Time and Dynamic Consistency of Risk Averse Stochastic Programs

Alexander Shapiro · Alois Pichler

October 2016

Abstract In various settings time consistency in dynamic programming has been addressed by many authors going all the way back to original developments by Richard Bellman. The basic idea of the involved dynamic principle is that a policy designed at the first stage, before observing realizations of the random data, should not be changed at the later stages of the decision process. This is a rather vague principle since this leaves a choice of optimality criteria at every stage of the process conditional on an observed realization of the random data. In this paper we discuss this from the point of view of modern theory of risk averse stochastic programming. In particular we discuss time consistent decision making by addressing risk measures which are recursive, nested, dynamically or time consistent. It turns out that the paradigm of time consistency is in conflict with various desirable, classical properties of general risk measures.

Keywords Risk averse stochastic programming, coherent risk measures, time consistency, dynamic programming

1 Introduction

A famous citation of Richard Bellman (1957): *"An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."* In other words this principle of optimality postulates that an optimal policy computed at the initial stage of the decision process, before any realization of the uncertainty data became available, remains optimal at the later stages. This formulation is quite vague since it is not clearly specified what optimality at the later stages does mean. In some situations this comes naturally and implicitly assumed. However, in more complex cases this could lead to a confusion and misunderstandings.

In this paper we discuss time consistency of multistage stochastic programs with relation to the modern theory of risk measures. Basically two main approaches were considered in the recent literature. In one approach optimality at every stage is defined by an objective function given by a conditional risk measure satisfying some basic assumptions. The other approach deals with optimal policies in a direct way and is closely related to Bellman's principle of optimality. One of the goals of this paper is to clarify a relation between these two approaches to time consistency. There is a large literature on time consistency dating years back. Rather than trying to give a

Research of the first author was partly supported by NSF grant 1633196 and DARPA EQUiPS program grant SNL 014150709.

A. Shapiro
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205, USA

A. Pichler
Fakultät für Mathematik
Technische Universität Chemnitz
D–09107 Chemnitz, Germany

comprehensive survey of this literature we aim at presenting a certain point of view. For a comprehensive literature review we can refer to the recent paper Bielecki et al. (2016).

Outline. We introduce the main terms, including time consistency, in the following Section 2. The main discussion on dynamic consistency is split in two parts. A special (additive) case is considered in Section 3 and a more general case in Section 4. Section 5 addresses the dynamic programming equations. Section 6 investigates essential properties of coherent risk measures related to time consistency. Section 7 discusses time consistency adjusted to optimal policies. For this case we derive special time consistent relations, while Section 8 concludes.

2 Basic formulation

We will use the following framework. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathfrak{F} := (\mathcal{F}_1, \dots \mathcal{F}_T)$ be a filtration (a sequence of increasing sigma algebras, $\mathcal{F}_1 \subset \dots \subset \mathcal{F}_T$) with[1] $\mathcal{F}_1 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. Let $\mathcal{Z}_1 \subset \dots \subset \mathcal{Z}_T$ be a sequence of linear spaces of measurable functions (random variables). We assume that $\mathcal{Z}_t := L_p(\Omega, \mathcal{F}_t, P)$ for some $p \in [1, \infty]$, although more general settings are possible. Since $\mathcal{F}_1$ is trivial, the space $\mathcal{Z}_1$ consists of constant functions and will be identified with $\mathbb{R}$. For $Z, Z' \in \mathcal{Z}_T$ we write $Z \succeq Z'$ to denote that $Z(\omega) \geq Z'(\omega)$ for a.e. $\omega \in \Omega$, and $Z \succ Z'$ to denote that $Z \succeq Z'$ and $P(Z > Z') > 0$. By $\mathbb{E}_{|\mathcal{F}_t}[\,\cdot\,]$ or $\mathbb{E}[\cdot|\mathcal{F}_t]$ we denote the conditional expectation with respect to $\mathcal{F}_t$. Note that since for $\mathcal{F}_s \subset \mathcal{F}_t$ for $s < t$, it follows that $\mathbb{E}_{|\mathcal{F}_s}[\mathbb{E}_{|\mathcal{F}_t}(\cdot)] = \mathbb{E}_{|\mathcal{F}_s}(\cdot)$; and since $\mathcal{F}_1$ is trivial, $\mathbb{E}_{|\mathcal{F}_1}$ is the corresponding unconditional expectation.

We denote by Z_t an element of the space $\mathcal{Z}_t$. Note that $Z_t : \Omega \to \mathbb{R}$ is $\mathcal{F}_t$ measurable and hence the process $Z_1, ..., Z_T$ is adapted to the filtration $\mathfrak{F}$. Preference relations between possible realizations of random data will be defined by a family of mappings

$$\mathcal{R}_{s,t} : \mathcal{Z}_{s,t} \to \mathcal{Z}_s, \quad 1 \leq s < t \leq T, \tag{2.1}$$

where $\mathcal{Z}_{s,t} := \mathcal{Z}_s \times \dots \times \mathcal{Z}_t$. We refer to each $\mathcal{R}_{s,t}$ as a *preference mapping* and to the family $\mathfrak{R} = \{\mathcal{R}_{s,t}\}_{1 \leq s < t \leq T}$ as a *preference system*. Sometimes we only deal with preference mappings $\mathcal{R}_{t,T} : \mathcal{Z}_{t,T} \to \mathcal{Z}_t$ and the corresponding preference system $\mathfrak{R} = \{\mathcal{R}_{t,T}\}_{t=1}^{T-1}$. Since $\mathcal{Z}_1$ is identified with $\mathbb{R}$, we view $\mathcal{R}_{1,T}(Z_1, ..., Z_T)$ as a real number for any $(Z_1, ..., Z_T) \in \mathcal{Z}_{1,T}$.

At the first stage for $\mathcal{R} = \mathcal{R}_{1,T}$ we consider the optimization problem

$$\begin{aligned} &\text{Min } \mathcal{R}\big[f_1(x_1), \dots, f_T(x_T, \omega)\big], \\ &\text{s.t. } \ x_1 \in \mathcal{X}_1, \ x_t \in \mathcal{X}_t(x_{t-1}, \omega), \ t = 2, ..., T, \end{aligned} \tag{2.2}$$

called the *reference* problem, where $f_1 : \mathbb{R}^{n_1} \to \mathbb{R}$, $\mathcal{X}_1 \subset \mathbb{R}^{n_1}$, $f_t : \mathbb{R}^{n_t} \times \Omega \to \mathbb{R}$ and $\mathcal{X}_t : \mathbb{R}^{n_{t-1}} \times \Omega \rightrightarrows \mathbb{R}^{n_t}$, $t = 2, ..., T$. It is assumed that $f_t(x_t, \cdot)$ and $\mathcal{X}_t(x_{t-1}, \cdot)$ are $\mathcal{F}_t$-measurable. A sequence $x_t : \Omega \to \mathbb{R}^{n_t}$, $t = 1, ..., T$, of mappings adapted to[2] $\mathfrak{F} = \{\mathcal{F}_t\}_{t=1}^T$ is called a policy. Since $\mathcal{F}_1$ is trivial, the first decision x_1 is deterministic. The policy $\pi = (x_1, x_2(\omega), ..., x_T(\omega))$ is feasible if it satisfies the feasibility constraints with probability 1. We denote by Π the set of feasible policies and such that $(f_1(x_1), f_2(x_2(\omega), \omega), \dots, f_T(x_T(\omega), \omega)) \in \mathcal{Z}_{1,T}$. Optimization in (2.2) is performed over $\pi \in \Pi$.

Remark 1 The constraints of problem (2.2) are formulated in such way that the t-th component x_t of a feasible policy depends on x_{t-1}, but not explicitly on all previous decisions $x_1, \dots, x_{t-1}$. This *Markovian* formulation is not a restriction, as one may incorporate the history of the decision process $(x_1, \dots, x_{t-1})$ in the formulation of the problem (2.2) instead of x_{t-1} (cf., e.g., Pflug and Römisch (2007, p.120)).

Definition 1 We say that an optimal policy[3] $(\bar{x}_1, ..., \bar{x}_T)$ of the reference problem (2.2) is *time consistent* if at stages $t = 2, ..., T$ the policy $(\bar{x}_t, ..., \bar{x}_T)$ is optimal for

$$\begin{aligned} &\text{Min } \mathcal{R}_{t,T}\big[f_t(x_t, \omega), \dots, f_T(x_T, \omega)\big], \\ &\text{s.t. } \ x_\tau \in \mathcal{X}_t(x_{\tau-1}, \omega), \ \tau = t, ..., T, \end{aligned} \tag{2.3}$$

conditional on $\mathcal{F}_t$ and $\bar{x}_{t-1}$.

[1] The sigma algebra consisting of only two sets, the empty set and the whole space Ω, is called trivial.
[2] It is said that the sequence is adapted to the filtration if $x_t(\cdot)$ is $\mathcal{F}_t$-measurable, $t = 1, ..., T$.
[3] When there is no ambiguity we write $\bar{x}_t = \bar{x}_t(\omega)$ for the corresponding random vectors.

Remark 2 The above formulation in terms of filtration could be convenient from a mathematical point of view but is not very intuitive. An alternative, more intuitive, approach is to present the problem in terms of a data process viewed as a scenario tree. That is, at stage $t = 1$ we have one root node denoted ξ_1. At stage $t = 2$ we have as many nodes (children nodes of the root node) as many different realizations of data may occur. Each of them is connected with the root node by an arc. A generic node at time $t = 2$ is denoted ξ_2, etc. at the later stages. A scenario, representing a realization (sample path or trajectory) of the data process, is a sequence $\xi_1, ..., \xi_T$ of nodes. A particular node of the tree represents the history of the data process up to this node. In case the number of children nodes of every node is finite, the tree is finite. However, one may also think about a process with an infinite (continuum) number of scenarios. Equipped with an appropriate probabilistic structure, $\xi_1, ..., \xi_T$ becomes a random (stochastic) process. With some abuse of the notation we use the same notation $\xi_1, ..., \xi_T$ for this random process and its particular realization (sample path), the exact meaning will be clear from the context. We use the notation $\xi_{[t]} := (\xi_1, ..., \xi_t)$ for the history of the process up to time $t = 1, ..., T$. For a more detailed discussion of such construction and a connection with the filtration approach we may refer to Pflug and Römisch (2007, Section 3.1.1) and Shapiro et al. (2014, Section 3.1). We can view the problem (2.3) at stage t as conditional on a realization of the data process, up to time t, and our decision $\bar{x}_{t-1}$. This is the meaning of "conditional on $\mathcal{F}_t$ and $\bar{x}_{t-1}$" in this framework.[4]

Definition 1 formalizes the meaning of optimality of a solution of the reference problem at the later stages of the decision process. Clearly this framework depends on a choice of the mappings $\mathcal{R}_{t,T}$ defining the respective conditional objective functions. This suggests the following basic questions: (i) what would be a 'natural' choice of mappings $\mathcal{R}_{t,T}$, (ii) what properties of mappings $\mathcal{R}_{t,T}$ are sufficient/ necessary to ensure that every (at least one) optimal solution of the reference problem is time consistent, (iii) how time consistency is related to dynamic programming equations.

It could be noted that we allow for the preference mapping $\mathcal{R}_{t,T}$, $t = 1, ..., T-1$, to depend on realizations of the data process up to time t, i.e., we assume that $\mathcal{R}_{t,T}(Z_1, ..., Z_t)$ is $\mathcal{F}_t$-measurable. However, we do not allow $\mathcal{R}_{t,T}$ to depend on our decisions. The above framework (2.2)–(2.3) is very general. In the following sections we consider some particular cases which on one hand are sufficiently general and on the other hand amendable for the analysis.

3 The additive case

In this section we assume that the objective of the reference problem is a function of the total cost. This may be the most common case considered in applications. That is, we consider preference system $\{\mathcal{R}_{t,T}\}_{t=1}^{T-1}$ with each $\mathcal{R}_{t,T}$ representable as a function of $Z_1 + ... + Z_t$, i.e., $\mathcal{R}_{t,T}(Z_1, ..., Z_t) = \rho_{t,T}(Z_1 + ... + Z_t)$ for some $\rho_{t,T} : \mathcal{Z}_T \to \mathcal{Z}_t$, $t = 1, ..., T-1$. The reference problem can be written here as

$$\begin{array}{ll} \text{Min } \varrho\big[f_1(x_1) + f_2(x_2, \omega) + \cdots + f_T(x_T, \omega)\big], \\ \text{s.t. } \ x_1 \in \mathcal{X}_1, \ x_t \in \mathcal{X}_t(x_{t-1}, \omega), \ t = 2, ..., T, \end{array} \tag{3.1}$$

with $\varrho = \rho_{1,T}$. The corresponding conditional problems (2.3) take here the form

$$\begin{array}{ll} \text{Min } \rho_{t,T}\big[f_t(x_t, \omega) + \cdots + f_T(x_T, \omega)\big], \\ \text{s.t. } \ x_\tau \in \mathcal{X}_t(x_{\tau-1}, \omega), \ \tau = t, ..., T. \end{array} \tag{3.2}$$

We refer to the following properties of preference mappings $\rho_{t,T}$, $t = 1, ..., T-1$.

(i) *Convexity*:
$$\rho_{t,T}(\alpha Z + (1-\alpha)Z') \preceq \alpha\rho_t(Z) + (1-\alpha)\rho_t(Z'),$$
for any $Z, Z' \in \mathcal{Z}_T$ and $\alpha \in [0, 1]$.

(ii) *Monotonicity*: if $Z, Z' \in \mathcal{Z}_T$ and $Z \succeq Z'$, then $\rho_{t,T}(Z) \succeq \rho_{t,T}(Z')$.

(iii) *Translation equivariance*: if $Z \in \mathcal{Z}_T$ and $Y \in \mathcal{Z}_t$, then
$$\rho_{t,T}(Z + Y) = \rho_{t,T}(Z) + Y. \tag{3.3}$$

[4] Note that for the considered Markovian formulation (2.2) of the reference problem, we need to condition only on the last decision in the sequence of decisions $\bar{x}_1, ..., \bar{x}_{t-1}$.

(iv) *Positive homogeneity*: if $\alpha \geq 0$ and $Z \in \mathcal{Z}_T$, then $\rho_{t,T}(\alpha Z) = \alpha \rho_{t,T}(Z)$.

It is said that mappings $\rho_{t,T}$ are *coherent* if they satisfy the above properties of convexity, monotonicity, translation equivariance and positive homogeneity (cf. Artzner et al. (1999)).

One choice of the reference functional is to take[5] $\varrho(\cdot) := \mathbb{E}(\cdot)$. In that case problem (3.1) is said to be risk neutral. It is natural then to define $\rho_{t,T}$ as the conditional expectation $\mathbb{E}_{|\mathcal{F}_t}$. Note that $\mathbb{E}_{|\mathcal{F}_1} = \mathbb{E}$ since $\mathcal{F}_1$ is trivial. Another example is the essential supremum operator $\varrho(Z) = \operatorname{ess\,sup}(Z)$, $Z \in L_\infty(\Omega, \mathcal{F}, P)$, with the corresponding preference mappings $\rho_{t,T}$ given by the respective conditional essential supremum[6] operators. We refer to this preference system as *max-type* preference system.

The following concept of dynamic consistency (also called time consistency by some authors) for preferences in slightly different forms was used by several authors (see Artzner et al. (2007), Cheridito et al. (2006), Riedel (2004), Ruszczyński (2010), Wang (1999) and references therein).

Definition 2 The preference system $\{\rho_{t,T}\}_{t=1}^{T-1}$ is said to be *dynamically consistent* if the implication

$$Z, Z' \in \mathcal{Z}_T \text{ and } \rho_{t,T}(Z) \succeq \rho_{t,T}(Z') \implies \rho_{s,T}(Z) \succeq \rho_{s,T}(Z'). \tag{3.4}$$

holds for all $1 \leq s < t \leq T-1$,

It turns out that the above 'forward' property of dynamic consistency is not always sufficient to ensure that every optimal policy is time consistent (see Example 1 below). We will need a stronger notion of dynamic consistency.

Definition 3 A dynamically consistent preference system $\{\rho_{t,T}\}_{t=1}^{T-1}$ is said to be *strictly dynamically consistent* if the implication

$$Z, Z' \in \mathcal{Z}_T \text{ and } \rho_{t,T}(Z) \succ \rho_{t,T}(Z') \implies \rho_{s,T}(Z) \succ \rho_{s,T}(Z') \tag{3.5}$$

holds for all $1 \leq s < t \leq T-1$ in addition to (3.4).

If $\rho_{t,T} := \mathbb{E}_{|\mathcal{F}_t}$, then for $s < t$ we have that $\rho_{s,T}(Z) = \mathbb{E}_{|\mathcal{F}_s}[\mathbb{E}_{|\mathcal{F}_t}(Z)]$ and hence this preference system is dynamically consistent. In fact it is not difficult to see that this preference system is strictly dynamically consistent. On the other hand the max-type preference system (given by the essential supremum mappings) is dynamically consistent but is not strictly dynamically consistent. We have the following results showing a relation of dynamic consistency of the preference system and time consistency of optimal policies (cf. Shapiro et al. (2014, Proposition 6.80)).

Proposition 1 *The following holds.* (i) *If the preference system is dynamically consistent and $\bar{\pi} \in \Pi$ is the unique optimal solution of the reference problem* (3.1)*, then $\bar{\pi}$ is time consistent.* (ii) *If the preference system is strictly dynamically consistent, then every optimal solution of the reference problem is time consistent.*

Proof We argue by a contradiction. Let $\bar{\pi} = (\bar{x}_1, ..., \bar{x}_T)$ be an optimal solution of the reference problem. Suppose that for some $t \geq 2$, conditional on $\mathcal{F}_t$ and $\bar{x}_{t-1}$ the policy $(\bar{x}_t, ..., \bar{x}_T)$ is not optimal for problem (3.2). That is, there exists a different from $(\bar{x}_t, ..., \bar{x}_T)$ policy $(\tilde{x}_t, ..., \tilde{x}_T)$ satisfying the feasibility constraints, conditional on $\mathcal{F}_t$ and $\bar{x}_{t-1}$, such that $\rho_{t,T}(\bar{Z}) \succ \rho_{t,T}(\tilde{Z})$, where

$$\bar{Z}(\omega) := f_t(\bar{x}_t, \omega) + \cdots + f_T(\bar{x}_T, \omega) \text{ and } \tilde{Z}(\omega) := f_t(\tilde{x}_t, \omega) + \cdots + f_T(\tilde{x}_T, \omega).$$

Consider the policy $\pi^* := (\bar{x}_1, ..., \bar{x}_{t-1}, \tilde{x}_t, ..., \tilde{x}_T)$. This policy is feasible for the reference problem (3.1) and by (3.4) it follows that $\rho_{2,T}(\bar{Z}) \succeq \rho_{2,T}(\tilde{Z})$. Hence the policy π^* is optimal for the reference problem, and it is different from policy $\bar{\pi}$. In case the optimal solution $\bar{\pi}$ is unique, this gives the desired contradiction and completes the proof of assertion (i).

Suppose further that the preference system is strictly dynamically consistent. Then it follows from $\rho_{t,T}(\bar{Z}) \succ \rho_{t,T}(\tilde{Z})$ that $\rho_{2,T}(\bar{Z}) > \rho_{2,T}(\tilde{Z})$. This contradicts optimality of $\bar{\pi}$, and completes proof of assertion (ii). □

[5] Unless stated otherwise the expectation operators are taken with respect to the reference probability measure P.

[6] For a mathematically rigorous introduction of the essential supremum, as well as the essential infimum, we refer to Karatzas and Shreve (1998, Appendix A).

Although dynamic consistency of the preference system is not sufficient to ensure time consistency of every optimal policy, under mild regularity conditions it ensures existence of at least one time consistent policy. More precisely we have the following result (cf. Shapiro et al. (2014, Proposition 6.83)). We give a proof for the sake of completeness.

Proposition 2 *Suppose that:* (i) *reference problem* (3.1) *has a nonempty set of optimal solutions,* (ii) *for every optimal policy* $\bar{\pi} = (\bar{x}_1, ..., \bar{x}_T)$ *and every* $t \geq 2$*, conditional on* $\mathcal{F}_t$ *and* $\bar{x}_{t-1}$*, problem* (3.2) *has an optimal solution,* (iii) *the preference system is dynamically consistent,* (iv) *the mappings* $\rho_{t,T}$*,* $t = 1, ..., T-1$*, are translation equivariant. Then at least one of the optimal solutions of the reference problem is time consistent.*

Proof Let $(\bar{x}_1^1, ..., \bar{x}_T^1)$ be an optimal solution of the reference problem (3.1). Conditional on $\bar{x}_1^1$ let $(\bar{x}_2^2, ..., \bar{x}_T^2)$ be an optimal solution of problem (3.2) for $t = 2$, such solution exists by the assumption (ii). Since $\varrho = \rho_{1,T}$ it follows that the policy $(\bar{x}_1^1, \bar{x}_2^2, ..., \bar{x}_T^2)$ is optimal for the reference problem. For $t = 3$ let $(\bar{x}_3^3, ..., \bar{x}_T^3)$ be an optimal solution of problem (3.2) conditional on $\bar{x}_1^1$ and $\bar{x}_2^2$. By optimality

$$\rho_{3,T}\big[f_3(\bar{x}_3^2, \omega) \cdots + f_T(\bar{x}_T^2, \omega)\big] \succeq \rho_{3,T}\big[f_3(\bar{x}_3^3, \omega) \cdots + f_T(\bar{x}_T^3, \omega)\big].$$

By assumption (iv) it follows that

$$\rho_{3,T}\big[f_2(\bar{x}_2^2, \omega) + f_3(\bar{x}_3^2, \omega) \cdots + f_T(\bar{x}_T^2, \omega)\big] \succeq \rho_{3,T}\big[f_2(\bar{x}_2^2, \omega) + f_3(\bar{x}_3^3, \omega) \cdots + f_T(\bar{x}_T^3, \omega)\big],$$

and hence by assumption (iii)

$$\rho_{2,T}\big[f_2(\bar{x}_2^2, \omega) + f_3(\bar{x}_3^2, \omega) \cdots + f_T(\bar{x}_T^2, \omega) \succeq \rho_{2,T}\big[f_2(\bar{x}_2^2, \omega) + f_3(\bar{x}_3^3, \omega) \cdots + f_T(\bar{x}_T^3, \omega)\big].$$

It follows that policy $(\bar{x}_1^1, \bar{x}_2^2, \bar{x}_3^3, ..., \bar{x}_T^3)$ is optimal for the reference problem. This policy is also optimal for problem (3.2), for $t = 2$ conditional on $\bar{x}_1^1$ and for $t = 3$ conditional on $\bar{x}_1^1$ and $\bar{x}_2^2$. By continuing this procedure we eventually construct the required optimal policy which is time consistent. □

Let us introduce now the following concepts.

Definition 4 It is said that $\rho_{t,T}$ is *strictly monotone* if $\rho_{t,T}$ is monotone and $Z, Z' \in \mathcal{Z}_T$ and $Z \succ Z'$ implies that $\rho_{t,T}(Z) \succ \rho_{t,T}(Z')$.

Definition 5 It is said that the preference system is *recursive* if for any $Z \in \mathcal{Z}_T$ and $1 \leq s < t \leq T-1$ it follows that

$$\rho_{s,T}(\rho_{t,T}(Z)) = \rho_{s,T}(Z). \tag{3.6}$$

Proposition 3 (i) *Suppose that preference mappings* $\rho_{t,T}$*,* $1 \leq t \leq T-1$*, are monotone (strictly monotone) and the preference system is recursive. Then the preference system is (strictly) dynamically consistent.* (ii) *Conversely, if the preference system is dynamically consistent, translation equivariant and* $\rho_{t,T}(0) = 0$*,* $t = 1, ..., T-1$*, then the preference system is recursive.*

Proof (i) Let $Z, Z' \in \mathcal{Z}_T$ be such that $\rho_{t,T}(Z) \succeq \rho_{t,T}(Z')$. Then by using (3.6) and monotonicity of preference functionals we have

$$\rho_{s,T}(Z) = \rho_{s,T}\big(\rho_{t,T}(Z)\big) \succeq \rho_{s,T}\big(\rho_{t,T}(Z')\big) = \rho_{s,T}(Z').$$

This shows that the preference system is dynamically consistent. Strict dynamic consistency can be shown by the same arguments using strict monotonicity of $\rho_{t,T}$.

(ii) Consider $1 \leq s < t \leq T-1$, $Z \in \mathcal{Z}_T$ and $Z' := \rho_{t,T}(Z)$. Since $Z' \in \mathcal{Z}_t$ and $\rho_{t,T}(0) = 0$ we have by (3.3) that

$$\rho_{t,T}(Z') = \rho_{t,T}(0 + Z') = \rho_{t,T}(0) + Z' = Z' = \rho_{t,T}(Z).$$

By (3.4) this implies $\rho_{s,T}(Z) = \rho_{s,T}(Z')$, that is equation (3.6) follows. This completes the proof. □

Part (ii) of the above proposition is due to Artzner et al. (2007, Theorem 5.1). The condition $\rho_{t,T}(0) = 0$, $t = 1, ..., T-1$, can be viewed as *normalization* of the preference mappings.

Recall that in the risk neutral case the preference mappings are given by conditional expectations $\rho_{t,T}(\cdot) = \mathbb{E}_{|\mathcal{F}_t}(\cdot)$. These mappings are decomposable in the sense that they can be represented as compositions of one-step mappings. That is, $\rho_{t,T}(\cdot) = \rho_{t+1}(\cdots \rho_T(\cdot))$, where $\rho_\tau := \mathbb{E}_{|\mathcal{F}_{\tau-1}}$.

Definition 6 It is said that preference mappings $\rho_{t,T}$ are *decomposable* via a family of one-step mappings $\rho_t : \mathcal{Z}_t \to \mathcal{Z}_{t-1}$, if they can be represented as compositions

$$\rho_{t,T}(\cdot) = \rho_{t+1}(\cdots \rho_T(\cdot)),$$

denoted $\rho_{t,T} = \rho_{t+1} \circ \cdots \circ \rho_T$, $t = 1, ..., T-1$.

Decomposable mappings $\rho_{t,T}$ inherit various properties of the corresponding one-step mappings ρ_t. In particular if every ρ_t is monotone (strictly monotone), then mappings $\rho_{t,T}$ are monotone (strictly monotone). If every ρ_t is monotone and convex, then $\rho_{t,T}$ are convex. If every ρ_t is translation equivariant, then $\rho_{t,T}$ are translation equivariant. If every ρ_t is positively homogeneous, then $\rho_{t,T}$ are positively homogeneous. That is, if every mapping ρ_t is coherent (i.e., satisfies the conditions of convexity, monotonicity, translation equivariance and positive homogeneity), then mappings $\rho_{t,T}$ are coherent.

Economic interpretation of decomposable risk measures. Suppose that $\rho_{1,T}$ is decomposable and the corresponding one-step mappings ρ_t are translation equivariant. Consider $Z = Z_1 + ... + Z_T$, where $Z_t \in \mathcal{Z}_t$ and hence is $\mathcal{F}_t$-measurable, $t = 1, ..., T$. Recall that $\mathcal{Z}_1$ is identified with $\mathbb{R}$ so that $Z_1 \in \mathbb{R}$. Then

$$\rho_{1,T}(Z) = \rho_2\big(\cdots \rho_T(Z_1 + ... + Z_T)\big) = Z_1 + \rho_2\big(Z_2 + \cdots + \rho_T(Z_T)\big). \tag{3.7}$$

The quantity (3.7) obeys the economic interpretation of an insurance premium. Indeed, associate each component Z_t of the total payoff Z with the outcome which appears at stage t. The first outcome Z_1 is certain, so it does not have to be insured (and appears linearly in (3.7)). The outcome Z_2 is random, and accepting for ρ_2 the interpretation of an insurance premium reveals that $\rho_2(Z_2)$ is the corresponding price. The following premium $\rho_3(Z_3)$ is the premium conditional on the observations Z_1 and Z_2. More generally, $\rho_t(Z_t)$ is the price for insurance, which is concluded at stage $t-1$ for the subsequent period until t. The total quantity (3.7) thus represents the total of all premiums, where insurance is bought at each stage to protect the following period.

Proposition 4 *Suppose that the preference mappings $\rho_{t,T}$, $t = 1, ..., T-1$, are decomposable with the corresponding one-step mappings $\rho_t : \mathcal{Z}_t \to \mathcal{Z}_{t-1}$. Then the following holds.* (i) *If the mappings ρ_t, $t = 2, ..., T$, are monotone (strictly monotone), then the preference system is (strictly) dynamically consistent.* (ii) *If the mappings ρ_t are translation equivariant and $\rho_t(0) = 0$, $t = 2, ..., T$, then the preference system is recursive.*

Proof (i) By the decomposability we have that

$$\rho_{s,T}(Z) = \rho_s \circ \cdots \circ \rho_{t-1}\big(\rho_{t,T}(Z)\big). \tag{3.8}$$

Then one can proceed in the same way as in the proof of Proposition 3 (i).

(ii) Note that because of the translation equivariance and since $\rho_t(0) = 0$, it follows for any $s < t$ and $Z \in \mathcal{Z}_s$ that $\rho_t(Z) = \rho_t(0 + Z) = \rho_t(0) + Z = Z$. Then for $Z \in \mathcal{Z}_T$ and $Z' := \rho_{t,T}(Z) \in \mathcal{Z}_t$ we have

$$\rho_{s,T}(Z') = \rho_s \circ \cdots \circ \rho_T(Z') = \rho_s \circ \cdots \circ \rho_{t-1}(Z') = \rho_s \circ \cdots \circ \rho_{t-1}\big(\rho_{t,T}(Z)\big) = \rho_{s,T}(Z).$$

This gives the required equation (3.6). □

Suppose now that the preference system is recursive. By the recursiveness we have

$$\rho_{t,T}(Z) = \rho_{t,T}\big(\rho_{t+1,T}(\cdots \rho_{T-1,T}(Z))\big), \tag{3.9}$$

which gives a decomposition of $\rho_{t,T}$ with the associate one-step mappings $\rho_\tau := \rho_{\tau-1,T}$ considered as mappings from $\mathcal{Z}_\tau$ to $\mathcal{Z}_{\tau-1}$. By Proposition 3 (ii) we have that if the preference system is dynamically consistent, translation equivariant and normalized, then it is recursive and hence decomposable in the sense of equation (3.9).

By combining Propositions 1, 2, 3 and 4 (i) we can formulate the main result of this section.

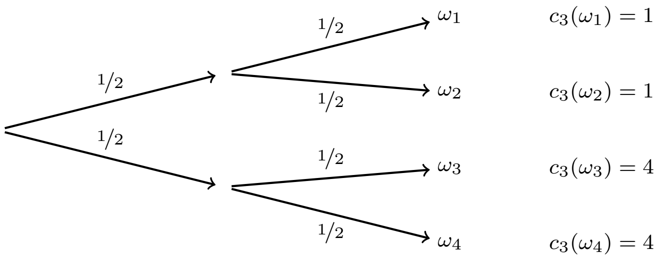

Fig. 1: The decision tree for the non-time consistent problem displayed in Example 1

Theorem 1 *Suppose that either the preference system is recursive and the preference mappings $\rho_{t,T}$, $t = 1, ..., T-1$, are monotone or the preference mappings are decomposable and the corresponding one-step mappings ρ_t are monotone. Then the following holds.* (i) *If $\bar{\pi} \in \Pi$ is the unique optimal solution of the reference problem (3.1), then $\bar{\pi}$ is time consistent.* (ii) *If moreover the preference mappings $\rho_{t,T}$ are strictly monotone (one-step mappings ρ_t are strictly monotone), then every optimal solution of the reference problem is time consistent.* (iii) *If the reference problem has a nonempty set of optimal solutions, the mappings $\rho_{t,T}$ (the mappings ρ_t) are translation equivariant and for every optimal policy $\bar{\pi} = \{\bar{x}_1, ..., \bar{x}_T\}$ and every $t \geq 2$, conditional on $\mathcal{F}_t$ and $\bar{x}_{t-1}$ problem (3.2) has an optimal solution, then at least one of optimal solutions of the reference problem is time consistent.*

In the proof of Proposition 1 (ii) that every optimal policy is time consistent, it was essential that the preference system is *strictly* dynamically consistent. In turn by Proposition 3 for recursive preference systems strict dynamic consistency required strict monotonicity of the preference mappings. The max-type preference system is recursive and monotone, but is not strictly monotone and is not strictly dynamically consistent. The following example (taken from Shapiro and Ugurlu (2016)) shows that indeed for a max-type preference system some of optimal policies can be time inconsistent.

Example 1 (Time inconsistency)
Consider a 3-stage optimization problem of the form (3.1) on the tree depicted in Figure 1, with nodes $\omega_1, \omega_2, \omega_3, \omega_4$ at stage $t = 3$, i.e., $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. Assume that decision variables are scalars (one dimensional), cost functions $f_t(x_t, w) = c_t(\omega)x_t$ are linear with zero coefficients ($c_t(\omega) = 0$) at stages $t = 1, 2$ and coefficients

$$c_3(\omega_1) = c_3(\omega_2) = 1, \ c_3(\omega_3) = c_3(\omega_4) = 4$$

at stage $t = 3$, with $\mathcal{X}_1 := \{0\}$, $\mathcal{X}_2(x_1, \omega)$ being identically $\{0\}$ and $\mathcal{X}_3(x_2, \omega)$ being identically the interval $[1, 2]$. Assume further that the preference system is of max-type (since here the probability space is finite, the essential supremum becomes the usual "max" operator). Since the feasible sets at first and second stage are $\{0\}$ and at the third stage it is the interval [1,2], it follows that for any feasible policy the first and second stage decision variables are 0 and that $x_3(\omega) \in [1, 2]$. Therefore we only need to consider the last stage cost $c_3(\omega)x_3(\omega)$ with the corresponding value

$$\max\{x_3(\omega_1), x_3(\omega_2), 4x_3(\omega_3), 4x_3(\omega_4)\}$$

of this policy.

Let us consider conditional problem of the form (3.2) at stage $t = 2$. There are two nodes at stage $t = 2$. Conditional on these nodes and first stage solution (which is fixed to be 0) for a feasible policy problem (3.2) has respective values $\max\{x_3(\omega_1), x_3(\omega_2)\}$ and $\max\{4x_3(\omega_3), 4x_3(\omega_4)\}$. Clearly policy $\bar{\pi}$ with $\bar{x}_3(\omega_i) = 1$, $i = 1, ..., 4$, is optimal for the reference problem with the corresponding optimal value 4. This policy is also conditionally optimal at stage $t = 2$, with respective conditional values of 1 and 4, and hence is time consistent. However the reference problem has many other optimal solutions. Any policy $\tilde{\pi}$ with $\tilde{x}_3(\omega_1), \tilde{x}_3(\omega_2) \in [1, 2]$ and $\tilde{x}_3(\omega_3) = \tilde{x}_3(\omega_4) = 1$, has value 4, and hence is optimal for the reference problem. Such policy has conditional values: $\max\{\tilde{x}_3(\omega_1), \tilde{x}_3(\omega_2)\}$ and 4, conditional on two nodes at stage $t = 2$. Hence policy $\tilde{\pi}$ is not time consistent if $\tilde{x}_3(\omega_1)$ or $\tilde{x}_3(\omega_2)$ is bigger than one. □

4 Dynamic consistency

In Section 3 we considered cases where the objective is a function of the sum $\sum_{t=1}^{T} Z_t$ of the costs $Z_t := f_t(x_t, \omega)$, $t = 1, ..., T$. In this section we outline a somewhat different approach dealing in a more direct way with the formulation (2.2) of the reference problem associated with preference system $\{\mathcal{R}_{t,T}\}_{t=1}^{T-1}$. We deal even with a more general setting of the preference system $\{\mathcal{R}_{s,t}\}_{1\le s<t\le T}$.

Together with the reference problem (2.2) consider the respective conditional problems (2.3). Definitions of monotonicity (strict monotonicity) are modified in an obvious way. The definition of dynamic consistency can be modified in the following way (cf. Cheridito et al. (2006); Ruszczyński (2010)). We also consider *strict* dynamic consistency.

Definition 7 A preference system $\{\mathcal{R}_{s,t}\}_{1\le s<t\le T}$ is said to be *dynamically consistent* if for $1 \le s < t < u \le T$ and $(Z_s, ..., Z_u), (Z'_s, ..., Z'_u) \in \mathcal{Z}_{s,u}$ such that $Z_\tau = Z'_\tau$, $\tau = s, ..., t-1$, the following 'forward' implication holds

$$\mathcal{R}_{t,u}(Z_t, ..., Z_u) \succeq \mathcal{R}_{t,u}(Z'_t, ..., Z'_u) \implies \mathcal{R}_{s,u}(Z_s, ..., Z_u) \succeq \mathcal{R}_{s,u}(Z'_s, ..., Z'_u). \tag{4.1}$$

If moreover the implication

$$\mathcal{R}_{t,u}(Z_t, ..., Z_u) \succ \mathcal{R}_{t,u}(Z'_t, ..., Z'_u) \implies \mathcal{R}_{s,u}(Z_s, ..., Z_u) \succ \mathcal{R}_{s,u}(Z'_s, ..., Z'_u) \tag{4.2}$$

holds, the system is said to be *strictly* dynamically consistent.

Note that it follows from (4.1) that

$$\mathcal{R}_{t,u}(Z_t, ..., Z_u) = \mathcal{R}_{t,u}(Z'_t, ..., Z'_u) \implies \mathcal{R}_{s,u}(Z_s, ..., Z_u) = \mathcal{R}_{s,u}(Z'_s, ..., Z'_u). \tag{4.3}$$

If $\mathcal{R}_{t,T}(Z_t, ..., Z_T)$ is given by $\rho_{t,T}(Z_t + ... + Z_T)$ as in Section 3, then condition (4.1) implies the corresponding condition (3.4) of Definition 2. Conversely, if moreover $\rho_{s,T}$ is translation equivariant, then we can write for $1 \le s < t \le T$,

$$\rho_{t,T}(Z_s + ... + Z_T) = Z_s + ... + Z_{t-1} + \rho_{t,T}(Z_t + ... + Z_T),$$

and hence condition (3.4) implies (4.1). Similar equivalence holds between conditions (4.2) and (3.5) for strict dynamic consistency.

Remark 3 Straightforward analogues of Propositions 1 and 2 hold here with basically the same proofs (cf. Shapiro et al. (2014, Propositions 6.80 and 6.83)).

Definition 8 A preference system $\{\mathcal{R}_{s,t}\}_{1\le s<t\le T}$ is said to be *recursive*, if

$$\mathcal{R}_{s,u}(Z_s, \dots, Z_u) = \mathcal{R}_{s,t}\big(Z_s, \dots, Z_{t-1}, \mathcal{R}_{t,u}(Z_t, \dots, Z_u)\big), \tag{4.4}$$

for $1 \le s < t < u \le T$ and $(Z_s, \dots, Z_u) \in \mathcal{Z}_{s,u}$.

If $\mathcal{R}_{t,T}(Z_s, ..., Z_T) = \rho_{t,T}(Z_t + ... + Z_T)$ and $\rho_{s,T}$ is translation equivariant, then equation (3.6) can be written as

$$\rho_{s,T}(Z_s + ... + Z_T) = \rho_{s,T}(Z_s + ... + Z_{t-1} + \rho_{t,T}(Z_t + ... + Z_T)),$$

which is equivalent to (4.4) for $u = T$.

Remark 4 A somewhat different approach was used in Ruszczyński (2010). Starting with a preference system $\{\mathcal{R}_{t,T}\}_{t=1}^{T-1}$, preference mappings $\mathcal{R}_{s,t}$, $1 \le s < t \le T$, were defined as

$$\mathcal{R}_{s,t}(Z_s, ..., Z_t) := \mathcal{R}_{s,T}(Z_s, ..., Z_t, 0, ..., 0). \tag{4.5}$$

The recursive relation (4.4) then takes the form

$$\mathcal{R}_{s,u}(Z_s, ..., Z_u) = \mathcal{R}_{s,u}(Z_s, ..., Z_{t-1}, \mathcal{R}_{t,u}(Z_t, ..., Z_u), 0, ..., 0)), \ 1 \le t < u \le T. \tag{4.6}$$

In the additive case, discussed in Section 3, we can define

$$\mathcal{R}_{s,t}(Z_s, ..., Z_t) := \rho_{s,T}(Z_s + ... + Z_t), \tag{4.7}$$

which is in line with equation (4.5). However, in more general settings this approach to recursiveness is not the same as the one of Definition 8 (see Example 2 below).

We have the following analogue of Proposition 3 (i) relating the concepts of dynamic consistency and recursiveness.

Proposition 5 *Suppose that preference mappings $\mathcal{R}_{s,t}$, $1 \le s < t \le T$, are monotone (strictly monotone) and recursive. Then $\{\mathcal{R}_{s,t}\}_{1\le s<t\le T}$ is dynamically consistent (strictly dynamically consistent).*

Proof We need to verify implication (4.1). By the recursiveness, for $1 \le s < t < u \le T$ we have

$$\mathcal{R}_{s,u}(Z_s,...,Z_u) = \mathcal{R}_{s,t}\big(Z_s,\dots,Z_{t-1},\mathcal{R}_{t,u}(Z_t,\dots,Z_u)\big),$$
$$\mathcal{R}_{s,u}(Z'_s,...,Z'_u), = \mathcal{R}_{s,t}\big(Z'_s,\dots,Z'_{t-1},\mathcal{R}_{t,u}(Z'_t,\dots,Z'_u)\big).$$

Since $Z_\tau = Z'_\tau$, $\tau = s,...,t-1$, by monotonicity (strict monotonicity) of $\mathcal{R}_{s,t}$ the implication (4.1) (the implication (4.2)) follows. □

The converse of Proposition 5 can be given in the framework of Remark 4 (cf. Ruszczyński (2010, Theorem 1)).

Proposition 6 *Suppose that:* (i) $\{\mathcal{R}_{s,t}\}_{1\le s<t\le T}$ *is dynamically consistent,* (ii)

$$\mathcal{R}_{s,t}(Z_s,...,Z_t) = \mathcal{R}_{s,u}(Z_s,...,Z_t,0,...,0),\ 1 \le s < t < u \le T, \tag{4.8}$$

(iii) *for $Z_t \in \mathcal{Z}_t$,*

$$\mathcal{R}_{t,u}(Z_t,0,...,0) = Z_t,\ 1 \le t < u \le T. \tag{4.9}$$

Then mappings $\mathcal{R}_{s,t}$, $1 \le s < t \le T$, are recursive.

Proof Consider sequences

$$Z = (Z_s,...,Z_{t-1},Z_t,...,Z_u) \in \mathcal{Z}_{s,u} \text{ and } Z' = (Z_s,...,Z_{t-1},\mathcal{R}_{t,u}(Z_t,...,Z_u),0,...,0) \in \mathcal{Z}_{s,u}.$$

By (4.9) we have

$$\mathcal{R}_{t,u}(Z'_t,...,Z'_u) = \mathcal{R}_{t,u}(\mathcal{R}_{t,u}(Z_t,...,Z_u),0,...,0) = \mathcal{R}_{t,u}(Z_t,...,Z_u).$$

Hence by dynamic consistency, $\mathcal{R}_{s,u}(Z_s,...,Z_u) = \mathcal{R}_{s,u}(Z'_s,...,Z'_u)$. Together with (4.8) This gives the required equation (4.4). □

Example 2 The following examples of preference systems are recursive and dynamically consistent.

(i) The preference system

$$\mathcal{R}_{s,t}(Z_s,\dots,Z_t) := \mathbb{E}_{|\mathcal{F}_s}[Z_s + \dots + Z_t] \tag{4.10}$$

is recursive, strictly monotone and hence strictly dynamically consistent. Here we can use spaces $\mathcal{Z}_t := L_1(\Omega,\mathcal{F}_t,P)$.

(ii) Let $\mathcal{Z}_t := L_\infty(\Omega,\mathcal{F}_t,P)$, $t = 1,...,T$, and define

$$\mathcal{R}_{s,t}(Z_s,\dots,Z_t) := \operatorname*{ess\,sup}_{\mathcal{F}_s}\{Z_s,\dots,Z_t\}, \tag{4.11}$$

where the conditional essential supremum is the smallest $\mathcal{F}_s$-random variable (X_s, say), so that $Z_\tau \le X_s$ (componentwise) for all τ with $s \le \tau \le t$.[7] This preference system is recursive and monotone and hence is dynamically consistent. However this system is not strictly monotone and is not strictly dynamically consistent. Also the relations (4.8) and (4.9) do not hold here.

(iii) Recursive is the system

$$\mathcal{R}_{s,t}(Z_s,\dots,Z_t) := \mathbb{E}_{|\mathcal{F}_s}(\zeta_{s+1}\cdot\dots\cdot\zeta_t\cdot Z_t), \tag{4.12}$$

where each ζ_τ is $\mathcal{F}_\tau$ measurable, $\zeta_s \ge 0$ with $\mathbb{E}[\zeta_s] = 1$ and the product $\zeta_{s+1}\cdot\dots\cdot\zeta_t$ is in the dual of $\mathcal{Z}_t$.

[7] See footnote 6 on page 4 for a helpful reference regarding the essential supremum.

Example 3 Consider one-step mappings $\rho_t : \mathcal{Z}_t \to \mathcal{Z}_{t-1}$, such that $\rho_t(0) = 0$, $t = 2..., T$, and define

$$\mathcal{R}_{t,T}(Z_t, ..., Z_T) := \mathbb{E}_{|\mathcal{F}_t}\left[Z_t + \sum_{\tau=t+1}^{T} \rho_\tau(Z_\tau)\right] = Z_t + \rho_{t+1}(Z_{t+1}) + \mathbb{E}_{|\mathcal{F}_t}\left[\sum_{\tau=t+2}^{T} \rho_\tau(Z_\tau)\right], \tag{4.13}$$

$t = 1, ..., T-1$. These risk mappings were considered in Pflug and Römisch (2007, Section 3).

This system can be decomposed in the following way (Homem-de-Mello and Pagnoncelli, 2016, Section 5.1). We have

$$\mathcal{R}_{t,T}(Z_t, ..., Z_T) = Z_t + \rho_{t+1}(Z_{t+1}) + \mathbb{E}_{|\mathcal{F}_t}\Big[\rho_{t+2}(Z_{t+2}) + \mathbb{E}_{|\mathcal{F}_{t+1}}\big[\rho_{t+3}(Z_{t+3}) + \cdots + \mathbb{E}_{|\mathcal{F}_{T-2}}\rho_T(Z_T)\big]\Big].$$

Suppose further that mappings ρ_t are translation equivariant. Then we have the following decomposition,

$$\mathcal{R}_{t,T}(Z_t, ..., Z_T) = Z_t + \rho_{t+1}\big(Z_{t+1} + \tilde{\rho}_{t+2}\big(Z_{t+2} + \cdots + \tilde{\rho}_T(Z_T)\big),\tag{4.14}$$

where $\tilde{\rho}_\tau(Z_\tau) := \mathbb{E}_{|\mathcal{F}_{\tau-2}}[\rho_\tau(Z_\tau)]$, $\tau = 3, ..., T$. Assuming further that mappings ρ_t are monotone (strictly monotone) it follows that this system is dynamically consistent (strictly dynamically consistent) in the sense of Definition 7. Note that $\tilde{\rho}_\tau$ maps $\mathcal{Z}_\tau$ into $\mathcal{Z}_{\tau-2}$ and the decomposition (4.14) is not of the form (3.7). □

5 Dynamic equations

In this section we derive the corresponding dynamic equations and discuss their relation to the decomposability and time consistency. The theorem presented below allows verification of a time consistent policy based on recursive preference relations. We start again with additive objective functions and extend the setting to general objective functions subsequently.

5.1 The additive objective function

Suppose that the objective function of the reference problem is a function of the sum of the costs and is decomposable in the form (3.7) with respective one-step mappings $\rho_t : \mathcal{Z}_t \to \mathcal{Z}_{t-1}$ being monotone and translation equivariant. Then *dynamic programming equations* for the reference problem can be written going backwards in time (cf. Ruszczyński and Shapiro (2006a)). That is, the cost-to-go (value) function at stage $t = T, ..., 2$ is

$$Q_t(x_{t-1}, \omega) := \operatorname*{ess\,inf}_{x_t \in \mathcal{X}_t(x_{t-1},\omega)} \big\{f_t(x_t, \omega) + \rho_{t+1}[Q_{t+1}(x_t, \omega)]\big\},\tag{5.1}$$

with the term $\rho_{T+1}[Q_{T+1}(x_T, \omega)]$, at stage $t = T$, omitted. At first stage the problem

$$\operatorname*{Min}_{x_1 \in \mathcal{X}_1} f_1(x_1) + \rho_2[Q_2(x_1, \omega)]\tag{5.2}$$

should be solved.

Consider policies $\bar{\pi} = (\bar{x}_1, ..., \bar{x}_T)$ with $\bar{x}_1$ being an optimal solution of problem (5.2) at the first stage and $\bar{x}_t = \bar{x}_t(\bar{x}_{t-1}, \omega)$, $t = 2, ..., T$, being an optimal solution of the problem

$$\operatorname*{Min}_{x_t \in \mathcal{X}_t(x_{t-1},\omega)} f_t(x_t, \omega) + \rho_{t+1}[Q_{t+1}(x_t, \omega)].\tag{5.3}$$

Provided that $\bar{\pi} \in \Pi$, we refer to such policy $\bar{\pi}$ as a *solution of dynamic programming equations* (5.1)–(5.2). If such minimizers $\bar{x}_t$ do exist, then the corresponding policy $\bar{\pi}$ is optimal for the reference problem, and moreover is time consistent (see below).

Example 4 (Solutions of the dynamic equations are time consistent) Consider the setting of Example 1 (and Fig. 1) with three stages and max-type preference system. Dynamic programming equations for that problem can be written as

$$Q_3(x_2,\omega) = \min_{x_3\in[1,2]} c_3(\omega)x_3 = \begin{cases} 1 & \text{if } \omega\in\{\omega_1,\omega_2\}, \\ 4 & \text{if } \omega\in\{\omega_3,\omega_4\}. \end{cases}$$

At the second stage $t=2$ we have two nodes, node whose children nodes at stage $t=3$ are nodes ω_1 and ω_2, and node whose children nodes at stage $t=3$ are nodes ω_3 and ω_4. We have that $Q_2(x_1,\omega)=\max\{1,4\}=4$ for both nodes. The policy $\bar\pi=(\bar x_1,\bar x_2,\bar x_3)$ that solves these dynamic equations is $\bar x_1=0$, $\bar x_2(\cdot)\equiv 0$ and $\bar x_3(\cdot)\equiv 1$, and is time consistent.

As it was shown in Example 1 there are many other policies which are optimal for the reference problem and are not time consistent. However, these time inconsistent policies are not solutions of the dynamic equations in the sense of condition (5.3). □

5.2 The general objective function

Consider the preference systems $\mathfrak{R}=\{\mathcal{R}_{s,t}\}_{1\le s<t\le T}$ with preference mappings of the general form (2.1). Based on the recursiveness property we derive dynamic programming equations. We also prove the verification theorem, which states that the policy solving the dynamic equations is time consistent.

Consider the following dynamic programming equations associated with the preference system $\mathfrak{R}$. At the last stage the cost-to-go function is defined as

$$Q_T(x_{T-1},\omega) := \operatorname*{ess\,inf}_{x_T\in\mathcal{X}_T(x_{T-1},\omega)} f_T(x_T,\omega), \tag{5.4}$$

and at stages $t=T-1,...,2$,

$$Q_t(x_{t-1},\omega) := \operatorname*{ess\,inf}_{x_t\in\mathcal{X}_t(x_{t-1},\omega)} \mathcal{R}_{t,t+1}\big(f_t(x_t,\omega), Q_{t+1}(x_t,\omega)\big). \tag{5.5}$$

At the first stage the problem

$$\operatorname*{Min}_{x_1\in\mathcal{X}_1} \mathcal{R}_{1,2}\big(f_1(x_1), Q_2(x_1,\omega)\big) \tag{5.6}$$

should be solved. In a rudimentary form such approach to writing dynamic equations with relation to time consistency was outlined in Shapiro (2009).

Remark 5 Recall that in the additive case $\mathcal{R}_{t,t+1}(Z_t,Z_{t+1})=\rho_{t,T}(Z_t+Z_{t+1})$ (see equation (4.7)). If moreover $\rho_{t,T}$ are translation equivariant, this becomes

$$\mathcal{R}_{t,t+1}(Z_t,Z_{t+1}) = Z_t + \rho_{t,T}(Z_{t+1}).$$

Suppose further that mappings $\rho_{t,T}$ are decomposable via a family of one-step mappings ρ_t (see Definition 6). In that case equations (5.4)–(5.6) coincide with the respective equations (5.1)–(5.2).

As in Section 5.1 we say that a policy $\bar\pi=(\bar x_1,...,\bar x_T)\in\Pi$ is a *solution of the dynamic programming equations* if each $\bar x_t$ is an optimal solution of the respective minimization problem in (5.4)–(5.6). The following verification theorem characterizes time consistent solutions by their dynamical equations, these equations constitute *necessary and sufficient* conditions to characterize a time consistent policy.

Theorem 2 (Verification theorem) *Suppose the preference system $\mathfrak{R}$ is recursive and monotone. Then the following holds. If $\bar\pi=(\bar x_1,\dots,\bar x_T)\in\Pi$ is a solution of dynamic equations (5.4)–(5.6), then $\bar\pi$ is time consistent, and the optimal value of reference problem (2.2) is given by the optimal value of the first stage problem (5.6).*

Conversely if $\bar\pi=(\bar x_1,\dots,\bar x_T)$ is a time consistent optimal solution of reference problem (2.2), then $\bar\pi$ is a solution of dynamic equations (5.4)–(5.6).

Proof To simplify notation we will write here $f_t(x_t)$ instead of $f_t(x_t,\omega)$. Let $\bar\pi=(\bar x_1,\dots,\bar x_T)\in\Pi$ be a solution of dynamic equations (5.4)–(5.6). By recursiveness and monotonicity we have that

$$\begin{aligned}&\operatorname*{ess\,inf}_{x_t,\dots,x_T}\mathcal{R}_{t,T}\big(f_t(x_t),\dots,f_T(x_T)\big)\\&\qquad=\operatorname*{ess\,inf}_{x_t,\dots,x_T}\mathcal{R}_{t,t+1}\Big(f_t(x_t),\mathcal{R}_{t+1,T}\big(f_{t+1}(x_{t+1}),\dots,f_T(x_T)\big)\Big)\\&\qquad\geq\operatorname*{ess\,inf}_{x_t}\mathcal{R}_{t,t+1}\Big(f_t(x_t),\operatorname*{ess\,inf}_{x_{t+1},\dots,x_T}\mathcal{R}_{t+1,T}\big(f_{t+1}(x_{t+1}),\dots,f_T(x_T)\big)\Big),\end{aligned}$$

and thus recursively

$$\begin{aligned}&\operatorname*{ess\,inf}_{x_t,\dots,x_T}\mathcal{R}_{t,T}\big(f_t(x_t),\dots,f_T(x_T)\big)\\&\qquad\geq\operatorname*{ess\,inf}_{x_t}\mathcal{R}_{t,t+1}\Big(f_t(x_t),\operatorname*{ess\,inf}_{x_{t+1}}\mathcal{R}_{t+1,t+2}\big(f_{t+1}(x_{t+1}),\operatorname*{ess\,inf}_{x_{t+2}}\mathcal{R}_{t+2,t+3}(\dots)\big)\Big).\end{aligned}\tag{5.7}$$

Now note that the optimal policy $\bar\pi=(\bar x_1,\dots,\bar x_T)$ corresponds exactly to the problem (5.7). We thus conclude that

$$\operatorname*{ess\,inf}_{x_t,\dots,x_T}\mathcal{R}_{t,T}\big(f_t(x_t),\dots,f_T(x_T)\big)\geq\mathcal{R}_{t,t+1}\Big(f_t(\bar x_t),\mathcal{R}_{t+1,t+2}\big(f_{t+1}(\bar x_{t+1}),\mathcal{R}_{t+2,t+3}(\dots)\big)\Big)$$

and by recursiveness of the preference system further that

$$\operatorname*{ess\,inf}_{x_t,\dots,x_T}\mathcal{R}_{t,T}\big(f_t(x_t),\dots,f_T(x_T)\big)\geq\mathcal{R}_{t,T}\big(f_t(\bar x_t),\dots,f_T(\bar x_T)\big).$$

It follows that the policy $\bar\pi=(\bar x_1,\dots,\bar x_T)$ is optimal and, in addition, time consistent.

Conversely let $\bar\pi=(\bar x_1,\dots,\bar x_T)$ be a time consistent optimal solution of reference problem (2.2). Time consistency means that $(\bar x_t,\dots,\bar x_T)$ is optimal for problem (2.3) conditional on $\mathcal{F}_t$ and $\bar x_{t-1}$, $t=2,...,T$ (see Definition 1). Now for $x_{t-1}=\bar x_{t-1}$ the right hand side of (5.5) gives the optimal value of problem (2.3). It follows that $\bar x_t$ is an optimal solution of the minimization problem (5.5). This completes the proof. □

Corollary 1 (Existence of a time consistent policy) *Suppose that the preference system $\mathfrak{R}$ is recursive and monotone, and the essential infimum in minimization problems (5.4)–(5.6) is attained. Then there exists an optimal policy which is time consistent.*

Example 5 Consider preference system (4.11) of Example 2. For this system $\mathcal{R}_{t,t+1}=\operatorname{ess\,sup}_{\mathcal{F}_t}\{Z_t,Z_{t+1}\}$ and hence dynamic equations (5.5) take the form

$$Q_t(x_{t-1},\omega):=\operatorname*{ess\,inf}_{x_t\in\mathcal{X}_t(x_{t-1},\omega)}\left\{\operatorname*{ess\,sup}_{\mathcal{F}_t}\big(f_t(x_t,\omega),Q_{t+1}(x_t,\omega)\big)\right\}.\tag{5.8}$$

6 Risk measures

In this section we assume that spaces $\mathcal{Z}_t:=L_p(\Omega,\mathcal{F}_t,P)$ with $p\in[1,\infty)$. Space $\mathcal{Z}_t$ is paired with its dual space $\mathcal{Z}_t^*:=L_q(\Omega,\mathcal{F}_t,P)$, where $q\in(1,\infty]$ is such that $1/p+1/q=1$. For the space $\mathcal{Z}_T=L_p(\Omega,\mathcal{F},P)$ we often drop the subscript and write it simply as $\mathcal{Z}=L_p(\Omega,\mathcal{F},P)$. A functional $\rho:\mathcal{Z}\to\mathbb{R}$ is said to be a *coherent risk measure* if it satisfies the conditions of convexity, monotonicity, translation equivariance and positive homogeneity (Artzner et al. (1999)). It is said that ρ is *law invariant* if it is a function of the distribution of $Z\in\mathcal{Z}$ rather than its particular realization, i.e., if $Z,Z'\in\mathcal{Z}$ are such that $P(Z\leq z)=P(Z'\leq z)$ for all $z\in\mathbb{R}$, then $\rho(Z)=\rho(Z')$.

6.1 Representation of Risk measures

A real valued convex and monotone functional $\rho : \mathcal{Z} \to \mathbb{R}$ is continuous in the norm topology of the space $L_p(\Omega, \mathcal{F}, P)$ (cf. Ruszczyński and Shapiro (2006b, Corollary 3.1)) and hence by the Fenchel-Moreau theorem has the dual representation

$$\rho(Z) = \sup_{\zeta \in \mathfrak{A}} \{\langle \zeta, Z\rangle - \rho^*(\zeta)\}, \tag{6.1}$$

where

$$\rho^*(\zeta) := \sup_{Z \in \mathcal{Z}} \{\langle \zeta, Z\rangle - \rho(Z)\}, \quad \zeta \in \mathcal{Z}^*,$$

is the conjugate of ρ,

$$\mathfrak{A} := \{\zeta \in \mathcal{Z}^* : \rho^*(\zeta) < \infty\}$$

is the domain of ρ^* and $\langle \zeta, Z\rangle := \int \zeta Z dP$ is the respective scalar product.

If ρ is convex, monotone and translation equivariant, then the domain $\mathfrak{A}$ consists of density functions, i.e., if $\zeta \in \mathfrak{A}$, then $\zeta \succeq 0$ and $\int \zeta dP = 1$. If moreover ρ is positively homogeneous, i.e., ρ is a coherent risk measure, then the conjugate of ρ is an indicator function, and ρ can be represented in the form

$$\rho(Z) = \sup_{\zeta \in \mathfrak{A}} \langle \zeta, Z\rangle, \tag{6.2}$$

where $\mathfrak{A} \subset \mathcal{Z}^*$ is a set of density functions. Since ρ is assumed to be real valued, the set $\mathfrak{A}$ is bounded. The set $\mathfrak{A}$ in representation (6.2) is not defined uniquely. The maximum in (6.2) is not changed if $\mathfrak{A}$ is replaced by weak* topological closure of its convex hull. Therefore unless stated otherwise we assume that the set $\mathfrak{A}$ in representation (6.2) is convex and weakly* compact.

If ρ is a law invariant coherent risk measure, then it can be considered as a function of cdf $F(z) = P(Z \le z)$. In that case, assuming that the probability space $(\Omega, \mathcal{F}, P)$ is nonatomic, the dual representation (6.2) takes the form

$$\rho(F) = \sup_{\sigma \in \Upsilon} \int_0^1 \sigma(t) F^{-1}(t) dt, \tag{6.3}$$

where $F^{-1}(t) = \inf\{z : F(z) \ge t\}$ is the respective (left side) quantile and Υ is a set of spectral functions. A function $\sigma : [0, 1) \to \mathbb{R}_+$ is called spectral if it is monotonically nondecreasing, right side continuous and $\int_0^1 \sigma(t) dt = 1$ (cf. Pichler and Shapiro (2015)). We will refer to Υ as a *generating set*. In particular if $\Upsilon = \{\sigma\}$ is a singleton, risk measure

$$\rho(F) = \int_0^1 \sigma(t) F^{-1}(t) dt \tag{6.4}$$

is called *spectral*. Note that here $\rho(F)$ is defined for cdfs F with finite p-th order moments, i.e., $\int_0^1 |F^{-1}(t)|^p dt < \infty$, and every $\sigma \in \Upsilon$ has finite q-th order moment, i.e., $\int_0^1 |\sigma(t)|^q dt < \infty$.

Recall that the subdifferential of a convex continuous functional $\rho : \mathcal{Z} \to \mathbb{R}$ is nonempty, convex and weakly* compact subset of $\mathcal{Z}^*$ and is given by

$$\partial\rho(Z) = \arg\max_{\zeta \in \mathcal{Z}^*} \{\langle \zeta, Z\rangle - \rho^*(\zeta)\}. \tag{6.5}$$

Moreover, if ρ is monotone, then $\zeta \succeq 0$ for any $\zeta \in \partial\rho(Z)$. The following condition for strict monotonicity is a straightforward extension of similar condition for coherent risk measures presented in Shapiro and Ugurlu (2016, Proposition 3.1):

$$\text{for any } Z \in \mathcal{Z} \text{ and any } \zeta \in \partial\rho(Z) \text{ it follows that } \zeta(\omega) > 0 \text{ for a.e. } \omega \in \Omega. \tag{6.6}$$

We give a proof for the sake of completeness.

Proposition 7 *Let $\rho : \mathcal{Z} \to \mathbb{R}$ be a convex monotone functional. Then ρ is strictly monotone iff condition (6.6) holds.*

Proof Recall that convexity and monotonicity of ρ implies its continuity. Thus for any $Z \in \mathcal{Z}$, subdifferential $\partial\rho(Z)$ is nonempty and (6.5) follows. Suppose that condition (6.6) holds. Let $Z' \succ Z$ and $\zeta \in \partial\rho(Z)$. Then

$$\rho(Z') - \rho(Z) \geq \langle \zeta, Z' - Z\rangle,$$

where $\langle \zeta, Z' - Z\rangle = \int \zeta(Z' - Z)dP$. Since $Z' \succ Z$ it follows by condition (6.6) that $\int \zeta(Z' - Z)dP > 0$, and hence $\rho(Z') > \rho(Z)$.

For the converse implication we argue by a contradiction. Suppose that there exist $\zeta \in \partial\rho(Z)$ and $A \in \mathcal{F}$ such that $P(A) > 0$ and $\zeta(\omega) = 0$ for all $\omega \in A$. Consider $Z' := Z - \mathbf{1}_A$. Clearly $Z \succ Z'$ and hence $\rho(Z) \geq \rho(Z')$. Moreover

$$\rho(Z) = \int_\Omega \zeta Z dP - \rho^*(\zeta) = \int_\Omega \zeta Z dP - \int_A \zeta dP - \rho^*(\zeta) = \int_\Omega \zeta Z' dP - \rho^*(\zeta) \leq \rho(Z').$$

It follows that $\rho(Z) = \rho(Z')$, a contradiction with strict monotonicity. □

It follows that spectral risk measure (6.4) is strictly monotone iff the corresponding spectral function $\sigma(t)$ is positive for all $t \in (0, 1)$ (cf. Shapiro et al. (2014, Proposition 6.37)). In particular the Average Value-at-Risk measure $\mathsf{AV@R}_\alpha(\cdot)$ (defined below) is not strictly monotone for $\alpha \in (0, 1)$. On the other hand convex combination $\lambda\mathbb{E}(\cdot) + (1 - \lambda)\mathsf{AV@R}_\alpha(\cdot)$ is strictly monotone for $\lambda \in (0, 1]$.

With every law invariant coherent risk measure $\rho(F)$ are associated its conditional analogues. That is, let $\mathcal{G}$ be a sigma subalgebra of $\mathcal{F}$. Then the corresponding conditional risk mapping can be defined by replacing cdf F with the respective conditional distribution $F_{|\mathcal{G}}$.

An important example of law invariant coherent risk measure is the Average Value-at-Risk measure (also called Conditional Value-at-Risk, Expected Shortfall and Expected Tail Loss). In various equivalent forms it was discovered and rediscovered by several authors. It can be written as

$$\mathsf{AV@R}_\alpha(F) = \frac{1}{1-\alpha}\int_\alpha^1 F^{-1}(t)dt,\ \alpha \in [0, 1), \tag{6.7}$$

or alternatively

$$\mathsf{AV@R}_\alpha(Z) = \inf_{u\in\mathbb{R}} \left\{u + (1-\alpha)^{-1}\mathbb{E}[Z - u]_+\right\},\ \alpha \in [0, 1). \tag{6.8}$$

The representation (6.7) shows that $\mathsf{AV@R}_\alpha$ is a spectral risk measure with the respective spectral function $\sigma(\cdot) = (1-\alpha)^{-1}\mathbf{1}_{[\alpha,1]}(\cdot)$. The variational representation (6.8) is due to Rockafellar and Uryasev (2000). The Average Value-at-Risk measure is naturally defined on the space $\mathcal{Z} = L_1(\Omega, \mathcal{F}, P)$, on which it is finite valued. As a function of α, $\mathsf{AV@R}_\alpha(Z)$ is monotonically nondecreasing, $\mathsf{AV@R}_0(\cdot) = \mathbb{E}(\cdot)$ and $\mathsf{AV@R}_\alpha(Z)$ tends to $\operatorname{ess\,sup}(Z)$ as $\alpha \uparrow 1$. The Average Value-at-Risk is a coherent risk measure. Its dual formulation following the general representation (6.1) is

$$\mathsf{AV@R}_\alpha(Z) = \sup\left\{\mathbb{E}[\zeta\, Z] : \zeta \in L_\infty(\Omega, \mathcal{F}, P),\ \mathbb{E}[\zeta] = 1 \text{ and } 0 \leq \zeta \leq (1-\alpha)^{-1} \text{ a.s.}\right\}. \tag{6.9}$$

6.2 Conditional risk mappings

Conditional on the sigma algebras $\mathcal{F}_t$, with $\mathsf{AV@R}_\alpha$ are associated its conditional analogues $\mathsf{AV@R}_{\alpha|\mathcal{F}_t} : \mathcal{Z}_T \to \mathcal{Z}_t$, where $\mathcal{Z}_t := L_1(\Omega, \mathcal{F}_t, P)$. This suggests to consider preference mappings $\rho_{t,T} := \mathsf{AV@R}_{\alpha|\mathcal{F}_{t-1}}$, $t = 2, ..., T$. Note however that for $\alpha \in (0, 1)$ this preference system is not recursive, i.e., it can happen that for a sigma algebra $\mathcal{G} \subset \mathcal{F}$,

$$\mathsf{AV@R}_\alpha\left(\mathsf{AV@R}_{\alpha|\mathcal{G}}(Z)\right) \neq \mathsf{AV@R}_\alpha(Z). \tag{6.10}$$

Remark 6 For $\alpha \in (0, 1)$ the nested risk measure $\mathsf{AV@R}_\alpha\big(\mathsf{AV@R}_{\alpha|\mathcal{G}}(\cdot)\big)$ is not strictly monotone. Also it is not dynamically consistent and not *law invariant*, although $\mathsf{AV@R}_\alpha$ is.

Remark 7 If $\mathcal{G}$ is independent of the sigma algebra generated by the random variable Z, then it holds that $\mathsf{AV@R}_{\alpha|\mathcal{G}}(Z) = \mathsf{AV@R}(Z)$. In this case, (6.10) holds with equality.

The left hand side of (6.10) is often bigger than the right hand side. However, this is not always the case and there are examples where the opposite inequality holds (cf. Pflug and Römisch (2007, pp. 157–158)).

Dynamic risk measures, which are law invariant. Let ρ be a law invariant, convex, monotone and translation equivariant risk measure. It is shown in Kupper and Schachermayer (2009) that (under certain regularity conditions) the recursion

$$\rho\left(\rho_{|\mathcal{G}}(\cdot)\right) = \rho(\cdot) \tag{6.11}$$

holds for a nontrivial sigma algebra $\mathcal{G} \subset \mathcal{F}$, $\mathcal{G} \neq \mathcal{F}$, only if

$$\rho(Z) = \frac{1}{\gamma} \log \mathbb{E}\left[\exp(\gamma Z)\right], \ \gamma \in [0, \infty]. \tag{6.12}$$

Consequently the preference system $\rho_{t,T} := \rho_{|\mathcal{F}_{t-1}}$, $t = 2, ..., T$, is recursive (dynamically consistent), in the sense of Section 3, only if ρ is of the form (6.12). The risk measures (6.12) are called *entropic*. Coherent risk measures $\rho(\cdot) = \mathbb{E}(\cdot)$ and $\rho(\cdot) = \operatorname{ess\,sup}(\cdot)$ correspond to the boundary case $\gamma = 0$ and $\gamma = \infty$, respectively. That is, for law invariant coherent risk measure ρ the preference system $\rho_{t,T} := \rho_{|\mathcal{F}_{t-1}}$, $t = 2, ..., T$, is recursive (dynamically consistent), in the sense of Section 3, only in two cases:

$$\rho(\cdot) = \mathbb{E}(\cdot) \text{ or } \rho(\cdot) = \operatorname{ess\,sup}(\cdot). \tag{6.13}$$

Nested risk measures An alternative is to consider nested risk measures. That is, let $\rho_t : \mathcal{Z}_t \to \mathcal{Z}_{t-1}$, $t = 2, ..., T$, be risk mappings satisfying the conditions of convexity, monotonicity and translation equivariance. Define $\rho_{t,T} := \rho_{t+1} \circ \cdots \circ \rho_T$, $t = 1, ..., T-1$. By the definition the obtained preference system $\{\rho_{t,T}\}_{t=1}^{T-1}$ is decomposable and hence, by Proposition 3, is recursive and thus is dynamically consistent. For example we can use $\rho_t := \mathsf{AV@R}_{\alpha|\mathcal{F}_t}$, $t = 1, ..., T-1$. The corresponding

$$\rho_{t,T}(\cdot) = \mathsf{AV@R}_{\alpha|\mathcal{F}_t}\left(\mathsf{AV@R}_{\alpha|\mathcal{F}_{t+1}}(\cdots \mathsf{AV@R}_{\alpha|\mathcal{F}_{T-1}}(\cdot))\right) \tag{6.14}$$

are called *nested Average Value-at-Risk* mappings. As it was pointed above, for $\alpha \in (0,1)$ these $\rho_{t,T} : \mathcal{Z}_T \to \mathcal{Z}_t$ are different from $\mathsf{AV@R}_{\alpha|\mathcal{F}_t} : \mathcal{Z}_T \to \mathcal{Z}_t$.

It was shown that a spectral risk measure is strictly monotone iff the corresponding spectral function is positive on the interval $(0,1)$. Spectral function of $\mathsf{AV@R}_\alpha$ is the step function $(1-\alpha)^{-1}\mathbf{1}_{[\alpha,1]}(\cdot)$. It follows that $\mathsf{AV@R}_\alpha$ is not strictly monotone for $\alpha \in (0,1)$. Therefore Theorem 1 cannot be applied to conclude that for the preference system given by nested Average Value-at-Risk mappings, every optimal solution of the corresponding reference problem is time consistent. And indeed it is possible to construct a counterexample where some of the optimal solutions are not time consistent although the preference mappings are decomposable into nested Average Value-at-Risk mappings. The settings of Example 1 can be also used for such counterexample (cf. Shapiro and Ugurlu (2016)). On the other hand a solution of the corresponding dynamic programming equations (if it exists) is time consistent (see Section 5).

7 Time consistent adaptation of the final risk measure

In this section we discuss the additive case considered in Section 3. We revisit the reference problem (3.1) with a given function $\varrho : \mathcal{Z}_T \to \mathbb{R}$ which is assumed to be a coherent risk measure. The risk measure ϱ constitutes a subjective risk preference of the decision maker. The respective policies $\pi \in \Pi$ are adapted to the filtration $\mathfrak{F}$. In this way the reference problem (3.1) is well-defined. However, the intermediate problems (3.2) at stages $t = 2, ..., T$ are not defined without explicitly specified preference mappings $\rho_{t,T}$. Consequently, it is not possible to talk about time consistency, as introduced in Definition 1, unless the preference system $\mathfrak{R} = \{\rho_{t,T}\}_{t=1}^{T}$ is clearly defined.

Based on an optimal policy $\bar{\pi} \in \Pi$ it is possible to specify a preference system $\mathfrak{R}$ *retrospectively*, so that the minimization problem appears to be time consistent. Such preference system would depend on the optimal policy $\bar{\pi}$. In this way the decision maker solves the entire optimization problem (3.1) first. Consequently the preference system, which is found retrospectively based on the optimal policy, can be used to manage the subsequent steps in a seemingly time consistent way. We include this approach and the following discussion to emphasize this conceptual difference from the concept of time consistency addressed in the previous sections.

The following decomposition theorem outlines how to adjust the risk preference in order to meet the goal $\varrho(\cdot) := \mathsf{AV@R}_\alpha(\cdot)$. It turns out that the risk profile has to be adjusted based on the information revealed, and the decision maker has to change its risk profile over time to meet the optimal policy. The following theorem provides an accordant statement for the Average Value-at-Risk. Pflug and Pichler (2016) provide the statement in more generality for general coherent risk measures.

Theorem 3 (Decomposition of the Average Value-at-Risk) *Let $\mathcal{G} \subset \mathcal{F}$ be a subalgebra of $\mathcal{F}$. The Average Value-at-Risk at level α has the decomposition*

$$\mathsf{AV@R}_\alpha(Z) = \sup_{\tilde{\zeta} \in \mathfrak{S}} \mathbb{E}\left[\tilde{\zeta} \cdot \mathsf{AV@R}_{1-(1-\alpha)\tilde{\zeta}}(Z|\,\mathcal{G})\right], \tag{7.1}$$

where $\mathfrak{S}$ is the set of random variables $\tilde{\zeta}$ which are $\mathcal{G}$ measurable and which satisfy $\mathbb{E}[\tilde{\zeta}] = 1$ and $0 \le \tilde{\zeta} \le \frac{1}{1-\alpha}$.

Remark 8 The constraints in the representation (7.1) are the same as for the unconditional dual representation (6.9), except that the dual random variable $\tilde{\zeta}$ is additionally adapted to the new information available, i.e., $\tilde{\zeta}$ is $\mathcal{G}$-measurable in Theorem 3. The optimal random variable in (7.1) and (6.9) are related by

$$\tilde{\zeta} = \mathbb{E}_{|\mathcal{G}}[\zeta]. \tag{7.2}$$

The identity (7.1) reveals the adapted risk preference. Indeed, the decision maker changes the risk level α and applies the new risk level

$$\alpha_t := 1 - (1-\alpha)\tilde{\zeta} \in [0,1] \tag{7.3}$$

conditionally on $\mathcal{G} := \mathcal{F}_t$ at stage t. Notice, that the new risk level α_t is adapted to $\mathcal{F}_t$ and does not involve future information.

Remark 9 (Economic interpretation) The adapted risk level α_t reflects the strategic manoeuver of the decision maker. He will accept higher risk after having observed encouraging returns already at the level t and tighten the risk level after observing discouraging conditional outcomes.

Time consistent decomposition of a coherent risk measure

First we discuss the decomposition of the Average Value-at-Risk and then extend the setting to general coherent risk measures.

Let $\bar{\pi}$ be an optimal policy of the reference problem (3.1) and ζ be the dual variable of the respective total cost $\bar{Z}(\cdot) = f_1(\bar{x}_1) + \ldots + f_T(\bar{x}_T(\cdot), \cdot)$. That is

$$\mathsf{AV@R}_\alpha(\bar{Z}) = \mathbb{E}[\zeta \bar{Z}],$$

where $\mathbb{E}[\zeta] = 1$ and $0 \le \zeta \le \frac{1}{1-\alpha}$. Define $\zeta_t := \frac{\mathbb{E}(\zeta|\mathcal{F}_t)}{\mathbb{E}(\zeta|\mathcal{F}_{t-1})}$ and the risk measure

$$\rho_t(Z) := \mathbb{E}_{|\mathcal{F}_t}[\zeta_t Z].$$

Then the risk measure is decomposable as

$$\mathsf{AV@R}_\alpha(Z) = \rho_2\big(\mathsf{AV@R}_{1-(1-\alpha)\zeta_2|\mathcal{F}_t}(Z)\big)$$

by Theorem 3. This procedure may be repeated iteratively to obtain

$$\varrho(Z) = \rho_2 \circ \cdots \circ \rho_T(Z)$$

(cf. Definition 6). The corresponding preference mappings are

$$\rho_{t,T}(Z) = \rho_t \circ \cdots \circ \rho_T(Z). \tag{7.4}$$

The preference mappings $\rho_{t,T}$ satisfy the axioms (i)–(iv) in Section 3 and the corresponding preference system $\mathfrak{R}$ is recursive in the sense of Definition 5.

More general coherent risk functionals allow a similar decomposition as the Average Value-at-Risk described in Theorem 3. The decomposition of the general risk measure results in a preference mapping of the type (7.4), although the intermediate random variables ζ_t is much more difficult to provide (we refer to Pflug and Pichler (2016) for details).

Remark 10 Note that the considered preference system depends on the optimal policy $\bar{\pi}$, as the mappings $\rho_{t,T}$ in (7.4) depend on the optimal total cost $\bar{Z}$. However, once the preference mappings $\rho_{t,T}$ are available, then the conditional problems (3.2) depend only on the data process and the considered problem is time consistent in this specific sense.

8 Summary and discussion

This paper discusses time and dynamic consistency of risk averse problems. We address properties and key ingredients of risk preference systems and associated preference mappings, which are sufficient to insure time consistency of optimal policies. Further we provide counterexamples demonstrating that time consistency could be absent even in comparatively simple optimization situations.

In various settings time consistency has been addressed by many authors going all the way back to original developments by Richard Bellman. We follow an approach of considering time consistency in a framework of modern theory of risk measures. As a central result we obtain a relation between dynamic consistency of preference systems and time consistency of optimal policies.

An often asked question is whether a formulated reference problem is time consistent. Of course this could be answered only given a precise definition of "time consistency". In our framework this depends on a choice of the preference system. By adjusting the preference system to a specifically considered problem any formulation could be made time consistent. This was already discussed in Section 7. Let us also consider the following simple example.

Example 6 Consider the reference problem (3.1) with the reference objective function $\varrho(Z) := \mathsf{AV@R}_\alpha(Z)$, $\alpha \in (0,1)$. As it was discussed in Section 6.2 this risk measure is not dynamically consistent in the sense of Section 3. On the other hand using the variational representation (6.8) of $\mathsf{AV@R}_\alpha$ we can write the corresponding reference problem as the minimization problem

$$\operatorname*{Min}_{u\in\mathbb{R}} \inf_{x_1,x_2(\cdot),\dots,x_T(\cdot)} \left\{u + (1-\alpha)^{-1}\mathbb{E}[f_1(x_1) + f_2(x_2,\omega) + \cdots + f_T(x_T,\omega) - u]_+\right\} \tag{8.1}$$

subject to the feasibility constraints and involving additional variable $u \in \mathbb{R}$. Let $\bar{u}$ be an optimal solution of problem (8.1). Then the inner minimization problem in (8.1) becomes the expected value (risk neutral) problem

$$\operatorname*{Min}_{x_1,x_2(\cdot),\dots,x_T(\cdot)} \mathbb{E}\left\{\bar{u} + (1-\alpha)^{-1}[f_1(x_1) + f_2(x_2,\omega) + \cdots + f_T(x_T,\omega) - \bar{u}]_+\right\}, \tag{8.2}$$

subject to the feasibility constraints.

For problem (8.2) one can use the standard preference mappings given by the respective conditional expectations. Of course such time consistent formulation involves an unknown value $\bar{u}$ and is adjusted to a particular optimization problem. Moreover, value $\bar{u}$ is decided before observing any realization of the underling data process and is the same at every stage of solution of problem (8.2). In that sense this approach violates the principle that conditional optimality at every stage should not depend on realizations of the data process which cannot happen in the future (cf. Shapiro (2009)). Of course there are many other ways of making this problem "time consistent" (compare with Section 7). The point is that by adjusting the preference system to a considered problem, any problem could be made time consistent. □

9 Acknowledgment

We want to thank Kerem Ugurlu (University of Washington) for pointing out some typos and suggesting improvements in a previous version of this paper.

References

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent Measures of Risk. *Mathematical Finance*, 9:203–228, 1999. doi:10.1111/1467-9965.00068. 4, 12

P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku. Coherent multiperiod risk adjusted values and Bellman's principle. *Annals of Operations Research*, 152:5–22, 2007. doi:10.1007/s10479-006-0132-6. 4, 6

R. E. Bellman. *Dynamic Programming*. Princeton University Press, Princton, NJ, 1957. 1

T. Bielecki, I. Cialenco, and M. Piterai. A survey of time consistency of dynamic risk measures and dynamic performance measures in discrete time: LM-measure perspective. *EconPapers*, 2016. URL `http://math.iit.edu/~bielecki,http://math.iit.edu/~igor,http://arxiv.org/abs/1603.09030`. 2

P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability*, (3):57–106, 2006. 4, 8

T. Homem-de-Mello and B. K. Pagnoncelli. Risk aversion in multistage stochastic programming: A modeling and algorithmic perspective. *European Journal of Operational Research*, 249(1):188–199, feb 2016. doi:10.1016/j.ejor.2015.05.048. 10

I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance*. Stochastic Modelling and Applied Probability. Springer, 1998. 4

M. Kupper and W. Schachermayer. Representation results for law invariant time consistent functions. *Mathematics and Financial Economics*, 2:189–210, 2009. doi:10.1007/s11579-009-0019-9. 15

G. Ch. Pflug and A. Pichler. Time-consistent decisions and temporal decomposition of coherent risk functionals. *Mathematics of Operations Research*, 41(2):682–699, 2016. doi:10.1287/moor.2015.0747. 16

G. Ch. Pflug and W. Römisch. *Modeling, Measuring and Managing Risk*. World Scientific, River Edge, NJ, 2007. doi:10.1142/9789812708724. 2, 3, 10, 14

A. Pichler and A. Shapiro. Minimal representations of insurance prices. *Insurance: Mathematics and Economics*, 62:184–193, 2015. doi:10.1016/j.insmatheco.2015.03.011. 13

F. Riedel. Dynamic coherent risk measures. *Stochastic Processes and their Applications*, 112(2):185–200, 2004. doi:10.1016/j.spa.2004.03.004. 4

R. T. Rockafellar and S. Uryasev. Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2(3):21–41, 2000. doi:10.21314/JOR.2000.038. 14

A. Ruszczyński. Risk-averse dynamic programming for Markov decision processes. *Math. Program., Ser. B*, 125: 235–261, 2010. doi:10.1007/s10107-010-0393-3. 4, 8, 9

A. Ruszczyński and A. Shapiro. Conditional risk mappings. *Mathematics of Operations Research*, 31(3):544–561, 2006a. doi:10.1287/moor.1060.0204. 10

A. Ruszczyński and A. Shapiro. Optimization of convex risk functions. *Mathematics of Operations Research*, 31: 433–452, 2006b. 13

A. Shapiro. On a time consistency concept in risk averse multi-stage stochastic programming. *Operations Research Letters*, 37:143–147, 2009. 11, 17

A. Shapiro and K. Ugurlu. Decomposability and time consistency of risk averse multistage programs. *Operations Research Letters*, 44(5):663–665, 2016. doi:10.1016/j.orl.2016.07.015. 7, 13, 15

A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming, Modeling and Theory*. MOS-SIAM Series on Optimization. SIAM, 2nd edition, 2014. 3, 4, 5, 8, 14

T. Wang. A class of dynamic risk measures. Technical report, University of British Columbia, 1999. 4