Elicitability and identifiability of tail risk measures

Tobias Fissler* Fangda Liu† Ruodu Wang‡ Linxiao Wei§

October 16, 2025

Abstract

Tail risk measures are fully determined by the distribution of the underlying loss beyond its quantile at a certain level, with Value-at-Risk, Expected Shortfall and Range Value-at-Risk being prime examples. They are induced by law-based risk measures, called their generators, evaluated on the tail distribution. This paper establishes joint identifiability and elicitability results of tail risk measures together with the corresponding quantile, provided that their generators are identifiable and elicitable, respectively. As an example, we establish the joint identifiability and elicitability of the tail expectile together with the quantile. The corresponding consistent scores constitute a novel class of weighted scores, nesting the known class of scores of Fissler and Ziegel for the Expected Shortfall together with the quantile. For statistical purposes, our results pave the way to easier model fitting for tail risk measures via regression and the generalized method of moments, but also model comparison and model validation in terms of established backtesting procedures.

Key-words: Tail risk measures, elicitability, identifiability, weighted scores, backtesting

MSC Classification: 91G70; 62F07; 62G32; 62P05

1 Introduction

Over the past three decades, tail risk measures, such as Value-at-Risk (VaR) and Expected Shortfall (ES), have been playing a prominent role as the standard risk metrics in global banking and insurance regulatory frameworks, such as Basel III/IV (BCBS, 2019) and Solvency II (EIOPA, 2011). Within the general framework of tail risk measures developed by Liu and Wang

---

*RiskLab, Department of Mathematics, ETH Zurich, Switzerland. ✉ `tobias.fissler@math.ethz.ch`

†Department of Statistics and Actuarial Science, University of Waterloo, Canada. ✉ `fangda.liu@uwaterloo.ca`

‡Department of Statistics and Actuarial Science, University of Waterloo, Canada. ✉ `wang@uwaterloo.ca`

§College of Science, Wuhan University of Technology, China. ✉ `lxwei@whut.edu.cn`

(2021) and Liu et al. (2022), there are many tail risk measures in the literature, such as the tail standard deviation (Furman and Landsman, 2006), the tail entropic risk measure (Tsanakas, 2009), the Glue VaR (Belles-Sampera et al., 2014), the Gini Shortfall (Furman et al., 2017), and the Range VaR (Cont et al., 2010; Embrechts et al., 2018), in addition to VaR and ES.

In view of the practical importance of tail risk measures, this paper studies their identifiability and elicitability. A risk measure is identifiable if it is the unique zero (root) of an expected identification function. Similarly, it is elicitable if it is the unique minimizer of an expected score; see Definitions 3.1 and 3.2. As such, identifiability and elicitability facilitate Z- and M-estimation (Huber and Ronchetti, 2009). More generally, model fitting can be pursued via the generalized method of moments (Newey and McFadden, 1994) or regression (Dimitriadis et al., 2024), also using recent tools from machine learning; see Fissler et al. (2023) and Cont et al. (2025) for examples. For model comparison and model validation, commonly subsumed under the umbrella term *backtesting*, identifiability and elicitability are crucial (Nolde and Ziegel, 2017; Fissler et al., 2016; Wang et al., 2025).

Elicitability, together with the closely related concept of identifiability, has received increasing attention in the statistics and risk management literature. In particular, characterization results on risk measures are obtained via elicitability or convex level sets; see Weber (2006), Bellini and Bignozzi (2015) and Delbaen et al. (2016) for convex risk measures, Ziegel (2016) for coherent risk measures, Kou and Peng (2016) and Wang and Ziegel (2015) for distortion risk measures, Wang and Wei (2020) for non-monotone risk measures, and Liu and Wang (2021) for tail risk measures. Elicitability can also be used indirectly to identify ES among coherent risk measures, as done by Embrechts et al. (2021).

A tail risk measure is determined by the tail of the loss distribution and a generating risk measure, called the generator. Our paper is motivated by the intriguing observation that the mean is elicitable, and so is the vector of VaR and ES at the same level (Fissler and Ziegel, 2016), where ES is the tail risk measure generated by the mean. This naturally leads to the question whether the elicitability of a tail risk measure is connected to the elicitability of its generator. This problem, as well as its variants, will be systemically studied in this paper.

Our main results, Theorems 4.3 and 5.3, establish the joint identifiability and elicitability of a tail risk measure together with the corresponding quantile, subject to identifiability and elicitability of its generator, respectively. For the elicitability result, a novel monotonicity condition on the consistent scoring function for the generator is required. Interestingly, the consistent scores of the latter theorem constitute a generalization of the weighted scores discussed by Holzmann and Klar (2017), but with a forecast dependent weight function. This class nests also the well known FZ-scores for ES and the quantile, characterized by Fissler and Ziegel (2016). Vice versa, Propositions 4.1 and 5.1 show that the identifiability and elicitability of the generator is necessary for the joint identifiability and elicitability results, respectively.

The paper is structured as follows. Section 2 collects preliminaries on tail risk measures. Section 3 defines the key concepts of elicitability, identifiability, and convex level sets, as well as some examples. Section 4 contains results on identifiability, and Section 5 contains results on elicitability. Some remarks on our results when changing from the right tail to the left tail, or to a body part of the distribution, are presented in Section 6. Section 7 concludes the paper. Some background on risk measures, technical assumptions, and omitted proofs are collected in three appendices.

2 Preliminaries

2.1 Notation

Let $\mathcal{M}^0$ be the set of all Borel probability distributions on $\mathbb{R}$ and let $\mathcal{M}^q \subseteq \mathcal{M}^0$, $q \in [1, \infty)$, be the subset of all distributions with a finite qth moment. Moreover, $\mathcal{M}^\infty$ denotes the set of all compactly supported distributions. For the ease of presentation, we identify distributions with the corresponding cumulative distribution functions; that is, $F(x)$ is understood as $F((-\infty, x])$ for $x \in \mathbb{R}$. For $F \in \mathcal{M}^0$, we define the left-continuous generalized inverse (left-quantile) as

$$F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geqslant t\}, \quad t \in (0, 1],$$

and in addition set $F^{-1}(0) = \inf\{x \in \mathbb{R} : F(x) > 0\}$. For a function F on a subset of $\mathbb{R}$, its left and right limits at x are given by $F(x-) = \lim_{y \uparrow x} F(y)$ and $F(x+) = \lim_{y \downarrow x} F(y)$, assuming that the limits are well-defined. For $x \in \mathbb{R}$, denote by δ_x the point-mass probability measure at x. For simplicity, $\int$ means integration on $\mathbb{R}$ unless otherwise specified. Throughout this article, if we do not specify $\mathcal{M}$, then the statements hold for any set $\mathcal{M} \subseteq \mathcal{M}^0$ such that the functional at consideration is well-defined and finite on $\mathcal{M}$.

2.2 Risk measures

A risk measure is a mapping $\rho\colon \mathcal{M} \to [-\infty, \infty]$, where the set $\mathcal{M} \subseteq \mathcal{M}^0$ is the domain of ρ. The typical interpretation of a risk measure is that its argument, $F \in \mathcal{M}$, represents the distribution of a random loss from some financial position, and that $\rho(F)$ represents the amount of capital required to hold to make the financial position acceptable. In the literature, risk measures are often defined as mappings from a set of random variables, instead of $\mathcal{M}$, to the extended real line. When such risk measures are law-based, they one-to-one correspond to risk measures under our definition. We explain in Appendix A details on risk measures and their desirable properties.

The two most popular classes of risk measures used in banking and insurance practice are the Value-at-Risk (VaR) and the Expected Shortfall (ES). For a confidence level p, the right-quantile (VaR_p^+) and the left-quantile (VaR_p^-), are defined for $F \in \mathcal{M}^0$ as

$$\begin{aligned}\mathrm{VaR}_p^+(F) &= \inf\{x \in \mathbb{R} : F(x) > p\} = F^{-1}(p+), \quad p \in [0,1);\\ \mathrm{VaR}_p^-(F) &= \inf\{x \in \mathbb{R} : F(x) \geqslant p\} = F^{-1}(p), \quad p \in (0,1].\end{aligned}$$

In addition, let $\text{ess-sup}(F) = \mathrm{VaR}_1^-(F)$ and $\text{ess-inf}(F) = \mathrm{VaR}_0^+(F)$ for $F \in \mathcal{M}^0$. In risk management practice, one typically does not distinguish between VaR_p^+ and VaR_p^- as they are identical for distributions with a continuous F^{-1} at p. However, to allow for the full generality of exposition, for $p \in (0,1)$, we shall mostly work with the interval-valued p-quantile

$$Q_p(F) = \{x \in \mathbb{R} : F(x-) \leqslant p \leqslant F(x)\} = \left[\mathrm{VaR}_p^-(F), \mathrm{VaR}_p^+(F)\right].$$

For a confidence level $p \in (0,1)$, ES_p is defined by

$$\mathrm{ES}_p(F) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_r^+(F)\,\mathrm{d}r \in \mathbb{R} \cup \{\infty\}, \quad F \in \mathcal{M}^0, \tag{2.1}$$

$\mathrm{ES}_1(F) = \mathrm{VaR}_1^-(F) = \text{ess-sup}(F) \in \mathbb{R} \cup \{\infty\}$ and $\mathrm{ES}_0(F)$ is the mean on $\mathcal{M}^1$. We remark that for $p \in (0,1)$ and $F \in \mathcal{M}^0$, $\mathrm{ES}_p(F)$ is finite if and only if $\int_0^\infty x\,\mathrm{d}F(x) < \infty$. Since $\mathrm{VaR}_r^-(F) \neq \mathrm{VaR}_r^+(F)$ only for at most countably many $r \in [0,1]$, we can also replace $\mathrm{VaR}_r^+(F)$ by $\mathrm{VaR}_r^-(F)$ in (2.1).

As a compromise between the robustness of VaR and and the coherence of ES, Cont et al.

(2010) introduced the so called Range Value-at-Risk (RVaR). For $0 < p < q < 1$, it is defined as

$$\mathrm{RVaR}_{p,q}(F) = \frac{1}{q-p}\int_p^q \mathrm{VaR}_r^+(F)\,\mathrm{d}r \in \mathbb{R}, \quad F \in \mathcal{M}^0. \tag{2.2}$$

Clearly, for the boundary cases of $q = 1$ or $p = q$, by taking limits, $\mathrm{RVaR}_{p,q}$ coincides with ES_p or VaR_p^+, respectively.

2.3 Tail distribution and tail risk measures

We follow the same definitions of tail distributions as in Rockafellar and Uryasev (2002). For a distribution $F \in \mathcal{M}^0$ and $p \in (0,1)$, let $F_p \in \mathcal{M}^0$ be the *tail distribution of* F beyond its p-quantile, that is,

$$F_p(x) = \frac{(F(x)-p)_+}{1-p}, \quad x \in \mathbb{R}.$$

Invoking our sign convention that a distribution $F \in \mathcal{M}^0$ is a loss distribution, it is clear that the right tail is the region of interest from a risk measurement and management perspective (McNeil et al., 2015). Clearly, with another sign convention and from a mathematical perspective, we could similarly consider the left tail of the distribution, and all the results presented in the article hold *mutatis mutandis* (this will be made more clear in Section 6). We impose the tacit condition on our generic classes $\mathcal{M} \subseteq \mathcal{M}^0$ that $F_p \in \mathcal{M}$ for each $F \in \mathcal{M}$ and $p \in (0,1)$. This assumption holds for common choices of $\mathcal{M}$, such as $\mathcal{M} = \mathcal{M}^q$, $q \in [1,\infty]$.[1] Note that $F_p = G_p$ if and only if $F(x) = G(x)$ for all $x \geqslant \mathrm{VaR}_p^+(F)$.

The following definition corresponds to Definition 3.1 of Liu and Wang (2021).

Definition 2.1 (Tail risk measures)**.** For $p \in (0,1)$, a risk measure $\rho\colon \mathcal{M} \to \mathbb{R}$ is a *p-tail risk measure* if $\rho(F) = \rho(G)$ for all $F, G \in \mathcal{M}$ satisfying $F_p = G_p$. A risk measure $\rho\colon \mathcal{M} \to \mathbb{R}$ is a *tail risk measure* if it is a p-tail risk measure for some $p \in (0,1)$.

It is immediate from Definition 2.1 that for $p, q \in (0,1)$ where $p < q$, VaR_p^+, ES_p and $\mathrm{RVaR}_{p,q}$ are r-tail risk measures for $r \in (0,p]$, but not for $r \in (p,1)$; VaR_p^- is an r-tail risk measure only for $r \in (0,p)$. The expectation is not a tail risk measure. An important property of tail risk measures is that they can be generated by other risk measures. Precisely, for any $p \in (0,1)$ and p-tail risk measure ρ on $\mathcal{M}$, there exists a risk measure ρ^* on $\mathcal{M}$, called the

---

[1] A notable exception, however, is $\mathcal{M} \setminus \mathcal{M}^\infty$.

p-generator of ρ, satisfying

$$\rho(F) = \rho^*(F_p), \qquad \text{for all } F \in \mathcal{M}, \tag{2.3}$$

and such a ρ^* is unique on the set $\mathcal{M}^*$ of distributions in $\mathcal{M}$ whose support is bounded from below (Proposition 1 of Liu and Wang, 2021). Conversely, to obtain a p-tail risk measure, $p \in (0,1)$, it suffices to specify a generic risk measure ρ^* and to define ρ via (2.3). We will define (ρ, ρ^*) as a pair, and study their joint properties.

Definition 2.2. For $p \in (0,1)$, a pair of risk measures (ρ, ρ^*) is called a *p-tail pair* on $\mathcal{M}$ if $\rho(F) = \rho^*(F_p)$ for all $F \in \mathcal{M}$, and (ρ, ρ^*) is called a *tail pair* on $\mathcal{M}$ if it is a p-tail pair on $\mathcal{M}$ for some $p \in (0,1)$.

A slightly different approach of measuring tail risk is through the loss part of the random variables instead of the part above a quantile level; see Jarrow (2002) and Cont et al. (2013). We focus on the formulation in Definition 2.1 as it includes the most popular risk measures VaR and ES and many other examples of interest.

3 Elicitability, identifiability and convex level sets

In the literature, the general discussion about elicitability and its importance due to its connection to comparative backtests and forecasts can be found in Lambert et al. (2008), Gneiting (2011a), Fissler and Ziegel (2016), and the references therein. In particular, Nolde and Ziegel (2017) elaborated in detail on its relevance in backtesting, and they provide calibration tests exploiting the notion of identifiability. More recently, Liu and Wang (2021) provided a characterization of VaR among tail risk measures based on elicitability. Moreover, both identifiability and elicitability can be used in estimation, and particularly elicitability also in regression (Huber and Ronchetti, 2009; Koenker, 2005; Dimitriadis et al., 2024).

We want to define the notions of identifiability, elicitability and two convex level sets (CxLS) properties in a unified manner, that is applicable for real-valued risk measures, interval-valued quantiles, and vectors consisting of (possibly multiple) real-valued risk measures and an interval-valued quantile. To this end, we study set-valued functionals $T\colon \mathcal{M} \to \mathcal{P}(\mathbb{R}^k)$, where $\mathcal{P}(\mathbb{R}^k)$ denotes the power set of $\mathbb{R}^k$. Moreover, we use the tacit convention to identify $x \in \mathbb{R}^k$ with the singleton $\{x\} \in \mathcal{P}(\mathbb{R}^k)$, such that the definitions apply also to $\mathbb{R}^k$-valued functionals.

Definition 3.1. For a functional $T\colon \mathcal{M} \to \mathcal{P}(\mathbb{R}^k)$, a function $V\colon \mathbb{R}^k \times \mathbb{R} \to \mathbb{R}^k$ is an $\mathcal{M}$-*identification function* for T if for all $F \in \mathcal{M}$

$$T(F) \subseteq \left\{ x \in \mathbb{R}^k : \int_{\mathbb{R}} V(x,y)\,\mathrm{d}F(y) = 0 \right\}, \tag{3.1}$$

assuming that the integral is well-defined. If (3.1) holds with an equality, then V is a *strict* $\mathcal{M}$-*identification function* for T. If T has a strict $\mathcal{M}$-identification function, we call it $\mathcal{M}$-*identifiable*.

Definition 3.2. (i) For a functional $T\colon \mathcal{M} \to \mathcal{P}(\mathbb{R}^k)$, a function $S\colon \mathbb{R}^k \times \mathbb{R} \to \mathbb{R}$ is an $\mathcal{M}$-*consistent score* if for all $F \in \mathcal{M}$

$$T(F) \subseteq \operatorname*{arg\,min}_{x \in \mathbb{R}^k} \int_{\mathbb{R}} S(x,y)\,\mathrm{d}F(y), \tag{3.2}$$

assuming that the integral is well-defined. If (3.2) holds with an equality, then S is a *strictly* $\mathcal{M}$-*consistent score* for T. If T has a strictly $\mathcal{M}$-consistent score, we call it $\mathcal{M}$-*elicitable*.

(ii) A functional $T\colon \mathcal{M} \to \mathcal{P}(\mathbb{R}^k)$, is $\mathcal{M}$-*conditionally elicitable* with a functional $T'\colon \mathcal{M} \to \mathcal{P}(\mathbb{R}^{k'})$ if T' is $\mathcal{M}$-elicitable and T is $\mathcal{M}(r, T')$-elicitable for any $r \in \mathbb{R}^{k'}$, where

$$\mathcal{M}(r, T') = \{F \in \mathcal{M} : r \in T'(F)\}.$$

The notion of conditional elicitability simplifies the original definition given in Emmer et al. (2015); see also Fissler and Hoga (2024).

Example 3.3. The mean functional is $\mathcal{M}^1$-identifiable with the strict $\mathcal{M}^1$-identification function $V(x,y) = x - y$, $x, y \in \mathbb{R}$. Under mild conditions, any other strict $\mathcal{M}^1$-identification function is of the form $h(x)(x-y)$, where $h \neq 0$ (Dimitriadis et al., 2024). Moreover, under similar conditions, any strictly $\mathcal{M}^1$-consistent scoring function for the mean is given by

$$S(x,y) = -\phi(x) + \phi'(x)(x-y) + a(y), \quad x, y \in \mathbb{R},$$

where $\phi\colon \mathbb{R} \to \mathbb{R}$ is strictly convex with subgradient ϕ', and $a\colon \mathbb{R} \to \mathbb{R}$ is an $\mathcal{M}$-integrable function (Gneiting, 2011a). This class nests the ubiquitous squared loss $S(x,y) = (x-y)^2$, which is strictly $\mathcal{M}^2$-consistent for the mean.

Example 3.4. For $p \in (0,1)$, the p-quantile Q_p is identifiable on

$$\mathcal{M}_{(p)} = \{F \in \mathcal{M}^0 : F(\mathrm{VaR}_p^-(F)) = p\}$$

with the strict $\mathcal{M}_{(p)}$-identification function $V(x,y) = \mathbb{1}_{\{y \leqslant x\}} - p,\ x, y \in \mathbb{R}$. Essentially, all other strict $\mathcal{M}_{(p)}$-identification functions are given by $h(x)(\mathbb{1}_{\{y \leqslant x\}} - p)$, where $h \neq 0$; see Dimitriadis et al. (2024) for details. Here, "essentially" means that for any other strict $\mathcal{M}_{(p)}$-identification function $\tilde{V}(x,y)$ it holds that $\int \tilde{V}(x,y)\,\mathrm{d}F(y) = h(x)(F(x) - p)$ for all $x \in \mathbb{R}$ and $F \in \mathcal{M}_{(p)}$. Interestingly, Q_p is elicitable on the entire $\mathcal{M}^0$. Subject to integrability and richness conditions, any strictly $\mathcal{M}^0$-consistent scoring function is given by

$$S(x,y) = \mathbb{1}_{\{y>x\}} g(y) + \big(\mathbb{1}_{\{y \leqslant x\}} - p\big) g(x) + a(y), \quad x, y \in \mathbb{R},$$

where g is strictly increasing (Gneiting, 2011b). This recovers the well-known pinball loss $S(x,y) = \big(\mathbb{1}_{\{y \leqslant x\}} - p\big)(x - y)$, which is strictly $\mathcal{M}^1$-consistent for Q_p.

Example 3.5. It is known that ES_p generally fails to be identifiable and elicitable on sufficiently rich classes $\mathcal{M}$ (Weber, 2006; Gneiting, 2011a). However, the pair (Q_p, ES_p) turns out to be $\mathcal{M}^1 \cap \mathcal{M}_{(p)}$-identifiable using the strict identification function

$$V(v,x,y) = \begin{pmatrix} \mathbb{1}_{\{y \leqslant v\}} - p \\ x - \frac{1}{1-p}\big[\mathbb{1}_{\{y>v\}} y + (\mathbb{1}_{\{y \leqslant v\}} - p) v\big] \end{pmatrix}, \quad v, x, y \in \mathbb{R}, \tag{3.3}$$

see Dimitriadis et al. (2024) for the full class of identification functions. Acerbi and Szèkely (2014) established the $\mathcal{M}$-elicitability of (Q_p, ES_p) under certain restrictive conditions on $\mathcal{M}$; and Fissler and Ziegel (2016, 2021b) showed that the pair is generally $\mathcal{M}^1$-elicitable. Strictly $\mathcal{M}^1$-consistent scoring functions are of the form

$$\begin{aligned} S(v,x,y) = {} & \mathbb{1}_{\{y>v\}} g(y) + \big(\mathbb{1}_{\{y \leqslant v\}} - p\big)\, g(v) \\ & + \phi'(x)\Big(x - \tfrac{1}{1-p}\big[\mathbb{1}_{\{y>v\}} y + (\mathbb{1}_{\{y \leqslant v\}} - p) v\big]\Big) - \phi(x) + a(y), \quad v, x, y \in \mathbb{R}, \end{aligned} \tag{3.4}$$

where $\phi\colon \mathbb{R} \to \mathbb{R}$ is strictly convex with subgradient ϕ', such that for all x the function $v \mapsto g(v) - \frac{1}{1-p}\phi'(x) v$ is strictly increasing, and where $a\colon \mathbb{R} \to \mathbb{R}$ is a function such that the expectation $\int S(v,x,y)\,\mathrm{d}F(y)$ is finite for all $F \in \mathcal{M}^1$.

Example 3.6. Similarly to ES_p, Fissler and Ziegel (2021b) showed that $\mathrm{RVaR}_{p,q}$, $0 < p < q < 1$, is not identifiable or elicitable on sufficiently rich classes $\mathcal{M}$. In the same vein as before, they established the $\mathcal{M}_{(p)} \cap \mathcal{M}_{(q)}$-identifiability of the triplet $(Q_p, Q_q, \mathrm{RVaR}_{p,q})$ via the identification function

$$V(v_1, v_2, x, y) = \begin{pmatrix} \mathbb{1}_{\{y \leqslant v_1\}} - p \\ \mathbb{1}_{\{y \leqslant v_2\}} - q \\ x - \frac{1}{q-p}\Big[\mathbb{1}_{\{v_1 < y \leqslant v_2\}} y + (\mathbb{1}_{\{y \leqslant v_1\}} - p)v_1 - (\mathbb{1}_{\{y \leqslant v_2\}} - q)v_2\Big] \end{pmatrix} \tag{3.5}$$

for $v_1, v_2, x, y \in \mathbb{R}$. Moreover, strictly $\mathcal{M}^0$-consistent scoring functions for this triplet are of the form

$$\begin{aligned} &S(v_1, v_2, x, y) \\ &\;= \mathbb{1}_{\{y > v_1\}} g_1(y) + \big(\mathbb{1}_{\{y \leqslant v_1\}} - p\big)\, g_1(v_1) + \mathbb{1}_{\{y > v_2\}} g_2(y) + \big(\mathbb{1}_{\{y \leqslant v_2\}} - q\big)\, g_2(v_2) \\ &\quad + \phi'(x)\Big(x - \tfrac{1}{q-p}\Big[\mathbb{1}_{\{v_1 < y \leqslant v_2\}} y + (\mathbb{1}_{\{y \leqslant v_1\}} - p)v_1 - (\mathbb{1}_{\{y \leqslant v_2\}} - q)v_2\Big]\Big) - \phi(x) + a(y) \end{aligned} \tag{3.6}$$

for $v_1, v_2, x, y \in \mathbb{R}$, where $\phi\colon \mathbb{R} \to \mathbb{R}$ is strictly convex with subgradient ϕ', such that for all x the functions $v_1 \mapsto g_1(v_1) - \frac{1}{q-p}\phi'(x)v_1$ and $v_2 \mapsto g_2(v_2) + \frac{1}{q-p}\phi'(x)v_2$ are strictly increasing, and where $a\colon \mathbb{R} \to \mathbb{R}$ is a function such that the expectation $\int S(v_1, v_2, x, y)\,\mathrm{d}F(y)$ is finite for all $F \in \mathcal{M}^0$.

We close this section by providing two versions of the CxLS property, defined in Fissler et al. (2021).

Definition 3.7. (i) A functional $T\colon \mathcal{M} \to \mathcal{P}(\mathbb{R}^k)$ satisfies the CxLS property on $\mathcal{M}$ if for all $F_0, F_1 \in \mathcal{M}$ and for all $\lambda \in (0,1)$ such that $(1-\lambda)F_0 + \lambda F_1 \in \mathcal{M}$ it holds that

$$T(F_0) \cap T(F_1) \subseteq T((1-\lambda)F_0 + \lambda F_1).$$

(ii) A functional $T\colon \mathcal{M} \to \mathcal{P}(\mathbb{R}^k)$ satisfies the CxLS* property on $\mathcal{M}$ if for all $F_0, F_1 \in \mathcal{M}$ and for all $\lambda \in (0,1)$ such that $(1-\lambda)F_0 + \lambda F_1 \in \mathcal{M}$ it holds that

$$T(F_0) \cap T(F_1) \neq \emptyset \implies T(F_0) \cap T(F_1) = T((1-\lambda)F_0 + \lambda F_1).$$

Note that Fissler et al. (2021) coined these two properties the *selective* CxLS and CxLS* properties. Since we do not make use of the counterpart – the *exhaustive* notion – we omit the qualifier "selective" in this paper.

Obviously, the CxLS* property implies the CxLS property. They both generalize the CxLS property for singleton-valued functionals, for which they coincide. Importantly for us, elicitability implies the CxLS* property (Fissler et al., 2021, Proposition 3.4), and similarly, identifiability implies the CxLS property (Proposition B.4 of the preprint version of Fissler and Hoga, 2024). This necessity has already partially been established by Osband (1985) and Gneiting (2011a). Under additional regularity conditions and for real-valued functionals, Steinwart et al. (2014) also showed the sufficiency of the CxLS property for elicitability and identifiability.

4 Identifiability results

We first define the following sets:

$$\begin{aligned}
\mathcal{M}_p^{\mathrm{c}} &= \left\{F \in \mathcal{M}^0 : F^{-1} \text{ is continuous at } p\right\}, \quad \text{for } p \in (0,1);\\
\mathcal{M}_{\geqslant r} &= \left\{F \in \mathcal{M}^0 : F^{-1}(0) \geqslant r\right\}, \quad \text{for } r \in \mathbb{R};\\
\mathcal{M}_{(p)} &= \left\{F \in \mathcal{M}^0 : F(F^{-1}(p)) = p\right\}, \quad \text{for } p \in [0,1).
\end{aligned}$$

The first result in this section discusses how the identifiability of the tail risk measure implies that of the corresponding generator.

Proposition 4.1. *Let (ρ, ρ^*) be a p-tail pair for some $p \in (0,1)$. The following statements hold.*

(i) For any $r \in \mathbb{R}$ it holds that if ρ is $\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}$-identifiable with a strict identification function V, then ρ^* is $\mathcal{M}_{\geqslant r} \cap \mathcal{M}$-identifiable with the strict identification function

$$V_r^*(x,y) = (1-p)V(x,y) + pV(x,r), \quad x, y \in \mathbb{R}. \tag{4.1}$$

(ii) For any $r \in \mathbb{R}$ it holds that if (Q_p, ρ) is $\mathcal{M}_{(p)} \cap \big(\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}\big)$-identifiable with a strict identification function

$$\left(\mathbb{1}_{\{y \leqslant v\}} - p, V(v,x,y)\right), \quad v, x, y \in \mathbb{R},$$

for some function $V : \mathbb{R}^3 \to \mathbb{R}$, then ρ^* is $\mathcal{M}_{\geqslant r} \cap \mathcal{M}_{(0)} \cap \mathcal{M}$-identifiable with the strict identification function

$$V_r^*(x,y) = (1-p)V(r,x,y) + pV(r,x,r), \quad x, y \in \mathbb{R}. \tag{4.2}$$

Proof. (i) Let V be a strict $\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}$-identification function for ρ and V_r^* as given in (4.1) for some $r \in \mathbb{R}$. Choose some $G \in \mathcal{M}_{\geqslant r} \cap \mathcal{M}$ and define

$$F(y) = \mathbb{1}_{\{y \geqslant r\}}\big((1-p)G(y) + p\big) = (1-p)G(y) + p\mathbb{1}_{\{y \geqslant r\}}.$$

Due to the assumed identifiability of ρ and since $F_p = G$, we get

$$\begin{aligned}
\rho^*(G) = \rho^*(F_p) = \rho(F) &= \left\{x \in \mathbb{R} : \int V(x,y)\,\mathrm{d}F(y) = 0\right\} \\
&= \left\{x \in \mathbb{R} : \int V(x,y)\,\mathrm{d}\left(\mathbb{1}_{\{y \geqslant r\}}(1-p)G(y)\right) + pV(x,r) = 0\right\} \\
&= \left\{x \in \mathbb{R} : \int \big((1-p)V(x,y) + pV(x,r)\big)\,\mathrm{d}G(y) = 0\right\} \\
&= \left\{x \in \mathbb{R} : \int V_r^*(x,y)\,\mathrm{d}G(y) = 0\right\},
\end{aligned}$$

where $V_r^*(x,y)$ is defined in (4.1). This shows the claim.

(ii) Let $\big(\mathbb{1}_{\{y \leqslant v\}} - p, V(v,x,y)\big)$, $v,x,y \in \mathbb{R}$, be a strict $\mathcal{M}_{(p)} \cap \big(\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}\big)$-identification function for (Q_p, ρ) and V_r^* given in (4.2) for some $r \in \mathbb{R}$. For any $F \in \mathcal{M}_{(p)} \cap \big(\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}\big)$ and $v \in Q_p(F)$ it holds that

$$\int V(v,x,y)\,\mathrm{d}F(y) = 0 \quad \text{if and only if} \quad x = \rho(F).$$

Let $G \in \mathcal{M}_{\geqslant r} \cap \mathcal{M}_{(0)} \cap \mathcal{M}$. Then, as in part (i), define $F(y) = \mathbb{1}_{\{y \geqslant r\}}\big((1-p)G(y) + p\big)$, resulting in $F \in \mathcal{M}_{(p)} \cap \big(\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}\big)$ and $r = \mathrm{VaR}_p^-(F) \in Q_p(F)$. Again, $F_p = G$, which implies

$$\begin{aligned}
\rho^*(G) = \rho^*(F_p) = \rho(F) &= \left\{x \in \mathbb{R} : \int V(r,x,y)\,\mathrm{d}F(y) = 0\right\} \\
&= \left\{x \in \mathbb{R} : \int V(r,x,y)\,\mathrm{d}\left(\mathbb{1}_{\{x \geqslant r\}}(1-p)G(x)\right) + pV(r,x,r) = 0\right\} \\
&= \left\{x \in \mathbb{R} : \int \big((1-p)V(r,x,y) + pV(r,x,r)\big)\,\mathrm{d}G(x) = 0\right\},
\end{aligned}$$

which shows the claim. □

Example 4.2. We illustrate the part (i) of Proposition 4.1. Let $\mathcal{M}$ be the class of continuous distribution functions which are strictly increasing on their support. For any $\alpha \in (0,1)$, $p \in (0,\alpha]$ and $r \in \mathbb{R}$, VaR_α^- is $\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}$-identifiable with the strict identification function $V(x,y) = \mathbb{1}_{\{y\leqslant x\}} - \alpha$. Moreover, VaR_α^- is a p-tail risk measure with generator $\mathrm{VaR}_{\alpha^*}^-$, where $\alpha^* = (\alpha - p)/(1-p)$. Then, (4.1) yields $V_r^*(x,y) = (1-p)\big(\mathbb{1}_{\{y\leqslant x\}} - \alpha\big) + p\big(\mathbb{1}_{\{y\leqslant r\}} - \alpha\big)$. Indeed, for any $F \in \mathcal{M}_{\geqslant r} \cap \mathcal{M}$ it holds that

$$\int V_r^*(x,y)\,\mathrm{d}F(y) = (1-p)(F(x)-\alpha) + p(1-\alpha) = 0 \quad \text{if and only if} \quad F(x) = \frac{\alpha - p}{1-p} = \alpha^*,$$

showing that V_r^* is a strict $\mathcal{M}_{\geqslant r}\cap\mathcal{M}$-identification function for $\mathrm{VaR}_{\alpha^*}^-$.

The next result addresses the inverse direction of Proposition 4.1, that is, how the identifiability of ρ^* implies that of ρ, together with the quantile interval.

Theorem 4.3. *Let (ρ,ρ^*) be a p-tail pair for some $p\in(0,1)$. If $\rho^*\colon \mathcal{M}\to\mathbb{R}$ is $\mathcal{M}$-identifiable with the strict identification function $V^*\colon \mathbb{R}\times\mathbb{R}\to\mathbb{R}$, then (Q_p,ρ) is $\mathcal{M}_{(p)}\cap\mathcal{M}$-identifiable with the strict $\mathcal{M}_{(p)}\cap\mathcal{M}$-identification function*

$$V(v,x,y) = \begin{pmatrix} \mathbb{1}_{\{y\leqslant v\}} - p \\ \mathbb{1}_{\{y>v\}} V^*(x,y) + (\mathbb{1}_{\{y\leqslant v\}} - p)V^*(x,v) \end{pmatrix}, \quad v,x,y\in\mathbb{R}. \tag{4.3}$$

Proof. Let $F \in \mathcal{M}_{(p)}\cap\mathcal{M}$. Clearly, $\int \big(\mathbb{1}_{\{y\leqslant v\}} - p\big)\,\mathrm{d}F(y) = 0$ if and only if $v\in Q_p(F)$. Take $v \in Q_p(F)$. Since $F_p(y) = \frac{(F(y)-p)^+}{1-p}$, we have

$$\begin{aligned}\int \big(\mathbb{1}_{\{y>v\}}V^*(x,y) + (\mathbb{1}_{\{y\leqslant v\}} - p)V^*(x,v)\big)\,\mathrm{d}F(y) &= \int \mathbb{1}_{\{y>v\}}V^*(x,y)\,\mathrm{d}F(y) \\ &= (1-p)\int V^*(x,y)\,\mathrm{d}F_p(y).\end{aligned}$$

Hence, $\int V^*(x,y)\,\mathrm{d}F_p(y) = 0$ if and only if $x = \rho^*(F_p) = \rho(F)$. This shows the claim. □

Remark 4.4. Theorem 4.3 holds also by using

$$V(v,x,y) = \begin{pmatrix} \mathbb{1}_{\{y\leqslant v\}} - p \\ \mathbb{1}_{\{y>v\}} V^*(x,y) \end{pmatrix}, \quad v,x,y\in\mathbb{R} \tag{4.4}$$

in (4.3). The reason for the seemingly more involved representation in (4.3) including the "correction term" $(\mathbb{1}_{\{y\leqslant v\}} - p)V^*(x,v)$ is the structural similarity to the form of the consistent scoring functions in Theorem 5.3; see the discussion in Remark 5.4.

Example 4.5. To illustrate Theorem 4.3, we let ρ^* be the expectation with a strict $\mathcal{M}^1$-identification function $V^*(x,y) = x - y$, $x,y\in\mathbb{R}$. The corresponding p-tail risk measure ρ

induced by ρ^* is ES_p. Indeed, a straightforward calculation yields that the second component of (4.3) is $(1-p)$ times the second component of (3.3).

Example 4.6. We know that $\mathrm{RVaR}_{p,q}$, $0<p<q<1$, is an r-tail risk measure for $r \in (0,p]$. For $r<p$, the generator ρ^* is again an RVaR. For $r=p$, the generator ρ^* is a lower expected shortfall. In both cases, ρ^* fails to be identifiable on reasonably large classes $\mathcal{M}$. As a consequence of Proposition 4.1, $(Q_r, \mathrm{RVaR}_{p,q})$, $0<r\leqslant p<q<1$ fails to be identifiable on reasonably large classes.

Corollary 4.7. *Let (ρ,ρ^*) be a p-tail pair for some $p\in(0,1)$. If $\rho^*\colon \mathcal{M}\to\mathbb{R}$ is $\mathcal{M}$-identifiable, then (Q_p,ρ) satisfies the CxLS* property on $\mathcal{M}_{(p)}\cap\mathcal{M}$.*

Proof. Due to Proposition B.4 of the preprint version of Fissler and Hoga (2024), Theorem 4.3 implies that (Q_p,ρ) satisfies the CxLS property on $\mathcal{M}_{(p)}\cap\mathcal{M}$. Since Q_p is $\mathcal{M}_{(p)}$-elicitable, it satisfies the CxLS* property on $\mathcal{M}_{(p)}$. And since ρ can be identified with a singleton-valued functional, (Q_p,ρ) satisfies the CxLS* property on $\mathcal{M}_{(p)}\cap\mathcal{M}$. ☐

Remark 4.8. Due to Corollary 10 of Steinwart et al. (2014), under some continuity assumptions on ρ^* and richness assumptions on $\mathcal{M}$, the risk measure ρ^* is $\mathcal{M}$-identifiable if and only if it satisfies the CxLS property on $\mathcal{M}$, which is in turn equivalent to its $\mathcal{M}$-elicitability.

5 Elicitability results

5.1 Elicitability relations and score functions

Corollary 4.7 and Remark 4.8 establish that for a p-tail pair (ρ,ρ^*) with an $\mathcal{M}$-elicitable ρ^*, the pair (Q_p,ρ) satisfies the CxLS* property on $\mathcal{M}_{(p)}\cap\mathcal{M}$, which is an important necessary condition for the elicitability of (Q_p,ρ). We emphasize that for multivariate functionals, the CxLS* property fails to be a sufficient condition for elicitability, as shown by Fissler et al. (2021) and Fissler and Hoga (2024). This section first establishes the conditional $\mathcal{M}$-elicitability of (Q_p,ρ) and then provides a simple sufficient condition for its $\mathcal{M}$-elicitability. The result below establishes the strictly consistent score function for ρ^* from that of ρ or (Q_p,ρ), if it exists, respectively.

Proposition 5.1. *Let (ρ,ρ^*) be a p-tail pair for some $p\in(0,1)$. The following statements hold.*

(i) For any $r \in \mathbb{R}$ it holds that if ρ is $\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}$-elicitable with a strictly consistent score $S(x, y)$, $x, y \in \mathbb{R}$, then ρ^* is $\mathcal{M}_{\geqslant r} \cap \mathcal{M}$-elicitable with the strictly consistent score

$$S_r^*(x, y) = (1-p)S(x, y) + pS(x, r), \quad x, y \in \mathbb{R}. \tag{5.1}$$

(ii) For any $r \in \mathbb{R}$ it holds that if (Q_p, ρ) is $\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}$-elicitable with a strictly consistent score $S(v, x, y)$, $v, x, y \in \mathbb{R}$, then ρ^* is $\mathcal{M}_{\geqslant r} \cap \mathcal{M}$-elicitable with the strictly consistent score

$$S_r^*(x, y) = (1-p)S(r, x, y) + pS(r, x, r), \quad x, y \in \mathbb{R}. \tag{5.2}$$

The proof is very similar to that of Proposition 4.1 and is presented in Appendix C.

Example 5.2 (Continuation of Example 4.2)**.** We now illustrate part (i) of Proposition 5.1. A strictly $\mathcal{M}^0$-consistent score for VaR_α^- is of the form $S(x, y) = \big(\mathbb{1}_{\{y \leqslant x\}} - \alpha\big)\big(g(x) - g(y)\big)$ for a strictly increasing and bounded function g (Gneiting, 2011b) (upon restricting the class $\mathcal{M} \subseteq \mathcal{M}^0$, we can also choose unbounded g). For any $p \in (0, \alpha]$ and $r \in \mathbb{R}$ it holds for any random variable Y with distribution $F \in \mathcal{M}_{\geqslant r}$ that the score S_r^* in (5.1) is almost surely

$$S_r^*(x, Y) = (1-p)\big(\mathbb{1}_{\{Y \leqslant x\}} - \alpha\big)\big(g(x) - g(Y)\big) + p\big(1-\alpha\big)\big(g(x) - g(Y)\big), \quad x \in \mathbb{R}.$$

A direct calculation verifies that this is a positive multiple of

$$\big(\mathbb{1}_{\{Y \leqslant x\}} - \alpha^*\big)\big(g(x) - g(Y)\big), \quad \alpha^* = \frac{\alpha - p}{1-p}.$$

The next result provides the score function for the tail risk measure ρ from that of ρ^* under some conditions. A list of technical assumptions adapted from those used by Fissler and Ziegel (2016) is needed for part (iii). Due to its length, this list is presented in Appendix B.

Theorem 5.3. *Let (ρ, ρ^*) be a p-tail pair for some $p \in (0, 1)$. If $\rho^* : \mathcal{M} \to \mathbb{R}$ is $\mathcal{M}$-elicitable with the strictly $\mathcal{M}$-consistent score $S^* : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Then the following statements hold.*

(i) ρ is $\mathcal{M}$-conditionally elicitable with Q_p. In particular, for any $v \in \mathbb{R}$, the score

$$S_v(x, y) = \mathbb{1}_{\{y > v\}} S^*(x, y) + \big(\mathbb{1}_{\{y \leqslant v\}} - p\big) S^*(x, v) + a(y), \quad x, y \in \mathbb{R}, \tag{5.3}$$

is strictly $\mathcal{M}(v, Q_p)$-consistent for ρ, where $\mathcal{M}(v, Q_p) = \{F \in \mathcal{M} : v \in Q_p(F)\}$.

(ii) If for all $x \in \mathbb{R}$, $S^*(x, y)$ is strictly increasing in y, then (Q_p, ρ) is $\mathcal{M}$-elicitable with a strictly $\mathcal{M}$-consistent score

$$S(v, x, y) = \mathbb{1}_{\{y>v\}} S^*(x, y) + \big(\mathbb{1}_{\{y\leqslant v\}} - p\big) S^*(x, v) + a(y), \quad v, x, y \in \mathbb{R}, \tag{5.4}$$

where $a\colon \mathbb{R} \to \mathbb{R}$ is some $\mathcal{M}$-integrable function.

(iii) Suppose $\mathcal{M} \subseteq \mathcal{M}_p^{\mathrm{c}} \cap \mathcal{M}_{(p)}$ and Assumption B.1 in Appendix B holds. Any (strictly) $\mathcal{M}$-consistent score $S\colon \mathbb{R}^3 \to \mathbb{R}$ for (Q_p, ρ) is necessarily of the form (5.4) for almost every $(v, x, y) \in \mathrm{int}(\mathsf{A}) \times \mathbb{R}$, where $\mathsf{A} = \{(\mathrm{VaR}_p^-(F), \rho(F)) : F \in \mathcal{M}\} \subseteq \mathbb{R}^2$ and $S^*\colon \mathbb{R}^2 \to \mathbb{R}$ is a (strictly) $\mathcal{M}$-consistent score for ρ^* such that $y \mapsto S^*(x, y)$ is (strictly) increasing for $x \in \mathbb{R}$.

Proof. For the part (i), let $F \in \mathcal{M}$ and $v \in Q_p(F)$. For S_v in (5.3), it holds that

$$\int S_v(x, y)\,\mathrm{d}F(y) = \int S^*(x, y)\,\mathrm{d}\big(F(y) - p\big)_+ + \int a(y)\,\mathrm{d}F(y).$$

Hence

$$\operatorname*{arg\,min}_{x\in\mathbb{R}} \int S_v(x, y)\,\mathrm{d}F(y) = \operatorname*{arg\,min}_{x\in\mathbb{R}} \int S^*(x, y)\,\mathrm{d}F_p(y) = \rho^*(F_p) = \rho(F),$$

which shows the claim.

For the part (ii), observe that, since $S^*(x, y)$ is strictly increasing in y for all $x \in \mathbb{R}$, the function $(v, y) \mapsto S(v, x, y)$ given in (5.4) is a generalized piecewise linear loss, which is strictly $\mathcal{M}$-consistent for Q_p, see Gneiting (2011b). Hence, for all $x \in \mathbb{R}$,

$$Q_p(F) = \operatorname*{arg\,min}_{v\in\mathbb{R}} \int S(v, x, y)\,\mathrm{d}F(y).$$

The rest follows from part (i). The proof of part (iii) can be found in Appendix C. ☐

Remark 5.4. It should be pointed out that for the elicitability result of Theorem 5.3 part (ii), we do not need the restriction to $\mathcal{M}_{(p)} \cap \mathcal{M}$ as for the corresponding identifiability result of Theorem 4.3. The reason for this is that, on the one hand, the p-quantile Q_p is elicitable on the entire $\mathcal{M}^0$ whereas it is identifiable only on $\mathcal{M}_{(p)}$. On the other hand, the presence of the "correction term"

$(\mathbb{1}_{\{y \leqslant v\}} - p)S^*(x,y)$ in (5.4) is crucial. In fact, part (i) of Theorem 5.3 would hold on $\mathcal{M} \cap \mathcal{M}_{(p)}$ without this correction term. That is, for any $v \in \mathbb{R}$, the score $S_v(x,y) = \mathbb{1}_{\{y>v\}}S^*(x,y)$ is strictly consistent for ρ on $\big\{F \in \mathcal{M} \cap \mathcal{M}_{(p)} : v \in Q_p(F)\big\}$. For part (ii), the correction term is also needed to render the score $(v,y) \mapsto \mathbb{1}_{\{y>v\}}S^*(x,y) + (\mathbb{1}_{\{y \leqslant v\}} - p)S^*(x,v)$ a strictly $\mathcal{M}$-consistent score for Q_p for any x. Interestingly, the score $S_v(x,y) = \mathbb{1}_{\{y>v\}}S^*(x,y)$ constitutes a weighted score in the sense of Holzmann and Klar (2017) – see also Gneiting and Ranjan (2011) – with weight function $w(y) = \mathbb{1}_{\{y>v\}}$. However, for the score in (5.4), the weight function obtains a variable threshold, given in terms of a quantile-forecast v. This extends the existing theory substantially, and it only works with the additional correction term and the monotonicity requirement on S^*.

Recall that if $S^*(x,y)$ is (strictly) $\mathcal{M}$-consistent for ρ^* and if $g\colon \mathbb{R} \to \mathbb{R}$ is $\mathcal{M}$-integrable, then $\tilde{S}^*$ given by

$$\tilde{S}^*(x,y) = S^*(x,y) + g(y)$$

is again (strictly) $\mathcal{M}$-consistent for ρ^*. Following Dawid (1998), we call S^* and $\tilde{S}^*$ *strongly equivalent*. Even though they coincide in terms of (strict) consistency for ρ^*, they can be clearly different in terms of their monotonicity behavior in y. The following lemma explores conditions for the existence of a strongly equivalent version of the score which is strictly increasing in its second argument.

Lemma 5.5. *Suppose that $S^*\colon \mathbb{R}^2 \to \mathbb{R}$ is partially differentiable with respect to its second argument. Then there is a differentiable function $g\colon \mathbb{R} \to \mathbb{R}$ such that $y \mapsto S^*(x,y) + g(y)$ is increasing for all x if and only if there is a function $h\colon \mathbb{R} \to \mathbb{R}$ such that*

$$h(y) \leqslant \partial_y S^*(x,y) \qquad \text{for all } x, y \in \mathbb{R}. \tag{5.5}$$

Proof. If (5.5) holds, then we can choose g to be an antiderivative of $-h$. Moreover, if g is an antiderivative of $-h$ plus a strictly increasing function, then $S^*(x,y) + g(y)$ is even strictly increasing in y. On the other hand, if $S^*(x,y) + g(y)$ is increasing in y, then (5.5) holds with $h = -g'$. □

In practice, the forecast for Q_p might rather play an auxiliary role in comparison to the forecast of the p-tail risk measure ρ. The following proposition establishes two relevant notions of *order-sensitivity* for strictly consistent scoring functions of the form (5.4). For a general discussion of order-sensitivity, we refer to Fissler and Ziegel (2019).

Proposition 5.6. *For the strictly $\mathcal{M}$-consistent score in* (5.4) *and for any $F \in \mathcal{M}$, if $v_2 < v_1 \leqslant \mathrm{VaR}_p^-(F)$ or $\mathrm{VaR}_p^+(F) \leqslant v_1 < v_2$, then*

(i) $\int S(v_1, x, y)\,\mathrm{d}F(y) < \int S(v_2, x, y)\,\mathrm{d}F(y)$ *for all* $x \in \mathbb{R}$;

(ii) $\min_{x\in\mathbb{R}} \int S(v_1, x, y)\,\mathrm{d}F(y) < \min_{x\in\mathbb{R}} \int S(v_2, x, y)\,\mathrm{d}F(y)$.

Proof. Part (i) follows from the fact that for any fixed $x \in \mathbb{R}$, the score $(v, y) \mapsto S(v, x, y)$ is strictly $\mathcal{M}$-consistent for Q_p. This readily implies the claimed order-sensitivity by Nau (1985, Proposition 3) and Bellini and Bignozzi (2015, Proposition 3.4). For part (ii) suppose that $x_1 \in \arg\min_{x\in\mathbb{R}} \int S(v_1, x, y)\,\mathrm{d}F(y)$ and $x_2 \in \arg\min_{x\in\mathbb{R}} \int S(v_2, x, y)\,\mathrm{d}F(y)$. Then $\int S(v_2, x_2, y)\,\mathrm{d}F(y) > \int S(v_1, x_2, y)\,\mathrm{d}F(y) \geqslant \int S(v_1, x_1, y)\,\mathrm{d}F(y)$, where the first inequality is due to part (i). $\square$

5.2 Special cases and discussions

We first provide the simple example of the pair of the expectation and ES.

Example 5.7. Take the p-tail pair $(\rho, \rho^*) = (\mathrm{ES}_p, \mathbb{E})$. If we choose the strictly $\mathcal{M}^1$-consistent score $S^*(x, y) = \frac{1}{1-p}\big(\phi'(x)(x-y) - \phi(x)\big) + g(y)$ with ϕ strictly convex, then the score in (5.4) coincides with the one in (3.4). The condition that $y \mapsto S^*(x, y)$ is strictly increasing is equivalent to the condition that $y \mapsto g(y) - \frac{1}{1-p}\phi'(x)y$ is strictly increasing. Invoking Lemma 5.5, a necessary and sufficient condition for the existence of such a g is that ϕ' is bounded from above.

Let us consider the τ-expectile ρ_τ, $\tau \in (0, 1)$. On the class $\mathcal{M}^1$ of distributions with finite mean, Newey and Powell (1987) introduced the τ-expectile as the unique solution $x = \rho_\tau(F)$ to the equation

$$\tau \int_x^\infty (y - x)\,\mathrm{d}F(y) = (1-\tau) \int_{-\infty}^x (x - y)\,\mathrm{d}F(y).$$

For $\tau = 1/2$, this family contains the usual expectation. Bellini et al. (2014) showed that for $\tau \geqslant 1/2$, the τ-expectile constitutes a coherent risk measure. Ziegel (2016) established that these are the only elicitable law-based coherent risk measures. The following elicitability result about induced p-tail risk measures is novel to the literature.

Proposition 5.8. *Let (ρ, ρ^*) be a p-tail pair and ρ^* the τ-expectile, $\tau \in (0, 1)$. The pair (Q_p, ρ) is $\mathcal{M}^1$-elicitable.*

Proof. Subject to mild regularity and integrability conditions, any strictly $\mathcal{M}^1$-consistent score for ρ^* is of the form

$$S^*(x,y) = |\mathbb{1}_{\{y\leqslant x\}} - \tau|\big(\phi(y) - \phi(x) + \phi'(x)(x-y)\big) + g(y),$$

where ϕ is strictly convex with subgradient ϕ' and g is arbitrary (Gneiting, 2011a, Theorem 10). For $g = 0$ and $y \neq x$, the partial derivative is

$$\partial_y S^*(x,y) = |\mathbb{1}_{\{y\leqslant x\}} - \tau|\big(\phi'(y) - \phi'(x)\big).$$

For $y = x$, the left-sided and right-sided partial derivative with respect to y exist, coincide and are both 0. Take ϕ such that $|\phi'| < C$ for some $C > 0$, for example we can use $\phi(x) = x^2/(1+|x|)$, then $|\partial_y S^*(x,y)| < 2C$. Therefore, we can apply Theorem 5.3 and Lemma 5.5 to construct a strictly $\mathcal{M}^1$-consistent score for the p-tail-τ-expectile and Q_p. □

Proposition 5.9. *Let $\rho^*\colon \mathcal{M} \to \mathbb{R}$ be the ratio of expectations, that is,*

$$\rho^*(F) = \frac{\int u(y)\,\mathrm{d}F(y)}{\int t(y)\,\mathrm{d}F(y)}$$

for two functions $u, t\colon \mathbb{R} \to \mathbb{R}$, where $\mathcal{M}$ is chosen such that ρ^* is well-defined and finite. Let (ρ, ρ^*) be a p-tail pair. If u and t are differentiable, then the pair (Q_p, ρ) is $\mathcal{M}$-elicitable.

Proof. Subject to mild regularity and integrability conditions, any strictly $\mathcal{M}$-consistent score for ρ^* is of the form

$$\begin{aligned} S^*(x,y) &= -\phi(x)t(y) + \phi'(x)\big(xt(y) - u(y)\big) + g(y) \\ &= t(y)\big(\phi'(x)x - \phi(x)\big) - t(y)\phi'(x) + g(y). \end{aligned}$$

where ϕ is strictly convex with subgradient ϕ' and g is arbitrary (Gneiting, 2011a, Theorem 8). For the strictly convex function $\phi(x) = x^2/(1+|x|)$, we obtain that $\phi'(x) \in [-1,1]$ and $\phi'(x)x - \phi(x) \in [0,1]$ for all $x \in \mathbb{R}$. If u and t are differentiable, then $|\partial_y S^*(x,y)| \leqslant |u'(y)| + |t'(y)|$ such that we can apply Lemma 5.5 and Theorem 5.3. □

Remark 5.10. We can further relax the differentiability condition on u and t in Proposition 5.9. Suppose that both u and t have a finite total variation on any compact interval, and we use the notation

$$\|u\|(y) = \mathrm{sign}(y) \sup_{P \in \mathcal{P}(y)} \sum_{i=0}^{n_P - 1} |u(z_{i+1}) - u(z_i)|, \qquad y \in \mathbb{R},$$

where $\mathcal{P}(y)$ denotes the set of all partitions $P = \{z_0, z_1, \ldots, z_{n_P} : z_i < z_{i+1}\}$ of the interval $[\min(0,y), \max(0,y)]$, then we obtain a strictly consistent score S in (5.4) upon choosing $g(y) = \|u\|(y) + \|t\|(y) + y$, $y \in \mathbb{R}$. For example, if $u = 1$ and $t(y) = y^2$, then $\|u\|(y) = 0$ and $\|t\|(y) = \mathrm{sign}(y)y^2$ for $y \in \mathbb{R}$. If $u(y) = \mathbb{1}_{[a,b)}(y)$ for $a < b$, then $\|u\|(y) = \mathbb{1}_{[a,\infty)}(y) + \mathbb{1}_{[b,\infty)}(y)$ for $y \in \mathbb{R}$.

Example 5.11. Consider ρ^* to be a shortfall risk measure induced by the loss function $\ell \colon \mathbb{R} \to \mathbb{R}$, which is left-continuous, non-decreasing and $\inf_{x\in\mathbb{R}} \ell(x) < 0 < \sup_{x\in\mathbb{R}} \ell(x)$. Then

$$\rho^*(F) = \inf\left\{ m \in \mathbb{R} : \int \ell(y - m)\, \mathrm{d}F(y) \leqslant 0 \right\}.$$

Further, suppose that we consider a class $\mathcal{M}$ of distributions such that for all $F \in \mathcal{M}$ and for all $m \in \mathbb{R}$

$$m = \rho^*(F) \quad \text{if and only if} \quad \int \ell(y - m)\, \mathrm{d}F(y) = 0.$$

This means that ρ^* is $\mathcal{M}$-identifiable with the strict $\mathcal{M}$-identification function $V^*(m, y) = \ell(y - m)$. We can apply Osband's principle (Fissler and Ziegel, 2016, Theorem 3.2 and Corollary 3.3) to see that – subject to regularity conditions – any strictly $\mathcal{M}$-consistent score S^* for ρ^* is of the form

$$S^*(x, y) = \int_0^x h(z)\ell(y - z)\, \mathrm{d}z + g(y)$$

for some non-positive function $h \colon \mathbb{R} \to \mathbb{R}$. The question as to what choices of h and g lead to a score which is strictly increasing in its second argument is hard to answer in full generality. We only provide a sufficient condition: If ℓ is bounded from below, then we can find some function g such that

$$S^*(x, y) = -\int_0^x \ell(y - z)\, \mathrm{d}z + g(y)$$

is increasing. Indeed, $\partial_y\big[S^*(x,y) - g(y)\big] = \ell(y - x) - \ell(y)$. Hence, by Lemma 5.5 and Theorem 5.3, the pair (Q_p, ρ) is $\mathcal{M}$-elicitable. We emphasize that the boundedness of ℓ is not necessary for the elicitability of (Q_p, ρ) as the case of $\rho^* = \mathbb{E}$ with $\ell(x) = x$ shows (or more generally the τ-expectile; see Proposition 5.8).

Remark 5.12. The CoVaR (Adrian and Brunnermeier, 2016) has become an important systemic risk measure. In its definition according to Girardi and Tolga Ergün (2013) and Nolde and Zhang (2020), it closely resembles a p-tail risk measure. For $\alpha \in (0,1)$ and $\beta \in [0,1)$, the $\mathrm{CoVaR}_{\alpha|\beta}$ of a two-dimensional random vector (X,Y) with joint distribution $F_{X,Y}$ is defined as $\mathrm{CoVaR}_{\alpha|\beta}(F_{X,Y}) = \mathrm{VaR}^-_\alpha(F_{Y|X \geqslant \mathrm{VaR}^-_\beta(F_X)})$, where $F_{Y|X \geqslant \mathrm{VaR}^-_\beta(F_X)}$ is the conditional distribution of Y on $\{X \geqslant \mathrm{VaR}^-_\beta(F_X)\}$. We assume that the marginal distributions are continuous and strictly increasing such that the distinction between the lower and upper quantile is inessential. Fissler and Hoga (2024) showed that $F_{X,Y} \mapsto (\mathrm{CoVaR}_{\alpha|\beta}(F_{X,Y}), \mathrm{VaR}_\beta(F_X))$ generally fails to be elicitable (even though it is identifiable and thus has convex level sets), where F_X is the distribution of X. The reason why it is not possible to leverage the construction principle of Theorem 5.3 is the fact that for CoVaR, the observation process is bivariate.

Example 5.13. Similarly to the discussion in Example 4.6, Proposition 5.1 implies that on reasonably large classes $\mathcal{M}$, the pair $(Q_r, \mathrm{RVaR}_{p,q})$, $0 < r \leqslant p < q < 1$, fails to be elicitable.

5.3 Elicitation of the tail distribution

In a risk management context, but also in other statistical contexts, one could be interested in the *entire* tail distribution F_p for some $F \in \mathcal{M}^0$, $p \in (0,1)$, instead of a given tail risk measure.

Consistent scoring functions are tailored to evaluate the accuracy of a point forecast for a certain functional of interest. Their counterpart for probabilisitic forecasts, specifying the entire distribution, are *proper scoring rules* (Gneiting and Raftery, 2007). Mathematically, a scoring rule is a map $s\colon \mathcal{M} \times \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ for some $\mathcal{M} \subseteq \mathcal{M}^0$. It is $\mathcal{M}$-*proper* if for all $F \in \mathcal{M}$

$$F \in \operatorname*{arg\,min}_{G \in \mathcal{M}} \int s(G,y) \,\mathrm{d}F(y).$$

It is *strictly* $\mathcal{M}$-*proper* if F is the unique element of the above argmin. In a similar vein, we define that s is $\mathcal{M}$-proper for the p-tail if for all $F \in \mathcal{M}$

$$\{H \in \mathcal{M} : H_p = F_p\} \subseteq \operatorname*{arg\,min}_{G \in \mathcal{M}} \int s(G,y) \,\mathrm{d}F(y).$$

Similarly, s is strictly $\mathcal{M}$-proper for the p-tail if we have equality above. Tail scores have been hinted at in the discussion at the end of Subsection 2.3 in Holzmann and Klar (2017).

The most prominent proper scoring rules are the log-score, $\mathrm{LogScore}(f,y) = -\log(f(y))$, where f is a predictive density, and the continuous ranked probability score (CRPS),

$$\mathrm{CRPS}(F, y) = \int \left(F(z) - \mathbb{1}_{\{y \leqslant z\}}\right)^2 \mathrm{d}z.$$

The CRPS is strictly $\mathcal{M}^1$-proper. Moreover, one can check that $\partial_y \mathrm{CRPS}(F, y) = 2F(y) - 1$ for all $F \in \mathcal{M}^1$. Hence, invoking Lemma 5.5 and using a similar construction as in Theorem 5.3 (ii), the map

$$S(v, F, y) = \mathbb{1}_{\{y > v\}}\big(\mathrm{CRPS}(F, y) + 2y\big) + \big(\mathbb{1}_{\{y \leqslant z\}} - p\big)\big(\mathrm{CRPS}(F, v) + 2v\big), \quad v, y \in \mathbb{R},\ F \in \mathcal{M}^1$$

is strictly $\mathcal{M}$-proper for $F \mapsto (Q_p(F), F_p)$.

While, to the best of our knowledge, the above score is novel to the statistical literature, an alternative strictly $\mathcal{M}^1$-proper score for the p-tail has been known in the form of the quantile weighted CRPS (Gneiting and Ranjan, 2011)

$$\mathrm{CRPS}_p(F, y) = \int_p^1 2\big(\mathbb{1}_{\{y \leqslant F^{-1}(r)\}} - r\big)\big(F^{-1}(r) - r\big)\,\mathrm{d}r.$$

This result hinges on the fact that for any $F \in \mathcal{M}^0$ and $p \in (0, 1)$, we have a bijection between F_p and $\big(F^{-1}(r)\big)_{r \in (p,1]}$.

6 Results on the left tail or the body of the distribution

We have been working with tail risk measures defined for the right tail distribution. Mathematically, it is also possible to translate all our results to the left tail distribution. This may be useful when the positions of gains and losses are switched, as in, e.g., the convention of Föllmer and Schied (2016), or when the best scenarios of random outcomes are of interest. We only mention a few useful points when converting to the left tail.

For $q \in (0, 1]$, we can define the left tail distribution of $F \in \mathcal{M}^0$ as

$$F^q(x) = \frac{\min(F(x), q)}{q}, \quad x \in \mathbb{R}.$$

To formulate the corresponding tail risk measure, for a risk measure $\rho^* \colon \mathcal{M} \to \mathbb{R}$ and a class $\mathcal{M}$

such that $F^q \in \mathcal{M}$ for each $F \in \mathcal{M}$, we can define

$$\rho^q(F) = \rho^*(F^q).$$

The pair (ρ^q, ρ^*) is called *left q-tail pair*. With the above definitions, we can reproduce our main results as well. We only state the most interesting and relevant counterpart, namely the one of Theorem 5.3 part (ii).

Theorem 6.1. *Let (ρ^q, ρ^*) be a left q-tail pair for some $q \in (0,1)$. If $\rho^* \colon \mathcal{M} \to \mathbb{R}$ is $\mathcal{M}$-elicitable with the strictly $\mathcal{M}$-consistent score $S^* \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and if for all $x \in \mathbb{R}$, $S^*(x,y)$ is strictly* decreasing *in y, then*

$$\begin{aligned} S(v,x,y) &= \mathbb{1}_{\{y \leqslant v\}} S^*(x,y) - \big(\mathbb{1}_{\{y \leqslant v\}} - q\big) S(x,v) + a(y), \\ &= -\mathbb{1}_{\{y > v\}} S^*(x,y) - \big(\mathbb{1}_{\{y \leqslant v\}} - q\big) S(x,v) + \tilde{a}(y), \quad v,x,y \in \mathbb{R}, \end{aligned}$$

is strictly $\mathcal{M}$-consistent for (Q_q, ρ^q), where $a \colon \mathbb{R} \to \mathbb{R}$ and $\tilde{a} \colon \mathbb{R} \to \mathbb{R}$ are some $\mathcal{M}$-integrable functions.

The proof of Theorem 6.1 is almost identical to the one of Theorem 5.3 (ii) and therefore omitted. While the left-tail results can be obtained directly form the right-tail results, more work is needed to combine these two tails. One relevant question is related to the body part of the distribution, which can be seen as the intersection of two tails. For $0 \leqslant p < q \leqslant 1$ and $F \in \mathcal{M}^0$, we can define $F^{[p,q]} \in \mathcal{M}^0$ by

$$F^{[p,q]}(x) = \frac{(\min(F(x), q) - p)_+}{q - p}, \quad x \in \mathbb{R},$$

and a risk measure $\rho^{[p,q]}$ via another generating risk measure ρ^* by

$$\rho^{[p,q]}(F) = \rho^*(F^{[p,q]}). \tag{6.1}$$

The most prominent example of a risk measure depending on the body of the distribution is RVaR. In particular, for $0 < p < q < 1$, $\mathrm{RVaR}_{p,q}$ arises in (6.1) with ρ^* as the mean.

We obtain the following counterpart to Theorem 5.3 part (ii) and Theorem 6.1.

Theorem 6.2. *Let $\rho^* \colon \mathcal{M} \to \mathbb{R}$ be $\mathcal{M}$-elicitable with strictly $\mathcal{M}$-consistent score $S^* \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$.*

For $0 < p < q < 1$, let $\rho^{[p,q]}$ be defined via (6.1). Then the function given by

$$\begin{aligned} &S(v_1, v_2, x, y) \\ &= \mathbb{1}_{\{v_1 < y \leqslant v_2\}} S^*(x, y) + \big(\mathbb{1}_{\{y \leqslant v_1\}} - p\big) S^*(x, v_1) - \big(\mathbb{1}_{\{y \leqslant v_2\}} - q\big) S^*(x, v_2) \\ &\quad + \big(\mathbb{1}_{\{y \leqslant v_1\}} - p\big) g_1(v_1) + \mathbb{1}_{\{y > v_1\}} g_1(y) + \big(\mathbb{1}_{\{y \leqslant v_2\}} - p\big) g_2(v_2) + \mathbb{1}_{\{y > v_2\}} g_2(y) + a(y), \end{aligned} \tag{6.2}$$

is strictly $\mathcal{M}$-consistent for $(Q_p, Q_q, \rho^{[p,q]})$, where $a\colon \mathbb{R} \to \mathbb{R}$ and $\tilde{a}\colon \mathbb{R} \to \mathbb{R}$ are some $\mathcal{M}$-integrable functions, and $g_1\colon \mathbb{R} \to \mathbb{R}$ and $g_2\colon \mathbb{R} \to \mathbb{R}$ are such that for each $x \in \mathbb{R}$, the functions

$$v_1 \mapsto g_1(v_1) + S^*(x, v_1) \quad \text{and} \quad v_2 \mapsto g_2(v_2) - S^*(x, v_2) \tag{6.3}$$

are strictly increasing.

Proof. Due to the monotonicity conditions in (6.3), for fixed $v_2, x \in \mathbb{R}$, the score $(v_1, y) \mapsto S(v_1, v_2, x, y)$ is a generalized piecewise linear loss, and hence $\mathcal{M}$-strictly consistent for Q_p. Similarly, for fixed $v_1, x \in \mathbb{R}$, the score $(v_2, y) \mapsto S(v_1, v_2, x, y)$ is strictly $\mathcal{M}$-consistent for Q_q. Finally, for $F \in \mathcal{M}$ and $v_1 \in Q_p(F)$, $v_2 \in Q_q(F)$ it holds that

$$\int S(v_1, v_2, x, y) \,\mathrm{d}F(y) = \int S^*(x, y) \,\mathrm{d}\big(\min(F(x), q) - p\big)_+ + \int \tilde{a}(y) \,\mathrm{d}F(y),$$

where $\tilde{a}$ does not depend on x. Hence,

$$\operatorname*{arg\,min}_{x \in \mathbb{R}} \int S(v_1, v_2, x, y) \,\mathrm{d}F(y) = \operatorname*{arg\,min}_{x \in \mathbb{R}} \int S^*(x, y) \,\mathrm{d}F^{[p,q]}(x) = \rho^*(F^{[p,q]}) = \rho^{[p,q]}(F),$$

which shows the claim. □

Identifiability results in the spirit of Theorem 4.3 can also be obtained.

Theorem 6.3. *Let $\rho^*\colon \mathcal{M} \to \mathbb{R}$ be $\mathcal{M}$-identifiable with strict $\mathcal{M}$-identification function $V^*\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. For $0 < p < q < 1$, let $\rho^{[p,q]}$ be defined via (6.1). Then $(Q_p, Q_q, \rho^{[p,q]})$ is $\mathcal{M}_{(p)} \cap \mathcal{M}_{(q)} \cap \mathcal{M}$-identifiable with a strict $\mathcal{M}_{(p)} \cap \mathcal{M}_{(q)} \cap \mathcal{M}$-identification function $V: \mathbb{R}^4 \to \mathbb{R}$ given by*

$$\begin{pmatrix} \mathbb{1}_{\{y \leqslant v_1\}} - p \\ \mathbb{1}_{\{y \leqslant v_2\}} - q \\ \mathbb{1}_{\{v_1 < y \leqslant v_2\}} V^*(x, y) + \big(\mathbb{1}_{\{y \leqslant v_1\}} - p\big) V^*(x, v_1) - \big(\mathbb{1}_{\{y \leqslant v_2\}} - q\big) V^*(x, v_2) \end{pmatrix} \tag{6.4}$$

for $(v_1, v_2, x, y) \in \mathbb{R}^4$.

The proof of Theorem 6.3 is completely analogous to the one of Theorem 4.3 and therefore omitted. In line with Remark 4.4, instead of the identification function in (6.4) we can also use the strict $\mathcal{M}_{(p)} \cap \mathcal{M}_{(q)} \cap \mathcal{M}$-identification function without "correction terms" in the form of

$$V(v_1, v_2, x, y) = \begin{pmatrix} \mathbb{1}_{\{y \leqslant v_1\}} - p \\ \mathbb{1}_{\{y \leqslant v_2\}} - q \\ \mathbb{1}_{\{v_1 < y \leqslant v_2\}} V^*(x, y) \end{pmatrix}.$$

Theorems 6.2 and 6.3 generalize results on the elicitability and identifiability of the triplet $(Q_p, Q_q, \mathrm{RVaR}_{p,q})$, $0 < p < q < 1$, from Fissler and Ziegel (2021b), which are summarized in Example 3.6. In fact, using the canonical identification function $V^*(x, y) = x - y$ for the mean ρ^*, the identification function in (6.4) equals the one in (3.5), except for a factor of $q - p$ in the third component. Similarly, the scoring function in (6.2) yields the strictly consistent score of Equation (3.6) upon choosing the strictly consistent score for the mean $S^*(x, y) = \frac{1}{q-p}\big(\phi'(x)(x - y) - \phi(x)\big)$, where ϕ is strictly convex. Also the monotonicity conditions in (6.3) are equivalent to $v_1 \mapsto g_1(v_1) - \frac{1}{q-p}\phi'(x)v_1$ and $v_2 \mapsto g_2(v_2) + \frac{1}{q-p}\phi'(x)v_2$ being strictly increasing for all $x \in \mathbb{R}$.

7 Concluding remarks

Our main results in Sections 4 and 5 establish an intimate connection between identifiability or elicitability of a tail risk measure ρ and that of its generator ρ^*, as well as their identification functions and score functions. To summarize briefly, under suitable conditions, the corresponding property of ρ^* implies that of (Q_p, ρ), and the corresponding property of ρ or (Q_p, ρ) implies that of ρ^*. Moreover, there is an explicit way to convert between their identification functions and between their score functions.

Some open questions remain. First, elicitability of risk measures defined for the intermediate region of the distribution, instead of the tail distributions, has been discussed briefly in Section 6, without being fully developed, as we only obtained a parallel result to Theorem 5.3 part (ii).

Second, we provided in Theorem 5.3 some sufficient conditions on the elicitable risk measure ρ^* that guarantee the elicitability of the p-tail risk measure ρ itself or the pair (Q_p, ρ). A full characterization of all elicitable risk measures ρ^* yielding the elicitability of ρ or (Q_p, ρ) is missing.

Theorem 5.3 does not provide a full answer to this question. In particular, one needs to rely on the knowledge of the existence of a strictly consistent scoring function for ρ^*, $S^*(x,y)$, that is strictly increasing in y. Third, the elicitation complexity (Frongillo and Kash, 2021) of tail risk measures remains unclear. Roughly (with suitable regularizing conditions), the elicitation complexity of ρ is a positive integer k that quantifies how many dimensions are needed to make ρ a function of an elicitable vector $(\rho_1,\dots,\rho_k)$. It may be tempting to conjecture that if ρ^* has elicitation complexity k, then ρ has elicitation complexity at most $k+1$, seeing from the example of $(\rho,\rho^*)=(\mathrm{ES}_p,\mathbb{E})$ with elicitation complexity $(2,1)$; see Frongillo and Kash (2021). This question seems to be very challenging to answer with current techniques.

Conflict of interest statement

The authors declare no conflict of interest.

Data availability statement

No datasets were generated or analyzed during the current study.

Acknowledgements

The authors thank the Editor, an Associate Editor, and the anonymous referees for helpful comments. F. Liu and R. Wang acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2024-03728, CRC-2022-00141, RGPIN-2020-04717, DGECR-2020-00340). L. Wei acknowledges financial support from the National Natural Science Foundation of China (No. 12271415).

Appendices

A Risk measures

We take the mapping $\rho:\mathcal{M}\to[-\infty,\infty]$ as the primary object in the paper. In the literature, a risk measure are often defined as a mapping $\hat{\rho}$ from a set of random variables $\mathcal{X}$ on an atomless probability space $(\Omega,\mathcal{F},\mathbb{P})$ to $[-\infty,\infty]$ or $\mathbb{R}$ (Föllmer and Schied, 2016). Commonly used $\hat{\rho}$ are *law-based* (also called *law-invariant*), meaning that for any $X,X'\in\mathcal{X}$ it holds that $\hat{\rho}(X)=\hat{\rho}(X')$ whenever the distributions of X and X' coincide. It is easy to see that a law-based $\hat{\rho}:\mathcal{X}\to[-\infty,\infty]$ one-to-one corresponds to a risk measure $\rho:\mathcal{M}\to[-\infty,\infty]$ with

$\mathcal{M} = \{F_X \in \mathcal{M}^0 : X \in \mathcal{X}\}$ via

$$\rho(F) = \hat{\rho}(X) \quad \text{where } X \text{ is distributed as } F. \tag{A.1}$$

Invoking the relation (A.1), our study can easily be translated into properties of $\hat{\rho}$. Here, we follow the sign convention of McNeil et al. (2015) to consider random variables in $\mathcal{X}$ as losses.

Some commonly used desirable properties of risk measures are listed below.

(A0) *Law-invariance*: if $X, Y \in \mathcal{X}$ and $X \overset{\mathrm{d}}{=} Y$, then $\hat{\rho}(X) = \hat{\rho}(Y)$.

(A1) *Monotonicity*: $\hat{\rho}(X) \leqslant \hat{\rho}(Y)$ if $X \leqslant Y$ a.s, $X, Y \in \mathcal{X}$.

(A2) *Translation-equivariance*: $\hat{\rho}(X - m) = \hat{\rho}(X) - m$ for $m \in \mathbb{R}$ and $X \in \mathcal{X}$.

(A3) *Convexity*: $\hat{\rho}(\lambda X + (1-\lambda)Y) \leqslant \lambda\hat{\rho}(X) + (1-\lambda)\hat{\rho}(Y)$ for $\lambda \in [0,1]$ and $X, Y \in \mathcal{X}$.

(A4) *Positive homogeneity*: $\hat{\rho}(\lambda X) = \lambda\hat{\rho}(X)$ for $\lambda > 0$ and $X \in \mathcal{X}$.

(A5) *Subadditivity*: $\hat{\rho}(X + Y) \leqslant \hat{\rho}(X) + \hat{\rho}(Y)$ for $X, Y \in \mathcal{X}$.

(A6) *Comonotonic additivity*: $\hat{\rho}(X + Y) = \hat{\rho}(X) + \hat{\rho}(Y)$ if $X, Y \in \mathcal{X}$ are comonotonic.[2]

It is well known that any pair of properties (A3), (A4) and (A5) implies the remaining third one. For economic interpretations of these properties, we refer to Artzner et al. (1999), Föllmer and Schied (2016) and Delbaen (2012).

Definition A.1. A risk measure ρ is a *monetary risk measure* if its corresponding $\hat{\rho}$ satisfies (A1) and (A2), it is a *convex risk measure* if $\hat{\rho}$ satisfies (A1)–(A3), and it is a *coherent risk measure* if $\hat{\rho}$ satisfies (A1)–(A4).

Some of the above terminologies appear in the main paper, but they are not essential for our results.

B Assumptions

The following list of assumptions is an adaptation of the assumptions for Proposition 1 in Fissler and Ziegel (2021a); see also Fissler and Ziegel (2016). This list of assumptions is needed for

---

[2] Two random variables X and Y are *comonotonic* if there exists $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ and $(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geqslant 0$ for all $\omega, \omega' \in \Omega_0$. Comonotonicity of X and Y is equivalent to the existence of a random variable $Z \in L^0$ and two non-decreasing functions f and g such that $X = f(Z)$ and $Y = g(Z)$ almost surely. We refer to Dhaene et al. (2002) for an overview on comonotonicity.

Theorem 5.3, part (iii). For an identification and scoring functions $V(x,y)$ and $S(x,y)$, $x \in \mathbb{R}^k$, $y \in \mathbb{R}$, we use the shorthands $\bar{V}(x,F) := \int V(x,y)\,\mathrm{d}F(y)$ and $\bar{S}(x,F) := \int S(x,y)\,\mathrm{d}F(y)$ for $F \in \mathcal{M}^0$, tacitly assuming that the integral is well-defined. In the sequel, we shall only consider $\mathcal{M} \subseteq \mathcal{M}_p^{\mathrm{c}} \cap \mathcal{M}_{(p)}$. On $\mathcal{M}_p^{\mathrm{c}}$, $Q_p(F)$ is a singleton and we identify it with its unique element, $\mathrm{VaR}_p^-(F)$. Moreover, here and in the proof of Section C, we work with an identification function of the form (4.4). We remark that we could equivalently work with an identification function of the form (4.3).

Assumption B.1. (i) Let $\mathsf{A} = \{(\mathrm{VaR}_p^-(F), \rho(F)) : F \in \mathcal{M}\} \subseteq \mathbb{R}^2$ and suppose that the interior of A, $\mathrm{int}(\mathsf{A})$ is simply connected.

(ii) The generator ρ^* is identifiable with a strict $\mathcal{M}$-identification function $V_{\rho^*} \colon \mathbb{R}^2 \to \mathbb{R}$, which is locally bounded.

(iii) The identification function $V \colon \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$V(v,x,y) = \begin{pmatrix} \mathbb{1}_{\{y \leqslant v\}} - p \\ \mathbb{1}_{\{y > v\}} V_{\rho^*}(x,y) \end{pmatrix}$$

is such that its expectation with respect to any $F \in \mathcal{M}$ is continuously differentiable. In particular, any $F \in \mathcal{M}$ has a continuous derivative f (which is also its Lebesgue density).

(iv) The expectation of the score S with respect to any $F \in \mathcal{M}$ is twice continuously differentiable.

(v) Suppose that $\mathcal{M}$ is convex and that for any $(v,x) \in \mathsf{A}$ there are $F_1, F_2, F_3 \in \mathcal{M}$ such that 0 is contained in the interior of the convex hull of $\{\bar{V}(v,x,F_1), \bar{V}(v,x,F_2), \bar{V}(v,x,F_3)\}$.

(vi) Suppose that for any $(v,x) \in \mathsf{A}$ there are $F_1, F_2 \in \mathcal{M}$ with derivatives f_1 and f_2 such that $(v,x) = (\mathrm{VaR}_p^-(F_1), \rho(F_1)) = (\mathrm{VaR}_p^-(F_2), \rho(F_2))$, $\partial_x \bar{V}_2(v,x,F_1) = \partial_x \bar{V}_2(v,x,F_2)$, but $f_1(v) \neq f_2(v)$.

(vii) Suppose that the complement of the set

$$\{(v,x,y) \in \mathsf{A} \times \mathbb{R} : V(v,x,\cdot) \text{ and } S(v,x,\cdot) \text{ are continuous at the point } y\}$$

has 3-dimensional Lebesgue measure zero.

(viii) For every $y \in \mathbb{R}$ there exists a sequence $(F_n)_{n\in\mathbb{N}}$ of distributions $F_n \in \mathcal{M}$ that converges weakly to the Dirac-measure δ_y such there is some $C > 0$ such that $F_n([-C, C]) = 1$ for all $n \in \mathbb{N}$.

C Proofs in Section 5

Proof of Proposition 5.1. (i) Let S be a strictly $\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}$-consistent score for ρ and S_r^* as given in (5.1) for some $r \in \mathbb{R}$. Choose some $G \in \mathcal{M}_{\geqslant r} \cap \mathcal{M}$ and define $F(y) = \mathbb{1}_{\{y\geqslant r\}}\big((1-p)G(y) + p\big)$. Due to the assumed elicitability of ρ and since $F_p = G$, we get

$$\begin{aligned}
\rho^*(G) = \rho^*(F_p) = \rho(F) &= \operatorname*{arg\,min}_{x\in\mathbb{R}} \int S(x,y)\,\mathrm{d}F(y) \\
&= \operatorname*{arg\,min}_{x\in\mathbb{R}} \int S(x,y)\,\mathrm{d}\left(\mathbb{1}_{\{y\geqslant r\}}(1-p)G(y)\right) + pS(x,r) \\
&= \operatorname*{arg\,min}_{x\in\mathbb{R}} \int \big((1-p)S(x,y) + pS(x,r)\big)\,\mathrm{d}G(y) \\
&= \operatorname*{arg\,min}_{x\in\mathbb{R}} \int S_r^*(x,y)\,\mathrm{d}G(y),
\end{aligned}$$

which shows the claim.

(ii) Let $S(v,x,y)$, $v,x,y \in \mathbb{R}$, be a strictly $\mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}$-consistent score and S_r^* given in (5.2) for some $r \in \mathbb{R}$. For any $F \in \mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}$ and $v \in Q_p(F)$ it holds that

$$\rho(F) = \operatorname*{arg\,min}_{x\in\mathbb{R}} \int_{\mathbb{R}} S(v,x,y)\,\mathrm{d}F(y).$$

Let $G \in \mathcal{M}_{\geqslant r} \cap \mathcal{M}$. Then, as in part (i), define $F(y) = \mathbb{1}_{\{y\geqslant r\}}\big((1-p)G(y) + p\big)$, resulting in $F \in \mathcal{M} \cup \{(1-p)G + p\delta_r : G \in \mathcal{M}\}$ and $r = \mathrm{VaR}_p^-(F) \in Q_p(F)$. Again, $F_p = G$, which implies

$$\begin{aligned}\rho^*(G) = \rho^*(F_p) = \rho(F) &= \operatorname*{arg\,min}_{x\in\mathbb{R}} \int S(r,x,y)\,\mathrm{d}F(y)\\ &= \operatorname*{arg\,min}_{x\in\mathbb{R}} \int S(r,x,y)\,\mathrm{d}\left(\mathbb{1}_{\{x\geqslant r\}}(1-p)G(x)\right) + pS(r,x,r)\\ &= \operatorname*{arg\,min}_{x\in\mathbb{R}} \int \left((1-p)S(r,x,y) + pS(r,x,r)\right)\,\mathrm{d}G(x)\\ &= \operatorname*{arg\,min}_{x\in\mathbb{R}} \int S_r^*(x,y)\,\mathrm{d}G(y),\end{aligned}$$

which shows the claim. □

For a proof of Theorem 5.3 part (iii), in the sequel, we shall only consider $\mathcal{M} \subseteq \mathcal{M}_p^{\mathrm{c}} \cap \mathcal{M}_{(p)}$. On $\mathcal{M}_p^{\mathrm{c}}$, $Q_p(F)$ is a singleton and we identify it with its unique element, $\mathrm{VaR}_p^-(F)$. Moreover, here and in the Assumption Section B, we work with an identification function of the form (4.4). We remark that we could equivalently work with an identification function of the form (4.3).

Proof of Theorem 5.3 part (iii). Using Theorem 4.3 and Remark 4.4, the function

$$V(v,x,y) = \begin{pmatrix} \mathbb{1}_{\{y\leqslant v\}} - p \\ \mathbb{1}_{\{y>v\}} V_{\rho^*}(x,y) \end{pmatrix}, \quad v,x,y \in \mathbb{R}$$

is a strict $\mathcal{M}$-identification function for (VaR_p^-, ρ); see also Fissler and Hoga (2024, Proposition 3.4). We obtain for $F \in \mathcal{M}$ with derivative (and density) f

$$\begin{aligned} &\bar{V}_1(v,x,F) = F(v) - p, \qquad \partial_v \bar{V}_1(v,x,F) = f(v), \qquad \partial_x \bar{V}_1(v,x,F) = 0,\\ &\bar{V}_2(v,x,F) = \int_v^\infty V_{\rho^*}(x,y)\,\mathrm{d}F(y), \qquad \partial_v \bar{V}_2(v,x,F) = -V_{\rho^*}(x,v)f(v).\end{aligned}$$

Due to Osband's principle Fissler and Ziegel (2016, Theorem 3.2), there is a matrix-valued function $h\colon \mathrm{int}(\mathsf{A}) \to \mathbb{R}^{2\times 2}$ such that

$$\nabla \bar{S}(v,x,F) = h(v,x)\nabla V(v,x,F) \qquad \text{for all } (v,x) \in \mathrm{int}(\mathsf{A}),\ F \in \mathcal{M}.$$

We obtain for the second order derivatives

$$\begin{aligned}\partial_x\partial_v\bar{S}(v,x,F) &= \partial_x h_{11}(v,x)\big(F(v)-p\big) + h_{12}(v,x)\partial_x\bar{V}_2(v,x,F) + \partial_x h_{12}(v,x)\bar{V}_2(v,x,F),\\ \partial_v\partial_x\bar{S}(v,x,F) &= h_{21}(v,x)f(v) + \partial_v h_{21}(v,x)\big(F(v)-p\big) - h_{22}(v,x)V_{\rho^*}(x,v)f(v)\\ &\qquad + \partial_v h_{22}(v,x)\bar{V}_2(v,x,F).\end{aligned}$$

The Hessian of the expected score $\bar{S}(v,x,F)$ needs to be symmetric for all $(v,x)\in\mathrm{int}(\mathsf{A})$ and for all $F\in\mathcal{M}$. Evaluating $\partial_x\partial_v\bar{S}(v,x,F) = \partial_v\partial_x\bar{S}(v,x,F)$ for $(v,x) = (\mathrm{VaR}_p^-(F),\rho(F))$, we get

$$h_{12}(v,x)\partial_x\bar{V}_2(v,x,F) = f(v)\Big(h_{21}(v,x) - h_{22}(v,x)V_{\rho^*}(x,v)\Big).$$

Using part (vi) of Assumption B.1, this implies that

$$h_{12}(v,x) = 0, \qquad h_{21}(v,x) = h_{22}(v,x)V_{\rho^*}(x,v).$$

Exploiting the surjectivity, this holds for all $(v,x)\in\mathrm{int}(\mathsf{A})$. We evaluate $\partial_x\partial_v\bar{S}(v,x,F) = \partial_v\partial_x\bar{S}(v,x,F)$ again, but now for a general $(v,x)\in\mathrm{int}(\mathsf{A})$, and obtain

$$\partial_x h_{11}(v,x)\big(F(v)-p\big) = \partial_v h_{21}(v,x)\big(F(v)-p\big) + \partial_v h_{22}(v,x)\bar{V}_2(v,x,F).$$

If $v = \mathrm{VaR}_p^-(F)$ and $x\neq\rho(F)$, we get that $\partial_v h_{22}(v,x) = 0$. Exploiting the surjectivity, we get

$$\partial_v h_{22}(v,x) = 0, \qquad \partial_x h_{11}(v,x) = \partial_v h_{21}(v,x), \qquad \text{for all } (v,x)\in\mathrm{int}(\mathsf{A}).$$

Hence, we can write h_{22} as a function of its second argument x only. Now, we integrate the partial derivatives of the expected score.

$$\begin{aligned}\int \partial_v\bar{S}(v,x,F)\,\mathrm{d}v &= \int h_{11}(v,x)\big(F(v)-p\big)\,\mathrm{d}v = \int h_{11}(v,x)\int\big(\mathbb{1}_{\{y\leqslant v\}}-p\big)\,\mathrm{d}F(y)\,\mathrm{d}v\\ &= \int G(y,x)\mathbb{1}_{\{y>v\}} + \big(\mathbb{1}_{\{y\leqslant v\}}-p\big)G(v,x) + a_1(y,x)\,\mathrm{d}F(y),\end{aligned}$$

where $\partial_v G(v,x) = h_{11}(v,x)$ and $a_1(y,x)$ is an integration constant not depending on v. Similarly,

$$\begin{aligned}
&\int \partial_x \bar{S}(v,x,F)\,\mathrm{d}x = \int h_{22}(x)\Big(V_{\rho^*}(x,v)\big(F(v)-p\big) + \bar{V}_2(v,x,F)\Big)\,\mathrm{d}x \\
&= \int h_{22}(x) \int V_{\rho^*}(x,v)\big(\mathbb{1}_{\{y\leqslant v\}} - p\big) + \mathbb{1}_{\{y>v\}} V_{\rho^*}(x,y)\,\mathrm{d}F(y)\,\mathrm{d}x \\
&= \int \int h_{22}(x)V_{\rho^*}(x,v)\,\mathrm{d}x\big(\mathbb{1}_{\{y\leqslant v\}} - p\big) + \mathbb{1}_{\{y>v\}} \int h_{22}(x)V_{\rho^*}(x,y)\,\mathrm{d}x\,\mathrm{d}F(y) \\
&= \int \Big(S_{\rho^*}(x,v) + a_2(v,y)\Big)\big(\mathbb{1}_{\{y\leqslant v\}} - p\big) + \mathbb{1}_{\{y>v\}}\Big(S_{\rho^*}(x,y) + a_3(v,y)\Big)\,\mathrm{d}F(y),
\end{aligned}$$

where $\partial_x S_{\rho^*}(x,y) = h_{22}(x)V_{\rho^*}(x,y)$ and $a_2(v,y)$ and $a_3(v,y)$ are integration constants not depending on x. A comparison of the two results and an application of Proposition 1 in Fissler and Ziegel (2021a) yields the form of S in (5.4). Since for any $F \in \mathcal{M}$, the function $(v,y) \mapsto S(v,\rho(F),y)$ needs to be (strictly) $\mathcal{M}$-consistent for VaR_p^-, $S_{\rho^*}(x,y)$ needs to be (strictly) increasing in y. On the other hand, $(x,y) \mapsto S(\mathrm{VaR}_p^-(F),x,y)$ needs to be (strictly) $\mathcal{M}$-consistent for ρ, which is why S_{ρ^*} needs to be (strictly) $\mathcal{M}$-consistent for ρ^*. ☐

References

Acerbi, C. and Szèkely, B. (2014). Back-testing expected shortfall. *Risk Magazine*, **27**, 76–81.

Adrian, T. and Brunnermeier, M. K. (2016). CoVaR. *American Economic Review*, **106**, 1705–1741.

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

BCBS (2019). *Minimum Capital Requirements for Market Risk. February 2019.* Basel Committee on Banking Supervision. Basel: Bank for International Settlements. https://www.bis.org/bcbs/publ/d457.htm

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

Bellini, F. and Bignozzi, V. (2015). On elicitable risk measures. *Quantitative Finance*, **15**(5), 725–733.

Bellini, F., Klar, B., Müller, A. and Rossaza Gianin, E. (2014). Generalized quantiles as risk measures. *Insurance: Mathematics and Economics*, **54**, 41–48.

Cont, R., Cucuringu, M., Xu, R., Zhang, C. (2025). Tail-GAN: Learning to simulate tail risk scenarios. *Management Science*, forthcoming.

Cont, R., Deguest, R. and He, X. D. (2013). Loss-based risk measures. *Statistics and Risk Modeling*, **30**, 133–167.

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

Dawid, A. P. (1998): *Coherent Measures of Discrepancy, Uncertainty and Dependence, With Applications to Bayesian Predictive Experimental Design.* Research Report 139, University College London, Dept. of Statistical Science.

Delbaen, F. (2012). *Monetary Utility Functions.* Osaka University Press, Osaka.

Delbaen, F., Bellini, F., Bignozzi, V. and Ziegel, J. (2016). Risk measures with convex level sets. *Finance and Stochastics*, **20**(2), 433–453.

Dhaene, J., Denuit, M., Goovaerts, M. J., Kaas, R. and Vyncke, D. (2002): The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics*, **31**(1), 3–33.

Dimitriadis, T., Fissler, T. and Ziegel, J. (2024). Osband's principle for identification functions. *Statistical Papers*, **65**, 1125–1132.

Dimitriadis, T., Fissler, T. and Ziegel, J. (2024). Characterizing M-estimators. *Biometrika*, **111**(1), 339–346.

EIOPA (2011). Equivalence assessment of the Swiss supervisory system in relation to articles 172, 227 and 260 of the Solvency II Directive, *EIOPA-BoS-11-028*.

Embrechts, P., Liu, H. and Wang, R. (2018). Quantile-based risk sharing. *Operations Research*, **66**(4), 936–949.

Embrechts, P., Mao, T., Wang, Q. and Wang, R. (2021). Bayes risk, elicitability, and the Expected Shortfall. *Mathematical Finance*, **31**, 1190–1217.

Emmer S., Kratz M. and Tasche D. (2015). What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk*, **18**, 31–60.

Fissler, T. Frongillo, R., Hlavinová, J. and Rudloff, B. (2021). Forecast evaluation of quantiles, prediction intervals, and other set-valued functionals. *Electronic Journal of Statistics*, **15**, 1034–1084.

Fissler, T. and Hoga, Y. (2024). Backtesting systemic risk forecasts using multi-objective elicitability. *Journal of Business & Economic Statistics*, **42**(2), 485–498.

Fissler, T. Merz, M. and Wüthrich, M. (2023). Deep quantile and deep composite model regression. *Insurance: Mathematics and Economics*. **109**, 94–112.

Fissler, T. and Ziegel, J. F. (2016). Higher order elicitability and Osband's principle. *Annals of Statistics*, **44**(4), 1680–1707.

Fissler, T. and Ziegel, J. F. (2019). Order-sensitivity and equivariance of scoring functions. *Electronic Journal of Statistics*, **13**(1), 1166–1211.

Fissler, T. and Ziegel, J. F. (2021a). Correction note: Higher order elicitability and Osband's principle. *Annals of Statistics*, **49**(1), 614–614.

Fissler, T. and Ziegel, J. F. (2021b). Elicitability of Range Value at Risk. *Statistics and Risk Modeling*, **38**(1-2), 25–46.

Fissler, T. and Ziegel, J. F. and Gneiting, T. (2016) Expected Shortfall is jointly elicitable with Value at Risk, Implications for backtesting. *Risk Magazine*, January 2016, 58–61.

Föllmer, H. and Schied, A. (2016). *Stochastic Finance: An Introduction in Discrete Time.* Walter de

Gruyter, Berlin, Fourth Edition.

Frongillo, R. and Kash, I. A. (2021). Elicitation complexity of statistical properties. *Biometrika*, **108**(4), 857–879.

Furman, E. and Landsman, Z. (2006). Tail variance premium with applications for elliptical portfolio of risks. *ASTIN Bulletin*, **36**, 433–462.

Furman, E., Wang, R. and Zitikis, R. (2017). Gini-type measures of risk and variability: Gini shortfall, capital allocation and heavy-tailed risks. *Journal of Banking and Finance*, **83**, 70–84.

Girardi, G. and Tolga Ergün, A. (2013). Systemic risk measurement: Multivariate GARCH estimation of CoVaR. *Journal of Banking and Finance*, **37**, 3169–3180.

Gneiting, T. (2011a). Making and evaluating point forecasts. *Journal of the American Statistical Association*, **106**(494), 746–762.

Gneiting, T. (2011b). Quantiles as optimal point forecasts. *International Journal of Forecasting*, **27**(2), 197–207.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, **102**(477), 359–378.

Gneiting, T. and Ranjan, R. (2011). Comparing density forecasts using threshold-and quantile-weighted scoring rules. *Journal of Business & Economic Statistics*, **29**(3), 411–422.

Holzmann, H. and Klar, B. (2017). Focusing on regions of interest in forecast evaluation. *The Annals of Applied Statistics*, **11**(4), 2404–2431.

Huber, P. J. and Ronchetti, E. M. (2009). *Robust Statistics*, 2nd ed., John Wiley & Sons, Hoboken.

Jarrow, R. (2002). Put option premiums and coherent risk measures. *Mathematical Finance*, **12**(2), 135–142.

Koenker, R. (2005). *Quantile Regression*, Cambridge University, Cambridge.

Kou, S. and Peng, X. (2016). On the measurement of economic tail risk. *Operations Research*, **64**(5), 1056–1072.

Lambert, N. S., Pennock, D. M. and Shoham, Y. (2008). Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, 129-138. ACM.

Liu, F., Mao, T., Wang, R. and Wei, L. (2022). Inf-convolution, optimal allocations, and model uncertainty for tail risk measures. *Mathematics of Operations Research*, , **47**(3), 2494–2519.

Liu, F. and Wang, R. (2021). A theory for measures of tail risk. *Mathematics of Operations Research*, **46**(3), 1109–1128.

McNeil, A. J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools*. Revised Edition. Princeton, NJ: Princeton University Press.

Nau, R. F. (1985). Should scoring rules be "effective"? *Management Science*, **31**, 527–535.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. In *Handbook of Econometrics,* vol. 4, R. F. Engle & D. McFadden, eds. Amsterdam: Elsevier, pp. 2111–245.

Newey, W. K. and Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica*,

55, 819–847.

Nolde, N., and Zhang, J. (2020). Conditional extremes in asymmetric financial markets. *Journal of Business & Economic Statistics*, **38**, 201–213.

Nolde, N., and Ziegel, J. F. (2017). Elicitability and backtesting: Perspectives for banking regulation. *The Annals of Applied Statistics*, **11**(4), 1833–1874.

Osband, K. H. (1985). Providing incentives for better cost forecasting. *PhD Thesis, University of California, Berkeley*. https://doi.org/10.5281/zenodo.4355667.

Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, **26**(7), 1443–1471.

Steinwart, I., Pasin, C., Williamson, R. and Zhang, S. (2014). Elicitation and identification of properties. *JMLR Workshop and Conference Proceedings*, **35**, 1–45.

Tsanakas, A. (2009). To split or not to split: Capital allocation with convex risk measures. *Insurance: Mathematics and Economics*, **44**(2), 268–277.

Wang, Q., Wang, R. and Ziegel, J. (2025). E-backtesting. *Management Science*, forthcoming.

Wang, R. and Wei, Y. (2020). Risk functionals with convex level sets. *Mathematical Finance*, **30**(4), 1337–1367.

Wang, R. and Ziegel, J. (2015). Elicitable distortion risk measures: A concise proof. *Statistics and Probability Letters*, **100**, 172–175.

Weber, S. (2006). Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, **16**, 419–441.

Ziegel, J. (2016). Coherence and elicitability. *Mathematical Finance*, **26**, 901–918.