Abstract

Bellman formulated a principle for optimization over time, which characterizes optimal policies by stating that a decision maker should not regret previous decisions retrospectively. This paper addresses time consistency and risk averse optimal stopping in stochastic optimization. The problem is stated in generality first. The paper provides a comprehensive discussion of time consistent decision-making by addressing risk measures which are recursive, nested, dynamically or time consistent and introduces stopping time risk measures. It turns out that the paradigm of time consistency is in conflict with various desirable, classical properties of general risk measures.

Time consistent dynamic formulations are obtained for risk averse optimal stopping problems. The numerical examples illustrate that risk aversion accounts for price differences, which are observed for optimally stopped options.

Keywords: Stochastic programming, coherent risk measures, time consistency, dynamic equations, optimal stopping time, Snell envelope.

*Research of this author was partly supported by NSF grant 1633196, and DARPA EQUiPS program grant SNL 014150709.
1 Introduction

Multistage problems arise naturally in environments of management or operations which proceed in successive, subsequent steps. The underlying managerial problem is very natural: a specified objective should be reached or achieved within a specific time horizon, while the corresponding costs should be as low as possible. Bellman has formulated a principle characterizing this situation of successive operations, at approximately the same time this principle was discovered by Pontryagin as well (Pontryagin’s maximum principle). Still today, Bellman’s principle is crucial in identifying the optimal policy to meet the objective.

It is a natural extension of the initial problem to incorporate aleatoric events, which are present during the execution of the implemented policy. The random realizations have a crucial impact on the initial objective. However, by involving statistics it is still possible to formulate objectives so that the problem is eligible for computational evaluations. The simplest of these statistics is the mean, or the expected value. Bellman’s principle applies to this situation after rather obvious modifications so that an adequate reformulation still allows efficient computations to identify the optimal risk neutral policy (cf. Fleming and Soner (2006)) even for a problem formulation in continuous time.

Unwanted and unfavorable random realizations can harm and jeopardize the initial objective. From a managerial perspective it is of course crucial to protect the operations against impacts as exploding costs in difficult environments. For this reasons, other objectives than the expectation are considered in multistage stochastic optimization. The modified objectives typically involve risk measures, which are designed to quantify the risk associated with random realizations. These risk measures have become the standard tool in stochastic optimization, portfolio optimization is an example where they are indispensable nowadays.

Several problem formulations involve risk measures even in multiple stages. In this risk averse setting it is natural to ask for dynamic equations, which facilitate policy evaluations and executions. Bellman’s principle, however, does not extend naturally to the situation involving risk. This problem has been realized very early in economics and management science and is still in consideration and is discussed. As well, the problem is intimately related to time consistency, that is, decisions should not be regretted from a later perspective with additional knowledge.

To highlight the conceptual difficulty regarding time consistency we elaborate and discuss different aspects and approaches here, which have been considered in the past. We shall distinguish the following three major lines of attack.

1.1 Consistency of risk measures

The discussion of consistency properties of risk measures has become popular in financial mathematics. Pioneers include Jobert and Rogers (2008), who introduce a concept of dynamic consistency of risk measures themselves. Other authors (e.g., Weber (2006); Cheridito and Kupper (2011)) take a similar approach by discussing axioms of risk measures. These publications try to identify properties of risk measures themselves, which are relevant in a
general, time consistent framework. Time consistency also respects increasing information, which can be investigated as essential feature of conditional risk measures (cf. Kovacevic and Pflug (2009)). Kupper and Schachermayer (2009) show that time consistent and law invariant risk measures have a very specific, entropy like representation which depends on not more than a single parameter.

1.2 Time consistency in multistage environments

Weller (1978) (cf. also Hammond (1989)) discusses time consistency in combination with expected utilities. He states:

I prove that under certain assumptions, consistency is equivalent to maximizing expected utility on the set of feasible plans, with a restricted set of utility functions and a tree of subjective probability distributions which satisfy the Bayesian updating rule. — Weller (1978, p. 263)

This problem setting puts risk measures in relation to stochastic optimization. In this context, however, it is to mention that Haviv (1996) gives a Markov decision process with constraints, so that the optimal solution does not satisfy Bellman’s principle.

Carpentier et al. (2012) propose another, general approach to time consistency in a multistage optimization framework. They state:

The sequence of optimization problems is said to be dynamically consistent if the optimal strategies obtained when solving the original problem at time $t_0$ remain optimal for all subsequent problems. In other words, dynamic consistency means that strategies obtained by solving the problem at the very first stage do not have to be questioned later on. — Carpentier et al. (2012, p. 249)

The paper Philpott et al. (2013) deserves to be mentioned as well in this concise overview, as it contains the perspective and the idea of insurance. Indeed, every amount exceeding the expected value can be attributed to risk, that is, the surplus amount constitutes an insurance premium. In a multistage framework these amounts of risk thus are insurance premiums of an insurance policy, which is renewed on a rolling horizon basis.

1.3 Time consistency ex post

In the context of stochastic optimization, time consistency can also be considered ex post, that is, in retrospect. To this end risk awareness is changed gradually, depending on the actual exposure of the objective. In situations with a high likelihood of severe danger, risk constraints are tightened, while risk constraints are relaxed whenever the danger is limited or even absent. However, the level of risk awareness or risk attitude is only available ex post, that is, the optimal policy has to be available. Details to this point of view are elaborated in Pflug and Pichler (2015, 2016), while Ruszczyński and Dentcheva (2018) present a new decomposition of risk forms.
1.4 Optimal stopping

Optimal stopping is intimately related to time consistency. These problems have been considered in statistics in Wald (1947, 1949) first. For a thorough discussion of optimal stopping we can refer to Shiryaev (1978). Optimal stopping together with (risk neutral) policy optimization was considered in Hordijk (1974) and Rieder (1976). The problem has been considered recently in a different context by Çavuş and Ruszczyński (2014), who define absorbing states in Markovian environments. Further, Belomestny and Krätschmer (2016) and Bayraktar et al. (2010) consider optimal stopping in a risk averse context.

It turns out that optimal stopping problems lead to dynamic equations in the classical situation. In this paper we introduce randomized risk measures and stopping time risk measures. Further, we relate them to risk averse stochastic optimization and give the optimal stopping times in risk averse situations explicitly. Dynamic equations are given explicitly as well.

Outline. This paper mainly relates the concepts of time consistency as discussed above and introduces risk averse optimal stopping. We provide the mathematical foundations and problem formulations in the following Section 2. The two stage problem is the simplest problem setting, where time consistency is an issue: Section 3 elaborates the topic. Of particular importance is the decomposition and dynamic equations, they are elaborated in Section 4, while Section 5 addresses time consistency from various perspectives with mathematical rigor. Section 6 introduces risk averse optimal stopping. Section 7 finally elaborates a typical optimal stopping problem and demonstrates that risk averse optimal stopping captures the differences of observed option prices compared to evaluations based on risk neutral assumptions.

2 Basic formulation

We will use the following framework. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathfrak{F} := (\mathcal{F}_0, \ldots, \mathcal{F}_T)$ be a filtration (a sequence of increasing sigma algebras, $\mathcal{F}_0 \subset \cdots \subset \mathcal{F}_T$) with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_T = \mathcal{F}$. Let $\mathcal{Z}_0 \subset \cdots \subset \mathcal{Z}_T$ be a sequence of linear spaces of functions $Z: \Omega \to \mathbb{R}$. 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. We denote by $Z_t$ an element of the space $\mathcal{Z}_t$. Note that an element $Z_t \in \mathcal{Z}_t$ actually is a class of $\mathcal{F}_t$-measurable functions which can be different from each other on a set of $P$-measure zero. Since $\mathcal{F}_0$ is trivial, the space $\mathcal{Z}_0$ consists of constant functions and will be identified with $\mathbb{R}$. Since $\mathcal{Z}_t: \Omega \to \mathbb{R}$ is $\mathcal{F}_t$-measurable, the process $Z_0, \ldots, Z_T$ is adapted to the filtration $\mathfrak{F}$.

We use the notation

$$Z_{t,u} := Z_t \times \cdots \times Z_u, \ 0 \leq t < u \leq T.$$ 

\(^1\)The sigma algebra consisting of only two sets, the empty set and the whole space $\Omega$, is called trivial.
For elements \( Z_{t,u} = (Z_t, \ldots, Z_u) \) and \( Z'_t, u = (Z'_t, \ldots, Z'_u) \) of \( Z_{t,u} \) we write \( Z_{t,u} \succeq Z'_{t,u} \) to denote that \( Z_\tau(\omega) \geq Z'_\tau(\omega) \) for \( \tau = t, \ldots, u \) and almost every (with respect to the reference probability measure \( \mu \)) \( \omega \in \Omega \), and write \( Z_{t,u} \succ Z'_{t,u} \) to denote that \( Z_{t,u} \succeq Z'_t,u \) and \( Z_{t,u} \neq Z'_t,u \). By \( \mathbb{E}_{\mathcal{F}_t} \), we denote the conditional expectation with respect to \( \mathcal{F}_t \). Note that since for \( \mathcal{F}_t \subset \mathcal{F}_u \) for \( t < u \), it follows that \( \mathbb{E}_{\mathcal{F}_t}[\mathbb{E}_{\mathcal{F}_u}(\cdot)] = \mathbb{E}_{\mathcal{F}_t}(\cdot) \) and since \( \mathcal{F}_0 \) is trivial, \( \mathbb{E}_{\mathcal{F}_0} \) is the corresponding unconditional expectation.

For a chosen preference functional\(^2\) \( \mathcal{R} : \mathcal{Z}_{0,T} \to \mathbb{R} \) we consider the optimization problem

\[
\text{Min } \mathcal{R}[f_0(x_0), f_1(x_1, \omega), \ldots, f_T(x_T, \omega)],
\]

s.t. \( x_0 \in \mathcal{X}_0, x_t \in \mathcal{X}_t(x_{t-1}, \omega), t = 1, \ldots, T, \)

called the reference problem, where \( f_0 : \mathbb{R}^n_0 \to \mathbb{R}, \mathcal{X}_0 \subset \mathbb{R}^n_0, f_t : \mathbb{R}^n_t \times \Omega \to \mathbb{R}, \mathcal{X}_t : \mathbb{R}^{n_t-1} \times \Omega \to \mathbb{R}^n_t, t = 1, \ldots, T \). It is assumed that \( f_t(x_{t-1}, \cdot) \) and \( \mathcal{X}_t(x_{t-1}, \cdot) \) are \( \mathcal{F}_t \)-measurable. A sequence \( \pi = \{x_0, x_1(\cdot), \ldots, x_T(\cdot)\} \) of mappings \( x_t : \Omega \to \mathbb{R}^n_t, t = 0, \ldots, T \) adapted to filtration\(^3\) \( \mathcal{F} = \{\mathcal{F}_t\}_{t=0}^T \) is called a policy or a decision. Since \( \mathcal{F}_0 \) is trivial, the first decision \( x_0 \) is deterministic. A policy \( \pi = \{x_0, x_1(\cdot), \ldots, x_T(\cdot)\} \) is feasible if it satisfies the feasibility constraints with probability 1 (w.p. 1). We denote by \( \Pi \) the set of feasible policies such that \( \left( f_0(x_0), f_1(x_1(\cdot), \cdot), \ldots, f_T(x_T(\cdot), \cdot) \right) \in \mathcal{Z}_{0,T} \). The optimization in (2.1) is over all feasible policies \( \pi \in \Pi \).

A famous quote 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. Therefore, in order to proceed, we consider a class of preference relations between possible realizations of random data defined by a family of mappings

\[ \mathcal{R}_{t,u} : \mathcal{Z}_{t,u} \to \mathcal{Z}_t, \quad 0 \leq t < u \leq T. \]

We refer to each \( \mathcal{R}_{t,u} \) as a preference mapping and to the family \( \mathfrak{R} = \{\mathcal{R}_{t,u}\}_{0 \leq t < u \leq T} \) as a preference system. Since \( \mathcal{Z}_0 \) is identified with \( \mathbb{R} \), we view \( \mathcal{R}_{0,T}(Z_0, \ldots, Z_T) \) as a real number for any \( (Z_0, \ldots, Z_T) \in \mathcal{Z}_{0,T} \).

\(^2\)Recall that \( \mathcal{Z}_{0,T} = \mathcal{Z}_0 \times \cdots \times \mathcal{Z}_T \).

\(^3\)We use bold notation \( x_t \) for (measurable) mappings in order to distinguish it from deterministic vector \( x_t \in \mathbb{R}^n_t \). Also by writing \( x_t(\cdot) \) we emphasize that this is a function of \( \omega \in \Omega \), i.e., is a random variable, rather than a deterministic vector. It is said that the sequence \( (x_0, x_1, \ldots, x_T) \) is adapted to the filtration if \( x_t(\cdot) \) is \( \mathcal{F}_t \)-measurable for every \( t = 1, \ldots, T \).
We assume that $\mathcal{R}_{0,T}$ coincides with the preference functional $\mathcal{R}$ of the reference problem (2.1), i.e., $\mathcal{R} = \mathcal{R}_{0,T}$.

**Definition 2.1 (Time consistent policies).** We say that an optimal policy $\bar{\pi} = \{\bar{x}_0, \ldots, \bar{x}_T\}$, solving the reference problem (2.1), is time consistent with respect to the preference system $\mathfrak{R}$, if at stages $t = 0, \ldots, T - 1$ the policy $\{\bar{x}_t, \ldots, \bar{x}_T\}$ is optimal for

$$
\min \mathcal{R}_{t,T} \left[ f_t(x_t, \omega), \ldots, f_T(x_{T-1}, \omega) \right],
$$

subject to $x_u \in \mathcal{X}_u(x_{u-1}, \omega)$, $u = t + 1, \ldots, T,$

(2.2)

conditional on $\mathcal{F}_t$.

In the above definition we only deal with preference mappings $\mathcal{R}_{t,T}: \mathcal{Z}_{t,T} \rightarrow \mathcal{Z}_t$ and the corresponding preference system $\{\mathcal{R}_{t,T}\}_{0 \leq t < T - 1}$. Nevertheless, considering preferences $\mathcal{R}_{t,u}$ for $u$ different from $T$ will be essential for our analysis. It could be noted that we allow for the preference mapping $\mathcal{R}_{t,T}$, $t = 1, \ldots, T - 1$, to depend on realizations of the data process up to time $t$, i.e., we have that $\mathcal{R}_{t,T}(Z_t, \ldots, Z_T)$ is $\mathcal{F}_t$-measurable. However, we do not allow $\mathcal{R}_{t,T}$ to depend on the decisions.

**Remark 2.2 (Scenario trees).** 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 $t = 0$ we have one root node denoted $\xi_0$. At stage $t = 1$ 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 $t = 1$ is denoted $\xi_1$, etc. at the later stages. A scenario, representing a realization (sample path or trajectory) of the data process, is a sequence $\xi_0, \ldots, \xi_T$ 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, $\xi_0, \ldots, \xi_T$ becomes a random (stochastic) process. With some abuse of the notation we use the same notation $\xi_0, \ldots, \xi_T$ for this random process and its particular realization (sample path), the exact meaning will be clear from the context. 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 (2.2) 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 $\mathcal{F}_t$ in this framework. Since $\bar{x}_{t-1}$ is $\mathcal{F}_t$ measurable, conditioning on $\mathcal{F}_t$ implies also conditioning on $\bar{x}_{t-1}$.

Definition 2.1 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 preference system $\mathfrak{R} = \{\mathcal{R}_{t,u}\}_{1 \leq t < u \leq T}$. This suggests the following basic questions: (i) what would be a ‘natural’ choice of mappings $\mathcal{R}_{t,u}$, (ii) what properties of mappings $\mathcal{R}_{t,u}$ 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. As we shall see the last question is closely related to decomposability of $R_{t,T}$ in terms of one-step mappings $\mathcal{R}_{t,t+1}$.

The minimal property that is required for the preference mappings is monotonicity.

**Definition 2.3** (Monotonicity). We say that preference mapping $R_{t,u}$ is **monotone** if for any $Z_{t,u}, Z'_{t,u} \in Z_{t,u}$ such that $Z_{t,u} \preceq Z'_{t,u}$ it follows that $R(Z_{t,u}) \geq R(Z'_{t,u})$. We say that preference mapping $R_{t,u}$ is **strictly monotone** if for any $Z_{t,u}, Z'_{t,u} \in Z_{t,u}$ such that $Z_{t,u} \succ Z'_{t,u}$ it follows that $R(Z_{t,u}) > R(Z'_{t,u})$. The preference system $\mathfrak{R} = \{\mathcal{R}_{t,u}\}_{0 \leq t < u \leq T}$ is said to be monotone (strictly monotone) if every preference mapping $R_{t,u}$ is monotone (strictly monotone).

**Remark 2.4** (Additive case). The above framework (2.1)–(2.2) is very general. The common case considered in the recent literature on risk averse stochastic optimization is when the reference problem is a function of the total cost. That is, when each $\mathcal{R}_{t,u}$ is representable as a function of $Z_t + \cdots + Z_u$, i.e.,

$$R_{t,u}(Z_t, \ldots, Z_u) := \rho_{t,T}(Z_t + \cdots + Z_u), \ 0 \leq t < u \leq T,$$

for some $\rho_{t,T}: Z_T \to Z_t, \ t = 0, \ldots, T$. We will refer to such framework as the additive case. We will also consider examples of natural and important preference systems which cannot be considered in the additive framework (2.3).

Following Artzner et al. (1999) we say that $\rho_{t,T}$ is a **coherent risk mapping** if it is subadditive, monotone, positively homogeneous and translation equivariant. The preference system (2.3) is monotone (strictly monotone) if $\rho_{t,T}$ are monotone (strictly monotone). In particular, let $\varrho: L_p(\Omega, \mathcal{F}, P) \to \mathbb{R}$ be a law invariant coherent risk measure and $\varrho|_{\mathcal{F}_t}$ be its conditional analogue. Then $\rho_{t,T} := \varrho|_{\mathcal{F}_t}$ is the corresponding coherent mapping. When $\varrho := \mathbb{E}$ is the expectation operator, the corresponding preference system is given by conditional expectations

$$R_{t,u}(Z_t, \ldots, Z_u) = \mathbb{E}|_{\mathcal{F}_t}[Z_t + \cdots + Z_u], \ 0 \leq t < u \leq T,$$

which corresponds to the risk neutral setting. As another example take $\varrho := \text{AV@R}_\alpha$, where the Average Value-at-Risk measure can be defined as

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

For $\alpha = 0$ the $\text{AV@R}_0$ is the expectation operator and for $\alpha = 1$ it becomes $\varrho = \text{ess.sup}$, the essential supremum operator. The risk measure $\text{AV@R}_\alpha$ is monotone but is not strictly monotone for $\alpha \neq 0$. The conditional analogue of the Average Value-at-Risk (2.4) is

$$\text{AV@R}_{\alpha|\mathcal{F}_t}(Z) = \text{ess inf}_{u_t \in L_{\infty}(\mathcal{F}_t)} \left\{ u_t + (1 - \alpha_t)^{-1}\mathbb{E}|_{\mathcal{F}_t}[Z - u_t]_+ \right\}, \ \alpha_t \in [0, 1).$$

\footnote{Note that since $Z_u \subset Z_T, \ u = t+1, \ldots, T$, the corresponding mapping $\rho_{t,T}: Z_u \to Z_t$ is defined as restriction of $\rho_{t,T}$ to $Z_u$.}

\footnote{Translation equivariance means that for any $Z_T \in Z_T$ and $Z_t \in Z_t$ it follows that $\rho_{t,T}(Z_T + Z_t) = \rho_{t,T}(Z_T) + Z_t$.}
Example 2.5 (The risk averse case). Let $Z_t := L_\infty(\Omega, \mathcal{F}_t, P)$ and\footnote{The conditional essential supremum is the smallest $\mathcal{F}_t$-random variable $(X_t, \text{say})$, so that $X_t \succeq Z_\tau$ for all $\tau$ with $t \leq \tau \leq u$. For a mathematically rigorous introduction of the essential supremum, as well as the essential infimum, we refer to Karatzas and Shreve (1998, Appendix A).}

$$R_{t,u}(Z_t, \ldots, Z_u) := \text{ess sup}_{\mathcal{F}_t} \{Z_t, \ldots, Z_u\}, \quad 0 \leq t < u \leq T,$$

The objective of the corresponding reference problem (2.1) is then given by the maximum of the essential supremum of the cost functions in the periods $t = 0, \ldots, T$.

In the additive case discussed in Remark 2.4, the value $R_{t,u}(Z_t, \ldots, Z_u)$ is a function of the sum $Z_t + \cdots + Z_u$ and is the same as value of $R_{t,T}$ applied to $Z_t + \cdots + Z_u$ for any $u = t + 1, \ldots, T$. That is, in that framework there is no point of considering preference mappings $R_{t,u}$ for $u$ different from $T$. On the other hand, the preference system of Example 2.5 is not additive and cannot be described in terms of the corresponding preference mappings $R_{t,T}$, $t = 0, \ldots, T - 1$, alone.

3 Two stage setting

It is informative at this point to discuss the two stage case, $T = 2$, since time inconsistency could already happen there. Consider the following two stage risk averse stochastic program

$$\begin{array}{ll}
\text{Min} & \mathcal{R}(f_0(x_0), f_1(x_1, \omega)), \\
\text{s.t.} & x_0 \in X_0, \ x_1(\omega) \in X_1(x_0, \omega),
\end{array} \tag{3.1}$$

where $X := \{x_1 : \Omega \to \mathbb{R}^n_1 | f_1(x_1, \cdot) \in Z\}$. The risk functional $\mathcal{R} : Z_{0,1} \to \mathbb{R}$ is defined on the space $Z_{0,1} = \mathbb{R} \times Z$ with $Z$ being a linear space of measurable functions $Z_1 : \Omega \to \mathbb{R}$. In order to deal with duality issues we consider the following two frameworks for defining the space $Z$. In one framework we use, as in the previous section, $Z := L_p(\Omega, \mathcal{F}, P)$, $p \in [1, \infty]$, with $P$ viewed as the reference probability measure (distribution). The space $Z$ is paired with the space $Z^* = L_q(\Omega, \mathcal{F}, P)$, where $q \in [1, \infty]$ is such that $1/p + 1/q = 1$, with the respective scalar product

$$\langle \zeta, Z \rangle = \int_\Omega \zeta(\omega)Z(\omega)P(d\omega), \ \zeta \in Z^*, \ Z \in Z.$$

This framework became standard in the theory of coherent risk measures. However, it is not applicable to the distributionally robust approach when the set of ambiguous distributions is defined by moment constraints and there is no reference probability measure. In that case we assume that $\Omega$ is a compact metric space and use the space $Z := C(\Omega)$ of continuous functions $Z : \Omega \to \mathbb{R}$. The dual $Z^*$ of this space is the space of finite signed Borel measures with the respective scalar product

$$\langle \mu, Z \rangle = \int_\Omega Z(\omega)\mu(d\omega), \ \mu \in Z^*, \ Z \in Z.$$
Note that in the first framework of $Z = L^p(\Omega, \mathcal{F}, P)$, an element $Z \in Z$ actually is a class of functions which can differ from each other on a set of $P$-measure zero.

The optimization in (3.1) is performed over $x_0 \in \mathcal{X}_0 \subset \mathbb{R}^{n_0}$ and $x_1 : \Omega \to \mathbb{R}^{n_1}$ such that $f_1(x_1(\cdot), \cdot) \in Z$. The feasibility constraints $x_1(\omega) \in \mathcal{X}_1(x_0, \omega)$ in problem (3.1) should be satisfied for $P$-almost every (a.e.) $\omega \in \Omega$ in the first framework, and for all $\omega \in \Omega$ in the second framework of $Z = C(\Omega)$. For $Z, Z' \in Z$ we use the notation $Z \succeq Z'$ to denote that $Z(\omega) \geq Z'(\omega)$ for a.e. $\omega \in \Omega$ in the first framework, and for all $\omega \in \Omega$ in the second framework. As before we write $Z \succ Z'$ to denote that $Z \succeq Z'$ and $Z \neq Z'$.

The common setting considered in the stochastic programming literature is to define $R(Z_0, Z_1) := \rho(Z_0 + Z_1)$, $(Z_0, Z_1) \in \mathbb{R} \times Z$, where $\rho : Z \to \mathbb{R}$ is a specified risk functional. This is the additive case discussed in Remark 2.4. In particular, when $\rho$ is the expectation operator this becomes the risk neutral formulation. However, there are many other possibilities to define risk functionals $R$ which are useful in various situations. For example, consider $R(Z_0, Z_1) := \max\{Z_0, \rho(Z_1)\}$. If moreover, in the framework of $Z = C(\Omega)$, we take $\rho(Z) := \sup_{\omega \in \Omega} Z(\omega)$, then the corresponding problem (3.1) can be viewed as a robust type problem with minimization of the worst possible outcome of the two stages. That is, if the second stage cost is bigger than the first stage cost for some scenarios, then the worst second stage cost is minimized. On the other hand, if the first stage cost is bigger for all scenarios, then the second stage problem is not considered. Similarly in the framework of $Z = L_\infty(\Omega, \mathcal{F}, P)$, we can take $\rho(Z) := \text{ess sup}(Z)$. This is the case of Example 2.5. As we shall discuss in Section 6 this is closely related to the problem of optimal stopping time.

In order to proceed we will need the following interchangeability result for functionals $\varrho : Z \to \mathbb{R}$. Consider a function $\psi : \mathbb{R}^n \times \Omega \to \mathbb{R} \cup \{+\infty\}$. Let

$$\Psi(\omega) := \inf_{y \in \mathbb{R}^n} \psi(y, \omega)$$

and

$$\mathcal{Y} := \{\eta : \Omega \to \mathbb{R}^n \mid \psi_\eta(\cdot) \in Z\},$$

where $\psi_\eta(\cdot) := \psi(\eta(\cdot), \cdot)$.

**Assumption 3.1.** In the framework of $Z = L^p(\Omega, \mathcal{F}, P)$, suppose that the function $\psi(y, \omega)$ is random lower semicontinuous, i.e., its epigraphical mapping is closed valued and measurable (such functions are called normal integrands in Rockafellar and Wets (1997)). In the framework of $Z = C(\Omega)$ suppose that the minimum of $\psi(y, \omega)$ over $y \in \mathbb{R}^n$ is attained for all $\omega \in \Omega$.

We have the following result about interchangeability of the minimum and risk operators (cf. Shapiro (2017)).

**Proposition 3.1.** Suppose that Assumption 3.1 holds, $\Psi \in Z$ and $\varrho : Z \to \mathbb{R}$ is monotone. Then

$$\varrho(\Psi) = \inf_{\eta \in \mathcal{Y}} \varrho(\psi_\eta) \quad (3.2)$$
and the implication

\[ \bar{\eta}(\cdot) \in \arg \min_{y \in \mathbb{R}^n} \psi(y, \cdot) \implies \bar{\eta} \in \arg \min_{\eta \in \mathbb{Y}} \varrho(\psi_\eta). \tag{3.3} \]

holds true. If moreover \( \varrho \) is strictly monotone, then the converse of (3.3) holds true as well, i.e.,

\[ \bar{\eta} \in \arg \min_{\eta \in \mathbb{Y}} \varrho(\psi_\eta) \implies \bar{\eta}(\cdot) \in \arg \min_{y \in \mathbb{R}^n} \psi(y, \cdot). \tag{3.4} \]

Equation (3.2) means that the minimization and risk operators can be interchanged, provided that the risk functional is monotone. Moreover, the pointwise minimizer in the left hand side of (3.3), if it exists, solves the corresponding minimization problem in the right hand side. In order to conclude the inverse implication (3.4), that the corresponding optimal functional solution is also the pointwise minimizer, the stronger condition of strict monotonicity is needed.

Consider now the problem (3.1), that depends on \( \omega \), and let

\[ V(x_0, \omega) := \inf_{x_1 \in \mathcal{X}_1(x_0, \omega)} \mathcal{R}(f_0(x_0), f_1(x_1, \omega)), \]

which can be viewed as value of the second stage problem. By Proposition 3.1 we have the following.

**Theorem 3.2.** Suppose that:

(i) the functional \( \varrho(\cdot) := \mathcal{R}(Z_0, \cdot) \) is monotone for any \( Z_0 \in \mathbb{R} \),

(ii) \( V(x_0, \cdot) \in \mathcal{Z} \) for all \( Z_0 \in \mathbb{R} \),

(iii) Assumption 3.1 holds for

\[ \psi(x_1, \omega) := \begin{cases} f_2(x_1, \omega) & \text{if } x_1 \in \mathcal{X}_1(x_0, \omega), \\ +\infty & \text{if } x_1 \notin \mathcal{X}_1(x_0, \omega). \end{cases} \]

Then the optimal value of the problem (3.1) is the same as the optimal value of

\[ \min_{x_0 \in \mathcal{X}_0} \mathcal{R}(f_1(x_0), V(x_0, \omega)). \tag{3.5} \]

Further, \((\bar{x}_0, \bar{x}_1)\) is an optimal solution of the two stage problem (3.1) if \( \bar{x}_0 \) is an optimal solution of the first stage problem (3.5) and\(^7\)

\[ \bar{x}_1(\cdot) \in \arg \min_{x_1 \in \mathcal{X}_1(\bar{x}_0, \cdot)} \mathcal{R}(f_0(\bar{x}_0), f_1(x_1, \cdot)). \tag{3.6} \]

Moreover, if \( \mathcal{R}(Z_0, \cdot) \) is strictly monotone, then \((\bar{x}_0, \bar{x}_1)\) is an optimal solution of problem (3.1) if and only if \( \bar{x}_0 \) is an optimal solution of the first stage problem and (3.6) holds.

\(^7\)By writing “\( \bar{x}_1(\cdot) \in \ldots \)” we mean that such inclusion holds for a.e. \( \omega \in \Omega \) in the setting of \( \mathcal{Z} = L_p(\Omega, \mathcal{F}, \mathbb{P}) \), and for all \( \omega \in \Omega \) in the setting of \( \mathcal{Z} = \mathcal{C}(\Omega) \).
Here, time consistency of a policy \((\tilde{x}_0, \tilde{x}_1)\), solving the two stage problem (3.1), means that \(\tilde{x}_1\) solves the respective second stage problem, i.e., condition (3.6) holds. That is, if \((\tilde{x}_0, \tilde{x}_1)\) is not time consistent, then there exists another feasible solution \(\tilde{x}_1\) such that \(f_2(\tilde{x}_1, \gamma) \succ f_2(\tilde{x}_1, \gamma')\). Without strict monotonicity of \(\mathcal{R}\), it could happen that problem (3.1) has optimal solutions which do not satisfy condition (3.6) and hence are not time consistent. That is, condition (3.6) is a sufficient but without strict monotonicity is not necessary for optimality. Such examples can be found, e.g., in Shapiro (2017) and for robust optimization were given in Delage and Iancu (2015). For example, for \(\mathcal{R}(Z_0, Z_1) := \max\{Z_0, \rho(Z_1)\}\) we have that if \(Z_1 \succ Z_1'\) are such that \(Z_0 > \rho(Z_1) > \rho(Z_1')\), then \(\mathcal{R}(Z_0, Z_1) = \mathcal{R}(Z_0, Z_1')\). That is, \(\mathcal{R}(Z_0, \cdot)\) is not strictly monotone. It could happen then that a second stage decision does not satisfy (3.6) and is not time consistent.

**Dual representation**

The space \(Z_{0,1} = \mathbb{R} \times Z\) can be equipped, for example, with the norm \(\|(Z_0, Z_1)\|_{0,1} := |Z_0| + \|Z_1\|\), where \(\|\cdot\|\) is the respective norm of the space \(Z\), and can be paired with the space \(Z_{0,1}^* = \mathbb{R} \times Z^*\) with the scalar product

\[
\langle (\zeta_0, \zeta_1), (Z_0, Z_1) \rangle := \zeta_0 Z_0 + \langle \zeta_1, Z_1 \rangle, \quad (\zeta_0, \zeta_1) \in Z_{0,1}^*, \quad (Z_0, Z_1) \in Z_{0,1}.
\]

Suppose that the functional \(\mathcal{R} : Z_{0,1} \to \mathbb{R}\) is convex and monotone. Then by the Klee-Nachbin-Namioka Theorem the functional \(\mathcal{R}\) is continuous in the norm topology of \(Z_{0,1}\) (cf. Ruszczyński and Shapiro (2006b, Proposition 3.1)). Suppose further that \(\mathcal{R}\) is positively homogeneous, i.e., \(\mathcal{R}(t Z_{0,1}) = t \mathcal{R}(Z_{0,1})\) for any \(t \geq 0\) and \(Z_{0,1} \in Z_{0,1}\). Then by the Fenchel-Moreau Theorem, \(\mathcal{R}\) has the dual representation

\[
\mathcal{R}(Z_0, Z_1) = \sup_{(\zeta_0, \zeta_1) \in \mathfrak{A}_{0,1}} \langle (\zeta_0, \zeta_1), (Z_0, Z_1) \rangle \tag{3.7}
\]

for some convex, bounded and weakly* closed set

\[
\mathfrak{A}_{0,1} \subset \{ (\zeta_0, \zeta_1) \in Z_{0,1}^*: \zeta_0 \geq 0, \ z_1 \geq 0 \}
\]

(cf. Ruszczyński and Shapiro (2006b, Theorem 2.2)). The subdifferential of \(\mathcal{R}\) is then given by

\[
\partial \mathcal{R}(Z_0, Z_1) = \text{arg max}_{(\zeta_0, \zeta_1) \in \mathfrak{A}_{0,1}} \langle (\zeta_0, \zeta_1), (Z_0, Z_1) \rangle.
\]

In particular \(\partial \mathcal{R}(0, 0) = \mathfrak{A}_{0,1}\).

**Example 3.3.** Consider risk functional of the form \(\mathcal{R}(Z_0, Z_1) := \varphi(Z_0, \rho(Z_1))\), where \(\rho : Z \to \mathbb{R}\) is a coherent risk measure and \(\varphi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}\) is a convex monotone positively homogeneous function. It follows then that \(\mathcal{R}(\cdot, \cdot)\) is convex monotone and positively homogeneous. Let \(\rho(Z) = \sup_{\zeta \in \mathfrak{A}} \langle \zeta, Z \rangle\) be the dual representation of \(\rho\), where \(\mathfrak{A}\) is a convex bounded weakly* closed subset of \(Z^*\). Then \(\partial \varphi(x_0, x_1)\) consists of vectors (subgradients) \((\gamma_0, \gamma_1)\) such that

\[
\varphi(y_0, y_1) - \varphi(x_0, x_1) \geq \gamma_0 (y_0 - x_0) + \gamma_1 (y_1 - x_1)
\]

11
for all \((y_0, y_1) \in \mathbb{R}^2\). Since \(\varphi\) is monotone, it follows that \(\partial \varphi(x_0, x_1) \subset \mathbb{R}_+^2\). Consequently the representation (3.7) holds with

\[ A_{0,1} = \partial \mathcal{R}(0, 0) = \{(\zeta_0, \zeta_1) \in \mathbb{R} \times \mathbb{Z}^*: \zeta_1 = \gamma \zeta, \ z \in A, (\zeta_0, \gamma) \in \partial \varphi(0, 0)\}. \]

For example let \(\varphi(x_0, x_1) := \max\{x_0, x_1\}\), and hence \(\mathcal{R}(Z_0, Z_1) = \max\{Z_0, \rho(Z_1)\}\). Then \(\partial \varphi(0, 0) = \{(t, 1-t): t \in [0, 1]\}\) and

\[ A_{0,1} = \{(t, (1-t)\zeta): t \in [0, 1], \ z \in A\}. \]

4 Decomposability and dynamic equations

Let us start with definition of the following basic decomposability concept.

**Definition 4.1** (Recursivity). The preference system \(\{\mathcal{R}_{t,u}\}_{0 \leq t < u \leq T}\) is said to be *recursive*, if

\[ \mathcal{R}_{t,u}(Z_t, \ldots, Z_u) = \mathcal{R}_{t,v}(Z_t, \ldots, Z_{v-1}, \mathcal{R}_{v,u}(Z_v, \ldots, Z_u)), \quad (4.1) \]

for any \(0 \leq t < v < u \leq T\) and \((Z_t, \ldots, Z_u) \in \mathcal{Z}_{t,u}\).

In the additive case (discussed in Remark 2.4), when \(\mathcal{R}_{t,T}(Z_t, \ldots, Z_T) = \rho_{t,T}(Z_t + \cdots + Z_T)\), \(t = 0, \ldots, T - 1\), the recursive property can be written as

\[ \rho_{t,T}(\rho_{v,T}(Z)) = \rho_{t,T}(Z), \ Z \in \mathcal{Z}_T, \ 0 \leq t < v \leq T - 1. \quad (4.2) \]

Note that since \(\rho_{v,T}(Z) \in \mathcal{Z}_v\), we have that

\[ \rho_{t,T}(\rho_{v,T}(Z)) = \rho_{t,v}(\rho_{v,T}(Z)). \]

By applying (4.2) recursively for \(v = t + 1, \ldots, T - 1\) means that \(\rho_{t,T}\) can be decomposed as

\[ \rho_{t,T}(Z) = \rho_{t,T}(\rho_{t+1,T}(\cdots \rho_{T-1,T}(Z))) \], \(Z \in \mathcal{Z}_T\).

If moreover \(\rho_{t,T}\) is translation equivariant, this becomes

\[ \rho_{t,T}(Z_t + \cdots + Z_T) = Z_t + \rho_{t,T}(Z_{t+1} + \rho_{t+1,T}(Z_{t+2}) + \cdots + \rho_{T-1,T}(Z_T)). \]

For law invariant coherent risk measure \(\varrho\) and \(\rho_{t,T} := \varrho_{[\mathcal{F}_t]}\), the recursive property (4.2) can hold only in two cases – for the ‘expectation’ and the ‘ess sup’ operators (cf. Kupper and Schachermayer (2009)). For example the Average Value-at-Risk preference system, \(\rho_{t,T} := \text{AV@R}_\alpha[\mathcal{F}_T]\), is not recursive for \(\alpha \in (0, 1)\). Recursive preference system, in the additive case, can be constructed in the nested form

\[ \rho_{t,T}(Z) := \phi_t(\phi_{t+1}(\cdots \phi_{T-1}(Z))), \ Z \in \mathcal{Z}_T, \quad (4.3) \]
where $\phi_s: Z_{s+1} \rightarrow Z_s$, $s = 1, \ldots, T - 1$, are one-step mappings. For example, taking $\rho := AV@R_\alpha$ this becomes

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

the so-called nested Average Value-at-Risk mappings. As it was pointed out above, for $\alpha \in (0,1)$ these nested Average Value-at-Risk mappings are different from the $AV@R_\alpha|_{\mathcal{F}_t}$.

Consider now the general case of the preference system $\{\mathcal{R}_{t,u}\}_{1 \leq t < u \leq T}$. The recursive property (4.1) implies that $\mathcal{R}_{t,u}$ can be decomposed in terms of one step mappings $\mathcal{R}_{s,s+1}$, $s = t, \ldots, u - 1$, as

$$\mathcal{R}_{t,u}(Z_t, \ldots, Z_u) = \mathcal{R}_{t,t+1}\left(Z_t, \mathcal{R}_{t+1,t+2}(Z_{t+1}, \cdots, \mathcal{R}_{u-1,u}(Z_{u-1}, Z_u))\right), \quad (4.4)$$

Conversely recursive preference mappings can be constructed in the form (4.4) by choosing one step mappings $\mathcal{R}_{s,s+1}: Z_s \times Z_{s+1} \rightarrow Z_s$, $s = 1, \ldots, T - 1$.

The recursive property (4.1) and monotonicity of the preference system allow to write the following dynamic programming equations for the reference problem (2.1), derivations are similar to the two stage case discussed in Section 3. At the last stage the cost-to-go function is defined as

$$V_T(x_{T-1}, \omega) := \text{ess inf}_{x_T \in X_T(x_{T-1}, \omega)} f_T(x_T, \omega). \quad (4.5)$$

Since $\mathcal{R} = \mathcal{R}_{0,T}$ and by recursiveness of the preference system we can write

$$\mathcal{R}\left[f_0(x_0), \ldots, f_T(x_T, \omega)\right] = \mathcal{R}_{0,T}\left[f_0(x_0), \ldots, f_{T-2}(x_{T-2}, \omega), \mathcal{R}_{T-1,T}(f_{T-1}(x_{T-1}, \omega), f_T(x_T, \omega))\right]. \quad (4.6)$$

For given $x_0, \ldots, x_T$ consider minimization of the right hand side of (4.6) with respect to $x_T \in X_T(x_{T-1}, \omega)$. By monotonicity of $\mathcal{R}_{0,T-1}$ this can be written as

$$\mathcal{R}_{0,T-1}\left[f_0(x_0), \ldots, f_{T-2}(x_{T-2}, \omega), \text{ess inf}_{x_T \in X_T(x_{T-1}, \omega)} \mathcal{R}_{T-1,T}(f_{T-1}(x_{T-1}, \omega), f_T(x_T, \omega))\right].$$

Moreover by using the interchangeability principle we have that

$$\text{ess inf}_{x_T \in X_T(x_{T-1}, \omega)} \mathcal{R}_{T-1,T}(f_{T-1}(x_{T-1}, \omega), f_T(x_T, \omega)) = \mathcal{R}_{T-1,T}(f_{T-1}(x_{T-1}, \omega), V_T(x_{T-1}, \omega)),$$

assuming that $V_T(x_{T-1}, \cdot) \in Z_T$. Continuing this backward in time we obtain at stages $t = T - 1, \ldots, 1$, the cost-to-go functions

$$V_t(x_{t-1}, \omega) := \text{ess inf}_{x_t \in X_t(x_{t-1}, \omega)} \mathcal{R}_{t,t+1}(f_t(x_t, \omega), V_{t+1}(x_t, \omega)), \quad (4.7)$$

representing the corresponding dynamic programming equations. Finally at the first stage the problem

$$\text{Min}_{x_0 \in X_0} \mathcal{R}_{0,1}(f_0(x_0), V_1(x_0, \omega)) \quad (4.8)$$

should be solved, with the optimal value of (4.8) equal to the optimal value of the multistage problem (2.1). In a rudimentary form such approach to writing dynamic equations with relation to time consistency was outlined in Shapiro (2009).
Definition 4.2 (Dynamic programming equations). We say that a policy \( \pi = (x_0, x_1, \ldots, x_T) \) satisfies the dynamic programming equations if

\[
\begin{align*}
x_T(\cdot) & \in \arg\min_{x_T \in \mathcal{X}_T(x_{T-1}, \cdot)} f_T(x_T, \cdot), \\
x_t(\cdot) & \in \arg\min_{x_t \in \mathcal{X}_t(x_{t-1}, \cdot)} \mathcal{R}_{t,t+1}(f_t(x_t, \cdot), V_{t+1}(x_t, \cdot)), \quad t = 1, \ldots, T - 1, \\
x_0 & \in \arg\min_{x_0 \in \mathcal{X}_0} \mathcal{R}_{0,1}(f_0(x_0), V_1(x_0, \cdot)).
\end{align*}
\] (4.9) (4.10) (4.11)

If a policy satisfies the dynamic programming equations, then it is optimal for the reference multistage problem (2.1) and is time consistent. Without strict monotonicity it could happen that a policy, which is optimal for the reference problem (2.1), does not satisfy the dynamic programming equations and is not time consistent. As it was discussed in Section 3 this could happen even in the two stage case and a finite number of scenarios.

Remark 4.3. Consider the additive case where \( \mathcal{R}_{t,t+1}(Z_t, Z_{t+1}) = \rho_{t,T}(Z_t + Z_{t+1}) \). If moreover mappings \( \rho_t \) 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 coherent risk mappings \( \phi_t \), as in (4.3). In that case equations (4.5)–(4.8) coincide with the respective equations of the additive case (cf. Ruszczyński and Shapiro (2006a)).

Example 4.4. Let us define one step mappings as

\[
\mathcal{R}_{s,s+1}(Z_s, Z_{s+1}) := \max \left\{ Z_s, \varrho_{|F_s}(Z_{s+1}) \right\}, \quad s = 0, \ldots, T - 1,
\]

and the corresponding preference risk mappings of the form (4.4), where \( \varrho \) is a law invariant coherent risk measure. In particular for \( \varrho := \text{ess sup} \) we obtain the preference system of Example 2.5. Here the dynamic equations (4.7) take the form

\[
V_t(x_t, \omega) = \text{ess inf}_{x_t \in \mathcal{X}_t(x_{t-1}, \omega)} \max \left\{ f_t(x_t, \omega), \varrho_{|F_t}[V_{t+1}(x_t, \omega)] \right\},
\]

and the reference problem can be viewed as minimization of the worst possible outcome over the considered period of time measured in terms of the risk measure \( \varrho \).

As we shall see in Section 6.1, this example is closely related to the stopping time risk aversion formulation of multistage programs.

5 Time consistency

The following concept of dynamic consistency (also called time consistency by some authors) 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 5.1 (Dynamical consistency). The preference system \( \{ R_{t,u} \}_{1 \leq t < u \leq T} \) is said to be dynamically consistent if for \( 1 \leq s < t < u \leq T \) and \( (Z_s, \ldots, Z_u), (Z'_s, \ldots, Z'_u) \in Z_{t,u} \) such that \( Z_\tau = Z'_\tau \), \( \tau = s, \ldots, t-1 \), the following ‘forward’ implication holds

\[
R_{t,u}(Z_t, \ldots, Z_u) \geq R_{t,u}(Z'_t, \ldots, Z'_u) \implies R_{s,u}(Z_s, \ldots, Z_u) \geq R_{s,u}(Z'_s, \ldots, Z'_u). \tag{5.1}
\]

It turns out that the above ‘forward’ property of dynamic consistency is not always sufficient to ensure that every optimal policy is time consistent. For that we need a stronger notion of dynamic consistency (cf. Shapiro et al. (2014, Section 6.8.5)).

Definition 5.2 (Strict dynamical consistency). A dynamically consistent preference system \( \{ R_{t,u} \}_{1 \leq t < u \leq T} \) is said to be strictly dynamically consistent if in addition to (5.1) the following implication holds

\[
R_{t,u}(Z_t, \ldots, Z_u) \succ R_{t,u}(Z'_t, \ldots, Z'_u) \implies R_{s,u}(Z_s, \ldots, Z_u) \succ R_{s,u}(Z'_s, \ldots, Z'_u) \tag{5.2}
\]

for all \( 1 \leq s < t < u \leq T \).

Note that it follows from (5.1) that

\[
R_{t,u}(Z_t, \ldots, Z_u) = R_{t,u}(Z'_t, \ldots, Z'_u) \implies R_{s,u}(Z_s, \ldots, Z_u) = R_{s,u}(Z'_s, \ldots, Z'_u).
\]

Recall that in the additive case, \( R_{t,T}(Z_t, \ldots, Z_T) \) is given by \( \rho_{t,T}(Z_t + \cdots + Z_T) \). In that case condition (5.1) implies that

\[
Z, Z' \in Z_T \text{ and } \rho_{t,T}(Z) \succeq \rho_{t,T}(Z') \implies \rho_{s,T}(Z) \succeq \rho_{s,T}(Z'), \quad 1 \leq s < t \leq T - 1. \tag{5.3}
\]

Conversely, if moreover \( \rho_{s,T} \) is translation equivariant, then we can write for \( 1 \leq s < t \leq T \),

\[
\rho_{t,T}(Z_s + \cdots + Z_T) = Z_s + \cdots + Z_{t-1} + \rho_{t,T}(Z_t + \cdots + Z_T),
\]

and hence condition (5.3) implies (5.1) for \( u = T \). If \( \rho_{t,T} := E_{|T_t} \), then for \( s < t \) we have that \( \rho_{s,T}(Z) = E_{|T_s} [E_{|T_t}(Z)] \) and hence this preference system is dynamically consistent. In fact it is not difficult to see that this preference system is strictly dynamically consistent.

We have the following relation between recursiveness and dynamic consistency.

**Proposition 5.3.** Suppose that preference mappings \( R_{t,u}, 1 \leq t < u \leq T \), are monotone (strictly monotone) and recursive. Then \( \{ R_{t,u} \}_{1 \leq t < u \leq T} \) is dynamically consistent (strictly dynamically consistent).

**Proof.** We need to verify implication (5.1). By the recursiveness, for \( 1 \leq s < t < u \leq T \) we have

\[
R_{s,u}(Z_s, \ldots, Z_u) = R_{s,t}(Z_s, \ldots, Z_{t-1}, R_{t,u}(Z_t, \ldots, Z_u)),
\]

\[
R_{s,u}(Z'_s, \ldots, Z'_u) = R_{s,t}(Z'_s, \ldots, Z'_{t-1}, R_{t,u}(Z'_t, \ldots, Z'_u)).
\]

Since \( Z_\tau = Z'_\tau \), \( \tau = s, \ldots, t-1 \), by monotonicity (strict monotonicity) of \( R_{s,t} \) the implication (5.1) (the implication (5.2)) follows.
In the additive case the converse of Proposition 5.3 can be given in the following framework.

**Proposition 5.4** (Ruszczyński (2010), Theorem 1). *Suppose the following hold true:
(i) \( \{\mathcal{R}_{t,u}\}_{1 \leq t < u \leq T} \) is dynamically consistent,
(ii) it holds that
\[
\mathcal{R}_{s,t}(Z_s, \ldots, Z_t) = \mathcal{R}_{s,u}(Z_s, \ldots, Z_t, 0, \ldots, 0), \quad 1 \leq s < t < u \leq T,
\]
(iii) for \( Z_t \in \mathcal{Z}_t \),
\[
\mathcal{R}_{t,u}(Z_t, 0, \ldots, 0) = Z_t, \quad 1 \leq t < u \leq T.
\]
Then the mappings \( \mathcal{R}_{s,t}, 1 \leq s < t \leq T \), are recursive.*

Note that conditions (5.4) and (5.5) are adjusted to the additive case. These conditions do not hold, for instance, in the settings of Examples 2.5 and 4.4.

Similar to the additive case we have the following result (cf. Shapiro et al. (2014, Propositions 6.80)).

**Proposition 5.5.** *The following holds true:
(i) If the preference system is dynamically consistent and \( \bar{\pi} \in \Pi \) is the unique optimal solution of the reference problem (2.1), 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.*

### 6 Optimal stopping

In this section we discuss a combination of the optimal stopping time and risk measures. Recall that a *stopping time*, adapted to the filtration \( \mathcal{F} = (\mathcal{F}_0, \ldots, \mathcal{F}_T) \), is a random variable \( \tau : \Omega \to \{0, \ldots, T\} \) such that \( \{\omega: \tau(\omega) = t\} \in \mathcal{F}_t \) for \( t = 0, \ldots, T \). We denote by \( \Omega_\tau \) the set of stopping times (adapted to the filtration \( \mathcal{F} \)). We have that \( \Omega \) is the union of the disjoint sets
\[
\Omega_\tau := \{\omega: \tau(\omega) = t\}, \quad t = 0, \ldots, T,
\]
and hence
\[
\mathbb{1}_\Omega = \sum_{t=0}^T \mathbb{1}_{\{\tau = t\}}.
\]
Note that \( \mathbb{1}_{\{\tau = t\}} \) are \( \mathcal{F}_t \) measurable for \( 0 \leq s \leq t \leq T \) and hence \( \mathbb{1}_{\{\tau \leq t\}} \) and \( \mathbb{1}_{\{\tau > t\}} = \mathbb{1}_\Omega - \mathbb{1}_{\{\tau \leq t\}} \) are also \( \mathcal{F}_t \) measurable for \( t = 0, \ldots, T \). Moreover \( \mathbb{1}_{\{\tau \leq t\}} Z_\tau = \mathbb{1}_{\{\tau = t\}} Z_t \) and thus for \( (Z_0, \ldots, Z_T) \in \mathcal{Z}_0,T \) it follows that
\[
Z_\tau = \sum_{t=0}^T \mathbb{1}_{\{\tau = t\}} Z_t = \sum_{t=0}^T \mathbb{1}_{\{\tau = t\}} Z_t.
\]
By replacing \( Z_t \) in (6.2) with \( Z_t' := Z_0 + \cdots + Z_t \) we obtain
\[
Z_0 + \cdots + Z_\tau = Z_\tau' = \sum_{t=0}^T \mathbb{1}_{\{\tau = t\}} Z_t' = \sum_{t=0}^T \sum_{i=0}^t \mathbb{1}_{\{\tau = i\}} Z_i
\]
\[
= \sum_{t=0}^T \sum_{i=t}^T \mathbb{1}_{\{\tau = i\}} Z_t = \sum_{t=0}^T \mathbb{1}_{\{\tau \geq t\}} Z_t.
\]
\(^8\)By \( \mathbb{1}_A \) we denote the indicator function of set \( A \).
Note that since \( \mathcal{F}_0 = \{ \emptyset, \Omega \} \) is trivial, we have that \( \{ \omega : \tau(\omega) = 0 \} \) is either \( \Omega \) or \( \emptyset \), and hence either \( 1_{\{ \tau = 0 \}} Z_0 = Z_0 \) or \( 1_{\{ \tau = 0 \}} Z_0 = 0 \).

Let \( \varrho_t|_{\mathcal{F}_t} : \mathcal{Z}_{t+1} \to \mathcal{Z}_t, t = 0, \ldots, T - 1 \), be translation equivariant\(^9\) mappings and consider the corresponding mappings \( \rho_{s,t} : \mathcal{Z}_t \to \mathcal{Z}_s \) represented in the nested form

\[
\rho_{s,t}(\cdot) := \varrho_{s|\mathcal{F}_s}\left(\varrho_{s+1|\mathcal{F}_{s+1}}(\cdots \varrho_{t-1|\mathcal{F}_{t-1}}(\cdot)\right), 0 \leq s < t \leq T. \tag{6.4}
\]

Since \( \mathcal{F}_0 \) is trivial, the corresponding functional \( \rho_{0,T} : \mathcal{Z}_T \to \mathbb{R} \) is real valued. By (6.2) and using the translation equivariance of the mappings \( \varrho_{t|\mathcal{F}_t} \) we can write

\[
\rho_{0,T}(Z_T) = 1_{\{ \tau = 0 \}} Z_0 + \varrho_{0|\mathcal{F}_0}\left(1_{\{ \tau = 1 \}} Z_1 + \cdots + \varrho_{T-1|\mathcal{F}_{T-1}}(1_{\{ \tau = T \}} Z_T)\right). \tag{6.5}
\]

In the risk neutral case when \( \varrho_{t|\mathcal{F}_t} := \mathbb{E}_{|\mathcal{F}_t} \), we have that \( \rho_{s,t} = \mathbb{E}_{|\mathcal{F}_s} \) for \( 0 \leq s < t \leq T \), in particular \( \rho_{0,T} = \mathbb{E} \), hence

\[
\begin{align*}
\mathbb{E}(Z_T) &= \mathbb{E}\left[ \sum_{t=0}^T 1_{\{ \tau = t \}} Z_t \right] \quad \text{and} \\
\mathbb{E}(Z_0 + \cdots + Z_T) &= \mathbb{E}\left[ \sum_{t=0}^T 1_{\{ \tau \geq t \}} Z_t \right].
\end{align*}
\]

**Remark 6.1 (Duality).** It is also possible to view the risk averse stopping time formulation (6.5) from the following distributionally robust point of view. Recall that a coherent risk measure \( \varrho : L_p(\Omega, \mathcal{F}, P) \to \mathbb{R} \) can be represented in the dual form

\[
\varrho(Z) = \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q[Z],
\]

where \( \mathfrak{M} \) is a set of probability measures absolutely continuous with respect to the reference probability measure \( P \) and such that the densities \( dQ/dP, Q \in \mathfrak{M} \), form a bounded convex weakly* closed set \( \mathfrak{A} \subset L_q(\Omega, \mathcal{F}, P) \) in the dual space \( L_q(\Omega, \mathcal{F}, P) = L_p(\Omega, \mathcal{F}, P)^* \). The functional \( \varrho \) is a coherent risk measure, it is law invariant iff the set \( \mathfrak{A} \) is law invariant in the sense that if \( \zeta \in \mathfrak{A} \) and \( \zeta' \) is a density distributionally equivalent\(^{10}\) to \( \zeta \), then \( \zeta' \in \mathfrak{A} \). This holds even if the reference probability measure \( P \) has atoms (cf. Shapiro (2017, Theorem 2.3)). Note that

\[
\sup_{Q \in \mathfrak{M}} \mathbb{E}_Q[\cdot] = \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q\left[ \mathbb{E}_Q[\cdot | \mathcal{F}_t] \right] \leq \sup_{Q \in \mathfrak{M}} \mathbb{E}_Q\left[ \text{ess sup}_{Q \in \mathfrak{M}} \mathbb{E}_Q[\cdot | \mathcal{F}_t] \right], \tag{6.6}
\]

where \( \varrho_{t|\mathcal{F}_t}(\cdot) = \text{ess sup}_{Q \in \mathfrak{M}} \mathbb{E}_Q[\cdot | \mathcal{F}_t] \) can be viewed as the conditional counterpart of the corresponding functional \( \varrho \). Equality in (6.6) holds only in rather specific cases. In general the functional \( \varrho \) has the recursive property \( \varrho(\cdot) = \varrho(\varrho_{t|\mathcal{F}_t}(\cdot)) \) only when the set \( \mathfrak{M} \) is a singleton or consists of all probability measures absolutely continuous with respect to the reference measure (cf. Kupper and Schachermayer (2009)).

---

\(^9\)Recall that translation equivariance means that \( \varrho_{t|\mathcal{F}_t}(Z_{t+1} + Z_t) = \varrho_{t|\mathcal{F}_t}(Z_{t+1}) + Z_t \) for any \( Z_{t+1} \in \mathcal{Z}_{t+1} \) and \( Z_t \in \mathcal{Z}_t \).

\(^{10}\)It is said that \( \zeta \) and \( \zeta' \) are distributionally equivalent if \( P(\zeta \leq z) = P(\zeta' \leq z) \) for all \( z \in \mathbb{R} \).
The nested functional $\rho_{0,T}$, defined in (6.4), is decomposable, i.e., has the recursive property (4.2). For not decomposable (law invariant) risk measure $\varrho$ the corresponding nested stopping objective $\rho_{0,T}(Z_\tau)$, of the form (6.5), is different from $\varrho(Z_\tau)$. As we shall see below the nested formulation of risk averse stopping time is suitable for writing dynamic programming equations and in the sense of nested decomposition is time consistent.

6.1 Multistage risk averse optimal stopping time

By employing nested stopping time risk measures $\rho_{0,T}$ of the form (6.4), and by using (6.2), the corresponding risk averse, multistage optimal stopping time problem can be written either as\footnote{Recall that $\Pi$ denotes the set of policies satisfying the feasibility constraints of problem (2.1).}

\begin{equation}
\min_{\pi \in \Pi} \sup_{t \in \Sigma} \rho_{0,T} (f_\tau(x_\tau, \omega))
\end{equation}

or

\begin{equation}
\min_{\pi \in \Pi} \inf_{t \in \Sigma} \rho_{0,T} (f_\tau(x_\tau, \omega)),
\end{equation}

with

$\rho_{0,T}(f_\tau(x_\tau, \omega)) = 1_{\{\tau=0\}} f_0(x_0) + \varrho_{0,T_0} \left( 1_{\{\tau=1\}} f_1(x_1, \omega) + \cdots + \varrho_{T-1,T_{T-1}} (1_{\{\tau=T\}} f_T(x_T, \omega)) \right)$.

Problems (6.7) and (6.8) can be viewed as the respective pessimistic (worst case) and optimistic formulations. Note that either with probability one $\tau = 0$, in which case $1_{\{\tau=0\}} f_0(x_0) = f_0(x_0)$, or with probability one $\tau > 0$, in which case $1_{\{\tau=0\}} f_0(x_0) = 0$.

Remark 6.2. It is also possible to consider optimization of $\rho_{0,T}(f_0(x_0) + \cdots + f_\tau(x_\tau, \omega))$ by using (6.3) rather than (6.2), i.e., by replacing the cost function $f_\tau(x_\tau, \omega)$ with the cumulative cost $f_0(x_0) + \cdots + f_\tau(x_\tau, \omega)$ (for the risk neutral case and fixed stopping time $\tau$ this type of problems were considered recently in Guigues (2018)). These formulations are equivalent and we concentrate below on the formulations (6.7) and (6.8) above.

The problems (6.7) and (6.8) are twofold, they consist in finding an optimal policy $\pi^* = (x_0^*, x_1^*, \ldots, x_T^*)$ and an optimal stopping time $\tau^* \in \mathcal{F}$. In the risk neutral case when $\varrho_{1,T} = E_{\mathcal{F}_T}$, for given (fixed) policy $\pi$, these problems become the classical problem of stopping time for the process

\begin{equation}
Z_t(\omega) := f_t(x_t(\omega), \omega).
\end{equation}

For a given stopping time $\tau \in \mathcal{F}$ we can write the corresponding dynamic programming equations, of the form (4.7), for the problems (6.7) and (6.8) (cf. Ruszczyński and Shapiro (2006a))

\begin{align}
V_T^T(x_{T-1}, \omega) &:= \essinf_{x_T \in \mathcal{X}_T(x_{T-1}, \omega)} 1_{\{\tau=T\}} f_T(x_T, \omega), \quad (6.10) \\
V_t^T(x_{t-1}, \omega) &:= \essinf_{x_t \in \mathcal{X}_t(x_{t-1}, \omega)} 1_{\{\tau=t\}} f_t(x_t, \omega) + \varrho_{t,T} (V_{t+1}^T(x_t, \omega)), \quad (6.11)
\end{align}

$t = 1, \ldots, T - 1$, and the first stage problem at $t = 0$ is

\begin{equation}
\min_{x_0 \in \mathcal{X}_0} f_0(x_0) + \varrho_{0,T_0} (V_1^T(x_0, \omega)).
\end{equation}
6.1.1 The min-max stopping problem

Let us consider the min-max problem (6.7) first. For a fixed policy \( \pi = (x_0, x_1, \ldots, x_T) \) we need to solve the optimal stopping time problem

\[
\begin{align*}
\max_{\tau \in \mathcal{T}} \left\{ \rho_{0,T}(Z_{\tau}) = 1_{\{\tau = 0\}} Z_0 + g_0, F_0 \left( 1_{\{\tau = 1\}} Z_1 + \cdots + g_{T-1}, F_{T-1} (1_{\{\tau = T\}} Z_T) \right) \right\},
\end{align*}
\]

with \( Z_t \) given in (6.9). The following is a natural extension of the classical results in the risk neutral case to the considered risk averse setting (6.12), e.g., Shiryaev (1978, Section 2.2).

**Definition 6.3 (Snell envelope).** Let \((Z_0, \ldots, Z_T) \in Z_{0,T}\) be a stochastic process. The risk-averse Snell envelope is the stochastic process

\[
E_T := Z_T,
\]

\[
E_t := \max \left\{ Z_t, \varrho_t | F_t \left( E_{t+1} \right) \right\}, \quad t = 0, \ldots, T - 1,
\]

defined in backwards recursive way.

For \( m = 0, \ldots, T \), consider \( \mathcal{X}_m := \{ \tau \in \mathcal{I}: \tau \geq m \} \), the optimization problem

\[
\max_{\tau \in \mathcal{X}_m} \rho_{0,T}(Z_{\tau}),
\]

and

\[
\tau^*_{m}(\omega) := \min \{ t: E_t(\omega) = Z_t(\omega), \ m \leq t \leq T \}, \ \omega \in \Omega.
\]

Denote by \( v_m \) the optimal value of the problem (6.14). Of course, for \( m = 0 \), the problem (6.14) coincides with problem (6.12) and \( v_0 \) is the optimal value of problem (6.12). Note that by the recursive property (4.2) we have that\(^{12}\) \( \rho_{0,T}(Z_{\tau}) = \rho_{0,m}(\rho_{m,T}(Z_{\tau})), \ m = 1, \ldots, T. \)

The following assumption was used by several authors under different names:

\[
\varrho_{t,F_t}(1_A \cdot Z) = 1_A \cdot \varrho_{t,F_t}(Z), \quad \text{for all } A \in F_t, \ t = 0, \ldots, T - 1.
\]

For coherent risk mappings \( \varrho_{t,F_t} \) it always holds (cf. Shapiro et al. (2014, Theorem 6.70)).

The following can be compared with classical results in the risk neutral case (cf., e.g., Bingham and Peskir (2008, Theorem 1)).

**Theorem 6.4 (Verification theorem).** Let \( \varrho_{t,F_t}: Z_{t+1} \to Z_t, \ t = 0, \ldots, T - 1 \), be monotone translation equivariant mappings possessing property (6.16) and \( \rho_{s,t}, \ 0 \leq s < t \leq T, \) be the corresponding nested functionals defined in (6.4). Then for \((Z_0, \ldots, Z_T) \in Z_{0,T}\) the following holds:

(i) for \( m = 0, \ldots, T, \)

\[
E_m \succeq \rho_{m,T}(Z_{\tau}), \ \forall \tau \in \mathcal{X}_m,
\]

\[
E_m = \rho_{m,T}(Z_{\tau^*_m}),
\]

\(^{12}\)By the definition \( \rho_{T,T}(Z_T) \equiv Z_T. \)
(ii) the stopping time $\tau_m^*$, from (6.15), is optimal for the problem (6.14),
(iii) if $\hat{\tau}_m$ is an optimal stopping time for the problem (6.14), then $\hat{\tau}_m \geq \tau_m^*$,
(iv) $u_m = g_{0,m}(E_m)$, $m = 1, \ldots, T$, and $u_0 = E_0$.

**Proof.** We use induction in $m$ going backwards in time. Recall that $E_T = Z_T$ and hence the
assertions follow for $m = T$. Now let $m = T - 1$ and $\tau \in \mathcal{F}_{T-1}$. Since $\rho_{T-1,T} = \rho_{T-1|F_{T-1}}$, by
using the translation equivariance and property (6.16), we can write

$$
\rho_{T-1,T}(Z_\tau) = \rho_{T-1|F_{T-1}} \left( \mathbb{1}_{\{\tau = T\}} Z_{T-1} + \mathbb{1}_{\{\tau = T\}} Z_T \right)
= \mathbb{1}_{\{\tau = T\}} Z_{T-1} + \mathbb{1}_{\{\tau = T\}} \rho_{T-1|F_{T-1}}(Z_T).
$$

We have that $\Omega$ is the union of the disjoint sets $\Omega^*_T$ and $\Omega_T$ (defined in (6.1)), and hence
(recall that $E_T = Z_T$)

$$
\mathbb{1}_{\{\tau = T\}} Z_{T-1} + \mathbb{1}_{\{\tau = T\}} \rho_{T-1|F_{T-1}}(Z_T) \leq \max\{Z_{T-1}, \rho_{T-1|F_{T-1}}(E_T)\} = E_{T-1}. \quad (6.17)
$$

It follows that $E_{T-1} \leq \rho_{T-1,T}(Z_\tau)$.

Conditional on the event $\{Z_{T-1} \geq \rho_{T-1|F_{T-1}}(E_T)\}$ we have: $E_{T-1} = Z_{T-1}$ and $\tau^*_{T-1} = T - 1$, and

$$
\rho_{T-1,T}(Z_{\tau^*_{T-1}}) = \rho_{T-1|F_{T-1}}(Z_{T-1}) = Z_{T-1} = E_{T-1},
$$

and $\tau^*_{T-1}$ is optimal for the corresponding problem (6.14). Otherwise conditional on the event
$\{Z_{T-1} < \rho_{T-1|F_{T-1}}(Z_T)\}$, we have that $E_{T-1} = \rho_{T-1|F_{T-1}}(Z_T)$, and $\tau^*_{T-1} = T$ is optimal for
the corresponding problem (6.14). In both cases the assertion 6.4 also holds.

Now let $m = T - 2$ and $\tau \in \mathcal{F}_{T-2}$. We have that $\rho_{T-2,T}(\cdot) = \rho_{T-2|F_{T-2}}(\rho_{T-1|F_{T-1}}(\cdot))$ and

$$
\rho_{T-2,T}(Z_\tau) = \rho_{T-2|F_{T-2}} \left( \rho_{T-1|F_{T-1}} \left( \mathbb{1}_{\{\tau = T-2\}} Z_{T-2} + \mathbb{1}_{\{\tau \geq T-1\}} Z_T \right) \right)
= \mathbb{1}_{\{\tau = T-2\}} Z_{T-2} + \mathbb{1}_{\{\tau = T-2\}} \rho_{T-2|F_{T-2}}(\mathbb{1}_{\{\tau = T-1\}} Z_{T-1} + \mathbb{1}_{\{\tau = T\}} \rho_{T-1|F_{T-1}}(Z_T))
= \mathbb{1}_{\{\tau = T-2\}} Z_{T-2} + \mathbb{1}_{\{\tau = T-2\}} \rho_{T-2|F_{T-2}}(\mathbb{1}_{\{\tau = T-1\}} Z_{T-1} + \mathbb{1}_{\{\tau = T\}} \rho_{T-1|F_{T-1}}(Z_T)),
$$

where the last equation holds since $\mathbb{1}_{\{\tau > T-2\}} \mathbb{1}_{\{\tau = T-1\}} = \mathbb{1}_{\{\tau = T-1\}}$ and $\mathbb{1}_{\{\tau > T-2\}} \mathbb{1}_{\{\tau = T\}} = \mathbb{1}_{\{\tau = T\}}$ and by (6.16). Then by (6.17) and monotonicity of $\rho_{T-2|F_{T-2}}$ we obtain

$$
\rho_{T-2,T}(Z_\tau) \leq \mathbb{1}_{\{\tau = T-2\}} Z_{T-2} + \mathbb{1}_{\{\tau > T-2\}} \rho_{T-2|F_{T-2}}(E_{T-1}) \leq \max\{Z_{T-2}, \rho_{T-2|F_{T-2}}(E_{T-1})\} = E_{T-2}.
$$

Conditional on $\{Z_{T-2} \geq \rho_{T-2|F_{T-2}}(E_{T-1})\}$, we have that $E_{T-2} = Z_{T-2}$ and $\tau^*_{T-2} = T - 2$, and

$$
\rho_{T-2,T}(Z_{\tau^*_{T-2}}) = \rho_{T-2|F_{T-2}}(\rho_{T-1|F_{T-1}}(Z_{T-2})) = Z_{T-2} = E_{T-2},
$$

and $\tau^*_{T-2}$ is optimal for the corresponding problem (6.14). Otherwise conditional on the event
$\{Z_{T-2} < \rho_{T-2|F_{T-2}}(E_{T-1})\}$, we have $E_{T-2} = \rho_{T-2|F_{T-2}}(E_{T-1})$ and $\tau^*_{T-2} \geq T - 1$. Conditioning further on $\{Z_{T-1} < \rho_{T-1|F_{T-1}}(E_T)\}$ we have that $E_{T-1} = \rho_{T-1|F_{T-1}}(Z_T)$, and $\tau^*_T = T$ is optimal for the corresponding problem (6.14). Otherwise conditional further on $\{Z_{T-1} \geq \rho_{T-1|F_{T-1}}(E_T)\}$ we have that $\tau^*_T = T - 1$ and the assertions are verified.

The assertion follows by going backwards in time for $m = T - 3, \ldots$. 

\[\Box\]
We have that $\tau^*_0 = \inf\{t: E_t = Z_t\}$ is an optimal solution of problem (6.12) and the optimal value of problem (6.12) is equal to $E_0$. That is, going forward the optimal stopping time $\tau^*_0$ stops at the first time $Z_t = E_t$. In particular it stops at $t = 0$ if $Z_0 = E_0$ (recall that $Z_0$ and $E_0$ are deterministic), i.e., if $Z_0 \geq \varrho_0|\mathcal{F}_0(E_1)$. As in the risk neutral case the time consistency (Bellman’s principle) is ensured here by the decomposable structure of the considered nested risk measure. That is, if it was not optimal to stop within the time set $\{0, \ldots, m - 1\}$, then starting the observation at time $t = m$ and being based on the information $\mathcal{F}_m$ (i.e., conditional on $\mathcal{F}_m$), the same stopping rule is still optimal for the problem (6.14).

Consider (compare with (4.12))

$$R_{s,s+1}(Z_s, Z_{s+1}) := \max \left\{ Z_s, \varrho_s|\mathcal{F}_s(Z_{s+1}) \right\}, \quad s = 0, \ldots, T - 1.$$  \(6.18\)

Then we can write $E_t$ in the following recursive form (compare with (4.4))

$$E_t = R_{t,t+1}(Z_t, R_{t+1,t+2}(Z_{t+1}, \ldots, R_{T-1,T}(Z_{T-1}, Z_T))), \quad t = 0, \ldots, T - 1.$$  

Consequently (recall that $v_0 = E_0$) problem (6.7) can be written in the form (2.1) with $\mathcal{R} := \mathcal{R}_{0,T}$, where $\mathcal{R}_{0,T}$ is given in the nested form discussed in Section 4 with $\mathcal{R}_{s,s+1}$ defined in (6.18). The dynamic equations (4.9)–(4.11) with $\mathcal{R}_{s,s+1}$ defined in (6.18) notably preserve convexity of $\varrho_s|\mathcal{F}_s$.

### 6.1.2 The min-min stopping problem

A similar analysis can be performed for the “optimistic” stopping time formulation (6.8) by defining

$$R_{s,s+1}(Z_s, Z_{s+1}) := \min \left\{ Z_s, \varrho_s|\mathcal{F}_s(Z_{s+1}) \right\}, \quad s = 0, \ldots, T - 1.$$  \(6.19\)

In contrast to (6.18) considered above, the mappings $\mathcal{R}_{s,s+1}$ defined in (6.19) do not preserve convexity of $\varrho_s|\mathcal{F}_s$. However, the corresponding dynamic equations (4.9)–(4.11) are slightly more simple with (6.19), as the essential infimum and the minimum can be interchanged here.

### 6.1.3 Supermartingales and delayed stopping

Recall that a sequence of random variables $\{X_t\}_{0 \leq t \leq T}$ is said to be supermartingale relative to the filtration $\mathcal{F}$, if $X_t \succeq \mathbb{E}_{t|\mathcal{F}_t}(X_{t+1})$, $t = 0, \ldots, T - 1$. By analogy we say that the sequence $X_t \in \mathcal{Z}_t$ is $\mathcal{P}$-supermartingale, with respect to the collection of mappings $\mathcal{P} = \{\varrho_t|\mathcal{F}_t\}_{t=0,\ldots,T-1}$, if

$$X_t \succeq \varrho_t|\mathcal{F}_t(X_{t+1}), \quad t = 0, \ldots, T - 1.$$  

It follows by the definition (6.13) that the Snell envelope sequence $\{E_t\}_{0 \leq t \leq T}$ is $\mathcal{P}$-supermartingale. It also follows from (6.13) that $E_t \succeq Z_t$, i.e., $E_t$ dominates $Z_t$. We have that $\{E_t\}_{0 \leq t \leq T}$ is the smallest $\mathcal{P}$-supermartingale which dominates the corresponding sequence $\{Z_t\}_{0 \leq t \leq T}$.
Remark 6.5 (Higher risk aversions delay stopping). Consider two collections \( \varphi_{t|\mathcal{F}_t} : Z_{t+1} \to Z_t \) and \( \varphi'_{t|\mathcal{F}_t} : Z_{t+1} \to Z_t, \ t = 0, \ldots, T - 1 \), of monotone translation equivariant mappings possessing property (6.16), with the respective Snell envelope sequences \( E_t \) and \( E'_t \) and stopping times \( \tau'_0 = \inf\{ t : E'_t = Z_t \} \) and \( \tau^*_0 = \inf\{ t : E_t = Z_t \} \), defined for \( (Z_0, \ldots, Z_T) \in Z_{0,T} \). Suppose that \( \varphi'_{t|\mathcal{F}_t}(\cdot) \succeq \varphi_{t|\mathcal{F}_t}(\cdot), \ t = 0, \ldots, T - 1 \). It follows then that \( E'_t \succeq E_t \) for \( t = 0, \ldots, T \), and hence \( \tau'_0 \geq \tau^*_0 \). That is, for larger risk mappings the optimal stopping time defined in (6.15) is more conservative.

7 Numerical illustration

To demonstrate the impact of risk aversion in the context of optimal stopping we consider the classical problem of pricing American put options. The problem statements by Bachelier considers a stock price \( S_t \) based on the random walk

\[
S_{t+1} = S_t + rS_0(t_{t+1} - t_t) + \sigma S_0(W_{t+1} - W_{t_t}),
\]

(7.1)

where \( 0 = t_0 < t_1 < \ldots \) are times, \( W_t \) is the Wiener process, \( r \in \mathbb{R} \) models the risk free interest rate and \( \sigma \geq 0 \) the constant volatility of the stock \( S \). The solution of the stochastic difference equation (7.1) is given explicitly by

\[
S_t = S_0e^{rt} + S_0e^{\sigma t},
\]

(7.2)

which shows that \( S_t \) has a normal distribution, \( S_t \sim \mathcal{N}(S_0e^{rt}, S_0^2\sigma^2 t) \).

7.1 Option pricing

The price of a classical American put option with strike \( K > 0 \) is

\[
\sup_{\tau \in \{0, \ldots, n\}} \mathbb{E}[(K - S_{\tau})_+ - r \cdot t_{\tau}],
\]

(7.3)

where the supremum is among all stopping times \( \tau \in \{0, \ldots, n\} \) with \( t_n := T \), the time of expiry of the option. The expectation \( \mathbb{E} \) in (7.3) is with respect to the risk free probability measure and \( x_+ = \max(0, x) \). In a risk averse approach the price of the option is

\[
\rho_{0,T}(Z_{\tau}),
\]

(7.4)

where \( \rho_{0,T}(\cdot) \) is the nested risk measure as introduced in (6.5).

7.2 Adapted risk functional

Consider the entropic risk measure

\[
\text{ENT}_u(Z) := u^{-1} \log \mathbb{E}[e^{uZ}], \ u > 0,
\]

(7.5)
which has the recursive property (4.2). By homogenization of ENT (cf. Shapiro et al. (2014, eq. (6.63))) we obtain its law invariant coherent risk measure companion, the entropic Value-at-Risk, for which we give the following definition below. Unlike ENT, the Entropic Value-at-Risk measure does not possess the corresponding recursive property. Nevertheless it has a surprising consistency property for the Wiener process, which exposes this risk functional for our purposes as the process $S_t$ in (7.2) has normally distributed increments.

**Definition 7.1** (Entropic Value-at-Risk). The Entropic Value-at-Risk at risk aversion rate $\beta$ ($\beta \geq 0$) of a random variable $Z$ is

$$
EV@R^\beta_{\Delta t}(Z) := \inf_{u > 0} \left\{ u^{-1} \left( \beta \cdot \Delta t + \log \mathbb{E} e^{uZ} \right) \right\}.
$$

(7.6)

**Remark 7.2.** Ahmadi-Javid and Pichler (2017) discuss the Entropic Value-at-Risk in detail. They employ the parametrization $\beta \Delta t = \log \frac{1}{1 - \alpha}$ for some risk level $\alpha \in [0, 1)$, but for the present context in multistage stopping times the parametrization in (7.6) is adapted and more convenient. It follows from this reference as well that the Entropic Value-at-Risk is a coherent risk measure. At rate $\beta = 0$, the Entropic Value-at-Risk is the expected value, i.e., the parameter setting $\beta = 0$ covers the risk neutral situation.

The corresponding conditional Entropic Value-at-Risk is

$$
EV@R^\beta_{\Delta t|\mathcal{F}_t}(Z) = \text{ess inf}_{u_t > 0} \left\{ u_t^{-1} \left( \beta \Delta t + \log \mathbb{E}_{\mathcal{F}_t} e^{u_t Z} \right) \right\},
$$

where the essential infimum is along all random variables $u_t$ which are $\mathcal{F}_t$-measurable and which satisfy $u_t > 0$ a.e. Following (4.4), the nested Entropic Value-at-Risk is defined by composition and we shall abbreviate

$$
nEV@R^\beta_{\Delta t_0, \ldots, \Delta t_{n-1}} (Z_0, Z_1, \ldots, Z_T) := EV@R^\beta_{\Delta t_0|Z_0} \left( \cdots EV@R^\beta_{\Delta t_{n-1}|Z_{t-1}} (Z_T) \right),
$$

where we set $\Delta t_i := t_{i+1} - t_i > 0$ for all $i = 0, \ldots, n - 1$.

We have the following explicit result for the nested Entropic Value-at-Risk, which highlights the predominant role of the Entropic Value-at-Risk in the context of the Wiener process.

**Proposition 7.3.** Let $t_i, i = 0, \ldots, n$, be a sequence with $t_0 := 0 < t_1 < \cdots < t_n = T$. Then the nested Entropic Value-at-Risk of the Wiener process $W_t$ is

$$
nEV@R^\beta_{\Delta t_0, \ldots, \Delta t_{n-1}} (W_{t_0}, W_{t_1}, \ldots, W_{t_n}) = EV@R^\beta_{\Delta t|Z_0} \left( \cdots EV@R^\beta_{\Delta t_{n-1}|Z_{t-1}} (W_T) \right),
$$

where $\Delta t_i = t_{i+1} - t_i$.

To verify the statement of the proposition we give the Entropic Value-at-Risk of a normally distributed random variable first.

**Lemma 7.4.** Let $Z \sim \mathcal{N}(\mu, \sigma^2)$ be normally distributed. Then the Entropic Value-at-Risk is

$$
EV@R^\beta_{\Delta t}(Z) = \mu + \sigma \sqrt{2 \beta \Delta t}.
$$

(7.7)
Proof. For $Z \sim N(\mu, \sigma^2)$, the objective in (7.6) is
\[
\frac{1}{u} \left( \beta \Delta t + \log e^{u\mu + \frac{1}{2}u^2\sigma^2} \right) = \frac{\beta \Delta t}{u} + \mu + u \frac{\sigma^2}{2},
\]
which attains its infimum at $u^* = \frac{1}{\sigma} \sqrt{2\beta \Delta t}$. At this point, the objective evaluates $\mu + \sigma \sqrt{2\beta \Delta t}$ and thus the result.

**Proof of Proposition 7.3.** The increments $W_{t+\Delta t} - W_t$ of the Wiener process have the distribution $N(0, \Delta t)$ so that
\[
\mathbb{E} \mathbb{V}_{\beta \Delta t|W_t}(W_{t+\Delta t}) = W_t + \Delta t \sqrt{2\beta \Delta t} = W_t + \Delta t \sqrt{2\beta}
\]
by (7.7). As the increments are independent the result of the nested Entropic Value-at-Risk follows by comparing the result in the latter display with (7.7), as $W_T \sim N(0, T)$ and $\sum_{i=0}^{n-1} \Delta t_i = T$.

To ensure weak convergence of the difference equation (7.1) towards the solution of the corresponding stochastic differential equation it is desirable to let $\Delta t_i \to 0$. The following results demonstrate that the binomial model is consistent with the Wiener process and the Entropic Value-at-Risk.

**Proposition 7.5** (Consistency of the binomial model). Let $Z_{\Delta t}$ be a random variable with $P(Z_{\Delta t} = \pm \sigma \sqrt{\Delta t}) = 1/2$. Then $\mathbb{E}[Z_{\Delta t}] = 0$, $\text{var}(Z_{\Delta t}) = \sigma^2 \Delta t$ and
\[
\mathbb{E} \mathbb{V}_{\beta \Delta t}(Z_{\Delta t}) = \sigma \Delta t \sqrt{2\beta} + o(\Delta t)
\]
for $\Delta t$ small.

**Proof.** It holds that $\mathbb{E}[e^{uZ_{\Delta t}}] = \frac{1}{2} e^{u\sigma \sqrt{\Delta t}} + \frac{1}{2} e^{-u\sigma \sqrt{\Delta t}} = 1 + \frac{1}{2} u^2 \sigma^2 \Delta t + o(\Delta t)$ for small $\Delta t$. Employing $\log(1 + x) = x + o(x)$ for small $x$, the objective in (7.6) is
\[
\frac{1}{u} \left( \beta \Delta t + \frac{1}{2} u^2 \sigma^2 \Delta t \right);
\]
this is (7.8) with $\sigma$ replaced by $\sigma \sqrt{\Delta t}$ and thus the result follows with (7.7).

### 7.3 Numerical Results

Figure 1 displays the prices of put options (7.4) for three different strike prices $K$ and a stock, which is traded at an initial price of $S_0 = 1$. Figure 1a displays the prices for the entropic risk measure ENT, while Figure 1b the prices for the coherent Entropic Value-at-Risk, EV@R. The risk aversion has the interpretation of an insurance premium so that the prices of risk averse American put options are more expensive compared to the risk-neutral setting.

Figure 2 displays the decision rules for varying risk aversion coefficients $\beta$. The option is not executed, as long as the price of the stock stays in the continuation region. Once the
price of the stock drops out of the continuation region into the stopping region, then it is optimal to execute the option. The optimal stopping time thus is

$$\tau = \inf \{ i \in \{0, \ldots, n\}: (t_i, S_{t_i}) \notin \text{continuation region} \}.$$  

In addition, the stopping rule can be restated as

$$S_t + p(t, S_t) \leq K,$$  \hspace{1cm} (7.9)

where $S_t$ is the current price of the stock and $p(t, S_t)$ the actual and updated price of the corresponding put option.

Figure 2a displays the continuation region and the stopping region for the entropy $\text{ENT}$, while Figure 2b employs the Entropic Value-at-Risk. The coefficient $\beta = 0$ corresponds to the risk neutral setting, while $\beta > 0$ indicates strict risk aversion. The regions do not vary significantly for the different risk measures, but they heavily depend on the risk aversion coefficient $\beta$.

It is well-known that American put option prices computed via Black–Scholes are less expensive, in general, than actual, observed prices. The Figures 1 and 2 provide an explanation. Indeed, the prices are more expensive for increasing risk levels $\beta$. Prices observed at stock markets thus can be interpreted as having an additional risk margin which corresponds to an insurance premium.

The Figures 2 show in addition that optimal stopping times execute later, if risk aversion increases. In a risk averse setting it is thus recommended to wait longer compared to a risk neutral environment. In view of (7.9) the stock has to drop to a lower level compared with the risk neutral setting. This is consistent with Remark 6.5, which provides the theoretical explanation of delayed stoppings in situations with increased risk aversions.
8 Summary

Time consistency has been discussed very actively in the recent literature on multistage stochastic optimization. A major effort has been made to introduce time consistent properties for risk measures. This paper proposes a different approach and regards time consistency as a property of the stochastic optimization problem itself. We relate these main features with dynamic equations and with optimal stopping in addition.

Dynamic equations appear in the analysis of optimal stopping as well. For this reason we introduce randomized and stopping time risk measures. With this means available we demonstrate that dynamic equations hold even for optimal stopping time problems. The conditional optimization problems relate naturally to optimal stopping time and we extend the classical theory to cover risk averse decision making.

References

A. Ahmadi-Javid and A. Pichler. An analytical study of norms and Banach spaces induced by the entropic value-at-risk. *Mathematics and Financial Economics*, 11(4):527–550, 9 2017. doi:10.1007/s11579-017-0197-9. 23

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. 7

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. 14
E. Bayraktar, I. Karatzas, and S. Yao. Optimal stopping for dynamic convex risk measures. *Illinois Journal of Mathematics*, 54(3):1025–1067, 2010. URL https://projecteuclid.org/euclid.ijm/1336049984.

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

D. Belomestny and V. Krätschmer. Optimal stopping under model uncertainty: Randomized stopping times approach. *The Annals of Applied Probability*, 26(2):1260–1295, 2016. doi:10.1214/15-AAP1116.

N. H. Bingham and G. Peskir. Optimal stopping and dynamic programming. In *Encyclopedia of Quantitative Risk Assessment and Analysis*, pages 1236–1243. John Wiley and Sons, 2008.

P. Carpentier, J.-P. Chancelier, G. Cohen, M. De Lara, and P. Girardeau. Dynamic consistency for stochastic optimal control problems. *Annals of Operations Research*, 200(1):247–263, 2012. doi:10.1007/s10479-011-1027-8.

Ö. Çavuş and A. Ruszczyński. Risk-averse control of undiscounted transient markov models. *SIAM Journal on Control and Optimization*, 52(6):3935–3966, jan 2014. doi:10.1137/13093902x.

P. Cheridito and M. Kupper. Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance*, 14(1):137–162, 2011. doi:10.1142/S0219024911006292.

P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability*, 11(3):57–106, 2006. doi:10.1214/EJP.v11-302.

E. Delage and D. Iancu. Robust multistage decision making. *Tutorials in Operations Research, INFORMS*, pages 20–46, 2015.

W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer, second edition, 2006. doi:10.1007/0-387-31071-1.

V. Guigues. Multistage stochastic programs with a random number of stages: dynamic programming equations, solution methods, and application to portfolio selection. Preprint, Cornell University Library, arXiv:1803.06034, 2018.

P. J. Hammond. Consistent plans, consequentialism, and expected utility. *Econometrica*, 57 (6):1445–1449, 1989. doi:10.2307/1913716.

M. Haviv. On constrained Markov decision processes. *Operations Research Letters*, 19(1):25–28, 1996. doi:10.1016/0167-6377(96)00003-x.
A. Hordijk. *Dynamic Programming and Markov Potential Theory*. Mathematical Centre Tracts No. 51, Amsterdam, 1974. 4

A. Jobert and L. C. G. Rogers. Valuations and dynamic convex risk measures. *Mathematical Finance*, 18(1):1–22, 2008. doi:10.1111/j.1467-9965.2007.00320.x. 2

I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance*. Stochastic Modelling and Applied Probability. Springer, 1998. doi:10.1007/b98840. 8

R. Kovacevic and G. Ch. Pflug. Time consistency and information monotonicity of multiperiod acceptability functionals. In H. Albrecher, W. J. Runggaldier, and W. Schachermayer, editors, *Advanced Financial Modelling*, number 8 in Radon Series on Computational and Applied Mathematics, pages 347–369. de Gruyter, 2009. 3

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. 3, 12, 17

G. Ch. Pflug and A. Pichler. Time-inconsistent multistage stochastic programs: Martingale bounds. *European Journal of Operational Research*, 249(1):155–163, 2015. doi:10.1016/j.ejor.2015.02.033. Initial working title: Multistage Optimization. 3

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. 3

G. Ch. Pflug and W. Römisch. *Modeling, Measuring and Managing Risk*. World Scientific, River Edge, NJ, 2007. doi:10.1142/9789812708724. 6

A. B. Philpott, V. L. de Matos, and E. Finardi. On solving multistage stochastic programs with coherent risk measures. *Operations Research*, 61(4):957–970, 2013. doi:10.1287/opre.2013.1175. 3

F. Riedel. Dynamic coherent risk measures. *Stochastic Processes and their Applications*, 112 (2):185–200, 2004. doi:10.1016/j.spa.2004.03.004. 14

U. Rieder. On optimal policies and martingales in dynamic programming. *Journal of Applied Probability*, 13:507–518, 1976. 4

R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer Verlag, 1997. doi:10.1007/978-3-642-02431-3. URL https://books.google.com/books?id=w-NdOE5fD8AC. 9

A. Ruszczynski. Risk-averse dynamic programming for Markov decision processes. *Math. Program., Ser. B*, 125:235–261, 2010. doi:10.1007/s10107-010-0393-3. 14, 16
A. Ruszczyński and D. Dentcheva. Risk forms: Representation, disintegration, and application to partially observable two-stage systems, 2018. URL https://arxiv.org/abs/1807.02300. unpublished. 3

A. Ruszczyński and A. Shapiro. Conditional risk mappings. Mathematics of Operations Research, 31(3):544–561, 2006a. doi:10.1287/moor.1060.0204. 14, 18

A. Ruszczyński and A. Shapiro. Optimization of convex risk functions. Mathematics of Operations Research, 31:433–452, 2006b. doi:10.1287/moor.1050.0186. 11

A. Shapiro. On a time consistency concept in risk averse multistage stochastic programming. Operations Research Letters, 37(37):143–147, 2009. doi:10.1016/j.orl.2009.02.005. 13

A. Shapiro. Distributionally robust stochastic programming. SIAM Journal on Optimization, 27(4):2258–2275, Jan 2017. doi:10.1137/16M1058297. URL http://www.optimization-online.org/DB_FILE/2015/12/5238.pdf. 9, 11, 17

A. Shapiro, D. Dentcheva, and A. Ruszczyński. Lectures on Stochastic Programming, Modeling and Theory. MOS-SIAM Series on Optimization. SIAM, 2nd edition, 2014. 6, 15, 16, 19, 23

A. N. Shiryaev. Optimal Stopping Rules. Springer New York, 1978. doi:10.1007/978-3-540-74011-7. 4, 19

A. Wald. Sequential Analysis. John Wiley & Sons, Inc., 1947. 4

A. Wald. Statistical decision functions. Ann. Math. Statist., 20(2):165–205, 1949. doi:10.1214/aoms/1177730030. 4

T. Wang. A class of dynamic risk measures. Technical report, University of British Columbia, 1999. 14

S. Weber. Distribution-invariant risk measures, information, and dynamic consistency. Mathematical Finance, 16(2):419–441, April 2006. doi:10.1111/j.1467-9965.2006.00277.x. 2

P. A. Weller. Consistent intertemporal decision making under uncertainty. Review of Economic Studies, 45:263–266, 1978. doi:10.2307/2297340. 3