Cascade Sensitivity Measures*

Silvana M. Pesenti†[1], Pietro Millossovich[2,3], and Andreas Tsanakas[2]

[1]Department of Statistical Sciences, University of Toronto

[2]The Business School (formerly Cass), City, University of London

[3]DEAMS, University of Trieste

November 15, 2020

Abstract

In risk analysis, sensitivity measures quantify the extent to which the probability distribution of a model output is affected by changes (stresses) in individual random input factors. For input factors that are statistically dependent, we argue that a stress on one input should also precipitate stresses in other input factors. We introduce a novel sensitivity measure, termed *cascade sensitivity*, defined as a derivative of a risk measure applied on the output, in the direction of an input factor. The derivative is taken after suitably transforming the random vector of inputs, thus explicitly capturing the direct impact of the stressed input factor, as well as indirect effects via other inputs. Furthermore, alternative representations of the cascade sensitivity measure are derived, allowing us to address practical issues, such as incomplete specification of the model and high computational costs. The applicability of the methodology is illustrated through the analysis of a commercially used insurance risk model.

Keywords Sensitivity analysis, importance measures, model uncertainty, risk measures, dependence, Rosenblatt transform.

*Earlier versions have been presented at the University of Trieste, at the University of Milano-Bicocca, at the *10th Conference in Actuarial Science & Finance* (Samos), at the *4th European Actuarial Journal Conference* (Leuven), at the *Workshop on Recent Developments in Dependence Modelling with Applications in Finance and Insurance* (Aegina), at the *KU Leuven & Cass Business School PhD colloquium* (Leuven), at the University of Liverpool, at the RiskLab, ETH Zurich, at the Vrije Universiteit Brussel and the University of Bologna. The authors are grateful for discussions with and comments from Vali Asimit, Patrick Cheridito, Umberto Cherubini, Paul Embrechts, Mike Ludkovski, Sabrina Mulinacci, Ludger Rüschendorf, Steven Vanduffel, Ruodu Wang, and Mario Wüthrich.

†E-mail: silvana.pesenti@utoronto.ca

1 Introduction

1.1 Overview and contribution

Sensitivity analysis is concerned with the attribution of the uncertainty of a model output to the uncertainties of model inputs (Saltelli et al., 2008). Principal tools in sensitivity analysis are sensitivity measures (also called 'importance measures'), which assign to each input factor a score, ranking inputs according to their ability to influence (a probabilistic summary of) the output; see Borgonovo and Plischke (2016) for an extensive review. Variance-based sensitivity measures, for example, distinguish input factors by their ability to affect the output's variance (Saltelli, 2002). In this paper, and as is typical in risk management applications, the output distribution is summarized through a quantile-based measure of risk. Specifically, we consider the class of distortion risk measures introduced by Wang (1996), which subsumes expected utilities and the two most common risk measures in financial risk management, Value-at-Risk (VaR) and Expected Shortfall (ES) (Belles-Sampera et al., 2014).

Sensitivity measures are often constructed via partial derivatives either of outputs with respect to inputs ('local' sensitivity measures, see Borgonovo and Plischke (2016) and references therein) or of the output risk measure in the direction of random inputs (Tsanakas and Millossovich, 2016; Antoniano-Villalobos et al., 2018). One drawback of such sensitivity measures is that they do not fully account for interactions among or statistical dependence between input factors. Extensions have so far focused on higher order derivatives (Mara and Tarantola, 2012; Borgonovo and Plischke, 2016). However, the dependence structure between input factors might substantially impact the sensitivities, as is illustrated in the following example.

Example 1 (*Non-linear insurance portfolio*)**.** Consider an insurance company with losses from three different lines of business, X_1, X_2 and X_3, each of which is subject to the same multiplicative factor X_4 arising from, e.g., inflation. The insurance company holds a reinsurance contract on the loss from the first two lines of business, $L = X_4(X_1 + X_2)$, with deductible d and limit l. This means that the aggregate loss faced by the insurer is

$$Y = L - \min\{(L - d)_+, l\} + X_3 X_4. \tag{1}$$

In this simple model, which we will use as a running example throughout the paper, we view (X_1, X_2, X_3, X_4) as the input factors and Y as the output. Assume that the input factors

(X_1, X_2, X_3, X_4) are dependent through a Gaussian copula with correlation matrix R:

$$R = \begin{pmatrix} 1 & 0.3 & 0 & 0.8 \\ 0.3 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0.8 & 0 & 0 & 1 \end{pmatrix}.$$

Each element r_{ij} of R closely approximates the Spearman rank correlation $\mathrm{Corr}^S(X_i, X_j)$ of the respective pair of input factors. Note that the aggregate portfolio loss Y is symmetric in X_1 and X_2 but the dependence structure is not; X_1 has a high rank correlation to X_4, $\mathrm{Corr}^S(X_1, X_4) = 0.8$, while $\mathrm{Corr}^S(X_2, X_4) = 0$. Measuring the sensitivity to X_1 and X_2 via a local method involving $\frac{\partial Y}{\partial X_1}$, $\frac{\partial Y}{\partial X_2}$ obviously fails to reflect the asymmetry in statistical dependence. More generally, as is demonstrated in the later Example 4, such asymmetry is also not fully reflected by global sensitivity measures based on partial derivatives, in the vein of Tsanakas and Millossovich (2016).

Example 2 (*London Insurance Market portfolio*)**.** Consider the situation of a model user or reviewer, who has only partial access to the model specifications. It is typical in risk management applications for models to be high dimensional, with calculation of the model's output distribution proceeding by Monte Carlo simulation (Arbenz et al., 2012; Choe et al., 2018; Risk and Ludkovski, 2018). A model user will often be supplied with a set of simulated scenarios from variables of interest (model inputs and outputs), without easy access to either (a) the distributional assumptions of inputs (which may themselves be outputs from sub-models) or (b) the model function mapping inputs to outputs (which may be highly non-linear and computationally expensive to evaluate). This situation is typical in the regulatory review of internal models in insurance (Cadoni, 2014).

For illustration of those points, we consider a proprietary model of a London Insurance Market portfolio, currently in use by a participant in that market. The model represents a portfolio with 72 input factors; the output is the portfolio loss. We do not have access to the marginal nor the joint distribution of the input vector; indeed, the input factors are themselves outputs of different sub-models. We were supplied by the model owner with a Monte Carlo sample of size $M = 500,000$, consisting of simulated observations from the model's inputs and corresponding output. We have no access to the data generating mechanism, hence we cannot

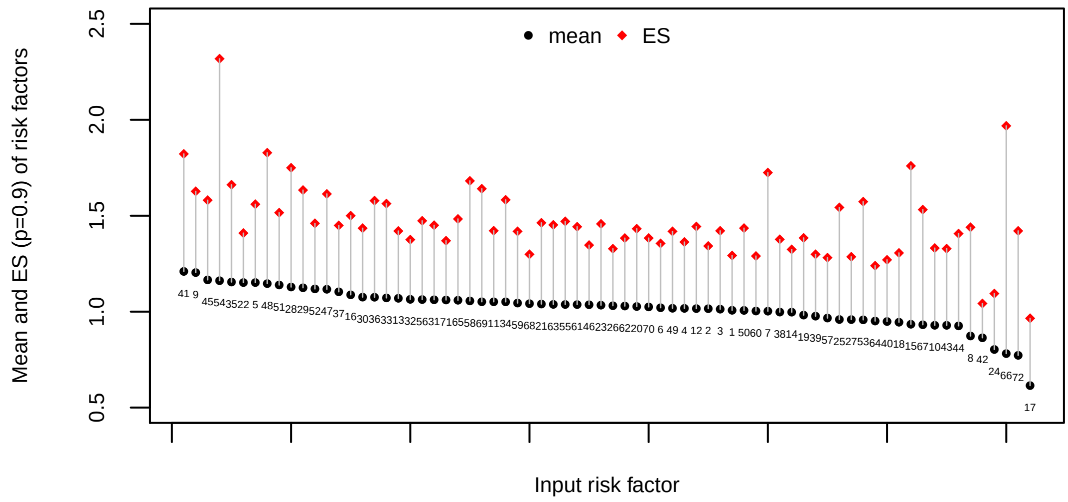

Figure 1: Expected value and $\mathrm{ES}_{0.9}$ risk measure of input risk factors in the London Market portfolio.

re-simulate under different model assumptions.

In Figure 1 we summarize the empirical distributions of individual input risk factors, by plotting the mean and the Expected Shortfall risk measure (at 90% level) of each. The question we aim to answer is: *with the given (lack of) information about the model, how can one assess the impact of stresses in a particular risk factor, taking fully into account the dependence between inputs?* We return to this example in Section 5.

In this paper, we aim to address the issues raised by the preceding two examples. We propose a novel sensitivity measure, termed *cascade sensitivity*, defined as the partial derivative of a distortion risk measure applied to the output, in the direction of a *stressed* input factor. This definition is closely related to the approaches of Hong (2009); Hong and Liu (2009); Tsanakas and Millossovich (2016). However, in our case the derivative is taken after a suitable transformation of the random vector of inputs, which enables cascade sensitivity measures to fully capture the impact of dependence between input factors, including indirect effects, such as those discussed in Example 1. Specifically, our cascade sensitivity framework is underpinned by a variation of the inverse Rosenblatt transform (Rosenblatt, 1952), which permits a stress on one input factor to propagate through the entire input vector, changing all its components according to the input vector's dependence structure. Thus, a stress on an input impacts on the output risk measure both directly and indirectly, via the generated cascade of stresses on other (dependent) inputs. In particular, the cascade sensitivity of an input factor decomposes into components, each reflecting the direct or indirect contribution of an input factor to the sensitivity of the output.

A sensitivity measure that fully reflects the dependence of the random vector of inputs is useful in applications where the dependence structure of the inputs is of particular interest, as in risk management applications (Glasserman and Xu, 2014; Lam, 2017). Challenges to practical application of sensitivity analysis methods include the incomplete specification of (black-box) models, as well as high computational costs (Saltelli et al., 2008; Lam and Qian, 2018). We demonstrate how the decomposition of cascade sensitivities can be calculated using only bivariate copula information. Furthermore, we extend our approach to situations where input factors may live on different scales, an issue identified by Antoniano-Villalobos et al. (2018). Finally, we provide representations of the cascade sensitivity which do not require the gradient of the aggregation function and allow for a straightforward implementation on a single Monte Carlo sample. In the case that input factors are independent, these representations allow for an alternative evaluation of marginal sensitivities studied by Hong (2009), without knowledge of the gradient of the model function. Hence, our proposed cascade sensitivity framework is practically useful, as illustrated through an application to the commercially used London Insurance Market portfolio model introduced above.

1.2 Relation to existing literature

Global sensitivity analysis, which assesses the importance of model inputs or parameters over a space of randomly generated outcomes, is established as a prominent tool for risk analysis in numerous fields. Recent examples include life cycle impact assessment (Cucurachi et al., 2016), land subsidence modelling (Sundell et al., 2019), seismic risk assessment (Foulser-Piggott et al., 2020), flood risk management (Oddo et al., 2020), and mortality modelling in insurance (Rabitti and Borgonovo, 2020). The success of those methods has led to extensions (Xiao et al., 2018) and higher-level conceptualisations of global sensitivity analysis methods (Borgonovo et al., 2016); for a wide-ranging review see Borgonovo and Plischke (2016).

In this paper, we offer a methodological contribution to the broad field of global sensitivity analysis. Specifically, we build on a strand of literature that is concerned with directional derivatives of expected utility (Antoniano-Villalobos et al., 2018) or risk measure (Hong, 2009; Tsanakas and Millossovich, 2016) functionals. However, in this stream of literature, sensitivity measures generally do not account explicitly for the impact of the dependence between input factors. This matters, given that the impact of (stochastic) dependencies is a persistent topic in risk modelling; see for example Mornet et al. (2015), Su et al. (2015), Wang et al. (2016), and Werner et al. (2018). Hence, we are concerned with reflecting the impact of the dependence between input factors in sensitivity metrics. For that purpose, we make use of the Rosenblatt transform, which is well known in statistics (e.g. in probability forecasting (Dawid, 1984)), but, to our knowledge, has not been utilised in the context of sensitivity analysis. Nonetheless, there are some conceptual parallels of our work with Mara and Tarantola (2012), who consider multivariate normal variables in a variance-based sensitivity framework, Mai et al. (2015), who study model robustness via a transformation of the input vector, and Kraus and Czado (2017), who carry out bank stress testing using graphical dependence models.

1.3 Structure of the paper

Section 2 introduces the necessary notation and discusses the choice of stresses on input factors. In Section 3 the cascade sensitivity measure is defined. The impact of input vectors' dependence structure on sensitivity is illustrated through numerical examples, including an application to factor prioritization, as well as a data-driven approach to model uncertainty. Section 4 is devoted to the calculation of the cascade sensitivity and its decomposition, using a bivariate copula approach. In Section 5, a detailed application of cascade sensitivity is presented, to the commercially used London Insurance Market portfolio discussed in Example 2. Technical assumptions are gathered in the Appendix A, proofs and further detailed analytical calculations are provided in the electronic companion.

2 Models and stresses

2.1 Preliminaries

Throughout the paper we work with a probability space $(\Omega, \mathcal{A}, P)$ and a random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ whose components, $X_1, \ldots, X_n$, represent *input factors*. The input factors $X_1, \ldots, X_n$ are assumed to be in $\mathcal{L}^1(\Omega, \mathcal{A}, P)$. We denote by F_j the marginal distribution function of the input X_j, $j = 1, \ldots, n$, and by $\boldsymbol{F}$ the joint distribution function of $\boldsymbol{X}$. It is assumed that the joint density $\boldsymbol{f}$ of $\boldsymbol{X}$ exists and we denote by f_j the marginal density of input factor $X_j,\ j = 1, \ldots, n$. The vector of input factors, $\boldsymbol{X}$, is mapped by an *aggregation function*, $g\colon \mathbb{R}^n \to \mathbb{R}$, assumed to be almost everywhere differentiable, to the (univariate) output $Y = g(\boldsymbol{X})$. We write H, h for the distribution function and the density of the output Y, respectively. Taken together, the input factors $\boldsymbol{X}$ and the aggregation function g constitute our *model*.

The left inverse of the distribution of any random variable $W \sim F_W$ is defined by $F_W^{-1}(u) = \inf\{x \in \mathbb{R} \mid F_W(x) \geq u\},\ u \in (0, 1]$. We use the notation U_W for a standard uniform random variable comonotonic to W, that is, $W = F_W^{-1}(U_W)$ a.s. In the case when W has a continuous distribution function, it holds that $U_W = F_W(W)$ a.s. Moreover, for an n-dimensional vector $\boldsymbol{W}$, we denote by $\boldsymbol{W}_{-j} = (W_1, \ldots, W_{j-1}, W_{j+1}, \ldots W_n)$ its sub-vector deprived of the j^{th} component.

The distribution of the output $Y = g(\boldsymbol{X}) \sim H$, representing a decision variable, is summarized through a risk measure. Risk measures are tools in financial risk management to assess different levels of risk severity (Artzner et al., 1999; Föllmer and Schied, 2011). Here we work with the class of *distortion risk measures* (Wang, 1996; Acerbi and Tasche, 2002; Belles-Sampera et al., 2014), which are defined through

$$\rho_\gamma(Y) = \int_0^1 H^{-1}(u)\gamma(u)\mathrm{d}u = E\left(H^{-1}(U_Y)\gamma(U_Y)\right),$$

where $\gamma\colon [0, 1] \to [0, \infty)$ is a normalized weight function such that $\int_0^1 \gamma(u)\mathrm{d}u = 1$. We assume the input factors $\boldsymbol{X}$ and the aggregation g to be such that the considered distortion risk measure $\rho(g(\boldsymbol{X}))$ is finite. The focus on distortion risk measures is not restrictive, as the proposed framework is also applicable for utility-type performance measures, see the remark at the end of Section 3.1. Examples of distortion risk measures include the two most widely used risk measures in risk management applications, the VaR and ES. The VaR of the random variable Y at level $\alpha \in (0, 1)$ is defined as the left α-quantile, $\mathrm{VaR}_\alpha(Y) = H^{-1}(\alpha)$ or through the weight function $\gamma(u) = \delta_\alpha(u)$, for the Dirac measure δ_α. The ES, also called Conditional Value-at-Risk, at level $\alpha \in [0, 1)$ arises from $\gamma(u) = \frac{1}{1-\alpha}\mathbb{1}_{\{u>\alpha\}}$, thus has representation $\mathrm{ES}_\alpha(Y) = \frac{1}{1-\alpha}\int_\alpha^1 H^{-1}(u)\mathrm{d}u$. Throughout this paper, the examples will be based on ES.

2.2 Stressing input factors

The objective of this paper lies in the study of the sensitivity of $\rho_\gamma(Y)$ to stressing input factor $X_i,\ 1 \leq i \leq n$. Specifically, the sensitivity measure introduced in this paper aims to capture how a stress on an input X_i propagates through both the vector of input factors $\boldsymbol{X}$ and, subsequently, the aggregation function, thus impacting on the risk measure of the stressed model output. In this section, we discuss the choice of stress on the input factor X_i. The question on how precisely a stress propagates through the model is considered in Section 3.2.

For simplicity, we fix $i \in \{1, \ldots, n\}$ for the rest of the paper, such that sensitivity to the same input is considered throughout. We call a *stress* on input factor X_i a family of random variables $X_{i,\varepsilon}(\omega) = K(X_i(\omega), \varepsilon)$, $\varepsilon \geq 0$, $\omega \in \Omega$, for some mapping $K(x, \varepsilon)$ that is differentiable in ε in a neighbourhood of 0 for all x. Moreover, K satisfies $K(x, 0) = x$, for all $x \in \mathbb{R}$. In particular, for any stress $X_{i,\varepsilon}$, it holds that $(X_1, \ldots, X_{i,\varepsilon}, \ldots, X_n)|_{\varepsilon=0} = \boldsymbol{X}$ a.s. We denote by $F_{i,\varepsilon}$, $\varepsilon \geq 0$, the distribution function of $X_{i,\varepsilon}$.

Here we discuss different options for constructing a stress on X_i. A first consideration is that, while in this paper we aim to reflect the dependence structure of $\boldsymbol{X}$ in assessing the impact of stresses on X_i, we do not stress the dependence structure itself. Thus, we require that the copula of $(X_1, \ldots, X_{i,\varepsilon}, \ldots, X_n)$ is the same as that of $\boldsymbol{X}$. This is achieved by choosing $K(x, \varepsilon)$ non-decreasing in x for all ε, which leads to $X_{i,\varepsilon}$ being comonotonic to X_i.

Further considerations for specifying a stress follow from the intended use of the sensitivity analysis. In this paper, we consider two applications, factor prioritization and model uncertainty; we refer to Borgonovo and Plischke (2016); Saltelli et al. (2008) for a comprehensive review on sensitivity analysis and its applications.

Factor prioritization. In a *factor prioritization* (Saltelli et al., 2008) setting, we aim to rank input factors depending on their relative importance in the model. For this purpose we limit ourselves to additive *shocks* on X_i of the form:

$$X_{i,\varepsilon} = X_i + \varepsilon k(X_i),$$

for a non-decreasing function $k \colon \mathbb{R} \to \mathbb{R}$. Thus, $\varepsilon k(X_i)$ specifies the additional risk exposure to be propagated through the model – factors are prioritized, which, after stressing, produce a higher impact on $\rho_\gamma(g(\boldsymbol{X}))$. An additive shock is consistent with the use of Gateaux (or directional) derivatives in assessing the sensitivity of statistical functionals, see e.g. Kalkbrener (2005) in the setting of capital allocation, Tsanakas and Millossovich (2016) in the field of risk analysis, and Kromer et al. (2016) in the context of systemic risk measurement.

Beyond the monotonicity requirement, the function k needs to satisfy continuity conditions;

specifically we require k to be Lipschitz continuous. Lipschitz continuity guarantees finiteness and robustness of $\rho_\gamma(g(\boldsymbol{X}))$ (in the sense of Pesenti et al. (2016)), with respect to substituting $X_{i,\varepsilon}$ for X_i; see Appendix A.2 for a more detailed argument. Further, when comparing sensitivities to input factors, the same type of stress should be applied to each input; thus, we will generally restrict to k having Lipschitz constant 1. Examples of stresses that satisfy these conditions are:

$$
\begin{aligned}
X_{i,\varepsilon} &= X_i + \varepsilon(X_i - m), \\
X_{i,\varepsilon} &= X_i + \varepsilon(X_i - t_1)\mathbb{1}_{\{X_i \le t_1\}} + \varepsilon(X_i - t_2)\mathbb{1}_{\{X_i \ge t_2\}}.
\end{aligned}
$$

In the first expression, m may be a measure of central tendency (e.g. mean) when one is interested in stressing the volatility of X_i (Tsanakas and Millossovich, 2016) or may be set to zero for a proportional stress. In the second, one only stresses the tails of X_i, consistently with a risk management perspective that places importance on the potential of extreme outcomes. Such distortions of tail quantiles are consistent with typical risk management considerations. This is the perspective we take in Example 4, where a stress on the right tail is adopted, following from the interpretation of X_i as an insurance loss. A proportional stress on (transformed) input factors is proposed in Section 5, where we deal with the problem of designing comparable stresses for input factors that may live on different scales.

Model uncertainty. A different situation emerges when there is uncertainty around the distribution of the input factor X_i. Here F_i, the distribution of X_i, gives an estimated baseline model, but with the understanding that an uncertainty set $\mathcal{F}$ of alternative technically plausible distributions for X_i exists. In this context, sensitivity analysis is meant to detect the potential impact of mis-specifying the distribution of X_i on the aggregate risk assessment $\rho_\gamma(g(\boldsymbol{X}))$. The stress is then constructed by a *perturbation* of the distribution of X_i, arising from a mixture of F_i with some distribution $\hat{F}_i \in \mathcal{F}$, where the set $\mathcal{F}$ contains distributions with finite absolute mean. The use of such perturbations is common in sensitivity analysis and in Bayesian and robust statistics (Hampel et al., 2011; Glasserman, 1991), as well as in financial risk management (Cont et al., 2010). Specifically, we propose the comonotonic perturbation:

$$
\begin{gathered}
X_{i,\varepsilon} = F_{i,\varepsilon}^{-1}(U_{X_i}), \quad \text{where} \\
F_{i,\varepsilon}(x) = (1-\varepsilon)F_i(x) + \varepsilon \hat{F}_i(x), \quad \hat{F} \in \mathcal{F}.
\end{gathered}
$$

We implement this approach to sensitivity analysis of model uncertainty in Example 5, where the set $\mathcal{F}$ arises from a bootstrapping procedure, leading to a data-driven specification of stresses.

3 Sensitivity measures

3.1 Marginal sensitivity

To assess the sensitivity of the output Y to the input X_i, sensitivity measures are defined. The approach we follow here is to take a directional derivative of the risk measure applied to the output distribution, in the direction of a stress to input X_i.

Definition 3.1. For a stress $X_{i,\varepsilon}$ and a distortion risk measure ρ_γ, we define the *marginal sensitivity* to input factor X_i by

$$\mathcal{S}_i(\boldsymbol{X}, g, \rho_\gamma) = \frac{\partial}{\partial \varepsilon} \rho_\gamma\big(g(X_1, \ldots, X_{i,\varepsilon}, \ldots, X_n)\big)\big|_{\varepsilon=0},$$

whenever the derivative exists.

The general form of the marginal sensitivity for distortion risk measures follows directly from Hong and Liu (2009) and is stated in the next proposition for completeness. It consists of an expectation involving the derivative of the stress, the partial derivative of the aggregation function in the direction of the stressed input factor and a weighting according to the chosen risk measure.

For the existence of the sensitivity measures considered in this paper, additional assumptions are needed on the distribution of the model output, following a stress on X_i. These are detailed in Assumption A.1 in the Appendix and can be summarized as follows: the derivative of the stressed output's distribution function and the derivative of its quantile function, with respect to ε, are, respectively, continuous and bounded in ε. We refer to Hong (2009) for a detailed discussion of those assumptions. In particular, the continuity of the stressed output is usually satisfied when g is continuous and the model is rich enough such that some of the elements of $\boldsymbol{X}$ are continuous.

Proposition 3.2. Given a stress $X_{i,\varepsilon}$ and under Assumptions A.1 in the appendix, the marginal sensitivity to input factor X_i is

$$\mathcal{S}_i(\boldsymbol{X}, g, \rho_\gamma) = E\Big(\frac{\partial}{\partial \varepsilon} X_{i,\varepsilon}\big|_{\varepsilon=0}\, g_i(\boldsymbol{X})\gamma(U_Y)\Big),$$

where $g_i(\boldsymbol{x}) = \frac{\partial}{\partial x_i} g(\boldsymbol{x})$ denotes the partial derivative of the aggregation function in the i^{th} component and $\frac{\partial}{\partial \varepsilon} X_{i,\varepsilon}(\omega) = \frac{\partial}{\partial \varepsilon} K(X_i(\omega), \varepsilon)$, for almost all $\omega \in \Omega$.

Remark. Definition 3.1 also includes sensitivity of expected utilities, considered in Cao and Wan (2017); Antoniano-Villalobos et al. (2018). Note that, for the trivial weight function $\gamma \equiv 1$, the distortion risk measure reduces to the expectation, $\rho_1(\cdot) \equiv E(\cdot)$. Thus, for a utility function $u \colon \mathbb{R} \to \mathbb{R}$, we can write

$$E(u(g(\boldsymbol{X})) = \rho_1\big((u \circ g)(\boldsymbol{X})\big),$$

implying that expected utilities are a special case of our framework, with aggregation function $u \circ g \colon \mathbb{R}^n \to \mathbb{R}$ and an expectation risk measure.

3.2 Inverse Rosenblatt transforms

The marginal sensitivity of Definition 3.1 does not fully account for interactions among or dependence between input factors, since, by its definition, when stressing one input factor, all other marginal input distributions remain unaltered; see also the discussion in Borgonovo and Plischke (2016). Note that the representation of the marginal sensitivity in Proposition 3.2 incorporates the derivative of the aggregation function solely in the direction of the stressed input factor.

In order to address the indirect effects induced by the dependence between the input factors, we utilize a representation of random vectors, termed *inverse Rosenblatt transform*[1] (Rosenblatt, 1952; Rüschendorf and de Valk, 1993).

Definition 3.3. An inverse Rosenblatt transform of an n-dimensional random vector $\boldsymbol{X}$, starting at X_i, is given by a differentiable function $\boldsymbol{\psi} = (\psi^{(1)}, \dots, \psi^{(n)})^\top \colon \mathbb{R}^n \to \mathbb{R}^n$ and a $(n-1)$-dimensional random vector $\boldsymbol{V} = (V_1, \dots, V_{n-1})$, consisting of independent standard uniform variables, independent of X_i, such that

$$\boldsymbol{X} = \boldsymbol{\psi}(X_i, \boldsymbol{V}) = \big(\psi^{(1)}(X_i, \boldsymbol{V}), \dots, \psi^{(n)}(X_i, \boldsymbol{V})\big) \text{ a.s.} \tag{2}$$

The set of inverse Rosenblatt transforms of $\boldsymbol{X}$, starting at X_i, is denoted by $\mathcal{R}_i = \{(\boldsymbol{\psi}, \boldsymbol{V}) \mid \boldsymbol{X} = \boldsymbol{\psi}(X_i, \boldsymbol{V})\}$.

[1] We call representation (2) the *inverse Rosenblatt transform* since, to be precise, the Rosenblatt transform is the transformation from a random vector to a vector consisting of uniform random variables.

It can be shown that for $\psi^{(j)},\ 1 \le j \le n$, to exist and be differentiable in the first component, it is sufficient that the joint density $\boldsymbol{f}$ is almost everywhere differentiable.

An inverse Rosenblatt transform can be explicitly constructed via the following process (Rüschendorf and de Valk, 1993; Rubinstein and Melamed, 1998). For $r = 1, \dots, n$ and $J \subseteq \{1, \dots, n\} \backslash \{r\}$, denote by $F_{r|J}(\cdot \mid x_j,\ j \in J)$ the conditional distribution function of X_r given $X_j = x_j,\ j \in J$. Then, it holds a.s. that

$$\begin{aligned}
X_1 &= F^{-1}_{1|i}(V_1 \,|\, X_i) = \psi^{(1)}(X_i, \boldsymbol{V}),\\
X_2 &= F^{-1}_{2|i,1}(V_2 \,|\, X_i, X_1) = \psi^{(2)}(X_i, \boldsymbol{V}),\\
X_3 &= F^{-1}_{3|i,1,2}(V_3 \,|\, X_i, X_1, X_2) = \psi^{(3)}(X_i, \boldsymbol{V}),\\
&\ \vdots\\
X_i &= \psi^{(i)}(X_i, \boldsymbol{V}),\\
&\ \vdots\\
X_n &= F^{-1}_{n|1,\dots,n-1}(V_{n-1} \,|\, X_1, \dots, X_{n-1}) = \psi^{(n)}(X_i, \boldsymbol{V}),
\end{aligned}$$

where $\psi^{(i)}$ is the identity function in the first argument. Note that in the above construction, each random variable X_j depends on X_i both directly and indirectly through $X_1, \dots, X_{j-1}$.

Deploying an inverse Rosenblatt transform of the vector $\boldsymbol{X} = \boldsymbol{\psi}(X_i, \boldsymbol{V}),\ (\boldsymbol{\psi}, \boldsymbol{V}) \in \mathcal{R}_i$, we can stress $\boldsymbol{X}$ through $X_{i,\varepsilon}$

$$\boldsymbol{X}_{i,\varepsilon} = \boldsymbol{\psi}(X_{i,\varepsilon}, \boldsymbol{V}) = \big(\psi^{(1)}(X_{1,\varepsilon}, \boldsymbol{V}), \dots, \psi^{(n)}(X_{1,\varepsilon}, \boldsymbol{V})\big). \tag{3}$$

Observe that the stress $X_{i,\varepsilon}$ is carried through the entire input vector, changing all factors according to their dependence on $X_{i,\varepsilon}$, resulting in a cascading effect.

Example 3 (*Inverse Rosenblatt transform for a Gaussian copula*)**.** Let $\boldsymbol{X}$ have a Gaussian copula with correlation matrix R, continuous marginals $F_j,\ j = 1, \dots, n$, and assume for simplicity, in this example, that $i = 1$. For a vector of independent standard uniforms $\boldsymbol{V} = (V_1, \dots, V_{n-1})$ independent of X_1, define the standard normal variables $Z_1 = \Phi^{-1}\big(F_1(X_1)\big)$ and $T_j = \Phi^{-1}(V_j),\ j = 1, \dots, n-1$, where Φ denotes the standard normal distribution function. Then, an inverse Rosenblatt transform starting from X_1, $(\boldsymbol{\psi}, \boldsymbol{V}) \in \mathcal{R}_1$, is derived by setting

$$\psi^{(1)}(X_1, \boldsymbol{V}) = X_1$$

$$\psi^{(2)}(X_1, \boldsymbol{V}) = F_2^{-1}\big(\Phi\left(d_{21}Z_1 + d_{22}T_1\right)\big)$$

$$\vdots$$

$$\psi^{(n)}(X_1, \boldsymbol{V}) = F_n^{-1}\big(\Phi\left(d_{n1}Z_1 + d_{n2}T_1 + \cdots + d_{nn}T_{n-1}\right)\big),$$

where $D = (d_{jr})_{1\leq j,r\leq n}$ is the lower triangular matrix resulting from the Cholesky decomposition of R. Following this representation, one can stress X_1 through substitution by $X_{1,\varepsilon}$, which implies replacing Z_1 by $Z_{1,\varepsilon} = \Phi^{-1}\big(F_1(X_{1,\varepsilon})\big)$, in the right hand-side of the above equations. It is then apparent how the stress on X_1 also produces a stress on $X_j,\ j = 2,\ldots,n$ (provided $d_{j1} \neq 0$).

3.3 Cascade sensitivity

The inverse Rosenblatt transform, by propagating stresses on one input factor across the vector of inputs, allows us to construct a sensitivity measure that fully reflects both the direct and the indirect impacts on the output.

Definition 3.4. For a stress $X_{i,\varepsilon}$, $(\boldsymbol{\psi}, \boldsymbol{V}) \in \mathcal{R}_i$ and a distortion risk measure ρ_γ, we define the *cascade sensitivity* to input factor X_i by

$$\mathcal{C}_i(\boldsymbol{X}, g, \rho_\gamma) = \frac{\partial}{\partial\varepsilon}\rho_\gamma\big(g\big(\boldsymbol{\psi}(X_{i,\varepsilon}, \boldsymbol{V})\big)\big)\big|_{\varepsilon=0},$$

whenever the derivative exists.

Hence, the cascade sensitivity framework directly extends approaches to sensitivity analysis that are based on partial derivatives in the direction of input factors. In particular, the cascade and the marginal sensitivities differ in the way a stress on an input propagates through the model. Though the current paper is focused on distortion risk measures (and expected utilities), the use of the inverse Rosenblatt transform for sensitivity analysis is a tool that could also be used in other (e.g. variance-based) sensitivity analysis approaches (Mara and Tarantola, 2012).

The cascade sensitivity to an input factor can be decomposed into the marginal sensitivity and additional components, each reflecting statistical, as well as functional, dependence between inputs.

Proposition 3.5. Given a stress $X_{i,\varepsilon}$, $(\boldsymbol{\psi}, \boldsymbol{V}) \in \mathcal{R}_i$, and under Assumptions A.1 in the appendix, the cascade sensitivity to input factor X_i is given by

$$\mathcal{C}_i(\boldsymbol{X}, g, \rho_\gamma) = \sum_{j=1}^{n} \mathcal{C}_{i,j}, \tag{4}$$

where

$$\mathcal{C}_{i,j} = E\Big(\frac{\partial}{\partial \varepsilon} X_{i,\varepsilon}\Big|_{\varepsilon=0} g_j(\boldsymbol{X}) \psi_1^{(j)}(X_i, \boldsymbol{V}) \gamma(U_Y)\Big) \tag{5}$$

and $\psi_1^{(j)}(x, \boldsymbol{v}) = \frac{\partial}{\partial x} \psi^{(j)}(x, \boldsymbol{v})$.

The set of inverse Rosenblatt transforms of a random vector, $\mathcal{R}_i$, is generally not a singleton, implying that the inverse Rosenblatt transform is not unique. For instance, in the last example, a different transform would be obtained for some permutation of $(X_2, \dots, X_n)$. However, as the next result shows, the cascade sensitivity does not depend on the particular choice of inverse Rosenblatt transform.

Proposition 3.6. For a stress $X_{i,\varepsilon}$, if the cascade sensitivity exists for one $(\boldsymbol{\psi}, \boldsymbol{V}) \in \mathcal{R}_i$, then it exists and admits the same value for all other transforms $(\boldsymbol{\phi}, \boldsymbol{U}) \in \mathcal{R}_i$. Moreover, each sensitivity $\mathcal{C}_{i,j}$ in (5) has the same value regardless of the choice of inverse Rosenblatt transform.

The decomposition of the cascade sensitivity $\mathcal{C}_i$ in Proposition 3.5 allows to quantify the contribution of each input factors' indirect effects, stemming from the dependence with the one being stressed, to $\mathcal{C}_i$. Specifically, $\mathcal{C}_{i,j}$ is the indirect contribution of input X_j to the sensitivity $\mathcal{C}_i$, when stressing input factor X_i. Moreover, note that $\mathcal{C}_{i,i} = \mathcal{S}_i$ is the marginal sensitivity; hence $\mathcal{C}_i - \mathcal{C}_{i,i}$ is the component of the cascade sensitivity which is solely due to the dependence structure between input factors. The components $\mathcal{C}_{i,1}, \dots, \mathcal{C}_{i,n}$, contributing to the cascade sensitivities $\mathcal{C}_i, i = 1, \dots, n$, can be visualized in Table 1. Each row shows a cascade sensitivity decomposed into its n-summands, $\mathcal{C}_i = \sum_{j=1}^{n} \mathcal{C}_{i,j},\ i = 1, \dots, n$. The diagonal contains the marginal sensitivities, while off-diagonal elements $\mathcal{C}_{i,j}$ reflect the indirect contribution of X_j to the cascade sensitivity of X_i.

Table 1: Illustration of cascade and marginal sensitivities in Proposition 3.5.

	X_1	X_2	$\dots$	X_n
$\mathcal{C}_1$	$\mathcal{S}_1$	$\mathcal{C}_{1,2}$	$\dots$	$\mathcal{C}_{1,n}$
$\mathcal{C}_2$	$\mathcal{C}_{2,1}$	$\mathcal{S}_2$	$\dots$	$\mathcal{C}_{2,n}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\mathcal{C}_n$	$\mathcal{C}_{n,1}$	$\mathcal{C}_{n,2}$	$\dots$	$\mathcal{S}_n$

Proposition 3.5 shows that the cascade sensitivity decomposes into the marginal sensitivity and components reflecting the dependence between the input factors. Thus, a natural question arises as to whether, in general, positive (negative) dependence among inputs results into a larger (smaller) cascade sensitivity compared to its marginal counterpart. We provide two results in this direction in Proposition A.3 in Appendix A.3. First, for independent input factors the cascade sensitivity reduces to the marginal sensitivity, irrespective of the aggregation function or the choice of distortion risk measure. Second, the cascade sensitivity dominates the marginal sensitivity, given positive dependence of the input vector, a non-decreasing aggregation function and e.g. an additive stress.

The following two examples illustrate the ideas discussed above.

Example 4 (*Non-linear insurance portfolio continued*)**.** We calculate the cascade sensitivity for the insurance portfolio example introduced in Section 1. The marginal distributions of input factors are summarized in Table 2 and we set the deductible $d = 380$ and the limit $l = 30$. Calculations are based on a simulated Monte Carlo sample of size $M = 100,000$.

Table 2: Distributional assumptions of input factors of the non-linear insurance portfolio example.

Input	Distribution	Mean	Std
X_1	Log-Normal $\left(4.98,\ 0.23^2\right)$	150	35.0
X_2	Log-Normal $\left(4.98,\ 0.23^2\right)$	150	35.0
X_3	Gamma $(100,\ 1)$	100	10.0
X_4	Log-Normal $\left(-0.005,\ 0.1^2\right)$	1	0.1

We consider an additive tail shock to the first two lines of business defined by $X_{i,\varepsilon} = X_i + \varepsilon(X_i - t_i)\mathbb{1}_{\{X_i>t_i\}}$, $i = 1, 2$, where $t_i = F_i^{-1}(0.9)$ is the 90% quantile of either. Table 3 reports the cascade sensitivities to inputs X_1 and X_2 for the risk measure $\mathrm{ES}_{0.9}$. By construction, inputs X_1 and X_2 represent directly comparable variables and are stressed in exactly the same fashion, allowing for a direct comparison of $\mathcal{C}_1$ and $\mathcal{C}_2$. In Section 5, we discuss how to consistently stress input factors for models whose input factors may live on different scales.

In the current example, the terms $\mathcal{C}_{i,j}$, reflecting the indirect effects of the dependence between the input factors, are proportional to their Spearman rank correlation with the stressed input factor, see also Proposition 4.1, case 1. Note that $\mathrm{Corr}^S(X_2, X_3) = \mathrm{Corr}^S(X_2, X_4) = 0$, hence $\mathcal{C}_{2,3} = \mathcal{C}_{2,4} = 0$, and thus $\mathcal{C}_2 = \mathcal{C}_{2,1} + \mathcal{S}_2$ only accounts for the positive dependence between X_1 and X_2. This is in contrast to $\mathcal{C}_1$, where $\mathcal{C}_{1,4}$, the indirect impact of X_4, constitutes

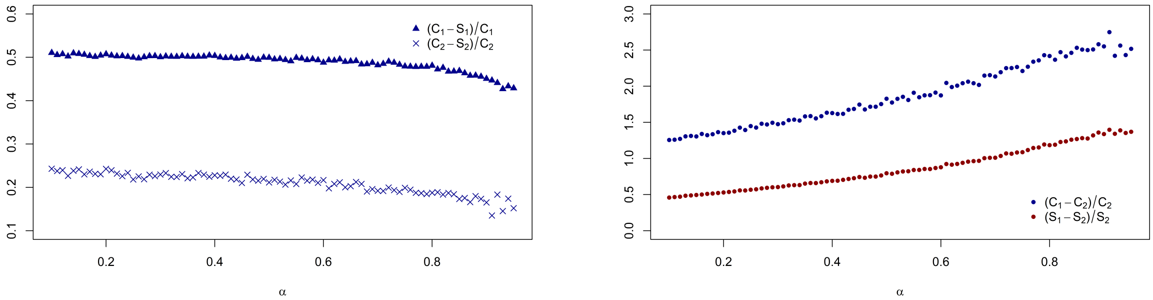

Figure 2: Sensitivities for the $\mathrm{ES}_{0.9}$. Left: $(\mathcal{C}_i - \mathcal{S}_i)/\mathcal{C}_i$, for $i = 1$ (triangles) and $i = 2$ (crosses), against different levels of tail-stresses α, with $t_i = F_i^{-1}(\alpha)$. Right: $(\mathcal{C}_1 - \mathcal{C}_2)/\mathcal{C}_2$ (blue) and $(\mathcal{S}_1 - \mathcal{S}_2)/\mathcal{S}_2$ (red) against different levels of tail-stresses α.

a substantial $38\% = \frac{19.43}{51.15}$ of the cascade sensitivity to X_1.

Table 3: The cascade sensitivities to inputs X_1, X_2 (along with standard errors) for $\mathrm{ES}_{0.9}$ and their decomposition into the direct effect of the stressed input ($\mathcal{C}_{i,i} = \mathcal{S}_i$) and the indirect effects of the other input factors ($\mathcal{C}_{i,j}$).

		$\mathcal{C}_{i,1}$		$\mathcal{C}_{i,2}$		$\mathcal{C}_{i,3}$	$\mathcal{C}_{i,4}$	
$\mathcal{C}_1 = 51.15$	(0.009)	26.02	(0.005)	5.70	(0.001)	0	19.43	(0.003)
$\mathcal{C}_2 = 13.96$	(0.004)	2.86	(0.001)	11.10	(0.003)	0	0	

Figure 2 (left) displays the quantities $(\mathcal{C}_i - \mathcal{S}_i)/\mathcal{C}_i$, $i = 1, 2$, reflecting the impact of the dependence on the cascade sensitivity of X_1 and X_2 for different levels of tail stress α using the risk measure $\mathrm{ES}_{0.9}$. It is seen that, consistently with Table 3, a substantial 50% of the cascade sensitivity to input factor X_1 stems from the dependence. The right-hand plot in Figure 2 shows the percentage increase of the sensitivity measures of X_1, compared to X_2. The plot indicates that input factor X_1 is more important both in the marginal and cascade sensitivity for all levels of tail-stresses, with the difference between X_1 and X_2 becoming more pronounced for high α (i.e. at the extreme right tail).

To further illustrate the effects of the dependence structure of the input vector on cascade sensitivities, we report the marginal and cascade sensitivities for $\mathrm{ES}_{0.9}$, with $\mathrm{Corr}^S(X_1, X_4) = 0.0, \ldots, 0.8$ in Table 4. Table 4 also states the percentage of the cascade sensitivity that stems solely from the effects of the dependence between inputs, $\frac{\mathcal{C}_j - \mathcal{S}_j}{\mathcal{C}_j}$, $j = 1, 2$. As seen in Table 4, both the marginal and the cascade sensitivity to X_1 increase with $\mathrm{Corr}^S(X_1, X_4)$, with $\mathcal{C}_1$ impacted more heavily. The marginal and cascade sensitivities to input X_2 are practically constant, implying that indirect effects of the dependence between X_1 and X_4 have a very minor impact on $\mathcal{C}_2$, which is in contrast to the cascade sensitivity to input X_1, which increases by 232% = (51.15 − 15.41)/15.41.

Table 4: Marginal and cascade sensitivity for the $\mathrm{ES}_{0.9}$ with $\mathrm{Corr}^S(X_1, X_4) = 0.0, \dots, 0.8$.

$\mathrm{Corr}^S(X_1, X_4)$	$\mathcal{C}_1$	$\mathcal{S}_1$	$\frac{\mathcal{C}_1-\mathcal{S}_1}{\mathcal{C}_1}$	$\mathcal{C}_2$	$\mathcal{S}_2$	$\frac{\mathcal{C}_2-\mathcal{S}_2}{\mathcal{C}_2}$
0.0	15.41	12.50	19%	15.41	12.50	19%
0.2	22.15	15.60	30%	14.95	12.06	19%
0.4	30.28	18.89	38%	14.58	11.70	20%
0.6	39.92	22.37	44%	14.24	11.37	20%
0.8	51.15	26.02	49%	13.96	11.10	20%

Example 5 (*Non-linear insurance portfolio continued (model uncertainty)*)**.** Here, we demonstrate how cascade sensitivities based on perturbations can be used to generate a data-driven measure of model uncertainty. We consider the same insurance portfolio model as in the last example, but now assuming that the parameters of the Log-Normal distributions of X_1 and X_2 are not known, but estimated from datasets of sizes M_1 and M_2 respectively, denoted by $\boldsymbol{x}_i = (x_{i1}, \dots, x_{iM_i})$, $i = 1, 2$. Our strategy is to generate model uncertainty sets for each of the two input factors by bootstrapping from the given datasets and re-fitting Log-Normal distribution parameters. Then, cascade sensitivities of input factors are calculated for perturbations with respect to every distribution in the respective uncertainty set.

Specifically, for each input $X_i,\ i = 1, 2$, we proceed as follows:

1. Estimate the parameter vector from the dataset $\boldsymbol{x}_i$ by its MLE, $\hat{\theta}_i = (\hat{\mu}(\boldsymbol{x}_i), \hat{\sigma}(\boldsymbol{x}_i))$ and denote by $F_i(\cdot|\hat{\theta}_i)$ the estimated distribution. Then, simulate 10^5 Monte Carlo samples of $(\boldsymbol{X}, Y = g(\boldsymbol{X}))$ using those estimated distributions.

2. Sample with replacement 1000 sets of M_i observations from $\boldsymbol{x}_i$, and denote the resulting vectors of bootstrapped observations by $\boldsymbol{x}_i^{(j)} = \big(x_{i1}^{(j)}, \dots, x_{iM_i}^{(j)}\big),\ j = 1, \dots, 1000$. For each bootstrapped sample, estimate the parameters $\hat{\theta}_i^{(j)} = \big(\hat{\mu}(\boldsymbol{x}_i^{(j)}), \hat{\sigma}(\boldsymbol{x}_i^{(j)})\big)$. Denote by $F_i(\cdot|\hat{\theta}_i^{(j)})$ the estimated distributions, using the bootstrapped parameters.

3. Evaluate the cascade sensitivities $\mathcal{C}_i(\boldsymbol{X}, g, \mathrm{ES}_{0.9})$ for each bootstrap sample, using the 10^5 Monte Carlo samples of step 1, with respect to perturbations $X_{i,\varepsilon}^{(j)} = F_{i,\varepsilon}^{-1,(j)}(U_{X_i})$, where $F_{i,\varepsilon}^{(j)}(\cdot) = (1-\varepsilon)F_i(\cdot|\hat{\theta}_i) + \varepsilon F_i(\cdot|\hat{\theta}_i^{(j)})$, $j = 1, \dots, 1000$. Denote the resulting cascade sensitivities for bootstrap sample j by $\mathcal{C}_i^{(j)}$.

For the datasets $\boldsymbol{x}_1, \boldsymbol{x}_2$ we consider three scenarios: (a) $M_1 = M_2 = 200$; (b) $M_1 = 100,\ M_2 = 1000$; (c) $M_1 = 1000,\ M_2 = 100$. In each scenario, the datasets $\boldsymbol{x}_1, \boldsymbol{x}_2$ are synthetically generated from a Log-Normal($\mu = 4.98,\ \sigma^2 = 0.23^2$) distribution, to maintain consistency with the previous example. The results are summarised in Table 5. For each scenario (a), (b) and (c), we report $\max_j \mathcal{C}_i^{(j)}$, the resulting maximum of the cascade sensitivities over all bootstrap iterations (i.e. over the space of alternative models in the uncertainty set). Furthermore, we report the average of each $\mathcal{C}_i^{(j)}$ over the its largest 5%, 10% and 50% (i.e. 50, 100, 500) outcomes.

Table 5: The cascade sensitivities to inputs X_1, X_2 for $\mathrm{ES}_{0.9}$, with respect to perturbations in the bootstrapped uncertainty set.

	$M_1 = 200,\ M_2 = 200$		$M_1 = 100,\ M_2 = 1000$		$M_1 = 1000,\ M_2 = 100$	
	X_1	X_2	X_1	X_2	X_1	X_2
Mean of 50% highest	10.37	4.13	15.67	1.68	3.96	4.25
Mean of 10% highest	27.12	8.77	38.54	3.57	8.78	10.88
Mean of 5% highest	33.65	10.25	46.20	4.15	10.50	13.26
Maximum	49.56	16.31	89.90	6.23	17.63	22.64

In scenario (a) $M_1 = M_2 = 200$, both X_1 and X_2 are subject to the same level of model uncertainty. The larger values of the cascade sensitivities for X_1 are due to the asymmetric dependence structure, consistently with Table 3. This phenomenon is exacerbated in the case $M_1 = 100,\ M_2 = 1000$, where X_1 is subject to larger model uncertainty than X_2, given the smaller dataset from which its parameters are estimated. Finally, when $M_1 = 1000,\ M_2 = 100$ the effect is reversed, with the cascade sensitivity of X_2 dominating that of X_1. Thus, the high model uncertainty that X_2 is exposed to in this scenario more than compensates for the impact of the dependence structure.

4 Evaluation of the cascade sensitivity

4.1 Direct evaluation of cascade sensitivity measures

Evaluating directly $\mathcal{C}_i$ and $\mathcal{C}_{i,j}$, as in Proposition 3.5, requires knowledge (or estimation) of the gradient of g and the functions $\psi^{(j)},\ j = 1, \ldots, n$, stemming from the inverse Rosenblatt transform. The gradient of g may be readily available, as is often assumed in the sensitivity analysis literature (e.g. Antoniano-Villalobos et al., 2018) or can be estimated using methods such as local regression (e.g. Tsanakas and Millossovich, 2016) and Gaussian process emulation (e.g. Bastos and O'Hagan, 2009; Risk and Ludkovski, 2018).

The function $\boldsymbol{\psi}$ requires the knowledge (or estimation) of the copula of $\boldsymbol{X}$. If a parametric form for the copula is given, the inverse Rosenblatt transform can often be obtained. Here, we present formulas for the cascade sensitivity, in the special cases of the popular Gaussian copula and t copula dependence models; these arise respectively by transforming the marginal distributions of a multivariate normal or multivariate t distribution (e.g. McNeil et al., 2015, Ch. 7). Inverse Rosenblatt transforms for the Archimedean and the elliptical copula families (Cambou et al., 2017, Sec. 3) are implemented in the **R** package *copula* (Hofert et al., 2017), and for the canonical and D-vine copulas (Aas et al., 2009) in the **R** package *VineCopula* (Nagler et al., 2019; Schepsmeier, 2015).

Proposition 4.1. Denote by Φ, ϕ, the distribution and density of a standard normal variable, and by t_ν, s_ν the distribution and density of a t-distributed random variable with ν degrees of freedom.

1. Let (X_i, X_j) have a Gaussian copula with correlation parameter r_{ij} and define $Z_i = \Phi^{-1}(U_{X_i})$, $Z_j = \Phi^{-1}(U_{X_j})$. Then, $\mathcal{C}_{i,j}$ becomes

$$\mathcal{C}_{i,j} = r_{ij} E\left[\left.\frac{\partial X_{i,\varepsilon}}{\partial \varepsilon}\right|_{\varepsilon=0} g_j(\mathbf{X}) \frac{f_i(X_i)\phi(Z_j)}{f_j(X_j)\phi(Z_i)} \gamma(U_Y)\right].$$

2. Let (X_i, X_j) have a t copula with correlation parameter r_{ij} and ν degrees of freedom and define $Z_i = t_\nu^{-1}(U_{X_i})$, $Z_j = t_\nu^{-1}(U_{X_j})$. Then, $\mathcal{C}_{i,j}$ becomes

$$\mathcal{C}_{i,j} = E\left[\left.\frac{\partial X_{i,\varepsilon}}{\partial \varepsilon}\right|_{\varepsilon=0} g_j(\mathbf{X}) \left(r_{ij} + \frac{Z_i Z_j - r_{ij} Z_i^2}{\nu + Z_i^2}\right) \frac{f_i(X_i) s_\nu(Z_j)}{f_j(X_j) s_\nu(Z_i)} \gamma(U_Y)\right].$$

It is seen that in the case of the Gaussian copula, $\mathcal{C}_{i,j}$ is directly proportional to r_{ij}. However, for t copulas the dependence effects are more complex, reflecting the tail dependence of the t copula model. We illustrate the calculation of $\mathcal{C}_{i,j}$ using the t copula in Section 5.

Remark. Note that the expressions for $\mathcal{C}_{i,j}$ in Proposition 4.1 require explicit knowledge of only the bivariate copula of (X_i, X_j), rather than the entire copula of $\boldsymbol{X}$. This is true also beyond the Gaussian and t models, as seen in the first few lines of the proof of Proposition 4.1.

4.2 Alternative representation of cascade sensitivity measures

In the case when only the aggregate cascade sensitivity is of interest and the additional insight of the decomposition $\mathcal{C}_i = \sum_{j=1}^n \mathcal{C}_{i,j}$ is not sought, an alternative representation of $\mathcal{C}_i$ may be used, which does not require explicit knowledge of (the gradient of) g and can be implemented efficiently on a single set of Monte Carlo samples. This representation of the cascade sensitivity is valid for both additive shocks (factor prioritisation setting) and perturbations (model uncertainty setting) discussed in Section 2.2 and implemented in Examples 4 and 5, respectively. In the former case, additional assumptions on the shock apply, in practice restricting the choice of shocks to tail stresses that are variations of

$$X_{i,\varepsilon} = X_i + \varepsilon(X_i - t_1)\mathbb{1}_{\{X_i \le t_1\}} + \varepsilon(X_i - t_2)\mathbb{1}_{\{X_i \ge t_2\}}$$

for $t_1 \le t_2 \in \mathbb{R}$ suitably chosen such that $f_i(\cdot)$ is non-decreasing on $(-\infty, t_1]$ and non-increasing on $[t_2, \infty)$.

Proposition 4.2. Let Assumptions A.1 in the appendix be satisfied, $(\boldsymbol{\psi}, \boldsymbol{V}) \in \mathcal{R}_i$ and a stress defined as either of the two cases:

1. An additive shock $X_{i,\varepsilon} = X_i + \varepsilon k(X_i)$, for a non-decreasing Lipschitz continuous function $k\colon \mathbb{R} \to \mathbb{R}$ with Lipschitz constant L, that satisfies $k(x) \le 0$ on the set where $f_i(x)$ is non-decreasing and $k(x) \ge 0$ on the set where $f_i(x)$ is non-increasing. Further define the distribution function $\hat{F}_i(x) = F_i(x) - \frac{k(x)}{L} f_i(x),\ x \in \mathbb{R}$.
2. A perturbation $X_{i,\varepsilon} = F_{i,\varepsilon}^{-1}(U_{X_i})$, where $F_{i,\varepsilon} = (1-\varepsilon)F_i + \varepsilon \hat{F}_i$, for a continuous distribution function $\hat{F}_i \in \mathcal{F}$, and set $L = 1$.

Then, the cascade sensitivity to input factor X_i has representations

$$\begin{aligned} \mathcal{C}_i(\boldsymbol{X}, g, \rho_\gamma) &= L \cdot E\Big[\frac{F_i(X_i) - \hat{F}_i(X_i)}{f_i(X_i)} (g \circ \boldsymbol{\psi})_1(X_i, \boldsymbol{V}) \gamma(U_Y)\Big] \\ &= L \cdot E\Big[\frac{H(Y) - \hat{H}(Y)}{h(Y)} \gamma(H(Y))\Big], \end{aligned} \tag{6}$$

where $\hat{H}$ denotes the distribution function of $\hat{Y} = g(\boldsymbol{\psi}(\hat{X}_i, \boldsymbol{V}))$, with $\hat{X}_i = \hat{F}_i^{-1}(U_{X_i})$.

While in equation (6) the expectation is evaluated only over Y, the calculation depends on the inverse Rosenblatt transform through the presence of $\boldsymbol{\psi}$ in $\hat{Y}$ and, thereby, in $\hat{H}$. We provide representation (6) of the cascade sensitivity for the two most common distortion risk measures used in practice, VaR and ES.

Corollary 4.3. Let Assumptions of Proposition 4.2 be fulfilled and, for simplicity, set $L = 1$. Then, the cascade sensitivities for the VaR and ES risk measures are

$$\mathcal{C}_i(\boldsymbol{X}, g, \mathrm{VaR}_\alpha) = \frac{\alpha - \hat{H}(H^{-1}(\alpha))}{h(H^{-1}(\alpha))},$$
$$\mathcal{C}_i(\boldsymbol{X}, g, \mathrm{ES}_\alpha) = \frac{1}{1-\alpha}\Big[E\Big(\big(\hat{Y} - H^{-1}(\alpha)\big)_+\Big) - E\Big(\big(Y - H^{-1}(\alpha)\big)_+\Big)\Big].$$

In representation (6) of the cascade sensitivity, the requirement for knowledge of the gradient of the aggregation function has been replaced with the need to evaluate $\hat{H}$, arising as a distorted distribution of the output, after substituting input X_i with $\hat{X}_i = \hat{F}_i^{-1}(U_{X_i})$. This can in itself be seen as a different kind of sensitivity test, in particular if $\hat{F}_i$ is more dispersed than F_i. For example, the formula for the sensitivity of ES in Corollary 4.3 involves the difference between two expectations over the right tail of the output, measuring the impact on the output of substituting X_i with its distorted version $\hat{X}_i$. In the case of additive shocks, $X_{i,\varepsilon} = X_i + \varepsilon k(X_i)$, the cascade sensitivity corresponds to comparing the output Y with the distorted output $\hat{Y} = g\big(\boldsymbol{\psi}(\hat{F}_i^{-1}(U_{X_i}), \boldsymbol{V})\big)$, where $\hat{F}_i(x) = F_i(x) - \frac{k(x)}{L} f_i(x),\ x \in \mathbb{R}$. Hence, a stochastic comparison of F_i and $\hat{F}_i$ might be of interest; indeed we show in Proposition A.4 in Appendix A.3, that the distribution $\hat{F}_i$ dominates F_i in increasing convex order.

If evaluation of the aggregation function g is computationally expensive, direct calculation of $\mathcal{C}_i$ via (6) can still be problematic. For example, in a Monte Carlo simulation setting with sample size M, the calculation of the cascade sensitivity to one input factor requires an inverse Rosenblatt transform and M evaluations of $\hat{Y} = g\big(\boldsymbol{\psi}(\hat{X}_i, \boldsymbol{V})\big)$. However, given the independence of X_i and $\boldsymbol{V}$, the distribution function of $\hat{Y}$ can be written as

$$\hat{H}(y) = E\Big(\mathbb{1}_{\{Y \leq y\}} \frac{\hat{f}_i(X_i)}{f_i(X_i)}\Big), \quad y \in \mathbb{R}.$$

Hence $\hat{H}$ can be computed on the same Monte Carlo sample without the need to explicitly calculate an inverse Rosenblatt transform. The ratio $\frac{\hat{f}_i(X_i)}{f_i(X_i)}$ can here be viewed as *importance weights*. Note that, in this context, importance sampling is not used as a variance reduction technique but as a way of re-weighting scenarios to reflect the cascading effect.

We illustrate this procedure for ES and a stress as in Proposition 4.2 (with $L = 1$):

1. Sample M multivariate scenarios $\boldsymbol{x}^{(1)} = \big(x_1^{(1)}, \ldots, x_n^{(1)}\big), \ldots, \boldsymbol{x}^{(M)} = \big(x_1^{(M)}, \ldots, x_n^{(M)}\big)$ from input vector $\boldsymbol{X}$ and calculate the corresponding realisations of the output $y^{(1)} =$

$g\left(\boldsymbol{x}^{(1)}\right), \ldots, y^{(M)} = g\left(\boldsymbol{x}^{(M)}\right)$.

2. Denote by v_α an estimate of $H^{-1}(\alpha)$, and estimate the cascade sensitivity for ES by

$$\mathcal{C}_i^{emp}(\mathrm{ES}_\alpha) = \frac{1}{M(1-\alpha)} \sum_{k=1}^{M} \left(\frac{\hat{f}_i(x_i^{(k)})}{f_i(x_i^{(k)})} - 1 \right) \left(y^{(k)} - v_\alpha\right)_+.$$

Estimators of the cascade sensitivity for a distortion risk measure different from ES can be obtained by utilizing its weight function γ and Proposition 4.2.

Remark. From Proposition A.3 it follows that, if X_i is independent of $\boldsymbol{X}_{-i}$, then $\mathcal{C}_i(\boldsymbol{X}, g, \rho_\gamma) = \mathcal{S}_i(\boldsymbol{X}, g, \rho_\gamma)$. Hence the methods developed in this section allow, in the case of independence, for an alternative evaluation of *marginal sensitivities* of the type studied by Hong (2009), without knowledge of the gradient of the function g. Specifically, in the case that X_i is independent of $\boldsymbol{X}_{-i}$, we have that $\boldsymbol{\psi}(\hat{X}_i, \boldsymbol{V}) = (X_1, \ldots, \hat{X}_i, \ldots, X_n)$, such that $\hat{Y} = g\big(\boldsymbol{\psi}(\hat{X}_i, \boldsymbol{V})\big)$ is the output after substituting $\hat{X}_i$ for X_i in $g(\boldsymbol{X})$. In that sense, the requirement for estimation of the gradient is replaced by the need to re-evaluate the model under a change of only the marginal distribution for X_i.

5 Application to a London Market portfolio

5.1 Aim of the case study

In this Section, we calculate cascade sensitivities for the London Insurance Market portfolio introduced in Section 1. Through this application, we aim to discuss sensitivity analysis in two related but distinct scenarios:

(a) The model has a large number of input factors, on potentially different scales. A possible criticism of any sensitivity measure defined via partial derivatives is that, for input factors on different scales, it is difficult to draw conclusions regarding their relative importance; in particular, such sensitivity measures are generally not invariant under monotone transformations of input factors (Antoniano-Villalobos et al., 2018). Hence, the consistency of stresses applied to different inputs is a point of interest.

(b) The model is a *black box*, that is, the model specification (joint distribution of input factors and aggregation function) is not available in full to an analyst, who is typically only presented with a table of simulated scenarios from input factors and corresponding outputs.

To address scenario (a) above, we propose applying a monotone transform to all input factors, such that the transformed vector of inputs has identical marginal distributions. We then proceed by stressing these transformed risk factors, in a consistent way. This can be seen as an adaptation of the cascade sensitivity measure, making it invariant to monotone transforms of the input factor being stressed. Regarding scenario (b), for a black box model, the cascade sensitivity $\mathcal{C}_i$ can be calculated without having access to the explicit form of g, as argued in Section 4.2. If in addition the form of g is known (or estimated), the decomposition of the cascade sensitivities $\mathcal{C}_{i,j}$ can be calculated following Section 4.1, subject to estimation of the joint distribution of (X_i, X_j) – a path we follow in the sequel.

5.2 Model features

The model we work with represents a London Insurance Market portfolio with 72 input factors. We do not have access to the marginal nor the joint distribution of the input vector in explicit form (the elements of $\boldsymbol{X}$ are themselves outputs of different sub-models). We work with a Monte Carlo sample of size $M = 500,000$, provided by the model owner, consisting of simulated observations from the input vector $\boldsymbol{X}$ and its corresponding output Y. In addition, we are given that the form of the aggregation function g is linear:

$$g(\boldsymbol{X}) = \sum_{j=1}^{72} w_j X_j, \quad \text{for some } w_1, \dots, w_{72} > 0.$$

An analysis of the simulated data shows that all input factors X_i are non-negative and right-skewed. As we do not have an explicit form of f_i, the density of each X_i is estimated using local linear likelihood methods, with the *locfit* package in **R**, see Loader (2006) and Loader (2013). We show densities of the first four input factors in Figure 3 (left), with the vertical axis on the log-scale, illustrating the variation in tail behaviour.

Furthermore, the input factors are positively correlated, with an average Kendall's τ of 0.32. A scatter plot of U_{X_2} against U_{X_1} is shown in Figure 3 (right). The dependence pattern (also observed between other pairs) matches that of a t copula (e.g. McNeil et al., 2015). We estimate the copula of $\boldsymbol{X}$ as a multivariate t copula, using the **R** package *copula* (Hofert et al., 2017). In particular, the estimated degrees of freedom are $\nu = 3.97$, providing evidence of substantial tail dependence.

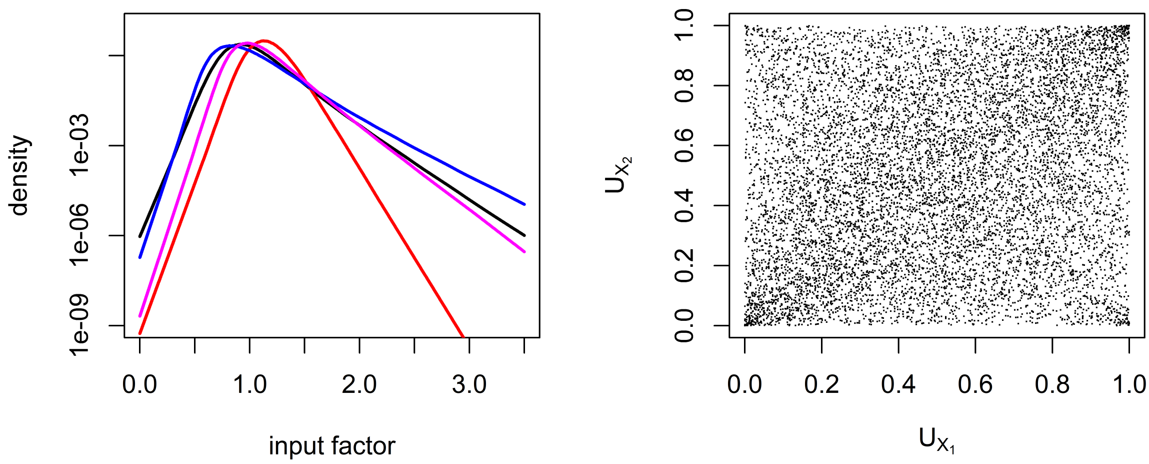

Figure 3: Left: estimated densities of input factors $X_1, \ldots, X_4$. Right: scatter plot of U_{X_2} against U_{X_1}.

5.3 Calculation of $\mathcal{C}_i$

To find a consistent way of stressing different input factors, we (monotonically) transform each $X_i,\ i = 1, \ldots, 72$, to the same marginal distribution. In this example, we define

$$Z_i = t_\nu^{-1}(F_i(X_i)),\ i = 1, \ldots, 72, \tag{7}$$

where t_ν is the Student t distribution function with ν degrees of freedom, such that $Z_i \sim t_\nu$ is comonotonic to X_i. Representation (7) allows to define a stress on the input X_i indirectly through distorting the corresponding variable Z_i. Thus, applying the same stress $Z_{i,\varepsilon}$ to all t_ν-distributed variables Z_i, results in a consistent and comparable stress for all input factors. Here, we let $Z_{i,\varepsilon} = Z_i(1 + \varepsilon)$, which can be interpreted as a direct stress on the volatility of Z_i. This choice of stress yields:

$$X_{i,\varepsilon} = F_i^{-1}(t_\nu(Z_i(1 + \varepsilon))).$$

The motivation for using the t distribution in conjunction with a proportional stress is fourfold: (a) as a location-scale family, proportional stresses indicate themselves as natural stresses on volatility; (b) the stressed variables $Z_{i,\varepsilon}$ have support within that of Z_i [2], allowing

[2] In (7), one can choose other transformations $W_i = F_{W_i}^{-1}(F_i(X_i))$. But care should be taken then that each

evaluation of $t_\nu(Z_i(1+\varepsilon))$; (c) the heavy tails of the t distribution ensure that extremal behaviour will be reflected; (d) for consistency with multivariate t copula used.

We evaluate the cascade sensitivity with respect to ES at level $\alpha = 90\%$. Note that the choice of stress and marginals for Z_i means that the assumptions of Proposition 4.2 are satisfied. Thus, following Section 4.2, we evaluate

$$\mathcal{C}_i(\boldsymbol{X}, g, \mathrm{ES}_\alpha) = \frac{1}{1-\alpha} E\left[\left(Y - H^{-1}(\alpha)\right)_+ \left(-Z_i \frac{s'_\nu(Z_i)}{s_\nu(Z_i)} - 1\right)\right]. \tag{8}$$

The Monte Carlo estimation of the cascade sensitivity via (8) makes no use of the given form of g, hence is applicable to black box models, and thus addresses scenario (b).

In Figure 4 the importance ranking of all 72 risk factors is displayed, according to the cascade sensitivity measure, along with 90% confidence intervals representing sampling error, calculated from a bootstrap sample of size 500. While all input factors display substantial sensitivity (none is close to zero), the ranking is apparent. Specifically, it is clear that the portfolio loss Y is significantly more sensitive to input factors 42 and 27, compared to other factors.

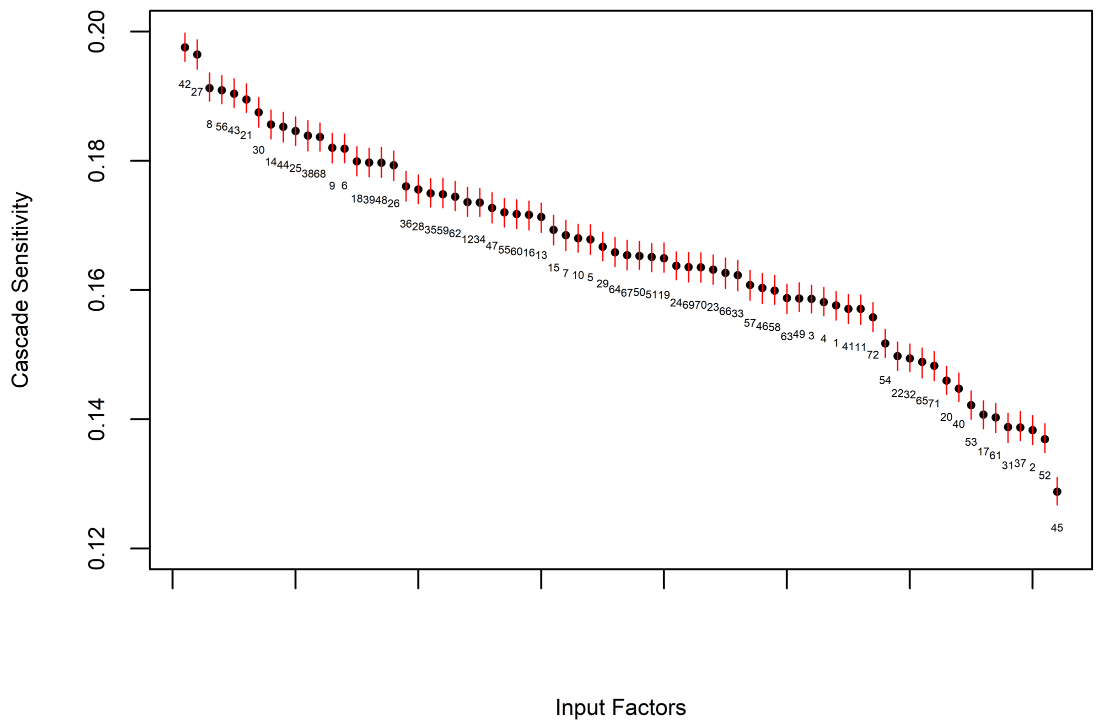

Figure 4: Importance ranking of cascade sensitivities of all 72 risk factors, with respect to a proportional stress on the t-transformed input factors. The red lines represent 90% confidence intervals.

stressed variable $W_{i,\epsilon}$ has support contained within that of W_i, to enable evaluation of $F_{W_i}(W_{i,\varepsilon})$.

5.4 Calculation of $\mathcal{C}_{i,j}$

Here we calculate the sensitivities $\mathcal{C}_{i,j}$, representing the contribution of the j^{th} input factor to the cascade sensitivity of the i^{th} input factor. The derivative of the stress applied on X_i is

$$\frac{\partial X_{i,\varepsilon}}{\partial \varepsilon}\bigg|_{\varepsilon=0} = \frac{\partial}{\partial \varepsilon} F_i^{-1}(t_\nu(Z_i(1+\varepsilon)))\bigg|_{\varepsilon=0} = \frac{s_\nu(Z_i)Z_i}{f_i(X_i)}.$$

Given the assumption of a bivariate t copula between (X_i, X_j) and the linearity of g, we obtain, using Proposition 4.1, case 2:

$$\mathcal{C}_{i,j} = \frac{1}{1-\alpha} w_j E\left[Z_i \left(r_{ij} + \frac{Z_i Z_j - r_{ij} Z_i^2}{\nu + Z_i^2} \right) \frac{s_\nu(Z_j)}{f_j(X_j)} \mathbb{1}_{\{Y > H^{-1}(\alpha)\}} \right].$$

The calculated values of $\mathcal{C}_{i,j}$ are represented in the heatmap of Figure 5. The i^{th} row displays the decomposition of $\mathcal{C}_i$, while the j^{th} column indicates the contributions of the j^{th} input factor to the cascade sensitivities of other factors. Marginal sensitivities are contained along the diagonal, which can be seen to be larger than the corresponding indirect effects $\mathcal{C}_{i,j}$, $i \neq j$, in the same row and column.

The vertical stripes indicate that some input factors (namely X_{42}, X_{43}, X_{55}, X_{68}) are making consistently high contributions to the cascade sensitivities of other risk factors. Comparing to Figure 4, we see that, while input factor X_{42} also has a very high cascade sensitivity, this is not the case for, say, X_{55}. This observation illustrates the additional information conveyed by the decomposition $\mathcal{C}_i = \sum_{j=1}^n \mathcal{C}_{i,j}$. In particular, as seen from (5), $\mathcal{C}_{i,j}$ depends on the partial derivative of g in the direction of X_j (which in this example is constant and equal to the weight w_j). Thus the columns capture the local effects on g of individual variables, while the row sums (cascade sensitivities $\mathcal{C}_i$) emphasize the impact of dependence.

6 Conclusion

We introduce a novel sensitivity measure, termed *cascade sensitivity*, which is defined as a directional derivative of a risk measure applied to the model output, in the direction of a stressed input factor. The derivative is taken after a suitable (inverse Rosenblatt) transformation, which results in capturing, not only the direct impact of the stressed input factor on the output, but also the indirect effects arising via dependence with other input factors. Through examples, we illustrate that the dependence between input factors may substantially contribute to the

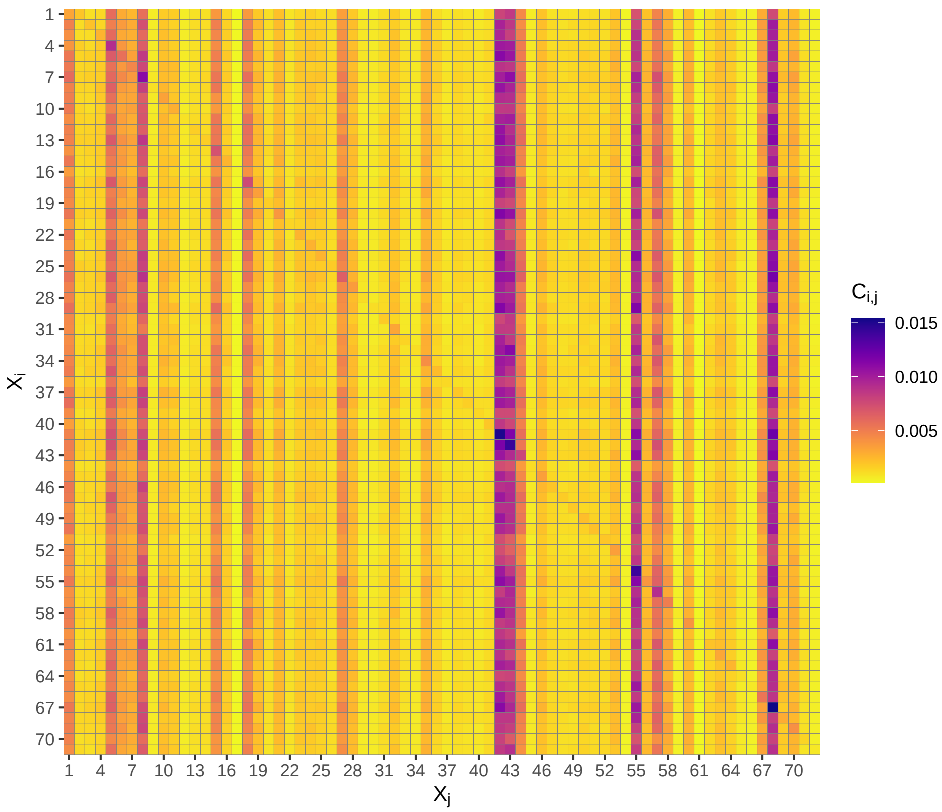

Figure 5: Heatmap of the decomposition of the cascade sensitivities $\mathcal{C}_{i,j}$, $i, j = 1, \ldots, n$. Rows correspond to input factors that are stressed, while columns correspond to the (indirect) contributions of other inputs factors.

cascade sensitivity of a particular input.

We show that the cascade sensitivity decomposes into components, with each component reflecting direct or indirect contributions of an input factor to the sensitivity of the output. If the decomposition is not sought, the cascade sensitivity admits a representation that allows for a straightforward calculation using a single Monte Carlo sample and does not require the knowledge of the gradient of the aggregation function. These representations of the cascade sensitivity make implementation of the proposed sensitivity measure numerically efficient and, thus, attractive for applications in the practice of risk analysis, as is demonstrated through a model of a London Insurance Market portfolio used in industry.

Acknowledgement

S. Pesenti would like to acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) with funding reference numbers DGECR-2020-00333 and RGPIN-2020-04289.

A Assumptions and technical details

A.1 Assumptions

Assumption A.1. Let ρ_γ be a distortion risk measure and $X_{i,\varepsilon}$, $\varepsilon > 0$, a stress on X_i. With abuse of notation, we denote the stressed input vector by $\boldsymbol{X}_{i,\varepsilon} = (X_1, \ldots, X_{i,\varepsilon}, \ldots, X_n)$ for the marginal sensitivity and $\boldsymbol{X}_{i,\varepsilon} = \boldsymbol{\psi}(X_{i,\varepsilon}, \boldsymbol{V})$, $(\boldsymbol{\psi}, \boldsymbol{V}) \in \mathcal{R}_i$, for the cascade sensitivity, respectively. We write $Y_{i,\varepsilon} = g(\boldsymbol{X}_{i,\varepsilon})$ for the stressed output and denote its distribution function by $H_{i,\varepsilon}$.

i) There exists a random variable $W \geq 0$ with $E(W) < +\infty$, such that $|Y_{i,\varepsilon_2} - Y_{i,\varepsilon_1}| \leq W|\varepsilon_2 - \varepsilon_1|$ for all $\varepsilon_1, \varepsilon_2$ in a neighbourhood of 0.

ii) The aggregation function g is invertible in at least one argument, say the j^{th}, and X_j has a conditional density given $\boldsymbol{X}_{-j}$.

Moreover, for all ε in a neighbourhood of 0:

i) The derivative of $Y_{i,\varepsilon}$ with respect to ε exists with probability 1.

ii) $Y_{i,\varepsilon}$ has a continuous density.

iii) $\frac{\partial}{\partial \varepsilon} H_{i,\varepsilon}$ exists and is continuous (in both arguments).

iv) The function $E\big(\frac{\partial}{\partial \varepsilon} Y_{i,\varepsilon} \,|\, Y_{i,\varepsilon} = y\big)$ is continuous in y.

v) $P(H_{i,\varepsilon}(Y_{i,\varepsilon}) \in D_\gamma) = 1$, where D_γ is the set where the weight function γ is differentiable.

vi) $\frac{\partial}{\partial \varepsilon} H_{i,\varepsilon}^{-1}$ exists and is bounded.

A.2 Robustness of stressed models

Here we show, that for a suitable risk measure and aggregation function, the stresses defined in Section 2.2 guarantee that the marginal and cascade sensitivities are well-defined.

Proposition A.2. Given a distortion risk measure ρ_γ, with γ non-decreasing, an aggregation function g satisfying a linear growth condition, i.e. $|g(\boldsymbol{x})| \leq a + b|\boldsymbol{x}|$, for all $\boldsymbol{x} > c$, with $a, b, c > 0$ (Pesenti et al., 2016). Further, consider either

1. a shock $X_{i,\varepsilon} = X_i + \varepsilon k(X_i)$, for a non-decreasing Lipschitz continuous function k, or

2. a perturbation $X_{i,\varepsilon} = (1-\varepsilon)F_i^{-1}(U_{X_i}) + \varepsilon\hat{F}_i^{-1}(U_{X_i})$, with $\hat{F}_i \in \mathcal{F}$, where $\mathcal{F}$ is a set of distribution functions with finite absolute mean.

Then, the composition $\rho_\gamma \circ g$ is robust, that is the functional $\rho_\gamma \circ g$ is continuous with respect to weak convergence. Moreover, the marginal and cascade sensitivities are well-defined.

Proof. Case 1: Denote by L the Lipschitz constant of k. By Lipschitz continuity of k and since $X_i \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$, it holds that the stressed input $X_{i,\varepsilon} \in \mathcal{L}^1(\Omega, \mathcal{A}, P)$. Moreover, for all $\varepsilon \geq 0$, we have a.s.

$$|X_{i,\varepsilon}| \leq |X_i| + \varepsilon|k(X_i) - k(0)| + \varepsilon|k(0)| \leq (1+\varepsilon L)|X_i| + \varepsilon|k(0)|.$$

Thus, for $0 \leq \varepsilon \leq \bar{\varepsilon}$, it holds a.s. $|X_{i,\varepsilon}| \leq (1+\bar{\varepsilon}L)|X_i| + \bar{\varepsilon}|k(0)|$ and the set $\{X_{i,\varepsilon} \mid X_{i,\varepsilon} = X_i + \varepsilon k(X_i),\ 0 \leq \varepsilon \leq \bar{\varepsilon}\}$ is uniformly integrable (by dominated convergence). Applying Theorem 4.8 by Pesenti et al. (2016), we obtain that the composition $\rho_\gamma \circ g$ is robust, that is $\rho_\gamma \circ g$ is continuous with respect to weak convergence.

Case 2: Note that $\hat{F}_i \in \mathcal{F}$ is a distribution function satisfying $\int |x| \mathrm{d}F(x) < \infty$. Using an argument similar to that in case 1, we have a.s.

$$|X_{i,\varepsilon}| \leq (1-\varepsilon)|X_i| + \varepsilon|\hat{X}_i|.$$

Thus, for all $0 \leq \varepsilon \leq \bar{\varepsilon}$, it holds a.s. $|X_{i,\varepsilon}| \leq (1-\bar{\varepsilon})|X_i| + \bar{\varepsilon}|\hat{X}_i|$ and the set $\{X_{i,\varepsilon} \mid X_{i,\varepsilon} = (1-\varepsilon)F_i^{-1}(U_{X_i}) + \varepsilon\hat{F}_i^{-1}(U_{X_i}),\ 0 \leq \varepsilon \leq \bar{\varepsilon}\}$ is uniformly integrable. Applying Theorem 4.8 by Pesenti et al. (2016), we obtain that the composition $\rho_\gamma \circ g$ is robust. □

A.3 Stochastic orders

A random vector $\boldsymbol{W}$ is said to be *conditionally increasing in sequence* (CIS) if, for all $j = 2, \ldots, n$, $E(l(W_j) \mid W_1 = w_1, \ldots, W_{j-1} = w_{j-1})$ is a non-decreasing function of $w_1, \ldots, w_{j-1}$, for all non-decreasing functions $l\colon \mathbb{R} \to \mathbb{R}$ for which the expectation exists (Müller and Stoyan, 2002).

Proposition A.3. Given a stress $X_{i,\varepsilon}$, $(\boldsymbol{\psi}, \boldsymbol{V}) \in \mathcal{R}_i$, and under Assumptions A.1, the following hold:

1. If X_i is independent of X_j, for $i \neq j$, then $\mathcal{C}_{i,j} = 0$. Hence, if X_i is independent of $\boldsymbol{X}_{-i}$, then $\mathcal{C}_i(\boldsymbol{X}, g, \rho_\gamma) = \mathcal{S}_i(\boldsymbol{X}, g, \rho_\gamma)$.

2. If the vector $(X_i, X_{\pi(1)}, \ldots, X_{\pi(n)})$ is CIS for a permutation π on $\{1, \ldots, n\} \backslash \{i\}$, the aggregation function is component-wise non-decreasing, and $\frac{\partial}{\partial \varepsilon} X_{i,\varepsilon}\big|_{\varepsilon=0} \geq 0$ a.s., then $\mathcal{C}_i(\boldsymbol{X}, g, \rho_\gamma) \geq \mathcal{S}_i(\boldsymbol{X}, g, \rho_\gamma)$.

Proof. Case 1: consider a permutation of $\{1, \ldots, n\}/\{i\}$ such that j appears in first position. An inverse Rosenblatt transform then satisfies $\psi^{(j)}(X_i, \boldsymbol{V}) = F_j^{-1}(V_j)$, thus $\psi^{(j)}$ does not depend on X_i and we obtain $\mathcal{C}_{i,j} = 0$. The conclusion follows from the uniqueness of the cascade sensitivity decomposition, see Proposition 3.6.

Case 2: let π be a permutation on $\{1, \ldots, n\}/\{i\}$. Then $(X_i, X_{\pi(1)}, \ldots, X_{\pi(n)})$ being CIS implies that the conditional distributions $F_{\pi(j)|i,\pi(1),\ldots,\pi(j-1)}(\cdot \mid X_i = x_i, X_{\pi(1)} = x_{\pi(1)}, \ldots, X_{\pi(j-1)} = x_{\pi(j-1)})$ are non-increasing in x_i (Müller and Stoyan, 2002). Therefore the quantile functions $F^{-1}_{\pi(j)|i,\pi(1),\ldots,\pi(j-1)}(\cdot \mid X_i = x_i, X_{\pi(1)} = x_{\pi(1)}, \ldots, X_{\pi(j-1)} = x_{\pi(j-1)})$ are non-decreasing in x_i and $\psi^{(j)}(X_i, \boldsymbol{V})$, $1 \leq j \leq n$, are non-decreasing functions of X_i and thus $\psi_1^{(j)}(X_i, \boldsymbol{V}) \geq 0$ for $1 \leq j \leq n$. The additional assumptions guarantee that all summands of the formula of the cascade sensitivity in Proposition 3.5 are non-negative. □

Examples of stresses with non-negative gradient include additive shocks $X_{i,\varepsilon} = X_i + \varepsilon k(X_i)$ as defined in Section 2.2. Examples of perturbations are $X_{i,\varepsilon} = F_{i,\varepsilon}^{-1}(U_{X_i})$ with $F_{i,\varepsilon} = (1 - \varepsilon)F_i + \varepsilon \hat{F}_i$, whenever the distribution $\hat{F}_i$ first order stochastically dominates F_i.

Note that, by Proposition 3.6, it is enough in Proposition A.3, case 2, that the vector $(X_i, X_{\pi(1)}, \ldots, X_{\pi(n)})$ is CIS for one permutation π. Examples of vectors that are CIS, which is a dependence concept of the copula alone (Müller and Scarsini, 2001, Prop. 3.5), include the multivariate normal distribution whose inverse covariance matrix contains non-positive off-diagonal elements, as well as the multivariate logistic, gamma, and negative binomial distributions (Müller and Scarsini, 2001; Karlin and Rinott, 1980). We also refer to Karlin and Rinott (1980) for further examples of *multivariate totally positive of order 2* distributions, a slightly stronger dependence concept than CIS.

We recall that a random variable W dominates Z in *increasing convex order*, $Z \preceq_{icx} W$, if $E(l(Z)) \leq E(l(W))$, for all increasing convex functions $l \colon \mathbb{R} \to \mathbb{R}$ such that the expectations exist (Müller and Stoyan, 2002).

Proposition A.4. Let X_i have finite expectation and define the random variable $\hat{X}_i$ with distribution function $\hat{F}_i(x) = F_i(x) - \frac{k(x)}{L} f_i(x),\ x \in \mathbb{R}$, as in Proposition 4.2. Then the following hold:

1. If $E(k(X_i)) \geq 0$, then $X_i \preceq_{icx} \hat{X}_i$.
2. If $0 < \operatorname{ess\,sup} k(X_i)$, then X_i does not dominate $\hat{X}_i$ in increasing convex order.

Proof. Applying Lemma B.1 we see that $\hat{F}_i$ is a distribution function. Note that $\hat{X}_i$ dominates X_i in increasing convex order, $X_i \preceq_{icx} \hat{X}_i$, if and only if $E((\hat{X}_i - t)_+) \geq E((X_i - t)_+)$ for all $t \in \mathbb{R}$. For case 1: let $t \in \mathbb{R}$ and apply Fubini,

$$
\begin{aligned}
E\big((\hat{X}_i - t)_+\big) - E\big((X_i - t)_+\big) &= \int_t^\infty (x - t)\big(\hat{f}_i(x) - f_i(x)\big)\mathrm{d}x \\
&= \int_t^\infty \int_t^\infty \big(\hat{f}_i(x) - f_i(x)\big)\mathbb{1}_{\{u \leq x\}}\mathrm{d}x\mathrm{d}u \\
&= \int_t^\infty \big(F_i(u) - \hat{F}_i(u)\big)\mathrm{d}u \\
&= \int_t^\infty \tfrac{k(u) f_i(u)}{L}\mathrm{d}u \\
&= \tfrac{1}{L}\, E(k(X_i)\mathbb{1}_{\{X_i > t\}}).
\end{aligned}
$$

Recall that k is a non-decreasing function, thus if $k(t) \geq 0$, the above expectation is non-negative. If $k(t) < 0$, we have that $E(k(X_i)\mathbb{1}_{\{X_i > t\}}) \geq E(k(X_i))$, which is non-negative by assumption.

To see case 2: note that $E(k(X_i)\mathbb{1}_{\{X_i > t\}})$ is negative if and only if

$$
E(k(X_i) \mid X_i > t) < 0 \text{ for all } t \in \mathbb{R},
$$

which is a contradiction to the assumption that $0 < \operatorname{ess\,sup} k(X_i)$. $\square$

Proposition A.4 case 1 is for example satisfied for a one-sided stress of an input factor, that is $X_{i,\varepsilon} = X_i + \varepsilon(X_i - t)_+,\ t > 0$. For an input factor that is symmetric around zero, consider the stress $X_{i,\varepsilon} = X_i + \varepsilon(X_i - t_1)\mathbb{1}_{\{X_i \leq t_1\}} + \varepsilon(X_i - t_2)\mathbb{1}_{\{X_i \geq t_2\}}$, for $t_1 < 0 < t_2$, such that the density of the input is non-decreasing for $x \leq t_1$ and non-increasing for $x \geq t_2$. Then, Proposition A.4 case 1, is fulfilled if $t_2 < |t_1|$.

B Proofs

B.1 Auxiliary results

Lemma B.1. Let Z be an integrable random variable with distribution function F_Z and right-continuous density f_Z whose support can be split into countably many intervals on which f_Z is monotonic. Let $k\colon \mathbb{R} \to \mathbb{R}$ a non-decreasing Lipschitz continuous function with Lipschitz constant $L > 0$, which satisfies $k(x) \leq 0$ on the set where $f_Z(x)$ is non-decreasing and $k(x) \geq 0$ on the set where $f_Z(x)$ is non-increasing. Then $\tilde{F}_Z(x) = F_Z(x) - \frac{k(x)}{L} f_Z(x),\ x \in \mathbb{R}$, defines a distribution function.

Proof. By Lipschitz continuity of k, it holds $-(k(y) - k(x)) \geq -L(y - x)$, for all $x < y$. Let $a, b \in \mathbb{R}$ such that f_Z is non-decreasing on $[a, b]$, then it holds for all $a \leq x < y \leq b$ that

$$
\begin{aligned}
\tilde{F}_Z(y) - \tilde{F}_Z(x) &= \int_x^y f_Z(u)\mathrm{d}u - \tfrac{k(y)}{L} f_Z(y) - \tfrac{k(y)-k(x)}{L} f_Z(x) + \tfrac{k(y)}{L} f_Z(x) \\
&\geq \int_x^y \big(f_Z(u) - f_Z(x)\big)\mathrm{d}u - \tfrac{k(y)}{L}\big(f_Z(y) - f_Z(x)\big) \\
&\geq 0.
\end{aligned}
$$

Similarly, for $a, b \in \mathbb{R}$ such that f_Z is non-increasing on $[a, b]$, we have for all $a \leq x < y \leq b$

$$
\begin{aligned}
\tilde{F}_Z(y) - \tilde{F}_Z(x) &= \int_x^y f_Z(u)\mathrm{d}u - \tfrac{k(x)}{L} f_Z(y) - \tfrac{k(y)-k(x)}{L} f_Z(y) + \tfrac{k(x)}{L} f_Z(x) \\
&\geq \int_x^y \big(f_Z(u) - f_Z(y)\big)\mathrm{d}u + \tfrac{k(x)}{L}\big(f_Z(x) - f_Z(y)\big) \\
&\geq 0.
\end{aligned}
$$

$\square$

Lemma B.2. Let $K\colon \mathbb{R} \to \mathbb{R}$ be an absolutely continuous function with representation $K(x) = \int_b^x \kappa(s)\mathrm{d}s$ for all $x \geq b$, where κ is a non-negative function and $b \geq -\infty$. Then, for any random variable $Z \geq b$ a.s. with $E(K(Z)) < \infty$ it holds that

$$
E(K(Z)) = \int_b^{+\infty} \kappa(s) P(Z > s)\mathrm{d}s.
$$

Proof. For any $b \geq -\infty$ we obtain, using Fubini,

$$E(K(Z)) = E\left(\int_b^Z \kappa(s)\mathrm{d}s\right) = E\left(\int_b^{+\infty} \kappa(s)\mathbb{1}_{\{Z\geq b\}}\mathrm{d}s\right) = \int_{-\infty}^{+\infty} \kappa(s)P(Z > s)\mathrm{d}s.$$

□

B.2 Proofs

Proof of Proposition 3.2. See Hong (2009); Hong and Liu (2009). □

Proof of Proposition 3.5. From Proposition 3.2 it holds that the cascade sensitivity to input X_i can be written as

$$\mathcal{C}_i(\boldsymbol{X}, g, \rho_\gamma) = \sum_{j=1}^n E\Big(\frac{\partial}{\partial\varepsilon} X_{i,\varepsilon}\Big|_{\varepsilon=0} g_j(\boldsymbol{X})\psi_1^{(j)}(X_i, \boldsymbol{V})\gamma(U_Y)\Big).$$

Recall that for $(\boldsymbol{\psi}, \boldsymbol{V}) \in \mathcal{R}_i$ and by construction of the inverse Rosenblatt transform $\psi^{(i)}$ is the identity function, thus $\psi_1^{(i)} = 1$ and $\mathcal{C}_{i,i} = \mathcal{S}_i$. □

Proof of Proposition 3.6. We first show that the cascade sensitivity is independent of the chosen Rosenblatt transform. Consider a stress $X_{i,\varepsilon}$ and $(\boldsymbol{\psi}, \boldsymbol{V}), (\boldsymbol{\phi}, \boldsymbol{U}) \in \mathcal{R}_i$. Note that $\boldsymbol{V}$ and $\boldsymbol{U}$ can be chosen to be independent of the stress $X_{i,\varepsilon}$. For a function $l\colon \mathbb{R}^n \to \mathbb{R}$ such that the following expectation exists, it holds that, for all $\varepsilon > 0$,

$$\begin{aligned}
E\Big((l \circ \boldsymbol{\psi})(X_{i,\varepsilon}, \boldsymbol{V})\Big) &= E\Big((l \circ \boldsymbol{\psi})(X_i, \boldsymbol{V})\frac{f_{X_{i,\varepsilon}}(X_i)}{f_i(X_i)}\Big) \\
&= E\Big((l \circ \boldsymbol{\phi})(X_i, \boldsymbol{U})\frac{f_{X_{i,\varepsilon}}(X_i)}{f_i(X_i)}\Big) \\
&= E\Big((l \circ \boldsymbol{\phi})(X_{i,\varepsilon}, \boldsymbol{U})\Big).
\end{aligned}$$

Thus, for all $\varepsilon > 0$, $\boldsymbol{\psi}(X_{i,\varepsilon}, \boldsymbol{V})$ and $\boldsymbol{\phi}(X_{i,\varepsilon}, \boldsymbol{U})$ follow the same distribution and therefore $\frac{\partial}{\partial\varepsilon}\rho_\gamma\big(g\big(\boldsymbol{\psi}(X_{i,\varepsilon}, \boldsymbol{V})\big)\big)\big|_{\varepsilon=0} = \frac{\partial}{\partial\varepsilon}\rho_\gamma\big(g\big(\boldsymbol{\phi}(X_{i,\varepsilon}, \boldsymbol{U})\big)\big)\big|_{\varepsilon=0}$.

To show that the decomposition of the cascade sensitivity in Proposition 3.5, (4), is unique, note that for $(\boldsymbol{\psi}, \boldsymbol{V}) \in \mathcal{R}_i$ and $j = 1, \ldots n$, we have

$$\mathcal{C}_{i,j} = \frac{\partial}{\partial\varepsilon}\rho_\gamma\big(g\big((X_1, \ldots, X_{j-1}, \psi^{(j)}(X_{i,\varepsilon}, \boldsymbol{V}), X_{j+1}, \ldots, X_n)\big)\big)\Big|_{\varepsilon=0}. \tag{9}$$

Define the random vector $\boldsymbol{X}^* = (X_1^*, \ldots, X_n^*)$, where X_k^* and X_k follow the same distribution, for

$k \in \{1, \dots, n\} \backslash \{j\}$, and the conditional distribution of X_j^* given X_i^* is equal to the conditional distribution of X_j given X_i. Further, let (X_i^*, X_j^*) be independent of $\boldsymbol{X}_{-i,j}^*$, where $\boldsymbol{X}_{-i,j}^*$ is the sub-vector of $\boldsymbol{X}^*$ deprived of its i^{th} and j^{th} components. Then, $\boldsymbol{X}^*$ admits an inverse Rosenblatt transform $(\boldsymbol{\psi}^*, \boldsymbol{V}^*)$ given by:

$$
\begin{aligned}
&\psi^{*(k)}(X_i^*, \boldsymbol{V}^*) = F_k^{-1}(V_k^*), \; k \in \{1, \dots, n\} \backslash \{i, j\}, \\
&\psi^{*(i)}(X_i^*, \boldsymbol{V}^*) = X_i, \\
&\psi^{*(j)}(X_i^*, \boldsymbol{V}^*) = \psi^{(j)}(X_i, \boldsymbol{V}^*).
\end{aligned}
$$

The cascade sensitivity of $\rho_\gamma(g(\boldsymbol{X}^*))$ to X_i^* is thus given by

$$
\mathcal{C}_i(\boldsymbol{X}^*, g, \rho_\gamma) = \mathcal{S}_i(\boldsymbol{X}^*, g, \rho_\gamma) + \mathcal{C}_{i,j},
$$

with $C_{i,j}$ given in (9). □

Proof of Proposition 4.1. Case 1: By invariance of $C_{i,j}$ on the choice of inverse Rosenblatt transform (Proposition 3.6), we can chose a transform starting from $X_i, X_j, \dots$, such that we represent input factors by

$$
\begin{aligned}
X_i &= F_i^{-1}(U_{X_i}) \\
X_j &= F_{j|i}^{-1}(V|X_i) = \psi^{(j)}(X_i, V) \\
&\vdots
\end{aligned}
$$

for some standard uniform random variable V. We further calculate

$$
\dot{F}_{j|i}^{-1}(v|x) := \frac{\partial}{\partial x} F_{j|i}^{-1}(v|x) = \psi_1^{(j)}(x, v).
$$

From the properties of the multivariate normal distribution, it holds that $(Z_j|Z_i = z) \sim N(r_{ij}z, 1 - r_{ij}^2)$. Consequently

$$
\begin{aligned}
F_{j|i}(y|x) &= \Phi\left(\frac{\Phi^{-1}(F_j(y)) - r_{ij}\Phi^{-1}(F_i(x))}{\sqrt{1 - r_{ij}^2}}\right), \\
F_{j|i}^{-1}(v|x) &= F_j^{-1} \circ \Phi\left(r_{ij}\Phi^{-1}(F_i(x)) + \sqrt{1 - r_{ij}^2}\Phi^{-1}(v)\right).
\end{aligned}
$$

Taking derivative with respect to x, we obtain

$$\dot{F}_{j|i}^{-1}(v|x) = \frac{\phi\left(r_{ij}\Phi^{-1}(F_i(x)) + \sqrt{1-r_{ij}^2}\Phi^{-1}(v)\right)}{f_j\left(F_j^{-1}\circ\Phi\left(r_{ij}\Phi^{-1}(F_i(x)) + \sqrt{1-r_{ij}^2}\Phi^{-1}(v)\right)\right)} \cdot \frac{r_{ij}f_i(x)}{\phi(\Phi^{-1}(F_i(x)))}.$$

Replacing $x = X_i$, $y = X_j$, and $v = V$, yields

$$\dot{F}_{j|i}^{-1}(V|X_i) = \frac{\phi\left(r_{ij}\Phi^{-1}(U_{X_i}) + \sqrt{1-r_{ij}^2}\Phi^{-1}(V)\right)}{f_j\left(F_j^{-1}\circ\Phi\left(r_{ij}\Phi^{-1}(U_{X_i}) + \sqrt{1-r_{ij}^2}\Phi^{-1}(V)\right)\right)} \cdot \frac{r_{ij}f_i(X_i)}{\phi(\Phi^{-1}(U_{X_i}))}.$$

Recall that, by definition,

$$V = F_{j|i}(X_j|X_i) = \Phi\left(\frac{\Phi^{-1}(U_{X_j}) - r_{ij}\Phi^{-1}(U_{X_i})}{\sqrt{1-r_{ij}^2}}\right).$$

Thus, we obtain that

$$\dot{F}_{j|i}^{-1}(V|X_i) = \frac{\phi\left(\Phi^{-1}(U_{X_j})\right)}{f_j\left(X_j\right)} \cdot \frac{r_{ij}f_i(X_i)}{\phi(\Phi^{-1}(U_{X_i}))} = r_{ij}\frac{\phi\left(Z_j\right)}{f_j\left(X_j\right)} \cdot \frac{f_i(X_i)}{\phi(Z_i)},$$

wherefrom the stated result follows.

Case 2: For the t copula the same steps are followed, with slightly lengthier calculations, not reported here. The key difference is the form of the conditional t distribution (Ding, 2016),

$$Z_j\big|_{Z_i=z} = r_{ij}z + \sqrt{\tfrac{\nu+z^2}{\nu+1}(1-r_{ij}^2)}\, W, \quad \text{for } z \in \mathbb{R},$$

where W is a t distributed random variable with $\nu+1$ degrees of freedom and the above equality holds in distribution. □

Proof of Proposition 4.2. Case 1 (additive shock): Consider an additive shock $X_{i,\varepsilon} = X_i + \varepsilon k(X_i)$, then by Lemma B.1, $\hat{F}_i(x) = F_i(x) - \frac{k(x)}{L}f_i(x)$, $x \in \mathbb{R}$, is indeed a distribution function and we denote its density by $\hat{f}_i$. We start with proving the first representation of $\mathcal{C}_i$ in Proposition 4.2. We have that the following equalities hold a.s.

$$\frac{\partial}{\partial\varepsilon}X_{i,\varepsilon}\big|_{\varepsilon=0} = k(X_i) = L\,\frac{F_i(X_i) - \hat{F}_i(X_i)}{f_i(X_i)}.$$

Thus, applying Proposition 3.5 results in the first formula for $\mathcal{C}_i$.

Next, we prove the second representation of $\mathcal{C}_i$. Define $\xi(y) = \gamma(H(y))$, $y \in \mathbb{R}$, and notate $(g \circ \boldsymbol{\psi})_1(x_i, \boldsymbol{v}) = \frac{\partial}{\partial x_i} g(\boldsymbol{\psi}(x_i, \boldsymbol{v}))$. Using independence of X_i and $\boldsymbol{V}$ and the definition of $\hat{F}_i$, the cascade sensitivity to stress $X_{i,\varepsilon}$ can be written as

$$
\begin{aligned}
\mathcal{C}_i(\boldsymbol{X}, g, \rho_\gamma) &= E(k(X_i)(g \circ \boldsymbol{\psi})_1(X_i, \boldsymbol{V})\xi(Y)) \\
&= E\Big(\int_{\mathbb{R}} k(x)(g \circ \boldsymbol{\psi})_1(x, \boldsymbol{V})\xi(g(\boldsymbol{\psi}(x, \boldsymbol{V})))f_i(x)\mathrm{d}x\Big) \\
&= L \cdot E\Big(\int_{\mathbb{R}} (F_i(x) - \hat{F}_i(x))\beta(x)\mathrm{d}x\Big),
\end{aligned}
$$

where we denote $\beta(x) = (g \circ \boldsymbol{\psi})_1(x, \boldsymbol{V})\xi(g(\boldsymbol{\psi}(x, \boldsymbol{V})))$ and $B(s) = \int_s^{\operatorname{ess\,sup} X_i} \beta(x)\mathrm{d}x$, hence suppressing the dependence on $\boldsymbol{V}$. Applying Fubini, we obtain

$$
\begin{aligned}
E\Big(\int_{\mathbb{R}} (F_i(x) - \hat{F}_i(x))\beta(x)\mathrm{d}x\Big) &= E\Big(\int_{\mathbb{R}}\int_{\mathbb{R}} \big(f_i(s) - \hat{f}_i(s)\big)\mathbb{1}_{\{s \le x\}}\mathrm{d}s\beta(x)\mathrm{d}x\Big) \\
&= E\Big(\int_{\mathbb{R}} \big(f_i(s) - \hat{f}_i(s)\big)B(s)\mathrm{d}s\Big) \\
&= E(B(X_i)) - E(B(\hat{X}_i)).
\end{aligned}
$$

Applying the change of variable $u = g(\boldsymbol{\psi}(t, \boldsymbol{V}))$, we obtain

$$
\begin{aligned}
B(X_i) &= \int_{X_i}^{\operatorname{ess\,sup} X_i} \beta(t)\mathrm{d}t \\
&= \int_{X_i}^{\operatorname{ess\,sup} X_i} (g \circ \boldsymbol{\psi})_1(t, \boldsymbol{V})\xi(g(\boldsymbol{\psi}(t, \boldsymbol{V})))\mathrm{d}t \\
&= \int_{Y}^{\operatorname{ess\,sup} Y} \xi(u)\mathrm{d}u,
\end{aligned}
$$

and similarly, $B(\hat{X}_i) = \int_{\hat{Y}}^{\operatorname{ess\,sup} \hat{Y}} \xi(u)\mathrm{d}u$. Thus, using Lemma B.2, the cascade sensitivity becomes

$$
\begin{aligned}
\mathcal{C}_i(\boldsymbol{X}, g, \rho_\gamma) &= L\int_{\mathbb{R}} (1 - \hat{H}(y))\xi(y)\mathrm{d}y - L\int_{\mathbb{R}} (1 - H(y))\xi(y)\mathrm{d}y \\
&= L\int_{\mathbb{R}} (H(y) - \hat{H}(y))\gamma(H(y))\mathrm{d}y \\
&= L \cdot E\Big(\frac{H(Y) - \hat{H}(Y)}{h(Y)}\gamma(H(Y))\Big).
\end{aligned}
$$

Case 2 (perturbation): Next we consider a perturbation $X_{i,\varepsilon} = F_{i,\varepsilon}^{-1}(U_{X_i})$, where $F_{i,\varepsilon} = (1 - \varepsilon)F_i + \varepsilon\hat{F}_i$ and show the first formula of $\mathcal{C}_i$ in Proposition 4.2. For all $0 < u < 1$ it holds that (Glasserman, 1991, Thm. 1.3)

$$\frac{\partial}{\partial \varepsilon} F_{i,\varepsilon}^{-1}(u)\big|_{\varepsilon=0} = \frac{u - \hat{F}_i(F_i^{-1}(u))}{f_i(F_i^{-1}(u))}$$

and we have almost surely

$$\frac{\partial}{\partial \varepsilon} X_{i,\varepsilon}\big|_{\varepsilon=0} = \frac{\partial}{\partial \varepsilon} F_{i,\varepsilon}^{-1}(U_{X_i})\big|_{\varepsilon=0} = \frac{F_i(X_i) - \hat{F}_i(X_i)}{f_i(X_i)}. \tag{10}$$

Thus, applying Proposition 3.5 gives the first representation.

To see the second representation, define, for all $\varepsilon \geq 0$, the random variable $\bar{X}_{i,\varepsilon} = X_i \mathbb{1}_A + \hat{X}_i \mathbb{1}_{A^c}$, where $\hat{X}_i = \hat{F}_i^{-1}(U_{X_i})$, $A \in \mathcal{A}$ is independent of $\boldsymbol{X}$ and $\boldsymbol{V}$, with $P(A) = 1 - \varepsilon$ and $A^c = \Omega \backslash A$. Then, $\bar{X}_{i,\varepsilon}$ and the stress $X_{i,\varepsilon}$ follow the same distribution function $F_{i,\varepsilon}$. By independence of $\hat{X}_i$ and $\boldsymbol{V}$, the random vectors $\boldsymbol{\psi}(\bar{X}_{i,\varepsilon}, \boldsymbol{V})$ and $\boldsymbol{\psi}(X_{i,\varepsilon}, \boldsymbol{V})$ are equal in distribution for all $\varepsilon > 0$. Thus, the cascade sensitivity to the stress $X_{i,\varepsilon}$ is equal to the cascade sensitivity to the stress $\bar{X}_{i,\varepsilon}$. To calculate the latter, note that the stressed output, $g(\boldsymbol{\psi}(\bar{X}_{i,\varepsilon}, \boldsymbol{V})) = Y \mathbb{1}_A + g(\boldsymbol{\psi}(\hat{X}_i, \boldsymbol{V}))\mathbb{1}_{A^c}$, follows the mixture distribution $(1-\varepsilon)H + \varepsilon \hat{H}$, where $\hat{H}$ denotes the distribution function of $g(\boldsymbol{\psi}(\hat{X}_i, \boldsymbol{V}))$. The representation of the cascade sensitivity to stress $\bar{X}_{i,\varepsilon}$ follows from a similar argument as in (10). □

Proof of Corollary 4.3. The representation of the cascade sensitivity for the VaR follows from the second representation of Proposition 4.2, noting that VaR_α corresponds to $\gamma(u) = \mathbb{1}_{\{u=\alpha\}}$.

The representation for ES follows by applying Fubini

$$\begin{aligned}
\mathcal{C}_i(\boldsymbol{X}, g, \mathrm{ES}_\alpha) &= E\Big[\frac{H(Y) - \hat{H}(Y)}{h(Y)} \gamma(H(Y))\Big] \\
&= \frac{1}{1-\alpha} \int_{H^{-1}(\alpha)}^{+\infty} (H(y) - \hat{H}(y)) \mathrm{d}y \\
&= \frac{1}{1-\alpha} \int_{H^{-1}(\alpha)}^{+\infty} \int_y^{\infty} \big(\hat{h}(z) - h(z)\big) \mathrm{d}z \mathrm{d}y \\
&= \frac{1}{1-\alpha} \int_{H^{-1}(\alpha)}^{+\infty} (z - H^{-1}(\alpha))\big(\hat{h}(z) - h(z)\big) \mathrm{d}z \\
&= \frac{1}{1-\alpha}\Big[E\Big(\big(\hat{Y} - H^{-1}(\alpha)\big)_+\Big) - E\Big(\big(Y - H^{-1}(\alpha)\big)_+\Big)\Big].
\end{aligned}$$

□

References

Aas, K., Czado, C., Frigessi, A. and Bakken, H. (2009), 'Pair-copula constructions of multiple dependence', *Insurance: Mathematics and Economics* **44**(2), 182–198.

Acerbi, C. and Tasche, D. (2002), 'Expected Shortfall: a natural coherent alternative to Value at Risk', *Economic Notes* **31**(2), 379–388.

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

Arbenz, P., Hummel, C. and Mainik, G. (2012), 'Copula based hierarchical risk aggregation through sample reordering', *Insurance: Mathematics and Economics* **51**(1), 122–133.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, E. (1999), 'Coherent measures of risk', *Mathematical Finance* **9**(3), 203–228.

Bastos, L. S. and O'Hagan, A. (2009), 'Diagnostics for Gaussian Process emulators', *Technometrics* **51**(4), 425–438.

Belles-Sampera, J., Guillén, M. and Santolino, M. (2014), 'Beyond value-at-risk: Gluevar distortion risk measures', *Risk Analysis* **34**(1), 121–134.

Borgonovo, E., Hazen, G. B. and Plischke, E. (2016), 'A common rationale for global sensitivity measures and their estimation', *Risk Analysis* **36**(10), 1871–1895.

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

Cadoni, P. (2014), *Internal models and Solvency II*, Risk Books, London.

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

Cao, X.-R. and Wan, X. (2017), 'Sensitivity analysis of nonlinear behavior with distorted probability', *Mathematical Finance* **27**(1), 115–150.

Choe, Y., Lam, H. and Byon, E. (2018), 'Uncertainty quantification of stochastic simulation for black-box computer experiments', *Methodology and Computing in Applied Probability* pp. 1–18.

Cont, R., Deguest, R. and Scandolo, G. (2010), 'Robustness and sensitivity analysis of risk measurement procedures', *Quantitative Finance* **10**(6), 593–606.

Cucurachi, S., Borgonovo, E. and Heijungs, R. (2016), 'A protocol for the global sensitivity analysis of impact assessment models in life cycle assessment', *Risk Analysis* **36**(2), 357–377.

Dawid, A. P. (1984), 'Present position and potential developments: Some personal views statistical theory the prequential approach', *Journal of the Royal Statistical Society: Series A (General)* **147**(2), 278–290.

Ding, P. (2016), 'On the conditional distribution of the multivariate t distribution', *The American Statistician* **70**(3), 293–295.

Föllmer, H. and Schied, A. (2011), *Stochastic Finance: An introduction in discrete time*, 3rd edn, Walter de Gruyte.

Foulser-Piggott, R., Bowman, G. and Hughes, M. (2020), 'A framework for understanding uncertainty in seismic risk assessment', *Risk analysis* **40**(1), 169–182.

Glasserman, P. (1991), *Gradient estimation via perturbation analysis*, Vol. 116, Springer Science & Business Media.

Glasserman, P. and Xu, X. (2014), 'Robust risk measurement and model risk', *Quantitative Finance* **14**(1), 29–58.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (2011), *Robust statistics: the approach based on influence functions*, Vol. 196, John Wiley & Sons.

Hofert, M., Kojadinovic, I., Maechler, M. and Yan, J. (2017), *Copula: Multivariate Dependence with Copulas*. R package version 0.999-18.
URL: *https://CRAN.R-project.org/package=copula*

Hong, L. J. (2009), 'Estimating quantile sensitivities', *Operations Research* **57**(1), 118–130.

Hong, L. J. and Liu, G. (2009), 'Simulating sensitivities of Conditional Value at Risk', *Management Science* **55**(2), 281–293.

Kalkbrener, M. (2005), *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* **15**(3), 425–437.

Karlin, S. and Rinott, Y. (1980), ‘Classes of orderings of measures and related correlation inequalities. I. multivariate totally positive distributions’, *Journal of Multivariate Analysis* **10**(4), 467–498.

Kraus, D. and Czado, C. (2017), ‘D-vine copula based quantile regression’, *Computational Statistics & Data Analysis* **110**, 1–18.

Kromer, E., Overbeck, L. and Zilch, K. (2016), ‘Systemic risk measures on general measurable spaces’, *Mathematical Methods of Operations Research* **84**(2), 323–357.

Lam, H. (2017), ‘Sensitivity to serial dependency of input processes: A robust approach’, *Management Science* **64**(3), 1311–1327.

Lam, H. and Qian, H. (2018), ‘Subsampling to enhance efficiency in input uncertainty quantification’, *arXiv preprint arXiv:1811.04500* .

Loader, C. (2006), *Local Regression and Likelihood*, Springer Science & Business Media.

Loader, C. (2013), *locfit: Local Regression, Likelihood and Density Estimation.* R package version 1.5-9.1.
URL: *https://4RAN.R-project.org/package=locfit*

Mai, J.-F., Schenk, S. and Scherer, M. (2015), ‘Analyzing model robustness via a distortion of the stochastic root: A Dirichlet prior approach’, *Statistics & Risk Modeling* **32**(3-4), 177–195.

Mara, T. A. and Tarantola, S. (2012), ‘Variance-based sensitivity indices for models with dependent inputs’, *Reliability Engineering & System Safety* **107**, 115–121.

McNeil, A. J., Frey, R. and Embrechts, P. (2015), *Quantitative risk management: concepts, techniques and tools-revised edition*, Princeton University Press.

Mornet, A., Opitz, T., Luzi, M. and Loisel, S. (2015), ‘Index for predicting insurance claims from wind storms with an application in france’, *Risk Analysis* **35**(11), 2029–2056.

Müller, A. and Scarsini, M. (2001), ‘Stochastic comparison of random vectors with a common copula’, *Mathematics of Operations Research* **26**(4), 723–740.

Müller, A. and Stoyan, D. (2002), *Comparison methods for stochastic models and risks*, John Wiley & Sons.

Nagler, T., Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B. and Erhardt, T. (2019), *VineCopula: Statistical Inference of Vine Copulas*. R package version 2.2.0.
URL: *https://CRAN.R-project.org/package=VineCopula*

Oddo, P. C., Lee, B. S., Garner, G. G., Srikrishnan, V., Reed, P. M., Forest, C. E. and Keller, K. (2020), 'Deep uncertainties in sea-level rise and storm surge projections: Implications for coastal flood risk management', *Risk Analysis* **40**(1), 153–168.

Pesenti, S. M., Millossovich, P. and Tsanakas, A. (2016), 'Robustness regions for measures of risk aggregation', *Dependence Modeling* **4**(1).

Rabitti, G. and Borgonovo, E. (2020), 'Is mortality or interest rate the most important risk in annuity models? a comparison of sensitivity analysis methods', *Insurance: Mathematics and Economics* **95**, 48 – 58.

Risk, J. and Ludkovski, M. (2018), 'Sequential design and spatial modeling for portfolio tail risk measurement', *SIAM Journal on Financial Mathematics* **9**(4), 1137–1174.

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

Rubinstein, R. Y. and Melamed, B. (1998), *Modern Simulation and Modeling*, Vol. 7, Wiley New York.

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

Saltelli, A. (2002), 'Sensitivity analysis for importance assessment', *Risk Analysis* **22**(3), 579–590.

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

Schepsmeier, U. (2015), 'Efficient information based goodness-of-fit tests for vine copula models with fixed margins: A comprehensive review', *Journal of Multivariate Analysis* **138**, 34–52.

Su, X., Mahadevan, S., Xu, P. and Deng, Y. (2015), 'Dependence assessment in human reliability analysis using evidence theory and ahp', *Risk Analysis* **35**(7), 1296–1316.

Sundell, J., Haaf, E., Norberg, T., Alén, C., Karlsson, M. and Rosén, L. (2019), 'Risk mapping of groundwater-drawdown-induced land subsidence in heterogeneous soils on large areas', *Risk Analysis* **39**(1), 105–124.

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

Wang, S. (1996), 'Premium calculation by transforming the layer premium density', *ASTIN Bulletin* **26**(I), 71–92.

Wang, T., Dyer, J. S. and Butler, J. C. (2016), 'Modeling correlated discrete uncertainties in event trees with copulas', *Risk Analysis* **36**(2), 396–410.

Werner, C., Bedford, T. and Quigley, J. (2018), 'Sequential refined partitioning for probabilistic dependence assessment', *Risk Analysis* **38**(12), 2683–2702.

Xiao, S., Lu, Z. and Wang, P. (2018), 'Multivariate global sensitivity analysis based on distance components decomposition', *Risk Analysis* **38**(12), 2703–2721.