating conditional tail expectation with actuarial applications in view. *Journal of Statistical Planning and Inference*, 138:3590–3604, 11 2008.

[BNH77] Ole Barndorff-Nielsen and Christian Halgreen. Infinite divisibility of the hyperbolic and generalized inverse gaussian distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 38:309–311, 1977.

[Cla21] Michael A. Clayton. Parameter estimation from overlapping observations. *SSRN Electronic Journal*, 2021. https://ssrn.com/abstract=2968896.

[Dav15] Carter Davis. *sgt: Skewed Generalized T Distribution Tree*, 2015. R package version 2.0.

[DJ66] A. R. DiDonato and M. P. Jarnagin. A method for computing the incomplete beta function ratio. NWL report No. 1949, U. S. Naval Weapons Laboratory, Dahlgren, Virginia, October 1966.

[DLMF] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.27 of 2020-06-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[DM92] Armido R. Didonato and Alfred H. Morris. Algorithm 708: Significant digit computation of the incomplete beta function ratios. *ACM Trans. Math. Softw.*, 18(3):360–373, September 1992.

[DZ16] Jón Daníelsson and Chen Zhou. Why risk is so hard to measure. Working Paper No. 494, De Nederlandsche Bank, 2016. Available at SSRN: https://ssrn.com/abstract=2597563.

[EBA19] EBA. EBA data collection on non-modellable risk factors. Data collection, European Banking Authority, Paris, France, June 2019. https://eba.europa.eu/eba-publishes-its-roadmap-for-the-new-market-and-counterparty-credit-risk-approaches-and-launches-consultation-on-technical-standards-on-the-ima-under.

[EBA20a] EBA. Draft Regulatory Technical Standards on the calculation of the stress scenario risk measure under Article 325bk(3) of Regulation (EU) No 575/2013. Consultation Paper EBA/CP/2020/10, European Banking Authority, Paris, France, June 2020. https://eba.europa.eu/calendar/consultation-paper-draft-regulatory-technical-standards-calculation-stress-scenario-risk.

[EBA20b] EBA. Final Draft Regulatory Technical Standards on the calculation of the stress scenario risk measure under Article 325bk(3) of Regulation (EU) No 575/2013 (Capital Requirements Regulation 2 - CRR2). Final Draft RTS EBA/RTS/2020/12, European Banking Authority, Paris, France, December 2020. https://www.eba.europa.eu/sites/default/documents/files/document_library/Publications/Draft Technical Standards/2020/RTS/961600/Final draft RTS on the calculation of stress scenario risk measure.pdf.

[EBA21] EBA. Supervisory benchmarking exercises. Regulatory activity overview, European Banking Authority, Paris, France, 2021. https://www.eba.europa.eu/regulation-and-policy/supervisory-benchmarking-exercises.

[Fal99] Michael Falk. A simple approach to the generation of uniformly distributed random variables with prescribed correlations. *Communications in Statistics - Simulation and Computation*, 28(3):785–791, 1999.

[Far93] Stanley J. Farlow. *Partial Differential Equations for Scientists and Engineers*. Dover Publications, 1993.

[GT71] John Gurland and Ram C. Tripathi. A simple approximation for unbiased estimation of the standard deviation. *The American Statistician*, 25(4):30–32, 1971. Equation 6.

[HB02] Ardian Harri and B. Wade Brorsen. The overlapping data problem. *SSRN Electronic Journal*, 2002. https://ssrn.com/abstract=76460.

[HMN10] Christian Hansen, James B. McDonald, and Whitney K. Newey. Instrumental variables estimation with flexible distributions. *Journal of Business and Economic Statistics*, 28(1):13–25, 2010.

[Hür02] Werner Hürlimann. Analytical bounds for two value-at-risk functionals. *ASTIN Bulletin*, 32(2):235–265, 2002. Theorem 3.1.

[Int20] Intesa Sanpaolo, Financial and market risk department. Response to EBA Consultation on Regulatory Technical Standards on the calculation of the stress scenario risk measure for NMRF. Public consultation response letter, Intesa Sanpaolo S.p.A., Turin, Italy, September 2020. https://www.eba.europa.eu/node/102208/submission/96851.

[Jen06] J. L. W. V. Jensen. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Math.*, 30:175–193, 1906.

[JG98] D. N. Joanes and C. A. Gill. Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 47(1):183–189, 1998.

[KM13] Sean C. Kerman and James B. McDonald. Skewness–kurtosis bounds for the skewed generalized t and related distributions. *Statistics & Probability Letters*, 83(9):2129 – 2134, 2013.

[LN20] Riu Li and Saralees Nadarajah. A review of Student's t distribution and its generalizations. *Empirical Economics*, 58:1461–1490, March 2020.

[MBP17] Adolfo Montoro, Tim Becker, and Lars Popken. Identification and capitalisation of non-modellable risk factors. risk.net, January 2017.

[MH05] B. John Manistre and G. Hancock. Variance of the CTE estimator. *North American Actuarial Journal*, 9:129 – 156, 2005.

[MM17] James B. McDonald and Richard A. Michelfelder. Partially adaptive and robust estimation of asset models: Accommodating skewness and kurtosis in returns. *Journal of Mathematical Finance*, 7:219–237, 2017.

[NZC14] Saralees Nadarajah, Bo Zhang, and Stephen Chan. Estimation methods for expected shortfall. *Quantitative Finance*, 14(2):271–291, 2014.

[PC13] European Parliament and Council. Regulation (EU) No 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms and amending Regulation (EU) No 648/2012 (Capital Requirements Regulation) as amended. *Official Journal of the European Union*, L 176, June 2013. http://data.europa.eu/eli/reg/2013/575/oj.

[PC19] European Parliament and Council. Regulation (EU) 2019/876 of the European Parliament and of the Council of 20 May 2019 amending Regulation (EU) No 575/2013 as regards the leverage ratio, the net stable funding ratio, requirements for own funds and eligible liabilities, counterparty credit risk, market risk, exposures to central counterparties, exposures to collective investment undertakings, large exposures, reporting and disclosure requirements, and Regulation (EU) No 648/2012 (Capital Requirements Regulation 2) as amended. *Official Journal of the European Union*, L 150, May 2019. http://data.europa.eu/eli/reg/2019/876/oj.

[R C19] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019.

[SNHG09] Heng Sun, Izzy Nelken, Guowen Han, and Jiping Guo. Error of VaR by overlapping intervals. *Risk*, 22(3):86–91, March 2009.

[TF18] Stephen Taylor and Ming Fang. Unbiased weighted variance and skewness estimators for overlapping returns. *Swiss Journal of Economics and Statistics*, 154(1):1–8, December 2018.

[The98] Panayiotis Theodossiou. Financial data and the skewed generalized t distribution. *Management Science*, 44:1650–1661, Dec 1998.

[The18a] Panayiotis Theodossiou. Risk measures for investment values and returns based on skewed-heavy tailed distributions: Analytical derivations and comparison. *SSRN Electronic Journal*, May 2018. https://ssrn.com/abstract=3194196.

[The18b] Panayiotis Theodossiou. Skewed generalized t and nested probability distributions: Specification and moments. *SSRN Electronic Journal*, May 2018. https://ssrn.com/abstract=3178548.

[TS16] Panayiotis Theodossiou and Christos S. Savva. Skewness and the relation between risk and return. *Management Science*, 62(6):1598–1609, 2016.

ACKNOWLEDGEMENTS

The methodology in the paper has benefited from discussions with several colleagues and different stakeholders. In particular, we would like to thank Svenja Israel, Philip Rau and Despo Malikkidou, all members of the EBA sub-group on Market Risk and its chair Stéphane Boivin, for their precious work in building up the methodology, as well as Lars Jul Overby and an anonymous referee for their comments. Any remaining errors are our own.

The opinions expressed are those of the authors and are not the responsibility of the institutions.

Martin Aichele:
Team Lead Banking Supervision, Directorate General Onsite & Internal Model Inspections, European Central Bank

Marco Giovanni Crotti:
Policy Expert, Directorate Prudential Regulation and Supervisory Policy, European Banking Authority

Benedikt Rehle*:
Parliamentary Assistant, European Parliament
(*) At the time of developing the methodology described in this paper, Mr. Rehle served as on-site inspector at Deutsche Bundesbank and was a representative at the EBA sub-group on Market Risk

EUROPEAN BANKING AUTHORITY

20 avenue André Prothin CS 30154
92927 Paris La Défense CEDEX, France

Tel. +33 1 86 52 70 00
E-mail: info@eba.europa.eu

https://eba.europa.eu/

ISBN 978-92-9245-747-1 doi:10.2853/859367
ISSN 2599-7831 DZ-AH-21-004-EN-N