

Conditional Risk Mappings

Andrzej Ruszczyński* Alexander Shapiro†

February 21, 2004

Abstract

We introduce an axiomatic definition of a conditional convex risk mapping and we derive its properties. In particular, we prove a representation theorem for conditional risk mappings in terms of conditional expectations. We also develop dynamic programming relations for multistage optimization problems involving conditional risk mappings.

Key words: Risk, Convex Analysis, Conjugate Duality, Stochastic Optimization, Dynamic Programming, Multi-stage Programming.

1 Introduction

Models of risk and optimization problems involving these models attracted a considerable attention in recent years. One direction of research associated with an axiomatic approach, was initiated in Artzner, Delbaen, Eber and Heath [1], where the concept of *coherent risk measures* was introduced. Subsequently, this approach was developed by Föllmer and Schied [4], Rockafellar, Uryasev and Zabarankin [11], and Ruszczyński and Shapiro [13]. In the discussion below we follow the general setting and terminology of [13].

We assume that Ω is a measurable space equipped with a sigma algebra \mathcal{F} of subsets of Ω , and that an uncertain outcome is represented by a measurable function $X : \Omega \rightarrow \mathbb{R}$. It is natural to assume, for example if X represents uncertain costs, that the smaller the values of X , the better it is. Of course, our constructions can be adapted to other situations as well.

If we introduce a space \mathcal{X} of measurable functions on Ω , we can talk of a *risk function* as a mapping $\rho : \mathcal{X} \rightarrow \mathbb{R}$ (we can also consider risk functions with values

*Rutgers University, Department of Management Science and Information Systems, Piscataway, NJ 08854, USA, e-mail: rusz@rutcor.rutgers.edu

†School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0205, USA, e-mail: ashapiro@isye.gatech.edu

in the extended real line). In our earlier work [13] we have refined and extended the analysis of [1, 4, 11] and we have derived, from a handful of axioms, fairly general properties of risk functions and of optimization problems involving such functions.

Our objective in this paper is to introduce models of risk in a dynamic setting. Here the main issue is our knowledge at the time when risk is evaluated. We consider, therefore, two sigma algebras $\mathcal{F}_1 \subset \mathcal{F}_2$ of subsets of Ω , with \mathcal{F}_1 representing our knowledge when risk is evaluated, and \mathcal{F}_2 representing all events under consideration. Together with that, we consider two vector spaces \mathcal{X}_1 and \mathcal{X}_2 of functions measurable with respect to \mathcal{F}_1 and \mathcal{F}_2 . A *conditional risk mapping* is defined in section 2 as a convex, monotone and translation equivariant mapping $\rho_{\mathcal{X}_2|\mathcal{X}_1} : \mathcal{X}_2 \rightarrow \mathcal{X}_1$. In section 3 we extend our insights from [13] to derive a duality representation theorem for conditional risk mappings. Section 4 is devoted to the analysis of relations of conditional risk mappings and conditional expectations. In section 5 we consider a sequence of sigma algebras $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_T$ and the corresponding linear spaces \mathcal{X}_t , $t = 1, \dots, T$, of measurable functions, and analyze compositions of risk mappings of the form $\rho_{\mathcal{X}_2|\mathcal{X}_1} \circ \dots \circ \rho_{\mathcal{X}_{t-1}|\mathcal{X}_{t-2}} \circ \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}$. Two practically important examples of conditional risk mappings are thoroughly analyzed in section 6. Finally, section 7 addresses the issue of risk measures for sequences, and develops a dynamic programming equation for associated optimization problems.

2 Axioms of Conditional Risk Mappings

In order to construct dynamic models of risk we need to extend the concept of risk functions. We proceed as follows. Let $\mathcal{F}_1 \subset \mathcal{F}_2$ be sigma algebras of subsets of a set Ω , and $\mathcal{X}_1 \subset \mathcal{X}_2$ be linear spaces of real valued functions $\phi(\omega)$, $\omega \in \Omega$, measurable with respect to \mathcal{F}_1 and \mathcal{F}_2 , respectively.

Definition 1 *We say that a mapping $\rho : \mathcal{X}_2 \rightarrow \mathcal{X}_1$ is a conditional risk mapping if the following properties hold:*

(A1) **Convexity:** *if $t \in [0, 1]$ and $X, Y \in \mathcal{X}_2$, then*

$$t\rho(X) + (1-t)\rho(Y) \succeq \rho[tX + (1-t)Y];$$

(A2) **Monotonicity:** *if $Y \succeq X$, then $\rho(Y) \succeq \rho(X)$;*

(A3) **Translation Equivariance:** *if $Y \in \mathcal{X}_1$ and $X \in \mathcal{X}_2$, then*

$$\rho(X + Y) = \rho(X) + Y.$$

The inequalities in (A1) and (A2) are understood componentwise, i.e., $Y \succeq X$ means that $Y(\omega) \geq X(\omega)$ for every $\omega \in \Omega$. Of course, the above definition depends on the choice of the spaces \mathcal{X}_1 and \mathcal{X}_2 . To emphasize this, we sometimes write $\rho_{\mathcal{X}_2|\mathcal{X}_1}$

for the conditional risk mapping. We show in section 4 that the concept of conditional risk mappings is closely related to the classical notion of conditional expectation. This provides the motivation for the use of the term *conditional* in the name of considered risk mappings.

For $\omega \in \Omega$, we associate with ρ the function

$$\rho_\omega(X) := [\rho(X)](\omega), \quad X \in \mathcal{X}_2. \quad (2.1)$$

Assumptions (A1) and (A2) mean that, for every $\omega \in \Omega$, the function $\rho_\omega : \mathcal{X}_2 \rightarrow \mathbb{R}$ is convex and monotone, respectively, and assumption (A3) implies that $\rho_\omega(X + a) = \rho_\omega(X) + a$ for any $X \in \mathcal{X}_2$ and $a \in \mathbb{R}$. That is, $\rho_\omega(\cdot)$ satisfies the axioms of convex risk functions, as given by Föllmer and Schied [4] and analyzed in our earlier paper [13]. In particular, if the sigma algebra \mathcal{F}_1 is trivial, i.e., $\mathcal{F}_1 = \{\emptyset, \Omega\}$, then any function $X \in \mathcal{X}_1$ is constant over Ω , and hence the space \mathcal{X}_1 can be identified with \mathbb{R} . In that case $\rho(\cdot)$ becomes real valued and assumptions (A1)–(A3) become the axioms of convex (real valued) risk functions.

We assume that with each space \mathcal{X}_i , $i = 1, 2$, is associated a linear space \mathcal{Y}_i of signed finite measures on (Ω, \mathcal{F}_i) such that $\mathcal{Y}_1 \subset \mathcal{Y}_2$, and¹ $\int_\Omega |X| d|\mu| < +\infty$ for every $X \in \mathcal{X}_i$ and $\mu \in \mathcal{Y}_i$. Then we can define the scalar product (bilinear form)

$$\langle \mu, X \rangle := \int_\Omega X(\omega) d\mu(\omega), \quad X \in \mathcal{X}_i, \mu \in \mathcal{Y}_i. \quad (2.2)$$

By $\mathcal{P}_{\mathcal{Y}_i}$ we denote the set of probability measures $\mu \in \mathcal{Y}_i$, i.e., $\mu \in \mathcal{P}_{\mathcal{Y}_i}$ if μ is nonnegative and $\mu(\Omega) = 1$. We assume that \mathcal{X}_i and \mathcal{Y}_i are *paired locally convex* topological vector spaces. That is, \mathcal{X}_i and \mathcal{Y}_i are equipped with respective topologies which make them locally convex topological vector spaces and these topologies are compatible with the scalar product (2.2), i.e., every linear continuous functional on \mathcal{X}_i can be represented in the form $\langle \mu, \cdot \rangle$ for some $\mu \in \mathcal{Y}_i$, and every linear continuous functional on \mathcal{Y}_i can be represented in the form $\langle \cdot, X \rangle$ for some $X \in \mathcal{X}_i$. In particular, we can equip each space \mathcal{X}_i and \mathcal{Y}_i with its weak topology induced by its paired space. This will make \mathcal{X}_i and \mathcal{Y}_i paired locally convex topological vector spaces provided that for any $X \in \mathcal{X}_i \setminus \{0\}$ there exists $\mu \in \mathcal{Y}_i$ such that $\langle \mu, X \rangle \neq 0$, and for any $\mu \in \mathcal{Y}_i \setminus \{0\}$ there exists $X \in \mathcal{X}_i$ such that $\langle \mu, X \rangle \neq 0$.

A natural choice of \mathcal{X}_i , $i = 1, 2$, is the space of all bounded \mathcal{F}_i -measurable functions $X : \Omega \rightarrow \mathbb{R}$. In that case we can take \mathcal{Y}_i to be the space of all signed finite measures on (Ω, \mathcal{F}_i) . Another possible choice is $\mathcal{X}_i := \mathcal{L}_p(\Omega, \mathcal{F}_i, \bar{\mu})$ for some positive (probability) measure $\bar{\mu}$ on (Ω, \mathcal{F}_2) and $p \in [1, +\infty]$. Note that since $\mathcal{F}_1 \subset \mathcal{F}_2$, $\bar{\mu}$ is also a positive measure on (Ω, \mathcal{F}_1) , and hence $\mathcal{X}_1 \subset \mathcal{X}_2$. We can take then \mathcal{Y}_i to be the linear space of measures ν absolutely continuous with respect to $\bar{\mu}$ and with density (Radon–Nikodym derivative) $h = d\nu/d\bar{\mu}$ belonging to the space $\mathcal{L}_q(\Omega, \mathcal{F}_i, \bar{\mu})$, where $q \geq 1$ is

¹For a signed measure μ we denote by $|\mu|$ the corresponding total variation measure, i.e., $|\mu| = \mu^+ + \mu^-$ where $\mu = \mu^+ - \mu^-$ is the Jordan decomposition of μ .

such that $1/p + 1/q = 1$. In that case we identify \mathcal{Y}_i with $\mathcal{L}_q(\Omega, \mathcal{F}_i, \bar{\mu})$. Note that an element $h \in \mathcal{L}_p(\Omega, \mathcal{F}_i, \bar{\mu})$ is a class of functions which are equal each other for almost every (a.e.) $\omega \in \Omega$ with respect to the measure $\bar{\mu}$. The space $\mathcal{X}_i := \mathcal{L}_p(\Omega, \mathcal{F}_i, \bar{\mu})$ is a Banach space and, for $p \in [1, +\infty)$, $\mathcal{Y}_i := \mathcal{L}_q(\Omega, \mathcal{F}_i, \bar{\mu})$ is its dual space of all continuous linear functionals on \mathcal{X}_i . When dealing with Banach spaces it is convenient to equip \mathcal{X}_i and $\mathcal{Y}_i := \mathcal{X}_i^*$ with the strong (norm) and weak* topologies, respectively. If \mathcal{X}_i is a reflexive Banach space, i.e., $\mathcal{X}_i^{**} = \mathcal{X}_i$, then \mathcal{X}_i and \mathcal{X}_i^* , both equipped with strong topologies, form paired spaces.

We assume throughout the paper that the space \mathcal{X}_2 is sufficiently large such that the following assumption holds (recall that for $X \in \mathcal{X}_2$, the notation $X \succeq 0$ means that $X(\omega) \geq 0$ for all $\omega \in \Omega$).

- (C) If $\mu \in \mathcal{Y}_2$ is not nonnegative, then there exists $X \in \mathcal{X}_2$ such that $X \succeq 0$ and $\langle \mu, X \rangle < 0$.

The above condition ensures that the cone of nonnegative valued functions in \mathcal{X}_2 and the cone of nonnegative measures in \mathcal{Y}_2 are dual to each other. We have that a measure μ is not nonnegative if $\mu(A) < 0$ for some $A \in \mathcal{F}_2$. Therefore, condition (C) holds, for example, if the space \mathcal{X}_2 contains all functions $\mathbb{1}_A(\cdot)$, $A \in \mathcal{F}_2$, where $\mathbb{1}_A(\omega) = 1$ for $\omega \in A$ and $\mathbb{1}_A(\omega) = 0$ for $\omega \notin A$. For technical reasons we assume that this property holds also for the space \mathcal{X}_1 .

- (C') For every $B \in \mathcal{F}_1$, the function $\mathbb{1}_B$ belongs to the space \mathcal{X}_1 .

We say that Y is an \mathcal{F}_1 -step function if it can be represented in the form $Y = \sum_{k=1}^K \alpha_k \mathbb{1}_{B_k}$, where $B_k \in \mathcal{F}_1$ and $B_k \cap B_l = \emptyset$, if $k \neq l$. Clearly, if $\alpha_k \geq 0$, then the step function Y is nonnegative. By the above assumption (C') we have that \mathcal{X}_1 contains all \mathcal{F}_1 -step functions, and in particular all constant functions.

It is said that the conditional risk mapping ρ is *positively homogeneous* if

$$\rho(tX) = t\rho(X), \text{ for all } X \in \mathcal{X}_2 \text{ and } t > 0. \quad (2.3)$$

In that case $\rho(0) = 0$, and for any $Y \in \mathcal{X}_1$, we have that

$$\rho(Y) = \rho(0 + Y) = \rho(0) + Y = Y,$$

and hence $\rho[\rho(X)] = \rho(X)$ for any $X \in \mathcal{X}_2$.

3 Conjugate Duality of Conditional Risk Mappings

We say that the mapping ρ is *lower semicontinuous* if, for every $\omega \in \Omega$, the corresponding function $\rho_\omega : \mathcal{X}_2 \rightarrow \mathbb{R}$ is lower semicontinuous (in the considered topology of \mathcal{X}_2). With ρ_ω is associated its conjugate function

$$\rho_\omega^*(\mu) := \sup_{X \in \mathcal{X}_2} \{ \langle \mu, X \rangle - \rho_\omega(X) \}. \quad (3.4)$$

We also use the notation $\rho^*(\mu, \omega)$ for the function $\rho_\omega^*(\mu)$ in order to emphasize that it is a function of two variables, i.e., $\rho^* : \mathcal{Y}_2 \times \Omega \rightarrow \overline{\mathbb{R}}$. It has the following properties: for every $\omega \in \Omega$ the function $\rho^*(\cdot, \omega)$ is convex and lower semicontinuous, and for any $\mu \in \mathcal{Y}_2$ the function $\rho^*(\mu, \cdot)$ is \mathcal{F}_1 -measurable. Note also that, since ρ_ω is real valued, it follows from (3.4) that $\rho^*(\mu, \omega) > -\infty$ for any $(\mu, \omega) \in \mathcal{Y}_2 \times \Omega$.

Recall that by $\mathcal{P}_{\mathcal{Y}_2}$ we denote the set of all probability measures on (Ω, \mathcal{F}_2) which are in \mathcal{Y}_2 . With each $\omega \in \Omega$ we associate a set of probability measures $\mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega) \subset \mathcal{P}_{\mathcal{Y}_2}$ as the set of all $\nu \in \mathcal{P}_{\mathcal{Y}_2}$ such that for every $B \in \mathcal{F}_1$ it holds that

$$\nu(B) = \begin{cases} 1, & \text{if } \omega \in B, \\ 0, & \text{if } \omega \notin B. \end{cases} \quad (3.5)$$

Note that ω is fixed here and B varies in \mathcal{F}_1 . Condition (3.5) simply means that for every ω and every $B \in \mathcal{F}_1$ we know whether B happened or not. In particular, if $\mathcal{F}_1 = \{\emptyset, \Omega\}$, then $\mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega) = \mathcal{P}_{\mathcal{Y}_2}$ for all $\omega \in \Omega$.

We can now formulate the basic duality result for conditional risk mappings.

Theorem 1 *Let $\rho = \rho_{\mathcal{X}_2|\mathcal{X}_1}$ be a lower semicontinuous conditional risk mapping satisfying assumptions (A1)–(A3). Then*

$$\rho_\omega(X) = \sup_{\mu \in \mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)} \{ \langle \mu, X \rangle - \rho^*(\mu, \omega) \}, \quad \omega \in \Omega, \quad X \in \mathcal{X}_2, \quad (3.6)$$

where $\mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)$ is the set of probability measures defined in (3.5), and $\rho^*(\mu, \omega)$ is defined in (3.4). Conversely, suppose that a mapping $\rho : \mathcal{X}_2 \rightarrow \mathcal{X}_1$ can be represented in form (3.6) for some $\rho^* : \mathcal{Y}_2 \times \Omega \rightarrow \overline{\mathbb{R}}$. Then ρ is lower semicontinuous and satisfies conditions (A1)–(A3).

Proof. We have that if assumptions (A1)–(A3) hold, then ρ_ω is a convex risk function. Since ρ_ω is lower semicontinuous, it follows (by applying the Fenchel-Moreau Theorem) that

$$\rho_\omega(X) = \sup_{\mu \in \mathcal{P}_{\mathcal{Y}_2}} \{ \langle \mu, X \rangle - \rho_\omega^*(\mu) \}, \quad X \in \mathcal{X}_2. \quad (3.7)$$

Conversely, if ρ_ω can be represented in form (3.6) for some ρ_ω^* , then ρ is lower semicontinuous and satisfies conditions (A1)–(A2). All these facts can be established by applying *verbatim* the proof of Theorem 2 in [13] to the function ρ_ω . Therefore, the only issue that needs to be clarified is the restriction of $\text{dom } \rho_\omega^*$ to $\mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)$.

Let $\omega \in \Omega$ be fixed and let $\mu_\omega \in \text{dom } \rho_\omega^*$, and hence $\rho_\omega^*(\mu_\omega)$ is finite. It follows from (A3) that for any $Y \in \mathcal{X}_1$ we have

$$\begin{aligned} \rho_\omega^*(\mu_\omega) &= \sup_{X \in \mathcal{X}_2} \{ \langle \mu_\omega, X + Y \rangle - \rho_\omega(X + Y) \} \\ &= \sup_{X \in \mathcal{X}_2} \{ \langle \mu_\omega, X \rangle + \langle \mu_\omega, Y \rangle - \rho_\omega(X) - Y(\omega) \} \\ &= \rho_\omega^*(\mu_\omega) + \langle \mu_\omega, Y \rangle - Y(\omega). \end{aligned}$$

Therefore $\langle \mu_\omega, Y \rangle = Y(\omega)$ for all $Y \in \mathcal{X}_1$. Setting $Y := \mathbb{1}_B$, where $B \in \mathcal{F}_1$, we conclude that $\mu_\omega \in \mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)$. It follows that $\text{dom } \rho_\omega^* \subseteq \mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)$.

To prove the converse we only need to verify assumption (A3). Suppose that (3.6) holds true. Then every $\mu_\omega \in \text{dom } \rho_\omega^*$ is an element of $\mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)$. Let $Y := \sum_{k=1}^K \alpha_k \mathbb{1}_{B_k}$, be an \mathcal{F}_1 -step function. By assumption (C') we have that $Y \in \mathcal{X}_1$. Then

$$\langle \mu_\omega, Y \rangle = \sum_{k=1}^K \alpha_k \mu_\omega(B_k) = Y(\omega).$$

Passing to the limit, we obtain $\langle \mu_\omega, Y \rangle = Y(\omega)$ for every \mathcal{F}_1 -measurable Y . Therefore (3.6) implies that for every $Y \in \mathcal{X}_1$ and all $\omega \in \Omega$ we have

$$\begin{aligned} [\rho(X + Y)](\omega) &= \sup_{\mu \in \mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)} \{ \langle \mu, X + Y \rangle - \rho^*(\mu, \omega) \} \\ &= \sup_{\mu \in \mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)} \{ \langle \mu, X \rangle - \rho^*(\mu, \omega) \} + Y(\omega). \end{aligned}$$

This is identical to (A3). ■

Recall that the space \mathcal{X}_2 is said to be a *lattice* if for any $X_1, X_2 \in \mathcal{X}_2$ the element $X_1 \vee X_2$, defined as

$$[X_1 \vee X_2](\omega) := \max \{ X_1(\omega), X_2(\omega) \}, \quad \omega \in \Omega,$$

belongs to \mathcal{X}_2 . For every $X \in \mathcal{X}_2$ we can then define $|X| \in \mathcal{X}_2$ in a natural way, i.e., $|X|(\omega) = |X(\omega)|$, $\omega \in \Omega$. The space \mathcal{X}_2 is a *Banach lattice* if it is a Banach space and $|X_1| \leq |X_2|$ implies $\|X_1\| \leq \|X_2\|$. For example, every space $\mathcal{X}_2 := \mathcal{L}_p(\Omega, \mathcal{F}_2, \bar{\mu})$, $p \in [1, +\infty]$, is a Banach lattice. We can remark here that the lower semicontinuity of ρ follows from conditions (A1)–(A2), if \mathcal{X}_2 has the structure of a Banach lattice. Note that $[\rho(X)](\cdot)$ is a finite valued function, and hence $\rho_\omega(X)$ is finite for all $\omega \in \Omega$ and all $X \in \mathcal{X}_2$. Direct application of [13, Proposition 1] yields the following result.

Proposition 1 *Suppose that \mathcal{X}_2 is a Banach lattice and $\rho : \mathcal{X}_2 \rightarrow \mathcal{X}_1$ satisfies assumptions (A1) and (A2). Then $\rho_\omega(\cdot)$ is continuous for all $\omega \in \Omega$.*

Clearly, if $\rho : \mathcal{X}_2 \rightarrow \mathcal{X}_1$ is positively homogeneous, then the corresponding function ρ_ω is also positively homogeneous. Therefore, if $\rho = \rho_{\mathcal{X}_2|\mathcal{X}_1}$ is a positively homogeneous, lower semicontinuous, conditional risk mapping, then $\rho^*(\cdot, \omega)$ is the indicator function of a closed convex set $\mathcal{A}(\omega) \subset \mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)$, and hence

$$\rho_\omega(X) = \sup_{\mu \in \mathcal{A}(\omega)} \langle \mu, X \rangle, \quad \omega \in \Omega, \quad X \in \mathcal{X}_2. \quad (3.8)$$

We view $\omega \mapsto \mathcal{A}(\omega)$ as a multifunction from Ω into the set $\mathcal{P}_{\mathcal{Y}_2}$ of probability measures, on (Ω, \mathcal{F}_2) , which are included in \mathcal{Y}_2 . Formula (3.8) extends to conditional risk mappings the risk envelope representation derived in [1, 11, 13].

The property of positive homogeneity can be strengthened substantially.

Proposition 2 *Let $\rho = \rho_{\mathcal{X}_2|\mathcal{X}_1}$ be a positively homogeneous, lower semicontinuous conditional risk mapping. Suppose that for any $B \in \mathcal{F}_1$ and $X \in \mathcal{X}_2$, it holds that $\mathbb{1}_B X \in \mathcal{X}_2$. Then for any nonnegative \mathcal{F}_1 -step function Y and $X \in \mathcal{X}_2$, we have that $YX \in \mathcal{X}_2$ and*

$$\rho(YX) = Y\rho(X). \quad (3.9)$$

Proof. Consider a set $B \in \mathcal{F}_1$ and any $X \in \mathcal{X}_2$. It follows from (3.8) that

$$\rho_\omega(X) = \sup_{\mu \in \mathcal{A}(\omega)} \{ \langle \mu, \mathbb{1}_B X \rangle + \langle \mu, \mathbb{1}_{\Omega \setminus B} X \rangle \}, \quad \omega \in \Omega. \quad (3.10)$$

If $\omega \in B$ then (3.5) implies that $\mu(\Omega \setminus B) = 0$ for all $\mu \in \mathcal{A}(\omega)$, and the second term in the right hand side of (3.10) vanishes, and hence

$$\rho_\omega(X) = \sup_{\mu \in \mathcal{A}(\omega)} \langle \mu, \mathbb{1}_B X \rangle = \rho_\omega(\mathbb{1}_B X). \quad (3.11)$$

By a similar argument we also obtain that $\rho_\omega(\mathbb{1}_B X) = 0$ for $\omega \notin B$. Thus

$$\rho(\mathbb{1}_B X) = \mathbb{1}_B \rho(X). \quad (3.12)$$

Consider now a nonnegative \mathcal{F}_1 -step function $Y := \sum_{k=1}^K \alpha_k \mathbb{1}_{B_k}$. Then $YX = \sum_{k=1}^K \alpha_k \mathbb{1}_{B_k} X \in \mathcal{X}_2$, and it follows from (3.12) that for $\omega \in B_k$ the following chain of equations holds true:

$$\rho_\omega(YX) = \mathbb{1}_{B_k}(\omega) \rho_\omega(YX) = \rho_\omega(\mathbb{1}_{B_k} YX) = \rho_\omega(\alpha_k \mathbb{1}_{B_k} X).$$

Using positive homogeneity and (3.12) again, we obtain that

$$\rho_\omega(YX) = \alpha_k \rho_\omega(\mathbb{1}_{B_k} X) = \alpha_k \rho_\omega(X), \text{ if } \omega \in B_k.$$

This means that

$$\rho_\omega(YX) = \sum_{k=1}^K \alpha_k \mathbb{1}_{B_k}(\omega) \rho_\omega(X) = Y(\omega) \rho_\omega(X), \quad \omega \in \Omega.$$

This completes the proof. ■

Remark 1 In order to pass in (3.9) from step functions to general functions $Y \in \mathcal{X}_1$ we need some additional assumptions about the spaces involved. For example, consider $\mathcal{X}_i := \mathcal{L}_p(\Omega, \mathcal{F}_i, \bar{\mu})$, $i = 1, 2$, where $\bar{\mu}$ is a probability measure on (Ω, \mathcal{F}_2) and $p \in [1, +\infty)$. Then for any \mathcal{F}_1 -step function Y and $X \in \mathcal{X}_2$, we have that $YX \in \mathcal{X}_2$. Moreover, the set of \mathcal{F}_1 -step functions is dense in \mathcal{X}_1 . Hence by “passing to the limit operation” and using Proposition 1 we obtain that (3.9) holds for any $Y \in \mathcal{X}_1$ and $X \in \mathcal{X}_2$, provided that $YX \in \mathcal{X}_2$.

4 Conditional Expectation Representation

In this section we discuss relations between conditional risk mappings and conditional expectations.

Definition 2 *We say that a multifunction $\mathcal{M} : \Omega \rightrightarrows \mathcal{P}_{\mathcal{Y}_2}$ is weakly* \mathcal{F}_1 -measurable if for every $X \in \mathcal{X}_2$ the multifunction $\mathcal{M}_X : \Omega \rightrightarrows \mathbb{R}$, defined as*

$$\mathcal{M}_X(\omega) := \{ \langle \mu, X \rangle : \mu \in \mathcal{M}(\omega) \},$$

is \mathcal{F}_1 -measurable. We say that a selection $\mu_\omega \in \mathcal{M}(\omega)$ is weakly \mathcal{F}_1 -measurable if for every $X \in \mathcal{X}_2$ the function $\omega \mapsto \langle \mu_\omega, X \rangle$ is \mathcal{F}_1 -measurable.*

The multifunction $\omega \mapsto \mathcal{A}(\omega)$, associated with representation (3.8), is weakly* \mathcal{F}_1 -measurable. Indeed,

$$\mathcal{A}_X(\omega) = [-\rho_\omega(-X), \rho_\omega(X)],$$

and hence \mathcal{F}_1 -measurability of $\mathcal{A}_X(\cdot)$ follows from the fact that $\rho(X) \in \mathcal{X}_1$. In the sequel, whenever speaking about measurability of multifunctions and their selections, we shall mean weak* measurability.

By Theorem 1 we have that, for every $\omega \in \Omega$, any measure $\nu \in \mathcal{A}(\omega)$ satisfies condition (3.5). Therefore, if $\mu_\omega = \mu(\omega)$ is a selection of $\mathcal{A}(\omega)$, then

$$[\mu(\cdot)](A) = \mathbb{1}_A(\cdot), \text{ for all } A \in \mathcal{F}_1. \quad (4.1)$$

Moreover, if $\mu(\omega)$ is weakly* \mathcal{F}_1 -measurable, then $[\mu(\cdot)](A)$ is \mathcal{F}_1 -measurable.

Example 1 Consider $\rho(X) := \mathbb{E}[X|\mathcal{F}_1]$, $X \in \mathcal{X}_2$, where the conditional expectation is taken with respect to a probability measure $\bar{\mu}$ on (Ω, \mathcal{F}_2) . It is assumed here that this conditional expectation is well defined for every $X \in \mathcal{X}_2$, and the space \mathcal{X}_1 is large enough such that it contains $\mathbb{E}[X|\mathcal{F}_1]$ for all $X \in \mathcal{X}_2$. Note that the function $\mathbb{E}[X|\mathcal{F}_1](\cdot)$ is defined up to a set of $\bar{\mu}$ -measure zero, i.e., two versions of $\mathbb{E}[X|\mathcal{F}_1](\cdot)$ can be different on a set of $\bar{\mu}$ -measure zero. The conditional expectation mapping ρ satisfies assumptions (A1)–(A3) and is a linear mapping. Representation (3.8) holds with $\mathcal{A}(\omega) = \{\mu(\omega)\}$ being a singleton and $\mu_\omega = \mu(\omega)$ being a probability measure on (Ω, \mathcal{F}_2) . By the definition of the conditional expectation we have that $\mathbb{E}[X|\mathcal{F}_1]$ is \mathcal{F}_1 -measurable, and hence μ_ω is weakly* \mathcal{F}_1 -measurable. Considering $X = \mathbb{1}_A$ for $A \in \mathcal{F}_2$, we see that

$$\mu_\omega(A) = \mathbb{E}[\mathbb{1}_A|\mathcal{F}_1](\omega) = [\bar{\mu}(A|\mathcal{F}_1)](\omega), \quad (4.2)$$

that is, $\mu(\cdot)$ is the conditional probability of $\bar{\mu}$ with respect to \mathcal{F}_1 (see, e.g., Billingsley [3, pp. 430–431]). Clearly it satisfies (3.5).

Let us note that the family of conditional risk mappings is closed under the operation of taking maximum. That is, let $\rho^\nu = \rho_{\mathcal{X}_2|\mathcal{X}_1}^\nu$, $\nu \in \mathcal{I}$, be a family of conditional risk mappings satisfying assumptions (A1)–(A3). Suppose, further, that for every $X \in \mathcal{X}_2$ the function

$$[\rho(X)](\cdot) := \sup_{\nu \in \mathcal{I}} [\rho^\nu(X)](\cdot) \quad (4.3)$$

belongs to the space \mathcal{X}_1 , and hence ρ maps \mathcal{X}_2 into \mathcal{X}_1 . It is then straightforward to verify that the max-function ρ also satisfies assumptions (A1)–(A3). Moreover, if ρ^ν , $\nu \in \mathcal{I}$, are lower semicontinuous, then ρ is also lower semicontinuous. In particular, let $\rho^\nu(X) := \mathbb{E}_\nu[X|\mathcal{F}_1]$, $\nu \in \mathcal{I}$, where \mathcal{I} is a family of probability measures on (Ω, \mathcal{F}_2) . Suppose that the corresponding max-function ρ is well defined, i.e., ρ maps \mathcal{X}_2 into \mathcal{X}_1 . Then ρ is a lower semicontinuous positively homogeneous conditional risk mapping. We show now that, under certain regularity conditions, the converse is also true, i.e., a positively homogeneous conditional risk mapping can be represented as maximum of a family of conditional expectations.

Let $\mu_\omega \in \mathcal{A}(\omega)$ be a weakly* \mathcal{F}_1 -measurable selection. We associate with $\mu(\omega) = \mu_\omega$ the following operator:

$$[\mathbb{Q}_\mu(\nu)](A) := \int_{\Omega} \mu_\omega(A) d\nu(\omega), \quad \nu \in \mathcal{Y}_2, \quad A \in \mathcal{F}_2. \quad (4.4)$$

It is not difficult to verify that, for every $\nu \in \mathcal{Y}_2$, the right hand side of (4.4) defines a finite signed measure on (Ω, \mathcal{F}_2) . In particular, if A_1, A_2, \dots , is a sequence of disjoint \mathcal{F}_2 -measurable sets and $A := A_1 \cup A_2 \cup \dots$, then by the Lebesgue Theorem,

$$\int_{\Omega} \mu_\omega(A) d\nu(\omega) = \int_{\Omega} \sum_i \mu_\omega(A_i) d\nu(\omega) = \sum_i \int_{\Omega} \mu_\omega(A_i) d\nu(\omega).$$

We make the following assumption for a considered, weakly* \mathcal{F}_1 -measurable, selection $\mu_\omega \in \mathcal{A}(\omega)$: (i) if $\nu \in \mathcal{Y}_2$, then $\mathbb{Q}_\mu(\nu) \in \mathcal{Y}_2$. The above assumption is not implied by the definition (4.4) and should be verified in a particular application. It is straightforward to verify that if ν is a probability measure, then $\mathbb{Q}_\mu(\nu)$ is also a probability measure. Therefore, under the above assumption (i), the operator \mathbb{Q}_μ maps $\mathcal{P}_{\mathcal{Y}_2}$ into itself.

A probability measure $\bar{\nu} \in \mathcal{P}_{\mathcal{Y}_2}$ is called a *fixed point* of \mathbb{Q}_μ , if $\mathbb{Q}_\mu(\bar{\nu}) = \bar{\nu}$. This means that

$$\int_{\Omega} \mu_\omega(A) d\bar{\nu}(\omega) = \bar{\nu}(A) \text{ for any } A \in \mathcal{F}_2. \quad (4.5)$$

Proposition 3 *Let $\mu_\omega \in \mathcal{A}(\omega)$ be a weakly* \mathcal{F}_1 -measurable selection and $\bar{\nu} \in \mathcal{P}_{\mathcal{Y}_2}$ be a fixed point of the operator \mathbb{Q}_μ , defined in (4.4). Then $\mu(\cdot)$ is the conditional probability of $\bar{\nu}$ with respect to \mathcal{F}_1 .*

Proof. Let $A \in \mathcal{F}_2$. By the weak* \mathcal{F}_1 -measurability of $\mu(\cdot)$, we have that $[\mu(\cdot)](A)$ is \mathcal{F}_1 -measurable. Therefore, we only need to verify that the following equation holds (see, e.g., Billingsley [3, p. 430]):

$$\int_S \mu_\omega(A) d\bar{\nu}(\omega) = \bar{\nu}(A \cap S), \text{ for all } S \in \mathcal{F}_1 \text{ and } A \in \mathcal{F}_2. \quad (4.6)$$

We have that $\mu_\omega(A) = \mu_w(A \cap S) + \mu_w(A \setminus S)$ and $\mu_w(A \setminus S) \leq \mu_w(\Omega \setminus S)$. Since $\Omega \setminus S \in \mathcal{F}_1$, it follows by (4.1) that $\mu_w(\Omega \setminus S) = 0$ for all $\omega \in S$. Hence $\mu_w(A \setminus S) = 0$, for any $\omega \in S$. Consequently,

$$\int_S \mu_\omega(A) d\bar{\nu}(\omega) = \int_\Omega \mu_\omega(A \cap S) d\bar{\nu}(\omega).$$

Substituting this to (4.5) we obtain (4.6). ■

There are classical results ensuring existence of a fixed point. For example, we can use Kakutani's Fixed Point Theorem. Recall that \mathcal{X}_2 and \mathcal{Y}_2 are paired locally convex topological vector spaces. We have that the set $\mathcal{P}_{\mathcal{Y}_2}$ is convex. Let us make the following assumption related to the operator \mathbb{Q}_μ , defined in (4.4), associated with a weakly* \mathcal{F}_1 -measurable selection $\mu_\omega \in \mathcal{A}(\omega)$.

- (K) The set $\mathcal{P}_{\mathcal{Y}_2}$ is compact and, for every weakly* \mathcal{F}_1 -measurable selection $\mu_\omega \in \mathcal{A}(\omega)$, the operator \mathbb{Q}_μ maps $\mathcal{P}_{\mathcal{Y}_2}$ into itself and is continuous (or, more generally, has a closed graph) in the considered topology of the space \mathcal{Y}_2 .

By Kakutani's Theorem we have that, under the above assumption (K), the operator \mathbb{Q}_μ has a fixed point $\bar{\nu} \in \mathcal{P}_{\mathcal{Y}_2}$. Assumption (K) holds, for example, if $\mathcal{X}_2 = \mathcal{L}_p(\Omega, \mathcal{F}_2, \bar{\mu})$ and $\mathcal{Y}_2 = \mathcal{L}_q(\Omega, \mathcal{F}_2, \bar{\mu})$, with $p \in (1, +\infty)$, and $\mathcal{X}_2, \mathcal{Y}_2$ are equipped with their respective weak topologies. Another case where assumption (K) holds, by Prohorov's Theorem, if Ω is a compact metric space, \mathcal{F}_2 is its Borel sigma algebra and \mathcal{Y}_2 is equipped with the corresponding weak topology. Note that although assumption (K) ensures the existence of a fixed point, it does not imply that such a fixed point is unique.

Corollary 1 *Suppose that representation (3.8) holds, $\mathcal{A}(\omega) = \{\mu_\omega\}$ is a singleton for every $\omega \in \Omega$, and assumption (K) is satisfied. Then there exists $\bar{\nu} \in \mathcal{P}_{\mathcal{Y}_2}$, satisfying equation (4.5), such that μ_ω is the conditional probability of $\bar{\nu}$ with respect to \mathcal{F}_1 , and $\rho(\cdot) = \mathbb{E}_{\bar{\nu}}[\cdot | \mathcal{F}_1]$.*

Recall that \mathcal{X}_i and \mathcal{Y}_i are assumed to be paired locally convex topological vector spaces. It is said that \mathcal{X}_i is *separable* if it has a countable dense subset.

Lemma 1 Suppose that the space \mathcal{X}_2 is separable and the representation (3.8) holds. Then there exists a countable family μ_ω^i , $i \in \mathbb{N}$, of weakly* \mathcal{F}_1 -measurable selections of $\mathcal{A}(\omega)$ such that

$$[\rho(X)](\omega) = \sup_{i \in \mathbb{N}} \langle \mu_\omega^i, X \rangle \quad (4.7)$$

for all $X \in \mathcal{X}_2$ and $\omega \in \Omega$.

Proof. Let $\varepsilon_k \downarrow 0$ be a sequence of positive numbers and $\{X_n\}_{n \in \mathbb{N}}$ be a dense subset of \mathcal{X}_2 . For every $k, n \in \mathbb{N}$ consider the multifunction

$$M_{k,n}(\omega) := \{ \nu \in \mathcal{A}(\omega) : \langle \nu, X_n \rangle \geq [\rho(X_n)](\omega) - \varepsilon_k \}.$$

This multifunction is weakly* \mathcal{F}_1 -measurable and nonempty valued. Since \mathcal{X}_2 is separable, the multifunction $M_{k,n}(\cdot)$ admits a weakly* \mathcal{F}_1 -measurable selection $\mu^{k,n}(\cdot)$ (see [5]). By the definition of $M_{k,n}$ we have then that

$$\langle \mu^{k,n}(\omega), X_n \rangle \geq [\rho(X_n)](\omega) - \varepsilon_k$$

for all $k, n \in \mathbb{N}$ and $\omega \in \Omega$. Since $\rho_\omega(\cdot)$ is lower semicontinuous for every $\omega \in \Omega$, it follows that

$$\sup_{k,n} \langle \mu^{k,n}(\omega), X \rangle \geq [\rho(X)](\omega), \quad \omega \in \Omega.$$

Because of (3.8) we also have that

$$[\rho(X)](\omega) \geq \sup_{k,n} \langle \mu^{k,n}(\omega), X \rangle, \quad \omega \in \Omega.$$

Thus representation (4.7) follows with $\mu_\omega^i := \mu^{k,n}(\omega)$, $i = (k, n) \in \mathbb{N} \times \mathbb{N}$. ■

Remark 2 Under the assumptions of Lemma 1, we can also write the following representation

$$[\rho(X)](\omega) = \sup_{\mu(\cdot) \in \mathcal{A}(\cdot)} \langle \mu(\omega), X \rangle, \quad (4.8)$$

where the supremum is taken over all weakly* \mathcal{F}_1 -measurable selections $\mu(\omega) \in \mathcal{A}(\omega)$.

Theorem 2 Let $\rho = \rho_{\mathcal{X}_2|\mathcal{X}_1}$ be a positively homogeneous, lower semicontinuous, conditional risk mapping. Suppose that the space \mathcal{X}_2 is separable and assumption (K) is satisfied. Then there exists a countable family $\nu^i \in \mathcal{P}_{\mathcal{Y}_2}$, $i \in \mathbb{N}$, of probability measures such that

$$\rho_\omega(\cdot) = \sup_{i \in \mathbb{N}} \mathbb{E}_{\nu^i} [\cdot | \mathcal{F}_1](\omega), \quad \omega \in \Omega. \quad (4.9)$$

Proof. By Theorem 1 we have here that representation (3.8) holds. The assertion then follows by Lemma 1 together with Proposition 3 and Kakutani's Fixed Point Theorem. ■

Remark 3 Assuming that the representation (4.9) holds, we have that for any $Y \in \mathcal{X}_1$, $Y \succeq 0$, and $X \in \mathcal{X}_2$ such that $YX \in \mathcal{X}_2$,

$$\rho_\omega(YX) = \sup_{i \in \mathbb{N}} \mathbb{E}_{\nu^i}[YX|\mathcal{F}_1](\omega) = Y(\omega) \sup_{i \in \mathbb{N}} \mathbb{E}_{\nu^i}[X|\mathcal{F}_1](\omega) = Y(\omega)\rho_\omega(X), \quad \omega \in \Omega.$$

That is, under the assumptions of Theorem 2, the result of Proposition 2 (i.e., equation (3.9)) holds for a general function $Y \in \mathcal{X}_1$ (compare with Remark 1).

5 Iterated Risk Mappings

Let $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3$ be sigma algebras, $\mathcal{X}_1 \subset \mathcal{X}_2 \subset \mathcal{X}_3$ be respective spaces of measurable functions, with dual spaces $\mathcal{Y}_1 \subset \mathcal{Y}_2 \subset \mathcal{Y}_3$, and let $\rho_{\mathcal{X}_3|\mathcal{X}_2} : \mathcal{X}_3 \rightarrow \mathcal{X}_2$ and $\rho_{\mathcal{X}_2|\mathcal{X}_1} : \mathcal{X}_2 \rightarrow \mathcal{X}_1$ be conditional risk mappings. (For any inclusion like $\mathcal{X}_2 \subset \mathcal{X}_3$, we assume that the topology of \mathcal{X}_2 is induced by the topology of \mathcal{X}_3 .) Then it can be easily verified that the composite mapping $\rho_{\mathcal{X}_3|\mathcal{X}_1} : \mathcal{X}_3 \rightarrow \mathcal{X}_1$, defined by

$$\rho_{\mathcal{X}_3|\mathcal{X}_1} := \rho_{\mathcal{X}_2|\mathcal{X}_1} \circ \rho_{\mathcal{X}_3|\mathcal{X}_2}, \quad (5.1)$$

is also a conditional risk mapping.

Suppose that both conditional risk mappings at the right hand side of (5.1) are positively homogeneous and lower semicontinuous. We have then

$$[\rho_{\mathcal{X}_3|\mathcal{X}_2}(X)](\tilde{\omega}) = \sup_{\mu_2 \in \mathcal{A}_2(\tilde{\omega})} \langle \mu_2, X \rangle, \quad X \in \mathcal{X}_3, \tilde{\omega} \in \Omega, \quad (5.2)$$

$$[\rho_{\mathcal{X}_2|\mathcal{X}_1}(Y)](\omega) = \sup_{\mu_1 \in \mathcal{A}_1(\omega)} \langle \mu_1, Y \rangle, \quad Y \in \mathcal{X}_2, \omega \in \Omega, \quad (5.3)$$

with the multifunctions $\mathcal{A}_2 : \Omega \rightrightarrows \mathcal{P}_{\mathcal{Y}_3}$ and $\mathcal{A}_1 : \Omega \rightrightarrows \mathcal{P}_{\mathcal{Y}_2}$ having closed convex values and weakly* measurable with respect to \mathcal{F}_2 and \mathcal{F}_1 , correspondingly. In order to analyze composition (5.1) it is convenient to consider weakly* measurable selections $\mu_i(\cdot)$ of the multifunctions $\mathcal{A}_i(\cdot)$, $i = 1, 2$.

Proposition 4 *Suppose that the space \mathcal{X}_3 is separable, and $\rho_{\mathcal{X}_2|\mathcal{X}_1}$ and $\rho_{\mathcal{X}_3|\mathcal{X}_2}$ are positively homogeneous, lower semicontinuous and satisfy conditions (A1)–(A3). Then the conditional risk mapping $\rho_{\mathcal{X}_3|\mathcal{X}_1}$ can be represented in the form*

$$[\rho_{\mathcal{X}_3|\mathcal{X}_1}(X)](\omega) = \sup_{\mu_1 \in \mathcal{A}_1(\omega)} \sup_{\mu_2(\cdot) \in \mathcal{A}_2(\cdot)} \int_{\Omega} \langle \mu_2(\tilde{\omega}), X \rangle d\mu_1(\tilde{\omega}), \quad (5.4)$$

where the second sup operation at the right hand side of (5.4) is taken with respect to weakly* \mathcal{F}_2 -measurable selections $\mu_2(\tilde{\omega}) \in \mathcal{A}_2(\tilde{\omega})$.

Proof. By (5.2) and (5.3) we have that, for every $\omega \in \Omega$,

$$[\rho_{\mathcal{X}_3|\mathcal{X}_1}(X)](\omega) = \sup_{\mu_1 \in \mathcal{A}_1(\omega)} \int_{\Omega} \sup_{\mu_2 \in \mathcal{A}_2(\tilde{\omega})} \langle \mu_2, X \rangle d\mu_1(\tilde{\omega}).$$

By Lemma 1 (see Remark 2), we also have that

$$\sup_{\mu_2 \in \mathcal{A}_2(\tilde{\omega})} \langle \mu_2, X \rangle = \sup_{\mu_2(\cdot) \in \mathcal{A}_2(\cdot)} \langle \mu_2(\tilde{\omega}), X \rangle, \quad (5.5)$$

where the supremum in the right hand side of (5.5) is taken over all weakly* \mathcal{F}_2 -measurable selections $\mu_2(\tilde{\omega}) \in \mathcal{A}_2(\tilde{\omega})$. Consequently

$$[\rho_{\mathcal{X}_3|\mathcal{X}_1}(X)](\omega) = \sup_{\mu_1 \in \mathcal{A}_1(\omega)} \int_{\Omega} \sup_{\mu_2(\cdot) \in \mathcal{A}_2(\cdot)} \langle \mu_2(\tilde{\omega}), X \rangle d\mu_1(\tilde{\omega}). \quad (5.6)$$

Similarly to the proof of Lemma 1, we can now interchange the integral and ‘sup’ operators at the right hand side of (5.6), and hence (5.4) follows. ■

Remark 4 By Lemma 1 we have that actually it suffices to take the second supremum at the right hand side of (5.4) with respect to a *countable* number of weakly* \mathcal{F}_2 -measurable selections $\mu_2(\tilde{\omega}) \in \mathcal{A}_2(\tilde{\omega})$.

Representation (5.4) means that $\rho_{\mathcal{X}_3|\mathcal{X}_1}$ can be written in form (3.8) with the set $\mathcal{A}(\omega)$ is formed by *all* measures $\mu \in \mathcal{Y}_3$ representable in the form

$$\mu(S) = \int_{\Omega} [\mu_2(\tilde{\omega})](S) d\mu_1(\tilde{\omega}), \quad S \in \mathcal{F}_3, \quad (5.7)$$

where $\mu_2(\cdot) \in \mathcal{A}_2(\cdot)$ is a weakly* \mathcal{F}_2 -measurable selection and $\mu_1 \in \mathcal{A}_1(\omega)$. We denote the multifunction \mathcal{A} by $\mathcal{A}_1 \circ \mathcal{A}_2$.

Consider now a sequence of sigma algebras (a filtration) $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_T$, with $\mathcal{F}_1 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. We define linear (locally convex topological vector) spaces $\mathcal{X}_1 \subset \dots \subset \mathcal{X}_T$ of real valued functions on Ω such that all functions in \mathcal{X}_t are \mathcal{F}_t -measurable. We also introduce the corresponding paired spaces $\mathcal{Y}_1 \subset \dots \subset \mathcal{Y}_T$ of measures, $t = 1, \dots, T$. Let $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$, $t = 2, \dots, T$, be conditional risk mappings. Note that since $\mathcal{F}_1 = \{\emptyset, \Omega\}$, the space \mathcal{X}_1 is formed by constant over Ω functions and can be identified with \mathbb{R} , and hence $\rho_{\mathcal{X}_2|\mathcal{X}_1}$ is an (unconditional) risk function.

With the above sequence of conditional risk mappings we associate the following (unconditional) risk functions

$$\rho_t := \rho_{\mathcal{X}_2|\mathcal{X}_1} \circ \dots \circ \rho_{\mathcal{X}_{t-1}|\mathcal{X}_{t-2}} \circ \rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}, \quad t = 2, \dots, T. \quad (5.8)$$

The recursive application of Proposition 4 renders the following result.

Theorem 3 Let $\rho_{\mathcal{X}_{t+1}|\mathcal{X}_t}$, $t = 1, \dots, T-1$, be positively homogeneous, lower semi-continuous, conditional risk mappings. Suppose that the spaces \mathcal{X}_t , $t = 2, \dots, T$, are separable. Then for every $X \in \mathcal{X}_t$, $t = 2, \dots, T$,

$$\rho_t(X) = \sup_{\mu \in \mathcal{A}_1 \circ \dots \circ \mathcal{A}_{t-1}} \langle \mu, X \rangle, \quad (5.9)$$

where each $\mathcal{A}_\tau : \Omega \rightrightarrows \mathcal{P}_{\mathcal{Y}_{\tau+1}}$ is weakly* \mathcal{F}_τ -measurable and such that

$$[\rho_{\mathcal{X}_{\tau+1}|\mathcal{X}_\tau}(X)](\omega) = \sup_{\mu \in \mathcal{A}_\tau(\omega)} \langle \mu, X \rangle. \quad (5.10)$$

Note that we always have $(\mathcal{A}_1 \circ \mathcal{A}_2) \circ \mathcal{A}_3 = \mathcal{A}_1 \circ (\mathcal{A}_2 \circ \mathcal{A}_3)$ and therefore there is no ambiguity in the notation $\mathcal{A}_1 \circ \dots \circ \mathcal{A}_{t-1}$.

6 Examples of Conditional Risk Mappings

In this section we discuss some examples of conditional risk mappings which can be considered as natural extensions of the corresponding examples of (real valued) risk functions (see [13]). We use the framework and notation of section 2, and take $\bar{\mu}$ to be a probability measure on (Ω, \mathcal{F}_2) . Unless stated otherwise, all expectations and probability statements in this section are made with respect to $\bar{\mu}$.

Example 2 Let $\mathcal{X}_i := \mathcal{L}_p(\Omega, \mathcal{F}_i, \bar{\mu})$ and $\mathcal{Y}_i := \mathcal{L}_q(\Omega, \mathcal{F}_i, \bar{\mu})$, $i = 1, 2$, for some $p \in [1, +\infty)$. Consider

$$\rho(X) := \mathbb{E}[X|\mathcal{F}_1] + c \sigma_p(X|\mathcal{F}_1), \quad X \in \mathcal{X}_2, \quad (6.1)$$

where $c \geq 0$ and $\sigma_p(\cdot|\mathcal{F}_1)$ is the conditional upper semi-deviation:

$$\sigma_p(X|\mathcal{F}_1) := \left(\mathbb{E} \left[(X - \mathbb{E}[X|\mathcal{F}_1])_+^p \mid \mathcal{F}_1 \right] \right)^{1/p}. \quad (6.2)$$

If the sigma algebra \mathcal{F}_1 is trivial, then $\mathbb{E}[\cdot|\mathcal{F}_1] = \mathbb{E}[\cdot]$ and $\sigma_p(X|\mathcal{F}_1)$ becomes the upper semi-deviation of X of order p . Thus ρ is the conditional counterpart of the mean-semideviation models of [6, 7].

Let us show that for $c \in [0, 1]$, the above mapping ρ satisfies assumptions (A1)–(A3). Assumption (A3) can be verified directly. That is, if $Y \in \mathcal{X}_1$ and $X \in \mathcal{X}_2$, then

$$\begin{aligned} \rho(X + Y) &= \mathbb{E}[X + Y|\mathcal{F}_1] + c \left(\mathbb{E} \left[(X + Y - \mathbb{E}[X + Y|\mathcal{F}_1])_+^p \mid \mathcal{F}_1 \right] \right)^{1/p} \\ &= \mathbb{E}[X|\mathcal{F}_1] + Y + c \left(\mathbb{E} \left[(X - \mathbb{E}[X|\mathcal{F}_1])_+^p \mid \mathcal{F}_1 \right] \right)^{1/p} = \rho(X) + Y. \end{aligned}$$

In order to verify assumptions (A1) and (A2) consider function ρ_ω defined in (2.1). For $\omega \in \Omega$ we can write

$$\mathbb{E}[\cdot|\mathcal{F}_1](\omega) = \mathbb{E}_{\mu_\omega}[\cdot], \quad (6.3)$$

where $\mu(\omega) = \mu_\omega$ is the conditional probability of $\bar{\mu}$ with respect to \mathcal{F}_1 (see Example 1). Therefore, for any $X \in \mathcal{X}_2$ and $\omega \in \Omega$,

$$\rho_\omega(X) = \mathbb{E}_{\mu_\omega}[X] + c \left(\mathbb{E}_{\mu_\omega} \left[(X - \mathbb{E}_{\mu_\omega}[X])_+^p \right] \right)^{1/p}. \quad (6.4)$$

We have that $\mu_\omega \in \mathcal{P}_{\mathcal{Y}_2|\mathcal{F}_1}(\omega)$ and its (conditional probability) density $f_\omega = d\mu_\omega/d\bar{\mu}$ has the following properties: $f_\omega \in \mathcal{Y}_2$, $f_\omega \geq 0$, for any $A \in \mathcal{F}_2$, the function $\omega \mapsto \int_A f_\omega d\bar{\mu}$ is \mathcal{F}_1 -measurable and, moreover, for any $B \in \mathcal{F}_1$, the following equality holds

$$\int_B \int_A f_\omega(\tilde{\omega}) d\bar{\mu}(\tilde{\omega}) d\bar{\mu}(\omega) = \bar{\mu}(A \cap B).$$

We see that for a fixed ω the function $\rho_\omega(X)$ is identical with the risk function analyzed in [13, Example 2]; the conditional measure μ_ω plays the role of the probability measure. It follows from the analysis in [13] that, for $c \in [0, 1]$, the function $\rho_\omega(\cdot)$ satisfies assumptions (A1) and (A2). Moreover, the representation

$$\rho_\omega(X) = \sup_{\gamma \in \mathcal{A}^*} \int_\Omega \gamma X d\mu_\omega$$

holds with

$$\mathcal{A}^* = \left\{ \gamma = 1 + h - \int_\Omega h d\mu_\omega : \int_\Omega h^q d\mu_\omega \leq c^q, h \succeq 0 \right\}.$$

Since $d\mu_\omega = f_\omega d\bar{\mu}$, we conclude that the representation (3.8) follows with

$$\mathcal{A}(\omega) = \left\{ g \in \mathcal{Y}_2 : g = f_\omega(1 + h - \mathbb{E}[f_\omega h]), h \in cB_q(\omega), h \succeq 0 \right\}, \quad (6.5)$$

where

$$B_q(\omega) := \{h \in \mathcal{Y}_2 : \mathbb{E}[h^q f_\omega] \leq 1\}.$$

Example 3 Let $\mathcal{X}_i := \mathcal{L}_1(\Omega, \mathcal{F}_i, \bar{\mu})$ and $\mathcal{Y}_i := \mathcal{L}_\infty(\Omega, \mathcal{F}_i, \bar{\mu})$, $i = 1, 2$. For constants $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$, consider

$$\rho(X) := \mathbb{E}[X|\mathcal{F}_1] + \Phi(X|\mathcal{F}_1), \quad X \in \mathcal{X}_2, \quad (6.6)$$

where

$$[\Phi(X|\mathcal{F}_1)](\omega) := \inf_{Z \in \mathcal{X}_1} \mathbb{E}\{\varepsilon_1[Z - X]_+ + \varepsilon_2[X - Z]_+ | \mathcal{F}_1\}(\omega). \quad (6.7)$$

It is straightforward to verify that assumption (A3) holds here. Indeed, for $X \in \mathcal{X}_2$ and $Y \in \mathcal{X}_1$ we have

$$[\Phi(X + Y|\mathcal{F}_1)](\omega) = \inf_{Z \in \mathcal{X}_1} \mathbb{E}\{\varepsilon_1[(Z - Y) - X]_+ + \varepsilon_2[X - (Z - Y)]_+ | \mathcal{F}_1\}(\omega).$$

By making change of variables $Z \mapsto Z - Y$, we obtain that $\Phi(X + Y|\mathcal{F}_1) = \Phi(X|\mathcal{F}_1)$, and hence assumption (A3) follows.

Because of (6.3), we can write, as in the previous example, that

$$\rho_\omega(X) = \mathbb{E}_{\mu_\omega}[X] + \inf_{Z \in \mathcal{X}_1} \mathbb{E}_{\mu_\omega} \{ \varepsilon_1[Z - X]_+ + \varepsilon_2[X - Z]_+ \} \quad (6.8)$$

$$= \mathbb{E}[f_\omega X] + \inf_{Z \in \mathcal{X}_1} \mathbb{E} \{ \varepsilon_1[f_\omega Z - f_\omega X]_+ + \varepsilon_2[f_\omega X - f_\omega Z]_+ \}, \quad (6.9)$$

where $f_\omega = d\mu_\omega/d\bar{\mu}$ is the conditional density. We can continue now in a way similar to the analysis of Example 3 in [13]. We have that

$$\mathbb{E} \{ \varepsilon_1[f_\omega Z - f_\omega X]_+ + \varepsilon_2[f_\omega X - f_\omega Z]_+ \} = \sup_{h \in \mathcal{M}} \mathbb{E} [h(f_\omega X - f_\omega Z)], \quad (6.10)$$

where

$$\mathcal{M} := \{ h \in \mathcal{Y}_2 : -\varepsilon_1 \leq h(\omega) \leq \varepsilon_2, \text{ a.e. } \omega \in \Omega \}. \quad (6.11)$$

By substituting the right hand side of (6.10) into (6.9), we obtain

$$\rho_\omega(X) = \mathbb{E}[f_\omega X] + \inf_{Z \in \mathcal{X}_1} \sup_{h \in \mathcal{M}} \mathbb{E} [h(f_\omega X - f_\omega Z)].$$

Since the set \mathcal{M} is compact in the weak* topology of \mathcal{Y}_2 , we can interchange the ‘inf’ and ‘sup’ operators in the right hand side of the above equation. Also we have that

$$\inf_{Z \in \mathcal{X}_1} \mathbb{E} [h f_\omega Z] = \inf_{Z \in \mathcal{X}_1} \mathbb{E} [Z \mathbb{E} [h f_\omega | \mathcal{F}_1]] = \begin{cases} 0, & \text{if } \mathbb{E} [h f_\omega | \mathcal{F}_1] = 0, \\ -\infty, & \text{otherwise.} \end{cases}$$

We obtain that

$$\rho_\omega(X) = \mathbb{E}[f_\omega X] + \sup \{ \mathbb{E} [h f_\omega X] : h \in \mathcal{M}, \mathbb{E} [h f_\omega | \mathcal{F}_1] = 0 \}.$$

It follows that for $\varepsilon_1 \in (0, 1]$ and $\varepsilon_2 > 0$, assumptions (A1) and (A2) are satisfied, and representation (3.8) holds with

$$\mathcal{A}(\omega) = \{ g \in \mathcal{Y}_2 : g = (1 + h)f_\omega, h \in \mathcal{M}, \mathbb{E} [h f_\omega | \mathcal{F}_1] = 0 \}, \quad (6.12)$$

where \mathcal{M} is defined in (6.11).

Since

$$\varepsilon_1[Z - X]_+ + \varepsilon_2[X - Z]_+ = \varepsilon_1(Z + (1 - p)^{-1}[X - Z]_+ - X),$$

where $p := \varepsilon_2/(\varepsilon_1 + \varepsilon_2)$, we have that

$$\rho(X) = (1 - \varepsilon_1)\mathbb{E}[X | \mathcal{F}_1] + \varepsilon_1 \text{CV@R}_{\mathcal{X}_2 | \mathcal{X}_1}[X], \quad (6.13)$$

where

$$\text{CV@R}_{\mathcal{X}_2 | \mathcal{X}_1}[X](\omega) := \inf_{Z \in \mathcal{X}_1} \mathbb{E} \{ Z + (1 - p)^{-1}[X - Z]_+ | \mathcal{F}_1 \} (\omega). \quad (6.14)$$

Clearly, for $\varepsilon_1 = 1$ we have that $\rho(\cdot) = \text{CV@R}_{\mathcal{X}_2 | \mathcal{X}_1}[\cdot]$. By the above analysis we obtain that for $p \in (0, 1)$, $\text{CV@R}_{\mathcal{X}_2 | \mathcal{X}_1}[\cdot]$ is a positively homogeneous, continuous risk mapping. If $\mathcal{F}_1 = \{\emptyset, \Omega\}$, then $\text{CV@R}_{\mathcal{X}_2 | \mathcal{X}_1}[\cdot]$ becomes the Conditional Value at Risk function analyzed in [10, 11, 14]. For a nontrivial \mathcal{F}_1 , the measure $\text{CV@R}_{\mathcal{X}_2 | \mathcal{X}_1}$ was analyzed in [9].

7 Multistage risk optimization problems

In order to construct risk models for multistage decision problems, we first introduce recursive risk models for sequences.

As in section 5, consider a sequence of sigma algebras $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_T$, with $\mathcal{F}_1 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$, and let $\mathcal{X}_1 \subset \dots \subset \mathcal{X}_T$ be a corresponding sequence of linear spaces of \mathcal{F}_t -measurable functions, $t = 1, \dots, T$. Let $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}} : \mathcal{X}_t \rightarrow \mathcal{X}_{t-1}$ be conditional risk mappings. Denote $\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_T$ and $X := (X_1, X_2, \dots, X_T)$, where $X_t \in \mathcal{X}_t$, $t = 1, \dots, T$, and define a function $\rho : \mathcal{X} \rightarrow \mathbb{R}$ as follows:

$$\begin{aligned} \rho(X) = X_1 + \rho_{\mathcal{X}_2|\mathcal{X}_1} \Big[X_2 + \rho_{\mathcal{X}_3|\mathcal{X}_2} \Big(X_3 + \dots \\ \dots + \rho_{\mathcal{X}_{T-1}|\mathcal{X}_{T-2}} [X_{T-1} + \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}(X_T)] \Big) \Big]. \end{aligned} \quad (7.1)$$

Note that since $\mathcal{F}_1 = \{\emptyset, \Omega\}$, the space \mathcal{X}_1 can be identified with \mathbb{R} , and hence $\rho(X)$ is real valued. Note also that by assumption (A3), we have

$$X_{T-1} + \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}(X_T) = \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}(X_{T-1} + X_T).$$

Applying this formula for $t = T, T-1, \dots, 2$ we obtain the equation:

$$\rho(X) = \rho_T(X_1 + \dots + X_T), \quad (7.2)$$

where, similarly to (5.8),

$$\rho_t := \rho_{\mathcal{X}_2|\mathcal{X}_1} \circ \dots \circ \rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}, \quad t = 2, \dots, T.$$

Since each conditional risk mapping $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$ satisfies (A1)–(A3), it follows that the function ρ_T satisfies (A1)–(A3) as well. Moreover, if conditional risk mappings $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$ are positively homogeneous, then ρ is positively homogeneous. Assuming further that the spaces \mathcal{X}_t are separable and $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$ are lower semicontinuous, we obtain by Theorem 3 that the following representation holds true

$$\rho(X) = \sup_{\mu \in \mathcal{A}} \mathbb{E}_\mu[X_1 + \dots + X_T], \quad (7.3)$$

where the set $\mathcal{A} := \mathcal{A}_1 \circ \dots \circ \mathcal{A}_{T-1}$ is given by the composition of the multifunctions $\mathcal{A}_\tau : \Omega \rightrightarrows \mathcal{P}_{\mathcal{Y}_{\tau+1}}$, $\tau = 1, \dots, T-1$, defined in equation (5.10) of Theorem 3.

It may be of interest to discuss the difference between our approach and a construction in Artzner et al [2]. In [2] an adapted sequence $\{X_t\}$, $t = 1, \dots, T$, is viewed as a measurable function on a new measurable space (Ω', \mathcal{F}') , with $\Omega' = \Omega \times \{1, \dots, T\}$, and with the sigma-field \mathcal{F}' generated by sets of form $B_t \times \{t\}$, for all $B_t \in \mathcal{F}_t$ and $t = 1, \dots, T$. Then representation (7.3), for some set \mathcal{A} , can be derived from axioms of *coherent risk measures* of [1]. In our setting these axioms correspond to the assumptions (A1)–(A3) for the trivial sigma algebra $\mathcal{F}_1 = \{\Omega', \emptyset\}$, and to the positive

homogeneity of the (unconditional) risk function $\rho(X)$. Our approach is via axioms of conditional risk mappings, which allows for a specific analysis of the structure of the set \mathcal{A} .

In applications, we frequently deal with random outcomes $X_t \in \mathcal{X}_t$ resulting from decisions z_t in some stochastic system. In order to model this situation, we introduce linear spaces \mathcal{Z}_t of \mathcal{F}_t -measurable functions² $Z_t : \Omega \rightarrow \mathbb{R}^{n_t}$ and consider functions $f_t : \mathbb{R}^{n_t} \times \Omega \rightarrow \mathbb{R}$, $t = 1, \dots, T$. With functions f_t we associate mappings $F_t : \mathcal{Z}_t \rightarrow \mathcal{X}_t$ defined as follows

$$[F_t(Z_t)](\omega) := f_t(Z_t(\omega), \omega), \quad Z_t \in \mathcal{Z}_t, \omega \in \Omega.$$

We assume that the functions $f_t(z_t, \omega)$ are *random lower semicontinuous*³, and that the mappings F_t are well defined, i.e., for every $Z_t \in \mathcal{Z}_t$ the function $f_t(Z_t(\cdot), \cdot)$ belongs to the space \mathcal{X}_t , $t = 1, \dots, T$. We say that the mapping F_t is convex if $[F_t(\cdot)](\omega)$ is convex for all $\omega \in \Omega$. Then for every conditional risk mapping $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$, satisfying (A1)–(A3), we have that $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}(F_t(\cdot))$ is convex in the sense that the function $[\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}(F_t(\cdot))](\omega)$ is convex for every $\omega \in \Omega$. This follows by assumptions (A1) and (A2) and can be shown in the same way as in [13, Proposition 2].

Let $\mathcal{Z} = \mathcal{Z}_1 \times \mathcal{Z}_2 \times \dots \times \mathcal{Z}_T$, and let $F : \mathcal{Z} \rightarrow \mathcal{X}$ be defined as

$$F(Z) := (F_1(Z_1), \dots, F_T(Z_T)).$$

With the risk function ρ , defined in (7.1), and the mapping F we can associate the function

$$\begin{aligned} \psi(Z) := \rho(F(Z)) = & F_1(Z_1) + \rho_{\mathcal{X}_2|\mathcal{X}_1} \left[F_2(Z_2) + \rho_{\mathcal{X}_3|\mathcal{X}_2} \left(F_3(Z_3) + \dots \right. \right. \\ & \left. \left. \dots + \rho_{\mathcal{X}_{T-1}|\mathcal{X}_{T-2}} [F_{T-1}(Z_{T-1}) + \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}(F_T(Z_T))] \right) \right]. \end{aligned}$$

As discussed above, by the recursive application of [13, Proposition 2], it can be easily shown that $\psi(\cdot)$ is a convex function. Also by using (7.2) and (7.3) we can write

$$\begin{aligned} \psi(Z) = & \rho_T(F_1(Z_1) + F_2(Z_2) + \dots + F_T(Z_T)) \\ = & \sup_{\mu \in \mathcal{A}} \int_{\Omega} [f_1(Z_1) + f_2(Z_2(\omega), \omega) \dots + f_T(Z_T(\omega), \omega)] d\mu(\omega). \end{aligned}$$

Suppose that we are given \mathcal{F}_t -measurable multifunctions $\mathcal{G}_t : \mathbb{R}^{n_1 + \dots + n_{t-1}} \times \Omega \rightrightarrows \mathbb{R}^{n_t}$. We define the set

$$\mathcal{S} := \{Z \in \mathcal{Z} : Z_t(\omega) \in \mathcal{G}_t(Z_1(\omega), \dots, Z_{t-1}(\omega), \omega), \omega \in \Omega, t = 1, \dots, T\},$$

²Note that since \mathcal{F}_1 is trivial, the space \mathcal{Z}_1 coincides with \mathbb{R}^{n_1} and elements $Z_1 \in \mathcal{Z}_1$ are n_1 -dimensional vectors.

³Random lower semicontinuous functions are also called *normal integrands* (see Definition 14.27 in [12, p.676]).

and consider the problem

$$\text{Min}_{Z \in \mathcal{S}} \psi(Z). \quad (7.4)$$

We shall derive a dynamic programming equation for this problem.

In order to accomplish that, we need some mild technical assumptions. We assume that the spaces \mathcal{X}_t are *solid* in the sense that for every two elements $\underline{X}, \overline{X} \in \mathcal{X}_t$ and every \mathcal{F}_t -measurable function X_t satisfying $\underline{X}(\cdot) \leq X_t(\cdot) \leq \overline{X}(\cdot)$, the function X_t is an element of \mathcal{X}_t . For example, the spaces $\mathcal{L}_p(\Omega, \mathcal{F}_t, \bar{\mu})$, $p \in [1, +\infty]$ are solid. Furthermore, we assume that there exist elements $\underline{X}_t \in \mathcal{X}_t$ such that for all $Z \in \mathcal{S}$ we have $F_t(Z_t) \succeq \underline{X}_t$, $t = 1, \dots, T$.

We introduce the notation

$$Z_{[1,t]} := (Z_1, \dots, Z_t), \quad t = 1, \dots, T.$$

If $Z_t(\cdot) \equiv z_t$ is a constant function, we write $(Z_{[1,t-1]}, z_t)$ for $Z_{[1,t]}$. Problem (7.4) can be written in a more explicit form as follows:

$$\text{Min}_{Z_1 \in \mathcal{G}_1} \text{Min}_{Z_2(\cdot) \in \mathcal{G}_2(Z_1, \cdot)} \dots \text{Min}_{Z_T(\cdot) \in \mathcal{G}_T(Z_{[1,T-1]}(\cdot), \cdot)} \rho_T [F_1(Z_1) + F_2(Z_2) + \dots + F_T(Z_T)]. \quad (7.5)$$

Consider the minimization with respect to Z_T in the above problem. Since the function ρ_T is monotone in the sense of (A2), and Z_T is required to be only \mathcal{F}_T -measurable, we can carry out this minimization inside the argument. We obtain the following formulation

$$\begin{aligned} & \text{Min}_{Z_1 \in \mathcal{G}_1} \text{Min}_{Z_2(\cdot) \in \mathcal{G}_2(Z_1, \cdot)} \dots \text{Min}_{Z_{T-1}(\cdot) \in \mathcal{G}_{T-1}(Z_{[1,T-2]}(\cdot), \cdot)} \rho_T [F_1(Z_1) + F_2(Z_2) + \dots \\ & \quad \dots + F_{T-1}(Z_{T-1}) + \inf_{z_T \in \mathcal{G}_T(Z_{[1,T-1]}(\cdot), \cdot)} f_T(z_T, \cdot)]. \end{aligned}$$

Note that because of \mathcal{F}_t -measurability of \mathcal{G}_T and random lower semicontinuity of $f_T(\cdot, \cdot)$, the pointwise infimum $\inf_{z_T \in \mathcal{G}_T(Z_{[1,T-1]}(\omega), \omega)} f_T(z_T, \omega)$ is \mathcal{F}_T -measurable (e.g., Rockafellar and Wets [12, Theorem 14.37]), and bounded because of the assumption that \mathcal{X}_T is solid. Consequently this infimum (as a function of ω) is an element of \mathcal{X}_T .

Using the fact that

$$\rho_t := \rho_{t-1} \circ \rho_{\mathcal{X}_t | \mathcal{X}_{t-1}}, \quad t = 2, \dots, T,$$

we can re-write the last problem as follows:

$$\begin{aligned} & \text{Min}_{Z_1 \in \mathcal{G}_1} \text{Min}_{Z_2(\cdot) \in \mathcal{G}_2(Z_1, \cdot)} \dots \text{Min}_{Z_{T-1}(\cdot) \in \mathcal{G}_{T-1}(Z_{[1,T-2]}(\cdot), \cdot)} \rho_{T-1} \left[F_1(Z_1) + F_2(Z_2) + \dots \right. \\ & \quad \left. \dots + F_{T-1}(Z_{T-1}) + \rho_{\mathcal{X}_T | \mathcal{X}_{T-1}} \left(\inf_{z_T \in \mathcal{G}_T(Z_{[1,T-1]}(\cdot), \cdot)} f_T(z_T, \cdot) \right) \right]. \end{aligned} \quad (7.6)$$

Define

$$\psi_T(Z_{[1,T-1]}) := \rho_{\mathcal{X}_T | \mathcal{X}_{T-1}} \left[\inf_{z_T \in \mathcal{G}_T(Z_{[1,T-1]}(\cdot), \cdot)} f_T(z_T, \cdot) \right]. \quad (7.7)$$

Note that, as it was discussed above, the pointwise infimum in the right hand side of (7.7) is an element of \mathcal{X}_T . Therefore we can evaluate $\rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}$ at this element. Thus $\psi_T(Z_{[1,T-1]}) \in \mathcal{X}_{T-1}$ and it can appear in the argument of ρ_{T-1} in (7.6). We obtain the formulation

$$\begin{aligned} \text{Min}_{Z_1 \in \mathcal{G}_1} \text{Min}_{Z_2(\cdot) \in \mathcal{G}_2(Z_1, \cdot)} \dots \text{Min}_{Z_{T-1}(\cdot) \in \mathcal{G}_{T-1}(Z_{[1,T-2]}(\cdot, \cdot))} \rho_{T-1} [F_1(Z_1) + F_2(Z_2) + \dots \\ \dots + F_{T-1}(Z_{T-1}) + \psi_T(Z_{[1,T-1]})]. \end{aligned}$$

Formally setting $\psi_{T+1}(Z_{[1,T]}) \equiv 0$, and proceeding in this way for $T-1, T-2, \dots, 1$ we can recursively define the sequence of functions

$$[\psi_t(Z_{[1,t-1]})](\cdot) := \rho_{\mathcal{X}_t|\mathcal{X}_{t-1}} \left[\inf_{z_t \in \mathcal{G}_t(Z_{[1,t-1]}(\cdot, \cdot))} \left\{ f_t(z_t, \cdot) + [\psi_{t+1}(Z_{[1,t-1]}, z_t)](\cdot) \right\} \right], \quad (7.8)$$

$t = T, T-1, \dots, 1$. We always have that the infima above are elements of the spaces \mathcal{X}_t , because the spaces \mathcal{X}_t are solid. Consequently, $\psi_t(Z_{[1,t-1]}) \in \mathcal{X}_{t-1}$ and the recursion can continue. Obviously, ψ_1 is the optimal value of problem (7.4).

We can interpret functions ψ_t as *cost-to-go functions* and equations (7.8) as *dynamic programming equations*, for the risk optimization problem (7.4). We can also write these functions in the form:

$$\begin{aligned} \psi_t(Z_{[1,t-1]}) = \inf_{Z_t(\cdot) \in \mathcal{G}_t(Z_{[1,t-1]}(\cdot, \cdot))} \dots \inf_{Z_T(\cdot) \in \mathcal{G}_T(Z_{[1,T-1]}(\cdot, \cdot))} \\ (\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}} \circ \dots \circ \rho_{\mathcal{X}_T|\mathcal{X}_{T-1}}) [F_t(Z_t) + \dots + F_T(Z_T)]. \end{aligned}$$

Suppose now that the conditional risk mappings $\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}$ are lower semicontinuous and positively homogeneous. Then it follows from (3.8) that there exists convex closed sets $\mathcal{A}_t(\omega) \subset \mathcal{P}_{\mathcal{Y}_t|\mathcal{F}_{t-1}}(\omega)$ such that

$$[\rho_{\mathcal{X}_t|\mathcal{X}_{t-1}}(X_t)](\omega) = \sup_{\mu_t \in \mathcal{A}_t(\omega)} \mathbb{E}_{\mu_t}[X_t], \quad \omega \in \Omega, \quad X_t \in \mathcal{X}_t.$$

Substitution into (7.8) yields, for $t = T, T-1, \dots, 1$,

$$[\psi_t(Z_{[1,t-1]})](\omega) = \sup_{\mu_t \in \mathcal{A}_t(\omega)} \mathbb{E}_{\mu_t} \left[\inf_{z_t \in \mathcal{G}_t(Z_{[1,t-1]}(\omega, \omega))} \left\{ f_t(z_t, \omega) + [\psi_{t+1}(Z_{[1,t-1]}, z_t)](\omega) \right\} \right]. \quad (7.9)$$

The dynamic programming equation (7.9) provides a framework for extending the theory of multistage stochastic optimization problems to risk functions.

In the special case when $\mathcal{X}_t = \mathcal{L}_p(\Omega, \mathcal{F}_t, \bar{\mu})$, $p \in [1, +\infty)$, we can identify the measures μ_t with their densities $h_t \in \mathcal{L}_q(\Omega, \mathcal{F}_t, \bar{\mu})$. We can observe that they form a process adapted to the filtration $\{\mathcal{F}_t\}$.

Acknowledgement. The authors are indebted to Darinka Dentcheva for helpful discussions regarding weakly measurable selections of multifunctions.

References

- [1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, Coherent measures of risk, *Mathematical Finance* 9 (1999) 203–228.
- [2] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath and H. Ku, Coherent multiperiod risk measurement, Manuscript, ETH Zürich, 2003.
- [3] P. Billingsley, *Probability and Measure*, Wiley, New York, 1995.
- [4] H. Föllmer and A. Schied, Convex measures of risk and trading constraints, *Finance and Stochastics* 6 (2002) 429–447.
- [5] K. Kuratowski and C. Ryll-Nardzewski, A general theorem on selectors, *Bull. Acad. Pol. Sc.* 13 (1965) 397–403.
- [6] W. Ogryczak and A. Ruszczyński, From stochastic dominance to mean–risk models: semideviations as risk measures, *European Journal of Operational Research*, 116 (1999), 33–50.
- [7] W. Ogryczak and A. Ruszczyński, On consistency of stochastic dominance and mean–semideviation models, *Mathematical Programming*, 89 (2001), 217–232.
- [8] W. Ogryczak and A. Ruszczyński, Dual stochastic dominance and related mean–risk models, *SIAM Journal on Optimization*, 13 (2002), 60–78.
- [9] G. Pflug and A. Ruszczyński, A risk measure for income processes, in: G. Szegö (Ed.), *Risk Measures for the 21st Century*, John Wiley & Sons, 2004.
- [10] R.T. Rockafellar and S.P. Uryasev, Optimization of conditional value-at-risk, *The Journal of Risk*, 2 (2000), 21–41.
- [11] R.T. Rockafellar, S. Uryasev and M. Zabarankin, Deviation measures in risk analysis and optimization, Research Report 2002-7, Department of Industrial and Systems Engineering, University of Florida.
- [12] R.T. Rockafellar and R.J-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.
- [13] A. Ruszczyński and A. Shapiro, Optimization of convex risk functions, *E-print available at: <http://www.optimization-online.org>*, 2004.
- [14] A. Shapiro and S. Ahmed, On a class of minimax stochastic programs, *SIAM Journal on Optimization*, accepted for publication.