Minimizing Risk Exposure when the Choice of a Risk Measure is Ambiguous

Erick Delage, Jonathan Y. Li

January 30, 2015

Abstract

Since the financial crisis of 2007-2009, there has been a renewed interest toward quantifying more appropriately the risks involved in financial positions. Popular risk measures such as variance and value-at-risk have been found inadequate as we now give more importance to properties such as monotonicity, convexity, translation invariance, scale invariance, and law invariance. Unfortunately, the challenge remains that it is unclear how to choose a risk measure that faithfully represents the decision maker's true risk attitude. In this work, we show that one can account precisely for (neither more nor less than) what we know of the risk preferences of an investor/policy maker when comparing and optimizing financial positions. We assume that the decision maker can commit to a subset of the above properties (the use of a law invariant convex risk measure for example) and that he can provide a series of assessments comparing pairs of potential risky payoffs. Given this information, we propose to seek financial positions that perform best with respect to the most pessimistic estimation of the level of risk potentially perceived by the decision maker. We present how this preference robust risk minimization problem can be solved numerically by formulating convex optimization problems of reasonable size. Numerical experiments on a portfolio selection problem, where the problem reduces to a linear program, will illustrate the advantage of accounting for the fact that the information about risk perception is limited.

1 Introduction

Since the financial crisis of 2007-2009, there has been a renewed interest toward quantifying more appropriately the amount of risk involved in financial positions. While it appears that not so long ago, value at risk (VaR) seemed like the most natural measure to quantify risk, there is now a growing dissatisfaction about this measure which is accused of leading to excessive risk-taking, to ignorance of outcomes in the tails of distributions and to indirectly create a false sense of security (see Basel Committee on Banking Supervision (2012) and Einhorn (2008) for some discussions about its use in the Basel II Accord).

In search for an alternative, there has been extensive recent work attempting to define the properties that should be satisfied by a reasonable measure of risk. In Föllmer and Schied (2002), the authors identified properties that have since then be considered legitimate by many investors, banking institutions and regulators, namely monotonicity, convexity, and translation invariance. The measures that satisfy these properties are known as convex risk measures. The measure becomes a coherent risk measure if it is in addition "positively homogeneous". Finally, in Kusuoka (2001), the author introduced the axiom of "law-invariance" which states that risk should depend solely on the distribution of future returns. These properties have led to a rise in popularity of a risk measure known under many names : tail conditional expectation (see Artzner et al. (1999)), expected shortfall, conditional value at risk (see Rockafellar and Uryasev (2000)), average value at risk, etc.

While conditional value at risk (CVaR) is now being used in a wide range of field of applications such as health care (see Chan et al. (2014)), supply chain (see Carneiro et al. (2010)), network design (see Babazadeh et al. (2011)), vehicle routing (see Toumazis and Kwon (2013)), energy (see Jabr (2005)), etc., there is still some resistance in the risk management community toward adopting such a measure as a golden standard for risk management. The issue that CVaR is not statistically elicitable is discussed in Cont et al. (2010), which refers to that it cannot be measured using statistically robust procedures and is not well fit for backtesting. CVaR is also too limiting to capture the risk attitude of the decision maker. In particular, this has raised the need for new measures that can express in simple terms a larger spectrum of risk attitudes. The importance of this need gave rise to the introduction of a generalization of CVaR measures known under the name of spectral risk measure (most likely coined in Acerbi (2002)) which employs a weighted combination of CVaR measures.

Along the new development of risk measures, to this day however, the community of risk management experts have still not engaged in a serious discussion about procedures that can be used by decision makers to guide them in identifying a measure that is well fit to characterize their own subjective perception of risks. Comparatively, the literature that covers the Von-Neuman Morgenstein expected utility framework is much richer on this topic perhaps due to its longer history. Most textbooks on decision analysis (see Clemen and Reilly (1999)) will for instance propose preference elicitation methods that are based on pairwise comparison of lotteries. More recently, efforts have even be made to account for incomplete preference information in such a framework. For instance, the authors in Chajewska et al. (2000) represent the knowledge of the decision maker's preferences using a probability distribution over the plausible risk preference relations in order to integrate this additional layer of uncertainty in the estimation of expected utility. Alternatively, in Armbruster and Delage (2015) and in Hu and Mehrotra (2015) the idea is to formulate a set of plausible risk preference relations and seek decisions that are optimal with respect to worst-case expected utility or worst-case certainty equivalent. Similar ideas have also been proposed in the context of multi-attribute utility theory. In particular, one might refer to Boutilier et al. (2006) where decisions were rather supported through the evaluation of regret experienced after implementation and once true preferences are revealed. The lack of similar literatures however in the context of risk measures directs us to the following question:

"How can one capture and account for limited information about the risk preferences of an investor/decision/policy maker (hereafter referred to as the decision maker) when using a convex risk measure to compare and optimize financial positions?"

In this work, we show that one can actually compare and optimize financial positions using a convex risk measure that accounts precisely for (neither more nor less than) what is known of his preferences regarding risk. We assume that the decision maker can agree with a subset of the axioms of convex risk measures (monotonicity, translation invariance, scale invariance, and law invariance) and that he can provide a series of assessment comparing pairs of potential risky payoffs. In practice, the latter might be obtained from a risk tolerance assessment survey such as in Grable and Lytton (1999) with questions like :

> "You are on a TV game show and can choose one of the following. Which would you take?
>
> A) \$1,000 in cash;
>
> B) A 50% chance at winning \$5,000;
>
> C) A 25% chance at winning \$10,000;
>
> D) A 5% chance at winning \$100,000."

Alternatively, if the perception of risk is dependent on the realization of specific events (i.e. the law invariance axiom is not entirely satisfied), risky payoffs might be compared in terms of these uncertain events.

> "You believe that the economy of the country where you are investing has 50% chance of falling in a recession in the coming year. Which investment would you take?
>
> A) An investment that is certain to yield a 5% return;
>
> B) An investment that loses 10% of its value in the case of a recession, but yields 30% return otherwise
>
> C) An investment that has 50% chance of losing 10% of its value and 50% chance of yielding 30% return no matter whether there is a recession or not."[1]

Given the information above, the framework we propose allows one to identify financial positions that perform best with respect to the most pessimistic estimation of the level of risk potentially perceived by such a decision maker. Given that the outcome space is finite and that the payoff function is concave with respect to the set of alternatives, we also present how this "robust" risk minimization problem can be solved efficiently by formulating convex optimization problems of reasonable size. Numerical experiments on a portfolio selection problem, where the optimization problem reduces to a linear program, will also illustrate the advantages of accounting for the fact that one's risk perception might be more sophisticated than a simple expected shortfall representation and illustrate how the information about risk perception can be used to improve the quality of the portfolio that is proposed.

[1] Here, with option B, we are able to identify decision makers that are more willing to accept some losses when the economy is doing poorly in general.

To the best of our knowledge, this novel "preference robust" risk minimization framework provides for the first time 1) a language that can be used to interact with decision makers who are not well trained in the use of quantitative measures to acquire information about their preference regarding risk, together with 2) the means of identifying and optimizing a risk measure that is specifically designed (or "personalized") for every different decision maker. Unlike work such as Armbruster and Delage (2015) that relies on expected utility theory, because our framework is inspired by convex risk measures, it will not require the decision maker to agree with the independence axiom which is subject to a non-negligible amount of controversy and known to raise important practical issues such as portrayed by Ellsberg's paradox (see Ellsberg (1961)). On the other hand, it will rely more heavily on the hypothesis of translation invariance which, although well accepted in a number of banking applications, might limit its application to contexts where absolute risk aversion appears relatively constant.[2] Finally, considering that communicating with a human is subject to misinterpretation and more importantly that human are not always perfectly logical beings (see the work on cognitive biases pioneered by Tversky and Kahneman (1974)), we indicate how inconsistencies in the decision maker's assessments might be identified and potentially corrected for.

In what follows, we start in section 2 by presenting the type of information that we assume can be obtained from the decision maker regarding his perception of risk. In particular, we will expose four fundamental hypotheses that need to be made regarding how he qualitatively compares the risk associated to different financial positions in order to employ a convex risk measure representation. The specificity of the risk measure will be obtained by specifying properties like scale invariance and law invariance, and most importantly by providing a list of pairs of uncertain payoffs ordered from least risky to most risky. We then introduce the preference robust risk minimization framework and establish that making decisions using this framework is consistent with each hypothesis that is made. In section 3 we present how the preference robust risk minimization framework formulated based on the available information can be reduced to a mathematical program of reasonable size. We complete the exposition of our framework by illustrating in section 4 how it can be applied to a portfolio optimization problem. Finally, we conclude in section 5 and briefly discuss how the framework could be modified to account for errors in the comparisons of payoffs due to the presence of cognitive biases.

Remark 1.1. While the six hypotheses that are made in section 2 about how risk is perceived will be shown equivalent to the popular axioms of convex risk measures, we believe essential that such hypotheses be expressed in terms of properties of a risk preference relation, i.e. statements that avoid the need to express how much riskier a position is compared to another. Indeed, from the point of view of decision making, the quantification of risk is an inherently subjective process and cannot be confirmed through any type of physical measurement. For this reason, following a similar philosophy as was followed in expected utility theory, we prefer presenting

[2]Note that the fact that risk measures such as CVaR have been successfully adopted by such a wide range of field of applications seems to testify that the translation invariance hypothesis is not so limiting in practice.

hypotheses that must be made about a risk preference relation in order to enable the use of a convex risk measures as a characterization. A side product of this conceptual choice is to facilitate the adoption of our risk management framework by individuals that are less comfortable formulating judgments about how a quantitative risk measure should behave. This being said, we welcome the readers that are experts in the use of convex risk measures to simply interpret our six hypotheses as their respective equivalent in terms of axiom, namely monotonicity, convexity, translation invariance, scale invariance, and law invariance.

2 Preference Robust Risk Minimization

Given an outcome space Ω with a finite number of outcomes, we consider a decision maker that is offered alternatives which will affect the payoff he will achieve. In particular, each alternative is associated to a random variable $Z : \Omega \to \mathbb{R}$ which captures the uncertainty he has in the payoff that will be achieved. Note that since Ω is finite, for simplicity of exposure we can represent any random variable Z as a vector $\vec{Z}$ in $\mathbb{R}^M$ with $M = |\Omega|$, and will let $\mathbb{R}^M$ be the set of all random variables on this finite outcome space.

Assuming that the decision maker is able for any pair $(\vec{Z}_1, \vec{Z}_2) \in \mathbb{R}^M \times \mathbb{R}^M$ of risky payoffs to establish whether $\vec{Z}_1$ is no more risky than $\vec{Z}_2$ or the opposite, we denote by $\succeq$ the risk preference relation that captures either of the assessment through $\vec{Z}_1 \succeq \vec{Z}_2$ or $\vec{Z}_2 \succeq \vec{Z}_1$ respectively, and by $\vec{Z}_1 \sim \vec{Z}_2$ when both are true (i.e. risk is exactly the same). Further assuming that this preference relation respects transitivity and continuity[3], then by the utility representation theorem in Debreu (1954) one can always identify a "risk measure" $\rho : \mathbb{R}^M \to \mathbb{R}$ that captures the decision maker's risk preferences, and formulate the risk minimization problem as

$$\min_{x \in \mathcal{X}} \rho(\vec{Z}(x)), \tag{1}$$

where $\mathcal{X} \subseteq \mathbb{R}^n$ is a convex set of feasible alternatives, $\vec{Z}(x)$ is the vector representing the random payoff that is achieved when implementing x. Throughout the following sections, we will make four fundamental hypotheses about the decision maker's risk preference relation $\succeq$:

1. (Monotonicity) If a random payoff Z_1 dominates Z_2 over Ω, i.e. $Z_1(\omega) \geq Z_2(\omega)$, $\forall\, \omega \in \Omega$, then Z_1 must be no more risky than Z_2.
2. (Strict monotonicity for certain payoffs) a certain amount is always strictly less risky than a certain strictly lower amount.
3. (Quasi-convexity) If two random payoffs Z_1 and Z_2 are considered no more risky than a third one Z_3, then any convex combination of Z_1 and Z_2 is also considered no more risky then Z_3.

[3]The transitivity property states that if Z_1 is preferred (in the sense of $\succeq$) to Z_2 and Z_2 is preferred to Z_3 then it must be that Z_1 is preferred to Z_3. The continuity property states that for any Z_1 and any sequence Z'_k preferred to Z_1 and converging to Z_2, it must be that Z_2 is preferred to Z_1, and similarly if the preference is in the opposite direction.

4. (Translation invariance) Adding or subtracting the same certain amount of cash to two random payoffs should not affect whether one is no more risky than another.

The first two hypotheses are fairly natural since we are talking about monetary amounts. The quasi-convexity hypothesis is also reasonable since it states that diversification should not be considered to increase risk. Although translation invariance is the hypothesis that can be most subject to controversy, it has been widely used in the literature on risk measures. The motivation might be considered two-fold. First, it is coherent with the common banking principle that if a certain amount needs to be reserved in order for the risk of a position to become acceptable, then a second position that guarantees to payoff exactly c\$ more than the first one should require c\$ less to be reserved. Secondly, from an operational point of view, it conveniently allows for decisions that do not involve uncertain outcomes to be treated separately (as they often are anyway) thus helping to reduce the complexity of problems.[4]

The set of risk measures that satisfy these properties can be shown to be the set of convex risk measures:

$$\mathfrak{R} := \left\{ \rho : \mathbb{R}^M \to \mathbb{R} \;\middle|\; \begin{array}{c} \vec{Z}_1 \geq \vec{Z}_2 \Leftrightarrow \rho(\vec{Z}_1) \leq \rho(\vec{Z}_2) \\ \theta\rho(\vec{Z}_1) + (1-\theta)\rho(\vec{Z}_2) \geq \rho(\theta\vec{Z}_1 + (1-\theta)\vec{Z}_2)\,,\, \forall\, \vec{Z}_1, \vec{Z}_2, 0 \leq \theta \leq 1 \\ \rho(\vec{Z} + c) = \rho(\vec{Z}) - c \\ \rho(0) = 0 \end{array} \right\},$$

where one can respectively recognize the popular monotonicity, convexity, and translation invariance axioms, and where the last constraint is a normalization constraint chosen, without loss of generality, such that each member of $\mathfrak{R}$ represents a unique risk preference relation and more importantly such that $\rho(\vec{Z})$ returns the negative of the certain amount considered to have equivalent risk to Z.[5] We refer the reader to appendix A for a detailed proof.

In order to solve problem (1), ones needs to have in hand complete information about the decision maker's risk preferences. Unfortunately, in practice it is unavoidable that the knowledge about these risk preferences is at best incomplete or even possibly inaccurate. Indeed although any measure ρ from the set $\mathfrak{R}$ might appear legitimate to use in problem (1), the possibility that the selected ρ might capture a notion of risk that is very different from what is intended can be worrisome as it might lead to severe underestimation of risk. This motivates the idea of accounting for ambiguity about the risk measure ρ through a robust optimization framework. In particular, given a set of risk measures $\mathcal{R} \subseteq \mathfrak{R}$ that are deemed to be potential candidates for representing one's risk preferences, we will seek the solution that minimizes the largest risk that might be perceived by the decision maker. In other words, we are interested in solving the following optimization problem :

$$\min_{x \in \mathcal{X}} \varrho_{\mathcal{R}}(\vec{Z}(x))\,, \tag{2}$$

[4]Indeed, given that a decision involves choosing between either random payoffs Z_1 or Z_2 and choosing among a set of deterministic payoffs $\mathcal{Y} = \{y_1, y_2, \ldots, y_m\}$. The choice between Z_1 and Z_2 can be made independently of the choice among $\mathcal{Y}$ since the difference between the total risk of $Z_1 + y$ and $Z_2 + y$ is the same as between Z_1 and Z_2 no matter what y is.

[5]This is without loss of generality since $\rho'(\vec{Z}) := \rho(\vec{Z}) - \rho(0)$ always captures the same attitude as $\rho(\cdot)$.

where $\varrho_{\mathcal{R}}(\vec{Z}(x)) := \sup_{\rho\in\mathcal{R}} \rho(\vec{Z}(x))$ and with $\mathcal{R} \subseteq \mathfrak{R}$. Note that problem (2) is a convex optimization problem when each term of $\vec{Z}(x)$ is a concave function of x, in other words when $Z(x, \omega)$ is a concave function of x for all $\omega \in \Omega$.

In fact, since we impose that $\rho(0) = 0$ for all $\rho \in \mathfrak{R}$, one can also interpret problem (2) as searching for the decision that maximizes the lowest certain payoff which might be perceived equivalent in terms of risk to the resulting payoff. Mathematically speaking, when choosing among all random payoffs $Z(x)$ achievable by adjusting $x \in \mathcal{X}$, our approach seeks to maximize the optimal value of :

$$\min_{z,\rho\in\mathcal{R}} \quad z \tag{3a}$$

$$\text{subject to} \quad z \sim_\rho Z(x) \ , \tag{3b}$$

where $Z_1 \sim_\rho Z_2$ captures to the condition that Z_1 is exactly as risky as Z_2 according to the plausible risk preference relation expressed by ρ.[6] This has an important implication that the financial decisions that our approach proposes will only engage the decision maker in a position that he considers no more risky than the position of earning the best-available certain payoff. Namely, given a solution $x^* \in \mathcal{X}$ that maximizes (3), for any $\bar{\rho} \in \mathcal{R}$ we have that $Z(x^*) \succeq_{\bar{\rho}} Z(\bar{x})$ for any $\bar{x} \in \mathcal{X}$ for which $Z(\bar{x}, \omega) = Z(\bar{x}, \omega')$ for any ω and ω' in Ω. [7]

Of course, problem (2) is only interesting if one formulates $\mathcal{R}$ in a way that accurately captures his knowledge about ρ. For instance, it is clear that if the risk perception has been fully and exactly specified, then $\mathcal{R}$ should reduce to a singleton $\{\bar{\rho}\}$ and problem (2) becomes equivalent to $\min_{x\in\mathcal{X}} \ \bar{\rho}(\vec{Z}(x))$. At the other end of the spectrum, since $\rho(\vec{Z}) := \max_i -\vec{Z}_i$ is a member of $\mathfrak{R}$ and clearly provides the maximum estimates of risk, simply using $\mathfrak{R}$ makes the problem reduce to the more classical robust optimization formulation $\min_{x\in\mathcal{X}} \ \max_{\omega\in\Omega} \ -Z(x, w)$ which will typically be considered overly conservative.

We are therefore in need of imposing additional structure on the risk measures of $\mathcal{R}$. Specifically we will consider injecting three types of information in the description of $\mathcal{R}$: scale invariance, law invariance, and elicited preference information. These take the following form:

5. (Scale invariance) If a random payoff Z_1 is considered no more risky than Z_2, then a scaled version λZ_1, with $\lambda \geq 0$, is considered no more risky than the scaled random payoff λZ_2.
6. (Law invariance) Given some probability measure F over the σ-algebra of Ω, if Z_1 and Z_2 have the same distribution based on F then Z_1 is considered exactly as risky as Z_2.

[6] The equivalence can be motivated as follow. First, the constraint indicates that $\rho(z) = \rho(\vec{Z}(x))$. By translation invariance and the fact that $\rho(0) = 0$, the constraint is further reduced to $-z = \rho(\vec{Z}(x))$. Finally, a simple replacement of $z := -\rho(\vec{Z}(x))$ demonstrates that the optimal value of problem (3) is indeed equal to $\inf_{\rho\in\mathcal{R}} -\rho(\vec{Z}(x))$. A random payoff $Z(x)$ that maximizes this measure over $x \in \mathcal{X}$ necessarily achieves optimality in problem (2).

[7] Letting $\mathbb{CE}_{\bar{\rho}}(Z(x))$ denote the optimal value of problem (3) when ρ is fixed to $\bar{\rho}$, one can establish that $\mathbb{CE}_{\bar{\rho}}(Z(\bar{x})) = \min_{\rho\in\mathcal{R}} \mathbb{CE}_\rho(Z(\bar{x})) \leq \min_{\rho\in\mathcal{R}} \mathbb{CE}_\rho(Z(x^*)) \leq \mathbb{CE}_{\bar{\rho}}(Z(x^*))$.

7. (Elicited comparisons) Given a set of pairs of random payoffs $\mathcal{E} := \{(W_k, Y_k)\}_{k=1}^{K}$, for each of these pairs, the random payoff W_k is known to be no more risky than Y_k.

When imposing scale invariance on measures in $\mathfrak{R}$, one can demonstrate (see appendix A) that the set of risk measures reduces to the set of coherent risk measures, where scale invariance plays the role of "cash invariance".

$$\mathcal{R}_{Coh} := \left\{ \rho \in \mathfrak{R} \,\middle|\, \rho(\lambda \vec{Z}) = \lambda \rho(\vec{Z}) \,, \forall\, \vec{Z} \in \mathbb{R}^M, \forall\, \lambda \geq 0 \right\} \,.$$

Similarly, making the law invariance hypothesis will reduce the set of convex risk measure to the popular set of law-invariant convex risk measures:

$$\mathcal{R}_{Law} := \left\{ \rho \in \mathfrak{R} \,\middle|\, \rho(\vec{Z}_1) = \rho(\vec{Z}_2) \,, \forall\, Z_1 \equiv_F Z_2 \right\} \,,$$

where $\equiv_F$ stands for equal in distribution. Finally, the final hypothesis allows one to specify in more refined details a list of statements that the risk measure should respect:

$$\mathcal{R}_{El}(\mathcal{E}) := \left\{ \rho \in \mathfrak{R} \,\middle|\, \rho(\vec{W}_k) \leq \rho(\vec{Y}_k) \,, \forall\, k \in \{1, 2, \ldots, K\} \right\} \,.$$

The concept of scale invariance (a.k.a. cash-invariance) and law-invariance are the most important properties that have been discussed in recent literature on convex risk measures. Given that it might be unclear whether decision makers truly feel comfortable with the notion of scale invariance and whether they can specify in a precise way the probability measure F, we leave it optional to impose either of them or both. Note that while it is true that these assumptions might not always be well justified, as we will see later in practice they can be very useful for narrowing down what are the potential ρ and identifying decision that perform well. Note that both of these assumptions can also be approximated by requiring the decision maker to compare a sufficient list of random payoffs and including this information in the form of elicited comparisons.

In summary, we will consider the following four sets of risk measures, each of which involves different combination of the three hypotheses above.

$$\begin{aligned}
&\mathcal{R}_{El}(\mathcal{E}) \,, \\
&\mathcal{R}_{CE}(\mathcal{E}) := \mathcal{R}_{Coh} \cap \mathcal{R}_{El}(\mathcal{E}) \,, \\
&\mathcal{R}_{LE}(\mathcal{E}) := \mathcal{R}_{Law} \cap \mathcal{R}_{El}(\mathcal{E}) \,, \\
&\mathcal{R}_{CLE}(\mathcal{E}) := \mathcal{R}_{Coh} \cap \mathcal{R}_{Law} \cap \mathcal{R}_{El}(\mathcal{E}) \,.
\end{aligned}$$

Proposition 2.1. *Given any set $\mathcal{R} \in \{\mathcal{R}_{El}, \mathcal{R}_{CE}, \mathcal{R}_{LE}, \mathcal{R}_{CLE}\}$, the risk measure implied by $\varrho_{\mathcal{R}}(\vec{Z}) := \sup_{\rho \in \mathcal{R}} \rho(\vec{Z})$ is itself a member of $\mathcal{R}$.*

Proof. We expose the proof for the case $\mathcal{R} = \mathcal{R}_{CLE}$ which can easily be modified to address the other cases. Let's consider $\varrho_{\mathcal{R}}(\vec{Z}) := \sup_{\rho \in \mathcal{R}_{CLE}} \rho(\vec{Z})$, in order to show that $\varrho_{\mathcal{R}} \in \mathcal{R}_{CLE}$ one needs to show that it satisfies all the axioms of a coherent law-invariant risk measure plus the set of elicited comparison constraints. Firstly, $\varrho_{\mathcal{R}}$ is monotone since given $\vec{Z}_1 \geq \vec{Z}_2$, we have that

$$\varrho_{\mathcal{R}}(\vec{Z}_1) = \sup_{\rho \in \mathcal{R}_{CLE}} \rho(\vec{Z}_1) \leq \sup_{\rho \in \mathcal{R}_{CLE}} \rho(\vec{Z}_2) = \varrho_{\mathcal{R}}(\vec{Z}_2) \,,$$

since every risk measure in $\mathcal{R}_{CLE}$ respects monotonicity. Secondly, $\varrho_{\mathcal{R}}$ respects translation invariance since

$$\varrho_{\mathcal{R}}(\vec{Z}+c) = \sup_{\rho\in\mathcal{R}_{CLE}} \rho(\vec{Z}+c) = \sup_{\rho\in\mathcal{R}_{CLE}} \rho(\vec{Z}) - c = \varrho_{\mathcal{R}}(\vec{Z}) - c\,,$$

again mostly due to the fact that each risk measure in $\mathcal{R}_{CLE}$ respects translation invariance. Convexity is a little more complicated to demonstrate:

$$\begin{aligned}\varrho_{\mathcal{R}}(\theta\vec{Z}_1+(1-\theta)\vec{Z}_2) &= \sup_{\rho\in\mathcal{R}_{CLE}} \rho(\theta\vec{Z}_1+(1-\theta)\vec{Z}_2) \leq \sup_{\rho\in\mathcal{R}_{CLE}} \theta\rho(\vec{Z}_1)+(1-\theta)\rho(\vec{Z}_2)\\ &\leq \sup_{\rho\in\mathcal{R}_{CLE}} \theta\rho(\vec{Z}_1) + \sup_{\rho\in\mathcal{R}_{CLE}} (1-\theta)\rho(\vec{Z}_2) = \theta\varrho_{\mathcal{R}}(\vec{Z}_1)+(1-\theta)\varrho_{\mathcal{R}}(\vec{Z}_2)\,,\end{aligned}$$

given that each member of $\mathcal{R}_{CLE}$ respects convexity and that the supremum of a sum is always smaller than the sum of the supremums. Finally, the last three properties follow using similar arguments:

Scale invariance: $\varrho_{\mathcal{R}}(\lambda\vec{Z}) = \sup_{\rho\in\mathcal{R}_{CLE}} \rho(\lambda\vec{Z}) = \sup_{\rho\in\mathcal{R}_{CLE}} \lambda\rho(\vec{Z}) = \lambda\varrho_{\mathcal{R}}(\vec{Z})$

Law invariance: $\varrho_{\mathcal{R}}(\vec{Z}_1) = \sup_{\rho\in\mathcal{R}_{CLE}} \rho(\vec{Z}_1) = \sup_{\rho\in\mathcal{R}_{CLE}} \rho(\vec{Z}_2) = \varrho_{\mathcal{R}}(\vec{Z}_2)\,,\ \forall\, Z_1 \equiv_F Z_2$

Elicited comparisons: $\varrho_{\mathcal{R}}(\vec{W}_k) = \sup_{\rho\in\mathcal{R}_{CLE}} \rho(\vec{W}_k) \leq \sup_{\rho\in\mathcal{R}_{CLE}} \rho(\vec{Y}_k) = \varrho_{\mathcal{R}}(\vec{Y}_k)\,,\ \forall\, k\,.$

□

This first result is interesting for two reasons. First, it ensures that the robust framework that we propose compares random payoffs in a way that is coherent with what is known of the decision maker's perception. Furthermore, the fact that $\sup_{\rho\in\mathcal{R}}\rho(\cdot)$ is itself a risk measure is already a good indicator that it is possible to capture this measurement using a tractable representation. We will specifically identify such representations in the next three sections.

3 Optimizing Preference Robust Risk Measures

Our focus will initially be on identifying tractable methods for evaluating the worst-case risk measure $\varrho_{\mathcal{R}}(\vec{Z})$ and for optimizing problem (2) in a context where $\mathcal{R} = \mathcal{R}_{El}$, i.e. that the risk measure that captures the decision maker's perception is only known to be a convex risk measure together with the information about a list of pairwise comparisons ($\mathcal{E}$). The main machinery behind our work will be the manipulation of acceptance sets that is intimately associated to any convex risk measure. Let us recall some of its relevant definitions here.

Definition 3.1. *(Föllmer and Schied (2002)) Any given convex risk measure $\rho \in \mathfrak{R}$ induces an acceptance set*

$$\mathcal{A}_\rho := \{\vec{Z} \in \mathbb{R}^M \,|\, \rho(\vec{Z}) \leq 0\}\,,$$

that is convex and monotone in the sense that

$$\vec{Z}_1 \in \mathcal{A}_\rho, \vec{Z}_2 \geq \vec{Z}_1 \Rightarrow \vec{Z}_2 \in \mathcal{A}_\rho \ ,$$

and completely determines ρ *such that*

$$\rho(\vec{Z}) = \inf_{m\in\mathbb{R}}\{m \mid \vec{Z} + m \in \mathcal{A}_\rho\} \ .$$

Conversely, given a convex and monotone acceptance set $\mathcal{A} \subset \mathbb{R}^M$ *containing all points that are considered acceptable, the risk measure*

$$\rho_{\mathcal{A}}(\vec{Z}) := \inf_{m\in\mathbb{R}}\{m \mid \vec{Z} + m \in \mathcal{A}\}$$

is a convex risk measure. It is also known that if ρ *is a coherent risk measure, then* $\mathcal{A}_\rho$ *is convex conique and monotone, and conversely that any such acceptance set can be used to construct a coherent risk measure* $\rho_{\mathcal{A}}(\vec{Z})$*.*

Intuitively, the points in the set $\mathcal{A}_\rho$ are considered acceptable because they do not require additional capital in order to be risk free (i.e. that they are less risky than the zero payoff). Moreover, the definition of a risk measure through $\rho_{\mathcal{A}}(\vec{Z})$ has played an important role in the adoption of convex risk measure by appealing to common sense which expects that a risk measure should inform of the minimum amount of capital required to render a random payoff acceptable in terms of risk.

The key of our analysis lies in studying the worst-case risk measure $\varrho_{\mathcal{R}}(\vec{Z})$ under the form

$$\varrho_{\mathcal{R}}(\vec{Z}) = \sup_{\mathcal{A}:\rho_{\mathcal{A}}\in\mathcal{R}} \ \inf_{m\in\mathbb{R}}\{m \,|\, \vec{Z} + m \in \mathcal{A}\} \ ,$$

where $\varrho_{\mathcal{R}}(\vec{Z})$ returns the largest amount of capital that could be required to guarantee that Z will be considered risk free. Since the optimal value of the above form can be attained by a convex risk measure (due to $\varrho_{\mathcal{R}} \in \mathcal{R}$), the problem can be equivalently stated as searching for a worst-case acceptance set $\mathcal{A}$ such that $\rho_{\mathcal{A}} = \varrho_{\mathcal{R}}$. All our results presented in the following sections are established based on the findings that such a set $\mathcal{A}$ can be efficiently constructed. We will initially address how to optimize problem (2) under convex risk measures and elicited comparisons, $\mathcal{R}_{El}$. We will then quickly address the case of coherent risk measures, $\mathcal{R}_{CE}$, and complete the story with the additional notion of law invariance, $\mathcal{R}_{LE}$ and $\mathcal{R}_{CLE}$.

3.1 The case of convex risk measures

We start this section by considering a special case of $\varrho_{\mathcal{R}}(\vec{Z})$ when $\mathcal{R} = \mathcal{R}_{El}(\mathcal{E})$ with $\mathcal{E} := \{(W_k, 0)\}_{k=1}^K$. We show how the value of $\varrho_{\mathcal{R}}(\vec{Z})$ in this case can be obtained by solving a finite dimensional linear program. This will provide valuable insights in order to obtain our results for the more general form.

Proposition 3.2. *Given a set of acceptable random payoffs* $\{W_k\}_{k=1}^K$ *and any random payoff* Z*, the value* $\varrho_{\mathcal{R}_{El}(\mathcal{E}_0)}(\vec{Z})$ *with* $\mathcal{E}_0 := \{(W_k, 0)\}_{k=1}^K$ *and* $\mathcal{R}_{El}(\mathcal{E}_0) \neq \emptyset$ *is obtained* *as the optimal value of the following linear program:*

$$\varrho_{\mathcal{R}_{El}(\{(W_k,0)\}_{k=1}^K)}(\vec{Z}) = \min_{t,\theta} \quad t, \tag{4a}$$

$$s.t. \quad \vec{Z} + t \geq \mathbb{W}\theta, \tag{4b}$$

$$\mathbf{1}^\top \theta \leq 1, \ \theta \geq 0, \tag{4c}$$

where $t \in \mathbb{R}$, $\theta \in \mathbb{R}^K$, $\vec{Z} \in \mathbb{R}^M$ is the vector representing the outcomes of Z, and $\mathbb{W}$ is a $M \times K$ matrix composed of the random payoff vectors $\{\vec{W}_k\}_{k=1,\ldots,K}$ as its column vectors, i.e. $\mathbb{W} = [\vec{W}_1 \vec{W}_2 \cdots \vec{W}_K]$. Furthermore, in order to verify that $\mathcal{R}_{El}(\{(W_k,0)\}_{k=1}^K) \neq \emptyset$, one can simply check that the optimal value of problem (4) *is zero when $\vec{Z} = \vec{0}$.*

Proof. Following Definition 3.1, all convex risk measures can be described through their acceptance sets. We can therefore focus on the set of acceptance set candidates

$$\mathbb{A} := \{\mathcal{A} \subseteq \mathbb{R}^M \,|\, \rho_{\mathcal{A}} \in \mathcal{R}_{El}(\{(W_k,0)\}_{k=1}^K)\} ,$$

and we wish to evaluate $\sup_{\mathcal{A}\in\mathbb{A}} \rho_{\mathcal{A}}(\vec{Z})$. It is not hard to see that for any two sets $\mathcal{A}_1 \in \mathbb{A}$ and $\mathcal{A}_2 \in \mathbb{A}$, if $\mathcal{A}_1 \subseteq A_2$ then, we have that $\rho_{\mathcal{A}_1}(\vec{Z}) \geq \rho_{\mathcal{A}_2}(\vec{Z})$. We will show that

$$\mathcal{A}^* := \{\vec{Z} \in \mathbb{R}^M \,|\, \exists\theta \in \mathbb{R}^K, \, \vec{Z} \geq \sum_{k=1}^K \theta_k \vec{W}_k, \ \sum_k \theta_k \leq 1, \ \ \theta \geq 0\}.$$

is both a subset of any other sets of $\mathbb{A}$ and a member of $\mathbb{A}$. From there one can easily conclude that

$$\rho_{\mathcal{A}^*} \leq \sup_{\mathcal{A}\in\mathbb{A}} \rho_{\mathcal{A}}(\vec{Z}) \leq \rho_{\mathcal{A}^*} \ .$$

Problem (4) follows from the definition of $\rho_{\mathcal{A}^*}$.

To show that $\mathcal{A}^* \subseteq \mathcal{A}$ for all $\mathcal{A} \in \mathbb{A}$, one simply needs to observe that $\mathcal{A}^*$ is the set of points that dominate some convex combinations of the set $\vec{0} \cup \{\vec{W}_k\}_{k=1}^K$ and must therefore be included in any convex monotone set containing $\vec{0} \cup \{\vec{W}_k\}_{k=1}^K$. Note that the zero vector $\vec{0}$ is implicitly part of all acceptable random payoff since $\mathcal{R}_{El}(\mathcal{E}) \subseteq \mathfrak{R}$ which imposes that $\rho(0) = 0$.

In the other direction, we need to verify that $\mathcal{A}^*$ contains the points $\{\vec{W}_k\}_{k=1}^K$, and is convex and monotone and evaluate $\rho_{\mathcal{A}^*}(0) = 0$. The first part is obvious. Convexity is verified as follows. Given $\vec{Z}_1 \in \mathcal{A}^*$ and $\vec{Z}_2 \in \mathcal{A}^*$, there must exist two convex combination θ^1 and θ^2 such that $\vec{Z}_i \geq \sum_k \theta_k^i \vec{W}_k$ when $i = 1, 2$. Therefore, given any $0 \leq \alpha \leq 1$, one can confirm that

$$\alpha\vec{Z}_1 + (1-\alpha)\vec{Z}_2 \geq \alpha \sum_k \theta_k^1 \vec{W}_k + (1-\alpha) \sum_k \theta_k^2 \vec{W}_k = \sum_k (\alpha\theta_k^1 + (1-\alpha)\theta_k^2)\vec{W}_k \ ,$$

thus that $\alpha\vec{Z}_1 + (1-\alpha)\vec{Z}_2$ is in $\mathcal{A}^*$.

Secondly, monotonicity is verified in a similar way. Given any $\vec{Z}_1 \in \mathcal{A}^*$, there exists a convex combination θ^1 such that $\vec{Z}_1 \geq \sum_k \theta_k^1 \vec{W}_k$. Hence, using the same θ^1, we can verify that any $\vec{Z}_2 \geq \vec{Z}_1$ satisfies

$$\vec{Z}_2 \geq \vec{Z}_1 \geq \sum_k \theta_k^1 \vec{W}_k \ ,$$

and is therefore also a member of $\mathcal{A}^*$.

Finally, by construction $\rho_{\mathcal{A}^*}(\vec{0}) \leq 0$ since $\vec{0} \in \mathcal{A}^*$. Also, since $\mathcal{A}^* \subseteq \mathcal{A}$ for all $\mathcal{A} \in \mathbb{A}$, it must be that $\rho_{\mathcal{A}^*}(0) \geq \rho_{\mathcal{A}}(0) = 0$, or otherwise one just established that $\mathcal{R}_{El}(\{(W_k, 0)\}_{k=1}^K) = \emptyset$. □

The intuition here is simple. Among all the candidate acceptance sets associated to convex risk measures and that are consistent with the stated acceptable points, the worst-case acceptance set is simply the smallest monotone polyhedron that covers the convex hull of these acceptance points.

We are now ready for a more interesting result which shows that the preference robust risk minimization problem can be solved efficiently although it involves a worst-case analysis over an infinite dimensional space $\mathcal{R}$, or alternatively over a set $\mathbb{A}$ of infinite size.

Proposition 3.3. *Given a set $\mathcal{E}$ of K comparisons, let the set $\{X_j\}_{j=1}^J := \vec{0} \cup \bigcup_{k=1}^K \{W_k, Y_k\}$ be the set of all random payoffs involved in one of the elicited comparison and the zero payoff which we identify as X_1. The preference robust risk minimization problem (2) with $\mathcal{R}_{El}(\mathcal{E})$ is equivalent to the optimization problem:*

$$\min_{x \in \mathcal{X}, t, \theta} \quad t, \tag{5a}$$

$$\text{subject to} \quad \vec{Z}(x) + t \geq \mathbb{X}\theta + \bar{\delta}^\top \theta, \tag{5b}$$

$$\vec{1}^\top \theta = 1, \; \theta \geq 0. \tag{5c}$$

where $t \in \mathbb{R}$, $\theta \in \mathbb{R}^J$, $\mathbb{X} := [\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_J]$, and where $\bar{\delta} \in \mathbb{R}^J$ is the optimal solution of the linear program

$$\max_{\delta, \{y_j\}_{j=1}^J} \quad \sum_{j=1}^J \delta_j, \tag{6a}$$

$$\text{subject to} \quad \delta_i \leq \delta_j \,, \forall\, (i,j) \in \bar{\mathcal{E}} \tag{6b}$$

$$(\vec{X}_i - \vec{X}_j)^\mathsf{T} y_j + \delta_i - \delta_j \geq 0 \,, \forall\, i \neq j \tag{6c}$$

$$\vec{1}^\top y_j = 1, \; y_j \geq 0, \; \forall j \tag{6d}$$

$$\delta_1 = 0 \,, \tag{6e}$$

where each $y_j \in \mathbb{R}^M$ and where $\bar{\mathcal{E}}$ is the set of edges in the partial ordering of $\{X_j\}_{j=1}^J$ described by the elicited comparisons: i.e.

$$\bar{\mathcal{E}} := \{(i,j) \in \{1, 2, \ldots, J\}^2 \,|\, (X_i, X_j) \in \mathcal{E}\} \,.$$

Moreover, problem (5) is a convex optimization problem when each term $(\vec{Z}(x))_i$ is a concave function of x, and problem (6) is feasible if and only if $\mathcal{R}_{El}(\mathcal{E})$ is non-empty.

While we defer the proof to appendix B, it is worth mentioning that the steps consist, as was the case for proposition 3.2, in showing that the acceptance set

$$\mathcal{A}(\mathbb{X}, \delta) := \{\vec{Z} \in \mathbb{R}^M \,|\, \exists \theta \in \mathbb{R}^K, \; \vec{Z} \geq \sum_{k=1}^K \theta_k (\vec{X}_k + \delta_k), \;\; \sum_k \theta_k = 1, \;\; \theta \geq 0\}$$

with $\delta = \bar{\delta}$ is always a worst-case acceptance set when evaluating the risk of any $Z(x)$. In particular, the set $\mathcal{A}(\mathbb{X}, \delta)$ takes the shape of the monotone convex hull of the points $\{\vec{X}_j + \delta_j\}_{j=1}^J$. In problem (6), the constraints (6c) and (6d) ensure that the values of δ satisfy $\rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_j) = \delta_j$, which together with the constraints $\delta_i \leq \delta_j$ in (6b) exactly impose that $\rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_i) \leq \rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_j)$. The constraint that $\delta_1 = 0$ ensures that the risk measure evaluated at zero payoff gives zero.

Example 3.1. In a portfolio optimization example, we have that $Z(x) = \xi^\top x$ with $\xi : \Omega \to \mathbb{R}^n$ the random vector of uncertain investment return rates and $x \in \mathbb{R}^n$ the wealth allocation vector. Hence, the preference robust risk minimization problem reduces to

$$\begin{aligned} \min_{x \in \mathcal{X}, t, \theta} \quad & t, \\ \text{subject to} \quad & \xi(\omega)^\top x + t \geq \textstyle\sum_j X_j(\omega)\theta_j + \bar{\delta}^\top \theta \,, \forall\, \omega \in \Omega \\ & \vec{1}^\top \theta = 1,\ \theta \geq 0, \end{aligned}$$

where $\xi(\omega)$ is the vector of returns achieved for each asset under outcome ω and $X_j(\omega)$ is the payoff returned for ω by the random payoff X_j used in the elicited comparisons. The set $\mathcal{X}$ is defined via linear portfolio constraints. For simplicity, we consider $\mathcal{X} := \{x \mid \vec{1}^\top x = W\}$, where W is the total wealth to be invested. One should note that this optimization problem takes the shape of a linear program with $O(n+K)$ decision variables and $O(M+K)$ constraints. Yet, in order to identify $\bar{\delta}$, problem (6) must first be solved which also takes the form of a linear program but with $O(MK)$ decision variables and $O(K^2)$ constraints. In both cases, the linear programs can be considered of reasonable sizes and solution times should scale reasonably well with the problem dimensions.

3.2 The case of coherent risk measures

Following closely the steps we took to show the tractability for the case of convex risk measures, we can prove quite straightforwardly the tractability for the case of coherent risk measures. In particular, we can carry over the same intuition regarding the construction of the worst-case acceptance set given in Proposition 3.2 to the case of coherent risk measures. In the former, given a set of acceptance points $\{W_k\}_{k=1}^K$ the worst-case acceptance set was shown to be the smallest monotone polyhedron that covers all the acceptance points. In the latter, as any candidate acceptance set $\mathcal{A}$ should additionally satisfies $\vec{Z} \in \mathcal{A} \Rightarrow \lambda\vec{Z} \in \mathcal{A}$, which is by definition a cone, the worst-case acceptance set would thus be the smallest monotone polyhedral cone that covers all the acceptance points $\{W_k\}_{k=1}^K$. Recall that such a polyhedral cone admits the representation of $\{\vec{Y} \in \mathbb{R}^M \mid \vec{Y} \geq \sum_{k=0}^K \vec{W}_k\theta_k,\ \theta_k \geq 0,\ \vec{W}_0 = \vec{0}\}$. A rigorous proof that validates this intuition is omitted here as it is almost identical to the proof of Proposition 3.2.

Once we confirmed the structure of the worst-case acceptance set, the preference robust risk minimization problem (2) with $\mathcal{R}_{CE}(\mathcal{E})$ can be found as tractable as with $\mathcal{R}_{El}(\mathcal{E})$. In particular, we arrive at the following result, which can be proved following identical steps as in appendix B.

Proposition 3.4. *Given a set $\mathcal{E}$ of K comparisons, let the set $\{X_j\}_{j=1}^J := \vec{0} \cup \bigcup_{k=1}^K \{W_k, Y_k\}$ be the support set of all random payoffs involved in one of the elicited comparison and the zero payoff which we identify as X_1. The preference robust risk minimization problem* (2) *with $\mathcal{R}_{CE}(\mathcal{E})$ is equivalent to the optimization problem:*

$$\min_{x\in\mathcal{X},t,\theta} \quad t, \tag{7a}$$

$$\text{subject to} \quad \vec{Z}(x) + t \geq \mathbb{X}\theta + \bar{\delta}^\top\theta, \tag{7b}$$

$$\theta \geq 0, \tag{7c}$$

where $t \in \mathbb{R}$, $\theta \in \mathbb{R}^J$, $\mathbb{X} := [\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_J]$, and where $\bar{\delta} \in \mathbb{R}^J$ is the optimal solution of the linear program

$$\max_{\delta,\{y_j\}_{j=1}^J} \quad \sum_{j=1}^J \delta_j, \tag{8a}$$

$$\text{subject to} \quad \delta_i \leq \delta_j \,,\, \forall\,(i,j) \in \bar{\mathcal{E}} \tag{8b}$$

$$-\vec{X}_j^\intercal y_j - \delta_j \geq 0 \,,\, \forall\, j = 1,\ldots,J \tag{8c}$$

$$\vec{X}_i^\intercal y_j + \delta_i \geq 0 \,,\, \forall\, i \neq j \tag{8d}$$

$$\vec{1}^\top y_j = 1,\ y_j \geq 0,\ \forall j \tag{8e}$$

$$\delta_1 = 0 \,, \tag{8f}$$

where each $y_j \in \mathbb{R}^M$, and where $\bar{\mathcal{E}}$ is the set of edges in the partial ordering of $\{X_j\}_{j=1}^J$ described by the elicited comparisons: i.e.

$$\bar{\mathcal{E}} := \{(i,j) \in \{1,2,\ldots,J\}^2 \,|\, (X_i, X_j) \in \mathcal{E}\} \,.$$

Moreover, problem (7) *is a convex optimization problem when each term $(\vec{Z}(x))_i$ is a concave function of x, and problem* (8) *is feasible if and only if $\mathcal{R}_{CE}(\mathcal{E})$ is non-empty.*

Comparing Proposition 3.3 and 3.4, one can see they differ in that the constraint $\mathbf{1}^\top\theta = 1$ in Proposition 3.3 is dropped in Proposition 3.4 due to the use of of a polyhedral cone, rather than a polyhedron, to represent the worst-case acceptance set.

Example 3.2. The portfolio optimization example in Example 3.1 can be reformulated as the following preference robust risk minimization problem based on Proposition 3.4:

$$\begin{aligned} \min_{x\in\mathcal{X},t,\theta} \quad & t, \\ \text{subject to} \quad & \xi(\omega)^T x + t \geq \textstyle\sum_j X_j(\omega)\theta_j + \bar{\delta}^\top\theta \,,\, \forall\,\omega \in \Omega \\ & \theta \geq 0. \end{aligned}$$

This is a linear program with $O(n+K)$ decision variables and $O(M+K)$ constraints. The parameter $\bar{\delta}$ can be generated from the linear program (8) that has $O(MK)$ decision variables and $O(K^2)$ constraints. Comparing with Example 3.1, one can see the sizes of the above linear program and the linear program (8) grow in the same order as the linear programs formulated based on convex risk measures.

3.3 The case of law-invariant risk measures

To incorporate further the law invariance hypothesis, we now consider the sets of risk measures

$$\mathcal{R}_{LE}(\mathcal{E}) := \mathcal{R}_{Law} \cap \mathcal{R}_{El}(\mathcal{E}) ,$$
$$\mathcal{R}_{CLE}(\mathcal{E}) := \mathcal{R}_{Coh} \cap \mathcal{R}_{Law} \cap \mathcal{R}_{El}(\mathcal{E}) .$$

We start by focusing on the following setting, which helps best understand the complexity involved in dealing with these sets and the techniques required to resolve the complexity.

Assumption 3.5. *The probability measure F associated to the discrete probability space is the uniform measure, i.e. $P(\{\omega_i\}) = 1/M$ for any $\omega_i \in \Omega$. Recall that $|\Omega| = M$.*

As would be shown later, in this setting the hypothesis of law invariance is tightly connected to the notion of permutation which we define in detail below.

Definition 3.6. *A permutation over M elements is a bijection function $\sigma : \{1, ..., M\} \to \{1, ..., M\}$. We call the random variable permutation operator σ, the operator that permutes the values associated to each outcome of a random variable according to the bijection σ. Mathematically speaking, we have that $(\sigma(\vec{X}))_i = X(\omega_{\sigma^{-1}(i)})$. We will denote with Σ the set of all random variable permutation operators in the Ω outcome space.*

Actually, under the law invariance assumption, one quickly realizes that any preference of the type $W \succeq Y$ implies that $\sigma(W) \succeq \sigma'(Y)$ for all pair (σ, σ') of permutation operators. This is simply due to the fact that $\sigma(Z) \equiv_F Z$ (i.e. in distribution) when F is the uniform distribution no matter what permutation operator is used. Hence we must have that $\sigma(W) \succeq W \succeq Y \succeq \sigma'(Y)$. This leads us to consider an augmented set of elicited comparison

$$\Sigma(\mathcal{E}) := \{(W, Y) | \exists \sigma' \in \Sigma, \sigma' \in \Sigma, (\sigma(W), \sigma'(Y)) \in \mathcal{E}\}.$$

Risk measures that are known to comply with the elicited comparisons $\mathcal{E}$ and to be law invariant should also respect preference orderings described in the augmented set $\Sigma(\mathcal{E})$. What is more interesting is actually the reverse statement, which we prove in the following lemma that robust risk measures if constructed directly based on the augmented set $\Sigma(\mathcal{E})$ would coincide with the law invariant measures based on $\mathcal{E}$. This result shifts the complexity of incorporating the hypothesis of law invariance from the set of law-invariant risk measures $\mathcal{R}_{LE}(\mathcal{E})$ to the set of convex risk measures $\mathcal{R}_{El}(\Sigma(\mathcal{E}))$ (or from $\mathcal{R}_{CLE}(\mathcal{E})$ to $\mathcal{R}_{CE}(\Sigma(\mathcal{E}))$).

Lemma 3.7. *The preference robust risk measure $\varrho_{\mathcal{R}_{LE}(\mathcal{E})}(\vec{Z}) := \sup_{\rho \in \mathcal{R}_{LE}(\mathcal{E})} \rho(\vec{Z})$ is equivalent to $\varrho_{\mathcal{R}_{LE}(\Sigma(\mathcal{E}))}(\vec{Z})$ and $\varrho_{\mathcal{R}_{El}(\Sigma(\mathcal{E}))}(\vec{Z})$. Similarly, the preference robust risk measure $\varrho_{\mathcal{R}_{CLE}(\mathcal{E})}(\vec{Z})$ is equivalent to $\varrho_{\mathcal{R}_{CLE}(\Sigma(\mathcal{E}))}(\vec{Z})$ and $\varrho_{\mathcal{R}_{CE}(\Sigma(\mathcal{E}))}(\vec{Z})$.*

Proof. In the case of robust convex risk measures, one can first establish that $\varrho_{\mathcal{R}_{LE}(\mathcal{E})}(\vec{Z}) = \varrho_{\mathcal{R}_{LE}(\Sigma(\mathcal{E}))}(\vec{Z})$ if it can be verified that $\mathcal{R}_{LE}(\mathcal{E}) = \mathcal{R}_{LE}(\Sigma(\mathcal{E}))$. This can be done with the following argument. Given that the later set imposes more constraints on the risk measure, it must be that $\mathcal{R}_{LE}(\mathcal{E}) \supseteq \mathcal{R}_{LE}(\Sigma(\mathcal{E}))$. In fact, given any $\rho \in \mathcal{R}_{LE}(\mathcal{E})$, one can confirm that it is a member of $\mathcal{R}_{LE}(\Sigma(\mathcal{E}))$ since for any pair $(W', Y') \in \Sigma(\mathcal{E})$, there must be a comparison $(W, Y) \in \mathcal{E}$ and a pair of permutation operators (σ_W, σ_Y) such that $W' = \sigma_W(W)$ and $Y' = \sigma_Y(Y)$, hence that

$$\rho(\vec{W}') = \rho(\sigma_W(\vec{W})) = \rho(\vec{W}) \geq \rho(\vec{Y}) = \rho(\sigma_Y(\vec{Y})) = \rho(\vec{Y}') \,,$$

where the second and third equalities are due to the fact that ρ is law invariant.

Next, we prove that $\varrho_{\mathcal{R}_{LE}(\Sigma(\mathcal{E}))}(\vec{Z})$ is equivalent to $\varrho_{\mathcal{R}_{El}(\Sigma(\mathcal{E}))}(\vec{Z})$. Since $\mathcal{R}_{LE}(\Sigma(\mathcal{E})) \subseteq \mathcal{R}_{El}(\Sigma(\mathcal{E}))$, we must have that $\varrho_{\mathcal{R}_{LE}(\Sigma(\mathcal{E}))}(\vec{Z}) \leq \varrho_{\mathcal{R}_{El}(\Sigma(\mathcal{E}))}(\vec{Z})$. To get the reverse, we can consider that given any $\rho \in \mathcal{R}_{El}(\Sigma(\mathcal{E}))$, one can construct a new risk measure $\rho_\Sigma(\vec{Z}) := \max_{\sigma\in\Sigma} \rho(\sigma(\vec{Z}))$. It is clear that $\rho_\Sigma \in \mathcal{R}_{LE}(\Sigma(\mathcal{E}))$ and that $\rho_\Sigma(\vec{Z}) \geq \rho(\vec{Z})$. [8] We can conclude that

$$\varrho_{\mathcal{R}_{El}(\Sigma(\mathcal{E}))}(\vec{Z}) = \sup_{\rho\in\mathcal{R}_{El}(\Sigma(\mathcal{E}))} \rho(\vec{Z}) \leq \sup_{\rho\in\mathcal{R}_{LE}(\Sigma(\mathcal{E}))} \rho(\vec{Z}) = \varrho_{\mathcal{R}_{LE}(\Sigma(\mathcal{E}))}(\vec{Z}) \,.$$

Hence, we know that

$$\varrho_{\mathcal{R}_{LE}(\Sigma(\mathcal{E}))}(\vec{Z}) \leq \varrho_{\mathcal{R}_{El}(\Sigma(\mathcal{E}))}(\vec{Z}) \leq \varrho_{\mathcal{R}_{LE}(\Sigma(\mathcal{E}))}(\vec{Z}) \,.$$

The proof for the case of robust coherent risk measures is similar. □

The complexity of constructing risk measures $\varrho_{\mathcal{R}_{El}(\Sigma(\mathcal{E}))}(\vec{Z})$ (resp. $\varrho_{\mathcal{R}_{CE}(\Sigma(\mathcal{E}))}(\vec{Z})$) lies in the size of the augmented set $\Sigma(\mathcal{E})$, which grows exponentially with respect to the number of elicited comparisons. Our next result is to show that the convex optimization problems formulated based on Proposition 3.3 (resp. Proposition 3.4) for $\varrho_{\mathcal{R}_{El}(\Sigma(\mathcal{E}))}(\vec{Z})$ (resp. $\varrho_{\mathcal{R}_{CE}(\Sigma(\mathcal{E}))}(\vec{Z})$) can be reduced to problems whose size no longer depend on the size of the set of permutations and grow polynomially with the number of elicited comparisons. In the following lemma, we present first the reduced formulation for the offline optimization problem (6) in Proposition 3.3 with the augmented set $\Sigma(\mathcal{E})$.

Lemma 3.8. *The optimization problem*

$$\max_{\delta, \{y_{j,\sigma}\}_{j=1,\sigma\in\Sigma}^{J}} \quad \sum_{j=1}^{J}\sum_{\sigma\in\Sigma} \delta_{j,\sigma}, \tag{9a}$$

$$\text{subject to} \quad \delta_{i,\sigma} \leq \delta_{j,\sigma'} \,,\, \forall\, (i,j) \in \bar{\mathcal{E}},\, \forall\, \sigma \in \Sigma,\, \forall\, \sigma' \in \Sigma \tag{9b}$$

$$(\sigma(\vec{X}_i) - \sigma'(\vec{X}_j))^\top y_{j,\sigma'} + \delta_{i,\sigma} - \delta_{j,\sigma'} \geq 0 \,,\, \forall\, i \neq j,\, \forall\, \sigma \in \Sigma,\, \forall\, \sigma' \in \Sigma \tag{9c}$$

$$\vec{1}^\top y_{j,\sigma} = 1,\; y_{j,\sigma} \geq 0,\, \forall\, j,\, \forall\, \sigma \in \Sigma \tag{9d}$$

$$\delta_{1,\sigma} = 0,\, \forall\, \sigma \in \Sigma \,, \tag{9e}$$

[8] We use the fact $W \equiv_F Y \Leftrightarrow \exists \sigma : \vec{Y} = \sigma(\vec{W})$ here. The direction $\Leftarrow$ is clear as discussed. To see $\Rightarrow$, note that following Assumption 3.5, any law F^X can be expressed by $P(X = x_k) = |\{\omega \in \Omega \mid X(\omega) = x_k\}|/M, \forall k$. $W \equiv_F Y$ implies that $|\{\omega \in \Omega \mid W(\omega) = x_k\}| = |\{\omega \in \Omega \mid Y(\omega) = x_k\}|, \forall k$. $\vec{Y} = \sigma(\vec{W})$ easily follows.

where $\bar{\mathcal{E}}$ *is the set of edges in the partial ordering of* $\{X_j\}_{j=1}^J$ *described by the elicited comparisons: i.e.*

$$\bar{\mathcal{E}} := \{(i,j) \in \{1,2,\ldots,J\}^2 \,|\, (X_i, X_j) \in \mathcal{E}\} \,,$$

has an optimal solution for which $\delta_{j,\sigma} = \delta_{j,\sigma'}$ *for all* j*, all* $\sigma \in \Sigma$*, and all* $\sigma' \in \Sigma$*. Furthermore, it reduces to solving the following linear programming problem*

$$\begin{aligned}
\max_{\delta, \{y_j\}_{j=1}^J, \{v_{i,j}, w_{i,j}\}_{i=1,j=1}^{i=J,j=J}} \quad & |\Sigma| \sum_{j=1}^J \delta_j, && \text{(10a)}\\
\text{subject to} \quad & \delta_i \leq \delta_j \,,\, \forall\,(i,j) \in \bar{\mathcal{E}} && \text{(10b)}\\
& \vec{1}^\top v_{i,j} + \vec{1}^\top w_{i,j} - \vec{X}_j^\top y_j + \delta_i - \delta_j \geq 0 \,,\, \forall\, i \neq j && \text{(10c)}\\
& \vec{X}_i y_j^\top - v_{i,j}\vec{1}^\top - \vec{1} w_{i,j} \geq 0 \,,\, \forall\, i \neq j && \text{(10d)}\\
& \vec{1}^\top y_j = 1,\ y_j \geq 0 \,,\, \forall\, j && \text{(10e)}\\
& \delta_1 = 0 \,, && \text{(10f)}
\end{aligned}$$

where each $y_j \in \mathbb{R}^M$*, each* $v_{i,j} \in \mathbb{R}^M$*, and each* $w_{i,j} \in \mathbb{R}^M$*.*

Proof. To demonstrate this lemma, we will prove that given any feasible solution $(\delta, \{y_{j,\sigma}\}_{j=1,\sigma\in\Sigma}^J)$, one can construct a feasible solution

$$\bar{\delta}_{i,\sigma} := \frac{1}{|\Sigma|} \sum_{\sigma' \in \Sigma} \delta_{i,\sigma'},\ \forall \sigma \in \Sigma,\ \ \forall i = 1, ..., J$$

$$\bar{y}_{i,\sigma} := \frac{1}{|\Sigma|} \sigma \left(\sum_{\sigma' \in \Sigma} {\sigma'}^{-1}(y_{i,\sigma'}) \right) \,,\ \forall \sigma \in \Sigma,\ \ \forall i = 1, ..., J$$

that achieves the same objective value and has the added property that the value of $\bar{\delta}_{i,\sigma}$ is the same for all permutations.

First, the objective is necessarily the same since

$$\begin{aligned}
\sum_{j=1}^J \sum_{\sigma\in\Sigma} \bar{\delta}_{j,\sigma} &= \sum_{j=1}^J \sum_{\sigma\in\Sigma} \frac{1}{|\Sigma|} \sum_{\sigma'\in\Sigma} \delta_{j,\sigma'} \\
&= \sum_{j=1}^J \sum_{\sigma'\in\Sigma} \sum_{\sigma\in\Sigma} \frac{1}{|\Sigma|} \delta_{j,\sigma'} \\
&= \sum_{j=1}^J \sum_{\sigma'\in\Sigma} \delta_{j,\sigma'} \,.
\end{aligned}$$

Next, we confirm one constraint at a time that each of constraints (9b) to (9e) are satisfied. For constraint (9b), we have that:

$$\bar{\delta}_{i,\sigma} = \frac{1}{|\Sigma|} \sum_{\sigma''\in\Sigma} \delta_{i,\sigma''} \leq \frac{1}{|\Sigma|} \sum_{\sigma'''\in\Sigma} \delta_{j,\sigma'''} = \bar{\delta}_{j,\sigma'} \,.$$

In the case of constraint (9c), the work is a bit more tedious

$$
\begin{aligned}
&(\sigma(\vec{X}_i) - \sigma'(\vec{X}_j))^\mathsf{T}\bar{y}_{j,\sigma'} + \bar{\delta}_{i,\sigma} - \bar{\delta}_{j,\sigma'} \\
&\qquad = (\sigma(\vec{X}_i) - \sigma'(\vec{X}_j))^\mathsf{T}\frac{1}{|\Sigma|}\sigma'\left(\sum_{\sigma''\in\Sigma}\sigma''^{-1}(y_{j,\sigma''})\right) + \bar{\delta}_{i,\sigma} - \bar{\delta}_{j,\sigma'} \\
&\qquad = \frac{1}{|\Sigma|}\sum_{\sigma''\in\Sigma}(\sigma'^{-1}(\sigma(\vec{X}_i)) - \vec{X}_j)^\mathsf{T}\sigma''^{-1}(y_{j,\sigma''}) + \bar{\delta}_{i,\sigma} - \bar{\delta}_{j,\sigma'} \\
&\qquad = \frac{1}{|\Sigma|}\sum_{\sigma''\in\Sigma}(\sigma''(\sigma'^{-1}(\sigma(\vec{X}_i))) - \sigma''(\vec{X}_j))^\mathsf{T}(y_{j,\sigma''}) + \bar{\delta}_{i,\sigma} - \bar{\delta}_{j,\sigma'} \\
&\qquad \geq \frac{1}{|\Sigma|}\sum_{\sigma''\in\Sigma}\delta_{j,\sigma''} - \delta_{i,\sigma''\circ\sigma'^{-1}\circ\sigma} + \bar{\delta}_{i,\sigma} - \bar{\delta}_{j,\sigma'} = 0 \;,
\end{aligned}
$$

where we used the fact that

$$
(\sigma''(\sigma'^{-1}(\sigma(\vec{X}_i))) - \sigma''(\vec{X}_j))^\mathsf{T}(y_{j,\sigma''}) \geq \delta_{j,\sigma''} - \delta_{i,\sigma''\circ\sigma'^{-1}\circ\sigma}
$$

and in the last step we used the fact the sum is over all possible permutations. Finally, constraints (9d) and (9e) can easily be verified.

$$
\vec{1}^\top\bar{y}_{j,\sigma} = \frac{1}{|\Sigma|}\vec{1}^\top\sigma\left(\sum_{\sigma'\in\Sigma}\sigma'^{-1}(y_{j,\sigma'})\right) = \frac{1}{|\Sigma|}\sum_{\sigma'\in\Sigma}\vec{1}^\top\sigma'^{-1}(y_{j,\sigma'}) = \frac{1}{|\Sigma|}\sum_{\sigma'\in\Sigma}\vec{1}^\top y_{j,\sigma'} = 1
$$

$$
\bar{\delta}_{1,\sigma} = \frac{1}{|\Sigma|}\sum_{\sigma'\in\Sigma}\delta_{1,\sigma'} = 0 \;.
$$

This completes the proof that problem (9) has an optimal solution for which $\delta_{j,\sigma} = \delta_{j,\sigma'}$ for all j, all $\sigma \in \Sigma$, and all $\sigma' \in \Sigma$.

Given the existence of an optimal solution with such structure it is possible to simplify the problem by optimizing only over all $\delta_{j,\sigma} = \delta_j$ and $y_{i,\sigma} := \sigma(y_i)$. This gives rise to the following problem:

$$
\begin{aligned}
\max_{\delta,\{y_j\}_{j=1}^J} \quad & \sum_{j=1}^J\sum_{\sigma\in\Sigma}\delta_j, \\
\text{subject to} \quad & \delta_i \leq \delta_j \;,\; \forall\,(i,j) \in \bar{\mathcal{E}} \\
& (\sigma(\vec{X}_i) - \vec{X}_j)^\mathsf{T}y_j + \delta_i - \delta_j \geq 0 \,,\; \forall\, i \neq j,\; \forall\,\sigma \in \Sigma \\
& \vec{1}^\top y_j = 1,\; y_j \geq 0,\; \forall\, j \\
& \delta_1 = 0 \;,
\end{aligned}
\tag{11}
$$

where we used the fact that

$$
\vec{1}^\top\sigma(y_j) = \vec{1}^\top y_j
$$

and the fact that

$$
(\sigma(\vec{X}_i) - \sigma'(\vec{X}_j))^\mathsf{T}\sigma'(y_j) + \delta_i - \delta_j = (\sigma'^{-1}(\sigma(\vec{X}_i)) - \vec{X}_j)^\mathsf{T}y_j + \delta_i - \delta_j \;.
$$

We are left with constraint (11) which can be stated as

$$\min_{\sigma\in\Sigma} \sigma(\vec{X}_i)^\intercal y_j - \vec{X}_j^\intercal y_j + \delta_i - \delta_j \geq 0 \,,\, \forall\, i \neq j \;, \tag{12}$$

and that we will reduce using duality theory.

Consider that $\min_{\sigma\in\Sigma} \sigma(\vec{X}_i)^\intercal y_j$ is equal to the optimal value of the optimization problem

$$\begin{aligned} \min_{Q} \quad & y^\top Q\vec{X} \\ \text{subject to} \quad & Q^\top\vec{1} = \vec{1} \\ & Q\vec{1} = \vec{1} \\ & Q_{k,l} \in \{0,1\}\,,\;\; k = 1,...,M,\;\; l = 1,...,M \end{aligned}$$

where $Q \in \mathbb{R}^{M\times M}$. Due to the result of Birkhoff (1946), the above problem, also known as linear assignment problem, can be solved exactly by relaxing the binary constraints into the constraint that each variable is real valued between 0 and 1. Since the relaxed form of this problem satisfies Slater conditions, we have that strict duality holds thus that the optimal value can also be obtained through the following dual problem :

$$\begin{aligned} \max_{v,w} \quad & \vec{1}^\top v + \vec{1}^T w \\ \text{subject to} \quad & \vec{X}y^\top - v\vec{1}^\top - \vec{1}w^\top \geq 0\,, \end{aligned}$$

where $v \in \mathbb{R}^M$ and $w \in \mathbb{R}^M$. When replacing the first term of the constraint stated in equation (12), we get that constraint (11) is satisfied as long as there exists a set of values for $v_{i,j}$ and $w_{i,j}$ that satisfy constraints (10c) and (10d). This completes our proof. □

We are left with showing that the problem (5) in Proposition 3.3 can as well be reduced a problem independent of the size of the set of permutations Σ. We prove this in the following proposition, which also wraps up our result for the case of law-invariant risk measures based on a uniform probability measure.

Proposition 3.9. *Given a set $\mathcal{E}$ of K comparisons, let the set $\{X_j\}_{j=1}^J := \vec{0} \cup \bigcup_{k=1}^K \{W_k, Y_k\}$ be the support set of all random payoffs involved in one of the elicited comparison and the zero payoff which we identify as X_1. When the distribution on the discrete probability space is uniform, the preference robust risk minimization problem (2) with $\mathcal{R}_{LE}(\mathcal{E})$ is equivalent to the optimization problem:*

$$\min_{x\in\mathcal{X},t,\theta,\{Q_j\}} \quad t, \tag{13a}$$

$$\text{subject to} \quad \vec{Z}(x) + t \geq \textstyle\sum_j Q_j\vec{X}_j + \vec{\delta}^\top\theta, \tag{13b}$$

$$Q_j\vec{1} = \theta_j,\, \forall\, j \tag{13c}$$

$$Q_j^\top\vec{1} = \theta_j,\, \forall\, j \tag{13d}$$

$$\vec{1}^\top\theta = 1 \tag{13e}$$

$$\theta \geq 0,\; Q_j \geq 0\,\forall\, j \tag{13f}$$

where $t \in \mathbb{R}$, $\theta \in \mathbb{R}^J$, and each $Q_j \in \mathbb{R}^{M\times M}$, and where $\bar{\delta} \in \mathbb{R}^J$ is the optimal solution of the linear program (10)*. Moreover, problem* (13) *is a convex optimization problem when each term $(\vec{Z}(x))_i$ is a concave function of x, and problem* (10) *is feasible if and only if $\mathcal{R}_{LE}(\mathcal{E})$ is non-empty.*

Proof. Based on Lemma 3.7, we have that the preference robust risk minimization problem (2) with $\mathcal{R}_{LE}(\mathcal{E})$ is equivalent to using the set of risk measure $\mathcal{R}_{El}(\Sigma(\mathcal{E}))$. According to Proposition 3.3, minimizing this preference robust risk measure can be achieved by solving:

$$\begin{array}{ll} \min\limits_{x\in\mathcal{X},t,\lambda} & t, \\ \text{subject to} & \vec{Z}(x) + t \geq \sum_{j,\sigma} \lambda_{j,\sigma}(\sigma(\vec{X}_j) + \bar{\delta}_{j,\sigma}) \\ & \sum_{j,\sigma} \lambda_{j,\sigma} = 1,\ \lambda \geq 0\ , \end{array}$$

where $\lambda \in \mathbb{R}^{J\times|\Sigma|}$, and where $\bar{\delta}_{j,\sigma}$ is an optimal solution to problem (9). Yet, according to Lemma 3.8, problem (9) reduces to problem (10) which always identifies solution for which $\bar{\delta}_{j,\sigma}$ is constant over $\sigma \in \Sigma$. This means that the optimization problem described above reduces to:

$$\begin{array}{ll} \min\limits_{x\in\mathcal{X},t,\lambda,\theta} & t, \\ \text{subject to} & \vec{Z}(x) + t \geq \sum\limits_{j,\sigma} \lambda_{j,\sigma}\sigma(\vec{X}_j) + \sum\limits_j \theta_j \bar{\delta}_j \\ & \theta_j = \sum\limits_{\sigma\in\Sigma} \lambda_{j,\sigma}\ ,\ \forall\, j \\ & \sum\limits_{j,\sigma} \lambda_{j,\sigma} = 1,\ \lambda \geq 0\ , \end{array}$$

which can easily be shown equivalent to

$$\begin{array}{ll} \min\limits_{x\in\mathcal{X},t,\lambda',\theta} & t, \\ \text{subject to} & \vec{Z}(x) + t \geq \sum\limits_j \theta_j (\sum\limits_\sigma \lambda'_{j,\sigma} Q_\sigma)\vec{X}_j + \sum\limits_j \theta_j \bar{\delta}_j \\ & \sum\limits_{\sigma\in\Sigma} \lambda'_{j,\sigma} = 1\ ,\ \forall\, j,\ \lambda' \geq 0\ , \\ & \sum\limits_j \theta_j = 1,\ \theta \geq 0\ , \end{array}$$

where $Q_\sigma \in \mathbb{R}^{M\times M}$ is the permutation matrix associated to the σ permutation operator. One can finally realize that for each j, the set captured by $\{\sum_\sigma \lambda'_{j,\sigma} Q_\sigma \mid \sum_{\sigma\in\Sigma} \lambda'_{j,\sigma} = 1,\ \lambda'_{j,\sigma} \geq 0\, \forall\, \sigma\ ,\}$ describes the convex hull of all permutation matrices which is known given the result of Birkhoff (1946) to be equivalently represented by $\{Q_j \in \mathbb{R}^{M\times M} \mid Q_j\vec{1} = \vec{1}, Q_j^\top \vec{1} = \vec{1}, Q_j \geq 0\}$. Using this simpler representation we obtain the following optimization problem

$$\begin{aligned}
\min_{x\in\mathcal{X},t,\theta,\{Q_j\}} \quad & t, \\
\text{subject to} \quad & \vec{Z}(x)+t \geq \sum_j \theta_j Q_j \vec{X}_j + \sum_j \theta_j \bar{\delta}_j \\
& Q_j\vec{1}=\vec{1}\,,\ Q_j^\top\vec{1}=\vec{1} \\
& \sum_j \theta_j = 1 \\
& Q_j \geq 0,\ \theta_j \geq 0\,, \forall j\,.
\end{aligned}$$

This last version of the problem reduces to problem (13) when replacing $Q'_j := \theta_j Q_j$. □

Up to this point, we have unraveled how to account in the optimization model for the law invariance hypothesis when the probability measure is uniform. Our next step is to show how this result can be generally applied in practice. In particular, let us consider how preference elicitation would actually work for a decision maker whose risk perception depends only on distribution of Z (i.e. its cumulative mass function), rather than how Z actually maps from Ω to $\mathbb{R}$. Given that when asked to compare W to Y, the decision maker really only compares their respective distribution F^W and F^Y, we can assume without loss of generality that preferences are stated in terms of distributions, which is also potentially more practical. Elicited comparisons therefore take the shape of the set $\{(F_k^W, F_k^Y)\}_{k=1}^K$ indicating that for each k the distribution F_k^W is perceived at most as risky as F_k^Y. For simplicity of exposure, we also make the following assumption.

Assumption 3.10. *The random variable $Z(x)$ can be expressed as $Z(x) := r(x,\xi)$ for some random vector $\xi : \Omega \to \mathbb{R}^m$ with finite support $\mathcal{S} \in \mathbb{R}^m$ such that $\mathcal{S} := \{\xi_1, \xi_2, \ldots, \xi_{M_\xi}\}$ and probability vector $p^\xi \in \mathbb{R}^{M_\xi}$ such that $p_i^\xi = P(\xi = \xi_i)$ for all i, and some function $r : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ that is concave in x over the set $\mathcal{X}$ for any $\xi_i \in \mathcal{S}$.*

Under the law-invariance assumption, the preference robust risk minimization problem (2) can then be presented as

$$\min_{x\in\mathcal{X}} \sup_{\rho\in\mathcal{R}_{LE}(\{(F_k^W,F_k^Y)\}_{k=1}^K)} \rho(\vec{Z}(x)), \tag{14}$$

where we overloaded the definition of $\mathcal{R}_{LE}$ so that

$$\mathcal{R}_{LE}(\{(F_k^W, F_k^Y)\}_{k=1}^K) := \mathcal{R}_{LE}(\{(W_k, Y_k)\}_{k=1}^K)$$

with W_k and Y_k chosen as any random variable that respectively have the distributions F_k^W and F_k^Y.

We show in the following proposition that the above problem can also be reformulated into a tractable convex optimization reformulation given that the following assumption holds.

Assumption 3.11. *All probability distributions employ rational numbers as probability values.*

The key idea behind the proof of the following result, which we defer to Appendix C, is that with Assumption 3.11 we can map any given distribution to a random payoff in a high-dimensional outcome space endowed with a uniform probability measure. In that space, the results of Lemma 3.8 and Proposition 3.9 can be applied. Thereafter, it is possible to reduce the problem formulated in that high-dimensional space to an equivalent optimization problem with dimensionality depending only on the sizes of the supports of the distributions that are involved.

Proposition 3.12. *Given a set $\mathcal{E}$ of K comparisons, let the set $\{F_j\}_{j=1}^J$ be the set of all distributions involved in one of the elicited comparison and the distribution of a zero payoff identified as F_1. Furthermore, for each $j \in \{1, 2, \ldots, J\}$, let the pair $(\vec{X}_j, p^j) \in \mathbb{R}^{M_j} \times \mathbb{R}^{M_j}$ capture the distribution of F_j such that $F_j = \sum_{i=1}^{M_j} p_i^j \boldsymbol{Dirac}((\vec{X}_j)_i)$, with $\boldsymbol{Dirac}((\vec{X}_j)_i)$ is the Dirac measure that puts all of its weight on $(\vec{X}_j)_i$. Given that assumptions 3.10 and 3.11 hold, the preference robust risk minimization problem* (2) *with $\mathcal{R}_{LE}(\{(F_k^W, F_k^Y)\}_{k=1}^K)$ is equivalent to the optimization problem:*

$$
\begin{aligned}
\min_{x \in \mathcal{X}, t, \theta, \{C_j\}} \quad & t, && \text{(15a)} \\
\text{subject to} \quad & \vec{Z}_\xi(x) + t \geq \textstyle\sum_j C_j \vec{X}_j + \bar{\delta}^\top \theta, && \text{(15b)} \\
& C_j \vec{1} = \theta_j, \; \forall j && \text{(15c)} \\
& (\Pi_j \circ C_j)^\top \vec{1} = \theta_j, \; \forall j && \text{(15d)} \\
& \vec{1}^\top \theta = 1 && \text{(15e)} \\
& \theta \geq 0, \; C_j \geq 0 \, \forall j && \text{(15f)}
\end{aligned}
$$

where $\circ$ is the Hadamard product [9]*, $t \in \mathbb{R}$, $\theta \in \mathbb{R}^J$, and each $C_j \in \mathbb{R}^{M_\xi \times M_j}$, and where $\vec{Z}_\xi(x)$ refers to the vector $[r(x, \xi_1) \;\; r(x, \xi_2) \;\; \ldots \;\; r(x, \xi_{M_\xi})]^\top$. Furthermore, for each j, the matrix $\Pi_j \in \mathbb{R}^{M_\xi \times M_j}$ is defined such that $(\Pi_j)_{m,n} = p_m^\xi / p_n^j$. Finally, the parameter $\bar{\delta} \in \mathbb{R}^J$ is the optimal solution of the following linear program.*

$$
\begin{aligned}
\max_{\delta, \{y_j\}_{j=1}^J, \{v_{i,j}, w_{i,j}\}_{i=1,j=1}^{i=J,j=J}} \quad & \sum_{j=1}^J \delta_j, && \text{(16a)} \\
\text{subject to} \quad & \delta_i \leq \delta_j \,, \forall (i,j) \in \bar{\mathcal{E}} && \text{(16b)} \\
& \vec{1}^\top v_{i,j} + \vec{1}^\top w_{i,j} - \vec{X}_j^\top y_j + \delta_i - \delta_j \geq 0 \,, \forall i \neq j && \text{(16c)} \\
& \Pi_{i,j} \circ (\vec{X}_i y_j^\top) - v_{i,j} \vec{1}^\top - \Pi_{i,j} \circ (\vec{1} w_{i,j}^\top) \geq 0 \,, \forall i \neq j && \text{(16d)} \\
& \vec{1}^\top y_j = 1, \; y_j \geq 0 \,, \forall j && \text{(16e)} \\
& \delta_1 = 0 \,, && \text{(16f)}
\end{aligned}
$$

where $\delta \in \mathbb{R}^J$, $y_j \in \mathbb{R}^{M_j}$, $w_{i,j} \in \mathbb{R}^{M_j}$, $v_{i,j} \in \mathbb{R}^{M_i}$, where the matrix $\Pi_{i,j} \in \mathbb{R}^{M_i \times M_j}$ is defined such that $(\Pi_{i,j})_{m,n} := p_m^i / p_n^j$, and finally where $\bar{\mathcal{E}}$ is the set of edges in the

[9]The Hadamard product of two matrices of same size is defined as $(A \circ B)_{i,j} := A_{i,j} B_{i,j}$.

partial ordering of $\{F_j\}_{j=1}^J$ described by the elicited comparisons. Namely,

$$\bar{\mathcal{E}} := \{(i,j) \in \{1,2,\ldots,J\}^2 \,|\, (F_i, F_j) \in \{(F_k^W, F_k^Y)\}_{k=1}^K\}\,.$$

Moreover, problem (15) *is a convex optimization problem when each term $r(x,\xi_i)$ is a concave function of x, and problem* (16) *is feasible if and only if $\mathcal{R}_{LE}(\{(F_k^W, F_k^Y)\}_{k=1}^K)$ is non-empty.*

Example 3.3. In a portfolio optimization example, we assume the vector of returns achieved for each asset follows a multivariate distribution such that $P(\xi = \xi_i) = p_i^\xi$ for each $i = 1, ..., M_\xi$. The preference robust risk minimization problem reduces to

$$\begin{aligned}
\min_{x\in\mathcal{X},t,\theta,\{C_j\}_{j=1}^J} \quad & t, && \text{(17a)}\\
\text{subject to} \quad & \xi_i^\top x + t \geq \sum_j (C_j)_{i,:}\vec{S}_j + \bar{\delta}^\top\theta, i = 1, ..., M_\xi && \text{(17b)}\\
& C_j\vec{1} = \theta, \forall\, j \in \{1, ..., J\} && \text{(17c)}\\
& (\Pi_j \circ C_j)^\top\vec{1} = \theta_j,\, \forall\, j \in \{1, ..., J\} && \text{(17d)}\\
& C_j \geq 0, \forall\, j \in \{1, ..., J\} && \text{(17e)}\\
& \vec{1}^\top\theta = 1,\;\; \theta \geq 0. && \text{(17f)}
\end{aligned}$$

This is a linear program with $O(n+\bar{M}^2K)$ decision variables and $O(\bar{M}^2K)$ constraints, where $\bar{M}$ denotes the size of the largest support of all distributions involved, i.e. $\bar{M} = \max\{M_\xi, M_1, M_2, \ldots, M_J\}$. The linear program (16) that generates the parameter $\bar{\delta}$ has $O(\bar{M}K^2)$ decision variables and $O(\bar{M}^2K^2)$ constraints. Note that while the computational burden associated to problem (16) grows quadratically with respect to the number of comparisons K, this problem needs only to be solved once after meeting with the investor. Once this is done, the portfolio can easily be re-optimized as new market information is received by solving problem (17) which size grows only linearly in K.

We end this section by considering the last set of measures $\mathcal{R}_{CLE}(\mathcal{E})$, which incorporates all the properties addressed earlier. Using the same notation for the case of $\mathcal{R}_{LE}(\mathcal{E})$, we can write the preference robust risk minimization problem (2) with the set $\mathcal{R}_{CLE}(\mathcal{E})$ as

$$\min_{x\in\mathcal{X}} \sup_{\rho\in\mathcal{R}_{CLE}(\{(F_k^W, F_k^Y)\})} \rho(\vec{Z}(x)), \tag{18}$$

where $\mathcal{R}_{CLE}$ is again overloaded to take as argument comparisons of distributions. We show in the following proposition that the optimization problem (18) is just as tractable as the problem (14).

The result below can be obtained by applying similar steps we took in Appendix C to prove Proposition 3.12.

Proposition 3.13. *Given a set $\mathcal{E}$ of K comparisons, let the set $\{F_j\}_{j=1}^J$ be the set of all distributions involved in one of the elicited comparison and the distribution of a zero payoff identified as F_1. Furthermore, for each $j \in \{1, 2, \ldots, J\}$, let the pair $(\vec{X}_j, p^j) \in$*

$\mathbb{R}^{M_j} \times \mathbb{R}^{M_j}$ capture the distribution of F_j such that $F_j = \sum_{i=1}^{M_j} p_i^j \boldsymbol{Dirac}((\vec{X}_j)_i)$, with $\boldsymbol{Dirac}((\vec{X}_j)_i)$ is the Dirac measure that puts all of its weight on $(\vec{X}_j)_i$. Given that assumptions 3.10 and 3.11 hold, the preference robust risk minimization problem (2) *with $\mathcal{R}_{CLE}(\{(F_k^W, F_k^Y)\}_{k=1}^K)$ is equivalent to the optimization problem* (15) *from which constraint* (15e) *is removed, and where $\bar{\delta}$ is the optimal solution of problem* (16) *with constraint* (16c) *replaced with*

$$\begin{aligned}
&-\vec{X}_j^\top y_j - \delta_j \geq 0, \forall j \\
&\vec{1}^\top v_{i,j} + \vec{1}^\top w_{i,j} + \delta_i \geq 0\,, \forall\, i \neq j\,.
\end{aligned}$$

Again, the modified version of problem (15) *is a convex optimization problem when each term $r(x, \xi_i)$ is a concave function of x, and the revised version of problem* (16) *is feasible if and only if $\mathcal{R}_{CLE}(\{(F_k^W, F_k^Y)\}_{k=1}^K)$ is non-empty.*

4 Numerical Experiments

In this section, we illustrate the use of robust risk measure framework on a portfolio optimization problem. Namely, we consider the situation of a financial adviser that is asked by his client to propose an investment strategy among a specific set of assets so that the risk exposure is as low as possible. A priori, the adviser is not aware of how client perceives risk and might simply assume that this client will be happy with a portfolio that minimizes an arbitrary conditional value at risk (CVaR) measure such as the 20% CVaR. Unfortunately, this might not be how this investor perceives risk hence the motivation for using our framework.

In this context, the investor is first asked whether he agrees with the monotonicity, strict monotonicity for certain payoffs, quasi-convexity, and translation invariance conditions, and can further be asked if he is comfortable with the scale or law invariance axioms. Once it is done, the set of plausible risk measures can be further refined by presenting to him a set of pairs of random payoffs to compare. If the client agreed with law invariance, then he can be asked to compare cumulative mass functions, otherwise it is necessary to describe to him what the payoffs are obtained for each potential future state of the world. Alternatively, the client can also be asked, from a given list of random payoffs, what the certain amount of payoff is that he considers equivalent to each payoff in the list, or the range that contains the certainty equivalent to each payoff. As an example, we present in Figure 1(a) three random returns $X_1(\omega)$, $X_2(\omega)$, and $X_3(\omega)$ constructed based on selected 13 weeks of returns, i.e. $|\Omega| = 13$, and present in Figure 1(b) the respective cumulative mass functions F^{X_1}, F^{X_2}, and F^{X_3}. One can see that comparing random returns is considerably more involved than comparing their cumulative functions, as the former is state-dependent. In addition, from (b) we can see that if the client agreed with law invariance, comparing X_1 and X_3 becomes unnecessary since F^{X_1} dominates F^{X_3}, which implies that there exists a permutation σ such that $X_1 \geq \sigma(X_3)$, and by monotonicity $\rho(X_1) \leq \rho(\sigma(X_3))$, $\forall \sigma$. Note that in the case that the client does not agree with the scale invariance, care has to be taken to clearly indicate to the client that Figure 1 is presented based on returns. His preferences could depend on the total amount of wealth he decides to invest.

Finally, the adviser must then solve the appropriate model presented in one of the previous sections in order to obtain a portfolio that is robust with respect to the existing ambiguity about the client's risk perception.

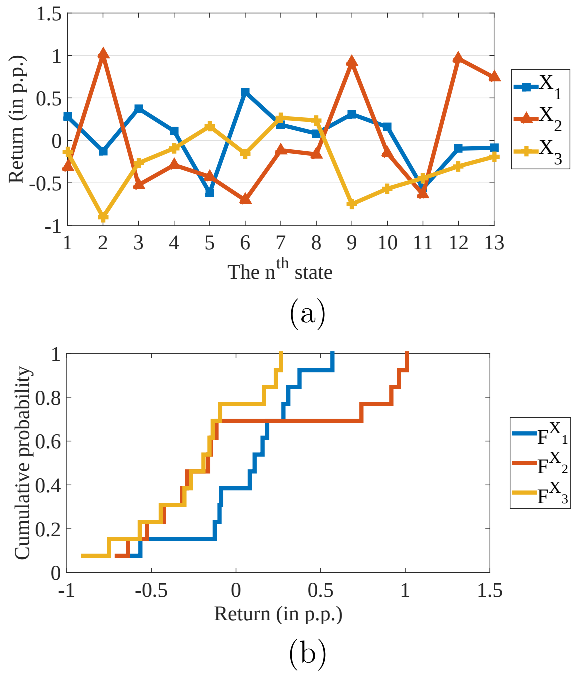


Figure 1: Three risky returns and their cumulative mass functions

The specifics of our experimentation are as follows. We consider a week long investment on a set of four assets. We simulate our client investor by assuming that his perception of risk is exactly captured by the spectral risk measure function:

$$\bar{\rho}(\vec{Z}) := 0.1 \cdot \text{CVaR}_{20\%}(\vec{Z}) + 0.9 \cdot E[-\vec{Z}] \ .$$

Note that while the simulated investor will always make comparisons of random payoffs according to this mathematical representation, neither the investor nor his adviser have the means of identifying precisely this function. In order to focus on issues related to the identification of the right risk exposure, we assume that both the investor and his adviser have reached a consensus on the probabilistic model that should be used. In particular, they assume that each of the last 13 weeks of the term provides an equiprobable scenario for the joint set of weekly returns generated by the assets. Since the true risk measure is not known by the adviser, it is plausible to think that if he was to commit to a certain risk measure, he might choose the 20%-CVaR or the expected loss. With our framework, the adviser instead would be able to ask the investor to estimate the certainty equivalent of a set of random payoffs, each constructed by sampling a random set of 13 weekly returns from a historical dataset for a range of reference years. It is expected that a larger number of these certainty equivalent estimates will lead to less conservative solutions. Specifically, we compare the performances of the portfolios that are obtained using the following risk measures: $\text{CVaR}_{20\%}(\vec{Z})$, expected loss, i.e. $E[-\vec{Z}]$, robust risk measure with $\mathcal{R}_{El}$, with $\mathcal{R}_{CE}$, with $\mathcal{R}_{LE}$, with $\mathcal{R}_{CLE}$, and investor's true risk measure: $\bar{\rho}(\vec{Z})$. The resulting performances will all be evaluated as truly perceived by the investor, i.e. evaluated using $\bar{\rho}(\vec{Z})$. Note that $\bar{\rho}(\vec{Z})$ is normalized so that it returns the certain loss that is considered equivalent to Z.

We run 5000 experiments using historical stock market data about 335 companies that are part of the S&P 500 index during the period from January 1994 until December 2013. Each experiment consists of drawing 4 assets randomly from the pool of 335 and a random week in the period from January 2004 to December 2013, and require each method to propose a portfolio using the distributional model constructed based on the latest 13 weekly returns, and subject to the constraints $\vec{1}^\top x = 1$ and $x \geq 0$. To simulate elicited comparisons, we require the investor to estimate the certainty equivalent of up to 500 random payoffs constructed based on historical data extracted from the period between January 1994 to December 2003. In each experiment, we evaluate the in-sample performance of each resulting portfolio using $\bar{\rho}(\vec{Z})$ based on the same 13 weekly returns used in optimizing the portfolio, and then calculate the realized return of each resulting portfolio based on the week that immediately follows. The average of the in-sample performance over 5000 experiments is presented in Figure 2, whereas in Figure 3 the statistics of realized returns in terms of average, 20%-CVaR, and $\bar{\rho}$ risk measure are presented as out-of-sample performance.

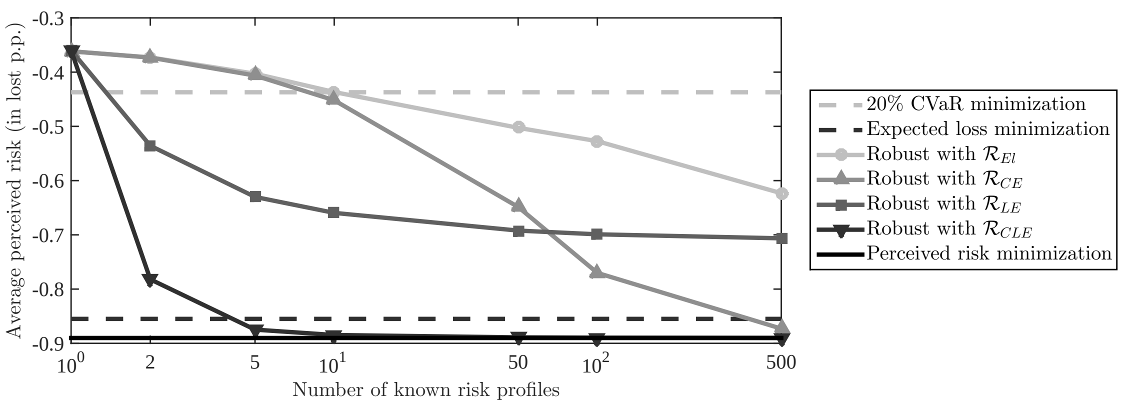


Figure 2: Comparison of the average perceived risk (in lost percentile points) for the portfolio obtained using either CVaR minimization, expected loss minimization or the minimization of a preference robust risk measure with certainty equivalent knowledge for up to 500 random payoffs (including the null payoff) in a set of 5000 experiments. We also report the best average perceived risk that could be obtained if the representation of this perception was exactly known.

When studying Figure 2, one should notice that when the wrong risk measure is used in the optimization model (specifically the expected loss or 20% CVaR measure), the performance of the portfolio that is obtained is necessarily biased. On the other hand, when a preference robust risk measure is used, the true perceived risk can quickly be improved by either making further hypotheses about the risk preference relation, like coherence or law invariance, or by increasing the knowledge about the risk measure through making the comparison of more uncertain payoffs. Based on this set of experiments, we observe that when the investor agrees with law invariance knowing the certainty equivalent of as little as 10 random payoffs is sufficient to approximate fairly accurately his attitude with respect to risk. Otherwise, the same accuracy can theoretically be achieved by assessing the certainty equivalent of a larger number of random payoffs. The results however seem to indicate that the price for doing so is high. Indeed, while one might not be entirely comfortable with the law invariance hypothesis, the reported performances for $\mathcal{R}_{El}$ and $\mathcal{R}_{CE}$ seem to indicate that an impracticable number of questions need to be answered by the investor in order to identify a legitimate portfolio. Note that in this regard we have good hope that preference elicitation strategies along the lines of what was presented in section 4.3.1 of Armbruster and Delage (2015) might significantly improve the convergence.

The out of sample statistics presented in Figure 3 do illustrate how methods that are based on expected loss, CVaR, and the $\bar{\rho}$ risk measure generally propose portfolios that perform better in terms of these specific statistics. More importantly, two key observations can be made in these figures. First, it appears in Figure 3 (c) that the convergence of the performance of preference robust risk minimization methods to the best perceivable risk is much improved in terms of the number of questions that need to be answered. We believe this to be due to the fact that the historical data somehow mis-estimates the odds of possible future payoffs thus reducing the value of reaching high accuracy in the risk measure that is employed. The second key observation lies in Figure 3 where one should notice that all preference robust risk minimization methods propose portfolios that are more conservative than the portfolios optimized based on the $\bar{\rho}$ risk measure in terms of 20% CVaR. We believe this is an important feature of these methods; namely, that in case of doubt about how risk should be measured, our framework will encourage investing conservatively. In other words, our framework will wait to have sufficient information indicating that the decision maker does not worry so much about tail events (i.e. low probability CVaR) before suggesting portfolios that achieve poorly for such events.

5 Conclusion

In this work, we propose for the first time a framework that can be used to interact with a decision maker who seeks to optimize a financial position, and to identify, based on information that is gathered about the decision maker's subjective perception of risk, a well-motivated convex risk measure to employ. Unlike previous discussions about what makes a good risk measure, we employ a language that is entirely based on assessing which random payoffs are considered less/as/more risky than the others, and resolve any ambiguity that is left regarding the final selection of an appropriate measure using a conservative point of view. In our opinion, the value of such a framework is two-fold. First, it should allow a better adoption of risk measures in fields of practice where the stake-holders are not comfortable with the statements that quantify how large the risk is (e.g. "returns of investment #1 are three times more risky then those of investment #2") or where stake-holders are unable for some payoffs to decide which one they would prefer being exposed to. Secondly, this framework has the potential to clarify what the responsibilities of a decision maker are when interacting with a risk management consultant. Namely, the decision maker should understand and decide which hypothesis he accepts among hypotheses 1 to 7, and only include in $\mathcal{E}$ the pairs of random payoffs that he feels comfortable with. Once this is confirmed, and possibly even signed off in the form of a contract, our framework will ensure that the financial decision that is made never engages the decision maker in a position that has risks that are unjustified when compared to any payoff that can be achieved with certainty (refer to section 2 for a discussion).

From a numerical point of view, we demonstrated that under reasonable concavity assumptions about the random mapping $\vec{Z}(x)$, the decision problem reduces to a convex optimization problem of reasonable size and can thus be solved efficiently using interior point algorithms. In case of a portfolio optimization problem, the problem further reduces to a linear program for which we illustrated that under the hypothesis of law invariance it is possible to assess fairly precisely the structure of the risk measure after only a few comparisons (possibly less than 10 questions). Of course, it will be interesting to confirm these insights on a more diverse set of decision problems and to investigate more seriously what schemes could guide the preference elicitation process.

Finally, given that in practice we cannot expect perfect accuracy when comparisons in $\mathcal{E}$ are made by human beings, it is possible that the information obtained about $\mathcal{R}$ be inconsistent thus making it impossible to identify a risk measure that is consistent with all the information. While it was shown in section 3 how one can diagnose such a situation, i.e. that $\mathcal{R}$ is effectively empty, it is worth briefly indicating how such an impasse can be resolved. The most reasonable hypothesis is potentially that comparisons that are described in $\mathcal{E}$ are being corrupted by some perception noise. Namely, this would imply that each pair $(W_k, Y_k) \in \mathcal{E}$ truly only indicates that $W_k + \varepsilon_k$ is less risky than Y_k for some small perturbation term ε_k. For this reason, it might be more appropriate to replace $\mathcal{R}_{El}(\mathcal{E})$ with

$$\mathcal{R}_{El}(\mathcal{E},\Gamma) := \left\{ \rho \in \mathfrak{R} \left| \exists\, \varepsilon \in \mathbb{R}^K, \begin{array}{c} \rho(\vec{W}_k) - \varepsilon_k \le \rho(\vec{Y}_k)\,,\ \forall\, k \in \{1,2,\ldots,K\} \\ \varepsilon_k \ge 0\,,\ \forall\, k \in \{1,2,\ldots,K\} \\ \|\varepsilon\| \le \Gamma \end{array} \right. \right\} ,$$

for some choice of norm $\|\cdot\|$ and total perturbation Γ.

Perhaps the simplest implementation of such an approach would be to use the infinity norm (i.e. $\|\varepsilon\| = \max_k \varepsilon_k$). In this case, the problem reduces to calibrating Γ by seeking the smallest positive perturbation $\bar{\varepsilon} \in \mathbb{R}$ which ensures that $\mathcal{R}_{El}(\{(W_k + \bar{\varepsilon}, Y_k)\}_{k=1}^K)$ is non-empty. This is equivalent to doing a search over the interval $[0, \sup_{k,\omega\in\Omega} Y_k(\omega) - W_k(\omega)]$ for the smallest value that makes the relevant problem among problems (6), (8), (16), or its revised version in proposition 3.13 feasible after replacing the constraints

$$\delta_i \le \delta_j\,,\ \forall\,(i,j) \in \bar{\mathcal{E}}$$

with

$$\delta_i - \varepsilon \le \delta_j\,,\ \forall\,(i,j) \in \bar{\mathcal{E}}\ .$$

One can then account for $\mathcal{R}_{El}(\mathcal{E}, \bar{\varepsilon})$ instead of $\mathcal{R}_{El}(\mathcal{E})$ in each formulation by simply employing $\mathcal{R}_{El}(\{(W_k + \bar{\varepsilon}, Y_k)\}_{k=1}^K)$.

References

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

Armbruster, B., E. Delage. 2015. Decision making under uncertainty when preference information is incomplete. *Management Science* **61**(1) 111–128.

Artzner, P., F. Delbaen, J. M. Eber, D. Heath. 1999. Coherent Measures of Risk. *Mathematical Finance* **9**(3).

Babazadeh, R., J. Razmi, M. S. Pishvaee. 2011. Robust facility location in supply chain network design under risk. *International Journal of Mechanic Systems Engineering* **1**(1) 35–47.

Basel Committee on Banking Supervision. 2012. Consultative document : Fundamental review of the trading book.

Birkhoff, G. 1946. Three observations on linear algebra. *Universidad Nacional de Tucumán Revista Serie A* **5** 147–151.

Boutilier, C., R. Patrascu, P. Poupart, D. Schuurmans. 2006. Constraint-based optimization and utility elicitation using the minimax decision criterion. *Artificial Intelligence* **170**(8-9) 686–713.

Carneiro, M. C., G. P. Ribas, S. Hamacher. 2010. Risk management in the oil supply chain: A cvar approach. *Industrial & Engineering Chemistry Research* **49**(7) 3286–3294.

Chajewska, U., D. Koller, R. Parr. 2000. Making rational decisions using adaptive utility elicitation. *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI)*. 363–369.

Chan, T. C.Y., H. Mahmoudzadeh, T. G. Purdie. 2014. A robust-cvar optimization approach with application to breast cancer therapy. *European Journal of Operational Research* **238**(3) 876 – 885.

Clemen, R. T., T. Reilly. 1999. *Making Hard Decisions with Decisiontools Suite*. 1st ed. Duxbury.

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

Debreu, G. 1954. Representation of a preference ordering by a numerical function. *Decision Processes* 159–165.

Einhorn, D. 2008. Private Profits and Socialized Risk. *Global Association of Risk Professionals Risk Review* (42) 10–18.

Ellsberg, D. 1961. Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics* **75**(4) 643–669.

Föllmer, Hans, Alexander Schied. 2002. Convex measures of risk and trading constraints. *Finance and Stochastics* **6**(4) 429–447.

Grable, J., R. H. Lytton. 1999. Financial risk tolerance revisited: the development of a risk assessment instrument. *Financial Services Review* **8**(3) 163–181.

Hu, J., S. Mehrotra. 2015. Robust decision making over a set of random targets or risk-averse utilities with an application to portfolio optimization. *IIE Transactions* (47) 1–15.

Jabr, R.A. 2005. Robust self-scheduling under price uncertainty using conditional value-at-risk. *IEEE Transactions on Power Systems* **20**(4) 1852–1858.

Kusuoka, S. 2001. On law invariant coherent risk measures. Shigeo Kusuoka, Toru Maruyama, eds., *Advances in Mathematical Economics*, *Advances in Mathematical Economics*, vol. 3. Springer Japan, 83–95.

Rockafellar, R. Tyrrell, Stanislav Uryasev. 2000. Optimization of conditional value-at-risk. *Journal of Risk* **2** 21–41.

Toumazis, T., C. Kwon. 2013. Routing hazardous materials on time-dependent networks using conditional value-at-risk. *Transportation Research Part C: Emerging Technologies* **37**(0) 73 – 92.

Tversky, A., D. Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science* **185**(4157) 1124–1131.

A Reduction to convex, coherent and law-invariant risk measures

In order to show that hypotheses 1 to 4 are equivalent to imposing that the risk measure be a convex risk measure, and that with hypotheses 5 and 6 the set of risk measure becomes equivalent to coherent and law invariant risk measures respectively, we start by demonstrating how each hypothesis about the risk preference relation implies that a certain constraint must be imposed on the risk measure that will be used. We then make sure that the risk measures that are circumscribed by these constraints represent risk preference relations that satisfy the spelled out hypotheses. Note that the equivalence between hypothesis 7 and $\mathcal{R}_{El}$ is trivial to verify.

One can easily verify, following the Debreu's representation theorem, that the monotonicity property as presented in hypothesis 1 is equivalent to the monotonicity axiom of convex risk measures:

$$Z_1(\omega) \geq Z_2(\omega)\,,\, \forall\, \omega \in \Omega \;\;\Rightarrow\;\; \rho(\vec{Z}_1) \leq \rho(\vec{Z}_2)\;.$$

Secondly, strict monotonicity for certain payoffs ensures that every random payoff has a well-defined and unique certainty equivalent, i.e. $\mathbb{CE}(Z)$ is the cash amount that is considered as risky as Z. Hence, without loss of generality we can normalize the risk measure such that $\rho(\vec{Z}) = -\mathbb{CE}(Z)$ simply by considering $\rho'(\vec{Z}) = f(\rho(\vec{Z}))$, with $f(y) = \sup\{z \in \mathbb{R} | \rho(-z) \leq y\}$. Indeed, we have that $\rho'(\vec{Z}) = f(\rho(\vec{Z})) = -\mathbb{CE}(Z)$, and $\rho'(\vec{Z})$ always captures the same preferences since $f(\cdot)$ is strictly increasing by strict monotonicity.

Using the notion that $\rho(\vec{Z})$ is the negative of the certainty equivalent of Z, one can show that translation invariance as presented in hypothesis 4 implies the cash invariance axiom of convex risk measures. Since, we have that

$$\rho(\vec{Z}) = \rho(\mathbb{CE}(Z)) \Rightarrow \rho(\vec{Z}) = \rho(-\rho(\vec{Z})) \tag{19}$$

it must be that

$$\rho(\vec{Z}) \leq \rho(-\rho(\vec{Z})) \;\Rightarrow\; \rho(\vec{Z}+c) \leq \rho(-\rho(\vec{Z})+c) = \rho(\vec{Z}) - c \;,$$

and similarly that,

$$\rho(\vec{Z}) \geq \rho(-\rho(\vec{Z})) \;\Rightarrow\; \rho(\vec{Z}+c) \geq \rho(-\rho(\vec{Z})+c) = \rho(\vec{Z}) - c \;,$$

where the last equalities of the above two are due to (19) again. Hence, we must have that $\rho(\vec{Z}+c) = \rho(\vec{Z}) - c$.

Together with translation invariance, quasi-convexity is actually equivalent to the convexity axiom of convex risk measures. Specifically, since we have that

$$\rho(\vec{Z}_i + \rho(\vec{Z}_i)) = \rho(\vec{Z}_i) - \rho(\vec{Z}_i) = 0 \leq \rho(0),\; i = 1, 2,$$

i.e. $\vec{Z}_1 + \rho(\vec{Z}_1)$ and $\vec{Z}_2 + \rho(\vec{Z}_2)$ are considered less risky than the zero payoff, we must have by quasi-convexity that

$$\rho(\theta(\vec{Z}_1 + \rho(\vec{Z}_1)) + (1-\theta)(\vec{Z}_2 + \rho(\vec{Z}_2))) \;\leq\; \rho(0) = 0 \;.$$

Hence, it must be the case that

$$\rho(\theta\vec{Z}_1 + (1-\theta)\vec{Z}_2 + \theta\rho(\vec{Z}_1) + (1-\theta)\rho(\vec{Z}_2)) = \rho(\theta\vec{Z}_1 + (1-\theta)\vec{Z}_2) - \theta\rho(\vec{Z}_1) - (1-\theta)\rho(\vec{Z}_2) \leq 0 \;,$$

so that $\rho(\cdot)$ must respect convexity.

In the case of scale invariance as presented in hypothesis 5, a similar argument as for translation invariance can be made. Since we have that $\rho(\vec{Z}) = -\mathbb{CE}(Z)$, it must be that $\rho(\vec{Z}) = \rho(-\rho(\vec{Z}))$ and hypothesis 5 is equivalent to saying that $\rho(\lambda\vec{Z}) \leq \rho(-\lambda\rho(\vec{Z})) = \lambda\rho(\vec{Z})$ and that $\rho(\lambda\vec{Z}) \geq \rho(-\lambda\rho(\vec{Z})) = \lambda\rho(\vec{Z})$.

Finally, as stated in hypothesis 6, law invariance implies that if Z_1 and Z_2 have the same distribution then $\rho(\vec{Z}_1) = \rho(\vec{Z}_2)$.

For completeness, we verify the other direction, in other words that the constraints we impose in the definitions of $\mathfrak{R}$, $\mathcal{R}_{Coh}$, and $\mathcal{R}_{Law}$ do imply that the risk measures ρ that are members of the different set of risk measures can only represent preference relations $Z_1 \succeq_\rho Z_2 \;\Leftrightarrow\; \rho(\vec{Z}_1) \leq \rho(\vec{Z}_2)$ that agree with hypotheses 1 to 7. First, the monotonicity and convexity axioms ensure that ρ is both finite, given that Ω is a finite outcome space, and concave hence continuous. Based on its construction, the risk preference relation $\succeq_\rho$ is therefore necessarily complete, transitive, and continuity. Furthermore, one easily verifies that the monotonicity axiom guarantees that if $\vec{Z}_1 \geq \vec{Z}_2$ then $\rho(\vec{Z}_1) \leq \rho(\vec{Z}_2)$ and thus $\vec{Z}_1$ is considered no more risky than $\vec{Z}_2$. Translation invariance ensures both strict monotonicity for certain payoffs and the stated translation invariance hypothesis 4. Finally, convexity of risk measures ensures that the preferences it captures satisfy the quasiconvexity hypothesis since:

$$\rho(\theta\vec{Z}_1 + (1-\theta)\vec{Z}_2) \leq \theta\rho(\vec{Z}_1) + (1-\theta)\rho(\vec{Z}_2) \leq \max(\rho(\vec{Z}_1); \rho(\vec{Z}_2)) \leq \rho(\vec{Z}_3) \;.$$

Regarding scale invariance, the argument is trivial when the constraint imposed in the definition of $\mathcal{R}_{Coh}$ is imposed. Regarding the law invariance axiom of risk measures, the constraint imposed in the definition of $\mathcal{R}_{Law}$ states that if all Z's that are identically distributed, then they must also have the same level of risk according to $\rho(\cdot)$.

B Proof of proposition 3.3

Lemma B.1. *Given a set $\mathcal{E}$ of K comparisons, let the set $\{X_j\}_{j=1}^J := \vec{0} \cup \bigcup_{k=1}^K \{W_k, Y_k\}$ be the support set of all random payoffs involved in one of the elicited comparison and the zero payoff which we identify as X_1. Given any random payoff Z, the worst-case risk measure of Z under $\mathcal{R}_{El}(\mathcal{E})$ is the optimal value of the following optimization problem:*

$$\varrho_{\mathcal{R}_{El}(\mathcal{E})}(\vec{Z}) \quad = \quad \max_{\delta \in \Delta} \ \rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{Z})$$

where $\mathcal{A}(\mathbb{X}, \delta)$ denotes the monotone convex hull of the set $\{\vec{X}_j + \delta_j\}_{j=1}^J$ which admits the following representation

$$\mathcal{A}(\mathbb{X}, \delta) = \{\vec{Z} \in \mathbb{R}^M \,|\, \exists \theta \in \mathbb{R}^J, \;\; \vec{Z} \geq \mathbb{X}\theta + \delta^\top \theta, \, \vec{1}^\top \theta = 1, \, \theta \geq 0\},$$

where the matrix $\mathbb{X} \in \mathbb{R}^{M \times J}$ is composed by reference points $\{\vec{X}_j\}_{j=1,...,J}$ as its column vectors, i.e. $\mathbb{X} = [\vec{X}_1, \vec{X}_2 \cdots \vec{X}_J]$, and where the set Δ is the set represented by

$$\Delta := \{\delta \in \mathbb{R}^J \,|\, \rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_i + \delta_i) \geq 0, \; i = 1, ..., J, \; \delta_1 = 0, \; \delta_i \leq \delta_j, \, \forall \, (i, j) \in \bar{\mathcal{E}}\} \, .$$

The intuition behind the lemma above is fairly straightforward. For a given random payoff, the search for the acceptance set that leads to the worst-case risk reduces to searching among the feasible risk values $\delta \in \Delta$ for the reference set of random payoffs $\{X_j\}_{j=1}^J$. Once this is done, the worst-case risk is obtained by considering the monotone convex hull of the points $\vec{X}_j + \delta_j$. Note that at this point, it is not clear whether the worst-case acceptance set is independent of the random payoff Z that is analyzed or not.

Proof. First, we decompose the worst-case analysis in two steps: a first search over how much risk might be evaluated at the points in $\{\vec{X}_j\}_{j=1}^J$

$$\begin{aligned} \varrho_{\mathcal{R}_{El}(\mathcal{E})}(\vec{Z}) \; = \; \max_{\delta} \quad & \sup_{\rho \in \mathcal{R}_\delta(\{(\vec{X}_j, \delta_j)\}_{j=1}^J)} \rho(\vec{Z}) && \text{(20a)} \\ \text{subject to} \quad & \delta_i \leq \delta_j \, \forall \, (i.j) \in \bar{\mathcal{E}} && \text{(20b)} \\ & \delta_1 = 0 && \text{(20c)} \\ & \mathcal{R}_\delta(\{(\vec{X}_j, \delta_j)\}_{j=1}^J) \neq \emptyset \, , && \text{(20d)} \end{aligned}$$

where $\mathcal{R}_\delta(\{(\vec{X}_j, \delta_j)\}_{j=1}^J) \subset \mathfrak{R}$ is the set of convex risk measures that consider the risk of each $\vec{X}_j$ to be respectively δ_j, i.e. that $\rho(\cdot) \in \mathcal{R}_\delta(\{(\vec{X}_j, \delta_j)\}_{j=1}^J)$ if and only if it is a convex risk measure that evaluates $\rho(\vec{X}_j) = \delta_j$. Hence if $\rho \in \mathcal{R}_\delta(\{(\vec{X}_j, \delta_j)\}_{j=1}^J)$, then it necessarily satisfies the elicited comparison.

In order to show that constraints (20b) to (20d) are equivalent to the set Δ, one needs to establish that constraint (20d) is equivalent to $\rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_j + \delta_j) \geq 0 \, , \forall \, j$. The statement that for all $\delta \in \mathbb{R}^J$,

$$\rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_j + \delta_j) \geq 0 \, , \forall \, j \quad \Rightarrow \quad \mathcal{R}_\delta(\{(\vec{X}_j, \delta_j)\}_{j=1}^J) \neq \emptyset$$

is somewhat trivial since $\mathcal{A}(\mathbb{X},\delta)$ is constructed to make each $\vec{X}_j+\delta_j$ acceptable and therefore

$$0 \geq \rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_j+\delta_j) \geq 0 \Rightarrow \rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_j+\delta_j) = 0 \Rightarrow \rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_j) = \delta_j \;,$$

so that $\rho_{\mathcal{A}(\mathbb{X},\delta)} \in \mathcal{R}_\delta(\{(\vec{X}_j,\delta_j)\}_{j=1}^J) \neq \emptyset$.

In the other direction, if $\mathcal{R}_\delta(\{(\vec{X}_j,\delta_j)\}_{j=1}^J) \neq \emptyset$. Let $\mathcal{A}'$ denote an acceptance set such that $\rho_{\mathcal{A}'} \in \mathcal{R}_\delta(\{(\vec{X}_j,\delta_j)\}_{j=1}^J)$. Since $\rho(\vec{X}_j)=\delta_j$ implies that $\rho(\vec{X}_j+\delta_j)=0$, each random payoff $X_j+\delta_j$ must therefore be members of $\mathcal{A}'$. Provided that $\mathcal{A}(\mathbb{X},\delta)$ is the smallest monotone convex set containing all $\vec{X}_j+\delta_j$ (see proof of proposition 3.2), it is necessarily the case that $\mathcal{A}(\mathbb{X},\delta) \subseteq \mathcal{A}'$ and that

$$\rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_j+\delta_j) \geq \rho_{\mathcal{A}'}(\vec{X}_j+\delta_j) = \rho_{\mathcal{A}'}(\vec{X}_j) - \delta_j = 0 \;,$$

where the first equality comes from tanslation invariance and the second from the definition of $\mathcal{A}'$.

We are left with showing that for any δ such that $\mathcal{R}_\delta(\{(\vec{X}_j,\delta_j)\}_{j=1}^J) \neq 0$, and any Z, we have that

$$\sup_{\rho \in \mathcal{R}_\delta(\{(\vec{X}_j,\delta_j)\}_{j=1}^J)} \rho(\vec{Z}) = \rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{Z}) \;.$$

Yet, in this regard, we just showed that $\rho_{\mathcal{A}(\mathbb{X},\delta)}(\cdot)$ is always a member of $\mathcal{R}_\delta(\{(\vec{X}_j,\delta_j)\}_{j=1}^J)$ when the latter in non-empty, and is a subset of any member of $\mathcal{R}_\delta(\{(\vec{X}_j,\delta_j)\}_{j=1}^J)$. Hence, it must be that

$$\rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{Z}) \geq \sup_{\rho \in \mathcal{R}_\delta(\{(\vec{X}_j,\delta_j)\}_{j=1}^J)} \rho(\vec{Z}) \geq \rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{Z}) \;.$$

This completes the proof. □

In order to establish that the worst-case $\delta \in \Delta$ does not depend on the random payoff $\vec{Z}$ that is evaluated, we will show through the following two lemmas that: first $\rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{Z})$ is non-decreasing in δ; second, that problem (6) returns a vector $\bar{\delta}$ for which each entry is at the maximum value that it can achieve. Together with lemma B.1, this completes the proof of proposition 3.3 since we can conclude that

$$\max_{\delta \in \Delta} \; \rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{Z}) = \rho_{\mathcal{A}(\mathbb{X},\bar{\delta})}(\vec{Z}) \;.$$

Lemma B.2. *Given any random payoff Z, the risk measure $\rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{Z})$ is non-decreasing in δ.*

Proof. Let $\delta_1 \geq \delta_2$, and let (t_1,θ_1) and (t_2,θ_2) be the optimal solutions of

$$\begin{aligned} \min_{t,\theta} \quad & t \\ \text{subject to} \quad & \vec{Z}+t \geq \mathbb{X}\theta + \delta^\top\theta \\ & \vec{1}^\top\theta = 1,\; \theta \geq 0 \end{aligned}$$

when $\delta = \delta_1$ and $\delta = \delta_2$ respectively. Since $\theta_1 \geq 0$, we have

$$\vec{Z} + t_1 \geq \mathbb{X}\theta_1 + \delta_1^\top \theta_1 \geq \mathbb{X}\theta_1 + \delta_2^\top \theta_1 \ ,$$

hence (t_1, θ_1) is a feasible solution when $\delta = \delta_2$. Since t_2 is the optimal solution when $\delta = \delta_2$, we have $t_2 \leq t_1$. □

Lemma B.3. *Let $\bar{\delta}$ be the optimal solution of $\max_{\delta \in \Delta} \ \sum_j \delta_j$, then each $\bar{\delta}_i$ is the optimal value of $\max_{\delta \in \Delta} \ \ \delta_i$. Furthermore, the problem $\max_{\delta \in \Delta} \ \sum_j \delta_j$ is equivalent to problem* (6)*.*

Proof. We first present problem $\max_{\delta \in \Delta} \sum_j \delta_j$ in details to help with the discussion:

$$\begin{aligned} \max_{\delta} \quad & \sum_{j=1}^{J} \delta_j, && \text{(21a)} \\ \text{subject to} \quad & \delta_i \leq \delta_j \ , \forall \, (i,j) \in \bar{\mathcal{E}} && \text{(21b)} \\ & \rho_{\mathcal{A}(\mathbb{X},\delta)}(\vec{X}_j + \delta_j) \geq 0 \, , \forall \, j = 1, \ldots, J && \text{(21c)} \\ & \delta_1 = 0 \ . && \text{(21d)} \end{aligned}$$

Next, we let $\hat{\delta}^{(i)}$ be the optimal solution of problem (21) when the objective replaced with the objective of maximizing δ_i. One can actually show that the solution $\bar{\delta}$ composed such that $\bar{\delta}_i = \hat{\delta}_i^{(i)}$ is feasible for problem (21) and therefore an optimal solution of this problem.

In the case of constraint (21b), by construction of $\bar{\delta}$ we have that $\bar{\delta} \geq \hat{\delta}^{(i)}$ for all i. Therefore, one can confirm that for a pair $(i,j) \in \bar{\mathcal{E}}$

$$\bar{\delta}_i = \hat{\delta}_i^{(i)} \leq \hat{\delta}_j^{(i)} \leq \hat{\delta}_j^{(j)} = \bar{\delta}_j \ .$$

As for constraint (21c), for any fixed j, the representation of this constraint which is based on optimization states that

$$\begin{aligned} \min_{t,\theta} \ \ & t && \geq 0 \\ \text{subject to} \ \ & \vec{X}_j + \delta_j + t \geq \mathbb{X}\theta + \delta^\top \theta \\ & \vec{1}^\top \theta = 1 \ , \ \theta \geq 0 \ . \end{aligned}$$

After replacing $t' := t + \delta_j - \delta^\top \theta$, one obtains the equivalent constraint:

$$\begin{aligned} \min_{t',\theta} \ \ & t' - \delta_j + \delta^\top \theta && \geq 0 && \text{(22)} \\ \text{subject to} \ \ & \vec{X}_j + t' \geq \mathbb{X}\theta \\ & \vec{1}^\top \theta = 1 \ , \ \theta \geq 0 \ . \end{aligned}$$

One can verify that $\bar{\delta}$ satisfies this constraint since for all feasible t' and θ, one has that

$$t' - \bar{\delta}_j + \bar{\delta}^\top \theta = t' - \hat{\delta}_j^{(j)} + \sum_i \hat{\delta}_i^{(i)} \theta_i \geq t' - \hat{\delta}_j^{(j)} + \sum_i \hat{\delta}_i^{(j)} \theta_i \geq 0 \ .$$

In order to show that the optimization problem $\max_{\delta \in \Delta} \sum_j \delta_j$ is equivalent to problem (6), we make use of duality to reformulate each of the constrains (21c). In particular, for a fixed j the constraint takes the explicit form presented in (22). One can verify that Slater condition is always respected for this minimization problem with $t' = 0$ and $\theta = \boldsymbol{e}_j$ where $\boldsymbol{e}_j$ is the vector of all zeros except for a one at the j-th entry. Duality theory therefore states that the following constraint is equivalent:

$$\begin{aligned} \max_{y,\phi} \quad & -\vec{X}_j^{\mathsf{T}} y - \delta_j + \phi & \geq 0 \\ \text{subject to} \quad & \vec{X}_i^{\mathsf{T}} y + \delta_i - \phi \geq 0 \,, \forall\, i = 1, \ldots, J & \\ & \vec{1}^\top y = 1,\ y \geq 0 \,, & \end{aligned}$$

where $y \in \mathbb{R}^M$ and $\phi \in \mathbb{R}$ are the respective dual variables of the two constraints. Necessarily, the maximization evaluation of this constraint can be merged to the outer maximization problem (21) as long as y and ϕ are properly indexed with j. In other words, for each j, add $y_j \in \mathbb{R}^M$ and $\phi_j \in \mathbb{R}$ as decision variables and replace the constraint with:

$$\begin{aligned} & -\vec{X}_j^{\mathsf{T}} y_j - \delta_j + \phi_j \geq 0 \\ & \vec{X}_i^{\mathsf{T}} y_j + \delta_i - \phi_j \geq 0 \,, \forall\, i = 1, \ldots, J \\ & \vec{1}^\top y_j = 1,\ y_j \geq 0 \,. \end{aligned}$$

Yet, one can even see that the constraints reduce to

$$\begin{aligned} & \vec{X}_j^{\mathsf{T}} y_j + \delta_j - \phi_j = 0 \\ & \vec{X}_i^{\mathsf{T}} y_j + \delta_i - \phi_j \geq 0 \,, \forall\, i \neq j \\ & \vec{1}^\top y_j = 1,\ y_j \geq 0 \,, \end{aligned}$$

so that decision variable ϕ_j can be replaced to get

$$\begin{aligned} & (\vec{X}_i - \vec{X}_j)^{\mathsf{T}} y_j + \delta_i - \delta_j \geq 0 \,, \forall\, i \neq j \\ & \vec{1}^\top y_j = 1,\ y_j \geq 0 \,, \end{aligned}$$

This completes the proof as we are obtaining problem (6). □

C Proof of Proposition 3.12

Given all the set of distributions $\{F_j\}_{j=1,...,J}$ involved in $\mathcal{E}$ and the distribution of ξ, our first step is to convert each distribution to a random payoff in an outcome space endowed with a uniform probability measure. We can achieve this by first representing each probability value using a common denominator M. This can be done since each probability value is a rational number and there are a finite number of them. Without loss of generality, we can then assume that all distributions are induced from an outcome space $\Omega' := \{\tilde{\omega}_d\}_{d=1,...,M}$ endowed with a uniform probability measure P, i.e. $P(\{\tilde{\omega}_d\}) = 1/M$, $d = 1, ..., M$. In this space, a distribution F_j given in the form of

$P(X_j = x_k) = p_k^j,\ k = 1, ..., M_j$ can be expressed as a random payoff in Ω' that takes the form

$$\vec{X}'_j = h(F_j) = [\underbrace{(\vec{X}_j)_1\ \cdots\ (\vec{X}_j)_1}_{\pi_1^{(j)}}\ \underbrace{(\vec{X}_j)_2\ \cdots\ (\vec{X}_j)_2}_{\pi_2^{(j)}}\ \cdots \underbrace{(\vec{X}_j)_{M_j}\ \cdots\ (\vec{X}_j)_{M_j}}_{\pi_{M_j}^{(j)}}]^\top, \tag{23}$$

where $\pi_k^{(j)} = p_k^j \cdot M$ for $k = 1, ..., M_j$, and we use $h(\cdot)$ to stand for an operator that maps a given distribution to a random payoff as described above. Each entry of this M-dimensional vector corresponds to the mapping from an outcome $\tilde{\omega} \in \Omega'$ to a real value. A similar mapping can be used for $\vec{Z}(x)$ with

$$\vec{Z}'(x) = [\underbrace{r(x,\xi_1)\ \cdots\ r(x,\xi_1)}_{\pi_1^\xi}\ \underbrace{r(x,\xi_2)\ \cdots\ r(x,\xi_2)}_{\pi_2^\xi}\ \cdots\ \underbrace{r(x,\xi_{M_\xi})\ \cdots\ r(x,\xi_{M_\xi})}_{\pi_{M_\xi}^\xi}]^\top,$$

To facilitate the exposition of the proof below, we use l_k^j to denote the set of indexes $l_k^j := \{k' \in \mathbb{N} | (\vec{X}'_j)_{k'} = (\vec{X}_j)_k\}$. Using this notation, we can say that $(\vec{X}'_j)_d = (\vec{X}_j)_k$ for all $d \in l_k^j$. Similarly, we will refer to $l_k^\xi := \{k' \in \mathbb{N} | (\vec{Z}'(x))_{k'} = r(x, \xi_k)\}$. By such conversions, we can reformulate the problem (14) into

$$\min_{x \in \mathcal{X}} \sup_{\rho \in \mathcal{R}_{LE}(\{h(F_k^W), h(F_k^Y)\}_{k=1}^K)} \rho(\vec{Z}'(x)),$$

which admits a convex optimization formulation following Lemma 3.8 and Proposition 3.9. In the next two corollaries, we show how the convex reformulation can be further reduced.

Corollary C.1. *The optimization problem (10) in Lemma 3.8 with* $\vec{X}'_i = h(F_i)$ *and* $\vec{X}'_j = h(F_j)$ *can be reduced to the problem of (16).*

Proof. Consider first the constraints (10c) and (10d)

$$\vec{1}^\top v_{i,j} + \vec{1}^\top w_{i,j} - \vec{X}_j'^\top y_j + \delta_i - \delta_j \geq 0 \quad i \neq j \tag{24}$$

$$\vec{X}'_i y_j^\top - v_{i,j} \vec{1}^\top - \vec{1} w_{i,j}^\top \geq 0 \qquad i \neq j \tag{25}$$

For any fixed i and $k \in \{1, ..., M_i\}$, we can re-write the constraint (25) as

$$(v_{i,j})_d \leq (\vec{X}_i)_k (y_j)_m - (w_{i,j})_m,\ \forall d \in l_k^i, \forall m \in \{1, ..., M\}.$$

Observe that each entry of $(v_{i,j})_d$ in this range is bounded above by the same value for any fixed y_j^* and $w_{i,j}^*$. Observe also that increasing $(v_{i,j})_d$ will not violate any other constraint in (10). Thus, for any given optimal solution $v_{i,j}^*$ with $(v_{i,j}^*)_{d_1} \neq (v_{i,j}^*)_{d_2}$ for some $d_1, d_2 \in l_k^i$, we can always increase each entry of $(v_{i,j}^*)_d$ up to the same value (the upper bound) and obtain a new optimal solution v^{**} that satisfies $(v_{i,j}^*)_{d_1} = (v_{i,j}^*)_{d_2}$ for any $d_1, d_2 \in l_k^i$.

Therefore, we can impose $(v_{i,j})_d = (\tilde{v}_{i,j})_k, \forall d \in l_k^i$, where $\tilde{v}_{i,j} \in \mathbb{R}^{M_i}$ and reformulate the constraints into

$$\vec{1}^\top(\pi^{(i)} \circ \tilde{v}_{i,j}) + \vec{1}^\top w_{i,j} - \vec{X}_j'^\top y_j + \delta_j - \delta_i \geq 0 \quad \forall i \neq j \tag{26}$$

$$\vec{X}_i y_j^\top - \tilde{v}_{i,j}\vec{1}^\top - \vec{1}w_{i,j}^\top \geq 0 \quad \forall i \neq j. \tag{27}$$

Next, suppose that $\tilde{v}_{i,j}^*, w_{i,j}^*, y_j^*, \delta_i^*, \delta_j^*$ are an optimal solution of the problem (10) with (10c) and (10d) replaced by the above two sets of constraints (26) and (27). We claim that the solution $\tilde{v}_{i,j}^*, \delta_i^*, \delta_j^*$ together with the newly constructed y_j^{**} and $w_{i,j}^{**}$:

$$(y_j^{**})_{\hat{d}} := (1/\pi^{(j)}_{k(j,\hat{d})}) \sum_{d \in l^j_{k(j,\hat{d})}} (y_j^*)_d$$

$$(w_{i,j}^{**})_{\hat{d}} = (1/\pi^{(j)}_{k(j,\hat{d})}) \sum_{d \in l^j_{k(j,\hat{d})}} (w_{i,j}^*)_d \ ,$$

where $k(j,\hat{d})$ refers to the only index such that $\hat{d} \in l_k^j$, will also be optimal. Substituting this new solution into the constraint (26), we have

$$\begin{aligned}
& \vec{1}^\top(\pi^{(i)} \circ \tilde{v}_{i,j}^*) + \vec{1}^\top w_{i,j}^{**} - \vec{X}_j'^\top y_j^{**} + \delta_i^* - \delta_j^* \\
= \quad & \vec{1}^\top(\pi^{(i)} \circ \tilde{v}_{i,j}^*) + \textstyle\sum_{k=1}^{M_j} \pi_k^{(j)}(1/\pi_k^{(j)}) \sum_{d\in l_k^j}(w_{i,j}^*)_d \\
& - \textstyle\sum_{k=1}^{M_j} (\vec{X}_j)_k \pi_k^{(j)}(1/\pi_k^{(j)}) \sum_{d\in l_k^j}(y_j^*)_d + \delta_i^* - \delta_j^* \\
= \quad & \vec{1}^\top(\pi^{(i)} \circ \tilde{v}_{i,j}^*) + \vec{1}^\top w_{i,j}^* - \vec{X}_j'^\top y_j^* + \delta_i^* - \delta_j^* \qquad \geq 0
\end{aligned}$$

To verify the feasibility of the second constraint (27), let us consider the following derivations for any fixed j and fixed $k \in \{1, 2, \ldots, M_j\}$:

$$\begin{aligned}
& \vec{X}_i y_j^{*\top} - \tilde{v}_{i,j}^*\vec{1}^\top - \vec{1}w_{i,j}^{*\ \top} \geq 0 \\
& \Rightarrow \vec{X}_i(\sum_{d\in l_k^j}(y_j^*)_d) - \tilde{v}_{i,j}^*\pi_k^{(j)} - \vec{1}(\sum_{d\in l_k^j}(w_{i,j}^*)_d) \geq 0 \\
& \Rightarrow \vec{X}_i(1/\pi_k^{(j)})(\sum_{d\in l_k^j}(y_j^*)_d) - \tilde{v}_{i,j}^* - \vec{1}(1/\pi_k^{(j)})(\sum_{d\in l_k^j}(w_{i,j}^*)_d) \geq 0 \\
& \Rightarrow \vec{X}_i y_j^{**\top} - \tilde{v}_{i,j}^*\vec{1}^\top - \vec{1}w_{i,j}^{**\top} \geq 0.
\end{aligned}$$

where the first step is obtained by summing the columns in the range l_k^j for the matrix on the lefthand side of the inequality.

It is straightforward to see $\vec{1}^\top y_j^{**} = \vec{1}^\top y_j^* = 1$, which verifies the feasibility of the constraint (10e).

Thus, we can impose that $(y_j)_d = (\tilde{y}_j)_{k(j,d)}$ for all j and d, where $\tilde{y}_j \in \mathbb{R}^{M_j}$, and that $(w_{i,j})_d = (\tilde{w}_{i,j})_k \ d \in l_k^j$, where $\tilde{w}_{i,j} \in \mathbb{R}^{M_j}$, and reformulate the constraints (26), (27) and (10e) into

$$\begin{aligned}
\vec{1}^\top(\pi^{(i)} \circ \tilde{v}_{i,j}) + \vec{1}^\top(\pi^{(j)} \circ \tilde{w}_{i,j}) - \vec{X}_j^\top(\pi^{(j)} \circ \tilde{y}_j) + \delta_j - \delta_i \geq 0 \quad & \forall i \neq j \\
\vec{X}_i\tilde{y}_j^\top - \tilde{v}_{i,j}\vec{1}^\top - \vec{1}\tilde{w}_{i,j}^\top \geq 0 \quad & \forall i \neq j \\
\vec{1}^\top(\pi^{(j)} \circ \tilde{y}_j) = 1. \quad &
\end{aligned}$$

Let $\hat{v}_{i,j} = \pi^{(i)} \circ \tilde{v}_{i,j}$, $\hat{w}_{i,j} = \pi^{(j)} \circ \tilde{w}_{i,j}$, and $\hat{y}_j = \pi^{(j)} \circ \tilde{y}_j$. The above second constraint becomes

$$\vec{X}_i((\pi^{(j)})^{-1} \circ \hat{y}_j)^\top - ((\pi^{(i)})^{-1} \circ \hat{v}_{i,j})\vec{1}^\top - \vec{1}((\pi^{(j)})^{-1} \circ \tilde{\omega}_{i,j})^\top \geq 0 \ ,$$

where $(\pi^{(i)})^{-1}$ satisfies $(\pi^{(i)})^{-1} \circ (\pi^{(i)}) = \vec{1}$. Finally, multiplying $(\pi^{(i)}\vec{1}^\top)$ to the inequality we have

$$\begin{aligned} &(\pi^{(i)}\vec{1}^\top) \circ \left(\vec{X}_i((\pi^{(j)})^{-1} \circ \hat{y}_j)^\top - ((\pi^{(i)})^{-1}\hat{v}_{i,j})\vec{1}^\top - \vec{1}((\pi^{(j)})^{-1} \circ \tilde{w}_{i,j})^\top\right) \geq 0 \\ &\Rightarrow \Pi_{i,j} \circ \vec{X}_i\hat{y}_j{}^\top - \hat{v}_{i,j}\vec{1}^\top - \Pi_{i,j} \circ \vec{1}\tilde{\omega}_{i,j}^\top \geq 0, \end{aligned}$$

where $\Pi \in \mathbb{R}^{M_i \times M_j}$ and

$$\Pi_{i,j} = (\pi^{(i)})((\pi^{(j)})^{-1})^\top,$$

as described in the corollary. This completes the proof. □

Note that in the proof below, for simplicity we will use the notation

$$V_{(a_1:a_2,b_1:b_2)} = \begin{bmatrix} V_{a_1,b_1} & \cdots & V_{a_1,b_2} \\ \vdots & \ddots & \vdots \\ V_{a_2,b_1} & \cdots & V_{a_2,b_2} \end{bmatrix}$$

to describe a submatrix of a matrix V. The notation $V_{(:,k)}$ (respectively $V_{(k,:)}$) will refer to the kth-column (respectively kth-row) of the matrix V.

Corollary C.2. *The problem (13) in Proposition 3.9 with $\vec{Z}'(x)$ and $\vec{X}'_j = h(F_j)$ can be reduced to the problem of (15)*

Proof. Suppose that $x^*, t^*, \theta^*, \{Q_j^*\}$ are an optimal solution of (13). We claim that the solution x^*, t^*, θ^* together with the newly constructed Q_j^{**} such that for all $k \in \{1, 2, \ldots, M_j\}$:

$$Q_j^{**}(:, l_k^j) := (1/\pi_k^{(j)}) \sum_{d \in l_k^j} Q_j^*(:, d)\vec{1}^\top$$

is also optimal.

Substituting into the constraint (13b), we have that for each $j \in \{1, ..., J\}$,

$$Q_j^{**}\vec{X}'_j = \sum_{k=1}^{M_j} \pi_k^{(j)} \cdot ((\vec{X}_j)_k/\pi_k^{(j)}) \left(\sum_{d \in l_k^j} (Q_j^*)_{(:,d)}\right) = \sum_{k=1}^{M_j} (\vec{X}_j)_k \left(\sum_{d \in l_k^j} (Q_j^*)_{(:,d)}\right) = Q_j^*\vec{X}'_j.$$

Substituting into the constraint (13c), we have for each $j \in \{1, ..., J\}$

$$Q_j^{**}\vec{1} = \sum_{k=1}^{M_j} \pi_k^{(j)} \cdot (1/\pi_k^{(j)}) \left(\sum_{d \in l_k^j} (Q_j^*)_{(:,d)}\right) = Q_j^*\vec{1} = \theta_j^*.$$

Substituting into the constraint (13d), we have for each $j \in \{1, ..., J\}$ and for each $\tilde{d} \in \{1, ..., M\}$ we have

$$
\begin{aligned}
(Q_j^{**\top}\vec{1})_{\tilde{d}} &= (1/\pi^{(j)}_{k(j,\tilde{d})}) \sum_{m=1}^{M} \sum_{d \in l^j_{k(j,\hat{d})}} (Q_j^*)_{m,d} \\
&= (1/\pi^{(j)}_{k(j,\tilde{d})}) \sum_{d \in d \in l^j_{k(j,\hat{d})}} \sum_{m=1}^{M} (Q_j^*)_{m,d} \\
&= (1/\pi^{(j)}_{k(j,\tilde{d})}) \sum_{d \in d \in l^j_{k(j,\hat{d})}} \theta_j^* = \theta_j^*.
\end{aligned}
$$

As $x^*, t^*, \theta^*, \{Q_j^{**}\}$ satisfy all constraints and t^* remains the same minimum value, $x^*, t^*, \theta^*, \{Q_j^{**}\}$ are also an optimal solution. Therefore, we can impose that for all j and all $k \in \{1, 2, \ldots, M_j\}$ we have that $(Q_j)_{(:,l_k^j)} = (\tilde{Q}_j)_{(:,k)}\vec{1}^\top$, where $\tilde{Q}_j \in \mathbb{R}^{M \times M_j}$, and reformulate (13) into

$$
\begin{aligned}
\min_{x \in \mathcal{X}, t, \theta, \{\tilde{Q}_j\}} \quad & t && \text{(28a)} \\
\text{subject to} \quad & \vec{Z}'(x) + t \geq \sum_j ((\vec{1}\pi^{(j)\top}) \circ \tilde{Q}_j)\vec{X}_j + \vec{\delta}^\top \theta && \text{(28b)} \\
& ((\vec{1}\pi^{(j)\top}) \circ \tilde{Q}_j)\vec{1} = \theta_j, \forall j && \text{(28c)} \\
& \tilde{Q}_j^{\top}\vec{1} = \theta_j, \forall j && \text{(28d)} \\
& \vec{1}^\top \theta = 1 && \text{(28e)} \\
& \theta \geq 0, \tilde{Q}_j \geq 0, \forall j. && \text{(28f)}
\end{aligned}
$$

Next, suppose that $x^*, t^*, \theta^*, \{\tilde{Q}_j^*\}$ is the optimal solution for the above problem. We claim that that the solution x^*, t^*, θ^* together with the newly constructed $\tilde{Q}_j^{**}$ such that for all $k = 1, ..., M_\xi$ and all $\tilde{d} \in l_k^\xi$ we have that

$$
(\tilde{Q}_j^{**})_{(\tilde{d},:)} = (1/\pi_k^\xi) \sum_{d \in l_k^\xi} (\tilde{Q}_j^*)_{(d,:)}
$$

is also optimal.

Substituting into the constraint (28b), we have for all $\tilde{d} \in \{1, ..., M\}$ we have that

$$
\begin{aligned}
\left(\sum_j ((\vec{1}\pi^{(j)\top}) \circ \tilde{Q}_j^{**})\vec{X}_j \right)_{\tilde{d}} &= (1/\pi^\xi_{k(j,\tilde{d})}) \left(\textstyle\sum_{d \in l^\xi_{k(j,\tilde{d})}} \sum_j (\pi^{(j)\top} \circ \tilde{Q}_j^*(d,:))\vec{X}_j \right) \\
&\leq (1/\pi^\xi_{k(j,\tilde{d})}) (\textstyle\sum_{d \in l^\xi_{k(j,\tilde{d})}} (\vec{Z}'(x))_d + t^* - \vec{\delta}^\top \theta^*) \\
&= (1/\pi^\xi_{k(j,\tilde{d})}) (\pi^\xi_{k(j,\tilde{d})} ((\vec{Z}'(x))_{\tilde{d}} + t^* - \vec{\delta}^\top \theta^*) \\
&= (\vec{Z}'(x))_{\tilde{d}} + t^* - \vec{\delta}^\top \theta^*.
\end{aligned}
$$

Substituting into the constraint (28c), we have for each j and for all $\tilde{d} \in \{1, ..., M\}$ that

$$\left(((\vec{1}\pi^{(j)\top}) \circ \tilde{Q}_j^{**})\vec{1}\right)_{\tilde{d}} = (1/\pi^{\xi}_{k(j,\tilde{d})}) \sum_{m=1}^{M_j} \pi_m^{(j)} \sum_{d \in l^{\xi}_{k(j,\tilde{d})}} (\tilde{Q}_j^*)_{(d,m)}$$

$$=(1/\pi^{\xi}_{k(j,\tilde{d})}) \sum_{d \in l^{\xi}_{k(j,\tilde{d})}} \sum_{m=1}^{M_j} (\pi_m^{(j)})(\tilde{Q}_j^*)_{(d,m)}$$

$$=(1/\pi^{\xi}_{k(j,\tilde{d})}) \sum_{d \in l^{\xi}_{k(j,\tilde{d})}} \theta_j^* = \theta_j^*.$$

Substituting into the constraint (28d), we have

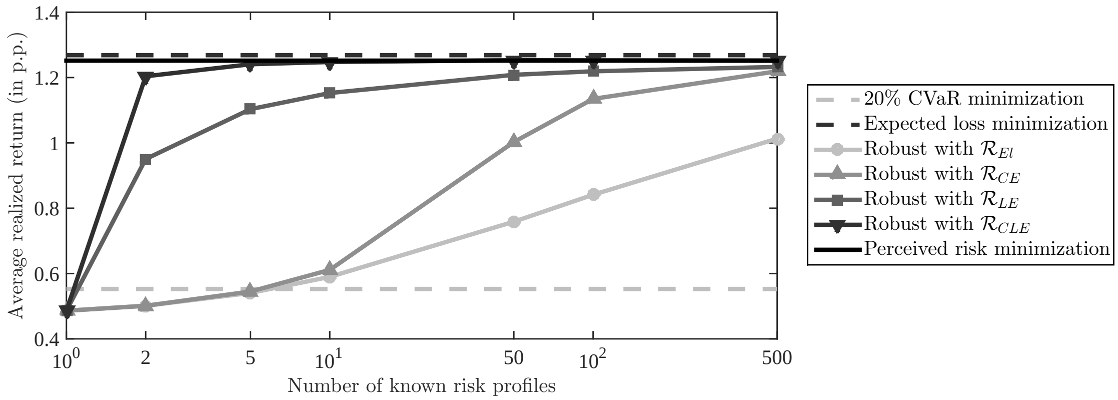
$$\tilde{Q}_j^{**\top}\vec{1} = \sum_{k=1}^{M_\xi} \pi_k^{\xi} \cdot (1/\pi_k^{\xi}) \sum_{d \in l_k^{\xi}} (\tilde{Q}_j^*)_{(d,:)} = \tilde{Q}_j^{*\top}\vec{1} = \theta_j^*.$$

As $x^*, t^*, \theta^*, \{\tilde{Q}_j^{**}\}$ satisfy all constraints and t^* remains the same minimum value, $x^*, t^*, \theta^*, \tilde{Q}_j^{**}$ are also an optimal solution. We can now impose that for all $k \in \{1, 2, \ldots, M_\xi\}$ we have that $(\tilde{Q}_j)_{(l_k^{\xi},:)} = \vec{1}(\hat{Q}_j)_{(k,:)}$, where $\hat{Q}_j \in \mathbb{R}^{M_\xi \times M_j}$, and reformulate the problem (28) into

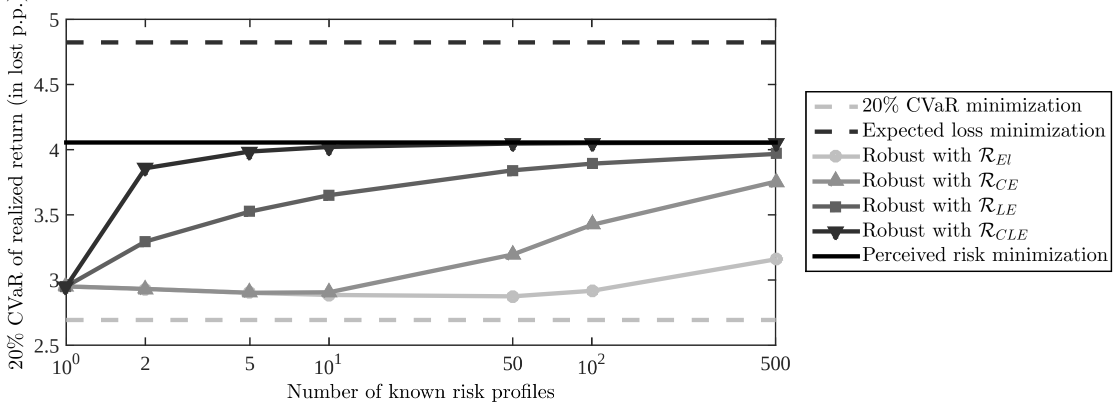
$$\begin{aligned}
\min_{x \in \mathcal{X}, t, \theta, \{\hat{Q}_j\}} \quad & t \\
\text{subject to} \quad & \vec{Z}_\xi(x) + t \geq \sum_j ((\vec{1}\pi^{(j)\top}) \circ \hat{Q}_j)\vec{X}_j + \vec{\delta}^\top \theta, \\
& ((\vec{1}\pi^{(j)\top}) \circ \hat{Q}_j)\vec{1} = \theta_j, \forall j \\
& ((\pi^{\xi}\vec{1}^\top) \circ \hat{Q}_j)^\top \vec{1} = \theta_j, \forall j \\
& \vec{1}^\top \theta = 1 \\
& \theta \geq 0, \hat{Q}_j \geq 0, \forall j.
\end{aligned}$$

Finally, Let $\hat{Q}'_j = (\vec{1}\pi^{(j)\top}) \circ \hat{Q}_j$. The left-hand-side of the above third constraint can be written as $((\pi^{\xi}\vec{1}^\top) \circ ((\vec{1}(\pi^{(j)})^{-1\top}) \circ \hat{Q}'_j) = (\pi^{\xi}((\pi^{(j)})^{-1})^\top) \circ \hat{Q}'_j$. Having $\Pi_j = \pi^{\xi}((\pi^{(j)})^{-1})^\top$, we arrive at the formulation.

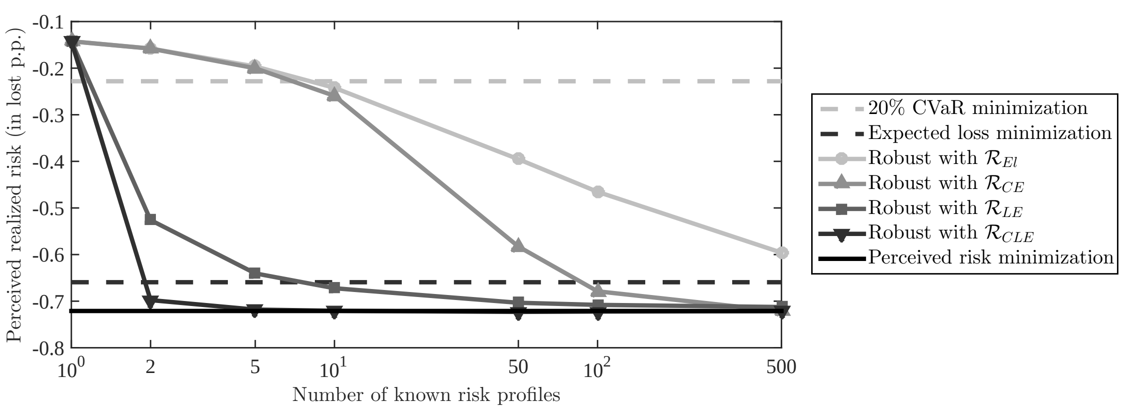
□



(a)



(b)



(c)

Figure 3: Statistic about the out of sample performance of the portfolios proposed by the different methods when implementing then on the week following the analysis. (a) compares the average return over the 5000 experiments. (b) compares the 20% CVaR of the empirical distribution described by the returns obtained in the 5000 experiments. (c) presents the risk as perceived by the investor for this distribution.