FULL LENGTH PAPER

Series B

Risk forms: representation, disintegration, and application to partially observable two-stage systems

Darinka Dentcheva[1] · Andrzej Ruszczyński[2]

© Springer-Verlag GmbH Germany, part of Springer Nature and Mathematical Optimization Society 2019

Abstract

We introduce the concept of a risk form, which is a real functional of two arguments: a measurable function on a Polish space and a measure on that space. We generalize the duality theory and the Kusuoka representation to this setting. For a risk form acting on a product of Polish spaces, we define marginal and conditional forms and we prove a disintegration formula, which represents a risk form as a composition of its marginal and conditional forms. We apply the proposed approach to two-stage stochastic programming problems with partial information and decision-dependent observation distribution.

1 Introduction

The theory of risk measures is one of the main directions of recent developments in stochastic optimization. It has found multitude of applications, far beyond the original motivation in finance. The main setting is the following: a probability space $(\Omega, \mathscr{F}, P)$ is given and a space $\mathscr{Z}$ of real-valued measurable functions on Ω is defined (usually,

✉ Andrzej Ruszczyński
rusz@rutgers.edu

Darinka Dentcheva
darinka.dentcheva@stevens.edu

1 Department of Mathematical Sciences, Stevens Institute of Technology, Hoboken, NJ 07030, USA

2 Department of Management Science and Information Systems, Rutgers University, Piscataway, NJ 08854, USA

$\mathscr{L}_p(\Omega, \mathscr{F}, P)$ with $p \in [1, \infty]$). A (convex) risk measure is a convex, monotonic, and translation-equivariant functional $\rho : \mathscr{Z} \to \overline{\mathbb{R}}$. We refer to [1,14,22,27] for initial contributions, and to [15,29,30,36,38] and to the survey [4] for detailed presentation, applications, and further references.

Two key results provide variational representation of risk measures. One of them, called *dual representation*, can be derived from the theory of conjugate duality, as shown in [36]. Another representation is known as *Kusuoka representation* of law invariant coherent measures of risk [24]. It is derived from the dual representation by employing the Hardy–Littlewood–Pólya inequality (see [17]) under several assumptions about the properties of the measures of risk. In all these developments, the original probability measure P is assumed fixed, which is essential for the use of convex analysis techniques in the spaces of integrable functions.

We propose a different approach. We fix a Polish space $\mathscr{X}$ with its Borel σ-algebra $\mathscr{B}(\mathscr{X})$, but we allow arbitrary probability measures on this space. In Sect. 2, we introduce real-valued functionals of two arguments, $\rho[Z, P]$, where Z is a bounded measurable function on $\mathscr{X}$ and P is a probability measure on $(\mathscr{X}, \mathscr{B}(\mathscr{X}))$. In analogy to the bilinear form $\mathbb{E}[Z, P] = \int_{\mathscr{X}} Z(x) \, P(dx)$, we call $\rho[Z, P]$ a *risk form*. Transition risk mappings, which arose in our recent research on risk-averse control [10–12,34], are special cases of risk forms.

Under less restrictive assumptions than in the fixed probability measure case, we prove a generalized Kusuoka representation of risk forms in Sect. 3. We establish the universal character of the risk representation; it remains valid for all probability measures.

The second contribution of the paper is the risk disintegration formula and its implications. In Sect. 4, we introduce the property of conditional consistency of risk forms. We prove that forms enjoying this property can be represented as compositions of two forms, which we call marginal and conditional forms. This result generalizes the decomposition of the bilinear form, resulting from the disintegration of probability measures. While our approach is related to the theory of dynamic and conditional risk measures (see [2,5,6,13,19,23,30,32,33,35,37] and the references therein), it allows for variable probability measures and does not have any time structure associated with it; the order of conditioning may be arbitrary. These results are generalized in Sect. 5, where we consider multi-step disintegration and prove the generalized tower property of conditional risk forms, as a counterpart of the tower property of conditional expectations.

Our final contribution is the application of the risk form theory to two-stage risk-averse optimization of models with partial observation (Sect. 6). Opposite to classical two-stage models, we assume that only partial information is available at the second stage, which allows for the update of the conditional distribution of the unobserved part. This setting was first considered in [18,25], in a special case, and with a postulated structure of the overall measure of risk. We generalize and justify the earlier contributions, by proving the equivalence of the overall risk optimization and two-stage optimization in this setting. We also allow for decision-dependent observation distribution and develop a risk-averse Bayes formula. In the risk-neutral case, stochastic programming models with *endogeneous* (decision-dependent) uncertainty have been

discussed in [20], where the probability distribution and the first stage decision are linked by a special constraint.

2 Risk models with variable probability measures

Let $\mathscr{P}(\mathscr{X})$ be the set of probability measures on $(\mathscr{X}, \mathscr{B}(\mathscr{X}))$. The space of all real-valued bounded measurable functions on $\mathscr{X}$ is denoted by $\mathbb{B}(\mathscr{X})$. We use x to denote an element of $\mathscr{X}$ and δ_x to denote the Dirac measure concentrated at x. The symbol $\mathbb{1}$ stands for the function in $\mathbb{B}(\mathscr{X})$, which is constantly equal to 1.

A *probabilistic model* is a pair $[Z, P] \in \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X})$. Here and elsewhere, the Borel σ-algebra $\mathscr{B}(\mathscr{X})$ is considered. For two probabilistic models $[V, P]$ and $[W, Q]$ the notation $[V, P] \sim [W, Q]$ means that $P\{V \leq \eta\} = Q\{W \leq \eta\}$ for all $\eta \in \mathbb{R}$ (both models have the same distribution). The inequality $Z \leq V$ between elements of $\mathbb{B}(\mathscr{X})$ is always understood pointwise.

Our goal is to propose a universal approach to risk evaluation of a family of probabilistic models.

Definition 1 A measurable functional $\rho : \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X}) \to \mathbb{R}$ is called a *risk form*.

 (i) It is *monotonic*, if $V \leq W$ implies $\rho[V, P] \leq \rho[W, P]$ for all $P \in \mathscr{P}(\mathscr{X})$;
 (ii) It is *normalized* if $\rho[0, P] = 0$ for all $P \in \mathscr{P}(\mathscr{X})$;
(iii) It is *translation equivariant* if for all $V \in \mathbb{B}(\mathscr{X})$, all $a \in \mathbb{R}$, and all $P \in \mathscr{P}(\mathscr{X})$, $\rho[a\mathbb{1} + V, P] = a + \rho[V, P]$;
(iv) It is *positively homogeneous*, if for all $V \in \mathbb{B}(\mathscr{X})$, all $\beta \in \mathbb{R}_+$, and all $P \in \mathscr{P}(\mathscr{X})$, $\rho[\beta V, P] = \beta\rho[V, P]$;
 (v) It is *law invariant* if $[V, P] \sim [W, Q]$ implies that $\rho[V, P] = \rho[W, Q]$;
(vi) It has the *support property*, if $\rho[\mathbb{1}_{\text{supp}(P)} V, P] = \rho[V, P]$ for all $(V, P) \in \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X})$.

An example of a risk form is the expected value, which is a well-understood bilinear form:

$$\mathbb{E}[Z, P] = \int_{\mathscr{X}} Z(x)\, P(dx).$$

In our analysis, we are interested mainly in forms depending on one or both arguments in a nonlinear way.

Our concept of law invariance is broader than that used in the literature, because it allows for the probability measure to vary. If the risk form is law invariant, then it has the support property, because $[V, P] \sim [\mathbb{1}_{\text{supp}(P)} V, P]$.

Lemma 1 *If a risk form $\rho : \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X}) \to \mathbb{R}$ has the normalization, translation equivariance, and support properties then for every $Z \in \mathbb{B}(\mathscr{X})$ and every $x \in \mathscr{X}$*

$$\rho[Z, \delta_x] = Z(x). \tag{1}$$

Proof Using the support property twice, the translation property, and the normalization property, we obtain the chain of equations:

$$\rho[Z, \delta_x] = \rho[\mathbb{1}_x Z, \delta_x] = \rho[Z(x)\mathbb{1}, \delta_x] = Z(x) + \rho[0, \delta_x] = Z(x). \qquad \square$$

This property was called *state-consistency* in [9].

Essential role in our analysis will be played by the increasing convex order (the counterpart of the second order stochastic dominance, when smaller outcomes are preferred).

Definition 2 A probabilistic model $[Z, P]$ is smaller than a probabilistic model $[Z', P']$ in the *increasing convex order*, written $[Z, P] \preceq [Z', P']$, if for all $\eta \in \mathbb{R}$

$$\int_{\mathscr{X}} [Z(x) - \eta]_+ P(dx) \leq \int_{\mathscr{X}} [Z'(x) - \eta]_+ P'(dx).$$

Here, $[a]_+ = \max(0, a)$.

This concept allows us to consider risk forms consistent with the increasing order.

Definition 3 A risk form $\rho : \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X}) \to \mathbb{R}$ is *consistent with the increasing convex order*, if

$$[Z, P] \preceq [Z', P'] \implies \rho[Z, P] \leq \rho[Z', P'].$$

Evidently, consistency with the increasing convex order implies monotonicity and law invariance.

We call two functions $Z, V \in \mathbb{B}(\mathscr{X})$ *comonotonic*, if

$$\big(Z(x') - Z(x)\big)\big(V(x') - V(x)\big) \geq 0, \quad \forall\, x, x' \in \mathscr{X}.$$

Definition 4 A risk form $\rho : \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X}) \to \mathbb{R}$ is *comonotonically convex*, if for all comonotonic functions $Z, V \in \mathbb{B}(\mathscr{X})$, all $P \in \mathscr{P}(\mathscr{X})$, and all $\lambda \in [0, 1]$,

$$\rho[\lambda Z + (1 - \lambda)V, P] \leq \lambda \rho[Z, P] + (1 - \lambda)\rho[V, P].$$

3 Dual and Kusuoka representations

In this section, we generalize the dual representation and the Kusuoka representation of law invariant coherent risk measures [24,36] to risk forms. In the extant literature, these representations is always derived under the assumption that the probability measure is fixed (see, e.g., [16,21,30]). We show that a more general result using variable probability measures is true.

With every probabilistic model $[Z, P]$, we associate its distribution function,

$$F[Z, P](z) = P[Z \leq z], \quad z \in \mathbb{R},$$

and its quantile function,

$$\Phi[Z, P](p) = \inf \{\eta : P[Z \leq \eta] \geq p\}, \quad p \in (0, 1].$$

The quantile functions are elements of the space $\mathbb{Q}_b$ of bounded, nondecreasing, and left-continuous functions on $(0, 1]$.

We first adapt the general duality result of [8] for risk models on the space of quantile functions. We denote by $\mathcal{M}$ the set of countably additive finite measures on $(0, 1]$. For every risk form $\rho : \mathbb{B}(\mathcal{X}) \times \mathcal{P}(\mathcal{X}) \to \mathbb{R}$, we define the conjugate functional $\rho^* : \mathcal{M} \to \mathbb{R} \cup \{+\infty\}$ as follows:

$$\rho^*(\mu) = \sup_{[Z,P] \in \mathbb{B}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})} \left\{ \int_0^1 \Phi[Z, P](p) \, \mu(dp) - \rho[Z, P] \right\}. \tag{2}$$

We recall that a *total preorder* $\trianglelefteq$ on the space $\mathbb{Q}_b$ is a binary relation, which is reflexive, transitive and complete. It is *directed* if it satisfies the following conditions:

(i) For any real numbers $\alpha < \beta$, the relation $\alpha \mathbb{1} \lhd \beta \mathbb{1}$ is true;
(ii) For every $\Psi \in \mathbb{Q}_b$, numbers α and β exist such that $\alpha \mathbb{1} \trianglelefteq \Psi \trianglelefteq \beta \mathbb{1}$.

In [8], we introduced the following properties of preorders:

Dual Translation: For all Ψ_1 and Ψ_2 in $\mathbb{Q}_b$ and all $c \in \mathbb{R}$

$$\Psi_1 \trianglelefteq \Psi_2 \implies \Psi_1 + c\mathbb{1} \trianglelefteq \Psi_2 + c\mathbb{1}.$$

Dual Monotonicity: For all Ψ_1 and Ψ_2 in $\mathbb{Q}_b$

$$\Psi_1 \leq \Psi_2 \text{ pointwise} \implies \Psi_1 \trianglelefteq \Psi_2.$$

We obtain the following dual representation of risk forms.

Theorem 1 *If the space $\mathcal{X}$ is uncountable and a risk form $\rho : \mathbb{B}(\mathcal{X}) \times \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ is normalized, translation equivariant, comonotonically convex, and consistent with the increasing convex order, then a uniquely defined closed convex set*

$$\mathcal{D}_\rho \subseteq \{\mu \in \mathcal{M} : \mu(0, \cdot] \text{ is nondecreasing and convex on } (0, 1], \ \mu(0, 1] = 1\} \tag{3}$$

exists, such that for all $[Z, P] \in \mathbb{B}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$

$$\rho[Z, P] = \sup_{\mu \in \mathcal{D}_\rho} \left\{ \int_0^1 \Phi[Z, P](p) \, \mu(dp) - \rho^*(\mu) \right\}. \tag{4}$$

If, additionally, ρ is positively homogeneneous, then $\rho^*(\mu) = 0$ for all $\mu \in \mathcal{D}_\rho$.

Proof First, we show that the risk form $\rho[\cdot, \cdot]$ defines a functional R on the space of quantile functions $\mathbb{Q}_b$ on $(0, 1]$ by the identity:

$$R(\Phi[Z, P]) = \rho[Z, P]. \tag{5}$$

Indeed, $\rho[\cdot, \cdot]$ is law invariant, due to its consistency with the increasing convex order. Therefore, if $\Phi[Z, P] = \Phi[Z', P']$ then $\rho[Z, P] = \rho[Z', P']$. Thus the functional R is well-defined on the set of quantile functions $\{\Phi[Z, P] : Z \in \mathbb{B}(\mathscr{X}), \ P \in \mathscr{P}(\mathscr{X})\}$.

Since $\mathscr{X}$ is an uncountable Polish space, it is isomorphic to [0,1] equipped with the Borel σ-algebra (see, *e.g.*, [28, Th. 2.8 and Th. 2.12]). Denote by $Z_0 : \mathscr{X} \to [0, 1]$ the said isomorphism. Since both Z_0 and Z_0^{-1} are measurable, we can define a probability measure P_0 on $\mathscr{B}(\mathscr{X})$ by $\lambda \circ Z_0$, where λ is the Lebesgue measure on [0,1]. By construction, $P_0[Z_0 \leq p] = p$ and thus $\Phi[Z_0, P_0](p) = p$, for all $p \in [0, 1]$. Moreover, for every function $\Psi \in \mathbb{Q}_b$, we can define $Z(x) = \Psi(Z_0(x))$, $x \in \mathscr{X}$, and $\Psi^{-1}(z) = \sup\{p \in [0, 1] : \Psi(p) \leq z\}$. Then, for every $z \in \mathbb{R}$,

$$P_0\{x : Z(x) \leq z\} = P_0\{x : \Psi(Z_0(x)) \leq z\} = P_0\{x : Z_0(x) \leq \Psi^{-1}(z)\} = \Psi^{-1}(z). \tag{6}$$

Consequently, the distribution function of Z under P_0 is the inverse of Ψ, and thus $\Phi[Z, P_0] = \Psi$. This means that the domain of R is the entire space $\mathbb{Q}_b$.

We verify the assumptions of Theorem 4 of [8]. We define a preference relation $\trianglelefteq$ on $\mathbb{Q}_b$ by setting

$$\Psi_1 \trianglelefteq \Psi_2 \ \text{ if and only if } \ R(\Psi_1) \leq R(\Psi_2).$$

Clearly, the relation $\trianglelefteq$ is a total preorder with R being its numerical representation. Since $\rho[\cdot, \cdot]$ is normalized and translation equivariant, the identity (5) implies that R is normalized and translation equivariant as well. We observe that R is monotonic, i.e., if $\Psi_1 \leq \Psi_2$ (pointwise), then $R(\Psi_1) \leq R(\Psi_2)$. Indeed, let $\Psi_1 \leq \Psi_2$ and set $Z_1(x) = \Psi_1(Z_0(x))$, $Z_2(x) = \Psi_2(Z_0(x))$ for all $x \in \mathscr{X}$. Similar to (6),

$$P_0(Z_1 \leq z) = \Psi_1^{-1}(z) \geq \Psi_2^{-1}(z) = P_0(Z_2 \leq z)$$

for all $z \in \mathbb{R}$. The last relation implies that $[Z_1, P_0] \preceq [Z_2, P_0]$. The consistency of ρ with the increasing convex order entails

$$R(\Psi_1) = \rho[Z_1, P_0] \leq \rho[Z_2, P_0] = R(\Psi_2),$$

which is the desired monotonicity. The properties of R further imply that the order $\trianglelefteq$ is directed, monotonic, and satisfies the dual translation property.

For any two comonotonic functions Z_1 and Z_2 in $\mathbb{B}(\mathscr{X})$, any $\lambda \in [0, 1]$, and any $P \in \mathscr{P}(\mathscr{X})$,

$$\Phi[\lambda Z_1 + (1 - \lambda)Z_2, P] = \lambda \Phi[Z_1, P] + (1 - \lambda)\Phi[Z_2, P].$$

The comonotonic convexity assumption implies that

$$\begin{aligned} R(\lambda \Phi[Z_1, P] + (1 - \lambda)\Phi[Z_2, P]) &= \rho[\lambda Z_1 + (1 - \lambda)Z_2, P] \\ &\leq \lambda \rho[Z_1, P] + (1 - \lambda)\rho[Z_2, P] = \lambda R(\Phi[Z_1, P]) + (1 - \lambda)R(\Phi[Z_2, P]). \end{aligned}$$

Since any two functions $\Psi_1, \Psi_2 \in \mathbb{Q}_b$ can be represented as

$$\Psi_1 = \Phi[Z_1, P_0], \quad \text{with} \quad Z_1(x) = \Psi_1(Z_0(x)), \ x \in \mathscr{X},$$
$$\Psi_2 = \Phi[Z_2, P_0], \quad \text{with} \quad Z_2(x) = \Psi_2(Z_0(x)), \ x \in \mathscr{X},$$

and the functions Z_1 and Z_2 are comonotonic by construction, the functional R is convex.

Consider a function $\Psi \in \mathbb{Q}_b$. For $(a, b] \subset (0, 1]$ we define

$$\Psi_{(a,b]}(p) = \begin{cases} \dfrac{1}{b-a} \displaystyle\int_a^b \Psi(s) \, ds & \text{if } p \in (a, b], \\ \Psi(p) & \text{otherwise.} \end{cases} \tag{7}$$

Directly from (7) we observe that for every $\alpha \in (0, 1]$

$$\int_{1-\alpha}^1 \Psi_{(a,b]}(p) \, dp \le \int_{1-\alpha}^1 \Psi(p) \, dp.$$

Therefore, for any $[Z, P]$ and $[V, Q]$ such that $\Psi_{(a,b]} = \Phi[Z, P]$ and $\Psi = \Phi[V, Q]$, we have $[Z, P] \preceq [V, Q]$. Due to the consistency of $\rho[\cdot, \cdot]$ with the increasing convex order,

$$R(\Psi_{(a,b]}) = \rho[Z, P] \le \rho[V, Q] = R(\Psi).$$

Therefore, the preorder $\trianglelefteq$ is risk averse in the sense of [8, Def. 2]. It follows from [8, Th. 4] that a set $\mathscr{D}_\rho$ satisfying (3) exists, such that

$$R(\Psi) = \sup_{\mu \in \mathscr{D}_\rho} \left\{ \int_0^1 \Psi(p) \, \mu(dp) - R^*(\mu) \right\},$$

$$\text{with} R^*(\mu) = \sup_{\Psi \in \mathbb{Q}_b} \left\{ \int_0^1 \Psi(p) \, \mu(dp) - R(\Psi) \right\}.$$

Moreover, $R^*(\mu) = 0$ for $\mu \in \mathscr{D}_\rho$, if R is positively homogeneous. The assertion of the theorem follows now from the substitution (5). □

The dual representation (4) can be written with the use of the Stjelties integral with respect to the distribution function $w(\cdot) = \mu(0, \cdot]$. It is particularly revealing in the homogeneous case:

$$\rho[Z, P] = \sup_{w \in \mathscr{W}_\rho} \int_0^1 \Phi[Z, P](p) \, dw(p), \tag{8}$$

where each $w(\cdot) \in \mathscr{W}_\rho$ can be interpreted as a dual (rank dependent) utility function. The set $\mathscr{W}_\rho$ is the set of distribution functions of measures $\mu \in \mathscr{D}_\rho$: a convex subset

of convex and nondecreasing functions from [0, 1] to [0, 1]. The relation (8) suggests an intriguing relation of law invariant risk forms and the dual utility theory of [31,39], as analyzed in [7].

Theorem 1 allows us to derive a generalization of the celebrated Kusuoka representation of law invariant coherent measures of risk (see [16,24], and [30, sec. 2.2.4] for an overview of relevant results).

Definition 5 The Average Value at Risk at level $\alpha \in [0, 1]$ of a probabilistic model $[Z, P]$ is defined as follows:

$$\text{AVaR}_\alpha[Z, P] = \begin{cases} \frac{1}{\alpha} \int_{1-\alpha}^1 \Phi[Z, P](p) \, dp & \text{if } \alpha \in (0, 1), \\ \Phi[Z, P](1) & \text{if } \alpha = 0, \\ \mathbb{E}[Z, P] & \text{if } \alpha = 1. \end{cases}$$

Then, repeating the considerations leading to [8, Cor. 1] *verbatim*, we obtain the following result.

Corollary 1 *Suppose the conditions of Theorem* 1 *are satisfied and the risk form* $\rho[\cdot, \cdot]$ *is positively homogeneous. Then a convex subset* Λ_ρ *of the set of probability measures on* [0, 1] *exists, such that for all* $[Z, P] \in \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X})$

$$\rho[Z, P] = \sup_{\lambda \in \Lambda_\rho} \int_0^1 \text{AVaR}_s[Z, P] \, \lambda(ds). \tag{9}$$

It it worth stressing that in the extant literature, the Kusuoka representations were derived for probabilistic models with a fixed atomless probability measure P. If P has atoms, additional conditions are needed, as discussed in [26]. Our approach proves the validity of the Kusuoka representations for probabilistic models with an arbitrary probability measure P. The set Λ_ρ is determined by the risk form alone; it is the same for all probabilistic models $[Z, P]$. This universal property is due to two key conditions: the consistency with the increasing convex order with *both* Z and P varying, and the requirement that $\mathscr{X}$ be Polish and uncountable.

4 The risk disintegration formula

Our interest in this section is measuring risk on product spaces. Consider two Polish spaces $\mathscr{X}$ and $\mathscr{Y}$ and their corresponding Borel σ-algebras $\mathscr{B}(\mathscr{X})$ and $\mathscr{B}(\mathscr{Y})$. We can disintegrate any $P \in \mathscr{P}(\mathscr{X} \times \mathscr{Y})$ into its marginal $P_{\mathscr{X}} \in \mathscr{P}(\mathscr{X})$ and a transition kernel $P_{\mathscr{Y}|\mathscr{X}} : \mathscr{X} \to \mathscr{P}(\mathscr{Y})$ as follows: $P(dx, dy) = P_{\mathscr{X}}(dx) \, P_{\mathscr{Y}|\mathscr{X}}(dy|x)$.

Let $\mathscr{Q}(\mathscr{Y}|\mathscr{X})$ be the space of all measurable mappings $Q : \mathscr{X} \to \mathscr{P}(\mathscr{Y})$ (transition kernels). For any measure $\lambda \in \mathscr{P}(\mathscr{X})$ and any kernel $Q \in \mathscr{Q}(\mathscr{Y}|\mathscr{X})$, the measure $\lambda \otimes Q$, defined as

$$\big[\lambda \otimes Q\big](dx, dy) = \lambda(dx) \, Q(dy|x),$$

is an element of $\mathscr{P}(\mathscr{X} \times \mathscr{Y})$.

Suppose the risk form $\rho : \mathbb{B}(\mathscr{X} \times \mathscr{Y}) \times \mathscr{P}(\mathscr{X} \times \mathscr{Y}) \to \mathbb{R}$ is monotonic, translation equivariant, and normalized. Then it induces a mapping $\rho_{\mathscr{Y}|\mathscr{X}} : \mathbb{B}(\mathscr{X} \times \mathscr{Y}) \times \mathscr{Q}(\mathscr{Y}|\mathscr{X}) \to \mathbb{B}(\mathscr{X})$ defined as follows:

$$\rho_{\mathscr{Y}|\mathscr{X}}[Z, Q](x) = \rho[Z, \delta_x \otimes Q], \quad x \in \mathscr{X}. \tag{10}$$

We call the mapping $\rho_{\mathscr{Y}|\mathscr{X}}[\cdot, \cdot]$ the *conditional risk operator* associated with $\rho[\cdot, \cdot]$.

To verify that the values are indeed elements of $\mathbb{B}(\mathscr{X})$, let $c \in \mathbb{R}$ be such that $Z \le c\mathbb{1}$. Then, by monotonicity, translation equivariance and normalization,

$$\rho[Z, \delta_x \otimes Q] \le \rho[c\mathbb{1}, \delta_x \otimes Q] = c.$$

The lower bound is similar, and thus the function $\rho_{\mathscr{Y}|\mathscr{X}}[Z, Q]$ is bounded. The function $\rho_{\mathscr{Y}|\mathscr{X}}[Z, Q](\cdot)$ is measurable, as a composition of measurable mappings.

If the risk form ρ has the support property, then for each $x \in \mathscr{X}$ the value of (10) depends only on the function $Z(x, \cdot) \in \mathbb{B}(\mathscr{Y})$ and the measure $Q(x) \in \mathscr{P}(\mathscr{Y})$. We can, therefore, define the functionals $\rho_{\mathscr{Y}|x} : \mathbb{B}(\mathscr{Y}) \times \mathscr{P}(\mathscr{Y}) \to \mathbb{R}$, $x \in \mathscr{X}$, as follows:

$$\rho_{\mathscr{Y}|x}[Z(x, \cdot), Q(x)] = \rho_{\mathscr{Y}|\mathscr{X}}[Z, Q](x), \quad x \in \mathscr{X}. \tag{11}$$

We call them *conditional risk forms* associated with $\rho[\cdot, \cdot]$. Observe that any function from $\mathbb{B}(\mathscr{Y})$ and any measure from $\mathscr{P}(\mathscr{Y})$ may feature as arguments of $\rho_{\mathscr{Y}|x}[\cdot, \cdot]$.

From now on, we always assume that the risk forms in question have the support property. The inequalities "$\le$" between functions are always understood point-wise.

Lemma 2 *If the risk form $\rho[\cdot, \cdot]$ is monotonic (normalized, translation equivariant), then, for every $x \in \mathscr{X}$, the conditional risk form $\rho_{\mathscr{Y}|x}$ is monotonic (normalized, translation equivariant).*

Proof All the properties follow directly from the equation

$$\rho_{\mathscr{Y}|x}[Z(x, \cdot), Q(x)] = \rho[Z, \delta_x \otimes Q],$$

which defines the conditional risk form. □

Definition 6 A risk form $\rho : \mathbb{B}(\mathscr{X} \times \mathscr{Y}) \times \mathscr{P}(\mathscr{X} \times \mathscr{Y}) \to \mathbb{R}$ is *conditionally consistent* if for all $Z, Z' \in \mathscr{B}(\mathscr{X} \times \mathscr{Y})$ and all $Q, Q' \in \mathscr{Q}(\mathscr{Y}|\mathscr{X})$ such that

$$\rho_{\mathscr{Y}|\mathscr{X}}[Z, Q] \le \rho_{\mathscr{Y}|\mathscr{X}}[Z', Q'],$$

we also have

$$\rho[Z, \lambda \otimes Q] \le \rho[Z', \lambda \otimes Q'], \quad \forall \lambda \in \mathscr{P}(\mathscr{X}).$$

The following result is the foundation of our further considerations.

Theorem 2 *Suppose a risk form $\rho : \mathbb{B}(\mathscr{X} \times \mathscr{Y}) \times \mathscr{P}(\mathscr{X} \times \mathscr{Y}) \to \mathbb{R}$ is monotonic, normalized, translation equivariant, has the support property, and is conditionally consistent. Then a risk form $\rho_{\mathscr{X}} : \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X}) \to \mathbb{R}$ exists, such that for all $[Z, P] \in \mathbb{B}(\mathscr{X} \times \mathscr{Y}) \times \mathscr{P}(\mathscr{X} \times \mathscr{Y})$ the following formula is true:*

$$\rho[Z, P] = \rho_{\mathscr{X}}\big[\rho_{\mathscr{Y}|\mathscr{X}}[Z, P_{\mathscr{Y}|\mathscr{X}}], P_{\mathscr{X}}\big]. \tag{12}$$

The risk form $\rho_{\mathscr{X}}$ is uniquely defined by the equation

$$\rho_{\mathscr{X}}[f, P_{\mathscr{X}}] = \rho[\underline{f}, P], \quad \text{with} \quad \underline{f}(x, y) \equiv f(x). \tag{13}$$

It is monotonic, normalized, translation equivariant, and has the support property.

Proof Let us verify (12). Suppose $[Z, P]$ and $[Z', P']$ are such that

$$\rho_{\mathscr{Y}|\mathscr{X}}[Z, P_{\mathscr{Y}|\mathscr{X}}] = \rho_{\mathscr{Y}|\mathscr{X}}[Z', P'_{\mathscr{Y}|\mathscr{X}}].$$

Then it follows from Definition 6 that $\rho[Z, \lambda \otimes P_{\mathscr{Y}|\mathscr{X}}] = \rho[Z', \lambda \otimes P'_{\mathscr{Y}|\mathscr{X}}]$ for all $\lambda \in \mathscr{P}(\mathscr{X})$. If, additionally, the marginal measures $P_{\mathscr{X}}$ and $P'_{\mathscr{X}}$ are identical, by setting $\lambda = P_{\mathscr{X}} = P'_{\mathscr{X}}$ we conclude that

$$\rho[Z, P] = \rho[Z, P_{\mathscr{X}} \otimes P_{\mathscr{Y}|\mathscr{X}}] = \rho[Z', P'_{\mathscr{X}} \otimes P'_{\mathscr{Y}|\mathscr{X}}] = \rho[Z', P'].$$

It follows that the value of $\rho[Z, P]$ is fully determined by the value of the conditional risk operator $\rho_{\mathscr{Y}|\mathscr{X}}[Z, P_{\mathscr{Y}|\mathscr{X}}]$ and the marginal measure $P_{\mathscr{X}}$. Therefore, the disintegration formula (12) is true.

It remains to verify the properties of $\rho_{\mathscr{X}}$. Set $Z(x, y) = \underline{f}(x, y) \equiv f(x)$ in (12). Then, by the support, translation equivariance, and normalization properties of $\rho[\cdot, \cdot]$, for every $x \in \mathscr{X}$ we obtain

$$\rho_{\mathscr{Y}|\mathscr{X}}[\underline{f}, P_{\mathscr{Y}|\mathscr{X}}](x) = \rho[f(x)\mathbb{1}, \delta_x \otimes P_{\mathscr{Y}|\mathscr{X}}] = f(x).$$

Combining with (12), we observe that the identity (13) is true. All the postulated properties of $\rho_{\mathscr{X}}[\cdot, \cdot]$ follow from the corresponding properties of $\rho[\cdot, \cdot]$. □

We call the identity (12) the *risk disintegration formula*. It represents the risk form $\rho[\cdot, \cdot]$ by its marginal risk form $\rho_{\mathscr{X}}[\cdot, \cdot]$ and its conditional risk operator $\rho_{\mathscr{Y}|\mathscr{X}}[\cdot, \cdot]$.

Example 1 Consider the risk form $\rho : \mathbb{B}(\mathscr{X} \times \mathscr{Y}) \times \mathscr{P}(\mathscr{X} \times \mathscr{Y}) \to \mathbb{R}$ defined as follows:

$$\rho[Z, P] = \min_{\eta(\cdot)} \mathbb{E}_P \Big[\eta(x) + \frac{1}{\alpha} \big(Z(x, y) - \eta(x) \big)_+ \Big],$$

where the minimization is over all measurable mappings $\eta : \mathscr{X} \to \mathbb{R}$. Directly from Definition 1 we verify that it is normalized, monotonic, translation invariant, and has the support property. To verify the conditional consistency, observe that

$$\rho_{\mathscr{Y}|x}[Z, Q] = \min_{\eta} \left[\eta + \frac{1}{\alpha} \mathbb{E}_{Q(x)} \left[(Z(x, y) - \eta)_+ \right] \right]. \tag{14}$$

Therefore, the relation

$$\rho[Z, \delta_x \otimes Q] \le \rho[Z', \delta_x \otimes Q'], \quad \forall \, x \in \mathscr{X},$$

means that

$$\min_{\eta} \left[\eta + \frac{1}{\alpha} \mathbb{E}_{Q(x)} \left[(Z(x, y) - \eta)_+ \right] \right] \le \min_{\eta} \left[\eta + \frac{1}{\alpha} \mathbb{E}_{Q'(x)} \left[(Z'(x, y) - \eta)_+ \right] \right],$$

for all $x \in \mathscr{X}$. By the interchangeability of $\min_{\eta(\cdot)}$ and $\mathbb{E}_\lambda$, for every $\lambda \in \mathscr{P}(\mathscr{X})$ we obtain:

$$\rho[Z, \lambda \otimes Q] = \min_{\eta(\cdot)} \mathbb{E}_\lambda \left[\eta(x) + \frac{1}{\alpha} \mathbb{E}_{Q(x)} \left[(Z(x, y) - \eta(x))_+ \right] \right]$$

$$= \mathbb{E}_\lambda \min_{\eta} \left[\eta + \frac{1}{\alpha} \mathbb{E}_{Q(x)} \left[(Z(x, y) - \eta)_+ \right] \right]$$

$$\le \mathbb{E}_\lambda \min_{\eta} \left[\eta + \frac{1}{\alpha} \mathbb{E}_{Q'(x)} \left[(Z'(x, y) - \eta)_+ \right] \right] = \rho[Z', \lambda \otimes Q'].$$

Therefore, the assumptions of Theorem 2 are satisfied, and the risk form $\rho[\cdot, \cdot]$ can be disintegrated into a marginal and conditional form. Using the dual representation of the Average Value at Risk in (14), we obtain the explicit formula:

$$\rho[Z, P] = \mathbb{E}_{P_{\mathscr{X}}} \left[\text{AVaR}_\alpha \left[Z(x, y), P_{\mathscr{Y}|\mathscr{X}} \right] \right].$$

This means that $\rho_{\mathscr{X}}[\cdot, P_{\mathscr{X}}] = \mathbb{E}_{P_{\mathscr{X}}}[\cdot]$ and $\rho_{\mathscr{Y}|\mathscr{X}}[\cdot, P_{\mathscr{Y}|\mathscr{X}}] = \text{AVaR}_\alpha[\cdot, P_{\mathscr{Y}|\mathscr{X}}]$.

The risk form $\rho[\cdot, \cdot]$ is not law invariant on the product space, because pairs $[Z, P]$ having identical distribution may have different conditional and marginal distributions. This can be seen on the case of $\mathscr{X} = \mathscr{Y} = [0, 1]$ with the Lebesgue measure P on $\mathscr{X} \times \mathscr{Y}$ and two functions: $Z(x, y) = x$ and $Z'(x, y) = y$. Then $[Z, P] \sim [Z', P]$. However,

$$\rho[Z, P] = \min_{\eta(\cdot)} \int_0^1 \int_0^1 \left[\eta(x) + \frac{1}{\alpha} (x - \eta(x))_+ \right] dy \, dx$$

$$= \int_0^1 \min_{\eta} \left[\eta + \frac{1}{\alpha} (x - \eta)_+ \right] dx = \int_0^1 x \, dx = \frac{1}{2},$$

and

$$\rho[Z', P] = \min_{\eta(\cdot)} \int_0^1 \int_0^1 \left[\eta(x) + \frac{1}{\alpha} (y - \eta(x))_+ \right] dy \, dx$$

$$\ge \int_0^1 \min_{\eta} \int_0^1 \left[\eta + \frac{1}{\alpha} (y - \eta)_+ \right] dy \, dx = \int_0^1 \text{AVaR}_\alpha(Y) \, dx = \text{AVaR}_\alpha(Y).$$

We observe that the minimizing η in the second line of the formula above is in fact constant and equal to $1 - \alpha$. Therefore, it can be used in the first line, the inequality becomes an equation, and $\rho[Z', P] = 1 - \alpha/2$. We conclude that $\rho[Z, P] \neq \rho[Z', P]$, unless $\alpha = 1$. This example illustrates the fact that the concept of law invariance on product spaces is excessively demanding. Therefore, we do not assume law invariance of risk forms on product spaces.

5 Composite disintegration

We now generalize our results to the product of multiple Polish spaces $\mathscr{X}_j$, $j = 1, \ldots, n$, with their corresponding Borel σ-algebras $\mathscr{B}(\mathscr{X}_j)$, $j = 1, \ldots, n$. For a nonempty set of indices $J \subset \{1, \ldots, n\}$, we write

$$\mathscr{X}_J = \underset{j \in J}{\bigtimes} \mathscr{X}_j, \quad \text{and} \quad J^c = \{1, \ldots, n\} \setminus J.$$

Let P be a probability measure on $\mathscr{X} = \bigtimes_{j=1}^n \mathscr{X}_j$. For every J such that $J^c \neq \emptyset$, we can disintegrate P into its marginal $P_{\mathscr{X}_J} \in \mathscr{P}(\mathscr{X}_J)$ and a transition kernel $P_{\mathscr{X}_{J^c} | \mathscr{X}_J} : \mathscr{X}_J \to \mathscr{P}(\mathscr{X}_{J^c})$ as follows:

$$P(dx_J, dx_{J^c}) = P_{\mathscr{X}_J}(dx_J) \, P_{\mathscr{X}_{J^c} | \mathscr{X}_J}(dx_{J^c} | x_J).$$

For the case $J^c = \emptyset$, trivially $P(dx_J) = P_{\mathscr{X}}(dx_J)$. We denote the set of transition kernels from $\mathscr{X}_J$ to $\mathscr{P}(\mathscr{X}_{J^c})$ by $\mathscr{Q}_{\mathscr{X}_{J^c} | \mathscr{X}_J}$.

Exactly as in Sect. 4, a monotonic, translation equivariant, and normalized risk form $\rho : \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X}) \to \mathbb{R}$ on the product space induces a family of conditional risk operators $\rho_{\mathscr{X}_{J^c} | \mathscr{X}_J} : \mathbb{B}(\mathscr{X}) \times \mathscr{Q}_{\mathscr{X}_{J^c} | \mathscr{X}_J} \to \mathbb{B}(\mathscr{X}_J)$, as follows:

$$\rho_{\mathscr{X}_{J^c} | \mathscr{X}_J}[Z, Q](x_J) = \rho[Z, \delta_{x_J} \otimes Q], \quad x_J \in \mathscr{X}_J. \tag{15}$$

If the risk form ρ has the support property, then for each $x_J \in \mathscr{X}_J$ the value of (10) depends only on the function $Z(x_J, \cdot) \in \mathbb{B}(\mathscr{X}_{J^c})$ and the measure $Q(x_J) \in \mathscr{P}(\mathscr{X}_{J^c})$. As in Sect. 4, we define the conditional risk forms functionals $\rho_{\mathscr{X}_{J^c} | x_J} : \mathbb{B}(\mathscr{X}_{J^c}) \times \mathscr{P}(\mathscr{X}_{J^c}) \to \mathbb{R}$, $x \in \mathscr{X}$, as follows:

$$\rho_{\mathscr{X}_{J^c} | x_J}[Z(x_J, \cdot), Q(x_J)] = \rho_{\mathscr{X}_{J^c} | \mathscr{X}_J}[Z, Q](x_J), \quad x_J \in \mathscr{X}_J. \tag{16}$$

Definition 7 A risk form $\rho : \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X}) \to \mathbb{R}$ is *conditionally consistent with respect to* $\mathscr{X}_J$, where $\emptyset \neq J \subset \{1, \ldots, n\}$, if for all $Z, Z' \in \mathscr{B}(\mathscr{X})$ and all $Q, Q' \in \mathscr{Q}_{\mathscr{X}_{J^c} | \mathscr{X}_J}$ the inequality

$$\rho_{\mathscr{X}_{J^c} | \mathscr{X}_J}[Z, Q] \leq \rho_{\mathscr{X}_{J^c} | \mathscr{X}_J}[Z', Q'],$$

implies that

$$\rho[Z, \lambda \otimes Q] \leq \rho[Z', \lambda \otimes Q'], \quad \forall \lambda \in \mathscr{P}(\mathscr{X}_J).$$

The following corollary results directly from Theorem 2.

Corollary 2 *If a risk form* $\rho : \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X}) \to \mathbb{R}$ *is monotonic, normalized, translation equivariant, and conditionally consistent with respect to* $\mathscr{X}_J$, *where* $\emptyset \neq J \subset \{1, \ldots, n\}$, *then a risk form* $\rho_{\mathscr{X}_J} : \mathbb{B}(\mathscr{X}_J) \times \mathscr{P}(\mathscr{X}_J) \to \mathbb{R}$ *exists, such that for all* $[Z, P] \in \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X})$ *the following formula holds:*

$$\rho[Z, P] = \rho_{\mathscr{X}_J} \big[\rho_{\mathscr{X}_{J^c} | \mathscr{X}_J}[Z, P_{\mathscr{X}_{J^c} | \mathscr{X}_J}], P_{\mathscr{X}_J} \big], \tag{17}$$

where the marginal risk form $\rho_{\mathscr{X}_J}$ *is uniquely defined by the Eq.* (13) *with* $(\mathscr{X}_j, \mathscr{X}_{J^c})$ *replacing* $(\mathscr{X}, \mathscr{Y})$. *It is monotonic, normalized, translation equivariant, and has the support property.*

A question arises what is the relation between the marginal and conditional risk forms for different subspaces.

Theorem 3 *Suppose* $\emptyset \neq J \subset L \subset \{1, \ldots, n\}$. *If a risk form* $\rho : \mathbb{B}(\mathscr{X}) \times \mathscr{P}(\mathscr{X}) \to \mathbb{R}$ *is monotonic, normalized, translation equivariant, and conditionally consistent with respect to both* $\mathscr{X}_J$ *and* $\mathscr{X}_L$, *then the following statements are true:*

(i) *For all* $x_J \in \mathscr{X}_J$ *the conditional risk forms* $\rho_{\mathscr{X}_{J^c} | x_J}$ *are monotonic, normalized, translation equivariant, and conditionally consistent with respect to* $\mathscr{X}_{L \setminus J}$;

(ii) *For all* $x_L \in \mathscr{X}_L$ *we have* $\big(\rho_{\mathscr{X}_{J^c} | x_J} \big)_{\mathscr{X}_{L^c} | x_{L \setminus J}} = \rho_{\mathscr{X}_{L^c} | x_L}$;

(ii) *The marginal risk form* $\rho_{\mathscr{X}_L}$ *is monotonic, normalized, translation equivariant, and conditionally consistent with respect to* $\mathscr{X}_J$;

(iv) *We have* $\rho_{\mathscr{X}_J} = \big(\rho_{\mathscr{X}_L} \big)_{\mathscr{X}_J}$.

Proof The monotonicity, normalization, and translation equivariance of the marginal and conditional forms follow from the corresponding properties of ρ via formula (17) by considering special classes of functions in $\mathbb{B}(\mathscr{X})$: the functions that depend only on x_J (for the marginal), and the functions that depend only on x_{J^c} (for the conditionals). The proof is identical to the last part of the proof of Theorem 2.

It remains to prove conditional consistency of the conditional and marginal forms and the tower formulae (ii) and (iv). For a fixed $x_J \in \mathscr{X}_J$, we verify Definition 7 for the conditional risk form $\rho_{\mathscr{X}_{J^c} | x_J}$ with respect to $\mathscr{X}_{L \setminus J}$. Let $K = L \setminus J$, and let $Z, Z' \in \mathbb{B}(\mathscr{X}_{J^c})$ and $Q, Q' \in \mathscr{Q}_{\mathscr{X}_{L^c} | \mathscr{X}_K}$. Suppose

$$\big(\rho_{\mathscr{X}_{J^c} | x_J} \big)_{\mathscr{X}_{L^c} | \mathscr{X}_K}[Z, Q] \leq \big(\rho_{\mathscr{X}_{J^c} | x_J} \big)_{\mathscr{X}_{L^c} | \mathscr{X}_K}[Z', Q'],$$

which means that

$$\rho_{\mathscr{X}_{J^c} | x_J}[Z, \delta_{x_K} \otimes Q] \leq \rho_{\mathscr{X}_{J^c} | x_J}[Z', \delta_{x_K} \otimes Q'], \quad \forall x_K \in \mathscr{X}_K. \tag{18}$$

We can formally extend the functions Z and Z' to the entire domain $\mathscr{X}$ by setting $\bar{Z}(x_J, x_{J^c}) = Z(x_{J^c})$ and $\bar{Z}'(x_J, x_{J^c}) = Z'(x_{J^c})$. We can also define the kernels $\bar{Q}$ and $\bar{Q}'$ in $\mathscr{Q}_{\mathscr{X}_{L^c}|\mathscr{X}_L}$ by setting $\bar{Q}(x_J, x_K) = Q(x_K)$ and $\bar{Q}'(x_J, x_K) = Q'(x_K)$. Then

$$\rho_{\mathscr{X}_{J^c}|x_J}[Z, \delta_{x_K} \otimes Q] = \rho[\bar{Z}, \delta_{x_L} \otimes \bar{Q}].$$

A similar equation is true for Z' and Q'. Then (18) can be written as follows:

$$\rho[\bar{Z}, \delta_{x_L} \otimes \bar{Q}] \le \rho[\bar{Z}', \delta_{x_L} \otimes \bar{Q}'].$$

By the conditional consistency of ρ,

$$\rho[\bar{Z}, \psi \otimes \bar{Q}] \le \rho[\bar{Z}', \psi \otimes \bar{Q}'], \quad \forall \psi \in \mathscr{P}(\mathscr{X}_J).$$

Let $\lambda \in \mathscr{P}(\mathscr{X}_K)$. By setting $\psi = \delta_{x_J} \otimes \lambda$ in the last displayed inequality, and using the fact that $\bar{Z}, \bar{Z}', \bar{Q}$, and $\bar{Q}'$ do not depend on x_J, we conclude that

$$\rho_{\mathscr{X}_{J^c}|x_J}[Z, \lambda \otimes Q] \le \rho_{\mathscr{X}_{J^c}|x_J}[Z', \lambda \otimes Q'].$$

This proves the conditional consistency of the conditional risk form $\rho_{\mathscr{X}_{J^c}|x_J}$.

To verify (ii), consider $f \in \mathbb{B}(\mathscr{X}_{L^c})$, $Q \in \mathscr{Q}_{\mathscr{X}_{L^c}|\mathscr{X}_L}$, and the natural extension $\bar{f}$ of f to the entire space $\mathscr{X}$, defined by $\bar{f}(x_L, x_{L^c}) = f(x_{L^c})$. We obtain the chain of equalities:

$$\rho_{\mathscr{X}_{L^c}|x_L}[f, Q] = \rho[\bar{f}, \delta_{x_L} \otimes Q] = \rho[\bar{f}, \delta_{x_J} \otimes \delta_{x_K} \otimes Q]$$
$$= \rho_{\mathscr{X}_{J^c}|x_J}[\bar{f}, \delta_{x_K} \otimes Q] = (\rho_{\mathscr{X}_{J^c}|x_J})_{\mathscr{X}_{L^c}|x_K}[f, Q].$$

Consider now the marginal risk form $\rho_{\mathscr{X}_L}$. Let $K = L \setminus J$ and $Q, Q' \in \mathscr{Q}_{\mathscr{X}_K|\mathscr{X}_J}$. Suppose

$$\rho_{\mathscr{X}_L}[Z, \delta_{x_J} \otimes Q] \le \rho_{\mathscr{X}_L}[Z', \delta_{x_J} \otimes Q'], \quad \forall x_J \in \mathscr{X}_J, \tag{19}$$

where $Z, Z' \in \mathbb{B}(\mathscr{X}_L)$. We can formally extend the functions Z and Z' to the entire domain $\mathscr{X}$ by setting $\bar{Z}(x_L, x_{L^c}) = Z(x_L)$ and $\bar{Z}'(x_L, x_{L^c}) = Z'(x_L)$. We can also define the kernels $\bar{Q}$ and $\bar{Q}'$ in $\mathscr{Q}_{\mathscr{X}_{J^c}|\mathscr{X}_J}$ by setting $\bar{Q} = Q \times M$ and $\bar{Q}' = Q' \times M$, where M is an arbitrary kernel in $\mathscr{Q}_{\mathscr{X}_{L^c}|\mathscr{X}_J}$. Since the functions $\bar{Z}$ and $\bar{Z}'$ do not depend on x_{L^c}, the relations (19) can be equivalently written as

$$\rho[\bar{Z}, \delta_{x_J} \otimes \bar{Q}] \le \rho[\bar{Z}', \delta_{x_J} \otimes \bar{Q}'], \quad \forall x_J \in \mathscr{X}_J.$$

By the conditional consistency of ρ with respect to $\mathscr{X}_J$,

$$\rho[\bar{Z}, \lambda \otimes \bar{Q}] \le \rho[\bar{Z}', \lambda \otimes \bar{Q}'], \quad \forall \lambda \in \mathscr{P}(\mathscr{X}_J).$$

Since $\bar{Z}$ and $\bar{Z}'$ do not depend on x_{L^c}, we conclude that

$$\rho_{\mathscr{X}_L}[Z, \lambda \otimes Q] \le \rho_{\mathscr{X}_L}[Z', \lambda \otimes Q'],$$

which proves the conditional consistency of the marginal risk form $\rho_{\mathscr{X}_L}$.

It remains to verify the tower property (iv). For $f \in \mathbb{B}(\mathscr{X}_J)$, we define

$$\underline{f}(x_J, x_{J^c}) = \bar{f}(x_L) = f(x_J).$$

From (13) and the tower property for marginal measures, we obtain the chain of equalities:

$$\rho_{\mathscr{X}_J}[f, P_{\mathscr{X}_J}] = \rho[\underline{f}, P] = \rho_{\mathscr{X}_L}[\bar{f}, P_{\mathscr{X}_L}] = (\rho_{\mathscr{X}_L})_{\mathscr{X}_J}[f, (P_{\mathscr{X}_L})_{\mathscr{X}_J}] = (\rho_{\mathscr{X}_L})_{\mathscr{X}_J}[f, P_{\mathscr{X}_J}],$$

which is (iv). □

6 Risk in two-stage partially observable systems

6.1 Fixed observation distribution

Let us start from the following simple setting. A random vector (X, Y) is distributed in the product of Polish spaces $\mathscr{X} \times \mathscr{Y}$ according to a measure P. For a bounded measurable function $c : \mathscr{X} \times \mathscr{Y} \to \mathbb{R}$, we can evaluate the risk of $c(X, Y)$ by a risk form $\rho[c, P]$.

However, we know that we shall be able to observe the value of X. After X is observed, we might refine our risk evaluation of $c(X, Y)$. Thus, a question arises: if a future possibility to observe X exists, what should be our present evaluation of the risk of $c(X, Y)$, *before* X is observed. Note that Y is never observed.

Let us start with the problem of risk evaluation after X is observed. We can disintegrate P into its marginal $P_{\mathscr{X}}$ on $\mathscr{X}$ and a transition kernel $P_{\mathscr{Y}|\mathscr{X}}$ from $\mathscr{X}$ to $\mathscr{P}(\mathscr{Y})$:

$$P(dx, dy) = P_{\mathscr{X}}(dx) P_{\mathscr{Y}|\mathscr{X}}(dy|x).$$

Suppose the risk form ρ is monotonic, normalized, has the translation property and the support property. Then the correct evaluation of the risk after $X = x$ is observed is

$$\rho_{\mathscr{Y}|x}[c(x, \cdot), P_{\mathscr{Y}|\mathscr{X}}(x)] = \rho[c, \delta_x \otimes P_{\mathscr{Y}|\mathscr{X}}].$$

This is nothing else, but the conditional risk form defined in (11). As a function of x, we obtain the conditional risk operator $\rho_{\mathscr{Y}|\mathscr{X}}[c, P_{\mathscr{Y}|\mathscr{X}}]$. Now, to evaluate the overall risk, we calculate $\rho_{\mathscr{X}}[\rho_{\mathscr{Y}|\mathscr{X}}[c, P_{\mathscr{Y}|\mathscr{X}}], P_{\mathscr{X}}]$. We thus arrive to the following conclusion

from Theorem 2. If the risk form $\rho : \mathbb{B}(\mathscr{X} \times \mathscr{Y}) \times \mathscr{P}(\mathscr{X} \times \mathscr{Y}) \to \mathbb{R}$ is normalized, translation equivariant, conditionally consistent, and has the support property, then

$$\rho[c, P] = \rho_{\mathscr{X}}\Big[\rho_{\mathscr{Y}|\mathscr{X}}\big[c, P_{\mathscr{Y}|\mathscr{X}}\big], P_{\mathscr{X}}\Big].$$

It follows that the two risk evaluations: without and with the perspective of inspection, are identical. The mere existence of inspection does not affect risk.

When a possibility of control exists, the situation is different. Suppose there are two Polish spaces $\mathscr{U}_1$ and $\mathscr{U}_2$, which we call *control spaces*. At stage one, a control $u \in U_1 \subset \mathscr{U}_1$ is chosen, where U_1 is a subset of $\mathscr{U}_1$. Then an observation of X is made. After observing the value of X, we choose control $u_2 \in U_2(X, u_1) \subset \mathscr{U}_2$ to minimize the risk of $c(X, Y, u_1, u_2)$. The risk is measured by the form $\rho[\cdot, \cdot]$. Here $U_2 : \mathscr{X} \times \mathscr{U}_1 \rightrightarrows \mathscr{U}_2$ is a measurable multifunction representing the feasible set at the second stage. We shall use the symbol $\pi(\cdot) \lessdot U_2(\cdot, u_1)$ to indicate that the function π is a measurable selection of $U_2(\cdot, u_1)$.

We may look at this problem from two perspectives. Let us start from the functional perspective. Since u_2 can be chosen after X is observed, we may represent it as a measurable function: $u_2 = \pi(x)$, $x \in \mathscr{X}$. Therefore, the overall cost has the form:

$$Z^{u_1, \pi}(x, y) = c(x, y, u_1, \pi(x)), \quad (x, y) \in \mathscr{X} \times \mathscr{Y}. \tag{20}$$

The distribution of (X, Y) is P. The problem takes on the form

$$\begin{aligned}
\min_{u_1, \pi} \ & \rho\big[Z^{u_1, \pi}, P\big] \\
\text{s.t. } & u_1 \in U_1, \\
& \pi(\cdot) \lessdot U_2(\cdot, u_1).
\end{aligned} \tag{21}$$

We now derive a two-stage representation of this problem.

Theorem 4 *We assume the following:*

(i) The risk form ρ is monotonic, normalized, translation equivariant, has the support property, and is conditionally consistent;

(ii) The multifunction U_2 is upper-semicontinuous and has nonempty and compact values;

(iii) The function c is uniformly bounded, measurable, and lower-semicontinuous with respect to its second argument.

Then problem (21) *is equivalent to the two-stage problem:*

$$\min_{u_1 \in U_1} \rho_{\mathscr{X}}\big[V(\cdot, u_1), P_{\mathscr{X}}\big], \tag{22}$$

where $V(\cdot, \cdot)$ is the optimal value of the following second stage problem:

$$V(x, u_1) = \min_{u_2 \in U_2(x, u_1)} \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, u_2), P_{\mathscr{Y}|\mathscr{X}}(x)\big], \quad x \in \mathscr{X}, \quad u_1 \in U_1. \tag{23}$$

Proof Since $\rho[\cdot, \cdot]$ satisfies the assumptions of Theorem 2, the risk of the function $Z^{u_1,\pi}$ can be calculated by the risk disintegration formula:

$$\rho\big[Z^{u_1,\pi}, P\big] = \rho_{\mathscr{X}}\Big[\rho_{\mathscr{Y}|\mathscr{X}}\big[Z^{u_1,\pi}, P_{\mathscr{Y}|\mathscr{X}}\big], P_{\mathscr{X}}\Big]$$
$$= \rho_{\mathscr{X}}\Big[x \mapsto \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, \pi(x)), P_{\mathscr{Y}|\mathscr{X}}(x)\big], P_{\mathscr{X}}\Big].$$

Then problem (21) takes on the form:

$$\min_{u_1,\pi} \rho_{\mathscr{X}}\Big[x \mapsto \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, \pi(x)), P_{\mathscr{Y}|\mathscr{X}}(x)\big], P_{\mathscr{X}}\Big],$$

subject to the same constraints. Due to the monotonicity of the marginal risk form $\rho_{\mathscr{X}}$, the smaller the values of the function $x \mapsto \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, \pi(x))\big]$, the smaller the value of $\rho_{\mathscr{X}}$. Since $u_2 = \pi(x)$ may depend on x, we may carry out the minimization with respect to u_2 inside the argument of $\rho_{\mathscr{X}}$:

$$\inf_{u_1 \in U_1} \inf_{\pi(\cdot) \prec U_2(\cdot, u_1)} \rho_{\mathscr{X}}\Big[x \mapsto \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, \pi(x)), P_{\mathscr{Y}|\mathscr{X}}(x)\big], P_{\mathscr{X}}\Big]$$
$$= \inf_{u_1 \in U_1} \rho_{\mathscr{X}}\Big[x \mapsto \inf_{u_2 \in U_2(x, u_1)} \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, u_2), P_{\mathscr{Y}|\mathscr{X}}(x)\big], P_{\mathscr{X}}\Big].$$

The only condition for the validity of this transformation is the measurability and boundedness of the optimal value function (23). This follows from Berge's theorem (see, e.g., [3, Th. 1.4.16]), which can be applied due to the assumptions (ii) and (iii). In fact, they also guarantee that the optimal value function is lower semicontinuous with respect to u_1.

We conclude that problem (21) reduces to the marginal risk optimization (22). □

Theorem 4 provides us with the second perspective on the problem. It has a hierarchical structure, similar to its expected-value full information counterpart: after $X = x$ is observed, the second stage problem (23) is to minimize the conditional risk. Then, the first stage problem takes on the form of minimizing the marginal risk (22) of the second-stage optimal value. The most important conclusion is that the risk disintegration formula allows us to write the overall problem in a hierarchical structure. The extended two-stage risk-averse model, which is introduced and analyzed in [25] (see also [18]) is a special case of this problem.

6.2 Controlled observation distribution

Now we consider a more complex situation. There are still two Polish control spaces $\mathscr{U}_1$ and $\mathscr{U}_2$. However, after a control $u_1 \in U_1 \subset \mathscr{U}_1$ is chosen, the distribution of the observation X depends on u_1. The dependence is described by a controlled kernel $K : \mathscr{Y} \times \mathscr{U}_1 \to \mathscr{P}(\mathscr{X})$. After observing X, we choose control $u_2 \in U_2(X, u_1) \subset \mathscr{U}_2$ to minimize the risk of $c(X, Y, u_1, u_2)$. The risk is measured by the form $\rho[\cdot, \cdot]$.

Assume the same conditions on U_1, U_2 and ρ as in the previous subsection. Let P_Y be the marginal distribution of Y. After the first decision u_1 will be chosen, the joint distribution of (Y, X) will become

$$M(u_1) = P_{\mathscr{Y}} \otimes K(\cdot, u_1),$$

that is, $M(dy, dx | u_1) = P_{\mathscr{Y}}(dy) K(dx | y, u_1)$. Therefore, denoting the second stage decision by $u_2 = \pi(x)$ (it may depend on x), our problem is to find

$$
\begin{aligned}
&\min_{u_1, \pi} \rho\big[Z^{u_1, \pi}, M(u_1)\big], \\
&\text{s.t. } u_1 \in U_1, \\
&\quad\quad \pi(\cdot) \prec U_2(\cdot, u_1).
\end{aligned}
\tag{24}
$$

where the function $Z^{u_1, \pi}(\cdot, \cdot)$ is given by (20).

Let us develop a two-stage version of the functional problem (24). The marginal distribution of the observation result is

$$M_{\mathscr{X}}(u_1) = \int_{\mathscr{Y}} K(y, u_1)\, P_{\mathscr{Y}}(dy)$$

where the integral is understood in the weak sense. Since the space $\mathscr{Y}$ is standard, the measure $M(u_1)$ can be disintegrated into the marginal $M_{\mathscr{X}}(u_1)$ and a transition kernel $\Gamma : \mathscr{X} \times \mathscr{U}_1 \to \mathscr{P}(\mathscr{Y})$ as follows

$$M(u_1) = M_{\mathscr{X}}(u_1) \otimes \Gamma(u_1),$$

which reads $M(dx, dy | u_1) = M_{\mathscr{X}}(dx | u_1) \Gamma(dy | x, u_1)$. The transition kernel Γ is called the *Bayes operator*.

Example 2 Assume that the joint distribution $M(u_1)$ of (X, Y) has a density $q(\cdot, \cdot \mid u_1)$ with respect to a finite product measure $\mu_{\mathscr{X}} \otimes \mu_{\mathscr{Y}}$ on $\mathscr{X} \times \mathscr{Y}$. Then the Bayes operator has the form

$$\Gamma(A | x, u_1) = \frac{\int_A \int_{\mathscr{Y}} q(x', y' \mid u_1)\, M_{\mathscr{Y}}(dy)\, \mu_{\mathscr{Y}}(dy')}{\int_{\mathscr{Y}} \int_{\mathscr{Y}} q(x', y' \mid u_1)\, M_{\mathscr{Y}}(dy)\, \mu_{\mathscr{Y}}(dy')}, \quad \forall A \in \mathscr{B}(\mathscr{Y}).$$

If the formula above has a zero denominator, we can formally define $\Gamma(x, u_1)$ to be an arbitrarily selected distribution on $\mathscr{Y}$. ☐

With the use of the Bayes operator, we can equivalently write problem (24) as a two-stage problem.

Theorem 5 *We assume the following:*

(i) The risk form ρ is monotonic, normalized, translation equivariant, conditionally consistent, and has the support property;

(ii) The multifunction U_2 is upper-semicontinuous and has nonempty and compact values;

(iii) The function c is uniformly bounded, measurable, and lower-semicontinuous with respect to its second argument.

Then problem (24) *is equivalent to the two-stage problem:*

$$\min_{u_1 \in U_1} \rho_{\mathscr{X}}\big[V(\cdot, u_1), P_{\mathscr{X}}\big],$$

where $V(\cdot, \cdot)$ is the optimal value of the following second stage problem:

$$V(x, u_1) = \min_{u_2 \in U_2(x, u_1)} \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, u_2), \Gamma(x, u_1)\big], \quad x \in \mathscr{X}, \quad u_1 \in U_1.$$

Proof With the use of the Bayes formula, we can write problem (24) as follows:

$$\min_{u_1 \in U_1} \min_{\pi(\cdot) \prec U_2(\cdot, u_1)} \rho\big[Z^{u_1, \pi}, P_{\mathscr{X}}(u_1) \otimes \Gamma(u_1)\big].$$

Since the risk form ρ satisfies the assumptions of Theorem 2, we can disintegrate it to obtain the following equivalent form:

$$\min_{u_1 \in U_1} \min_{\pi(\cdot) \prec U_2(\cdot, u_1)} \rho_{\mathscr{X}}\Big[x \mapsto \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, \pi(x)), \Gamma(x, u_1)\big], P_{\mathscr{X}}(u_1)\Big].$$

The remaining considerations are the same as in the proof of Theorem 4. Due to the monotonicity of the marginal risk form $\rho_{\mathscr{X}}$, the smaller the values of the function $x \mapsto \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, \pi(x)), \Gamma(x, u_1)\big]$, the smaller the value of $\rho_{\mathscr{X}}$. Since $u_2 = \pi(x)$ may depend on x, we may carry out the minimization with respect to u_2 inside the argument of $\rho_{\mathscr{X}}$:

$$\inf_{u_1 \in U_1} \inf_{\pi(\cdot) \prec U_2(\cdot, u_1)} \rho_{\mathscr{X}}\Big[x \mapsto \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, \pi(x)), \Gamma(x, u_1)\big], P_{\mathscr{X}}(u_1)\Big]$$

$$= \inf_{u_1 \in U_1} \rho_{\mathscr{X}}\Big[x \mapsto \inf_{u_2 \in U_2(x, u_1)} \rho_{\mathscr{Y}|x}\big[c(x, \cdot, u_1, u_2), \Gamma(x, u_1)\big], P_{\mathscr{X}}(u_1)\Big].$$

The only difference is that the new marginal distribution and the Bayes operator are the disintegration components of the probability measure and feature in the risk disintegration formula. The "inf" operation in the second stage problem can be replaced by "min" because of conditions (ii) and (iii). □

7 Conclusions

Our work initiates systematic research of risk measures considered as functionals of two arguments: a function on a space $\mathscr{X}$ and an underlying probability measure on the σ-algebra of Borel subsets of $\mathscr{X}$. Such functionals, which we call risk forms, occur in two- and multi-stage optimization models, in which the probability measure

depends on decisions, and in models, in which only partial observation is available and Bayesian updates of the probability measure are employed.

Two main results: the Kusuoka representation and the risk disintegration formula, generalize the classical properties of risk measures to the new setting. The derivation of the dual and Kusuoka representations hinges on novel duality theory for functionals of quantile functions [8]. The risk disintegration formula uses a new concept of conditional consistency.

For both contributions, essential is the boundedness of the functions under consideration. It allows us to consider arbitrary probability measures on the spaces involved. In the first group of results, duality between bounded functions and finitely additive measures plays a role; consistency with the increasing convex order allows to pass to countably additive measures. In the second group of results, the very concept of conditional consistency uses Dirac delta measures (in products with conditional measures), which are natural in a space of bounded functions.

The theory of risk measures is well-developed in the spaces of integrable functions. A fixed underlying probability measure is essential for defining the space of functions. A fundamental challenge is to extend the theory of risk forms to important classes of unbounded functions. It may require to precisely define both domains of the risk form: a broader class of functions and a narrower class of probability measures, so that properties of the forms can be preserved. In particular, one could conjecture that a generalized Kusuoka representation could be developed in such a broader setting, but this requires further research.

References

1. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. Math. Financ. **9**, 203–228 (1999)
2. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., Ku, H.: Coherent multiperiod risk adjusted values and Bellman's principle. Ann. Oper. Res. **152**, 5–22 (2007)
3. Aubin, J.-P., Frankowska, H.: Set-Valued Analysis. Birkhäuser, Boston (2009)
4. Bielecki, T.R., Cialenco, I., Pitera, M.: A survey of time consistency of dynamic risk measures and dynamic performance measures in discrete time: Lm-measure perspective. Probab., Uncertain. Quant. Risk **2**(1), 3 (2017)
5. Cheridito, P., Delbaen, F., Kupper, M.: Dynamic monetary risk measures for bounded discrete-time processes. Electron. J. Probab. **11**, 57–106 (2006)
6. Cheridito, P., Kupper, M.: Composition of time-consistent dynamic monetary risk measures in discrete time. Int. J. Theor. Appl. Financ. **14**(01), 137–162 (2011)
7. Dentcheva, D., Ruszczyński, A.: Common mathematical foundations of expected utility and dual utility theories. SIAM J. Optim. **23**(1), 381–405 (2013)
8. Dentcheva, D., Ruszczyński, A.: Risk preferences on the space of quantile functions. Math. Program. **148**(1–2), 181–200 (2014)
9. Dentcheva, D., Ruszczyński, A.: Risk-averse control of continuous-time Markov chains. In: 2017 Proceedings of the Conference on Control and its Applications, pp. 78–85. SIAM, Philadelphia (2017)
10. Dentcheva, D., Ruszczyński, A.: Time-coherent risk measures for continuous-time Markov chains. SIAM J. Financ. Math. **9**(2), 690–715 (2018)
11. Fan, J., Ruszczyński, A.: Process-based risk measures and risk-averse control of discrete-time systems. Math. Program. (2018). https://doi.org/10.1007/s10107-018-1349-2

12. Fan, J., Ruszczyński, A.: Risk measurement and risk-averse control of partially observable discrete-time Markov systems. Math. Methods Oper. Res. **88**, 161–184 (2018)
13. Föllmer, H., Penner, I.: Convex risk measures and the dynamics of their penalty functions. Stat. Decis. **24**(1/2006), 61–96 (2006)
14. Föllmer, H., Schied, A.: Convex measures of risk and trading constraints. Financ. Stoch. **6**, 429–447 (2002)
15. Föllmer, H., Schied, A.: Stoch. Financ. Walter de Gruyter & Co., Berlin (2011)
16. Frittelli, M., Rosazza Gianin, E.: Law invariant convex risk measures. In: Advances in Mathematical Economics, vol. 7, pp. 33–46. Springer, Tokyo (2005)
17. Littlewood, J.E., Hardy, G.H., Pólya, G.: Inequalities. Cambridge University Press, Cambridge (1934)
18. Gülten, S., Ruszczyński, A.: Two-stage portfolio optimization with higher-order conditional measures of risk. Ann. Oper. Res. **229**(1), 409–427 (2015)
19. Jobert, A., Rogers, L.C.G.: Valuations and dynamic convex risk measures. Math. Financ. **18**(1), 1–22 (2008)
20. Jonsbråten, T.W., Wets, R.J.-B., Woodruff, D.L.: A class of stochastic programs with decision dependent random elements. Ann. Oper. Res. **82**, 83–106 (1998)
21. Jouini, E., Schachermayer, W., Touzi, N.: Law invariant risk measures have the Fatou property. In: Advances in Mathematical Economics, vol. 9, pp. 49–71. Springer, Tokyo (2006)
22. Kijima, M., Ohnishi, M.: Mean-risk analysis of risk aversion and wealth effects on optimal portfolios with multiple investment opportunities. Ann. Oper. Res. **45**, 147–163 (1993)
23. Klöppel, S., Schweizer, M.: Dynamic indifference valuation via convex risk measures. Math. Financ. **17**(4), 599–627 (2007)
24. Kusuoka, S.: On law-invariant coherent risk measures. In: Kusuoka, S., Maruyama, T. (eds.) Advances in Mathematical Economics, vol. 3, pp. 83–95. Springer, Tokyo (2001)
25. Miller, N., Ruszczyński, A.: Risk-averse two-stage stochastic linear programming: modeling and decomposition. Oper. Res. **59**(1), 125–132 (2011)
26. Noyan, N., Rudolf, G.: Kusuoka representations of coherent risk measures in general probability spaces. Ann. Oper. Res. **229**(1), 591–605 (2015)
27. Ogryczak, W., Ruszczyński, A.: From stochastic dominance to mean-risk models: semideviations as risk measures. Eur. J. Oper. Res. **116**(1), 33–50 (1999)
28. Parthasarathy, K.R.: Probability measures on metric spaces. American Mathematical Society, Providence (RI) (2005)
29. Pflug, G.C., Pichler, A.: Multistage Stochastic Optimization. Springer, Berlin (2016)
30. Pflug, G.C., Römisch, W.: Modeling, Measuring and Managing Risk. World Scientific, Singapore (2007)
31. Quiggin, J.: A theory of anticipated utility. J. Econ. Behav. Organ. **3**, 323–343 (1982)
32. Riedel, F.: Dynamic coherent risk measures. Stoch. Process. Appl. **112**, 185–200 (2004)
33. Roorda, B., Schumacher, J.M., Engwerda, J.: Coherent acceptability measures in multiperiod models. Math. Financ. **15**(4), 589–612 (2005)
34. Ruszczyński, A.: Risk-averse dynamic programming for Markov decision processes. Math. Program. **125**(2), 235–261 (2010)
35. Ruszczyński, A., Shapiro, A.: Conditional risk mappings. Math. Oper. Res. **31**, 544–561 (2006)
36. Ruszczyński, A., Shapiro, A.: Optimization of convex risk functions. Math. Oper. Res. **31**, 433–542 (2006)
37. Scandolo, G.: Risk Measures in a Dynamic Setting. PhD thesis, Università degli Studi di Milano, Milan (2003)
38. Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on Stochastic Programming. SIAM, Philadelphia (2009)
39. Yaari, M.E.: The dual theory of choice under risk. Econometrica **55**, 95–115 (1987)