

Article

Reverse Sensitivity Analysis for Risk Modelling

Silvana M. Pesenti 

Department of Statistical Sciences, University of Toronto, Toronto, ON M5S 3G3, Canada;
silvana.pesenti@utoronto.ca

Abstract: We consider the problem where a modeller conducts sensitivity analysis of a model consisting of random input factors, a corresponding random output of interest, and a baseline probability measure. The modeller seeks to understand how the model (the distribution of the input factors as well as the output) changes under a stress on the output's distribution. Specifically, for a stress on the output random variable, we derive the unique stressed distribution of the output that is closest in the Wasserstein distance to the baseline output's distribution and satisfies the stress. We further derive the stressed model, including the stressed distribution of the inputs, which can be calculated in a numerically efficient way from a set of baseline Monte Carlo samples and which is implemented in the R package SWIM on CRAN. The proposed reverse sensitivity analysis framework is model-free and allows for stresses on the output such as (a) the mean and variance, (b) any distortion risk measure including the Value-at-Risk and Expected-Shortfall, and (c) expected utility type constraints, thus making the reverse sensitivity analysis framework suitable for risk models.

Keywords: distortion risk measures; expected utility; Wasserstein distance; robustness and sensitivity analysis; model uncertainty



Citation: Pesenti, Silvana M. 2022. Reverse Sensitivity Analysis for Risk Modelling. *Risks* 10: 141. <https://doi.org/10.3390/risks10070141>

Academic Editors: Annamaria Olivieri and Ermanno Pitacco

Received: 26 May 2022

Accepted: 7 July 2022

Published: 18 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sensitivity analysis is indispensable for model building, model interpretation, and model validation, as it provides insight into the relationship between model inputs and outputs. A key tool used for sensitivity analysis are sensitivity measures, that assign to each model input a score, representing an input factor's ability to explain the variability of a model output's summary statistic; see [Saltelli et al. \(2008\)](#) and [Borgonovo and Plischke \(2016\)](#) for an in-depth review. One of the most widely used output summary statistic is the variance, which gives rise to sensitivity measures, e.g., the Sobol indices, that apportion the uncertainty in the output's variance to input factors. In many applications, such as reliability management and financial and insurance risk management, however, the variance is not the output statistic of concern and instead quantile-base measures are used; indicatively, see [Asimit et al. \(2019\)](#); [Fissler and Pesenti \(2022\)](#); [Maume-Deschamps and Niang \(2018\)](#); [Tsanakas and Millossovich \(2016\)](#). Furthermore, typical for financial risk management applications is that model inputs are subject to distributional uncertainty. Probabilistic (or global) sensitivity measures, however, tacitly assume that the model's distributional assumptions are correctly specified; indeed, sensitivity measures based on the difference between conditional (on a model input) and unconditional densities (of the output) are termed "common rationale" [Borgonovo et al. \(2016\)](#). Examples include indices, such as Borgonovo's sensitivity measures [Borgonovo \(2007\)](#), the f -sensitivity index [Rahman \(2016\)](#), and sensitivity indices based on the Cramér–von Mises distance [Gamboa et al. \(2018\)](#), we also refer to [Plischke and Borgonovo \(2019\)](#) for a detailed overview and to [Gamboa et al. \(2020\)](#) for estimation of these sensitivity measures. Recently, [Plischke and Borgonovo \(2019\)](#) define sensitivity measures that depend only on the copula between input factors, whereas [Pesenti et al. \(2021\)](#) propose a sensitivity measure based on directional derivatives that take dependence between input factors into account. Estimating these sensitivities, however, may render difficult in application where joint observations are

scarce, e.g., insurance portfolios, and their interpretation may be limited as dependence structures are commonly specified by expert opinions [Denuit et al. \(2006\)](#).

We consider an alternative sensitivity analysis framework proposed in [Pesenti et al. \(2019\)](#) that (a) considers statistical summaries relevant to risk management, (b) applies to models subject to distributional uncertainty, thus instead of relying on correctly specified distributions from which to calculate sensitivity measures we derive alternative distributions that fulfil a specific probabilistic stress and are “closest” to the baseline distribution; and (c) studies reverse sensitivity measures. Differently to the framework proposed in [Pesenti et al. \(2019\)](#) who use the Kullback–Leibler divergence to quantify the closedness of probability measures, in this work we consider the Wasserstein distance of order two to measure the distance between distribution functions. The Wasserstein distance allows for more flexibility in the choice of stresses including survival probabilities (via quantiles) used in reliability analysis, risk measures employed in finance and insurance, and utility functions relevant for decision under ambiguity.

Central to the reverse sensitivity analysis framework is a *baseline model*, the 3-tuple $(\mathbf{X}, g, \mathbb{P})$, consisting of random input factors $\mathbf{X} = (X_1, \dots, X_n)$, an aggregation function $g: \mathbb{R}^n \rightarrow \mathbb{R}$ mapping input factors to a univariate output $Y = g(\mathbf{X})$, and a probability measure \mathbb{P} . The methodology has been termed *reverse sensitivity analysis* by [Pesenti et al. \(2019\)](#) since it proceeds in a reverse fashion to classical sensitivity analysis where input factors are perturbed and the corresponding altered output is studied. Indeed, in the reverse sensitivity analysis proposed by [Pesenti et al. \(2019\)](#) a stress on the output’s distribution is defined and changes in the input factors are monitored. The quintessence of the sensitivity analysis methodology is, however, not confined to stressing the output’s distribution, it is also applicable to stressing an input factor and observing the changes in the model output and in the other inputs. Throughout the exposition, we focus on the reverse sensitivity analysis that proceeds via the following steps:

- (i) Specify a stress on the baseline distribution of the output;
- (ii) Derive the unique stressed distribution of the output that is closest in the Wasserstein distance and fulfils the stress;
- (iii) The stressed distribution induces a canonical Radon–Nikodym derivative $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$; a change of measures from the baseline \mathbb{P} to the stressed probability measure \mathbb{Q}^* ;
- (iv) Calculate sensitivity measures that reflect an input factors’ change in distribution from the baseline to the stressed model.

Sensitivity testing using divergence measures—in the spirit of the reverse sensitivity methodology—has been studied by [Cambou and Filipović \(2017\)](#) using f -divergences on a finite probability space; by [Pesenti et al. \(2019\)](#) and [Pesenti et al. \(2021\)](#) using the Kullback–Leibler divergence; and [Makam et al. \(2021\)](#) consider a discrete sample space combined with the χ^2 -divergence. It is however known that the set of distribution functions with finite f -divergence, e.g., the Kullback–Leibler and χ^2 divergence—around a baseline distribution function depends on the baseline’s tail-behaviour, thus the choice of f -divergence should be chosen dependent on the baseline distribution [Kruse et al. \(2019\)](#). The Wasserstein distance on the contrary, automatically adapts to the baseline distribution function in that the Wasserstein distance penalises dissimilar distributional features such as different tail behaviour [Bernard et al. \(2020\)](#). The Wasserstein distance has enjoyed numerous applications to quantify distributional uncertainty, see, e.g., [Blanchet and Murthy \(2019\)](#) and [Bernard et al. \(2020\)](#) for applications to financial risk management. In the context of uncertainty quantification, [Moosmüller et al. \(2020\)](#) utilise the Wasserstein distance to elicit the (uncertain) aggregation map g from the distributional knowledge of the inputs and outputs. [Fort et al. \(2021\)](#) utilises the Wasserstein distance to introduce global sensitivity indices for computer codes whose output is a distribution function. In this manuscript we use the Wasserstein distance as it allows for different stresses compared to the Kullback–Leibler divergence. Indeed, the Wasserstein distance allows for stresses on any distortion risk measures, while the Kullback–Leibler divergence only allow for stresses on risk measures which are Value-at-Risk (VaR) and VaR and Expected Shortfall jointly, see [Pesenti et al. \(2019\)](#).

This paper is structured as follows: In Section 2, we state the notation and definitions necessary for the exposition. Section 3 introduces the optimisation problems and we derive the unique stressed distribution function of the output which has minimal Wasserstein distance to the baseline output's distribution and satisfies a stress. The considered stresses include constraints on risk measures, quantiles, expected utilities, and combinations thereof. In Section 4, we characterise the canonical Radon–Nikodym derivative, induced by the stressed distribution function, and study how input factors' distributions change when moving from the baseline to the stressed model. An application of the reverse sensitivity analysis is demonstrated on a mixture model in Section 5.

All proofs are delegated to Appendix A.

2. Preliminaries

Throughout we work on a measurable space (Ω, \mathcal{A}) and denote the sets of distribution functions with finite second moment by

$$\mathcal{M} = \left\{ G: \mathbb{R} \rightarrow [0, 1] \mid \begin{array}{l} G \text{ non-decreasing, right-continuous, } \lim_{x \searrow -\infty} G(x) = 0, \\ \lim_{x \nearrow +\infty} G(x) = 1, \text{ and } \int x^2 dG(x) < +\infty \end{array} \right\},$$

and the corresponding set of square-integrable (left-continuous) quantile functions by

$$\check{\mathcal{M}} = \left\{ \check{G} \in \mathbb{L}^2([0, 1]) \mid \check{G} \text{ non-decreasing \& left-continuous} \right\}.$$

or any distribution function $G \in \mathcal{M}$, we denote its corresponding (left-continuous) quantile function by $\check{G} \in \check{\mathcal{M}}$, that is $\check{G}(u) = \inf\{y \in \mathbb{R} \mid G(y) \geq u\}$, $u \in [0, 1]$, with the convention that $\inf \emptyset = +\infty$. We measure the discrepancy between distribution functions on the real line using the Wasserstein distance of order 2, defined as follows.

Definition 1 (Wasserstein Distance). *The Wasserstein distance (of order 2) between two distribution functions F_1 and F_2 is defined as Villani (2008)*

$$W_2(F_1, F_2) = \inf_{\pi \in \Pi(F_1, F_2)} \left\{ \left(\int_{\mathbb{R}^2} |z_1 - z_2|^2 \pi(dz_1, dz_2) \right)^{\frac{1}{2}} \right\},$$

where $\Pi(F_1, F_2)$ denotes the set of all bivariate probability measures with marginal distributions F_1 and F_2 , respectively.

The Wasserstein distance is the minimal quadratic cost associated with transporting the distribution F_1 to F_2 using all possible couplings (bivariate distributions) with fixed marginals F_1 and F_2 . The Wasserstein distance admits desirable properties to quantify model uncertainty such as the comparison of distributions with differing support, e.g., with the empirical distribution function. Moreover it is symmetric and forms a metric on the space of probability measures; we refer to Villani (2008) for an overview and properties of the Wasserstein distance. It is well known (Dall'Aglio 1956) that for distributions on the real line, the Wasserstein distance admits the representation

$$W_2(F_1, F_2) = \left(\int_0^1 |\check{F}_1(u) - \check{F}_2(u)|^2 du \right)^{\frac{1}{2}}.$$

3. Deriving the Stressed Distribution

Throughout this section we assume that the modeller's *baseline model* is the 3-tuple $(\mathbf{X}, g, \mathbb{P})$ consisting of a random vector of input factors $\mathbf{X} = (X_1, \dots, X_n)$, an aggregation function $g: \mathbb{R}^n \rightarrow \mathbb{R}$ mapping input factors to a (for simplicity) univariate output $Y = g(\mathbf{X})$,

and a probability measure \mathbb{P} . The baseline probability measure \mathbb{P} reflects the modeller's (statistical and expert) knowledge of the distribution of X and we denote the distribution function of the output by $F(y) = \mathbb{P}(Y \leq y)$. The modeller then performs reverse sensitivity analysis, that is tries to understand how prespecified stresses/constraints on the output distribution F , e.g., an increase in jointly its mean and standard deviation or a risk measures such as the Value-at-Risk (VaR) or Expected Shortfall (ES), affects the baseline model, e.g., the joint distribution of the input factors. For this, we first define the notion of a *stressed distribution*. Specifically, for given constraints we call a solution to the optimisation problem

$$\arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to stresses/constraints on } G, \quad (1)$$

a stressed distribution. In problem (1), the baseline distribution F is fixed and we seek over all alternative distributions $G \in \mathcal{M}$ the one who satisfies the stress(es) and which has smallest Wasserstein distance to F . The solution to problem (1)—the stressed distribution—may be interpreted as the most “plausible” distribution function arising under adverse circumstances. Examples of stresses and constraints considered in this work include an increase (decrease), compared to their corresponding values under the reference probability \mathbb{P} , in e.g., the mean, mean and standard deviation, distortion risk measures, and utility functions, and combinations thereof.

Next, we recall the concept of weighted isotonic projection which is intrinsically connected to the solution of optimisation problem (1); indeed the stressed quantile functions can be uniquely characterised via weighted isotonic projections.

Definition 2 (Weighted Isotonic Projection [Barlow et al. \(1972\)](#)). *The weighted isotonic projection $\ell^{\uparrow w}$ of a function $\ell \in \mathbb{L}^2([0, 1])$ with weight function $w: [0, 1] \rightarrow [0, +\infty)$, $w \in \mathbb{L}^2([0, 1])$, is its weighted projection onto the set of non-decreasing and left-continuous functions in $\mathbb{L}^2([0, 1])$. That is, the unique function satisfying*

$$\ell^{\uparrow w} = \arg \min_{h \in \check{\mathcal{M}}} \int_0^1 (\ell(u) - h(u))^2 w(u) du.$$

When the weight function is constant, i.e., $w(x) \equiv c$, $c > 0$, we write $\ell^{\uparrow}(\cdot) = \ell^{\uparrow c}(\cdot)$, as in this case the isotonic projection is indeed independent of c . The weighted isotonic projection admits not only a graphical interpretation as the non-decreasing function that minimises the weighted \mathbb{L}^2 -distance from ℓ but has also a discrete counterpart: the weighted isotonic regression [Barlow et al. \(1972\)](#). Numerically efficient algorithms for calculating weighted isotonic regressions are available, e.g., the R package `isotone` [De Leeuw et al. \(2010\)](#).

In the next sections, we solve problem (1) for different choices of constraints. Specifically, for risk measures constraints (Section 3.1), integral constraints (Section 3.2), Value-at-Risk constraints (Section 3.3), and expected utility constraint (Section 3.4), and in Section 3.5 we consider ways to smooth stressed distributions. Using these stressed distributions, we derive the stressed probability measures in Section 4 and study how a stress on the output is reflected on the input distribution(s).

3.1. Risk Measure Constraints

This section considers stresses on distortion risk measures, that is we derive the unique stressed distribution that satisfies an increase and/or decrease of distortion risk measures while minimising the Wasserstein distance to the baseline distribution F .

Definition 3 (Distortion Risk Measures). Let $\gamma \in \mathbb{L}^2([0, 1])$ be a square-integrable function with $\gamma: [0, 1] \rightarrow [0, +\infty)$ and $\int_0^1 \gamma(u) du = 1$. Then the distortion risk measure ρ_γ with distortion weight function γ is defined as

$$\rho_\gamma(G) = \int_0^1 \check{G}(u)\gamma(u) du \quad \text{for } G \in \mathcal{M}. \tag{2}$$

The above definition of distortion risk measures makes the assumption that positive realisations are undesirable (losses) while negative realisations are desirable (gains). The class of distortion risk measures includes one of the most widely used risk measures in financial risk management, the Expected Shortfall (ES) at level $\alpha \in [0, 1)$ (also called Tail Value-at-Risk), with $\gamma(u) = \frac{1}{1-\alpha} \mathbb{1}_{\{u>\alpha\}}$, see, e.g., Acerbi and Tasche (2002). The often used risk measure Value-at-Risk (VaR), while admitting a representation given in (2), has a corresponding weight function γ that is not square-integrable. We derive the solution to optimisation problem (1) with a VaR constraint in Section 3.3.

Theorem 1 (Distortion Risk Measures). Let $r_k \in \mathbb{R}$, ρ_{γ_k} be a distortion risk measure with weight function γ_k and assume there exists a distribution function $\check{G} \in \mathcal{M}$ satisfying $\rho_{\gamma_k}(\check{G}) = r_k$ for all $k \in \{1, \dots, d\}$. Then, the optimisation problem

$$\arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to} \quad \rho_{\gamma_k}(G) = r_k \quad k = 1, \dots, d, \tag{3}$$

has a unique solution given by

$$\check{G}^*(u) = \left(\check{F}(u) + \sum_{k=1}^d \lambda_k \gamma_k(u) \right)^\uparrow, \tag{4}$$

where the Lagrange multipliers λ_k are such that the constraints are fulfilled, that is $\rho_{\gamma_k}(\check{G}^*) = r_k$ for all $k = 1, \dots, d$.

In the above theorem, and also in later results, we assume that there exists a distribution function which satisfies all constraints. This assumption is not restrictive and requires that, particularly, multiple constraints are chosen carefully, e.g., imposing that $\int_0^1 \check{G}(u) du > \frac{1}{1-\alpha} \int_\alpha^1 \check{G}(u) du$ for $\alpha \in (0, 1)$, i.e., the mean being larger than the ES_α , cannot be fulfilled by any distribution function; thus, a combination of stresses not of interest to a modeller.

We observe that the optimal quantile function is the isotonic projection of a weighted linear combination of the baseline’s quantile function \check{F} and the distortion weight functions of the risk measures. A prominent group of risk measures is the class of coherent risk measures, that are risk measures fulfilling the properties of monotonicity, positive homogeneity, translation invariance, and sub-additivity; see Artzner et al. (1999) for a discussion and interpretation. It is well-known that a distortion risk measure is coherent, if and only if, its distortion weight function $\gamma(\cdot)$ is non-decreasing Kusuoka (2001). For the special case of a constraint on a coherent distortion risk measure that results in a larger risk measure compared to the baseline’s, we obtain an analytical solution without the need to calculate an isotonic projection.

Proposition 1 (Coherent Distortion Risk Measure). If ρ_γ is a coherent distortion risk measure and $r \geq \rho_\gamma(F)$, then optimisation problem (3) with $d = 1$ has a unique solution given by

$$\check{G}^*(u) = \check{F}(u) + \frac{r - \rho_\gamma(F)}{\int_0^1 (\gamma(u))^2 du} \gamma(u).$$

We illustrate the stressed distribution functions for constraints on distortion risk measures in the next example. Specifically, we look at the α - β risk measures which are a parametric family of distortion risk measures.

Example 1 (α - β Risk Measure). The α - β risk measure, $0 < \beta \leq \alpha < 1$, is defined by

$$\gamma(u) = \frac{1}{\eta} \left(p \mathbb{1}_{\{u < \beta\}} + (1 - p) \mathbb{1}_{\{u \geq \alpha\}} \right),$$

where $p \in [0, 1]$ and $\eta = p\beta + (1 - p)(1 - \alpha)$ is the normalising constant. This parametric family contains several notable risk measures as special cases: for $p = 0$ we obtain ES_α , and for $p = 1$ the conditional lower tail expectation (LTE) at level β .

Moreover, if $p < \frac{1}{2}$ ($p > \frac{1}{2}$) the α - β risk measure emphasises losses (gains) relative to gains (losses). For $\alpha = \beta$ and $p < \frac{1}{2}$, the risk measure is equivalent to $\kappa(ES_\alpha[Y] - \lambda \mathbb{E}[Y])$, where $\kappa = \frac{(1-2p)(1-\alpha)}{\eta}$ and $\lambda = \frac{p}{\kappa\eta}$.

Figure 1 displays the baseline \check{F}_Y and the stressed \check{G}_Y^* quantile functions of a random variable Y under a 10% increase on the α - β risk measure with $\beta = 0.1$, $\alpha = 0.9$, and various $p \in \{0.25, 0.5, 0.75\}$. The baseline distribution is chosen to be F_Y is Lognormal(μ, σ^2) with parameters $\mu = \frac{7}{8}$ and $\sigma = 0.5$. We observe in Figure 1 that the stressed quantile functions \check{G}_Y^* have, in all three plots, a flat part which straddles $\beta = 0.1$ and a jump at $\alpha = 0.9$. The length of the flat part is increasing with increasing p while the size of the jump is decreasing with increasing p . This can also be seen in the stressed densities g_Y^* which have, for all values of p , a much heavier right albeit a much lighter left tail than the density of the baseline model. Thus, under this stress, both tails of the baseline distribution are altered.

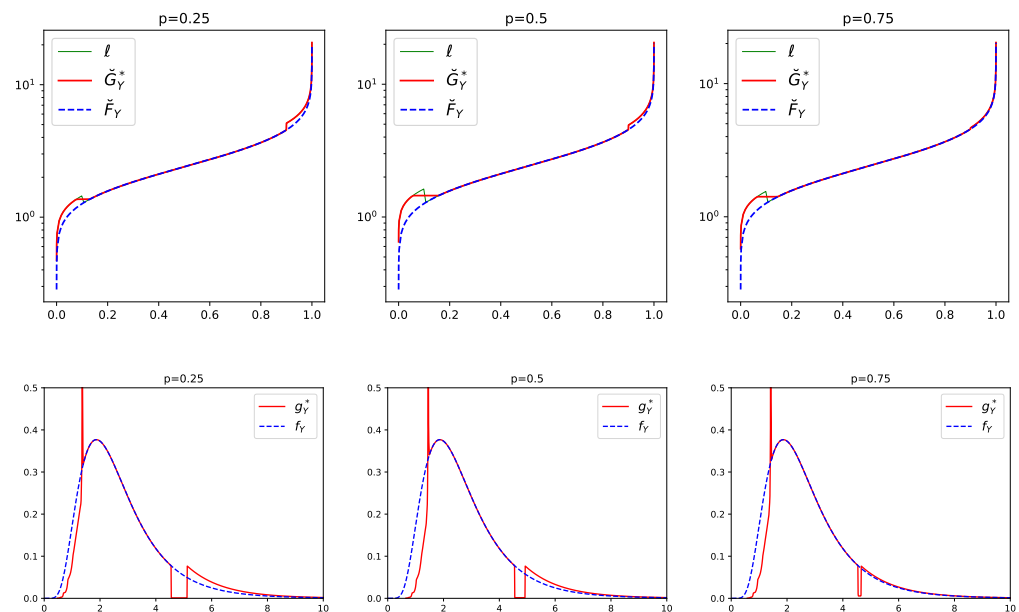


Figure 1. Top panels: Baseline quantile function \check{F}_Y (blue dashed) compared to the stressed quantile function \check{G}_Y^* (red solid) for a 10% increase on the α - β risk measure with $\beta = 0.1$, $\alpha = 0.9$, and various values of p . The green line $\ell(\cdot)$ is the function, whose isotonic projection equals $\check{G}_Y^*(\cdot)$. Bottom panels: corresponding baseline f_Y and stressed g_Y^* densities.

3.2. Integral Constraints

The next results are generalisations of stresses on distortion risk measures to integral constraints, and include as a special case a stress jointly on the mean, the variance, and distortion risk measures.

Theorem 2 (Integral). Let $h_k, \tilde{h}_l: [0, 1] \rightarrow [0, \infty)$ be square-integrable functions and assume there exists a distribution function $\tilde{G} \in \mathcal{M}$ satisfying $\int_0^1 h_k(u)\tilde{G}(u) du \leq c_k$ and $\int_0^1 \tilde{h}_l(u)(\tilde{G}(u))^2 du \leq \tilde{c}_l$ for all $k = 1, \dots, d$, and $l = 1, \dots, \tilde{d}$. Then the optimisation problem

$$\begin{aligned} \arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to} \quad & \int_0^1 h_k(u)\tilde{G}(u) du \leq c_k, \quad k = 1, \dots, d, \\ & \int_0^1 \tilde{h}_l(u)(\tilde{G}(u))^2 du \leq \tilde{c}_l, \quad l = 1, \dots, \tilde{d}, \end{aligned}$$

has a unique solution given by

$$\check{G}^*(u) = \left(\frac{1}{\tilde{\Lambda}(u)} \left(\check{F}(u) + \sum_{k=1}^d \lambda_k h_k(u) \right) \right)^{\uparrow_{\tilde{\Lambda}}},$$

where $\tilde{\Lambda}(u) = 1 + \sum_{k=1}^{\tilde{d}} \tilde{\lambda}_k \tilde{h}_k(u)$ and the Lagrange multipliers $\lambda_1, \dots, \lambda_d$ and $\tilde{\lambda}_1, \dots, \tilde{\lambda}_{\tilde{d}}$ are non-negative and such that the constraints are fulfilled.

A combination of the above theorems provides stresses jointly on the mean, the variance, and on multiple distortion risk measures.

Proposition 2 (Mean, Variance, and Risk Measures). Let $m' \in \mathbb{R}, \sigma' > 0, r_k \in \mathbb{R}$, and distortion risk measures $\rho_{\gamma_k}, k = 1, \dots, d$. Assume there exists a distribution function $\tilde{G} \in \mathcal{M}$ with mean m' , standard deviation σ' , and which satisfies $\rho_{\gamma_k}(\tilde{G}) = r_k$, for all $k = 1, \dots, d$. Then the optimisation problem

$$\begin{aligned} \arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to} \quad & \int x dG(x) = m', \\ & \int (x - m')^2 dG(x) = (\sigma')^2 \quad \text{and} \\ & \rho_{\gamma_k}(G) = r_k, \quad k = 1, \dots, d, \end{aligned}$$

has a unique solution given by

$$\check{G}^*(u) = \left(\frac{1}{1 + \lambda_2} \left(\check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right) \right)^{\uparrow},$$

and the Lagrange multipliers $\lambda_1, \dots, \lambda_{d+2}$ with $\lambda_2 \neq -1$ are such that the constraints are fulfilled.

Example 2 (Mean, Variance, and ES). Here, we illustrate Proposition 2 with the ES risk measure and three different stresses. The top panels of Figure 2 display the baseline quantile function \check{F}_Y and the stressed quantile function \check{G}_Y^* of Y , where the baseline distribution F_Y of Y is again Lognormal(μ, σ^2) with parameters $\mu = \frac{7}{8}$ and $\sigma = 0.5$. The bottom panels display the corresponding baseline and stressed densities. The left panels correspond to a stress, where, under the stressed model, the $ES_{0.95}$ and the mean are kept fixed at their corresponding values under the baseline model, while the standard deviation is increased by 20%. We observe, both in the quantile and density plot, that the stressed distribution is more spread out indicating a larger variance. Furthermore, at $y \approx 5.77$ the stressed density $g_Y^*(y)$ drops to ensure that $ES_{0.95}(G_Y^*) = ES_{0.95}(F_Y)$. This drop is due to the fact that a stress composed of a 20% increase in the standard deviation while fixing the mean (i.e., without a constraint on ES) results in an ES that is larger compared to the baseline's. Indeed, under this alternative stress (without a constraint on ES) we obtain that $ES_{0.95}(G_Y^*) \approx 7.70$ compared to $ES_{0.95}(F_Y) \approx 6.87$.

The middle panels correspond to a 10% increase in $ES_{0.95}$ and a 10% decrease in the mean, while keeping the standard deviation fixed at its value under the baseline model. The density plot clearly indicates a general shift of the stressed density to the left, stemming from the decrease in the

mean, and a single trough which is induced by the increase in ES. The right panels correspond to a 10% increase in $ES_{0.95}$, a 10% increase in the mean, and a 10% decrease in the standard deviation. The stressed density still has the trough from the increase in ES; however, the density is less spread out (reduction in the standard deviation) and generally shifted to the right (increase in the mean).

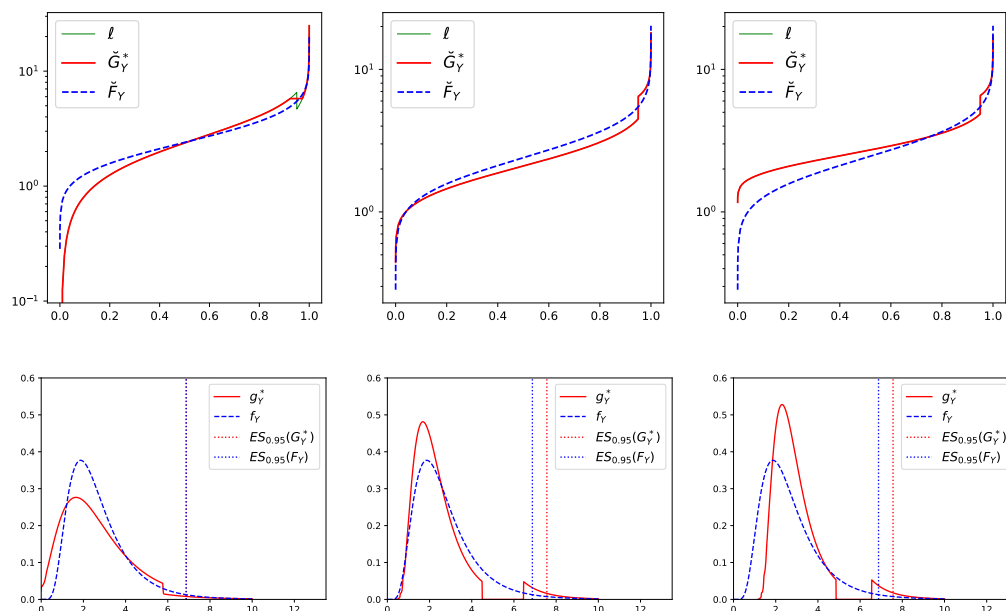


Figure 2. Top: Baseline quantile function \check{F}_Y compared to the stressed quantile function \check{G}_Y^* . Bottom: corresponding baseline f_Y and stressed g_Y^* densities. Left: $ES_{0.95}$ and the mean being fixed and a 20% increase in the standard deviation. Middle: 10% increase in $ES_{0.95}$, 10% decrease in the mean, and fixed standard deviation. Right: 10% increase in $ES_{0.95}$, 10% increase in the mean, and 10% decrease in standard deviation. Note that in the middle and right panel the green lines are equal to the red lines.

3.3. Value-at-Risk Constraints

In this section we study stresses on the risk measure Value-at-Risk (VaR). The VaR at level $\alpha \in (0, 1)$ of a distribution function $G \in \mathcal{M}$ is defined as its left-continuous quantile function evaluated at α , that is

$$VaR_\alpha(G) = \check{G}(\alpha).$$

We further define the right-continuous VaR^+ , that is the right-continuous quantile function of $G \in \mathcal{M}$ evaluated at α , by

$$VaR_\alpha^+(G) = \check{G}^+(\alpha) = \inf \{y \in \mathbb{R} \mid F(y) > \alpha\}.$$

Theorem 3 (VaR). Let $q \in \mathbb{R}$ and consider the optimisation problem

$$\begin{aligned} \arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to} \quad & (a) \quad VaR_\alpha(G) = q \quad \text{or} \\ & (b) \quad VaR_\alpha^+(G) = q, \end{aligned}$$

and define α_F such that $VaR_{\alpha_F}(F) = q$. Then, the following holds

- (i) under constraint (a), if $q \leq VaR_\alpha(F)$, then the unique solution is given by

$$\check{G}^*(u) = \check{F}(u) + (q - \check{F}(u)) \mathbf{1}_{\{u \in (\alpha_F, \alpha]\}};$$

if $q > VaR_\alpha(F)$, then there does not exist a solution.

(ii) under constraint (b), if $q \geq \text{VaR}_\alpha^+(F)$, then the unique solution is given by

$$\check{G}^*(u) = \check{F}(u) + (q - \check{F}(u)) \mathbb{1}_{\{u \in (\alpha, \alpha_F]\}};$$

if $q < \text{VaR}_\alpha^+(F)$, then there does not exist a solution.

The above theorem states that if the optimal quantile function exists it is either the baseline quantile function \check{F} or constant equal to q . Moreover, the stressed quantile function (if it exists) jumps at α which implies that the existence of a solution hinges on the careful choice of the stress. For a stress on VaR (constraint (a)) for example, a solution exists if and only if the constraint satisfies $q \leq \text{VaR}_\alpha(F)$; a decrease in the VaR_α from the baseline to the stressed model. The reason for the non-existence of a solution when stressing VaR upwards is that the unique increasing function that minimises the Wasserstein distance and satisfies the constraint is not left-continuous and thus not a quantile function.

Alternatively to stressing VaR or VaR^+ , and in particularly in the case when a desired stressed solution does not exist, one may stress instead the distortion risk measure Range-Value-at-Risk (RVaR) Cont et al. (2010). The RVaR at levels $0 \leq \alpha < \beta \leq 1$ is defined by

$$\text{RVaR}_{\alpha, \beta}(G) = \frac{1}{\beta - \alpha} \int_\alpha^\beta \check{G}(u) du, \quad \text{for } G \in \mathcal{M},$$

and belongs to the class of distortion risk measures. The RVaR attains as limiting cases the VaR and VaR^+ . Indeed, for any $G \in \mathcal{M}$ it holds

$$\text{VaR}_\alpha(G) = \lim_{\alpha' \nearrow \alpha} \text{RVaR}_{\alpha', \alpha}(G) \quad \text{and} \quad \text{VaR}_\alpha^+(G) = \lim_{\beta \searrow \alpha} \text{RVaR}_{\alpha, \beta}(G).$$

The solution to stressing RVaR is provided in Theorem 1.

3.4. Expected Utility Constraint

This section considers the change from the baseline to the stressed distribution under an increase of an expected utility constraint. In the context of utility maximisation, the next theorem provides a way to construct stressed models with a larger utility compared to the baseline.

Theorem 4 (Expected Utility and Risk Measures). *Let $u: \mathbb{R} \rightarrow \mathbb{R}$ be a differentiable concave utility function, $r_k \in \mathbb{R}$, and ρ_{γ_k} be distortion risk measures, for $k = 1, \dots, d$. Assume there exists a distribution function \check{G} satisfying $\int_{\mathbb{R}} u(x) d\check{G}(x) \geq c$ and $\rho_{\gamma_k}(\check{G}) = r_k$ for all $k = 1, \dots, d$. Then the optimisation problem*

$$\arg \min_{G \in \mathcal{M}} W_2(G, F) \quad \text{subject to} \quad \int_{\mathbb{R}} u(x) dG(x) \geq c \quad \& \quad \rho_{\gamma_k}(G) = r_k, \quad k = 1, \dots, d$$

has a unique solution given by

$$\check{G}^*(u) = \check{v}_{\lambda_1} \left(\left(\check{F}(u) + \sum_{k=1}^d \lambda_{k+1} \gamma_k(v) \right)^\uparrow \right), \tag{5}$$

where \check{v}_{λ_1} is the left-inverse of $v_{\lambda_1}(x) = x - \lambda_1 u'(x)$, and $\lambda_1 \geq 0, (\lambda_2, \dots, \lambda_{d+1}) \in \mathbb{R}^d$ are such that the constraints are fulfilled.

The utility function in Theorem 4 need not be monotone, indeed the theorem applies to any differentiable concave function, without the need of an utility interpretation. Moreover, Theorem 4 also applies to differentiable convex (disutility) functions \check{u} and constraint $\int_{\mathbb{R}} \check{u}(x) dG(x) \leq c$; a situation of interest in insurance premium calculations. In this case, the solution is given by (5) with $u(x) = -\check{u}(x)$.

Example 3 (HARA Utility & ES). The Hyperbolic absolute risk aversion (HARA) utility function is defined by

$$u(x) = \frac{1 - \eta}{\eta} \left(\frac{ax}{1 - \eta} + b \right)^\eta,$$

with parameters $a > 0$, $\frac{ax}{1 - \eta} + b > 0$, and where $\eta \leq 1$ guarantees concavity.

We again choose the baseline distribution F_Y of Y to be Lognormal(μ, σ^2) with $\mu = \frac{7}{8}$ and $\sigma = 0.5$ and consider utility parameters $a = 1$, $b = 5$, and $\eta = 0.5$. Figure 3 displays the baseline and the stressed quantile functions \check{F}_Y and \check{G}_Y^* , respectively, for a combined stress on the HARA utility and on ES at levels 0.8 and 0.95. Specifically, for all three stresses we decreasing $ES_{0.8}$ by 10% and increasing $ES_{0.95}$ by 10% compared to their values under the baseline model. Moreover, the HARA utility is increased by 0%, 1%, and 3%, respectively, corresponding to the panels from the left to the right. The flat part in the stressed quantile function $\check{G}^*(u)$ around $u = 0.8$, visible in all top panels of Figure 3, is induced by the decrease in $ES_{0.8}$ while the jump at $u = 0.95$ is due to the increase in $ES_{0.95}$. From the left to the right panel in Figure 3, we observe that the larger the stress on the HARA utility, the more the stressed quantile function shifts away from the baseline quantile function \check{F}_Y .

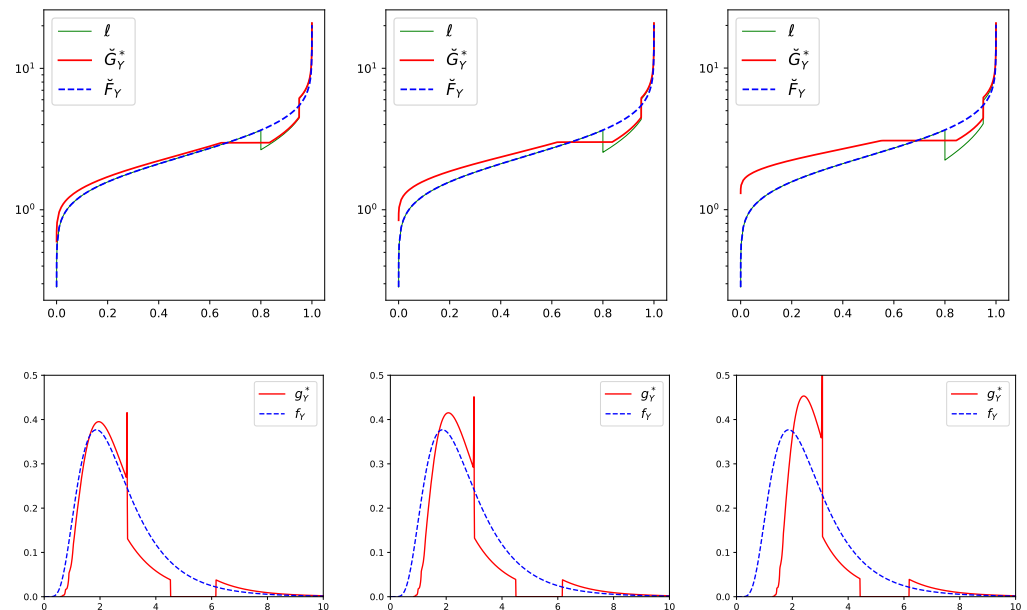


Figure 3. Top panels: Baseline quantile function \check{F}_Y compared to the stressed quantile function \check{G}_Y^* , for a 10% decrease in $ES_{0.8}$, and a 10% increase in $ES_{0.95}$, and, from left to right, a 0%, 1%, and 3% increase in the HARA utility, respectively. The function $\ell(\cdot)$ (solid green) is the function whose isotonic projection equals $\check{G}^*(\cdot)$. Bottom panels: Corresponding baseline f_Y and stressed g_Y^* densities.

3.5. Smoothing of the Stressed Distribution

We observe that the stressed quantile functions derived in Section 3 generally contain jumps and/or flat parts even if the baseline distribution is absolutely continuous. In situation where this is not desirable, one may consider a smoothed version of the stressed distributions. For this, we recall that the isotonic regression, the discrete counterpart of the weighted isotonic projection, of a function ℓ evaluated at u_1, \dots, u_n with positive weights w_1, \dots, w_n , is the solution to

$$\min_{u_1, \dots, u_n} \sum_{i=1}^n (u_i - \ell(u_i))^2 w_i, \quad \text{subject to } u_i \leq u_j, \quad i \leq j. \tag{6}$$

There are numerous efficient algorithms that solve (6) most notably the pool-adjacent-violators (PAV) algorithm Barlow et al. (1972). It is well-known that the solution to the

isotonic regression contains flat parts and jumps. A smoothed isotonic regression algorithm, termed smooth pool-adjacent-violators (SPAV) algorithm, using an \mathbb{L}^2 regularisation was recently proposed by [Sysoev and Burdakov \(2019\)](#). Specifically, they consider

$$\min_{u_1, \dots, u_n} \sum_{i=1}^n (u_i - \ell(u_i))^2 w_i + \sum_{i=1}^n \zeta_i (u_{i+1} - u_i)^2, \quad \text{subject to } u_i \leq u_j, \quad i \leq j,$$

where $\zeta_i \geq 0, i = 0, \dots, n-1$, are prespecified smoothing parameters. Using a probabilistic reasoning, [Sysoev and Burdakov \(2019\)](#) argue that ζ_i may be chosen proportional to a (e.g., quadratic) kernel evaluated at u_i and u_{i+1} , that is

$$\zeta_i = \zeta K(u_i, u_{i+1}) \quad \text{with} \quad K(u_i, u_{i+1}) = \frac{1}{|u_i - u_{i+1}|^2} \quad \text{and} \quad \zeta \geq 0.$$

The choice of smoothing parameter $\zeta = 0$ correspond to the original isotonic regression larger values of ζ correspond to a greater degree of smoothness of the solution. ζ can either be prespecified or estimated using cross-validation, see e.g., [Sysoev and Burdakov \(2019\)](#).

To guarantee that the smoothed quantile function still fulfils the constraint, one may replace in every step of the optimisation for finding the Lagrange parameter the PAV with the SPAV algorithm. Thus, the Lagrange parameter are indeed found such that the constraints are fulfilled.

Remark 1. *There are numerous works proposing smooth versions of isotonic regressions. Approaches include kernel smoothers, e.g., [Hall and Huang \(2001\)](#), and spline techniques, e.g., [Meyer \(2008\)](#). These algorithms, however, are computationally heavy in that their computational cost is $O(n^2)$, where n is the number of data points. Furthermore, these algorithm require a careful choice of the kernel or the spline basis which is in contrast to the SPAV. We refer the reader to [Sysoev and Burdakov \(2019\)](#) for a detailed discussion and references to smooth isotonic regression algorithms.*

4. Analysing the Stressed Model

Recall that a modeller is equipped with a baseline model, the 3-tuple (X, g, \mathbb{P}) , consisting of a set of input factors $X = (X_1, \dots, X_n)$, a univariate output random variable of interest, $Y = g(X)$, and a probability measure \mathbb{P} . For a stress on the output's baseline distribution F_Y , we derived in Section 3 the corresponding unique stressed distribution function, denoted here by G_Y^* . Thus, to fully specify the stressed model we next define a stressed probability measure \mathbb{Q}^* that is induced by G_Y^* .

4.1. The Stressed Probability Measures

A stressed distribution G_Y^* induces a canonical change of measure that allows the modeller to understand how the baseline model including the distributions of the inputs changes under the stress. The Radon–Nikodym (RN) derivative of the baseline to the stressed model is

$$\frac{d\mathbb{Q}^*}{d\mathbb{P}} = \frac{g_Y^*(Y)}{f_Y(Y)},$$

where f_Y and g_Y^* denote the densities of the baseline and stressed output distribution, respectively. The RN derivative is well-defined since $f_Y(Y) > 0$, \mathbb{P} -a.s. The distribution functions of input factors under the stressed model – the stressed distributions – are then given, e.g., for input $X_i, i \in \{1, \dots, n\}$, by

$$\mathbb{Q}^*(X_i \leq x_i) = \mathbb{E} \left[\mathbb{1}_{\{X_i \leq x_i\}} \frac{d\mathbb{Q}^*}{d\mathbb{P}} \right] = \mathbb{E} \left[\mathbb{1}_{\{X_i \leq x_i\}} \frac{g_Y^*(Y)}{f_Y(Y)} \right], \quad x_i \in \mathbb{R},$$

and for multivariate inputs \mathbf{X} by

$$\mathbb{Q}^*(\mathbf{X} \leq \mathbf{x}) = \mathbb{E} \left[\mathbb{1}_{\{\mathbf{X} \leq \mathbf{x}\}} \frac{g_Y^*(Y)}{f_Y(Y)} \right], \quad \mathbf{x} \in \mathbb{R}^n,$$

where $\mathbb{E}[\cdot]$ denotes the expectation under \mathbb{P} . Note that under the stressed probability measure \mathbb{Q}^* , the input factors' marginal and joint distributions may be altered.

Example 4 (HARA Utility & ES continued). We continue Example 3 and illustrate the RN-densities $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ for the following three stresses (from the left to the right panel): a 10% decrease in $ES_{0.8}$ and a 10% increasing $ES_{0.95}$ for all three stresses, and a 0%, 1%, and 3% increase in the HARA utility, respectively.

We observe in Figure 4, that for all three stresses large realisations of Y obtain a larger weight under the stressed probability measures \mathbb{Q}^* compared to the baseline probability \mathbb{P} . Indeed, for all three stresses it holds that $\frac{d\mathbb{Q}^*}{d\mathbb{P}}(\omega) > 1$ whenever $Y(\omega) > 6$ and $\omega \in \Omega$. This is in contrast to small realisations of Y which obtain a weight smaller than 1. The impact of the different levels of stresses of the HARA utility (0%, 1%, and 3%, from the left to the right panel) can be observed in the left tail of $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$; a larger stress on the utility induces larger weights. The length of the trough of $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ is increasing from the left panel (approx. (4.53, 6.15)) to the right panel (approx. (4.43, 6.18)), and corresponds in all cases to the constant part in G_Y^* (see Figure 3, top panels) which is induced by the decrease in $ES_{0.8}$ under the stressed model.

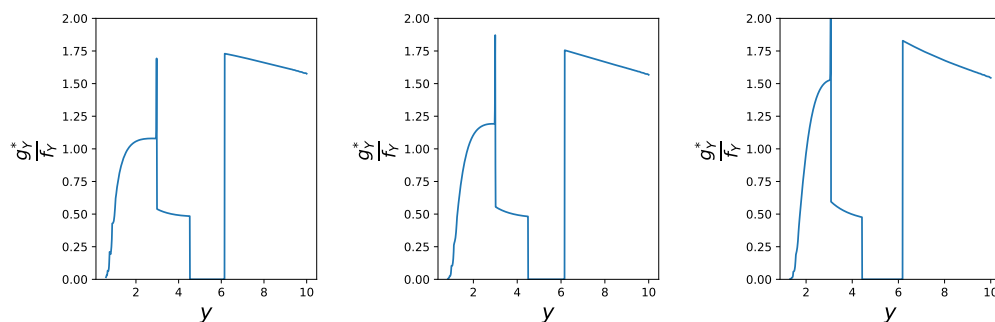


Figure 4. RN-densities for following stresses: a 10% decrease in both $ES_{0.8}$ and $ES_{0.95}$, and an increase in the HARA utility. The change in HARA utility is 0%, 1%, and 3%, respectively, from left to right.

4.2. Reverse Sensitivity Measures

Comparison of the baseline and a stressed model can be conducted via different approaches depending on the modeller's interest. While probabilistic sensitivity measures underlie the assumption of a fixed probability measure and quantify the divergence between the conditional (on a model input) and the unconditional output density Saltelli et al. (2008), the proposed framework compares a baseline and a stressed model, i.e., distributions under different probability measures. Therefore, to quantify the distributional change in input factor X_i from the baseline \mathbb{P} to the stressed \mathbb{Q}^* probability, a sensitivity measure introduced by Pesenti et al. (2019) may be suitable which quantifies the variability of an input factor's distribution from the baseline to the stressed model. A generalisation of the reverse sensitivity measure is stated here.

Definition 4 (Marginal Reverse Sensitivity Measure [Pesenti et al. \(2019\)](#)). For a function $s: \mathbb{R} \rightarrow \mathbb{R}$, the reverse sensitivity measure to input X_i with respect to a stressed probability measure \mathbb{Q}^* is defined by

$$\mathcal{S}_i^{\mathbb{Q}^*} = \begin{cases} \frac{\mathbb{E}^{\mathbb{Q}^*}[s(X_i)] - \mathbb{E}[s(X_i)]}{\max_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X_i)] - \mathbb{E}[s(X_i)]} & \mathbb{E}^{\mathbb{Q}^*}[s(X_i)] \geq \mathbb{E}[s(X_i)], \\ -\frac{\mathbb{E}^{\mathbb{Q}^*}[s(X_i)] - \mathbb{E}[s(X_i)]}{\min_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X_i)] - \mathbb{E}[s(X_i)]} & \text{otherwise,} \end{cases}$$

where $\mathcal{Q} = \{\mathbb{Q} \mid \mathbb{Q} \text{ probability measure with } \frac{d\mathbb{Q}}{d\mathbb{P}} \stackrel{\mathbb{P}}{=} \frac{d\mathbb{Q}^*}{d\mathbb{P}}\}$ is the set of all probability measures whose RN-derivative have the same distribution as $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ under \mathbb{P} . We adopted the convention that $\pm \frac{\infty}{\infty} = \pm 1$ and $\frac{0}{0} = 0$.

The sensitivity measure is called “reverse”, as the stress is applied to the output random variable Y and the sensitivity monitors the change in input X_i . The definition of 4 applies, however, also to stresses on input factors, in which case the RN-density $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ is a function of the stressed input factor and we refer to [Pesenti et al. \(2019\)](#) for a discussion. Note, that the reverse sensitivity measure can be viewed as a normalised covariance measure between the input $s(X_i)$ and the Radon Nikodym derivative $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$.

The next proposition provides a collection of properties that the reverse sensitivity measure possesses, we also refer to [Pesenti et al. \(2019\)](#) for a detailed discussion of these properties. For this, we first recall the definition of comonotonic and counter-monotonic random variables.

Definition 5. Two random variables Y_1 and Y_2 are comonotonic under \mathbb{P} , if and only if, there exists a random variable W and non-decreasing functions $h_1, h_2: \mathbb{R} \rightarrow \mathbb{R}$, such that the following equalities hold in distribution under \mathbb{P}

$$Y_1 = h_1(W) \quad \text{and} \quad Y_2 = h_2(W). \tag{7}$$

The random variables Y_1 and Y_2 are counter-monotonic under \mathbb{P} , if and only if, (7) holds with one of the functions $h_1(\cdot), h_2(\cdot)$ being non-increasing, and the other non-decreasing.

If two random variables are (counter) comonotonic under one probability measure, then they are also (counter) comonotonic under any other absolutely continuous probability measure, see, e.g., Proposition 2.1 of [Cuestaalbertos et al. \(1993\)](#). Thus, we omit the specification of the probability measure when discussing counter- and comonotonicity.

Proposition 3 (Properties of Reverse Sensitivity Measure). The reverse sensitivity measure possesses the following properties:

- (i) $\mathcal{S}_i^{\mathbb{Q}^*} \in [-1, 1]$;
- (ii) $\mathcal{S}_i^{\mathbb{Q}^*} = 0$ if $(s(X_i), \frac{d\mathbb{Q}^*}{d\mathbb{P}})$ are independent under \mathbb{P} ;
- (iii) $\mathcal{S}_i^{\mathbb{Q}^*} = 1$ if and only if $(s(X_i), \frac{d\mathbb{Q}^*}{d\mathbb{P}})$ are comonotonic;
- (iv) $\mathcal{S}_i^{\mathbb{Q}^*} = -1$ if and only if $(s(X_i), \frac{d\mathbb{Q}^*}{d\mathbb{P}})$ are counter-comonotonic.

The function $s(\cdot)$ provides the flexibility to create sensitivity measures that quantify changes in moments, e.g., via $s(x) = x^k, k \in \mathbb{N}$, or in the tail of distributions, e.g., via $s(x) = \mathbb{1}_{\{x > \text{VaR}_\alpha(X_i)\}}$, for $\alpha \in (0, 1)$.

Next, we generalise Definition 4 to a sensitivity measure that accounts for multiple input factors. While $\mathcal{S}_i^{\mathbb{Q}^*}$ measures the change of the distribution of X_i from the base-line to the stressed model the sensitivity $\mathcal{S}_{i,j}^{\mathbb{Q}^*}$ introduced below, quantifies how the joint distribution of (X_i, X_j) changes when moving from \mathbb{P} to \mathbb{Q}^* .

Definition 6 (Bivariate Reverse Sensitivity Measure). For a function $s: \mathbb{R}^2 \rightarrow \mathbb{R}$, the reverse sensitivity measure to inputs (X_i, X_j) with respect to a stressed probability measure \mathbb{Q}^* is defined by

$$\mathcal{S}_{i,j}^{\mathbb{Q}^*} = \begin{cases} \frac{\mathbb{E}^{\mathbb{Q}^*}[s(X_i, X_j)] - \mathbb{E}[s(X_i, X_j)]}{\max_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X_i, X_j)] - \mathbb{E}[s(X_i, X_j)]} & \mathbb{E}^{\mathbb{Q}^*}[s(X_i, X_j)] \geq \mathbb{E}[s(X_i, X_j)], \\ -\frac{\mathbb{E}^{\mathbb{Q}^*}[s(X_i, X_j)] - \mathbb{E}[s(X_i, X_j)]}{\min_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X_i, X_j)] - \mathbb{E}[s(X_i, X_j)]} & \text{otherwise,} \end{cases}$$

where \mathcal{Q} is given in Definition 4.

The bivariate sensitivity measure satisfies all the properties in Proposition 3 when $s(X_i)$ is replaced by $s(X_i, X_j)$. The bivariate sensitivity $\mathcal{S}_{i,j}^{\mathbb{Q}^*}$ can also be generalised to k input factors by choosing a function $s: \mathbb{R}^k \rightarrow \mathbb{R}$.

Remark 2. Probabilistic sensitivity measures are typically used for importance measurement and take values in $[0, 1]$; with 1 being the most important input factor and 0 being (desirably) independent from the output [Borgonovo et al. \(2021\)](#). This is in contrast to our framework where $\mathcal{S}_i^{\mathbb{Q}^*}$ lives in $[-1, 1]$ and e.g., a negative dependence, such as negative quadrant dependence between $s(X_i)$ and $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ implies that $\mathcal{S}_i^{\mathbb{Q}^*} < 0$, see [Pesenti et al. \(2019\)](#) [Proposition 4.3]. Thus, the proposed sensitivity measure is different in that it allows for negative sensitivities where the sign of $\mathcal{S}_i^{\mathbb{Q}^*}$ indicates the direction of the distributional change.

5. Application to a Spatial Model

We consider a spatial model for modelling insurance portfolio losses where each individual loss occurs at different locations and the dependence between individual losses is a function of the distance between the locations of the losses. Mathematically, denote the locations of the insurance losses by z_1, \dots, z_{10} , where $z_m = (z_m^1, z_m^2)$ are the coordinates of location z_m , $m = 1, \dots, 10$. The insurance loss at location m , denoted by L_m , follows a $\text{Gamma}(5, \frac{0.2}{m})$ distribution with location parameter 25. Thus, the minimum loss at each location is 25 and locations with larger mean also exhibit larger standard deviations. The losses L_1, \dots, L_m have, conditionally on $\Theta = \theta$, a Gaussian copula with correlation matrix given by $\rho_{i,j} = \text{Cor}(L_i, L_j) = e^{-\theta \|z_i - z_j\|}$, where $\|\cdot\|$ denotes the Euclidean distance. Thus, the further apart the locations z_i and z_j are the smaller the correlation between L_i and L_j . The parameter Θ takes values $(0, 0.4, 5)$ with probabilities $(0.05, 0.6, 0.35)$ that represent different regimes. Indeed, $\Theta = 0$ corresponds to a correlation of 1 between all losses, independently of their location. Larger values of Θ correspond to smaller albeit still positive correlation. Thus, regime with $\Theta = 0$ can be viewed as, e.g., circumstances suitable for natural disasters. We further define the total loss of the insurance company by $Y = \sum_{m=1}^{10} L_m$.

We perform two different stresses on the total loss Y detailed in Table 1. Specifically, we consider as a first stress a 0% change in HARA utility, a 0% change in $\text{ES}_{0.8}(Y)$, and a 1% increase in $\text{ES}_{0.95}(Y)$ from the baseline to the stressed model. The second stress is composed of a 1% increase in HARA utility, a 1% increase in $\text{ES}_{0.8}(Y)$, and a 3% increase in $\text{ES}_{0.95}(Y)$ compared to the baseline model. As the second stress increases all three metrics it may be viewed as a more severe distortion of the baseline model.

Table 1. Summary of the stresses applied to the portfolio loss Y represented in relative increases of the stressed model from the baseline model.

	HARA Utility	ES _{0.8} (Y)	ES _{0.95} (Y)
Stress 1: Q_1^*	0%	0%	1%
Stress 2: Q_2^*	1%	1%	3%

Next, we calculate reverse sensitivity measures for the losses L_1, \dots, L_{10} for both stresses Q_1^* and Q_2^* . Figure 5 displays the reverse sensitivity measures for functions $s(x) = x$, $s(x) = \mathbb{1}_{\{x > \check{F}_i(0.8)\}}$, and $s(x) = \mathbb{1}_{\{x > \check{F}_i(0.95)\}}$, from the left to the right panel, and where \check{F}_i denotes the \mathbb{P} -quantile function of L_i , $i = 1, \dots, 10$.

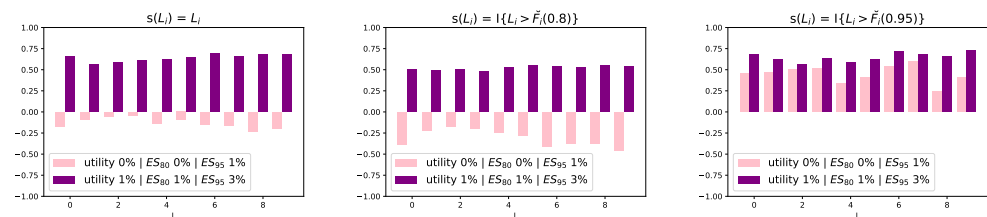


Figure 5. Reverse sensitivity measures with $s(x) = x$, $s(x) = \mathbb{1}_{\{x > \check{F}_i(0.8)\}}$, and $s(x) = \mathbb{1}_{\{x > \check{F}_i(0.95)\}}$ (left to right), for two different stresses on the output Y . First stress (salmon) is keeping the HARA utility and $ES(Y)_{0.8}$ fixed and increasing the $ES(Y)_{0.85}$ by 1%. Second stress (violet) is an increase of 1% in HARA utility, 1% $ES(Y)_{0.8}$, and 3% in $ES(Y)_{0.85}$.

We observe that for stress 2, the reverse sensitivities to all losses L_i and all choices of function $s(\cdot)$ are positive. This contrasts the reverse sensitivities for stress 1. Indeed, for stress 1 the reverse sensitivities with both $s(x) = x$ and $s(x) = \mathbb{1}_{\{x > \check{F}_i(0.8)\}}$ are negative, with the former values being smaller indicating a smaller change in the distributions of the L_i 's. By definition of the reverse sensitivity, the left panel corresponds to the (normalised) difference between the expectation under the stressed and baseline model. The middle and right panels correspond to the (normalised) change in the probability of exceeding $\check{F}_i(0.8)$ and $\check{F}_i(0.95)$, respectively. Thus, as seen in the plots, while the expectations and probabilities of exceeding the 80% \mathbb{P} -quantile are smaller under the stressed model, the probabilities of exceeding the 95% \mathbb{P} -quantile are increased substantially. The first stress increases the ES at level 0.95 while simultaneously fixes the utility and ES at level 0.8 to its values under the baseline model. This induces under the stressed probability measure a reduction of the mean and of the probability of exceeding the 80% \mathbb{P} -quantile while the probability of exceeding the 95% \mathbb{P} -quantile increases. Thus, the reverse sensitivity measures provide a spectrum of measures to analyse the distributional change of the losses L_i from the baseline to the stressed model.

Next, for a comparison we calculate the delta sensitivity measure of introduced by [Borgonovo \(2007\)](#). For a probability measure \mathbb{Q} the delta measure of L_i is defined by

$$\zeta^{\mathbb{Q}}(L_i) = \frac{1}{2} \int \int \left| f_Y^{\mathbb{Q}}(y) - f_{Y|L_i}^{\mathbb{Q}}(y|z) \right| f_i^{\mathbb{Q}}(z) dy dz,$$

where $f_Y^{\mathbb{Q}}(\cdot)$ and $f_i^{\mathbb{Q}}(\cdot)$ are the densities of Y and L_i under \mathbb{Q} , respectively, and where $f_{Y|L_i}^{\mathbb{Q}}(\cdot|\cdot)$ is the conditional density of the total portfolio loss Y given L_i under \mathbb{Q} .

Table 2 reports the delta measures under the baseline model $\zeta^{\mathbb{P}}$ and the two stresses, i.e., $\zeta^{Q_1^*}$ and $\zeta^{Q_2^*}$. We observe that the delta measures are similar for all losses L_i and do not change significantly under the different probability measures. As the delta sensitivity measure quantifies the importance of input factors under a probability measure, having similar values for $\zeta^{\mathbb{P}}$, $\zeta^{Q_1^*}$, and $\zeta^{Q_2^*}$, means that the importance ranking of the L_i 's under different stresses does not change. We also report, in the first two columns of Table 2, the

reverse sensitivity measures with $s(x) = \mathbb{1}_{\{x > \check{F}(0.95)\}}$. The reverse sensitivity measures provide, in contrast to the delta measure, insight into the change in the distributions of the L_i 's from \mathbb{P} and \mathbb{Q}^* .

Table 2. Comparison of different sensitivity measures: First two columns correspond to the reverse sensitivity measures with $s(x) = \mathbb{1}_{\{x > \check{F}(0.95)\}}$ and stressed models \mathbb{Q}_1^* and \mathbb{Q}_2^* , respectively. The last three columns are the delta measure under \mathbb{P} , \mathbb{Q}_1^* and \mathbb{Q}_2^* , respectively.

	$\mathcal{S}_i^{\mathbb{Q}_1^*}$	$\mathcal{S}_i^{\mathbb{Q}_2^*}$	$\zeta^{\mathbb{P}}$	$\zeta^{\mathbb{Q}_1^*}$	$\zeta^{\mathbb{Q}_2^*}$
L_1	0.45	0.68	0.38	0.38	0.38
L_2	0.47	0.62	0.29	0.29	0.29
L_3	0.51	0.57	0.30	0.30	0.29
L_4	0.52	0.63	0.30	0.30	0.29
L_5	0.34	0.58	0.33	0.34	0.33
L_6	0.41	0.62	0.34	0.34	0.32
L_7	0.54	0.72	0.40	0.40	0.38
L_8	0.60	0.69	0.38	0.39	0.39
L_9	0.24	0.66	0.40	0.40	0.38
L_{10}	0.41	0.73	0.39	0.38	0.37

Alternatively to considering the change in the marginal distributions L_i from the baseline to the stressed model, we can study the change in the dependence between the losses when moving from the baseline to a stressed model. For this, we consider the bivariate reverse sensitivity measures for the pairs (L_5, L_{10}) and (L_9, L_{10}) for the second stress \mathbb{Q}_2^* , that is a 1% increase in HARA utility and $ES_{0.8}$, and a 3% increase in $ES_{0.95}$. Specifically, we look at the function $s(L_i, L_j) = \mathbb{1}_{\{L_i > \check{F}_i(0.95)\}} \mathbb{1}_{\{L_j > \check{F}_j(0.95)\}}$, where $\check{F}_i(\cdot)$ and $\check{F}_j(\cdot)$ are the \mathbb{P} -quantile functions of L_i and L_j respectively. This bivariate sensitivity measure quantifies the impact a stress has on the probability of joint exceedances with values $\mathcal{S}_{5,10}^{\mathbb{Q}_2^*} = 0.78$ and $\mathcal{S}_{9,10}^{\mathbb{Q}_2^*} = 0.81$ indicating that the probabilities of joint exceedances increase more for stress 2. This can also be seen in Figure 6 which shows the bivariate copulae contours of (L_5, L_{10}) (top panels) and (L_9, L_{10}) (bottom panels). The left contour plots correspond to the baseline model \mathbb{P} whereas the right panels display the contours under the stress model \mathbb{Q}_2^* (solid lines) together with the baseline contours (reported using partially transparent lines). The red dots are the simulated realisations of the losses (L_5, L_{10}) and (L_9, L_{10}) , respectively (which are the same for the baseline and stressed model). We observe that for both pairs (L_5, L_{10}) and (L_9, L_{10}) the copula under the stressed model admits larger probabilities of joint large events, which is captured by the bivariate reverse sensitivity measure admitting positive values close to 1.

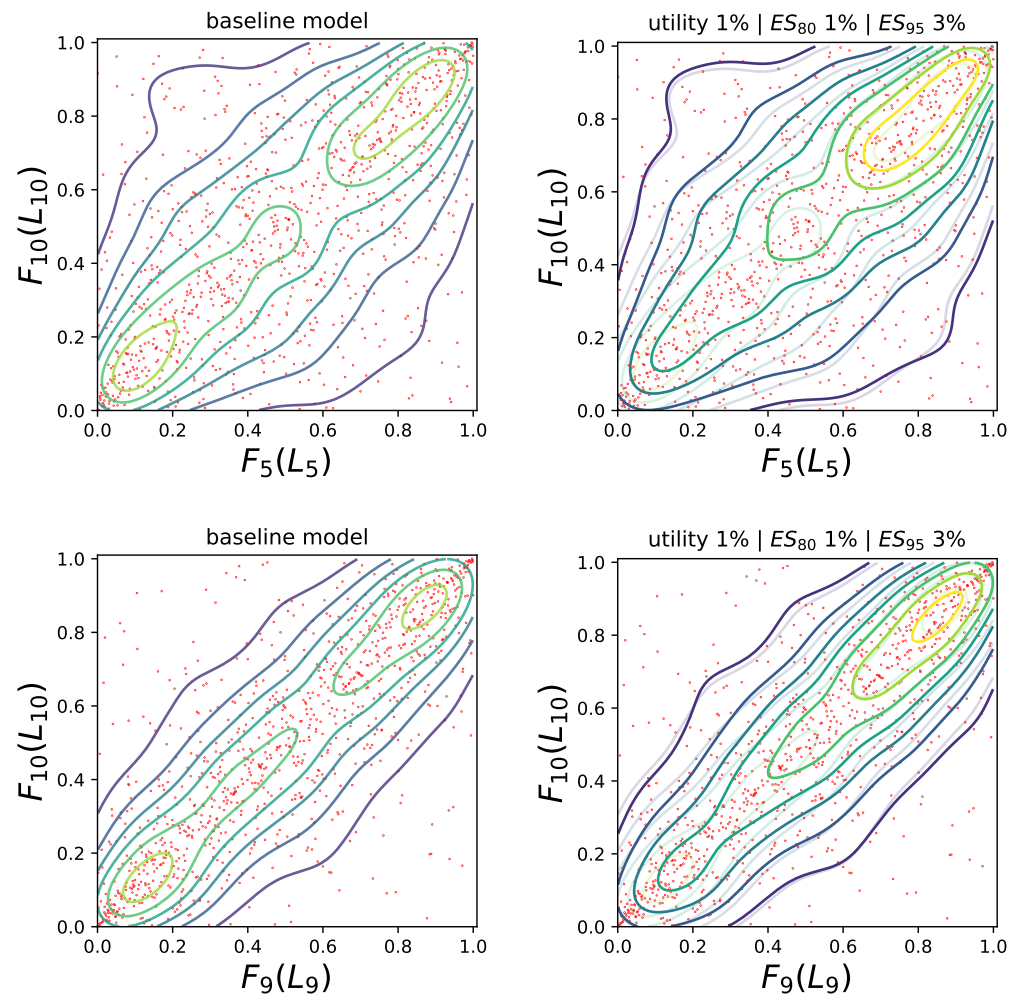


Figure 6. Contour plots of the bivariate copulae of (L_5, L_{10}) (**top** panels) and (L_9, L_{10}) (**bottom** panels) under different models. The left contour plots correspond to the baseline model and the right panels to the stress \mathbb{Q}_2^* (solid lines) with the baseline contours reported using partially transparent lines. Red points are simulated realisations.

6. Concluding Remarks

We extend the reverse sensitivity analysis proposed by [Pesenti et al. \(2019\)](#) which proceeds as follows. Equipped with a baseline model which comprises of input and output random variables and a baseline probability measure, one derives a unique stressed model such that the output (or input) under the stressed model satisfies a prespecified stress and is closest to the baseline distribution. While [Pesenti et al. \(2019\)](#) consider the Kullback–Leibler divergence to measure the difference between the baseline and stressed models we utilise Wasserstein distance of order two. Compared to [Pesenti et al. \(2019\)](#) the Wasserstein distance allows for additional and different stresses on the output including the mean and variance, any distortion risk measure including the Value-at-Risk and Expected-Shortfall, and expected utility type constraints, thus making the reverse sensitivity analysis framework suitable for models used in financial and insurance risk management. We further discuss reverse sensitivity measures which quantify the change the inputs’ distribution when moving from the baseline to a stressed model and illustrate our results on a spatial insurance portfolio application. The reverse sensitivity analysis framework (including the results from this work and from [Pesenti et al. \(2019\)](#)) are implemented in the R package SWIM which is available on CRAN.

Funding: This research was funded by the Connaught Fund, the Canadian Statistical Sciences Institute (CANSSI), and the Natural Sciences and Engineering Research Council of Canada (NSERC) with funding reference numbers DGEGR-2020-00333 and RGPIN-2020-04289.

Data Availability Statement: Not applicable.

Acknowledgments: S.M.P. would like to thank Judy Mao for her help in implementing the numerical examples.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proofs

Proof of Theorem 1. We solve the optimisation on the set of quantile functions $\check{\mathcal{M}}$ and define the Lagrangian with Lagrange multipliers $\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{R}^d$

$$\begin{aligned} L(\check{G}, \lambda) &= \int_0^1 (\check{G}(u) - \check{F}(u))^2 - 2 \sum_{k=1}^d \lambda_k (\check{G}(u) \gamma_k(u) - r_k) \, du \\ &= \int_0^1 \left(\check{G}(u) - \left(\check{F}(u) + \sum_{k=1}^d \lambda_k \gamma_k(u) \right) \right)^2 \\ &\quad - 2 \sum_{k=1}^d \lambda_k (\check{F}(u) \gamma_k(u) - r_k) - \left(\sum_{k=1}^d \lambda_k \gamma_k(u) \right)^2 \, du. \end{aligned}$$

Thus, the optimisation problem (3) is equivalent to first solving, for fixed λ , the optimisation problem

$$\arg \min_{\check{G} \in \check{\mathcal{M}}} L(\check{G}, \lambda) \tag{A1}$$

and then finding λ such that the constraints are fulfilled. For fixed λ , the solution to (A1) is equal to the solution to

$$\arg \min_{\check{G} \in \check{\mathcal{M}}} \int_0^1 \left(\check{G}(u) - \left(\check{F}(u) + \sum_{k=1}^d \lambda_k \gamma_k(u) \right) \right)^2 \, du,$$

which is given by the isotonic projection of $\check{F}(u) + \sum_{k=1}^d \lambda_k \gamma_k(u)$ onto the set $\check{\mathcal{M}}$ and the Lagrange multipliers are such that the constraints are satisfied. Existence of the Lagrange multipliers follows since the set \mathcal{M} is non-empty. Uniqueness follows by convexity of the Wasserstein distance, by convexity of the constraints on the set of quantile functions. \square

Proof of Proposition 1. For coherent distortion risk measures the corresponding weight function γ is non-decreasing. Moreover the optimal quantile function is given by Theorem 1 and is of the form $\check{G}_\lambda(u) = (\check{F}(u) + \lambda \gamma(u))^\uparrow$ for some λ such that \check{G}_λ fulfils the constraint. The choice

$$\lambda^* = \frac{r - \rho_\gamma(F)}{\int_0^1 (\gamma(u))^2 \, du} \geq 0$$

implies that $\check{G}_{\lambda^*}(u) = \check{F}(u) + \lambda^* \gamma(u)$ is a quantile function of the form (4) that fulfils the constraint. By uniqueness of Theorem 1, \check{G}_{λ^*} is indeed the unique solution. \square

Proof of Theorem 2. Since both constraints are convex in \check{G} the Lagrangian with parameters $\lambda = (\lambda_1, \dots, \lambda_d)$ and $\tilde{\lambda} = (\tilde{\lambda}_1, \dots, \tilde{\lambda}_{\bar{d}}) \geq 0$ becomes

$$\begin{aligned} L(\check{G}, \lambda, \tilde{\lambda}) &= \int_0^1 (\check{G}(u) - \check{F}(u))^2 + 2 \sum_{k=1}^d \lambda_k (h_k(u)\check{G}(u) - c_k) du \\ &\quad + \sum_{k=1}^{\bar{d}} \tilde{\lambda}_k (\tilde{h}_k(u)(\check{G}(u))^2 - \tilde{c}_k) du \\ &= \int_0^1 \tilde{\Lambda}(u) \left(\check{G}(u) - \frac{1}{\tilde{\Lambda}(u)} \left(\check{F}(u) - \sum_{k=1}^d \lambda_k h_k(u) \right) \right)^2 \\ &\quad - \frac{1}{\tilde{\Lambda}(u)} \left(\check{F}(u) - \sum_{k=1}^d \lambda_k h_k(u) \right)^2 + (\check{F}(u))^2 - 2 \sum_{k=1}^d \lambda_k c_k - \sum_{k=1}^{\bar{d}} \tilde{\lambda}_k \tilde{c}_k, \end{aligned}$$

where $\tilde{\Lambda}(u) = 1 + \sum_{k=1}^{\bar{d}} \tilde{\lambda}_k \tilde{h}_k(u)$. Since $\tilde{\lambda} \geq 0$ by the KKT-condition, we obtain that $\tilde{\Lambda}(u) \geq 0$ for all $u \in [0, 1]$. Therefore, for fixed $\lambda, \tilde{\lambda}$, using an argument similar to the proof of Theorem 1, the solution (as a function of $\lambda, \tilde{\lambda}$) is given by the weighted isotonic projection of $\frac{1}{\tilde{\Lambda}(u)} (\check{F}(u) - \sum_{k=1}^d \lambda_k h_k(u))$, with weight function $\tilde{\Lambda}(\cdot)$. \square

Proof of Proposition 2. The mean and variance constraint can be rewritten as

$$\begin{aligned} m' &= \int x dG(x) = \int_0^1 \check{G}(u) du \quad \text{and} \\ (\sigma')^2 &= \int (x - m')^2 dG(x) = \int_0^1 (\check{G}(u) - m')^2 du. \end{aligned}$$

Thus, the Lagrangian with Lagrange multipliers $\lambda = (\lambda_1, \dots, \lambda_{k+2})$ is, if $\lambda_2 \neq -1$,

$$\begin{aligned} L(\check{G}, \lambda) &= \int_0^1 (\check{G}(u) - \check{F}(u))^2 du - 2\lambda_1 \left(\int_0^1 \check{G}(u) du - m' \right) \\ &\quad + \lambda_2 \left(\int_0^1 (\check{G}(u) - m')^2 du - (\sigma')^2 \right) \\ &\quad - 2 \sum_{k=1}^d \lambda_{k+2} \left(\int_0^1 \check{G}(u) \gamma_k(u) du - r_k \right) \\ &= (1 + \lambda_2) \int_0^1 \left(\check{G}(u) - \frac{1}{1 + \lambda_2} \left(\check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right) \right)^2 \\ &\quad - \frac{1}{1 + \lambda_2} \left(\check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right)^2 + (\check{F}(u))^2 du \\ &\quad + 2\lambda_1 m' + \lambda_2 \left((m')^2 - (\sigma')^2 \right) + 2 \sum_{k=1}^d \lambda_{k+2} r_k. \end{aligned}$$

For fixed Lagrange multipliers λ with $\lambda_2 \neq -1$, the optimal quantile function is characterised by the isotonic projection and given by (using an analogous argument to the proof of Theorem 1)

$$\begin{aligned}
 \check{G}^*(u) &= \left(\frac{1}{1 + \lambda_2} \left(\check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right) \right)^\uparrow & (A2) \\
 &= \frac{1}{|1 + \lambda_2|} \left(\operatorname{sgn}(1 + \lambda_2) \left(\check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right) \right)^\uparrow \\
 &= \frac{1}{|1 + \lambda_2|} \check{H}(u),
 \end{aligned}$$

where we define $\check{H}(u) = \left(\operatorname{sgn}(1 + \lambda_2) \left(\check{F}(u) + \lambda_1 + \lambda_2 m' + \sum_{k=1}^d \lambda_{k+2} \gamma_k(u) \right) \right)^\uparrow \in \check{\mathcal{M}}$, and $\operatorname{sgn}(\cdot)$ denotes the sign function. Next we show that λ_2 cannot be in a neighbourhood of -1 . It holds that for $\lambda_2 \neq -1$,

$$\int_0^1 (\check{G}^*(u))^2 du = \frac{1}{(1 + \lambda_2)^2} \int_0^1 (\check{H}(u))^2 du. \tag{A3}$$

Since the rhs of (A3) is increasing for $|\lambda_2 + 1| \searrow 0$, there exists a $\varepsilon_0 > 0$ such that for all $\varepsilon < \varepsilon_0$ and $\lambda_2 \in (-1 - \varepsilon, -1 + \varepsilon)$, it holds that

$$\frac{1}{(1 + \lambda_2)^2} \int_0^1 (\check{H}(u))^2 du > (\sigma')^2 + (m')^2,$$

which is a contradiction to the optimality of \check{G}^* . Thus, λ_2 is indeed bounded away from -1 and the unique solution is given in (A2). \square

Proof of Theorem 3. We split this proof into the two cases (i), that is constraint (a) and (ii), i.e., constraint (b).

Case (i): For constraint a), i.e., $\operatorname{VaR}_\alpha(G) = q$, we first assume that $q \leq \operatorname{VaR}_\alpha(F)$ which implies $\check{F}(\alpha_F) = q \leq \check{F}(\alpha)$ and thus $\alpha_F \leq \alpha$. Therefore, $\check{G}^*(u) = \check{F}(u) + (q - \check{F}(u)) \mathbb{1}_{\{u \in (\alpha_F, \alpha]\}}$ is a quantile function which satisfies the constraint. Next, we show that G^* has a smaller Wasserstein distance to F than any other distribution function satisfying the constraint. For this, let \check{H} be a quantile function satisfying the constraint and $\check{H}(u) \neq \check{G}^*(u)$ on a measurable set of non-zero measure. Then

$$\begin{aligned}
 W_2(H, F) &= \int_0^{\alpha_F} (\check{H}(u) - \check{F}(u))^2 du + \int_{\alpha_F}^\alpha (\check{H}(u) - \check{F}(u))^2 du + \int_\alpha^1 (\check{H}(u) - \check{F}(u))^2 du \\
 &\geq \int_{\alpha_F}^\alpha (\check{H}(u) - \check{F}(u))^2 du.
 \end{aligned}$$

By non-decreasingness of \check{H} and \check{F} and by the constraint it holds for all $u \in [\alpha_F, \alpha]$ that $\check{H}(u) \leq \check{H}(\alpha) = q = \check{F}(\alpha_F) \leq \check{F}(u)$. Thus, on the interval $[\alpha_F, \alpha]$, we obtain $(\check{H}(u) - \check{F}(u))^2 \geq (q - \check{F}(u))^2$ and therefore

$$W_2(H, F) \geq \int_{\alpha_F}^\alpha (\check{H}(u) - \check{F}(u))^2 du \geq \int_{\alpha_F}^\alpha (q - \check{F}(u))^2 du = W_2(G^*, F),$$

where at least one inequality is strict since $\check{H}(u) \neq \check{G}^*(u)$ on a measurable set of non-zero measure. Uniqueness follows by the strict convexity of the Wasserstein distance and since the constraint is convex on the set of quantile functions.

Second, we assume that $q > \operatorname{VaR}_\alpha(F)$ and show that there does not exist a solution. Assume by contradiction that \check{G} is an optimal quantile function satisfying the constraint. By definition of α_F , we have that $q = \check{F}(\alpha_F) > \check{F}(\alpha)$ and thus $\alpha_F \geq \alpha$. We apply a similar argument to the first part of the proof using non-decreasingness of \check{G} , $\check{G}(\alpha) = q$, and optimality of \check{G} , to obtain that \check{G} is constant equal to q on $[\alpha, \alpha_F]$ and equal to $\check{F}(u)$ for $u > \alpha_F$. Specifically, it holds that

$$\check{G}(u) = \check{F}(u) + (q - \check{F}(u))\mathbb{1}_{\{u \in (\alpha, \alpha_F]\}}, \quad \text{for all } u > \alpha.$$

Moreover, since the optimal quantile function minimises the Wasserstein distance to F , it holds that, for all $\varepsilon > 0$, \check{G} satisfies

$$\check{G}(u) = \check{F}(u), \quad \text{for all } u \leq \alpha - \varepsilon.$$

Thus, we can define for all $\varepsilon \in (0, \alpha)$ the family of quantile functions

$$\check{H}_\varepsilon(u) = \check{F}(u) + (q - \check{F}(u))\mathbb{1}_{\{u \in (\alpha - \varepsilon, \alpha_F]\}},$$

which satisfies $W_2(H_{\varepsilon_1}, F) < W_2(H_{\varepsilon_2}, F)$ for all $0 \leq \varepsilon_1 < \varepsilon_2$, and $\check{H}_\varepsilon(\alpha) = q$ for all $\varepsilon > 0$. However, $\lim_{\varepsilon \searrow 0} \check{H}_\varepsilon(\alpha) = \check{F}(\alpha) < q$ and thus the quantile function $\lim_{\varepsilon \searrow 0} \check{H}_\varepsilon(u)$ does not fulfil the constraint. Hence, we obtain a contradiction to the optimality of \check{G} .

Case (ii): First, we assume that $q \geq \text{VaR}_\alpha^+(F)$ which implies that $\check{F}(\alpha_F) = q \geq \check{F}^+(\alpha) \geq \check{F}(\alpha)$ and thus $\alpha_F \geq \alpha$. Therefore $\check{G}^*(u) = \check{F}(u) + (q - \check{F}(u))\mathbb{1}_{\{u \in (\alpha, \alpha_F]\}}$ is a quantile function. Moreover, \check{G}^* satisfies the constraint since by right-continuity of \check{G}^* , we have that

$$\check{G}^{*+}(\alpha) = \lim_{\varepsilon \searrow 0} \check{G}^{*+}(\alpha + \varepsilon) = q.$$

The proof that \check{G}^* has the smallest Wasserstein distance to F compared to any other distribution function satisfying the constraint is analogous to the one in case (i).

For the case when $q > \text{VaR}_\alpha^+(F)$, the argument of non-existence of the solution follows using similar arguments as those in case (i). \square

Proof of Theorem 4. By concavity of the utility function, the constraint is convex and can be written as $-\int_0^1 u(\check{G}(v)) dv + c \leq 0$. Thus, we can define the Lagrangian with $\lambda_1 \geq 0$ and $(\lambda_2, \dots, \lambda_{d+1}) \in \mathbb{R}^d$ by

$$\begin{aligned} L(\check{G}, \lambda) &= \frac{1}{2} \int_0^1 (\check{G}(v) - \check{F}(v))^2 - \lambda_1 (u(\check{G}(v)) - c) - \sum_{k=1}^d \lambda_{k+1} (\check{G}(v) \gamma_k(v) - r_k) dv \\ &= \int_0^1 T(\check{G}(v)) - \check{G}(v) \left(\check{F}(v) + \sum_{k=1}^d \lambda_{k+1} \gamma_k(v) \right) \\ &\quad + \frac{1}{2} (\check{F}(v))^2 + \lambda_1 c + \sum_{k=1}^d \lambda_{k+1} r_k dv, \end{aligned}$$

where $T(x) = \frac{1}{2}x^2 - \lambda_1 u(x)$. Therefore, for fixed $\lambda_1, \dots, \lambda_{d+1}$, we apply Theorem 3.1 by Barlow and Brunk (1972) and obtain the unique optimal quantile function (as a function of $\lambda_1, \dots, \lambda_{d+1}$), that is $\check{G}^*(v) = \check{\nu}_{\lambda_1} \left((\check{F}(v) + \sum_{k=1}^d \lambda_{k+1} \gamma_k(v))^\uparrow \right)$, where $\check{\nu}_{\lambda_1}$ is the left-inverse of $\nu_{\lambda_1}(x) = x - \lambda_1 u'(x)$.

Next, we show that if $d = 0$, the utility constraint is binding, that is $\lambda_1 > 0$. For this, assume by contradiction that the $\lambda_1 = 0$, then the optimal quantile function becomes $\check{G}^*(u) = \check{\nu}_0(\check{F}(u))$. Since $\nu_0(x) = x$, we obtain that $\check{G}^*(u) = \check{F}(u)$. \check{F} , however, does not fulfil the constraint, which is a contradiction to the optimality of \check{G}^* . \square

Proof of Proposition 3. We prove the properties one-by-one:

- (i) We first define for a random variable Z with \mathbb{P} -distribution F_Z the random variable $U_Z := F_Z(Z)$. Then, U_Z and Z are comonotonic and U_Z has a uniform distribution under \mathbb{P} . Next, recall that for any random variables Y_1, Y_2 it holds that Rüschemdorf (1983)

$$\mathbb{E} \left[Y_1 F_{Y_2}^{-1}(1 - U_{Y_1}) \right] \leq \mathbb{E}[Y_1 Y_2] \leq \mathbb{E} \left[Y_1 F_{Y_2}^{-1}(U_{Y_1}) \right]. \tag{A4}$$

where $F_{Y_2}^{-1}(U_{Y_1})$ is the random variable that is comonotonic to Y_1 and has the same \mathbb{P} -distribution as Y_2 . Similarly, $F_{Y_2}^{-1}(1 - U_{Y_1})$ is the random variable that is counter-monotonic to Y_1 and has the same \mathbb{P} -distribution as Y_2 . The left (right) inequality in (A4) become equality if and only if the random variables Y_1 and Y_2 are counter-monotonic (comonotonic).

Thus, we can rewrite the maximum in the normalising constant of the reverse sensitivity measure as follows

$$\max_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X)] = \max_{Z \stackrel{\mathbb{P}}{=} \frac{d\mathbb{Q}^*}{d\mathbb{P}}} \mathbb{E}[s(X) Z] = \mathbb{E} \left[s(X) F_{\frac{d\mathbb{Q}^*}{d\mathbb{P}}}^{-1} \left(U_{s(X)} \right) \right]$$

and the minimum in the normalising constant is

$$\min_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[s(X)] = \min_{Z \stackrel{\mathbb{P}}{=} \frac{d\mathbb{Q}^*}{d\mathbb{P}}} \mathbb{E}[s(X) Z] = \mathbb{E} \left[s(X) F_{\frac{d\mathbb{Q}^*}{d\mathbb{P}}}^{-1} \left(1 - U_{s(X)} \right) \right].$$

The reverse sensitivity for the case $\mathbb{E}^{\mathbb{Q}^*}[s(X_i)] \geq \mathbb{E}[s(X_i)]$ then becomes

$$S_i^{\mathbb{Q}^*} = \frac{\mathbb{E}[s(X_i) \frac{d\mathbb{Q}^*}{d\mathbb{P}}] - \mathbb{E}[s(X_i)]}{\mathbb{E} \left[s(X) F_{\frac{d\mathbb{Q}^*}{d\mathbb{P}}}^{-1} \left(U_{s(X)} \right) \right] - \mathbb{E}[s(X_i)]},$$

which satisfies $0 \leq S_i^{\mathbb{Q}^*} \leq 1$ using again (A4). For the case $\mathbb{E}^{\mathbb{Q}^*}[s(X_i)] \leq \mathbb{E}[s(X_i)]$, it holds that

$$S_i^{\mathbb{Q}^*} = - \frac{\mathbb{E}[s(X_i) \frac{d\mathbb{Q}^*}{d\mathbb{P}}] - \mathbb{E}[s(X_i)]}{\mathbb{E} \left[s(X) F_{\frac{d\mathbb{Q}^*}{d\mathbb{P}}}^{-1} \left(1 - U_{s(X)} \right) \right] - \mathbb{E}[s(X_i)]},$$

which satisfies $-1 \leq S_i^{\mathbb{Q}^*} \leq 0$.

- (ii) Assume that $s(X_i)$ and $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ are independent under \mathbb{P} , then

$$\mathbb{E}[s(X_i) \frac{d\mathbb{Q}^*}{d\mathbb{P}}] = \mathbb{E}[s(X_i)] \mathbb{E} \left[\frac{d\mathbb{Q}^*}{d\mathbb{P}} \right] = \mathbb{E}[s(X_i)],$$

and the reverse sensitivity measure is indeed zero.

- (iii) From property (i) we observe that $s(X_i)$ and $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ are comonotonic, if and only if, $S_i^{\mathbb{Q}^*} = 1$ since in this case the right inequality in Equation (A4) becomes equality.
- (iv) From property (i) we observe that $s(X_i)$ and $\frac{d\mathbb{Q}^*}{d\mathbb{P}}$ are counter-comonotonic, if and only if, then $S_i^{\mathbb{Q}^*} = -1$ as in this case left inequality in Equation (A4) becomes equality.

The proof that the joint reverse sensitivity $S_{i,j}^{\mathbb{Q}^*}$ also fulfils the above properties follows using analogous arguments and replacing $s(X_i)$ with $s(X_i, X_j)$. \square

References

- Acerbi, Carlo, and Dirk Tasche. 2002. On the coherence of Expected Shortfall. *Journal of Banking & Finance* 26: 1487–503.
- Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. Coherent measures of risk. *Mathematical Finance* 9: 203–28. [\[CrossRef\]](#)
- Asimit, Vali, Liang Peng, Ruodu Wang, and Alex Yu. 2019. An efficient approach to quantile capital allocation and sensitivity analysis. *Mathematical Finance* 29: 1131–56. [\[CrossRef\]](#)
- Barlow, Richard E., and Hugh D. Brunk. 1972. The isotonic regression problem and its dual. *Journal of the American Statistical Association* 67: 140–47. [\[CrossRef\]](#)
- Barlow, Richard E., David J. Bartholomew, John. M. Bremner, and Hugh D. Brunk. 1972. *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression*. Hoboken: Wiley.
- Bernard, Carole, Silvana M. Pesenti, and Steven Vanduffel. 2020. Robust distortion risk measures. *arXiv arXiv:2205.08850*.

- Blanchet, Jose, and Karthyek Murthy. 2019. Quantifying distributional model risk via optimal transport. *Mathematics of Operations Research* 44: 565–600. [\[CrossRef\]](#)
- Borgonovo, Emanuele. 2007. A new uncertainty importance measure. *Reliability Engineering & System Safety* 92: 771–84.
- Borgonovo, Emanuele, and Elmar Plischke. 2016. Sensitivity analysis: A review of recent advances. *European Journal of Operational Research* 48: 869–87. [\[CrossRef\]](#)
- Borgonovo, Emanuele, Gordon B. Hazen, and Elmar Plischke. 2016. A common rationale for global sensitivity measures and their estimation. *Risk Analysis* 36: 1871–95. [\[CrossRef\]](#)
- Borgonovo, Emanuele, Gordon B. Hazen, Victor Richmond R. Jose, and Elmar Plischke. 2021. Probabilistic sensitivity measures as information value. *European Journal of Operational Research* 289: 595–610. [\[CrossRef\]](#)
- Cambou, Mathieu, and Damir Filipović. 2017. Model uncertainty and scenario aggregation. *Mathematical Finance* 27: 534–67. [\[CrossRef\]](#)
- Cont, Rama, Romain Deguest, and Giacomo Scandolo. 2010. Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance* 10: 593–606. [\[CrossRef\]](#)
- Cuestaalbertos, Juan-Alberto, Ludger Rüschendorf, and Araceli Tuerodiaz. 1993. Optimal coupling of multivariate distributions and stochastic processes. *Journal of Multivariate Analysis* 46: 335–61. [\[CrossRef\]](#)
- Dall’Aglia, Giorgio. 1956. Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze* 10: 35–74.
- De Leeuw, Jan, Kurt Hornik, and Patrick Mair. 2010. Isotone optimization in R: Pool-adjacent-violators algorithm (pava) and active set methods. *Journal of Statistical Software* 32: 1–24. [\[CrossRef\]](#)
- Denuit, Michel, Jan Dhaene, Marc Goovaerts, and Rob Kaas. 2006. *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. Hoboken: John Wiley & Sons.
- Fissler, Tobias, and Silvana M. Pesenti. 2022. Sensitivity measures based on scoring functions. *arXiv* arXiv:2203.00460.
- Fort, Jean-Claude, Thierry Klein, and Agnès Lagnoux. 2021. Global sensitivity analysis and wasserstein spaces. *SIAM/ASA Journal on Uncertainty Quantification* 9: 880–921. [\[CrossRef\]](#)
- Gamboa, Fabrice, Pierre Gremaud, Thierry Klein, and Agnès Lagnoux. 2020. Global sensitivity analysis: A new generation of mighty estimators based on rank statistics. *arXiv* arXiv:2003.01772.
- Gamboa, Fabrice, Thierry Klein, and Agnès Lagnoux. 2018. Sensitivity analysis based on Cramér–von Mises distance. *SIAM/ASA Journal on Uncertainty Quantification* 6: 522–48. [\[CrossRef\]](#)
- Hall, Peter, and Li-Shan Huang. 2001. Nonparametric kernel regression subject to monotonicity constraints. *The Annals of Statistics* 29: 624–47. [\[CrossRef\]](#)
- Kruse, Thomas, Judith C. Schneider, and Nikolaus Schweizer. 2019. The joint impact of f-divergences and reference models on the contents of uncertainty sets. *Operations Research* 67: 428–35. [\[CrossRef\]](#)
- Kusuoka, Shigeo. 2001. On law invariant coherent risk measures. In *Advances in Mathematical Economics*. Berlin and Heidelberg: Springer, pp. 83–95.
- Makam, Vaishno Devi, Pietro Millosovich, and Andreas Tsanakas. 2021. Sensitivity analysis with χ^2 -divergences. *Insurance: Mathematics and Economics* 100: 372–83. [\[CrossRef\]](#)
- Maume-Deschamps, Véronique, and Ibrahima Niang. 2018. Estimation of quantile oriented sensitivity indices. *Statistics & Probability Letters* 134: 122–27.
- Meyer, Mary C. 2008. Inference using shape-restricted regression splines. *The Annals of Applied Statistics* 2: 1013–33. [\[CrossRef\]](#)
- Moosmüller, Caroline, Felix Dietrich, and Ioannis G. Kevrekidis. 2020. A geometric approach to the transport of discontinuous densities. *SIAM/ASA Journal on Uncertainty Quantification* 8: 1012–35. [\[CrossRef\]](#)
- Pesenti, Silvana M., Alberto Bettini, Pietro Millosovich, and Andreas Tsanakas. 2021. Scenario weights for importance measurement (SWIM)—An R package for sensitivity analysis. *Annals of Actuarial Science* 15: 458–83. [\[CrossRef\]](#)
- Pesenti, Silvana M., Pietro Millosovich, and Andreas Tsanakas. 2019. Reverse sensitivity testing: What does it take to break the model? *European Journal of Operational Research* 274: 654–70. [\[CrossRef\]](#)
- Pesenti, Silvana M., Pietro Millosovich, and Andreas Tsanakas. 2021. Cascade sensitivity measures. *Risk Analysis* 31: 2392–414. [\[CrossRef\]](#)
- Plischke, Elmar, and Emanuele Borgonovo. 2019. Copula theory and probabilistic sensitivity analysis: Is there a connection? *European Journal of Operational Research* 277: 1046–59. [\[CrossRef\]](#)
- Rahman, Sharif. 2016. The f-sensitivity index. *SIAM/ASA Journal on Uncertainty Quantification* 4: 130–62. [\[CrossRef\]](#)
- Rüschendorf, Ludger. 1983. Solution of a statistical optimization problem by rearrangement methods. *Metrika* 30: 55–61. [\[CrossRef\]](#)
- Saltelli, Andrea, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. 2008. *Global Sensitivity Analysis: The Primer*. Hoboken: John Wiley & Sons.
- Sysoev, Oleg, and Oleg Burdakov. 2019. A smoothed monotonic regression via l2 regularization. *Knowledge and Information Systems* 59: 197–218. [\[CrossRef\]](#)
- Tsanakas, Andreas, and Pietro Millosovich. 2016. Sensitivity analysis using risk measures. *Risk Analysis* 36: 30–48. [\[CrossRef\]](#) [\[PubMed\]](#)
- Villani, Cédric. 2008. *Optimal transport: Old and New*. Berlin and Heidelberg: Springer Science & Business Media, vol. 338.