International Journal of Theoretical and Applied Finance
Vol. 17, No. 1 (2014) 1450003 (33 pages)

DOI: 10.1142/S0219024914500034

ARE TIME CONSISTENT VALUATIONS INFORMATION MONOTONE?

RAIMUND M. KOVACEVIC*

Department of Statistics and Operations Research (ISOR)
University of Vienna, Universitaetsstrasse 5
A-1010 Wien-Vienna, Austria
raimund.kovacevic@univie.ac.at

GEORG CH PFLUG

ISOR and IIASA
Laxenburg, Austria
georg.pflug@univie.ac.at

Received 7 February 2012
Accepted 10 February 2013
Published 30 January 2014

Multi-period risk functionals assign a risk value to discrete-time stochastic processes. While convexity and monotonicity extend in straightforward manner from the single-period case, the role of information is more problematic in the multi-period situation. In this paper, we define multi-period functionals in such a way that the development of available information over time (expressed as a filtration) enters explicitly the definition of the functional. This allows to define and study the property of information monotonicity, i.e. monotonicity w.r.t. increasing filtrations. On the other hand, time consistency of valuations is a favorable property and it is well-known that this requirement essentially leads to compositions of conditional mappings. We demonstrate that generally spoken the intersection of time consistent and information monotone valuation functionals is rather sparse, although both classes alone are quite rich. In particular, the paper gives a necessary and sufficient condition for information monotonicity of additive compositions of positively homogeneous risk/acceptability mappings. Within the class of distortion functionals only compositions of expectation or essential infima are information monotone. Furthermore, we give a sufficient condition and examples for compositions of nonhomogeneous mappings exhibiting information monotonicity.

Keywords: Risk functional; acceptability functional; multi-period, conditional risk mapping; average value-at-risk; dual representation; information monotonicity; value of information.

AMS 2000 Classification: 91G99, 91B06

JEL Classification: G32

*This author was funded by FWF.

1. Introduction

Information is an important issue in economic literature for more than 40 years, see e.g., the important book by [34]. Information economics is based on the fundamental consent that information has a value or price. The processing, transfer and use of information, as well as the consequences of information costs on classical economic results, such as the "fundamental theorem of welfare economics" or the "law of the single price", was deeply analyzed over the last decades.

The idea that information has a value is also important in the field of stochastic optimization, where it expresses the amount, a rational decision maker would be willing to pay (in terms of his objective function) to reveal or partly reveal uncertain quantities before taking a decision (see [29, 33]). Clearly, this is closely related to the view of information economics (e.g., [15]).

Consequently, the effects of information should be taken into account by any quantification of risk or acceptability: if a functional measuring the acceptability of economic undertakings gives higher values in more informed situations, it will be called *information monotone*. Violation of this principle results in the strange situation that the same process is acceptable for some information pattern, but not acceptable, if more information is available at some point in time. Or to put it differently: if a functional is not information monotone, then hiding or neglecting information could make acceptable an *a priori* unacceptable undertaking.

On the other hand, many papers see *time consistency* as an important issue. This property basically means that valuations of an outcome process at different times do not contradict each other. Different definitions have been proposed in literature, but throughout this paper we follow [6, 7, 9, 20, 30, 35]. In this approach time consistency is practically equivalent to *recursivity*, which means that the functional is a composition of conditional mappings (for details see [1, 2, 20]).

Both properties, time consistency and information monotonicity are present in large classes of multi-period valuation functionals: recursive constructions lead to time consistency, while — as a typical example — the optimal values of information constrained stochastic optimization problems are information monotone. Although both classes are rather fundamental, we will demonstrate in this paper that it is hard to achieve information monotonicity for time consistent valuation functionals.

The main results of this paper are a necessary and sufficient condition for information monotonicity of compositions of positively homogeneous acceptability mappings and a sufficient condition for compositions of nonhomogeneous acceptability mappings. The method of proofs relies heavily on conjugate representations of concave functionals. Particular interesting examples of the homogeneous case will be compositions of distortion functionals, a class of functionals which is important in the context of valuation in insurance.

The paper is organized as follows: In Sec. 2, information monotonicity and time consistency is discussed informally, in particular we analyze the basic question of the paper and an important argument used later in the proofs. The following sections are written in a more rigorous style. Section 3 reviews the fundamental properties of multi-period functionals and conditional mappings and formally defines information monotonicity, valuation sequences, and time consistent. Additionally, some basic properties related to information monotonicity and time consistency are discussed. The main results of the paper are contained in Secs. 4 and 5: Section 4 analyzes information monotonicity of compositions of positively homogeneous acceptability mappings, with compositions of distortion functionals as an important special case. Section 5 deals with information monotonicity of compositions of nonhomogeneous mappings. Finally, the appendices give some additional technical background.

2. Information Monotonicity and Time Consistency

Suppose we want to valuate a stochastic process $\mathbf{Y} = (Y_1, \dots Y_T)$ — e.g., representing future prices of some commodities or securities, related cash flows, or a wealth process — with respect to its acceptability or risk. Acceptability and risk are the different sides of the same coin: risk is negative acceptability and we may use these two terms practically simultaneously. We are in particular interested in two aspects, information monotonicity and time consistency, of such valuations. Both notions will be rigorously defined and analyzed in Secs. 3–5, but for the moment we will start with an informal discussion.

2.1. *A basic relaxation argument*

A very general method for defining information monotone functionals is based on the observation that in real world decision problems the valuation of stochastic processes influencing the success of a business line or a project is undissolvably amalgamated with the decisions that can be taken in order to control the process and its wealth or risk. Therefore, it is natural to base the valuation on optimal decisions in the following way.

Suppose that in addition to observing the process $\mathbf{Y}$, it is possible to take actions $x = (x_0, \dots, x_{T-1})$, considered as (random) decisions with values in $\mathbb{R}^m$. Furthermore, let $H(x_0, Y_1, x_1, Y_2, \dots, x_{T-1}, Y_T)$ denote a profit function, which is concave w.r.t. the decisions x. In this framework, constraints on x can be included in the function H, defining $H(x_0, Y_1, x_1, Y_2, \dots, x_{T-1}, Y_T) = -\infty$ for values of x that are not feasible. For some filtration, i.e. an increasing sequence of σ-fields $\mathfrak{F} = (\mathcal{F}_0, \mathcal{F}_1, \dots, \mathcal{F}_T)$ representing the observable relevant information, define a multi-period valuation functional by

$$\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) = \sup_x \{\mathbb{E}[H(x_0, Y_1, x_1, Y_2, \dots, x_{T-1}, Y_T)]; x_t \lhd \mathcal{F}_t, t = 0, \dots, T-1\}. \tag{2.1}$$

Here, $x_t \lhd \mathcal{F}_t$ — the so called *nonanticipativity constraints* — denotes measurability of x_t with respect to the σ-field $\mathcal{F}_t$. This assumption refers to the fact that decisions at time t have to be based on information available at or before time t.

Any decision problem — e.g., optimal production, transportation, asset allocation or hedging — can be related to an optimization problem of the form (2.1). If H is monotone and concave in the Y_t's then the functional $\mathcal{A}$ is also monotone and concave, hence a multi-period valuation functional in the sense of Sec. 3.

2.1.1. *Relaxation of nonanticipativity constraints*

Replace now the filtration $\mathfrak{F}$ in (2.1) by a finer filtration $\mathfrak{F}' = (\mathcal{F}'_0, \mathcal{F}'_1, \ldots, \mathcal{F}'_T)$, i.e. use $\mathcal{F}_t \subseteq \mathcal{F}'_t$, but keep the underlying probability space the same. This leads to

$$\mathcal{A}(\mathbf{Y} \mid \mathfrak{F}') = \sup_x \{\mathbb{E}[H(x_0, Y_1, x_1, Y_2, \ldots, x_{T-1}, Y_T)]; x_t \lhd \mathcal{F}'_t, t = 0, \ldots, T-1\}, \tag{2.2}$$

which is just a relaxation of (2.1), i.e. a problem with a larger feasible set.

Let us illustrate this for the case that all σ-fields are finite: denote a set $\Gamma \subseteq \mathcal{F}_t$ as *atom* if no proper subset is contained in $\mathcal{F}$. The nonanticipativity constraints $x_t \lhd \mathcal{F}'_t$ can be reformulated by requiring that the decisions at stage t must be identical on the atoms, i.e. the constraints

$$x_t(\omega_1) = x_t(\omega_2) \quad (\omega_1, \omega_2 \in \Gamma) \tag{2.3}$$

must hold for all atoms Γ. In other words: if the decision maker has not enough information to distinguish two different scenarios, then the decision must be the same for both of them. For a more informative filtration $\mathfrak{F}'$ some σ-field $\mathcal{F}'_t$ will contain all atoms of the σ-field $\mathcal{F}_t$ but also at least two additional disjoint atoms Γ'_1, Γ'_2, such that $\Gamma'_1 \subset \Gamma$ and $\Gamma'_2 \subset \Gamma$ for some atom Γ of $\mathcal{F}_t$. When nonanticipativity constraints are formulated for the new atom, the constraints (2.3) related to Γ are replaced by the weaker constraints related to Γ'_1 and Γ'_2 separately, which leads to a relaxation.

Clearly, the optimal value of a relaxed problem is larger than or equal to the optimal value of the original problem and hence the optimal value functional $\mathcal{A}$ in (2.1), interpreted as a function of the filtration $\mathcal{F}$, is information monotone. The difference $\mathcal{A}(\mathbf{Y} \mid \mathfrak{F}') - \mathcal{A}(\mathbf{Y} \mid \mathfrak{F})$ is nonnegative and can be seen as the *value of additional information*. The idea of information monotonicity was introduced, e.g., in [15]. At that time it was called *information dominance* and was formulated without recourse to filtrations.

The most relaxed version of the original problem is the so-called clairvoyant-problem

$$\hat{\mathcal{A}}(\mathbf{Y} \mid \mathfrak{F}') = \sup_x \{\mathbb{E}[H(x_0, Y_1, x_1, Y_2, \ldots, x_{T-1}, Y_T)]; x_t \lhd \mathcal{F}_T, t = 0, \ldots, T-1\}. \tag{2.4}$$

Here, the decisions are only restricted by the information available at the end of the planning horizon. The difference $\hat{\mathcal{A}}(\mathbf{Y} \mid \mathfrak{F}') - \mathcal{A}(\mathbf{Y} \mid \mathfrak{F})$ measures the value of having the full information available at any point in time. Differences of this kind were introduced by [29] under the name of *expected value of perfect information* and have a long history in stochastic optimization (see e.g., [33]).

2.1.2. *Some caveats*

While the statement of the relaxation argument — that more information leads to an improved optimal value — is fundamental, some caveats should be kept in mind: it is important to observe the above condition, that both, the state space and the probability measure should not be changed when introducing the refined filtration and that $\mathcal{F}_t \subseteq \mathcal{F}'_t$.

Violating this principle may easily destroy the relaxation argument, because introducing new states (together with related values), and reassigning probabilities leads to a completely different optimization problem that is not a relaxation any more. Hence, the "refined" filtration may worsen the optimal value in this case. For an early paper that analyzes the effect of information without using the concept of finer filtrations see [21].

Note also that the relaxation argument only works if applied to an optimization problem (2.1) and its *global* optimum. Local optima or heuristic solutions easily can become worse, if a refined σ-field is introduced.

We finally remark that the relaxation argument is built on a basic property of optimization problems and cannot be extended to game theoretic situations, or more generally to problems related to saddle points. Relaxed constraints can lead to equilibria that are worse for both participants. If the basic probability space can be extended, examples similar to the Braess-paradoxon may be found (see [14]).

2.2. *The basic question*

In the main part of the paper, we will not refer to decision problems but will focus on the properties of generic multi-period acceptability functionals, that can be used independently of a special decision context. In particular, our main results will be formulated in terms of (concave, convex, and even translation equivariant) *acceptability functionals and mappings* (see [26, 31]), also known as *monetary utility functions* (e.g., [16]). It should be kept in mind that up to sign, such functionals are identical with convex risk measures ([12, 13]). Despite this restriction to generic valuation functionals we will see that the main arguments will build heavily on a generalization of the relaxation argument discussed above.

While information monotonicity is a property of multi-period valuation functionals, time consistency is a property of valuation sequences, i.e. valuations of the future development of a process at each point in time. Based on the literature discussed in the introduction, the technical definition is given in Definition 3.3. Basically, time consistent valuation sequences prefer a process X over a process Y (both defined on the same probability space with identical information structure (filtration)) at time $t-1$, if it prefers X to Y at time t and $X \geq Y$ a.s. at time $t-1$.

It is well-known that time consistency is (under quite general circumstances) equivalent to the presence of a recursive structure, i.e. the valuation at any time t can be reformulated as the composition of a valuation at time t and a valuation at time $t+1$ (see [2] or [20]). In the context of stochastic control, time consistent objective functions fulfill the Bellman-principle.

This paper analyzes the intersection of time consistent valuation sequences (compositions) and information monotone valuations. Clearly compositions of conditional expectation operators lead to time consistency and the resulting valuation over the whole planning horizon is information monotone. The same holds true, when conditional expectation is applied to a utility process, which means that expected utility also fulfills both criteria. Unfortunately, expectation is information monotone only in a trivial sense: If two processes are distinguishable only by their respective filtrations, their expectations are equal, hence the inequality in Definition 3.3 holds always with equality. Therefore, we extend the analysis to the broader class of acceptability functionals.

The following simple counterexamples show that it is easy to construct information monotone valuations that are not time consistent, and time consistent valuation sequences that are not information monotone. Both examples use the average value at risk, resp. its conditional version, which will be an important building block throughout the rest of this paper.

The *average value at risk* with parameter α is a special case of (2.1) for $T = 1$. with profit function $H(x, Y) = x - \frac{1}{\alpha}[Y - x]_-$, where $[a]_- = -\min(a, 0)$, which results in the definition

$$\mathbb{A}\mathrm{V@R}_\alpha(Y) = \sup_x \{\mathbb{E}[H(x, Y)] : x \in \mathbb{R}\}. \tag{2.5}$$

It is well-known (see [27]) that $\mathbb{A}\mathrm{V@R}_\alpha(Y) = \frac{1}{\alpha}\int_0^\alpha G_Y^{-1}(u)du$, where $G_Y(\cdot)$ is the distribution function of Y and $G_Y^{-1}(u) = \inf\{x : G_Y(x) > u\}$. The name average value-at-risk is due to Foellmer and Schied, but the same functional is also called *expected shortfall* or *conditional value-at-risk*. It is the concave minorant of the (nonconvex) value at risk, i.e. the largest concave functional dominated by the value at risk functional.

The first counterexample shows a sequence of perfectly information monotone valuation functionals that nevertheless is not time consistent.

Example 2.1. In view of (2.5) the valuation of the full tree, using the average value at risk $\mathbb{A}\mathrm{V@R}_{0.1}$ for valuing random payments, is information monotone by the relaxation argument. For a full proof see [26], Proposition 3.10. However, we will use Fig. 1 to show that valuation sequences consisting of (conditional) versions of $\mathbb{A}\mathrm{V@R}$ are not time consistent.

We will discuss conditional versions of the $\mathbb{A}\mathrm{V@R}$ in Sec. 3.2. For the moment it suffices to identify filtrations with the tree structure of probability trees. In this setup, we may calculate the conditional $\mathbb{A}\mathrm{V@R}$ of the random variables X and Y at each node of the respective tree by calculating the (unconditional) $\mathbb{A}\mathrm{V@R}$ of

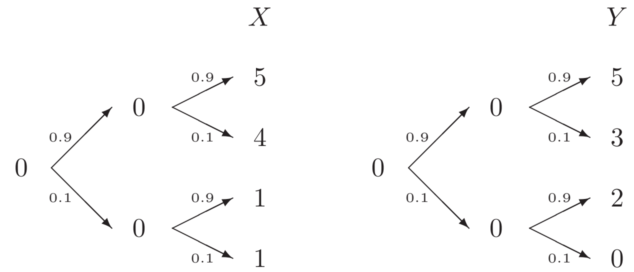

Fig. 1. The left process X is preferred over the right process Y at time 1, but the opposite preference holds at time 0.

these random variables, restricted to the respective subtree. For this the related conditional probabilities have to be used.

In this manner one sees that at time 1

$$\mathbb{AV@R}_{0.1}(X \mid \mathcal{F}_1) = (4; 1) \geq (3; 0) = \mathbb{AV@R}_{0.1}(Y \mid \mathcal{F}_1),$$

that is X is preferred over Y. Nevertheless at time 0, since

$$\mathbb{AV@R}_{0.1}(X) = 1 < 1.8 = \mathbb{AV@R}_{0.1}(Y),$$

Y is preferred over X, which contradicts time consistency.

Further examples can be fornulated easily. For instance, multistage models which minimize the final variance under a mean constraint (e.g., [23, 36]) are information-monotone but time inconsistent.

While we already know that time consistency can easily be achieved by compositions of conditional mappings, the second counterexample shows that this construction may immediately violate information monotonicity.

Example 2.2. (see [20]) Compositions of conditional mappings will be introduced formally in Sec. 3.3. In our simple tree setup without intermediate payoffs, and based on the $\mathbb{AV@R}_\alpha$, compositions are calculated as follows: for each leaf-node the value of the composition is just the value of the random variable related to this node. For any other node the value of the composition is given by the $\mathbb{AV@R}$ of the compositional values calculated for all the successor nodes of the node under consideration. Again, the calculation uses the related conditional probabilities.

This is a recursive construction and hence the resulting composition in the root node is time consistent. However, such compositions are not information monotone in general: consider Fig. 2: a final random payoff Y at time $t = 2$ is shown under two alternative filtrations $\mathfrak{F}$ and $\mathfrak{F}'$, represented by the two probability trees. Both processes are valuated under the same probability measure, defined by the probabilities assigned to the leaf nodes, but the information structure (filtration) is different: for the second case (the finer filtration $\mathcal{F}'$) a decision maker has more information available (i_1,i_2) about the final outcome at stage 1 than in the first case (the coarser

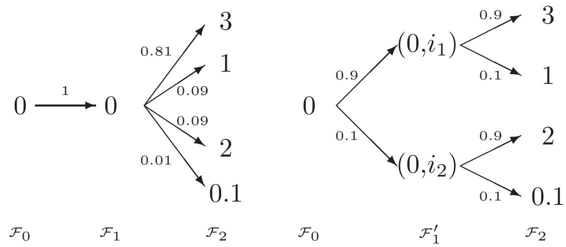

Fig. 2. Two trees with identical final payoff (assigned to the leaf nodes) and consistent conditional probabilities (assigned to the edges). The left tree however has a coarser (less informative) filtration than the right one.

filtration $\mathcal{F}$). This may result from an additional process, observable in the second case but not in the first.

If one calculates the acceptability given by the composition of $AV@R_{0.1}$, one gets

$$AV@R_{0.1}(AV@R_{0.1}(Y \mid \mathcal{F}_1')) = 0.1$$

and

$$AV@R_{0.1}(AV@R_{0.1}(Y \mid \mathcal{F}_1)) = 0.91.$$

This contradicts information monotonicity, because the coarser filtration leads to the larger valuation.

Within the discussed framework the main question of the paper can be restated: which compositions of acceptability mappings are also information monotone?

The classical relaxation arguments, is based purely on the relaxation of nonanticipativity constraints. The key arguments of Secs. 4 and 5 will generalize this by analyzing the concave Fenchel–Moreau conjugate representations of compositions and by observing the behavior of the related constraint sets with respect to changes in the underlying filtration. In the context of positively homogeneous acceptability functionals (Sec. 4) it will show that information monotonicity is valid if and only if a reformulation as an optimization problem is possible, such that a finer filtration leads to a relaxation. For nonhomogeneous functionals the picture is more complex, but nevertheless the relaxation argument will be part of the necessary conditions derived in Sec. 5.

3. Valuation Functionals

Throughout this paper we will use (possibly indexed) upper case characters to denote random variables, e.g., Y, Y_t, and bold face upper case characters to denote random vectors, or even stochastic processes, e.g., $\mathbf{Y}, \mathbf{Y}^{(i)}$. Let now $\mathcal{Y} = \times_{t=1}^{T} \mathcal{Y}_t$ be a normed linear space of stochastic processes $\mathbf{Y} = (Y_1, \ldots, Y_T)$, defined on some probability space $(\Omega, \mathcal{F}, P)$, where $\mathcal{Y}_t \subseteq L_1(\Omega, \mathcal{F}_t, P)$ are normed linear spaces with respective topological duals $\mathcal{Z}_t$. Typical examples are $\mathcal{Y}_t = L_p(\Omega, \mathcal{F}_t, P)$ and $\mathcal{Z}_t = L_q(\Omega, \mathcal{F}_t, P)$ with $\frac{1}{p} + \frac{1}{q} = 1$ and $p \in [1, \infty)$. In two later examples (see Examples 4.3 and 5.1), the dual pair is formed by the Zygmund spaces L_{exp} and $L \log^+ L$ (see [3, Theorem 6.5, p. 247]). Since we are interested in possibly unbounded processes, we do not require $\mathcal{Y} \subseteq \times_{t=1}^{T} L_\infty(\Omega, \mathcal{F}, P)$. The dual space of $\mathcal{Y}$ is given by $\mathcal{Z} = \times_{t=1}^{T} \mathcal{Z}_t$ and the related dual pairing on $\mathcal{Y} \times \mathcal{Z}$ is $\sum_{t=1}^{T} \mathbb{E}[Y_t \cdot Z_t]$.

Here, $\mathbb{E}[\cdot]$ denotes the expectation operator. We will use the short notation $\langle \mathbf{Y}, \mathbf{Z} \rangle := \sum_{t=1}^{T} \mathbb{E}[Y_t \cdot Z_t[$.

In order to model the evolution of information over time, we define $\mathbb{F}$ as the family of all filtrations of length $T + 1$ in $\mathcal{F}$, i.e. $\mathfrak{F} = (\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_T) \in \mathbb{F}$ iff $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$ and $\mathcal{F}_T \subseteq \mathcal{F}$. In addition, we assume $\mathcal{F}_0 = \{\emptyset, \Omega\}$. An observable process $\mathbf{Y}$ is adapted to the filtration $\mathfrak{F}$ (in symbol $\mathbf{Y} \lhd \mathfrak{F}$) iff Y_t is measurable with respect to $\mathcal{F}_t$ (in symbol $Y_t \lhd \mathcal{F}_t$) for $t = 1, \ldots, T$.

Typically, $\mathcal{F}_t$ can be larger than the σ-field generated by $(Y_1, \ldots, Y_t)$. This is evident if one considers *final* processes of the form $(0, \ldots, 0, Y_T)$. Here, the σ-field generated by $\mathbf{Y}$ is trivial up to time $T - 1$, but it might be possible to collect nontrivial information at times $0, \ldots, T - 1$. For this reason, we will conceptually separate the available information from the observed values, and define acceptability functionals $\mathcal{A}$ on pairs $(\mathbf{Y} \,|\, \mathfrak{F})$ consisting of an observable process $\mathbf{Y}$ and a filtration $\mathfrak{F}$.

3.1. *Multi-period valuation functionals and information monotonicity*

Consider functionals $\mathcal{A}$ that map the elements $\mathbf{Y} \in \mathcal{Y}$ and filtrations $\mathfrak{F}$ to the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\}$. We will assume that those functionals are proper (i.e. for every filtration $\mathfrak{F}$ it holds $\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) < +\infty$ for all $\mathbf{Y} \in \mathcal{Y}$ and $\mathcal{A}(\mathbf{Y}) > -\infty$ for some $\mathbf{Y} \in \mathcal{Y}$). If $\mathbf{Y}$ is not adapted to $\mathfrak{F}$, i.e. $\mathbf{Y} \ntriangleleft \mathfrak{F}$ we agree that $\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) = -\infty$. Such functionals will be called *probability functionals* in the following.

A probability functional $\mathcal{A} : \mathcal{Y} \times \mathbb{F} \to \overline{\mathbb{R}}$ with values $\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F})$ is called *multi-period valuation functional* if it satisfies the following properties for any filtration $\mathfrak{F} \in \mathbb{F}$:

(MA1) **Concavity.** The functional $\mathbf{Y} \mapsto \mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F})$ is concave.
(MA2) **Monotonicity.** If $\mathbf{X}, \mathbf{Y} \in \mathcal{Y}$ and $X_t \leq Y_t$ holds a.s. for $t = 1, \ldots, T$, then

$$\mathcal{A}(\mathbf{X} \,|\, \mathfrak{F}) \leq \mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}).$$

Remark 3.1. By definition we have $\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) > -\infty$ implies $\mathbf{Y} \lhd \mathfrak{F}$. This property may also be defined by looking at the conjugate $\mathcal{A}^+$, see Lemma C.1 in Appendix C.

In order to define the property of information monotonicity we introduce a partial order $\leq$ between filtrations $\mathfrak{F}' = (\mathcal{F}_0, \mathcal{F}'_1, \ldots, \mathcal{F}'_T)$ and $\mathfrak{F} = (\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_T)$:

$$\mathfrak{F}' \preceq \mathfrak{F} \quad \text{iff } \mathcal{F}'_t \subseteq \mathcal{F}_t \quad \text{for all } t. \tag{3.1}$$

We write $\mathfrak{F}' \prec \mathfrak{F}$ if at least one inclusion is strict in (3.1).

Within this setup and following [26] we define the property of *information monotonicity:*

Definition 3.1. A multi-period valuation functional $\mathcal{A} : \mathcal{Y} \times \mathbb{F} \to \overline{\mathbb{R}}$ is called *information monotone*, if $\mathfrak{F}' \preceq \mathfrak{F}$ implies

$$\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}') \leq \mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) \tag{3.2}$$

for all $\mathbf{Y} \in \mathcal{Y}$. It is called *strictly information monotone*, if $\mathfrak{F}' \prec \mathfrak{F}$ implies $\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}') < \mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F})$ for at least some $\mathbf{Y} \in \mathcal{Y}$.

Functionals which do not depend on the filtration (like $\mathcal{A}(\mathbf{Y}) = \sum_{t=1}^{T} \mathbb{E}[Y_t]$) are information monotone in a trivial sense. However, they fail to be strictly monotone.

If $\mathcal{A} = \mathcal{A}(\cdot \,|\, \mathfrak{F})$ is a multi-period valuation functional with nonempty domain, its *conjugate* $\mathcal{A}^+(\cdot \,|\, \mathfrak{F}) : \mathcal{Z} \to \overline{\mathbb{R}}$ obtained from the Legendre–Fenchel transform is given by

$$\mathcal{A}^+(\mathbf{Z} \,|\, \mathfrak{F}) := \inf\{\langle \mathbf{Y}, \mathbf{Z} \rangle - \mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) : \mathbf{Y} \in \mathcal{Y}\}. \tag{3.3}$$

Note that $\mathcal{A}^+(\cdot \,|\, \mathfrak{F})$ is again proper and concave. If $\mathcal{A}(\cdot \,|\, \mathfrak{F})$ is upper semicontinuous, the Rockafellar–Fenchel–Moreau theorem implies

$$\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) = \inf\{\langle \mathbf{Y}, \mathbf{Z} \rangle - \mathcal{A}^+(\mathbf{Z} \,|\, \mathfrak{F}) : \mathbf{Z} \in \mathcal{Z}\}. \tag{3.4}$$

Furthermore, if the infimum is attained in (3.4) i.e.

$$\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) = \langle \mathbf{Y}, \mathbf{Z} \rangle - \mathcal{A}^+(\mathbf{Z} \,|\, \mathfrak{F}),$$

then $\mathbf{Z}$ is a *supergradient* of $\mathcal{A}$ at $\mathbf{Y}$.

Denote by $\mathcal{S}(\mathfrak{F})$ the domain of $\mathcal{A}^+$, i.e. $\mathcal{S}(\mathfrak{F}) = \mathrm{dom}\mathcal{A}^+(\mathfrak{F}) = \{\mathbf{Z} \in \mathcal{Z} : \mathcal{A}^+(\mathbf{Z} \,|\, \mathfrak{F}) > -\infty\}$, which is closed and convex. We call $\mathcal{S} = \mathrm{dom}\mathcal{A}^+$ the *supergradient hull* of $\mathcal{A}$, since it is the convex hull of the supergradient set. It is well-known that property (MA2) implies that $\mathcal{S}(\mathfrak{F}) \subset \{\mathbf{Z} \in \mathcal{Z} : Z_t \geq 0\}$.

Remark 3.2. Notice that $\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F})$ is information monotone, if and only if its conjugate $\mathcal{A}^+(\mathbf{Y} \,|\, \mathfrak{F})$ is information antitone, i.e. fulfills (3.2) with reversed inequality sign. This can be seen easily from the relations (3.3) and (3.4).

The following lemma shows that the operation of sup-convolution has a special role for information monotone multi-period valuation functionals. Recall the definition of sup-convolution $\mathcal{A}_1 \# \mathcal{A}_2$ of functionals $\mathcal{A}_1$, $\mathcal{A}_2$,

$$(\mathcal{A}_1 \# \mathcal{A}_2)(\mathbf{Y} \,|\, \mathfrak{F}) = \sup\{\mathcal{A}_1(\mathbf{Y}^{(1)} \,|\, \mathfrak{F}) + \mathcal{A}_2(\mathbf{Y}^{(2)} \,|\, \mathfrak{F}) : \mathbf{Y}^{(1)} + \mathbf{Y}^{(2)} = \mathbf{Y}\}. \tag{3.5}$$

Lemma 3.1. *The family of all information monotone valuation functionals forms a convex cone, which is closed under sup-convolution.*

Proof. Notice that if both functionals $\mathcal{A}_1$ and $\mathcal{A}_2$ fulfill (MA1) and (MA2), the same is true for their sup-convolution. The convex cone property follows directly from the defining properties of acceptability functionals. Closedness under sup-convolution follows from the basic property

$$(\mathcal{A}_1 \# \mathcal{A}_2)^+(\mathbf{Z} \,|\, \mathfrak{F}) = \mathcal{A}_1^+(\mathbf{Z} \,|\, \mathfrak{F}) + \mathcal{A}_2^+(\mathbf{Z} \,|\, \mathfrak{F}), \tag{3.6}$$

since the sum is information antitone, if both summands are information antitone. □

3.2. *Conditional acceptability mappings and information monotonicity*

Conditional acceptability mappings will be used as the basic building blocks for constructing an important class of time consistent multi-period valuation functionals. In the following we use spaces $\mathcal{Y}$ and $\mathcal{Y}'$, where both $\mathcal{Y}$ and $\mathcal{Y}'$ are closed subspaces of some space $\boldsymbol{L}_1(\Omega, \mathcal{F}, P)$ with $\mathcal{F}' \subseteq \mathcal{F}$ and $\mathcal{Y} \subseteq \mathcal{Y}'$. All equations and inequalities involving random variables are understood in the sense of "valid almost surely" (a.s.), and **inf** denotes the infimum with respect to the partial order based on $\leq$ a.s.

Following the basic setup in [11, 17, 18, 26] *conditional acceptability mappings* with observable information $\mathcal{F}'$ are upper semicontinuous (in the sense of ([24], [25, Definition 5.3])) mappings $\mathcal{A}(\cdot \,|\, \mathcal{F}') : \mathcal{Y} \to \mathcal{Y}'$ such that the following conditions are satisfied for all $X, Y \in \mathcal{Y}$:

(CA1) **Concavity.** $\mathcal{A}(\lambda \cdot X + (1-\lambda) \cdot Y \,|\, \mathcal{F}') \geq \lambda \cdot \mathcal{A}(X \,|\, \mathcal{F}') + (1-\lambda) \cdot \mathcal{A}(Y \,|\, \mathcal{F}')$ holds for any $\lambda \in [0, 1]$.
(CA2) **Monotonicity.** $X \leq Y$ a.s. implies $\mathcal{A}(X \,|\, \mathcal{F}') \leq \mathcal{A}(Y \,|\, \mathcal{F}')$.
(CA3) **Predictable Translation Equivariance.** $\mathcal{A}(Y + Y' \,|\, \mathcal{F}') = \mathcal{A}(Y \,|\, \mathcal{F}') + Y'$ if $Y' \lhd \mathcal{F}'$.

Conditional acceptability mappings are closely related to conditional risk measures: functionals $\rho(Y \,|\, \mathcal{F}') := -\mathcal{A}(Y \,|\, \mathcal{F}')$ are conditional risk mappings. If in addition $\mathcal{A}(\lambda Y \,|\, \mathcal{F}') = \lambda \mathcal{A}(Y \,|\, \mathcal{F}')$ holds for all $\lambda \geq 0$ and $Y \in \mathcal{Y}$ then the risk mapping $\rho(\cdot \,|\, \mathcal{F}')$ is *coherent*.

In the following, we restrict ourself to conditional acceptability mappings with conjugate mapping

$$\mathcal{A}^+(Z \,|\, \mathcal{F}') = \mathbf{inf}\{\mathbb{E}[Y \cdot Z \,|\, \mathcal{F}'] - \mathcal{A}(Y \,|\, \mathcal{F}') : Y \in \mathcal{Y}\}, \tag{3.7}$$

representable by a dual representation

$$\mathcal{A}(Y \,|\, \mathcal{F}') = \mathbf{inf}\{\mathbb{E}[Y \cdot Z \,|\, \mathcal{F}'] - \mathcal{A}^+(Z \,|\, \mathcal{F}') : Z \in \mathcal{S}(\mathcal{F})\}, \tag{3.8}$$

where **inf** denotes the infimum w.r.t. the ordering based on $\leq$ a.s.

Example 3.1. The conditional version of $\mathbb{A}$V@R (see Eq. (2.5)) is given by

$$\mathbb{A}\mathrm{V@R}_\alpha(Y \,|\, \mathcal{F}') = \sup_x \{\mathbb{E}[H(x, Y) \,|\, \mathcal{F}'] : x \lhd \mathcal{F}'\}.$$

It has the conjugate representation

$$\mathbb{A}\mathrm{V@R}_\alpha(Y \,|\, \mathcal{F}') = \mathbf{inf}\{\mathbb{E}[Y \cdot Z] : \mathbb{E}[Z \,|\, \mathcal{F}'] = 1; 0 \leq Z \leq 1/\alpha\}.$$

Any minimizer of (3.8) is called a *supergradient* of $\mathcal{A}$ at Y. See [11, 17] for the theoretical background and for technical details of this construction. In particular, if $\mathcal{A}$ maps $\boldsymbol{L}_p$ to $\boldsymbol{L}_{p'}$ with $p \geq p'$, the dual space has to be restricted such that the supergradients of $\mathcal{A}$ lie in $\boldsymbol{L}_s$ with $s = \frac{p \cdot p'}{p - p'}$. Again we call the domain of the conjugate mapping, $\mathcal{S}(\mathcal{F}') = \{Z : \mathcal{A}^+(Z \,|\, \mathcal{F}') > -\infty\}$ the *supergradient hull*.

Definition 3.2. A conditional acceptability mapping $\mathcal{A}(\cdot \,|\, \cdot)$ is called *information monotone*, if $\mathcal{F}' \subseteq \mathcal{F}$ implies that

$$\mathbb{E}[\mathcal{A}(Y \,|\, \mathcal{F}')\mathbb{1}_B] \leq \mathbb{E}[\mathcal{A}(Y \,|\, \mathcal{F})\mathbb{1}_B],$$

for all $B \in \mathcal{F}'$ and $Y \in \mathcal{Y}$.

A necessary and sufficient condition for this property for positive homogeneous mappings is given by the following lemma.

Lemma 3.2. *A positively homogeneous conditional acceptability mapping is information monotone iff its supergradient hull satisfies*

$$\mathcal{F}' \subseteq \mathcal{F} \Rightarrow \mathcal{S}(\mathcal{F}') \supseteq \mathcal{S}(\mathcal{F}). \tag{3.9}$$

Proof. Let $\mathcal{F}' \subseteq \mathcal{F}$. Suppose $\mathcal{S}(\mathcal{F}') \supseteq \mathcal{S}(\mathcal{F})$. Then information monotonicity follows from

$$\begin{aligned}
\mathbb{E}[\mathcal{A}(Y \,|\, \mathcal{F}')\mathbb{1}_B] &= \mathbb{E}[\mathbf{inf}\{\mathbb{E}(YZ \,|\, \mathcal{F}') : Z \in \mathcal{S}(\mathcal{F}')\}\mathbb{1}_B] \\
&= \inf\{\mathbb{E}[Y \cdot Z\mathbb{1}_B \,|\, \mathcal{F}'] : Z \in \mathcal{S}(\mathcal{F}')\} \leq \inf\{\mathbb{E}[Y \cdot Z\mathbb{1}_B \,|\, \mathcal{F}'] : Z \in \mathcal{S}(\mathcal{F})\} \\
&= \mathbb{E}[\mathcal{A}(Y \,|\, \mathcal{F})\mathbb{1}_B].
\end{aligned}$$

Conversely, suppose that $\mathcal{S}(\mathcal{F}') \not\supseteq \mathcal{S}(\mathcal{F})$. Then there is a $Z \in \mathcal{S}(\mathcal{F}) \backslash \mathcal{S}(\mathcal{F}')$. By convex separation, one may find a $Y \in \mathcal{Y}$ such that $\mathbb{E}[Y \cdot Z] < \inf\{\mathbb{E}[Y \cdot Z'] : Z' \in \mathcal{S}(\mathcal{F}')\}$. Therefore, $\inf\{\mathbb{E}[Y \cdot Z'] : Z' \in \mathcal{S}(\mathcal{F})\} < \inf\{\mathbb{E}[Y \cdot Z'] : Z' \in \mathcal{S}(\mathcal{F}')\}$ and $\mathcal{A}$ cannot be information monotone. $\square$

The notion of sup-convolution extends in an obvious manner (e.g., [22]) to conditional acceptability mappings:

$$(\mathcal{A}_1 \# \mathcal{A}_2)(Y \,|\, \mathcal{F}) = \mathbf{sup}\{\mathcal{A}_1(Y_1 \,|\, \mathcal{F}) + \mathcal{A}_2(Y_2 \,|\, \mathcal{F}) : Y_1 + Y_2 = Y\},$$

where the supremum is understood in the almost sure sense.

Lemma 3.3. *The family of positively homogeneous, information monotone conditional acceptability mappings forms a convex cone, which is closed under the operation of sup-convolution.*

Proof. The first assertion is trivial, the second follows from the fact that

$$(\mathcal{A}_1 \# \mathcal{A}_2)^+(Z \,|\, \mathcal{F}) = \mathcal{A}_1^+(Z \,|\, \mathcal{F}) + \mathcal{A}_2^+(Z \,|\, \mathcal{F})$$

holds for the conjugate representation $(\mathcal{A}_1 \# \mathcal{A}_2)^+(Z \mid \mathcal{F})$ (see e.g., [22, 3.7.6]). In particular for positively homogeneous functionals, setting $\mathcal{S}_i = \mathrm{dom}\mathcal{A}_i^+$ we have

$$(\mathcal{A}_1 \# \mathcal{A}_2)^+(Z \mid \mathcal{F}) = \mathcal{J}_{\mathcal{S}_1} + \mathcal{J}_{\mathcal{S}_2} = \mathcal{J}_{\mathcal{S}_1 \cap \mathcal{S}_2},$$

where

$$\mathcal{J}_{\mathcal{S}}(Z) = \begin{cases} 0 & \text{if } Z \in \mathcal{S}, \\ -\infty & \text{otherwise.} \end{cases}$$

But if $\mathcal{F} \mapsto \mathcal{S}_i(\mathcal{F})$ is antitone (3.9) for $i = 1, 2$, then so is $\mathcal{F} \mapsto \mathcal{S}_1(\mathcal{F}) \cap \mathcal{S}_1(\mathcal{F})$. $\square$

Example 3.2. Supergradient hulls of the form

$$\mathcal{S}(\mathcal{F}) = \{Z : Z \geq 0, \mathbb{E}[Z \mid \mathcal{F}] = 1, Z \in A, a_i \leq \mathbb{E}[\phi_i(Z) \mid \mathcal{F}] \leq b_i \text{ for } i = 1, 2, \ldots\},$$

for some set A and measurable function ϕ_i, lead to information monotone positive homogeneous mappings. In particular, the conditional $\mathbb{AV@R}_\alpha$ (see Example 2.2) is information monotone, as well as any convex combination of $\mathbb{AV@R}_{\alpha_i}$'s.

Example 3.3. Conditional versions of regular distortion functionals (see Appendix A) can be defined as

$$\mathcal{A}_h(Y \mid \mathcal{F}) = \mathbf{inf}\left\{\mathbb{E}[Y \cdot Z \mid \mathcal{F}] : Z \geq 0, \mathbb{E}[Z \mid \mathcal{F}] = 1, \mathbb{E}[Z^k \mid \mathcal{F}] \leq \int_0^1 [h(u)]^k du, \text{ for } k = 2, 3, \ldots\right\} \tag{3.10}$$

for some nonnegative, monotonically decreasing function h. Regular distortions form a broad subclass of conditional positively homogeneous information monotone acceptability functionals.

3.3. *Time consistency of valuation functionals*

As above, let $\mathbf{Y} = (Y_1, \ldots, Y_T) \in \mathcal{Y}$ be a stochastic process with $Y_t \in \mathcal{Y}_t$ and $\mathfrak{F}$ a filtration, which is fixed for the moment. We denote by $\mathbf{Y}^{(t)} = (Y_t, \ldots, Y_T)$ and $\mathfrak{F}^{(t)} = (\mathcal{F}_t, \ldots, \mathcal{F}_T)$ the truncated sequences. Consider now sequences of functionals $\mathcal{A}^{(t)}(\cdot \mid \mathfrak{F}^{(t)})$ mapping $\times_{j=t+1}^T \mathcal{Y}_j$ to $\mathcal{Y}_t$, assigning a $\mathcal{F}_t$-measurable random variable to the truncated process of length $T - t$, which results in a stochastic process of valuations. If the mappings are jointly concave and monotone, we call them multi-period valuation mappings and the whole sequence a valuation sequence.

While information monotonicity is a property of multi-period valuation functionals, time consistency is a property of valuation sequences and can be defined as follows (see also the related definitions in [2, 7], for weaker concepts: [35]):

Definition 3.3. A valuation sequence $(\mathcal{A}^{(t)}(\cdot \mid \mathfrak{F}^{(t)}))_{t=0,\ldots,T-1}$ is called *time consistent*, if

$$\mathcal{A}^{(t)}(\mathbf{X}^{(t+1)} \mid \mathfrak{F}^{(t)}) \geq \mathcal{A}^{(t)}(Y^{(t+1)} \mid \mathfrak{F}^{(t)}) \quad \text{and} \quad X_t \geq Y_t$$

implies

$$\mathcal{A}^{(t-1)}(\mathbf{X}^{(t)} \,|\, \mathfrak{F}^{(t-1)}) \geq \mathcal{A}^{(t-1)}(Y^{(t)} \,|\, \mathfrak{F}^{(t-1)}),$$

for all $\mathbf{X}, \mathbf{Y} \in \times_{j=1}^{T} \mathcal{Y}_j$ and all $t = 1, \ldots, T-1$. A valuation sequence $(\mathcal{A}^{(t)}(\cdot \,|\, \mathfrak{F}^{(t)}))_{t=0,\ldots,T-1}$ is called *recursive*, if

$$\mathcal{A}^{(t-1)}(\mathbf{Y}^{(t)} \,|\, \mathfrak{F}^{(t-1)}) = \mathcal{A}^{(t-1)}(Y_t, \mathcal{A}^{(t)}(\mathbf{Y}^{(t+1)} \,|\, \mathfrak{F}^{(t)}), 0, \ldots, 0 \,|\, \mathfrak{F}^{(t-1)})$$

for all t and any $\mathbf{Y}^{(t)} \in \times_{j=t}^{T} \mathcal{Y}_t$.

Time consistency and recursivity are basically two sides of the same coin: we say that a mapping $\mathcal{A}^{(t)}(\cdot \,|\, \mathfrak{F}^{(t)})$ has the *weak projection property* if $Y_{t+1} \lhd \mathcal{F}_t$ implies that

$$\mathcal{A}^{(t)}(Y_{t+1}, 0, \ldots, 0 \,|\, \mathfrak{F}^{(t)}) = Y_{t+1}. \tag{3.11}$$

It has been shown (see [2, 20]) that a valuation sequence $(\mathcal{A}^{(t)}(\cdot \,|\, \mathfrak{F}^{(t)}))_{t=0,\ldots,T-1}$ with all components $\mathcal{A}^{(t)}(\cdot \,|\, \mathfrak{F}^{(t)})$ fulfilling the weak projection property is time consistent if and only if it is recursive.

A special case of recursive — and therefore time consistent — sequences are *additive acceptability compositions* (the idea goes back to ([32])).

Definition 3.4. Let $\mathbf{Y} = (Y_1, \ldots, Y_T)$ be a stochastic process with $Y_t \in \mathcal{Y}_t$ and let $\mathcal{A}_t(\cdot \,|\, \mathcal{F}_t)$, $t = 0, \ldots, T-1$ be a sequence of conditional acceptability mappings. We define the composition $\overline{\mathcal{A}}_t(\mathbf{Y}^{(t+1)} \,|\, \mathfrak{F}^{(t)})$ as

$$\overline{\mathcal{A}}^{(t)}(\mathbf{Y}^{(t+1)} \,|\, \mathfrak{F}^{(t)}) = \mathcal{A}_t \circ \mathcal{A}_{t+1} \circ \cdots \circ \mathcal{A}_{T-1}\left(\sum_{k=t+1}^{T} Y_k\right), \tag{3.12}$$

and for the complete composition (which is an unconditional functional) $\overline{\mathcal{A}}(\mathbf{Y} \,|\, \mathfrak{F}) := \overline{\mathcal{A}}^{(0)}(\mathbf{Y}^{(1)} \,|\, \mathfrak{F}^{(0)})$. If all of the mappings $\mathcal{A}_t(\cdot \,|\, \mathcal{F}_t)$ are conditional acceptability mappings we call the composition $\overline{\mathcal{A}}(\mathbf{Y} \,|\, \mathfrak{F})$ an *additive acceptability composition.*

Remark 3.3. By predictable translation equivariance (CA3), compositions can be equivalently defined in a stepwise manner

$$\overline{\mathcal{A}}^{(T-1)}(Y_T \,|\, \mathfrak{F}^{(T-1)}) = \mathcal{A}_{T-1}(Y_T \,|\, \mathcal{F}_{T-1}),$$

and

$$\overline{\mathcal{A}}^{(t-1)}(\mathbf{Y}^{(t)} \,|\, \mathfrak{F}^{(t-1)}) = \mathcal{A}_{t-1}(Y_t + \overline{\mathcal{A}}^{(t)}(\mathbf{Y}^{(t+1)} \,|\, \mathfrak{F}^{(t)}) \,|\, \mathcal{F}_{t-1})$$

for $t < T-1$.

Remark 3.4. Notice that meaningful values will only occur if all random variables Y_t belong to the respective domains and all mappings $\mathcal{A}_t$ map into the correct spaces.

An alternative way of ensuring meaningful results is to start with the dual representation (3.8) which, for $\mathbf{Y} \notin \mathcal{Y}$ may yield as possible value $-\infty$ with positive probability. In this case we set the function value equal to $-\infty$ almost surely, and require $\mathcal{A}(-\infty \,|\, \mathcal{F}) = -\infty$. In this way the composition leads to finite results only if functionals with appropriate domains and ranges were chosen.

The following representation (see [17] and [18], Theorem 5.2.6, Corollary 5.2.7) holds for the additive composition of acceptability mappings and will be important throughout the rest of this paper. Its proof is repeated in Appendix B.

Proposition 3.1. *Let $(\mathcal{A}_t)_{t\in\{1,\dots,T-1\}}$ be a collection of monotone, concave probability mappings with $T \geq 2$. Then for $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}$ the related acceptability composition $\overline{\mathcal{A}}(\mathbf{Y} \,|\, \mathfrak{F})$ can be represented in the following way*:

$$
\begin{aligned}
\overline{\mathcal{A}}(\mathbf{Y} \,|\, \mathfrak{F}) &= \inf_{Z_1, Z_2, \dots, Z_T} \Bigg\{ \sum_{t=1}^{T} \mathbb{E}\left[Y_t \cdot \prod_{i=1}^{t} Z_i \right] - \mathcal{A}_0^+(Z_1) \\
&\qquad - \sum_{t=1}^{T-1} \mathbb{E}\left[\mathcal{A}_t^+(Z_{t+1}) \prod_{i=1}^{t} Z_i \right] : Z_t \in \mathcal{S}_{t-1}(\mathcal{F}_{t-1}) \Bigg\} \\
&= \inf_{M_1, M_2, \dots, M_T} \Bigg\{ \langle \mathbf{Y}, \mathbf{M} \rangle - \mathcal{A}_0^+(M_1) \\
&\qquad - \sum_{t=1}^{T-1} \mathbb{E}\left[\mathcal{A}_t^+\left(\frac{M_{t+1}}{M_t} \right) M_t \right] : M_0 \equiv 1, \frac{M_t}{M_{t-1}} \in \mathcal{S}_{t-1}(\mathcal{F}_{t-1}) \Bigg\},
\end{aligned} \tag{3.13}
$$

where $\mathbf{M} = (M_1, \dots, M_T)$ with $M_t = \prod_{i=1}^{t} Z_i$.

Remark 3.5. Notice that (M_t) is a nonnegative martingale, since $\mathbb{E}[Z_t \,|\, \mathcal{F}_{t-1}] = 1$. The functional

$$
\mathcal{A}_0^+(M_1) + \sum_{t=1}^{T-1} \mathbb{E}\left[\mathcal{A}_t^+\left(\frac{M_{t+1}}{M_t} \middle| \mathcal{F}_t \right) M_t \right] \tag{3.14}
$$

appearing in (3.13) is not necessarily identical to the conjugate functional of $\overline{\mathcal{A}}$: The latter is the concave majorant of (3.14), i.e. the smallest concave functional that dominates (3.14).

Using the martingale property, it is possible to rewrite the representation in the following way.

$$
\begin{aligned}
\overline{\mathcal{A}}(\mathbf{Y} \,|\, \mathfrak{F}) = \inf_{M_T} \Bigg\{ & \langle \mathbf{Y}, \mathbb{E}[M_T \,|\, \mathcal{F}_t] \rangle - \mathcal{A}_0^+(\mathbb{E}[M_T \,|\, \mathcal{F}_1]) \\
& - \sum_{t=1}^{T-1} \mathbb{E}\left[\mathcal{A}_t^+\left(\frac{\mathbb{E}[M_T \,|\, \mathcal{F}_{t+1}]}{\mathbb{E}[M_T \,|\, \mathcal{F}_t]} \right) \mathbb{E}[M_T \,|\, \mathcal{F}_t] \right] : M_T \in \mathcal{M}_T, \\
& \frac{\mathbb{E}[M_T \,|\, \mathcal{F}_t]}{\mathbb{E}[M_T \,|\, \mathcal{F}_{t-1}]} \in S_{t-1}(\mathcal{F}_{t-1}) \Bigg\},
\end{aligned} \tag{3.15}
$$

where $\mathcal{M}_T = \{M_T \mid M_T \geq 0, \mathbb{E}[M_T] = 1\}$ is the family of all probability densities w.r.t. P.

Remark 3.6. In (3.15) we define $0/0 = 0$, which does not harm, since by nonnegativity of M_T

$$\{\mathbb{E}[M_T \mid \mathcal{F}_t] = 0\} \subseteq \{\mathbb{E}[M_T \mid \mathcal{F}_{t+1}] = 0\}.$$

Hence, $\frac{M_{t+1}}{M_t} = \frac{\mathbb{E}[M_T \mid \mathcal{F}_{t+1}]}{\mathbb{E}[M_T \mid \mathcal{F}_t]}$ is well defined.

4. Information Monotonicity of Compositions of Positively Homogeneous Mappings

We start our analysis of compositions by studying compositions of positively homogeneous information monotone acceptability mappings. Notice that positive homogeneous acceptability mappings are fully characterized by the mappings $\mathcal{F}_t \mapsto \mathcal{S}_t(\mathcal{F}_t)$. From Proposition 3.1 it follows that such compositions can be represented as

$$\overline{\mathcal{A}}(\mathbf{Y} \mid \mathfrak{F}) = \inf_{M_1, M_2, \ldots, M_T} \left\{ \langle \mathbf{Y}, \mathbf{M} \rangle : \frac{M_t}{M_{t-1}} = Z_t \in \mathcal{S}_{t-1}(\mathcal{F}_{t-1}); M_0 = 1 \right\}. \quad (4.1)$$

For a sequence $(\mathcal{A}_t)$ of conditional mappings with supergradient hulls $\mathcal{S}_t$ and a filtration $\mathfrak{F}$ we define the following related sequence $(\mathcal{M}_t)_{t=0,\ldots,T}$ of sets:

$$\mathcal{M}_0 = \mathcal{S}_0(\mathcal{F}_0) = \{1\}, \quad (4.2)$$

$$\mathcal{M}_t = \mathcal{M}_{t-1} \cdot \mathcal{S}_t(\mathcal{F}_t), \quad (4.3)$$

for $1 \leq t \leq T-1$, where for two sets of random variables V, W the Minkowski-type (pointwise) product $V \cdot W$ is defined as

$$V \cdot W = \{X : X = Z \cdot Z' : Z \in V, Z' \in W\}. \quad (4.4)$$

Remark 4.1. Throughout this paper we will use $\times$ for the cartesian product of sets (and $\times_i$ for its repetition) and $\cdot$ for the ordinary multiplication between reals or random variables, as well as for the Minkowski-type product. For random vectors $\cdot$ refers to the scalar product. Repetition of the ordinary or Minkowski-type multiplication is denoted by $\prod_i$.

Note that the Z_t's from representation (3.13) are elements of $\mathcal{S}_{t-1}$, while the $M_t = \prod_{i=1}^t Z_i$ are elements of $\mathcal{M}_{t-1}$. We may rewrite (4.1) in simpler form

$$\overline{\mathcal{A}}(\mathbf{Y} \mid \mathfrak{F}) = \inf_{\mathbf{M}} \{\langle \mathbf{Y}, \mathbf{M} \rangle : M_t \in \mathcal{M}_{t-1}\}. \quad (4.5)$$

Within our framework, information monotonicity can be characterized in the following way:

Proposition 4.1. *Let $\overline{\mathcal{A}}(\mathbf{Y} \mid \mathfrak{F})$ be a composition of information monotone, positively homogeneous conditional acceptability mappings $\mathcal{A}_t$ with supergradient hulls*

$\mathcal{S}_t(\cdot)$. *The composition is information monotone if and only if*

$$\mathcal{S}_{t-1}(\mathcal{F}') \cdot \mathcal{S}_t(\mathcal{F}) \subseteq \mathcal{S}_t(\mathcal{F}') \tag{4.6}$$

for every $t, 1 \leq t \leq T-1$ *and every pair of* σ*-fields* $\mathcal{F}'$, $\mathcal{F}$ *such that* $\mathcal{F}' \subseteq \mathcal{F}$.

Strict information monotonicity is achieved if and only if $\subset$ *holds instead of* $\subseteq$ *for some* t *in* (4.6).

Proof. We show first that condition (4.6) is necessary. Consider a filtration $\mathfrak{F} = (\mathcal{F}_t)_{t=0,\ldots,T}$. Then fix a point in time t_0, and define a modified filtration $\mathfrak{F}'$, where $\mathcal{F}_{t_0}$ is replaced by $\mathcal{F}'_{t_0} = \mathcal{F}_{t_0-1}$, and all other σ-fields $\mathcal{F}_t$, $t \neq t_0$ remain unchanged.

A representation of the mapping with modified σ-fields $\mathfrak{F}'$ is given by

$$\overline{\mathcal{A}}(\mathbf{Y} \,|\, \mathfrak{F}') = \inf_{\mathbf{M}} \{\langle \mathbf{Y}, \mathbf{M}' \rangle : M'_t \in \mathcal{M}'_{t-1}\}, \tag{4.7}$$

where $\mathcal{M}'_t = \mathcal{M}_t$ for $t < t_0$, and

$$\mathcal{M}'_t = \mathcal{M}_{t_0-2} \cdot \mathcal{S}_{t_0}(\mathcal{F}_{t_{0-1}}) \cdot \prod_{k=t_0+1}^{t} \mathcal{S}_k(\mathcal{F}_k), \tag{4.8}$$

for $t \geq t_0$.

Notice that for closed convex sets $\mathcal{M}_1, \mathcal{M}_2$ the relation

$$\inf_{\mathbf{M}} \{\langle \mathbf{Y}, \mathbf{M} \rangle : \mathbf{M} \in \mathcal{M}_1\} \leq \inf_{\mathbf{M}} \{\langle \mathbf{Y}, \mathbf{M} \rangle : \mathbf{M} \in \mathcal{M}_2\}$$

holds if and only if $\mathcal{M}_1 \supseteq \mathcal{M}_2$. The if part is obvious. For the only if part assume that $\mathcal{M}_1 \not\supseteq \mathcal{M}_2$. Then there is an element $\mathbf{M}^* \in \mathcal{M}_2$ with $\mathbf{M}^* \notin \mathcal{M}_1$: $\mathcal{M}_2$ can be separated from $\mathcal{M}_1$ by a continuous linear functional, i.e. there exists a $\mathbf{Y}$ such that $\langle \mathbf{Y}, \mathbf{M}^* \rangle < \inf_{\mathbf{M}} \{\langle \mathbf{Y}, \mathbf{M} \rangle : \mathbf{M} \in \mathcal{M}_1\}$, hence $\inf_{\mathbf{M}} \{\langle \mathbf{Y}, \mathbf{M} \rangle : \mathbf{M} \in \mathcal{M}_1\} > \inf_{\mathbf{M}} \{\langle \mathbf{Y}, \mathbf{M} \rangle : \mathbf{M} \in \mathcal{M}_2\}$, a contradiction.

Using this argument, and comparing (4.5) and (4.7), we see that information monotonicity

$$\overline{\mathcal{A}}(\mathbf{Y} \,|\, \mathfrak{F}) = \inf_{\mathbf{M}} \{\langle \mathbf{Y}, \mathbf{M} \rangle : M_t \in \mathcal{M}_{t-1}\} \geq \inf_{\mathbf{M}} \{\langle \mathbf{Y}, \mathbf{M} \rangle : M_t \in \mathcal{M}'_{t-1}\} = \overline{\mathcal{A}}(\mathbf{Y} \,|\, \mathfrak{F}') \tag{4.9}$$

implies that

$$\begin{aligned} &\mathcal{M}_{t_0-2} \cdot \mathcal{S}_{t_0-1}(\mathcal{F}_{t_0-1}) \cdot \mathcal{S}_{t_0}(\mathcal{F}_{t_0}) \cdot \prod_{k=t_0+1}^{t} \mathcal{S}_k\left(\mathcal{F}_k\right) \\ &\quad \subseteq \mathcal{M}_{t_0-2} \cdot \mathcal{S}_{t_0}(\mathcal{F}_{t_0-1}) \cdot \prod_{k=t_0+1}^{t} \mathcal{S}_k(\mathcal{F}_k)\} \end{aligned} \tag{4.10}$$

holds for all $t \geq t_0$. Furthermore, (4.10) is equivalent to

$$\mathcal{S}_{t_0-1}(\mathcal{F}_{t_0-1}) \cdot \mathcal{S}_{t_0}(\mathcal{F}_{t_0}) \subseteq \mathcal{S}_{t_0}(\mathcal{F}_{t_0-1}) \tag{4.11}$$

as we show now: (4.10) follows trivially from (4.11). For the other direction choose $\{1\} \in \mathcal{M}_{t_0-2}$, $Z_1 \in \mathcal{S}_{t_0-1}(\mathcal{F}_{t_0-1})$, $Z_2 \in \mathcal{S}_{t_0}(\mathcal{F}_{t_0})$, $Z_3 \in \prod_{k=t_0+1}^{t} \mathcal{S}_k(\mathcal{F}_k)$.

Inequality (4.10) then implies that there are $Z_0' \in \mathcal{M}_{t_0-2}$, $Z_1' \in \mathcal{S}_{t_0}(\mathcal{F}_{t_0-1})$ and $Z_3' \in \prod_{k=t_0+1}^{t} \mathcal{S}_k(\mathcal{F}_k)$ such that

$$Z_1 \cdot Z_2 \cdot Z_3 = Z_0' \cdot Z_1' \cdot Z_3'.$$

Taking the conditional expectation with respect to $\mathcal{F}_{t_0}$ we arrive at

$$Z_1 \cdot Z_2 = Z_0' \cdot Z_1',$$

since all Z_i are conditional densities. Taking conditional expectation with respect to $\mathcal{F}_{t_0-2}$ we get $Z_0' = 1$, hence

$$Z_1 \cdot Z_2 = Z_1',$$

and we see that property (4.11) is fulfilled.

Regarding sufficiency, it has been shown so far that if condition (4.11) is valid for all t, then information monotonicity is not disturbed when $\mathcal{F}_{t_0}$ is replaced by $\mathcal{F}_{t_0-1}$. If in addition all the mappings $\mathcal{A}_t$ are information monotone in the sense of Definition 3.2, then all the factors in the products, and hence the products themselves are information antitone. Hence, if $\mathcal{F}_{t_0}$ is replaced by $\mathcal{F}'$ with $\mathcal{F}_{t_0-1} \subset \mathcal{F}' \subseteq \mathcal{F}_{t_0}$ then

$$M_{t+1} \in \mathcal{M}_{t_0-1} \cdot \mathcal{S}_{t_0}(\mathcal{F}') \cdot \prod_{k=t_0+1}^{t} \mathcal{S}_k(\mathcal{F}_k).$$

This results in

$$\overline{\mathcal{A}}^{(0)}(\mathbf{Y} \mid \mathfrak{F}) = \inf_{\mathbf{M}}\{\langle \mathbf{Y}, \mathbf{M}\rangle : M_t \in \mathcal{M}_{t-1}\} \geq \inf_{\mathbf{M}}\{\langle \mathbf{Y}, \mathbf{M}\rangle : M_t \in \mathcal{M}_{t-1}'\} = \overline{\mathcal{A}}^{(0)}(\mathbf{Y} \mid \mathfrak{F}),$$

for any $\mathbf{Y}$, which completes the proof that condition (4.11) is sufficient for information monotonicity. □

Two simple but fundamental cases of compositions show how the above results can be used.

Example 4.1. An additive composition of expectations

$$\mathbb{E}[Y_1 + \cdots + Y_T] = \mathbb{E}[Y_1 + \mathbb{E}[Y_2 + \cdots \mathbb{E}[\cdots + \mathbb{E}[Y_T \mid \mathcal{F}_{T-1}] \cdots \mid \mathcal{F}_t] \mid \mathcal{F}_1] \mid \mathcal{F}_0]$$

is information monotone in a trivial way. Although this is evident, one may also evoke Proposition 4.1, since the conditional expectation has the conjugate representation

$$\mathbb{E}[Y \mid \mathcal{F}_t] = \inf\{\mathbb{E}[Y \cdot Z \mid \mathcal{F}_t] : Z \equiv 1\},$$

i.e. $\mathcal{S}_t(\mathcal{F}_t) = \{1\}$. Therefore, condition (4.11) is fulfilled.

Example 4.2. The composition of (identical) positively homogeneous acceptability mappings on L^∞ with conjugate representation

$$\mathcal{A}_t(Y \mid \mathcal{F}') = \mathbf{inf}\{\mathbb{E}[Y \cdot Z \mid \mathcal{F}'] : 0 \leq Z, \mathbb{E}[Z \mid \mathcal{F}'] = 1\},$$

which is the conditional version of the essential infimum, is information monotone: let $\mathcal{F}' \subseteq \mathcal{F}$. Then we have

$$S_{t-1}(\mathcal{F}') = \{Z : 0 \leq Z, \mathbb{E}[Z \,|\, \mathcal{F}'] = 1\},$$
$$S_t(\mathcal{F}) = \{Z : 0 \leq Z, \mathbb{E}[Z \,|\, \mathcal{F}] = 1\}.$$

Consider $Z_1 \in S_{t-1}(\mathcal{F}'), Z_2 \in S_t(\mathcal{F})$. Then $Z_1 \cdot Z_2 \geq 0$, and $\mathbb{E}[Z_1 \cdot Z_2 \,|\, \mathcal{F}'] = \mathbb{E}[Z_1 \mathbb{E}(Z_2 \,|\, \mathcal{F}) \,|\, \mathcal{F}'] = 1$, i.e. $Z_1 \cdot Z_2 \in \mathcal{S}_t(\mathcal{F}')$.

Note that in both cases, Examples 4.1 and 4.2, information monotonicity is not strict. The expectation and compositions of essential infima are extreme cases of a broad spectrum of compositions with conjugate representation

$$\mathcal{A}_t[Y \,|\, \mathcal{F}'] = \mathbf{inf}\{\mathbb{E}\mathbb{E}[Y \cdot Z \,|\, \mathcal{F}'] : 0 \leq a_t \leq Z \leq b_t, \mathbb{E}[Z \,|\, \mathcal{F}'] = 1\}, \tag{4.12}$$

where $a_t \in \mathbb{R}, b_t \in \mathbb{R} \cup \{+\infty\}$. These mappings can also be represented as

$$\mathcal{A}_t[Y \,|\, \mathcal{F}'] = a_t \mathbb{E}[Y \,|\, \mathcal{F}' + (1 - a_t)\mathbb{AV@R}_{(\alpha t)}(Y \,|\, \mathcal{F}')],$$

where $\alpha_t = \frac{1-a_t}{b_t - a_t}$. For the expectation we have $a_t = b_t = 1$, while the essential infimum is given by $a_t = 0$, $b_t = \infty$. Clearly, with $a_t = 0$ and $b_t = \frac{1}{\alpha}$ this class also contains any $\mathbb{AV@R}_\alpha$.

One could ask the question, whether it is possible to compose any member of this family. The answer is no, as the following corollary shows.

Corollary 4.1. *Consider a filtration $\mathfrak{F}$, and a sequence of conditional acceptability mappings with conjugate representation* (4.12)*. An additive composition based on those mappings is information monotone if the following conditions are fulfilled*:

(i) *Any occurence of a conditional essential infimum can be preceded by any mapping $\mathcal{A}_{t-1}$ that fulfills* (4.12)*.*
(ii) *Any other mapping $\mathcal{A}_t$ than the essential infimum must be preceded by a conditional expectation.*

Proof. The individual mappings are information monotone, hence Proposition 4.1 can be applied. We have

$$\mathcal{S}_{t-1}(\mathcal{F}_{t-1}) = \{Z : 0 \leq a_{t-1} \leq Z_t \leq b_{t-1}, \mathbb{E}[Z_t \,|\, \mathcal{F}_{t-1}] = 1\}, \tag{4.13}$$

$$\mathcal{S}_t(\mathcal{F}_t) = \{Z : 0 \leq a_t \leq Z_{t+1} \leq b_t, \mathbb{E}[Z_{t+1} \,|\, \mathcal{F}_t] = 1\}, \tag{4.14}$$

$$\mathcal{S}_t(\mathcal{F}_{t-1}) = \{M : 0 \leq a_t \leq M \leq b_t, \mathbb{E}[M \,|\, \mathcal{F}_{t-1}] = 1\}. \tag{4.15}$$

If the composition is information monotone then (4.6) is valid. Hence, the following two equations,

$$a_{t-1} \cdot a_t \geq a_t \tag{4.16}$$

and

$$b_{t-1} \cdot b_t \leq b_t \tag{4.17}$$

must hold. In addition $\mathbb{E}[Z_{t+1} \,|\, \mathcal{F}_t] = 1$ and $\mathbb{E}[Z_t \,|\, \mathcal{F}_{t-1}] = 1$ imply

$$0 \leq a_t \leq 1 \leq b_t \tag{4.18}$$

and

$$0 \leq a_{t-1} \leq 1 \leq b_{t-1}. \tag{4.19}$$

If we consider mappings with finite $b_t < \infty$, the four Eqs. (4.16)–(4.19) reduce to the alternative

$$(0 \leq a_{t-1} < 1 \wedge b_{t-1} = 1 \wedge a_t = 0 \wedge b_t \geq 1) \tag{4.20}$$

or

$$(a_{t-1} = 1 \wedge b_{t-1} = 1 \wedge 0 \leq a_t \leq 1 \wedge b_t \geq 1), \tag{4.21}$$

and, bearing in mind $\mathbb{E}[Z_t \,|\, \mathcal{F}_{t-1}] = 1$, we have in both cases $\mathcal{S}_{t-1}(\mathcal{F}_{t-1}) = \{1\}$. This shows $\mathcal{A}_{t-1}(\cdot \,|\, \cdot) = \mathbb{E}[\cdot \,|\, \cdot]$, see case (ii) above. If $b_t = \infty$, inequality 4.17 becomes obsolete and the other inequalities reduce to the alternative

$$(0 \leq a_{t-1} < 1 \wedge b_{t-1} \geq 1 \wedge a_t = 0) \tag{4.22}$$

or

$$(a_{t-1} = 1 \wedge b_{t-1} \geq 1 \wedge 0 \leq a_t \leq 1). \tag{4.23}$$

The first case ($a_t = 0$) shows that a_{t-1}, b_{t-1} can be chosen arbitrarily, if $\mathcal{A}_t$ is the conditional essential infimum (see (i)). The second case shows that $\mathcal{S}_{t-1}(\mathcal{F}_{t-1}) = \{1\}$ is allowed as the supergradient hull of $\mathcal{A}_{t-1}$ for arbitrary a_t (see (ii)).

We show now that information monotonicity is preserved under conditions (1), (2). Select any Z_{t+1} from $\mathcal{S}_t(\mathcal{F}_t)$ and Z_t from $\mathcal{S}_{t-1}(\mathcal{F}_t - 1)$. Then

$$0 \leq a_{t-1} \cdot a_t \leq Z_{t+1} \cdot Z_{t+1} \leq b_{t-1} \cdot b_t \tag{4.24}$$

holds for the product. In the first case (i) we have $a_t = 0$ and $b_t = \infty$, and (4.24) reduces to $0 \leq Z_t \cdot Z_{t+1}$. In the second case (ii) we have $a_{t-1} = 1$ and $b_{t-1} = 1$. Here, (4.24) reduces to $0 \leq a_t \leq Z_t \cdot Z_{t+1} \leq b_t$. Hence, in both cases the first constraint of $\mathcal{S}_t(\mathcal{F}_{t-1})$ is fulfilled for the product $M = Z_t \cdot Zt + 1$.

Furthermore, due to $\mathbb{E}[Z_t \,|\, \mathcal{F}_{t-1}] = 1$, it can be seen easily that the product also fulfills the second constraint of $\mathcal{S}_t(\mathcal{F}_{t-1})$, namely $\mathbb{E}[Z_t \cdot Z_{t+1} \,|\, \mathcal{F}_{t-1}] = 1$. □

Also within the following important class, compositions of conditional distortion mappings, information monotonicity cannot be attained easily.

Corollary 4.2. *Let $\overline{\mathcal{A}}$ be a composition of regular distortion functionals $\mathcal{A}_1, \ldots, \mathcal{A}_{T-1}$. If at least one of the functionals $\mathcal{A}_t$, $t < T - 2$, is not the conditional expectation, then $\overline{\mathcal{A}}$ is not information monotone.*

Proof. As is shown in Appendix A, (A.2) for regular distortion functionals

$$\mathcal{S}_t(\mathcal{F}) = \left\{ Z : \mathbb{E}[Z^k \,|\, \mathcal{F}] \le \int_0^1 [h_t(u)]^k \, du, \text{for } k = 0, 1, 2, \dots \right\}. \tag{4.25}$$

Since supergradients exist in the interior of the domain, there are functions Z such that the equality sign holds for every k. Suppose that $\mathcal{A}_{t-1}$ is not the conditional expectation. We show that condition (4.6) is violated. Let $c_{k,t} = \int_0^1 [h_t(u)]^k du$. Notice that $c_{k,t-1} > 1$ for $k > 0$, since $h_{t-1} \neq 1$.

Let $Z_{t-1} \in \mathcal{S}_{t-1}(\mathcal{F}_{t-1})$ and $Z_t \in \mathcal{S}_t(\mathcal{F}_t)$ such that the equality sign holds in (4.25). Then $\mathbb{E}[Z_{t-1}^k \,|\, \mathcal{F}_{t-1}] = c_{k,t-1}$, $\mathbb{E}[Z_t^k \,|\, \mathcal{F}_t] = c_{k,t}$ and thus

$$\begin{aligned} \mathbb{E}[(Z_{t-1} \cdot Z_t)^k \,|\, \mathcal{F}_{t-1}] &= \mathbb{E}[Z_{t-1}^k \cdot \mathbb{E}[Z_t^k \,|\, \mathcal{F}_t] \,|\, \mathcal{F}_{t-1}] \\ &= \mathbb{E}[Z_{t-1}^k c_{k,t} \,|\, \mathcal{F}_{t-1}) = c_{k,t} \cdot c_{k,t-1} > c_{k,t}. \end{aligned}$$

Hence, condition (4.6) is violated. □

From the previous corollaries it can be seen that a composition of $\mathbb{AV@R}_{\alpha_t}$'s can only be information information monotone if the vector $(\alpha_1, \dots, \alpha_T)$ is of the form $(1, \dots, 1, \alpha_t, 0, \dots, 0)$. Functionals of the form $\mathbb{E}[\mathcal{A}_{T-1}[\sum_{t=1}^T Y_t \,|\, \mathcal{F}_{T-1}]]$ play a special role with respect to information monotonicity. Clearly, such functionals are information monotone, if the mapping $\mathcal{A}_{T-1}$ is information monotone in the sense of Definition 3.2. In the end of Sec. 5, we will give a sufficient condition for general compositions of similar type type (separable expected conditional (SEC) functionals).

Another interesting example is the following mapping, which is closely related to MaxLoss-functional defined in [4]: the MinWin-functional can be interpreted as the minimum expected revenue under a generalized stress testing scenario — the worst distribution which is close to a given base line distribution with respect to the Kullback–Leibler divergence.

Example 4.3. Consider the positively homogeneous conditional functional

$$\mathrm{MinWin}_K(Y \,|\, \mathcal{F}) = \mathbf{inf}\{\mathbb{E}[Y \cdot Z \,|\, \mathcal{F}] : \mathbb{E}[Z \log Z \,|\, \mathcal{F}] \le K : \mathbb{E}[Z \,|\, \mathcal{F}] = 1, Z \ge 0\}, \tag{4.26}$$

where $K > 0$.

Its unconditional version can be written as

$$\mathrm{MinWin}(Y) = \inf\{\mathbb{E}_Q[Y] : \mathrm{KL}(P, Q) \le K\},$$

where $\mathrm{KL}(Q) = \mathbb{E}[\frac{dQ}{dP} \cdot \log(\frac{dQ}{dQ}]$ is the Kullback–Leibler divergence. The negative of this functional, applied to loss instead of profit, was introduced under the name of MaxLoss in [4], see also [19]. The dual pairing for the MinWin functional is given by the Zygmund spaces $L_{\exp}$ respectively $L \log^+ L$ (see [3, Theorem 6.5, p. 247]). MinWin is concave, positively homogeneous and translation-equivariant. Using the cumulant-generating function $\Lambda(\vartheta) = \log \mathbb{E}[\exp(\vartheta Y)]$, it is known that

$\mathrm{MinWin}_K(Y) = \Lambda'(\vartheta_K^-(Y))$, where $\vartheta_K^-(Y)$ is — under certain regularity conditions (see [4] for details) — the negative solution of

$$\vartheta\Lambda'(\vartheta) - \Lambda(\vartheta) = K. \tag{4.27}$$

The supergradient at Y is

$$Z = \frac{\exp(\vartheta_K^- Y)}{\mathbb{E}[\exp(\vartheta_K^- Y)]}. \tag{4.28}$$

For the nesting of two MinWin functionals, one has to look at condition (4.6). Notice that

$$\mathcal{S}_{t-1}(\mathcal{F}') = \{Z_1 : \mathbb{E}[Z_1 \log Z_1 \,|\, \mathcal{F}'] \leq K_1, \mathbb{E}[Z_1 \,|\, \mathcal{F}'] = 1, Z_1 \geq 0\},$$
$$\mathcal{S}_t(\mathcal{F}) = \{Z_2 : \mathbb{E}[Z_2 \log Z_2 \,|\, \mathcal{F}] \leq K_2, \mathbb{E}[Z_2 \,|\, \mathcal{F}] = 1, Z_2 \geq 0\}.$$

In the light (4.28) it is possible to choose densities Z_1^* and Z_2^* in a way such that

$$\mathbb{E}[Z_1^* \log Z_1^* \,|\, \mathcal{F}'] = K_1$$

and

$$\mathbb{E}[Z_2^* \log Z_2^* \,|\, \mathcal{F}] = K_2.$$

The product $M^* = Z_1^* \cdot Z_2^*$ lies in set $\mathcal{S}_{t-1}(\mathcal{F}') \cdot \mathcal{S}_t(\mathcal{F})$.

It follows that

$$\mathbb{E}[M^* \log M^* \,|\, \mathcal{F}'] = \mathbb{E}[K_2 Z_2^* + Z_2^* \log Z_2^* \,|\, \mathcal{F}'] = K_1 + K_2.$$

This result shows that $M^* \notin \mathcal{S}_t(\mathcal{F}')$: the inequality

$$\mathbb{E}[M \log M \,|\, \mathcal{F}'] \leq K_1$$

is not satisfied, because of $K_2 > 0$. Therefore, the composition of MinWin functionals is not information monotone.

5. Compositions of Nonhomogeneous Acceptability Functionals

Consider now additive compositions $\overline{\mathcal{A}}$ of mappings $\mathcal{A}_0, \ldots, \mathcal{A}_{T-1}$ of the form

$$\mathcal{A}_t(Y \,|\, \mathcal{F}) = \mathbf{inf}\{\mathbb{E}[Y \cdot Z \,|\, \mathcal{F}] - \mathcal{A}_t^+(Z \,|\, \mathcal{F}) : Z \in \mathcal{S}_t(\mathcal{F})\}, \tag{5.1}$$

such that the conjugate mappings $\mathcal{A}_t^+(Z \,|\, \mathcal{F}')$ take possibly other values than 0 and $-\infty$, at least for some t.

A general characterization for information monotonicity of such compositions is not known so far: the interplay between the mappings $\mathbb{E}[Y \cdot Z \,|\, \mathcal{F}] - \mathcal{A}_t^+(Z \,|\, \mathcal{F})$ and the sets $\mathcal{S}_t(\mathcal{F})$ can be quite complex. The following Proposition 5.1 is a generalization of the sufficiency part of Proposition 4.1, while generalization of the necessity part is yet an open problem.

Note that we call a conditional mapping $\mathcal{A}$ *compound concave*, if for all $Y \in \mathrm{dom}(\mathcal{A})$ and all $\mathcal{F}' \subseteq \mathcal{F}$

$$\mathcal{A}(Y \,|\, \mathcal{F}') \geq \mathbb{E}[\mathcal{A}(Y \,|\, \mathcal{F}) \,|\, \mathcal{F}'].$$

Proposition 5.1. *Let $\overline{\mathcal{A}}$ be a composition of acceptability mappings $\mathcal{A}_0, \ldots, \mathcal{A}_{T-1}$ such that*

$$\mathcal{A}_t(Y \,|\, \mathcal{F}) = \mathbf{inf}\{\mathbb{E}[Y \cdot Z \,|\, \mathcal{F}] - \mathcal{A}_t^+(Z \,|\, \mathcal{F}) : Z \in \mathcal{S}_t(\mathcal{F})\}$$

with supergradient hull $\mathcal{S}_t(\mathcal{F})$ for a given σ-field $\mathcal{F}$. Assume that each positively homogeneous mapping

$$\mathcal{A}_t(Y \,|\, \mathcal{F}) = \mathbf{inf}\{\mathbb{E}[Y \cdot Z \,|\, \mathcal{F}] : Z \in \mathcal{S}_t(\mathcal{F})\} \tag{5.2}$$

is information monotone, and the sets $\mathcal{S}_t(\cdot)$ fulfill the condition

$$\mathcal{S}_{t-1}(\mathcal{F}') \cdot \mathcal{S}_t(\mathcal{F}) \subseteq \mathcal{S}_t(\mathcal{F}'), \tag{5.3}$$

for every $t, 1 \leq t \leq T-1$ and every pair of σ-fields $\mathcal{F}'$, $\mathcal{F}$ such that $\mathcal{F}' \subseteq \mathcal{F}$.

Assume further that there are representations

$$\mathcal{A}_t^+ \left(\frac{M_{t+1}}{M_t} \middle| \mathcal{F}_t \right) M_t = \widetilde{\mathcal{A}}_t(f_t(M_{t+1}, M_t) \,|\, \mathcal{F}_t), \tag{5.4}$$

for $t = 1, \ldots, T$ such that all f_t are measurable functions, and all $\widetilde{\mathcal{A}}_t$ are compound concave conditional mappings. If in addition the inequality

$$\begin{aligned} &\mathbb{E}\left[\mathcal{A}_{t-1}^+ \left(\frac{M_t}{M_{t-1}} \middle| \mathcal{F}_{t-1} \right) M_{t-1} \right] + \mathbb{E}\left[\mathcal{A}_t^+ \left(\frac{M_{t+1}}{M_t} \middle| \mathcal{F}_t \right) M_t \right] \\ &\quad \leq \mathbb{E}\left[\mathcal{A}_t^+ \left(\frac{M_{t+1}}{M_{t-1}} \middle| \mathcal{F}_{t-1} \right) M_{t-1} \right] \end{aligned} \tag{5.5}$$

holds for any nonnegative martingale (M_{t-1}, M_t, M_{t+1}) with $E[M_t] = 1$ and for $t = 2, \ldots, T$, then the composition $\overline{\mathcal{A}}$ is information monotone. If any of the inequalities involved is strict for at least one $t = 1, \ldots, T$, then the mapping is strictly information monotone.

Proof. By (3.13), we have

$$\begin{aligned} \overline{\mathcal{A}}(\mathbf{Y} \,|\, \mathfrak{F}) = \inf_{M_1, M_2, \ldots, M_T} \Bigg\{ &\langle \mathbf{Y}, \mathbf{M} \rangle - \mathcal{A}_0^+(M_1) \\ &- \sum_{t=1}^{T-1} \mathbb{E}\left[\mathcal{A}_t^+ \left(\frac{M_{t+1}}{M_t} \middle| \mathcal{F}_t \right) M_t \right] : M_0 \equiv 1, \frac{M_{t+1}}{M_t} \in \mathcal{S}_t(\mathcal{F}_t) \Bigg\} \end{aligned}$$

for a given process $\mathbf{Y}$ related to the filtration $\mathfrak{F}$.

Now we distinguish two cases. (i) Consider a filtration $\mathfrak{F}' = (\mathcal{F}_0', \ldots, \mathcal{F}_T')$, where (at some point in time t_0) $\mathcal{F}_{t_0}$ is replaced by a σ-field $\mathcal{F}_{t_0}'$ with $\mathcal{F}_{t_{0-1}} \subset \mathcal{F}_{t_0}' \subseteq \mathcal{F}_{t_0}$, and all other σ-fields $\mathcal{F}_t$, $t \neq t_0$ remain unchanged. Define a modified optimization problem

$$A^* = \inf_{M_1, M_2, \ldots, M_T} \Bigg\{ \langle \mathbf{Y}, \mathbf{M} \rangle - \mathcal{A}_0^+(M_1) - \sum_{t=1}^{T-1} \mathbb{E}\left[\mathcal{A}_t^+ \left(\frac{M_{t+1}}{M_t} \middle| \mathcal{F}_t \right) M_t \right] : M_0 \equiv 1, \frac{M_{t+1}}{M_t} \in \mathcal{S}_t(\mathcal{F}'_t) \Bigg\}.$$

Note that the objective function remains unchanged, while the sets $\mathcal{S}_t$ are built w.r.t. $\mathfrak{F}'$. Using (5.3) and information monotonicity of (5.2) it is easily possible to extend the arguments related to the sufficiency part of Proposition 4.1, which results in

$$\overline{\mathcal{A}}(\mathbf{Y} \mid \mathfrak{F}) \geq A^*.$$

Because A^* is an infimum, for any $\varepsilon > 0$ it is possible to select a $\mathbf{M}^{(\epsilon)}$ such that $\frac{M_{t+1}^{(\varepsilon)}}{M_t^{\varepsilon}} \in \mathcal{S}_t(\mathcal{F}'_t)$ and

$$\langle \mathbf{Y}, \mathbf{M}^{(\varepsilon)} \rangle - \mathcal{A}_0^+(M_1^{(\varepsilon)}) - \sum_{t=1}^{T-1} \mathbb{E}\left[\mathcal{A}_t^+ \left(\frac{M_{t+1}^{(\varepsilon)}}{M_t^{(\varepsilon)}} \middle| \mathcal{F}_t \right) M_t^{(\varepsilon)} \right] \leq A^* + \varepsilon.$$

Applying compound concavity we get

$$\begin{aligned}
\mathbb{E}\left[\mathcal{A}_t^+ \left(\frac{M_{t_0+1}^{(\varepsilon)}}{M_{t_0}^{(\varepsilon)}} \middle| \mathcal{F}_{t_0} \right) M_{t_0}^{(\epsilon)} \right] &= \mathbb{E}[\widetilde{\mathcal{A}}_{t_0}(f_{t_0}(M_{t_0+1}^{(\epsilon)}, M_{t_0}^{(\epsilon)}) \mid \mathcal{F}_t)] \\
&= \mathbb{E}[\mathbb{E}[\widetilde{\mathcal{A}}_{t_0}(f_{t_0}(M_{t_0+1}^{(\epsilon)}, M_{t_0}^{(\epsilon)}) \mid \mathcal{F}_{t_0}) \mid \mathcal{F}'_{t_0}]] \\
&\leq \mathbb{E}[\widetilde{\mathcal{A}}_{t_0}(f_{t_0}(M_{t_0+1}^{(\epsilon)}, M_{t_0}^{(\epsilon)}) \mid \mathcal{F}'_{t_0})] \\
&= \mathbb{E}\left[\mathcal{A}_{t_0}^+ \left(\frac{M_{t_0+1}^{(\epsilon)}}{M_{t_0}^{(\epsilon)}} \middle| \mathcal{F}'_{t_0} \right) M_{t_0}^{(\epsilon)} \right],
\end{aligned}$$

which implies

$$\langle \mathbf{Y}, \mathbf{M}^{(\varepsilon)} \rangle - \mathcal{A}_0^+(M_1^{(\varepsilon)}) - \sum_{t=1}^{T-1} \mathbb{E}\left[\mathcal{A}_t^+ \left(\frac{M_{t+1}^{(\varepsilon)}}{M_t^{(\varepsilon)}} \middle| \mathcal{F}'_t \right) M_t^{(\varepsilon)} \right] \leq A^* + \varepsilon,$$

for all $\varepsilon > 0$. It follows that

$$\inf \left\{ \langle \mathbf{Y}, \mathbf{M} \rangle - \mathcal{A}_0^+(M_1) - \sum_{t=1}^{T-1} \mathbb{E}\left[\mathcal{A}_t^+ \left(\frac{M_{t+1}}{M_t} \middle| \mathcal{F}'_t \right) M_t^{(\varepsilon)} \right] : \frac{M_{t+1}^{(\varepsilon)}}{M_t^{\varepsilon}} \in \mathcal{S}_t(\mathcal{F}'_t) \right\} \leq A^*,$$

hence

$$\overline{\mathcal{A}}(\mathbf{Y} \mid \mathfrak{F}) \geq \overline{\mathcal{A}}(\mathbf{Y} \mid \mathfrak{F}').$$

(ii) Consider now a filtration $\mathfrak{F}' = (\mathcal{F}'_0, \ldots, \mathcal{F}'_T)$, where — at some time t_0 – $\mathcal{F}_{t_0}$ is replaced by $\mathcal{F}'_{t_0} = \mathcal{F}_{t_{0-1}}$. All other σ-fields $\mathcal{F}_t$, $t \neq t_0$ remain unchanged. This only works for processes where Y_{t_0} is measurable w.r.t. $\mathcal{F}_{t_{0-1}}$. Otherwise $\overline{\mathcal{A}}(\mathbf{Y},\mathfrak{F}')=-\infty$ and information monotonicity is fulfilled in a trivial way. Hence, the composition can be represented as

$$\overline{\mathcal{A}}(\mathbf{Y}\,|\,\mathfrak{F}) = \inf_{\mathbf{M}_1,\mathbf{M}_2,\ldots,\mathbf{M}_T}\Bigg\{\langle \mathbf{Y},\mathbf{M}\rangle - \mathcal{A}_0^+(M_1) - \sum_{t=1}^{T-1}\mathbb{E}\left[\mathcal{A}_t^+\left(\frac{M_{t+1}}{M_t}\middle|\mathcal{F}_t\right)M_t\right] : M_0=1, \frac{M_{t+1}}{M_t}\in\mathcal{S}_t(\mathcal{F}_t)\Bigg\}$$

with $\langle \mathbf{Y},\mathbf{M}\rangle = \sum_{t=1}^{t_0-2}\mathbb{E}[Y_tM_t] + \mathbb{E}[(Y_{t_0-1}+Y_{t_0})M_{t_0-1}] + \sum_{t=t_0+1}^{T-1}\mathbb{E}[Y_tM_t]$.

Applying the same arguments as above to the modified optimization problem

$$A^{**} = \inf_{M_1,M_2,\ldots,M_T}\Bigg\{\langle \mathbf{Y},\mathbf{M}\rangle - \mathcal{A}_0^+(M_1) - \sum_{t=1}^{T-1}\mathbb{E}\left[\mathcal{A}_t^+\left(\frac{M_{t+1}}{M_t}\middle|\mathcal{F}_t\right)M_t\right] : M_0\equiv 1, \frac{M_{t+1}}{M_t}\in\mathcal{S}_t(\mathcal{F}_t')\Bigg\}$$

we get

$$\langle \mathbf{Y},\mathbf{M}^{(\varepsilon)}\rangle - \mathcal{A}_0^+(M_1^{(\varepsilon)}) - \sum_{t=1}^{T-1}\mathbb{E}\left[\mathcal{A}_t^+\left(\frac{M_{t+1}^{(\varepsilon)}}{M_t^{(\varepsilon)}}\middle|\mathcal{F}_t\right)M_t^{(\varepsilon)}\right] \le A^{**}+\varepsilon,$$

and by assumption (5.5)

$$\langle \mathbf{Y},\mathbf{M}^{(\varepsilon)}\rangle - \sum_{t\ne t_0-1}\mathbb{E}\left[\mathcal{A}_t^+\left(\frac{M_{t+1}^{(\varepsilon)}}{M_t^{(\varepsilon)}}\middle|\mathcal{F}_t'\right)M_t^{(\varepsilon)}\right] \le A^{**}+\varepsilon.$$

Hence

$$\inf_{\mathbf{M}}\Bigg\{\langle \mathbf{Y},\mathbf{M}\rangle - \sum_{t=1}^{t_0-2}\mathbb{E}\left[\mathcal{A}_t^+\left(\frac{M_{t+1}}{M_t}\middle|\mathcal{F}_t'\right)M_t\right] + \mathbb{E}\left(\mathcal{A}_{t_0}^+\left(\frac{M_{t_0+1}}{M_{t_0-1}}\middle|\mathcal{F}_{t_0-1}'\right)M_{t_0-1}\right) + \sum_{t=t_0+1}^{T-1}\mathbb{E}\left[\mathcal{A}_t^+\left(\frac{M_{t+1}}{M_t}\middle|\mathcal{F}_t'\right)M_t\right] : \frac{M_{t+1}}{M_t}\in\mathcal{S}_t(\mathcal{F}_t')\Bigg\} \le A^{**}.$$

The left side of this inequality is a representation of the composition $\overline{\mathcal{A}}(\mathbf{Y}\,|\,\mathfrak{F})$, where any occurrence of time t_0-1 has been removed due to measurability of Y_{t_0} and translation equivariance. Hence, again

$$\overline{\mathcal{A}}(\mathbf{Y}\,|\,\mathfrak{F}) \ge \overline{\mathcal{A}}(\mathbf{Y}\,|\,\mathfrak{F}'). \qquad \square$$

The following examples show applications of Proposition 5.1:

Example 5.1. Nested entropic acceptability functionals are defined as compositions $\mathcal{A}_0\circ\mathcal{A}_1\circ\cdots\circ\mathcal{A}_{T-1}$, where $\mathcal{A}_t(Y\,|\,\mathcal{F}_t) = -\frac{1}{\gamma_t}\log\mathbb{E}[\exp(-\gamma_tY)\,|\,\mathcal{F}_t]$, and $\gamma_t\ge 0$

for $\mathbf{Y} \in \mathcal{Y}$. Entropic risk functionals — with reversed sign — were introduced in [12] and further studied in [9]. The conjugate representation of $\mathcal{A}_t$ is

$$\mathcal{A}_t(Y \mid \mathcal{F}_t) = \mathbf{inf}\left\{\mathbb{E}[Y \cdot Z \mid \mathcal{F}_t] + \frac{1}{\gamma_t}\mathbb{E}[Z \log Z \mid \mathcal{F}_t] : \mathbb{E}[Z \mid \mathcal{F}_t] = 1, Z \geq 0\right\},$$

where $0 \log 0$ is defined as 0. The correct dual pairing for the entropic functional is given by the Zygmund spaces $L_{\exp}$ and $L \log^+ L$. We will show that nested entropic functionals are information monotone, if $\gamma_0 \leq \gamma_1 \leq \cdots \leq \gamma_{T-1}$.

The composition has the representation

$$\overline{\mathcal{A}}(\mathbf{Y} \mid \mathfrak{F}) = \inf\left\{\langle \mathbf{Y}, \mathbf{M}\rangle + \sum_{t=0}^{T-1} \frac{1}{\gamma_t}\mathbb{E}\left[\mathbb{E}\left[\frac{M_{t+1}}{M_t}\log\frac{M_{t+1}}{M_t}\middle|\mathcal{F}_t\right] M_t\right] : \mathbf{M} \in \mathcal{S}\right\},$$

with $M_0 = 1$. This can be rewritten as

$$\overline{\mathcal{A}}(\mathbf{Y}|\mathfrak{F}) = \inf\left\{\langle \mathbf{Y}, \mathbf{M}\rangle + \sum_{t=0}^{T-1} \frac{1}{\gamma_t}\mathbb{E}[\mathbb{E}[M_{t+1}\log M_{t+1} - M_{t+1}\log M_t|\mathcal{F}_t]] : \mathbf{M} \in \mathcal{S}\right\},$$

and shows that with $f(M_t, M_{t+1}) = M_{t+1}\log M_{t+1} - M_{t+1}\log M_t$, and $\widetilde{\mathcal{A}}_t(\cdot|\mathcal{F}) = \frac{1}{\gamma_t} \cdot \mathbb{E}[\cdot|\mathcal{F}]$ the composition $\overline{\mathcal{A}}(\mathbf{Y}|\mathfrak{F})$ has the compounding property of Theorem 5.1, because $\mathbb{E}[\cdot|\mathcal{F}]$ is compound linear. In addition, the sets $\mathcal{S}_t$ fulfill (5.3).

Furthermore, using $\mathbb{E}[M_{t+1}|\mathcal{F}_t] = M_t$ leads to $\mathbb{E}[M_{t+1}\log M_t|\mathcal{F}_t] = M_t \log M_t$, hence

$$\begin{aligned}
&\mathbb{E}\left[\mathcal{A}_{t-1}^+\left(\frac{M_t}{M_{t-1}}\middle|\mathcal{F}_{t-1}\right) M_{t-1}\right] + \mathbb{E}\left[\mathcal{A}_t^+\left(\frac{M_{t+1}}{M_t}\middle|\mathcal{F}_t\right) M_t\right] \\
&\quad = -\frac{1}{\gamma_{t-1}}\mathbb{E}[M_t\log M_t - M_t \log M_{t-1}] - \frac{1}{\gamma_t}\mathbb{E}[(M_{t+1}\log M_{t+1} - M_{t+1}\log M_t] \\
&\quad = -\frac{1}{\gamma_{t-1}}\mathbb{E}[M_t\log M_t - M_{t-1} \log M_{t-1}] - \frac{1}{\gamma_t}\mathbb{E}[M_{t+1}\log M_{t+1} - M_t\log M_t] \\
&\quad = -\left(\frac{1}{\gamma_{t-1}} - \frac{1}{\gamma_t}\right)\mathbb{E}\left[M_t \log M_t\right] - \frac{1}{\gamma_t}\mathbb{E}\left[M_{t+1}\log M_{t+1}\right] \\
&\qquad + \frac{1}{\gamma_{t-1}}\mathbb{E}\left[M_{t-1}\log M_{t-1}\right] \\
&\quad \leq -\left(\frac{1}{\gamma_{t-1}} - \frac{1}{\gamma_t}\right)\mathbb{E}\left[M_{t-1} \log M_{t-1}\right] - \frac{1}{\gamma_t}\mathbb{E}\left[M_{t+1}\log M_{t+1}\right] \\
&\qquad + \frac{1}{\gamma_{t-1}}\mathbb{E}\left[M_{t-1}\log M_{t-1}\right] = -\frac{1}{\gamma_t}\mathbb{E}[\mathbb{E}[M_{t+1}\log M_{t+1} - M_{t+1}\log M_{t-1}|\mathcal{F}_{t-1}]] \\
&\quad = \mathbb{E}\left[\mathcal{A}_t^+\left(\frac{M_{t+1}}{M_{t-1}}\middle|\mathcal{F}_{t-1}\right) M_{t-1}\right],
\end{aligned}$$

if $(\frac{1}{\gamma_{t-1}} - \frac{1}{\gamma_t}) \geq 0$. Notice that $\mathbb{E}[M_{t-1}\log M_{t-1}] \leq \mathbb{E}[M_t \log M_t]$, since $\mathbb{E}[M_t|\mathcal{F}_{t-1}] = M_{t-1}$ and $u \mapsto u \log u$ is convex.

It follows that compositions of entropic mappings are information monotone, if $\gamma_t \geq \gamma_{t-1}$ for all t.

An important class of multi-period valuation functionals are SEC functionals. It shows that they are both, time consistent and information monotone.

Based on conditional acceptability mappings $\mathcal{A}_t(\cdot|\mathcal{F}_t)$ with conjugates $\mathcal{A}_t^+(\cdot|\mathcal{F}_t)$ they are constructed ([26]) as follows:

$$\mathcal{A}(\mathbf{Y}|\mathfrak{F}) := \sum_{t=0}^{T-1} \mathbb{E}[\mathcal{A}_t(Y_{t+1}|\mathcal{F}_t)]. \tag{5.6}$$

Such functionals were already shortly mentioned in Sec. 4. It has been shown (see [26, Proposition 3.27]) that SEC functionals (5.6) have the conjugate representation

$$\mathcal{A}(\mathbf{Y}|\mathfrak{F}) = \inf\left\{\langle \mathbf{Y}, \mathbf{Z}\rangle - \sum_{t=0}^{T-1} \mathbb{E}[\mathcal{A}_t^+(Z_{t+1}|\mathcal{F}_t)] : Z_{t+1} \geq 0,\right.$$
$$\left.\mathbb{E}[Z_{t+1}|\mathcal{F}_t] = 1, t = 0, \ldots, T-1\right\},$$

hence the conjugate of SEC functionals is also SEC.

Although they are no additive compositions, SEC functionals can be constructed recursively and hence are time consistent. It should be noted that the family of information monotone functionals is closed under addition, because it forms a convex cone (see Lemma 3.1). Hence, SEC functionals are information monotone, if all compositions $\mathbb{E}[\mathcal{A}_t(Y_{t+1}|\mathcal{F}_t)]$ are information monotone.

For positive homogeneous conditional mappings $\mathcal{A}_t$ this is the case, if (3.9) holds. For nonhomogeneous mappings sufficient conditions are more complicated. It shows that SEC-functionals are information monotone if the $\mathcal{A}_t^+$ are compound concave and their supergradient hulls fulfill the conditions of Proposition 5.1. This can be seen as follows:

Consider the simple acceptability compositions

$$\overline{\mathcal{A}}^{(t)}(\mathbf{Y}|\mathfrak{F}) = \mathbb{E}[\mathcal{A}_t(Y_{t+1}|\mathcal{F}_t)]. \tag{5.7}$$

We may apply our theorem with

$$f(M_t, M_{t-1}) = f(M_t, 1) = M_t,$$

since $\mathcal{S}_{t-1}(\mathcal{F}_t) = \{1\}$ for the expectation, and $\mathcal{A}_t^+$ is compound concave. For the same reason the sets $\mathcal{S}_t$ — information monotone by assumption — can easily seen to fulfill condition (5.3). Finally inequality (5.5) holds if one formally replaces $t-1$ by 0 and considers $M_0 = M_t = 1$ and $\mathcal{A}_0^+(1|\mathcal{F}_{t-1}) = 0$ for the expectation.

Surprisingly, another example for SEC-functionals are compositions of entropic functionals, already analyzed in Example 5.1: by the arguments of Proposition 5.1

such mappings can be represented as

$$\overline{\mathcal{A}}(\mathbf{Y}|\mathfrak{F}) = \inf\left\{\langle \mathbf{Y}, \mathbf{M}\rangle + \sum_{t=1}^{T-1}\left(\frac{1}{\gamma_{t-1}} - \frac{1}{\gamma_t}\right)\mathbb{E}[M_t \log M_t]\right.$$

$$\left. + \frac{1}{\gamma_{T-1}}\mathbb{E}[M_T \log M_T] : \mathbf{M} \in \mathcal{S}\right\}.$$

This is the representation as a SEC functional only if $(\frac{1}{\gamma_{t-1}} - \frac{1}{\gamma_t}) \geq 0$. Otherwise the expresssion $-(\frac{1}{\gamma_{t-1}} - \frac{1}{\gamma_t})\mathbb{E}[M_t \log M_t|\mathcal{F}_{t-1}]$ would not be concave. If the requirements on all the $\mathcal{S}_t$ are fulfilled, information monotonicity must hold, as was already shown in Example 5.1 in another way.

6. Conclusions

We defined the notion of information monotonicity for multi-period acceptability type valuation functionals and argued for the importance of this concept, especially in economic decision making. After discussing the connections to multi-stage stochastic programming we analyzed the application of the definition to the broad class of additive compositions of acceptability mappings. Such nested functionals have another important property: time consistency, which builds on valuations over time that do not contradict each other.

The main results — Propositions 4.1 and 5.1 — answer the question, under which additional requirements additive compositions can be information monotone. It shows that the domains of their conjugate mappings play a key role for answering this question.

- For compositions of positively homogeneous mappings with monotone behavior of their supergradient hulls, the restrictive condition (4.6) is necessary and sufficient. An important example — showing the severity of this condition — is given by compositions of distortion functionals, which reduce to the only feasible case of iterative application of conditional expectations or essential infima.
- If the composition contains mappings that are not positively homogeneous, only a sufficient condition is known so far. The same conditions as in the positively homogeneous case hold for the supergradient hull. In addition the conjugate functional has to fulfill a condition, representable by a certain compound concave construction (5.4) for the case $\mathcal{F}_{t-1} \subset \mathcal{F}'_t \subset \mathcal{F}_t$, and a monotonicity condition (5.5) for the case $\mathcal{F}'_t = \mathcal{F}_{t-1}$. Only certain compositions of mappings representable as SEC functionals are known to fulfill this requirements so far.

The results show that time consistency and information monotonicity cannot easily be combined for multi-period-valuation functionals, although a nontrivial intersection, in particular SEC functionals, exists. Clearly, this comes from the fact that both time consistency and information monotonicity are considerably strong requirements. In concrete situations, one always has to consider which property might be more important. Despite the fact that time consistency (recursivity) allows for dynamic programming, practitioners often might prefer information monotonicity, because usually, their estimated models and the related decisions have to be revised in the light of new data.

Acknowledgment

We thank the anonymous referee for valuable suggestions and hints to the literature, which helped to improve the paper.

Appendix A. Distortion Functionals

Distortion functionals (see e.g, [28]) are a very broad subclass of positively homogeneous acceptability mappings, and we will analyze compositions of distortion functionals with respect to information monotonicity in Sec. 4.

Let $H(u)$ be a concave, monotone, right continuous function $[0-,1] \to [0,1]$ such that $H(0-) = 0$, $H(0) = 1$. Concavity implies that H is continuous in $(0,1]$ and that a jump may only occur at zero. Moreover, H is representable as $H(u) = H(0) + \int_0^u h(v)dv$ for a monotonically decreasing nonnegative function h.

Definition A.1. The functional $\mathcal{A}_H(Y) = \int_{0-}^1 G_Y(u)^{-1} dH(u)$ is called *distortion functional* with distortion H. The derivative h is called *distortion density*.

Example A.1. All admissible distortions lie between $H(u) = u$ and $H(u) = 1$. The extreme cases are $H(u) = u$ leading to $\mathcal{A}_H(Y) = \mathbb{E}[Y]$ and $H(u) = 1$ leading to $\mathcal{A}_H(Y) = \operatorname{essinf}(Y)$. The choice of $H(u) = \min(u/\alpha, 1)$ leads to the unconditional average value-at-risk $\mathbb{AV@R}_\alpha(Y)$. Another often used functional is $\mathcal{A}(Y) = \mathbb{E}[Y] - \operatorname{Gini}(Y)$, where $\operatorname{Gini}(Y) = \mathbb{E}|Y - Y'|$, with Y' being an independent copy of Y. A calculation shows that this functional equals

$$2\int_0^1 \alpha \mathbb{AV@R}_\alpha d\alpha = \int_0^1 G_Y^{-1}(u) d(2u - u^2) = 2\int_0^1 G_Y^{-1}(u)(1-u)du,$$

i.e. it is of distortion type.

If $H(0) = 0$, the conjugate representation of the distortion functional is

$$\mathcal{A}_h(Y) = \inf\{\mathbb{E}[Y \cdot Z] : \mathbb{E}[Z] = 1, Z = h(U), \text{where } U \sim \text{Uniform } [0,1]\} \qquad \text{(A.1)}$$

(see [28]). The infimum is attained, the supergradients are given by $Z = G_Y(Y) - V_Y$, where $V_Y = V[(G(Y) - G(Y-)]$ with V being a uniform [0, 1] variable, independent of Y. This generalized quantile transform (see [10]) can be constructed if the probability space is rich enough.

The essential infimum has the representation

$$\operatorname{essinf}(Y) = \inf\{\mathbb{E}[Y \cdot Z] : \mathbb{E}[Z] = 1, Z \geq 0\}.$$

Note that the infimum is not attained in this case.

Let us call a distortion regular, if $H(0) = 0$ and $\int h^k(u)du < \infty$ for all k. Notice that $Z \sim h(U)$ iff $\mathbb{E}[Z^k] = \int_0^1 [h(u)]^k du$, for $k = 1, 2, \ldots,$ since, by the Carleman condition (see [5]) the integer moments determine fully a bounded distribution. It is well-known that

$$\begin{aligned}&\mathrm{conv}\{Z : \mathbb{E}[Z] = 1, Z = h(U), \text{ where } U \sim \text{Uniform } [0,1]\}\\ &\qquad = \{Z : \mathbb{E}[Z] = 1, Z \prec_{CXD} h(U), \text{ where } U \sim \text{Uniform } [0,1]\},\end{aligned}$$

where $\prec_{CDX}$ denotes convex dominance (see [8]). (A.1) is equivalent to

$$\mathcal{A}_h(Y) = \inf \left\{ \mathbb{E}[Y \cdot Z] : \mathbb{E}[Z] = 1, \mathbb{E}[Z^k] \leq \int_0^1 [h(u)]^k du, \text{ for } k = 2, 3, \ldots \right\}. \tag{A.2}$$

Appendix B. Proof of Proposition 3.1

Following [17], we show the formula for a two-period composition, the generalization to arbitrary T is obvious. In the following, all the infima must be understood with respect to the constraints $Z_2 \geq 0$, $\mathbb{E}[Z_2 \,|\, \mathcal{F}_1] = 1$, and $Z_1 \geq 0$; $\mathbb{E}[Z_1 \,|\, \mathcal{F}_0] = 1$. For simplicity of notation, we write $\mathcal{A}_t(\cdot)$ for $\mathcal{A}_t(\cdot \,|\, \mathcal{F}_t)$, $t = 1, 2$. Based on

$$\mathcal{A}_1(Y_2) = \inf_{Z_2} \{\mathbb{E}[Y_2 Z_2 \,|\, \mathcal{F}_1] - \mathcal{A}_1^+(Z_2)\},$$

$$\mathcal{A}_0(Y_1) = \inf_{Z_1} \{\mathbb{E}[Y_1 Z_1 \,|\, \mathcal{F}_0] - \mathcal{A}_0^+(Z_1)\},$$

we get

$$\begin{aligned}\mathcal{A}_0(Y_1 + \mathcal{A}_1(Y_2)) &= \inf_{Z_1} \{\mathbb{E}[[Y_1 + \mathcal{A}_1(Y_2)] \cdot Z_1 \,|\, \mathcal{F}_0] - \mathcal{A}_0^+(Z_1)\}\\ &= \inf_{Z_1} \left\{ \mathbb{E}[Y_1 Z_1 \,|\, \mathcal{F}_0] + \mathbb{E}\left[\inf_{Z_2} \{\mathbb{E}[Y_2 Z_2 \,|\, \mathcal{F}_1] - \mathcal{A}_1^+(Z_2)\} Z_1 \,|\, \mathcal{F}_0 \right] - \mathcal{A}_0^+(Z_1) \right\}\\ &= \inf_{Z_1, Z_2} \{\mathbb{E}[Y_1 \cdot Z_1 \,|\, \mathcal{F}_0] + \mathbb{E}[Y_2 Z_1 Z_2 \,|\, \mathcal{F}_0] - \mathcal{A}_0^+(Z_1) - \mathbb{E}[\mathcal{A}_1^+(Z_2) Z_1 \,|\, \mathcal{F}_0]\}.\end{aligned}$$

Here, we have used that

$$\begin{aligned}\mathbb{E}\left[\inf_{Z_2} \{\mathbb{E}[Y_2/, Z_2 \,|\, \mathcal{F}_1] - \mathcal{A}_1(Z_2)\} Z_1 \,|\, \mathcal{F}_0 \right] &= \inf_{Z_2} \{\mathbb{E}[\mathbb{E}[Y_2 Z_2 \,|\, \mathcal{F}_1] Z_1 - \mathcal{A}_1(Z_2) Z_1 \,|\, \mathcal{F}_0]\}\\ &= \inf_{Z_2} \{\mathbb{E}[Y_2 Z_2 Z_1 \,|\, \mathcal{F}_0] - \mathbb{E}[\mathcal{A}_1(Z_2) Z_1 \,|\, \mathcal{F}_0]\}.\end{aligned} \tag{B.1}$$

To prove (B.1) notice that

$$\begin{aligned}&\mathbb{E}\left[\inf_{Z_2} \{\mathbb{E}[Y_2/, Z_2 \,|\, \mathcal{F}_1] - \mathcal{A}_1(Z_2)\} Z_1 \,|\, \mathcal{F}_0 \right]\\ &\qquad \leq \inf_{Z_2} \{\mathbb{E}[(\mathbb{E}[Y_2 Z_2 \,|\, \mathcal{F}_1] - \mathcal{A}_1(Z_2)) Z_1 \,|\, \mathcal{F}_0]\}.\end{aligned}$$

Since the infimum is attained, there is a Z_{Y_2} such that

$$\mathbb{E}\left[\inf_{Z_2}\{(\mathbb{E}[Y_2 Z_2 \,|\, \mathcal{F}_1] - \mathcal{A}_1(Z_2))Z_1\} \,|\, \mathcal{F}_0\right] = \mathbb{E}[(\mathbb{E}[Y_2/, Z_{Y_2} \,|\, \mathcal{F}_1] - \mathcal{A}_1(Y_{Z_2}))Z_1 \,|\, \mathcal{F}_0]$$

$$\geq \inf_{Z_2}\{\mathbb{E}[(\mathbb{E}[Y_2 Z_2 Z_1 \,|\, \mathcal{F}_1] - \mathcal{A}_1(Z_2))Z_1 \,|\, \mathcal{F}_0]\}$$

which shows (B.1). □

Appendix C. Probability Functionals of Adapted Processes

Lemma C.1. *Let $\mathcal{A}$ be a proper multi-period acceptability functional and fix some filtration $\mathfrak{F}$. Then $\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) = -\infty$ for $\mathbf{Y} \ntriangleleft \mathfrak{F}$ if and only if $\mathcal{A}^+(\mathbf{Z} \,|\, \mathfrak{F}) = \mathcal{A}^+(\mathbb{E}[\mathbf{Z} \,|\, \mathfrak{F}] \,|\, \mathfrak{F})$ for all $\mathbf{Z} \in \mathcal{Z}$, where $\mathbb{E}[\mathbf{Z} \,|\, \mathfrak{F}] = (\mathbb{E}[Z_1 \,|\, \mathcal{F}_1], \ldots, \mathbb{E}[Z_T \,|\, \mathcal{F}_T])$.*

Proof. Suppose that $\mathcal{A}^+(\mathbf{Z}) = \mathcal{A}^+(\mathbb{E}[\mathbf{Z} \,|\, \mathfrak{F}])$ for all $\mathbf{Z} \in \mathcal{Z}$ and assume that $\mathbf{Y} \ntriangleleft \mathfrak{F}$. Let $Y_t^{(1)} = \mathbb{E}[Y_t \,|\, \mathcal{F}_t]$ and $Y_t^{(2)} = Y - Y_t^{(1)}$. Notice that by $\mathbf{Y} \ntriangleleft \mathfrak{F}$ it follows $\sum_t \mathbb{E}[|Y_t|] > 0$. Let $\mathbf{Z}^{(1)} = (Z_1^{(1)}, \ldots, Z_T^{(1)}) \triangleleft \mathfrak{F}$ such that $\mathcal{A}^+(\mathbf{Z}^{(1)}) > -\infty$. Let now $Z_t^{(K)} = -K \operatorname{sgn} Y_t^{(2)} + K\mathbb{E}[\operatorname{sgn} Y_t^{(2)} \,|\, \mathcal{F}_t]$. Notice that $\mathbb{E}[\mathbf{Z}^{(K)} \,|\, \mathfrak{F}] = 0$. Set $\mathbf{Z} = \mathbf{Z}^{(1)} + \mathbf{Z}^{(K)}$. Then

$$\begin{aligned}
\mathcal{A}(\mathbf{Y}) &= \inf\{\langle \mathbf{Y}, \mathbf{Z}\rangle - \mathcal{A}^+(\mathbb{E}[\mathbf{Z} \,|\, \mathfrak{F}]) : \mathbf{Z} \in \mathcal{Z}\} \\
&\leq \inf\{\langle \mathbf{Y}^{(1)}, \mathbf{Z}^{(1)}\rangle + \langle \mathbf{Y}^{(2)}, \mathbf{Z}^{(K)}\rangle - \mathcal{A}^+(\mathbf{Z}^{(1)}) : K \in \mathbb{R}\} \\
&= \inf\left\{\langle \mathbf{Y}^{(1)}, \mathbf{Z}^{(1)}\rangle - K\sum_{t=1}^{T} \mathbb{E}[|Y_t^{(2)}|] - \mathcal{A}^+(\mathbf{Z}^{(1)}) : K \in \mathbb{R}\right\} \\
&= -\infty,
\end{aligned}$$

since $\sum_{t=1}^{T} \mathbb{E}[|Y_t^{(2)}|] > 0$. Conversely, if $\mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) = -\infty$ for $\mathbf{Y} \ntriangleleft \mathfrak{F}$, then

$$\begin{aligned}
\mathcal{A}^+(\mathbf{Z} \,|\, \mathfrak{F}) &= \inf\{\langle \mathbf{Y}, \mathbf{Z}\rangle - \mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) : \mathbf{Y} \in \mathcal{Y}\} \\
&= \inf\{\langle \mathbf{Y}, \mathbf{Z}\rangle - \mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) : \mathbf{Y} \triangleleft \mathfrak{F}, \mathbf{Y} \in \mathcal{Y}\} \\
&= \inf\{\langle \mathbf{Y}, \mathbb{E}[\mathbf{Z} \,|\, \mathfrak{F}]\rangle - \mathcal{A}(\mathbf{Y} \,|\, \mathfrak{F}) : \mathbf{Y} \triangleleft \mathfrak{F}, \mathbf{Y} \in \mathcal{Y}\} \\
&= \mathcal{A}^+(\mathbb{E}[\mathbf{Z} \,|\, \mathfrak{F}]).
\end{aligned}$$

□

References

[1] B. Acciaio, H. Föllmer and I. Penner, Risk assessment for uncertain cash flows: Model ambiguity, discounting ambiguity, and the role of bubbles, *Finance Stoch.* **16** (2012) 669–709.

[2] P. Artzner, F. Delbaen, D. Heath and H. Ku, Coherent multiperiod risk adjusted values and Bellman's principle, *Annals of Operations Research* **152**(1) (2007) 5–22.

[3] C. Benett and R. Shapley, *Interpolation of Operators* (Academic Press, Orlando, 1988).

[4] T. Breuer and I. Csiszár, Systematic stress tests with entropic plausibility constraints, *J. Bank. Finance* **37**(5) (2003) 1552–1559.

[5] T. Carleman, *Les Fonctions Quasi-Analytiques*, Collection Borel (Gauthier-Villars, Paris, 1926).
[6] P. Cheridito, F. Delbaen and M. Kupper, Coherent and convex monetary risk measures for unbounded cdlg processes, *Finance Stoch.* **9**(3) (2005) 369–387.
[7] P. Cheridito and M. Kupper, Composition of time-consistent dynamic monetary risk measures in discrete time, *Int. J. Theor. Appl. Finance* **14**(1) (2011) 137–162.
[8] D. Dentcheva and A. Ruszczynski, Convexification of stochastic ordering, *C. R. Acad. Bulgare Sci.* **57**(4) (2004) 11–16.
[9] K. Detlefsen and G. Scandolo, Conditional and dynamic convex risk measures, *Finance Stoch.* **9** (2005) 539–561.
[10] T. Ferguson, *Mathematical Statistics*: *A Decision Theoretic Approach* (Academic Press, New York, 1967).
[11] D. Filipovic, M. Kupper and N. Vogelpoth, Approaches to conditional risk, Working Paper 671, Swiss Finance Institute Research Paper No. 11-02 (2011). Available at: http://ssrn.com/abstract=1752851 or http://dx.doi.org/10.2139/ssrn.1752851.
[12] H. Föllmer and A. Schied, Convex measures of risk and trading constraints, *Finance Stoch.* **6** (2002) 429–447.
[13] M. Fritelli and G. E. Rosazza, Putting order in risk measures, *J. Bank. Finance* **26** (2002) 1473–1486.
[14] R. Hassin and M. Haviv, *To Queue or not to Queue: Equilibrium Behavior in Queuing Systems* (Springer, New York, 2003).
[15] J. Hirshleifer and J. Riley, *The Analytics of Uncertainty and Information* (Cambridge University Press, Cambridge, UK, 1992).
[16] E. Jouini, W. Schachermayer and N. Touzi, Optimal risk sharing for law invariant monetary utility functions, *Math. Finance* **18**(2) (2008) 269–292.
[17] R. Kovacevic, Conditional risk and acceptability mappings as banach-lattice valued mappings, *Statistics and Risk Modeling* **29**(1) (2012) 1–18.
[18] R. M. Kovacevic, Conditional acceptability mappings: Convex analysis in banach lattices, Dissertation, University of Vienna (2008).
[19] R. M. Kovacevic, Maximum-loss, minimum-win and the Esscher pricing principle, *IMA Journal of Management Mathematics* **23**(4) (2011) 325–340.
[20] R. M. Kovacevic and G. C. Pflug, Time consistency and information monotonicity of multiperiod acceptability functionals, in *Advanced Financial Modelling*, Radon Series on Computational and Applied Mathematics, eds. H. Albrecher, W. Runggaldier and W. Schachermayer, Vol. 8 (de Gruyter, Berlin, 2009), pp. 347–370.
[21] J. Krawczyk and G. Karacaoglu, On repetitive control and the behaviour of a middle-aged consumer, *Eur. J. Oper. Res.* **66**(1) (1993) 89–99.
[22] S. S. Kutadeladze, Convex operators, *Uspekhi Mat. Nauk* **34**(1) (1979) 167–196.
[23] X. Li, X. Y. Zhou and A. E. B. Lim, Dynamic mean-variance portfolio selection with no-shorting constraints, *SIAM Journal of Control and Optimization* **40**(5) (2002) 1540–1555.
[24] M. A. Mansour, C. Malivert and M. Théra, Semicontinuity of vector-valued mappings, *Optimization* **56**(1) (2007) 241–252.
[25] N. S. Papageorgiou, Nonsmooth analysis on partially ordered vector spaces: Part 1 — Convex case, *Pacific J. Math.* **107**(2) (1983) 403–458.
[26] G. Pflug and W. Rmisch, *Modeling, Measuring and Managing Risk* (World Scientific, Singapore, 2007).
[27] G. Pflug and A. Ruszczyński, Measuring risk for income streams, *Comput. Optim. Appl.* **32** (2005) 161–178.
[28] G. C. Pflug, On distortion functionals, *Statist. Decisions* **24** (2006) 45–60.

[29] H. Raiffa and R. Schlaifer, *Applied Statistical Decision Theory* (Harvard University Press, 1961).
[30] F. Riedel, Dynamic coherent risk measures, *Stochastic Process. Appl.* **112** (2004) 185–200.
[31] B. Roorda and J. Schumacher, Time consistency conditions for acceptability measures, with an application to tail value at risk, *Insurance. Math. Econom.* **40** (2007) 209–230.
[32] A. Ruszczyński and A. Shapiro, Conditional risk mappings, *Math. Oper. Res.* **31**(3) (2006) 544–561.
[33] W. A. Spurr and C. Bonini, *Statistical Analysis of Business Decisions* (Richard D. Irwin Publishing House, Homewood, 1973).
[34] J. E. Stiglitz, Information and economic analysis, in *Current Economic Problems*, eds. M. Parkin and A. R. Nobay (Cambridge University Press, Cambridge, UK, 1975).
[35] S. Weber, Distribution-invariant dynamic risk measures, information and dynamic consistency, *Math. Finance* **16** (2006) 419–414.
[36] X. Zhou and D. Li, Continuous time mean-variance portfolio selection: A stochastic lq framework, *Applied Mathematical Optimization* **42** (2000) 19–33.