International Journal of Theoretical and Applied Finance
Vol. 20, No. 5 (2017) 1750026 (48 pages)
© World Scientific Publishing Company
DOI: 10.1142/S0219024917500261

World Scientific
www.worldscientific.com

SET-VALUED SHORTFALL AND DIVERGENCE RISK MEASURES

ÇAĞIN ARARAT
Department of Industrial Engineering
Bilkent University, Ankara 06800, Turkey
cararat@bilkent.edu.tr

ANDREAS H. HAMEL
School for Economics and Management
Free University Bozen
Bozen-Bolzano 39031, Italy
andreas.hamel@unibz.it

BIRGIT RUDLOFF
Institute for Statistics and Mathematics
Vienna University of Economics and Business
Vienna 1020, Austria
brudloff@wu.ac.at

Received 30 August 2016
Accepted 28 March 2017
Published 16 May 2017

Risk measures for multivariate financial positions are studied in a utility-based framework. Under a certain incomplete preference relation, shortfall and divergence risk measures are defined as the optimal values of specific set minimization problems. The dual relationship between these two classes of multivariate risk measures is constructed via a recent Lagrange duality for set optimization. In particular, it is shown that a shortfall risk measure can be written as an intersection over a family of divergence risk measures indexed by a scalarization parameter. Examples include set-valued versions of the entropic risk measure and the average value at risk. As a second step, the minimization of these risk measures subject to trading opportunities is studied in a general convex market in discrete time. The optimal value of the minimization problem, called the market risk measure, is also a set-valued risk measure. A dual representation for the market risk measure that decomposes the effects of the original risk measure and the frictions of the market is proved.

Keywords: Optimized certainty equivalent; shortfall risk; divergence; relative entropy; entropic risk measure; average value at risk; set-valued risk measure; multivariate risk; incomplete preference; transaction cost; solvency cone; liquidity risk; infimal convolution; Lagrange duality; set optimization.

1. Introduction

Risk measures for random vectors have recently gained interest in the financial mathematics community. Introduced in the pioneering work Jouini *et al.* (2004), *set-valued risk measures* have been used to quantify *financial risk* in markets with frictions such as transaction costs or illiquidity effects. These risk measures are functions which assign to an m-dimensional random vector X a set $R(X) \subseteq \mathbb{R}^m$ whose elements can be used as risk compensating portfolios. Here, X denotes a financial position in m assets whose components are in terms of physical units rather than values with respect to a specific numéraire. More recently, set-valued risk measures have also been used to quantify *systemic risk* in financial networks; see Feinstein *et al.* (2015), Ararat & Rudloff (2016). In this case, m is the number of financial institutions and the components of X denote the corresponding magnitudes of a random shock (equity/loss) for these institutions.

The coherent set-valued risk measures in Jouini *et al.* (2004) have been extended to the convex case in Hamel & Heyde (2010) and to random market models in Hamel *et al.* (2011). These extensions were possible by an application of the duality theory and, in particular, the Moreau–Fenchel biconjugation theorem for set-valued functions developed in Hamel (2009). Extensions to the dynamic framework have been studied in Feinstein & Rudloff (2013, 2015a, 2015b), Ben Tahar & Lepinette (2014) and to set-valued portfolio arguments in Cascos & Molchanov (2013). Scalar risk measures for multivariate random variables, which can be interpreted as scalarizations of set-valued risk measures (see Feinstein & Rudloff 2015b, Sec. 2.4), have been studied in Jouini & Kallal (1995), Burgert & Rüschendorf (2006), Weber *et al.* (2013) (financial risk) as well as in Chen *et al.* (2013), Biagini *et al.* (2015) (systemic risk).

Set-valued generalizations of some well-known scalar coherent risk measures have already been studied such as the set-valued version of the average value at risk in Hamel *et al.* (2013), Feinstein & Rudloff (2015a), Hamel *et al.* (2014), or the set of superhedging portfolios in markets with transaction costs in Hamel *et al.* (2011), Löhne & Rudloff (2013), Feinstein & Rudloff (2013). Other examples of coherent risk measures for multivariate claims can be found in Ben Tahar (2006), Cascos & Molchanov (2013). To the best of our knowledge, apart from superhedging with certain trading constraints in markets with frictions, which leads to set-valued convex risk measures (see Hamel *et al.* 2014), no other examples have been studied in the convex case yet.

This paper introduces utility-based convex risk measures for random vectors. The basic assumption is that the investor has a complete risk preference towards each asset which has a numerical representation in terms of a von Neumann–Morgenstern loss (utility) function. However, her risk preference towards multivariate positions is incomplete and it can be represented in terms of the vector of individual loss functions. Based on this incomplete preference, the *shortfall risk* of the random vector X is defined as the collection of all portfolios $z \in \mathbb{R}^m$ for

which $X + z$ is preferred to a benchmark portfolio $z^0 \in \mathbb{R}^m$. As an example, when the individual loss functions are exponential, we obtain set-valued versions of the well-known *entropic risk measure* (see Föllmer & Schied 2002, 2011).

We formulate the computation of the shortfall risk measure as a constrained set optimization problem and apply recent tools from the set optimization literature to obtain a dual formulation. In particular, using the Lagrange duality in Hamel & Löhne (2014), another type of convex risk measures, called *divergence risk measures*, are obtained in the dual problem. A divergence risk measure is defined based on the trade-off between consuming a deterministic amount $z \in \mathbb{R}^m$ of the position today and realizing the expected loss of the remaining amount $X - z$ at terminal time. The decision making problem is bi-objective: The investor wants to choose a portfolio z so as to *maximize* z and *minimize* the (vector-valued) expected loss due to $X - z$ at the same time, both of which are understood in the sense of set optimization (see Sec. 3.3). The two objectives are combined by means of a relative weight (scalarization) parameter $r \in \mathbb{R}^m_+$ and the divergence risk of X is defined as an unconstrained set optimization problem over the choices of the deterministic consumption z. As special cases, we obtain the definition of the set-valued average value at risk given in Hamel *et al.* (2013) as well as a convex version of it.

One of the main results of this paper is that a shortfall risk measure can be written as the intersection of divergence risk measures indexed by their relative weights and, in general, the intersection is not attained by a unique relative weight. Hence, the shortfall risk measure is a (much) more conservative risk measure than a divergence risk measure. While the shortfall risk measure is more difficult to compute as a constrained optimization problem, we show that the computation of a divergence risk measure can be reduced to the computation of scalar divergence risk measures (*optimized certainty equivalents* in Ben Tal & Teboulle 1986, 2007). On the flip side, and in contrast to shortfall risk measures, to be able to use a divergence risk measure, the decision maker has to specify the relative weight of her loss with respect to her consumption for each asset.

While shortfall and divergence risk measures are defined based on the preferences of the investor, they do not take into account how the market frictions affect the riskiness of a position. In Sec. 5, we propose a method for incorporating these frictions in the computation of risk. We generalize the notion of *market risk measure* (see Hamel *et al.* 2013 with the name *market-extension*) by including trading constraints modeled by convex random sets, and considering issues of liquidation into a certain subcollection of the assets. In contrast to Hamel *et al.* (2013), we allow for a convex (and not necessarily conical) market model to include temporary illiquidity effects in which the bid-ask prices depend on the magnitude of the trade, and thus, are given by the shape of the limit order book; see Astic & Touzi (2007), Pennanen & Penner (2010), for instance. Letting R be a (market-free) risk measure such as a shortfall or divergence risk measure, its induced market risk measure is defined as the minimized value of R over the set of all financial positions that are

attainable by trading in the market. As the second main result of the paper, we prove a dual representation theorem for the market risk measure (Theorem 5.1). In particular, we show that the penalty (Fenchel conjugate) function of the market risk measure is the sum of the penalty function of the base risk measure R and the supporting halfspaces of the convex regions of the market.

The rest of this paper is organized as follows. In Sec. 2, we review the scalar theory of shortfall and divergence risk measures. However, we generalize the standard results in the literature as we allow for extended real-valued loss functions and we do not impose any growth conditions on the loss functions as in Föllmer & Schied (2002), Ben Tal & Teboulle (2007). The main part of the paper is Sec. 3, where set-valued shortfall and divergence risk measures are studied. In Sec. 4, set-valued entropic risk measures are studied as examples of shortfall risk measures and set-valued average value at risks are recalled as examples of divergence risk measures. Market risk measures in a general convex market model with liquidation and trading constraints are studied in Sec. 5. All proofs are collected in Sec. 6.

2. Scalar Shortfall and Divergence Risk Measures

In this section, we summarize the theory of (utility/loss-based) shortfall and divergence risk measures for univariate financial positions. Shortfall risk measures are introduced in Föllmer & Schied (2002). Divergence risk measures are introduced in Ben Tal & Teboulle (1986), and analyzed further in Ben Tal & Teboulle (2007) with the name *optimized certainty equivalents* and in Cherny & Kupper (2007) with the name *divergence utilities* for their negatives. The dual relationship between shortfall and divergence risk measures is pointed out in Schied (2007) and Ben Tal & Teboulle (2007). In terms of the assumptions on the underlying loss function, we generalize the results of these papers by dropping growth conditions; see Sec. 6.2 for a comparison.

The proofs of the results of this section are given in Sec. 6.1 and most of them inherit the convex duality arguments in Ben Tal & Teboulle (2007) rather than the analytic arguments in Föllmer & Schied (2002).

Definition 2.1. A lower semicontinuous, convex function $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ with effective domain $\operatorname{dom} f = \{x \in \mathbb{R} \,|\, f(x) < +\infty\}$ is said to be a *loss function* if it satisfies the following properties:

(1) f is nondecreasing with $\inf_{x\in\mathbb{R}} f(x) > -\infty$.
(2) $0 \in \operatorname{dom} f$.
(3) f is not identically constant on $\operatorname{dom} f$.

Throughout this section, let $\ell : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a loss function. Definition 2.1 above guarantees that $\operatorname{int} \ell(\mathbb{R}) \neq \emptyset$, where int denotes the interior operator. Let us fix a threshold level $x^0 \in \operatorname{int} \ell(\mathbb{R})$ for expected loss values. Without loss of generality, we assume $x^0 = 0$. Based on the loss function ℓ, we define the shortfall

risk measure on the space L^∞ of essentially bounded real-valued random variables of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where random variables are identified up to almost sure equality.

Definition 2.2. The function $\rho_\ell : L^\infty \to \mathbb{R} \cup \{\pm\infty\}$ defined by

$$\rho_\ell(X) = \inf\{s \in \mathbb{R} \,|\, \mathbb{E}[\ell(-X - s)] \leq 0\} \tag{2.1}$$

is called the *shortfall risk measure*.

Proposition 2.1. *The function ρ_ℓ is a (weak*-)lower semicontinuous convex risk measure in the sense of Föllmer & Schied* (2011, *Definitions* 4.1 *and* 4.4)*. In particular, ρ_ℓ takes values in $\mathbb{R}$.*

Remark 2.1. Since $\inf_{x\in\mathbb{R}} \ell(x) > -\infty$, it holds $\mathbb{E}[\ell(-X - s)] > -\infty$ for every $X \in L^\infty, s \in \mathbb{R}$. Hence, the expectation in (2.1) is always well-defined. Moreover, the assumption $x^0 = 0 \in \operatorname{int} \ell(\mathbb{R})$ is essential for the finiteness of $\rho_\ell(X)$ as shown in the proof of Proposition 2.1; see Sec. 6.1.

According to Definition 2.2, the number, $\rho_\ell(X)$ can be seen as the optimal value of a convex minimization problem. The next proposition computes $\rho_\ell(X)$ as the optimal value of the corresponding Lagrangian dual problem. Its proof in Sec. 6.1 is an easy application of strong duality.

Proposition 2.2. *For every $X \in L^\infty$,*

$$\rho_\ell(X) = \sup_{\lambda\in\mathbb{R}_+} \delta_{\ell,\lambda}(X), \tag{2.2}$$

where

$$\begin{aligned} \delta_{\ell,\lambda}(X) &:= \inf_{s\in\mathbb{R}:\mathbb{E}[\ell(-X-s)]<+\infty} (s + \lambda\mathbb{E}[\ell(-X - s)]) \\ &= \begin{cases} \inf_{s\in\mathbb{R}}(s + \lambda\mathbb{E}[\ell(-X - s)]), & \text{if } \lambda > 0, \\ -\operatorname{ess\,inf} X - \sup \operatorname{dom} \ell, & \text{if } \lambda = 0. \end{cases} \end{aligned} \tag{2.3}$$

Note that $\delta_{\ell,\lambda}$ is a monotone and translative function on L^∞ for each $\lambda \in \mathbb{R}_+$. Our aim is to determine the values of λ for which this function is a lower semicontinuous convex risk measure with values in $\mathbb{R}$. To that end, we define the Legendre–Fenchel conjugate $g : \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ of the loss function by

$$g(y) := \ell^*(y) = \sup_{x\in\mathbb{R}}(xy - \ell(x)). \tag{2.4}$$

In the following, we will adopt the convention $(+\infty) \cdot 0 = 0$ as usual in convex analysis, see Rockafellar & Wets (1998). We will also use $\frac{1}{+\infty} = 0$ as well as $\frac{1}{0} = +\infty$.

Definition 2.3. A proper, convex, lower semicontinuous function $\varphi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ with effective domain $\operatorname{dom}\varphi = \{y \in \mathbb{R} \,|\, \varphi(y) < +\infty\}$ is said to be a

divergence function if it satisfies the following properties:

(1) $0 \in \operatorname{dom}\varphi \subseteq \mathbb{R}_+$.
(2) φ attains its infimum over $\mathbb{R}$.
(3) φ is not of the form $y \mapsto +\infty \cdot 1_{\{y<0\}} + (ay+b)\cdot 1_{\{y\geq 0\}}$ with $a \in \mathbb{R}_+ \cup \{+\infty\}$ and $b \in \mathbb{R}$.

Proposition 2.3. *Legendre–Fenchel conjugation furnishes a bijection between loss and divergence functions.*

Remark 2.2. Let $\lambda > 0$. If f is a loss function, then λf is also a loss function. If φ is a divergence function, then the function $y \mapsto \varphi_\lambda(y) := \lambda\varphi(\frac{y}{\lambda})$ on $\mathbb{R}$ is also a divergence function. The functions f and φ are conjugates of each other if and only if λf and φ_λ are. In this case, we also define the *recession function* $\varphi_0 : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ of φ by

$$\varphi_0(y) := \sup_{\lambda>0}(\varphi_\lambda(y) - \lambda\varphi(0)) = \lim_{\lambda\downarrow 0}\varphi_\lambda(y) = \begin{cases} y \sup \operatorname{dom} f, & \text{if } y \geq 0, \\ +\infty, & \text{if } y < 0, \end{cases} \tag{2.5}$$

for each $y \in \mathbb{R}$. Here, $\lambda \mapsto \varphi_\lambda(y) - \lambda g(0)$ is a nonincreasing convex function on $\mathbb{R}_{++}$ for each $y \in \mathbb{R}$. Moreover, the second equality holds thanks to the assumption $0 \in \operatorname{dom}\varphi$, see Rockafellar (1970, Theorem 8.5, Corollary 8.5.2). The last equality is due to the fact that the support function of the effective domain of the proper convex function f coincides with the recession function φ_0 of its conjugate, see Rockafellar (1970, Theorem 13.3).

We next recall the notion of divergence. To that end, let $\mathcal{M}(\mathbb{P})$ be the set of all probability measures on $(\Omega, \mathcal{F})$ that are absolutely continuous with respect to $\mathbb{P}$.

Definition 2.4. Let φ be a divergence function with the corresponding loss function f. For $\lambda \in \mathbb{R}_+$ and $\mathbb{Q} \in \mathcal{M}(\mathbb{P})$, the quantity

$$I_{\varphi,\lambda}(\mathbb{Q}\,|\,\mathbb{P}) := \mathbb{E}\left[\varphi_\lambda\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] = \begin{cases} \lambda\mathbb{E}\left[\varphi\left(\frac{1}{\lambda}\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] & \text{if } \lambda > 0, \\ \sup \operatorname{dom} f & \text{if } \lambda = 0 \end{cases} \tag{2.6}$$

is called the (φ,λ)-*divergence* of $\mathbb{Q}$ with respect to $\mathbb{P}$.

Remark 2.3. $I_{\varphi,1}$ is the usual φ-divergence in the sense of Csiszár (1967). It is a notion of "distance" between probability measures and includes the well-known *relative entropy* as a special case, see (4.12).

Note that $g = \ell^*$ is a divergence function, and $\operatorname{dom} g$ is an interval of the form $[0,\beta)$ or $[0,\beta]$ for some $\beta \in \mathbb{R}_{++} \cup \{+\infty\}$. Here, we have $\operatorname{dom} g \neq \{0\}$ since otherwise g would be of the form $y \mapsto +\infty \cdot 1_{\{y<0\}} + (ay+b)\cdot 1_{\{y\geq 0\}}$ for $a = +\infty$ and $b = g(0)$. For each $\lambda > 0$, $y \mapsto g_\lambda(y) := \lambda g(\frac{y}{\lambda})$ on $\mathbb{R}$ is a divergence function

with $\operatorname{dom} g_\lambda = [0, \lambda\beta)$ or $\operatorname{dom} g_\lambda = [0, \lambda\beta]$ by Remark 2.2, and the corresponding (g, λ)-divergence is defined according to Definition 2.4. In the case $\lambda = 0$, $y \mapsto g_0(y) = +\infty \cdot 1_{\{y<0\}} + (\sup \operatorname{dom} \ell) y \cdot 1_{\{y \geq 0\}}$ on $\mathbb{R}$ is not a divergence function. Moreover, we have $\operatorname{dom} g_0 = \{0\}$ if $\operatorname{dom} \ell = \mathbb{R}$, and $\operatorname{dom} g_0 = \mathbb{R}_+$ if $\operatorname{dom} \ell \neq \mathbb{R}$.

Theorem 2.1. *For every $\lambda \in \mathbb{R}_+$ and $X \in L^\infty$,*

$$\delta_{\ell,\lambda}(X) = \sup_{\mathbb{Q} \in \mathcal{M}(\mathbb{P})} (\mathbb{E}^{\mathbb{Q}}[-X] - I_{g,\lambda}(\mathbb{Q} \,|\, \mathbb{P})). \tag{2.7}$$

Moreover, $\delta_{\ell,\lambda}$ is a lower semicontinuous convex risk measure if $1 \in \operatorname{dom} g_\lambda$, and $\delta_{\ell,\lambda}(X) = -\infty$ for every $X \in L^\infty$ otherwise. Hence,

$$\rho_\ell(X) = \sup_{\lambda \in \mathbb{R}_+ : 1 \in \operatorname{dom} g_\lambda} \delta_{\ell,\lambda}(X). \tag{2.8}$$

In particular, if $\operatorname{dom} \ell = \mathbb{R}$, then

$$\rho_\ell(X) = \sup_{\lambda > 0 : \frac{1}{\lambda} \in \operatorname{dom} g} \delta_{\ell,\lambda}(X). \tag{2.9}$$

Definition 2.5. For $\lambda \in \mathbb{R}_+$ with $1 \in \operatorname{dom} g_\lambda$, the function $\delta_{\ell,\lambda} : L^\infty \to \mathbb{R}$ is called the *divergence risk measure* with weight λ.

In (2.7), note that a divergence risk measure is represented in terms of probability measures. More generally, by Föllmer & Schied (2011, Theorem 4.33), every lower semicontinuous convex risk measure $\rho : L^\infty \to \mathbb{R}$ has a *dual representation* in the sense that it is characterized by its so-called *penalty function* $\alpha_\rho : \mathcal{M}(\mathbb{P}) \to \mathbb{R} \cup \{+\infty\}$ by the following relationships:

$$\rho(X) = \sup_{\mathbb{Q} \in \mathcal{M}(\mathbb{P})} (\mathbb{E}^{\mathbb{Q}}[-X] - \alpha_\rho(\mathbb{Q})), \quad \alpha_\rho(\mathbb{Q}) = \sup_{X \in L^\infty} (\mathbb{E}^{\mathbb{Q}}[-X] - \rho(X)). \tag{2.10}$$

In Proposition 2.4, we check that (2.7) is indeed the dual representation of the divergence risk measure $\delta_{\ell,\lambda}$. We also compute the penalty function of the shortfall risk measure in terms of the penalty functions of divergence risk measures.

Proposition 2.4. *Let $\lambda \in \mathbb{R}_+$ with $1 \in \operatorname{dom} g_\lambda$. For each $\mathbb{Q} \in \mathcal{M}(\mathbb{P})$, it holds*

$$\alpha_{\delta_{\ell,\lambda}}(\mathbb{Q}) = I_{g,\lambda}(\mathbb{Q} \,|\, \mathbb{P}), \tag{2.11}$$

and moreover,

$$\alpha_{\rho_\ell}(\mathbb{Q}) = \inf_{\lambda \in \mathbb{R}_+} I_{g,\lambda}(\mathbb{Q} \,|\, \mathbb{P}) = \inf_{\lambda \in \mathbb{R}_+ : 1 \in \operatorname{dom} g_\lambda} \alpha_{\delta_{\ell,\lambda}}(\mathbb{Q}). \tag{2.12}$$

3. Set-Valued Shortfall and Divergence Risk Measures

In this section, we introduce utility-based shortfall and divergence risk measures for multivariate financial positions, the central objects of this paper. The proofs are presented in Sec. 6.4.

Let us introduce some notation that will be used frequently throughout the rest of the paper. Let $m \geq 1$ be an integer and $|\cdot|$ an arbitrary fixed norm on $\mathbb{R}^m$.

By $\mathbb{R}^m_+$ and $\mathbb{R}^m_{++}$, we denote the set of elements of $\mathbb{R}^m$ with nonnegative and strictly positive components, respectively.

Throughout, we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $L^0_m := L^0_m(\Omega, \mathcal{F}, \mathbb{P})$ the linear space of random variables taking values in $\mathbb{R}^m$, where two elements are identified if they are equal $\mathbb{P}$-almost surely; and we define

$$\begin{aligned} L^1_m &= \{X \in L^0_m \,|\, \mathbb{E}[|X|] < +\infty\}, \\ L^\infty_m &= \{X \in L^0_m \,|\, \operatorname{ess\,sup} |X| < +\infty\}, \\ L^p_{m,+} &= \{X \in L^p_m \,|\, \mathbb{P}\{X \in \mathbb{R}^m_+\} = 1\}, \quad p \in \{1, +\infty\}. \end{aligned} \tag{3.1}$$

Componentwise ordering of vectors is denoted by $\leq$, that is, for $x, z \in \mathbb{R}^m$, it holds $x \leq z$ if and only if $x_i \leq z_i$ for each $i \in \{1, \ldots, m\}$. The Hadamard product of $x, z \in \mathbb{R}^m$ is defined by $x \cdot z := (x_1 z_1, \ldots, x_m z_m)^\mathsf{T}$. We denote by $\mathcal{P}(\mathbb{R}^m)$ the power set of $\mathbb{R}^m$, that is, the set of all subsets of $\mathbb{R}^m$ including the empty set $\emptyset$. On $\mathcal{P}(\mathbb{R}^m)$, the Minkowski addition and multiplication with scalars are defined by $A + B = \{a + b \,|\, a \in A,\ b \in B\}$ and $sA = \{sa \,|\, a \in A\}$ for $A, B \subseteq \mathbb{R}^m$ and $s \in \mathbb{R}$ with the conventions $A + \emptyset = \emptyset + B = \emptyset + \emptyset = \emptyset$, $s\emptyset = \emptyset$ $(s \neq 0)$, and $0\emptyset = \{0\} \subseteq \mathbb{R}^m$. We also use the shorthand notations $A - B = A + (-1)B$ and $z + A = \{z\} + A$. For $x \in \mathbb{R}^m$ and a nonempty set $A \subseteq \mathbb{R}^m$, we set $x \cdot A := \{x \cdot a \,|\, a \in A\}$. These operations can be defined on the power set $\mathcal{P}(L^p_m)$ of L^p_m, $p \in \{0, 1, \infty\}$, in a similar way. (In)equalities between random variables are understood in the $\mathbb{P}$-almost sure sense.

3.1. *The incomplete preference relation*

Let $m \geq 1$ be an integer denoting the number of assets in a financial market. The linear space $\mathbb{R}^m$ is called the space of eligible portfolios. This means that every $z \in \mathbb{R}^m$ is a potential deposit to be used at initial time in order to compensate for the risk of a financial position.

We model a financial position as an element $X \in L^\infty_m$, where $X_i(\omega)$ represents the number of physical units in the ith asset for $i \in \{1, \ldots, m\}$ when the state of the world $\omega \in \Omega$ occurs. We assume that the investor has a (possibly) incomplete preference relation for multivariate financial positions in L^∞_m. Its numerical representation is in terms of the individual loss functions for the assets and a comparison rule for the vectors of expected losses:

(1) **Loss functions for assets:** We assume that the investor has a complete preference relation $\succeq_i$ on L^∞ corresponding to each asset $i \in \{1, \ldots, m\}$ and this preference relation has a von Neumann–Morgenstern representation given by a (scalar) loss function $\ell_i : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ (see Definition 2.1). That is, for $X_i, Z_i \in L^\infty$,

$$X_i \succeq_i Z_i \Leftrightarrow \mathbb{E}[\ell_i(-X_i)] \leq \mathbb{E}[\ell_i(-Z_i)]. \tag{3.2}$$

(2) **Vector loss function:** Let $\ell : \mathbb{R}^m \to \mathbb{R}^m \cup \{+\infty\}$ be the *vector loss function* defined by

$$\ell(x) = \begin{cases} (\ell_1(x_1), \ldots, \ell_m(x_m))^{\mathsf{T}}, & \text{if } x \in \bigtimes_{i=1}^m \operatorname{dom} \ell_i, \\ +\infty, & \text{else}, \end{cases} \tag{3.3}$$

for $x \in \mathbb{R}^m$. Similarly, the expected loss vector corresponding to a random position $X \in L_m^\infty$ is $\mathbb{E}[\ell(-X)] := (\mathbb{E}[\ell_1(-X_1)], \ldots, \mathbb{E}[\ell_m(-X_m)])^{\mathsf{T}}$ if $\mathbb{P}\{-X_i \in \operatorname{dom} \ell_i\} = 1$ for each $i \in \{1, \ldots, m\}$, and $\mathbb{E}[\ell(-X)] := +\infty$ otherwise.

(3) **Comparison rule:** Let $C \subseteq \mathbb{R}^m$ be a closed convex set such that $C + \mathbb{R}_+^m \subseteq C$ and $0 \in \mathbb{R}^m$ is a boundary point of C. Expected loss vectors will be compared according to the relation $\leq_C$ on $\mathbb{R}^m$ defined by

$$x \leq_C z \Leftrightarrow z \in x + C. \tag{3.4}$$

As $\mathbb{R}_+^m \subseteq C$, the relation $\leq_C$ provides a definition for a "smaller" expected loss vector by generalizing the componentwise comparison of expected loss vectors with $\leq_{\mathbb{R}_+^m}$. Some examples of the set C are discussed in Example 3.1 below.

Finally, the incomplete preference relation $\succeq$ of the investor on L_m^∞ is assumed to have the following numerical representation: For $X, Z \in L_m^\infty$,

$$X \succeq Z \Leftrightarrow \mathbb{E}[\ell(-X)] \leq_C \mathbb{E}[\ell(-Z)]. \tag{3.5}$$

Remark 3.1. In (3.4) and (3.5), the element $+\infty$ is added to $\mathbb{R}^m$ as a top element with respect to $\leq_C$, that is, $z \leq_C +\infty$ for every $z \in \mathbb{R}^m \cup \{+\infty\}$. The addition on $\mathbb{R}^m$ is extended to $\mathbb{R}^m \cup \{+\infty\}$ by $z + (+\infty) = (+\infty) + z = +\infty$ for every $z \in \mathbb{R}^m \cup \{+\infty\}$.

Remark 3.2. Note that $\leq_C$ (and hence $\succeq$) is reflexive (since $0 \in C$), transitive if $C + C \subseteq C$ and antisymmetric if $C \cap (-C) = \{0\}$ (C is "pointed."). In particular, if C is a pointed convex cone, then $\leq_C$ is a partial order which is compatible with the linear structure on $\mathbb{R}^m$.

Remark 3.3. It is easy to check that $\succeq$ respects the complete preferences $\succeq_1, \ldots, \succeq_m$ on individual assets described in (1). In other words, for every $i \in \{1, \ldots, m\}$, $X \in L_m^\infty$, and $Z_i \in L^\infty$,

$$X_i \succeq_i Z_i \Rightarrow X \succeq (X_1, \ldots, X_{i-1}, Z_i, X_{i+1}, \ldots, X_m). \tag{3.6}$$

This is thanks to $\mathbb{R}_+^m \subseteq C$.

Remark 3.4. The choice of the componentwise structure of the vector loss function in (3.3) is justified by the following reasons.

(1) For each asset i, one could consider a more general loss function ℓ_i that depends on the vector $x \in \mathbb{R}^m$ but not only on the component x_i. However, the interconnectedness of the components of a portfolio $x = X(\omega)$ at time t will be modeled

in Sec. 5 by the prevailing exchange rates $\mathcal{C}_t(\omega)$ and trading constraints $\mathcal{D}_t(\omega)$, and thus, will be included in the market risk measure.

(2) One could also consider vector loss functions $\ell : \mathbb{R}^d \to \mathbb{R}^m$ with $d > m$. The dimension reduction, which is motivated by allowing only m of the d assets to be used as eligible assets for risk compensation, is modeled by forcing liquidation into L^∞_m in Definition 5.2 of the market risk measure. This includes the case where a large number of assets d are denoted in a few ($m < d$) currencies, the currencies are used as eligible assets, and the loss functions are just defined for each of the m currencies (but not for each asset individually).

(3) On the other hand, the use of a vector loss function in the present paper is already more general than working under the assumption that there is a complete risk preference for multivariate positions (as in Burgert & Rüschendorf 2006) which even has a von Neumann–Morgenstern representation given by a real-valued loss function on $\mathbb{R}^m$ as in Campi & Owen (2011) (see Example 3.1(2)).

Example 3.1. We consider the following examples of comparison rules for different choices of C.

(1) If $C = \mathbb{R}^m_+$, then $\leq_C = \leq$ corresponds to the componentwise ordering of the expected loss vectors. In this case, we simply have $\succeq = \succeq_1 \times \cdots \times \succeq_m$.

(2) If C is a halfspace of the form $C = \{x \in \mathbb{R}^m \mid w^\mathsf{T} x \geq 0\}$ for some $w \in \mathbb{R}^m_+ \backslash \{0\}$, then

$$X \succeq Z \Leftrightarrow \mathbb{E}[L(-X)] \leq \mathbb{E}[L(-Z)], \tag{3.7}$$

where $x \mapsto L(x) := \sum_{i=1}^m w_i \ell_i(x_i)$ is a multivariate real-valued loss function as in Campi & Owen (2011, Example 2.10). In this case, $\succeq$ is a complete preference relation.

(3) If C is a polyhedral convex set of the form $C = \{x \in \mathbb{R}^m \mid Ax \geq b\}$ for some $A \in \mathbb{R}^{n \times m}_+$, $b \in \mathbb{R}^n$, $n \geq 1$ (with $b_j = 0$ for some $j \in \{1, \ldots, n\}$), then

$$X \succeq Z \Leftrightarrow \mathbb{E}[A(\ell(-X) - \ell(-Z))] \leq_{\mathbb{R}^n_+} b \tag{3.8}$$

which is a system of linear inequalities.

3.2. *The shortfall risk measure and its set optimization formulation*

Based on the incomplete preference relation $\succeq$ described in Sec. 3.1, we define the shortfall risk measure next. To that end, for each $i \in \{1, \ldots, m\}$, let $z^0_i \in \mathbb{R}$ such that $x^0_i := \ell_i(-z^0_i) \in \operatorname{int} \ell_i(\mathbb{R})$. The point $z^0 = (z^0_1, \ldots, z^0_m)^\mathsf{T}$ will be used as a *deterministic benchmark* for multivariate random positions and $x^0 = (x^0_1, \ldots, x^0_m)^\mathsf{T}$ is the corresponding *threshold value* for expected losses. Throughout, we assume that $x^0 = 0$. This is without loss of generality as otherwise one can shift the loss

function and work with $x \mapsto \tilde{\ell}(x) = \ell(x) - x^0$. Recalling (3.4) and (3.5), note that

$$X \succeq z^0 \Leftrightarrow \mathbb{E}[\ell(-X)] \leq_C 0 \Leftrightarrow 0 \in \mathbb{E}[\ell(-X)] + C \Leftrightarrow \mathbb{E}[\ell(-X)] \in -C. \quad (3.9)$$

The shortfall risk of a multivariate position $X \in L_m^\infty$ is defined as the set of all deterministic portfolio vectors $z \in \mathbb{R}^m$ that make $X + z$ preferable to the benchmark z^0.

Definition 3.1. The function $R_\ell : L_m^\infty \to \mathcal{P}(\mathbb{R}^m)$ defined by

$$R_\ell(X) = \{z \in \mathbb{R}^m \mid X + z \succeq z^0\} = \{z \in \mathbb{R}^m \mid \mathbb{E}[\ell(-X - z)] \in -C\} \quad (3.10)$$

is called the *shortfall risk measure* (on L_m^∞ with comparison rule C).

In other words, the shortfall risk of $X \in L_m^\infty$ is the set of all vectors $z \in \mathbb{R}^m$ for which $X + z$ has a "small enough" expected loss vector.

Proposition 3.1. *The shortfall risk measure R_ℓ satisfies the following properties*:

(1) ***Monotonicity*:** *$Z \geq X$ implies $R_\ell(Z) \supseteq R_\ell(X)$ for every $X, Z \in L_m^\infty$.*
(2) ***Translativity*:** *$R_\ell(X + z) = R_\ell(X) - z$ for every $X \in L_m^\infty$, $z \in \mathbb{R}^m$.*
(3) ***Finiteness at* 0:** *$R_\ell(0) \notin \{\emptyset, \mathbb{R}^m\}$.*
(4) ***Convexity*:** *$R_\ell(\lambda X + (1-\lambda)Z) \supseteq \lambda R_\ell(X) + (1-\lambda)R_\ell(Z)$ for every $X, Z \in L_m^\infty$, $\lambda \in (0,1)$.*
(5) ***Weak*-closedness*:** *The set* $\operatorname{graph} R_\ell := \{(X, z) \in L_m^\infty \times \mathbb{R}^m \mid z \in R_\ell(X)\}$ *is closed with respect to the product of the weak* topology $\sigma(L_m^\infty, L_m^1)$ and the usual topology on $\mathbb{R}^m$.*

Remark 3.5. The properties in Proposition 3.1 make R_ℓ a sensible measure of risk for multivariate positions in the sense that every portfolio $z \in R_\ell(X)$ can compensate for the risk of $X \in L_m^\infty$. Let us comment on the financial interpretations of these properties. Monotonicity ensures that a larger position (with respect to componentwise ordering) is *less* risky, that is, it has a *larger* set of risk compensating portfolios. Translativity is the requirement that a deterministic increment to a position reduces each of its risk compensating portfolios by the same amount. Finiteness at 0 guarantees that the risk of the zero position can be compensated by at least one but not all eligible portfolios. With the former two properties, it even guarantees *finiteness everywhere* in the sense that $R_\ell(X) \notin \{\emptyset, \mathbb{R}^m\}$ for every $X \in L_m^\infty$. Convexity can be interpreted as the reduction of risk by diversification. Finally, weak*-closedness is the set-valued version of the weak*-lower semicontinuity property (as in Föllmer & Schied 2011) for scalar risk measures.

Set-valued functions satisfying the properties in Proposition 3.1 are called (*set-valued*) (*weak**-) *closed convex risk measures* and are studied in Hamel & Heyde (2010), Hamel *et al.* (2011). An immediate consequence of these properties is that

the values of a closed convex risk measure belong to the collection

$$\mathcal{G}_m := \mathcal{G}(\mathbb{R}^m, \mathbb{R}^m_+) := \{A \subseteq \mathbb{R}^m \mid A = \operatorname{cl}\operatorname{co}(A + \mathbb{R}^m_+)\}, \tag{3.11}$$

where cl and co denote the closure and convex hull operators, respectively. It turns out that $\mathcal{G}_m$ is a convenient image space[a] to study set optimization, see Hamel (2009). In particular, it is an order complete lattice when equipped with the usual superset relation $\supseteq$. We have the following infimum and supremum formulae for every nonempty subset $\mathcal{A}$ of $\mathcal{G}_m$:

$$\inf_{(\mathcal{G}_m,\supseteq)} \mathcal{A} = \operatorname{cl}\operatorname{co} \bigcup_{A\in\mathcal{A}} A, \qquad \sup_{(\mathcal{G}_m,\supseteq)} \mathcal{A} = \bigcap_{A\in\mathcal{A}} A. \tag{3.12}$$

The infimum formula is motivated by the fact that the union of closed (convex) sets is not necessarily closed (convex). We also use the conventions $\inf_{(\mathcal{G}_m,\supseteq)} \emptyset = \emptyset$ and $\sup_{(\mathcal{G}_m,\supseteq)} \emptyset = \mathbb{R}^m$.

Note that $C \in \mathcal{G}_m$ with 0 being a boundary point of it. If $C = \mathbb{R}^m_+$, then the shortfall risk measure R_ℓ becomes a trivial generalization of the scalar shortfall risk measure (see Definition 2.2) in the sense that

$$R_\ell(X) = (\rho_{\ell_1}(X_1), \ldots, \rho_{\ell_m}(X_m))^{\mathsf{T}} + \mathbb{R}^m_+, \tag{3.13}$$

for every $X \in L^\infty_m$. In general, such an explicit representation of R_ℓ may not exist. However, given $X \in L^\infty_m$, one may write

$$R_\ell(X) = \inf_{(\mathcal{G}_m,\supseteq)} \{z + \mathbb{R}^m_+ \mid 0 \in \mathbb{E}[\ell(-X - z)] + C, z \in \mathbb{R}^m\}, \tag{3.14}$$

that is, $R_\ell(X)$ is the optimal value of the set minimization problem

$$\text{minimize } \Phi(z) \quad \text{subject to } 0 \in \Psi(z), \ \ z \in \mathbb{R}^m, \tag{3.15}$$

where $\Phi : \mathbb{R}^m \to \mathcal{G}_m$ and $\Psi : \mathbb{R}^m \to \mathcal{G}_m$ are the (set-valued) objective function and constraint function, respectively, defined by

$$\Phi(z) = z + \mathbb{R}^m_+, \quad \Psi(z) = \mathbb{E}[\ell(-X - z)] + C. \tag{3.16}$$

Here, it is understood that $\Psi(z) = \emptyset$ whenever $\mathbb{E}[\ell(-X - z)] = +\infty$.

A Lagrange duality theory for problems of the form (3.15) is developed in Hamel & Löhne (2014). Using this theory, we will compute the Lagrangian dual problem for $R_\ell(X)$. It will turn out that, after a change of variables, the dual objective function gives rise to another class of set-valued convex risk measures, called *divergence risk measures*. We introduce these risk measures separately in Sec. 3.3 first, and the duality results are deferred to Sec. 3.4.

[a]The phrase "image space" for a set-valued function refers to the set (subset of a power set) where the function maps into. This set is not a linear space in general. In particular, $\mathcal{G}_m$ is a conlinear space in the sense of Hamel (2009). It is closed under the closure of the Minkowski addition, and it is closed under multiplication by nonnegative scalars (with the convention $0\emptyset = \mathbb{R}^m_+$).

3.3. *Divergence risk measures*

In this section, we introduce divergence risk measures as decision-making problems of the investor about the level of consumption of a multivariate random position. The relationship between shortfall and divergence risk measures will be formulated in Sec. 3.4.

Suppose that the investor with random portfolio $X \in L_m^\infty$ wants to choose a deterministic portfolio $z \in \mathbb{R}^m$ to be received at initial time. Hence, she will hold $X - z$ at terminal time. She has two competing objectives:

(1) **Maximizing consumption:** The investor wants to maximize her immediate consumption z, or equivalently, minimize $-z$. The optimization is simply with respect to the componentwise ordering of portfolio vectors.
(2) **Minimizing loss:** The investor wants to minimize the expected loss $\mathbb{E}[\ell(-X + z)]$ of the remaining random position $X - z$. In this case, the expected loss vectors are compared with respect to the set C as in (3) of Sec. 3.1.

These two objectives can be summarized as the following set minimization problem where the objective function maps into $\mathcal{G}_{2m}$

$$\text{minimize} \begin{pmatrix} -z + \mathbb{R}_+^m \\ \mathbb{E}[\ell(-X + z)] + C \end{pmatrix} \quad \text{subject to } z \in \mathbb{R}^m. \tag{3.17}$$

Here and in Definition 3.2, the value of the objective function is understood to be $\emptyset$ if $\mathbb{E}[\ell(-X - z)] \notin \mathbb{R}^m$. On the other hand, the investor combines these competing objectives by means of a *relative weight vector* $r \in \mathbb{R}_+^m$: For each asset $i \in \{1, \ldots, m\}$, r_i is the relative weight of the expected loss $\mathbb{E}[\ell_i(-X_i + z_i)]$ with respect to $-z_i$. As a result, she solves the "partially scalarized" problem

$$\text{minimize} \quad -z + \mathbb{R}_+^m + r \cdot (\mathbb{E}[\ell(-X + z)] + C) \quad \text{subject to } z \in \mathbb{R}^m. \tag{3.18}$$

The optimal value of this problem is defined as the divergence risk of X.

Definition 3.2. Let $r \in \mathbb{R}_+^m$. The function $D_{\ell,r} : L_m^\infty \to \mathcal{G}_m$ defined by

$$D_{\ell,r}(X) = \inf_{(\mathcal{G}_m, \supseteq)} \{-z + r \cdot (\mathbb{E}[\ell(-X + z)] + C) \,|\, z \in \mathbb{R}^m\}, \tag{3.19}$$

is called the *divergence risk measure* with relative weight vector r.

Apparently, for some values of $r \in \mathbb{R}_+^m$, the optimization problem is unbounded and one has $D_{\ell,r}(X) = \mathbb{R}^m$ for every $X \in L_m^\infty$. In Proposition 3.2, we will characterize the set of all values of r for which $D_{\ell,r}$ has finite values and indeed is a closed convex risk measure, that is, it satisfies the five properties in Proposition 3.1.

Remark 3.6. In the one-dimensional (1D) case $m = 1$, one has $D_{\ell,r}(X) = \delta_{\ell,r}(X) + \mathbb{R}_+$, where $\delta_{\ell,r}(X) = \inf_{z \in \mathbb{R}: \mathbb{E}[\ell(-X+z)] < +\infty}(-z + r\mathbb{E}[\ell(-X + z)])$ is the divergence risk measure as in Definition 2.5. In the literature (see Ben Tal & Teboulle 1986, 2007), only the case $r = 1$ is considered in the definition

of divergence risk measure (or *optimized certainty equivalent* for $-\delta_{\ell,1}$). In Ben Tal & Teboulle (2007), the general case $r > 0$ is simply treated with a scaled loss function $r\ell$ since $\delta_{\ell,r} = \delta_{r\ell,1}$. (This simplification is not possible in the multi-dimensional case $m > 1$ as $r \in \mathbb{R}^m_+$ is multiplied by the set $\mathbb{E}[\ell(-X + z)] + C$ but not only the vector $\mathbb{E}[\ell(-X + z)]$.)

However, in our treatment, $\delta_{\ell,r}$ is interpreted as a weighted sum scalarization of a bi-objective optimization problem and this problem is, in turn, characterized by the whole family $(\delta_{\ell,r})_{r\in\mathbb{R}_+}$ of divergence risk measures. This interpretation is an original contribution of the present paper to the best of our knowledge.

3.4. *The Lagrange dual formulation of the shortfall risk measure*

This section formulates one of the main results of the paper, Theorem 3.1. The shortfall risk measure can be written as the intersection, that is, the set-valued supremum (see (3.12)), of divergence risk measures indexed by their relative weight vectors.

The result is derived in Sec. 6.4 using the recent Lagrange duality in Hamel & Löhne (2014) applied to the shortfall risk measure as the primal problem. The Lagrange duality is reviewed in Sec. 6.3. The result of its direct application to the shortfall risk measure is stated in Lemma 6.1, followed by a change of variables provided in Lemma 6.2. This additional latter step is essential in obtaining divergence risk measures in the (reformulated) dual problem.

Recall that the conjugate function of ℓ_i is denoted by g_i, which is a divergence function in the sense of Definition 2.3. The *vector divergence function* $g : \mathbb{R}^m \to \mathbb{R}^m \cup \{+\infty\}$ is defined by

$$g(y) = \begin{cases} (g_1(y_1), \dots, g_m(y_m))^\mathsf{T}, & \text{if } y \in \operatorname{dom} g := \bigtimes_{i=1}^m \operatorname{dom} g_i, \\ +\infty, & \text{else.} \end{cases} \tag{3.20}$$

In view of Remark 2.2, given $r \in \mathbb{R}^m_+$, we define

$$g_r(y) = ((g_1)_{r_1}(y_1), \dots, (g_m)_{r_m}(y_m))^\mathsf{T} \tag{3.21}$$

for each $y \in \mathbb{R}^m$ and set $\operatorname{dom} g_r = \bigtimes_{i=1}^m \operatorname{dom}(g_i)_{r_i}$. Moreover, for $r \in \mathbb{R}^m_{++}$, we write $\frac{1}{r} := (\frac{1}{r_1}, \dots, \frac{1}{r_m})^\mathsf{T}$.

Theorem 3.1. *For every* $X \in L^\infty_m$,

$$R_\ell(X) = \sup_{(\mathcal{G}_m,\supseteq)} \{D_{\ell,r}(X) \,|\, r \in \mathbb{R}^m_+, 1 \in \operatorname{dom} g_r\} = \bigcap_{r\in\mathbb{R}^m_+ : 1\in\operatorname{dom} g_r} D_{\ell,r}(X). \tag{3.22}$$

In particular, if $\operatorname{dom}\ell := \bigtimes_{i=1}^m \operatorname{dom}\ell_i = \mathbb{R}^m$, *then*

$$R_\ell(X) = \sup_{(\mathcal{G}_m,\supseteq)} \left\{ D_{\ell,r}(X) \,|\, r \in \mathbb{R}^m_{++}, \frac{1}{r} \in \operatorname{dom} g \right\} = \bigcap_{r\in\mathbb{R}^m_{++} : \frac{1}{r}\in\operatorname{dom} g} D_{\ell,r}(X). \tag{3.23}$$

Remark 3.7. Theorem 3.1 shows that the shortfall risk measure can be computed as a set-valued supremum over divergence risk measures. However, in general, there is no single $r \in \mathbb{R}^m_+$ with $1 \in \operatorname{dom} g_r$ which yields this supremum, that is, the supremum is not attained at a single argument. Instead, one could look for a set $\Gamma \subseteq \mathbb{R}^m_+$ with $1 \in \operatorname{dom} g_r$ for every $r \in \Gamma$ that satisfies the following two conditions: (3.22) holds with the intersection running through all $r \in \Gamma$, and each $D_{\ell,r}(X)$ with $r \in \Gamma$ is a maximal element of the set $\{D_{\ell,r}(X) \,|\, r \in \mathbb{R}^m_+, 1 \in \operatorname{dom} g_r\}$ with respect to $\supseteq$. This corresponds to the solution concept for set optimization problems introduced in Heyde & Löhne (2011) (see also Hamel & Löhne 2014, Definition 3.3) and will be discussed for the entropic risk measure in Sec. 4.1 together with the precise definition of a maximal element.

For every $r \in \mathbb{R}^m_+$, define a function $\delta_{\ell,r} : L^\infty_m \to \mathbb{R}^m \cup \{-\infty\}$ by

$$\delta_{\ell,r}(X) = (\delta_{\ell_1,r_1}(X_1), \ldots, \delta_{\ell_m,r_m}(X_m))^{\mathsf{T}} \tag{3.24}$$

whenever the right-hand side is in $\mathbb{R}^m$ and $\delta_{\ell,r}(X) = -\infty$ otherwise. Recall that δ_{ℓ_i,r_i} is given by

$$\delta_{\ell_i,r_i}(X_i) = \inf_{z_i \in \mathbb{R}} (z_i + r_i \mathbb{E}[\ell_i(-X_i - z_i)]). \tag{3.25}$$

If $r_i > 0$, and we have $\delta_{\ell_i,0}(X_i) = -\operatorname{ess\,inf} X_i - \sup \operatorname{dom} \ell_i$; see (2.3). If $1 \in \operatorname{dom}(g_i)_{r_i}$, then δ_{ℓ_i,r_i} is the scalar (ℓ_i, r_i)-divergence risk measure according to Definition 2.5.

As a byproduct of Theorem 3.1, we show that a divergence risk measure has a much simpler form in terms of scalar divergence risk measures.

Proposition 3.2. *Let $r \in \mathbb{R}^m_+$.*

(1) *If $1 \in \operatorname{dom} g_r$, then $D_{\ell,r}$ is a closed convex risk measure with the representation*

$$D_{\ell,r}(X) = \delta_{\ell,r}(X) + r \cdot C. \tag{3.26}$$

(2) *Otherwise, $D_{\ell,r}(X) = \mathbb{R}^m$ for every $X \in L^\infty_m$.*

In particular, if $\operatorname{dom} \ell = \mathbb{R}^m$, then $D_{\ell,r}$ is a closed convex risk measure if and only if $r \in \mathbb{R}^m_{++}$ with $\frac{1}{r} \in \operatorname{dom} g$.

Note that, in the representation in Proposition 3.2, the dependence on $X \in L^\infty_m$ is only through the vector part $\delta_{\ell,r}(X)$; however, the choice of the relative weight vector r still affects the distortion on the set C through $r \cdot C$.

Remark 3.8. Let us comment on the trade-off between using a shortfall risk measure and a divergence risk measure. According to the representation in Proposition 3.2(1), the divergence risk measure with relative weight vector $r \in \mathbb{R}^m_+$ (with $1 \in \operatorname{dom} g_r$) has the simple structure $D_{\ell,r}(X) = \delta_{\ell,r}(X) + r \cdot C$, where the dependence on X is solely on the vector $\delta_{\ell,r}(X)$ of scalar divergence risk measures. Hence, the computation of the divergence risk measure reduces to the computation of m

scalar risk measures. However, the shortfall risk measure does not possess such a simple representation as a constrained optimization problem. It is a (much) more conservative notion of set-valued risk as $R_\ell(X) \subseteq D_{\ell,r}(X)$. On the other hand, the divergence risk measure has the additional parameter r. For each asset, the investor has to specify how many units of her expected loss is comparable with one unit of the consumption at initial time.

3.5. *Dual representations*

In this section, we state representations of shortfall and divergence risk measures in terms of *vector probability measures* and *weight vectors*. Such representations of convex risk measures are called *dual representations*.

In Hamel & Heyde (2010) and Hamel *et al.* (2011), it is shown that a closed convex risk measure is indeed characterized by a halfspace-valued function that shows up in its dual representation. We recall this result first. To that end, let $\mathbb{Q} = (\mathbb{Q}_1, \ldots, \mathbb{Q}_m)^\mathsf{T}$ be an *m-dimensional vector probability measure* in the sense that $\mathbb{Q}_i$ is a probability measure on $(\Omega, \mathcal{F})$ for each $i \in \{1, \ldots, m\}$. We define $\mathbb{E}^{\mathbb{Q}}[X] = (\mathbb{E}^{\mathbb{Q}_1}[X_1], \ldots, \mathbb{E}^{\mathbb{Q}_m}[X_m])^\mathsf{T}$ for every $X \in L^0_m$ such that the components exist in $\mathbb{R}$. We denote by $\mathcal{M}_m(\mathbb{P})$ the set of all m-dimensional vector probability measures on $(\Omega, \mathcal{F})$ whose components are absolutely continuous with respect to $\mathbb{P}$. For $\mathbb{Q} \in \mathcal{M}_m(\mathbb{P})$, we set $\frac{d\mathbb{Q}}{d\mathbb{P}} = (\frac{d\mathbb{Q}_1}{d\mathbb{P}}, \ldots, \frac{d\mathbb{Q}_m}{d\mathbb{P}})^\mathsf{T}$, where, for each $i \in \{1, \ldots, m\}$, $\frac{d\mathbb{Q}_i}{d\mathbb{P}}$ denotes the Radon–Nikodym derivative of $\mathbb{Q}_i$ with respect to $\mathbb{P}$. For $w \in \mathbb{R}^m_+ \backslash \{0\}$, we define the halfspace

$$G(w) := \{z \in \mathbb{R}^m \mid w^\mathsf{T} z \geq 0\}. \tag{3.27}$$

Proposition 3.3 (Hamel *et al.* 2011, Theorem 4.2). *A function $R : L^\infty_m \to \mathcal{G}_m$ is a closed convex risk measure if and only if, for every $X \in L^\infty_m$,*

$$R(X) = \bigcap_{(\mathbb{Q},w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}^m_+ \backslash \{0\})} (-\alpha_R(\mathbb{Q}, w) + \mathbb{E}^{\mathbb{Q}}[-X]), \tag{3.28}$$

where $-\alpha_R : \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}^m_+ \backslash \{0\}) \to \mathcal{G}_m$ is the penalty function of R defined by

$$-\alpha_R(\mathbb{Q}, w) = \operatorname{cl} \bigcup_{X \in L^\infty_m} [R(X) + (\mathbb{E}^{\mathbb{Q}}[X] + G(w))], \tag{3.29}$$

for each $(\mathbb{Q}, w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}^m_+ \backslash \{0\})$.

As in the scalar case, the penalty function of a closed convex risk measure basically coincides with its Fenchel conjugate. In the set-valued case, the transformation from the set-valued conjugate with dual variables $L^1_m \times (\mathbb{R}^m_+ \backslash \{0\})$ to a penalty function with dual variables in $\mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}^m_+ \backslash \{0\})$ requires some extra care; this procedure is described in detail in Hamel & Heyde (2010), Hamel *et al.* (2011).

In Propositions 3.4 and 3.5, we present the penalty functions of divergence and shortfall risk measures, respectively. To that end, we define a divergence for a vector probability measure.

Definition 3.3. Let $r \in \mathbb{R}^m_+$ and $\mathbb{Q} \in \mathcal{M}_m(\mathbb{P})$. For each $i \in \{1, \ldots, m\}$, let

$$I_{g_i,r_i}(\mathbb{Q}_i \,|\, \mathbb{P}) := \begin{cases} r_i \mathbb{E}\left[g_i\left(\frac{1}{r_i} \frac{d\mathbb{Q}_i}{d\mathbb{P}} \right) \right], & \text{if } r_i > 0, \\ \sup \operatorname{dom} \ell_i, & \text{if } r_i = 0. \end{cases} \tag{3.30}$$

The element $I_{g,r}(\mathbb{Q} \,|\, \mathbb{P}) \in \mathbb{R}^m \cup \{+\infty\}$ defined by

$$I_{g,r}(\mathbb{Q} \,|\, \mathbb{P}) := (I_{g_1,r_1}(\mathbb{Q}_1 \,|\, \mathbb{P}), \ldots, I_{g_m,r_m}(\mathbb{Q}_m \,|\, \mathbb{P}))^{\mathsf{T}} \tag{3.31}$$

if $I_{g_i,r_i}(\mathbb{Q}_1 \,|\, \mathbb{P}) \in \mathbb{R}$ for each $i \in \{1, \ldots, m\}$, and by $I_{g,r}(\mathbb{Q} \,|\, \mathbb{P}) := +\infty$ otherwise is called the *vector* (g, r)*-divergence* of $\mathbb{Q}$ with respect to $\mathbb{P}$.

Note that $I_{g_i,r_i}(\mathbb{Q}_i \,|\, \mathbb{P})$ is the (scalar) (g_i, r_i)-divergence of $\mathbb{Q}_i$ with respect to $\mathbb{P}$, see Definition 2.4.

Proposition 3.4. *Let $r \in \mathbb{R}^m_+$ with $1 \in \operatorname{dom} g_r$. The penalty function of the divergence risk measure $D_{\ell,r}$ is given by*

$$-\alpha_{D_{\ell,r}}(\mathbb{Q}, w) = -I_{g,r}(\mathbb{Q} \,|\, \mathbb{P}) + r \cdot C + G(w) \tag{3.32}$$

for each $(\mathbb{Q}, w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}^m_+ \backslash \{0\})$ with the convention $-\alpha_{D_{\ell,r}}(\mathbb{Q}, w) = \mathbb{R}^m$ if $I_{g,r}(\mathbb{Q} \,|\, \mathbb{P}) = +\infty$.

Proposition 3.5. *The penalty function of the shortfall risk measure R_ℓ is given by*

$$\begin{aligned} -\alpha_{R_\ell}(\mathbb{Q}, w) &= \left\{ z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq \sup_{r \in \mathbb{R}^m_+} \left(-w^{\mathsf{T}} I_{g,r}(\mathbb{Q} \,|\, \mathbb{P}) + \inf_{x \in C} w^{\mathsf{T}}(r \cdot x) \right) \right\} \\ &= \bigcap_{r \in \mathbb{R}^m_+ : 1 \in \operatorname{dom} g_r} -\alpha_{D_{\ell,r}}(\mathbb{Q}, w), \end{aligned} \tag{3.33}$$

for each $(\mathbb{Q}, w) \in \mathcal{M}_m(\mathbb{P}) \times \mathbb{R}^m_{++}$ with the convention $-\alpha_{D_{\ell,r}}(\mathbb{Q}, w) = \mathbb{R}^m$ if $I_{g,r}(\mathbb{Q} \,|\, \mathbb{P}) = +\infty$. In particular, if $\operatorname{dom} \ell = \mathbb{R}^m$, then

$$-\alpha_{R_\ell}(\mathbb{Q}, w) = \bigcap_{r \in \mathbb{R}^m_{++} : \frac{1}{r} \in \operatorname{dom} g} -\alpha_{D_{\ell,r}}(\mathbb{Q}, w). \tag{3.34}$$

4. Examples

4.1. *Set-valued entropic risk measures*

In this section, we assume that the vector loss function ℓ of Sec. 3 is the *vector exponential loss function* with constant risk aversion vector $\beta \in \mathbb{R}^m_{++}$, that is, for

each $i \in \{1, \dots, m\}$ and $x \in \mathbb{R}$,

$$\ell_i(x) = \frac{e^{\beta_i x} - 1}{\beta_i}, \tag{4.1}$$

which satisfies the conditions in Definition 2.1. The corresponding vector divergence function g is given by

$$g_i(y) = \frac{y}{\beta_i} \log y - \frac{y}{\beta_i} + \frac{1}{\beta_i}, \tag{4.2}$$

for each $i \in \{1, \dots, m\}$ and $y \in \mathbb{R}$. Here and elsewhere, we make the convention $\log y = -\infty$ for every $y \leq 0$.

For convenience, we sometimes use the notation $[x_i]_{i=1}^m$ for $x = (x_1, \dots, x_m)^\mathsf{T} \in \mathbb{R}^m$. Let us also define $x^{-1} := (x_1^{-1}, \dots, x_m^{-1})^\mathsf{T}$ and $\log x := (\log x_1, \dots, \log x_m)^\mathsf{T}$ for $x \in \mathbb{R}^m_{++}$, and $\log[A] := \{\log x \,|\, x \in A\}$ for $A \subseteq \mathbb{R}^m_{++}$. We will also use $1 := (1, \dots, 1)^\mathsf{T}$ as an element of $\mathbb{R}^m$.

Note that $\operatorname{int} \ell(\operatorname{dom} \ell) = \ell(\operatorname{dom} \ell) = \ell(\mathbb{R}) = -\beta^{-1} + \mathbb{R}^m_{++}$ so that $0 \in \operatorname{int} \ell(\operatorname{dom} \ell)$. Let $C \in \mathcal{G}_m$ with $0 \in \mathbb{R}^m$ being a boundary point of C. We call the corresponding shortfall risk measure $R^{\mathrm{ent}} := R_\ell$ the *entropic risk measure*. The next proposition shows that R^{ent} has the simple form of "a vector-valued function plus a fixed set", which is, in general, not the case for an arbitrary loss function. Note that the functional ρ^{ent} in Proposition 4.1 is the vector of scalar entropic risk measures.

Proposition 4.1. *For every $X \in L^\infty_m$,*

$$R^{\mathrm{ent}}(X) = \rho^{\mathrm{ent}}(X) + C^{\mathrm{ent}}, \tag{4.3}$$

where

$$\begin{aligned} \rho^{\mathrm{ent}}(X) &:= \left[\frac{1}{\beta_i} \log \mathbb{E}[e^{-\beta_i X_i}] \right]_{i=1}^m, \\ C^{\mathrm{ent}} &:= -\beta^{-1} \cdot \log[(1 - \beta \cdot C) \cap \mathbb{R}^m_{++}]. \end{aligned} \tag{4.4}$$

Note that the set $\operatorname{dom} g$ defined in (3.20) becomes $\mathbb{R}^m_+$. Since $\operatorname{dom} \ell = \mathbb{R}^m$, by Proposition 3.2, $D^{\mathrm{ent}}_r := D_{\ell,r}$ is a closed convex risk measure (divergence risk measure) if $r \in \mathbb{R}^m_{++}$ and $D_{\ell,r}(X) = \mathbb{R}^m$ for every $X \in L^\infty_m$ if $r \in \mathbb{R}^m_+ \backslash \mathbb{R}^m_{++}$.

Proposition 4.2. *For every $r \in \mathbb{R}^m_{++}$ and $X \in L^\infty_m$,*

$$D^{\mathrm{ent}}_r(X) = \rho^{\mathrm{ent}}(X) + \beta^{-1} \cdot (1 - r + \log r) + r \cdot C, \tag{4.5}$$

where $\rho^{\mathrm{ent}}(X)$ is defined by (4.4).

Recall from (3.23) that $R^{\mathrm{ent}}(\cdot)$ is the supremum of all $D^{\mathrm{ent}}_r(\cdot)$ with $r \in \mathbb{R}^m_{++}$, that is, for $X \in L^\infty_m$,

$$R^{\mathrm{ent}}(X) = \sup_{r \in \mathbb{R}^m_{++}} D^{\mathrm{ent}}_r(X) = \bigcap_{r \in \mathbb{R}^m_{++}} D^{\mathrm{ent}}_r(X). \tag{4.6}$$

If $m = 1$, then the only choice for C is $\mathbb{R}_+$. In this case, one can check that, for $X \in L^\infty$,

$$R^{\text{ent}}(X) = D_1^{\text{ent}}(X) = \rho^{\text{ent}}(X) + \mathbb{R}_+. \tag{4.7}$$

In other words, the supremum in (4.6) is *attained* at $r = 1$. In general, when $m \geq 2$, we may not be able to find some $\bar{r} \in \mathbb{R}_{++}$ for which $R^{\text{ent}}(X) = D_{\bar{r}}^{\text{ent}}(X)$. Instead, we will compute a *solution* to this set maximization problem in the sense of Hamel & Löhne (2014, Definition 3.3), that is, we will find a set $\Gamma \subseteq \mathbb{R}^m_{++}$ such that

(1) $R^{\text{ent}}(X) = \bigcap_{r\in\Gamma} D_r^{\text{ent}}(X)$,
(2) for each $\bar{r} \in \Gamma$, $D_{\bar{r}}^{\text{ent}}(X)$ is a *maximal element* of the collection $\{D_r^{\text{ent}}(X) \,|\, r \in \mathbb{R}^m_{++}\}$ in the following sense:

$$\forall\, r \in \mathbb{R}^m_{++} : \quad D_r^{\text{ent}}(X) \subseteq D_{\bar{r}}^{\text{ent}}(X) \;\Rightarrow\; r = \bar{r}. \tag{4.8}$$

Moreover, the set Γ will be independent of the choice of X. To that end, by Proposition 4.2, we can rewrite $D_r^{\text{ent}}(X)$ as

$$D_r^{\text{ent}}(X) = \rho^{\text{ent}}(X) + \bigcap_{w\in\mathbb{R}^m_+\setminus\{0\}} \{z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq -(f_w(r) + h_w(r))\}, \tag{4.9}$$

where, for $w \in \mathbb{R}^m_+\setminus\{0\}$, $r \in \mathbb{R}^m_{++}$,

$$\begin{aligned} f_w(r) &:= w^{\mathsf{T}}(-\beta^{-1} \cdot (1 - r + \log r)), \\ h_w(r) &:= -\inf_{x\in C} w^{\mathsf{T}}(r \cdot x) = \sup_{x\in -C} w^{\mathsf{T}}(r \cdot x). \end{aligned} \tag{4.10}$$

Lemma 4.1. *Let $w \in \mathbb{R}^m_+\setminus\{0\}$. The function $f_w + h_w$ on $\mathbb{R}^m_{++}$ is either identically $+\infty$ or else it attains its infimum at a unique point $r^w \in \mathbb{R}^m_{++}$ which is determined by the following property: r^w is the only vector $r \in \mathbb{R}^m_{++}$ for which C is supported at the point $\beta^{-1} \cdot (1 - r^{-1})$ by the hyperplane with normal direction $r \cdot w$.*

Proposition 4.3. *Using the notation in Lemma* 4.1, *the set*

$$\Gamma := \{r^w \,|\, w \in \mathbb{R}^m_+\setminus\{0\}, f_w + h_w \text{ is proper}\} \tag{4.11}$$

is a solution to the maximization problem in (4.6) *for every $X \in L^\infty_m$.*

Finally, we compute the penalty function of R^{ent} in terms of the *vector relative entropies*

$$H(\mathbb{Q} \,|\, \mathbb{P}) := \left[\mathbb{E}^{\mathbb{Q}_i}\left[\log \frac{d\mathbb{Q}_i}{d\mathbb{P}}\right]\right]_{i=1}^m \tag{4.12}$$

of vector probability measures $\mathbb{Q} \in \mathcal{M}_m(\mathbb{P})$. Thus, the penalty function for the entropic risk measure is of the form "negative vector relative entropy plus a nonhomogeneous halfspace" (except for the trivial case).

Proposition 4.4. *For every* $(\mathbb{Q}, w) \in \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}^m_+ \backslash \{0\})$, *we have* $-\alpha_{R^{\text{ent}}}(\mathbb{Q}, w) = \mathbb{R}^m$ *if* $h_w(r) = +\infty$ *for every* $r \in \mathbb{R}^m_{++}$, *and*

$$-\alpha_{R^{\text{ent}}}(\mathbb{Q}, w) = -\beta^{-1} \cdot H(\mathbb{Q} \,|\, \mathbb{P}) - \beta^{-1} \cdot \log[(1 - \beta \cdot C) \cap \mathbb{R}^m_{++}] + G(w) \tag{4.13}$$

if h_w *is a proper function.*

4.2. *Set-valued average values at risks*

In this section, we assume that the vector loss function ℓ of Sec. 3 is the (*vector*) *scaled positive part function* with scaling vector $\alpha \in (0, 1]^m$, that is, for each $i \in \{1, \ldots, m\}$ and $x \in \mathbb{R}$,

$$\ell_i(x) = \frac{x^+}{\alpha_i} \tag{4.14}$$

which satisfies the conditions in Definition 2.1. The corresponding vector divergence function g is given by

$$g_i(y) = \begin{cases} 0 & \text{if } y \in \left[0, \dfrac{1}{\alpha_i}\right], \\ +\infty & \text{else,} \end{cases} \tag{4.15}$$

for each $i \in \{1, \ldots, m\}$ and $y \in \mathbb{R}$.

Note that $0 \notin \operatorname{int} \ell(\operatorname{dom} \ell) = \mathbb{R}^m_{++}$ in this example. Hence, let us fix $x^0 \in \mathbb{R}^m_{++}$ and $C \in \mathcal{G}_m$ with 0 being a boundary point of C. We will apply the definitions and results of Sec. 3 to the shifted loss function $\tilde{\ell}(x) = \ell(x) - x^0$. The corresponding shortfall risk measure is given by

$$R_{\tilde{\ell}}(X) = \{z \in \mathbb{R}^m \,|\, \mathbb{E}[(z - X)^+] \in \alpha \cdot (x^0 - C)\}, \tag{4.16}$$

where the positive part function is applied componentwise.

Note that the set $\operatorname{dom} g$ defined in (3.20) becomes $\bigtimes_{i=1}^m [0, \frac{1}{\alpha_i}]$. Since $\operatorname{dom} \tilde{\ell} = \mathbb{R}^m$, by Proposition 3.2, $D_{\tilde{\ell},r}$ is a closed convex risk measure (divergence risk measure) if $r \in \bigtimes_{i=1}^m [\alpha_i, +\infty)$ and $D_{\tilde{\ell},r}(X) = \mathbb{R}^m$ for every $X \in L^\infty_m$ if $r \in \mathbb{R}^m_+ \backslash \bigtimes_{i=1}^m [\alpha_i, +\infty)$. In the former case, the divergence risk measure with relative weight vector $r \in \bigtimes_{i=1}^m [\alpha_i, +\infty)$ is given by $D_{\tilde{\ell},r}(X) = \delta_{\tilde{\ell},r}(X) + r \cdot C$ for $X \in L^\infty_m$, where, for each $i \in \{1, \ldots, m\}$,

$$\delta_{\tilde{\ell}_i, r_i}(X_i) = \inf_{z_i \in \mathbb{R}} \left(-z_i + \frac{r_i}{\alpha_i} \mathbb{E}[(z_i - X_i)^+] \right) - r_i x^0_i. \tag{4.17}$$

When $r = (1, \ldots, 1)^\mathsf{T}$ and $C = \mathbb{R}^m_+$, we obtain the set-valued average value at risk in the sense of Hamel *et al.* (2013, Definition 2.1 for $M = \mathbb{R}^m$), which is given by

$$\text{AV@R}_\alpha(X) := D_{\tilde{\ell},1}(X) + x^0 = \left[\inf_{z_i \in \mathbb{R}} \left(-z_i + \frac{1}{\alpha_i} \mathbb{E}[(z_i - X_i)^+] \right) \right]_{i=1}^m + \mathbb{R}^m_+. \tag{4.18}$$

Hence, our framework offers the following generalization of the set-valued average value at risk as a convex risk measure:

$$\begin{aligned}\mathrm{AV@R}_{\alpha,r}(X) &:= D_{\tilde{\ell},r}(X) + r \cdot x^0 \\ &= \left[\inf_{z_i \in \mathbb{R}} \left(-z_i + \frac{r_i}{\alpha_i}\mathbb{E}[(z_i - X_i)^+]\right)\right]_{i=1}^{m} + r \cdot C. \qquad (4.19)\end{aligned}$$

As in the scalar case, this definition even works for $X \in L^1_m$.

5. Market Risk Measures

The purpose of this section is to propose a method to incorporate the frictions of the market into the quantification of risk. As the first step of the method, it is assumed that there is a "pure" risk measure R that represents the attitude of the investor towards the assets of the market. This could be one of the utility-based risk measures introduced in Sec. 3. Since the risk measure R does not take into account the frictions of the market, the second step consists of minimizing risk subject to the trading opportunities of the market. More precisely, we minimize (in the sense of set optimization) the value of R over the set of financial positions that can be reached with the given position by trading in the so-called *convex market model*. The result of the risk minimization, as a function of the given position, is called the *market risk measure* induced by R.

In the literature, minimization of scalar risk measures subject to trading constraints are considered in Barrieu & El Karoui (2008). In the multivariate case, market risk measures are introduced in Hamel *et al.* (2011) and Hamel *et al.* (2013) for the special case of a conical market model. Here, this notion is considered for an arbitrary convex risk measure with the more general convex market model of Pennanen & Penner (2010) and the possibility of trading constraints and liquidation into fewer assets. The market is described in Sec. 5.1. The dual representation result, Theorem 5.1 in Sec. 5.2, is one of the main contributions of this paper. Finally, in Sec. 5.3, we present sufficient conditions under which Theorem 5.1 can be applied to shortfall and divergence risk measures.

5.1. *The convex market model with trading constraints*

Consider a financial market with $d \in \{1, 2, \ldots\}$ assets. We assume that the market has convex transaction costs or nonlinear illiquidities in finite discrete time. Following Pennanen & Penner (2010), we use convex solvency regions to model such frictions. To that end, let $T \in \{1, 2, \ldots\}$, $\mathbb{T} = \{0, \ldots, T\}$, and $(\mathcal{F}_t)_{t\in\mathbb{T}}$ a filtration of $(\Omega, \mathcal{F}, \mathbb{P})$ augmented by the $\mathbb{P}$-null sets of $\mathcal{F}$. The number T denotes the time horizon, and $(\mathcal{F}_t)_{t\in\mathbb{T}}$ represents the evolution of information over time. We suppose that there is no information at time 0, that is, every $\mathcal{F}_0$-measurable function is deterministic $\mathbb{P}$-almost surely; and there is full information at time T, that is, $\mathcal{F}_T = \mathcal{F}$. For $p \in \{0, 1, +\infty\}$ and $t \in \mathbb{T}$, we denote by $L^p_d(\mathcal{F}_t)$ the linear subspace of all $\mathcal{F}_t$-measurable random variables in L^p_d.

Let $t \in \mathbb{T}$. By the $\mathcal{F}_t$-measurability of a set-valued function $D : \Omega \to \mathcal{P}(\mathbb{R}^d)$, it is meant that the graph $\{(\omega, y) \in \Omega \times \mathbb{R}^d \,|\, y \in D(\omega)\}$ is $\mathcal{F}_t \otimes \mathcal{B}(\mathbb{R}^d)$-measurable, where $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel σ-algebra on $\mathbb{R}^d$. For such function D, define the set $L^p_d(\mathcal{F}_t, D) := \{Y \in L^p_d(\mathcal{F}_t) \,|\, \mathbb{P}\{\omega \in \Omega \,|\, Y(\omega) \in D(\omega)\} = 1\}$ for $p \in \{0, 1, +\infty\}$.

We recall the convex market model of Pennanen & Penner (2010) next. For each $t \in \mathbb{T}$, let $\mathcal{C}_t : \Omega \to \mathcal{G}_d$ be an $\mathcal{F}_t$-measurable function such that $\mathbb{R}^d_+ \subseteq \mathcal{C}_t(\omega)$ and $-\mathbb{R}^d_+ \cap \mathcal{C}_t(\omega) = \{0\}$ for each $t \in \{0, \ldots, T\}$ and $\omega \in \Omega$. The set $\mathcal{C}_t$ is called the (*random*) *solvency region* at time t; see Astic & Touzi (2007), Pennanen & Penner (2010). It models the bid and ask prices as a function of the magnitude of a trade, for instance, as in Çetin *et al.* (2004), Çetin & Rogers (2007), Rogers & Singh (2010); and thus, directly relates to the shape of the order book. More precisely, $\mathcal{C}_t(\omega)$ is the set of all portfolios which can be exchanged into ones with nonnegative components at time t when the outcome is ω. Convex solvency regions allow for the modeling of temporary illiquidity effects in the sense that they cover nonlinear illiquidities; however, they assume that agents have no market power, and thus, their trades do not affect the costs of subsequent trades.

Example 5.1. An important special case is the conical market model introduced in Kabanov (1999). Suppose that $\mathcal{C}_t(\omega)$ is a (closed convex) cone for each $t \in \mathbb{T}$ and $\omega \in \Omega$. In this case, the transaction costs are proportional to the size of the orders.

From a financial point of view, it is possible to have additional constraints on the trading opportunities at intermediate times. For instance, trading may be allowed only up to a (possibly state- and time-dependent) threshold level for the assets (Example 5.2), or it may be the case that a certain linear combination of the trading units should not exceed a threshold level (Example 5.3). Such constraints are modeled via convex random sets. Given $t \in \{0, \ldots, T-1\}$, let $\mathcal{D}_t : \Omega \to \mathcal{P}(\mathbb{R}^d)$ be an $\mathcal{F}_t$-measurable function such that $\mathcal{D}_t(\omega)$ is a closed convex set and $0 \in \mathcal{C}_t(\omega) \cap \mathcal{D}_t(\omega)$ for every $\omega \in \Omega$. Note that $\mathcal{D}_t$ does not necessarily map into $\mathcal{G}_d$, and this is why we prefer to work with $\mathcal{C}_t \cap \mathcal{D}_t$ instead of replacing $\mathcal{C}_t$ by $\mathcal{C}_t \cap \mathcal{D}_t$. For convenience, let us also set $\mathcal{D}_T(\omega) = \mathbb{R}^d$ for every $\omega \in \Omega$.

Example 5.2. For each $t \in \{0, \ldots, T-1\}$, suppose that $\mathcal{D}_t = \bar{Y}_t - \mathbb{R}^d_+$, for some $\bar{Y}_t \in L^0_d(\mathcal{F}_t, \mathbb{R}^d_+)$. In this case, trading in asset $i \in \{1, \ldots, d\}$ at time $t \in \{0, \ldots, T-1\}$ may not exceed the level $(\bar{Y}_t)_i$.

Example 5.3. For each $t \in \{0, \ldots, T-1\}$, suppose that $\mathcal{D}_t = \{y \in \mathbb{R}^d \,|\, A_t^{\mathsf{T}} y \leq B_t\}$, for some $A_t \in L^0_d(\mathcal{F}_t, \mathbb{R}^d_+ \setminus \{0\})$ and $B_t \in L^0_1(\mathcal{F}_t, \mathbb{R}_+)$. In this case, trading in each asset is unlimited but the linear combination of the trading units with the weight vector A_t cannot exceed the level B_t.

The set of all financial positions that can be obtained by trading in the market starting with the zero position is

$$\mathcal{K} := -\sum_{t=0}^{T} L_d^\infty(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t). \tag{5.1}$$

Hence, an investor with a financial position $Y \in L_d^\infty$ can ideally reach any element of the set $Y + \mathcal{K}$ by trading in the market. However, it may be the case that the risk of the resulting position is evaluated only through a (small) selection of the d assets, in other words, trading has to be done in such a way that the only possibly nonzero components of the resulting position can be in some selected subset of the d assets. Without loss of generality, suppose that liquidation is made into the first $m \leq d$ of the assets. The idea of liquidation is made precise by the notion of liquidation function introduced in Definition 5.1. Let us introduce the linear operator $B : \mathbb{R}^m \to \mathbb{R}^d$ defined by

$$Bz = (z_1, \ldots, z_m, 0, \ldots, 0)^\mathsf{T}. \tag{5.2}$$

We will use the composition of B with random variables in L_m^0. Given $X \in L_m^0$, BX denotes the element in L_d^0 defined by $(BX)(\omega) = B(X(\omega))$ for $\omega \in \Omega$. The adjoint $B^* : \mathbb{R}^d \to \mathbb{R}^m$ of B is given by

$$B^* y = (y_1, \ldots, y_m)^\mathsf{T}. \tag{5.3}$$

Similarly, B^* can be composed with random variables in L_d^0. With a slight abuse of notation, we will also use B^* in the context of vector probability measures. Given $\mathbb{Q} \in \mathcal{M}_d(\mathbb{P})$, we define $B^*\mathbb{Q} = (\mathbb{Q}_1, \ldots, \mathbb{Q}_m)^\mathsf{T} \in \mathcal{M}_m(\mathbb{P})$.

Definition 5.1. The function $\Lambda_m : L_d^\infty \to \mathcal{P}(L_m^\infty)$ defined by

$$\Lambda_m(Y) = \{X \in L_m^\infty \,|\, BX \in Y + \mathcal{K}\} \tag{5.4}$$

is called the *liquidation function* associated with $\mathcal{K}$.

Hence, given $Y \in L_d^\infty$, the set $\Lambda_m(Y)$ consists of all possible resulting positions in $Y + \mathcal{K}$ that are already liquidated into the first m assets.

5.2. *Market risk measures and their dual representations*

Let us consider a closed convex risk measure $R : L_m^\infty \to \mathcal{G}_m$ which is used for risk evaluation after liquidating the resulting positions into the first m assets. As all the positions in $\Lambda_m(Y)$ are accessible to the investor with position $Y \in L_d^\infty$, the value of R is to be minimized over the set $\Lambda_m(Y)$ as the following definition suggests.

Definition 5.2. The function $R^{\mathrm{mar}} : L_d^\infty \to \mathcal{P}(\mathbb{R}^m)$ defined by

$$R^{\mathrm{mar}}(Y) := \inf_{(\mathcal{G}_m, \supseteq)} \{R(X) \,|\, X \in \Lambda_m(Y)\} = \mathrm{cl}\,\mathrm{co} \bigcup_{X \in \Lambda_m(Y)} R(X) \tag{5.5}$$

is called the *market risk measure* induced by R.

Remark 5.1. In the case of the conical market model described in Example 5.1, when $\mathcal{D}_t(\omega) = \mathbb{R}^d$ for each $\omega \in \Omega$ and $t \in \{0, \dots, T\}$, and no liquidation at $t = T$ is considered ($m = d$), Definition 5.2 recovers the notion of market-extension (with closed values) given in Hamel *et al.* (2013, Definition 2.8, Remark 2.9).

Recall that a closed convex risk measure $R : L^\infty_m \to \mathcal{G}_m$ is defined by the five properties in Proposition 3.1. For the market risk measure, these properties need to be rewritten with obvious changes as the function is now defined on L^∞_d. (For instance, the translativity of R^{mar} reads as $R^{\mathrm{mar}}(Y + Bz) = R^{\mathrm{mar}}(Y) - z$ for every $Y \in L^\infty_d$ and $z \in \mathbb{R}^m$.) The next proposition shows that the market risk measure is a closed convex risk measure except for a finiteness condition and weak*-closedness.

Proposition 5.1. *The market risk measure R^{mar} is monotone, translative and convex, and it satisfies $R^{\mathrm{mar}}(0) \neq \emptyset$. In addition, the convex hull operator can be dropped from Definition* 5.2*, that is, for $Y \in L^\infty_d$,*

$$R^{\mathrm{mar}}(Y) = \mathrm{cl} \bigcup_{X \in \Lambda_m(Y)} R(X). \tag{5.6}$$

To recover weak*-closedness, we define the closed version of R^{mar} via the notion of *closed hull*.

Definition 5.3. The *closed hull* $\mathrm{cl}\, F$ of a function $F : L^\infty_d \to \mathcal{G}_m$ is the pointwise greatest weak*-closed function minorizing it, that is, if $F' : L^\infty_d \to \mathcal{G}_m$ is a weak*-closed function such that $F(Y) \subseteq F'(Y)$ for all $Y \in L^\infty_d$, then we have $(\mathrm{cl}\, F)(Y) \subseteq F'(Y)$ for every $Y \in L^\infty_d$. The closed hull $\mathrm{cl}\, R^{\mathrm{mar}}$ of R^{mar} is called the *closed market risk measure* induced by R.

One can check that monotonicity, translativity and convexity are preserved under taking the closed hull. Hence, in view of Proposition 5.1, the closed market risk measure induced by a closed convex risk measure $R : L^\infty_m \to \mathcal{G}_m$ is a closed convex risk measure if $(\mathrm{cl}\, R^{\mathrm{mar}})(0) \neq \mathbb{R}^m$. Theorem 5.1 below gives a dual representation of the closed market risk measure in terms of the penalty function of the original risk measure R under the assumption of finiteness at zero. The special case of no trading constraints in a convex (conical) market model is given in Corollary 5.1 (Corollary 5.2). The set of dual variables to be used in the results below is given by

$$\mathcal{W}_{m,d} := \mathcal{M}_d(\mathbb{P}) \times ((\mathbb{R}^m_+ \backslash \{0\}) \times \mathbb{R}^{d-m}_+). \tag{5.7}$$

We will also make use of the homogeneous halfspaces $G(w) := \{y \in \mathbb{R}^d \,|\, w^{\mathsf{T}} y \geq 0\}$ for $w \in \mathbb{R}^d_+ \backslash \{0\}$.

Theorem 5.1. *Suppose that $R : L^\infty_m \to \mathcal{G}_m$ is a closed convex risk measure with penalty function $-\alpha_R : \mathcal{M}_m(\mathbb{P}) \times (\mathbb{R}^m_+ \backslash \{0\}) \to \mathcal{G}_m$, see Proposition* 3.3*. Assume that $(\mathrm{cl}\, R^{\mathrm{mar}})(0) \neq \mathbb{R}^m$. Then the closed market risk measure $\mathrm{cl}\, R^{\mathrm{mar}} : L^\infty_d \to \mathcal{G}_m$*

is also a closed convex risk measure, and it has the following dual representation: For every $Y \in L_d^\infty$,

$$(\operatorname{cl} R^{\max})(Y) = \bigcap_{(\mathbb{Q},w)\in\mathcal{W}_{m,d}} [-\alpha_{\operatorname{cl} R^{\max}}(\mathbb{Q},w) + B^*((\mathbb{E}^{\mathbb{Q}}[-Y] + G(w)) \cap B(\mathbb{R}^m))], \tag{5.8}$$

where $-\alpha_{\operatorname{cl} R^{\max}} : \mathcal{W}_{m,d} \to \mathcal{G}_m$ *is defined by*

$$\alpha_{\operatorname{cl} R^{\max}}(\mathbb{Q},w) = -\alpha_R(B^*\mathbb{Q}, B^*w) + \sum_{t=0}^{T} \operatorname{cl} \bigcup_{U^t \in L_d^\infty(\mathcal{F}_t, \mathcal{C}_t\cap\mathcal{D}_t)} B^*((\mathbb{E}^{\mathbb{Q}}[U^t] + G(w)) \cap B(\mathbb{R}^m)). \tag{5.9}$$

Recall that the *recession cone* of a nonempty convex set $C \subseteq \mathbb{R}^d$ is the convex cone $0^+C := \{y \in \mathbb{R}^d \,|\, y + C \subseteq C\}$ and the *positive dual cone* of a nonempty convex cone $K \subseteq \mathbb{R}^d$ is the convex cone $K^+ := \{y \in \mathbb{R}^d \,|\, \forall\, k \in K : y^{\mathsf{T}}k \geq 0\}$; see Rockafellar (1970, Sec. 8, p. 61) and Zalinescu (2002, Sec. 1.1, p. 7), for instance.

Corollary 5.1. *Under the assumptions of Theorem* 5.1, *suppose that* $\mathcal{D}_t(\omega) = \mathbb{R}^d$ *for each* $\omega \in \Omega$ *and* $t \in \mathbb{T}$. *Then* $-\alpha_{\operatorname{cl} R^{\max}}$ *given by* (5.9) *is concentrated on the set*

$$\mathcal{W}_{m,d}^{\text{convex}} := \left\{(\mathbb{Q},w) \in \mathcal{W}_{m,d} \,|\, \forall\, t \in \mathbb{T} : w \cdot \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_t\right] \in L_d^1(\mathcal{F}_t, (0^+\mathcal{C}_t)^+)\right\}, \tag{5.10}$$

where, for each $t \in \mathbb{T}$, $(0^+\mathcal{C}_t)^+ : \Omega \to \mathcal{G}_d$ *is the measurable function defined by* $(0^+\mathcal{C}_t)^+(\omega) := (0^+\mathcal{C}_t(\omega))^+$.

In other words, we have $-\alpha_{\operatorname{cl} R^{\max}}(\mathbb{Q},w) = \mathbb{R}^m$ for $(\mathbb{Q},w) \in \mathcal{W}_{m,d}\backslash\mathcal{W}_{m,d}^{\text{convex}}$ within the setting of the previous result.

Corollary 5.2. *Under the assumptions of Theorem* 5.1 *suppose that* $\mathcal{D}_t(\omega) = \mathbb{R}^d$ *for each* $\omega \in \Omega$ *and* $t \in \mathbb{T}$, *and that the market model is conical as in Example* 5.1. *Consider the set*

$$\mathcal{W}_{m,d}^{\text{cone}} := \left\{(\mathbb{Q},w) \in \mathcal{W}_{m,d} \,|\, \forall\, t \in \mathbb{T} : w \cdot \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\middle|\mathcal{F}_t\right] \in L_d^1(\mathcal{F}_t, \mathcal{C}_t^+)\right\}, \tag{5.11}$$

where, for each $t \in \mathbb{T}$, $\mathcal{C}_t^+ : \Omega \to \mathcal{G}_d$ *is the measurable function defined by* $\mathcal{C}_t^+(\omega) := (\mathcal{C}_t(\omega))^+$. *Then,* (5.9) *reduces to*

$$-\alpha_{\operatorname{cl} R^{\max}}(\mathbb{Q},w) = \begin{cases} -\alpha_R(B^*\mathbb{Q}, B^*w) & \text{if } (\mathbb{Q},w) \in \mathcal{W}_{m,d}^{\text{cone}}, \\ \mathbb{R}^m & \text{else}, \end{cases} \tag{5.12}$$

for each $(\mathbb{Q},w) \in \mathcal{W}_{m,d}$; *hence, for every* $Y \in L_d^\infty$,

$$(\operatorname{cl} R^{\max})(Y) = \bigcap_{(\mathbb{Q},w)\in\mathcal{W}_{m,d}^{\text{cone}}} [-\alpha_R(B^*\mathbb{Q}, B^*w) + B^*((\mathbb{E}^{\mathbb{Q}}[-Y] + G(w)) \cap B(\mathbb{R}^m))]. \tag{5.13}$$

The proofs of Theorem 5.1, Corollaries 5.1 and 5.2 above are given in Sec. 6.6. They rely on the observation that, roughly speaking, the market risk measure is the (*set-valued*) *infimal convolution* of the original risk measure and the (set-valued) *indicator functions* of the convex sets $L^\infty_d(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t)$, $t \in \mathbb{T}$. This technical observation is discussed in Sec. 6.6, where the definitions of these notions are also given.

5.3. *Market risk measures induced by shortfall and divergence risk measures*

In this section, we present sufficient conditions that guarantee the finite-valuedness condition $(\operatorname{cl} R^{\mathrm{mar}})(0) \neq \mathbb{R}^m$ for the closed market risk measures induced by shortfall and divergence risk measures. Once this property is established, these closed market risk measures are closed convex risk measures and their dual representations are provided by Theorem 5.1. For simplicity, we assume that the market model is conical in the sense of Example 5.1.

Assumption 5.1. Suppose that the solvency cones of the market model share a common supporting halfspace in the sense that there exists $\bar{w} \in \mathbb{R}^d_+ \backslash \{0\}$ such that for $\mathbb{P}$-almost every $\omega \in \Omega$ and every $t \in \mathbb{T}$, $\inf_{y \in \mathcal{C}_t(\omega)} \bar{w}^{\mathsf{T}} y > -\infty$, or equivalently, $\bar{w} \in (\mathcal{C}_t(\omega))^+$.

Remark 5.2. Assumption 5.1 states the existence of a halfspace $G(\bar{w}) = \{z \in \mathbb{R}^d \,|\, \bar{w}^{\mathsf{T}} z \geq 0\}$ for some $\bar{w} \in \mathbb{R}^d_+$ which satisfies $G(\bar{w}) \supseteq \mathcal{C}_t(\omega)$ for $\mathbb{P}$-almost every $\omega \in \Omega$ and $t \in \mathbb{T}$. In particular, when the solvency cones are constructed from bid-ask prices (see Kabanov 1999), this is equivalent to the ask prices having a uniform (in time and outcome) lower bound, or equivalently, the bid prices having a uniform (in time and outcome) upper bound. That is, $\bar{w}_j \leq \pi_{ij}(\omega, t)\bar{w}_i$ for every $i, j \in \{1, \ldots, d\}$, every $t \in \mathbb{T}$, and $\mathbb{P}$-almost every $\omega \in \Omega$, where $\pi_{ij}(\omega, t)$ is the number of units of asset i for which an agent can buy one unit of asset j at time t and state ω, and thus, denotes the ask price of asset j in terms of asset i.

Proposition 5.2. *Suppose that Assumption* 5.1 *holds and* $\operatorname{dom} \ell = \mathbb{R}^m$.

(1) *Let* $r \in \mathbb{R}^m_{++}$ *with* $\frac{1}{r} \in \operatorname{dom} g$. *If*

$$\inf_{x \in C} \bar{w}^{\mathsf{T}}(r \cdot x) > -\infty, \tag{5.14}$$

then $(\operatorname{cl} D^{\mathrm{mar}}_{\ell,r})(0) \neq \mathbb{R}^m$. *In particular,* $\operatorname{cl} D^{\mathrm{mar}}_{\ell,r}$ *is a closed convex risk measure with a dual representation provided by Theorem* 5.1.

(2) *If there exists* $r \in \mathbb{R}^m_{++}$ *with* $\frac{1}{r} \in \operatorname{dom} g$ *such that* (5.14) *holds, then* $(\operatorname{cl} R^{\mathrm{mar}}_{\ell})(0) \neq \mathbb{R}^m$. *In particular,* $\operatorname{cl} R^{\mathrm{mar}}_{\ell}$ *is a closed convex risk measure with a dual representation provided by Theorem* 5.1.

6. Proofs and Technical Remarks

6.1. *Proofs of the results in Sec.* 2

Proof of Proposition 2.1. Monotonicity, translativity and convexity are trivial. Let $X \in L^\infty$. It holds $\ell(-\operatorname{ess\,sup} X - s) \leq \mathbb{E}[\ell(-X-s)] \leq \ell(-\operatorname{ess\,inf} X - s)$ for every $s \in \mathbb{R}$. Note that ℓ is strictly increasing on $\ell^{-1}(\operatorname{int} \ell(\mathbb{R})) := \{x \in \mathbb{R} \,|\, \ell(x) \in \operatorname{int} \ell(\mathbb{R})\} = (a,b)$, where $a := \inf\{x \in \mathbb{R} \,|\, \ell(x) > \inf_{y\in\mathbb{R}} \ell(y)\} \in \mathbb{R} \cup \{-\infty\}$ and $b := \sup\{x \in \mathbb{R} \,|\, \ell(x) < +\infty\} \in \mathbb{R} \cup \{+\infty\}$. Hence, the inverse ℓ^{-1} is well-defined as a function from $\operatorname{int} \ell(\mathbb{R})$ to (a,b). It holds $\mathbb{E}[\ell(-X-s)] \leq 0$ for each $s \geq -\operatorname{ess\,inf} X - \ell^{-1}(0)$, and $\mathbb{E}[\ell(-X-s)] > 0$ for each $s < -\operatorname{ess\,sup} X - \ell^{-1}(0)$. So $\rho_\ell(X) \in \mathbb{R}$. Besides, $\mathbb{E}[\ell(-X-\rho_\ell(X))] \leq 0$ since the restriction of ℓ on $\operatorname{dom} \ell$ is a continuous function. To show (weak*-)lower semicontinuity, let $(X^n)_{n\in\mathbb{N}}$ be a bounded sequence in L^∞ converging to some $X \in L^\infty$ $\mathbb{P}$-almost surely. Then, using Fatou's lemma together with the fact that the restriction of ℓ on $\operatorname{dom} \ell$ is nondecreasing and continuous, we have

$$\begin{aligned}\mathbb{E}\Big[\ell\Big(-X - \liminf_{n\to\infty} \rho_\ell(X^n)\Big)\Big] &= \mathbb{E}\Big[\ell\Big(\liminf_{n\to\infty} (-X^n - \rho_\ell(X^n))\Big)\Big] \\ &\leq \liminf_{n\to\infty} \mathbb{E}[\ell(-X^n - \rho_\ell(X^n))] \leq 0. \qquad (6.1)\end{aligned}$$

This implies the so-called *Fatou property* of ρ_ℓ, namely, that $\rho_\ell(X) \leq \liminf_{n\to\infty} \rho_\ell(X^n)$. By Föllmer & Schied (2011, Theorem 4.33), this is equivalent to the lower semicontinuity of ρ_ℓ. □

Proof of Proposition 2.2. Note that $s \mapsto \mathbb{E}[\ell(-X-s)]$ is a proper convex function on $\mathbb{R}$. Hence, by Definition 2.2, $\rho_\ell(X)$ is the optimal value of a convex minimization problem. The corresponding Lagrangian dual objective function h on $\mathbb{R}_+$ is given by

$$h(\lambda) = \inf_{s\in\mathbb{R}:\mathbb{E}[\ell(-X-s)]<+\infty} (s + \lambda \mathbb{E}[\ell(-X-s)]). \qquad (6.2)$$

Clearly, $h(\lambda) = \delta_{\ell,\lambda}(X)$ if $\lambda > 0$ since $\lambda\mathbb{E}[\ell(-X-s)] = +\infty$ if $\mathbb{E}[\ell(-X-s)] = +\infty$. On the other hand, note that $\mathbb{E}[\ell(-X-s)] < +\infty$ if and only if $\mathbb{P}\{-X - s \in \operatorname{dom} \ell\} = 1$. It follows that

$$h(0) = \inf\{s \in \mathbb{R} \,|\, \mathbb{E}[\ell(-X-s)] < +\infty\} = -\operatorname{ess\,inf} X - \sup \operatorname{dom} \ell. \qquad (6.3)$$

Therefore, the optimal value of the dual problem equals the right-hand side of (2.2). Finally, the two sides of (2.2) are equal since the usual *Slater's condition* holds: There exists $\bar{s} \in \mathbb{R}$ such that $\mathbb{E}[\ell(-X-\bar{s})] < 0$. This is because we have $\mathbb{E}[\ell(-X-s)] < 0$ for each $s > -\operatorname{ess\,inf} X - \ell^{-1}(0)$, where ℓ^{-1} is the inverse function on $\operatorname{int} \ell(\mathbb{R})$ as in the proof of Proposition 2.1. □

Proof of Proposition 2.3. Let f be a loss function and $f^* : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ its conjugate function. Note that $\operatorname{dom} f^* \subseteq \mathbb{R}_+$ since, for each $y < 0$,

we have

$$f^*(y) \geq \sup_{n\in\mathbb{N}}(-ny - f(-n)) \geq \sup_{n\in\mathbb{N}}(-ny) - f(0) = +\infty, \tag{6.4}$$

where we use the monotonicity of f for the second inequality. Moreover, $0 \in \operatorname{dom} f^*$ since $f^*(0) = -\inf_{x\in\mathbb{R}} f(x) < +\infty$. Clearly, $f^*(y) \geq -f(0)$ for each $y \in \mathbb{R}$. Besides, by Rockafellar (1970, Theorem 23.3), the subdifferential $\partial f(0)$ of f at 0 is nonempty and, by Rockafellar (1970, Theorem 23.5), we have $f^*(y) = -f(0)$ for every $y \in \partial f(0)$. Hence, f^* attains its infimum. Finally, f^* is not of the form $y \mapsto +\infty \cdot 1_{\{y<0\}} + (ay + b) \cdot 1_{\{y\geq 0\}}$ for some $a \in \mathbb{R}_+ \cup \{+\infty\}$ and $b \in \mathbb{R}$ as otherwise we would get $f(x) = (f^*)^*(x) = +\infty \cdot 1_{\{x>a\}} - b \cdot 1_{\{x\leq a\}}$, $x \in \mathbb{R}$, so that f would be identically constant on $\operatorname{dom} f$. Hence, f^* is a divergence function. Conversely, let φ be a divergence function and $\varphi^* : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ its conjugate function. Let $x^1, x^2 \in \mathbb{R}$ with $x^1 \geq x^2$. Since $\operatorname{dom}\varphi \subseteq \mathbb{R}_+$, we have $x^1 y - \varphi(y) \geq x^2 y - \varphi(y)$ for each $y \in \operatorname{dom}\varphi$ so that $\varphi^*(x^1) \geq \varphi^*(x^2)$. Hence, φ^* is nondecreasing. Moreover, $\inf_{x\in\mathbb{R}} \varphi^*(x) = -\varphi(0) > -\infty$ since $\varphi = (\varphi^*)^*$ and $0 \in \operatorname{dom}\varphi$. Clearly, $\varphi^*(0) = -\inf_{y\in\mathbb{R}} \varphi(y) \in \mathbb{R}$ so that $0 \in \operatorname{dom}\varphi^*$. Finally, φ^* is not identically constant on $\operatorname{dom}\varphi^*$ as otherwise $\varphi = (\varphi^*)^*$ would fail to satisfy property (3) in Definition 2.3. Hence, φ^* is a loss function. □

Proof of Theorem 2.1. If $\lambda = 0$, then $I_{g,\lambda}(\mathbb{Q} \,|\, \mathbb{P}) = \sup \operatorname{dom}\ell$ for every $\mathbb{Q} \in \mathcal{M}(\mathbb{P})$ and we have

$$\delta_{\ell,0}(X) = -\operatorname{ess\,inf} X - \sup \operatorname{dom}\ell = \sup_{\mathbb{Q}\in\mathcal{M}(\mathbb{P})} \mathbb{E}^{\mathbb{Q}}[-X] - \sup \operatorname{dom}\ell \tag{6.5}$$

by the dual representation of the worst-case risk measure $X \mapsto -\operatorname{ess\,inf} X$; see Föllmer & Schied (2011, Example 4.39), for instance. Hence, (2.7) holds in the case $\lambda = 0$. Moreover, by Remark 2.2, $\sup \operatorname{dom}\ell < +\infty$ if and only if $1 \in \operatorname{dom} g_0$. Hence, $\delta_{\ell,0}$ is a lower semicontinuous convex risk measure if $1 \in \operatorname{dom} g_0$, and $\delta_{\ell,0}(X) = -\infty$ for every $X \in L^\infty$ otherwise.

Suppose $\lambda > 0$. Note that the right-hand side of (2.7) can be rewritten as a maximization problem on the space L^1_+ of integrable nonnegative real-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ (identified up to almost sure equality):

$$\begin{aligned}&\sup_{\mathbb{Q}\in\mathcal{M}(\mathbb{P})} (\mathbb{E}^{\mathbb{Q}}[-X] - I_{g,\lambda}(\mathbb{Q} \,|\, \mathbb{P})) \\ &\qquad = \sup_{V\in L^1_+} \left\{ \mathbb{E}[-XV] - \lambda \mathbb{E}\left[g\left(\frac{1}{\lambda} V\right)\right] \,\middle|\, \mathbb{E}[V] = 1 \right\}. \end{aligned} \tag{6.6}$$

The optimal value of the corresponding Lagrangian dual problem is computed as

$$\begin{aligned} q_X &:= \inf_{s\in\mathbb{R}} \sup_{V\in L^1_+} \left(\mathbb{E}[-XV] - \lambda \mathbb{E}\left[g\left(\frac{1}{\lambda} V\right)\right] + s(1 - \mathbb{E}[V]) \right) \\ &= \inf_{s\in\mathbb{R}} \left(s + \sup_{V\in L^1_+} \mathbb{E}\left[(-X - s)V - \lambda g\left(\frac{1}{\lambda} V\right)\right] \right) \end{aligned}$$

$$= \inf_{s\in\mathbb{R}} \left(s + \mathbb{E}\left[\sup_{z\in\mathbb{R}_+} \left((-X-s)z - \lambda g\left(\frac{1}{\lambda} z\right)\right)\right]\right)$$
$$= \inf_{s\in\mathbb{R}} (s + \mathbb{E}[g_\lambda^*(-X-s)]), \tag{6.7}$$

where the third equality is due to Rockafellar & Wets (1998, Theorem 14.60), and g_λ^* is the conjugate of the divergence function g_λ; see Remark 2.2. Hence, $g_\lambda^* = \lambda\ell$ and q_X equals the left-hand side of (2.7). Finally, to conclude (2.7), we consider the following cases:

(1) Suppose that $1 \in \operatorname{int}\operatorname{dom} g_\lambda$, that is, $\frac{1}{\lambda} < \beta$. (Recall that $\operatorname{int}\operatorname{dom} g_\lambda = (0, \lambda\beta)$, see Definition 2.4 *et seq.*) Then the following constraint qualification holds, for instance, with $\bar{V} \equiv 1$:

$$\exists\, \bar{V} \in L^1_+ : \mathbb{E}[\bar{V}] = 1, \quad \bar{V} \in \operatorname{int}\operatorname{dom} g_\lambda \ \mathbb{P}\text{-almost surely.} \tag{6.8}$$

By Borwein & Lewis (1992, Corollary 4.8), (6.8) suffices to conclude (2.7). Note that we have

$$-\mathbb{E}[X] - \lambda g\left(\frac{1}{\lambda}\right) \le \sup_{V\in L^1_+} \left\{ \mathbb{E}[-XV] - \lambda\mathbb{E}\left[g\left(\frac{1}{\lambda} V\right)\right] \,\middle|\, \mathbb{E}[V] = 1\right\}$$
$$\le -\operatorname{ess\,inf} X - \lambda \inf_{x\in\mathbb{R}} g(x) \tag{6.9}$$

so that both sides of (2.7) are in $\mathbb{R}$.

(2) Suppose that $\lambda\beta = 1$ and $\operatorname{dom} g_\lambda = [0, \lambda\beta] = [0, 1]$, that is, $\operatorname{dom} g = [0, \beta] = [0, \frac{1}{\lambda}]$. In this case, the only $V \in L^1_+$ with $\mathbb{E}[V] = 1$ and $\mathbb{P}\{V \in \operatorname{dom} g_\lambda\} = 1$ is $V \equiv 1$, and hence, the right-hand side of (2.7) gives $-\mathbb{E}[X] - \lambda g(\frac{1}{\lambda}) \in \mathbb{R}$. Note that (6.8) fails to hold here. Using (2.7) for the previous case, we have

$$\inf_{s\in\mathbb{R}} (s + \lambda\mathbb{E}[\ell(-X-s)]) = \lim_{\varepsilon\downarrow 0} \inf_{s\in\mathbb{R}} (s + (\lambda+\varepsilon)\mathbb{E}[\ell(-X-s)])$$
$$= \lim_{\varepsilon\downarrow 0} \sup_{\mathbb{Q}\in\mathcal{M}(\mathbb{P})} \left(\mathbb{E}^{\mathbb{Q}}[-X] - (\lambda+\varepsilon)\mathbb{E}\left[g\left(\frac{1}{\lambda+\varepsilon} \frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]\right), \tag{6.10}$$

where the first equality follows since the proper, concave, upper semicontinuous function

$$\mathbb{R} \ni \gamma \mapsto \inf_{s\in\mathbb{R}} (s + \gamma\mathbb{E}\ell(-X-s)) \in \mathbb{R}\cup\{-\infty\} \tag{6.11}$$

is right-continuous at $\gamma = \lambda$. Finally, we have

$$\lim_{\varepsilon\downarrow 0} \sup_{\mathbb{Q}\in\mathcal{M}(\mathbb{P})} \left(\mathbb{E}^{\mathbb{Q}}[-X] - (\lambda+\varepsilon)\mathbb{E}\left[g\left(\frac{1}{\lambda+\varepsilon} \frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]\right)$$
$$= \lim_{\varepsilon\downarrow 0} \sup_{\mathbb{Q}\in\mathcal{M}(\mathbb{P})} \left(\mathbb{E}^{\mathbb{Q}}[-X] - (\lambda+\varepsilon)\mathbb{E}\left[g\left(\frac{1}{\lambda+\varepsilon} \frac{d\mathbb{Q}}{d\mathbb{P}}\right) - g(0)\right]\right)$$
$$- \lim_{\varepsilon\downarrow 0} (\lambda+\varepsilon) g(0) \tag{6.12}$$

$$= \inf_{\gamma\in[\lambda,\lambda+\varepsilon']} \sup_{\mathbb{Q}\in\mathcal{M}(\mathbb{P})} \left(\mathbb{E}^{\mathbb{Q}}[-X] - \gamma\mathbb{E}\left[g\left(\frac{1}{\gamma}\frac{d\mathbb{Q}}{d\mathbb{P}}\right) - g(0)\right]\right) - \lambda g(0)$$

$$= \sup_{\mathbb{Q}\in\mathcal{M}(\mathbb{P})} \inf_{\gamma\in[\lambda,\lambda+\varepsilon']} \left(\mathbb{E}^{\mathbb{Q}}[-X] - \gamma\mathbb{E}\left[g\left(\frac{1}{\gamma}\frac{d\mathbb{Q}}{d\mathbb{P}}\right) - g(0)\right]\right) - \lambda g(0)$$

$$= \sup_{\mathbb{Q}\in\mathcal{M}(\mathbb{P})} \left(\mathbb{E}^{\mathbb{Q}}[-X] - \lambda\mathbb{E}\left[g\left(\frac{1}{\lambda}\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]\right)$$

$$= -\mathbb{E}[X] - \lambda g\left(\frac{1}{\lambda}\right) \in \mathbb{R}, \tag{6.13}$$

where $\varepsilon' > 0$ is some fixed number. Here, the second equality follows since the function $\gamma \mapsto \gamma(g(\frac{y}{\gamma}) - g(0))$ is a nonincreasing function on $\mathbb{R}_{++}$ for each $y \in \mathbb{R}$; see Remark 2.2. The third equality is due to a classical minimax theorem and it uses the compactness of the interval $[\lambda, \lambda+\varepsilon']$, see Sion (1958, Corollary 3.3). The fourth equality follows by monotone convergence theorem and the monotonicity of the function $\gamma \mapsto \gamma(g(\frac{y}{\gamma}) - g(0))$ on $\mathbb{R}_{++}$. The last equality is already discussed above. Finally, the first equality follows since the two limits in (6.13) are shown to be finite by the succeeding equalities. Hence, we obtain (2.7).

(3) Suppose $1 \notin \operatorname{dom} g_\lambda$, that is, either $\operatorname{dom} g_\lambda = [0, \lambda\beta)$ and $\lambda\beta \geq 1$, or, $\operatorname{dom} g_\lambda = [0, \lambda\beta]$ and $\lambda\beta > 1$. In this case, there is no $Y \in L^1_+$ with $\mathbb{E}[Y] = 1$ and $\mathbb{P}\{Y \in \operatorname{dom} g_\lambda\} = 1$. Hence, the right-hand side of (2.7) gives $-\infty$. On the other hand, we have

$$\begin{aligned}\inf_{s\in\mathbb{R}}(s + \lambda\mathbb{E}[\ell(-X-s)]) &\leq \inf_{s\in\mathbb{R}}(s + \lambda\ell(-\operatorname{ess\,inf} X - s)) \\ &= -\operatorname{ess\,inf} X - \sup_{s\in\mathbb{R}}(s - \lambda\ell(s)) \\ &= -\operatorname{ess\,inf} X - \lambda g\left(\frac{1}{\lambda}\right) = -\infty. \end{aligned} \tag{6.14}$$

Hence, (2.7) holds. In the first two cases where $1 \in \operatorname{dom} g_\lambda$, we observe that $\delta_{\ell,\lambda}(0) \in \mathbb{R}$. Moreover, (2.3) guarantees monotonicity, translativity, convexity and lower semicontinuity directly, which makes $\delta_{\ell,\lambda}$ a lower semicontinuous convex risk measure. In the last case where $1 \notin \operatorname{dom} g_\lambda$, $\delta_{\ell,\lambda}(X) = -\infty$ for every $X \in L^\infty$. □

Proof of Proposition 2.4. Let $\mathbb{Q} \in \mathcal{M}(\mathbb{P})$ and $\lambda \in \mathbb{R}_+$ with $1 \in \operatorname{dom} g_\lambda$. If $\lambda = 0$, then it follows from Definition 2.3 and the proof of Theorem 2.1 that $\alpha_{\delta_{\ell,0}}(\mathbb{Q}) = \sup \operatorname{dom} \ell = I_{g,0}(\mathbb{Q}\,|\,\mathbb{P})$. Suppose $\lambda > 0$. Using (2.3) and the definition of penalty function in (2.10),

$$\begin{aligned}\alpha_{\delta_{\ell,\lambda}}(\mathbb{Q}) &= \sup_{s\in\mathbb{R}}\left(-s + \sup_{X\in L^\infty}\mathbb{E}\left[-\frac{d\mathbb{Q}}{d\mathbb{P}}X - \lambda\ell(-X-s)\right]\right) \\ &= \sup_{s\in\mathbb{R}}\left(-s + \mathbb{E}\left[\sup_{x\in\mathbb{R}}\left(-\frac{d\mathbb{Q}}{d\mathbb{P}}x - \lambda\ell(-x-s)\right)\right]\right)\end{aligned}$$

$$= \sup_{s\in\mathbb{R}} \left(-s + \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} s + g_\lambda \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) \right] \right)$$
$$= I_{g,\lambda}(\mathbb{Q} \,|\, \mathbb{P}), \tag{6.15}$$

where the second equality follows from Rockafellar & Wets (1998, Theorem 14.60) and the third equality follows from Remark 2.2. For the penalty function of ρ_ℓ, note that

$$\begin{aligned} \alpha_{\rho_\ell}(\mathbb{Q}) &= \sup_{X\in L^\infty} \left(\mathbb{E}^{\mathbb{Q}}[-X] - \inf_{s\in\mathbb{R}} (s + I_{(-\infty,0]}(\mathbb{E}[\ell(-X-s)])) \right) \\ &= \sup_{X\in L^\infty} (\mathbb{E}^{\mathbb{Q}}[X] - I_{(-\infty,0]}(\mathbb{E}[\ell(X)])) \\ &= \sup_{X\in L^\infty} \{\mathbb{E}^{\mathbb{Q}}[X] \,|\, \mathbb{E}[\ell(X)] \le 0\}. \end{aligned} \tag{6.16}$$

For the last maximization problem, the corresponding Lagrangian dual objective function h on $\mathbb{R}_+$ is given by

$$h(\lambda) = \sup_{X\in L^\infty : \mathbb{E}[\ell(X)] < +\infty} (\mathbb{E}^{\mathbb{Q}}[X] - \lambda \mathbb{E}[\ell(X)]). \tag{6.17}$$

Note that $\mathbb{E}[\ell(X)] < +\infty$ if and only if $\mathbb{P}\{X \in \operatorname{dom}\ell\} = 1$. If $\lambda = 0$, then

$$h(0) = \sup_{X\in L^\infty : \mathbb{E}[\ell(X)] < +\infty} \mathbb{E}^{\mathbb{Q}}[X] = \sup \operatorname{dom}\ell = I_{g,0}(\mathbb{Q} \,|\, \mathbb{P}). \tag{6.18}$$

On the other hand, if $\lambda > 0$, then

$$\begin{aligned} h(\lambda) &= \sup_{X\in L^\infty : \mathbb{E}[\ell(X)] < +\infty} (\mathbb{E}^{\mathbb{Q}}[X] - \lambda \mathbb{E}[\ell(X)]) \\ &= \mathbb{E}\left[\sup_{x\in\mathbb{R}} \left(\frac{d\mathbb{Q}}{d\mathbb{P}} x - \lambda \ell(x) \right) \right] = I_{g,\lambda}(\mathbb{Q} \,|\, \mathbb{P}), \end{aligned} \tag{6.19}$$

where we use Rockafellar & Wets (1998, Theorem 14.60) for the second equality and Remark 2.2 for the third equality. Hence, the optimal value of the dual problem is given by

$$q(\mathbb{Q}) := \inf_{\lambda\in\mathbb{R}_+} h(\lambda) = \inf_{\lambda\in\mathbb{R}_+} I_{g,\lambda}(\mathbb{Q} \,|\, \mathbb{P}). \tag{6.20}$$

Note that Slater's condition holds, that is, there exists $\bar{X} \in L^\infty$ such that $\mathbb{E}[\ell(\bar{X})] < 0$; take, for example, $\bar{X} \equiv \ell^{-1}(0) - 1$, where ℓ^{-1} is the inverse function on $\operatorname{int}\ell(\mathbb{R})$ as in the proof of Proposition 2.1. Therefore, $\alpha_{\rho_\ell}(\mathbb{Q}) = q(\mathbb{Q})$. Note that $I_{g,\lambda}(\mathbb{Q} \,|\, \mathbb{P}) = +\infty$ for every $\lambda \in \mathbb{R}_+$ with $1 \notin \operatorname{dom} g_\lambda$, see case (3) in the proof of Theorem 2.1. Hence, we also have $q(\mathbb{Q}) = \inf_{\lambda\in\mathbb{R}_+ \,:\, 1\in\operatorname{dom} g_\lambda} \alpha_{\delta_{\ell,\lambda}}(\mathbb{Q})$. □

6.2. *A remark about the scalar loss functions*

In Föllmer & Schied (2002, Theorem 10) and Föllmer & Schied (2011, Theorem 4.115), the second part of Proposition 2.4 is proved with the additional assumption that ℓ maps into $\mathbb{R}$. This assumption implies that the ℓ-shortfall risk

measure is continuous from below and the first supremum in (2.10) is attained (Föllmer & Schied 2011, Proposition 4.113). Besides, the same assumption implies the so-called *superlinear growth condition* on g, namely, that $\lim_{y\to\infty} \frac{g(y)}{y} = +\infty$ (Föllmer & Schied 2002, Lemma 11). The analytic proof for Proposition 2.4 in Föllmer & Schied (2002) makes use of this property instead of the dual relationship with divergence risk measures. Using this proposition and assuming that $1 \in \operatorname{dom} g$, Theorem 2.1 is proved for $\lambda = 1$ (Föllmer & Schied 2011, Theorem 4.122), in which case $\delta_{\ell,1}$ is guaranteed to be a risk measure (it has finite values). In our treatment, while 1 may not be in $\operatorname{dom} g$, there exists some $\bar{\lambda} > 0$ with $1 \in \operatorname{dom} g_{\bar{\lambda}}$ and hence $\delta_{\ell,\bar{\lambda}}$ is a risk measure.

In Ben Tal & Teboulle (2007), on the other hand, the divergence function g is of central importance. In addition to the assumptions here, it is assumed in Ben Tal & Teboulle (2007) that g attains its infimum at 1 with value 0, which is equivalent to assuming that $\ell(0) = 0$ and $1 \in \partial\ell(0)$. These assumptions make g a natural divergence function in the sense that the function $\mathbb{Q} \mapsto \mathbb{E}[g(\frac{d\mathbb{Q}}{d\mathbb{P}})]$ on $\mathcal{M}(\mathbb{P})$ has nonnegative values and takes the value 0 if $\mathbb{Q} = \mathbb{P}$; $\mathbb{E}[g(\frac{d\mathbb{Q}}{d\mathbb{P}})]$ can be interpreted as the distance between some "subjective" measure $\mathbb{Q} \in \mathcal{M}(\mathbb{P})$ and the physical measure $\mathbb{P}$. On the other hand, the additional assumptions on the loss function ℓ may be considered as restrictive. Here, we take ℓ as the central object by dropping these assumptions and use the convex duality approach as in Ben Tal & Teboulle (2007). Note that Theorem 2.1 (Ben Tal & Teboulle 2007, Theorem 4.2) and the first part of Proposition 2.4 (Ben Tal & Teboulle 2007, Theorem 4.4) are proved in Ben Tal & Teboulle (2007) for the case $\lambda = 1$. Here, we generalize this proof, basically, by considering the cases where the constraint qualification (6.8), which is also used in the proof of Ben Tal & Teboulle (2007, Theorem 4.2), fails to hold.

6.3. *Lagrange duality for set optimization: A quick review*

The proofs of Theorem 3.1 and Proposition 3.5 rely on the application of the recent Lagrange duality in Hamel & Löhne (2014). We quickly review the definition of the dual problem here. Let $\mathcal{X}$ be a locally convex topological linear space. Consider a set minimization problem of the form (3.15), where $\Phi : \mathcal{X} \to \mathcal{G}_m$ is an arbitrary objective function and $\Psi : \mathcal{X} \to \mathcal{G}_m$ is an arbitrary constraint function. The optimal value of this problem is $p := \inf_{(\mathcal{G}_m,\supseteq)}\{\Phi(x) \,|\, 0 \in \Psi(x), x \in \mathcal{X}\}$.

The halfspace-valued functions $S_{\lambda,v} : \mathbb{R}^m \to \mathcal{G}_m$ for $\lambda \in \mathbb{R}^m$, $v \in \mathbb{R}^m_+\backslash\{0\}$ defined by

$$S_{\lambda,v}(z) = \{\eta \in \mathbb{R}^m \,|\, v^{\mathsf{T}}\eta \geq \lambda^{\mathsf{T}} z\} \tag{6.21}$$

will be used as set-valued substitutes for the (continuous) linear functionals of the scalar duality theory as in Hamel (2009), Hamel & Löhne (2014). Here, there are two types of dual variables: The variable $\lambda \in \mathbb{R}^m$ is the usual vector of *Lagrange multipliers* which is used to scalarize the values of Ψ whereas the variable $v \in \mathbb{R}^m_+$ is the *weight vector* which is used to scalarize the values of Φ. The set-valued

Lagrangian $L : \mathcal{X} \times \mathbb{R}^m \times (\mathbb{R}^m_+ \backslash \{0\}) \to \mathcal{G}_m$ and the objective function $H : \mathbb{R}^m \times (\mathbb{R}^m_+ \backslash \{0\}) \to \mathcal{G}_m$ of the dual problem for (3.15) are defined by

$$\begin{aligned} L(x, \lambda, v) &= \mathrm{cl}\Big(\Phi(x) + \inf_{(\mathcal{G}_m, \supseteq)} \{S_{\lambda,v}(z) \,|\, z \in \Psi(x)\}\Big), \\ H(\lambda, v) &= \inf_{(\mathcal{G}_m, \supseteq)} \{L(x, \lambda, v) \,|\, x \in \mathcal{X}\}. \end{aligned} \tag{6.22}$$

The optimal value q of the dual problem is the supremum of the dual objective function over the dual variables:

$$q := \sup_{(\mathcal{G}_m, \supseteq)} \{H(\lambda, v) \,|\, \lambda \in \mathbb{R}^m, v \in \mathbb{R}^m_+ \backslash \{0\}\}. \tag{6.23}$$

Proposition 6.1 (Hamel & Löhne 2014, Theorem 6.1). *Assume that* Φ *and* Ψ *are convex functions and* $p \neq \mathbb{R}^m$. *Strong duality holds, that is,* $p = q$ *if the following set-valued version of Slater's condition holds*: *There exists* $\bar{x} \in \mathcal{X}$ *such that* $\Phi(\bar{x}) \neq \emptyset$ *and* $\Psi(\bar{x}) \cap -\mathbb{R}^m_{++} \neq \emptyset$.

6.4. *Proofs of the results in Sec.* 3

Proof of Proposition 3.1. Monotonicity, translativity and convexity are trivial. To show finiteness at 0, using the proof of Proposition 2.1, we can find $z^1 \in \mathbb{R}^m$ with $\ell(-z^1) \in -\mathbb{R}^m_+$ and $z^2 \in \mathbb{R}^m$ with $\ell(-z^2) \in \mathbb{R}^m_{++}$. By the properties of C, it follows that $R_\ell(0) \notin \{\emptyset, \mathbb{R}^m\}$. To show weak*-closedness, let $(X^n)_{n\in\mathbb{N}}$ be a bounded sequence in L^∞_m converging to some $X \in L^\infty_m$ $\mathbb{P}$-almost surely. Let $z \in \mathbb{R}^m$ and suppose that there exists $z^n \in R_\ell(X^n)$, for each $n \in \mathbb{N}$, such that $(z^n)_{n\in\mathbb{N}}$ converges to z. Using dominated convergence theorem, the closedness of $-C$, and the fact that the restriction of ℓ on $\operatorname{dom} \ell := \{x \in \mathbb{R}^m \,|\, \ell(x) \in \mathbb{R}^m\} = \bigtimes_{i=1}^m \operatorname{dom} \ell_i \subseteq \mathbb{R}^m$ is continuous, we have

$$\mathbb{E}[\ell(-X - z)] = \mathbb{E}\Big[\ell\Big(\lim_{n\to\infty} (-X^n - z^n)\Big)\Big] = \lim_{n\to\infty} \mathbb{E}[\ell(-X^n - z^n)] \in -C, \tag{6.24}$$

that is, $z \in R_\ell(X)$. This shows the so-called *Fatou property* of R_ℓ, namely, that

$$\liminf_{n\to\infty} R_\ell(X^n) := \Big\{z \in \mathbb{R}^m \,|\, \forall\, n \in \mathbb{N}\ \exists\, z^n \in R_\ell(X^n) : \lim_{n\to\infty} z^n = z\Big\} \subseteq R_\ell(X). \tag{6.25}$$

By Hamel & Heyde (2010, Theorem 6.2), this is equivalent to the weak*-closedness of R_ℓ. □

For the proof of Theorem 3.1, we will need the following lemmata.

Lemma 6.1. *For every* $X \in L^\infty_m$,

$$R_\ell(X) = \bigcap_{\substack{\lambda\in\mathbb{R}^m_+, \\ v\in\mathbb{R}^m_+\backslash\{0\}}} \left\{\eta \in \mathbb{R}^m \,|\, v^\mathsf{T}\eta \geq \inf_{\substack{z\in\mathbb{R}^m: \\ \mathbb{E}[\ell(-X-z)]\in\mathbb{R}^m}} f_{\lambda,v}(z) + \inf_{x\in C} \lambda^\mathsf{T} x\right\}, \tag{6.26}$$

where

$$f_{\lambda,v}(z) := v^{\mathsf{T}} z + \lambda^{\mathsf{T}} \mathbb{E}[\ell(-X - z)]. \tag{6.27}$$

Proof. Let $X \in L^{\infty}_m$. Using (6.22) and (3.12), the Lagrangian for the problem (3.14) is computed as

$$\begin{aligned} L(z, \lambda, v) &= \mathrm{cl}\left(z + \mathbb{R}^m_+ + \mathrm{cl} \bigcup_{x \in (\mathbb{E}[\ell(-X-z)]+C) \cap \mathbb{R}^m} S_{\lambda,v}(x) \right) \\ &= \begin{cases} \{\eta \in \mathbb{R}^m \mid v^{\mathsf{T}} \eta \geq f_{\lambda,v}(z) + \inf_{x \in C} \lambda^{\mathsf{T}} x\} & \text{if } \mathbb{E}[\ell(-X-z)] \in \mathbb{R}^m, \\ \emptyset & \text{if } \mathbb{E}[\ell(-X-z)] \notin \mathbb{R}^m \end{cases} \end{aligned} \tag{6.28}$$

for $z \in \mathbb{R}^m, \lambda \in \mathbb{R}^m, v \in \mathbb{R}^m_+ \backslash \{0\}$. Hence, the dual objective function is given by

$$H(\lambda, v) = \left\{ \eta \in \mathbb{R}^m \mid v^{\mathsf{T}} \eta \geq \inf_{z \in \mathbb{R}^m : \mathbb{E}[\ell(-X-z)] \in \mathbb{R}^m} f_{\lambda,v}(z) + \inf_{x \in C} \lambda^{\mathsf{T}} x \right\} \tag{6.29}$$

for $\lambda \in \mathbb{R}^m$, $v \in \mathbb{R}^m_+ \backslash \{0\}$. Suppose $\lambda \notin \mathbb{R}^m_+$. Since $C + \mathbb{R}^m_+ \subseteq C$, there exists $\bar{x} \in C$ such that, for every $n \in \mathbb{N}$, we have $n\bar{x} \in C$ and $\lambda^{\mathsf{T}} \bar{x} < 0$. Hence, $\inf_{x \in C} \lambda^{\mathsf{T}} x = -\infty$ and $H(\lambda, v) = \mathbb{R}^m$ for every $v \in \mathbb{R}^m_+ \backslash \{0\}$. Therefore, by (6.23) and (3.12), the optimal value of the dual problem is given by the right-hand side of (6.26). Finally, the two sides of (6.26) are equal by Proposition 6.1 since *Slater's condition* holds: There exists $\bar{z} \in \mathbb{R}^m$ such that $(\mathbb{E}[\ell(-X - \bar{z})] + C) \cap -\mathbb{R}^m_{++} \neq \emptyset$. This follows as for the scalar version, see the proof of Proposition 2.2. □

Lemma 6.2. *Set* $w^{\mathsf{T}}(-\infty) = -\infty$ *whenever* $w \in \mathbb{R}^m_+ \backslash \{0\}$. *Then, for every* $X \in L^{\infty}_m$,

$$R_\ell(X) = \bigcap_{r \in \mathbb{R}^m_+, w \in \mathbb{R}^m_+ \backslash \{0\}} \left\{ z \in \mathbb{R}^m \mid w^{\mathsf{T}} z \geq w^{\mathsf{T}} \delta_{\ell,r}(X) + \inf_{x \in C} w^{\mathsf{T}}(r \cdot x) \right\}. \tag{6.30}$$

Proof. With (6.29) in view, for $r, w \in \mathbb{R}^m_+ \backslash \{0\}$, we define

$$M(r, w) := \left\{ \eta \in \mathbb{R}^m \mid w^{\mathsf{T}} \eta \geq w^{\mathsf{T}} \delta_{\ell,r}(X) + \inf_{x \in C} w^{\mathsf{T}}(r \cdot x) \right\}, \tag{6.31}$$

and we will show

$$\bigcap_{\substack{\lambda \in \mathbb{R}^m_+, \\ v \in \mathbb{R}^m_+ \backslash \{0\}}} H(\lambda, v) = \bigcap_{\substack{r \in \mathbb{R}^m_+, \\ w \in \mathbb{R}^m_+ \backslash \{0\}}} M(r, w). \tag{6.32}$$

First, if $r \in \mathbb{R}^m_+, w \in \mathbb{R}^m_+ \backslash \{0\}$, then we define $\lambda_i = r_i w_i$ and $v_i = w_i$ for $i \in \{1, \ldots, m\}$. Then, $\lambda \in \mathbb{R}^m_+, v \in \mathbb{R}^m_+ \backslash \{0\}$ as well as $H(\lambda, v) = M(r, w)$; see (3.24) and (3.25). This means that the intersection on the left-hand side runs over at least

as many sets as the one on the right-hand side; hence, "$\subseteq$" holds true. Conversely, if $\lambda \in \mathbb{R}^m_+, v \in \mathbb{R}^m_+ \backslash \{0\}$, then we define, for each $n \in \mathbb{N}$ and $i \in \{1, \ldots, m\}$,

$$(r_i^n, w_i^n) := \begin{cases} \left(\dfrac{\lambda_i}{v_i}, v_i\right) & \text{if } v_i > 0, \\ (1, v_i) & \text{if } v_i = 0, \lambda_i = 0, \\ \left(n\lambda_i, \dfrac{1}{n}\right) & \text{if } v_i = 0, \lambda_i > 0. \end{cases} \tag{6.33}$$

Then $r^n \in \mathbb{R}^m_+, w^n \in \mathbb{R}^m_+ \backslash \{0\}$ and $\lambda_i = r_i^n w_i^n$. Let η be a point in the right-hand side of (6.32). If there is no $i \in \{1, \ldots, m\}$ satisfying $v_i = 0$ and $\lambda_i > 0$, then $v = w^n$ and $H(\lambda, v) = M(r^n, w^n)$ for every $n \in \mathbb{N}$; hence, $\eta \in H(\lambda, v)$. Next, assume there is some $j \in \{1, \ldots, m\}$ with $v_j = 0$, $\lambda_j > 0$. Since $\eta \in M(r^n, w^n)$ for every $n \in \mathbb{N}$, it follows

$$\begin{aligned} &\sum_{\substack{i:v_i>0 \\ i:v_i=\lambda_i=0}} v_i\eta_i + \sum_{i:v_i=0,\lambda_i>0} \frac{\eta_i}{n} \\ &\quad \geq \sum_{\substack{i:v_i>0 \\ i:v_i=\lambda_i=0}} \inf_{z_i \in \mathbb{R}: \mathbb{E}[\ell(-X_i - z_i)] < +\infty} (v_i z_i + \lambda_i \mathbb{E}[\ell_i(-X_i - z_i)]) \\ &\qquad + \sum_{i:v_i=0,\lambda_i>0} \frac{1}{n} \inf_{z_i \in \mathbb{R}} (z_i + n\lambda_i \mathbb{E}[\ell_i(-X_i - z_i)]) + \inf_{x \in C} \lambda^{\mathsf{T}} x. \end{aligned} \tag{6.34}$$

If $j \in \{1, \ldots, m\}$ such that $v_i = 0, \lambda_j > 0$, then we obtain, for each $n \in \mathbb{N}$,

$$\begin{aligned} -\operatorname{ess\,sup} X_j - n\lambda_j g_j\left(\frac{1}{n\lambda_j}\right) &\leq \inf_{z_j \in \mathbb{R}} (z_j + n\lambda_j \mathbb{E}[\ell_j(-X_j - z_j)]) \\ &\leq -\operatorname{ess\,inf} X_j - n\lambda_j g_j\left(\frac{1}{n\lambda_j}\right). \end{aligned} \tag{6.35}$$

This can be checked by a similar calculation to the one in (6.14). Since g_j is convex and lower semicontinuous, the restriction of g_j to $\operatorname{cl}\operatorname{dom} g_j$ is a continuous function, see Zalinescu (2002, Proposition 2.1.6), so that

$$\lim_{n\to\infty} \frac{1}{n} \inf_{z_j \in \mathbb{R}} (z_j + n\lambda_j \mathbb{E}[\ell_j(-X_j - z_j]) = -\lim_{n\to\infty} \lambda_j g_j\left(\frac{1}{n\lambda_j}\right) = -\lambda_j g_j(0). \tag{6.36}$$

On the other hand,

$$\begin{aligned} \inf_{z_j \in \mathbb{R}} (v_j z_j + \lambda_i \mathbb{E}[\ell_i(-X_j - z_j)]) &= \lambda_j \inf_{z_j \in \mathbb{R}} \mathbb{E}[\ell_j(-X_j - z_j)] \\ &= \lambda_j \inf_{y \in \mathbb{R}} \ell_j(y) = -\lambda_j g_j(0) \end{aligned} \tag{6.37}$$

since ℓ_j is nondecreasing and $X_j \in L^\infty$. Taking the limit in (6.34) as $n \to \infty$, we finally obtain

$$v^{\mathsf{T}}\eta \geq \sum_{i=1}^{m} \inf_{z_i \in \mathbb{R}: \mathbb{E}[\ell_i(-X_i - z_i)] < +\infty} (v_i z_i + \lambda_i \mathbb{E}[\ell_i(-X_i - z_i)]) + \inf_{x \in C} \lambda^{\mathsf{T}} x, \tag{6.38}$$

that is, $\eta \in H(\lambda, v)$. Hence, (6.32) follows. □

Proof of Theorem 3.1. Let $r \in \mathbb{R}^m_+$ and define

$$g_{w,r}(z) := w^{\mathsf{T}}(-z + r \cdot \mathbb{E}[\ell(-X + z)]) \tag{6.39}$$

for each $w \in \mathbb{R}^m_+ \backslash \{0\}$ and $z \in \mathbb{R}^m$ with $\mathbb{E}[\ell(-X - z)] \in \mathbb{R}^m$. Note that

$$\begin{aligned} D_{\ell,r}(X) &= \bigcap_{w \in \mathbb{R}^m_+ \backslash \{0\}} \left\{ \eta \in \mathbb{R}^m \mid w^{\mathsf{T}}\eta \geq \inf_{\substack{z \in \mathbb{R}^m: \\ \mathbb{E}[\ell(-X-z)] \in \mathbb{R}^m}} g_{w,r}(z) + \inf_{x \in C} w^{\mathsf{T}}(r \cdot x) \right\} \\ &= \bigcap_{w \in \mathbb{R}^m_+ \backslash \{0\}} \left\{ \eta \in \mathbb{R}^m \mid w^{\mathsf{T}}\eta \geq w^{\mathsf{T}}\delta_{\ell,r}(X) + \inf_{x \in C} w^{\mathsf{T}}(r \cdot x) \right\}, \end{aligned} \tag{6.40}$$

which follows from Definition 3.2, (3.24), (3.25), and the fact that a closed convex set is the intersection of all of its supporting halfspaces; see Hamel & Löhne (2014, (5.2)). By Theorem 2.1, we have $\delta_{\ell,r}(X) \in \mathbb{R}^m$ if and only if $1 \in \operatorname{dom} g_r$. Hence, $D_{\ell,r}(X) = \mathbb{R}^m$ if and only if $1 \notin \operatorname{dom} g_r$. The result follows directly from Lemma 6.2. □

Proof of Proposition 3.2. If $1 \in \operatorname{dom} g_r$, then $\delta_{\ell,r}(X) \in \mathbb{R}^m$ and the computation in the proof of Theorem 3.1 can be concluded as

$$\begin{aligned} D_{\ell,r}(X) &= \delta_{\ell,r}(X) + \bigcap_{w \in \mathbb{R}^m_+ \backslash \{0\}} \left\{ z \in \mathbb{R}^m \mid w^{\mathsf{T}}z \geq \inf_{x \in C} w^{\mathsf{T}}(r \cdot x) \right\} \\ &= \delta_{\ell,r}(X) + r \cdot C. \end{aligned} \tag{6.41}$$

With this representation, it is easy to check that $D_{\ell,r}$ is a closed convex risk measure since δ_{ℓ_i,r_i} is a lower semicontinuous convex scalar risk measure for each $i \in \{1, \ldots, m\}$. If $1 \notin \operatorname{dom} g_r$, then $\delta_{\ell,r}(X) = -\infty$ and hence $D_{\ell,r}(X) = \mathbb{R}^m$ due to the convention $w^{\mathsf{T}}(-\infty) = -\infty$ in Lemma 6.2. □

Proof of Proposition 3.4. Let $w \in \mathbb{R}^m_+ \backslash \{0\}$ and $M(r, w)$ as in (6.31). For the moment, let us denote by $-\alpha$ the function defined by the right-hand side of (3.32). Using the dual representation of scalar divergence risk measures provided by Theorem 2.1, we have

$$\begin{aligned} &M(r, w) \\ &= \left\{ z \in \mathbb{R}^m \mid w^{\mathsf{T}}z \geq \sum_{i=1}^{m} \sup_{\mathbb{Q}_i \in \mathcal{M}(\mathbb{P})} w_i(\mathbb{E}^{\mathbb{Q}_i}[-X_i] - I_{g_i,r_i}(\mathbb{Q}_i \mid \mathbb{P})) + \inf_{x \in C} w^{\mathsf{T}}(r \cdot x) \right\} \end{aligned}$$

$$= \bigcap_{\mathbb{Q}\in\mathcal{M}_m(\mathbb{P})} \left\{ z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq w^{\mathsf{T}}(\mathbb{E}^{\mathbb{Q}}[-X] - I_{g,r}(\mathbb{Q}\,|\,\mathbb{P})) + \inf_{x\in C} w^{\mathsf{T}}(r\cdot x) \right\}$$

$$= \bigcap_{\mathbb{Q}\in\mathcal{M}_m(\mathbb{P})} (-\alpha(\mathbb{Q}, w) + \mathbb{E}^{\mathbb{Q}}[-X]). \tag{6.42}$$

Hence,

$$D_{\ell,r}(X) = \bigcap_{w\in\mathbb{R}^m_+\setminus\{0\}} M(r,w)$$

$$= \bigcap_{(\mathbb{Q},w)\in\mathcal{M}_m(\mathbb{P})\times(\mathbb{R}^m_+\setminus\{0\})} (-\alpha(\mathbb{Q}, w) + \mathbb{E}^{\mathbb{Q}}[-X]). \tag{6.43}$$

Finally, we show that $-\alpha = -\alpha_{D_{\ell,r}}$. Using (3.29) for $\mathbb{Q} \in \mathcal{M}_m(\mathbb{P})$, $w \in \mathbb{R}^m_+\setminus\{0\}$, we obtain

$$\begin{aligned} -\alpha_{D_{\ell,r}}(\mathbb{Q}, w) &= \operatorname{cl} \bigcup_{X\in L^\infty_m} (\mathbb{E}^{\mathbb{Q}}[X] + \delta_{\ell,r}(X) + r\cdot C + G(w)) \\ &= \left\{ z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq \inf_{X\in L^\infty_m} w^{\mathsf{T}}(\mathbb{E}^{\mathbb{Q}}[X] + \delta_{\ell,r}(X)) + \inf_{x\in C} w^{\mathsf{T}}(r\cdot x) \right\} \\ &= \left\{ z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq -\sum_{i=1}^m w_i I_{g_i,r_i}(\mathbb{Q}_i\,|\,\mathbb{P}) + \inf_{x\in C} w^{\mathsf{T}}(r\cdot x) \right\} \\ &= -\alpha(\mathbb{Q}, w), \end{aligned} \tag{6.44}$$

where the third equality follows from the analogous scalar result established in Proposition 2.4. Hence, $-\alpha_{D_{\ell,r}} = -\alpha$ and (3.32) holds. □

Proof of Proposition 3.5. Using (3.29) for $\mathbb{Q} \in \mathcal{M}_m(\mathbb{P})$, $w \in \mathbb{R}^m_{++}$, we obtain

$$\begin{aligned} &-\alpha_{R_\ell}(\mathbb{Q}, w) \\ &= \operatorname{cl} \bigcup_{X\in L^\infty_m} \left(\mathbb{E}^{\mathbb{Q}}[X] + G(w) + \operatorname{cl} \bigcup_{\substack{z\in\mathbb{R}^m: \\ \mathbb{E}[\ell(-X-z)]\in\mathbb{R}^m}} \{z + \mathbb{R}^m_+ \,|\, 0 \in \mathbb{E}[\ell(-X-z)] + C\} \right) \\ &= \operatorname{cl} \bigcup_{z\in\mathbb{R}^m} \bigcup_{\substack{X\in L^\infty_m: \\ \mathbb{E}[\ell(-X-z)]\in\mathbb{R}^m}} \{z + \mathbb{E}^{\mathbb{Q}}[X] + G(w) \,|\, 0 \in \mathbb{E}[\ell(-X-z)] + C\} \\ &= \operatorname{cl} \bigcup_{X\in L^\infty_m} \{\mathbb{E}^{\mathbb{Q}}[-X] + G(w) \,|\, 0 \in \mathbb{E}[\ell(X)] + C\} \\ &= \inf_{(\mathcal{G}_m,\supseteq)} \{\mathbb{E}^{\mathbb{Q}}[-X] + G(w) \,|\, 0 \in \mathbb{E}[\ell(X)] + C, X \in L^\infty_m\}, \end{aligned} \tag{6.45}$$

where $\mathbb{E}[\ell(X)] + C$ is understood to be $\emptyset$ whenever $\mathbb{E}[\ell(X)] = +\infty$. Next, we compute the optimal value of the dual problem for this convex set-valued minimization

problem. By (6.22), for $X \in L^\infty_m$, $\lambda \in \mathbb{R}^m_+$, $v \in \mathbb{R}^m_+ \backslash \{0\}$, we have $L(X, \lambda, v) = \mathbb{R}^m$ if $v \notin \{sw \,|\, s > 0\}$. Moreover, if $v = sw$ for some $s > 0$, then

$$\begin{aligned} L(X, \lambda, v) &= \mathbb{E}^{\mathbb{Q}}[-X] + G(sw) + \left\{ z \in \mathbb{R}^m \,|\, sw^{\mathsf{T}} z \geq \lambda^{\mathsf{T}} \mathbb{E}[\ell(X)] + \inf_{x \in C} \lambda^{\mathsf{T}} x \right\} \\ &= \left\{ z \in \mathbb{R}^m \,|\, sw^{\mathsf{T}} z \geq sw^{\mathsf{T}} \mathbb{E}^{\mathbb{Q}}[-X] + \lambda^{\mathsf{T}} \mathbb{E}[\ell(X)] + \inf_{x \in C} \lambda^{\mathsf{T}} x \right\}, \end{aligned} \tag{6.46}$$

whenever $\mathbb{E}[\ell(X)] \in \mathbb{R}^m$ and $L(X, \lambda, v) = \emptyset$ otherwise. Observe $G(sw) = G(w)$ for every $s > 0$. Hence,

$$\begin{aligned} H(\lambda, sw) &= \left\{ z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq \inf_{\substack{X \in L^\infty_m : \\ \mathbb{E}[\ell(X)] \in \mathbb{R}^m}} \left(w^{\mathsf{T}} \mathbb{E}^{\mathbb{Q}}[-X] + \frac{1}{s} \lambda^{\mathsf{T}} \mathbb{E}[\ell(X)] \right) + \inf_{x \in C} \frac{1}{s} \lambda^{\mathsf{T}} x \right\} \\ &= H\left(\frac{\lambda}{s}, w \right) \end{aligned} \tag{6.47}$$

for $\lambda \in \mathbb{R}^m_+$, $s > 0$. The optimal value of the dual problem is

$$\sup \left\{ H\left(\frac{\lambda}{s}, w \right) \,\middle|\, s > 0, \lambda \in \mathbb{R}^m_+ \right\} = \sup \{ H(\lambda, w) \,|\, \lambda \in \mathbb{R}^m_+ \}. \tag{6.48}$$

Since $w_i > 0$ for every $i \in \{1, \ldots, m\}$ by assumption, we have

$$\begin{aligned} &H(\lambda, w) \\ &\quad = \left\{ z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq \inf_{X \in L^\infty_m} (w^{\mathsf{T}} \mathbb{E}^{\mathbb{Q}}[-X] + w^{\mathsf{T}} (r \cdot \mathbb{E}[\ell(X)])) + \inf_{x \in C} w^{\mathsf{T}} (r \cdot x) \right\}, \end{aligned} \tag{6.49}$$

where $r_i := \frac{\lambda_i}{w_i}$, $i \in \{1, \ldots, m\}$. Note that

$$\begin{aligned} &\inf_{X \in L^\infty_m} (w^{\mathsf{T}} \mathbb{E}^{\mathbb{Q}}[-X] + w^{\mathsf{T}} (r \cdot \mathbb{E}[\ell(X)])) \\ &\quad = \sum_{i=1}^m w_i \inf_{X_i \in L^\infty} \left(w_i \mathbb{E}\left[-\frac{d\mathbb{Q}_i}{d\mathbb{P}} X_i \right] + r_i \mathbb{E}[\ell_i(X_i)] \right) \\ &\quad = \sum_{i=1}^m w_i \mathbb{E}\left[\inf_{x_i \in \mathbb{R}} \left(-\frac{d\mathbb{Q}_i}{d\mathbb{P}} x_i + r_i \ell_i(x_i) \right) \right] \\ &\quad = w^{\mathsf{T}} I_{g,r}(\mathbb{Q} \,|\, \mathbb{P}). \end{aligned} \tag{6.50}$$

Therefore, the optimal value of the dual problem equals the middle term in (3.33). Note that Slater's condition holds, that is, there exists $\bar{X} \in L^\infty_m$ such that $(\mathbb{E}[\ell(\bar{X})] + C) \cap -\mathbb{R}^m_{++} \neq \emptyset$. This is immediate from the scalar version as in the proof of Proposition 2.4. Hence, the first equality in (3.33) holds by Hamel & Löhne (2014, Theorem 6.6). Since $I_{g,r}(\mathbb{Q} \,|\, \mathbb{P}) \notin \mathbb{R}^m$ if $1 \notin \operatorname{dom} g_r$, we also have the second equality in (3.33). □

6.5. *Proofs of the results in Sec. 4*

Proof of Proposition 4.1. Using the definitions, we have

$$\begin{aligned} R^{\mathrm{ent}}(X) &= \left\{ z \in \mathbb{R}^m \,|\, \exists\, c \in C\ \forall\, i \in \{1,\dots,m\} : \frac{\mathbb{E}[e^{\beta_i(-X_i - z_i)}] - 1}{\beta_i} = -c_i \right\} \\ &= \left\{ z \in \mathbb{R}^m \,|\, \exists\, c \in C\ \forall\, i \in \{1,\dots,m\} : z_i = \frac{1}{\beta_i} \log \frac{\mathbb{E}[e^{-\beta_i X_i}]}{1 - \beta_i c_i},\ 1 > \beta_i c_i \right\} \\ &= \rho^{\mathrm{ent}}(X) + C^{\mathrm{ent}}. \end{aligned} \tag{6.51}$$

□

Proof of Proposition 4.2. For each $i \in \{1,\dots,m\}$, note that

$$\begin{aligned} \delta_{\ell_i, r_i}(X_i) &= \inf_{z_i \in \mathbb{R}} (z_i + r_i \mathbb{E}[\ell_i(-X_i - z_i)]) \\ &= \frac{1}{\beta_i} \log \mathbb{E}[e^{-\beta_i X_i}] + \frac{1}{\beta_i}(1 - r_i + \log r_i) \in \mathbb{R}. \end{aligned} \tag{6.52}$$

The result follows from Proposition 3.2. □

Proof of Lemma 4.1. First, we extend f_w and h_w from $\mathbb{R}^m_{++}$ to $\mathbb{R}^m$ with their original definitions so that we have $\inf_{r \in \mathbb{R}^m_{++}}(f_w(r) + h_w(r)) = \inf_{r \in \mathbb{R}^m}(f_w(r) + h_w(r))$. Note that f_w is a proper, strictly convex, continuous function and has a unique minimum point. Hence, by Rockafellar (1970, Theorem 27.1(d)), f_w has no directions of recession, that is, the recession function $f_w 0^+$ of f_w always takes strictly positive values; see Rockafellar (1970, pp. 66 and 69) for definitions. Besides, h_w is a proper, convex, lower semicontinuous function. If $h_w \equiv +\infty$, then the infimum of $f_w + h_w$ is $+\infty$. Suppose that h_w is a proper function. Since 0 is a boundary point of $-C$, h_w always takes nonnegative values. Hence, the infimum of h_w is finite. By Rockafellar (1970, Theorem 27.1(a), (i)), this implies that the recession function $h_w 0^+$ of h_w always takes nonnegative values. Therefore, $f_w + h_w$ has no directions of recession since $(f_w + h_w)0^+ = f_w 0^+ + h_w 0^+$ by Rockafellar (1970, Theorem 9.3). Hence, by Rockafellar (1970, Theorem 27.1(b), (d)) and the strict convexity of $f_w + h_w$, this function has a unique minimum point $r^w \in \mathbb{R}^m_{++}$ which is determined by the first-order condition

$$\begin{aligned} 0 &\in \partial(f_w + h_w)(r^w) \\ &= \left[\frac{w_i}{\beta_i} - \frac{w_i}{\beta_i r_i^w} \right]_{i=1}^m + \left\{ w \cdot \bar{x} \,|\, \bar{x} \in -C,\ \sup_{x \in -C} w^{\mathsf{T}}(r^w \cdot x) = w^{\mathsf{T}}(r^w \cdot \bar{x}) \right\}, \end{aligned} \tag{6.53}$$

that is,

$$\left[\frac{1}{\beta_i} \left(1 - \frac{1}{r_i^w} \right) \right]_{i=1}^m \in C, \quad \inf_{x \in C} w^{\mathsf{T}}(r^w \cdot x) = \sum_{i=1}^m \frac{w_i r_i^w}{\beta_i} \left(1 - \frac{1}{r_i^w} \right), \tag{6.54}$$

which is the claimed property of r^w. □

Proof of Proposition 4.3. By Lemma 4.1, it is clear that, for each $w \in \mathbb{R}^m_+\backslash\{0\}$, we have $\inf_{r\in\mathbb{R}^m_{++}}(f_w(r) + h_w(r)) = \inf_{r\in\Gamma}(f_w(r) + h_w(r)) = f_w(r^w) + h_w(r^w)$. Hence,

$$\begin{aligned} R^{\mathrm{ent}}(X) &= \rho^{\mathrm{ent}}(X) + \bigcap_{\substack{w\in\mathbb{R}^m_+\backslash\{0\},\\ r\in\mathbb{R}^m_{++}}} \{z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq -(f_w(r) + h_w(r))\} \\ &= \rho^{\mathrm{ent}}(X) + \bigcap_{w\in\mathbb{R}^m_+\backslash\{0\}} \Big\{z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq -\inf_{r\in\Gamma}(f_w(r) + h_w(r))\Big\} \\ &= \bigcap_{r\in\Gamma} D^{\mathrm{ent}}_r(X). \end{aligned} \tag{6.55}$$

Let $w \in \mathbb{R}^m_+\backslash\{0\}$ such that $f_w + h_w$ is proper and let $r \in \mathbb{R}^m_{++}$. Suppose that $D^{\mathrm{ent}}_r(X) \subseteq D^{\mathrm{ent}}_{r^w}(X)$. Then, $-(f_w(r) + h_w(r)) = \inf_{z\in D^{\mathrm{ent}}_r(X)} w^{\mathsf{T}} z \geq \inf_{z\in D^{\mathrm{ent}}_{r^w}(X)} w^{\mathsf{T}} z = -(f_w(r^w) + h_w(r^w))$, that is, $f_w(r) + h_w(r) \leq f_w(r^w) + h_w(r^w)$. By Lemma 4.1, this implies that $r = r^w$. □

Proof of Proposition 4.4. Proposition 3.5 and Lemma 4.1 give

$$\begin{aligned} &-\alpha_{R_\ell}(\mathbb{Q}, w) \\ &= \bigcap_{r\in 1/\mathrm{dom}\, g} \Big\{z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq -w^{\mathsf{T}} I_{g,r}(\mathbb{Q}\,|\,\mathbb{P}) + \inf_{x\in C} w^{\mathsf{T}}(r \cdot x)\Big\} \\ &= -\beta^{-1} \cdot H(\mathbb{Q}\,|\,\mathbb{P}) \\ &\quad + \bigcap_{r\in\mathbb{R}^m_{++}} \left\{z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq \sum_{i=1}^m \frac{w_i}{\beta_i}(1 - r_i + \log r_i) + \inf_{x\in C} w^{\mathsf{T}}(r \cdot x)\right\} \\ &= -\beta^{-1} \cdot H(\mathbb{Q}\,|\,\mathbb{P}) + \bigcap_{r\in\mathbb{R}^m_{++}} \{z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq -(f_w(r) + h_w(r))\} \\ &= -\beta^{-1} \cdot H(\mathbb{Q}\,|\,\mathbb{P}) + \{z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq -(f_w(r^w) + h_w(r^w))\} \\ &= -\beta^{-1} \cdot H(\mathbb{Q}\,|\,\mathbb{P}) \\ &\quad + \left\{z \in \mathbb{R}^m \,|\, w^{\mathsf{T}} z \geq \sum_{i=1}^m \frac{w_i}{\beta_i}(1 - r^w_i + \log r^w_i) + \inf_{x\in C} w^{\mathsf{T}}(r^w \cdot x)\right\}, \end{aligned} \tag{6.56}$$

assuming that h_w is not identically $+\infty$ (otherwise $-\alpha_{R_\ell}(\mathbb{Q}, w) = \mathbb{R}^m$). The passage from the last line to the claimed formula is by (6.54). □

6.6. *Proofs of the results in Sec.* 5

Proof of Proposition 5.1. Clearly, $R^{\mathrm{mar}}(0) \neq \emptyset$ since $0 \in \Lambda_m(0)$ and $R(0) \neq \emptyset$. We prove the monotonicity and translativity of the function $Y \mapsto \tilde{R}(Y) := \bigcup_{X\in\Lambda_m(Y)} R(X)$ first. For monotonicity, consider $Y^1, Y^2 \in L^\infty_d$ with $Y^1 \leq Y^2$.

Let $X \in \Lambda_m(Y^1)$. With $\tilde{Y} := Y^2 - Y^1 \in L^\infty_{d,+}$, it holds

$$\begin{aligned} BX \in Y^1 + \mathcal{K} &= Y^2 - \tilde{Y} + \mathcal{K} \\ &= Y^2 - \sum_{t=0}^{T-1} L^\infty_d(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t) - (\tilde{Y} + L^\infty_d(\mathcal{F}_T, \mathcal{C}_T)) \\ &\subseteq Y^2 - \sum_{t=0}^{T-1} L^\infty_d(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t) - (L^\infty_{d,+} + L^\infty_d(\mathcal{F}_T, \mathcal{C}_T)) \\ &\subseteq Y^2 + \mathcal{K}, \end{aligned} \tag{6.57}$$

where the last inclusion holds since $L^\infty_{d,+} + L^\infty_d(\mathcal{F}_T, \mathcal{C}_T) = L^\infty_d(\mathcal{F}_T, \mathbb{R}^d_+) + L^\infty_d(\mathcal{F}_T, \mathcal{C}_T) = L^\infty_d(\mathcal{F}_T, \mathcal{C}_T)$ due to $\mathcal{C}_T(\omega) \in \mathcal{G}_d$ for every $\omega \in \Omega$. Hence, $X \in \Lambda_m(Y^2)$. Therefore, $\Lambda_m(Y^1) \subseteq \Lambda_m(Y^2)$, which implies $\tilde{R}(Y^1) \subseteq \tilde{R}(Y^2)$. To prove translativity, let $Y \in L^\infty_d$, $z \in \mathbb{R}^m$. For every $X \in L^\infty_m$, it holds

$$\begin{aligned} X \in \Lambda_m(Y + Bz) &\Leftrightarrow BX \in Y + Bz + \mathcal{K} \\ &\Leftrightarrow B(X - z) \in Y + \mathcal{K} \\ &\Leftrightarrow X - z \in \Lambda_m(Y). \end{aligned} \tag{6.58}$$

Hence,

$$\begin{aligned} \tilde{R}(Y + Bz) &= \bigcup_{X \in \Lambda_m(Y + Bz)} R(X) = \bigcup_{X - z \in \Lambda_m(Y)} R(X) \\ &= \bigcup_{X \in \Lambda_m(Y)} R(X + z) = \tilde{R}(Y) - z \end{aligned} \tag{6.59}$$

from which translativity follows. It is easy to check that the last two properties are preserved under the closure and convex hull operators. Hence, R^{mar} is monotone and translative. It is also easy to check that $\tilde{R}$ and R^{mar} are convex since R is convex. Finally, since $\tilde{R}$ has convex values and this property is preserved under the closure operator, (5.6) follows. □

As a preparation for the proof of Theorem 5.1, we establish a link between the notions of market risk measure and set-valued infimal convolution. We begin by introducing two key concepts from (complete lattice-based) set-valued convex analysis, the reader is referred to Hamel (2009) for details.

Definition 6.1 (Hamel 2009, Example 1). Let $\mathcal{Y} \subseteq L^\infty_d$. The *indicator function* of the set $\mathcal{Y}$ is the function $\mathcal{I}^m_{\mathcal{Y}} : L^\infty_d \to \mathcal{G}_m$ defined by

$$\mathcal{I}^m_{\mathcal{Y}}(Y) = \begin{cases} \mathbb{R}^m_+ & \text{if } Y \in \mathcal{Y}, \\ \emptyset & \text{else.} \end{cases} \tag{6.60}$$

Definition 6.2 (Hamel 2009, Sec. 4.4(C)). Let $N \geq 1$ be an integer. For each $n \in \{1, \ldots, N\}$, let $F^n : L_d^\infty \to \mathcal{G}_m$ be a function. The function $\square_{n=1}^N F^n : L_d^\infty \to \mathcal{G}_m$ defined by

$$(\square_{n=1}^N F^n)(Y) = \operatorname{cl}\operatorname{co} \bigcup_{Y^1,\ldots,Y^N \in L_d^\infty} \left\{ \sum_{n=1}^N F^n(Y^n) \,|\, Y^1 + \cdots + Y^N = Y \right\}, \quad (6.61)$$

is called the *infimal convolution* of $F^1, \ldots, F^N$.

Recall the linear operator $B : \mathbb{R}^m \to \mathbb{R}^d$ defined by (5.2): $Bx = (x_1, \ldots, x_m, 0, \ldots, 0)^\mathsf{T}$ for $x \in \mathbb{R}^m$. Its adjoint $B^* : \mathbb{R}^d \to \mathbb{R}^m$ is defined by (5.3): $B^* y = (y_1, \ldots, y_m)^\mathsf{T}$ for $y \in \mathbb{R}^d$.

The next lemma shows that the market risk measure is basically the infimal convolution of the original risk measure and the indicator function of the negative of the set $\mathcal{K}$ of all freely available portfolios defined by (5.1).

Lemma 6.3. *Let $R : L_m^\infty \to \mathcal{G}_m$ be a closed convex risk measure and define $\tilde{R} : L_d^\infty \to \mathcal{G}_m$ by*

$$\tilde{R}(Y) = \begin{cases} R(B^*Y), & \text{if } Y \in B(L_m^\infty), \\ \emptyset, & \text{else}. \end{cases} \quad (6.62)$$

Then, for each $Y \in L_d^\infty$,

$$\begin{aligned} R^{\mathrm{mar}}(Y) &= (\tilde{R} \,\square\, \mathcal{I}^m_{-\mathcal{K}})(Y) \\ &= (\tilde{R} \,\square\, \mathcal{I}^m_{L_d^\infty(\mathcal{F}_0, \mathcal{C}_0 \cap \mathcal{D}_0)} \,\square \cdots \square\, \mathcal{I}^m_{L_d^\infty(\mathcal{F}_T, \mathcal{C}_T \cap \mathcal{D}_T)})(Y). \end{aligned} \quad (6.63)$$

Proof. For each $Y \in L_d^\infty$, we have

$$\begin{aligned} &R^{\mathrm{mar}}(Y) \\ &\quad = \operatorname{cl} \bigcup_{\{X \in L_m^\infty \,|\, BX \in Y + \mathcal{K}\}} R(X) = \operatorname{cl} \bigcup_{U \in Y + \mathcal{K}} \tilde{R}(U) \\ &\quad = \operatorname{cl} \bigcup_{U, U' \in L_d^\infty} \{\tilde{R}(U) + \mathcal{I}^m_{-\mathcal{K}}(U') \,|\, U + U' = Y\} \\ &\quad = \operatorname{cl} \bigcup_{U, U^0, \ldots, U^T \in L_d^\infty} \left\{ \tilde{R}(U) + \sum_{t=0}^T \mathcal{I}^m_{L_d^\infty(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t)}(U^t) \,|\, U + U^0 + \cdots + U^T = Y \right\}. \end{aligned} \quad (6.64)$$

Since each of the functions in the infimal convolution is convex, we can omit the convex hull operator in Definition 6.2; and the result follows. □

By Lemma 6.3, the market risk measure can be formulated as an infimal convolution. As in the scalar theory, the Legendre–Fenchel conjugate of the infimal convolution of finitely many convex functions is the sum of the Legendre–Fenchel

conjugates of these convex functions; see Hamel (2009, Lemma 2). The application of this result is the main step of the proof of Theorem 5.1. For completeness, we begin with the definition of conjugate for set-valued functions.

Definition 6.3 (Hamel 2009, Definition 5). Let $F : L_d^\infty \to \mathcal{G}_m$ be a function. The (*Fenchel*) *conjugate* of F is the function $-F^* : L_d^1 \times (\mathbb{R}_+^m \backslash \{0\})$ defined by

$$-F^*(V,v) = \operatorname{cl} \bigcup_{Y \in L_d^\infty} (F(Y) + \{z \in \mathbb{R}^m \mid v^{\mathsf{T}} z \geq \mathbb{E}[-V^{\mathsf{T}} Y]\}). \tag{6.65}$$

Proof of Theorem 5.1. Since $\operatorname{cl} R^{\mathrm{mar}}$ has closed values, Lemma 6.3 implies that, for each $Y \in L_d^\infty$,

$$(\tilde{R} \,\square\, \mathcal{I}^m_{L_d^\infty(\mathcal{F}_0, \mathcal{C}_0 \cap \mathcal{D}_0)} \,\square \cdots \square\, \mathcal{I}^m_{L_d^\infty(\mathcal{F}_T, \mathcal{C}_T \cap \mathcal{D}_T)})(Y) = R^{\mathrm{mar}}(Y) \subseteq (\operatorname{cl} R^{\mathrm{mar}})(Y). \tag{6.66}$$

By Hamel (2009, Remark 6, Lemma 2), R^{mar} and $\operatorname{cl} R^{\mathrm{mar}}$ have the same conjugate on $L_d^1 \times (\mathbb{R}_+^m \backslash \{0\})$ given by

$$\begin{aligned} &-(\tilde{R} \,\square\, \mathcal{I}^m_{L_d^\infty(\mathcal{F}_0, \mathcal{C}_0 \cap \mathcal{D}_0)} \,\square\, \cdots \,\square\, \mathcal{I}^m_{L_d^\infty(\mathcal{F}_T, \mathcal{C}_T \cap \mathcal{D}_T)})^* \\ &\quad = -\tilde{R}^* + \sum_{t=0}^{T} -(\mathcal{I}^m_{L_d^\infty(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t)})^*. \end{aligned} \tag{6.67}$$

Note that this is the set-valued version of the rule "the conjugate of the infimal convolution of finitely many convex functions is the sum of their conjugates". Let $(V,v) \in L_d^1 \times (\mathbb{R}_+^m \backslash \{0\})$. By Hamel *et al.* (2011, Proposition 6.7) on the conjugate of a risk measure, for every $(V,v) \in L_d^1 \times (\mathbb{R}_+^m \backslash \{0\})$, we have $-(\operatorname{cl}(R^{\mathrm{mar}}(\cdot)))^*(V, v) = \mathbb{R}^m$ unless we have $V \in -L_{d,+}^1$ and $v = \mathbb{E}[-B^* V]$.

Next, we pass from $L_d^1 \times (\mathbb{R}_+^m \backslash \{0\})$ to $\mathcal{W}_{m,d} = \mathcal{M}_d(\mathbb{P}) \times ((\mathbb{R}_+^m \backslash \{0\}) \times \mathbb{R}_+^{d-m})$ using the "change of variables formula" (Hamel *et al.* 2011, Lemma 3.4). One obtains that for every $V \in -L_{d,+}^1$ with $v = \mathbb{E}[-B^* V]$, there exists $(\mathbb{Q}, w) \in \mathcal{W}_{m,d}$ such that, for every $Y \in L_d^\infty$,

$$\{z \in \mathbb{R}^m \mid v^{\mathsf{T}} z \geq \mathbb{E}[(-V)^{\mathsf{T}} Y]\} = B^*((\mathbb{E}^{\mathbb{Q}}[Y] + G(w)) \cap B(\mathbb{R}^m)), \tag{6.68}$$

and conversely, every $(\mathbb{Q}, w) \in \mathcal{W}_{m,d}$ can be obtained by some $V \in -L_{d,+}^1$ with $v = \mathbb{E}[-B^* V]$ such that (6.68) holds for every $Y \in L_d^\infty$. Note that $B^*(\mathbb{R}^d) = \mathbb{R}^m \times \{0 \in \mathbb{R}^{d-m}\}$. For such corresponding pairs (V,v) and $(\mathbb{Q}, w)$, using (6.68), we first observe that

$$\begin{aligned} -\tilde{R}^*(V,v) &= \operatorname{cl} \bigcup_{Y \in L_d^\infty} (\tilde{R}(Y) + \{z \in \mathbb{R}^m \mid v^{\mathsf{T}} z \geq \mathbb{E}[-V^{\mathsf{T}} Y]\}) \\ &= \operatorname{cl} \bigcup_{Y \in L_d^\infty} (\tilde{R}(Y) + B^*((\mathbb{E}^{\mathbb{Q}}[Y] + G(w)) \cap B(\mathbb{R}^m))) \end{aligned}$$

$$= \operatorname{cl} \bigcup_{Y \in B(L^\infty_m)} (R(B^*Y) + \mathbb{E}^{B^*\mathbb{Q}}[B^*Y] + G(B^*w))$$
$$= \operatorname{cl} \bigcup_{X \in L^\infty_m} (R(X) + \mathbb{E}^{B^*\mathbb{Q}}[X] + G(B^*w)) = -\alpha_R(B^*\mathbb{Q}, B^*w). \tag{6.69}$$

Next, let $t \in \mathbb{T}$. For the same pairs (V, v) and $(\mathbb{Q}, w)$, by Definitions 6.1 and 6.3, we have

$$-(\mathcal{I}^m_{L^\infty_d(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t)})^*(V, v) = \operatorname{cl} \bigcup_{U^t \in L^\infty_d(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t)} \{z \in \mathbb{R}^m \mid v^\mathsf{T} z \geq \mathbb{E}[(-V)^\mathsf{T} U^t]\}$$
$$= \operatorname{cl} \bigcup_{U^t \in L^\infty_d(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t)} B^*((\mathbb{E}^{\mathbb{Q}}[U^t] + G(w)) \cap B(\mathbb{R}^m)). \tag{6.70}$$

Finally, note that $\operatorname{cl} R^{\mathrm{mar}}$ is a closed convex set-valued function that is finite at zero by assumption. Hence, by biconjugation for set-valued functions, see (Hamel 2009, Theorem 2), we have

$$(\operatorname{cl} R^{\mathrm{mar}})(Y) = \bigcap_{\substack{V \in -L^1_{d,+}, \\ v = \mathbb{E}[-B^*V]}} [-(\operatorname{cl} R^{\mathrm{mar}})^*(V, v) + \{z \in \mathbb{R}^m \mid v^\mathsf{T} z \geq \mathbb{E}[V^\mathsf{T} Y]\}], \tag{6.71}$$

for every $Y \in L^\infty_d$, and the above calculations allow for a passage to vector probability measures:

$$(\operatorname{cl} R^{\mathrm{mar}})(Y) = \bigcap_{(\mathbb{Q}, w) \in \mathcal{W}_{m,d}} [-\alpha_{\operatorname{cl} R^{\mathrm{mar}}}(\mathbb{Q}, w) + B^*((\mathbb{E}^{\mathbb{Q}}[-Y] + G(w)) \cap B(\mathbb{R}^m))], \tag{6.72}$$

where, for $(\mathbb{Q}, w) \in \mathcal{W}_{m,d}$,

$$-\alpha_{\operatorname{cl} R^{\mathrm{mar}}}(\mathbb{Q}, w) = -\alpha_R(B^*\mathbb{Q}, B^*w)$$
$$+ \sum_{t=0}^{T} \operatorname{cl} \bigcup_{U^t \in L^\infty_d(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t)} B^*((\mathbb{E}^{\mathbb{Q}}[U^t] + G(w)) \cap B(\mathbb{R}^m)). \tag{6.73}$$

□

Proof of Corollary 5.1. Let $(\mathbb{Q}, w) \in \mathcal{W}_{m,d} \backslash \mathcal{W}^{\mathrm{convex}}_{m,d}$. So there exist $t \in \mathbb{T}$ and $A \in \mathcal{F}_t$ such that $\mathbb{P}(A) > 0$ and $w \cdot \mathbb{E}[\frac{d\mathbb{Q}}{d\mathbb{P}} \mid \mathcal{F}_t](\omega) \notin (0^+\mathcal{C}_t(\omega))^+$ for each $\omega \in A$. Using the fact that the effective domain of the support function of a nonempty closed convex set in $\mathbb{R}^d$ is a subset of its recession cone, which is an easy consequence of

Rockafellar (1970, Corollary 14.2.1), we see that $\inf_{y^t \in \mathcal{C}_t(\omega)} (w \cdot \mathbb{E}[\frac{d\mathbb{Q}}{d\mathbb{P}} \,|\, \mathcal{F}_t](\omega))^{\mathsf{T}} y^t = -\infty$ for each $\omega \in A$. Note that

$$\begin{aligned} &\operatorname{cl} \bigcup_{U^t \in L_d^\infty(\mathcal{F}_t, \mathcal{C}_t)} B^*((\mathbb{E}^{\mathbb{Q}}[U^t] + G(w)) \cap B(\mathbb{R}^m)) \\ &\quad = \left\{ z \in \mathbb{R}^m \,|\, w^{\mathsf{T}}(Bz) \geq \inf_{U^t \in L_d^\infty(\mathcal{F}_t, \mathcal{C}_t)} w^{\mathsf{T}} \mathbb{E}^{\mathbb{Q}}[U^t] \right\} \\ &\quad = \left\{ z \in \mathbb{R}^m \,|\, (B^* w)^{\mathsf{T}} z \geq \mathbb{E}\left[\inf_{y^t \in \mathcal{C}_t} \left(w \cdot \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathcal{F}_t \right] \right)^{\mathsf{T}} y^t \right] \right\}, \end{aligned} \tag{6.74}$$

where the last equality is by Rockafellar & Wets (1998, Theorem 14.60). Note that the passage to conditional expectations in the third line is necessary for the application of this theorem. Since $\mathbb{P}(A) > 0$, this implies $\operatorname{cl} \bigcup_{U^t \in L_d^\infty(\mathcal{F}_t, \mathcal{C}_t)} B^*((\mathbb{E}^{\mathbb{Q}}[U^t] + G(w)) \cap B(\mathbb{R}^m)) = \mathbb{R}^m$. By the computation in the proof of Proposition 5.1, it follows that $-\alpha_{\operatorname{cl} R^{\max}}(\mathbb{Q}, w) = \mathbb{R}^m$. □

Proof of Corollary 5.2. Let $t \in \mathbb{T}$. For each $\omega \in \Omega$, we have

$$\inf_{y^t \in \mathcal{C}_t(\omega)} \left(w \cdot \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathcal{F}_t \right](\omega) \right)^{\mathsf{T}} y^t = \begin{cases} 0 & \text{if } w \cdot \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathcal{F}_t \right](\omega) \in (\mathcal{C}_t(\omega))^+, \\ -\infty & \text{else} \end{cases} \tag{6.75}$$

since $\mathcal{C}_t(\omega)$ is a nonempty closed convex cone. Similar to the calculation in the proof of Corollary 5.1, we have

$$\begin{aligned} &\operatorname{cl} \bigcup_{U^t \in L_d^\infty(\mathcal{F}_t, \mathcal{C}_t)} B^*((\mathbb{E}^{\mathbb{Q}}[U^t] + G(w)) \cap B(\mathbb{R}^m)) \\ &\quad = \left\{ z \in \mathbb{R}^m \,|\, (B^* w)^{\mathsf{T}} z \geq \mathbb{E}\left[\inf_{y^t \in \mathcal{C}_t} \left(w \cdot \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathcal{F}_t \right] \right)^{\mathsf{T}} y^t \right] \right\} \end{aligned} \tag{6.76}$$

from which the result follows immediately. □

Proof of Proposition 5.2. For the first part, let $i \in \{1, \ldots, m\}$. From Remark 2.2, recall that $r_i \ell_i(s) = \sup_{y \in \mathbb{R}}(sy - r_i g_i(\frac{y}{r_i})) \geq s - r_i g_i(\frac{1}{r_i})$ for every $s \in \mathbb{R}$. Hence, given $X \in L_m^\infty$,

$$\delta_{\ell_i, r_i}(X_i) = \inf_{y \in \mathbb{R}} (y + r_i \mathbb{E}[\ell_i(-X_i - y)]) \geq -\mathbb{E}[X_i] - r_i g_i\left(\frac{1}{r_i}\right) \tag{6.77}$$

for every $i \in \{1, \ldots, m\}$. Then,

$$
\begin{aligned}
&\inf_{z\in D^{\mathrm{mar}}_{\ell,r}(0)} (B^*\bar{w})^{\mathsf{T}} z \\
&\quad = \inf_{X\in\Lambda_m(0)} \inf_{z\in D_{\ell,r}(X)} (B^*\bar{w})^{\mathsf{T}} z \\
&\quad = \inf_{X\in\Lambda_m(0)} (B^*\bar{w})^{\mathsf{T}} \delta_{\ell,r}(X) + \inf_{x\in C} (B^*\bar{w})^{\mathsf{T}} (r\cdot x) \\
&\quad \geq \inf_{X\in\Lambda_m(0)} (B^*\bar{w})^{\mathsf{T}} \mathbb{E}[-X] - \sum_{i=1}^m \bar{w}_i r_i g_i\left(\frac{1}{r_i}\right) + \inf_{x\in C} (B^*\bar{w})^{\mathsf{T}} (r\cdot x) \\
&\quad = \inf_{X\in\Lambda_m(0)} \bar{w}^{\mathsf{T}} \mathbb{E}[-BX] - \sum_{i=1}^m \bar{w}_i r_i g_i\left(\frac{1}{r_i}\right) + \inf_{x\in C} (B^*\bar{w})^{\mathsf{T}} (r\cdot x) \\
&\quad \geq \inf_{Y\in\mathcal{K}} \bar{w}^{\mathsf{T}} \mathbb{E}[-Y] - \sum_{i=1}^m \bar{w}_i r_i g_i\left(\frac{1}{r_i}\right) + \inf_{x\in C} (B^*\bar{w})^{\mathsf{T}} (r\cdot x) \\
&\quad = \sum_{t=0}^T \inf_{U\in L^\infty_d(\mathcal{F}_t,\mathcal{C}_t\cap\mathcal{D}_t)} \mathbb{E}[\bar{w}^{\mathsf{T}} U] - \sum_{i=1}^m \bar{w}_i r_i g_i\left(\frac{1}{r_i}\right) + \inf_{x\in C} (B^*\bar{w})^{\mathsf{T}} (r\cdot x) \\
&\quad \geq \sum_{t=0}^T \inf_{U\in L^\infty_d(\mathcal{F}_t,\mathcal{C}_t)} \mathbb{E}[\bar{w}^{\mathsf{T}} U] - \sum_{i=1}^m \bar{w}_i r_i g_i\left(\frac{1}{r_i}\right) + \inf_{x\in C} (B^*\bar{w})^{\mathsf{T}} (r\cdot x) =: a,
\end{aligned}
\tag{6.78}
$$

where the first inequality follows from (6.77), the second inequality follows since $\Lambda_m(0) = \{X \in L^\infty_m \mid BX \in \mathcal{K}\}$, and the last inequality follows since $L^\infty_d(\mathcal{F}_t, \mathcal{C}_t \cap \mathcal{D}_t) \subseteq L^\infty_d(\mathcal{F}_t, \mathcal{C}_t)$ for each $t \in \{0,\dots,T\}$. By the same arguments as in the proofs of Corollary 5.1 and Corollary 5.2, the hypotheses guarantee that $a > -\infty$. Hence,

$$
\begin{aligned}
D^{\mathrm{mar}}_{\ell,r}(0) &\subseteq \left\{\eta\in\mathbb{R}^m \mid (B^*\bar{w})^{\mathsf{T}}\eta \geq \inf_{z\in D^{\mathrm{mar}}_{\ell,r}(0)} (B^*\bar{w})^{\mathsf{T}} z\right\} \\
&\subseteq \{\eta\in\mathbb{R}^m \mid (B^*\bar{w})^{\mathsf{T}}\eta \geq a\} \neq \mathbb{R}^m.
\end{aligned}
\tag{6.79}
$$

Note that $L^\infty_m \ni X \mapsto \{\eta \in \mathbb{R}^m \mid (B^*\bar{w})^{\mathsf{T}}\eta \geq a\} \in \mathcal{G}_m$ is a weak*-closed convex function. Hence, the desired finiteness condition follows since Remark 5.3 yields

$$
(\mathrm{cl}\, D^{\mathrm{mar}}_{\ell,r})(0) \subseteq \{\eta\in\mathbb{R}^m \mid (B^*\bar{w})^{\mathsf{T}}\eta \geq a\} \neq \mathbb{R}^m. \tag{6.80}
$$

For the second part, (3.23) yields $R_\ell(X) \subseteq D_{\ell,r}(X)$ for every $X \in L^\infty_m$; hence, by Definition 5.3, $(\mathrm{cl}\, R^{\mathrm{mar}}_\ell)(Y) \subseteq (\mathrm{cl}\, D^{\mathrm{mar}}_{\ell,r})(Y)$ for every $Y \in L^\infty_d$. The result follows now from the previous part. □

Acknowledgments

The authors are grateful to two anonymous referees whose comments were very helpful in improving the paper. The authors would like to thank Zachary Feinstein and Samuel Drapeau as well for useful comments on set-valued entropic risk measures and optimized certainty equivalents.

References

Ç. Ararat & B. Rudloff (2016) Dual representations for systemic risk measures. arXiv:1607.03430.

F. Astic & N. Touzi (2007) No arbitrage conditions and liquidity, *Journal of Mathematical Economics* **43** (6), 692–708.

P. Barrieu & N. El Karoui (2008) Pricing, hedging and optimally designing derivatives via minimization of risk measures. In: *Volume on Indifference Pricing* (R. Carmona, ed.), 77–146. Princeton, New Jersey: Princeton University Press.

I. Ben Tahar (2006) Tail conditional expectation for vector-valued risks, *SFB 649 Discussion Papers* 2006-029, Humboldt University, Collaborative Research Center 649.

I. Ben Tahar & E. Lepinette (2014) Vector-valued coherent risk measure processes, *International Journal of Theoretical and Applied Finance* **17** (2), 1450011.

A. Ben-Tal & M. Teboulle (1986) Expected utility, penalty functions and duality in stochastic nonlinear programming, *Management Science* **32** (11), 1445–1466.

A. Ben-Tal & M. Teboulle (2007) An old-new concept of convex risk measures: The optimized certainty equivalent, *Mathematical Finance* **17** (3), 449–476.

F. Biagini, J.-P. Fouque, M. Frittelli & T. Meyer-Brandis (2015) A unified approach to systemic risk measures via acceptance sets. arXiv:1503.06354.

J. M. Borwein & A. S. Lewis (1992) Partially finite convex programming, Part I: Quasi relative interiors and duality theory, *Mathematical Programming* **57** (1), 15–48.

C. Burgert & L. Rüschendorf (2006) Consistent risk measures for portfolio vectors, *Insurance*: *Mathematics and Economics* **38** (2), 289–297.

L. Campi & M. P. Owen (2011) Multivariate utility maximization with proportional transaction costs, *Finance and Stochastics* **15** (3), 461–499.

I. Cascos & I. Molchanov (2016) Multivariate risk measures: A constructive approach based on selections, *Mathematical Finance* **26** (4), 867–900.

U. Çetin, R. A. Jarrow & P. Protter (2004) Liquidity risk and arbitrage pricing theory, *Finance and Stochastics* **8** (3), 311–341.

U. Çetin & L. C. G. Rogers (2007) Modeling liquidity effects in discrete time, *Mathematical Finance* **17** (1), 15–29.

C. Chen, G. Iyengar & C. Moallemi (2013) An axiomatic approach to systemic risk, *Management Science* **59** (6), 1373–1388.

A. Cherny & M. Kupper (2007) Divergence utilities. SSRN:1023525.

I. Csiszár (1967) On topological properties of f-divergence, *Studia Scientiarum Mathematicarum Hungarica* **2** (1), 329–339.

Z. Feinstein & B. Rudloff (2013) Time consistency of dynamic risk measures in markets with transaction costs, *Quantitative Finance* **13** (9), 1473–1489.

Z. Feinstein & B. Rudloff (2015a) Multi-portfolio time consistency for set-valued convex and coherent risk measures, *Finance and Stochastics* **19** (1), 67–107.

Z. Feinstein & B. Rudloff (2015b) A comparison of techniques for dynamic risk measures with transaction costs. In: *Set Optimization and Applications — The State of the Art* (A. H. Hamel, F. Heyde, A. Löhne, B. Rudloffs, C. Schrage, eds.), 3–41. Berlin, Heidelberg: Springer-Verlag.

Z. Feinstein, B. Rudloff & S. Weber (2015) Measures of systemic risk. Forthcoming in *SIAM Journal on Financial Mathematics*, arXiv:1502.07961.

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

H. Föllmer & A. Schied (2011) Stochastic finance: An introduction in discrete time, third revised and extended edition. Berlin, New York: Walter de Gruyter.

A. H. Hamel (2009) A duality theory for set-valued functions I: Fenchel conjugation theory, *Set-Valued and Variational Analysis* **17** (2), 153–182.
A. H. Hamel & F. Heyde (2010) Duality for set-valued measures of risk, *SIAM Journal on Financial Mathematics* **1** (1), 66–95.
A. H. Hamel, F. Heyde & B. Rudloff (2011) Set-valued risk measures for conical market models, *Mathematics and Financial Economics* **5** (1), 1–28.
A. H. Hamel & A. Löhne (2014) Lagrange duality in set optimization, *Journal of Optimization Theory and Applications* **161** (2), 368–397.
A. H. Hamel, A. Löhne & B. Rudloff (2014) Benson type algorithms for linear vector optimization and applications, *Journal of Global Optimization* **59** (4), 811–836.
A. H. Hamel, B. Rudloff & M. Yankova (2013) Set-valued average value at risk and its computation, *Mathematics and Financial Economics* **7** (2), 229–246.
F. Heyde & A. Löhne (2011) Solution concepts in vector optimization: A fresh look at an old story, *Optimization* **60** (12), 1421–1440.
E. Jouini & H. Kallal (1995) Martingales and arbitrage in securities markets with transaction costs, *Journal of Economic Theory* **66** (1), 178–197.
E. Jouini, M. Meddeb & N. Touzi (2004) Vector-valued coherent risk measures, *Finance and Stochastics* **8** (4), 531–552.
Y. M. Kabanov (1999) Hedging and liquidation under transaction costs in currency markets, *Finance and Stochastics* **3** (2), 237–248.
A. Löhne & B. Rudloff (2014) An algorithm for calculating the set of superhedging portfolios in markets with transaction costs, *International Journal of Theoretical and Applied Finance* **17** (2), 1450012.
T. Pennanen & I. Penner (2010) Hedging of claims with physical delivery under convex transaction costs, *SIAM Journal on Financial Mathematics* **1** (1), 158–178.
R. T. Rockafellar (1970) *Convex Analysis*. Princeton, New Jersey: Princeton University Press.
R. T. Rockafellar & R. J.-B. Wets (1998) *Variational Analysis*. Berlin, Heidelberg: Springer-Verlag.
L. C. G. Rogers & S. Singh (2010) The cost of illiquidity and its effects on hedging, *Mathematical Finance* **20** (4), 597–615.
A. Schied (2007) Optimal investments for risk- and ambiguity-averse preferences: A duality approach, *Finance and Stochastics* **11** (1), 107–129.
M. Sion (1958) On general minimax theorems, *Pacific Journal of Mathematics* **8** (1), 171–176.
S. Weber, W. Anderson, A.-M. Hamm, T. Knispel, M. Liese & T. Salfeld (2013) Liquidity-adjusted risk measures, *Mathematics and Financial Economics* **7** (1), 69–91.
C. Zalinescu (2002) *Convex Analysis in General Vector Spaces*. Singapore: World Scientific.