Bonn Econon Discussion Papers

Discussion Paper 20/2009

Optimal Stopping with Dynamic Variational Preferences

by

Daniel Engelage

August 2009

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Kaiserstrasse 1
D-53113 Bonn

Financial support by the
Deutsche Forschungsgemeinschaft (DFG)
through the
Bonn Graduate School of Economics (BGSE)
is gratefully acknowledged.

Deutsche Post World Net is a sponsor of the BGSE.

Optimal Stopping with Dynamic Variational Preferences

Daniel Engelage*

August 24, 2009

Abstract

We consider optimal stopping problems in uncertain environments for an agent assessing utility by virtue of dynamic variational preferences as in [15] or, equivalently, assessing risk by dynamic convex risk measures as in [4]. The solution is achieved by generalizing the approach in [20] introducing the concept of variational supermartingales and an accompanying theory. To illustrate results, we consider prominent examples: dynamic entropic risk measures and a dynamic version of generalized average value at risk introduced in [5].

Keywords: Optimal Stopping, Uncertainty, Dynamic Variational Preferences, Dynamic Convex Risk Measures, Dynamic Penalty, Time-Consistency, Entropic Risk, Average Value at Risk
JEL-Classification: C61, C65, D81

*This article has been written while the author was visiting ORFE, Princeton University, on a scholarship granted by the German Academic Exchange Service. Fruitful discussions with Prof. Frank Riedel, Institute for Mathematical Economics, Bielefeld University, Prof. Patrick Cheridito, ORFE, Princeton University, and Monika Bier, Bielefeld University, are gratefully acknowledged. E-Mail: daniel.engelage@uni-bonn.de

1 Introduction

In our everyday life we face a broad variety of *optimal stopping problems*: We accept bids for our used car to sell or stop the process of potential marriage partners not knowing whether a more appropriate partner is still to come. On financial markets, agents try to maximize profits from American options. Hence, optimal stopping problems are not just of value for theoretical considerations but of great virtue in applications. All examples have in common that, on an abstract level, an agent has to find an optimal stopping time for some stochastic payoff process.

The classical solution to this problem, as inter alia given in [17], assumes the agent to possess a unique subjective prior ruling the payoff process and to maximize expected payoff. In an *uncertain* environment however, there might not be a unique prior distribution: On incomplete financial markets, we might be faced with multiple equivalent martingale measures not being sure which one is ruling the world. Hence, with multiple possible distributions, a solution to the problem by virtue of simple expected utility maximization with respect to some subjective prior cannot be eligible: An alternative notion of "expected reward" has to be used. In this article, we hereto choose *dynamic variational preferences*.

Equivalently, a risk manager runs the danger of high model risk when assuming a particular probabilistic model. An alternative route is given by *dynamic convex risk measures*: the robust representation explicitly mirrors multiplicity of possible distributions and hence reduces model risk. As will be motivated below, both approaches, dynamic convex risk measures and dynamic variational preferences, are equivalent in mathematical terms for our model. Only the economic interpretation differs.

In [20], the problem to optimally stop an adapted payoff process $(X_t)_{t\in\mathbb{N}}$ facing uncertainty is considered when expected reward is induced by *dynamic multiple prior preferences* introduced in [8]. By virtue of a *robust representation* theorem, expected reward at time 0 from stopping time τ is then given by minimal expectation of the form

$$\inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[X_\tau]$$

for a fixed set $\mathcal{Q}$ of *prior distributions* of the payoff process. In this sense, an uncertainty averse agent, not able to determine a unique subjective prior, considers a set $\mathcal{Q}$ of distributions to be possible and equally likely. Equivalently, the above minimized expectation is, modulo a minus sign, the robust representation of *coherent risk measures* introduced in [1] and applied to a dynamic setting in [19]: Risk is assessed as maximal expected loss with respect to all distributions that are considered likely. Hence, model risk is significantly reduced as no specific probability distribution is assumed. However, limitations of coherent risk measures are stated in [10]: Due to *homogeneity* coherent risk measures do not account for *liquidity risk*. Secondly, the robust representation shows coherent risk measures to assess risk quite conservatively. To overcome these shortcomings, the coherent approach is generalized to *convex risk measures* relaxing homogeneity and sub-additivity to a convexity condition; in a dynamic context elaborately discussed in [11] and [4]. Furthermore, the fundermental *entropic risk measure* is not coherent but convex.

Equivalently, multiple prior preferences are generalized to so called *variational preferences* in [14] and to *dynamic variational preferences* in [15]. In a more general setup, dynamic risk adjusted values or concave utilities are introduced in [4] for stochastic processes. Under the assumption of risk neutrality but uncertainty aversion, a discount factor of unity and without intermediate payoffs, expected reward π_t at time t for stopping the process $(X_t)_t$ with stopping strategy τ induced by dynamic variational preferences is given by a robust representation of the form

$$\pi_t(X_\tau) = \operatorname*{ess\,inf}_{\mathbb{Q}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t(\mathbb{Q}) \right),$$

for some *dynamic penalty* $(\alpha_t)_t$. Intuitively, expected reward $(\pi_t)_t$ is given by minimal penalized expectation – penalized in the sense that nature has to compensate the agent for choosing a distribution. Again, dynamic variational preferences and dynamic convex risk measures are equivalent as the robust representations coincide up to a minus sign: The equivalent dynamic convex risk measure is then given as $\rho_t := -\pi_t$. Hence, in terms of the above robust representation, assing risk by virtue of dynamic convex risk measures amounts to maximal penalized expected loss. It is immediate that dynamic multiple prior preferences are a special case of dynamic variational preferences when the penalty is trivial, i.e. only achieves values null and infinity. In the same token, this holds for dynamic coherent risk measures as a special case of dynamic convex ones. It is beyond the scope of this article to discuss the axioms of variational preferences or convex risk measures, respectively. We just take the robust representation as given.

The dynamic penalty $(\alpha_t)_t$, formally derived by a Fenchel-Legendre transform, might be interpreted as *ambiguity index* as in [14] and [15]. From a preference based point of view, $(\alpha_t)_t$ is a measure for uncertainty aversion: Given two agents assessing utility in terms of dynamic variational preferences, one with penalty $(\alpha_t^1)_t$, the other with $(\alpha_t^2)_t$. If $(\alpha_t^1)_t \geq (\alpha_t^2)_t$, then agent 1 is less uncertainty averse than agent 2. In other terms, risk measure 1 is less conservative than 2. Equivalently, we might think of $(\alpha_t)_t$ as an inverse likelihood of a distribution to be the ruling one: the larger the penalty, the less likely the agent assumes the respective distribution. Thus, nature has to compensate the agent more, the less likely the chosen distribution. In the prominent example of entropic risk measures or multiplier preferences, penalty is given by relative entropy. The "further away" a chosen distribution in terms of entropy from the reference distribution, the higher the penalty nature has to pay. Throughout this article we make use of robust representation in terms of the *minimal dynamic penalty* $(\alpha_t^{\min})_t$, uniquely characterizing the underlying dynamic variational preference.

For dynamic models, the first question is how conditional preferences π_t at distinct time periods are interrelated. An assumption that serves as a link between time periods is *time-consistency*, defined by virtue of $\pi_t = \pi_t(\pi_{t+1})$. Robust representation results showing equivalence of time-consistency and a condition on minimal dynamic penalty $(\alpha_t^{\min})_t$, called *no-gain condition*, are obtained in [4], [11], and [15]: Minimal penalty can be rephrased as a sum of contingent penalties and a one-step-ahead penalty. Hence, time-consistency leads to a *recursive robust representation* in terms of minimal penalized expected utility. As shown in [15], the no-gain condition on $(\alpha_t^{\min})_t$ reduces to *stability* of the set $\mathcal{Q}$ of priors for dynamic multiple prior preferences.

Results in this article constitute a generalization of results in [20] by applying optimal stopping to dynamic variational preferences. By virtue of the recursion formula for robust representation, we obtain a worst-case distribution among those with finite penalty. However we do not obtain the elegant intuition in [20] that the agent behaves as expected utility maximizer with respect to the worst-case distribution as the penalty is not trivial and, hence, does not necessarily vanish for the worst-case distribution. As in [17] and [20], we make use of a *Snell envelope approach* to solve the optimal stopping problem under dynamic variational preferences by showing equality of the *value function* and an appropriately generalized Snell envelope, called *variational Snell envelope*, for a finite horizon. In the infinite horizon case, we show the *Bellman principle* to hold for the value function. These results allow us to obtain an optimal stopping strategy recursively. We observe that the smallest optimal stopping time obeys well-known characteristics: Stop when the payoff process equals the problem's value. A further result is a minimax theorem. We introduce the notion of *variational (super-, sub-) martingales* and an accompanying theory: We obtain a Doob decomposition and an optional sampling theorem.

To illustrate our results, we consider two prominent examples: dynamic *entropic risk measures* (or dynamic multiplier preferences) and a dynamic convex generalization of *average value at risk* (AVaR) introduced in [5]. Examples are stated in terms of dynamic convex risk measures instead of dynamic variational preferences. Due to mathematical equivalence of both approaches, the reason is merely owed to topicality of appropriate risk measures for financial markets in face of the current financial crisis questioning the core of financial practice. In the first example on dynamic entropic risk measures, we obtain quite intuitive results on the worst-case measure for a specific kind of payoff processes. Thereafter, we consider generalized average value at risk (gAVaR) as introduced in [4]. As the natural dynamic extension of these risk measures is not time-consistent, we achieve a time-consistent version by virtue of a recursive construction in terms of the minimal penalty. As we see in the examples, when considering non-trivial penalty functions applications become more complex: in particular, independence, inevitably used in simple examples in [20], does not hold any longer. Nevertheless, the second example constitutes a tangible alternative to widely used VaR taking into account liquidity risk, satisfying time-consistency, and avoiding the problem of risk accumulation caused by VaR.

In [23] an approach to optimal behavior on financial markets is applied without time-consistency. Agents maximize minimal penalized intertemporal utility as given above. Making use of convex conjugates, a minimax theorem similar to ours is achieved but without constructive recursion for worst-case measures. However, we are convinced that time-consistency is not only a crucial property from a theoretical perspective but also intuitively justifiable.

The article is structured as follows: The next section defines the model, gathers the relevant assumptions and then states the optimal stopping problem. This directly leads to the definition of variational supermartingales and an accompanying theory in Section 3. Section 4 contains the main results. Section 5 discusses examples. Thereafter, we conclude. Elaborate proofs are given in the Appendix.

2 The Model

Let $T \in \mathbb{N} \cup \{\infty\}$ and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \leq T}, \mathbb{P}_0)$ be an arbitrary but fixed underlying filtered probability space with $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F} = \sigma\left(\bigcup_{t \leq T} \mathcal{F}_t\right)$. Intuitively, the filtration $(\mathcal{F}_t)_{t \leq T}$ models the information process for the agent.

Let $(X_t)_{t \leq T}$ be an adapted essentially bounded payoff process that the agent aims to optimally stop in an uncertain environment, i.e. to find a *stopping time* τ in order to maximize expected reward. The specific form of expected reward used in this article emerging from dynamic variational preferences will be encountered below.

Equalities are meant to hold $\mathbb{P}_0$-a.s. Let $\mathcal{M}^e(\mathbb{P}_0)$ denote the set of all probability distributions on $(\Omega, \mathcal{F})$ that are *locally equivalent* to $\mathbb{P}_0$, i.e.

$$\mathcal{M}^e(\mathbb{P}_0) := \{\mathbb{P}\,|\forall t \leq T, \forall F \in \mathcal{F}_t, \mathbb{P}(F) = 0 \Leftrightarrow \mathbb{P}_0(F) = 0\}\ .$$

As we see in [11], the assumption to only consider locally equivalent distributions is justified as the robust representation of dynamic variational preferences is only based on these. Intuitively, the reference distribution $\mathbb{P}_0$ fixes the null sets, i.e. sure and impossible events. Recall that a stopping time τ is an integer valued random variable such that $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \leq T$. For $\omega \in \Omega$, we set $X_\tau(\omega) := X_{\tau(\omega)}(\omega)$. Let $L^\infty := L^\infty(\Omega, \mathcal{F}, \mathbb{P}_0)$ be the space of all essentially bounded $\mathcal{F}$-measurable random variables. Analog, for $t \leq T$, let $L^\infty_t := L^\infty(\Omega, \mathcal{F}_t, \mathbb{P}_0)$ be the space of all essentially bounded $\mathcal{F}_t$-measurable random variables.

2.1 Robust Representation of Time-Consistent Dynamic Variational Preferences

Given a stopping time τ, we first have to answer how agents assess utility in uncertain environments? More elaborately, given the agent is not able to entirely assess the ruling distribution of the payoff process and is uncerteinty averse but risk neutral, how does expected reward look like? In *expected utility theory* the agent is assumed to possess a unique subjective probability distribution, say $\mathbb{Q}$, and assesses expected reward by $\mathbb{E}^{\mathbb{Q}}[X_\tau]$. In [20] the agent is not sure about the distribution of $(X_t)_{t\leq T}$ but assumes the relevant distributions in some convex set $\mathcal{Q} \subset \mathcal{M}^e(\mathbb{P}_0)$ all being equally likely. Then, *multiple prior* expected reward is given by $\inf_{\mathbb{Q}\in\mathcal{Q}} \mathbb{E}^{\mathbb{Q}}[X_\tau]$.

In this article, we go a step further by assuming that an agent determines expected reward in terms of *dynamic variational preferences* as introduced in [15] or, equivalently, by a *dynamic convex risk measure* as in [11] assuming risk neutrality and no discounting. As shown in [15] and [4], the agent then assesses *conditional variational expected reward* $\pi_t(X_\tau)$ at time t from stopping at τ by virtue of a robust representation in terms of minimal penalized expected utility. This is obtained from the axioms of dynamic variational preferences. As we do not consider the axioms, we pose assumptions that imply this robust representation.

Notation 2.1. *Throughout this article, we denote by $(\rho_t)_{t\leq T}$ a dynamic convex risk measure as introduced in [11] or, equivalently, by $(\pi_t)_{t\leq T}$ the robust representation of a dynamic variational preference as in [15]. Moreover, we identify the preference with its robust representation.*

We now state rigorous definitions obtained from [11] and [15]:

Definition 2.2 (Dynamic Penalty & Time-Consistency)**.** (a) *We call a family* $(\alpha_t)_{t\leq T}$ *a* dynamic penalty *if each* α_t *satisfies:*

- α_t *is a mapping* $\alpha_t : \mathcal{M}^e(\mathbb{P}_0) \to L^1_+(\mathcal{F}_t)$*: For each* $\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)$*,* $\alpha_t(\mathbb{Q})$ *is an* $\mathcal{F}_t$*-measurable random variable with values in* $\mathbb{R}_+ \cup \{\infty\}$*.*[1]
- *For all* $t \geq 0$*,* α_t *is* grounded*, i.e.* ess $\inf_{\mathbb{Q}\in\mathcal{M}^e(\mathbb{P}_0)} \alpha_t(\mathbb{Q}) = 0$.
- α_t *is closed and convex.*[2]

(b) *Given a dynamic convex risk measure* $(\rho_t)_{t\leq T}$*. At* t*, define the* acceptance set *by* $\mathcal{A}_t := \{X \in L^\infty | \rho_t(X) \leq 0\}$*. Then, we define the* minimal penalty $(\alpha_t^{\min})_{t\leq T}$ *by*

$$\alpha_t^{\min}(\mathbb{Q}) := \operatorname*{ess\,sup}_{X\in\mathcal{A}_t} \mathbb{E}^{\mathbb{Q}}[-X|\mathcal{F}_t].$$

for all $\mathbb{Q} \in \mathcal{M}$*.*[3]

(c) *Let* $(p_t)_{t\leq T}$ *(resp.* $(q_t)_{t\leq T}$*) denote the density process of* $\mathbb{P}$ *(resp.* $\mathbb{Q}$*) in* $\mathcal{M}^e(\mathbb{P}_0)$ *with respect to* $\mathbb{P}_0$*, i.e.* $p_t := \frac{d\mathbb{P}}{d\mathbb{P}_0}\Big|_{\mathcal{F}_t}$*, where* $\frac{d\mathbb{P}}{d\mathbb{P}_0}$ *denotes the Radon-Nikodym derivative. For a stopping time* θ *define the "pasted distribution"* $\mathbb{P} \otimes_\theta \mathbb{Q}$ *by*

$$\frac{d(\mathbb{P} \otimes_\theta \mathbb{Q})}{d\mathbb{P}_0}\Bigg|_{\mathcal{F}_t} := \begin{cases} p_t & \textit{if } t \leq \theta, \\ \frac{p_\theta q_t}{q_\theta} & \textit{else.} \end{cases}$$

(d) $(\alpha_t)_{t\leq T}$ *satisfies the* no-gain *condition if for all* $t \geq 0$ *and* $\mathbb{Q}$ *we have*

$$\alpha_t(\mathbb{Q}) = \mathbb{E}^{\mathbb{Q}}\left[\alpha_{t+1}(\mathbb{Q})|\mathcal{F}_t\right] + \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t(\mathbb{Q} \otimes_{t+1} \mathbb{P}). \tag{1}$$

(e) $(\pi_t)_{t\leq T}$ *is called* time-consistent *if it satisfies* $\pi_t = \pi_t(\pi_{t+1})$ *for all* $t < T$*. Equivalently,* $\rho_t = \rho_t(-\rho_{t+1})$*.*[4]

---

[1] More elaborately, for all $\omega \in \Omega$, $\alpha_t(\cdot)(\omega)$ is a function on the $\mathcal{F}_t$-bayesian updated distributions in $\mathcal{M}^e(\mathbb{P}_0)$, i.e. the *effective domain* satisfies $\text{effdom}(\alpha_t(\cdot)(\omega)) \subset \{\mathbb{Q}(\cdot|F_t) : \mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0), \omega \in F_t \in \mathcal{F}_t\}$. Hence, when writing $\alpha_t(\mathbb{Q})$ we actually have in mind $\alpha_t(\mathbb{Q}(\cdot|\mathcal{F}_t))$.

[2] This assumption is well defined by [10], Remark 4.16.

[3] $(\alpha_t^{\min})_{t\leq T}$ is a penalty function in terms of (a).

[4] In general, time-consistency is defined as: $\rho_t = \rho_t(-\rho_{t+s})$, $t, s \leq T$, $t + s \leq T$. In this sense, our definition of time-consistency is a special case, called "one-step time-consistency" in [4]. However, for the proofs in this article, our definition is sufficient and, of course, always satisfied in the general case of time-consistency. On the other hand, one-step time-consistency implies general time-consistency under our continuity assumptions by Proposition 4.5 in [4]. Hence, our definition of time-consistency in terms of "one-step time-consistency" is equivalent to the general notion of time-consistency.

Notation 2.3. *Define the set $\mathcal{M}$ of distributions in $\mathcal{M}^e(\mathbb{P}_0)$ by*[5]

$$\mathcal{M} := \{\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0) \quad | \quad \alpha_0(\mathbb{Q}) < \infty\}.$$

Given the distribution $\mathbb{Q} \in \mathcal{M}$, $\mathbb{Q}|_{\mathcal{F}_t}$ denotes the restriction of $\mathbb{Q}$ to $\mathcal{F}_t$ given $\mathcal{F}_{t-1}$. As usual $\mathbb{Q}(\cdot|\mathcal{F}_t)$ denotes the conditional probability distribution of the process given history up to time t.

Taking into account that α_t only depends on bayesian updates, we simplify notation when appropriate and write $\alpha_t(\mathbb{Q} \otimes_{t+1} \mathbb{P}) = \alpha_t(q_{t+1}p_{t+2}\ldots)$.

Assumption 2.4 (Main Assumption)**.** *Throughout this article we assume the agent to assess risk in terms of a relevant time-consistent dynamic convex risk measure $(\rho_t)_{t\leq T}$ on the set of essentially bounded $\mathcal{F}$-measurable random variables as in [11] or, equivalently, assess utility in terms of time-consistent dynamic variational preferences $(\pi_t)_{t\leq T}$ for end-period payoffs as in [15] with no-discounting and risk neutrality. Furthermore, we assume continuity from below for $(\rho_t)_{t\leq T}$, i.e. for all $(X_n)_n \subset L^\infty$ such that $X_n \nearrow X$ for some $X \in L^\infty$, we have $\rho_t(X_n) \searrow \rho_t(X)$. Equivalently, we assume continuity from below of $(\pi_t)_{t\leq T}$, i.e. $\pi_t(X_n) \nearrow \pi_t(X)$ for the above sequence.*

Remark 2.5. *Given the payoff process $(X_t)_{t\leq T}$ and stopping time $\tau \leq T$, [4] and [11] show that, under Assumption 2.4, $(\rho_t)_{t\leq T}$ and $(\pi_t)_{t\leq T}$ have a robust representation of the form*

$$\pi_t(X_\tau) = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\} = -\rho_t(X_\tau),$$

with $(\alpha_t^{\min})_{t\leq T}$, the minimal penalty, assumed to satisfy the no-gain condition.

Consider a distribution $\mathbb{Q} \in \mathcal{M}^e(\mathbb{P}_0)$ such that, for all t, $\alpha_t(\mathbb{Q}) = 0$ and ∞ else: We achieve expected utility with subjective prior $\mathbb{Q}$. As shown in [15], multiple prior expected reward with $\mathcal{Q} \subset \mathcal{M}^e(\mathbb{P}_0)$ is a special case of variational expected reward with $\alpha_t = 0$ on $\mathcal{Q}$ and ∞ else. Hence, the present article is a generalization of the approach in [20].

By the Fenchel-Legendre Transform, minimal penalty can be written as

$$\alpha_t^{\min}(\mathbb{Q}) = \operatorname*{ess\,sup}_{X\in L^\infty}(\mathbb{E}^{\mathbb{Q}}[-X|\mathcal{F}_t] - \rho_t(X))$$

[5]It can be seen in [11], Lemma 3.5, that this domain of a penalty is well defined in case of relevant time-consistent dynamic convex risk measures as relevance allows to only consider the set of locally equivalent distributions in the robust representation and time-consistency in conjunction with relevance implies $\alpha_t(\mathbb{Q}) < \infty$ for all t. We call a dynamic convex risk measure $(\rho_t)_{t\leq T}$ *relevant*, if $\mathbb{P}_0[\rho_t(-\epsilon\mathbb{I}_A) > 0] > 0$ for all t, $\epsilon > 0$ and $A \in \mathcal{F}$ such that $\mathbb{P}_0[A] > 0$.

for all $\mathbb{Q} \in \mathcal{M}$. The term "minimal" is justified as the robust representation allows for multiple penalties $(\alpha_t)_{t\leq T}$, but the minimal one satisfies $\alpha_t^{\min}(\mathbb{Q}) \leq \alpha_t(\mathbb{Q})$ for all $\mathbb{Q} \in \mathcal{M}$. The minimal penalty uniquely characterizes the agent's preferences or, equivalently, risk attitude. *Throughout, we assume a robust representation in terms of minimal penalty* for technical reasons.

The intuition of equation (1) is the following: Nature has to pay a penalty for choosing a specific distribution at time t: α_t. On the left hand side of equation (1), nature uses the time-consistent way by just choosing a probability $\mathbb{Q}$, pay the appropriate amount and do nothing in the next period but go with the conditional distribution $\mathbb{Q}(\cdot|\mathcal{F}_t)$. However, the right hand side describes the possibly time-inconsistent way of choosing a probability: It chooses today a distribution $\mathbb{P}$ that inuces the same distribution today as $\mathbb{Q}$ but may differ from tomorrow on and pays the amount $\alpha_t(\mathbb{Q} \otimes_{t+1} \mathbb{P})$. In the second step, i.e. after realization of $\mathcal{F}_{t+1}$, nature may deviate and, conditionally on $\mathcal{F}_t$, choose a distribution $\mathbb{Q}$. If this time-inconsistent way of choosing a distribution is not less costly, $(\alpha_t)_t$ satisfies equation (1). In particular, the cost of choosing $\mathbb{Q}$ at time t can be decomposed into the sum of expected cost of choosing $\mathbb{Q}$'s conditionals at time $t+1$ and the cost of inducing $\mathbb{Q}|_{\mathcal{F}_{t+1}}$ as a so-called *one-period-ahead* marginal distribution of the payoff process at time t. The no-gain condition on $(\alpha_t)_t$ is the generalization of the time-consistency condition in [20]: As shown in [15], if (α_t) is trivial, the no-gain condition is equivalent to stability of the set of priors. This also holds true in the not necessarily finite case as shown in e.g. in [4]. In course of this section, we explicitly show time-consistency results.

Remark 2.6. (a) *As motivated in Remark 2.5, the no-gain condition on $(\alpha_t^{\min})_t$ is equivalent to time-consistency. We will make this explicit later.*

(b) *As stated in [12], Remark* 1.1, *continuity from below of π_t or ρ_t implies continuity from above of either one. Continuity from above is equivalent to the existence of a robust representation of π_t (or ρ_t) in terms of minimal penalized expected payoff. Continuity from below induces the worst case distribution to be achieved. We will make this explicit in Proposition 3.2. We hence could change the* inf *into a* min *but stick to the notion above as this seems common in the literature.*

Remark 2.7. *The following assumption is equivalent to π_t (or equivalently ρ_t) being continuous from below:*

$$\left\{ \frac{d\mathbb{P}}{d\mathbb{P}_0}\bigg|_{\mathcal{F}_t} \;\middle|\; \mathbb{P} \in \mathcal{M}, \alpha_t^{\min}(\mathbb{P}) < c \right\},$$

for each $c \in \mathbb{R}$, $t \in \mathbb{N}$, being relatively weakly compact in $L^1(\Omega, \mathcal{F}, \mathbb{P}_0)$.

Proof. Theorem 1.2 in [12] states the assertion in an unconditional setting. Due to the properties of conditional expectations, the assertion also holds in our dynamic set-up. □

Remark 2.8 (Conditional Cash Invariance). *One of the axioms of dynamic variational preferences (and dynamic convex risk measures) is* conditional cash invariance. *In conjunction with a normalization assumption, this property becomes: for all $t \leq T$ and $\mathcal{F}_t$-measurable X_t, we have $\pi_t(X_t) = X_t$. As we do not consider the axiomatic approach, we immediately derive this property from the robust representation as α_t is assumed to be grounded:*

$$\pi_t(X_t) = X_t + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}) = X_t.$$

The next result explicitly states the connection of time-consistency and the no-gain condition. The proof is a special case of the proof of Theorem 4.22 in [4]. It is stated as it generates fruitful insights.

Proposition 2.9. *Equation (1) implies time-consistency of $(\pi_t)_{t\leq T}$. More precisely, we have for all $(X_t)_{t\leq T}$ and $\tau \leq T$*

$$\pi_t(X_\tau) = X_\tau \mathbb{I}_{\{\tau\leq t\}} + \pi_t(\pi_{t+1}(X_\tau))\mathbb{I}_{\{\tau\geq t+1\}} = \pi_t(\pi_{t+1}(X_\tau)).$$

Proof. See Appendix A.1 □

As in [15], we have the following result on the recursive structure of variational expected reward π_t at time t. However, we achieve this result for more general probability spaces but under the assumption of end-period payoffs, risk neutrality and a discount factor of unity.

Corollary 2.10. *Given equation (1), it holds*

$$\pi_t(X_\tau) = X_\tau \mathbb{I}_{\{\tau\leq t\}} + \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\int \pi_{t+1}(X_\tau) d\mu + \gamma_t(\mu) \right) \mathbb{I}_{\{\tau\geq t+1\}},$$

where

$$\gamma_t(\mu) := \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{Q}) \quad \forall \mu \in \mathcal{M}|_{\mathcal{F}_{t+1}},$$

and $\mathcal{M}|_{\mathcal{F}_{t+1}}$ denotes the set of all distributions in $\mathcal{M}$ restricted on $\mathcal{F}_{t+1}$ conditional on $\mathcal{F}_t$. To have this expression well-defined, we set $\operatorname{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mu\otimes_{t+1} \mathbb{P}) := \operatorname{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q} \otimes_{t+1} \mathbb{P})$ *with $\mathbb{Q} \in \mathcal{M}$ such that $\mathbb{Q}|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t) = \mu$.*

Proof. See Appendix A.1 □

γ_t might be viewed as nature's penalty when choosing the one-period-ahead marginal μ. Hence, it is called *one-period-ahead penalty* in analogy to [15]. In terms of γ_t, equation (1) becomes

$$\alpha_t^{\min}(\mathbb{Q}) = \mathbb{E}^{\mathbb{Q}}[\alpha_{t+1}^{\min}(\mathbb{Q})|\mathcal{F}_t] + \gamma_t(\mathbb{Q}|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t)). \tag{2}$$

Remark 2.11 (Bellman Principle for Nature). *Given $\tau \leq T$, Corollary 2.10 can be rephrased as*

$$\pi_t(X_\tau) \quad = \quad \underset{\mathbb{Q}|_{\mathbb{F}_{t+1}} \in \mathcal{M}|_{\mathcal{F}_{t+1}}}{\text{ess inf}} \left(\mathbb{E}^{\mathbb{Q}|_{\mathbb{F}_{t+1}}}[\pi_{t+1}(X_\tau)|\mathcal{F}_t] + \gamma_t(\mathbb{Q}|_{\mathbb{F}_{t+1}}) \right).$$

Intuitively, this constitutes a Bellman principle for nature's choice of a worst-case distribution:[6] *Given the optimal (worst-case) distribution from time $t+1$ on, represented by its value π_{t+1}, nature chooses a minimizing one-period ahead conditional distribution $\mathbb{Q}|_{\mathbb{F}_{t+1}}$. Note, that the above expression is basically the same as the robust representation but in terms of a one-step-ahead problem. This insight is particularly adjuvant when constructing a worst-case distribution in Proposition 3.2 in terms of pasted one-period ahead conditional distributions.*

2.2 The Agent's Problem

Given $(X_t)_{t\leq T}$, $T \in \mathbb{N}\cup\{\infty\}$, the agent has to maximize variational expected reward $(\pi_t(X_\tau))_{t\leq T}$, i.e. the agent solves the following problem by finding an appropriate stopping time τ with respect to $(\mathcal{F}_t)_{t\leq T}$:

$$\sup_{0\leq\tau\leq T} \inf_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_0] + \alpha_0^{\min}(\mathbb{Q}) \right)$$

among all stopping times that are universally finite, i.e.

$$\inf_{\mathbb{Q}\in\mathcal{M}} \mathbb{Q}[\tau < \infty] = 1.$$

Definition 2.12 (Value Function, Snell Envelope). (a) *For the problem at hand, the* value (function) $(V_t)_{t\leq T}$ *at time $t \leq T$ is given by*

$$V_t := \underset{T\geq\tau\geq t}{\text{ess sup}}\ \underset{\mathbb{Q}\in\mathcal{M}}{\text{ess inf}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \right). \tag{3}$$

[6] This should not be mixed up with the Bellman principle in the next chapter's theorems on optimal stopping: there, we achieve Bellman equations for the optimal stopping decision of the agent, not for the worst-case distribution decision of nature.

(b) *For finite T, define the* variational Snell envelope $(U_t)_{t\leq T}$ *of* $(X_t)_{t\leq T}$ *with respect to dynamic minimal penalty* $(\alpha_t^{\min})_{t\leq T}$ *recursively by* $U_T := X_T$ *and*

$$U_t := \max\left\{X_t, \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\} \quad \textit{for } t < T. \tag{4}$$

(c) *Define the stopping time*

$$\tau^* := \inf\{t \geq 0 | U_t = X_t\}. \tag{5}$$

By time-consistency the variational Snell envelope can also be written as:

$$U_t = \max\left\{X_t; \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}}\left(\int U_{t+1}d\mu + \gamma_t(\mu)\right)\right\}$$

Subsequently, we show that the value function and the variational Snell envelope coincide when T is finite. In the infinite time-horizon case, we show the Bellman principle to hold for the value function allowing for recursive solutions. Furthermore, it follows that τ^* is an optimal stopping time, i.e. a solution to the initial problem. Note, that the variational Snell envelope coincides with the multiple prior Snell envelope in case of multiple prior preferences as introduced in [20]. It coincides with the "good old" Snell envelope as e.g. set out in [17] in case of a unique subjective prior.

3 Variational Supermartingales

From the approach to optimal stopping in terms of Snell envelopes or more generally with multiple prior Snell envelopes as in [20], we know that the value function satisfies some kind of martingale property until optimal stopping and some kind of supermartingale property thereafter. We now come up with an appropriate notion of martingale for dynamic variational preferences generalizing the notion of multiple prior (sub-, super-) martingales in [20]:

Definition 3.1. *Given dynamic minimal penalty $(\alpha_t^{\min})_{t\in\mathbb{N}}$ satisfying equation (1). Let $(M_t)_{t\in\mathbb{N}}$ be an $(\mathcal{F}_t)_{t\in\mathbb{N}}$-adapted process with $\mathbb{E}^{\mathbb{Q}}[|M_t|] < \infty$ for all $t \in \mathbb{N}$ and all $\mathbb{Q} \in \mathcal{M}$. $(M_t)_{t\in\mathbb{N}}$ is called a* variational (sub-, super-) martingale *with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$ if the following relation holds for $t \in \mathbb{N}$:*

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = (\geq, \leq) M_t.$$

[20], Lemma 6, shows an elegant way to characterize the concepts of multiple prior (sub-, super-) martingales in terms of (sub-, super-) martingales with respect to a worst-case distribution $\mathbb{P}^* \in \mathcal{Q}$. However, this result is owed to the simple structure of $(\alpha_t^{\min})_{t\in\mathbb{N}}$ in the multiple priors case. Under variational preferences we can state a similar result for variational supermartingales as being a supermartingale "modulo penalty" with respect to some worst-case distribution $\mathbb{Q}^* \in \mathcal{M}$. This non-vanishing penalty is the reason why the intuition of an agent behaving as expected utility maximizer under the worst case distribution does *not* carry over from [20]. The worst-case distribution is achieved recursively: At each time t, the *worst-case conditional one-step-ahead distribution* is chosen. In [20], time-consistency is needed to ensure the recursively pasted distribution to be again in the set priors $\mathcal{Q}$. By definition of $\mathcal{M}$ and equation (1), we obviously have that pasted distributions are again in $\mathcal{M}$: $\alpha_{t+1}^{\min}(\mathbb{Q}) < \infty$ implies $\alpha_t^{\min}(\mathbb{Q}) < \infty$. The most important part in our construction is that, given equation (1), pasting of worst-case one-step-ahead distributions is consistent with being of worst-case type: *Having achieved a worst-case distribution from $t+1$ onwards, we paste this with the one-step-ahead worst-case conditional distribution from t to $t+1$ and achieve the worst-case distribution from time t onwards.*

Proposition 3.2. *Let $(M_t)_{t\in\mathbb{N}}$ be an adapted process and $(\alpha_t^{\min})_{t\in\mathbb{N}}$ satisfy equation (1).*
(a) *If $(M_t)_{t\in\mathbb{N}}$ is a $\mathbb{Q}$-submartingale for all $\mathbb{Q} \in \mathcal{M}$, then $(M_t)_{t\in\mathbb{N}}$ is a variational submartingale with respect to $(\alpha_t^{\min})_t$.*
(b) *$(M_t)_{t\in\mathbb{N}}$ is a variational supermartingale with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$ if and only if there exists a $\mathbb{Q}^* \in \mathcal{M}$ such that $(M_t)_{t\in\mathbb{N}}$ is a $\mathbb{Q}^*$-supermartingale "modulo penalty", i.e.*

$$\mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*) \leq M_t.$$

In particular, $(M_t)_{t\in\mathbb{N}}$ is a $\mathbb{Q}^*$-supermartingale, i.e. $\mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] \leq M_t$.

Proof. See Appendix A.2 □

Remark 3.3. *By lemmata in Appendix A.2, the foregoing assertion can be generalized to: $\exists\ \mathbb{Q}^* \in \mathcal{M}$ such that $\forall t, s$ we have*

$$\mathbb{E}^{\mathbb{Q}^*}[M_s|\mathcal{F}_t] + \alpha_t(\mathbb{Q}^*)\mathbb{I}_{\{s>t\}} \leq M_t.$$

In the same token as in [20], we generalize standard results for supermartingales to our notion of variational supermartingales.

Proposition 3.4 (Doob Decomposition). *Let $(S_t)_{t\in\mathbb{N}}$ be a variational supermartingale with respect to dynamic penalty $(\alpha_t^{\min})_{t\in\mathbb{N}}$ satisfying equation (1).*

Then there exists a variational martingale $(M_t)_{t\in\mathbb{N}}$ *with respect to* $(\alpha_t^{\min})_{t\in\mathbb{N}}$ *and a predictable non-decreasing process* $(A_t)_{t\in\mathbb{N}}$, $A_0 = 0$, *such that* $S_t = M_t - A_t$ *for all* t *and this decomposition is unique.*

Proof. See Appendix A.3 □

In [20], the proof of optional sampling is immediate as the minimal penalty vanishes; here we mimic the proof of the original optional sampling theorem.

Proposition 3.5 (Optional Sampling)**.** *Let* $(S_t)_{t\in\mathbb{N}}$ *be a variational supermartingale with respect to dynamic minimal penalty* $(\alpha_t^{\min})_{t\in\mathbb{N}}$ *satisfying equation (1) and* $\sigma \le \tau$ *be universally finite stopping times. Then*

$$S_\sigma \ge \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right).$$

Proof. See Appendix A.3 □

Corollary 3.6 (from Propsition 3.5)**.** *Let* $(S_t)_{t\in\mathbb{N}}$ *be a variational supermartingale with respect to dynamic minimal penalty* $(\alpha_t^{\min})_{t\in\mathbb{N}}$ *satisfying equation (1). Then we have for every universally finite stopping time* τ

$$S_{\tau\wedge t} \ge \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{\tau\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right).$$

Proof. See Appendix A.3 □

4 Main Results

We are now enabled to state and prove the main results of this article. These directly generalize the results in [20] to dynamic variational preferences. Throughout, we assume $(\alpha_t^{\min})_{t\le T}$ to satisfy equation (1).

4.1 Finite Horizon

Let $T < \infty$. The following result extends the fundamental Propositions VI-1-2 and VI-1-3 in [17] to dynamic variational preferences:

Theorem 4.1. (a) *The variational Snell envelope* $(U_t)_{t\le T}$ *defined in equation (4) is the smallest variational supermartingale with respect to* $(\alpha_t^{\min})_{t\le T}$ *that dominates* $(X_t)_{t\le T}$.

(b) *We have* $U_t = V_t$ *for all* $t \le T$, *i.e. the variational Snell envelope, equation (4), equals the problem's value function, equation (3).*

(c) τ^* *from equation (5) is the smallest optimal stopping time, i.e. solves the optimal stopping problem stated in Remark* **??**.

Proof. See Appendix A.4 □

We now state a minimax-theorem allowing to interchange the "inf" and "sup" in the formulation of the problem: It does not matter if nature chooses a worst case distribution first and then the agent maximizes or vice versa.

Theorem 4.2 (Minimax-Theorem)**.** *For every $t \leq T$, we have*

$$\operatorname*{ess\,sup}_{T\geq\tau\geq t} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\operatorname*{ess\,sup}_{T\geq\tau\geq t} \mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right).$$

Proof. See Appendix A.4 □

Remark 4.3. *Posed in another way, we have*

$$U_t = \operatorname*{ess\,sup}_{T\geq\tau\geq t} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(U_t^{\mathbb{Q}} + \alpha_t^{\min}(\mathbb{Q})\right),$$

where $U_t^{\mathbb{Q}}$ denotes the Snell envelope of the expected-utility optimal stopping problem with subjective prior $\mathbb{Q}$. Hence, we do not *have the elegant result as in [20] that the variational Snell envelope $(U_t)_{t\leq T}$ is the lower envelope of the individual Snell envelopes $(U_t^{\mathbb{Q}})_{t\leq T}$ as the penalty is not necessarily zero.*

Remark 4.4. *Set $\mathbb{Q}^{\mathcal{M}}$ the worst-case distribution for dynamic variational preferences and $\mathbb{Q}^{\mathcal{Q}}$ the worst-case distribution for multiple priors in $\mathcal{Q}$ assuming $\mathcal{M} = \mathcal{Q}$, i.e. the sets of distributions with finite penalty coincide. Let $(V_t)_{t\leq T}$ denote the value function for dynamic variational preferences and $(V_t^{\mathbb{Q}})_{t\leq T}$ the value of the optimal stopping problem with subjective prior $\mathbb{Q}$ for an expected utility maximizer. We then have*

$$V_t = \operatorname*{ess\,sup}_{T\geq\tau\geq t} \left(\mathbb{E}^{\mathbb{Q}^{\mathcal{M}}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^{\mathcal{M}})\right) \geq V_t^{\mathbb{Q}^{\mathcal{M}}}.$$

In particular, smallest optimal stopping times differ. Furthermore, we see

$$V_t \geq V_t^{\mathbb{Q}^{\mathcal{Q}}}.$$

In other words, sophistication of $(\alpha_t^{\min})_{t\leq T}$ increases expected reward. Intuitively: The agent has more information on the likelihood of distributions available under variational preferences than under multiple priors and hence values the problem more. Stated in other terms more important to applications in risk management: Convex risk measures assess risk in a more liberal fashion than coherent ones given the sets of considered distributions coincide.

4.2 Infinite Horizon

Let $T = \infty$. We now show the value function to satisfy the Bellman principle:

Theorem 4.5. (a) *The value process* $(V_t)_{t\in\mathbb{N}}$ *as defined in equation (3) is the smallest variational supermartingale with respect to* $(\alpha_t^{\min})_{t\in\mathbb{N}}$ *that dominates the payoff process* $(X_t)_{t\in\mathbb{N}}$.
(b) *The value process* $(V_t)_{t\in\mathbb{N}}$ *satisfies the Bellman principle, i.e.*

$$V_t = \max\left\{X_t, \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\} \quad \text{for all } t \geq 0.$$

(c) $\tau^* := \inf\{t \geq 0 | V_t = X_t\}$ *is the smallest optimal stopping time.*
(d) *Let* $(U_t^T)_{t\leq T}$ *denote the variational Snell envelope with respect to* $(\alpha_t^{\min})_{t\leq T}$ *for the optimal stopping problem of* $(X_t)_{t\leq T}$ *truncated to finite horizon* $T < \infty$. *Let* $(V_t)_{t\in\mathbb{N}}$ *denote the value of the infinite problem. Then we have* $\lim_{T\to\infty} U_t^T = V_t$ *for all* $t \geq 0$.

Proof. See Appendix A.5. □

The last part of the foregoing theorem is particularly valuable for constructive solutions of infinite models in terms of limiting solutions of truncated ones.

5 Examples

In this section, we consider optimal stopping problems for prominent examples of dynamic variational preferences. First, we consider stopping with dynamic multiplier preferences or, equivalently, dynamic entropic risk measures. Secondly, we apply our theory to a generalized version of average value at risk (gAVaR) particularly paying attention to time-consistency issues.

In [20], simplicity of examples is due to triviality of the dynamic minimal penalty for multiple prior preferences. In particular, for monotone problems, this fact allows to obtain a worst-case distribution by virtue of stochastic dominance for the expectation operator. Then, the agent behaves as expected utility maximizer with respect to this worst-case distribution. As the penalty is not trivial here, we might have a trade off between stochastic dominance on the payoff process and the penalty. Hence, the worst-case distribution cannot be attained any longer by stochastic dominance for the payoff process even in the monotone case. Furthermore, we observe that correlation is introduced even in quite simple contexts.

5.1 Dynamic Entropic Risk Measures

As first example we consider *dynamic entropic risk measures* as in [4] and [11] or, equivalently, *dynamic multiplier preferences* as in as in [15].

Definition 5.1. *For $\mathbb{P} \ll \mathbb{Q}$, locally, we define the* conditional relative entropy *of $\mathbb{P}$ with respect to $\mathbb{Q}$ at time $t \geq 0$ as*

$$\hat{H}_t(\mathbb{P}|\mathbb{Q}) := \mathbb{E}^{\mathbb{P}}\left[\ln\left(\frac{Z_T}{Z_t}\right)\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\frac{Z_T}{Z_t}\ln\left(\frac{Z_T}{Z_t}\right)\middle|\mathcal{F}_t\right]\mathbb{I}_{\{Z_t>0\}},$$

where $Z_t := \frac{d\mathbb{P}}{d\mathbb{Q}}\Big|_{\mathcal{F}_t}$.

Basic properties of relative entropy are stated in [7]. As we assume local equivalence, the indicator function in the last equation vanishes.

We now formally introduce dynamic multiplier preferences:

Definition 5.2. *Let $\theta > 0$. We say that dynamic variational expected reward $(\pi_t^e(X_\tau))_{t\leq T}$ is obtained by* dynamic multiplier preferences *given reference model $\mathbb{Q}$ or, equivalently, by* dynamic entropic risk measures*, if its robust representation is of the form*

$$\pi_t^e(X_\tau) = \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{P}}[X_\tau|\mathcal{F}_t] + \theta\hat{H}_t(\mathbb{P}|\mathbb{Q})\right). \tag{6}$$

Intuitively, the agent expects a reference distribution $\mathbb{Q} \in \mathcal{M}$ most likely and distributions further away – in the sense of relative entropy – seem to be more and more unlikely.

Remark 5.3. *The variational formula for relative entropy implies*

$$\pi_t^e(X_\tau) = -\theta\ln(\mathbb{E}^{\mathbb{Q}}[e^{-\frac{1}{\theta}X_\tau}|\mathcal{F}_t]).$$

Proposition 5.4. *Dynamic multiplier preferences with reference distribution $\mathbb{Q} \in \mathcal{M}$ are time-consistent: Its robust representation has minimal penalty $\alpha_t^{\min}(\mathbb{P}) = \theta\hat{H}_t(\mathbb{P}|\mathbb{Q})$ for $t \leq T$, $\mathbb{P} \in \mathcal{M}$. Hence, we have*

$$\pi_t^e(X_\tau) = X_t\mathbb{I}_{\{\tau=t\}} + \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}}\left(\int\pi_{t+1}^e(X_\tau)d\mu + \theta\hat{H}_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t))\right)\mathbb{I}_{\{\tau\geq t+1\}},$$

where we set $\hat{H}_{t+1}(\mu|\mathbb{Q}(\cdot|\mathcal{F}_t)) := \mathbb{E}^\mu[\ln(\frac{d\mu}{d\mathbb{Q}(\cdot|\mathcal{F}_t)|_{\mathcal{F}_{t+1}}})]$ which, by abuse of notation, we write as $\mathbb{E}^\mu[\ln(\frac{d\mu}{d\mathbb{Q}(\cdot|\mathcal{F}_t)}\Big|_{\mathcal{F}_{t+1}})]$, $\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}$.

Proof. The specific form of minimal penalty is shown in [11], Lemma 6.2; time-consistency in [11], p.92. The intuitive representation in terms of one-step ahead penalty can straightforwardly be achieved by Corollary 2.10; as the calculations are simple but extensive, we do not state them here. □

As we want to achieve explicit solutions, we further confine ourselves:

Assumption 5.5. *Let the underlying probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\leq T}, \mathbb{P}_0)$ *be given as the independent product of the time-t state space,* $(S, \mathcal{S}, \nu_0)$*,* $S \subset \mathbb{R}$*. Then* $\mathbb{P}_0 = \otimes_{t=1}^T \nu_o$ *and* $\mathcal{F}_s$ *is generated by the projection mappings* $\epsilon_t : \Omega \mapsto S$*,* $t \leq s$*. In particular, the* ϵ_t*'s are i.i.d. with* ν_0 *under* $\mathbb{P}_0$*.*

As in [20], we confine ourselves to the set

$$\mathcal{M}^{[a,b]} := \left\{ \mathbb{P}^\beta \approx \mathbb{P}_0 \quad : \quad \frac{d\mathbb{P}^\beta}{d\mathbb{P}_0}\bigg|_{\mathcal{F}_t} = D_t^\beta \quad \forall t,\ (\beta_t)_t \subset [a,b],\ predictable \right\},$$

$D_t^\beta := \exp(\sum_{s=1}^t \beta_s \epsilon_s - \sum_{s=1}^t L(\beta_s))$ *for some predictable process* $(\beta_t)_{t\leq T} \subset [a,b] \subset \mathbb{R}$ *and* $L(\beta_t) := \ln \int_S e^{\beta_t x} \nu_0(dx)$*.*

Notation 5.6. *The reference distribution of the entropic penalty write as* $\mathbb{Q} := \mathbb{P}^{\beta^1}$*, i.e.* $(\beta_t^1)_{t\leq T}$ *denotes the process defining the penalty's reference distribution. Other distributions in* $\mathcal{M}^{[a,b]}$ *write as* $\mathbb{P} := \mathbb{P}^{\beta^2}$*. Then, the entropic penalty with reference distribution* $\mathbb{Q}$ *is given by*

$$\alpha_t^{\min}(\mathbb{P}^{\beta^2}) = \theta \mathbb{E}^{\mathbb{P}^{\beta^2}} \left[\sum_{s=t+1}^T (\beta_s^2 - \beta_s^1)\epsilon_s - \sum_{s=t+1}^T [L(\beta_s^2) - L(\beta_s^1)] \,\middle|\, \mathcal{F}_t \right].$$

We write $\mathbb{E}^\beta := \mathbb{E}^{\mathbb{P}^\beta}$ and $\hat{H}_t(\beta^2|\beta^1) := \hat{H}_t(\mathbb{P}^{\beta^2}|\mathbb{P}^{\beta^1})$ as well as $\alpha_t^{\min}(\beta^2)$. Note, in case $\mathbb{Q} = \mathbb{P}_0$, we have $(\beta_t^1)_{t\leq T} = 0$ and hence for $\mathbb{P} = \mathbb{P}^{\beta^2}$: $\alpha_t^{\min}(\beta^2) = \theta \mathbb{E}^{\beta^2} \left[\sum_{s=t+1}^T \beta_s^2 \epsilon_s - \sum_{s=t+1}^T L(\beta_s^2) \middle| \mathcal{F}_t \right]$.

Hence, the value function is achieved by

$$\begin{aligned} V_t &= \operatorname*{ess\,sup}_{t\leq\tau\leq T} \operatorname*{ess\,inf}_{\beta^2\subset[a,b]} \mathbb{E}^{\beta^2} \Bigg[X_\tau + \theta \Bigg(\sum_{s=t+1}^T (\beta_s^2 - \beta_s^1)\epsilon_s \\ &\qquad - \sum_{s=t+1}^T [L(\beta_s^2) - L(\beta_s^1)] \Bigg) \Bigg| \mathcal{F}_t \Bigg] \\ &= \max \Big\{ X_t \quad ; \quad \operatorname*{ess\,inf}_{\beta_{t+1}^2\in[a,b]} \mathbb{E}^{\beta_{t+1}^2} \big[V_{t+1} + \theta \big((\beta_{t+1}^2 - \beta_{t+1}^1)\epsilon_{t+1} \\ &\qquad -(L(\beta_{t+1}^2) - L(\beta_{t+1}^1)) \big) \big] \Big\}. \end{aligned} \tag{7}$$

In particular, we see that the value of the problem – and hence the worst case distribution – depends on the reference distribution $\mathbb{Q} = \mathbb{P}^{\beta^1}$ of the penalty.

To further solve problems under entropic risk, we constraint ourselves to monotone problems:

Assumption 5.7. *$X_t := f(t, \epsilon_t)$, $t \leq T$, where f is a bounded measurable function that is strictly monotone in the state variable ϵ_t.*

For monotone payoff processes under multiple priors it is shown in [20] that U_t is increasing in ϵ_t. However, having a look at the proof therein, we see that this crucially depends on ϵ_t being independent of $\mathcal{F}_{t-1}$ which does not hold in case of dynamic variational preferences. Furthermore, in [20] the calculation of a worst case measure is done by virtue of stochastic dominance on the payoff process. It is intuitive that this cannot work as elegant under variational preferences as the penalty is not trivial. In particular, in the entropic case, the worst-case measure depends on the reference distribution $\mathbb{Q}$: there might be a trade off between stochastic dominance on $(X_t)_t$ and the penalty: The penalty increases the further nature moves away from $\mathbb{Q}$ and in direction of a distribution minimizing the expectation of the payoff process.

Example 5.8. *Let f be increasing and the reference distribution be $\mathbb{P}^a$. We encounter for the first term in the value function, $\mathbb{E}^{\beta^2}[f(\tau, \epsilon_\tau)|\mathcal{F}_t]$: $\mathbb{P}^a$ is stochastically dominated, i.e. minimizes that term on $\mathcal{M}^{[a,b]}$. $\mathbb{P}^a$ also minimizes the penalty: $\hat{H}_t(\beta^2|a)$ is increasing in β^2 on $[a, b]$, $\hat{H}_t \geq 0$ and zero if and only if $\mathbb{P}^{\beta^2} = \mathbb{P}^a$. Hence we have equivalence of the problem under dynamic multiplier preferences and the expected utility problem under the worst case distribution $\mathbb{P}^a$ as in Theorem 5 in [20]:*

Proposition 5.9. *Let f be increasing, $T < \infty$, and τ^a denote the optimal stopping time for the classical optimal stopping problem of $(X_t)_{t \leq T}$ under subjective distribution $\mathbb{P}^a$, i.e. τ^a solves $\max_{0 \leq \tau \leq T} \mathbb{E}^a[X_\tau]$. Let $\mathbb{Q} = \mathbb{P}^a$ be the reference distribution for the penalty. Then, τ^a is the solution to the robust problem with dynamic multiplier preferences $(\pi^e_t)_{t \leq T}$.*

Proof. For all increasing bounded measurable functions $h : \Omega \to \mathbb{R}$ and all $t \geq 1$, we have by Lemma 13 in [20]

$$
\begin{aligned}
\mathbb{E}^a[h(\epsilon_t)|\mathcal{F}_{t-1}] &= \operatorname*{ess\,inf}_{\beta^2 [a,b]} \mathbb{E}^{\beta^2}[h(\epsilon_t)|\mathcal{F}_{t-1}] + \underbrace{\min_{\beta^2 \in [a,b]} \theta \hat{H}_{t-1}(\beta^2|a)}_{=\hat{H}_t(a|a)=0} \\
&= \operatorname*{ess\,inf}_{\beta^2 \in [a,b]} \left(\mathbb{E}^{\beta^2}[h(\epsilon_t)|\mathcal{F}_{t-1}] + \theta \hat{H}_{t-1}(\beta^2|a) \right),
\end{aligned}
$$

where the last equation follows as the joint minimizer of both summands is $\mathbb{P}^a$. Given this result, we can mimic the proof of Theorem 5 in [20]: Let $(U_t)_{t\leq T}$ denote the variational Snell envelope of the problem with multiplier preferences and reference distribution $\mathbb{P}^a$ and $(U_t^a)_{t\leq T}$ the classical Snell envelope with respect to subjective prior $\mathbb{P}^a$. For $t=T$, we have $U_T = X_T = f(T,\epsilon_T) = U_T^a$ and hence increasing in ϵ_T. As by induction hypothesis U_{t+1} is an increasing function of ϵ_{t+1}, say $U_{t+1} = u(\epsilon_{t+1})$ for some bounded measurable increasing u, we have for all $t<T$

$$\begin{aligned} U_t &:= \max\left\{f(t,\epsilon_t), \quad \operatorname*{ess\,inf}_{\beta^2\in\mathcal{M}^{[a,b]}} \left(\mathbb{E}^{\beta^2}[U_{t+1}|\mathcal{F}_t] + \theta \hat{H}_t(\beta^2|a)\right)\right\} \\ &= \max\left\{f(t,\epsilon_t), \quad \mathbb{E}^a[U_{t+1}|\mathcal{F}_t] + \underbrace{\theta\hat{H}_t(a|a)}_{=0}\right\} \\ &= \max\{f(t,\epsilon_t), \quad \mathbb{E}^a[U_{t+1}|\mathcal{F}_t]\} =: U_t^a. \end{aligned}$$

□

The argument in the foregoing proof for the case $\mathbb{Q}=\mathbb{P}^a$ is that $\mathbb{P}^a$ minimizes $\mathbb{E}^{\mathbb{P}}[f(t,\epsilon_t)]$ as well as $\hat{H}_t(\mathbb{P}|a)$. Of course, this does not hold true if the reference distribution $\mathbb{Q}=\mathbb{P}^{\beta^1}$ is such that β_t^1 is not identical a. Then, we have a trade off between a decrease in the first term, $\mathbb{E}^{\mathbb{P}}[f(t,\epsilon_t)]$, which is independent of $\mathbb{P}^{\beta^1}$, and an increase of the penalty in the second term, $\hat{H}_t(\mathbb{P}|\beta^1)$, the further nature deviates from the reference distribution $\mathbb{P}^{\beta^1}$.However, moving from $\mathbb{P}^{\beta^1}$ in direction of the upper extremal distribution $\mathbb{P}^b$, both terms increase:

Proposition 5.10. *Let $\mathbb{Q}=\mathbb{P}^{\beta^1}\in\mathcal{M}^{[a,b]}$ be the reference distribution of the entropic penalty, and f be increasing. Then, the worst-case distribution $\mathbb{P}^{\bar\beta^2}$ satisfies $\bar\beta_t^2\in[a,\beta_t^1]$ $\forall t$.*

Proof. For h as above, we have

$$\operatorname*{ess\,inf}_{\beta\in[a,b]}\left\{\mathbb{E}^{\beta}[h(\epsilon_t)|\mathcal{F}_{t-1}] + \hat{H}_{t-1}(\beta|\beta^1)\right\} \leq \mathbb{E}^{\beta^2}[h(\epsilon_t)|\mathcal{F}_{t-1}] + \hat{H}_{t-1}(\beta^2|\beta^1)$$

for all $\beta_t^2\in[\beta_t^1,b]$ for all t as $\hat{H}_{t-1}(\beta^1|\beta^1)=0$ and ≥ 0 else and furthermore $\mathbb{E}^{\beta^2}[h(\epsilon_t)|\mathcal{F}_{t-1}]$ is increasing in β^2 as seen in the proof of Lemma 13 in [20]. As $\hat{H}_t(\cdot|\beta^1)$ is strictly increasing on $[\beta_t^1,b]$, we have strict inequality on $]\beta_t^1,b]$. □

Remark 5.11. *In particular, we see that the worst case distribution depends on the specific form of f, not just on f being increasing. This has severe consequences for the complexity of calculations: Let us for example take the* *case of an American call as considered in [20]. When out of the money, nature cannot just apply a distribution low enough to likely staying out of the money but also has to take care of being close enough to $\mathbb{Q}$ not to increase the penalty too much. In particular, the one step ahead worst case distribution depends on the current state:* In case of dynamic variational preferences, correlation is already introduced for the call that has independent rewards under multiple priors as shown in [20].

In general, we obtain a *negation* of Theorem 5 in [20] for our approach:

Remark 5.12. *Let $(\bar{\beta}^2_t)_t$ denote the process of the worst-case distribution for the monotone problem under dynamic multiplier preferences $(\pi^e_t)_{t\leq T}$. Then,*

$$\begin{aligned} U_t &= \max\left\{X_t; \mathbb{E}^{\bar{\beta}^2_{t+1}}[U_{t+1}|\mathcal{F}_t] + \theta H_{t+1}(\bar{\beta}^2_{t+1}|\mathbb{P}^{\beta^1}(\cdot|\mathcal{F}_t))\right\} \\ &\geq \max\left\{X_t; \mathbb{E}^{\bar{\beta}^2_{t+1}}[U_{t+1}|\mathcal{F}_t]\right\} = U^{\bar{\beta}^2}_t, \end{aligned}$$

where $U^{\bar{\beta}^2}_t$ denotes the classical Snell envelope of the optimal stopping problem under subjective prior given by $\bar{\beta}^2$. In particular, we see that

$$\tau^* = \inf_t\{X_t = U_t\} \geq \inf_t\{X_t = U^{\bar{\beta}^2}_t\} = \tau^{\bar{\beta}^2 *}.$$

The intuition in [20] is not valid anymore: The agent does not behave as the expected utility maximizer under the worst case distribution. However, sophistication of the penalty has increased the problem's value.

Example 5.13 (American Options in CRR-Model)**.** *Consider an agent with expected reward $(\pi^e_t)_{t\leq T}$ given by parameter $\theta = 1$ and reference distribution $\mathbb{P}^b$, i.e. the agent to consider the market as "emerging". We consider American options for the Cox-Ross-Rubinstein (CRR) model: Let $\Omega := \{0,1\}^T$, $T < \infty$. Let $\epsilon_t : \Omega \to \{0,1\}$, $t \leq T$, be the projection mappings and $\mathbb{P}_0$ such that ϵ_t's are i.i.d. under $\mathbb{P}_0$ with $\mathbb{P}_0[\epsilon_t = 1] = \mathbb{P}_0[\epsilon_t = 0] = \frac{1}{2}$. Let $\mathcal{M}^{[a,b]}$ be given as in Assumption 5.5. As in [20], we then have for all $\beta := (\beta_t)_t$ that $\mathbb{P}^\beta[\epsilon_t = 1|\mathcal{F}_{t-1}] \in [\underline{p}; \bar{p}]$, where $\underline{p} := \frac{e^a}{1+e^a}$ and $\bar{p} := \frac{e^b}{1+e^b}$. Let $\mathbb{P}^a$ be again the distribution induced by the constant process with $\beta_t = a$ for all t and equivalently for $\mathbb{P}^b$. Then, under $\mathbb{P}^a$, ϵ_t's are i.i.d. with $\mathbb{P}^a[\epsilon_t] = \underline{p}$ and equivalently for $\mathbb{P}^b$ with $\mathbb{P}^b[\epsilon_t] = \bar{p}$.*

The "ingredients" of the CRR-model are given by a risk-less asset with value process $B_t = (1+r)^t$ for some fixed interest rate $r > -1$ and a risky asset with value process S_t at t such that $S_0 = 1$ and

$$S_{t+1} = S_t \cdot \begin{cases} (1+d) & \text{if } \epsilon_{t+1} = 1, \\ (1+c) & \text{if } \epsilon_{t+1} = 0, \end{cases}$$

where we assume the constants to satisfy $-1 < c < r < d$ for the model not to allow for arbitrage opportunities.

Let $A^p(t, S_t)$ bei an American put and, hence, decreasing in S_t for all t. Let $(U_t^b)_{t \le T}$ denote the classical Snell envelope of $A^p(t, S_t)$ under subjective probability $\mathbb{P}^b$, i.e.

$$U_t^b(t, S_t) = \max\left\{A^p(t, S_t); \bar{p}U_t^b(t+1, S_t(1+d)) + (1-\bar{p})U_t^b(t+1, S_t(1+c))\right\}.$$

The following assertion holds: The variational Snell envelope $(U_t)_{t \le T}$ of the American put problem with dynamic multiplier preferences $(\pi_t^e)_{t \le T}$ and reference distribution $\mathbb{P}^b$ satisfies $(U_t)_{t \le T} = (U_t^b)_{t \le T}$. In particular, the worst case distribution is given by $\mathbb{P}^b$ and, as the penalty vanishes for this distribution, the optimal stopping time is given by $\tau^* = \inf\{t \ge 0 | A^p(t, S_t) = U_t^b\} = \tau^{b*}$, i.e. the optimal stopping time τ^{b*} of the problem under subjective prior $\mathbb{P}^b$.

The proof of this assertion is immediate by virtue of stochastic dominance: As in Appendix H in [20], we show for the variational Snell envelope $(U_t)_{t \le T}$ that $U_t = u(t, S_t) = U_t^b$, $t \le T$, for a function u that is decreasing in the second variable: First, we have $U_T = A^p(T, S_T) = U_T^b$ by definition. For an inductive proof, we write with a slight but intuitively understandable misuse of notation $\hat{H}_t(p_{t+1} \otimes p_{t+2} \otimes \ldots | \mathbb{P}^b)$[7] for $p_i \in [\underline{p}; \bar{p}]$ and note that $\hat{H}_t(\bar{p} \otimes \bar{p} \otimes \ldots | \mathbb{P}^b) = 0$ and ≥ 0 else. From the induction hypothesis, we have $u(t+1, S_t(1+d)) \le u(t+1, S_t(1+c))$ and hence

$$\begin{aligned} U_t = \max\Big\{A^p(t, S_t) \quad ; \quad & \min_{p_{t+1} \in [\underline{p}; \bar{p}]} \{p_{t+1}u(t+1, S_t(1+d)) \\ & \qquad +(1-p_{t+1})u(t+1, S_t(1+c)) \\ & \qquad +H_t(p_{t+1} \otimes \bar{p} \otimes \ldots | \mathbb{P}^b)\}\Big\} \\ = \max\Big\{A^p(t, S_t) \quad ; \quad & \bar{p}u(t+1, S_t(1+d)) + (1-\bar{p})u(t+1, S_t(1+c)) \\ & + \underbrace{H_t(\bar{p} \otimes \bar{p} \otimes \ldots | \mathbb{P}^b)}_{=0}\Big\} = U_t^b. \end{aligned}$$

In a way, the result in the example is more like a self fulfilling prophecy as the agent assumes the worst-case distribution to be the most likely one. The same holds true for an American call with reference distribution $\mathbb{P}^a$: In that case, the reference distribution is also the worst-case one. However, due to the tradeoff effects, $\mathbb{P}^a$ is not the worst-case distribution for the American call when $\mathbb{P}^b$ is the reference distribution; as $\mathbb{P}^b$ is not worst-case distribution for the American put when $\mathbb{P}^a$ is the reference one.

[7] Formally: $\hat{H}_t(p_{t+1} \otimes p_{t+2} \otimes \ldots | \mathbb{P}^b) := \hat{H}_t(\mathbb{P}^\beta | \mathbb{P}^b)$ with $(\beta_t)_{t \le T}$ such that $\mathbb{P}^\beta[\epsilon_t = 1 | \mathcal{F}_{t-1}] = p_t$ for $t \le T$; well defined as $p_1, \ldots, p_t$ drops by general definition of $\hat{H}_t$.

5.2 Dynamic Generalized AVaR

In the financial industry *value at risk (VaR)* still is a standard method for risk quantification and risk management. Prominence of VaR is due to its simplicity and intuitive appeal. Though widely used, VaR is not convex: Applying VaR, a risk officer runs the danger or accumulating a highly risky portfolio. A standard example is inter alia given in [16]. Being aware of VaR's shortcomings, *average value at risk ($AVaR$)* is introduced taking into account not only loss probabilities in terms of quantiles, as VaR does, but also the amount of possible loss. Nevertheless, $AVaR$ is still intuitive and easily implemented by virtue of

$$AVaR_\lambda(X_T) := \frac{1}{\lambda}\int_0^\lambda VaR_m(X_T)dm$$

for some level $\lambda \in]0,1[$. It can be shown that AVaR satisfies a robust representation with minimal penalty

$$\alpha^{\min}(\mathbb{Q}) = \begin{cases} 0 & \text{if } \frac{d\mathbb{Q}}{d\mathbb{P}_0} \leq \frac{1}{\lambda}, \\ \infty & \text{else.} \end{cases}$$

Hence $AVaR$ is a coherent risk measure. Elaborate discussions on AVaR can be found in [16]. A generalization, called *utility based shortfall risk measure*, is introduced in [12]. A convenient representation for $AVaR$ which has an immediate generalization to a convex risk measure, called *generalized $AVaR$ ($gAVaR$)* here, is given in [12]. This convex risk measure gives raise to a variational preference in the canonical way.

As shown in [6] as well as [2] the natural dynamic extension of $AVaR$, and hence of $gAVaR$, in terms of conditional expectations is *not* time-consistent. We thus define a time-consistent dynamic version of $gAVaR$, directly inducing a time-consistent dynamic variational preference, recursively in terms of the penalty function as in [15] by composing one period ahead penalties.

To introduce a dynamic version, we start with the static convex risk measure $gAVaR$ for some end period payoff $X_T \in L^\infty_T$ as in [5]:

Definition 5.14. *For $(\theta, \beta, p) \in]0,\infty[\times]1,\infty[\times[1,\infty[$, define the risk measure gAVaR for $X_T \in L^\infty_T$, called* generalized Average Value at Risk (gAVaR)*:*

$$gAVaR^{\beta,p}_\theta(X_T) := \min_{s\in\mathbb{R}}\left\{\frac{1}{\theta}\left\|(s-X_T)^+\right\|_p^\beta - s\right\},$$

where $\|\cdot\|_p := (\mathbb{E}^{\mathbb{P}_0|\mathcal{F}_T}[|\cdot|^p])^{\frac{1}{p}}$ denotes the usual p-norm.

For ease of notation, we do not explicitly state the parameters but just write $gAVaR$ instead of $gAVaR_\theta^{\beta,p}$ when these are obvious. We have:

Proposition 5.15. *(a) For* $(\theta,\beta,p)\in\ ]0,\infty[\times]1,\infty[\times[1,\infty[$*,* $gAVaR_\theta^{\beta,p}$ *is a* convex *risk measure with minimal penalty* $\alpha^{\min^{\text{gAVaR}}}(\mathbb{Q}) := c\left\|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}\right\|_q^d$*, where* $q:=\frac{p}{p-1}$*,* $d:=\frac{\beta}{\beta-1}$ *and* $c=\theta^{d-1}\beta^{1-d}d^{-1}$*. Hence*

$$gAVaR_\theta^{\beta,p}(X_T)=\sup_{\mathbb{Q}\in\mathcal{M}}\left\{\mathbb{E}^{\mathbb{Q}|_{\mathcal{F}_T}}[-X_T]-c\left\|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}\right\|_q^d\right\}.$$

(b) For $\theta\in]0,1[$*,* $\beta=p=1$*, we have* $\|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}\|_\infty=\operatorname{ess\,sup}|\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}|$ *and hence the robust representation becomes*

$$\begin{aligned}gAVaR_\theta^{1,1}(X_T) &= \sup_{\mathbb{Q}\in\mathcal{M}}\left\{\mathbb{E}^{\mathbb{Q}|_{\mathcal{F}_T}}[-X_T]\,\middle|\,0\le\frac{d\mathbb{Q}|_{\mathcal{F}_T}}{d\mathbb{P}_0|_{\mathcal{F}_T}}\le\frac{1}{\theta}\right\}\\ &= AVaR_\theta(X_T).\end{aligned}$$

Proof. cp. [5]. □

A time-consistent dynamic version of $AVaR$ for end period payoff X_T is recursively achieved in [6]. Mimicking this approach by virtue of the definition of time-consistency for dynamic convex risk measures, i.e. $\rho_t=\rho_t(-\rho_{t+1})$, we would obtain a time-consistent dynamic version of $gAVaR_\theta^{\beta,p}$. However, this would not be in terms of a robust representation needed to achieve explicit solutions in terms of worst-case distributions. Hence, we use the minimal penalty $\alpha^{\min^{\text{gAVaR}}}$ of the static $gAVaR$ as defined in Proposition 5.15: We apply the recursive procedure from [15], Theorem 2, in terms of one period ahead penalties $(\gamma_t^{gAVaR})_{t\le T}$ to achieve a time-consistent dynamic minimal penalty $(\alpha_t^{\min^{\text{gAVaR}}})_{t\le T}$. Define γ_t^{gAVaR} on $\mathcal{M}|_{\mathcal{F}_{t+1}}$ by

$$\gamma_t^{gAVaR}(\mathbb{Q}|_{\mathcal{F}_{t+1}}(\cdot|F_t)):=c\left(\mathbb{E}^{\mathbb{P}_0}\left[\left(\left.\frac{d\mathbb{Q}}{d\mathbb{P}_0}\right|_{\mathcal{F}_{t+1}}\right)^q\,\middle|\,\mathcal{F}_t\right]\right)^{\frac{d}{q}},$$

and recursively $(\alpha_t^{\min^{\text{gAVaR}}})_{t\le T}$:

Definition 5.16. *Let* $F_t\in\mathcal{F}_t$*. Set*

$$\begin{aligned}\alpha_T^{\min^{\text{gAVaR}}}(\mathbb{Q})(\omega) &:= \begin{cases}0 & \text{if } \mathbb{Q}=\mathbb{I}_{\{\omega\}},\\ \infty & \text{else}\end{cases}\qquad \text{for } \omega\in\Omega,\\ \alpha_t^{\min^{\text{gAVaR}}}(\mathbb{Q})(F_t) &:= \int\alpha_{t+1}^{\min^{\text{gAVaR}}}(\mathbb{Q}(\cdot|\mathcal{F}_{t+1}))d\mathbb{Q}(\cdot|F_t)+\gamma_t^{gAVaR}(\mathbb{Q}(\cdot|F_t)|_{\mathcal{F}_{t+1}})\\ &\qquad \text{if } \mathbb{Q}(F_t)>0,\\ \alpha_t^{\min^{\text{gAVaR}}}(\mathbb{Q})(F_t) &:= \infty\quad \text{if } \mathbb{Q}(F_t)=0,\end{aligned}$$

for $t < T$. Hence, for $X_T \in L_T^\infty$, we define $(\pi_t^{\alpha^{\min^{\mathrm{gAVaR}}}})_{t \le T}$ by

$$\pi_t^{\alpha^{\min^{\mathrm{gAVaR}}}}(X_T) := \operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left\{ \mathbb{E}^{\mathbb{Q}}[X_T | \mathcal{F}_t] + \alpha_t^{\min^{\mathrm{gAVaR}}}(\mathbb{Q}) \right\}.$$

Remark 5.17. $(\pi_t^{\alpha^{\min^{\mathrm{gAVaR}}}})_{t \le T}$ *is a time-consistent dynamic variational preference: It is a dynamic variational preference by virtue of its definition in terms of a robust representation. Time-consistency follows by Proposition 2.9 as* $(\alpha_t^{\min^{\mathrm{gAVaR}}})_{t \le T}$ *is defined recursively in terms of the no-gain condition.*

Thus, we have a recursive representation for the variational Snell envelope of time-consistent dynamic variational preferences $(\pi_t^{\alpha^{\min^{\mathrm{gAVaR}}}})_{t \le T}$. This representation enables us, given an explicit structure of $(X_t)_{t \le T}$, to solve the problem for an optimal stopping time τ^* as in Theorem 4.1.

6 Conclusions

We have generalized the theory of optimal stopping with multiple priors as set out in [20] to dynamic variational preferences introduced in [15] or, equivalently, dynamic convex risk measures in [11]. To achieve our results, we have introduced the notion of variational supermartingales as a generalization of the usual notion of supermartingales. For this concept, we have obtained results including a Doob decomposition and optional sampling. These enabled us to generalize the classical optimal stopping approach for an expected utility maximizer in [17] (Section VI.1) in terms of Snell envelopes to the case of dynamic variational preferences by virtue of variational Snell envelopes. We have achieved minimal optimal stopping times and an explicit characterization of worst-case distributions. We have shown that the solution to the infitite horizon problem can be approximated by a sequence of solutions for an approximating sequence of finite horizon problems. A further insight is a minimax theorem similar to a minimax result in [23] but making use of time-consistency.

Our results were applied to prominent examples: dynamic entropic risk and dynamic generalized average value at risk. For the latter, we are not aware of any reference having considered this notion in a dynamic context.

To conclude, the virtue of the present article is that optimal stopping problems are now solved for dynamic variational preferences or, equivalently, dynamic convex risk measures. This is important for applications on financial markets: coherent risk measures, as a robust approach reducing model risk, are too conservative. Convex risk measures are a comprehensive vehicle to more liberally assess risk while still being robust, as no specific probabilistic model is assumed, and satisfying the "margin of conservatism" required in the Basel II accord.

Our approach leaves a realm for further generalizations. It seems possible to achieve the results in this article for general time-consistent (monotone) monetary risk measures, i.e. relaxing the convexity assumption. Of course, in that case, the robust representation in terms of penalty α does not hold anymore. However, as explicitly stated, the variational Snell envelope does not need a robust representation and can hence be generalized to more general risk measures as done in [4], Chapter 5.3. It is shown that the value function is time-consistent and again a monetary risk measure. Due to a missing robust representation, the solution is not explicit. The next direction in which theory might be generalized is to relax the assumption of the payoff process being essentially bounded. Several of the cited references consider convex risk measures for L^p processes or, as in [5], risk measures defined on Orlicz spaces.

Besides these theoretical considerations, further examples and concrete applications might be elaborated: dynamic convex risk measures based on expected shortfall, inter alia elaborated in [12] and [9], are a generalization of dynamic entropic risk measures or dynamic multiplier preferences when loss is not exponential.

At last, the problem might be considered in a continuous time setting. Several approaches to convex risk measures in a time-continuous framework are available: In [3], dynamic convex risk measures are achieved by virtue of BMO martingales. A special case of this approach is given in [22] via BSDE resulting in g-expectations as introduced in [18].

A Proofs

A.1 Proof of Proposition 2.9

Proof of Proposition 2.9. (i) $\tau \leq t$: In this case, X_τ is $\mathcal{F}_t$-measurable and in particular $\mathcal{F}_{t+1}$-measurable. Hence, by conditional cash invariance, we have

$$\pi_t(X_\tau) = X_\tau = \pi_{t+1}(X_\tau)$$

and hence $\pi_t(X_\tau) = \pi_t(\pi_{t+1}(X_\tau))$.

(ii) $\tau \geq t+1$: "$\leq$": If, for all $\mathbb{Q} \in \mathcal{A}$, we have

$$\alpha_t^{\min}(\mathbb{Q}) \leq \mathbb{E}^{\mathbb{Q}}\left[\alpha_{t+1}^{\min}(\mathbb{Q})|\mathcal{F}_t\right] + \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q} \otimes_{t+1} \mathbb{P}),$$

then also

$$\alpha_t^{\min}(\mathbb{Q} \otimes_{t+1} \mathbb{P}) \leq \mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}}\left[\alpha_{t+1}^{\min}(\mathbb{Q} \otimes_{t+1} \mathbb{P})|\mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q}).$$

It is immediate that

$$\left\{\mathbb{E}^{\mathbb{P}}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{P}) : \mathbb{P} \in \mathcal{M}\right\}$$

is downward directed. Hence, there exists a sequence $(\mathbb{P}_n)_n \subset \mathcal{M}$ such that

$$\mathbb{E}^{\mathbb{P}_n}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{P}_n) \searrow \pi_{t+1}(X_\tau).$$

As $\mathcal{M}$ is closed under pasting, we obtain for all $\mathbb{Q} \in \mathcal{M}$ and such $\mathbb{P}_n$:

$$\begin{aligned}
\pi_t(X_\tau) &= \operatorname*{ess\,inf}_{\mathbb{P},\mathbb{Q}} \left(\mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P})\right) \\
&\leq \mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}_n}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P}_n) \\
&\leq \underbrace{\mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}_n}[X_\tau|\mathcal{F}_t]}_{=\mathbb{E}^{\mathbb{Q}}[\mathbb{E}^{\mathbb{P}_n}[X_\tau|\mathcal{F}_{t+1}]|\mathcal{F}_t]} + \mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}_n}\left[\alpha_{t+1}^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P}_n)|\mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q}) \\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{P}_n}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{P}_n)\middle|\,\mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q}).
\end{aligned}$$

Hence, letting $n \to \infty$, we achieve for all $\mathbb{Q} \in \mathcal{M}$

$$\pi_t(X_\tau) \leq \mathbb{E}^{\mathbb{Q}}\left[\pi_{t+1}(X_\tau)|\,\mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q}).$$

Applying the essential infimum to this expression yields

$$\pi_t(X_\tau) \leq \pi_t(\pi_{t+1}(X_\tau)).$$

"$\geq$": Assuming

$$\alpha_t^{\min}(\mathbb{Q}) \geq \mathbb{E}^{\mathbb{Q}}\left[\alpha_{t+1}^{\min}(\mathbb{Q})|\mathcal{F}_t\right] + \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P})$$

for all $\mathbb{Q} \in \mathcal{M}$, we obtain

$$\begin{aligned}
& \mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \\
\geq\ & \mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \mathbb{E}^{\mathbb{Q}}\left[\alpha_{t+1}^{\min}(\mathbb{Q})|\mathcal{F}_t\right] + \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P}) \\
\geq\ & \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}}\left[\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q})\middle|\,\mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P})\right) \\
\geq\ & \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}\otimes_{t+1}\mathbb{P}}\left[\pi_{t+1}(X_\tau)|\,\mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q}\otimes_{t+1}\mathbb{P})\right) \\
\geq\ & \pi_t(\pi_{t+1}(X_\tau)).
\end{aligned}$$

Applying the essential infimum, we achieve

$$\pi_t(X_\tau) \geq \pi_t(\pi_{t+1}(X_\tau)).$$

□

Proof of Corollary 2.10. By conditional cash invariance, we have

$$\pi_t(X_\tau) = X_\tau \mathbb{I}_{\{\tau \le t\}} + \pi_t(\pi_{t+1}(X_\tau))\mathbb{I}_{\{\tau \ge t+1\}}.$$

As π_{t+1} is $\mathcal{F}_{t+1}$-measurable we have, whenever $\tau \ge t+1$,

$$\begin{aligned}
\pi_t(\pi_{t+1}(X_\tau)) &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[\pi_{t+1}(X_\tau)|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\
&= \operatorname*{ess\,inf}_{\mathbb{R},\mathbb{P}\in\mathcal{M}} \left(\underbrace{\mathbb{E}^{\mathbb{R}\otimes_{t+1}\mathbb{P}}[\pi_{t+1}(X_\tau)|\mathcal{F}_t]}_{\mathbb{E}^{\mathbb{R}|_{\mathcal{F}_{t+1}}}[\pi_{t+1}(X_\tau)|\mathcal{F}_t]} + \alpha_t^{\min}(\mathbb{R}\otimes_{t+1}\mathbb{P}) \right) \\
&= \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}},\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mu}[\pi_{t+1}(X_\tau)|\mathcal{F}_t] + \alpha_t^{\min}(\mu\otimes_{t+1}\mathbb{P})\right) \\
&= \operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\mathbb{E}^{\mu}[\pi_{t+1}(X_\tau)|\mathcal{F}_t] + \underbrace{\operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \alpha_t^{\min}(\mu\otimes_{t+1}\mathbb{P})}_{=:\gamma_t(\mu)} \right).
\end{aligned}$$

□

A.2 Proof of Proposition 3.2

The following lemmata directly generalize Lemmata 9 and 10 in [20] to dynamic variational preferences applying interim results from [11].

Lemma A.1. *Let $Z \in L^\infty_T$. Then, for any stopping time τ, the set*

$$\left\{\mathbb{E}^{\mathbb{Q}}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}) : \mathbb{Q}\in\mathcal{M}, \mathbb{P}^\tau|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}\right\}$$

is downward directed, i.e. for any $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathcal{M}$ with $\mathbb{Q}_1|_{\mathcal{F}_\tau} = \mathbb{Q}_2|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}$, there exists $\mathbb{Q}_3 \in \mathcal{M}$ with $\mathbb{Q}_3|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}$ such that

$$\mathbb{E}^{\mathbb{Q}_3}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_3) = \min\left\{\mathbb{E}^{\mathbb{Q}_1}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_1); \mathbb{E}^{\mathbb{Q}_2}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_2)\right\}.$$

Proof. Let $\mathbb{Q}_1$ and $\mathbb{Q}_2$ be chosen as above. Consider some arbitrary set $B \in \mathcal{F}_\tau$ and define $\mathbb{Q}_3$ by virtue of

$$\frac{d\mathbb{Q}_3}{d\mathbb{P}_0} := \mathbb{I}_B \frac{d\mathbb{Q}_1}{d\mathbb{P}_0} + \mathbb{I}_{B^C}\frac{d\mathbb{Q}_2}{d\mathbb{P}_0}.$$

We have $\mathbb{Q}_3 \in \mathcal{M}$, $\mathbb{Q}_3|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}$, and by [11], Lemma 3.3, we have the *local propery* of dynamic minimal penalty:

$$\alpha_\tau^{\min}(\mathbb{Q}_3) = \mathbb{I}_B \alpha_\tau^{\min}(\mathbb{Q}_1) + \mathbb{I}_{B^C}\alpha_\tau^{\min}(\mathbb{Q}_2) < \infty.$$

Now, define $B \in \mathcal{F}_\tau$ by

$$B := \left\{ \mathbb{E}^{\mathbb{Q}_2}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_2) \geq \mathbb{E}^{\mathbb{Q}_1}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_1) \right\}.$$

Then, by definition of $\mathbb{Q}_3$ and the local property, we have

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}_3}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_3) \\ = \ & \left(\mathbb{E}^{\mathbb{Q}_1}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_1)\right) \mathbb{I}_B + \left(\mathbb{E}^{\mathbb{Q}_2}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_2)\right) \mathbb{I}_{B^C} \\ = \ & \min \left\{ \mathbb{E}^{\mathbb{Q}_1}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_1); \mathbb{E}^{\mathbb{Q}_2}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q}_2) \right\}, \end{aligned}$$

which completes the proof. □

Lemma A.2. *For all $\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}$ there exists $\mathbb{P}^* \in \mathcal{M}(\cdot|\mathcal{F}_{t+1})$ such that $\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}^*) = \operatorname{ess\,inf}_{\mathbb{P} \in \mathcal{M}(\cdot|\mathcal{F}_{t+1})} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P})$.*

Proof. By the weak compactness of the set of density processes, it is sufficient to show that there exists a sequence $(\mathbb{P}_n)_n \subset \mathcal{M}(\cdot|\mathcal{F}_{t+1})$ such that

$$\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_n) \searrow \operatorname*{ess\,inf}_{\mathbb{P} \in \mathcal{M}(\cdot|\mathcal{F}_{t+1})} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}).$$

Hence, it suffices to show that for all $\mu \in \mathcal{M}|_{\mathcal{F}_{t+1}}$, the set

$$\{\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_n) : \mathbb{P} \in \mathcal{M}(\cdot|\mathcal{F}_{t+1})\}$$

is downward directed, i.e. for every $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}(\cdot|\mathcal{F}_{t+1})$, there exists a $\mathbb{P}_3 \in \mathcal{M}(\cdot|\mathcal{F}_{t+1})$ sucht that

$$\min \left\{ \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_1), \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_2) \right\} = \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_3). \tag{8}$$

Indeed, set $A := \{\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_1) < \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_2)\}$ and define $\mathbb{P}_3$ by

$$\frac{d\mathbb{P}_3}{d\mathbb{P}_0} := \mathbb{I}_A \frac{d\mathbb{P}_1}{d\mathbb{P}_0} + \mathbb{I}_{A^C} \frac{d\mathbb{P}_2}{d\mathbb{P}_0}.$$

By Lemma 3.3 in [11], we have $\alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_3) = \mathbb{I}_A \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_1) + \mathbb{I}_{A^C} \alpha_t^{\min}(\mu \otimes_{t+1} \mathbb{P}_2)$ since $\mu \otimes_{t+1} \mathbb{P}_3 = (\mu \otimes_{t+1} \mathbb{P}_1)\mathbb{I}_A + (\mu \otimes_{t+1} \mathbb{P}_2)\mathbb{I}_{A^C}$. Hence, equation (8) to holds. □

Lemma A.3. *Let $Z \in L_s^\infty$, $s \leq T$, and τ a stopping time. Then there exists $\mathbb{P}^\tau \in \mathcal{M}$ sucht that $\mathbb{P}^\tau|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}$ and*

$$\operatorname*{ess\,inf}_{\mathbb{Q} \in \mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q})\right) = \mathbb{E}^{\mathbb{P}^\tau}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{P}^\tau)\mathbb{I}_{\{s>\tau\}}.$$

Proof. In case $\tau \geq s$, the assertion obviously holds true by conditional cash invariance: Both sides of the equation equal Z.

Hence, we consider the case $\tau < s$. To show: $\exists\ (\mathbb{P}_m)_m \subset \mathcal{M}$ with $\mathbb{P}_m|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}$ such that

$$\begin{aligned} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{Q})\right) &= \lim_{m\to\infty} \mathbb{E}^{\mathbb{P}_m}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{P}_m). \\ &= \mathbb{E}^{\mathbb{P}_\infty}[Z|\mathcal{F}_\tau] + \alpha_\tau^{\min}(\mathbb{P}_\infty) \end{aligned}$$

for some $\mathbb{P}_\infty \in \mathcal{M}$ by weak closeness assumption as $\mathbb{P}_m \to_{m\to\infty} \mathbb{P}_\infty$ weakly. Setting $\mathbb{P}_\infty =: \mathbb{P}^\tau$ then concludes the proof.

It leaves to prove existence of a sequence $(\mathbb{P}_m)_m \subset \mathcal{M}$ with the above properties: As in the proof of Lemma 10 in [20], Bayes rule as well as the dependence of α_τ only on the $\mathcal{F}_\tau$-conditional distribution allows us to restrict attention to $\mathbb{Q} \in \mathcal{M}$ such that $\mathbb{Q} = \mathbb{P}_0$ on $\mathcal{F}_t$. This is made explicit in Corollary 2.4 in [11]. Hence, existence of the sequence is assured by Lemma A.1 showing the set $\left\{\mathbb{E}^{\mathbb{Q}}[Z|\mathcal{F}_\tau] + \alpha_\tau(\mathbb{Q}) \quad : \quad \mathbb{Q} \in \mathcal{M}, \quad \mathbb{P}^\tau|_{\mathcal{F}_\tau} = \mathbb{P}_0|_{\mathcal{F}_\tau}\right\}$. to be downward directed. □

Corollary A.4 (from Lemma A.3)**.** *For all* $Z \in L^\infty_{t+1}$, $\exists \mu^* \in \mathcal{M}|_{\mathcal{F}_{t+1}}$ *s.t.*

$$\operatorname*{ess\,inf}_{\mu\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\mathbb{E}^\mu[Z|\mathcal{F}_t] + \gamma_t(\mu)\right) = \mathbb{E}^{\mu^*}[Z|\mathcal{F}_t] + \gamma_t(\mu^*).$$

Lemma A.2 and Corollary A.4 prove Proposition 3.2:

Proof of Proposition 3.2. ad (a): Let $(M_t)_{t\in\mathbb{N}}$ be a submartingale for every $\mathbb{Q} \in \mathcal{M}$, i.e.

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] \geq M_t \quad \forall \mathbb{Q} \in \mathcal{M} \\ \Rightarrow\ & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\} \\ & \geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}) \\ & = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] \geq M_t. \end{aligned}$$

ad (b): "$\Leftarrow$" Let $\mathbb{Q}^* \in \mathcal{M}$ be such that $M_t \geq \mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*)$. Then obviously, $M_t \geq \operatorname{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\}$ and hence $(M_t)_{t\in\mathbb{N}}$ is a variational supermartingale w.r.t. $(\alpha_t^{\min})_{t\in\mathbb{N}}$ and a $\mathbb{Q}^*$-supermartingale: $M_t \geq \mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*) \geq \mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t]$.

"$\Rightarrow$" By making use of Corollary 2.10, we will explicitly construct a *worst-case distribution* $\mathbb{Q}^* \in \mathcal{M}$ that satisfies

$$\begin{aligned} M_t &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ &= \mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*) \end{aligned}$$

for $t < T$, attained due to continuity from below. Let $\mathcal{M}(\cdot|\mathcal{F}_t)$ denote the set of all distributions in $\mathcal{M}$ conditional on $\mathcal{F}_t$. We have

$$\begin{aligned}
M_t &\geq \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{P}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P})\right) \\
&= \operatorname*{ess\,inf}_{\mathbb{Q}|_{\mathcal{F}_{t+1}}\in\mathcal{M}|_{\mathcal{F}_{t+1}}} \left(\mathbb{E}^{\mathbb{Q}|_{\mathcal{F}_{t+1}}}[\underbrace{\pi_{t+1}(M_{t+1})}_{M_{t+1}}|\mathcal{F}_t] + \gamma_t(\mathbb{Q}|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t))\right) \\
&\quad \text{by Corollary 2.10} \\
&= \mathbb{E}^{\mathbb{Q}^*|_{\mathcal{F}_{t+1}}}[M_{t+1}|\mathcal{F}_t] + \gamma_t(\mathbb{Q}^*|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t)) \\
&\quad \text{with } \mathbb{Q}^* \text{ as achieved in Corollary A.4} \\
&= \mathbb{E}^{\mathbb{Q}^*|_{\mathcal{F}_{t+1}}}[M_{t+1}|\mathcal{F}_t] + \operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}(\cdot|\mathcal{F}_{t+1})} \alpha_t^{\min}(\mathbb{Q}^*|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t)\otimes_{t+1}\mathbb{P}) \\
&= \mathbb{E}^{\mathbb{Q}^*|_{\mathcal{F}_{t+1}}\otimes_{t+1}\mathbb{Q}^*(\cdot|\mathcal{F}_{t+1})}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t)\otimes_{t+1}\mathbb{Q}^*(\cdot|\mathcal{F}_{t+1})) \\
&\quad \text{by Lemma A.2 and Bayes rule on the first summand} \\
&= \mathbb{E}^{\mathbb{Q}^*(\cdot|\mathcal{F}_t)}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*(\cdot|\mathcal{F}_t)) \\
&= \mathbb{E}^{\mathbb{Q}^*}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*),
\end{aligned}$$

where $\mathbb{Q}^*(\cdot|\mathcal{F}_t) := \mathbb{Q}^*|_{\mathcal{F}_{t+1}}(\cdot|\mathcal{F}_t)\otimes_{t+1}\mathbb{Q}^*(\cdot|\mathcal{F}_{t+1})$ is the pasting of the $\mathbb{Q}^*|_{\mathcal{F}_s}$'s, $s \geq t$, and $\mathbb{Q}^*$ the respective recursive pasting. The last equality makes use of the fact that the dynamic minimal penalty only depends on conditionals – hence justifies our intuitive notation – and that the conditional expectation is the unconditional one with respect to the conditional distribution. □

A.3 Proofs of Propositions 3.4 & 3.5

Proof of Proposition 3.4. (a) Uniqueness: Let $S = M - A$ as in the assertion:

$$\begin{aligned}
&\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{t+1} - S_t + A_{t+1} - A_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\
= &\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_{t+1} - M_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\
= &\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) - M_t = 0,
\end{aligned}$$

as M is assumed to be a variational martingale. By uniqueness of $\alpha_t^{\min}$, due to the relevance assumption, and as A is assumed to be predictable, we have

$$A_{t+1} = A_t - \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{t+1} - S_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right).$$

This shows uniqueness of A and hence of M.

(b) Existence: Define $(A_t)_{t\in\mathbb{N}}$ by virtue of $A_0 = 0$ and

$$A_{t+1} := A_t - \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{t+1} - S_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right).$$

Then, $A_{t+1} \in \mathcal{F}_t$, i.e. $(A_t)_{t\in\mathbb{N}}$ is predictable and non-decreasing as $(S_t)_{t\in\mathbb{N}}$ is a variational supermartingale. Set $M_t := S_t + A_t$, then

$$\begin{aligned}
& \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) - M_t \\
= \; & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{t+1} - S_t + A_{t+1} - A_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\
= \; & A_{t+1} - A_t + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{t+1} - S_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = 0.
\end{aligned}$$

Thus, $(M_t)_{t\in\mathbb{N}}$ is a variational martingale with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$. □

Proof of Proposition 3.5. We know from Proposition 3.2 that there exists a worst-case distribution $\mathbb{P}^* \in \mathcal{M}$ such that

$$S_t \geq \mathbb{E}^{\mathbb{P}^*}[S_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P}^*).$$

(i) First, we show that for fixed $N \in \mathbb{N}$ a stopped "supermartingale modulo penalty" $(S_{N\wedge t})_{t\in\mathbb{N}}$ is again one such. I.e.

$$S_{N\wedge t} \geq \mathbb{E}^{\mathbb{P}^*}[S_{N\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>t\}}. \tag{9}$$

Indeed, we have

$$\begin{aligned}
S_{N\wedge t} \; = \; & S_0 + \sum_{k=1}^{t} \mathbb{I}_{\{N\geq k\}}(S_k - S_{k-1}) \\
\geq \; & S_0 + \sum_{k=1}^{t} \mathbb{I}_{\{N\geq k\}}(S_k - S_{k-1}) \\
& + \mathbb{I}_{\{N\geq t+1\}}(\mathbb{E}^{\mathbb{P}^*}[S_{t+1} - S_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P}^*)) \\
= \; & \mathbb{E}^{\mathbb{P}^*}\left[S_0 + \sum_{k=1}^{t} \mathbb{I}_{\{N\geq k\}}(S_k - S_{k-1}) + \mathbb{I}_{\{N\geq t+1\}}(S_{t+1} - S_t)|\mathcal{F}_t\right] \\
& + \alpha_t^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>t\}} \\
= \; & \mathbb{E}^{\mathbb{P}^*}[S_{N\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>t\}}.
\end{aligned}$$

(ii) Note: By (i), we have for a variational martingale $(M_t)_{t\in\mathbb{N}}$

$$\mathbb{E}^{\mathbb{P}^*}[M_{N\wedge t}] = \mathbb{E}^{\mathbb{P}^*}[M_{N\wedge(t+1)} + \alpha_t^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>t\}}|\mathcal{F}_t]$$

and in particular

$$\mathbb{E}^{\mathbb{P}^*}[M_0] = \mathbb{E}^{\mathbb{P}^*}\left[M_{N\wedge t} + \sum_{i=0}^{t-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>i\}}\right] \quad \forall N, t.$$

Moreover, it holds

$$\lim_{t\to\infty} \mathbb{E}^{\mathbb{P}^*}[M_{N\wedge t} + \sum_{i=0}^{t-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>i\}}] = \mathbb{E}^{\mathbb{P}^*}[M_N] + \mathbb{E}^{\mathbb{P}^*}[\sum_{i=0}^{\infty} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>i\}}].$$

Hence,

$$\mathbb{E}^{\mathbb{P}^*}[M_0] = \mathbb{E}^{\mathbb{P}^*}[M_N] + \mathbb{E}^{\mathbb{P}^*}\left[\sum_{i=0}^{\infty} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>i\}}\right].$$

We set $\sum_{i=0}^{\infty} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{\{N>i\}} =: \sum_{i=0}^{N-1} \alpha_i^{\min}(\mathbb{P}^*)$. Now, let $B \in \mathcal{F}_\sigma$ and define

$$S^B := \sigma\mathbb{I}_B + \kappa\mathbb{I}_{B^C}, \quad T^B := \tau\mathbb{I}_B + \kappa\mathbb{I}_{B^C},$$

where $\kappa := \sup N$. Then S^B and T^B are stopping times. By equation (1):

$$\begin{aligned}
& \mathbb{E}^{\mathbb{P}^*}\left[M_\sigma\mathbb{I}_B + \sum_{i=0}^{\sigma-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_B\right] + \mathbb{E}^{\mathbb{P}^*}\left[M_\kappa\mathbb{I}_{B^c} + \sum_{i=0}^{\kappa-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{B^c}\right] \\
= \; & \mathbb{E}^{\mathbb{P}^*}\left[M_{S^B} + \sum_{i=0}^{S^B-1} \alpha_i^{\min}(\mathbb{P}^*)\right] = \mathbb{E}^{\mathbb{P}^*}[M_0] = \mathbb{E}^{\mathbb{P}^*}\left[M_{T^B} + \sum_{i=0}^{T^B-1} \alpha_i^{\min}(\mathbb{P}^*)\right] \\
= \; & \mathbb{E}^{\mathbb{P}^*}\left[M_\tau\mathbb{I}_B + \sum_{i=0}^{\tau-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_B\right] + \mathbb{E}^{\mathbb{P}^*}\left[M_\kappa\mathbb{I}_{B^c} + \sum_{i=0}^{\kappa-1} \alpha_i^{\min}(\mathbb{P}^*)\mathbb{I}_{B^c}\right],
\end{aligned}$$

and hence

$$\mathbb{E}^{\mathbb{P}^*}[M_\sigma\mathbb{I}_B] = \mathbb{E}^{\mathbb{P}^*}\left[(M_\tau + \sum_{i=\sigma}^{\tau-1} \alpha_i^{\min}(\mathbb{P}^*))\mathbb{I}_B\right].$$

Since this holds true for all $B \in \mathcal{F}_\sigma$, we have

$$\mathbb{E}^{\mathbb{P}^*}[M_\sigma|\mathcal{F}_\sigma] = \mathbb{E}^{\mathbb{P}^*}[M_\tau + \sum_{i=\sigma}^{\tau-1} \alpha_i^{\min}(\mathbb{P}^*)|\mathcal{F}_\sigma],$$

i.e.

$$M_\sigma = \mathbb{E}^{\mathbb{P}^*}\left[M_\tau + \sum_{i=\sigma+1}^{\tau-1} \alpha_i^{\min}(\mathbb{P}^*)\middle| \mathcal{F}_\sigma\right] + \alpha_\sigma^{\min}(\mathbb{P}^*)\mathbb{I}_{\{\tau>\sigma\}}.$$

Summing up, we have shown for $\tau > \sigma$

$$M_\sigma \geq \mathbb{E}^{\mathbb{P}^*}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{P}^*) \geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right)$$

for a variational martingale M; for $\tau = \sigma$

$$M_\sigma = M_\tau = \mathbb{E}^{\mathbb{P}^*}[M_\tau|\mathcal{F}_\sigma] = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right)$$

as α_σ is grounded and $M_\tau \in \mathcal{F}_\sigma$. Hence, for $\tau \geq \sigma$

$$\begin{aligned} M_\sigma &\geq \mathbb{E}^{\mathbb{P}^*}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{P}^*)\mathbb{I}_{\{\tau>\sigma\}} \\ &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right). \end{aligned}$$

For $(S_t)_{t\in\mathbb{N}}$ being a variational supermartingale, the conjecture then follows from Proposition 3.4:

$$\begin{aligned} &\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_\tau - S_\sigma|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right) \\ = &\underbrace{\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[M_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right) - M_\sigma}_{\leq 0} + \underbrace{A_\sigma - A_\tau}_{\leq 0} \leq 0. \end{aligned}$$

Hence,

$$S_\sigma \geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_\tau|\mathcal{F}_\sigma] + \alpha_\sigma^{\min}(\mathbb{Q})\right).$$

□

Proof of Corollary 3.6. From the first part of the proof of Proposition 3.5, we have

$$\begin{aligned} S_{\tau\wedge t} &\geq \mathbb{E}^{\mathbb{P}^*}[S_{\tau\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{P}^*)\mathbb{I}_{\{\tau>t\}} \\ &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_{\tau\wedge(t+1)}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right). \end{aligned}$$

In case $\tau \leq t$ we have

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[S_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = S_\tau + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}) = S_\tau.$$

□

A.4 Proof of Theorems 4.1 & 4.2

The following proof is analog to the respective one in [20].

Proof of Theorem 4.1. ad (a): By definition we have $U_t \geq X_t$, $t \leq T$, and

$$U_t \geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}|\mathcal{F}_t] + \alpha_t(\mathbb{Q})\right)$$

for all $t \leq T-1$. Hence, $(U_t)_{t\leq T}$ is a variational supermartingale with respect to $(\alpha_t^{\min})_{t\leq T}$ exceeding $(X_t)_{t\leq T}$. Let $(Z_t)_{t\leq T}$ be another such variational supermartingale with respect to $(\alpha_t^{\min})_{t\leq T}$. We show $(Z_t)_{t\leq T} \geq (U_t)_{t\leq T}$ inductively: By definition $Z_T \geq X_T = U_T$. Assuming $Z_{t+1} \geq U_{t+1}$, we achieve

$$\begin{aligned} Z_t &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[\underbrace{Z_{t+1}}_{\geq U_{t+1}}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ &\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right). \end{aligned}$$

Thus, as by assumption $Z_t \geq X_t$, we have hence shown (a):

$$Z_t \geq \max\left\{X_t, \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\} = U_t.$$

ad (b): We first show "$\geq$": By Proposition 3.5, we have for the variational supermartingale $(U_t)_{t\leq T} \geq (X_t)_{t\leq T}$ and all $t \leq \tau \leq T$:

$$U_t \geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right).$$

Hence, we have

$$U_t \geq \operatorname*{ess\,sup}_{t\leq\tau\leq T} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = V_t.$$

To show "$\leq$", we define the stopping rule

$$\tau_t^* := \inf\{s \geq t : U_s = X_s\}.$$

Now, fix $t \leq T$. If we can show the stopped variational supermartingale $(U_{s\wedge\tau_t^*})_{t\leq s\leq T}$ to be a variational martingale with respect to $(\alpha_s^{\min})_{t\leq s\leq T}$, we are done: Indeed, in that case we have, as $\tau_t^* \geq t$,

$$\begin{aligned} U_t &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{\tau_t^*}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_t^*}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ &\leq \operatorname*{ess\,sup}_{t\leq\tau\leq T} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = V_t. \end{aligned}$$

Hence, it leaves to show the variational martingale property of the stopped variational Snell envelope $(U_{s\wedge\tau_t^*})_{t\leq s\leq T}$: Let $t \leq s < T$.
(i) Whenever $\tau_t^* \leq s$, we have $U_{(s+1)\wedge\tau_t^*} = U_{\tau_t^*} = U_{s\wedge\tau_t^*}$ and hence

$$\begin{aligned}\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{(s+1)\wedge\tau_t^*}|\mathcal{F}_s] + \alpha_s^{\min}(\mathbb{Q})\right) &= \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{s\wedge\tau_t^*}|\mathcal{F}_s] + \alpha_s^{\min}(\mathbb{Q})\right) \\ &= U_{s\wedge\tau_t^*} + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \alpha_s^{\min}(\mathbb{Q}) = U_{s\wedge\tau_t^*}.\end{aligned}$$

(ii) For $\tau_t^* > s$, we have (by (a) and the definition of τ_t^*) $U_s > X_s$ and hence

$$\begin{aligned}U_{s\wedge\tau_t^*} &= U_s = \max\left\{X_s, \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{s+1}|\mathcal{F}_s] + \alpha_s^{\min}(\mathbb{Q})\right)\right\} \\ &= \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[\underbrace{U_{s+1}}_{=U_{(s+1)\wedge\tau_t^*}}|\mathcal{F}_s] + \alpha_s^{\min}(\mathbb{Q})\right).\end{aligned}$$

(i) and (ii) show the stopped variational martingale property.

ad (c): Let $t = 0$. Then by definition

$$\tau^* = \tau_0^* = \inf\{s \geq 0 : U_s = X_s\}$$

and $(U_{s\wedge\tau^*})_{s\leq T}$ is a variational martingale with respect to $(\alpha_s^{\min})_{s\leq T}$. Hence

$$\begin{aligned}\operatorname*{ess\,sup}_{0\leq\tau\leq T} \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau] + \alpha_0^{\min}(\mathbb{Q})\right) &= V_0 = U_0 \\ &= \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[U_{\tau^*}|\mathcal{F}_0] + \alpha_0^{\min}(\mathbb{Q})\right) \\ &= \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau^*}] + \alpha_0^{\min}(\mathbb{Q})\right).\end{aligned}$$

Hence, τ^* is optimal. Morover, any stopping time such that $\mathbb{P}_0[\tau^{**} < \tau^*] > 0$ cannot be optimal as in that case, by definition of τ^* and part (b),

$$V_0 > \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau^{**}}] + \alpha_0^{\min}(\mathbb{Q})\right).$$

□

Proof of Theorem 4.2. "$\leq$": This inequality is shown in [21] for general minimax-problems.

"$\geq$": By virtue of Proposition 3.2 there exists a $\mathbb{Q}^* \in \mathcal{M}$ such that

$$\begin{aligned}&\operatorname*{ess\,sup}_{T\geq\tau\geq t} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = \;&\operatorname*{ess\,sup}_{T\geq\tau\geq t} \left(\mathbb{E}^{\mathbb{Q}^*}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}^*)\mathbb{I}_{\{\tau>t\}}\right) \\ \geq \;&\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \operatorname*{ess\,sup}_{T\geq\tau\geq t} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\mathbb{I}_{\{\tau>t\}}\right) \\ = \;&\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \operatorname*{ess\,sup}_{T\geq\tau\geq t} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\end{aligned}$$

as $\alpha_t^{\min}$ is grounded, i.e. on $\{\tau = t\}$, we have

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_t|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) = X_t + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \alpha_t^{\min}(\mathbb{Q}) = X_t.$$

□

A.5 Proof of Theorem 4.5

Again, the proof follows the lines of [20].

Lemma A.5. *Let $(\alpha_t^{\min})_{t\in\mathbb{N}}$ be a dynamic minimal penalty satisfying equation (1). For $t \in \mathbb{N}$, the set*

$$\left\{ \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \middle| \tau \geq t \right\}$$

is upward directed, i.e. for any two stopping times τ_1, τ_2, there exists a stopping time, say, $\tau_3 \geq t$ such that

$$\begin{aligned} & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_3}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = \; & \max\left\{ \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right);\ \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \right\}. \end{aligned}$$

Proof. Define $A \in \mathcal{F}_t$ by

$$A := \left\{ \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) > \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \right\}$$

and the stopping time $\tau_3 := \tau_1 \mathbb{I}_A + \tau_2 \mathbb{I}_{A^C}$.

"$\geq$": By Lemma A.3, there exists $\mathbb{Q}_3 \in \mathcal{M}$ such that

$$\begin{aligned} & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_3}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = \; & \mathbb{E}^{\mathbb{Q}_3}[X_{\tau_3}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}_3)\mathbb{I}_{\{\tau_3>t\}} \\ = \; & \mathbb{E}^{\mathbb{Q}_3}[X_{\tau_1}|\mathcal{F}_t]\mathbb{I}_A + \mathbb{E}^{\mathbb{Q}_3}[X_{\tau_2}|\mathcal{F}_t]\mathbb{I}_{A^C} + \alpha_t^{\min}(\mathbb{Q}_3)\mathbb{I}_{\{\tau_3>t\}\cap A} + \alpha_t^{\min}(\mathbb{Q}_3)\mathbb{I}_{\{\tau_3>t\}\cap A^c} \\ = \; & \left(\mathbb{E}^{\mathbb{Q}_3}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}_3)\mathbb{I}_{\{\tau_1>t\}}\right)\mathbb{I}_A + \left(\mathbb{E}^{\mathbb{Q}_3}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}_3)\mathbb{I}_{\{\tau_2>t\}}\right)\mathbb{I}_{A^C} \\ \geq \; & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\}\mathbb{I}_A + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\}\mathbb{I}_{A^C} \\ = \; & \max\left\{ \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\};\ \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\} \right\}, \end{aligned}$$

where the second equality follows from the *local property* of minimal penalty, [11] Lemma 3.3.

"$\leq$": Since

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}}[X_{\tau_3}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q}) \\ = & \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \mathbb{I}_A + \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \mathbb{I}_{A^C}, \end{aligned}$$

we have

$$\begin{aligned} & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_{\tau_3}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\} \\ = & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\} \mathbb{I}_A + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\} \mathbb{I}_{A^C} \\ \leq & \max\left\{\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_{\tau_1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\};\ \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left\{\mathbb{E}^{\mathbb{Q}}[X_{\tau_2}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right\}\right\}. \end{aligned}$$

□

Proof of Theorem 4.5. ad (b): "$\geq$": By Lemma A.5, there exists a sequence $(\tau_k)_k$ of stopping times, such that

$$\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_k}|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q})\right) \nearrow_k V_{t+1}.$$

Hence, making use of time-consistency and continuity from below, we have

$$\begin{aligned} & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = & \lim_{k\to\infty} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}\left[\operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{P}}[X_{\tau_k}|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{P})\right) \middle| \mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = & \lim_{k\to\infty} \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_{\tau_k}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \leq V_t. \end{aligned}$$

Furthermore, by definition of $(V_t)_{t\in\mathbb{N}}$, we have $V_t \geq X_t$ and hence $\forall t \geq 0$

$$V_t \geq \max\left\{X_t, \operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\}.$$

"$\leq$": Given τ, t, set $\sigma := \max\{\tau, t+1\}$. Then, by conditional cash invariance,

$$\begin{aligned} & \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \\ = & X_t \mathbb{I}_{\{\tau\leq t\}} + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[X_\sigma|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right) \mathbb{I}_{\{\tau\geq t+1\}} \\ = & X_t \mathbb{I}_{\{\tau\leq t\}} \\ & + \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}\left[\operatorname*{ess\,inf}_{\mathbb{P}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{P}}[X_\sigma|\mathcal{F}_{t+1}] + \alpha_{t+1}^{\min}(\mathbb{Q})\right) \middle| \mathcal{F}_t\right] + \alpha_t^{\min}(\mathbb{Q})\right) \mathbb{I}_{\{\tau\geq t+1\}} \\ \leq & \max\left\{X_t, \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}} \left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t] + \alpha_t^{\min}(\mathbb{Q})\right)\right\}, \end{aligned}$$

as $\operatorname{ess\,inf}_{\mathbb{P}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{P}}[X_\sigma|\mathcal{F}_{t+1}]+\alpha_{t+1}^{\min}(\mathbb{Q})\right)\leq V_{t+1}$.

This shows "$\leq$" since the above inequality holds for all $\tau\geq t$ and hence for the $\operatorname{ess\,sup}_{\tau\geq t}$. Hence (b) is achieved.

ad (a): By (b) we have for all t

$$V_t\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[V_{t+1}|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\quad\text{and}\quad V_t\geq X_t.$$

Hence, $(V_t)_{t\in\mathbb{N}}$ is a variational supermartingale with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$ and $V_t\geq X_t$. Let $(W_t)_{t\in\mathbb{N}}$ be another variational supermartingale with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$ exceeding $(X_t)_{t\in\mathbb{N}}$. By Proposition 3.5 we have for all $\tau\geq t\in\mathbb{N}$

$$W_t\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[W_\tau|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\geq \operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right),$$

as $W_\tau\geq X_\tau$ and, hence,

$$W_t\geq \operatorname*{ess\,sup}_{\tau\geq t}\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)=V_t.$$

ad (c): As in the proof of Theorem 4.1, we can show $(V_{s\wedge\tau^*})_{s\in\mathbb{N}}$ to be a variational martingale. By our continuity assumption, we hence achieve

$$\begin{aligned}&\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[V_{\tau^*}|\mathcal{F}_0]+\alpha_0^{\min}(\mathbb{Q})\right)\\ =\;&\lim_{s\to\infty}\operatorname*{ess\,inf}_{\mathbb{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[V_{s\wedge\tau^*}|\mathcal{F}_0]+\alpha_0^{\min}(\mathbb{Q})\right)=V_0.\end{aligned}$$

ad (d): Since $(X_t)_{t\in\mathbb{N}}$ is assumed to be bounded, $(U_t^T)_{t\leq T}$ is bounded, too. Furthermore, enlarging the set of stopping times when considering the process up to $T+1$ instead of T, we have $U_t^T\leq U_t^{T+1}$. Hence, the limit $U_t^\infty:=\lim_{T\to\infty}U_t^T$ is well-defined for all t. Hence, by continuity from below

$$\begin{aligned}U_t^\infty &= \lim_{T\to\infty}\overbrace{\max\left\{X_t,\operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}^T|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\right\}}^{U_t^T}\\ &= \max\left\{X_t,\operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[U_{t+1}^\infty|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\right\}.\end{aligned}$$

Thus, $(U_t^\infty)_{t\in\mathbb{N}}$ is a variational supermartingale with respect to $(\alpha_t^{\min})_{t\in\mathbb{N}}$ exceeding $(X_t)_{t\in\mathbb{N}}$. We now show $(V_t)_{t\in\mathbb{N}}=(U_t^\infty)_{t\in\mathbb{N}}$, where $(V_t)_{t\in\mathbb{N}}$ is the infinite horizon problem's value function: By (a) and $(U_t)_{t\in\mathbb{N}}$ being a variational supermartingale exceeding $(X_t)_{t\in\mathbb{N}}$, we have $(U_t^\infty)_{t\in\mathbb{N}}\geq(V_t)_{t\in\mathbb{N}}$. From the finite horizon problem, we have for all T and t

$$\begin{aligned}U_t^T &= \operatorname*{ess\,sup}_{t\leq\tau\leq T}\operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)\\ &\leq \operatorname*{ess\,sup}_{t\leq\tau\leq\infty}\operatorname*{ess\,inf}_{\mathcal{Q}\in\mathcal{M}}\left(\mathbb{E}^{\mathbb{Q}}[X_\tau|\mathcal{F}_t]+\alpha_t^{\min}(\mathbb{Q})\right)=V_t.\end{aligned}$$

Hence, for all t, it holds

$$U_t^\infty = \lim_{T\to\infty} U_t^T \le V_t.$$

This shows $(V_t)_{t\in\mathbb{N}} = (U_t^\infty)_{t\in\mathbb{N}}$ and completes the proof. □

References

[1] P. Artzner, F. Delbaen, J.M. Eber, D. Heath, Coherent Measures of Risk, Math. Finance 9 (1999), 203-228.

[2] P. Artzner, F. Delbaen, J.M. Eber, D. Heath, H. Ku, Coherent Multiperiod Risk Adjusted Values and Bellman's Principle, Ann. Oper. Res. 152 (2007), 5-22.

[3] J. Bion-Nadal, Dynamic Risk Measures: Time Consistency and Risk Measures from BMO Martingales, Finance Stochastics 12 (2008), 219-244.

[4] P. Cheridito, F. Delbaen, M. Kupper, Dynamic Monetary Risk Measures for Bounded Discrete-Time Processes, Electr. J. Prob. 11 (2006), 57-106.

[5] P. Cheridito, L. Tianhui Risk Measures on Orlicz Hearts, Math. Finance 19 (2009), 189-214.

[6] P. Cheridito, M. Stadje, Time-Inconsistency of VaR and Time-Consistent Alternatives, Finan. Res. Letters 1 (2009), 40-46.

[7] J. Csiszar, I-Divergence Geometry of Probability Distributions an Minimization Problems, Ann. Prob. 3 (1975), 146-158.

[8] L. Epstein, M. Schneider, Recursive Multiple Priors, J. Econ. Theory, 113 (2003), 1-13.

[9] H. Föllmer, A. Schied, Convex Measures of Risk and Trading Constraints, Finance Stochastics 6 (2002), 429-447.

[10] H. Föllmer, A. Schied, Stochastic Finance: An Introduction in Discrete Time, 2nd edition, Walter De Gruyter, Berlin, 2004.

[11] H. Föllmer, I. Penner, Convex Risk Measures and the Dynamics of their Penalty Functions, Statistics Decisions 24 (2006), 61-96, 2006.

[12] H. Föllmer, A. Schied, S. Weber, Stefan, Robust Preferences and Robust Portfolio Choice, in: P. Ciarlet, A. Bensoussan, Q. Zhang (eds.), Handbook of Numerical Analysis: Mathematical Modelling and Numerical Methods in Finance 15 (2009), 29-89.

[13] I. Gilboa, D. Schmeidler, Maximin Expected Utility with non-unique Prior, J. Math. Econ. 18 (1989), 141-153.

[14] F. Maccheroni, M. Marinacci, A. Rustichini, Ambiguity Aversion, Robustness, and the Variational Representation of Preferences, Econometrica 74 (2006), 1447-1498.

[15] F. Maccheroni, M. Marinacci, A. Rustichini, Dynamic Variational Preferences, J. Econ. Theory 128 (2006), 4-44.

[16] A. McNeil, R. Frey, P. Embrechts, Quantitative Risk Management : Concepts, Techniques, and Tools, Princeton University Press, Princeton, NJ, 2005.

[17] J. Neveu, Discrete-Parameter Martingales, North-Holland, Oxford, 1975.

[18] S. Peng, Backward SDE and related g-expectations, in: N. El Karoui, L. Mazliak, L. (eds.), Backward Stochastic Differential Equations, Pitman Res. Notes Math. Ser. 364 (1997), 141-159.

[19] F. Riedel, Dynamic Coherent Risk Measures, Stoch. Proc. Appl. 112 (2004), 185-200.

[20] F. Riedel, Optimal Stopping with Multiple Priors, Econometrica, forthcoming, 2009.

[21] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, NJ, 1970.

[22] E. Roszza Gianin, Risk Measures via g-Expectations, Ins.: Mathematics Econ. 39 (2006), 19-34.

[23] A. Schied, Optimal Investment for Risk- and Ambiguity-Averse Preferences: A Duality Approach, Finance Stochastics 11 (2007), 107-129.