Differential Sensitivity in Discontinuous Models

Silvana M. Pesenti
Department of Statistical Sciences, University of Toronto, silvana.pesenti@utoronto.ca

Pietro Millossovich
Bayes Business School (formerly Cass), City, University of London,
DEAMS, University of Trieste

Andreas Tsanakas
Bayes Business School (formerly Cass), City, University of London

11 October 2023

Differential sensitivity measures provide valuable tools for interpreting complex computational models used in applications ranging from simulation to algorithmic prediction. Taking the derivative of the model output in direction of a model parameter can reveal input-output relations and the relative importance of model parameters and input variables. Nonetheless, it is unclear how such derivatives should be taken when the model function has discontinuities and/or input variables are discrete. We present a general framework for addressing such problems, considering derivatives of quantile-based output risk measures, with respect to distortions to random input variables (risk factors), which impact the model output through step-functions. We prove that, subject to weak technical conditions, the derivatives are well-defined and derive the corresponding formulas. We apply our results to the sensitivity analysis of compound risk models and to a numerical study of reinsurance credit risk in a multi-line insurance portfolio.

Key words: Sensitivity analysis, importance measurement, differential sensitivity measures, simulation, risk measures, credit risk.

---

1. Introduction

The interpretability of complex computational models is of fundamental importance across areas of applications, with sensitivity analysis providing tools for understanding the importance of risk factors, their interactions and their impact on a model's output (Saltelli et al. 2008, Borgonovo and Plischke 2016, Razavi et al. 2021, Fissler and Pesenti 2023). In recent years, the field received renewed impetus by the widespread adoption of machine learning and artificial intelligence models for prediction tasks, which are usually opaque and thus require additional work to illuminate input/output relationships. Contributions in this field range from the development of general model-agnostic model interpretation procedures (Ribeiro et al. 2016, Borgonovo et al. 2023), to those tailored to a class of models, such as tree ensembles (Lundberg et al. 2018) and neural networks (Merz et al. 2022), or to specific applications, such as image recognition (Chen et al. 2019) and credit scoring (Chen et al. 2023). Furthermore, the interest in model interpretation is amplified by the requirement for models' behaviour to be fair, in the sense that it does not generate discriminatory impacts on protected groups (Frees and Huang 2021, Kozodoi et al. 2022, Lindholm et al. 2022) – such concerns have generated further research at the interface of sensitivity analysis and algorithmic fairness (Bénesse et al. 2022, Hiabu et al. 2023).

As part of sensitivity analysis, metrics are often used to assess the importance of model inputs. A broad class of such metrics is that of differential sensitivity measures, which rely on derivatives of (a statistical functional of) the model output, in the direction of a perturbation of a (random) input factor (Antoniano-Villalobos et al. 2018). Recent advances in sensitivity analysis pertain to perturbing quantile-based risk measures of the model output (Tsanakas and Millossovich 2016, Browne et al. 2017, Pesenti et al. 2021, Merz et al. 2022). In that context, fundamental technical requirements for differential sensitivity measures include differentiability of the model function and Lipschitz continuity of the model output in the perturbation (Broadie and Glasserman 1996, Hong 2009, Hong and Liu 2009). These requirements are stringent, as many computational models map input factors to outputs in a discontinuous manner; examples include credit risk models (Chen and Glasserman 2008), financial derivatives and insurance contracts (Albrecher et al. 2017), and tree-based predictive models (Chen and Guestrin 2016).

In this work, we overcome such strong conditions and derive, under rather mild assumptions, formulas for differential quantile-based sensitivity measures, in models where the input-output relationship contains step functions. This is a general setting, since many functions with a finite number of jump discontinuity points can be written via a sum of step functions. We focus on the two most common quantile-based risk measures, Value-at-Risk (VaR) and Expected Shortfall (ES), although the expressions can be generalised for the broader case of distortion risk measures and rank dependent expected utilities. We consider two types of differential sensitivity measure, marginal sensitivities and cascade sensitivities. The marginal sensitivity quantifies an input factor's sole effect on a model output's risk measure (Hong 2009, Tsanakas and Millossovich 2016). In contrast, in the cascade sensitivity setting (Pesenti et al. 2021) a perturbation of a risk factor affects other dependent risk factors, which in turn impact the output risk measure. To prove the derived sensitivity formulas we use quantile differentiation and weak convergence of generalised functions. We find that stresses propagated via step functions naturally lead to delta functions, which in turn allow for representation as conditional expectations. Hence, our framework allows estimation of differential sensitivity measures by standard simulation-based methods (Glasserman 2005, Fu et al. 2009, Koike et al. 2022).

Key to our framework is the choice of perturbation or *stress* on the random input factor. In particular, the technical conditions we require pertain to the continuity of the stressing mechanism rather than the underlying random input factor. Consequently, our methods can also be applied to the calculation of differential sensitivity measures with respect to discrete random inputs for a suitably chosen stress. Sensitivity to discrete or categorical input factors is of importance in a variety of fields, such as modelling biological systems (Gunawan et al. 2005), chemical processes (Plyasunov and Arkin 2007), and insurance claims (Wüthrich and Merz 2023).

The manuscript is organised as follows. Section 2 introduces the discontinuous loss model and discusses choices of stresses on a risk factor. Following that, expressions are derived for differential (marginal) sensitivities, with respect to the VaR and ES risk measures. The next two sections contain extensions within that framework. Section 3 deals with cascade sensitivities, which reflect indirect effects via risk factors' dependence structure. Section 4 provides differential sensitivities when the considered input random variables are discrete, along with an application to compound distributions. Finally, a detailed numerical study of a reinsurance credit risk portfolio is given in Section 5. Differential sensitivity formulas for a more general model function can be found in Appendix A. All proofs are delegated to Appendix B. Finally, Appendix C contains additional details on the reinsurance credit risk portfolio model used in Section 5.

2. Differential Sensitivity Measures

2.1. Portfolio Loss Model

We work on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and consider a discontinuous model of the form

$$L := \sum_{j=1}^{m} g_j(\boldsymbol{Z}) \mathbb{1}_{\{X_j \leq d_j\}}, \tag{1}$$

where:

- The random vectors $\boldsymbol{X} := (X_1, \ldots, X_m)$, $\boldsymbol{Z} := (Z_1, \ldots, Z_n)$, $m, n \in \mathbb{N}$ are model inputs or *risk factors*;
- L is the (univariate) random model output, which we typically interpret as a *loss*;
- Discontinuities emerge at those states where elements of $\boldsymbol{X}$ cross the *thresholds* $d_1, \ldots, d_m \in \mathbb{R}$;
- The functions $g_j \colon \mathbb{R}^n \to \mathbb{R}$, $j \in \mathcal{M} := \{1, \ldots, m\}$ represent the (random) jump of the model output at the points of discontinuity.

We assume throughout that the marginal distribution functions of X_j, $j \in \mathcal{M}$ and Z_k, $k \in \mathcal{N} := \{1, \ldots, n\}$, denoted by $F_j(x) := \mathbb{P}(X_j \leq x)$ and $F_{m+k}(z) := \mathbb{P}(Z_k \leq z)$, respectively, are absolutely continuous and strictly increasing on their support and denote their corresponding (strictly positive, a.e. on their support) densities by f_j and f_{m+k} respectively. We further denote by $F(l) := \mathbb{P}(L \leq l)$ the distribution of the loss L. The functions $g_j \colon \mathbb{R}^n \to \mathbb{R}$, $j \in \mathcal{M}$ are differentiable.

A standard example of a discontinuous loss (1) is a structural model of a credit risk portfolio (e.g. McNeil et al. 2015, Ch. 11), where $\{X_j \leq d_j\}$ represents the default event of obligor j and $g_j(\boldsymbol{Z})$ the corresponding loss given default. Applications to credit risk modelling are further discussed in

Example 1 and Section 5. We note that the model (1) is formulated such that g_j are functions of $\boldsymbol{Z}$ only. We make this assumption throughout the paper to simplify exposition; the general case of g_j depending on both $\boldsymbol{Z}$ and $\boldsymbol{X}$ is treated in Appendix A.2.

The risk of a loss is assessed via a risk measure $\rho\colon \mathcal{L}^1 \to \mathbb{R}$, where $\mathcal{L}^1$ denotes the set of integrable random variables on $(\Omega, \mathcal{A}, \mathbb{P})$. The two most widely used risk measures in practice are the *Value-at-Risk* (VaR) and the *Expected Shortfall* (ES). The VaR at level $\alpha \in [0,1]$ of the portfolio loss L is defined as the (left-) quantile function of L evaluated at α, that is

$$\mathrm{VaR}_\alpha(L) := F^{-1}(\alpha) = \inf\{y \in \mathbb{R} \,|\, F(y) \geq \alpha\}\,,$$

with the convention that $\inf \emptyset = +\infty$. The Expected Shortfall at level $\alpha \in [0,1)$ of the portfolio loss L is defined by

$$\mathrm{ES}_\alpha(L) := \frac{1}{1-\alpha}\int_\alpha^1 F^{-1}(u)\,du\,.$$

While we focus on VaR and ES, the sensitivities can be generalised to other quantile-based functionals, such as rank dependent expected utilities or spectral risk measures (Acerbi 2002) – in the interest of concision we do not pursue this further.

In Section 2.3 we introduce the *marginal sensitivity measure*, and derive expressions in the context of the VaR/ES risk measures and the discontinuous model (1). The sensitivity measure is defined via a partial derivative of a risk measure in the direction of a *stressed* version of a risk factor; hence we first introduce ways of stressing risk factors.

2.2. Stressing Risk Factors

Throughout the paper, we fix the index i of the risk factor with respect to which sensitivity is calculated, such that stresses are applied to either X_i, with $i \in \mathcal{M}$, or to Z_i, with $i \in \mathcal{N}$. We define a *stress* on X_i or Z_i as a deformation of the risk factor given by

$$X_{i,\varepsilon} := \kappa_\varepsilon(X_i)\,, \quad \text{respectively,} \quad Z_{i,\varepsilon} := \kappa_\varepsilon(Z_i)\,,$$

where $\kappa_\varepsilon\colon \mathbb{R} \to \mathbb{R}$ is a *stress function* defined as follows.

Definition 1 (Stress Function). A family of functions $\kappa_\varepsilon\colon A \to A$, $A \subseteq \mathbb{R}$, where $\varepsilon \in [0, +\infty)$, is called a *stress function*, if it satisfies the following properties:

i) For all ε in a neighbourhood of 0, the function $\kappa_\varepsilon(x)$ is invertible in $x \in A$, denoted by $\kappa_\varepsilon^{-1}(\cdot)$;

ii) $\lim_{\varepsilon \searrow 0} \kappa_\varepsilon(x) = x$, for all $x \in A$;

iii) $\lim_{\varepsilon \searrow 0} \kappa_\varepsilon^{-1}(x) = x$, for all $x \in A$;

iv) One of the following holds:

  (a) for all ε in a neighbourhood of 0 and all $x \in A$, it holds that $\kappa_\varepsilon(x) \geq x$; or

(*b*) for all ε in a neighbourhood of 0 and all $x \in A$, it holds that $\kappa_\varepsilon(x) \le x$;

v) $\kappa_\varepsilon(x)$ is differentiable in ε at $\varepsilon = 0$, and we denote its derivative by

$$\mathfrak{K}(x) := \lim_{\varepsilon \to 0} \frac{\kappa_\varepsilon(x) - x}{\varepsilon}, \qquad x \in A;$$

vi) $\kappa_\varepsilon^{-1}(x)$ is differentiable in ε at $\varepsilon = 0$, and we denote its derivative by

$$\mathfrak{K}^{-1}(x) := \lim_{\varepsilon \to 0} \frac{\kappa_\varepsilon^{-1}(x) - x}{\varepsilon}, \qquad x \in A.$$

We further define

$$c(\kappa) := \begin{cases} +1, & \text{if} \quad \kappa_\varepsilon \text{ fulfils } iv)\ (a), \\ -1, & \text{if} \quad \kappa_\varepsilon \text{ fulfils } iv)\ (b). \end{cases} \tag{2}$$

The requirements on the stress function are assumptions on its continuity. First, if stressing X_i, we typically assume that the domain of the stress function A is equal to the support of X_i. This guarantees that X_i and its stressed version $\kappa_\varepsilon(X_i)$ have the same support. Second, properties *i)* to *iii)* provide that the stressed risk factor converges $\mathbb{P}$-a.s. to its unstressed form as $\varepsilon \searrow 0$. Property *iv)* means that the stress, e.g. $X_{i,\varepsilon}$, either approaches X_i $\mathbb{P}$-a.s. from above or below. The last two properties imply that the stress function and its inverse are differentiable, so that the sensitivities, introduced in Sections 2.3 and 3, exist.

Different stress functions may be used, depending on the context of the problem investigated and what type of deformation of a risk factor is interpretable within that context. For example additive and proportional stresses can be seen as modifications of location and scale respectively; tail stresses may reflect risk management objectives; mixture stresses are used to represent model uncertainty. Some stress functions and related quantities are summarised in Table 1. The additive and proportional stresses with $\beta > 0$ are such that property *iv) (a)* is satisfied and the stress stochastically increases the risk factor; this is easily modified by choosing $\beta < 0$. For the mixture and tail stresses both increasing (*iv) (a)*) and decreasing (*iv) (b)*) versions of the stresses are stated. The functions $\mathfrak{K}$, $\mathfrak{K}^{-1}$ are easily worked out; some additional detail for mixture stresses is given in Appendix B.1. Stress functions should be designed with particular investigations in mind; we explore this further in Example 1.

2.3. Marginal Sensitivity

For a stress $Z_{i,\varepsilon}$ or a stress $X_{i,\varepsilon}$, we denote the corresponding marginally stressed loss model by, respectively

$$L_\varepsilon(Z_i) := \sum_{j \in \mathcal{M}} g_j(\boldsymbol{Z}_{-i}, Z_{i,\varepsilon})\, \mathbb{1}_{\{X_j \le d_j\}} \quad \text{and}$$

$$L_\varepsilon(X_i) := \sum_{\substack{j \ne i \\ j \in \mathcal{M}}} g_j(\boldsymbol{Z})\, \mathbb{1}_{\{X_j \le d_j\}} + g_i(\boldsymbol{Z})\, \mathbb{1}_{\{X_{i,\varepsilon} \le d_i\}},$$

Table 1 Types of stress functions and related quantities.

Type of stress	κ_ε	$\mathfrak{K}$	$\mathfrak{K}^{-1}$	$c(\kappa)$
Additive	$x+\beta\varepsilon$	β	$-\beta$	$\text{sgn}(\beta)$
Proportional	$x(1+\beta\varepsilon)$	βx	$-\beta x$	$\text{sgn}(\beta)$
Probability	$F_i^{-1}\big(F_i(x)+\beta x\big)$	$\frac{\beta}{f_i(x)}$	$\frac{-\beta}{f_i(x)}$	$\text{sgn}(\beta)$
Mixture	$F_{i,\varepsilon}^{-1}\circ F_i(x)$, where $F_{i,\varepsilon}(x):=(1-\varepsilon)\,F_i(x)+\varepsilon\,G(x)$	$\frac{F_i(x)-G(x)}{f_i(x)}$	$\frac{G(x)-F_i(x)}{f_i(x)}$	$\text{sgn}\big(F_i(x)-G(x)\big)$
Tail	$x+\varepsilon\,(x-t)\,\mathbb{1}_{\{x\geq t\}}$	$(x-t)_+$	$-(x-t)_+$	1
	$x+\varepsilon\,(x-t)\,\mathbb{1}_{\{x\leq t\}}$	$-(t-x)_+$	$(t-x)_+$	-1

where $(\boldsymbol{Z}_{-i}, Z_{i,\varepsilon})$ is the vector $\boldsymbol{Z}$ whose i^{th} component is replaced by $Z_{i,\varepsilon}$. We call L_ε, denoting either $L_\varepsilon(Z_i)$ or $L_\varepsilon(X_i)$, the marginally stressed loss, since only the marginal distribution of Z_i or X_i is altered, leaving all other input factors fixed. We denote by $F_\varepsilon(\cdot)$ the distribution function of L_ε and by $q_\varepsilon(\cdot):=F_\varepsilon^{-1}(\cdot)$ the quantile function of L_ε for any $\varepsilon\geq 0$.

For the sensitivities to exist, we require two assumptions on the stressed loss model.

Assumption 1. *Let $0\leq\alpha\leq 1$. For all ε in a neighbourhood of 0 the distribution function F_ε of the marginally stressed loss L_ε is continuously differentiable at $q_\alpha:=F^{-1}(\alpha)$ and let $f_\varepsilon(\cdot)$ be the derivative of F_ε. For $\varepsilon=0$, we simply write $F:=F_0$ and $f:=f_0$.*

Assumption 2. *Let $0\leq\alpha\leq 1$. The quantile function at level α of the stressed loss L_ε, $q_\varepsilon(\alpha)$, is differentiable with respect to ε, that is $\frac{\partial}{\partial\varepsilon}q_\varepsilon(\alpha)$ exists.*

Definition 2 (Marginal Sensitivity). The *marginal sensitivity* to the risk factor Z_i and X_i for a risk measure ρ is defined by, respectively,

$$\mathcal{S}_{Z_i}[\rho]:=\frac{\partial}{\partial\varepsilon}\rho\left(L_\varepsilon(Z_i)\right)\Big|_{\varepsilon=0}\quad\text{and}\quad\mathcal{S}_{X_i}[\rho]:=\frac{\partial}{\partial\varepsilon}\rho\left(L_\varepsilon(X_i)\right)\Big|_{\varepsilon=0},\tag{3}$$

whenever the derivatives exists.

Theorem 1 **(Marginal Sensitivity VaR)**. *Let Assumptions 1 and 2 be fulfilled for a given $\alpha\in(0,1)$. Then, the marginal sensitivity for VaR_α to input factor Z_i for a stress with stress function κ_ε is*

$$\mathcal{S}_{Z_i}[\text{VaR}_\alpha]=\sum_{j\in\mathcal{M}}\mathbb{E}\left[\mathfrak{K}(Z_i)\partial_i g_j(\boldsymbol{Z})\mathbb{1}_{\{X_j\leq d_j\}}\;\Big|\;L=q_\alpha\right],$$

where $\partial_i g_j(\boldsymbol{z}):=\frac{\partial}{\partial z_i}g_j(\boldsymbol{z})$ is the partial derivative in the i^{th} component. The marginal sensitivity to input factor X_i is given by

$$\mathcal{S}_{X_i}[\text{VaR}_\alpha]=c(\kappa)\mathfrak{K}^{-1}(d_i)\,\frac{f_i(d_i)}{f\,(q_\alpha)}\;\mathbb{E}\left[\left(\mathbb{1}_{\{L\leq q_\alpha+c(\kappa)g_i(\boldsymbol{Z})\}}-\mathbb{1}_{\{L\leq q_\alpha\}}\right)\;\Big|\;X_i=d_i\right].$$

Theorem 2 **(Marginal Sensitivity ES)**. *Let Assumptions 1 and 2 be fulfilled for a given* $\alpha \in (0,1)$*. Then, the marginal sensitivity for* ES_α *to input factor* Z_i *for a stress with stress function* κ_ε *is*

$$\mathcal{S}_{Z_i}[\mathrm{ES}_\alpha] = \sum_{j \in \mathcal{M}} \mathbb{E}\left[\mathfrak{K}(Z_i)\, \partial_i\, g_j(\boldsymbol{Z}) \mathbb{1}_{\{X_j \le d_j\}} \;|\; L \ge q_\alpha \right].$$

The marginal sensitivity to input factor X_i *for a stress with stress function* κ_ε *is*

$$\mathcal{S}_{X_i}[\mathrm{ES}_\alpha] = \frac{-c(\kappa)\mathfrak{K}^{-1}(d_i)\, f_i(d_i)}{1-\alpha}\, \mathbb{E}\left[\left(L - c(\kappa) g_i(\boldsymbol{Z}) - q_\alpha\right)_+ - \left(L - q_\alpha\right)_+ \;\middle|\; X_i = d_i \right].$$

The marginal sensitivity measures to Z_i for both the VaR and ES generalise the sensitivities derived in Hong (2009), Hong and Liu (2009) to loss functions L that are not Lipschitz continuous and to general types of stresses. Related, Fu et al. (2009) proposes a conditional Monte-Carlo approach to estimate quantile sensitivities. Note however, that their key assumption is that the perturbed distribution function of L_ε can be written as $F_{L_\varepsilon}(t) = \mathbb{E}[G(t, \varepsilon, Y(\varepsilon))]$, where G is $\mathbb{P}$-a.s. continuous w.r.t. ε and $Y(\varepsilon)$ is an arbitrary random variable. This assumption does not hold in our setting as can be seen in, e.g., Equation (15) of the Proof of Theorem 1. Furthermore, one could derive the marginal sensitivities of ES_α – as well as those of other spectral risk measures (Acerbi 2002) – using its representation as the integral of VaR_α. Interchanging the limit and the integral, however, requires that the sensitivities for VaR_β to exist, for all $\beta \in [\alpha, 1)$. This in particular would imply that Assumptions 1 and 2 need to hold for all $\beta \in (\alpha, 1)$, which is in contrast Theorem 2 which requires Assumptions 1 and 2 to hold for α only.

We now provide an expression for the marginal sensitivity of the mean. While this could be obtained as a special case of Expected Shortfall with $\alpha = 0$, it is simpler to derive Corollary 1 as a direct consequence of Lemma 1 in Appendix B.

Corollary 1 **(Marginal Sensitivity Mean)**. *Let* κ_ε *be a stress function, then the marginal sensitivity for the mean* ($\mathbb{E}$) *to input factor* Z_i *respective* X_i *are*

$$\mathcal{S}_{Z_i}[\mathbb{E}] = \sum_{j \in \mathcal{M}} \mathbb{E}\left[\mathfrak{K}(Z_i)\, \partial_i\, g_j(\boldsymbol{Z}) \mathbb{1}_{\{X_j \le d_j\}} \right] \quad \textit{and}$$

$$\mathcal{S}_{X_i}[\mathbb{E}] = \mathfrak{K}^{-1}(d_i)\, f_i(d_i)\, \mathbb{E}[g_i(\boldsymbol{Z}) \;|\; X_i = d_i].$$

We conclude the section with an example of how the marginal sensitivity measure can be applied in the context of a standard portfolio credit risk model, with two different analysis objectives in mind.

Example 1. Consider a credit risk setting, where $\boldsymbol{Z}$ has the same dimension as $\boldsymbol{X}$, with $g_j(\boldsymbol{Z}) = Z_j,\ j \in \mathcal{M}$ representing the loss given default and $\{X_j \le d_j\}$ the default events with corresponding probabilities $F_j(d_j)$. Hence we have that

$$L = \sum_{j \in \mathcal{M}} Z_j \mathbb{1}_{\{X_j \le d_j\}}.$$

A first analysis pertains to the calculation of the sensitivity of the portfolio ES with respect to the probability of the i-th default event. To achieve this, we need to formulate an appropriate stress function. Consider the probability stress from Table 1, $\kappa_\varepsilon(x) = F_i^{-1}(F_i(x) - \varepsilon)$, leading to

$$X_{i,\varepsilon} = F_i^{-1}(F_i(X_i) - \varepsilon) \quad \text{and}$$
$$\mathbb{P}(X_{i,\varepsilon} \leq d_i) = \mathbb{P}\left(F_i(X_i) \leq F_i(d_i) + \varepsilon\right) = F_i(d_i) + \varepsilon.$$

Hence the chosen stress function gives an additive stress on the default probability, such that the sensitivity $\mathcal{S}_{X_i}[\mathrm{ES}_\alpha]$ becomes precisely the derivative of the portfolio risk in direction of the default probability of the i-th obligor. Using $c(\kappa) = -1$ and $\mathfrak{K}^{-1}(x) = \frac{1}{f_i(x)}$; Theorem 2 yields:

$$\mathcal{S}_{X_i}[\mathrm{ES}_\alpha] = \frac{1}{1-\alpha}\,\mathbb{E}\left[\left(L - (q_\alpha - Z_i)\right)_+ - \left(L - q_\alpha\right)_+ \;\Big|\; X_i = d_i\right].$$

The resulting sensitivity can thus be understood as the difference between two expectations, each representing the excess portfolio loss over a threshold, conditioned on the least adverse outcome of X_i that gives a default of the i-th obligor. The difference between the two terms lies in the lower threshold used in the first term, which is reduced by the loss given default Z_i.

Second, we consider the sensitivity to a proportional increase in the loss given default Z_i, that is, using $\kappa_\varepsilon(z) = z(1+\varepsilon)$. Application of Theorem 2 then gives us:

$$\mathcal{S}_{Z_i}[\mathrm{ES}_\alpha] = \mathbb{E}\left[\,Z_i \mathbb{1}_{\{X_i \leq d_i\}} \mid L \geq q_\alpha\right].$$

Note that this is precisely the Euler allocation of the risk $\mathrm{ES}_\alpha(L)$ to the loss $Z_i \mathbb{1}_{\{X_i \leq d_i\}}$ to the i-th obligor (Tasche 1999).

Finally, within the same model, we turn our attention to assessment of the relative importance of common factors that drive dependence between defaults. The dependence of the critical variables X_j is often modelled via factor models (McNeil et al. 2015, Ch. 6.4, 11) and a question of interest is the relative importance of underlying factors for portfolio risk. Consider the following representation

$$X_j := \sum_{t=1}^{\tau} \beta_{j,t} W_t + V_j, \quad j \in \mathcal{M},$$

where $W_t,\ t = 1, \dots, \tau,\ \tau \in \mathbb{N}$, are the common factors, and V_j are idiosyncratic error terms. We are interested in the sensitivity of the portfolio loss to the factor W_s. To that effect, define:

$$\tilde{X}_{j,\varepsilon} := \sum_{t \neq s} \beta_{j,t} W_t + \beta_{j,s}(W_s - \varepsilon) + V_j = X_j - \beta_{j,s}\varepsilon\,,$$
$$\tilde{\kappa}_\varepsilon(x) := x - \beta_{j,s}\varepsilon\,,$$
$$\tilde{L}_\varepsilon := \sum_{j \in \mathcal{M}} Z_j \mathbb{1}_{\{\tilde{\kappa}_\varepsilon(X_j) \leq d_j\}}\,.$$

The sensitivity of the portfolio risk to the factor W_s can then be written as

$$\frac{\partial}{\partial\varepsilon}\mathrm{ES}_\alpha(\tilde{L}_\varepsilon)\Big|_{\varepsilon=0} = \sum_{j\in\mathcal{M}} S_{X_j}[\mathrm{ES}_\alpha]\,,$$

where the sensitivities $S_{X_j}[\mathrm{ES}_\alpha]$ are now calculated with the stress functions $\tilde{\kappa}_\varepsilon$ above. Applying again Theorem 2 leads to

$$\frac{\partial}{\partial\varepsilon}\mathrm{ES}_\alpha(\tilde{L}_\varepsilon)\Big|_{\varepsilon=0} = \sum_{j\in\mathcal{M}} \frac{f_j(d_j)\beta_{j,s}}{1-\alpha}\,\mathbb{E}\left[\big(L-(q_\alpha - Z_j)\big)_+ - \big(L-q_\alpha\big)_+ \;\Big|\; X_i = d_i\right].$$

Hence, intuitively, the sensitivity to the common factor W_s is expressed as sum of sensitivities for each obligor, weighted by the factor loadings $\beta_{j,s}$.

3. Measuring Cascading Effects

The marginal sensitivity introduced in Section 2.3 quantifies the differential impact of stressing a risk factor on the portfolio loss. Here, we provide the first generalisation/adjustment of the framework, by considering *cascade sensitivity* measures, as discussed in Pesenti et al. (2021). These sensitivity measures quantify not only the sensitivity to an individual input X_i, but also consider (joint) perturbation of all other risk factors X_j, $j \neq i$, and Z_k, $k\in\mathcal{N}$, induced by their statistical dependence on X_i. This is achieved by using the *inverse Rosenblatt transform*, recalled next (Rosenblatt 1952, Rüschendorf and de Valk 1993).

Definition 3 (Inverse Rosenblatt Transform). An inverse Rosenblatt transform of an m-dimensional random vector $\boldsymbol{Y}$, starting at Y_i, for fixed $i\in\mathcal{M}$, is given by a function $\boldsymbol{\Psi} = (\Psi^{(1)},\dots,\Psi^{(m)})^\top : \mathbb{R}^m \to \mathbb{R}^m$ and an $(m-1)$-dimensional random vector $\boldsymbol{V} = (V_1,\dots,V_{m-1})$, consisting of independent standard uniform variables, independent of Y_i, such that

$$\boldsymbol{Y} = \boldsymbol{\Psi}(Y_i,\boldsymbol{V}) = \big(\Psi^{(1)}(Y_i,\boldsymbol{V}),\,\dots,\,\Psi^{(m)}(Y_i,\boldsymbol{V})\big)\,,\quad \mathbb{P}\text{-a.s.}.$$

In particular, $Y_k = \Psi^{(k)}(Y_i,\boldsymbol{V})$ $\mathbb{P}$-a.s. for all $k\in\mathcal{M}$.

To construct the cascade sensitivity to input X_i, we proceed via the following steps. First, we write the vector of risk factors $(\boldsymbol{X},\boldsymbol{Z})$ via its inverse Rosenblatt transform starting from X_i. From this we obtain $(\boldsymbol{X},\boldsymbol{Z}) = \boldsymbol{\Psi}(X_i,\boldsymbol{V})$ $\mathbb{P}$-a.s., that is $X_j = \Psi^{(j)}(X_i,\boldsymbol{V})$, for all $j\in\mathcal{M}$, and $Z_k = \Psi^{(m+k)}(X_i,\boldsymbol{V})$ for all $k\in\mathcal{N}$. Second, we stress X_i to $X_{i,\varepsilon}$, such that the stressed vector of risk factors becomes $\boldsymbol{\Psi}(X_{i,\varepsilon},\boldsymbol{V})$. Thus, using the inverse Rosenblatt transform, all other risk factors are perturbed according to their dependence on X_i. Third, we apply the marginal sensitivity to the portfolio loss as a function of the stressed input vector $\boldsymbol{\Psi}(X_{i,\varepsilon},\boldsymbol{V})$. When stressing Z_i the process is analogous. For simplicity of notation, we use $\boldsymbol{\Psi}$ for the inverse Rosenblatt transform regardless of whether the transform starts at X_i or Z_i. We further write $\boldsymbol{\Psi}^{(\boldsymbol{Z})} = (\Psi^{(m+1)}, \dots, \Psi^{(m+n)})$, so that $\boldsymbol{Z} = \boldsymbol{\Psi}^{(\boldsymbol{Z})}(X_i, \boldsymbol{V})$ or $\boldsymbol{Z} = \boldsymbol{\Psi}^{(\boldsymbol{Z})}(Z_i, \boldsymbol{V})$ (depending on whether the inverse Rosenblatt transform starts at X_i or Z_i). Specifically, following the above process, we can define the stressed loss models for stresses $Z_{i,\varepsilon}$ and $X_{i,\varepsilon}$ respectively by

$$L_\varepsilon^{\boldsymbol{\Psi}}(Z_i) := \sum_{j\in\mathcal{M}} g_j\left(\boldsymbol{\Psi}^{(\boldsymbol{Z})}\left(Z_{i,\varepsilon}, \boldsymbol{V}\right)\right) \mathbb{1}_{\{\Psi^{(j)}(Z_{i,\varepsilon}, \boldsymbol{V}) \le d_j\}} \quad \text{and}$$
$$L_\varepsilon^{\boldsymbol{\Psi}}(X_i) := \sum_{j\in\mathcal{M}} g_j\left(\boldsymbol{\Psi}^{(\boldsymbol{Z})}\left(X_{i,\varepsilon}, \boldsymbol{V}\right)\right) \mathbb{1}_{\{\Psi^{(j)}(X_{i,\varepsilon}, \boldsymbol{V}) \le d_j\}}\,.$$

With these building blocks in place, we can now define the cascade sensitivity measure in the specific context of this paper.

Definition 4 (Cascade Sensitivity). The *cascade sensitivity* to the risk factor Z_i and X_i for a risk measure ρ is defined by, respectively,

$$\mathcal{C}_{Z_i}[\rho] := \frac{\partial}{\partial\varepsilon}\rho\big(L_\varepsilon^{\boldsymbol{\Psi}}(Z_i)\big)\Big|_{\varepsilon=0}, \quad \text{and} \quad \mathcal{C}_{X_i}[\rho] := \frac{\partial}{\partial\varepsilon}\rho\big(L_\varepsilon^{\boldsymbol{\Psi}}(X_i)\big)\Big|_{\varepsilon=0}, \tag{5}$$

whenever the derivative exists.

Note that if the cascade sensitivity is exists, it is independent of the choice of Rosenblatt transform, see Prop. 3.6 in Pesenti et al. (2021). In order to establish existence, in this section we make the assumption that the inverse Rosenblatt transforms are differentiable and locally monotone in their first argument. This means that stressing a model input leads to perturbation of elements of $\boldsymbol{X}$ that makes them $\mathbb{P}$-a.s. greater (or smaller) than the original input X_j.

Assumption 3. *Let κ_ε be a stress function and $Y,\ Y_\varepsilon$ be such that either $Y := Z_i,\ Y_{i,\varepsilon} := Z_{i,\varepsilon}$ or $Y := X_i,\ Y_{i,\varepsilon} := X_{i,\varepsilon}$. Let $\boldsymbol{\Psi}$ be a differentiable inverse Rosenblatt transform starting at Y, such that $(\boldsymbol{X}, \boldsymbol{Z}) = \boldsymbol{\Psi}(Y, \boldsymbol{V})$. Then, for each $j \in \mathcal{M}$, one of the following holds*

(a) for all ε in a neighbourhood of 0, it holds $\Psi^{(j)}\left(Y_{i,\varepsilon}, \boldsymbol{V}\right) \ge X_j$ $\mathbb{P}$-a.s.; or

(b) for all ε in a neighbourhood of 0, it holds $\Psi^{(j)}\left(Y_{i,\varepsilon}, \boldsymbol{V}\right) \le X_j$ $\mathbb{P}$-a.s.

In the case (a) we denote $c(\kappa;\, j) = 1$ and in the case (b) $c(\kappa;\, j) = -1$.

With these assumptions in place, we can now obtain explicit formulas for the cascade sensitivity measure of Definition 4. In Theorems 3 and 4 below we deal with the case of ES, while formulas for VaR are given in Appendix A.1. We observe that the cascade sensitivity to both X_i and Z_i entails a decomposition, reflecting the indirect contribution of the vector being stressed via the other inputs $X_j,\ Z_k$.

Theorem 3 **(Cascade Sensitivity ES to X_i)**. *Let Assumptions 1, 2 and 3 (for $Y = X_i$) be fulfilled for the stressed model $L_\varepsilon^{\mathbf{\Psi}}(X_i)$ and $\alpha \in (0,1)$. Denote $\Psi_1^{(j)}(x,\boldsymbol{v}) := \frac{\partial}{\partial x}\Psi^{(j)}(x,\boldsymbol{v})$. Then, the cascade sensitivity for ES_α to input X_i is given by*

$$\mathcal{C}_{X_i}[\,\mathrm{ES}_\alpha\,] = \sum_{j\in\mathcal{M}} \mathcal{C}_{X_i,X_j} + \sum_{k\in\mathcal{N}} \mathcal{C}_{X_i,Z_k},$$

where, for all $k \in \mathcal{N}$,

$$\mathcal{C}_{X_i,Z_k} = \sum_{j\in\mathcal{M}} \mathbb{E}\left[\mathfrak{K}(X_i)\,\partial_k\, g_j(\boldsymbol{Z})\Psi_1^{(m+k)}(X_i,\boldsymbol{V})\mathbb{1}_{\{X_j\le d_j\}} \;\middle|\; L \ge q_\alpha\right],$$

and for all $j \in \mathcal{M}$,

$$\mathcal{C}_{X_i,X_j} = -\frac{c(\kappa;j)f_j(d_j)}{1-\alpha}\mathbb{E}\left[\mathfrak{K}^{-1}(X_i)\Psi_1^{(j)}(X_i,\boldsymbol{V})\left(\left(L - c(\kappa;j)g_j(\boldsymbol{Z}) - q_\alpha\right)_+ - \left(L-q_\alpha\right)_+\right) \;\middle|\; X_j = d_j\right].$$

Theorem 4 **(Cascade Sensitivity ES to Z_i)**. *Let Assumptions 1, 2 and 3 (for $Y = Z_i$) be fulfilled for the stressed model $L_\varepsilon^{\mathbf{\Psi}}(Z_i)$ and $\alpha \in (0,1)$. Then, the cascade sensitivity for ES_α to input Z_i is given by*

$$\mathcal{C}_{Z_i}[\,\mathrm{ES}_\alpha\,] = \sum_{j\in\mathcal{M}} \mathcal{C}_{Z_i,X_j} + \sum_{k\in\mathcal{N}} \mathcal{C}_{Z_i,Z_k},$$

where, for all $k \in \mathcal{N}$,

$$\mathcal{C}_{Z_i,Z_k} = \sum_{j\in\mathcal{M}} \mathbb{E}\left[\mathfrak{K}(Z_i)\,\partial_k\, g_j(\boldsymbol{Z})\Psi_1^{(m+k)}(Z_i,\boldsymbol{V})\mathbb{1}_{\{X_j\le d_j\}} \;\middle|\; L \ge q_\alpha\right],$$

and for $j \in \mathcal{M}$,

$$\mathcal{C}_{Z_i,X_j} = -\frac{c(\kappa;j)f_j(d_j)}{1-\alpha}\mathbb{E}\left[\mathfrak{K}^{-1}(Z_i)\Psi_1^{(j)}(Z_i,\boldsymbol{V})\left(\left(L - c(\kappa;j)g_j(\boldsymbol{Z}) - q_\alpha\right)_+ - \left(L-q_\alpha\right)_+\right) \;\middle|\; X_j = d_j\right].$$

To calculate the cascade sensitivities, we need the derivative of the inverse Rosenblatt transform. This calculation is simplified by noting that the value of the cascade sensitivity is independent of the specific choice of Rosenblatt transform. Hence, when calculating, for example, $\Psi_1^{(j)}(X_i,\boldsymbol{V})$, we can without loss of generality use the standard construction (Rüschendorf and de Valk 1993) $\Psi^{(j)}(X_i,\boldsymbol{V}) = F^{-1}_{X_j|X_i}(V_1|X_i)$ in Theorem 3 – analogously if Z_i is being stressed (Theorem 4). As a result, it is sufficient to consider the derivatives of inverse Rosenblatt transforms corresponding to the bivariate dependence structure of, e.g., (X_i,X_j). If the bivariate copula between the risk factors are known, analytical expressions for the required derivatives may be available. We refer also to Pesenti et al. (2021), where the formulas given below for the Gaussian and t copulas are derived.

For simplicity of presentation, we only provide the expressions for $\Psi_1^{(j)}(X_i,V)$, where V is a suitably defined random variable such that $X_j = \Psi^{(j)}(X_i,V)$. The formulas for $\Psi_1^{(j)}(Z_i,V)$, $\Psi_1^{(m+k)}(X_i,V)$, and $\Psi_1^{(m+k)}(Z_i,V)$, for $j\in\mathcal{M}$, $k\in\mathcal{N}$ follow analogously.

Proposition 1 **(Bivariate Inverse Rosenblatt Transform)**. *Denote by Φ, ϕ, the distribution function and density of a standard normal variable, and by t_ν, s_ν the distribution function and density of a t-distributed random variable with ν degrees of freedom.*

1. *Assume (X_i, X_j) follows a Gaussian copula with correlation parameter r_{ij} and define $Y_i := \Phi^{-1}(F_i(X_i))$ and $Y_j := \Phi^{-1}(F_j(X_j))$. Then,*

$$\Psi_1^{(j)}(X_i, V) = r_{ij} \, \frac{f_i(X_i)}{\phi(Y_i)} \, \frac{\phi(Y_j)}{f_j(X_j)} \,,$$

2. *Assume (X_i, X_j) follows a t copula with correlation parameter r_{ij} and ν degrees of freedom and define $Y_i := t_\nu^{-1}(F_i(X_i))$ and $Y_j := t_\nu^{-1}(F_j(X_j))$. Then,*

$$\Psi_1^{(j)}(X_i, V) = \left(r_{ij} + \frac{Y_i Y_j - r_{ij} Y_i^2}{\nu + Y_i^2} \right) \frac{f_i(X_i)}{s_\nu(Y_i)} \frac{s_\nu(Y_j)}{f_j(X_j)} \,.$$

3. *Assume (X_i, X_j) follows a Archimedean copula with generator $\psi\colon [0, +\infty] \to [0, 1]$, i.e., the copula is given by*

$$C(u_1, u_2) = \psi\left(\psi^{-1}(u_1) + \psi^{-1}(u_2)\right) , \quad u_1, u_2 \in [0, 1] \,,$$

where ψ^{-1} denotes the inverse of the generator ψ. Then, for $i \neq j$

$$\Psi_1^{(j)}(X_i, V) = \frac{\dot\psi\left(\psi^{-1}(U_j)\right)}{\dot\psi\left(\psi^{-1}(U_i)\right)} \left(\frac{\dot\psi\left(\psi^{-1}(U_i) + \psi^{-1}(U_j)\right)}{\ddot\psi\left(\psi^{-1}(U_i) + \psi^{-1}(U_j)\right)} \, \frac{\ddot\psi\left(\psi^{-1}(U_i)\right)}{\dot\psi\left(\psi^{-1}(U_i)\right)} - 1 \right) \frac{f_i(X_i)}{f_j(X_j)} \,,$$

where $U_j := F_j(X_j)$, $\dot\psi(x) := \frac{\partial}{\partial x}\psi(x)$, and $\ddot\psi(x) := \frac{\partial}{\partial x}\dot\psi(x)$.

4. Sensitivity to Discrete Random Variables

In this section, we adapt the techniques developed so far, to calculate differential sensitivities to discrete risk factors. Given the different portfolio structure we consider here, we change notation to avoid confusion with previous sections. We consider the loss model

$$T := h\,(W, \boldsymbol{Y}) \,, \tag{6}$$

where $\boldsymbol{Y} := (Y_1, \ldots, Y_d)$, the function $h\colon \mathbb{R}^{d+1} \to \mathbb{R}$ is differentiable, and W is a discrete random variable which sensitivity we aim to assess. Such a sensitivity calculation presents both technical and conceptual challenges. While h is differentiable, the corresponding differential (or infinitesimal increment) in its first argument is hard to interpret, given the discreteness of W. Here we propose to calculate the differential sensitivity with respect to a continuous variable, from which W is obtained via a (discontinuous) transformation. In other words, we exchange the problem of discreteness with the one of non-differentiability, which we have established.

We assume that W takes values $w_1 < \cdots < w_r$ with $\mathbb{P}(W \leq w_k) = p_k$, $k = 1, \dots, r$, such that $0 =: p_0 < p_1 < \cdots < p_r = 1$. Next, we rewrite the loss model (6) into a form analogous to (1). For this, let $V \sim \mathrm{U}(0,1)$ be independent of $(W, \boldsymbol{Y})$ and define the uniform random variable U by

$$U := \tilde{F}_W(W; V),$$

where $\tilde{F}_W(w; \lambda) := \mathbb{P}(W < w) + \lambda \mathbb{P}(W = w)$ is the generalised distributional transform of W (Rüschendorf 2013). It then follows that $U \sim \mathrm{U}(0,1)$, U is comonotonic to W, and

$$W = F_W^{-1}(U) = \sum_{k=1}^{r} w_k \mathbb{1}_{\{p_{k-1} < U \leq p_k\}}, \quad \mathbb{P}\text{-a.s.}.$$

Then, following some manipulations, the loss model admits the form:

$$T = \sum_{k=1}^{r} h(w_k, \boldsymbol{Y}) \mathbb{1}_{\{p_{k-1} < U \leq p_k\}} = \sum_{k=1}^{r} \Delta_k h(W, \boldsymbol{Y}) \mathbb{1}_{\{U \leq p_k\}},$$

where $\Delta_k h(W, \boldsymbol{Y}) := h(w_k, \boldsymbol{Y}) - h(w_{k+1}, \boldsymbol{Y})$, for $k = 1, \dots r-1$, and $\Delta_r h(W, \boldsymbol{Y}) := h(w_r, \boldsymbol{Y})$.

We next stress the portfolio loss T with respect to W by applying a stress function to U. Hence, we write the stressed model as follows:

$$T_{W,\varepsilon} := \sum_{k=1}^{r} \Delta_k h(W, \boldsymbol{Y}) \mathbb{1}_{\{\kappa_\varepsilon(U) \leq p_k\}}. \tag{7}$$

Stressing the uniform variable that generates W allows for a cohesive stress, given the comonotonicity of (W, U). Next, we define the differentiable sensitivity to W via a stress on U by:

$$\tilde{\mathcal{S}}_W[\rho] := \frac{\partial}{\partial \varepsilon} \rho(T_{W,\varepsilon}) \Big|_{\varepsilon=0}.$$

Formulas for this sensitivity are given in the following result.

THEOREM 5 **(Marginal Sensitivity – Discrete)**. *Let Assumptions 1 and 2 be fulfilled for the loss model (6) and for a fixed $\alpha \in (0,1)$. Then the sensitivity for VaR to the discrete input W is*

$$\tilde{\mathcal{S}}_W[\mathrm{VaR}_\alpha] = \frac{c(\kappa)}{f(q_\alpha)} \sum_{k=1}^{r} \mathfrak{K}^{-1}(p_k) \mathbb{E}\left[\left(\mathbb{1}_{\{T \leq q_\alpha + c(\kappa)\Delta_k h(W, \boldsymbol{Y})\}} - \mathbb{1}_{\{T \leq q_\alpha\}}\right) \mid W = w_k\right],$$

where, for simplicity of notation, q_α is the α-quantile of T and f its density. The sensitivity for ES to the discrete input W is

$$\tilde{\mathcal{S}}_W[\mathrm{ES}_\alpha] = -\frac{c(\kappa)}{1-\alpha} \sum_{k=1}^{r} \mathfrak{K}^{-1}(p_k) \mathbb{E}\left[\left(T - c(\kappa)\Delta_k h(W, \boldsymbol{Y}) - q_\alpha\right)_+ - (T - q_\alpha)_+ \mid W = w_k\right].$$

In the next example, we apply Theorem 5 for the ES-sensitivity calculation of the frequency variable in a compound loss model. Compound distributions are canonical tools in modelling insurance claims, as well as credit and operational risk losses, and the impact of the choice of frequency distribution is well attested, see e.g., McNeil et al. (2015).

Example 2. We represent by $T = h(W, \boldsymbol{Y})$ a compound random variable. Specifically, we set $r = d+1$ and assume that W is a discrete loss frequency, taking values in $\{w_1 = 0, \dots, w_{d+1} = d\}$, while the d elements of $\boldsymbol{Y} = (Y_1, \dots, Y_d)$ are loss severities. The variable W has distribution $\mathbb{P}(W \le k-1) = p_k,\ k = 1, \dots, d+1$. Furthermore, we assume that $Y_1, \dots, Y_d$ are i.i.d., continuously distributed with $Y_1 \sim F_Y$, and independent of W. The portfolio loss is:

$$T = h(W, \boldsymbol{Y}) = \sum_{\ell=1}^{W} Y_\ell \,,$$

with the understanding that for $W = 0$ we have $T = 0$. Our aim is to calculate the sensitivity of the portfolio's ES to the frequency variable, i.e. to evaluate the quantity $\tilde{\mathcal{S}}_W[\mathrm{ES}_\alpha]$, and to compare this with the impact of the vector of loss severities, $\tilde{\mathcal{S}}_{\boldsymbol{Y}}[\mathrm{ES}_\alpha]$, which are defined below.

The sensitivity $\tilde{\mathcal{S}}_W[\mathrm{ES}_\alpha]$ is evaluated by application of Theorem 5. As before let $W = F_W^{-1}(U)$. To stress U we need to specify a stress function $\kappa_\varepsilon(u) : (0,1) \to (0,1)$. Let $\kappa_\varepsilon(u) := \Phi\left(\Phi^{-1}(u) + \varepsilon\right)$, where Φ is the standard normal distribution. This choice is consistent with the well-known *Wang Transform* (Wang 2000) in risk measure theory and satisfies the conditions of Definition 1, with $c(\kappa) = 1$. Then, for $U_\varepsilon := \Phi\left(\Phi^{-1}(U) + \varepsilon\right)$ we obtain, using Theorem 5 and after a few manipulations not documented here, that

$$\tilde{\mathcal{S}}_W[\mathrm{ES}_\alpha] = \sum_{k=1}^{d} \big(v(p_k) - v(p_{k+1})\big)\, \mathbb{E}\left[\left(\sum_{\ell=1}^{k} Y_\ell - q_\alpha\right)_+\right],$$

where $v(p) := \frac{\phi\left(\Phi^{-1}(p)\right)}{1-\alpha}$, $p \in [0,1]$. Hence, the sensitivity becomes a linear combination of the stop-loss terms $\mathbb{E}[(\sum_{\ell=1}^{k} Y_\ell - q_\alpha)_+]$, with the coefficient weights driven by the distribution of the loss frequency W.

We now turn our attention to stressing the loss severities $\boldsymbol{Y}$. We choose to stress all elements of $\boldsymbol{Y}$ at the same time, using a stress function consistent with the one used for W, i.e. the same κ_ε. Specifically, for $U_\ell := F_Y(Y_\ell),\ \ell = 1, \dots, d$, we define the stressed portfolio

$$T_{\boldsymbol{Y},\varepsilon} := \sum_{\ell=1}^{W} F_Y^{-1}\left(\kappa_\varepsilon(U_\ell)\right) = \sum_{k=1}^{d} \mathbb{1}_{\{W=k\}} \sum_{\ell=1}^{k} F_Y^{-1}\left(\kappa_\varepsilon(U_\ell)\right)$$

and calculate the sensitivity

$$\tilde{\mathcal{S}}_{\boldsymbol{Y}}[\mathrm{ES}_\alpha] := \frac{\partial}{\partial \varepsilon} \mathrm{ES}_\alpha(T_{\boldsymbol{Y},\varepsilon})\Big|_{\varepsilon=0}.$$

By the pointwise continuity of the mapping $\varepsilon \mapsto T_{\boldsymbol{Y},\varepsilon}$ we can calculate $\tilde{\mathcal{S}}_{\boldsymbol{Y}}[\mathrm{ES}_\alpha]$ by standard methods (Hong and Liu 2009), yielding:

$$\tilde{\mathcal{S}}_{\boldsymbol{Y}}[\mathrm{ES}_\alpha] := \sum_{k=1}^{d} \mathbb{P}(W=k) \sum_{\ell=1}^{k} \mathbb{E}\left[\mathbb{1}_{\{T > q_\alpha\}} \frac{v(U_\ell)}{f_Y(Y_\ell)} \mid W = k\right].$$

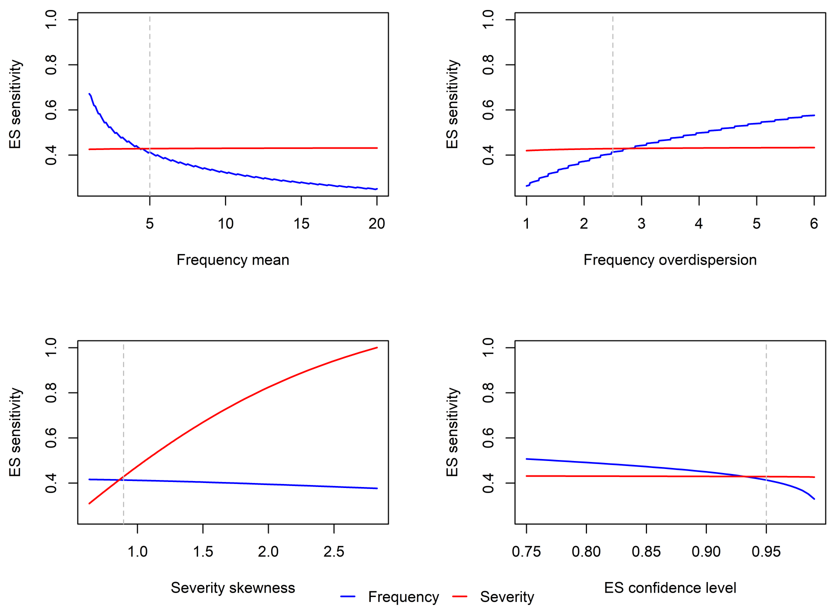

Figure 1 Changes in the scaled ES-sensitivity to the frequency (blue) and severity (red) of a compound Negative Binomial-Gamma distribution. The vertical dashed line in each plot represent baseline assumptions.

We now calculate the sensitivities $\tilde{\mathcal{S}}_W\,[\mathrm{ES}_\alpha\,]$ and $\tilde{\mathcal{S}}_{\boldsymbol{Y}}\,[\mathrm{ES}_\alpha\,]$ for a model with the following baseline assumptions. The confidence level of the risk measure is $\alpha = 0.95$. The frequency W follows a Negative Binomial distribution with mean $\mathbb{E}[W] = 5$ and over-dispersion $\mathrm{Var(W)}/\mathbb{E}[\mathrm{W}] = 2.5$, truncated at the 99.9th percentile. The severities Y_ℓ follow Gamma distributions with shape parameter $\theta = 5$, corresponding to a skewness coefficient of 0.894. With these choices, we find that the sensitivities, scaled by the portfolio risk, take values $\frac{\tilde{\mathcal{S}}_W\,[\mathrm{ES}_\alpha\,]}{\mathrm{ES}_\alpha(T)} = 0.414$ and $\frac{\tilde{\mathcal{S}}_{\boldsymbol{Y}}\,[\mathrm{ES}_\alpha\,]}{\mathrm{ES}_\alpha(T)} = 0.429$. This indicates that the compound sum T is approximately equally sensitive to the loss frequency and severity.

In Figure 1 we depict how the scaled sensitivities change after varying the baseline assumptions, one at a time, regarding frequency mean, frequency over-dispersion, the skewness of the severity distribution, and the confidence level of the ES risk measure. In each plot the baseline assumption is indicated by a vertical dashed line. We observe that, as the frequency mean increases, the importance of severities dominates, given the larger overall number of individual losses. On the other hand, when the frequency over-dispersion increases, the importance of frequency dominates, since the variance of the frequency distribution becomes the key risk driver. Furthermore, as one would expected, the sensitivity of the severities $\boldsymbol{Y}$ increases in the skewness, which reflects a riskier loss profile. Finally, as the confidence level increases, severities become more important than the frequency W, representing a more pronounced impact on the extreme tail of the portfolio loss.

5. Application to reinsurance credit risk modelling

Reinsurance credit risk modelling represents a prominent example where credit risk exposures are non-granular and inhomogeneous. Insurers buy reinsurance products in order to transfer some of the risk of (typically) higher than expected claim amounts to a third party. By taking on insurers' excess liabilities, the reinsurance market thus operates as an industry-wide risk pooling arrangement (Albrecher et al. 2017). Credit risk then arises from the possibility that, in the event of high (industry) losses, reinsurers will not be able to make good on their obligations to insurers.

Reinsurance credit risk has two features particularly relevant to our setting. First, dependence is of primary importance. Different reinsurers' ability to fulfil obligations is highly dependent on each other, given the systemic impact of overall (re)insurance market conditions and industry shocks. As a result, reinsurers' default indicators should also be considered dependent on insurers' gross (i.e. before-reinsurance) losses; hence one needs to account for the event that reinsurers default precisely at those times when insurers need them most. Second, reinsurance credit risk exposures are highly inhomogeneous. Different reinsurers often reinsure different lines of business at different levels of extreme loss. Furthermore, the credit rating of reinsurers varies and insurers typically transfer the most extreme layers of their gross losses to a small number of highly rated reinsurers – while this is a rational strategy, it creates non-trivial concentration effects. The concern with the risk of reinsurance defaults, and particularly with their dependence, has been thoroughly reflected in actuarial modelling practice (Ter Berg 2008, Britt and Krvavych 2009).

Here we present a numerical example of differential sensitivity analysis to reinsurance defaults, working with an illustrative model of reinsurance credit risk. In equation (1), we interpret the terms as follows:

- L is the total reinsurance credit risk loss for an insurer.
- $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ are the gross losses of the insurer, from its $n = 12$ lines of business (LoB).
- $g_j(\boldsymbol{Z}),\ j \in \mathcal{M}$ are the reinsurance recoveries expected from each of $m = 8$ reinsurers.
- $\{X_j \leq d_j\}$ is the event that the j-th reinsurer defaults.

The 12 LoB are marginally Lognormal distributed with the same mean and coefficient of variation (CoV) given in Table 2, and which are consistent with the Solveny II standard formula parameters (Lloyd's 2022). In specifying the form of the g_js we make the simplifying assumption that all reinsurance contracts bought consist of reinsurance layers on the gross losses $Z_1, \ldots, Z_{12}$.

We assume that each of the first 6 reinsurers covers a layer from two LoBs, with deductibles $s_{j,k}$ and limit $t_{j,k}$. Each of reinsurers 7 and 8 covers a higher layer from 6 LoBs. Specifically, we have:

$$g_j(\boldsymbol{Z}) = \sum_{k=2j-1}^{2i} \min\left\{(Z_k - s_{j,k})_+ \,,\, t_{j,k}\right\}, \quad \text{for} \quad j = 1, \ldots, 6,$$

Table 2 Name of Lines of business (LoB) and Coefficient of variation (CoV) (Source: Lloyd's (2022)).

LOB Name	LoB	CoV
Direct and Proportional Motor Vehicle Liability	Z_1	0.1
Direct and Proportional Other Motor	Z_2	0.08
Direct and Proportional Marine, Aviation and Transportation	Z_3	0.15
Direct and Proportional Fire & Other Damage to Property	Z_4	0.08
Direct and Proportional General Liability	Z_5	0.14
Direct and Proportional Credit & Suretyship	Z_6	0.19
Direct and Proportional Legal Expenses	Z_7	0.083
Direct and Proportional Assistance	Z_8	0.064
Direct and Proportional Miscellaneous Financial Loss	Z_9	0.13
Non-Proportional Casualty Reinsurance	Z_{10}	0.17
Non-Proportional Marine, Aviation and Transportation Reinsurance	Z_{11}	0.17
Non-Proportional Property Reinsurance	Z_{12}	0.17

$$g_7(\boldsymbol{Z}) = \sum_{k=1}^{6} \min\left\{(Z_k - s_{7,k})_+\,,\, t_{7,k}\right\}, \quad \text{and}$$
$$g_8(\boldsymbol{Z}) = \sum_{k=7}^{12} \min\left\{(Z_k - s_{8,k})_+\,,\, t_{8,k}\right\}.$$

The deductibles and limits are such that the first six reinsurers cover losses between the 55% and 85% quantile, whereas the last two reinsurers cover the losses between the 85% and the 95% quantile, i.e.,

$$s_{j,k} = F^{-1}_{Z_k}(0.55) \quad \text{and} \quad t_{j,k} = F^{-1}_{Z_k}(0.85) - s_{j,k}\,, \quad \text{for} \quad j = 1, \dots, 6\,,$$
$$s_{j',k} = F^{-1}_{Z_k}(0.85) \quad \text{and} \quad t_{j',k} = F^{-1}_{Z_k}(0.95) - s_{j,k}\,, \quad \text{for} \quad j' = 7, 8\,.$$

Finally, the default probabilities are set at 1.5% for the first 6 reinsurers and 1% for reinsurers 7 and 8. We assume that the random vector $(\boldsymbol{X}, \boldsymbol{Z})$ is dependent with a t-copula with 4 degrees of freedom, such that the correlation matrix of $\boldsymbol{Z}$ satisfies Solvency II assumptions (Lloyd's 2022), while the random vector $\boldsymbol{X}$ has a homogeneous correlation matrix such that $\text{Corr}(X_i, X_j) = 0.05$. The joint dependence of $(\boldsymbol{X}, \boldsymbol{Z})$ is effected via a t-distribution factor model; further details are given in Appendix C.

The distribution of the total credit risk loss L is evaluated by Monte-Carlo simulation. Specifically, since almost all scenarios of $(\boldsymbol{X}, \boldsymbol{Z})$ result in a credit loss of zero, i.e., a realisation $\{L = 0\}$, we generated a dataset of size 500,000 (keeping track of the total number of simulations), in which all realisations satisfy $L > 0$. The probability that $L > 0$ is 5.044% in our dataset. Figure 2 depicts a histogram of the insurer's total credit risk loss L conditional that a loss occurred. We also report the unconditional VaR and ES at level $\alpha = 0.975$. The skewness and multimodality of the loss distribution, driven by the portfolio's lack of homogeneity, are apparent.

We apply stresses on each of the risk factors X_i and Z_i. Specifically, we apply left-tail stresses (see Table 1) on the risk factors driving defaults, i.e., $X_{i,\varepsilon} := X_i + \varepsilon\left(X_i - F^{-1}_{X_i}(0.2)\right)\mathbb{1}_{\{X_i \le F^{-1}_{X_i}(0.2)\}}$,

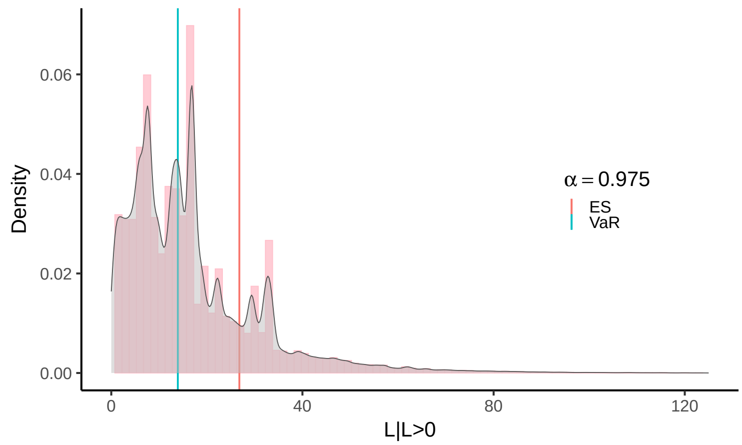

Figure 2 Histogram of the insurer's total reinsurance credit risk loss conditional on a positive loss, i.e., $L|L>0$. Vertical lines are the unconditional VaR and ES at level $\alpha = 0.975$.

$i=1,\dots,8$. These stresses increase the probability of reinsurance defaults, though in a more complex way compared to Example 1. For each LoB, we consider a right-tailed stress $Z_{i,\varepsilon} := Z_i + \varepsilon\left(Z_i - F^{-1}_{Z_i}(0.8)\right)\mathbb{1}_{\{Z_i \geq F^{-1}_{Z_i}(0.8)\}}$, $i=1,\dots,12$, which increases the loss quantiles of Z_i, beyond its 80% quantile.

We calculate the sensitivities with respect to the VaR and ES risk measures at level $\alpha = 0.975$, according to Theorems 1 and 2. To calculate the sensitivities to Z_is, we require estimates of expectation conditioned on the event $\{L = q_\alpha\}$ and $\{L \geq q_\alpha\}$. For estimating the expectation conditional on the event of zero probability $\{L = q_\alpha\}$, we use the δ-estimator (Glasserman 2005). Specifically, for $\delta > 0$ with $0 < \alpha - \delta$, and $\alpha + \delta < 1$, we approximate the sensitivity of VaR to Z_i by

$$\widehat{\mathcal{S}}_{Z_i}\left[\mathrm{VaR}_\alpha\right] = \frac{1}{2\delta}\sum_{j=1}^{8}\mathbb{E}\Big[\mathfrak{K}(Z_i)\partial_i\, g_j(\boldsymbol{Z})\mathbb{1}_{\{X_j \leq d_j\}}\mathbb{1}_{\{L \in (F^{-1}(\alpha-\delta),\, F^{-1}(\alpha+\delta))\}}\Big].$$

Mathematically, we replace the conditioning event $\{L = q_\alpha\}$ by an event of probability 2δ, i.e. by $\{L \in (F^{-1}(\alpha-\delta),\, F^{-1}(\alpha+\delta)\}$. A value of $\delta = 0.005$ was used throughout. We use our sample of $(\boldsymbol{X}, \boldsymbol{Z}, L \mid L > 0)$, which contains 500,000 simulated scenarios, and estimate the sensitivities using bootstrap with replacement and a bootstrap size of 450,000. The reported sensitivities are averaged over 100 bootstrap estimates.

For estimating the sensitivities to each X_i, a different dataset is simulated. Specifically, for each $j = 1,\dots,8$, we generate a dataset of size 500,000, in which all realisations of $(\boldsymbol{X}, \boldsymbol{Z})$ satisfy $X_j \in \left(F^{-1}_{X_j}(d_j - \delta),\, F^{-1}_{X_j}(d_j + \delta)\right)$, for small $\delta > 0$. Again, sensitivities were estimated by bootstrapping 100 times with replacement and bootstrap size 450,000. Figures 3 and 4 display violin plots of the sensitivities to Z_i and to X_i for both VaR and ES. Again a value of $\delta = 0.005$ is used as a baseline; the effect of this choice on sensitivity estimates is discussed in the sequel (Figure 6).

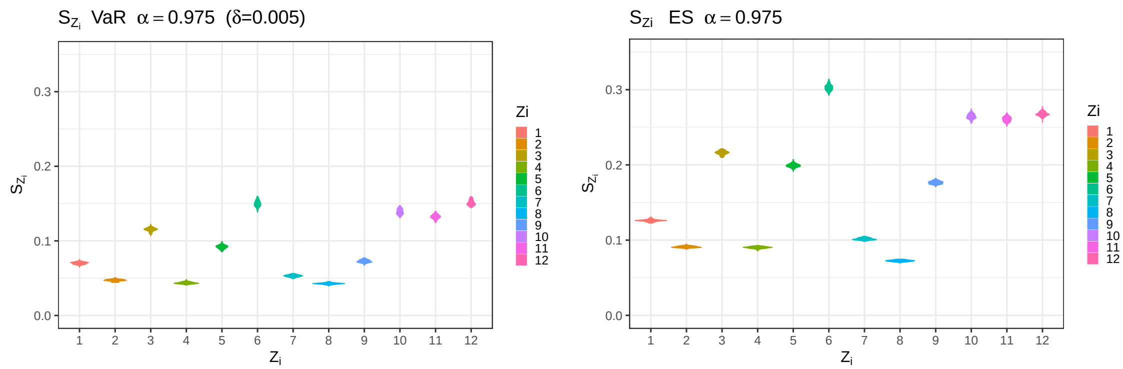

Figure 3 Marginal sensitivity to Z_is of VaR (left; with $\delta = 0.005$) and ES (right) with $\alpha = 0.975$.

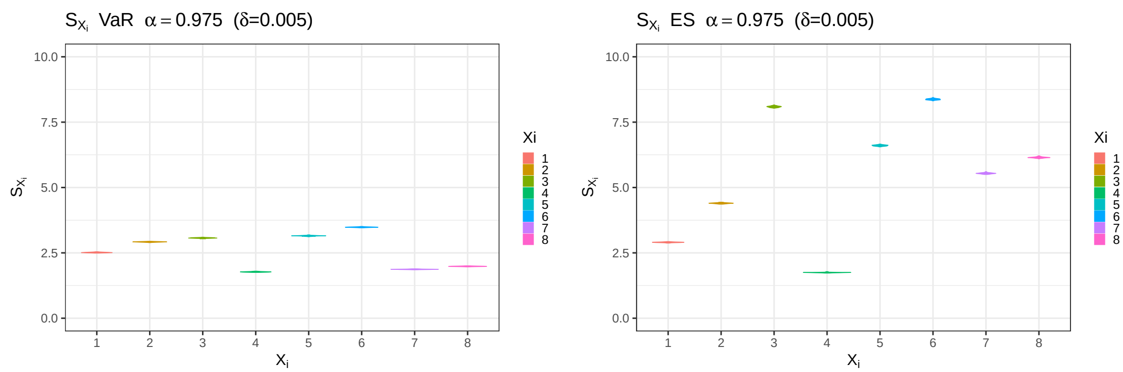

Figure 4 Marginal sensitivity to X_is of VaR (left; with $\delta = 0.005$) and ES (right; with $\delta = 0.005$) with $\alpha = 0.975$.

In Figure 3, where the sensitivities to the Z_is are plotted, we observe that business line 6 has a large sensitivity for VaR, the LoB with the largest CoV, see Table 2. In the right panel we observe that for ES the sensitivities are ordered similarly to the CoV of the business lines, but with a larger spread compared to the case of VaR – this could be attributed to the higher tail-sensitivity of the ES measure. Indeed, LoB 6 has the largest sensitivity, followed by 10, 11, and 12, which all have the same, second largest, CoV. Furthermore, LoB 2, 4, 7, and 8, which have the smallest CoVs, have small sensitivities for both VaR and ES.

In Figure 4, we depict the sensitivities to the X_is. A similar picture emerges, with the sensitivities for VaR (left panel) being very close together and for ES (right panel) being more spread-out. For ES, we observe that reinsurer 3, which has a layer on LoBs 5 and 6, and reinsurer 6, which has a layer on LoBs 11 ad 12, have the largest sensitivities. These LoBs have large sensitivities for ES, as seen in Figure 3 (right panel). Thus, a default of these reinsurers would naturally have a large impact on the ES of the total loss. We also see that reinsurers 7 and 8 have large sensitivities for ES. This is in line with expectations, since reinsurer 7 and 8 take on the highest layers (between the 85% and 95% quantile) of 6 business lines each. Nonetheless, this concentration effect is not picked up by the VaR sensitivity, as in the left panel the sensitivities to X_7, X_8 are rather low. This points to the importance of selecting an appropriately tail-sensitive risk measure.

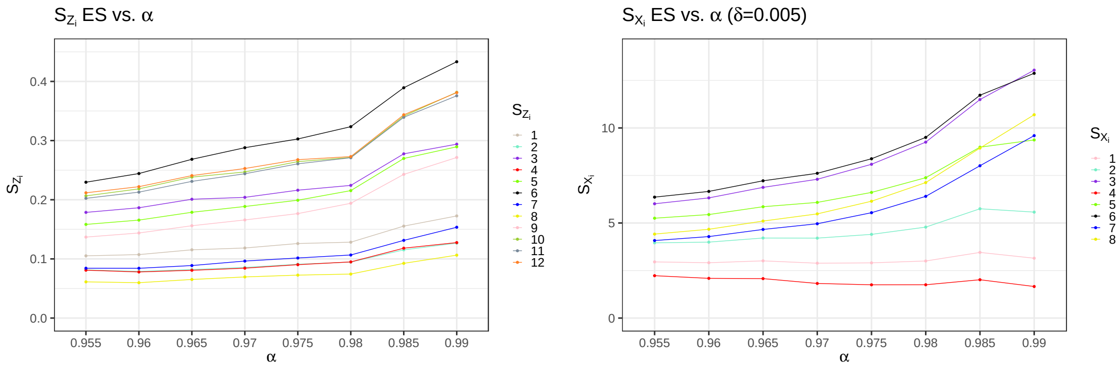

Figure 5 Marginal sensitivities of ES and different choices of α between 0.955 and 0.99. Left: sensitivity to Z_is. Right: sensitivity to X_is.

Figure 5 depicts the sensitivities of ES for different choices of α, from 0.955 to 0.99. The left panel contains the sensitivities to the LoBs (Z_i) and the right panel the sensitivities to the reinsurers (X_i). We observe that the ordering of the risk factors is mostly consistent with respect to changes in confidence level. An exception to this are the sensitivities to X_7 and X_8 which increase faster (relative to others) with α, as seen by the line crossings on the right panel. Once again this demonstrates the increased impact of default risk concentration at high loss quantiles.

Finally, in Figure 6 top panels, we show the sensitivities to X_is for VaR (left panel) and ES (right panel) with $\alpha = 0.975$, using different choices of δ for approximating the expectation conditional on $\{X_i = d_i\}$. We observe that the estimates are very stable for different choices of δ. Furthermore, in the bottom panels of Figure 6 we plot the standard deviation of the sensitivity estimators, thus choosing $\delta = 0.005$ provides a suitable bias and variance trade-off.

6. Conclusion

Taking derivatives of model outputs in the direction of inputs is a foundational process for interpreting complex computational models. However, differential sensitivity measures typically require stringent assumptions on differentiability and Lipschitz continuity of the model function. This severely limits the scope of current methods of differential sensitivity analysis. We address the problem by noting that, when inputs are uncertain – as is the case in settings ranging from Monte Carlo simulation to algorithmic prediction – a global view can be more appropriate than a local one. For a global assessment, differentiation is required across the entire input space; but then, it is not the derivative of the model function as such that is of primary interest, but rather the

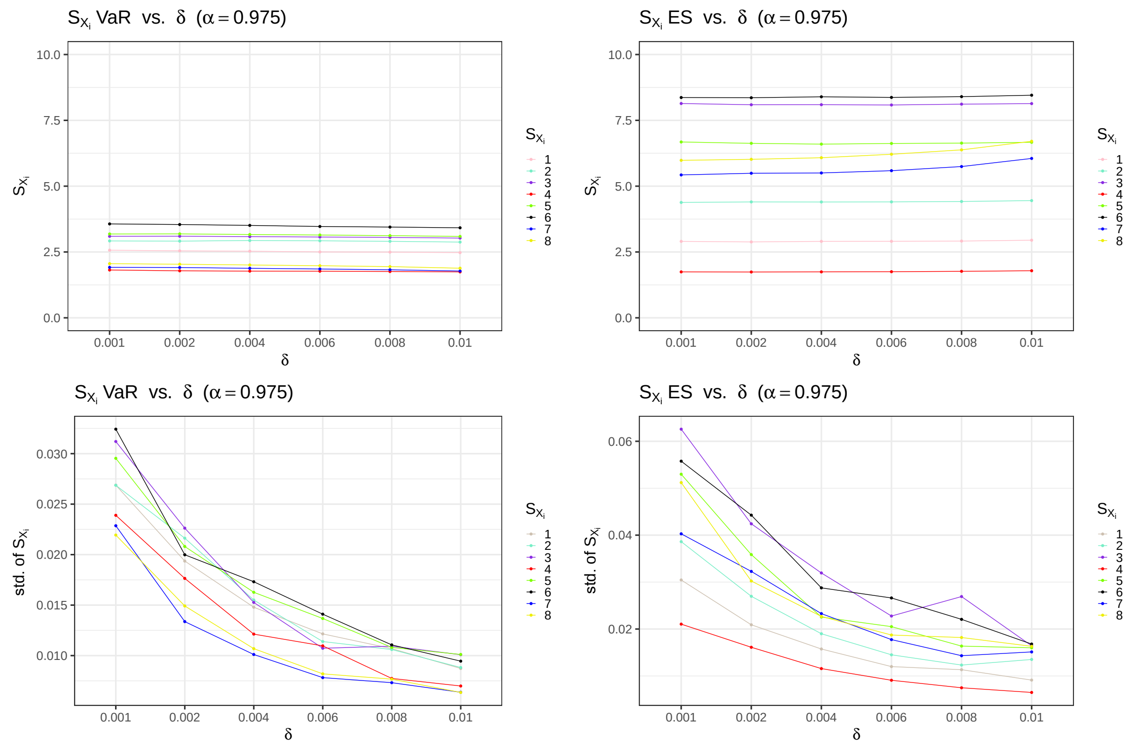

Figure 6 Marginal sensitivities to X_i and their sample standard deviations for different choices of δ. Top panels: Marginal sensitivity for VaR (left) and for ES (right) both with $\alpha = 0.975$. Bottom panels: Sample standard deviation of the sensitivity estimators for VaR (left) and ES (right).

derivative of a statistical functional of the output. Still, extent literature on sensitivity analysis of risk measures typically requires differentiability of the model aggregation function.

In this paper, we overcome current limitations in the literature and derive expressions for derivatives of quantile-based risk measures of model outputs, in a general setting where aggregation functions contain step functions and thus are not Lipschitz continuous. The conditions we require are rather weak and the sensitivity measures obtained admit representations as conditional expected values, which allows their estimation by standard methods. There are multiple potential applications of our methodology. We demonstrate applications in the area of credit risk modelling, but also in assessing sensitivity with respect to discrete random inputs. While our work is applicable in principle to discontinuous (e.g., tree-based) predictive models, addressing the idiosyncratic challenges of such exercises remains a topic for future work.

Acknowledgments

The authors would like to thank Wen Yuan (William) Tang for his help in implementing the numerical examples. SP gratefully acknowledges the support of the Canadian Statistical Sciences Institute (CANSSI) and the Natural Sciences and Engineering Research Council of Canada (NSERC) with funding reference numbers DGECR-2020-00333 and RGPIN-2020-04289.

Appendix A: Additional Sensitivity Formulas

Here we provide additional results for marginal and cascade sensitivities, which are omitted from the main body of the text for reasons of concision. In Section A.1 we deal with cascade sensitivities of VaR, while in Section A.2 we present results for a more general model than (1).

A.1. Cascade Sensitivities to VaR

Here we report the cascade sensitivity formulas for VaR.

THEOREM 6 **(Cascade Sensitivity VaR to X_i).** *Let Assumptions 1, 2 and 3 (for $Y = X_i$) be fulfilled for the stressed model $L_\varepsilon^{\mathbf{\Psi}}(X_i)$ and given $\alpha \in (0,1)$. Then, the cascade sensitivity for* VaR_α *to input X_i is given by,*

$$\mathcal{C}_{X_i}[\mathrm{VaR}_\alpha] = \sum_{j\in\mathcal{M}} \mathcal{C}_{X_i,X_j} + \sum_{k\in\mathcal{N}} \mathcal{C}_{X_i,Z_k},$$

where for all $k \in \mathcal{N}$,

$$\mathcal{C}_{X_i,Z_k} = \sum_{j\in\mathcal{M}} \mathbb{E}\left[\mathfrak{K}(X_i)\,\partial_k g_j(\boldsymbol{Z})\,\Psi_1^{(m+k)}(X_i,\boldsymbol{V})\mathbb{1}_{\{X_j\le d_j\}} \;\Big|\; L = q_\alpha\right],$$

and for $j \in \mathcal{M}$,

$$\mathcal{C}_{X_i,X_j} = \frac{c(\kappa;j) f_j(d_j)}{f(q_\alpha)}\;\mathbb{E}\left[\mathfrak{K}^{-1}(X_i)\Psi_1^{(j)}(X_i,\boldsymbol{V})\left(\mathbb{1}_{\{L\le q_\alpha + c(\kappa;j)g_j(\boldsymbol{Z})\}} - \mathbb{1}_{\{L\le q_\alpha\}}\right) \;\Big|\; X_j = d_j\right].$$

THEOREM 7 **(Cascade Sensitivity VaR to Z_i).** *Let Assumptions 1, 2 and 3 (for $Y = Z_i$) be fulfilled for the stressed model $L_\varepsilon^{\mathbf{\Psi}}(Z_i)$ and given $\alpha \in (0,1)$. Then, the cascade sensitivity for* VaR_α *to input Z_i is given by,*

$$\mathcal{C}_{Z_i}[\mathrm{VaR}_\alpha] = \sum_{j\in\mathcal{M}} \mathcal{C}_{Z_i,X_j} + \sum_{k\in\mathcal{N}} \mathcal{C}_{Z_i,Z_k},$$

where for all $k \in \mathcal{N}$,

$$\mathcal{C}_{Z_i,Z_k} = \sum_{j\in\mathcal{M}} \mathbb{E}\left[\mathfrak{K}(Z_i)\,\partial_k g_j(\boldsymbol{Z})\,\Psi_1^{(m+k)}(Z_i,\boldsymbol{V})\mathbb{1}_{\{X_j\le d_j\}} \;\Big|\; L = q_\alpha\right],$$

and for $j \in \mathcal{N}$,

$$\mathcal{C}_{Z_i,X_j} = \frac{c(\kappa;j) f_j(d_j)}{f(q_\alpha)}\;\mathbb{E}\left[\mathfrak{K}^{-1}(Z_i)\Psi_1^{(j)}(Z_i,\boldsymbol{V})\left(\mathbb{1}_{\{L\le q_\alpha + c(\kappa;j)g_j(\boldsymbol{Z})\}} - \mathbb{1}_{\{L\le q_\alpha\}}\right) \;\Big|\; X_j = d_j\right].$$

A.2. Sensitivity to General Loss Models

In this section, we generalise the loss model (1) to include cases where the functions g_j depends on both $\boldsymbol{Z}$ and $\boldsymbol{X}$. Specifically, we let

$$L := \sum_{j \in \mathcal{M}} g_j(\boldsymbol{X}, \boldsymbol{Z}) \mathbb{1}_{\{X_j \leq d_j\}}, \tag{9}$$

where a stress on X_i results in the stressed loss model

$$L_\varepsilon(X_i) := \sum_{\substack{j \neq i \\ j \in \mathcal{M}}} g_j(X_{i,\varepsilon}, \boldsymbol{X}_{-i}, \boldsymbol{Z}) \mathbb{1}_{\{X_j \leq d_j\}} + g_i(X_{i,\varepsilon} \boldsymbol{X}_{-i}, \boldsymbol{Z}) \mathbb{1}_{\{X_{i,\varepsilon} \leq d_j\}}\,. \tag{10}$$

We only present the sensitivities to X_i, as the sensitivities to Z_i are not impacted by the model generalisation. We observe in the next result that the marginal sensitivity to X_i of the loss model (9) accounts for both the stress in the indicator and the stress via the functions g_j, $j \in \mathcal{M}$.

THEOREM 8 **(Marginal Sensitivity – General Loss Model)**. *Let Assumptions 1 and 2 be fulfilled for the loss model (9) and for fixed $\alpha \in (0,1)$. Then, the marginal sensitivity for* VaR *to X_i for loss model (9) is*

$$\begin{aligned} \mathcal{S}_{X_i}[\mathrm{VaR}_\alpha] = \sum_{j \in \mathcal{M}} \Big(\mathbb{E}\Big[\mathfrak{K}(X_i) \partial_i g_j(\boldsymbol{X}, \boldsymbol{Z}) \mathbb{1}_{\{X_j \leq d_j\}} \;\Big|\; L = q_\alpha \Big] \Big) \\ + c(\kappa) \mathfrak{K}^{-1}(d_i) \frac{f_i(d_i)}{f(q_\alpha)} \mathbb{E}\Big[\Big(\mathbb{1}_{\{L \leq q_\alpha + c(\kappa) g_i(\boldsymbol{X}, \boldsymbol{Z})\}} - \mathbb{1}_{\{L \leq q_\alpha\}} \Big) \;\Big|\; X_i = d_i \Big]\,. \end{aligned}$$

The marginal sensitivity for ES to X_i for the loss model (9) is

$$\begin{aligned} \mathcal{S}_{X_i}[\mathrm{ES}_\alpha] = \sum_{j \in \mathcal{M}} \Big(\mathbb{E}\Big[\mathfrak{K}(X_i)\, \partial_i g_j(\boldsymbol{X}, \boldsymbol{Z})\, \mathbb{1}_{\{X_j \leq d_j\}} \;\Big|\; L \geq q_\alpha \Big] \Big) \\ - \frac{c(\kappa) \mathfrak{K}^{-1}(d_i) f_i(d_i)}{1 - \alpha} \mathbb{E}\Big[(L - c(\kappa) g_i(\boldsymbol{X}, \boldsymbol{Z}) - q_\alpha)_+ - (L - q_\alpha)_+ \;\Big|\; X_i = d_i \Big]\,. \end{aligned}$$

THEOREM 9 **(Cascade Sensitivity VaR – General Loss Model)**. *Let Assumptions 1, 2, and 3 (for $Y = X_i$) be fulfilled for the stressed model (10) and given $\alpha \in (0,1)$. Then, the cascade sensitivity for* VaR_α *to input X_i is given by*

$$\mathcal{C}_{X_i}[\mathrm{VaR}_\alpha] = \sum_{j \in \mathcal{M}} \mathcal{C}_{X_i, X_j} + \sum_{k \in \mathcal{N}} \mathcal{C}_{X_i, Z_k},$$

where for all $k \in \mathcal{N}$,

$$\mathcal{C}_{X_i, Z_k} = \sum_{j \in \mathcal{M}} \mathbb{E}\Big[\mathfrak{K}(X_i)\, \partial_{m+k} g_j(\boldsymbol{X}, \boldsymbol{Z})\, \Psi_1^{(m+k)}(X_i, \boldsymbol{V}) \mathbb{1}_{\{X_j \leq d_j\}} \;\Big|\; L = q_\alpha \Big]\,,$$

and for $j \in \mathcal{M}$,

$$\begin{aligned} \mathcal{C}_{X_i, X_j} = \Big(\sum_{r=1}^{m} \mathbb{E}\Big[\mathfrak{K}(X_i) \partial_j g_r(\boldsymbol{X}, \boldsymbol{Z}) \Psi_1^{(j)}(X_i, \boldsymbol{V}) \mathbb{1}_{\{X_r \leq d_r\}} \;\Big|\; L = q_\alpha \Big] \Big) \\ + c(\kappa; j) \frac{f_j(d_j)}{f(q_\alpha)} \mathbb{E}\Big[\mathfrak{K}^{-1}(X_i) \Psi_1^{(j)}(X_i, \boldsymbol{V}) \left(\mathbb{1}_{\{L \leq q_\alpha + c(\kappa; j) g_i(\boldsymbol{X}, \boldsymbol{Z})\}} - \mathbb{1}_{\{L \leq q_\alpha\}} \right) \;\Big|\; X_j = d_j \Big]\,. \end{aligned}$$

Theorem 10 **(Cascade Sensitivity ES – General Loss model)**. *Let Assumptions 1, 2, and 3 (for $Y = X_i$) be fulfilled for the stressed model (10) and given $\alpha \in (0,1)$. Then, the cascade sensitivity for ES_α to input X_i is given by*

$$\mathcal{C}_{X_i}\left[\mathrm{ES}_\alpha\right] = \sum_{j\in\mathcal{M}} \mathcal{C}_{X_i,X_j} + \sum_{k\in\mathcal{N}} \mathcal{C}_{X_i,Z_k},$$

where for all $k \in \mathcal{N}$,

$$\mathcal{C}_{X_i,Z_k} = \sum_{j\in\mathcal{M}} \mathbb{E}\left[\mathfrak{K}(X_i)\, \partial_{m+k} g_j(\boldsymbol{X},\boldsymbol{Z})\, \Psi_1^{(m+k)}(X_i,\boldsymbol{V}) \mathbb{1}_{\{X_j\le d_j\}} \;\middle|\; L \ge q_\alpha\right],$$

and for $j \in \mathcal{M}$,

$$\begin{aligned}\mathcal{C}_{X_i,X_j} = &\sum_{r=1}^{m}\left(\mathbb{E}\left[\mathfrak{K}(X_i)\partial_j g_r(\boldsymbol{X},\boldsymbol{Z})\Psi_1^{(j)}(X_i,\boldsymbol{V})\mathbb{1}_{\{X_r\le d_r\}} \;\middle|\; L\ge q_\alpha\right]\right)\\ &- c(\kappa;j)\frac{f_j(d_j)}{1-\alpha}\,\mathbb{E}\left[\mathfrak{K}^{-1}(X_i)\Psi_1^{(j)}(X_i,\boldsymbol{V})\Big((L-c(\kappa;j)g_j(\boldsymbol{X},\boldsymbol{Z})-q_\alpha)_+ - (L-q_\alpha)_+\Big) \;\middle|\; X_j = d_j\right].\end{aligned}$$

Appendix B: Proofs

B.1. Proof of Mixture Stress

Proof of mixture stress stress properties. We prove that the mixture stress in Table 1 is well-defined. First, the stress function and its inverse are given by

$$\kappa_\varepsilon(x) = F_{i,\varepsilon}^{-1}\left(F_i(x)\right) \quad \text{and} \quad \kappa_\varepsilon^{-1}(x) = F_i^{-1}\left(F_{i,\varepsilon}(x)\right),$$

where $F_{i,\varepsilon} = (1-\varepsilon)F_i + \varepsilon G$. By construction, it holds that $F_{i,\varepsilon}(x)$ is continuous and strictly increasing in x for all $\varepsilon \ge 0$. Furthermore, $F_{i,\varepsilon}$ and $F_{i,\varepsilon}^{-1}$ converge pointwise, as $\varepsilon \searrow 0$, to F_i and F_i^{-1} respectively, thus the stress function fulfils *ii)* and *iii)*. Next, if $G(x) \le F_i(x)$ for all $x \in \mathbb{R}$, then $F_{i,\varepsilon}(x) = F_i(x) + \varepsilon\left(G(x) - F_i(x)\right) \le F_i(x)$, and therefore $F_{i,\varepsilon}^{-1}(u) \ge F_i^{-1}(u)$ for all $u \in [0,1]$. Thus $\kappa_\varepsilon(x) = F_{i,\varepsilon}^{-1}\left(F_i(x)\right) \ge x$ and the stress function fulfils *iv) (a)*. The case when the stress function fulfils *iv) (b)* if $G(x) \ge F_i(x)$ for all $x \in \mathbb{R}$, follows similarly.

Now, to check property *v)*, note that

$$\mathfrak{K}(x) = \lim_{\varepsilon\searrow 0} \tfrac{1}{\varepsilon}\left(\kappa_\varepsilon(x) - x\right) = \lim_{\varepsilon\searrow 0}\tfrac{1}{\varepsilon}\left(F_{i,\varepsilon}^{-1}\big(F_i(x)\big) - F_i^{-1}\big(F_i(x)\big)\right) = \frac{\partial}{\partial\varepsilon}F_{i,\varepsilon}^{-1}\left(F_i(x)\right)\Big|_{\varepsilon=0}.$$

Further, from the relation $\frac{\partial}{\partial\varepsilon}F_{i,\varepsilon}^{-1}(x)|_{\varepsilon=0} = -\frac{\frac{\partial}{\partial\varepsilon}F_{i,\varepsilon}(y)}{f_i(y)}|_{y=F_i^{-1}(x)} = \frac{F_i(y)-G(y)}{f_i(y)}|_{y=F_i^{-1}(x)}$, it follows that

$$\mathfrak{K}(x) = \frac{F_i(x) - G(x)}{f_i(x)}.$$

To obtain the expression for *vi)*, note that the stress function $\kappa_\varepsilon(x)$ is differentiable in x, for all ε in a neighbourhood of 0, and $\frac{\partial}{\partial x}\kappa_\varepsilon(x)|_{\varepsilon=0} \ne 0$, then property *vi)* is satisfied with $\mathfrak{K}^{-1}(x) = \frac{-\mathfrak{K}(x)}{\frac{\partial}{\partial x}\kappa_\varepsilon(x)|_{\varepsilon=0}}$, which follows immediately from an application of the chain rule applied to the identity $\kappa_\varepsilon^{-1}(\kappa_\varepsilon(x)) = x$. Moreover,

$$\frac{\partial}{\partial x}\kappa_\varepsilon(x) = -\frac{f_i(x)}{\frac{\partial}{\partial x}F_{i,\varepsilon}(F_{i,\varepsilon}^{-1}(F_i(x)))|_{\varepsilon=0}} = -1,$$

which concludes the proof. ∎

B.2. Proofs of Marginal Sensitivity: Theorems 1 and 2

For the proof of the marginal sensitivities to VaR and ES, we need the following lemma concerning sequences of functions that converge weakly to a Dirac delta function. For this, we first write the marginally stressed portfolios as

$$L(Z_{i,\varepsilon}) = L + \sum_{k=1}^{m} \Delta_\varepsilon g_k \quad \text{and}$$
$$L(X_{i,\varepsilon}) = L + g_i(\boldsymbol{Z}) \left(\mathbb{1}_{\{X_{i,\varepsilon} \leq d_i\}} - \mathbb{1}_{\{X_i \leq d_i\}} \right),$$

where we define $\Delta_\varepsilon g_k := (g_k(\boldsymbol{Z}_{-i}, \kappa_\varepsilon(Z_i)) - g_k(\boldsymbol{Z}))\mathbb{1}_{\{X_k \leq d_k\}}$. When the stress is clear from the context, we write $L_\varepsilon = L(Z_{i,\varepsilon})$ and $L_\varepsilon = L(X_{i,\varepsilon})$.

Lemma 1. *For fixed $d \in \mathbb{R}$, define the family of functions*

$$\delta_\varepsilon(x) = \frac{\left| \mathbb{1}_{\{\kappa_\varepsilon(x) \leq d\}} - \mathbb{1}_{\{x \leq d\}} \right|}{\varepsilon}, \quad x \in \mathbb{R}, \quad \varepsilon > 0.$$

Then, δ_ε converges weakly to a scaled Dirac delta function at d for $\varepsilon \searrow 0$. Moreover, for any family of measurable functions $h_\varepsilon\colon \mathbb{R}^{m+n} \to \mathbb{R}$ such that $\lim_{\varepsilon \searrow 0} \mathbb{E}[|h_\varepsilon(\boldsymbol{X}, \boldsymbol{Z})|] < \infty$, the following holds:

$$\lim_{\varepsilon \searrow 0} \mathbb{E}\left[\delta_\varepsilon(X_i)\, h_\varepsilon(\boldsymbol{X}, \boldsymbol{Z})\right] = -c(\kappa) \mathfrak{K}^{-1}(d)\, f_i(d)\, \mathbb{E}\left[\, h_0(\boldsymbol{X}, \boldsymbol{Z}) \mid X_i = d\,\right],$$

where $c(\kappa)$ is given in (2), and $h_0(\boldsymbol{x}, \boldsymbol{z}) = \lim_{\varepsilon \searrow 0} h_\varepsilon(\boldsymbol{x}, \boldsymbol{z})$.

Proof of Lemma 1. First note that

$$\left| \mathbb{1}_{\{\kappa_\varepsilon(x) \leq d\}} - \mathbb{1}_{\{x \leq d\}} \right| = -c(\kappa) \left(\mathbb{1}_{\{\kappa_\varepsilon(x) \leq d\}} - \mathbb{1}_{\{x \leq d\}} \right). \tag{11}$$

Let ξ be an infinitely often differentiable function. Using the change of variable $y = \kappa_\varepsilon(x)$, we obtain

$$\begin{aligned}\int_{-\infty}^{+\infty} \xi(x)\delta_\varepsilon(x)\, dx &= -\frac{c(\kappa)}{\varepsilon} \int_{-\infty}^{+\infty} \xi(x) \left(\mathbb{1}_{\{\kappa_\varepsilon(x) \leq d\}} - \mathbb{1}_{\{x \leq d\}} \right) dx \\ &= -\frac{c(\kappa)}{\varepsilon} \int_{-\infty}^{+\infty} \frac{\xi(z)}{\frac{\partial}{\partial x}\kappa_\varepsilon(z)} \Big|_{z = \kappa_\varepsilon^{-1}(y)} \mathbb{1}_{\{y \leq d\}}\, dy - \frac{1}{\varepsilon}\int_{-\infty}^{d} \xi(x)\, dx\,.\end{aligned}$$

Letting Ξ be a primitive of ξ vanishing at $-\infty$, then

$$\begin{aligned}\int_{-\infty}^{+\infty} \xi(x)\delta_\varepsilon(x)\, dx &= -\frac{c(\kappa)}{\varepsilon} \left(\int_{-\infty}^{d} \frac{d}{dy} \Xi(\kappa_\varepsilon^{-1}(y))\, dy - \Xi(d) \right) \\ &= -\frac{c(\kappa)}{\varepsilon} \left(\Xi(\kappa_\varepsilon^{-1}(d)) - \Xi(d) \right).\end{aligned}$$

Taking the limit as $\varepsilon \to 0$, we obtain that

$$\lim_{\varepsilon \searrow 0} \int_{-\infty}^{+\infty} \xi(x)\delta_\varepsilon(x)\, dx = -c(\kappa)\, \xi(d) \mathfrak{K}^{-1}(d)\,.$$

For the second part of the statement, note that

$$\begin{aligned}\lim_{\varepsilon\searrow 0}\mathbb{E}\left[\delta_\varepsilon(X_i)\,h_\varepsilon(\boldsymbol{X},\boldsymbol{Z})\right]&=\lim_{\varepsilon\searrow 0}\int_{\mathbb{R}^{m+n}}\delta_\varepsilon(x_i)\,h_\varepsilon(\boldsymbol{x},\boldsymbol{z})\,f_{\boldsymbol{X},\boldsymbol{Z}}(\boldsymbol{x},\boldsymbol{z})\,d\boldsymbol{x}\,d\boldsymbol{z}\\&=-c(\kappa)\,\mathfrak{K}^{-1}(d)\int_{\mathbb{R}^{m+n-1}}h_\varepsilon(\boldsymbol{x}_{-i},d,\boldsymbol{z})\,f_{\boldsymbol{X},\boldsymbol{Z}}(\boldsymbol{x}_{-i},d,\boldsymbol{z})\frac{f_i(d)}{f_i(d)}\,d\boldsymbol{x}_{-i}\,d\boldsymbol{z}\\&=-c(\kappa)\,\mathfrak{K}^{-1}(d)\,f_i(d)\mathbb{E}\left[\,h_0(\boldsymbol{X},\boldsymbol{Z})\mid X_i=d\,\right].\end{aligned}$$

□

Lemma 2. *For fixed $0<\alpha<1$ and $\boldsymbol{z}\in\mathbb{R}^n$, define the sequence of functions*

$$\delta_\varepsilon(l)=\frac{\mathbb{1}_{\left\{l\leq q_\alpha-\sum_{k=1}^m\Delta_\varepsilon g_k\right\}}-\mathbb{1}_{\{l\leq q_\alpha\}}}{\varepsilon}\qquad l\in\mathbb{R},\quad \varepsilon>0.$$

where $\Delta_\varepsilon g_k=(g_k(\boldsymbol{z}_{-i},\kappa_\varepsilon(z_i))-g_k(\boldsymbol{z}))\mathbb{1}_{\{x_k\leq d_k\}}$, $\boldsymbol{z}\in\mathbb{R}^n$, and $l\geq 0$. Then, δ_ε converges weakly to a scaled Dirac delta function at q_α for $\varepsilon\searrow 0$. Moreover, for any family of measurable functions $h_\varepsilon\colon\mathbb{R}^{m+n}\to\mathbb{R}$ such that $\lim_{\varepsilon\searrow 0}\mathbb{E}\left[|h_\varepsilon(\boldsymbol{X},\boldsymbol{Z})|\right]<\infty$, the following holds:

$$\lim_{\varepsilon\searrow 0}\mathbb{E}\left[\delta_\varepsilon(L)h_\varepsilon(\boldsymbol{X},L)\right]=-f(q_\alpha)\sum_{k=1}^m\mathbb{E}\left[\mathfrak{K}(Z_i)\partial_i\,g_k(\boldsymbol{Z})\mathbb{1}_{\{X_k\leq d_k\}}\,h_0(\boldsymbol{X},L)\,\middle|\,L=q_\alpha\right].\tag{12}$$

Proof of Lemma 2. Let $\xi(\cdot)$ be an infinitely often differentiable function. Applying Taylor's Theorem of g_k around z_i, and then using $\kappa_\varepsilon(z_i)=z_i+\varepsilon\mathfrak{K}(z_i)+o(\varepsilon)$, we obtain for all $k=1,\ldots,n$, that

$$\begin{aligned}g_k(\boldsymbol{z}_{-i},\kappa_\varepsilon(z_i))-g_k(\boldsymbol{z})&=\left(\kappa_\varepsilon(z_i)-z_i\right)\partial_i\,g_k(\boldsymbol{z})+o\left(\kappa_\varepsilon(z_i)-z_i\right)\\&=\varepsilon\mathfrak{K}(z_i)\,\partial_i\,g_k(\boldsymbol{z})+o\left(\varepsilon\right),\end{aligned}\tag{13}$$

where $\partial_i\,g_k(\boldsymbol{z})=\frac{\partial}{\partial z_i}g_k(\boldsymbol{z})$ is the derivative in the i^{th} component. Thus, we have that for all $\boldsymbol{z}\in\mathbb{R}^n$, using the Mean Value Theorem for some $l^*\in(q_\alpha,q_\alpha-\Delta_\varepsilon g]$ (or $l^*\in(q_\alpha-\Delta_\varepsilon g\,,q_\alpha]$) in the second equation, and then (13) that

$$\begin{aligned}\int_{-\infty}^{+\infty}\xi(l)\delta_\varepsilon(l)\,dl&=\frac{1}{\varepsilon}\int_{q_\alpha}^{q_\alpha-\sum_{k=1}^m\Delta_\varepsilon g_k}\xi(l)\,dl=-\frac{1}{\varepsilon}\sum_{k=1}^m\Delta_\varepsilon g_k\,\xi(l^*)\\&=-\left(\sum_{k=1}^m\mathfrak{K}(z_i)\,\partial_i\,g_k(\boldsymbol{z})\mathbb{1}_{\{x_k\leq d_k\}}+o\left(1\right)\right)\xi(l^*).\end{aligned}$$

Taking the limit for $\varepsilon\searrow 0$, we have

$$\lim_{\varepsilon\searrow 0}\int_{-\infty}^{+\infty}\xi(l)\delta_\varepsilon(l)\,dl=-\mathfrak{K}(z_i)\sum_{k=1}^m\partial_i\,g_k(\boldsymbol{z})\mathbb{1}_{\{x_k\leq d_k\}}\xi\left(q_\alpha\right).$$

Equation (12) follows using a similar argument as in Lemma 1. □

Proof of Theorem 1 (Marginal Sensitivity VaR) By Proposition 2.3 in Embrechts and Hofert (2013) it holds for all $\varepsilon \geq 0$ that $F_\varepsilon\left(q_\varepsilon(\alpha)\right) = \alpha$. Taking derivative with respect to ε and evaluating at $\varepsilon = 0$ (note that Assumptions 1 and 2 are fulfilled), we obtain

$$f\left(q_\alpha\right) \left.\tfrac{\partial}{\partial \varepsilon} q_\varepsilon(\alpha)\right|_{\varepsilon=0} + \tfrac{\partial}{\partial \varepsilon} F_\varepsilon(q_\alpha) = 0 \quad \text{and} \quad \left.\frac{\partial}{\partial \varepsilon} q_\varepsilon(\alpha)\right|_{\varepsilon=0} = -\frac{1}{f\left(q_\alpha\right)} \frac{\partial}{\partial \varepsilon} F_\varepsilon(q_\alpha)\,, \tag{14}$$

whenever $\frac{\partial}{\partial \varepsilon} F_\varepsilon(q_\alpha(\alpha))$ exists. Next, we show that the derivative of F_ε with respect to ε exists.
Part 1: We first consider the case of stressing X_i and calculate

$$\begin{aligned} F_\varepsilon(q_\alpha) - F(q_\alpha) &= \mathbb{P}\left(L_\varepsilon \leq q_\alpha\right) - \mathbb{P}\left(L \leq q_\alpha\right) & \text{(15a)} \\ &= \mathbb{E}\left[\mathbb{1}_{\left\{L \leq q_\alpha - g_i(\boldsymbol{Z})\left(\mathbb{1}_{\{\kappa_\varepsilon(X_i) \leq d_i\}} - \mathbb{1}_{\{X_i \leq d_i\}}\right)\right\}} - \mathbb{1}_{\{L \leq q_\alpha\}}\right] & \text{(15b)} \\ &= \mathbb{E}\left[\left|\mathbb{1}_{\{\kappa_\varepsilon(X_i) \leq d_i\}} - \mathbb{1}_{\{X_i \leq d_i\}}\right| \left(\mathbb{1}_{\{L \leq q_\alpha + c(\kappa) g_i(\boldsymbol{Z})\}} - \mathbb{1}_{\{L \leq q_\alpha\}}\right)\right], & \text{(15c)} \end{aligned}$$

where the last equality follows from (11). Invoking Lemma 1 we obtain

$$\frac{\partial}{\partial \varepsilon} F_\varepsilon(q_\alpha) = -c(\kappa) \mathfrak{K}^{-1}(d_i)\, f_i(d_i) \;\; \mathbb{E}\left[\left(\mathbb{1}_{\{L \leq q_\alpha + c(\kappa) g_i(\boldsymbol{Z})\}} - \mathbb{1}_{\{L \leq q_\alpha\}}\right) \;\middle|\; X_i = d_i\right].$$

Combining with Equation (14) concludes the first part.

Part 2: Next, we consider the case of stressing Z_i. For this, it holds that

$$F_\varepsilon(q_\alpha) - F(q_\alpha) = \mathbb{E}\left[\left(\mathbb{1}_{\left\{L \leq q_\alpha - \sum_{k=1}^m \Delta_\varepsilon g_k\right\}} - \mathbb{1}_{\{L \leq q_\alpha\}}\right)\right].$$

Applying Lemma 2 and Equation (14) conclude the proof. □

Proof of Theorem 2 (Marginal Sensitivity ES). We first calculate the sensitivity to X_i, and in a second step to Z_i.

Part 1: To calculate the sensitivity to X_i, we observe that

$$\begin{aligned} \frac{\mathrm{ES}_\alpha(L_\varepsilon) - \mathrm{ES}_\alpha(L)}{\varepsilon} &= \frac{1}{\varepsilon(1-\alpha)} \mathbb{E}\left[(L_\varepsilon - q_\varepsilon)_+ - (L - q_\alpha)_+\right] + \frac{q_\varepsilon - q_\alpha}{\varepsilon} \\ &= \underbrace{\frac{1}{\varepsilon(1-\alpha)} \mathbb{E}\left[(L_\varepsilon - q_\varepsilon)_+ - (L_\varepsilon - q_\alpha)_+\right]}_{:=A(\varepsilon)} + \underbrace{\mathbb{E}\left[(L_\varepsilon - q_\alpha)_+ - (L - q_\alpha)_+\right]}_{:=B(\varepsilon)} + \underbrace{\frac{q_\varepsilon - q_\alpha}{\varepsilon}}_{:=C(\varepsilon)}. \end{aligned} \tag{16}$$

To calculate the expectation in $A(\varepsilon)$, we use integration by parts in the third equation, and interpret $\int_b^a h(x)\,dx = -\int_a^b h(x)\,dx$, if $a < b$.

$$\begin{aligned} A(\varepsilon)\,\varepsilon(1-\alpha) &= \int_{q_\varepsilon}^{+\infty} (y - q_\varepsilon)\, dF_\varepsilon(y) - \int_{q_\alpha}^{+\infty} (y - q_\alpha)\, dF_\varepsilon(y) \\ &= \int_{q_\varepsilon}^{q_\alpha} y\, dF_\varepsilon(y) - q_\varepsilon\,(1-\alpha) + q_\alpha\,(1 - F_\varepsilon(q_\alpha)) \\ &= q_\alpha F_\varepsilon(q_\alpha) - q_\varepsilon\,\alpha - \int_{q_\varepsilon}^{q_\alpha} F_\varepsilon(y)\, dy - q_\varepsilon\,(1-\alpha) + q_\alpha\,(1 - F_\varepsilon(q_\alpha)) \\ &= (q_\alpha - q_\varepsilon) - \int_{q_\varepsilon}^{q_\alpha} F_\varepsilon(y)\, dy\,. \end{aligned}$$

Next, we collect parts $A(\varepsilon)$ and $C(\varepsilon)$, and use the Mean Value Theorem, that is there exists a $q^* \in (q_\varepsilon, q_\alpha]$ (or $q^* \in (q_\alpha, q_\varepsilon]$, if $q_\alpha < q_\varepsilon$) such that $\int_{q_\varepsilon}^{q_\alpha} F_\varepsilon(y)\, dy = (q_\alpha - q_\varepsilon)\, F_\varepsilon(q^*)$. Thus, we obtain

$$A(\varepsilon) + C(\varepsilon) = \frac{1}{\varepsilon(1-\alpha)}\big((q_\alpha - q_\varepsilon)\,(1 - F_\varepsilon(q^*)) + \frac{q_\varepsilon - q_\alpha}{\varepsilon}$$
$$= \frac{(q_\varepsilon - q_\alpha)}{\varepsilon}\left(1 - \frac{1 - F_\varepsilon(q^*)}{1-\alpha}\right).$$

Talking the limit for $\varepsilon \searrow 0$, and noting that the derivative of the quantile function with respect to ε exists by Theorem 1, we obtain

$$\lim_{\varepsilon \searrow 0} A(\varepsilon) + C(\varepsilon) = 0\,.$$

Next, for part $B(\varepsilon)$ we obtain using Equation (11)

$$B(\varepsilon) = \frac{1}{\varepsilon(1-\alpha)}\mathbb{E}\left[\left(L + g_i(\boldsymbol{Z})\,\left(\mathbb{1}_{\{X_{i,\varepsilon} \le d_i\}} - \mathbb{1}_{\{X_i \le d_i\}}\right) - q_\alpha\right)_+ - (L - q_\alpha)_+\right]$$
$$= \frac{1}{\varepsilon(1-\alpha)}\mathbb{E}\left[\left|\mathbb{1}_{\{X_{i,\varepsilon} \le d_i\}} - \mathbb{1}_{\{X_i \le d_i\}}\right|\Big(\,(L - c(\kappa) g_i(\boldsymbol{Z})\, - q_\alpha)_+ - (L - q_\alpha)_+\Big)\right].$$

Applying Lemma 1, we obtain

$$\lim_{\varepsilon \searrow 0} B(\varepsilon) = \frac{-c(\kappa)\mathfrak{K}^{-1}(d_i)\, f_i(d_i)}{1-\alpha}\mathbb{E}\left[(L - c(\kappa) g_i(\boldsymbol{Z})\, - q_\alpha)_+ - (L - q_\alpha)_+ \;\mid\; X_i = d_i\right].$$

Part 2: For the sensitivity to Z_i, we write similarly to part 1,

$$\frac{\mathrm{ES}_\alpha(L_\varepsilon) - \mathrm{ES}_\alpha(L)}{\varepsilon} = A(\varepsilon) + B(\varepsilon) + C(\varepsilon)\,,$$

where $A(\varepsilon)$ and $C(\varepsilon)$ are the same as in (16), while $B(\varepsilon)$ is

$$B(\varepsilon) = \frac{1}{\varepsilon(1-\alpha)}\,\mathbb{E}\left[\left(L + \sum_{k=1}^{m} \Delta_\varepsilon g_k \, - q_\alpha\right)_+ - (L - q_\alpha)_+\right] \tag{17a}$$
$$= \frac{1}{\varepsilon(1-\alpha)}\,\mathbb{E}\left[(L - q_\alpha)\left(\mathbb{1}_{\{L \le q_\alpha\}} - \mathbb{1}_{\{L \le q_\alpha - \sum_{k=1}^m \Delta_\varepsilon g_k\}}\right) + \sum_{k=1}^{m} \Delta_\varepsilon g_k \mathbb{1}_{\{L \ge q_\alpha - \sum_{k=1}^m \Delta_\varepsilon g_k\}}\right], \tag{17b}$$

where in the last equality we used that $\mathbb{1}_{\{L > q_\alpha\}} = 1 - \mathbb{1}_{\{L \le q_\alpha\}}$. Note that the argument that $A(\varepsilon) + C(\varepsilon)$ converges to 0 for $\varepsilon \searrow 0$ only depends on the fact that F_ε converges to F for $\varepsilon \searrow 0$. Thus, also here, it holds that $\lim_{\varepsilon \searrow 0} A(\varepsilon) + C(\varepsilon) = 0$. To calculate the limit of $B(\varepsilon)$, we apply Lemma 2 to the first term, which turns out to be equal to zero. For the second term, note that $\frac{1}{\varepsilon}\Delta_\varepsilon g_k$ converges to $\mathfrak{K}(Z_i)\partial_i\, g_k(\boldsymbol{Z})\mathbb{1}_{\{X_k \le d_k\}}$ $\mathbb{P}$-a.s. for $\varepsilon \searrow 0$, see also Equation (13). Thus,

$$\lim_{\varepsilon \searrow 0} B(\varepsilon) = \frac{1}{1-\alpha}\sum_{k=1}^{m} \mathbb{E}\left[\mathfrak{K}(Z_i)\partial_i\, g_k(\boldsymbol{Z})\mathbb{1}_{\{X_k \le d_k\}}\mathbb{1}_{\{L \ge q_\alpha\}}\right]$$
$$= \sum_{k=1}^{m} \mathbb{E}\left[\mathfrak{K}(Z_i)\partial_i\, g_k(\boldsymbol{Z})\mathbb{1}_{\{X_k \le d_k\}} \;\mid\; L \ge q_\alpha\right].$$

□

B.3. Proof of Cascade Sensitivity: Theorems 3, 4, 6, and 7.

For the proofs of the cascade sensitivities to VaR and ES, we need the following lemmas concerning sequences of functions that converge weakly to Dirac delta functions. For this, we first provide a representation of the stressed loss, when stressing X_i.

For a stress function κ_ε and a Rosenblatt transform Ψ, we define for all $j \in \mathcal{M}$ and fixed $\mathbf{v}$,

$$a_{\varepsilon,j}(x) := |\mathbb{1}_{\{\eta_{\varepsilon,j}(x) \le d_j\}} - \mathbb{1}_{\{x \le d_j\}}|,$$

where $\eta_{\varepsilon,j}(x) := \Psi^{(j)}\left(\kappa_\varepsilon\left(\Psi^{(j),-1}(x,\mathbf{v})\right), \mathbf{v}\right)$ and $\Psi^{(j),-1}$ denotes the inverse in the first component of $\Psi^{(j)}$. Further, we let

$$A_{\varepsilon,j} := a_{\varepsilon,j}(X_j),$$

where it is implicit that $\mathbf{v}$ is replaced by $\boldsymbol{V}$. Note that $X_j = \Psi^{(j)}(X_i, \boldsymbol{V})$ $\mathbb{P}$-a.s., and therefore

$$\Psi^{(j)}\left(X_{i,\varepsilon}, \boldsymbol{V}\right) = \Psi^{(j)}\left(\kappa_\varepsilon(X_i), \boldsymbol{V}\right) = \Psi^{(j)}\left(\kappa_\varepsilon\left(\Psi^{(j),-1}(X_j, \boldsymbol{V})\right), \boldsymbol{V}\right) = \eta_{\varepsilon,j}(X_j) \quad \mathbb{P}\text{-a.s.}.$$

Lemma 3 **(Stressed Portfolio Loss)**. *For a stress $X_{i,\varepsilon}$, the stressed portfolio admits representation*

$$L^{\boldsymbol{\Psi}}(X_{i,\varepsilon}) = L + \sum_{k=1}^m \tilde{\Delta}_\varepsilon g_k - \sum_{j=1}^m c(\kappa;\, j) g_j(\boldsymbol{Z})\, A_{\varepsilon,j},$$

where $\tilde{\Delta}_\varepsilon\, g_k = \left(g_k\left(\Psi^{(\boldsymbol{Z})}(X_{i,\varepsilon}, \boldsymbol{V})\right) - g_k(\boldsymbol{Z})\right) \mathbb{1}_{\{\Psi^{(k)}(X_{i,\varepsilon}, \boldsymbol{V}) \le d_k\}}$.

Proof of Lemma 3. We obtain

$$\begin{aligned} L^{\boldsymbol{\Psi}}(X_{i,\varepsilon}) &= L + \sum_{k=1}^m \tilde{\Delta}_\varepsilon g_k + \sum_{j=1}^m g_j(\boldsymbol{Z})\, \left(\mathbb{1}_{\{\eta_{\varepsilon,j}(X_j) \le d_j\}} - \mathbb{1}_{\{X_j \le d_j\}}\right) \\ &= L + \sum_{k=1}^m \tilde{\Delta}_\varepsilon g_k - \sum_{j=1}^m c(\kappa;\, j) g_j(\boldsymbol{Z})\, A_{\varepsilon,j}, \end{aligned}$$

since by Assumption 3 it holds that $\mathbb{1}_{\{\eta_{\varepsilon,j}(x) \le d_j\}} - \mathbb{1}_{\{x \le d_j\}} = -c(\kappa;\, j) a_{\varepsilon,j}(x)$ for all $j \in \mathcal{M}$. ☐

Lemma 4. *Let $\mathcal{K} \subset \mathcal{M}$ and its complement $\mathcal{K}^\complement = \mathcal{M}/\mathcal{K}$ and define the sequence of functions*

$$\delta_\varepsilon^{\mathcal{K}}(\boldsymbol{x}) = \frac{1}{\varepsilon} \prod_{k \in \mathcal{K}} a_{\varepsilon,k}(x_k) \prod_{l \in \mathcal{K}^\complement} a^\complement_{\varepsilon,l}(x_l), \qquad \varepsilon > 0,$$

where $a^\complement_{\varepsilon,k}(x) = 1 - a_{\varepsilon,k}(x)$.

Then, for all functions $h_\varepsilon \colon \mathbb{R}^{m+n} \to \mathbb{R}$ such that $\lim_{\varepsilon \searrow 0} \mathbb{E}\left[|h_\varepsilon(\boldsymbol{X}, \boldsymbol{Z})|\right] < \infty$, the following holds:

i) if $\mathcal{K}$ contains one element, $\mathcal{K} = \{k\}$, then

$$\lim_{\varepsilon \searrow 0} \mathbb{E}\left[\delta_\varepsilon^{\mathcal{K}}(\boldsymbol{X})\, h_\varepsilon(\boldsymbol{X}, \boldsymbol{Z})\right] = -c(\kappa;\, k)\, f_k(d_k)\, \mathbb{E}\left[\mathfrak{K}^{-1}(X_i)\; \Psi_1^{(k)}(X_i, \boldsymbol{V})\, h_0(\boldsymbol{X}, \boldsymbol{Z}) \;\middle|\; X_k = d_k\right].$$

ii) if $\mathcal{K}$ contains two or more elements, then

$$\lim_{\varepsilon \searrow 0} \mathbb{E}\left[\delta_\varepsilon^{\mathcal{K}}(\boldsymbol{X})\, h_\varepsilon(\boldsymbol{X}, \boldsymbol{Z})\right] = 0\,.$$

Proof of Lemma 4. First, let $\mathcal{K} = \{k\}$ and note that $\lim_{\varepsilon \searrow 0} a_{\varepsilon,j}^{\complement}(x) = \lim_{\varepsilon \searrow 0} 1 - a_{\varepsilon,j}(x) = 1$, for all $j \in \mathcal{M}$ and $x \in \mathbb{R}$. Next, we calculate the inverse of $\eta_{\varepsilon,k}(x)$ in x, which is given by

$$\eta_{\varepsilon,k}^{-1}(x) = \Psi^{(k)}\left(\kappa_\varepsilon^{-1}\left(\Psi^{(k),-1}(x, \mathbf{v})\right), \mathbf{v}\right) = \Psi^{(k)}\left(\kappa_\varepsilon^{-1}(x), \mathbf{v}\right).$$

Its derivative is, noting that $\eta_{0,k}(x) = \eta_{0,k}^{-1}(x) = x$,

$$\lim_{\varepsilon \searrow 0} \tfrac{1}{\varepsilon}\left(\eta_{\varepsilon,k}^{-1}(x) - x\right) = \Psi_1^{(k)}(x, \mathbf{v})\, \mathfrak{K}^{-1}(x)\,.$$

Using similar arguments as in the proof of Lemma 1, replacing κ_ε^{-1} with $\eta_{\varepsilon,k}^{-1}$, we obtain that

$$\lim_{\varepsilon \searrow 0} \mathbb{E}[\delta_\varepsilon^k(\boldsymbol{X}) h_\varepsilon(\boldsymbol{X}, \boldsymbol{Z})] = -c(\kappa; k) f_i(d_k)\, \mathbb{E}\left[\mathfrak{K}^{-1}(X_i)\, \Psi_1^{(k)}(X_i, \boldsymbol{V})\, h_0(\boldsymbol{X}, \boldsymbol{Z}) \mid X_i = d_k\right].$$

Next, assume that $\mathcal{K} = \{k, j\}$ contains two indices and let $\xi\colon \mathbb{R}^2 \to \mathbb{R}$ be an infinitely often differentiable function. Then, using (11) and the following change of variable $y_j = \eta_{\varepsilon,j}(x_j)$ in the first equation

$$\begin{aligned}
&\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \xi(x_j, x_k)\delta_\varepsilon^{\mathcal{K}}(x_j, x_k)\, dx_j dx_k \\
&= -c(\kappa)\frac{1}{\varepsilon}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \xi(x_j, x_k)\left(\mathbb{1}_{\{\eta_{\varepsilon,j}(x_j) \le d_j\}} - \mathbb{1}_{\{x_j \le d_j\}}\right) dx_j\, a_{\varepsilon,k}(x_k) \prod_{l \ne j,k} a_{\varepsilon,l}^{\complement}(x_l)\, dx_k \\
&= -c(\kappa)\int_{-\infty}^{+\infty}\left(\frac{1}{\varepsilon}\left(\int_{-\infty}^{+\infty} \frac{\xi(\eta_{\varepsilon,j}^{-1}(y_j), x_k)}{\eta_{\varepsilon,j}'(\eta_{\varepsilon,j}^{-1}(y_j))}\mathbb{1}_{\{y_j \le d\}}\, dy_j - \int_{-\infty}^{d_j} \xi(x_j, x_k)\, dx_j\right)\right) a_{\varepsilon,k}(x_k) \prod_{l \ne j,k} a_{\varepsilon,l}^{\complement}(x_l)\, dx_k\,.
\end{aligned}$$

Define the function $\Xi(x, y)$, such that $\frac{d}{dx}\Xi(x, y) = \xi(x, y)$, so that

$$\frac{1}{\varepsilon}\int_{-\infty}^{+\infty} \frac{\xi(\eta_{\varepsilon,j}^{-1}(y_j), x_k)}{\eta_{\varepsilon,j}'(\eta_{\varepsilon,j}^{-1}(y_j))}\mathbb{1}_{\{y_j \le d\}}\, dy_j - \int_{-\infty}^{d} \xi(x_j, x_k)\, dx_j = \frac{1}{\varepsilon}\left(\Xi(\eta_{\varepsilon,j}^{-1}(d, x_k)) - \Xi(d, x_k)\right). \tag{18}$$

The limit of (18) for $\varepsilon \searrow 0$ exists, moreover $a_{\varepsilon,k}(x)$ converges to 1, for $\varepsilon \searrow 0$, while $a_{\varepsilon,l}^{\complement}(x)$, $l \ne \{j, k\}$, converge to 0 for $\varepsilon \searrow 0$. Thus, we obtain that $\delta_\varepsilon^{\mathcal{K}}(\cdot)$ converges weakly to 0, for $\varepsilon \searrow 0$.

The cases when $\mathcal{K}$ contains more than two indices follow analogous. □

LEMMA 5. *Define the sequence of functions*

$$\delta_\varepsilon(l) = \frac{\mathbb{1}_{\left\{l \le q_\alpha - \sum_{k=1}^m \tilde{\Delta}_\varepsilon g_k\right\}} - \mathbb{1}_{\{l \le q_\alpha\}}}{\varepsilon},$$

where $\tilde{\Delta}_\varepsilon g_k = \left(g_k\left(\Psi^{(\boldsymbol{Z})}(\kappa_\varepsilon(x_i), \boldsymbol{v})\right) - g_k(\boldsymbol{z})\right)\mathbb{1}_{\{\Psi^{(k)}(\kappa_\varepsilon(x_i), \boldsymbol{v}) \le d_k\}}$, $\boldsymbol{z} \in \mathbb{R}^n$, $x_i \in \mathbb{R}$, *and* $l \ge 0$. *Then,* δ_ε *converges weakly to a scaled Dirac delta function at* q_α *for* $\varepsilon \searrow 0$. *Moreover, for any function* $h_\varepsilon\colon \mathbb{R}^{m+n} \to \mathbb{R}$ *such that* $\lim_{\varepsilon \searrow 0} \mathbb{E}[|h_\varepsilon(\boldsymbol{X}, \boldsymbol{Z})|] < \infty$, *the following holds:*

$$\lim_{\varepsilon \searrow 0} \mathbb{E}[\delta_\varepsilon(L) h_\varepsilon(\boldsymbol{X}, L)] = -f(q_\alpha)\sum_{k=1}^m\sum_{l=1}^n \mathbb{E}\left[\mathfrak{K}(X_i)\partial_l\, g_k(\boldsymbol{Z})\Psi_1^{(m+l)}(X_i, \boldsymbol{V})\mathbb{1}_{\{X_k \le d_k\}}\, h_0(\boldsymbol{X}, L)\,\Big|\, L = q_\alpha\right].$$

Proof of Lemma 5. This proof follows along the lines of the proof of Lemma 2. Note that $z_l = \Psi^{(m+l)}(x_i, \boldsymbol{v})$, and that the Taylor approximation of $g_k\left(\Psi^{(\boldsymbol{z})}(\kappa_\varepsilon(x_i), \boldsymbol{v})\right)$ around $\varepsilon = 0$, becomes, using first an approximation of g_k around $\boldsymbol{z}$, then of $\boldsymbol{\Psi}^{(m+l)}$ around x_i, for all $l = 1, \dots, n$, and finally for κ_ε around $\varepsilon = 0$

$$\begin{aligned}
g_k\left(\Psi^{(\boldsymbol{Z})}(\kappa_\varepsilon(x_i), \boldsymbol{v})\right) - g_k(\boldsymbol{z}) &= \sum_{l=1}^{n} \partial_l\, g_k(\boldsymbol{z})\left(\Psi^{(m+l)}(\kappa_\varepsilon(x_i), \boldsymbol{v}) - z_l\right) + o\left(\Psi^{(m+l)}(\kappa_\varepsilon(x_i), \boldsymbol{v}) - z_l\right) \\
&= \sum_{l=1}^{n} \partial_l\, g_k(\boldsymbol{z})\, \Psi_1^{(m+l)}(x_i, \boldsymbol{v})\left(\kappa_\varepsilon(x_i) - x_i\right) + o\left(\kappa_\varepsilon(x_i)\right) \\
&= \varepsilon \sum_{l=1}^{n} \partial_l\, g_k(\boldsymbol{z})\, \Psi_1^{(m+l)}(x_i, \boldsymbol{v})\, \mathfrak{K}(x_i) + o(\varepsilon)\,.
\end{aligned}$$

The reminder of the proof follows analogous steps to those in the proof of Lemma 2. ☐

Proof of Theorem 6 (Cascade Sensitivity VaR to X_i). Analogous to the proof of Theorem 1, we use Equation (14) and, thus, we only need to calculate $\frac{\partial}{\partial\varepsilon} F_\varepsilon(q_\alpha)|_{\varepsilon=0}$. Using Lemma 3, we obtain

$$F_\varepsilon(q_\alpha) - F(q_\alpha) = \mathbb{E}\left[\mathbb{1}_{\left\{L \le q_\alpha - \sum_{k=1}^{m} \tilde{\Delta}_\varepsilon g_k + \sum_{j=1}^{m} c(\kappa;j) g_j(\boldsymbol{Z}) A_{\varepsilon,j}\right\}} - \mathbb{1}_{\{L \le q_\alpha\}}\right],$$

where we recall that $A_{\varepsilon,j} = |\mathbb{1}_{\{\eta_{\varepsilon,j}(X_j) \le d_j\}} - \mathbb{1}_{\{X_j \le d_j\}}|$ and denote its complement by $A^{\complement}_{\varepsilon,j}$, i.e., $A^{\complement}_{\varepsilon,j} = 1 - A_{\varepsilon,j}$. Next, as $A_{\varepsilon,j}$ are indicators, we can rewrite the expectation and split it into multiple sums, as follow: The first expectation corresponding to all $A^{\complement}_{\varepsilon,j}$ (19a), and then we sum over all possible combinations of $A_{\varepsilon,j}$ and $A^{\complement}_{\varepsilon,k}$.

$$\begin{aligned}
F_\varepsilon(q_\alpha) - F(q_\alpha) = \mathbb{E}&\left[\prod_{i=1}^{m} A^{\complement}_{\varepsilon,i}\left(\mathbb{1}_{\left\{L \le q_\alpha - \sum_{k=1}^{m} \tilde{\Delta}_\varepsilon g_k\right\}} - \mathbb{1}_{\{L \le q_\alpha\}}\right)\right] \qquad (19a) \\
&+ \sum_{k=1}^{m} \sum_{\substack{i_1, \dots, i_k = 1 \\ i_1 < \dots < i_k}}^{m} \mathbb{E}\left[\prod_{j=1}^{k} A_{\varepsilon,i_j} \prod_{\substack{l=1 \\ l \notin \{i_1, \dots, i_k\}}}^{m} A^{\complement}_{\varepsilon,l}\right. \\
&\times \left.\left(\mathbb{1}_{\left\{L \le q_\alpha - \sum_{r=1}^{m} \tilde{\Delta}_\varepsilon g_r + \sum_{j=1}^{k} c(\kappa;j) g_{i_j}(\boldsymbol{Z})\right\}} - \mathbb{1}_{\{L \le q_\alpha\}}\right)\right].
\end{aligned}$$

For the first expectation above (Equation (19a)), we apply Lemma 5 and that $\lim_{\varepsilon \searrow 0} A^{\complement}_{\varepsilon,k} = 1$ for all $k = 1, \dots, m$. For the other terms, we apply Lemma 4. Specifically, we observe that only the summands that contains exactly one $A_{\varepsilon,k}$ do not converge to 0. Thus, we obtain the limit, noting that for all $k = 1, \dots, m$, $\tilde{\Delta}_\varepsilon\, g_k$ converges to 0, for $\varepsilon \searrow 0$,

$$\begin{aligned}
&\lim_{\varepsilon \searrow 0} \frac{F_\varepsilon(q_\alpha) - F(q_\alpha)}{\varepsilon} \\
&= -\sum_{j=1}^{m} \sum_{l=1}^{n} f(q_\alpha)\mathbb{E}\left[\mathfrak{K}(X_i)\partial_l\, g_j(\boldsymbol{Z})\Psi_1^{(m+l)}(X_i, \boldsymbol{V})\mathbb{1}_{\{X_j \le d_j\}}\,\Big|\, L = q_\alpha\right] \\
&\quad - \sum_{j=1}^{m} c(\kappa;j) f_j(d_j)\mathbb{E}\left[\mathfrak{K}^{-1}(X_i)\Psi_1^{(j)}(X_i, \boldsymbol{V})\left(\mathbb{1}_{\left\{L \le q_\alpha + c(\kappa;j) g_j(\boldsymbol{Z})\right\}} - \mathbb{1}_{\{L \le q_\alpha\}}\right) \mid X_j = d_j\right].
\end{aligned}$$

Combining with Equation (14) concludes the proof. ☐

Proof of Theorem 3 (Cascade Sensitivity ES to X_i). We write analogous to the proof of Theorem 2

$$\lim_{\varepsilon\searrow 0}\tfrac{1}{\varepsilon}\left(\mathrm{ES}_\alpha(L_\varepsilon)-\mathrm{ES}_\alpha(L)\right)=\lim_{\varepsilon\searrow 0}A(\varepsilon)+B(\varepsilon)+C(\varepsilon)=\lim_{\varepsilon\searrow 0}B(\varepsilon)\,.$$

For part $B(\varepsilon)$, we proceed similar to the proof of Theorem 6 and write, using the notation from the proof of Theorem 6 and Lemma 3

$$\begin{aligned}
B(\varepsilon)(1-\alpha)\varepsilon &= \mathbb{E}\left[\left(L+\sum_{r=1}^m\tilde\Delta_\varepsilon g_r-\sum_{k=1}^m c(\kappa;\,k)g_k(\boldsymbol{Z})\,A_{\varepsilon,k}-q_\alpha\right)_+-(L-q_\alpha)_+\right]\\
&=\mathbb{E}\left[\prod_{i=1}^m A^{\complement}_{\varepsilon,i}\left(\left(L+\sum_{r=1}^m\tilde\Delta_\varepsilon g_r-q_\alpha\right)_+-(L-q_\alpha)_+\right)\right] &\text{(20a)}\\
&\quad+\sum_{k=1}^m\sum_{\substack{i_1,\dots,i_k=1\\ i_1<\dots<i_k}}^m\mathbb{E}\Bigg[\prod_{j=1}^k A_{\varepsilon,i_j}\prod_{\substack{l=1\\ l\notin\{i_1,\dots,i_k\}}}^m A^{\complement}_{\varepsilon,l}\\
&\quad\times\left(\left(L+\sum_{r=1}^m\tilde\Delta_\varepsilon g_r-\sum_{j=1}^k c(\kappa;\,j)g_{i_j}(\boldsymbol{Z})\,A_{\varepsilon,i_j}-q_\alpha\right)_+-\left(L-q_\alpha\right)_+\right)\Bigg]\,. &\text{(20b)}
\end{aligned}$$

To calculate the limit of the expectation in Equation (20a), we rewrite similar to (17)

$$\begin{aligned}
\left(L+\sum_{r=1}^m\tilde\Delta_\varepsilon g_r-q_\alpha\right)_+-(L-q_\alpha)_+ &= \left(L-q_\alpha\right)_+\left(\mathbb{1}_{\{L\le q_\alpha\}}-\mathbb{1}_{\{L\le q_\alpha-\sum_{k=1}^m\tilde\Delta_\varepsilon g_k\}}\right) &\text{(21a)}\\
&\quad+\sum_{r=1}^m\tilde\Delta_\varepsilon g_r\mathbb{1}_{\{L\ge q_\alpha-\sum_{k=1}^m\tilde\Delta_\varepsilon g_k\}}\,. &\text{(21b)}
\end{aligned}$$

For the term (21a) we apply Lemma 5, noting that $A^{\complement}_{\varepsilon,k}$ converges to 1, for all $k=1,\dots,m$, as $\varepsilon\searrow 0$. For the term (21b), we note that for all $k=1,\dots,m$, it holds $\mathbb{P}$-a.s. (see the Proof of Lemma 5) that

$$\lim_{\varepsilon\searrow 0}\frac{\tilde\Delta_\varepsilon g_k}{\varepsilon}=\sum_{l=1}^n\partial_l\, g_k(\boldsymbol{Z})\Psi_1^{(m+l)}(X_i,\boldsymbol{V})\mathfrak{K}(X_i)\mathbb{1}_{\{X_k\le d_k\}}\,.$$

For all the other summands in Equation (20b) we apply Lemma 4. Collecting, we obtain that

$$\begin{aligned}
&(1-\alpha)\lim_{\varepsilon\searrow 0}B(\varepsilon)=\\
&\sum_{k=1}^m\sum_{l=1}^n f(q_\alpha)\mathbb{E}\left[\mathfrak{K}(X_i)\partial_l\, g_k(\boldsymbol{Z})\Psi_1^{(m+l)}(X_i,\boldsymbol{V})\mathbb{1}_{\{X_k\le d_k\}}\left(L-q_\alpha\right)_+\Big|\,L=q_\alpha\right] &\text{(22a)}\\
&+\sum_{j=1}^m\sum_{l=1}^n\mathbb{E}\left[\partial_l\, g_j(\boldsymbol{Z})\Psi_1^{(m+l)}(X_i,\boldsymbol{V})\mathfrak{K}(X_i)\mathbb{1}_{\{X_j\le d_j\}}\mathbb{1}_{\{L\ge q_\alpha\}}\right]\\
&-\sum_{j=1}^m c(\kappa;\,j)f_j(d_j)\mathbb{E}\left[\mathfrak{K}^{-1}(X_i)\Psi_1^{(j)}(X_i,\boldsymbol{V})\left(\left(L-c(\kappa;\,j)g_j(\boldsymbol{Z})-q_\alpha\right)_+-\left(L-q_\alpha\right)_+\right)\;\Big|\;X_j=d_j\right]\,.
\end{aligned}$$

Due to the conditioning event, (22a) is equal to 0. □

Proof of Theorems 7 and 4 (Cascade Sensitivities to Z_i). The proof follows by noting that the stressed portfolio for a stress on Z_i, admits an analogous representation as when stressing X_i, with the only difference that the inverse Rosenblatt transform starts at Z_i instead of X_i, see Equation (4). □

Proof of Proposition 1. The first two cases follow from Proposition 4.2 in Pesenti et al. (2021). Assume (X_i, X_j) follow a Archimedean copula. By independence of the cascade sensitivity to the choice of Rosenblatt transform, we have can choose (Rüschendorf and de Valk 1993)

$$\Psi_1^{(j)}(X_i, \boldsymbol{V}) = F_{j|i}^{-1}(V_j|X_i) \quad \mathbb{P}\text{-a.s.},$$

where a.s. $V_j = F_{X_j\,|\,X_i}(X_j\,|\,X_i)$ and $F_{X_j\,|\,X_i}(x\,|\,y) = \mathbb{P}(X_j \le x\,|\,X_i = y)$ denotes the conditional distribution of X_j given $X_i = y$. We observe that $\Psi_1^{(j)}(X_i, \boldsymbol{V})$ only depends on V_j and we may write $\Psi_1^{(j)}(X_i, V_j)$ instead of $\Psi_1^{(j)}(X_i, \boldsymbol{V})$.

We use Sklar's theorem to write the conditional distribution and quantile functions as

$$F_{j|i}(x_j\,|\,x_i) = \mathrm{C}_{j|i}(F_j(x_j)\,|\,F_i(x_i)) \quad \text{and} \quad F_{j|i}^{-1}(v\,|\,x_i) = F_j^{-1}\left(\mathrm{C}_{j|i}^{-1}(v\,|\,F_i(x_i))\right). \tag{23}$$

Taking derivative of the conditional quantile function with respect to the conditioning variable

$$\Psi_1^{(j)}(y, v) = \frac{\partial}{\partial y} F_{j|i}^{-1}(v\,|\,y) = \frac{f_i(y)}{f_j\left(F_{j|i}^{-1}(v\,|\,y)\right)} \frac{\partial}{\partial z} \mathrm{C}_{j|i}^{-1}(v\,|\,z)\Big|_{z=F_i(y)}.$$

By definition of V_j, it holds $\mathbb{P}$-a.s. that $F_{j|i}^{-1}(V_j\,|X_i) = X_j$, thus

$$\Psi_1^{(j)}(X_i, V_j) = \frac{f_i(X_i)}{f_j\,(X_j)} \frac{\partial}{\partial z} \mathrm{C}_{j|i}^{-1}(V_j\,|\,z)\Big|_{z=F_i(X_i)}. \tag{24}$$

Next, we calculate the derivative (with respect to the conditioning argument) of the inverse of an conditional Archimedean copula with generator ψ. The conditional Archimedean copula and its inverse are given by (Cambou et al. 2017)

$$\mathrm{C}_{j|i}(x\,|\,y) = \frac{\dot{\psi}\left(\psi^{-1}(y) + \psi^{-1}(x)\right)}{\dot{\psi}\left(\psi^{-1}(y)\right)} \quad \text{and} \tag{25a}$$

$$\mathrm{C}_{j|i}^{-1}(v\,|\,y) = \psi\left[(\dot{\psi})^{-1}\left\{v\,\dot{\psi}\left(\psi^{-1}(y)\right)\right\} - \psi^{-1}(y)\right],$$

where $\dot{\psi}(x) = \frac{d}{dx}\psi(x)$. Taking derivative

$$\frac{\partial}{\partial y}\mathrm{C}_{j|i}^{-1}(v\,|\,y) = \dot{\psi}\left[(\dot{\psi})^{-1}\left\{v\,\dot{\psi}\left(\psi^{-1}(y)\right)\right\} - \psi^{-1}(y)\right] \frac{1}{\dot{\psi}\left(\psi^{-1}(y)\right)} \left\{\frac{v\,\ddot{\psi}\left(\psi^{-1}(y)\right)}{\ddot{\psi}\left((\dot{\psi})^{-1}\{v\,\dot{\psi}\left(\psi^{-1}(y)\right)\}\right)} - 1\right\}.$$

Next, we use the definition of V_j and Equations (23) and (25a) to write

$$V_j = \mathrm{C}_{j|i}(F_j(X_j)\,|\,F_i(X_i)) = \frac{\dot{\psi}\left(\psi^{-1}(U_i) + \psi^{-1}(U_j)\right)}{\dot{\psi}\left(\psi^{-1}(U_i)\right)},$$

where $U_j = F_j(X_j)$ and $U_i = F_i(X_i)$. Using the above, we obtain

$$\frac{\partial}{\partial y} \mathrm{C}_{j|i}^{-1}(V_j \,|\, y)\bigg|_{y=U_i} = \frac{\dot\psi\left(\psi^{-1}(U_j)\right)}{\dot\psi\left(\psi^{-1}(U_i)\right)} \left\{ \frac{\ddot\psi\left(\psi^{-1}(U_i)\right)}{\ddot\psi\left(\psi^{-1}(U_i)+\psi^{-1}(U_j)\right)} \frac{\dot\psi\left(\psi^{-1}(U_i)+\psi^{-1}(U_j)\right)}{\dot\psi\left(\psi^{-1}(U_i)\right)} - 1 \right\}.$$

Combining with Equation (24) concludes the proof. □

B.4. Proof of Sensitivity to General Loss Model: Theorems 8, 9, and 10

Proof of Theorem 8 (Marginal Sensitivity - General Loss Model). The stressed loss model has representation (using Equation (11))

$$L_\varepsilon(X_i) = L + \sum_{j=1}^m \Delta_\varepsilon g_j - c(\kappa) g_i\left(X_{i,\varepsilon}, \boldsymbol{X}_{-i}, \boldsymbol{Z}\right) A_\varepsilon\,,$$

where $\Delta_\varepsilon g_j := \left(g_j(X_{i,\varepsilon}, \boldsymbol{X}_{-i}, \boldsymbol{Z}) - g_j(\boldsymbol{X}, \boldsymbol{Z})\right) \mathbb{1}_{\{X_j \le d_j\}}$ and $A_\varepsilon = |\mathbb{1}_{\{X_{i,\varepsilon} \le d_i\}} - \mathbb{1}_{\{X_i \le d_i\}}|$. To prove the case for VaR, note that

$$F_\varepsilon(q_\alpha) - F(q_\alpha) = \mathbb{E}\left[A_\varepsilon^\complement \left(\mathbb{1}_{\{L+\sum_{j=1}^m \Delta_\varepsilon g_j \le q_\alpha\}} - \mathbb{1}_{\{L \le q_\alpha\}}\right)\right] \tag{26a}$$

$$+\mathbb{E}\left[A_\varepsilon \left(\mathbb{1}_{\{L+\sum_{j=1}^m \Delta_\varepsilon g_j - c(\kappa) g_i(X_{i,\varepsilon}, \boldsymbol{X}_{-i}, \boldsymbol{Z}) \le q_\alpha\}} - \mathbb{1}_{\{L \le q_\alpha\}}\right)\right], \tag{26b}$$

Applying Lemma 2 to (26a), noting that $\lim_{\varepsilon\searrow 0} A_\varepsilon^\complement = 1$, and Lemma 1 to (26b), noting that $\lim_{\varepsilon\searrow 0} \sum_{j=1}^m \Delta_\varepsilon g_j = 0$ concludes the proof for VaR.

To prove the case of ES, we have from the proof of Theorem 2 that

$$\begin{aligned}\lim_{\varepsilon\searrow 0} \frac{\mathrm{ES}_\alpha(L_\varepsilon) - \mathrm{ES}_\alpha(L)}{\varepsilon} &= \lim_{\varepsilon\searrow 0} \frac{1}{\varepsilon(1-\alpha)} \mathbb{E}[(L_\varepsilon - q_\alpha)_+ - (L - q_\alpha)_+] \\ &= \lim_{\varepsilon\searrow 0} \frac{1}{\varepsilon(1-\alpha)} \Bigg\{ \mathbb{E}\Big[A_\varepsilon^\complement \underbrace{\Big((L + \sum_{j=1}^m \Delta_\varepsilon g_j - q_\alpha)_+ - (L - q_\alpha)_+ \Big)}_{=B_\varepsilon}\Big] \\ &\quad + \mathbb{E}\left[A_\varepsilon \left(\Big(L + \sum_{j=1}^m \Delta_\varepsilon g_j - c(\kappa) g_i(X_{i,\varepsilon}, \boldsymbol{X}_{-i}, \boldsymbol{Z}) - q_\alpha\Big)_+ - (L - q_\alpha)_+ \right)\right] \Bigg\}.\end{aligned} \tag{27a}$$

Next, we see that

$$B_\varepsilon = (L - q_\alpha)_+ \left(\mathbb{1}_{\{L \le q_\alpha\}} - \mathbb{1}_{\{L \le q_\alpha - \sum_{j=1}^m \Delta_\varepsilon g_j\}}\right) + \sum_{j=1}^m \Delta_\varepsilon g_j \mathbb{1}_{\{L \ge q_\alpha - \sum_{j=1}^m \Delta_\varepsilon g_i\}}\,.$$

Using similar arguments as in the proof of Theorem 3 (in particular applying Lemma 5), we observe that

$$\lim_{\varepsilon\searrow 0} \frac{1}{\varepsilon(1-\alpha)} \mathbb{E}[A_\varepsilon^\complement B_\varepsilon] = \sum_{j=1}^m \mathbb{E}\left[\partial_i g_j(\boldsymbol{X}, \boldsymbol{Z}) \mathfrak{K}(X_i) \mathbb{1}_{\{X_j \le d_j\}} \;|\; L \ge q_\alpha\right].$$

Applying Lemma 1 to (27a) concludes the proof. □

The proofs of Theorems 9, and 10 follow along the lines of the proofs of Theorems 6 and 3.

B.5. Proof of Sensitivity to Discrete Random Variable: Theorem 5.

Proof of Theorem 5 (Marginal Sensitivity - Discrete). The stressed loss can be written as

$$T_{W,\varepsilon} = T - c(\kappa) \sum_{k=1}^{r} \Delta_k \, h(W, \boldsymbol{Y}) A_{\varepsilon,k} \,,$$

where $A_{\varepsilon,k} = |\mathbb{1}_{\{\kappa_\varepsilon(U) \le p_k\}} - \mathbb{1}_{\{U \le p_k\}}|$ for $k \in \{1, \dots, r\}$. To prove the formula for VaR, we calculate similarly to the proof of Theorem 6

$$\begin{aligned}
&\mathbb{P}\left(T_{W,\varepsilon} \le q_\alpha\right) - \mathbb{P}\left(T \le q_\alpha\right) \\
&\qquad = \mathbb{E}\left[\mathbb{1}_{\{T \le q_\alpha + c(\kappa)\sum_{k=1}^{r} \Delta_k \, h(W,\boldsymbol{Y}) A_{\varepsilon,k}\}} - \mathbb{1}_{\{T \le q_\alpha\}}\right] \\
&\qquad = \sum_{k=1}^{r} \sum_{\substack{i_1,\dots,i_k=1 \\ i_1 < \dots < i_k}}^{r} \mathbb{E}\left[\prod_{j=1}^{k} A_{\varepsilon,i_j} \prod_{\substack{l=1 \\ l \notin \{i_1,\dots i_k\}}}^{r} A^{\complement}_{\varepsilon,l} \left(\mathbb{1}_{\{T \le q_\alpha + c(\kappa)\sum_{j=1}^{k} \Delta_j \, h(W,\boldsymbol{Y})\}} - \mathbb{1}_{\{T \le q_\alpha\}}\right)\right].
\end{aligned}$$

Applying Lemma 4 we obtain

$$\begin{aligned}
\frac{\partial}{\partial \varepsilon} F_\varepsilon(q_\alpha) &= -\sum_{k=1}^{r} c(\kappa) \mathbb{E}\left[\mathfrak{K}^{-1}(U)\left(\mathbb{1}_{\{T \le q_\alpha + c(\kappa)\Delta_k \, h(W,\boldsymbol{Y})\}} - \mathbb{1}_{\{T \le q_\alpha\}}\right) \;|\; U = p_k\right] \\
&= -c(\kappa) \sum_{k=1}^{r} \mathfrak{K}^{-1}(p_k) \mathbb{E}\left[\left(\mathbb{1}_{\{T \le q_\alpha + c(\kappa)\Delta_k \, h(W,\boldsymbol{Y})\}} - \mathbb{1}_{\{T \le q_\alpha\}}\right) \;|\; W = w_k\right].
\end{aligned}$$

Combining with Equation (14) concludes the proof for VaR.

Second, we prove the case ES. From the proof of Theorem 2, we have

$$\begin{aligned}
\mathrm{ES}_\alpha(T_{W,\varepsilon}) - \mathrm{ES}_\alpha(T) &= \tfrac{1}{(1-\alpha)} \mathbb{E}[(T_{W,\varepsilon} - q_\alpha)_+ - (T - q_\alpha)_+] \\
&= \tfrac{1}{(1-\alpha)} \mathbb{E}\left[\left(T - c(\kappa)\sum_{k=1}^{r} \Delta_k \, h(W, \mathbf{Y}) A_{\varepsilon,k} - q_\alpha\right)_+ - (T - q_\alpha)_+\right] \\
&= \tfrac{1}{(1-\alpha)} \sum_{k=1}^{r} \sum_{\substack{i_1,\dots,i_k=1 \\ i_1 < \dots < i_k}}^{m} \mathbb{E}\left[\prod_{j=1}^{k} A_{\varepsilon,i_j} \prod_{\substack{l=1 \\ l \notin \{i_1,\dots i_k\}}}^{m} A^{\complement}_{\varepsilon,l}\right. \\
&\qquad \times \left.\left(\left(T - c(\kappa)\sum_{j=1}^{k} \Delta_{i_j} h(W, \mathbf{Y}) - q_\alpha\right)_+ - (T - q_\alpha)_+\right)\right].
\end{aligned}$$

Next, we apply Lemma 4 and obtain

$$\lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon}\left(\mathrm{ES}_\alpha(L_\varepsilon) - \mathrm{ES}_\alpha(L)\right) = -\frac{c(\kappa)}{1-\alpha} \sum_{k=1}^{r} \mathfrak{K}^{-1}(p_k) \mathbb{E}\left[\left(T - c(\kappa)\Delta_k h(W, \mathbf{Y}) - q_\alpha\right)_+ - (T - q_\alpha)_+ \;\middle|\; U = p_k\right],$$

which concludes the proof. $\square$

Table 3 Correlation matrix $\boldsymbol{R}$ of Z (source: Lloyd's (2022)).

	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8	Z_9	Z_{10}	Z_{11}	Z_{12}
Z_1	1	0.5	0.5	0.25	0.5	0.25	0.5	0.25	0.5	0.25	0.25	0.25
Z_2	0.5	1	0.25	0.25	0.25	0.25	0.5	0.5	0.5	0.25	0.25	0.25
Z_3	0.5	0.25	1	0.25	0.25	0.25	0.25	0.5	0.5	0.25	0.5	0.25
Z_4	0.25	0.25	0.25	1	0.25	0.25	0.25	0.5	0.5	0.25	0.5	0.5
Z_5	0.5	0.25	0.25	0.25	1	0.5	0.5	0.25	0.5	0.5	0.25	0.25
Z_6	0.25	0.25	0.25	0.25	0.5	1	0.5	0.25	0.5	0.5	0.25	0.25
Z_7	0.5	0.5	0.25	0.25	0.5	0.5	1	0.25	0.5	0.5	0.25	0.25
Z_8	0.25	0.5	0.5	0.5	0.25	0.25	0.25	1	0.5	0.25	0.25	0.5
Z_9	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1	0.25	0.5	0.25
Z_{10}	0.25	0.25	0.25	0.25	0.5	0.5	0.5	0.25	0.25	1	0.25	0.25
Z_{11}	0.25	0.25	0.5	0.5	0.25	0.25	0.25	0.25	0.5	0.25	1	0.25
Z_{12}	0.25	0.25	0.25	0.5	0.25	0.25	0.25	0.5	0.25	0.25	0.25	1

Appendix C: Model specification for the numerical example of Section 5

The insurance company has $n = 12$ lines of business (LoB) each following a LogNormal distribution, $Z_k \sim \mathcal{LN}(\mu_k, \sigma_k^2)$, $k \in \mathcal{N}$. The parameters μ_k and σ_k are such that $\mathbb{E}[Z_i] = e^{\mu_i + \frac{1}{2}\sigma_i^2} = 100$ (reflecting the business volume) and $\text{CoV}_i = \sqrt{\text{var}(Z_i)}/\mathbb{E}[Z_i] = \sqrt{e^{\sigma_i^2} - 1}$, where CoV denotes the coefficients of variation. The considered CoV and lines of business are reported in Table 2 and the correlation matrix $\mathbf{R}$ in Table 3. These figures are taken from the Solvency II Standard Formula parameters (Lloyd's 2022).

We assume that the reinsurers' critical variables follow a standardised student t distributions with $\nu = 4$ degrees of freedom, i.e. $X_i \sim t(4)$, for all $i = 1, \dots, 8$. (Note that the choice of marginal distribution for X_i is irrelevant since we consider only the event $\{X_i \leq d_i\}$.) The default probabilities $\mathbb{P}(X_i \leq d_i) = q_i$ are set to $q_i = 0.015$, $i = 1, \dots, 6$ and $q_i = 0.01$, $i = 7, 8$.

We assume that $(\boldsymbol{X}, \boldsymbol{Z})$ has a multivariate t copula with $\nu = 4$ degrees of freedom and correlation parameter matrix $\mathbf{\Sigma} = \{\sigma_{i,j}\}_{i,j=1,\dots,m+n}$ (McNeil et al. 2015, Sec. 7.3). The elements of $\mathbf{\Sigma}$ comprise the pairwise Pearson correlations of multivariate t_ν-distributed random vector arising from monotone transformations of elements of $(\boldsymbol{X}, \boldsymbol{Z})$, so that each has t_ν marginals. As the (multivariate) margins of multivariate t_ν distributions are again multivariate t_ν, we start by specifying the correlation parameters of the vectors $\boldsymbol{X}$ and $\boldsymbol{Z}$ separately and then consider the dependence across the two vectors' elements. First, $\boldsymbol{X}$ has standardised t_ν margins and hence follows a multivariate t_ν distribution. Furthermore, we assume that the dependence structure is homogeneous, such that $\sigma_{i,j} = \text{Corr}(X_i, X_j) = \lambda > 0$, for all $i \neq j$, $i, j \in \mathcal{M}$. Second, $\boldsymbol{Z}$ has a multivariate t_ν copula with correlation parameter matrix $\mathbf{R} = \{r_{k,l}\}_{k,l \in \mathcal{N}}$, given in Table 2, thus $\sigma_{m+k,m+l} := r_{k,l}$, $k, l \in \mathcal{N}$. Third, to specify the elements of $\mathbf{\Sigma}$ characterising the dependence of (X_j, Z_k), i.e. $\sigma_{j,m+k}$, $j \in \mathcal{M}$, $k \in \mathcal{N}$, we build a dependence model that links gross losses to reinsurance defaults using a single factor model, reflecting the homogeneity in the dependence of $\boldsymbol{X}$. The common factor is a function of the gross losses and acts as a proxy for industry effects.

By the representation of multivariate t distributions as normal mixtures (McNeil et al. 2015, Sec. 6.2), we can represent each Z_k as

$$Z_k = F_{Z_k}^{-1}\left(t_\nu\big(\sqrt{W}\tilde{Z}_k\big)\right), \quad k \in \mathcal{N},$$

where $W \sim \text{InvGamma}(\nu/2, \nu/2-1)$, such that $\mathbb{E}[W]=1$, and $(\tilde{Z}_1, \dots, \tilde{Z}_n)$ are multivariate standard normal, with correlation matrix $\mathbf{R}$. Then the random variables $(\sqrt{W}\tilde{Z}_1, \dots, \sqrt{W}\tilde{Z}_n)$ are multivariate t_ν distributed, with correlation matrix $\mathbf{R}$ and margins standardised to have unit variance. (Note that this is slightly different to the standard t_ν construction, which has margins with variance $\nu/(\nu-2)$. This choice, which does not affect the dependence model, is made to simplify moment calculations.) Define:

$$\beta := var\Big(\sum_{k=1}^n \tilde{Z}_k\Big) = \sum_{k,l\in\mathcal{N}} r_{k,l}\,, \quad \text{and} \quad \Psi := \frac{1}{\sqrt{\beta}} \sum_{k\in\mathcal{N}} \tilde{Z}_k \sim \text{N}(0,1)\,.$$

Then the factor model becomes:

$$X_j = \sqrt{W}\left(\sqrt{\lambda}\Psi + \sqrt{1-\lambda}\Theta_j\right), \quad j \in \mathcal{M},$$

where $\Theta_1, \dots, \Theta_m$ are i.i.d. standard normal variables, independent of $(\tilde{Z}_1, \dots, \tilde{Z}_n, W)$. It follows easily that $\mathbb{E}[X_i]=0$, $var(X_i)=1$ and $\text{Corr}(X_i, X_j)=\lambda$ is fulfilled. Furthermore, the cross-correlation values:

$$\begin{aligned}\sigma_{i,m+k} &= \text{Corr}(X_i, t_\nu^{-1}\left(F_{Z_k}(Z_k)\right) = \text{Corr}\Big(\sqrt{W}\left(\sqrt{\lambda}\Psi + \sqrt{1-\lambda}\Theta_i\right)\ ,\ \sqrt{W}\tilde{Z}_k\Big) \\ &= \sqrt{\lambda}\ \text{Corr}\big(\Psi\ ,\ \tilde{Z}_k\big) = \sqrt{\tfrac{\lambda}{\beta}}\ \text{Corr}\Big(\sum_{l\in\mathcal{N}} \tilde{Z}_l\ ,\ \tilde{Z}_k\Big) = \sqrt{\tfrac{\lambda}{\beta}} \sum_{l\in\mathcal{N}} r_{k,l}\,.\end{aligned}$$

This completes the dependence model specification.

References

Acerbi C (2002) Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance* 26(7):1505–1518.

Albrecher H, Beirlant J, Teugels JL (2017) *Reinsurance: Actuarial and Statistical Aspects* (John Wiley & Sons).

Antoniano-Villalobos I, Borgonovo E, Siriwardena S (2018) Which parameters are important? Differential importance under uncertainty. *Risk Analysis* 38(11):2459–2477.

Bénesse C, Gamboa F, Loubes JM, Boissin T (2022) Fairness seen as global sensitivity analysis. *Machine Learning* 1–28.

Borgonovo E, Ghidini V, Hahn R, Plischke E (2023) Explaining classifiers with measures of statistical association. *Computational Statistics & Data Analysis* 182:107701.

Borgonovo E, Plischke E (2016) Sensitivity analysis: A review of recent advances. *European Journal of Operational Research* 248(3):869–887.

Britt S, Krvavych Y (2009) Reinsurance credit risk modelling. *Proceedings of the 2009 ASTIN Conference.*

Broadie M, Glasserman P (1996) Estimating security price derivatives using simulation. *Management Science* 42(2):269–285.

Browne T, Fort JC, Iooss B, Le Gratiet L (2017) Estimate of quantile-oriented sensitivity indices. *Preprint* URL https://hal.archives-ouvertes.fr/hal-01450891.

Cambou M, Hofert M, Lemieux C (2017) Quasi-random numbers for copula models. *Statistics and Computing* 27(5):1307–1329.

Chen C, Li O, Tao D, Barnett A, Rudin C, Su JK (2019) This looks like that: deep learning for interpretable image recognition. *Advances in neural information processing systems* 32.

Chen T, Guestrin C (2016) XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794.

Chen Y, Calabrese R, Martin-Barragan B (2023) Interpretable machine learning for imbalanced credit scoring datasets. *European Journal of Operational Research* 312(1):357–372.

Chen Z, Glasserman P (2008) Sensitivity estimates for portfolio credit derivatives using Monte Carlo. *Finance and Stochastics* 12(4):507–540.

Embrechts P, Hofert M (2013) A note on generalized inverses. *Mathematical Methods of Operations Research* 77(3):423–432.

Fissler T, Pesenti SM (2023) Sensitivity measures based on scoring functions. *European Journal of Operational Research* 307(3):1408–1423.

Frees EW, Huang F (2021) The discriminating (pricing) actuary. *North American Actuarial Journal* 1–23.

Fu MC, Hong LJ, Hu JQ (2009) Conditional Monte Carlo estimation of quantile sensitivities. *Management Science* 55(12):2019–2027.

Glasserman P (2005) Measuring marginal risk contributions in credit portfolios. *Journal of Computational Finance* 2(9):1–41.

Gunawan R, Cao Y, Petzold L, Doyle FJ (2005) Sensitivity analysis of discrete stochastic systems. *Biophysical Journal* 88(4):2530–2540.

Hiabu M, Meyer JT, Wright MN (2023) Unifying local and global model explanations by functional decomposition of low dimensional structures. *International Conference on Artificial Intelligence and Statistics*, 7040–7060 (PMLR).

Hong LJ (2009) Estimating quantile sensitivities. *Operations Research* 57(1):118–130.

Hong LJ, Liu G (2009) Simulating sensitivities of Conditional Value at Risk. *Management Science* 55(2):281–293.

Koike T, Saporito Y, Targino R (2022) Avoiding zero probability events when computing Value at Risk contributions. *Insurance: Mathematics and Economics* 106:173–192.

Kozodoi N, Jacob J, Lessmann S (2022) Fairness in credit scoring: Assessment, implementation and profit implications. *European Journal of Operational Research* 297(3):1083–1094.

Lindholm M, Richman R, Tsanakas A, Wüthrich MV (2022) Discrimination-free insurance pricing. *ASTIN Bulletin: The Journal of the IAA* 52(1):55–89.

Lloyd's (2022) Standard formula for the solvency capital requirement (scr). https://www.lloyds.com/resources-and-services/capital-and-reserving/capital-guidance/standard-formula-scr/, accessed: February 21, 2023.

Lundberg SM, Erion GG, Lee SI (2018) Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888* .

McNeil AJ, Frey R, Embrechts P (2015) *Quantitative Risk Management: Concepts, techniques and tools-revised edition* (Princeton university press).

Merz M, Richman R, Tsanakas A, Wüthrich MV (2022) Interpreting deep learning models with marginal attribution by conditioning on quantiles. *Data Mining and Knowledge Discovery* 36(4):1335–1370.

Pesenti SM, Millossovich P, Tsanakas A (2021) Cascade sensitivity measures. *Risk Analysis* 41(12):2392–2414.

Plyasunov S, Arkin AP (2007) Efficient stochastic sensitivity analysis of discrete event systems. *Journal of Computational Physics* 221(2):724–738.

Razavi S, Jakeman A, Saltelli A, Prieur C, Iooss B, Borgonovo E, Plischke E, Piano SL, Iwanaga T, Becker W, et al. (2021) The future of sensitivity analysis: An essential discipline for systems modeling and policy support. *Environmental Modelling & Software* 137:104954.

Ribeiro MT, Singh S, Guestrin C (2016) Model-agnostic interpretability of machine learning. *arXiv preprint arXiv:1606.05386* .

Rosenblatt M (1952) Remarks on a multivariate transformation. *The Annals of Mathematical Statistics* 23(3):470–472.

Rüschendorf L (2013) Mathematical risk analysis. *Springer Ser. Oper. Res. Financ. Eng. Springer, Heidelberg* .

Rüschendorf L, de Valk V (1993) On regression representations of stochastic processes. *Stochastic Processes and their Applications* 46(2):183–198.

Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S (2008) *Global sensitivity analysis: the primer* (John Wiley & Sons).

Tasche D (1999) Risk contributions and performance measurement. *Report of the Lehrstuhl für mathematische Statistik, TU München* .

Ter Berg P (2008) Portfolio modelling of counterparty reinsurance default risk. *Life & Pensions* 29–33.

Tsanakas A, Millossovich P (2016) Sensitivity analysis using risk measures. *Risk Analysis* 36(1):30–48.

Wang SS (2000) A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance* 15–36.

Wüthrich MV, Merz M (2023) *Statistical foundations of actuarial learning and its applications* (Springer Nature).