Time Consistent Decisions and Temporal Decomposition of Coherent Risk Functionals

Georg Ch. Pflug,* Alois Pichler †

March 29, 2014

Abstract

In management and planning it is a daily reality that more and more information is available gradually over time. It is well known that most risk measures (risk functionals) are time *in*consistent in this situation in the following sense: it may happen that today some loss distribution appears to be less risky than another, 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 allows a temporal decomposition of the initial risk functional in a way, which is consistent with the past and the future. The central result is a decomposition theorem, which allows recomposing the initial coherent risk functional by compounding the conditional risk functionals without loosing 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 different policies, even based on incomplete information.

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

Keywords: Risk Measures, Time Consistency, Dual Representation
Classification: 90C15, 60B05, 62P05

1 Introduction

Risk measures have been introduced in the pioneering paper by Artzner et al. [3], and they have become increasingly important since then. They are supposed to measure and quantify the risk, which is associated with a particular random outcome.

In various situations the random outcome is not revealed suddenly, but gradually, such that the information available is growing steadily: if additional information is available, then a conscientious decision maker will question 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.

Consider a random variable Y describing the financial loss associated with a certain activity. Profits are treated as negative losses. Denote by $\mathcal{R}$ some risk functional, for instance the (upper) Average Value-at-Risk ($\mathsf{AV@R}$, see Section 2 below).

The simple example in Figure 1 demonstrates the fact that sticking to the same, genuine risk functional at all decision stages (in this case to the $\mathsf{AV@R}$ at the same level) may lead to contradicting, conflicting, and even wrong decisions. So how does the decision maker have to change his objective given the information that was already revealed?

*University of Vienna. Department of Statistics and Operations Research.
International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria.
†Norwegian University of Science and Technology, NTNU
Contact: aloisp@ntnu.no

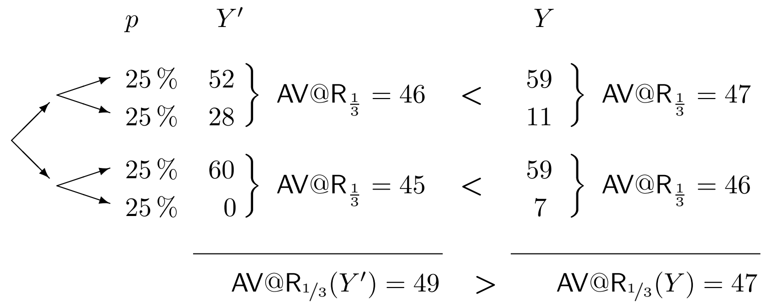

Figure 1: This example addresses the Average Value-at-Risk at level $\alpha = 1/3$. The random variable Y is preferred over Y', as it has *lower* Average Value-at-Risk at level $\alpha = 1/3$, $\mathsf{AV@R}_{1/3}(Y') > \mathsf{AV@R}_{1/3}(Y)$. But both conditional observations support the opposite result: it holds that $\mathsf{AV@R}_{1/3}(Y'|\mathcal{F}_1) < \mathsf{AV@R}_{1/3}(Y|\mathcal{F}_1)$, the conditional observations thus suggest that Y' should be preferred over Y. Details are elaborated in Section 7.

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, which allows a re-composition of the initial (positively homogeneous) risk functional. The decomposition is given for the Average Value-at-Risk first, and 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 [8, 9], Pflug and Kovacevic [19, 13], and other authors). The results by Kupper and Schachermayer [27], as well as by Shapiro [31] 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 (Section 7) investigates a further concept of consistency, which leads to the same type of negative results.

To obtain time consistency desirable properties thus have to be relaxed (for example law invariance (see below) in the papers [15] by Ruszczyński et al.), or different concepts have to be investigated. This paper demonstrates a method which allows re-assembling 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 which is already available. The conditional risk functional differs from the initial risk functional, but it keeps its focus on the original risk functional and retains 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 Section 2. The dual representation of risk functionals (Theorem 9), which is elaborated in Section 3, is the first result of this paper and an essential component to defining the conditional risk functional in Section 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 [3, 5, 4]. The definitions are stated here, as they are not used consistently in the literature.

Definition 1. Let $L^\infty(\mathcal{F}, P)$ (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 $\mathcal{R}\colon L^\infty \to \mathbb{R}$ with the following properties:

(i) Monotonicity: $\mathcal{R}(Y_1) \le \mathcal{R}(Y_2)$ whenever $Y_1 \le Y_2$ almost surely;

(ii) Convexity: $\mathcal{R}((1-\lambda)Y_0 + \lambda Y_1) \le (1-\lambda)\mathcal{R}(Y_0) + \lambda\mathcal{R}(Y_1)$ for $0 \le \lambda \le 1$;

(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 $\rho\colon Y \mapsto \mathcal{R}(-Y)$ 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.

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 ($p \ge 1$) we refer to [20].

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

$$\mathsf{AV@R}_\alpha(Y) := \frac{1}{1-\alpha}\int_\alpha^1 \mathsf{V@R}_p(Y)\,\mathrm{d}p \quad (0 \le \alpha < 1)$$

and

$$\mathsf{AV@R}_1(Y) := \lim_{\alpha \nearrow 1} \mathsf{AV@R}_\alpha(Y) = \operatorname{ess\,sup}(Y),$$

where

$$\mathsf{V@R}_\alpha(Y) := \inf\{y\colon P(Y \le y) \ge \alpha\}$$

is the *Value-at-Risk* (the left-continuous, lower semi-continuous *quantile* or *lower inverse* cdf) at level α, often denoted $\mathsf{V@R}_\alpha(Y) = F_Y^{-1}(\alpha)$ as well ($F_Y(y) = P(Y \le y)$ 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, $P(Y_1 \le y) = P(Y_2 \le y)$ for all $y \in \mathbb{R}$.

[1] In an economic or monetary environment this is often called Cash invariance instead.
[2] often also *law invariant* or *distribution based*.

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

Theorem 3 (Kusuoka's representation)**.** *Let $\mathcal{B}$ collect the Borel sets on $[0,1]$ and the probability measure P be atom-less. Then any version independent, coherent risk functional $\mathcal{R}$ on $L^\infty([0,1],\mathcal{B},P)$ has the representation*

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

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

Proof. Cf. Kusuoka's original paper [14] or [30] by Shapiro, in combination with the paper [11] by Jouini et al.. □

Spectral Risk Functionals

A subclass of risk functionals is given when the set of measures $\mathscr{M}$ in (1) reduces to a singleton, $\mathscr{M} = \{\mu\}$. This class can be equivalently described as the class of comonotone additive functionals. In this paper, we call them (following Acerbi, cf. [2, 1]) *spectral risk functionals*, although the term distortion risk functional 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 $\mathscr{M} = \{\delta_\alpha\}$. Assuming that

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

is well-defined,[3] it follows by integration by parts that[4]

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

Definition 4. The risk functional with representation

$$\mathcal{R}_\sigma(Y) := \int_0^1 \mathsf{V@R}_p(Y)\,\sigma(p)\,\mathrm{d}p$$

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 non-decreasing probability density on $[0,1]$ (details are elaborated in Pflug [18]).

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 $\mathsf{AV@R}$s.

Corollary 5 (Kusuoka representation in terms of spectral risk functionals)**.** *For every version independent risk functional $\mathcal{R}$ as in Theorem 3 there is a set $\mathcal{S}$ of spectral densities such that*

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

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

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

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

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

[4] The inverse of this operation is given by the measure with distribution function $\mu_\sigma(\alpha) := (1-\alpha)\,\sigma(\alpha) + \int_0^\alpha \sigma(u)\mathrm{d}u$.

3 The Dual of Risk Functionals

The decomposition of risk functionals with respect to incomplete information employs the dual, or convex conjugate function (cf. the book [32] and paper [26] by Ruszczyński et al.). As any risk functional $\mathcal{R}: L^\infty \to \mathbb{R}$ is Lipschitz continuous (with Lipschitz constant 1) the Fenchel–Moreau duality theorem can be stated in the following way:

Theorem 6 (Fenchel–Moreau duality theorem)**.** *The risk functional $\mathcal{R}: L^\infty \to \mathbb{R}$ has the representation*

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

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

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

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 [30], and for the concept of stochastic dominance relations in general we refer to the books by Stoyan and Müller [16] and Shaked et al. [28], as well as to the papers [17] and [25] by Ruszczyiński et al..

Definition 7 (Second order stochastic dominance, or convex ordering)**.** Let $\tau, \sigma : (0,1) \to [0,\infty)$ be integrable functions. It is said that σ majorizes τ (denoted $\sigma \succcurlyeq \tau$ or $\tau \preccurlyeq \sigma$) iff

$$\int_\alpha^1 \tau(p)\,\mathrm{d}p \le \int_\alpha^1 \sigma(p)\,\mathrm{d}p \;\text{ for all } \alpha \in [0,1] \text{ and } \int_0^1 \tau(p)\,\mathrm{d}p = \int_0^1 \sigma(p)\,\mathrm{d}p = 1.$$

Remark 8. For the inverse quantile function F_Z^{-1} the relation $F_Z^{-1} \preccurlyeq \sigma$ is notably equivalent to

$$\mathbb{E}(Z) = \int_0^1 \sigma(p)\,\mathrm{d}p = 1, \quad Z \ge 0$$

and

$$(1-\alpha)\mathsf{AV@R}_\alpha(Z) \le \int_\alpha^1 \sigma(p)\,\mathrm{d}p \quad \text{ for all } \alpha \in [0,1].$$

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 (cf. Ruszczyński and Ogryczak [17]).

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*

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

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

Proof. The dual representation

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

for spectral risk functionals can be derived from Shapiro [30] and is contained in [21]. The assertion for general risk functionals $\mathcal{R}_{\mathcal{S}}$ is immediate now from the definition, as

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

by using (7). □

4 Convex Duality

The dual representation (Equation (6)) suggests to consider the following function as a candidate for the dual representation,

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

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

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

The inequality $\mathcal{R}_{\mathcal{S}}^{*}(Z)\le\tilde{\mathcal{R}}_{\mathcal{S}}(Z)$ 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 $\mathcal{R}_{\mathcal{S}}^{*}$ precisely by an augmented spectrum $\mathcal{S}^{*}$, such that $\mathcal{R}_{\mathcal{S}}^{*}(Z)=\tilde{\mathcal{R}}_{\mathcal{S}^{*}}(Z)$ holds true.

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

$$\begin{aligned}\mathcal{S}^{*} &= \operatorname{conv}\{\tau\colon [0,1]\to[0,\infty)\,|\tau \text{ is non-decreasing, lower semi-continuous, and } \exists\sigma\in\mathcal{S}:\tau\preccurlyeq\sigma\}\\ &= \operatorname{conv}\left\{\tau\colon [0,1]\to[0,\infty)\ \middle|\ \begin{array}{l}\tau \text{ is non-decreasing, lower semi-continuous, } \int_0^1\tau(p)\,\mathrm{d}p=1 \text{ and}\\ \exists\sigma\in\mathcal{S}\colon \int_{\alpha}^{1}\tau(p)\,\mathrm{d}p\le\int_{\alpha}^{1}\sigma(p)\,\mathrm{d}p \text{ for all } 0\le\alpha\le 1\end{array}\right\}\end{aligned}$$

of spectral functions.

Proof. Define

$$\mathcal{Z}_{\mathcal{S}} := \left\{Z\in L^{1}\colon \exists\sigma\in\mathcal{S}\colon F_Z^{-1}\preccurlyeq\sigma\right\}$$

and observe that $s_{\mathcal{Z}_{\mathcal{S}}}(Y)=\mathcal{R}_{\mathcal{S}}(Y)$, 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 (cf. Rockafellar [22]) that $s_{\mathcal{Z}_{\mathcal{S}}}^{*}=\mathbb{I}_{\overline{\operatorname{conv}\mathcal{Z}_{\mathcal{S}}}}$, where $\mathbb{I}$ is the indicator function of its index set,[5] and hence $\mathcal{R}_{\mathcal{S}}^{*}=\mathbb{I}_{\overline{\operatorname{conv}\mathcal{Z}_{\mathcal{S}}}}$ and $\mathcal{R}_{\mathcal{S}}(Y)=\sup_{Z\in\overline{\operatorname{conv}\mathcal{Z}_{\mathcal{S}}}}\mathbb{E}(YZ)$. We show that $\mathcal{Z}_{\mathcal{S}}=\mathcal{Z}_{\mathcal{S}^{*}}$.

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

Moreover $\mathcal{Z}_{\mathcal{S}}\subset\mathcal{Z}_{\mathcal{S}^{*}}$ by considering the lower semi-continuous and increasing function $\sigma_Z(\alpha)=\mathsf{V@R}_{\alpha}(Z)$ for any $Z\in\mathcal{Z}_{\mathcal{S}}$.

The assertion finally follows from $\operatorname{conv}\mathcal{Z}_{\mathcal{S}}=\mathcal{Z}_{\operatorname{conv}\mathcal{S}}$, which completes the proof. □

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

Remark 11. In the present context of increasing functions the requirement *lower semi-continuous* for $\tau \in \mathcal{S}^*$ is equivalent to τ *continuous from the left*. Moreover, the constant function $\tau \equiv \mathbb{1}$ is always contained in the set $\mathcal{S}^*$; the constant function $\tau \equiv \mathbb{1}$ is associated with the expected value.

Remark 12. The closure of the convex set $\operatorname{conv} \mathcal{S}^*$ coincides (by Mazur's Theorem on convex sets, cf. Wojtaszczyk [35, II.A.4]) with the $\sigma\left(L^1, L^\infty\right)$ (or weak) closure.

Feasible Dual Variables

The following two remarks prepare for the main theorem of this section. We shall write $Z \lhd \mathcal{F}$ 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

$$\mathcal{R}\left(\mathbb{E}\left(Y|\mathcal{F}_t\right)\right) \le \mathcal{R}\left(Y\right). \tag{8}$$

The proof relies on the conditional Jensen inequality (cf. [6, Section 34] or [34, Chapter 9]) and is contained in [32, Corollary 6.30].

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

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

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

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

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

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

Theorem 15 (Bochner representation of feasible dual variables)**.** *Let $Z_\alpha \lhd \mathcal{F}_t$ satisfy $0 \le Z_\alpha \le \frac{1}{1-\alpha}$, $\mathbb{E}\left(Z_\alpha\right) = 1$, and $\mu \in \mathscr{M}$ be a probability measure on $[0,1]$. Then the Bochner-Integral $\int_0^1 Z_\alpha \mu\left(\mathrm{d}\alpha\right)$ (provided that $\alpha \mapsto Z_\alpha$ is measurable and the integral exists) is feasible for $\mathcal{R}_{\sigma_\mu}$, that is*

$$(1-\alpha')\mathsf{AV@R}_{\alpha'}\left(\int_0^1 Z_\alpha \mu\left(\mathrm{d}\alpha\right)\right) \le \int_{\alpha'}^1 \sigma_\mu\left(p\right)\mathrm{d}p \qquad (0 \le \alpha' \le 1),$$

where σ_μ is the spectral function associated with μ (cf. (3)). A fortiori it is feasible for $\sup_{\sigma \in \mathcal{S}} \mathcal{R}_\sigma$, provided that $\sigma_\mu \in \mathcal{S}^*$.

Proof. The Average Value-at-Risk is convex, hence

$$\mathsf{AV@R}_{\alpha'}\left(\int_0^1 Z_\alpha \mu\left(\mathrm{d}\alpha\right)\right) \le \int_0^1 \mathsf{AV@R}_{\alpha'}\left(Z_\alpha\right)\mu\left(\mathrm{d}\alpha\right).$$

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

$$\mathsf{AV@R}_{\alpha'}\left(Z_\alpha\right) = \frac{1}{1-\alpha'}\int_{\alpha'}^1 \mathsf{V@R}_p\left(Z_\alpha\right)\mathrm{d}p \le \frac{1}{1-\alpha'}\int_0^1 \mathsf{V@R}_p\left(Z_\alpha\right)\mathrm{d}p = \frac{\mathbb{E}\left(Z_\alpha\right)}{1-\alpha'} = \frac{1}{1-\alpha'},$$

implying that

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

and hence $\int_0^1 Z_\alpha\mu\left(\mathrm{d}\alpha\right)$ is feasible by Theorem 9. ☐

5 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 $Y\in L^\infty\left(\mathcal{F},P\right)$ and consider a sub-sigma-algebra $\mathcal{F}_t$ ($\mathcal{F}_t\subset\mathcal{F}$). As above we shall write $Z\lhd\mathcal{F}_t$ 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 conditional random variables $Y|\mathcal{F}$. We extend this concept by allowing a change-of-measure through a probability density of the random variable Z', which is chosen in an optimal way.

Definition 16 (Conditional risk functional)**.** Let $\mathcal{R}_\mathcal{S}$ be a coherent risk functional induced by the spectrum $\mathcal{S}$. For all feasible duals $Z_t\lhd\mathcal{F}_t$, the *conditional risk functional* is defined as [6]

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

which can be rewritten as

$$\mathcal{R}_{\mathcal{S};Z_t}\left(Y|\mathcal{F}_t\right)=\operatorname{ess\,sup}\left\{\mathbb{E}\left(YZ'|\mathcal{F}_t\right)\,\middle|\,\begin{array}{l}\mathbb{E}\left(Z'|\mathcal{F}_t\right)=\mathbb{1},\text{ and}\\ \exists\sigma\in\mathcal{S}^*:\ F^{-1}_{Z_t\cdot Z'}\preccurlyeq\sigma\end{array}\right\}.$$

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

$$\mathsf{AV@R}_{\alpha;Z_t}\left(Y|\mathcal{F}_t\right)=\operatorname{ess\,sup}\left\{\mathbb{E}\left(YZ'|\mathcal{F}_t\right)\,\middle|\,\begin{array}{l}\mathbb{E}\left(Z'|\mathcal{F}_t\right)=\mathbb{1},\ Z'\ge 0\text{ and}\\ (1-\alpha)Z_tZ'\le\mathbb{1}\end{array}\right\}\tag{9}$$

whenever $\alpha\le 1$. This representation inspires an equivalent definition of the Average Value-at-Risk 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* $\alpha\lhd\mathcal{F}_t$ ($0\le\alpha\le\mathbb{1}$ a.s.) is the $\mathcal{F}_t$-random variable[7]

$$\mathsf{AV@R}_\alpha\left(Y|\mathcal{F}_t\right):=\operatorname{ess\,sup}\left\{\mathbb{E}\left(YZ'|\mathcal{F}_t\right)|\,\mathbb{E}\left(Z'|\mathcal{F}_t\right)=\mathbb{1},\ Z'\ge 0\text{ and }(1-\alpha)Z'\le\mathbb{1}\right\}.\tag{10}$$

[6] See e.g. Appendix A in [12] or [10] for a rigorous definition and discussion of the essential supremum of a family of random variables.

[7] Cf. also [9, Section 2.3.1]

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

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

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

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

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

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

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

$$\mathsf{AV@R}_0(Y|\mathcal{F}_t) \le \mathcal{R}_{\mathcal{S};Z}(Y|\mathcal{F}_t) \le \mathsf{AV@R}_1(Y|\mathcal{F}_t)$$

in Theorem 20 below.

5.2 Elementary Properties of Conditional Risk Functionals

The conditional risk functional inherits, on conditional basis, all essential properties of the original risk functional. Notice that our 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* $Y \lhd \mathcal{F}_t$*;*

(ii) Translation equivariance: $\mathcal{R}_Z(Y+c|\mathcal{F}_t) = \mathcal{R}_Z(Y|\mathcal{F}_t) + c$ *if* $c \lhd \mathcal{F}_t$*;*

(iii) Positive homogeneity: $\mathcal{R}_Z(\lambda Y|\mathcal{F}_t) = \lambda \mathcal{R}_Z(Y|\mathcal{F}_t)$ *whenever* $\lambda \lhd \mathcal{F}_t$*,* $\lambda \ge 0$ *and* λ *bounded;*

(iv) Monotonicity: $\mathcal{R}_Z(Y_1|\mathcal{F}_t) \le \mathcal{R}_Z(Y_2|\mathcal{F}_t)$ *whenever* $Y_1 \le Y_2$*;*

(v) Convexity: *The mapping* $Y \mapsto \mathcal{R}_Z(Y|\,\mathcal{F}_t)$ *is convex, that is*

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

for $\lambda \lhd \mathcal{F}_t$ *and* $0 \le \lambda \le 1$*, almost surely;*

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

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

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 $Y \lhd \mathcal{F}_t$,

(ii) Translation Equivariance 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 from $\operatorname{ess\,sup} \mathbb{E}(\lambda Y Z'|\mathcal{F}_t) = \lambda \operatorname{ess\,sup} \mathbb{E}(Z'|\mathcal{F}_t)$, as $0 \le \lambda \lhd \mathcal{F}_t$.

(iv) Monotonicity is inherited from the conditional expected value, as $\mathbb{E}\left(Y_1 Z'|F_t\right) \leq \mathbb{E}\left(Y_2 Z'|F_t\right)$ whenever $Y_1 \leq Y_2$ and $Z' \geq 0$.

(v) For Convexity observe that

$$\begin{aligned}
&(1-\lambda)\,\mathcal{R}_{Z_t}\left(Y_0|\mathcal{F}_t\right)+\lambda\mathcal{R}_{Z_t}\left(Y_1|\mathcal{F}_t\right)\\
&\qquad=(1-\lambda)\operatorname*{ess\,sup}_{\mathcal{R}^*\left(Z_t Z_0'\right)=0}\mathbb{E}\left(Y_0 Z_0'|F_t\right)+\lambda\operatorname*{ess\,sup}_{\mathcal{R}^*\left(Z_t Z_1'\right)=0}\mathbb{E}\left(Y Z_1'|F_t\right)\\
&\qquad\geq\operatorname*{ess\,sup}_{\mathcal{R}^*\left(Z_t Z'\right)=0}(1-\lambda)\,\mathbb{E}\left(Y_0 Z'|F_t\right)+\lambda\mathbb{E}\left(Y Z'|F_t\right)\\
&\qquad=\mathcal{R}_{Z_t}\left(\left(1-\lambda\right)Y_0+\lambda Y_1|\mathcal{F}_t\right).
\end{aligned}$$

(vi) For Concavity let $Z_0,\, Z_1 \lhd \mathcal{F}_t$ be feasible and Z_0' and Z_1' be chosen such that $\mathcal{R}^*\left(Z_0 Z_0'\right) < \infty$ and $\mathcal{R}^*\left(Z_1 Z_1'\right) < \infty$. 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 $\mathbb{E}\left(Z_\lambda'|\mathcal{F}_t\right) = \begin{cases}\frac{(1-\lambda)Z_0+\lambda Z_1}{Z_\lambda} & \text{if } Z_\lambda > 0\\ 1 & \text{else}\end{cases} = \mathbb{1}$. Then, by convexity of $\mathcal{Z}_{\mathcal{S}^*}$, $Z_\lambda := (1-\lambda)\,Z_0 + \lambda Z_1 \lhd \mathcal{F}_t$ is feasible as well, and $\mathcal{R}^*\left(Z_\lambda Z_\lambda'\right) \leq (1-\lambda)\,\mathcal{R}^*\left(Z_0 Z_0'\right) + \lambda\mathcal{R}^*\left(Z_1 Z_1'\right) < \infty$, such that $Z_\lambda Z_\lambda'$ as well is feasible. It follows that

$$Z_\lambda\cdot\mathcal{R}_{Z_\lambda}\left(Y|\,\mathcal{F}_t\right)\geq Z_\lambda\cdot\mathbb{E}\left(Y Z_\lambda'|\mathcal{F}_t\right)=\left(1-\lambda\right)Z_0\mathbb{E}\left(Y Z_0'|\mathcal{F}_t\right)+\lambda Z_1\mathbb{E}\left(Y Z_1'|\mathcal{F}_t\right)$$

Taking the essential supremum (with respect to Z_0' and Z_1') reveals that

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

which is the assertion.

□

Theorem 20 (Lower and upper bounds)**.** *Let $\mathcal{F}_\tau \supset \mathcal{F}_t$ be sigma algebras. Then the following inequalities hold true:*

(i) $\mathbb{E}\left(Y|\mathcal{F}_t\right) \leq \mathcal{R}_Z\left(\mathbb{E}\left(Y|\mathcal{F}_\tau\right)|\,\mathcal{F}_t\right) \leq \mathcal{R}_Z\left(Y|\mathcal{F}_t\right) \leq \mathsf{AV@R}_1\left(Y|\,\mathcal{F}_t\right) \leq \mathsf{AV@R}_1\left(Y\right) = \operatorname{ess\,sup}(Y)$,

(ii) $\mathbb{E}\left(Y\right) \leq \mathbb{E}\left(\mathcal{R}_{Z_t}\left(Y|\mathcal{F}_t\right)\right) \leq \mathcal{R}\left(Y\right) \leq \mathsf{AV@R}_1\left(Y\right) = \operatorname{ess\,sup}(Y)$.

Proof. By definition

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

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

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

which is the first inequality in (i).

Next,

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

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

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

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

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 $\mathcal{R}=\mathcal{R}_\mathcal{S}$ be a version independent risk functional.*

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

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

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

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

$$\mathcal{R}_\mathcal{S}\left(Y|\mathcal{F}_t\right)=\operatorname*{ess\,sup}\mathbb{E}\Big[Z_\tau\cdot\mathcal{R}_{\mathcal{S};Z_\tau}\left(Y|\mathcal{F}_\tau\right)\Big|\,\mathcal{F}_t\Big],\tag{13}$$

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

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)**.** *The decomposition for the Average Value-at-Risk is*

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

where the supremum is over all random variables $Z\lhd\mathcal{F}_t$ satisfying $\mathbb{E}(Z)=1$, $Z\geq0$ 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 $\mathsf{AV@R}_{1-(1-\alpha)Z}\left(Y|\mathcal{F}_t\right)$ (cf. (9)), that is they satisfy $Z'_i\geq0$, $\mathbb{E}\left(Z'_i|\mathcal{F}_t\right)=\mathbb{1}$ and $(1-\alpha)ZZ'_i=\left(1-\left[1-(1-\alpha)Z\right]\right)Z'_i\leq\mathbb{1}$.

Define $Z':=\sum_i\mathbb{1}_{B_i}Z'_i$. It is immediate that $(1-\alpha)ZZ'\leq\mathbb{1}$ and $ZZ'\geq0$. Moreover $\mathbb{E}\left(Z'|\mathcal{F}_t\right)=\sum_i\mathbb{1}_{B_i}=\mathbb{1}$, from which follows that

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

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

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

By the dual representations of the Average Value-at-Risk and $\mathsf{AV@R}(Z) \geq \mathsf{AV@R}(\mathbb{E}(Z|F_t))$ (cf. (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

$$\mathsf{AV@R}_\alpha(Y) = \mathbb{E}(Y\tilde{Z}) = \mathbb{E}(ZYZ') = \mathbb{E}\left(Z\,\mathbb{E}\left(YZ'|\mathcal{F}_t\right)\right) \leq \mathbb{E}\left(Z\,\mathsf{AV@R}_{1-(1-\alpha)Z}(Y|\mathcal{F}_t)\right),$$

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

For the remaining situation $\alpha = 1$ choose $Z^\varepsilon \geq 0$ with $\mathbb{E}(Z^\varepsilon Y) \geq \mathsf{AV@R}_1(Y) - \varepsilon$, where $\varepsilon > 0$. By the conditional $L^1 - L^\infty$-Hölder inequality

$$\begin{aligned}\mathsf{AV@R}_1(Y) - \varepsilon &\leq \mathbb{E}(Z_\varepsilon Y) \leq \mathbb{E}\left(\mathbb{E}\left[Z^\varepsilon|\mathcal{F}_t\right]\mathsf{AV@R}_1(Y|\mathcal{F}_t)\right) \\ &\leq \mathbb{E}\left(\mathbb{E}\left[Z^\varepsilon|\mathcal{F}_t\right]\right)\cdot \mathsf{AV@R}_1(Y) = \mathsf{AV@R}_1(Y),\end{aligned}$$

hence

$$\mathsf{AV@R}_1(Y) \leq \mathbb{E}\left(Z_t^\varepsilon \mathsf{AV@R}_{1-(1-\alpha)Z_t^\varepsilon}(Y|\mathcal{F}_t)\right) + \varepsilon.$$

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

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 $(1-\alpha)\mathsf{AV@R}_\alpha(Z_t Z') \leq \int_\alpha^1 \sigma(p)\,\mathrm{d}p$ for all $\alpha \in [0,1]$ with

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

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 $\sigma_\alpha(U)$ (cf. (3)) is nonnegative. Furthermore, $\sigma_\alpha(U) \geq 0$ and $\mathbb{E}(\sigma_\alpha(U)) = \int_0^1 \sigma_\alpha(p)\,\mathrm{d}p = 1$, and moreover

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

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

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

Using the setting

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

it follows that

$$\int_0^1 Z_\alpha Z'_\alpha \mu(\mathrm{d}\alpha) = \int_0^1 \sigma_\alpha(U)\,\mu(\mathrm{d}\alpha) = Z_t Z' \tag{15}$$

and

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

By construction of the random variables we have moreover the properties

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

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

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

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

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

(The latter inequality in fact holds with equality by the interchangeability principle ([23, Theorem 14.60]), the Bochner-integral in (16) and the essential supremum may be exchanged.) Hence

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

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

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

and

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

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, $(1-\alpha)\mathsf{AV@R}_\alpha(Z) \leq \int_\alpha^1 \sigma(p)\,\mathrm{d}p$. Define $Z_t := \mathbb{E}\left(Z|\mathcal{F}_t\right)$ and $Z' := \begin{cases}\frac{Z}{Z_t} & \text{if } Z_t > 0\\ 1 & \text{else}\end{cases}$ and, by Lemma 13, Z_t is feasible as well; that is $(1-\alpha)\mathsf{AV@R}_\alpha(Z_t) \leq \int_\alpha^1 \sigma(p)\,\mathrm{d}p$. With this choice it is obvious that

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

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

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

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

$$\mathcal{R}_Z\left(Y|\mathcal{F}_t\right)=\frac{\int_0^1 Z_\alpha \mathsf{AV@R}_{1-(1-\alpha)Z_\alpha}\left(Y|\mathcal{F}_t\right)\mu\left(\mathrm{d}\alpha\right)}{\int_0^1 Z_\alpha\mu\left(\mathrm{d}\alpha\right)}, \tag{17}$$

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

An exemplary 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 introductory example in Figure 1 illustrates that applying the same risk functional at the previous stage may reverse the preference, the conditional observations give a false order relation of the considered random variables Y and Y'. This counterexample (Figure 1) particularly demonstrates that

$$\mathsf{AV@R}_{\alpha\cdot\mathbb{1}}\left(Y'|\mathcal{F}_t\right)\le\mathsf{AV@R}_{\alpha\cdot\mathbb{1}}\left(Y|\mathcal{F}_t\right)\not\Longrightarrow\mathsf{AV@R}_\alpha\left(Y'\right)\le\mathsf{AV@R}_\alpha\left(Y\right), \tag{18}$$

if the same risk level α is employed conditionally.

7.1 Ranking based on Conditional Information

The decomposition theorem (Theorem 21) often allows a ranking of random variables Y and Y' based on incomplete information, i.e., based on $\mathcal{F}_t$.

To this end suppose that $\mathcal{R}(Y)=\mathbb{E}(YZ)$, where Z is the optimal feasible random variable satisfying $\mathcal{R}^*(Z)=0$ according the Fenchel–Moreau theorem (Theorem 6). Suppose further that

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

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

$$\mathcal{R}(Y')\ge\mathbb{E}\left(Z_t\cdot\mathcal{R}_{Z_t}\left(Y'|\mathcal{F}_t\right)\right)\ge\mathbb{E}\left(Z_t\cdot\mathcal{R}_{Z_t}\left(Y|\mathcal{F}_t\right)\right)=\mathcal{R}(Y). \tag{19}$$

This is a comparison, 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 $Z=\mathbb{1}$ is not adequate.

Figure 2 continues the example given in Figure 1, it displays the correction, which is provided by the decomposition theorem for the Average Value-at-Risk (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 $\mathsf{AV@R}_{1/3}$ already earlier, 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.

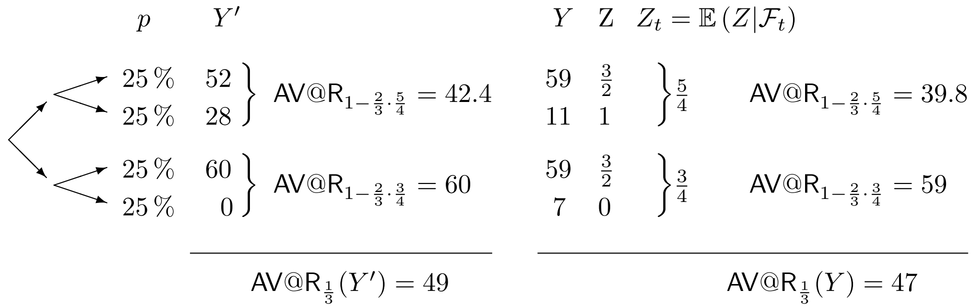

Figure 2: The same processes as in Figure 1. Applying the *adjusted* risk level to compare the random variables Y and Y' it holds that $\mathsf{AV@R}_{\frac{1}{3}}(Y') > \mathsf{AV@R}_{\frac{1}{3}}(Y)$, *because* $\mathsf{AV@R}_{1-(1-\alpha)Z_t}(Y'|\mathcal{F}_t) > \mathsf{AV@R}_{1-(1-\alpha)Z_t}(Y|\mathcal{F}_t)$, where the critical risk profile Z is chosen as outlined in Section 7.1.

7.2 Consistent Decisions

Adapting the risk profile conditionally is in practice often forgotten or neglected. To facilitate decision making it is thus of interest to know, if there are other coherent risk functionals, which allow a consistent decision already at an earlier stage without changing the risk profile.

The general answer to this question is negative. 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.

Definition 25. A risk functional $\mathcal{R}$ allows *(time) consistent decisions* if

$$\mathcal{R}(Y|\mathcal{F}_t) \le \mathcal{R}(Y'|\mathcal{F}_t) \implies \mathcal{R}(Y) \le \mathcal{R}(Y')$$

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

Remark 26 (The relation to time consistency). 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 in [33, Section 4.1].)

Time consistency, in contrast to (time) consistent decisions, is a desirable property of multistage or dynamic optimization problems themselves (cf. Shapiro [31, p. 437]): the 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 (cf. Carpentier at al. [7], as well as Shapiro [29]) and Markov decision processes (cf. Ruszczyński [24]).

Obviously,

(i) the expectation $\mathbb{E}(\cdot) = \mathsf{AV@R}_0(\cdot)$ allows time consistent decisions, as

$$\mathbb{E}(Y'|\mathcal{F}_t) \le \mathbb{E}(Y|\mathcal{F}_t) \Longrightarrow \mathbb{E}(Y') \le \mathbb{E}(Y);$$

(ii) moreover the max-risk functional $\operatorname{ess\,sup}(\cdot) = \mathsf{AV@R}_1(\cdot)$ is consistent as well, as

$$\operatorname{ess\,sup}(Y'|\mathcal{F}_t) \le \operatorname{ess\,sup}(Y|\mathcal{F}_t) \Longrightarrow \operatorname{ess\,sup}(Y') \le \operatorname{ess\,sup}(Y)$$

clearly holds true.

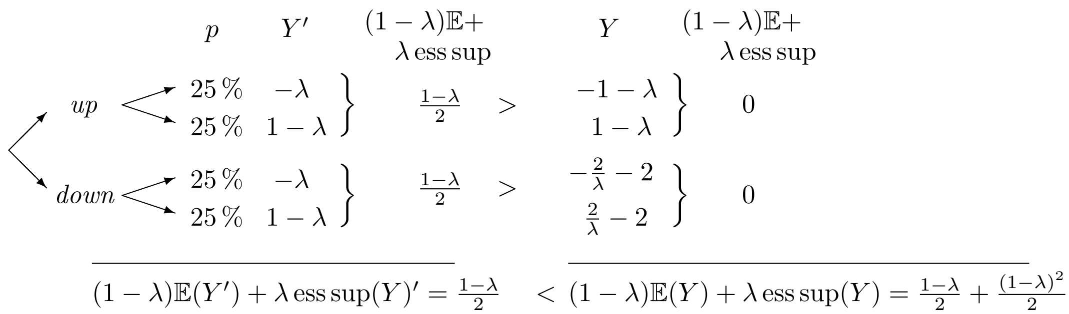

Figure 3: This example addresses the risk functional $\mathcal{R}(Y) := (1-\lambda)\cdot\mathbb{E}(Y) + \lambda\cdot\operatorname{ess\,sup}(Y)$ for an arbitrary $\lambda \in (0,1)$. This risk functional has higher conditional outcomes for Y' in comparison to Y ($\frac{1-\lambda}{2} > 0$), but taking all outcomes into account simultaneously the result reverses.

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

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

does *not* allow consistent decisions whenever $0 < \lambda < 1$, as Figure 3 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 risk functional

$$\mathcal{R}(Y) = \sup_{\mu\in\mathscr{M}} \int_0^1 \mathsf{AV@R}_\alpha(Y)\,\mu(\mathrm{d}\alpha), \tag{20}$$

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

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

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

and

(ii)

$$\sup_{\mu\in\mathscr{M}} \mu\left([\gamma, 1]\right) \to 0$$

whenever $\gamma \to 1$.

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

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

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

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

The random variable Y' is given by

$$Y' = \begin{cases} Y_1' & \text{w. pr. } q \\ Y_2' & \text{w. pr. } 1-q, \end{cases}$$

where

$$Y_1' = \begin{cases} -1 & \text{w. pr. } p, \\ 0 & \text{w. pr } 1-p \end{cases} \qquad \text{and} \qquad Y_2' = 0 \text{ w. pr } 1.$$

Notice that Y (Y', resp.) have the distribution

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

Calculating the Average Value-at-Risk we find that

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

and

$$\mathsf{AV@R}_\alpha(Y_1') = \begin{cases} 1 & \text{for } 1-p \le \alpha \\ \frac{p}{1-\alpha} & \text{for } \alpha \le 1-p, \end{cases}$$

moreover

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

and

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

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

$$\mathcal{R}(Y_1) > \mathcal{R}(Y_1') \text{ and } \mathcal{R}(Y_2) \ge \mathcal{R}(Y_2'),$$

but the unconditional random variables show the opposite inequality,

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

For this notice that

$$\begin{aligned} \mathsf{AV@R}_\alpha(Y_1) &\ge \mathsf{AV@R}_\alpha(Y_1') && \text{if } \alpha \le \varepsilon, \\ \mathsf{AV@R}_\alpha(Y_1) &\le \mathsf{AV@R}_\alpha(Y_1') && \text{if } \alpha \ge \varepsilon, \end{aligned}$$

and

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

Therefore

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

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

$$\mathcal{R}\left(Y_1\right) < \mathcal{R}\left(Y_1'\right). \tag{21}$$

On the other hand,

$$\int_0^1 \mathsf{AV@R}_\alpha\left(Y_2\right) - \mathsf{AV@R}_\alpha\left(Y_2'\right)\mu\left(\mathrm{d}\alpha\right) = 0,$$

such that

$$\mathcal{R}\left(Y_2\right) \le \mathcal{R}\left(Y_2'\right). \tag{22}$$

For the unconditional variables, however,

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

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

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

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.

8 Summary And Outlook

The present paper introduces a new, general concept of conditional risk functionals. The conditional risk functional respects the history of already available information, but in addition it reflects the initial risk functional without any modification. The conditional risk functional presented is consistent with the past *and* the future in a way, which is clarified by the central decomposition theorem, Theorem 21 (the central result of this paper). 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 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 further outline and illustrate the results achieved.

Stochastic Optimization

A key driver for the present investigations is stochastic optimization. The relation to stochastic optimization is very important, in particular for applications. Especially the nested decomposition (13) is important for multistage stochastic optimization. However, stochastic optimization is beyond the scope of the present paper and we leave it for a separate discussion.

9 Acknowledgment

We wish to thank two anonymous referees and the associate editor in their commitment to carefully reading the manuscript and providing suggestions to improve the paper.

References

[1] C. Acerbi. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26:1505–1518, 2002.

[2] C. Acerbi and P. Simonetti. Portfolio optimization with spectral measures of risk. *EconPapers*, 2002.

[3] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent Measures of Risk. *Mathematical Finance*, 9:203–228, 1999.

[4] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku. Coherent multiperiod risk adjusted values and Bellman's principle. *Annals of Operations Research*, 152:5–22, 2007.

[5] P. Artzner, F. Delbaen, and D. Heath. Thinking coherently. *Risk*, 10:68–71, November 1997.

[6] P. Billingsley. *Probability and Measure*. Wiley series in probability and mathematicals statistics. Jon Wiley & Sons, New York, 1979.

[7] P. Carpentier, J.-P. Chancelier, G. Cohen, M. de Lara, and P. Girardeau. Dynamic consistency for stochastic optimal control problems. *Annals of Operations Research*, 200(1):247–263, 2012.

[8] P. Cheridito and M. Kupper. Recursiveness of indifference prices and translation-invariant preferences. *Mathematics and Financial Economics*, 2(3):173–188, 2009.

[9] P. Cheridito and M. Kupper. Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance (IJTAF)*, 14(1):137–162, 2011.

[10] N. Dunford and J. T. Schwartz. *Linear Operators. Part I. General Theory*. Wiley-Interscience, New York, 1957.

[11] E. Jouini, W. Schachermayer, and N. Touzi. Law invariant risk measures have the Fatou property. In S. Kusuoka and A. Yamazaki, editors, *Advances in Mathematical Economics*, volume 9 of *Kusuoka, Shigeo and Yamazaki, Akira*, pages 49–71. Springer Japan, 2006.

[12] I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance*. Stochastic Modelling and Applied Probability. Springer, 1998.

[13] R. Kovacevic and G. Ch. Pflug. Time consistency and information monotonicity of multiperiod acceptability functionals. Number 8 in Radon Series on Computational and Applied Mathematics, pages 347–369. de Gruyter, 2009.

[14] S. Kusuoka. On law invariant coherent risk measures. *Advances in Mathematical Economics*, 3:83–95, 2001.

[15] N. Miller and A. Ruszczyński. Risk-averse two-stage stochastic linear programming: Modeling and decomposition. *Operations Research*, 59:125–132, 2011.

[16] A. Müller and D. Stoyan. *Comparison methods for stochastic models and risks.* Wiley series in probability and statistics. Wiley, Chichester, 2002.

[17] W. Ogryczak and A. Ruszczyński. Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization*, 13(1):60–78, 2002.

[18] G. Ch. Pflug. On distortion functionals. *Statistics and Risk Modeling (formerly: Statistics and Decisions)*, 24:45–60, 2006.

[19] G. Ch. Pflug and W. Römisch. *Modeling, Measuring and Managing Risk.* World Scientific, River Edge, NJ, 2007.

[20] A. Pichler. The natural Banach space for version independent risk measures. *Insurance: Mathematics and Economics*, 53(2):405–415, 2013.

[21] A. Pichler. Premiums and reserves, adjusted by distortions. *Scandinavian Actuarial Journal*, 2013.

[22] R. T. Rockafellar. *Conjugate Duality and Optimization*, volume 16. CBMS-NSF Regional Conference Series in Applied Mathematics. 16. Philadelphia, Pa.: SIAM, Society for Industrial and Applied Mathematics. VI, 74 p., 1974.

[23] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis.* Springer Verlag, 1997.

[24] A. Ruszczyński. Risk-averse dynamic programming for Markov decision processes. *Math. Program., Ser. B*, 125:235–261, 2010.

[25] A. Ruszczyński and D. Dentcheva. Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance*, 30:433–451, 2006.

[26] A. Ruszczyński and A. Shapiro. Optimization of convex risk functions. *Mathematics of operations research*, 31:433–452, 2006.

[27] W. Schachermayer and M. Kupper. Representation results for law invariant time consistent functions. *Mathematics and Financial Economics*, 2:189–210, 2009.

[28] M. Shaked and J. G. Shanthikumar. *Stochastic Order.* Springer Series in Statistics. Springer, 2007.

[29] A. Shapiro. On a time consistency concept in risk averse multistage stochastic programming. *Operations Research Letters*, 37(37):143–147, 2009.

[30] A. Shapiro. On Kusuoka representations of law invariant risk measures. *Mathematics of Operations Research*, November 2012.

[31] A. Shapiro. Time consistency of dynamic risk measures. *Operations Research Letters*, 40(6):436–439, 2012.

[32] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming.* MOS-SIAM Series on Optimization. MQS-SIAM Series on Optimization 9, 2009.

[33] T. Wang. A class of dynamic risk measures. *University of British Columbia*, 1999.

	p	Y	Z	$\mathbb{E}(Z\|\mathcal{F})$	$Z'_{0.7}\cdot Z_{0.7}$		$Z'_{0.4}\cdot Z_{0.4}$		$\mathcal{R}_Z(Y\|\mathcal{F})$
u	21%	1	$\frac{2}{7}$	1	0	1	$\frac{5}{7}$	1	$6.07=\frac{3}{5}\mathsf{AV@R}_{0.7}+\frac{2}{5}\mathsf{AV@R}_{0.4}$
	6%	7	2.66		3.33		1.67		
	3%	8	2.66		3.33		1.67		
m	28%	0	0	0.2	0	0	0	0.5	$3=0\cdot\mathsf{AV@R}_{1}+1\cdot\mathsf{AV@R}_{0.7}$
	12%	3	$\frac{2}{3}$		0		1.67		
d	21%	4	$\frac{38}{21}$	$\frac{31}{15}$	$\frac{40}{21}$	2.33	1.67	1.67	$5.94=\frac{21}{31}\mathsf{AV@R}_{0.3}+\frac{10}{31}\mathsf{AV@R}_{0}$
	9%	9	2.66		3.33		1.67		

Figure 4: Nested decomposition of $\mathcal{R}=\frac{3}{5}\mathsf{AV@R}_{0.7}(Y)+\frac{2}{5}\mathsf{AV@R}_{0.4}(Y)$. As the dual variable Z is optimal for (23) the conditional risk functional has the representation (17), $\mathcal{R}_Z(Y|\mathcal{F})=\frac{\int Z_\alpha \mathsf{AV@R}_{1-(1-\alpha)Z_\alpha}(Y|\mathcal{F})\mu(\mathrm{d}\alpha)}{\int Z_\alpha\mu(\mathrm{d}\alpha)}$, which is indicated on the very right.

[34] D. Williams. *Probability with Martingales.* Cambridge University Press, Cambridge, 1991.

[35] P. Wojtaszczyk. *Banach Spaces for Analysts.* Cambridge University Press, Cambridge, UK, 1991.

Appendix

Exemplary Decomposition of a Risk Measure

Figure 4 addresses the risk functional

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

the conditional probabilities are displayed in the tree's nodes. We demonstrate that the optimal conditional risk functionals are [8]

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

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 (cf. (2))

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

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

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

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

[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 $\mathcal{F}_t=\{u,m,d\}$.

The respective dual variables $Z_{0.7}\cdot Z'_{0.7}$ ($Z_{0.4}\cdot Z'_{0.4}$, respectively), as well as its conditional version $Z_{0.7}=\mathbb{E}\left(Z_{0.7}Z'_{0.7}|\mathcal{F}_t\right)$ ($\mathbb{E}\left(Z_{0.4}|\mathcal{F}_t\right)$, resp.) with respect to the filtration induced by the tree, are displayed in Figure 4.

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

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

and

$$\mathsf{AV@R}_{1-(1-0.4)\cdot Z_{0.4}}\left(Y|\mathcal{F}_t\right)=\begin{cases}\mathsf{AV@R}_{0.4}\left(Y|u\right) & =25/6\approx 4.17\\ \mathsf{AV@R}_{0.7}\left(Y|m\right) & =3\\ \mathsf{AV@R}_{0}\left(Y|d\right) & =5.5.\end{cases}$$

Notice now that

$$\mathbb{E}\left(Z_{0.7}\cdot\mathsf{AV@R}_{1-(1-0.7)\cdot Z_{0.7}}\left(Y|\mathcal{F}_t\right)\right)=6.5=\mathsf{AV@R}_{0.7}\left(Y\right)\text{ and}$$
$$\mathbb{E}\left(Z_{0.4}\cdot\mathsf{AV@R}_{1-(1-0.4)\cdot Z_{0.4}}\left(Y|\mathcal{F}_t\right)\right)=4.6=\mathsf{AV@R}_{0.4}\left(Y\right),$$

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 $Z=\int Z_\alpha\mu\left(\mathrm{d}\alpha\right)=\frac{3}{5}Z_{0.7}+\frac{2}{5}Z_{0.4}$ and $ZZ'=\int Z_\alpha Z'_\alpha\mu\left(\mathrm{d}\alpha\right)=\frac{3}{5}Z_{0.7}Z'_{0.7}+\frac{2}{5}Z_{0.4}Z'_{0.4}$ – built according to (15) and depicted in Figure 4. ZZ' is feasible and it holds that

$$\mathcal{R}\left(Y\right)=\tfrac{3}{5}\mathsf{AV@R}_{0.7}\left(Y\right)+\tfrac{2}{5}\mathsf{AV@R}_{0.4}\left(Y\right)=5.74=\mathbb{E}\,YZZ'. \tag{25}$$

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

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

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); they just have in common that 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

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

which is in accordance with (25).