



Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Time-Consistent Decisions and Temporal Decomposition of Coherent Risk Functionals

Georg Ch. Pflug, Alois Pichler

To cite this article:

Georg Ch. Pflug, Alois Pichler (2016) Time-Consistent Decisions and Temporal Decomposition of Coherent Risk Functionals. Mathematics of Operations Research 41(2):682-699. <https://doi.org/10.1287/moor.2015.0747>

Full terms and conditions of use: <http://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2016, INFORMS

Please scroll down for article—it is on subsequent pages



INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Time-Consistent Decisions and Temporal Decomposition of Coherent Risk Functionals

Georg Ch. Pflug

University of Vienna, A-1090 Vienna, Austria; and International Institute for Applied Systems Analysis (IIASA),
A-2361 Laxenburg, Austria, georg.pflug@univie.ac.at

Alois Pichler

Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway, aloisp@NTNU.no

In management and planning it is commonplace for additional information to become available gradually over time. It is well known that most risk measures (risk functionals) are time *inconsistent* in the following sense: it may happen that at a given time period, some loss distribution appears to be less risky than another one, but looking at the conditional distribution at a later time, the opposite relation holds almost surely.

The extended conditional risk functionals introduced in this paper enable a temporal decomposition of the initial risk functional that can be used to ensure consistency between past and future preferences. The central result is a decomposition theorem, which allows recomposing the initial coherent risk functional by compounding the conditional risk functionals without losing information or preferences. It follows from our results that the revelation of partial information in time must change the decision maker's preferences—for consistency reasons—among the remaining courses of action. Further, in many situations, the extended conditional risk functional allows ranking of different policies, even based on incomplete information.

In addition, we use counterexamples to show that without change-of-measures, the only time-consistent risk functionals are the expectation and the essential supremum.

Keywords: risk measures; time consistency; convex duality

MSC2000 subject classification: 90C15, 60B05, 62P05

OR/MS subject classification: Primary: dynamic programming/optimal control: deterministic; secondary: statistics: nonparametric

History: Received October 17, 2012; revised July 25, 2013, February 4, 2014, December 29, 2014, May 20, 2015.

Published online in *Articles in Advance* January 22, 2016.

1. Introduction. Coherent risk measures were introduced in the pioneering paper by Artzner et al. [4], and they have become increasingly important since then. They are used to measure and quantify risk associated with random outcomes.

In many problems some of the relevant information is not available a priori, and it becomes known gradually over time as the overall amount information steadily increases. When new information becomes available, an intelligent decision maker might wish to adjust his policy in order to keep track of the original goal. But how can this be accomplished? Is this possible by sticking to the initial objective? How to update the objective relative to the new information is the main topic of this paper.

There are simple examples showing that sticking to the same, genuine risk functional at all decision stages may lead to contradicting, conflicting, and even wrong decisions. Such an example is contained in Figure 1(a). So how does the decision maker have to change his objective given the information that was already revealed?

In this paper we introduce an extended notion of conditional risk functionals, which reflects the information that is already revealed. In many situations the extended conditional risk functional allows ranking random variables based on incomplete information, i.e., based on the conditional observation. This is derived from a key property of the conditional risk functionals that allows a recomposition of the initial (positively homogeneous) risk functional. The decomposition is given for the Average Value-at-Risk first, and then extended to general risk functionals. The presented concept is an alternative to dynamic extensions or artificial compositions of risk functionals, which are difficult to interpret and thus difficult to justify in many practical situations.

Relation to other work and outline. Conditional versions of risk measures have been investigated earlier in a context related to time consistency (see, among others, the papers by Cheridito and Kupper [9, 10]; Kovacevic and Pflug [16], Pflug and Römisch [23]; and other authors). The results by Kupper and Schachermayer [17] as well as by Shapiro [33] indicate that usual notions of time consistency are probably too restrictive: risk functionals, which are time consistent in the strict sense imposed by these papers, are only the expectation and the max-risk functional. The latter section of this paper (§7) investigates a further concept of consistency that leads to the same type of negative results.

To obtain time consistency, desirable properties thus have to be relaxed (for example, law invariance in the paper by Miller and Ruszczyński [19]) or different concepts have to be investigated. This paper demonstrates

a method that allows reassembling the risk given the information observed in a way, which is consistent with the past and the initial objective. The correction, which consists in passing to a new conditional risk functional, reflects the information already available. The conditional risk functional differs from the initial risk functional, but it retains its general characteristics. The conditional risk functionals, however, depend on the problem itself, and they can be applied a posteriori. They do not constitute an a priori rule.

Outline of the paper. The decomposition of the Average Value-at-Risk is elaborated first and employed to establish the general theory. The strategy to develop the respective decomposition of version-independent risk functionals involves Kusuoka's representation of version-independent risk functionals.

We develop the essential terms in §2. The dual representation of risk functionals (Theorem 9), which is elaborated in §3, is the first result of this paper and an essential component to defining the conditional risk functional in §5. The main result is the decomposition theorem (Theorem 21) for general risk functionals.

The last section finally addresses the aspect of consistency in the context of decision making and adds a negative statement on restrictive concepts of consistency. The appendix contains an instructive example to illustrate and outline the results.

2. Kusuoka's representation and spectral risk functionals. Throughout the paper, we shall investigate coherent risk functionals, as they were conceptually introduced and discussed first by Artzner et al. in the papers Artzner et al. [3, 4, 5]. The definitions are stated here because they are not used consistently in the literature.

DEFINITION 1. Let $L^\infty(\Omega, \mathcal{F}, P)$ (or simply L^∞) be the space of all essentially bounded, \mathbb{R} -valued random variables on a probability space with probability measure P and sigma algebra \mathcal{F} .

A coherent *risk functional* is a mapping $\mathcal{R}: L^\infty \rightarrow \mathbb{R}$ with the following properties:

- (i) **MONOTONICITY:** $\mathcal{R}(Y_1) \leq \mathcal{R}(Y_2)$ whenever $Y_1 \leq Y_2$ almost surely;
- (ii) **CONVEXITY:** $\mathcal{R}((1 - \lambda)Y_0 + \lambda Y_1) \leq (1 - \lambda)\mathcal{R}(Y_0) + \lambda\mathcal{R}(Y_1)$ for $0 \leq \lambda \leq 1$;
- (iii) **TRANSLATION EQUIVARIANCE:**¹ $\mathcal{R}(Y + c) = \mathcal{R}(Y) + c$ if $c \in \mathbb{R}$;
- (iv) **POSITIVE HOMOGENEITY:** $\mathcal{R}(\lambda Y) = \lambda \cdot \mathcal{R}(Y)$ whenever $\lambda > 0$.

In the literature the term coherent is related to condition (iv), positive homogeneity. Moreover, the mapping $\rho: Y \mapsto \mathcal{R}(-Y)$ is often called *coherent risk functional* instead of \mathcal{R} , and Y is associated with a profit rather than a loss: whereas \mathcal{R} is natural and more frequent in an actuarial (insurance) context, ρ is typically used in a banking context. In this paper we focus on coherent risk functionals and we use the terms *risk functional* and *coherent risk functional* synonymously.

The term *acceptability functional* is related to risk measures as well; it is frequently employed to identify acceptable strategies in a decision or optimization process: the acceptability functional is the concave mapping $\mathcal{A}: Y \mapsto -\mathcal{R}(-Y)$. Moreover, the domain for the risk measure considered here is L^∞ . This is for the simplicity of presentation; for extensions to a larger domain than L^p ($p \geq 1$), we refer to Pichler [24].

Average Value-at-Risk. The most well-known risk functional satisfying all axioms of Definition 1 is the (upper) *Average Value-at-Risk* at level α , which is given by

$$AV@R_\alpha(Y) := \frac{1}{1 - \alpha} \int_\alpha^1 V@R_p(Y) dp, \quad (0 \leq \alpha < 1)$$

and

$$AV@R_1(Y) := \lim_{\alpha \nearrow 1} AV@R_\alpha(Y) = \text{ess sup}(Y),$$

where

$$V@R_\alpha(Y) := \inf\{y: P(Y \leq y) \geq \alpha\}$$

is the *Value-at-Risk* (the left-continuous, lower semicontinuous *quantile* or *lower inverse cdf*) at level α , often denoted $V@R_\alpha(Y) = F_Y^{-1}(\alpha)$ as well as $(F_Y(y) = P(Y \leq y))$, is Y 's cdf).

DEFINITION 2. A risk functional \mathcal{R} is *version independent*² if $\mathcal{R}(Y_1) = \mathcal{R}(Y_2)$ whenever Y_1 and Y_2 share the same law; that is, $P(Y_1 \leq y) = P(Y_2 \leq y)$ for all $y \in \mathbb{R}$.

The Average Value-at-Risk is an elementary, version-independent risk functional:

¹ In an economic or monetary environment, this is often called CASH INVARIANCE instead.

² Often it is also *law invariant* or *distribution based*.

THEOREM 3 (KUSUOKA'S REPRESENTATION). *Let \mathcal{B} collect the Borel sets on $[0, 1]$ and the probability measure P be atomless. Then any version independent, coherent risk functional \mathcal{R} on $L^\infty([0, 1], \mathcal{B}, P)$ has the representation*

$$\mathcal{R}(Y) = \sup_{\mu \in \mathcal{M}} \int_0^1 \text{AV@R}_\alpha(Y) \mu(d\alpha), \quad (1)$$

where \mathcal{M} is a set of probability measures on $[0, 1]$.

PROOF. See Kusuoka's [18] original paper or Shapiro [34], in combination with the paper by Jouini et al. [14]. \square

Spectral risk functionals. A subclass of risk functionals is given when the set of measures \mathcal{M} in (1) reduces to a singleton, $\mathcal{M} = \{\mu\}$. This class can be equivalently described as the class of comonotone additive functionals. In this paper, we call them (following Acerbi; see Acerbi [1], Acerbi and Simonetti [2]) *spectral risk functionals*, although the term *distortion risk functionals* is in frequent use as well.

An example is the Average Value-at-Risk at level $\alpha < 1$, which is provided by the Dirac measure $\mathcal{M} = \{\delta_\alpha\}$. Assuming that

$$\sigma_\mu(u) := \int_0^u \frac{\mu(d\alpha)}{1-\alpha} \quad (2)$$

is well defined,³ we can use integration by parts to show that⁴

$$\int_0^1 \text{AV@R}_\alpha(Y) \mu(d\alpha) = \int_0^1 \text{V@R}_p(Y) \sigma_\mu(p) dp.$$

DEFINITION 4. The risk functional with representation

$$\mathcal{R}_\sigma(Y) := \int_0^1 \text{V@R}_p(Y) \sigma(p) dp$$

is called a *spectral risk functional*. The function σ is called the *spectral density*.

It follows from the axioms imposed on a risk functional in Definition 1 that σ is necessarily a nondecreasing probability density on $[0, 1]$ (details are elaborated in Pflug [22]).

It is essential to note that every risk functional \mathcal{R} has a representation in terms of spectral risk functionals, in a similar way as Kusuoka's representation involves mixtures of AV@Rs.

COROLLARY 5 (KUSUOKA REPRESENTATION IN TERMS OF SPECTRAL RISK FUNCTIONALS). *For every version-independent, coherent risk functional \mathcal{R} as in Theorem 3 there is a set \mathcal{S} of spectral densities such that*

$$\mathcal{R}(Y) = \sup_{\sigma \in \mathcal{S}} \mathcal{R}_\sigma(Y). \quad (4)$$

\mathcal{S} may be assumed to consist of strictly increasing, bounded, and continuous functions only.

If the relation (4) holds, we write $\mathcal{R} = \mathcal{R}_\sigma$ and call \mathcal{R} the risk functional induced by \mathcal{S} . Notice that \mathcal{S} is not uniquely determined by \mathcal{R}_σ because, for example, $\text{AV@R}_\alpha(Y) = \sup_{\alpha' < \alpha} \text{AV@R}_{\alpha'}(Y)$.

3. The dual of risk functionals. The decomposition of risk functionals with respect to incomplete information employs the dual or convex conjugate function (see the book by Shapiro et al. [35] and paper by Ruszczyński and Shapiro [30]). As any convex, version-independent risk functional $\mathcal{R}: L^\infty \rightarrow \mathbb{R}$ is Lipschitz continuous (thus lower semicontinuous; see Jouini et al. [14]) the Fenchel–Moreau duality theorem can be stated in the following way:

³ For the Average Value at risk at level α , particularly,

$$\sigma(\cdot) = \frac{1}{1-\alpha} \mathbb{1}_{[\alpha, 1]}(\cdot). \quad (3)$$

⁴ The inverse of this operation is given by the measure with distribution function $\mu_\sigma(\alpha) := (1-\alpha)\sigma(\alpha) + \int_0^\alpha \sigma(u) du$.

THEOREM 6 (FENCHEL–MOREAU DUALITY THEOREM). *The risk functional $\mathcal{R}: L^\infty \rightarrow \mathbb{R}$ has the representation*

$$\mathcal{R}(Y) = \sup\{\mathbb{E}(YZ) - \mathcal{R}^*(Z): Z \in L^1\}, \quad (5)$$

where the convex conjugate $\mathcal{R}^*: L^1 \rightarrow \mathbb{R} \cup \{\infty\}$ is defined by

$$\mathcal{R}^*(Z) = \sup\{\mathbb{E}(YZ) - \mathcal{R}(Y): Y \in L^\infty\}, \quad (Z \in L^1).$$

Positive homogeneity of the risk functional \mathcal{R} is moreover equivalent to the fact that \mathcal{R}^* takes only the values 0 and ∞ ; i.e., Z is a feasible dual variable iff $\mathcal{R}^*(Z) = 0$.

Corollary 5 builds an arbitrary risk functional \mathcal{R} by means of spectral risk functionals. In order to formulate the decomposition result in a stringent way, we recall the duality of spectral risk functionals. For this the notion of *convex ordering* is useful. We follow the terms introduced in Shapiro [34]; for the concept of stochastic dominance relations in general, we refer to the books by Stoyan and Müller [20] and Shaked and Shanthikumar [31] as well as to the papers by Ogryczak and Ruszczyński [21] and Ruszczyński and Dentcheva [29].

DEFINITION 7 (SECOND ORDER STOCHASTIC DOMINANCE, OR CONVEX ORDERING). Let $\tau, \sigma: (0, 1) \rightarrow [0, \infty)$ be integrable functions. It is said that σ majorizes τ (denoted $\sigma \succcurlyeq \tau$ or $\tau \preccurlyeq \sigma$) iff

$$\int_\alpha^1 \tau(p) dp \leq \int_\alpha^1 \sigma(p) dp \quad \text{for all } \alpha \in [0, 1] \quad \text{and} \quad \int_0^1 \tau(p) dp = \int_0^1 \sigma(p) dp = 1.$$

REMARK 8. For the inverse quantile function F_Z^{-1} of a nonnegative random variable Z with expectation 1, i.e., satisfying

$$\mathbb{E}(Z) = \int_0^1 \sigma(p) dp = 1, \quad Z \geq 0,$$

the relation $F_Z^{-1} \preccurlyeq \sigma$ is notably equivalent to

$$(1 - \alpha)AV@R_\alpha(Z) \leq \int_\alpha^1 \sigma(p) dp \quad \text{for all } \alpha \in [0, 1].$$

The following theorem characterizes the risk functional by employing the stochastic order relation. It can be interpreted as a consequence of consistency of version-independent risk measures with respect to second order stochastic dominance (see Ogryczak and Ruszczyński [21]).

THEOREM 9 (DUAL REPRESENTATION OF VERSION-INDEPENDENT RISK FUNCTIONALS). *Let $\mathcal{R}_\mathcal{S}$ be a version-independent, positively homogeneous risk functional induced by the set \mathcal{S} of spectral functions. Then*

$$\mathcal{R}_\mathcal{S}(Y) = \sup\{\mathbb{E}(YZ): \text{there exists a spectral density } \sigma \in \mathcal{S} \text{ such that } F_Z^{-1} \preccurlyeq \sigma\}; \quad (6)$$

i.e., the supremum is among all variables $Z \geq 0$ for which there is a $\sigma \in \mathcal{S}$ with $F_Z^{-1} \preccurlyeq \sigma$.

PROOF. The dual representation

$$\begin{aligned} \mathcal{R}_\sigma(Y) &= \sup\{\mathbb{E}(YZ): F_Z^{-1} \preccurlyeq \sigma\} \\ &= \sup\left\{\mathbb{E}(YZ): \mathbb{E}(Z) = 1, (1 - \alpha)AV@R_\alpha(Z) \leq \int_\alpha^1 \sigma(p) dp, 0 \leq \alpha < 1\right\} \end{aligned} \quad (7)$$

for spectral risk functionals can be derived from Shapiro [34] and is contained in Pichler [25]. The assertion for general risk functionals $\mathcal{R}_\mathcal{S}$ is immediate now from the definition as

$$\begin{aligned} \mathcal{R}_\mathcal{S}(Y) &= \sup_{\sigma \in \mathcal{S}} \mathcal{R}_\sigma(Y) \\ &= \sup_{\sigma \in \mathcal{S}} \sup_Z \left\{\mathbb{E}(YZ): \mathbb{E}(Z) = 1, (1 - \alpha)AV@R_\alpha(Z) \leq \int_\alpha^1 \sigma(p) dp, 0 \leq \alpha < 1\right\} \\ &= \sup\left\{\mathbb{E}(YZ): \mathbb{E}(Z) = 1, \exists \sigma \in \mathcal{S} \forall \alpha \in [0, 1]: (1 - \alpha)AV@R_\alpha(Z) \leq \int_\alpha^1 \sigma(p) dp\right\} \end{aligned}$$

by using (7). \square

4. Convex duality. The dual representation (Equation (6)) suggests considering the following function as a candidate for the dual representation:

$$\tilde{\mathcal{R}}_{\mathcal{F}}(Z) := \begin{cases} 0 & \text{if } \exists \sigma \in \mathcal{S}: (1 - \alpha)AV@R_{\alpha}(Z) \leq \int_{\alpha}^1 \sigma(p) dp \text{ for all } 0 < \alpha \leq 1, \\ +\infty & \text{otherwise,} \end{cases}$$

for which $\mathcal{R}_{\mathcal{F}}(Y) = \sup_Z \{\mathbb{E}(YZ) - \tilde{\mathcal{R}}_{\mathcal{F}}(Z)\}$. For this function it follows that

$$\begin{aligned} \mathcal{R}_{\mathcal{F}}^*(Z) &= \sup_Y \{\mathbb{E}(YZ) - \mathcal{R}_{\mathcal{F}}(Y)\} = \sup_Y \left\{ \mathbb{E}(YZ) - \left(\sup_{Z'} \{\mathbb{E}YZ' - \tilde{\mathcal{R}}_{\mathcal{F}}(Z')\} \right) \right\} \\ &\leq \sup_Y \{\mathbb{E}(YZ) - (\mathbb{E}(YZ) - \tilde{\mathcal{R}}_{\mathcal{F}}(Z))\} = \tilde{\mathcal{R}}_{\mathcal{F}}(Z). \end{aligned}$$

The inequality $\mathcal{R}_{\mathcal{F}}^*(Z) \leq \tilde{\mathcal{R}}_{\mathcal{F}}(Z)$ may also be strict. Equality, however, can be obtained by augmenting the spectrum \mathcal{S} in an appropriate way. The following corollary characterizes the dual $\mathcal{R}_{\mathcal{F}}^*$ precisely by an augmented spectrum \mathcal{S}^* such that $\mathcal{R}_{\mathcal{F}}^*(Z) = \tilde{\mathcal{R}}_{\mathcal{F}^*}(Z)$ holds true.

COROLLARY 10. *It holds that $\mathcal{R}_{\mathcal{F}}^*(Z) = \tilde{\mathcal{R}}_{\mathcal{F}^*}(Z)$, where \mathcal{S}^* is the set*

$$\begin{aligned} \mathcal{S}^* &= \text{conv} \left\{ \tau: [0, 1] \rightarrow [0, \infty) \mid \tau \text{ is nondecreasing, lower semicontinuous, and } \exists \sigma \in \mathcal{S}: \tau \preceq \sigma \right\} \\ &= \text{conv} \left\{ \tau: [0, 1] \rightarrow [0, \infty) \mid \begin{array}{l} \tau \text{ is nondecreasing, lower semicontinuous, } \int_0^1 \tau(p) dp = 1, \text{ and } \exists \sigma \in \mathcal{S}: \\ \int_{\alpha}^1 \tau(p) dp \leq \int_{\alpha}^1 \sigma(p) dp \text{ for all } 0 \leq \alpha \leq 1, \end{array} \right\} \end{aligned}$$

of spectral functions.

PROOF. Define

$$\mathcal{Z}_{\mathcal{F}} := \{Z \in L^1: (\exists \sigma \in \mathcal{S}): F_Z^{-1} \preceq \sigma\}$$

and observe that $s_{\mathcal{Z}_{\mathcal{F}}}(Y) = \mathcal{R}_{\mathcal{F}}(Y)$, where $s_C(Y) := \sup_{Z^* \in C} Z^*(Y)$ is the *support function* of the set C (a subset of the dual). It follows by the Rockafellar–Fenchel–Moreau-duality Theorem (see Rockafellar [26]) that $s_{\mathcal{Z}_{\mathcal{F}}}^* = \mathbb{I}_{\overline{\text{conv } \mathcal{Z}_{\mathcal{F}}}}$, where \mathbb{I} is the indicator function of its index set,⁵ and hence $\mathcal{R}_{\mathcal{F}}^* = \mathbb{I}_{\overline{\text{conv } \mathcal{Z}_{\mathcal{F}}}}$ and $\mathcal{R}_{\mathcal{F}}(Y) = \sup_{Z \in \overline{\text{conv } \mathcal{Z}_{\mathcal{F}}}} \mathbb{E}(YZ)$. We show that $\mathcal{Z}_{\mathcal{F}} = \mathcal{Z}_{\mathcal{F}^*}$.

Indeed, suppose that $Z \in \mathcal{Z}_{\mathcal{F}^*}$, that is there is a function $\tau \in \mathcal{S}^*$ such that $F_Z^{-1} \preceq \tau$. But as $\tau \in \mathcal{S}^*$ there is $\sigma \in \mathcal{S}$ with $F_Z^{-1} \preceq \tau \preceq \sigma$, which shows that $Z \in \mathcal{Z}_{\mathcal{F}}$ and hence $\mathcal{Z}_{\mathcal{F}} \supset \mathcal{Z}_{\mathcal{F}^*}$.

Moreover $\mathcal{Z}_{\mathcal{F}} \subset \mathcal{Z}_{\mathcal{F}^*}$ by considering the lower semicontinuous and increasing function $\sigma_Z(\alpha) = V@R_{\alpha}(Z)$ for any $Z \in \mathcal{Z}_{\mathcal{F}}$.

The assertion finally follows from $\text{conv } \mathcal{Z}_{\mathcal{F}} = \mathcal{Z}_{\text{conv } \mathcal{F}}$, which completes the proof. \square

REMARK 11. In the present context of increasing functions the requirement *lower semicontinuous* for $\tau \in \mathcal{S}^*$ is equivalent to τ *continuous from the left*. Moreover, the constant function $\tau \equiv \mathbb{I}$ is always contained in the set \mathcal{S}^* ; the constant function $\tau \equiv \mathbb{I}$ is associated with the expected value.

REMARK 12. The closure of the convex set $\text{conv } \mathcal{S}^*$ coincides (by Mazur's Theorem on convex sets; see Wojtaszczyk [38, II.A.4]) with the $\sigma(L^1, L^\infty)$ (or weak) closure.

Feasible dual variables. The following two remarks prepare for the main theorem of this section. We shall write $Z \triangleleft \mathcal{F}$ to express that Z is measurable with respect to the sigma algebra \mathcal{F} .

REMARK 13. Let \mathcal{R} be a version-independent risk functional and let \mathcal{F}_t be a sub-sigma algebra of \mathcal{F} . Then

$$\mathcal{R}(\mathbb{E}(Y | \mathcal{F}_t)) \leq \mathcal{R}(Y). \quad (8)$$

The proof relies on the conditional Jensen inequality (see Billingsley [7, §34] or Williams [37, Chapter 9]) and is contained in Shapiro et al. [35, Corollary 6.30].

⁵ $\mathbb{I}_B(x) = 0$ if $x \in B$ and $\mathbb{I}_B(x) = \infty$ else.

REMARK 14. It follows in particular from the previous Remark 13 that $AV@R_\alpha(\mathbb{E}(Z|\mathcal{F}_t)) \leq AV@R_\alpha(Z)$ for all $\alpha \in (0, 1)$. From (7) it follows in particular that $\mathbb{E}(Z|\mathcal{F}_t)$ is a feasible dual variable, provided that Z is a feasible dual variable.

As the conditional expectation is further self-adjoint; that is

$$\mathbb{E}(YZ) = \mathbb{E}(\mathbb{E}(Y|\mathcal{F}_t) \cdot Z) = \mathbb{E}(Y \cdot \mathbb{E}(Z|\mathcal{F}_t))$$

whenever $Y \triangleleft \mathcal{F}_t$, it is sufficient to consider \mathcal{F}_t -measurable dual variables in representation (5), that is,

$$\mathcal{R}(Y) = \sup\{\mathbb{E}(YZ): Z \in L^\infty(\mathcal{F}_t), \mathcal{R}^*(Z) = 0\}, \quad (Y \triangleleft \mathcal{F}_t).$$

The next theorem provides an additional recipe to construct useful, feasible dual random variables.

THEOREM 15 (BOCHNER REPRESENTATION OF FEASIBLE DUAL VARIABLES). *Let, for $\alpha \in (0, 1)$, $Z_\alpha \triangleleft \mathcal{F}_t$ satisfy $0 \leq Z_\alpha \leq 1/(1 - \alpha)$, $\mathbb{E}(Z_\alpha) = 1$, and $\mu \in \mathcal{M}$ be a probability measure on $[0, 1]$. Then the Bochner integral $\int_0^1 Z_\alpha \mu(d\alpha)$ (provided that $\alpha \mapsto Z_\alpha$ is measurable and the integral exists) is feasible for \mathcal{R}_{σ_μ} ; that is,*

$$(1 - \alpha')AV@R_{\alpha'}\left(\int_0^1 Z_\alpha \mu(d\alpha)\right) \leq \int_{\alpha'}^1 \sigma_\mu(p) dp, \quad (0 \leq \alpha' \leq 1),$$

where σ_μ is the spectral function associated with μ (see (2)). A fortiori it is feasible for $\sup_{\sigma \in \mathcal{F}} \mathcal{R}_\sigma$, provided that $\sigma_\mu \in \mathcal{F}^*$.

PROOF. The Average Value-at-Risk is convex; hence,

$$AV@R_{\alpha'}\left(\int_0^1 Z_\alpha \mu(d\alpha)\right) \leq \int_0^1 AV@R_{\alpha'}(Z_\alpha) \mu(d\alpha).$$

Note that if $Z \geq 0$ and $\mathbb{E}(Z) = 1$, then

$$AV@R_{\alpha'}(Z_\alpha) = \frac{1}{1 - \alpha'} \int_{\alpha'}^1 V@R_p(Z_\alpha) dp \leq \frac{1}{1 - \alpha'} \int_0^1 V@R_p(Z_\alpha) dp = \frac{\mathbb{E}(Z_\alpha)}{1 - \alpha'} = \frac{1}{1 - \alpha'},$$

implying that

$$\begin{aligned} AV@R_{\alpha'}\left(\int_0^1 Z_\alpha \mu(d\alpha)\right) &\leq \int_0^1 \min\left\{\frac{1}{1 - \alpha}, \frac{1}{1 - \alpha'}\right\} \mu(d\alpha) \\ &= \int_0^{\alpha'} \frac{1}{1 - \alpha} \mu(d\alpha) + \int_{\alpha'}^1 \frac{1}{1 - \alpha'} \mu(d\alpha) \\ &= \sigma_\mu(\alpha') + \frac{\mu(\alpha', 1)}{1 - \alpha'} = \sigma_\mu(\alpha') + \frac{1}{1 - \alpha'} \int_{\alpha'}^1 (1 - p) d\sigma(p) \\ &= \sigma_\mu(\alpha') + \frac{1}{1 - \alpha'} \left(-(1 - \alpha')\sigma_\mu(\alpha') + \int_{\alpha'}^1 \sigma_\mu(p) dp \right) = \frac{1}{1 - \alpha'} \int_{\alpha'}^1 \sigma_\mu(p) dp, \end{aligned}$$

and hence $\int_0^1 Z_\alpha \mu(d\alpha)$ is feasible by Theorem 9. \square

5. Extended conditional risk functionals. In the previous sections we have discussed the dual representations of positively homogeneous (i.e., coherent) risk functionals \mathcal{R} in a sufficiently broad context. This allows us to introduce a *conditional version* of a coherent risk functional.

5.1. Definition of extended conditional risk functionals. Let $Y \in L^\infty(\mathcal{F}, P)$ and consider a sub-sigma algebra \mathcal{F}_t ($\mathcal{F}_t \subset \mathcal{F}$). As above we shall write $Z \triangleleft \mathcal{F}_t$ to express that a random variable Z is measurable with respect to the sigma algebra \mathcal{F}_t .

The usual conditional risk functionals are defined as the functionals applied to the random variables, which result from conditioning Y on the fibers of \mathcal{F} ; see Graf and Mauldin [12]. We extend this concept by allowing a change-of-measure through a probability density of the random variable Z' that is chosen in an optimal way.

DEFINITION 16 (EXTENDED CONDITIONAL RISK FUNCTIONAL). Let $\mathcal{R}_{\mathcal{F}}$ be a coherent risk functional induced by the spectrum \mathcal{S} . For all feasible duals $Z_t \triangleleft \mathcal{F}_t$, the *extended conditional risk functional* is defined as⁶

$$\mathcal{R}_{\mathcal{F}; Z_t}(Y | \mathcal{F}_t) := \text{ess sup} \{ \mathbb{E}(YZ' | \mathcal{F}_t) | \mathbb{E}(Z' | \mathcal{F}_t) = \mathbb{1}, \text{ and } \mathcal{R}^*(Z_t Z') < \infty \},$$

which can be rewritten as

$$\mathcal{R}_{\mathcal{F}; Z_t}(Y | \mathcal{F}_t) = \text{ess sup} \{ \mathbb{E}(YZ' | \mathcal{F}_t) | \mathbb{E}(Z' | \mathcal{F}_t) = \mathbb{1} \text{ and } (\exists \sigma \in \mathcal{S}^*) F_{Z_t, Z'}^{-1} \preceq \sigma \}.$$

Notice that by $\mathcal{R}^*(Z_t) < \infty$, the essential supremum is formed over a nonempty set since $Z' \equiv \mathbb{1}$ is always a possible choice and $\mathcal{R}_{\mathcal{F}; Z_t}(Y | \mathcal{F}_t)$ thus is well defined.

The conditional Average Value-at-Risk is an important example of a conditional risk functional. In view of the dual representation of the Average Value-at-Risk, we have the representation

$$\text{AV@R}_{\alpha; Z_t}(Y | \mathcal{F}_t) = \text{ess sup} \{ \mathbb{E}(YZ' | \mathcal{F}_t) | \mathbb{E}(Z' | \mathcal{F}_t) = \mathbb{1}, Z' \geq 0 \text{ and } (1 - \alpha)Z_t Z' \leq \mathbb{1}, \} \quad (9)$$

whenever $\alpha \leq 1$. Notice that $Z_t Z' \geq 0$ since by feasibility $Z_t \geq 0$. Notice also that conditional risk functionals in the sense of Definition 16 differ from the usual definition since they depend not only on Y and the sigma algebra \mathcal{F}_t , but also on the random variable Z_t . For the Average Value-at-Risk, the notion of the extended conditional functional is equivalent to the notion of AV@R at random level, where the level α is considered to be a \mathcal{F}_t -measurable random variable.

DEFINITION 17 (CONDITIONAL AV@R AT RANDOM LEVEL). The conditional Average Value-at-Risk at random level $\alpha \triangleleft \mathcal{F}_t$ ($0 \leq \alpha \leq \mathbb{1}$ a.s.) is the \mathcal{F}_t random variable⁷

$$\text{AV@R}_{\alpha}(Y | \mathcal{F}_t) := \text{ess sup} \{ \mathbb{E}(YZ' | \mathcal{F}_t) | \mathbb{E}(Z' | \mathcal{F}_t) = \mathbb{1}, Z' \geq 0 \text{ and } (1 - \alpha)Z' \leq \mathbb{1} \}. \quad (10)$$

REMARK 18. The central identity combining (9) and the Average Value-at-Risk at random level (10) is

$$\text{AV@R}_{\alpha; Z_t}(Y | \mathcal{F}_t) = \text{AV@R}_{1 - (1 - \alpha)Z_t}(Y | \mathcal{F}_t).$$

For the particular choice $Z_t = \mathbb{1}$ thus

$$\text{AV@R}_{\alpha; \mathbb{1}}(Y | \mathcal{F}_t) = \text{AV@R}_{\alpha}(Y | \mathcal{F}_t),$$

which is, in view of the defining Equation (10), the traditional conditional Average Value-at-Risk for a fixed nonrandom level α .

Moreover, for $\alpha \equiv 0$, just the random variable $Z' = \mathbb{1}$ is feasible in (10), which makes the particular case

$$\text{AV@R}_0(Y | \mathcal{F}_t) = \mathbb{E}(Y | \mathcal{F}_t)$$

evident. We will see that this is an extreme case in the sense that

$$\text{AV@R}_0(Y | \mathcal{F}_t) \leq \mathcal{R}_{\mathcal{F}; Z}(Y | \mathcal{F}_t) \leq \text{AV@R}_1(Y | \mathcal{F}_t)$$

in Theorem 20 below.

5.2. Elementary properties of extended conditional risk functionals. The extended conditional risk functional inherits, on conditional basis, all essential properties of the original risk functional. Notice that the extended conditional functional is defined for pairs (Y, Z) , where $Y \in L^\infty$ and $\mathcal{R}^*(Z) < \infty$.

Here and in what follows we drop the subscript \mathcal{S} and write the conditional functional simply as \mathcal{R}_Z .

THEOREM 19 (PROPERTIES OF THE CONDITIONAL RISK FUNCTIONALS). *The conditional risk functional obeys the following properties:*

- (i) PREDICTABILITY: $\mathcal{R}_Z(Y | \mathcal{F}_t) = Y$ if $Y \triangleleft \mathcal{F}_t$;
- (ii) TRANSLATION EQUIVARIANCE: $\mathcal{R}_Z(Y + c | \mathcal{F}_t) = \mathcal{R}_Z(Y | \mathcal{F}_t) + c$ if $c \triangleleft \mathcal{F}_t$;

⁶ See e.g., Karatzas and Shreve [15, Appendix A] or Dunford and Schwartz [11] for a rigorous definition and discussion of the essential supremum of a family of random variables.

⁷ See also Cheridito and Kupper [10, §2.3.1].

- (iii) POSITIVE HOMOGENEITY: $\mathcal{R}_Z(\lambda Y | \mathcal{F}_t) = \lambda \mathcal{R}_Z(Y | \mathcal{F}_t)$ whenever $\lambda \triangleleft \mathcal{F}_t$, $\lambda \geq 0$ and λ bounded;
- (iv) MONOTONICITY: $\mathcal{R}_Z(Y_1 | \mathcal{F}_t) \leq \mathcal{R}_Z(Y_2 | \mathcal{F}_t)$ whenever $Y_1 \leq Y_2$;
- (v) CONVEXITY: The mapping $Y \mapsto \mathcal{R}_Z(Y | \mathcal{F}_t)$ is convex; that is,

$$\mathcal{R}_Z((1 - \lambda)Y_0 + \lambda Y_1 | \mathcal{F}_t) \leq (1 - \lambda)\mathcal{R}_Z(Y_0 | \mathcal{F}_t) + \lambda \mathcal{R}_Z(Y_1 | \mathcal{F}_t)$$

for $\lambda \triangleleft \mathcal{F}_t$ and $0 \leq \lambda \leq 1$, almost surely;

- (vi) CONCAVITY: The mapping $Z \mapsto Z \cdot \mathcal{R}_Z(Y | \mathcal{F}_t)$ is concave; more specifically,

$$Z_\lambda \cdot \mathcal{R}_{Z_\lambda}(Y | \mathcal{F}_t) \geq (1 - \lambda)Z_0 \cdot \mathcal{R}_{Z_0}(Y | \mathcal{F}_t) + \lambda Z_1 \cdot \mathcal{R}_{Z_1}(Y | \mathcal{F}_t)$$

almost everywhere, where $Z_\lambda = (1 - \lambda)Z_0 + \lambda Z_1$ and $\lambda \in [0, 1]$.

PROOF. (i) PREDICTABILITY follows from $\mathbb{E}(YZ' | \mathcal{F}_t) = Y \cdot \mathbb{E}(Z' | \mathcal{F}_t) = Y$ whenever $Y \triangleleft \mathcal{F}_t$,

(ii) TRANSLATION EQUIVARIANCE follows from $\mathbb{E}((Y + c)Z' | \mathcal{F}_t) = \mathbb{E}(YZ' | \mathcal{F}_t) + c \cdot \mathbb{E}(Z' | \mathcal{F}_t) = \mathbb{E}(YZ' | \mathcal{F}_t) + c$ and

(iii) POSITIVE HOMOGENEITY follows from $\text{ess sup } \mathbb{E}(\lambda YZ' | \mathcal{F}_t) = \lambda \text{ess sup } \mathbb{E}(YZ' | \mathcal{F}_t)$, as $0 \leq \lambda \triangleleft \mathcal{F}_t$.

(iv) MONOTONICITY is inherited from the conditional expected value as $\mathbb{E}(Y_1 Z' | \mathcal{F}_t) \leq \mathbb{E}(Y_2 Z' | \mathcal{F}_t)$ whenever $Y_1 \leq Y_2$ and $Z' \geq 0$.

- (v) For CONVEXITY observe that

$$\begin{aligned} & (1 - \lambda)\mathcal{R}_{Z_0}(Y_0 | \mathcal{F}_t) + \lambda \mathcal{R}_{Z_1}(Y_1 | \mathcal{F}_t) \\ &= (1 - \lambda) \text{ess sup}_{\mathcal{R}^*(Z_0 Z'_0)=0} \mathbb{E}(Y_0 Z'_0 | \mathcal{F}_t) + \lambda \text{ess sup}_{\mathcal{R}^*(Z_1 Z'_1)=0} \mathbb{E}(Y_1 Z'_1 | \mathcal{F}_t) \\ &\geq \text{ess sup}_{\mathcal{R}^*(Z_\lambda Z'_\lambda)=0} (1 - \lambda)\mathbb{E}(Y_0 Z'_0 | \mathcal{F}_t) + \lambda \mathbb{E}(Y_1 Z'_1 | \mathcal{F}_t) \\ &= \mathcal{R}_{Z_\lambda}((1 - \lambda)Y_0 + \lambda Y_1 | \mathcal{F}_t). \end{aligned}$$

(vi) For CONCAVITY let $Z_0, Z_1 \triangleleft \mathcal{F}_t$ be feasible and Z'_0 and Z'_1 be chosen such that $\mathcal{R}^*(Z_0 Z'_0) < \infty$ and $\mathcal{R}^*(Z_1 Z'_1) < \infty$. Define

$$Z'_\lambda := \begin{cases} \frac{(1 - \lambda)Z_0 Z'_0 + \lambda Z_1 Z'_1}{Z_\lambda} & \text{if } Z_\lambda > 0, \\ 1 & \text{if } Z_\lambda \leq 0, \end{cases}$$

and observe that

$$\mathbb{E}(Z'_\lambda | \mathcal{F}_t) = \begin{cases} \frac{(1 - \lambda)Z_0 + \lambda Z_1}{Z_\lambda} & \text{if } Z_\lambda > 0, \\ 1 & \text{else,} \end{cases} = \mathbb{1}.$$

Then by convexity of $\mathcal{L}_{\mathcal{F}^*}$, $Z_\lambda := (1 - \lambda)Z_0 + \lambda Z_1 \triangleleft \mathcal{F}_t$ is feasible as well, and $\mathcal{R}^*(Z_\lambda Z'_\lambda) \leq (1 - \lambda)\mathcal{R}^*(Z_0 Z'_0) + \lambda \mathcal{R}^*(Z_1 Z'_1) < \infty$, such that $Z_\lambda Z'_\lambda$ as well is feasible. It follows that

$$Z_\lambda \cdot \mathcal{R}_{Z_\lambda}(Y | \mathcal{F}_t) \geq Z_\lambda \cdot \mathbb{E}(YZ'_\lambda | \mathcal{F}_t) = (1 - \lambda)Z_0 \mathbb{E}(YZ'_0 | \mathcal{F}_t) + \lambda Z_1 \mathbb{E}(YZ'_1 | \mathcal{F}_t).$$

Taking the essential supremum (with respect to Z'_0 and Z'_1) yields

$$Z_\lambda \cdot \mathcal{R}_{Z_\lambda}(Y | \mathcal{F}_t) \geq (1 - \lambda)Z_0 \cdot \mathcal{R}_{Z_0}(Y | \mathcal{F}_t) + \lambda Z_1 \cdot \mathcal{R}_{Z_1}(Y | \mathcal{F}_t),$$

which is the assertion. \square

THEOREM 20 (LOWER AND UPPER BOUNDS). Let $\mathcal{F}_\tau \supset \mathcal{F}_t$ be sigma algebras. Then the following inequalities hold true:

- (i) $\mathbb{E}(Y | \mathcal{F}_t) \leq \mathcal{R}_Z(\mathbb{E}(Y | \mathcal{F}_\tau) | \mathcal{F}_t) \leq \mathcal{R}_Z(Y | \mathcal{F}_t) \leq \text{AV@R}_1(Y | \mathcal{F}_t) \leq \text{AV@R}_1(Y) = \text{ess sup}(Y)$,
- (ii) $\mathbb{E}(Y) \leq \mathbb{E}(\mathcal{R}_{Z_t}(Y | \mathcal{F}_t)) \leq \mathcal{R}(Y) \leq \text{AV@R}_1(Y) = \text{ess sup}(Y)$.

PROOF. By definition

$$\mathcal{R}_Z(\mathbb{E}(Y | \mathcal{F}_\tau) | \mathcal{F}_t) = \text{ess sup}_{\mathcal{R}^*(ZZ')=0} \mathbb{E}(Z' \cdot \mathbb{E}(Y | \mathcal{F}_\tau) | \mathcal{F}_t),$$

and one may choose the dual variable \mathcal{F}_τ -measurable, $ZZ' \triangleleft \mathcal{F}_\tau$, by Theorem 13; that is, $Z' \triangleleft \mathcal{F}_\tau$, as $Z \triangleleft \mathcal{F}_\tau$. As the operation of conditional expectation is self-adjoint it follows that

$$\begin{aligned}\mathcal{R}_Z(\mathbb{E}(Y | \mathcal{F}_\tau) | \mathcal{F}_t) &= \operatorname{ess\,sup}_{\mathcal{R}^*(ZZ')=0, Z' \triangleleft \mathcal{F}_\tau} \mathbb{E}(Z' \cdot \mathbb{E}(Y | \mathcal{F}_\tau) | \mathcal{F}_t) \\ &= \operatorname{ess\,sup}_{\mathcal{R}^*(ZZ')=0, Z' \triangleleft \mathcal{F}_\tau} \mathbb{E}(\mathbb{E}(Z' | \mathcal{F}_\tau) \cdot Y | \mathcal{F}_t) \\ &= \operatorname{ess\,sup}_{\mathcal{R}^*(ZZ')=0, Z' \triangleleft \mathcal{F}_\tau} \mathbb{E}(Z' Y | \mathcal{F}_t) \\ &\leq \operatorname{ess\,sup}_{\mathcal{R}^*(ZZ')=0} \mathbb{E}(Z' \cdot Y | \mathcal{F}_t) = \mathcal{R}_Z(Y | \mathcal{F}_t)\end{aligned}$$

which is the first inequality in (i).

Next,

$$\mathcal{R}_Z(Y | \mathcal{F}_t) = \operatorname{ess\,sup}_{\mathcal{R}^*(ZZ')=0} \mathbb{E}(Z' Y | \mathcal{F}_t) \leq \mathbb{E}(Y | \mathcal{F}_t) \quad (11)$$

as $Z = Z \cdot \mathbb{1}$ ($Z' = \mathbb{1}$) is feasible. Replacing Y by $\mathbb{E}(Y | \mathcal{F}_\tau)$ in (11) reveals that

$$\mathcal{R}_Z(\mathbb{E}(Y | \mathcal{F}_\tau) | \mathcal{F}_t) \geq \mathbb{E}(\mathbb{E}(Y | \mathcal{F}_\tau) | \mathcal{F}_t) = \mathbb{E}(Y | \mathcal{F}_t).$$

The other inequalities in (i) are obvious. The proof of the first part of (ii) is a consequence of the decomposition Theorem 21 below. \square

6. The decomposition theorem. The decomposition of a risk functional with respect to incomplete information is accomplished by the conditional risk functional. The following theorem, which is the main theorem of this paper, elaborates that the initial risk functional can be recovered from its conditional dissections by applying the appropriate spectral density.

THEOREM 21 (DECOMPOSITION THEOREM). *Let $\mathcal{R} = \mathcal{R}_{\mathcal{F}}$ be a version-independent, coherent risk functional; then the following hold true:*

(i) $\mathcal{R}_{\mathcal{F}}$ obeys the decomposition

$$\mathcal{R}_{\mathcal{F}}(Y) = \sup \mathbb{E}[Z \cdot \mathcal{R}_{\mathcal{F}; Z}(Y | \mathcal{F}_t)], \quad (12)$$

where the supremum is among all feasible, nonnegative random variables $Z \triangleleft \mathcal{F}_t$ satisfying $\mathbb{E}(Z) = 1$ and $F_Z^{-1} \preceq \sigma$ for an $\sigma \in \mathcal{F}^*$ —that is, $\mathcal{R}_{\mathcal{F}}^*(Z) < \infty$ for the associated spectrum \mathcal{F} .

(ii) Let $\mathcal{F}_t \subset \mathcal{F}_\tau$. The risk conditional functional obeys the nested decomposition

$$\mathcal{R}_S(Y | \mathcal{F}_t) = \operatorname{ess\,sup} \mathbb{E}[Z_\tau \cdot \mathcal{R}_{\mathcal{F}; Z_\tau}(Y | \mathcal{F}_\tau) | \mathcal{F}_t], \quad (13)$$

where the essential supremum is taken among all feasible dual random variables $Z_\tau \triangleleft \mathcal{F}_\tau$.

The proof presented here builds on the respective statement of the Average Value-at-Risk.

LEMMA 22 (DECOMPOSITION THEOREM FOR THE AVERAGE VALUE-AT-RISK). *For all $\alpha \in [0, 1]$, the decomposition of the Average Value-at-Risk is*

$$\operatorname{AV@R}_\alpha(Y) = \sup \mathbb{E}[Z \cdot \operatorname{AV@R}_{1-(1-\alpha)Z}(Y | \mathcal{F}_t)],$$

where the supremum is over all random variables $Z \triangleleft \mathcal{F}_t$ satisfying $\mathbb{E}(Z) = 1$, $Z \geq 0$ and $(1 - \alpha)Z \leq \mathbb{1}$.

PROOF OF LEMMA 22. Let $B_i \in \mathcal{F}_t$ be a finite tessellation such that $B_i \cap B_j = \emptyset$ and $\bigcup_i B_i = \Omega$, and let Z'_i be feasible for $\operatorname{AV@R}_{1-(1-\alpha)Z}(Y | \mathcal{F}_t)$ (see (9)); that is, they satisfy $Z'_i \geq 0$, $\mathbb{E}(Z'_i | \mathcal{F}_t) = \mathbb{1}$ and $(1 - \alpha)ZZ'_i = (1 - [1 - (1 - \alpha)Z])Z'_i \leq \mathbb{1}$.

Define $Z' := \sum_i \mathbb{1}_{B_i} Z'_i$. It is immediate that $(1 - \alpha)ZZ' \leq \mathbb{1}$ and $ZZ' \geq 0$. Moreover $\mathbb{E}(Z' | \mathcal{F}_t) = \sum_i \mathbb{1}_{B_i} = \mathbb{1}$, from which follows that

$$\mathbb{E}(ZZ') = \mathbb{E}(Z\mathbb{E}(Z' | \mathcal{F}_t)) = \mathbb{E}(Z) = 1.$$

Hence ZZ' is feasible for the $\operatorname{AV@R}_\alpha$, and $\operatorname{AV@R}_\alpha(Y) \geq \mathbb{E}(YZZ') = \mathbb{E}(Z \cdot \mathbb{E}(YZ' | \mathcal{F}_t))$. As Z' is composed to represent the essential supremum, it follows that $\operatorname{AV@R}_\alpha(Y) \geq \mathbb{E}(Z \cdot \operatorname{AV@R}_{1-(1-\alpha)Z}(Y | \mathcal{F}_t))$.

As for the converse choose $\tilde{Z} \geq 0$, feasible for the Average Value-at-Risk and satisfying $AV@R_\alpha(Y) = \mathbb{E}(Y\tilde{Z})$ (which exists for $\alpha < 1$). Define $Z := \mathbb{E}(\tilde{Z} | \mathcal{F}_t)$ and

$$Z' := \begin{cases} \frac{\tilde{Z}}{Z} & \text{if } Z > 0, \\ 1 & \text{else.} \end{cases}$$

Then $Z' \geq 0$ and $(1 - (1 - (1 - \alpha)Z))Z' \leq \mathbb{1}$, such that Z' is feasible for the conditional $AV@R_\alpha(Y | \mathcal{F}_t)$.

By the dual representations of the Average Value-at-Risk and $AV@R(Z) \geq AV@R(\mathbb{E}(Z | \mathcal{F}_t))$ (see (8) in Proposition 13) it follows that Z is feasible for the Average Value-at-Risk, that is $(1 - \alpha)Z \leq \mathbb{1}$ and $\mathbb{E}(Z) = 1$. Hence

$$AV@R_\alpha(Y) = \mathbb{E}(Y\tilde{Z}) = \mathbb{E}(ZY Z') = \mathbb{E}(Z\mathbb{E}(YZ' | \mathcal{F}_t)) \leq \mathbb{E}(ZAV@R_{1-(1-\alpha)Z}(Y | \mathcal{F}_t)),$$

which is the assertion, provided that $\alpha < 1$.

For the remaining situation $\alpha = 1$, choose $Z^\varepsilon \geq 0$ with $\mathbb{E}(Z^\varepsilon Y) \geq AV@R_1(Y) - \varepsilon$, where $\varepsilon > 0$. By the conditional (L^1 - L^∞)-Hölder inequality

$$\begin{aligned} AV@R_1(Y) - \varepsilon &\leq \mathbb{E}(Z^\varepsilon Y) \leq \mathbb{E}(\mathbb{E}[Z^\varepsilon | \mathcal{F}_t] AV@R_1(Y | \mathcal{F}_t)) \\ &\leq \mathbb{E}(\mathbb{E}[Z^\varepsilon | \mathcal{F}_t]) \cdot AV@R_1(Y) = AV@R_1(Y), \end{aligned}$$

hence,

$$AV@R_1(Y) \leq \mathbb{E}(Z_t^\varepsilon AV@R_{1-(1-\alpha)Z_t^\varepsilon}(Y | \mathcal{F}_t)) + \varepsilon.$$

This proves the converse assertion for $\alpha = 1$, as $\varepsilon > 0$ is arbitrary. \square

Proof of the decomposition theorem (Theorem 21).

PROOF. Let Z_t and Z' , with $\mathbb{E}(Z' | \mathcal{F}_t) = \mathbb{1}$, be fixed such that $Z_t Z'$ is a feasible random variable satisfying $(1 - \alpha)AV@R_\alpha(Z_t Z') \leq \int_\alpha^1 \sigma(p) dp$ for all $\alpha \in [0, 1]$ with

$$\sigma \in \mathcal{S}^*, \quad \text{where } \sigma(\alpha) := F_{Z_t Z'}^{-1}(\alpha). \quad (14)$$

By Corollary 5 one may assume without loss of generality that σ is strictly increasing and bounded, hence invertible.

Moreover, let U be a uniformly distributed random variable (i.e., $P(U \leq u) = u$), coupled in a *co-monotone* way with $Z_t Z'$. The random variable $\sigma_\alpha(U)$ (see (3)) is nonnegative. Furthermore, $\sigma_\alpha(U) \geq 0$ and $\mathbb{E}(\sigma_\alpha(U)) = \int_0^1 \sigma_\alpha(p) dp = 1$, and, moreover,

$$\begin{aligned} P\left(\int_0^1 \sigma_\alpha(U) \mu(d\alpha) \leq \sigma(u)\right) &= P\left(\int_U^1 \frac{1}{1-\alpha} \mu(d\alpha) \leq \sigma(u)\right) \\ &= P(\sigma(U) \leq \sigma(u)) = P(U \leq u) = u. \end{aligned}$$

Hence $\int_0^1 \sigma_\alpha(U) \mu(d\alpha)$ has law σ^{-1} , which is the same law as $Z_t Z'$. By the co-monotone coupling we thus have that

$$Z_t Z' = \int_0^1 \sigma_\alpha(U) \mu(d\alpha) \quad \text{a.s.}$$

Using the setting

$$Z_\alpha := \mathbb{E}(\sigma_\alpha(U) | \mathcal{F}_t) \quad \text{and} \quad Z'_\alpha := \frac{\sigma_\alpha(U)}{Z_\alpha}$$

it follows that

$$\int_0^1 Z_\alpha Z'_\alpha \mu(d\alpha) = \int_0^1 \sigma_\alpha(U) \mu(d\alpha) = Z_t Z' \quad (15)$$

and

$$\begin{aligned} \int_0^1 Z_\alpha \mu(d\alpha) &= \int_0^1 \mathbb{E}(\sigma_\alpha(U) | \mathcal{F}_t) \mu(d\alpha) = \mathbb{E}\left(\int_0^1 \sigma_\alpha(U) \mu(d\alpha) | \mathcal{F}_t\right) \\ &= \mathbb{E}(Z_t Z' | \mathcal{F}_t) = Z_t. \end{aligned}$$

By construction of the random variables we have moreover the properties

$$0 \leq (1 - \alpha)Z_\alpha Z'_\alpha \leq \mathbb{1} \quad \text{and} \quad \mathbb{E}(Z'_\alpha | \mathcal{F}_t) = \mathbb{1}.$$

Then it follows from Lemma 22 that $AV@R_\alpha(Y) \geq \mathbb{E}(Z_\alpha AV@R_{1-(1-\alpha)Z_\alpha}(Y | \mathcal{F}_t))$, and hence

$$\mathcal{R}(Y) \geq \int_0^1 \mathbb{E}(Z_\alpha AV@R_{1-(1-\alpha)Z_\alpha}(Y | \mathcal{F}_t)) \mu(d\alpha) = \mathbb{E}\left(\int_0^1 Z_\alpha AV@R_{1-(1-\alpha)Z_\alpha}(Y | \mathcal{F}_t) \mu(d\alpha)\right).$$

By the definition of the Average Value-at-Risk at random level this is

$$\begin{aligned} \mathcal{R}(Y) &\geq \mathbb{E}\left(\int_0^1 Z_\alpha \operatorname{ess\,sup}\{\mathbb{E}(YZ'_\alpha | \mathcal{F}_t): 0 \leq (1 - \alpha)Z_\alpha Z'_\alpha \leq \mathbb{1}, \mathbb{E}(Z'_\alpha | \mathcal{F}_t) = \mathbb{1}\} \mu(d\alpha)\right) \\ &= \mathbb{E}\left(\int_0^1 \operatorname{ess\,sup}\{\mathbb{E}(YZ'_\alpha | \mathcal{F}_t): 0 \leq (1 - \alpha)Z_\alpha Z'_\alpha \leq \mathbb{1}, \mathbb{E}(Z'_\alpha | \mathcal{F}_t) = \mathbb{1}\} \mu(d\alpha)\right) \\ &\geq \mathbb{E}\left(\operatorname{ess\,sup}\left\{\int_0^1 \mathbb{E}(YZ'_\alpha | \mathcal{F}_t) \mu(d\alpha): 0 \leq (1 - \alpha)Z_\alpha Z'_\alpha \leq \mathbb{1}, \mathbb{E}(Z'_\alpha | \mathcal{F}_t) = \mathbb{1}\right\}\right). \end{aligned} \quad (16)$$

(The latter inequality in fact holds with equality by the interchangeability principle (Rockafellar and Wets [27, Theorem 14.60]), the Bochner-integral in (16) and the essential supremum may be exchanged.) Hence,

$$\mathcal{R}(Y) \geq \mathbb{E}\left(Z_t \operatorname{ess\,sup}\left\{\mathbb{E}\left(Y \frac{1}{Z_t} \int_0^1 Z_\alpha Z'_\alpha \mu(d\alpha) \mid \mathcal{F}_t\right): 0 \leq (1 - \alpha)Z_\alpha Z'_\alpha \leq \mathbb{1}, \mathbb{E}(Z'_\alpha | \mathcal{F}_t) = \mathbb{1}\right\}\right).$$

Now recall the identities (22) and (15), as well as (14), such that we may continue with

$$\mathcal{R}(Y) \geq \mathbb{E}\left(Z_t \operatorname{ess\,sup}\left\{\mathbb{E}\left(Y \frac{1}{Z_t} Z_t Z' \mid \mathcal{F}_t\right): (1 - \alpha)AV@R_\alpha(Z_t Z') = \int_\alpha^1 \sigma(p) dp, \mathbb{E}(Z' | \mathcal{F}_t) = \mathbb{1}\right\}\right)$$

and

$$\begin{aligned} \mathcal{R}(Y) &\geq \mathbb{E}\left(Z_t \operatorname{ess\,sup}\left\{\mathbb{E}(YZ' | \mathcal{F}_t): \mathbb{E}(Z' | \mathcal{F}_t) = \mathbb{1}, \exists \sigma \in \mathcal{S}^*: (1 - \alpha)AV@R_\alpha(Z_t Z') \leq \int_\alpha^1 \sigma(p) dp\right\}\right) \\ &= \mathbb{E}\left(Z_t \operatorname{ess\,sup}\left\{\mathbb{E}(YZ' | \mathcal{F}_t): \mathbb{E}(Z' | \mathcal{F}_t) = \mathbb{1}, \exists \sigma \in \mathcal{S}: (1 - \alpha)AV@R_\alpha(Z_t Z') \leq \int_\alpha^1 \sigma(p) dp\right\}\right) \\ &= \mathbb{E}(Z_t \mathcal{R}_{Z_t}(Y | \mathcal{F}_t)), \end{aligned}$$

where we have used the same reasoning as in the proof of Theorem 9. This establishes the first inequality “ \geq .”

As for the converse inequality, let $\sigma \in \mathcal{S}$ be chosen such that $\mathcal{R}_\sigma(Y) \geq \sup_{\sigma' \in \mathcal{S}} \mathcal{R}_{\sigma'}(Y) - \varepsilon$, and let a feasible Z be chosen such that $\mathbb{E}(YZ) > \mathcal{R}_\sigma(Y) - \varepsilon$; by feasibility, $(1 - \alpha)AV@R_\alpha(Z) \leq \int_\alpha^1 \sigma(p) dp$. Define $Z_t := \mathbb{E}(Z | \mathcal{F}_t)$ and

$$Z'_t := \begin{cases} \frac{Z}{Z_t} & \text{if } Z_t > 0 \\ 1 & \text{else} \end{cases}$$

and, by Lemma 13, Z_t is feasible as well; that is, $(1 - \alpha)AV@R_\alpha(Z_t) \leq \int_\alpha^1 \sigma(p) dp$. With this choice it is obvious that

$$\mathbb{E}(YZ) = \mathbb{E}(Z_t \cdot \mathbb{E}(YZ'_t | \mathcal{F}_t)) \leq \mathbb{E}(Z_t \cdot \mathcal{R}_{Z_t}(Y | \mathcal{F}_t))$$

and hence $\mathbb{E}(Z_t \cdot \mathcal{R}_{Z_t}(Y | \mathcal{F}_t)) \geq \mathcal{R}_\sigma(Y) - \varepsilon$, which finally completes the proof of the first statement.

The nested decomposition for an intermediate sigma algebra \mathcal{F}_τ reads along the same lines as the preceding proof, but conditioned on \mathcal{F}_t and \mathcal{F}_t replaced by \mathcal{F}_τ . \square

REMARK 23. It follows from the previous proof that the optimal dual variable in the decomposition (12) is unique, if the dual variable Z for \mathcal{R} at Y is unique as well. Further, the conditional risk functional is explicitly given by

$$\mathcal{R}_Z(Y | \mathcal{F}_t) = \frac{\int_0^1 Z_\alpha AV@R_{1-(1-\alpha)Z_\alpha}(Y | \mathcal{F}_t) \mu(d\alpha)}{\int_0^1 Z_\alpha \mu(d\alpha)}, \quad (17)$$

provided that $Z = \int_0^1 Z_\alpha \mu(d\alpha)$ and $Z_\alpha = \mathbb{E}(\tilde{Z}_\alpha | \mathcal{F}_t)$, where \tilde{Z}_α is the optimal dual variable to compute the Average Value-at-Risk at level α (that is $\mathbb{E}(Y\tilde{Z}_\alpha) = AV@R_\alpha(Y)$ and $0 \leq \tilde{Z}_\alpha \leq 1/(1 - \alpha)$).

An example of a decomposition of a risk functional is provided in the appendix.

REMARK 24. The representation (13) extends the decomposition (12) to the case of a filtered probability space with increasing sigma algebras: the risk functional thus can be nested in the way described by Equation (13).

7. Time-consistent decision making. The illustrating example in Figure 1(a) shows that applying the same risk functional in its conditional form and in unconditional form at the previous stage may reverse the preference relation. It particularly demonstrates that

$$AV@R_{\alpha\mathbb{I}}(Y' | \mathcal{F}_t) \leq AV@R_{\alpha\mathbb{I}}(Y | \mathcal{F}_t) \not\Rightarrow AV@R_{\alpha}(Y') \leq AV@R_{\alpha}(Y), \quad (18)$$

if the same risk level α is employed conditionally and also unconditionally. In contrast, our decomposition theorem (Theorem 21) allows a time-consistent comparison of random variables Y' and Y on the basis of incomplete information.

7.1. Ranking based on conditional information. Suppose that $\mathcal{R}(Y) = \mathbb{E}(YZ)$, where Z is the optimal feasible random variable satisfying $\mathcal{R}^*(Z) = 0$ according to the Fenchel–Moreau theorem (Theorem 6). Suppose further that

$$\mathcal{R}_{Z_t}(Y' | \mathcal{F}_t) \geq \mathcal{R}_{Z_t}(Y | \mathcal{F}_t),$$

where $Z_t = \mathbb{E}(Z | \mathcal{F}_t)$. It follows then from the decomposition theorem and its proof that

$$\mathcal{R}(Y') \geq \mathbb{E}(Z_t \cdot \mathcal{R}_{Z_t}(Y' | \mathcal{F}_t)) \geq \mathbb{E}(Z_t \cdot \mathcal{R}_{Z_t}(Y | \mathcal{F}_t)) = \mathcal{R}(Y). \quad (19)$$

This is a ranking of the random variables Y and Y' based on incomplete information \mathcal{F}_t . Note, that a special risk profile Z is necessary in (19) to allow the conclusion; the choice $Z = \mathbb{I}$ is not adequate.

Although Figure 1(a) shows an example of inconsistency of fixed level $AV@R$, the resolution of the inconsistency by random level $AV@R$ is demonstrated in Figure 1(b), which is based on Lemma 22. When applying the adjusted risk level according the decomposition theorem for the Average Value-at-Risk, it turns out that the displayed random variables can be ordered with respect to $AV@R_{1/3}$ at an earlier stage, after revelation of the partial information relative to \mathcal{F}_t . A decision, which is consistent with the final decision criterion, thus is available already at an earlier stage.

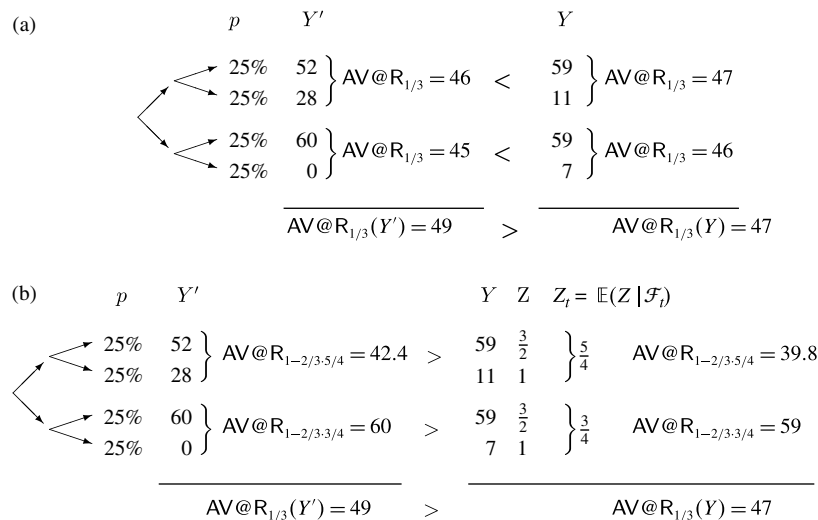


FIGURE 1. An example of inconsistency of the Average Value-at-Risk at fixed level and its resolution. (a) The random variable Y is preferred over Y' because it has lower Average Value-at-Risk at level $\alpha = \frac{1}{3}$, $AV@R_{1/3}(Y') > AV@R_{1/3}(Y)$. But both conditional observations support the opposite result: it holds that $AV@R_{1/3}(Y' | \mathcal{F}_t) < AV@R_{1/3}(Y | \mathcal{F}_t)$, and the conditional observations thus suggest that Y' should be preferred over Y . (b) The same processes as in Figure 1(a). Applying the adjusted risk level to compare the random variables Y and Y' , it holds that $AV@R_{1/3}(Y') > AV@R_{1/3}(Y)$ because $AV@R_{1-(1-\alpha)Z_t}(Y' | \mathcal{F}_t) > AV@R_{1-(1-\alpha)Z_t}(Y | \mathcal{F}_t)$, where the critical risk profile Z is chosen as outlined in §7.1.

7.2. Time-consistent decisions. In this section we discuss the question which coherent risk functionals allow consistent decisions already at an earlier stage when the usual conditioning (and not the extended) is used.

In what follows we elaborate that (18) is a general pattern of practically relevant risk functionals \mathcal{R} : in general, the order relation incorporated by \mathcal{R} , when applied at a previous stage without modification, is destroyed. Here is the formal definition of time-consistent decisions.

DEFINITION 25. A version-independent risk functional \mathcal{R} allows *time-consistent decisions* if

$$\mathcal{R}(Y | \mathcal{F}) \leq \mathcal{R}(Y' | \mathcal{F}) \text{ a.s.} \implies \mathcal{R}(Y) \leq \mathcal{R}(Y') \quad (20)$$

for all Y and Y' and all sigma algebras \mathcal{F} , where the same risk functional \mathcal{R} is repeated on conditional basis.

REMARK 26. Notice that we require here that the conditional risk functionals are the usual conditional versions of the same functional \mathcal{R} and not the extended conditional risk functionals in the sense of Definition 16. Risk functionals allowing time consistent decisions, as addressed in Definition 25, aim at ranking a strategy Y higher than Y' already after revelation of partial, conditional outcomes. (A similar type of consistency as considered here has been introduced by Wang [36, §4.1].)

Property (20) is a necessary property for risk measures to qualify for objectives in time-consistent multi-stage optimization problems (see Shapiro [33, p. 437]): the decision problem is said to be time consistent if a partial solution can be fixed, and this solution will never be subject to changes when considering the problem conditioned on the partial, fixed solution. Time consistency has been further studied in the context of dynamic programming (see Carpentier et al. [8] as well as Shapiro [32]) and Markov decision processes (see Ruszczyński [28]). Whereas the latter papers consider nested structures of conditional risk functionals, we investigate property (20) here for just one measure \mathcal{R} and its conditional version $\mathcal{R}(\cdot | \mathcal{F})$. It is well known that the following functionals have property (20):

- (i) the expectation operator $\mathbb{E}(\cdot) = \text{AV} @ \mathcal{R}_0(\cdot)$ as

$$\mathbb{E}(Y' | \mathcal{F}_t) \leq \mathbb{E}(Y | \mathcal{F}_t) \implies \mathbb{E}(Y') \leq \mathbb{E}(Y);$$

- (ii) the max-risk functional $\text{ess sup}(\cdot) = \text{AV} @ \mathcal{R}_1(\cdot)$ as

$$\text{ess sup}(Y' | \mathcal{F}_t) \leq \text{ess sup}(Y | \mathcal{F}_t) \implies \text{ess sup}(Y') \leq \text{ess sup}(Y).$$

One might think that functionals, which allow (time) consistent decisions, form a convex set. This is not the case: even the simple functional

$$\mathcal{R}(Y) := (1 - \lambda) \cdot \mathbb{E}(Y) + \lambda \cdot \text{ess sup}(Y)$$

does *not* allow consistent decisions whenever $0 < \lambda < 1$, as Figure 2 shows.

We give a final example to elaborate that even general, positively homogeneous risk functionals do not allow consistent decisions in the sense specified by Definition 25. For this, consider the version-independent risk functional

$$\mathcal{R}(Y) = \sup_{\mu \in \mathcal{M}} \int_0^1 \text{AV} @ \mathcal{R}_\alpha(Y) \mu(d\alpha),$$

where the expectation and the essential supremum are excluded by the following assumptions:

- (i) There is an $\varepsilon > 0$ such that

$$\inf_{\mu \in \mathcal{M}} \mu([\varepsilon, 1 - \varepsilon]) > 0$$

and

- (ii)

$$\sup_{\mu \in \mathcal{M}} \mu([\gamma, 1]) \rightarrow 0$$

whenever $\gamma \rightarrow 1$.

We demonstrate that \mathcal{R} does not allow (time) consistent decision making.

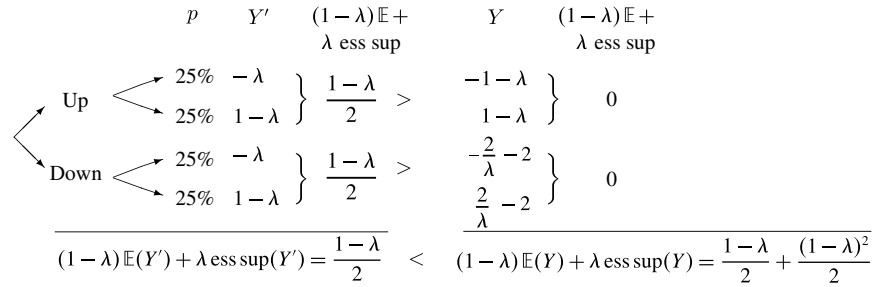


FIGURE 2. This example addresses the risk functional $\mathcal{R}(Y) := (1-\lambda) \cdot \mathbb{E}(Y) + \lambda \cdot \text{ess sup}(Y)$ for an arbitrary $\lambda \in (0, 1)$. This risk functional has higher conditional outcomes for Y' in comparison to Y ($(1-\lambda)/2 > 0$). However, if we take into account all outcomes simultaneously, then the result reverses.

Indeed, choose $\varepsilon > 0$ according to (i) and let $0 < p < \varepsilon$. Let the random variable Y be given by

$$Y = \begin{cases} Y_1 & \text{with probability } q, \\ Y_2 & \text{with probability } 1-q, \end{cases}$$

where

$$Y_1 = \begin{cases} \frac{p}{1-\varepsilon} & \text{w. p. } 1-\varepsilon, \\ \frac{2\eta}{\varepsilon} & \text{w. p. } \frac{\varepsilon}{2}, \\ -\frac{2\eta}{\varepsilon} & \text{w. p. } \frac{\varepsilon}{2}, \end{cases} \quad \text{and} \quad Y_2 = 0 \text{ w. p. } 1,$$

(the constants q and η ($\eta < \varepsilon p/2(1-\varepsilon)$) are specified later).

The random variable Y' is given by

$$Y' = \begin{cases} Y'_1 & \text{w. p. } q, \\ Y'_2 & \text{w. p. } 1-q, \end{cases}$$

where

$$Y'_1 = \begin{cases} -1 & \text{w. p. } p, \\ 0 & \text{w. p. } 1-p, \end{cases} \quad \text{and} \quad Y'_2 = 0 \text{ w. p. } 1.$$

Notice that Y (Y' , respectively) have the distribution

$$Y = \begin{cases} \frac{p}{1-\varepsilon} & \text{w. p. } (1-\varepsilon)q, \\ \frac{2\eta}{\varepsilon} & \text{w. p. } \frac{\varepsilon q}{2}, \\ 0 & \text{w. p. } 1-q, \\ -\frac{2\eta}{\varepsilon} & \text{w. p. } \frac{\varepsilon q}{2}, \end{cases} \quad \text{and} \quad Y' = \begin{cases} 0 & \text{w. p. } 1-pq, \\ 1 & \text{w. p. } pq. \end{cases}$$

Calculating the Average Value-at-Risk we find that

$$AV@R_\alpha(Y_1) = \begin{cases} \frac{p}{1-\varepsilon} & \text{for } 1-\varepsilon \leq \alpha, \\ \frac{p\varepsilon + 2\eta(\varepsilon - \alpha)}{\varepsilon(1-\alpha)} & \text{for } \frac{\varepsilon}{2} \leq \alpha \leq \varepsilon, \\ \frac{p\varepsilon + 2\eta\alpha}{\varepsilon(1-\alpha)} & \text{for } \varepsilon \leq \alpha, \end{cases}$$

and

$$AV@R_\alpha(Y'_1) = \begin{cases} 1 & \text{for } 1-p \leq \alpha, \\ \frac{p}{1-\alpha} & \text{for } \alpha \leq 1-p; \end{cases}$$

moreover,

$$\text{AV@R}_\alpha(Y) = \begin{cases} \frac{p}{1-\varepsilon} & \text{for } 1-q(1-\varepsilon) \leq \alpha, \\ \frac{pq\varepsilon + 2\eta(1-\alpha-q(1-\varepsilon))}{\varepsilon(1-\alpha)} & \text{for } 1-q\left(1-\frac{\varepsilon}{2}\right) \leq \alpha \leq 1-q(1-\varepsilon), \\ \frac{q(p+\eta)}{1-\alpha} & \text{for } q \leq \alpha \leq 1-q\left(1-\frac{\varepsilon}{2}\right), \\ \frac{pq\varepsilon + 2\alpha\eta}{\varepsilon(1-\alpha)} & \text{for } \alpha \leq q, \end{cases}$$

and

$$\text{AV@R}_\alpha(Y') = \begin{cases} 1 & \text{for } 1-qp \leq \alpha, \\ \frac{qp}{1-\alpha} & \text{for } \alpha \leq 1-qp. \end{cases}$$

We show that $\eta > 0$ and $q > 0$ can be chosen in such way that

$$\mathcal{R}(Y_1) > \mathcal{R}(Y'_1) \quad \text{and} \quad \mathcal{R}(Y_2) \geq \mathcal{R}(Y'_2),$$

but the unconditional random variables show the opposite inequality,

$$\mathcal{R}(Y) < \mathcal{R}(Y').$$

It holds that

$$\begin{aligned} \text{AV@R}_\alpha(Y_1) &\geq \text{AV@R}_\alpha(Y'_1), & \text{if } \alpha \leq \varepsilon, \\ \text{AV@R}_\alpha(Y_1) &\leq \text{AV@R}_\alpha(Y'_1), & \text{if } \alpha \geq \varepsilon, \end{aligned}$$

and

$$\begin{aligned} \text{AV@R}_\alpha(Y) &\leq \text{AV@R}_\alpha(Y'), & \text{if } \alpha \leq 1-q(1-\varepsilon), \\ \text{AV@R}_\alpha(Y) &\geq \text{AV@R}_\alpha(Y'), & \text{if } \alpha \geq 1-q(1-\varepsilon). \end{aligned}$$

Therefore,

$$\begin{aligned} &\int_0^1 \text{AV@R}_\alpha(Y_1) - \text{AV@R}_\alpha(Y'_1) \mu(d\alpha) \\ &= \left(\int_0^\varepsilon + \int_\varepsilon^{2\varepsilon} + \int_{2\varepsilon}^1 \right) \text{AV@R}_\alpha(Y_1) - \text{AV@R}_\alpha(Y'_1) \mu(d\alpha) \\ &\leq \frac{\eta}{1-\alpha} + 0 - \mu([2\varepsilon, 1]) \frac{p\varepsilon}{(1-\varepsilon)(1-2\varepsilon)}. \end{aligned}$$

By assumption (i) $\eta > 0$ can be chosen small enough such that the latter expression is strictly negative. Therefore,

$$\mathcal{R}(Y_1) < \mathcal{R}(Y'_1). \quad (21)$$

On the other hand,

$$\int_0^1 \text{AV@R}_\alpha(Y_2) - \text{AV@R}_\alpha(Y'_2) \mu(d\alpha) = 0,$$

is such that

$$\mathcal{R}(Y_2) \leq \mathcal{R}(Y'_2). \quad (22)$$

For the unconditional variables, however,

$$\begin{aligned} &\int_0^1 \text{AV@R}_\alpha(Y) - \text{AV@R}_\alpha(Y') \mu(d\alpha) \\ &= \left(\int_0^{q\varepsilon/2} + \int_{q\varepsilon/2}^{1-q} + \int_{1-q}^1 \right) \text{AV@R}_\alpha(Y) - \text{AV@R}_\alpha(Y') \mu(d\alpha) \\ &\geq q \left(0 + \frac{2\eta}{2-q\eta} \mu\left(\left[\frac{q\varepsilon}{2}, 1-q\right]\right) - \mu([1-q, 1]) \right), \end{aligned}$$

which again, by (ii), can be made strictly positive by choosing q small enough. It follows that

$$\mathcal{R}(Y) > \mathcal{R}(Y').$$

Together with (21) and (22) it becomes apparent that the risk functional \mathcal{R} reverses the preference for the two random variables Y and Y' : the risk functional \mathcal{R} thus does not allow (time) consistent decisions.

In summary, one may say that time-consistent decisions can only be expected in the following situations

- (i) if the risk measure is either the expectation or the essential supremum;
- (ii) if the risk measure is a nested composition of conditional measures in the sense of Ruszczyński [28];
- (iii) if the risk measure is randomly adapted according to the available information, i.e., if it is based on the extended conditional versions, as it is shown in this paper.

8. Summary and outlook. The present paper introduces a new, general concept of conditional risk functionals. The extended conditional risk functional defined in the new sense respects the history of already available information, but in addition it reflects the initial risk functional without any modification. The extended conditional risk functional is consistent with the past *and* the future in a way, which is clarified by the main decomposition result, Theorem 21. This is a positive result for time-consistent decision making when involving coherent risk functionals.

The theory is elaborated with the help of convex conjugate (or dual) functions. The first part of the paper characterizes version-independent risk functionals by use of dual representations. This representation is used then to define extended conditional risk functionals in a sufficiently broad context, which is key for the decomposition result. The presented, new concept appears to be able to substitute dynamic extensions or artificial compositions of risk functionals.

Finally, numerical examples are used to further outline and illustrate the results of the proposed methodology. It is demonstrated that the usual (i.e., nonextended) conditional risk functionals lead to time-inconsistent decisions, a problem which disappears when using the extended form.

A key motivation for the present investigations is the use of risk functionals as decision criteria in stochastic optimization problems. The relation of properties of risk functionals to properties of stochastic optimization problems formulated with the help of these functionals is very important for applications. However, the implications for stochastic optimization is beyond the scope of the present paper and we leave them for a separate discussion.

Acknowledgments. The authors thank two anonymous referees and the associate editor in their commitment to carefully reading the manuscript and providing suggestions to improve the paper. Parts of this paper are contained in their book *Multistage Stochastic Optimization* (2014, Springer), which addresses and summarizes many more topics in multistage stochastic optimization. The book cites a working version of this paper because it was completed before final acceptance of this paper.

Appendix. An example of a decomposition of a risk measure. Figure A.1 addresses the risk functional

$$\mathcal{R}(Y) := \frac{3}{5} \cdot \text{AV@R}_{0.7}(Y) + \frac{2}{5} \cdot \text{AV@R}_{0.4}(Y), \quad (23)$$

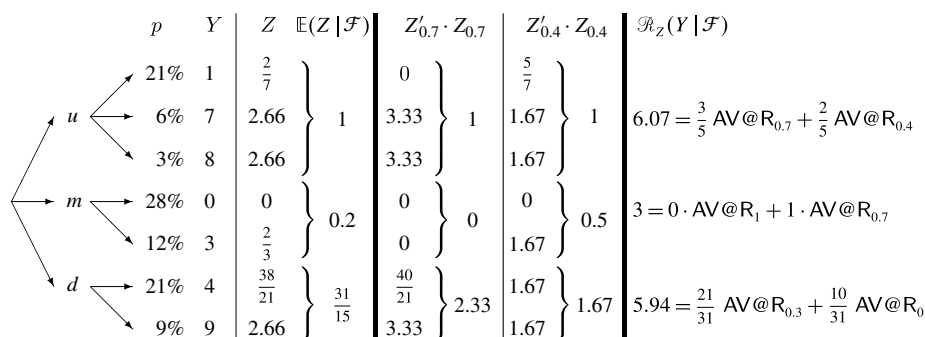


FIGURE A.1. Nested decomposition of $\mathcal{R} = \frac{3}{5} \text{AV@R}_{0.7}(Y) + \frac{2}{5} \text{AV@R}_{0.4}(Y)$. As the dual variable Z is optimal for (23), the conditional risk functional has the representation (17), $\mathcal{R}_Z(Y|\mathcal{F}) = (\int Z_\alpha \text{AV@R}_{1-(1-\alpha)Z_\alpha}(Y|\mathcal{F}) \mu(d\alpha)) / (\int Z_\alpha \mu(d\alpha))$, which is indicated on the very right.

the conditional probabilities are displayed in the tree's nodes. We demonstrate that the optimal conditional risk functionals are⁸

$$\begin{aligned}\mathcal{R}(\cdot | u) &= \frac{3}{5} \cdot \text{AV@R}_{0.7}(\cdot) + \frac{2}{5} \cdot \text{AV@R}_{0.4}(\cdot), \\ \mathcal{R}(\cdot | m) &= 0 \cdot \text{AV@R}_1(\cdot) + 1 \cdot \text{AV@R}_{0.7}(\cdot) = \text{AV@R}_{0.7}(\cdot), \\ \mathcal{R}(\cdot | d) &= \frac{21}{31} \cdot \text{AV@R}_{0.3}(\cdot) + \frac{10}{31} \cdot \text{AV@R}_0(\cdot) = \frac{21}{31} \cdot \text{AV@R}_{0.3}(\cdot) + \frac{10}{31} \cdot \text{ess sup}(\cdot),\end{aligned}\quad (24)$$

which is a different risk functional at each node. Note that not just the risk levels differ in between and from the original risk functional (23), but the weights differ as well.

The associated spectral density for Kusuoka's measure $\mu = \frac{3}{5}\delta_{0.7} + \frac{2}{5}\delta_{0.4}$ is (see (2))

$$\sigma(\alpha) = 2 \cdot \mathbb{I}_{[0.7, 1]}(\alpha) + \frac{2}{3} \cdot \mathbb{I}_{[0.4, 1]}(\alpha).$$

We consider first the Average Values-at-Risk for both levels, which evaluate to

$$\text{AV@R}_{0.7}(Y) = 6.5 \quad \text{and} \quad \text{AV@R}_{0.4}(Y) = 4.6,$$

such that $\mathcal{R}(Y) = 5.74$ by (23).

The corresponding dual variables $Z_{0.7} \cdot Z'_{0.7}$ ($Z_{0.4} \cdot Z'_{0.4}$, respectively), as well as its conditional version $Z_{0.7} = \mathbb{E}(Z_{0.7}Z'_{0.7} | \mathcal{F}_t)$ ($\mathbb{E}(Z_{0.4}Z'_{0.4} | \mathcal{F}_t)$, resp.) with respect to the filtration induced by the tree, are displayed in Figure A.1.

Next we compute both conditional Average Value-at-Risk at random level $1 - (1 - 0.7) \cdot Z_{0.7}$ and $1 - (1 - 0.4) \cdot Z_{0.4}$ for Y , which have the outcomes

$$\text{AV@R}_{1-(1-0.7) \cdot Z_{0.7}}(Y | \mathcal{F}_t) = \begin{cases} \text{AV@R}_{0.7}(Y | u) = 7.33, \\ \text{AV@R}_1(Y | m) = 3, \\ \text{AV@R}_{0.3}(Y | d) = 43/7, \end{cases}$$

and

$$\text{AV@R}_{1-(1-0.4) \cdot Z_{0.4}}(Y | \mathcal{F}_t) = \begin{cases} \text{AV@R}_{0.4}(Y | u) = 25/6 \approx 4.17, \\ \text{AV@R}_{0.7}(Y | m) = 3, \\ \text{AV@R}_0(Y | d) = 5.5. \end{cases}$$

Notice now that

$$\begin{aligned}\mathbb{E}(Z_{0.7} \cdot \text{AV@R}_{1-(1-0.7) \cdot Z_{0.7}}(Y | \mathcal{F}_t)) &= 6.5 = \text{AV@R}_{0.7}(Y) \quad \text{and} \\ \mathbb{E}(Z_{0.4} \cdot \text{AV@R}_{1-(1-0.4) \cdot Z_{0.4}}(Y | \mathcal{F}_t)) &= 4.6 = \text{AV@R}_{0.4}(Y),\end{aligned}$$

which is the content of the decomposition Theorem 21 for the Average Value-at-Risk at its respective levels $\alpha = 0.7$ and $\alpha = 0.4$.

Next consider the random variables $Z = \int Z_\alpha \mu(d\alpha) = \frac{3}{5}Z_{0.7} + \frac{2}{5}Z_{0.4}$ and $ZZ' = \int Z_\alpha Z'_\alpha \mu(d\alpha) = \frac{3}{5}Z_{0.7}Z'_{0.7} + \frac{2}{5}Z_{0.4}Z'_{0.4}$ —built according to (15) and depicted in Figure A.1. ZZ' is feasible and it holds that

$$\mathcal{R}(Y) = \frac{3}{5}\text{AV@R}_{0.7}(Y) + \frac{2}{5}\text{AV@R}_{0.4}(Y) = 5.74 = \mathbb{E}YZZ'. \quad (25)$$

According to the proof of Theorem 21 and (17) one needs to consider

$$\begin{aligned}\mathcal{R}_Z(Y | \mathcal{F}_t) &= \frac{\int Z_\alpha \text{AV@R}_{1-(1-\alpha)Z_\alpha}(Y | \mathcal{F}_t) \mu(d\alpha)}{\int Z_\alpha \mu(d\alpha)} \\ &= \begin{cases} \frac{3}{5} \cdot \text{AV@R}_{0.7}(Y | u) + \frac{2}{5} \cdot \text{AV@R}_{0.4}(Y | u) = 91/15 \approx 6.07, \\ 0 \cdot \text{AV@R}_1(Y | m) + 1 \cdot \text{AV@R}_{0.7}(Y | m) = 3, \\ \frac{21}{31} \cdot \text{AV@R}_{0.3}(Y | d) + \frac{10}{31} \cdot \text{AV@R}_0(Y | d) = 184/31 \approx 5.94. \end{cases}\end{aligned}\quad (26)$$

At each node in (26) the corresponding risk functionals in its respective Kusuoka representation are given by (24). These risk functionals in general *differ* from the initial risk functional (23). However, any of these risk functionals is built of at most two AV@R's, as is (23), and they are version independent (law invariant).

It is evident that the risk functionals (24) have different risk levels α , but they have different weights as well (Kusuoka representation).

The representation theorem (Theorem 21) finally ensures that

$$\mathcal{R}(Y) = \mathbb{E}(Z \cdot \mathcal{R}_Z(Y | \mathcal{F}_t)) = 5.74,$$

which is in accordance with (25).

⁸ u (up) stands for $Y \in \{1, 7, 8\}$, m (mid) for $Y \in \{0, 3\}$ and d (down) for $Y \in \{4, 9\}$ such that $\mathcal{F}_t = \{u, m, d\}$.

References

- [1] Acerbi C (2002) Spectral measures of risk: A coherent representation of subjective risk aversion. *J. Banking Finance* 26:1505–1518.
- [2] Acerbi C, Simonetti P (2002) Portfolio optimization with spectral measures of risk. arXiv:cond-mat/0203607v1.
- [3] Artzner P, Delbaen F, Heath D (1997) Thinking coherently. *Risk* 10:68–71.
- [4] Artzner P, Delbaen F, Eber J-M, Heath D (1999) Coherent measures of risk. *Math. Finance* 9:203–228.
- [5] Artzner P, Delbaen F, Eber J-M, Heath D, Ku H (2007) Coherent multiperiod risk adjusted values and Bellman's principle. *Ann. Oper. Res.* 152:5–22.
- [6] Asamov T, Ruszczyński A (2014) Time-consistent approximations of risk-averse multistage stochastic optimization problems. *Math. Programming* 153(2):1–35.
- [7] Billingsley P (1979) *Probability and Measure*, Wiley series in probability and mathematical statistics (Jon Wiley & Sons, New York).
- [8] Carpentier P, Chancelier J-P, Cohen G, de Lara M, Girardeau P (2012) Dynamic consistency for stochastic optimal control problems. *Ann. Oper. Res.* 200(1):247–263.
- [9] Cheridito P, Kupper M (2009) Recursiveness of indifference prices and translation-invariant preferences. *Math. Financial Econom.* 2(3):173–188.
- [10] Cheridito P, Kupper M (2011) Composition of time-consistent dynamic monetary risk measures in discrete time. *Internat. J. Theoret. Appl. Finance (IJTAF)* 14(1):137–162.
- [11] Dunford N, Schwartz JT (1957) *Linear Operators. Part I. General Theory* (John Wiley & Sons, New York).
- [12] Graf S, Mauldin RD (1989) A classification of disintegrations of measures. *Contemporary Math.* 94:147–158.
- [13] Iancu DA, Petrik M, Subramanian D (2015) Tight approximations of dynamic risk measures. *Math. Oper. Res.* 40(3):655–682.
- [14] Jouini E, Schachermayer W, Touzi N (2006) Law invariant risk measures have the Fatou property. Kusuoka S, Yamazaki A, eds., *Advances in Mathematical Economics, vol. 9* (Springer, Japan), 49–71.
- [15] Karatzas I, Shreve SE (1998) *Methods of Mathematical Finance: Stochastic Modelling and Applied Probability* (Springer, New York).
- [16] Kovacevic R, Pflug GCh (2009) Time consistency and information monotonicity of multiperiod acceptability functionals. Albrecher H, Runggaldier WJ, Schachermayer W, eds. *Advanced Financial Modelling* (de Gruyter, Berlin), 347–369.
- [17] Kupper M, Schachermayer W (2009) Representation results for law invariant time consistent functions. *Math. Financial Econom.* 2:189–210.
- [18] Kusuoka S (2001) On law invariant coherent risk measures. *Ad. Math. Econom.* 3:83–95.
- [19] Miller N, Ruszczyński A (2011) Risk-averse two-stage stochastic linear programming: Modeling and decomposition. *Oper. Res.* 59(1):125–132.
- [20] Müller A, Stoyan D (2002) *Comparison Methods for Stochastic Models and Risks*, Wiley series in probability and statistics (John Wiley & Sons, Chichester, UK).
- [21] Ogryczak W, Ruszczyński A (2002) Dual stochastic dominance and related mean-risk models. *SIAM J. Optim.* 13(1):60–78.
- [22] Pflug GCh (2006) On distortion functionals. *Statist. Risk Modeling* 24:45–60.
- [23] Pflug GCh, Römisch W (2007) *Modeling, Measuring and Managing Risk* (World Scientific, River Edge, NJ).
- [24] Pichler A (2013) The natural Banach space for version independent risk measures. *Insurance: Math. Econom.* 53(2):405–415, <http://dx.doi.org/10.1016/j.insmatheco.2013.07.005>.
- [25] Pichler A (2015) Premiums and reserves, adjusted by distortions. *Scandinavian Actuarial J.* 2015(4):332–351.
- [26] Rockafellar RT (1974) *Conjugate Duality and Optimization*, Vol. 16. CBMS-NSF Regional Conference Series in Applied Mathematics (SIAM, Philadelphia).
- [27] Rockafellar RT, Wets RJ-B (1997) *Variational Analysis* (Springer, Berlin).
- [28] Ruszczyński A (2010) Risk-averse dynamic programming for Markov decision processes. *Math. Program., Ser. B* 125:235–261.
- [29] Ruszczyński A, Dentcheva D (2006) Portfolio optimization with stochastic dominance constraints. *J. Banking Finance* 30:433–451.
- [30] Ruszczyński A, Shapiro A (2006) Optimization of convex risk functions. *Math. Oper. Res.* 31(3):433–452.
- [31] Shaked M, Shanthikumar JG (2007) *Stochastic Orders*. Springer Series in Statistics (Springer, New York).
- [32] Shapiro A (2009) On a time consistency concept in risk averse multistage stochastic programming. *Oper. Res. Lett.* 37(3):143–147.
- [33] Shapiro A (2012) Time consistency of dynamic risk measures. *Oper. Res. Lett.* 40(6):436–439.
- [34] Shapiro A (2013) On Kusuoka representation of law invariant risk measures. *Math. Oper. Res.* 38(1):142–152.
- [35] Shapiro A, Dentcheva D, Ruszczyński A (2009) *Lectures on Stochastic Programming*. MOS-SIAM Series on Optimization, <http://dx.doi.org/10.1137/1.9780898718751>.
- [36] Wang T (1999) A class of dynamic risk measures. Working paper, University of British Columbia, Vancouver.
- [37] Williams D (1991) *Probability with Martingales* (Cambridge University Press, Cambridge, UK).
- [38] Wojtaszczyk P (1991) *Banach Spaces for Analysts* (Cambridge University Press, Cambridge, UK).