The Annals of Applied Probability
2019, Vol. 29, No. 1, 577–612
https://doi.org/10.1214/18-AAP1428

EXPONENTIAL UTILITY MAXIMIZATION UNDER MODEL UNCERTAINTY FOR UNBOUNDED ENDOWMENTS

BY DANIEL BARTL[1]

University of Konstanz

We consider the robust exponential utility maximization problem in discrete time: An investor maximizes the worst case expected exponential utility with respect to a family of nondominated probabilistic models of her endowment by dynamically investing in a financial market, and statically in available options.

We show that, for any measurable random endowment (regardless of whether the problem is finite or not) an optimal strategy exists, a dual representation in terms of (calibrated) martingale measures holds true, and that the problem satisfies the dynamic programming principle (in case of no options). Further, it is shown that the value of the utility maximization problem converges to the robust superhedging price as the risk aversion parameter gets large, and examples of nondominated probabilistic models are discussed.

1. Introduction. In this article, we study the problem of robust exponential utility maximization in discrete time. Here, the term robust reflects uncertainty about the true probabilistic model and the consideration of a whole family of models as a consequence. This is not a new concept and since the seminal papers [23] and [31] it has gained a lot of attention; see, for example, [1, 5, 6, 10, 11, 13, 18, 25, 35, 38, 41] and [14] for an overview. To state our problem more precisely, given the exponential utility function

$$U(x) := -\exp(-\gamma x)$$

with risk-aversion parameter $\gamma > 0$, a possibly nondominated set of probabilistic models $\mathcal{P}$ and the agent's random endowment X, we are interested in the optimization problem

$$\sup_{(\vartheta,\alpha)\in\Theta\times\mathbb{R}^e} \inf_{P\in\mathcal{P}} E_P\big[U\big(X + (\vartheta\cdot S)_T + \alpha(g - g_0)\big)\big]. \tag{1.1}$$

Here, $g^1, \dots, g^e$ are traded options available for buying and selling at time 0 for the prices $g_0^1, \dots, g_0^e$, the set Θ consists of all predictable dynamic trading strategies for the (discounted) stock S, and $(\vartheta\cdot S)_T + \alpha(g - g_0)$ is the outcome of a semistatic trading strategy $(\vartheta, \alpha) \in \Theta \times \mathbb{R}^e$.

Received June 2017; revised January 2018.

[1]Supported by Austrian Science Fund (FWF) Grant Y00782.

MSC2010 subject classifications. Primary 91B16, 49L20; secondary 60G42.

Key words and phrases. Utility maximization, robust finance, duality, dynamic programming.

The first immediate question when investigating the optimization problem (1.1) is whether an optimal strategy (ϑ, α) (which should be defined simultaneously under all models $P \in \mathcal{P}$) exits. Due to the absence of a measure capturing all zero sets and the failure of classic arguments such as Komlos' theorem as a consequence, this is nontrivial. Our second interest lies in the validity of a dual representation with respect to linear pricing measures, namely if (1.1) is equal to

$$-\exp\Big(-\inf_{Q\in\mathcal{M}}\big(\gamma E_Q[X] + H(Q,\mathcal{P})\big)\Big),$$

where $\mathcal{M}$ denotes the set of all martingale measures Q for the stock S calibrated to the options (i.e., $E_Q[g^i] = g_0^i$ for $1 \le i \le e$) under which the robust entropy $H(Q,\mathcal{P})$ is finite. Finally, we study if (1.1) satisfies the dynamic programming principle (in case without options), meaning that it is possible to analyze the problem locally and later "glue" everything together. In particular, this implies that a strategy which is optimal at time 0, will be optimal again, if one starts to solve the optimization problem at some positive time t.

The main contribution of this paper is to show that positive answers to all three questions, namely the existence of an optimal strategy, duality and dynamic programming can be given under weak assumptions; see Theorem 2.1 and Theorem 2.2. Further, it is shown that a scaled version of (1.1) converges to the minimal superhedging price of X if the risk-aversion parameter γ tends to infinity; see Theorem 2.4. In fact, we adopt the setting suggested by Bouchard and Nutz in the milestone paper [10] and show by means of optimal control, that for any unbounded measurable (lower semianalytic) random endowment X (regardless of whether the optimization problem (1.1) is finite or not), existence, duality and the dynamic programming principle hold true.

Needless to say, utility maximization is an important topic in mathematical finance starting with [28, 34]. In case of exponential utility function (though in a continuous-time and nonrobust setting), [22] and [15] were the first to prove duality and existence, which lead to further analysis, for example, a BSDE characterization of the optimal value and solution in an incomplete market and under trading constraints is given in [26], and the dynamics and asymptotics in the risk-aversion parameter γ are studied in [32]. In the presence of uncertainty, starting with [42] and [43], most results are obtained under the assumption that $\mathcal{P}$ is dominated; see, for example, [2, 24, 40]. The literature focusing on a nondominated set $\mathcal{P}$ is still comprehensible and in continuous-time results are given in [18, 33, 35].

In the present setting (i.e., discrete-time and a nondominated set $\mathcal{P}$), the dynamic programming principle and the existence of an optimal strategy are first shown in [38], where the author considers a random utility function U defined on $\Omega \times \mathbb{R}_+$ satisfying a certain boundedness (which would correspond to a random endowment that is bounded from below in our setting). More recently, there are three papers generalizing the result of [38]. In [9], the boundedness of the random utility (still defined on the positive real line) is replaced by a certain integrability

condition and dynamic programming as well as the existence of an optimal strategy is shown. In [36, 37], the random utility function (which may be nonconcave in the second work) is no longer defined on the positive real line, but satisfies certain boundedness similar to [38]. Moreover, the market is more general and includes, for example, trading constraints or proportional transaction cost. Convergence of the utility indifference prices (to the superheding price) is shown in [8]. Duality on the other hand is shown in Section 4.2 of [14] under a compactness condition on the set $\mathcal{P}$ and (semi)continuity of the random endowment X.

In order to lighten notation, we will assume without loss of generality that the prices of the traded options are 0 and, instead of (1.1), consider the equivalent problem

$$\inf_{(\vartheta,\alpha)\in\Theta\times\mathbb{R}^e} \sup_{P\in\mathcal{P}} \log E_P\big[\exp\big(X + (\vartheta\cdot S)_T + \alpha g\big)\big].$$

It is clear that both problems are one to one, except all results for an endowment X in the original problem hold for $-X$ in the transformed one, and vice versa.

The remainder of this paper is organized as follows: Section 2 contains the setting, all main results, a discussion of the assumptions and some examples. Sections 3 and Section 4 are devoted to the proofs for the one period and general case, respectively. Finally, technical proofs are given in Appendix A and a brief introduction to the theory of analytic sets is given in Appendix B.

2. Main results.

2.1. *Setting.* Up to a change regarding the no-arbitrage condition (discussed in Remark 2.5), we work in the setting proposed by Bouchard and Nutz [10], which is briefly summarized below. Analytic sets and the general terminology are shortly discussed in Appendix B. Let Ω_0 be a singleton and Ω_1 be a Polish space. Fix $d, T \in \mathbb{N}$, let $\Omega_t := \Omega_1^t$, and define $\mathcal{F}_t$ to be the universal completion of the Borel σ-field on Ω_t for each $0 \le t \le T$. To simplify notation, we denote $(\Omega, \mathcal{F}) = (\Omega_T, \mathcal{F}_T)$ and often consider Ω_t as a subset of Ω. For $s < t$, some fixed $\omega \in \Omega_s$, and a function X with domain Ω_t, we consider $X(\omega, \cdot)$ as a function with domain Ω_{t-s}, that is, $\Omega_{t-s} \ni \omega' \mapsto X(\omega, \omega')$. For each $0 \le t \le T-1$ and $\omega \in \Omega_t$, there is a given convex and nonempty set of probabilities $\mathcal{P}_t(\omega) \subset \mathfrak{P}(\Omega_1)$, which can be seen as all possible probability scenarios for the price of the stock at time $t+1$, given the history ω. The assumption throughout is that the stock $S_t : \Omega_t \to \mathbb{R}^d$ is Borel and that the set-valued mapping $\mathcal{P}_t$ has analytic graph. The latter in particular ensures that

$$\mathcal{P} := \big\{P = P_0 \otimes \cdots \otimes P_{T-1} : P_t(\cdot) \in \mathcal{P}_t(\cdot)\big\} \tag{2.1}$$

is not empty. Here, each P_t is a selector of $\mathcal{P}_t$, that is, a universally measurable function $P_t : \Omega_t \to \mathfrak{P}(\Omega_1)$ satisfying $P_t(\omega) \in \mathcal{P}_t(\omega)$ for each ω, and the probability P on Ω is defined by

$$P(A) := \int_{\Omega_1} \cdots \int_{\Omega_1} 1_A(\omega_1, \ldots, \omega_T) P_{T-1}(\omega_1, \ldots, \omega_{T-1}, d\omega_T) \cdots P_0(d\omega_1).$$

The set of all dynamic trading strategies is denoted by Θ and an element $\vartheta \in \Theta$ is a vector $\vartheta = (\vartheta_1, \dots, \vartheta_T)$ consisting of $\mathcal{F}_{t-1}$-measurable mappings $\vartheta_t \colon \Omega_{t-1} \to \mathbb{R}^d$. The outcome at time t of trading according to the dynamic strategy ϑ starting at time $s \le t$ is given by

$$(\vartheta \cdot S)_s^t := \vartheta_{s+1}\Delta S_{s+1} + \cdots + \vartheta_t \Delta S_t \qquad \text{where } \Delta S_u := S_u - S_{u-1}$$

and $\vartheta_u \Delta S_u := \sum_{i=1}^d \vartheta_u^i \Delta S_u^i$ is the inner product. As $\mathcal{P}$ has a dynamic form, one can consider both a local and a global no arbitrage condition: The global NA($\mathcal{P}$) condition is satisfied if $(\vartheta \cdot S)_0^T \ge 0$ $\mathcal{P}$-q.s. implies $(\vartheta \cdot S)_0^T = 0$ $\mathcal{P}$-q.s. for every $\vartheta \in \Theta$, and the local NA($\mathcal{P}_t(\omega)$) condition (for fixed $0 \le t \le T-1$ and $\omega \in \Omega$) is satisfied if $h\Delta S_{t+1}(\omega, \cdot) \ge 0$ $\mathcal{P}_t(\omega)$-q.s. implies $h\Delta S_{t+1}(\omega, \cdot) = 0$ $\mathcal{P}_t(\omega)$-q.s. for every $h \in \mathbb{R}^d$. Throughout this article we assume that

$$\mathrm{NA}\big(\mathcal{P}_t(\omega)\big) \qquad \text{holds for every } 0 \le t \le T-1 \text{ and } \omega \in \Omega_t. \tag{2.2}$$

Note that this assumption is purely technical as NA($\mathcal{P}$) holds true if and only if the set of all ω such that NA($\mathcal{P}_t(\omega)$) fails for some t is a zero set under all $P \in \mathcal{P}$, see [10], Theorem 4.5, and Remark 2.5 for a discussion. Finally, define

$$\mathcal{M} = \big\{Q \in \mathfrak{P}(\Omega) : S \text{ is a martingale under } Q \text{ and } H(Q, \mathcal{P}) < +\infty\big\},$$

to be the set of martingale measures with finite robust entropy

$$H(Q, \mathcal{P}) := \inf_{P \in \mathcal{P}} H(Q, P)$$
$$\text{where } H(Q, P) := \begin{cases} E_P\left[\dfrac{dQ}{dP} \log \dfrac{dQ}{dP}\right] & \text{if } Q \ll P, \\ +\infty & \text{else.} \end{cases}$$

Throughout, the convention $E_P[X] := E_P[X^+] - E_P[X^-]$ with $E_P[X] := -\infty$ if $E_P[X^-] = +\infty$ is in force; in particular X is integrable with respect to P if and only if $E_P[X] \in \mathbb{R}$.

2.2. *Main results.*

THEOREM 2.1 (Without options). *Let $X \colon \Omega \to (-\infty, +\infty]$ be upper semi-analytic. Then*

$$\inf_{\vartheta \in \Theta} \sup_{P \in \mathcal{P}} \log E_P\big[\exp\big(X + (\vartheta \cdot S)_0^T\big)\big] = \sup_{Q \in \mathcal{M}} \big(E_Q[X] - H(Q, \mathcal{P})\big) \tag{2.3}$$

and both terms are not equal to $-\infty$. Moreover, the infimum over $\vartheta \in \Theta$ is attained and the optimization problem satisfies the dynamic programming principle; see Theorem 4.1 for the precise formulation of the last statement.

In addition to the previous setting, assume that there are $e \in \mathbb{N} \cup \{0\}$ options ($e = 0$ corresponding to the case without options), that is, Borel functions $g^1, \dots, g^e \colon \Omega \to \mathbb{R}$, available at time $t = 0$ for price zero. The outcome of a semistatic trading strategy $(\vartheta, \alpha) \in \Theta \times \mathbb{R}^e$ equals $(\vartheta \cdot S)_0^T + \alpha g$, where $\alpha g := \sum_{i=1}^e \alpha_i g^i$ again denotes the inner product. In addition to the already imposed no arbitrage condition, assume that $(\vartheta \cdot S)_0^T + \alpha g \geq 0$ $\mathcal{P}$-q.s. implies $(\vartheta \cdot S)_0^T + \alpha g = 0$ $\mathcal{P}$-q.s. for every strategy $(\vartheta, \alpha) \in \Theta \times \mathbb{R}^e$.

THEOREM 2.2 (With options). *Fix a Borel function $Z \colon \Omega \to [0, +\infty)$ such that $|g^i| \leq Z$ for $1 \leq i \leq e$ and let $X \colon \Omega \to \mathbb{R}$ be an upper semianalytic function satisfying $|X| \leq Z$. Then it holds*

$$\inf_{(\vartheta,\alpha)\in\Theta\times\mathbb{R}^e} \sup_{P\in\mathcal{P}} \log E_P\big[\exp\big(X + (\vartheta \cdot S)_0^T + \alpha g\big)\big] = \sup_{Q\in\mathcal{M}_g} \big(E_Q[X] - H(Q, \mathcal{P})\big),$$

where $\mathcal{M}_g$ denotes the set of all $Q \in \mathcal{M}$ with $E_Q[Z] < +\infty$ and $E_Q[g^i] = 0$ for $1 \leq i \leq e$. Moreover, the infimum over $(\vartheta, \alpha) \in \Theta \times \mathbb{R}^e$ is attained.

REMARK 2.3. 1. The no-arbitrage condition NA($\mathcal{P}$) is essential. Indeed, even if both sides in (2.3) do not take the value $-\infty$, the condition NA($\mathcal{P}$) does not need to hold—nor does equation (2.3); see Appendix A.

2. If X is allowed to take the value $-\infty$ in Theorem 2.1, then neither duality nor the existence of an optimal strategy hold true; see Appendix A.

3. In general, due the supremum over $P \in \mathcal{P}$, the minimizer ϑ in (2.3) is not unique and the supremum over Q is not attained.

4. In Theorem 2.1, the set $\mathcal{M}$ can be replaced by $\mathcal{M}(Y) := \{Q \in \mathcal{M} : E_Q[Y] < +\infty\}$, where $Y \colon \Omega \to [0, +\infty)$ is an arbitrary function such that $-Y$ is upper semianalytic. The same holds true for Theorem 2.2, that is, one can replace $\mathcal{M}_g$ by $\mathcal{M}_g(Y) := \{Q \in \mathcal{M}_g : E_Q[Y] < +\infty\}$.

Another interesting problem is the study of asymptotic behavior of the optimization problem in the risk-aversion parameter γ; see, for example, [32]. Let us give some motivation: Typically the superhedging price

$$\pi(X) := \inf\big\{m \in \mathbb{R} : m + (\vartheta \cdot S)_0^T + \alpha g \geq X \mathcal{P}\text{-q.s. for some } (\vartheta, \alpha) \in \Theta \times \mathbb{R}^e\big\}$$

is extremely high. A natural way of shrinking π is to allow $m + (\vartheta \cdot S)_0^T + ug < X$ with positive probability in a "controlled" way; see, for example, [14, 20]. More precisely, define

$$\pi_\gamma(X) := \inf\left\{m : \begin{array}{l} \sup_{P\in\mathcal{P}} \dfrac{1}{\gamma} \log E_P\big[\exp\big(\gamma\big(X - m - (\vartheta \cdot S)_0^T - \alpha g\big)\big)\big] \leq 0 \\ \text{for some } (\vartheta, \alpha) \in \Theta \times \mathbb{R}^e \end{array}\right\}$$

for each risk-aversion parameter $\gamma > 0$. Then $\pi_\gamma(X) \leq \pi(X)$ by definition and since $\exp(\gamma x)/\gamma \to +\infty 1_{(0,+\infty]}(x)$ as $\gamma \to +\infty$, an evident question is whether the same holds true for the superhedging prices.

THEOREM 2.4 (Entropic hedging). *In the setting of Theorem 2.2, it holds*

$$\pi(X) = \lim_{\gamma \to +\infty} \pi_\gamma(X)$$

and the limit in γ is a supremum over $\gamma > 0$.

REMARK 2.5. The reason why we require $\mathrm{NA}(\mathcal{P}_t(\omega))$ to hold for every ω and not only for $\mathcal{P}$-quasi every ω as in Bouchard and Nutz [10], is the following. In order to apply the one-period results (i.e., duality and existence of an optimal strategy) to the local problem $\mathcal{E}_t(\omega, x)$ (see (4.1) for the precise definition), one needs that $\mathcal{E}_{t+1}(\omega \otimes_t \omega', x) > -\infty$ for $\mathcal{P}_t(\omega)$-quasi all $\omega' \in \Omega_1$, see point 2) in Remark 2.3. However, to ensure the latter, $\mathrm{NA}(\mathcal{P}_{t+1}(\omega \otimes_t \omega'))$ needs to holds for $\mathcal{P}_t(\omega)$-quasi all $\omega' \in \Omega_1$. Due to the fact that the set $N_{t+1} := \{\tilde{\omega} \in \Omega_{t+1} : \mathrm{NA}(\mathcal{P}_t(\tilde{\omega}))$ fails$\}$ is merely universally measurable, it is not clear that this condition holds true for "sufficiently many" $\omega \in \Omega_t$.

In case of only one measure (i.e., $\mathcal{P} = \{P = P_0 \otimes \cdots \otimes P_{T-1}\}$ is a singleton), this problem has an easy solution: Since $P(N_t) = 0$ for every t, one can redefine P_t by $\tilde{P}_t := P_t 1_{N_t^c} + \delta_{S_t} 1_{N_t}$. Then $P = \tilde{P}_0 \otimes \cdots \otimes \tilde{P}_{T-1}$ and $\tilde{\mathcal{P}}_t(\omega) := \{\tilde{P}_t(\omega)\}$ has analytic graph (a proof is given in Appendix A). The same can be done in the setting of general $\mathcal{P}$, as long as one requires the sets N_t to be Borel; otherwise the graph of $\tilde{\mathcal{P}}_t := \mathcal{P}_t 1_{N_t^c} + \{\delta_{S_t}\} 1_{N_t}$ needs not to be analytic.

Finally, note that $\mathcal{P}$ is defined though the sets $\mathcal{P}_t$ (and not the other way around), which means that the assumption for $\mathrm{NA}(\mathcal{P}_t(\omega))$ to hold for every ω does not seem restrictive regarding applications; see also Section 2.3.

Recently, using a different approach, [17] were able to get rid of this assumption.

REMARK 2.6. The (technical but crucial) assumption that the graph of $\mathcal{P}_t$ is analytic has two consequences: It allows for measurable selection arguments and enables to define pointwise conditional sublinear expectations, that is, ensure that

$$\mathcal{E}(X|\Omega_t)(\omega) := \sup_{P \in \mathcal{P}_t(\omega)} E_P\big[X(\omega, \cdot)\big] \qquad \text{for } \omega \in \Omega_t \text{ and } X\colon \Omega_{t+1} \to \mathbb{R} \tag{2.4}$$

is upper semianalytic as a mapping of ω whenever X is [10, 39]. The converse holds true as well: Given an arbitrary sublinear conditional expectation $\mathcal{E}(\cdot|\Omega_t)$ (satisfying some continuity), there always exists a set-valued mapping $\mathcal{P}_t$ with analytic graph such that (2.4) holds true [3], Theorem 1.1. Similarly, the "time-consistency" (2.1) of $\mathcal{P}$ is equivalent to the tower-property [3], Theorem 1.2.

2.3. *Examples*. In this section, we discuss a general method of robustifying a given probability and also give applications to financial models. All nontrivial claims are proven at the beginning of Appendix A.

In many cases, the physical measure is not known a priori, but rather a result of collecting data and estimation. In particular, the estimator is not equal to, but only

"converges" (as the data grow richer) to the actual unknown physical measure. A canonical way of taking this into account therefore consists of adding some sort of "neighborhood" to the estimator $P^* = P_0^* \otimes \cdots \otimes P_{T-1}^*$, that is, to define

$$\mathcal{P}_t(\omega) := \{P \in \mathfrak{P}(\Omega_1) : \operatorname{dist}(P, P_t^*(\omega)) \le \varepsilon_t(\omega)\}. \tag{2.5}$$

Here, as the name suggests,

$$\operatorname{dist}\colon \Omega_1 \times \Omega_1 \to [0, +\infty]$$

can be thought of a distance and $\varepsilon_t\colon \Omega_t \to [0, +\infty]$ as the size of the neighborhood. If dist, ε_t (and P_t^*) are Borel—from now on a standing assumption—then $\mathcal{P}_t$ has analytic graph. If dist is in fact a metric or at least fulfills $\operatorname{dist}(P, P) = 0$, the values of $\mathcal{P}_t$ are also nonempty. Since the distance should be compatible with estimation, natural choices include the Wasserstein distances of order p or, more generally, the cost of transportation, that is,

$$\operatorname{dist}(Q, P) := \inf\left\{\int_{\Omega_1 \times \Omega_1} c(x, y)\Pi(dx, dy) : \Pi\right\},$$

where the infimum is taken over all measures on the product $\Pi \in \mathfrak{P}(\Omega_1 \times \Omega_1)$ with $\Pi(\cdot \times \Omega_1) = Q$ and $\Pi(\Omega_1 \times \cdot) = P$, and $c\colon \Omega_1 \times \Omega_1 \to [0, +\infty]$ is a given lower semicontinuous function (the "cost"). This includes the Wasserstein distance of order p; then the cost c equals a metric on Ω_1 to the power p; see, for example, Chapters 5 and 6 in [45]. This traceable distance has many advantages, for example, that besides metrizing weak convergence, it controls the integrability of the tails. In this case, $\mathcal{P}_t$ has convex values. Moreover, [4] provide a finite dimensional formula how to compute the worst case expectation over all probabilities in a Wasserstein neighborhood of a baseline distribution.

The above method can also be applied when a certain model for the dynamics of the underlying is fixed and only the parameters are uncertain. For simplicity, assume that $\Omega = \mathbb{R}^T$, $S_t(\omega) = \omega_t$ is the canonical space of a one-dimensional stock. We illustrate in two concrete examples: the binomial model, which reads as

$$S_{t+1}(\omega, \cdot) = S_t(\omega) + B(\cdot) \tag{2.6}$$

for every t and $\omega \in \Omega_t$ where $B\colon \Omega_1 \to \mathbb{R}$ is binomially distributed, and a discrete version of the Black–Scholes model, which reads as

$$S_{t+1}(\omega, \cdot) = S_t(\omega)\big(\mu \Delta t + \sigma \Delta W(\cdot)\big), \tag{2.7}$$

where $\mu \in \mathbb{R}$, $\sigma, \Delta t > 0$, and $\Delta W\colon \Omega_1 \to \mathbb{R}$ is normally distributed with mean 0 and variance Δt; we write $\Delta W \sim N(0, \Delta t)$. Defining $f_t(\omega, x) := S_t(\omega) + x$, $X := B$ in case of the binomial, and $f_t(\omega, x) := S_t(\omega)x$, $X := \mu \Delta t + \sigma \Delta W$ in case of the Black–Scholes model, it follows that both can be written in the more general form

$$S_{t+1}(\omega, \cdot) = f_t\big(\omega, X(\cdot)\big), \tag{2.8}$$

where $f_t\colon \Omega_t \times \mathbb{R} \to \mathbb{R}$ and $X\colon \Omega_1 \to \mathbb{R}$ are Borel. In terms of distributions, (2.8) means nothing but

$$\operatorname{law} S_{t+1}(\omega, \cdot) = R \circ f(\omega, \cdot)^{-1} \qquad \text{where } R := \operatorname{law} X.$$

Therefore, a canonical way of robustifying a given model of the form (2.8) is to replace R in the equation above by a set $\mathcal{R}_t(\omega) \subset \mathfrak{P}(\Omega_1)$, and to define

$$\hat{\mathcal{P}}_t(\omega) := \{\operatorname{law} S_{t+1}(\omega, \cdot) = R \circ f(\omega, \cdot)^{-1} : R \in \mathcal{R}_t(\omega)\}.$$

For example, in line with the first part of this section, one can take some neighborhood

$$\mathcal{R}_t(\omega) = \{R \in \mathfrak{P}(\Omega_1) : \operatorname{dist}(R, \operatorname{law} X) \le \varepsilon_t(\omega)\}, \tag{2.9}$$

or, if there are even less data, one might argue that

$$\mathcal{R}_t(\omega) := \{R \in \mathfrak{P}(\Omega_1) : E_R[\phi_t^i(\omega, \cdot)] \le 1 \text{ for } 1 \le i \le n\} \tag{2.10}$$

for some given Borel functions $\phi_t^1, \dots, \phi_t^n\colon \Omega_t \times \mathbb{R}^d \to \mathbb{R}$ is a good choice. Here, if $\inf_x \phi_t^i(\omega, x) \le 0$ for all i and f_t in (2.8) is such that $S_t(\omega)$ lies in the relative interior of $f_t(\omega, \mathbb{R}^d)$ for every $\omega \in \Omega_t$—an assumption which is usually fulfilled—then the resulting model of (2.10) satisfies $\mathrm{NA}(\mathcal{P}_t(\omega))$ for every t and ω. The same holds true for $\mathcal{R}_t$ defined by (2.9) under the mentioned assumption on f_t if, for example, dist is the Wasserstein distance of order p and X has a finite pth moment.

On a technical level, $\mathcal{R}_t$ defined by (2.10) has analytic graph and so do $\hat{\mathcal{P}}_t$ and $\mathcal{P}_t$, the latter begin defined as $\mathcal{P}_t(\omega) := \operatorname{conv} \hat{\mathcal{P}}_t(\omega)$ the convex hull of $\hat{\mathcal{P}}_t$. The same holds true for $\mathcal{R}_t$ defined by (2.9).

EXAMPLE 2.7 (Binomial model). Besides what was mentioned above, another natural generalization of the Binomal model is to allow for the jump size and probability to take values in some intervals (which may depend on the time t and past $\omega \in \Omega_t$). This corresponds to

$$\mathcal{R}_t(\omega) := \left\{ p\delta_a + (1-p)\delta_b : \begin{array}{l} p \in [\underline{p}_t(\omega), \overline{p}_t(\omega)] \text{ and} \\ a \in [\underline{a}_t(\omega), \overline{a}_t(\omega)], b \in [\underline{b}_t(\omega), \overline{b}_t(\omega)] \end{array} \right\},$$

where $0 < \underline{p}_t \le \overline{p}_t < 1$, $\underline{a}_t \le \overline{a}_t < 0 < \underline{b}_t \le \overline{b}_t$ are Borel functions. Here, δ_a denotes the Dirac measure at point a. Note that $\mathrm{NA}(\mathcal{P}_t(\omega))$ is trivially satisfied for every t and ω.

Regarding the Black–Scholes model in continuous time, there is a popular and well-studied way of robustification (see, e.g., [41]): Consider all models (2.7) with μ and volatility σ in some given intervals. This can be done as in the previous example, however, then each $\mathcal{P}_t(\omega)$ and, therefore, also the resulting family $\mathcal{P}$ is dominated (by the Lebesgue measure). In the present discrete-time setting, it seems more interesting to discard the assumption of normality of ΔW in (2.7).

EXAMPLE 2.8 (Black–Scholes). Fix two Borel functions $\mu_t\colon \Omega_1 \to \mathbb{R}$ and $\sigma_t\colon \Omega_1 \to (0, +\infty)$, and let ε_t and dist be as above. Now define

$$\mathcal{R}_t(\omega) := \{R * \delta_{\mu\Delta t} : \mu \in [\underline{\mu}_t(\omega), \overline{\mu}_t(\omega)] \text{ and } \operatorname{dist}(R, N(0, \sigma_t^2(\omega)\Delta t)) \le \varepsilon_t(\omega)\},$$

where $R * \delta_{\mu\Delta t}$ denotes the convolution $R * \delta_{\mu\Delta t}(A) := R(A - \mu\Delta t)$. The set $\hat{\mathcal{P}}_t$ therefore corresponds to the Black–Scholes model with drift and volatility uncertainty in the sense that one considers all models

$$S_{t+1}(\omega, \cdot) = S_t(\omega)(\mu\Delta t + Y), \qquad \begin{array}{l}\mu \in [\underline{\mu}_t(\omega), \overline{\mu}_t(\omega)] \text{ and the law of} \\ Y \text{ is } \varepsilon_t(\omega) \text{ close to } N(0, \sigma_t^2(\omega)\Delta t)\end{array}$$

simultaneously. To be more in line with the original model, one can also require that R (resp., Y) has mean 0 in the definition of $\mathcal{R}_t$. Note that for any reasonable choice for the distance (e.g., Wasserstein), the set $\mathcal{P}_t(\omega)$ satisfies all of our assumptions.

3. Proof for the one-period setting. Let $(\Omega, \mathcal{F})$ be a measurable space armed with a family of probability measures $\mathcal{P} \subset \mathfrak{P}(\Omega)$. Further, let $S_0 \in \mathbb{R}$ and $S_1\colon \Omega \to \mathbb{R}$ be measurable. We write $h \in \Theta = \mathbb{R}^d$ for trading strategies and assume the no-arbitrage NA($\mathcal{P}$), that is, $h\Delta S \ge 0$ $\mathcal{P}$-q.s. implies $h\Delta S = 0$ $\mathcal{P}$-q.s. for every $h \in \mathbb{R}^d$. Given some random variable $Z\colon \Omega \to [0, +\infty)$, denote by

$$\mathcal{M}(Z) = \{Q \in \mathfrak{P}(\Omega) : E_Q[|\Delta S| + Z] + H(Q, \mathcal{P}) < +\infty \text{ and } E_Q[\Delta S] = 0\}$$

the set of martingale measures that have finite entropy and integrate Z. The following is the main result of this section.

THEOREM 3.1. *Fix a random variable $X\colon \Omega \to (-\infty, +\infty]$. Then one has*

$$\inf_{h\in\mathbb{R}^d} \sup_{P\in\mathcal{P}} \log E_P[\exp(X + h\Delta S)] = \sup_{Q\in\mathcal{M}(Y)} (E_Q[X] - H(Q, \mathcal{P})) \tag{3.1}$$

for every random variable $Y\colon \Omega \to [0, +\infty)$ and both terms are not equal to $-\infty$. Moreover, the infimum over $h \in \mathbb{R}^d$ is attained.

The following lemma, which turns out to be be helpful in the multiperiod case, is shown in the course of the proof of Theorem 3.1.

LEMMA 3.2. *Let $X_n\colon \Omega \to (-\infty, +\infty]$ be a sequence of random variables increasing point-wise to X. Then it holds*

$$\sup_n \inf_{h\in\mathbb{R}^d} \sup_{P\in\mathcal{P}} \log E_P[\exp(X_n + h\Delta S)] = \inf_{h\in\mathbb{R}^d} \sup_{P\in\mathcal{P}} \log E_P[\exp(X + h\Delta S)],$$

that is, the optimization problem is continuous from below.

LEMMA 3.3. *Fix a random variable $X\colon \Omega \to \mathbb{R}$. Then one has*

$$\sup_{P\in\mathcal{P}} \log E_P\big[\exp(X + h\Delta S)\big] = \sup_{Q\in\mathcal{C}}\big(E_Q[X + h\Delta S] - H(Q,\mathcal{P})\big)$$

for every $h \in \mathbb{R}^d$, where

$$\mathcal{C} := \big\{Q \in \mathfrak{P}(\Omega) : E_Q\big[|X| + |\Delta S| + Y\big] + H(Q,\mathcal{P}) < +\infty\big\}$$

and $Y\colon \Omega \to [0,+\infty)$ is an arbitrary random variable.

PROOF. (a) Define $Z := X + h\Delta S$ and fix a measure $P \in \mathcal{P}$. It follows from the well-known representation of expected exponential utility and the monotone convergence theorem that

$$\log E_P\big[\exp(Z)\big] = \sup_{Q\in\mathcal{A}_P}\big(E_Q[Z] - H(Q,P)\big), \tag{3.2}$$

where

$$\mathcal{A}_P := \big\{Q \in \mathfrak{P}(\Omega) : E_Q\big[Z^-\big] + H(Q,P) < +\infty\big\}.$$

For the sake of completeness, a proof is provided in Lemma A.1. We claim that one can replace $\mathcal{A}_P$ with $\mathcal{C}_P$ in (3.2) without changing the value of the supremum, where

$$\mathcal{C}_P := \big\{Q \in \mathfrak{P}(\Omega) : E_Q\big[|X| + |\Delta S| + Y\big] + H(Q,P) < +\infty\big\}.$$

Since $\mathcal{C}_P$ is a subset of $\mathcal{A}_P$, it suffices to show that for any $Q \in \mathcal{A}_P$, there exists a sequence $Q_n \in \mathcal{C}_P$ such that $E_{Q_n}[Z] - H(Q_n,P)$ converges to $E_Q[Z] - H(Q,P)$. To that end, fix some $Q \in \mathcal{A}_P$ and define

$$Q_n := Q(\cdot|B_n) \qquad \text{where } B_n := \big\{|X| + |\Delta S| + Y \le n\big\}$$

for all n large enough such that $Q(B_n) > 0$. Then it holds

$$\frac{dQ_n}{dP} = \frac{1_{B_n}}{Q(B_n)}\frac{dQ}{dP}$$

and since $B_n \uparrow \Omega$, a straightforward computation shows that

$$H(Q_n,P) = E_P\left[\frac{1_{B_n}}{Q(B_n)}\frac{dQ}{dP}\log\frac{dQ}{dP}\right] - \log Q(B_n) \to H(Q,P).$$

In particular, $H(Q_n,P) < +\infty$ and since X, ΔS, and Y are integrable with respect to Q_n, it follows that $Q_n \in \mathcal{C}_P$. Further, the integrability of Z^- with respect to Q guarantees the convergence of $E_{Q_n}[Z]$ to $E_Q[Z]$ and, therefore,

$$E_Q[Z] - H(Q,P) = \lim_n\big(E_{Q_n}[Z] - H(Q_n,P)\big) \le \sup_{Q\in\mathcal{C}_P}\big(E_Q[Z] - H(Q,P)\big).$$

Taking the supremum over all $Q \in \mathcal{A}_P$ yields the claim.

(b) To conclude the proof, make the simple observation that $\mathcal{C}$ equals the union over $\mathcal{C}_P$ where P runs trough $\mathcal{P}$. This implies that

$$\begin{aligned}\sup_{P\in\mathcal{P}} \log E_P\big[\exp(Z)\big] &= \sup_{P\in\mathcal{P}} \sup_{Q\in\mathcal{C}_P} \big(E_Q[Z] - H(Q,P)\big)\\ &= \sup_{Q\in\mathcal{C}} \big(E_Q[Z] - H(Q,\mathcal{P})\big),\end{aligned}$$

where the first equality follows from step (a). □

LEMMA 3.4. *The relative entropy H is jointly convex. Moreover, the function $H(\cdot,\mathcal{P})$ and the set $\mathcal{C}$ defined in Lemma 3.3 are convex.*

PROOF. It follows from [21], Lemma 3.29, that

$$H(Q,P) = \sup\big\{E_Q[Z] - \log E_P\big[\exp(Z)\big] : Z \text{ is a bounded random variable}\big\}.$$

For any such Z, the function $(Q,P) \mapsto E_Q[Z] - \log E_P[\exp(Z)]$ is convex. Thus H, as the supremum over convex functions, is itself convex. Furthermore, the convexity of $\mathcal{P}$ yields that $H(\cdot,\mathcal{P})$ and $\mathcal{C}$ are convex. □

In the proof of Theorem 3.1, it will be important that $0 \in \operatorname{ri}\{E_Q[\Delta S] : Q \in \mathcal{C}\}$ where $\mathcal{C}$ was defined in Lemma 3.3 and ri denotes the relative interior. To get the idea why this is true, assume for simplicity that $d = 1$ and that ΔS is not $\mathcal{P}$-quasi surely equal to 0. Then, by the no-arbitrage condition, there exist two measures $P^{\pm}$ such that $P^{\pm}(\pm\Delta S > 0) > 0$. Now define

$$Q_\lambda := \lambda P^+\big(\cdot|0 < \Delta S, |X|, Y < n\big) + (1-\lambda)P^-\big(\cdot|-n < \Delta S, -|X|, -Y < 0\big)$$

for n large enough and every $\lambda \in [0,1]$. Then X, ΔS, and Y are integrable with respect to Q_λ and since $E_{Q_0}[\Delta S] < 0$, $E_{Q_1}[\Delta S] > 0$ it follows that $0 \in \operatorname{int}\{E_{Q_\lambda}[\Delta S] : \lambda \in [0,1]\}$. As the density of Q_λ with respect to $(P^+ + P^-)/2 \in \mathcal{P}$ is bounded, it holds $H(Q_\lambda,\mathcal{P}) < +\infty$, and thus $Q_\lambda \in \mathcal{C}$.

LEMMA 3.5 ([10], Lemma 3.3). *Let $X, Y : \Omega \to \mathbb{R}$ be random variables and assume that Y is nonnegative. Then one has $0 \in \operatorname{ri}\{E_Q[\Delta S] : Q \in \mathcal{C}\}$ where $\mathcal{C}$ was defined in Lemma 3.3.*

PROOF. Even though [10], Lemma 3.3, states that $0 \in \operatorname{ri}\{E_Q[\Delta S] : Q \in \Theta\}$ for the set $\Theta = \{Q : E_Q[|X| + |\Delta S| + Y] < +\infty \text{ and } Q \ll P \text{ for some } P \in \mathcal{P}\}$, the constructed measures Q have bounded densities dQ/dP with respect to some $P \in \mathcal{P}$, in particular $H(Q,\mathcal{P})$ is finite. The proof can be copied word by word. □

Before being ready for the proof of the main theorem, one last observation on the decomposition of $\mathbb{R}^d$ into relevant and irrelevant strategies h needs to be

made. Denote by $\operatorname{supp}_{\mathcal{P}} \Delta S$ the smallest closed subset of $\mathbb{R}^d$ such that $\Delta S(\omega) \in \operatorname{supp}_{\mathcal{P}} \Delta S$ for $\mathcal{P}$-quasi every ω; see [10], Lemma 4.2. Further, write $\operatorname{lin} A$ for the smallest linear space which contains a given set $A \subset \mathbb{R}^d$, and $L^\perp := \{h \in \mathbb{R}^d : hl = 0 \text{ for all } l \in L\}$ for the orthogonal complement of a linear space $L \subset \mathbb{R}^d$.

LEMMA 3.6 ([38], Lemma 2.6). *Define $L := \operatorname{lin} \operatorname{supp}_{\mathcal{P}} \Delta S$. Then one has $h \in L^\perp$ if and only if $h \Delta S = 0$ $\mathcal{P}$-quasi surely.*

PROOF OF THEOREM 3.1 AND LEMMA 3.2. In step (a), duality is shown under the assumption that X is bounded from above. The existence of an optimizer $h \in \mathbb{R}^d$ as well as continuity from below are proven simultaneously in step (b). Finally, the results from (a) and (b) are combined to extend to unbounded random endowment X in step (c).

(a) Throughout this step, assume that X is bounded from above, meaning that there exists some constant k such that $X(\omega) \le k$ for every ω. The goal is to show the following dual representation:

$$\inf_{h \in \mathbb{R}^d} \sup_{P \in \mathcal{P}} \log E_P\big[\exp(X + h\Delta S)\big] = \sup_{Q \in \mathcal{M}(|X|+Y)} \big(E_Q[X] - H(Q, \mathcal{P})\big). \tag{3.3}$$

By Lemma 3.3, it holds

$$\inf_{h \in \mathbb{R}^d} \sup_{P \in \mathcal{P}} \log E_P\big[\exp(X + h\Delta S)\big] = \inf_{h \in \mathbb{R}^d} \sup_{Q \in \mathcal{C}} \big(E_Q[X + h\Delta S] - H(Q, \mathcal{P})\big),$$

where

$$\mathcal{C} := \big\{Q \in \mathfrak{P}(\Omega) : E_Q\big[|X| + |\Delta S| + Y\big] + H(Q, \mathcal{P}) < +\infty\big\}.$$

Thus, if interchanging the infimum over $h \in \mathbb{R}^d$ and the supremum over $Q \in \mathcal{C}$ were possible, (3.3) would follow since $\inf_{h \in \mathbb{R}^d} E_Q[h\Delta S] = -\infty$ whenever Q is not a martingale measure. In what follows, we argue why one can in fact interchange the infimum and the supremum. Define

$$\Gamma := \operatorname{lin}\big\{E_Q[\Delta S] : Q \in \mathcal{C}\big\}$$

and notice that if $\Gamma = \{0\}$, then $\mathcal{C} = \mathcal{M}(|X| + Y)$ and $E_Q[h\Delta S] = 0$ for all $h \in \mathbb{R}^d$ and $Q \in \mathcal{C}$ so that there is nothing to prove. Therefore, assume in the sequel that $\Gamma \neq \{0\}$ and let

$$\{e_1, \dots, e_r\} \text{ be an orthonormal basis of } \Gamma.$$

Further, to simplify notation, define the function $J : \mathcal{C} \times \mathbb{R}^d \to \mathbb{R}$,

$$J(Q, h) := hE_Q[\Delta S] + E_Q[X] - H(Q, \mathcal{P}).$$

By Lemma 3.4, the set $\mathcal{C}$ and the function $H(\cdot, \mathcal{P})$ are convex, which shows that $J(\cdot, h)$ is concave for all $h \in \mathbb{R}^d$. Further, $J(Q, \cdot)$ is convex for all $Q \in \mathcal{C}$. Therefore, [44], Theorem 4.1, gives a sufficient condition for

$$\inf_{h \in \mathbb{R}^d} \sup_{Q \in \mathcal{C}} J(Q, h) = \sup_{Q \in \mathcal{C}} \inf_{h \in \mathbb{R}^d} J(Q, h)$$

to hold true, namely that

$$\text{(3.4)}\quad \begin{cases} \text{for every } c < \inf\limits_{h\in\mathbb{R}^d} \sup\limits_{Q\in\mathcal{C}} J(Q,h) \text{ one can find a finite set } F \subset \mathcal{C} \\ \qquad \text{such that for every } h \in \mathbb{R}^d \text{ there exists } Q \in F \text{ satisfying} \\ \qquad J(Q,h) > c. \end{cases}$$

To prove (3.4), fix such c and notice that

$$\{h \in \mathbb{R}^d : J(Q,h) > c\} = \{h \in \Gamma : J(Q,h) > c\} + \Gamma^{\perp}$$

since $hE_Q[\Delta S] = 0$ for every $h \in \Gamma^{\perp}$ and $Q \in \mathcal{C}$. Therefore, we can assume without loss of generality that $h \in \Gamma$ in the sequel. In fact, we shall distinguish between elements in Γ with large and small (Euclidean) length. From Lemma 3.5, it follows that

$$0 \in \operatorname{ri}\{E_Q[\Delta S] : Q \in \mathcal{C}\}$$

which implies that there exist $a_i^{\pm} > 0$ and $Q_i^{\pm} \in \mathcal{C}$ satisfying

$$E_{Q_i^{\pm}}[\Delta S] = \pm a_i^{\pm} e_i \qquad \text{for } 1 \le i \le r.$$

We claim that

$$\text{(3.5)}\quad \begin{cases} \max\{J(h, Q_i^{\pm}) : 1 \le i \le r\} > c + 1 > c \\ \qquad \text{for all } h \in \Gamma \text{ such that } |h| > m\sqrt{r}/\delta, \end{cases}$$

where

$$m := \max\{c + 1 - E_{Q_i^{\pm}}[X] + H(Q_i^{\pm}, \mathcal{P}) : 1 \le i \le r\} \in \mathbb{R}$$

and

$$\delta := \min\{a_i^{\pm} : 1 \le i \le r\} > 0.$$

Indeed, since $\sum_{i=1}^r (he_i)^2 = |h|^2 > r(m/\delta)^2$, it follows that $|he_j| > m/\delta$ for some $1 \le j \le r$. If $he_j > m/\delta$, it holds

$$hE_{Q_j^+}[\Delta S] = ha_j^+ e_j > \frac{ma_j^+}{\delta} \ge m \ge c + 1 - E_{Q_j^+}[X] + H(Q_j^+, \mathcal{P})$$

and a rearrangement of the appearing terms yields $J(h, Q_j^+) > c + 1$. If $he_j < -m/\delta$, the same argumentation shows that $J(h, Q_j^-) > c + 1$. Further, as

$$J(Q, \cdot) \text{ is continuous and } c < \inf_{h\in\Gamma} \sup_{Q\in\mathcal{C}} J(Q,h),$$

the collection

$$U_Q := \{h \in \Gamma : J(Q,h) > c\},$$

where $Q \in \mathcal{C}$, forms an open cover of Γ. By compactness of the set $\{h \in \Gamma : |h| \leq m\sqrt{r}/\delta\}$, there exists a finite family $F' \subset \mathcal{C}$ such that

$$\{h \in \Gamma : |h| \leq m\sqrt{r}/\delta\} \subset \bigcup \{U_Q : Q \in F'\}.$$

Then $F := F' \bigcup \{Q_i^{\pm} : 1 \leq i \leq r\}$ is still finite and it holds

$$\Gamma = \bigcup \{U_Q : Q \in F\},$$

which is a reformulation of (3.4). Putting everything together, it follows that

$$\begin{aligned}
&\inf_{h \in \mathbb{R}^d} \sup_{P \in \mathcal{P}} \log E_P\big[\exp(X + h\Delta S)\big] \\
&\quad = \inf_{h \in \mathbb{R}^d} \sup_{Q \in \mathcal{C}} J(Q, h) \\
&\quad = \sup_{Q \in \mathcal{C}} \inf_{h \in \mathbb{R}^d} J(Q, h) = \sup_{Q \in \mathcal{M}(|X|+Y)} \big(E_Q[X] - H(Q, \mathcal{P})\big).
\end{aligned}$$

In particular, since $\mathcal{M}(|X| + Y)$ is not empty by Lemma 3.5, it follows that the optimization problem does not take the value $-\infty$.

(b) We proceed to show that the optimization problem is continuous from below (Lemma 3.2) and that an optimal strategy $h \in \mathbb{R}^d$ exists. Recall that X_n is a sequence increasing point-wise to X. For the existence of an optimal strategy for a fixed function X, consider the constant sequence $X_n := X$ in the following argumentation. For each natural number n, let $h_n \in \mathbb{R}^d$ such that

$$\begin{aligned}
&\inf_{h \in \mathbb{R}^d} \sup_{P \in \mathcal{P}} \log E_P\big[\exp(X_n + h\Delta S)\big] \\
&\qquad \geq \sup_{P \in \mathcal{P}} \log E_P\big[\exp(X_n + h_n\Delta S)\big] - \frac{1}{n}.
\end{aligned} \tag{3.6}$$

By step (a), this is possible, that is, the left-hand side of (3.6) is not equal to $-\infty$. By Lemma 3.6, we may assume without loss of generality that every h_n is an element of $L := \operatorname{lin}\operatorname{supp}_{\mathcal{P}} \Delta S$.

First assume that the sequence h_n is unbounded, that is, $\sup_n |h_n| = +\infty$. Then, possibly after passing to a subsequence, $h_n/|h_n|$ converges to some limit h^*. Since $|h^*| = 1$ and $h^* \in L$, it follows from Lemma 3.6 and the NA($\mathcal{P}$)-condition, that $P'(A) > 0$ for some $P' \in \mathcal{P}$ where $A := \{h^*\Delta S > 0\}$. However, since

$$\exp(X_n + h_n\Delta S)1_A \to +\infty 1_A,$$

an application of Fatou's lemma yields

$$\inf_{h \in \mathbb{R}^d} \sup_{P \in \mathcal{P}} \log E_P\big[\exp(X_n + h\Delta S)\big] \geq \log E_{P'}\big[\exp(X_n + h_n\Delta S)\big] - \frac{1}{n} \to +\infty.$$

But then, since the sequence X_n is increasing, it follows that

$$\begin{aligned}&\inf_{h\in\mathbb{R}^d}\sup_{P\in\mathcal{P}}\log E_P\big[\exp(X+h\Delta S)\big]\\&\qquad\geq\lim_n\inf_{h\in\mathbb{R}^d}\sup_{P\in\mathcal{P}}\log E_P\big[\exp(X_n+h\Delta S)\big]=+\infty.\end{aligned}$$

Hence the optimization problem is continuous from below and every $h\in\mathbb{R}^d$ is optimal for X.

If the sequence h_n is bounded, again possibly after passing to a subsequence, h_n converges to some limit $h^*\in\mathbb{R}^d$. Now it follows that

$$\begin{aligned}&\sup_{P\in\mathcal{P}}\log E_P\big[\exp(X+h^*\Delta S)\big]\\&\qquad\leq\liminf_n\bigg(\sup_{P\in\mathcal{P}}\log E_P\big[\exp(X_n+h_n\Delta S)\big]-\frac{1}{n}\bigg)\\&\qquad\leq\liminf_n\inf_{h\in\mathbb{R}^d}\sup_{P\in\mathcal{P}}\log E_P\big[\exp(X_n+h\Delta S)\big]\\&\qquad\leq\inf_{h\in\mathbb{R}^d}\sup_{P\in\mathcal{P}}\log E_P\big[\exp(X+h\Delta S)\big],\end{aligned}$$

where the first inequality follows from Fatou's lemma, the second one since h_n was chosen optimal up to an error of $1/n$, and the last one since X_n is an increasing sequence. This shows both that the optimization problem is continuous from below and that h^* is optimal for X.

(c) In the final step, the duality established in (a) is extended to general random endowment. Let $X\colon\Omega\to(-\infty,+\infty]$ be measurable and observe that

$$\mathcal{M}\big(X^-+Y\big)=\mathcal{M}\big(|X\wedge n|+Y\big)\qquad\text{for all } n\in\mathbb{N}$$

since X^- is integrable if and only if $(X\wedge n)^-$ is. Moreover, for any $Q\in\mathcal{M}(X^-+Y)$ the monotone convergence theorem applies and yields that $\sup_n E_Q[X\wedge n]=E_Q[X]$. But then it follows that

$$\begin{aligned}&\inf_{h\in\mathbb{R}^d}\sup_{P\in\mathcal{P}}\log E_P\big[\exp(X+h\Delta S)\big]\\&\qquad=\sup_n\inf_{h\in\mathbb{R}^d}\sup_{P\in\mathcal{P}}\log E_P\big[\exp(X\wedge n+h\Delta S)\big]\\&\qquad=\sup_n\sup_{Q\in\mathcal{M}(X^-+Y)}\big(E_Q[X\wedge n]-H(Q,\mathcal{P})\big)\\&\qquad=\sup_{Q\in\mathcal{M}(X^-+Y)}\big(E_Q[X]-H(Q,\mathcal{P})\big),\end{aligned}$$

where the first and second equality follow from step (b) and (a), respectively, and the last one by interchanging two suprema. □

4. Proofs for the multiperiod case.

4.1. *The case without options.* In this section, measurable selection arguments are used to show that the global analysis can be reduced to a local one wherein the results of the one-period case are used. For each $0 \le t \le T-1$ and $\omega \in \Omega_t$, define

$$\mathcal{P}_t^T(\omega) = \{P_t \otimes \cdots \otimes P_{T-1} : P_s(\cdot) \in \mathcal{P}_s(\omega, \cdot) \text{ for } t \le s \le T-1\},$$

where each $P_s(\cdot)$ is a universally measurable selector of $\mathcal{P}_s(\omega, \cdot)$. Thus $\mathcal{P}_t^T(\omega)$ corresponds to the set of all possible probability scenarios for the future stock prices $S_{t+1}, \dots, S_T$, given the past $\omega \in \Omega_t$. In particular, it holds $\mathcal{P} = \mathcal{P}_0^T$ in line with Bouchard and Nutz. In order to keep the indices to a minimum, fix two functions

$$X : \Omega \to (-\infty, +\infty] \quad \text{and} \quad Y : \Omega \to [0, +\infty)$$

such that X and $-Y$ are upper semianalytic, and define the set of all martingale measures for the future stock prices $S_{t+1}, \dots, S_T$ given the past $\omega \in \Omega_t$ by

$$\mathcal{M}_t^T(\omega) = \left\{ Q \in \mathfrak{P}(\Omega_{T-t}) : \begin{array}{l} (S_s(\omega, \cdot))_{t \le s \le T} \text{ is a } Q\text{-martingale and} \\ E_Q[X(\omega, \cdot)^- + Y(\omega, \cdot)] + H(Q, \mathcal{P}_t^T(\omega)) < +\infty \end{array} \right\}$$

for each $0 \le t \le T-1$ and $\omega \in \Omega_t$. It is shown in Lemma 4.5 that $\mathcal{M}_t^T$ has analytic graph and within the proof of Theorem 4.1 that its values are not empty. Note that $\mathcal{M}_0^T = \mathcal{M}(Y) = \{Q \in \mathcal{M} : E_Q[Y + X^-] < +\infty\}$, where $\mathcal{M}$ was defined in Section 2. Further, introduce the dynamic version of the optimization problem: Define

$$\mathcal{E}_T(\omega, x) := X(\omega) + x$$

for $(\omega, x) \in \Omega \times \mathbb{R}$ and recursively

$$\mathcal{E}_t(\omega, x) := \inf_{h \in \mathbb{R}^d} \sup_{P \in \mathcal{P}_t(\omega)} \log E_P[\exp(\mathcal{E}_{t+1}(\omega \otimes_t \cdot, x + h \Delta S_{t+1}(\omega, \cdot)))] \tag{4.1}$$

for $(\omega, x) \in \Omega_t \times \mathbb{R}$. Here, we write $\omega \otimes_t \omega' := (\omega, \omega') \in \Omega_{t+s}$ for $\omega \in \Omega_t$ and $\omega' \in \Omega_s$ instead of $(\omega, \cdot)$ to avoid confusion. It will be shown later that $\mathcal{E}_t$ is well defined, that is, that the term inside the expectation is appropriately measurable.

The following theorem is the main result of this section and includes Theorem 2.1 as a special case (corresponding to $t = 0$).

THEOREM 4.1. *For every $0 \le t \le T-1$ and $\omega \in \Omega_t$, it holds*

$$\begin{aligned} \mathcal{E}_t(\omega, x) - x &= \inf_{\vartheta \in \Theta} \sup_{P \in \mathcal{P}_t^T(\omega)} \log E_P[\exp(X(\omega, \cdot) + (\vartheta \cdot S)_t^T(\omega, \cdot))] \\ &= \sup_{Q \in \mathcal{M}_t^T(\omega)} (E_Q[X(\omega, \cdot)] - H(Q, \mathcal{P}_t^T(\omega))) \end{aligned}$$

and both terms are not equal to $-\infty$. Moreover, the infimum over $\vartheta \in \Theta$ is attained.

We start by investigating properties of the (robust) relative entropy and the graph of $\mathcal{M}_t$, which will ensure that measurable selection arguments can be applied. We then focus on deriving a duality for $\mathcal{E}_t$ and last prove the dynamic programming principle.

LEMMA 4.2 ([19], Lemma 1.4.3.b). *The relative entropy H is Borel.*

PROOF. Any Borel function can be approximated in measure by continuous functions, so it follows as in the proof of Lemma 3.4 that

$$H(Q, P) = \sup\{E_Q[Z] - \log E_P[\exp(Z)] : Z \text{ is bounded and continuous}\}.$$

Therefore, $\{H \le c\}$ is closed for any real number c showing that H is Borel. □

The so-called chain rule for the relative entropy is well known; a proof can be found, for example, in Appendix C3 of the book by Dupuis and Ellis [19]. However, since we are dealing with universally measurable kernels and also in order to be self-contained, a proof is given in the Appendix. For the link between dynamic risk measures and this chain rule, see, for example, [12] in the dominated, and [3, 29] in the nondominated setting.

LEMMA 4.3. *Let $0 \le t \le T-1$ and $P, Q \in \mathfrak{P}(\Omega_{T-t})$. Then*

$$H(Q, P) = \sum_{s=t}^{T-1} E_Q[H(Q_s(\cdot), P_s(\cdot))],$$

where Q_s and P_s are universally measurable kernels such that $Q = Q_t \otimes \cdots \otimes Q_{T-1}$ and $P = P_t \otimes \cdots \otimes P_{T-1}$.

LEMMA 4.4. *For any $0 \le t \le T-1$, the function*

$$\Omega_t \times \mathfrak{P}(\Omega_{T-t}) \to [-\infty, 0], \qquad (\omega, Q) \mapsto -H(Q, \mathcal{P}_t^T(\omega))$$

is upper semianalytic. Moreover, it holds

$$H(Q, \mathcal{P}_t^T(\omega)) = H(Q_t, \mathcal{P}_t(\omega)) + E_Q[H(Q'(\cdot), \mathcal{P}_{t+1}^T(\omega, \cdot))],$$

where Q' is a universally measurable kernel such that $Q = Q_t \otimes Q'$.

PROOF. Every probability $Q \in \mathfrak{P}(\Omega_{T-t})$ can be written as $Q = Q_t \otimes \cdots \otimes Q_{T-1}$ where Q_s are the kernels from Remark B.2, that is, such that

$$\Omega_{s-t} \times \mathfrak{P}(\Omega_{T-t}) \to \mathfrak{P}(\Omega_1), \qquad (\bar{\omega}, Q) \mapsto Q_s(\bar{\omega})$$

is Borel.

(a) We start by showing that

$$\begin{aligned} &\Omega_t \times \mathfrak{P}(\Omega_{T-t}) \to [-\infty, 0], \\ &(\omega, Q) \mapsto \sum_{s=t}^{T-1} -E_Q\big[H\big(Q_s(\cdot), \mathcal{P}_s(\omega, \cdot)\big)\big] \end{aligned} \tag{4.2}$$

is upper semianalytic. Fix some $t \le s \le T-1$. In the sequel, ω will refer to elements in Ω_t and $\bar{\omega}$ to elements in Ω_{s-t}. Since $(\bar{\omega}, Q) \mapsto Q_s(\bar{\omega})$ is Borel by construction and the entropy H is Borel by Lemma 4.2, the composition

$$\begin{aligned} &\Omega_t \times \Omega_{s-t} \times \mathfrak{P}(\Omega_{T-t}) \times \mathfrak{P}(\Omega_1) \to [-\infty, 0], \\ &(\omega, \bar{\omega}, Q, R) \mapsto -H\big(Q_s(\bar{\omega}), R\big) \end{aligned}$$

is Borel as well. As the graph of $\mathcal{P}_s$ is analytic, it follows from [7], Proposition 7.47, that

$$\begin{aligned} &\Omega_t \times \Omega_{s-t} \times \mathfrak{P}(\Omega_{T-t}) \to [-\infty, 0], \\ &(\omega, \bar{\omega}, Q) \mapsto -H\big(Q_s(\bar{\omega}), \mathcal{P}_s(\omega, \bar{\omega})\big) \end{aligned} \tag{4.3}$$

is upper semianalytic. Moreover, [7], Proposition 7.50, guarantees that for any $\varepsilon > 0$, there exists a universally measurable kernel P_s^ε such that

$$\begin{cases} P_s^\varepsilon(\omega, \bar{\omega}, Q) \in \mathcal{P}_s(\omega, \bar{\omega}), \\ H\big(Q_s(\bar{\omega}), P_s^\varepsilon(\omega, \bar{\omega}, Q)\big) \le H\big(Q_s(\bar{\omega}), \mathcal{P}_s(\omega, \bar{\omega})\big) + \varepsilon \end{cases} \tag{4.4}$$

for all $(\omega, \bar{\omega}, Q)$. This will be used in part (b). Further, since

$$\Omega_t \times \mathfrak{P}(\Omega_{T-t}) \to [-\infty, 0], \qquad (\omega, Q) \mapsto -E_Q\big[H\big(Q_s(\cdot), \mathcal{P}_s(\omega, \cdot)\big)\big]$$

is just (4.3) integrated with respect to $Q(d\bar{\omega})$, an application of Lemma B.1 shows that this mapping is upper semianalytic. Finally, the fact that sums of upper semianalytic functions are again upper semianalytic (see [7], Lemma 7.30) implies that (4.2) is upper semianalytic as was claimed.

(b) Fix some $\omega \in \Omega_t$ and $Q \in \mathfrak{P}(\Omega_{T-t})$. From Lemma 4.3, it follows that

$$\begin{aligned} H\big(Q, \mathcal{P}_t^T(\omega)\big) &= \inf_{P \in \mathcal{P}_t^T(\omega)} \sum_{s=t}^{T-1} E_Q\big[H\big(Q_s(\cdot), P_s(\cdot)\big)\big] \\ &\ge \sum_{s=t}^{T-1} E_Q\big[H\big(Q_s(\cdot), \mathcal{P}_s(\omega, \cdot)\big)\big]. \end{aligned}$$

For the other inequality, let $\varepsilon > 0$ be arbitrary and P_s^ε be the kernels from (4.4). Recall that Q and ω are fixed so that

$$P_s' \colon \Omega_{s-t} \to \mathfrak{P}(\Omega_1), \qquad \bar{\omega} \mapsto P_s^\varepsilon(\omega, \bar{\omega}, Q)$$

is still universally measurable by [7], Lemma 7.29. Then it follows that

$$P' := P'_t \otimes \cdots \otimes P'_{T-1} \in \mathcal{P}_t^T(\omega)$$

and, using Lemma 4.3 once more, that

$$\begin{aligned}
&\sum_{s=t}^{T-1} E_Q\big[H\big(Q_s(\cdot), \mathcal{P}_s(\omega, \cdot)\big)\big] \\
&\qquad \geq \sum_{s=t}^{T-1} E_Q\big[H\big(Q_s(\cdot), P'_s(\cdot)\big) - \varepsilon\big] \\
&\qquad = H\big(Q, P'\big) - (T-t)\varepsilon \geq H\big(Q, \mathcal{P}_t^T(\omega)\big) - (T-t)\varepsilon.
\end{aligned}$$

As ε was arbitrary, this shows the desired inequality.

(c) Finally, kernels are almost-surely unique so that

$$Q' = Q_{t+1} \otimes \cdots \otimes Q_{T-1}, \qquad Q_t\text{-almost surely.}$$

Hence it follows that

$$H\big(Q'(\cdot), \mathcal{P}_{t+1}(\omega, \cdot)\big) = \sum_{s=t+1}^{T-1} E_{Q'(\cdot)}\big[H\big(Q_s(\cdot), \mathcal{P}_s(\omega, \cdot)\big)\big], \qquad Q_t\text{-almost surely.}$$

It only remains to integrate this equation with respect to Q_t. $\square$

Fix a measure $Q = Q_t \otimes \cdots \otimes Q_{T-1} \in \mathfrak{P}(\Omega_{T-t})$ and $\omega \in \Omega_t$. An elementary computation shows that Q is a martingale measure for $(S_s(\omega, \cdot))_{t \leq s \leq T}$ if and only if $E_Q[|\Delta S_{s+1}(\omega, \cdot)|] < +\infty$ and

$$E_{Q_s(\bar{\omega})}\big[\Delta S_{s+1}(\omega, \bar{\omega}, \cdot)\big] = 0 \qquad \text{for } Q_t \otimes \cdots \otimes Q_{s-1}\text{-almost every } \bar{\omega} \in \Omega_{s-t}$$

and every $t \leq s \leq T-1$. This is used in the sequel without reference.

LEMMA 4.5. *The graph of $\mathcal{M}_t^T$ is analytic.*

PROOF. First notice that $Z := X \wedge 0 - Y$ is upper semianalytic. This follows from the fact that $\{X \wedge 0 \geq a\}$ equals $\varnothing$ if $a > 0$ and $\{X \geq a\}$ else and that the sum of upper semianalytic functions remains upper semianalytic. Therefore, an application of Lemma B.1 shows that

$$\Omega_t \times \mathfrak{P}(\Omega_{T-t}) \to [-\infty, 0], \qquad (\omega, Q) \mapsto E_Q\big[Z(\omega, \cdot)\big]$$

is upper semianalytic. Then, since $(\omega, Q) \mapsto -H(Q, \mathcal{P}_t^T(\omega))$ is upper semianalytic by Lemma 4.4 and the sum of upper semianalytic mappings is again upper semianalytic, it follows that

$$A := \big\{(\omega, Q) : E_Q\big[Z(\omega, \cdot)\big] - H\big(Q, \mathcal{P}_t^T(\omega)\big) > -\infty\big\}$$

is an analytic set. The missing part now is the martingale property. First notice that

$$\Omega_t \times \mathfrak{P}(\Omega_{T-t}) \to [0, +\infty], \qquad (\omega, Q) \mapsto E_Q\big[\big|\Delta S_{s+1}(\omega, \cdot)\big|\big]$$

is Borel by Lemma B.1. As before, for every $Q \in \mathfrak{P}(\Omega_{T-t})$, we will write $Q = Q_t \otimes \cdots \otimes Q_{T-1}$ for the kernels Q_s from Remark B.2. Then, since $(\bar{\omega}, Q) \mapsto Q_s(\bar{\omega})$ is Borel, a twofold application of Lemma B.1 shows that

$$\Omega_t \times \mathfrak{P}(\Omega_{T-t}) \to [0, +\infty], \qquad (\omega, Q) \mapsto E_Q\big[\big|E_{Q_s(\cdot)}\big[\Delta S_{s+1}(\omega, \cdot)\big]\big|\big]$$

is Borel. Thus

$$B_s := \big\{(\omega, Q) : E_Q\big[\big|\Delta S_{s+1}(\omega, \cdot)\big|\big] < +\infty \text{ and } E_Q\big[\big|E_{Q_s(\cdot)}\big[\Delta S_{s+1}(\omega, \cdot)\big]\big|\big] = 0\big\}$$

is Borel which implies that

$$\operatorname{graph} \mathcal{M}_t^T = \bigcap \{B_s : t \le s \le T-1\} \cap A,$$

as the finite intersection of analytic sets, is itself analytic (see [7], Corollary 7.35.2). $\square$

Define

$$\mathcal{D}_t(\omega) := \sup_{Q \in \mathcal{M}_t^T(\omega)} \big(E_Q\big[X(\omega, \cdot)\big] - H\big(Q, \mathcal{P}_t^T(\omega)\big)\big)$$

for all $0 \le t \le T-1$ and $\omega \in \Omega_t$, and recall that

$$\mathcal{E}_t(\omega, x) := \inf_{h \in \mathbb{R}^d} \sup_{P \in \mathcal{P}_t(\omega)} \log E_P\big[\exp\big(\mathcal{E}_{t+1}\big((\omega, \cdot), x + h\Delta S_{t+1}(\omega, \cdot)\big)\big)\big]$$

for $(\omega, x) \in \Omega_t \times \mathbb{R}$.

PROOF OF THEOREM 4.1—DUALITY. We claim that

$$\begin{cases} \mathcal{E}_t(\omega, x) = \mathcal{D}_t(\omega) + x \text{ and } \mathcal{D}_t(\omega) \in (-\infty, +\infty] \\ \qquad \text{for all } \omega \in \Omega_t,\ x \in \mathbb{R} \text{ and } 0 \le t \le T-1. \end{cases} \tag{4.5}$$

The proof will be a backward induction. For $t = T-1$, (4.5) is just the statement of Theorem 3.1.

Now assume that (4.5) holds true for $t+1$. First we artificially bound X from above and then pass to the limit. More precisely, define

$$\mathcal{D}_s^n(\omega) := \sup_{Q \in \mathcal{M}_s^T(\omega)} \big(E_Q\big[X(\omega, \cdot) \wedge n\big] - H\big(Q, \mathcal{P}_s^T(\omega)\big)\big)$$

for $s = t, t+1$ and $\omega \in \Omega_s$, and notice that $\mathcal{D}_s^n$ is upper semianalytic. Indeed, since $X(\omega, \cdot) \wedge n$ is upper semianalytic, the mapping $(\omega, Q) \mapsto E_Q[X(\omega, \cdot) \wedge n]$

is upper semianalytic by Lemma B.1. Then Lemma 4.4 and the fact that the sum of upper semianalytic functions stays upper semianalytic implies that

$$(\omega, Q) \mapsto E_Q\big[X(\omega, \cdot) \wedge n\big] - H\big(Q, \mathcal{P}_s^T(\omega)\big)$$

is upper semianalytic. Since the graph of $\mathcal{M}_s^T$ is analytic by Lemma 4.5, it follows from [7], Proposition 7.47, that $\mathcal{D}_s^n$ is upper semianalytic. Moreover, [7], Proposition 7.50, guarantees that for any $\varepsilon > 0$ there exists a universally measurable kernel $Q^\varepsilon(\cdot) \in \mathcal{M}_{t+1}^T(\omega, \cdot)$ such that

$$\mathcal{D}_{t+1}^n(\omega \otimes_t \cdot) \le E_{Q^\varepsilon(\cdot)}\big[X(\omega, \cdot) \wedge n\big] - H\big(Q^\varepsilon(\cdot), \mathcal{P}_{t+1}^T(\omega, \cdot)\big) + \varepsilon. \tag{4.6}$$

By interchanging two suprema, it holds $\mathcal{D}_s = \sup_n \mathcal{D}_s^n$ (for more details, see part (c) of the proof of Theorem 3.1). In particular, $\mathcal{D}_s$ is upper semianalytic, as the countable supremum over upper semianalytic functions. Therefore, it follows from Lemma 3.2 that $\mathcal{E}_t = \sup_n \mathcal{E}_t^n$ where

$$\mathcal{E}_t^n(\omega, x) := \inf_{h \in \mathbb{R}^d} \sup_{P \in \mathcal{P}_t(\omega)} \log E_P\big[\exp\big(\mathcal{D}_{t+1}^n(\omega \otimes_t \cdot) + h\Delta S_{t+1}(\omega, \cdot) + x\big)\big].$$

The goal now is to show that $\mathcal{E}_t^n$ equals $\mathcal{D}_t^n$ for all n, from which it follows that

$$\mathcal{E}_t(\omega, x) = \sup_n \mathcal{E}_t^n(\omega, x) = \sup_n \mathcal{D}_t^n(\omega) + x = \mathcal{D}_t(\omega) + x$$

and the proof is complete. To show that indeed $\mathcal{E}_t^n(\omega, x) = \mathcal{D}_t^n(\omega) + x$, fix some n, x, and $\omega \in \Omega_t$. By Theorem 3.1, it holds

$$\mathcal{E}_t^n(\omega, x) = \sup_{Q_t \in \mathcal{M}_t(Z)} \big(E_{Q_t}\big[\mathcal{D}_{t+1}^n(\omega \otimes_t \cdot)\big] - H\big(Q_t, \mathcal{P}_t(\omega)\big)\big) + x > -\infty,$$

where

$$\mathcal{M}_t(Z) := \left\{ Q \in \mathfrak{P}(\Omega_1) : \begin{array}{l} E_Q\big[\mathcal{D}_{t+1}^n(\omega \otimes_t \cdot)^- + \big|\Delta S_{t+1}(\omega, \cdot)\big| + Z\big] < +\infty, \\ H\big(Q, \mathcal{P}_t(\omega)\big) < +\infty \text{ and } E_Q\big[\Delta S_{t+1}(\omega, \cdot)\big] = 0 \end{array} \right\}$$

and $Z\colon \Omega_1 \to [0, +\infty)$ is an arbitrary universally measurable function.

We start by showing that $\mathcal{E}_t^n(\omega, x) \le \mathcal{D}_t^n(\omega) + x$. Fix some $\varepsilon > 0$, let $Q^\varepsilon(\cdot) \in \mathcal{M}_{t+1}^T(\omega, \cdot)$ be the kernel from (4.6), and define $Z\colon \Omega_1 \to [0, +\infty)$,

$$Z := E_{Q^\varepsilon(\cdot)}\left[X(\omega, \cdot)^- + Y(\omega, \cdot) + \sum_{s=t+2}^T \big|\Delta S_s(\omega, \cdot)\big|\right] + H\big(Q^\varepsilon(\cdot), \mathcal{P}_{t+1}^T(\omega, \cdot)\big).$$

Then Z is real-valued by the definition of $\mathcal{M}_{t+1}^T(\omega, \cdot)$ and it follows from Lemma B.1, Lemma 4.4 and [7], Proposition 7.44, that Z is universally measurable. Moreover,

$$Q_t \otimes Q^\varepsilon \in \mathcal{M}_t^T(\omega) \qquad \text{for any } Q_t \in \mathcal{M}_t(Z). \tag{4.7}$$

To show this, fix some $Q_t \in \mathcal{M}_t(Z)$ and define $Q := Q_t \otimes Q^\varepsilon$. Then an application of Lemma 4.4 yields

$$\begin{aligned} H\big(Q, \mathcal{P}_t^T(\omega)\big) &= H\big(Q_t, \mathcal{P}_t(\omega)\big) + E_{Q_t}\big[H\big(Q^\varepsilon(\cdot), \mathcal{P}_{t+1}^T(\omega, \cdot)\big)\big] \\ &\le H\big(Q_t, \mathcal{P}_t(\omega)\big) + E_{Q_t}[Z] < +\infty. \end{aligned}$$

Moreover, it holds

$$E_Q\left[X(\omega, \cdot)^- + Y(\omega, \cdot) + \sum_{s=t+1}^T \big|\Delta S_s(\omega, \cdot)\big|\right] \le E_{Q_t}\big[\big|\Delta S_{t+1}(\omega, \cdot)\big| + Z\big] < +\infty$$

so that indeed $Q \in \mathcal{M}_t^T(\omega)$ and, therefore,

$$\begin{aligned} &E_{Q_t}\big[\mathcal{D}_{t+1}^n(\omega \otimes_t \cdot)\big] - H\big(Q_t, \mathcal{P}_t(\omega)\big) \\ &\quad \le E_{Q_t}\big[E_{Q^\varepsilon(\cdot)}\big[X(\omega, \cdot) \wedge n\big] - H\big(Q^\varepsilon(\cdot), \mathcal{P}_{t+1}^T(\omega, \cdot)\big) + \varepsilon\big] - H\big(Q_t, \mathcal{P}_t(\omega)\big) \\ &\quad = E_Q\big[X(\omega, \cdot) \wedge n\big] - H\big(Q, \mathcal{P}_t^T(\omega)\big) + \varepsilon \le \mathcal{D}_t^n(\omega) + \varepsilon. \end{aligned}$$

As $Q_t \in \mathcal{M}_t(Z)$ and $\varepsilon > 0$ were arbitrary, it follows that $\mathcal{E}_t^n(\omega, x) \le \mathcal{D}_t^n(\omega) + x$.

To show the other inequality, that is, $\mathcal{E}_t^n(\omega, x) \ge \mathcal{D}_t^n(\omega) + x$, fix some measure $Q \in \mathcal{M}_t^T(\omega)$ which we write as

$$Q = Q_t \otimes Q'$$

for a measure Q_t on Ω_1 and a Borel kernel $Q' \colon \Omega_1 \to \mathfrak{P}(\Omega_{T-t-1})$. Then

$$Q_t \in \mathcal{M}_t(0) \quad \text{and} \quad Q'(\cdot) \in \mathcal{M}_{t+1}^T(\omega, \cdot) \qquad Q_t\text{-almost surely},$$

where $\mathcal{M}_t(0) = \mathcal{M}_t(Z)$ for the function $Z \equiv 0$. Indeed, first notice that

$$H\big(Q_t, \mathcal{P}_t(\omega)\big) + E_{Q_t}\big[H\big(Q'(\cdot), \mathcal{P}_{t+1}^T(\omega, \cdot)\big)\big] = H\big(Q, \mathcal{P}_t^T(\omega)\big) < +\infty$$

by Lemma 4.4, so that

$$H\big(Q'(\cdot), \mathcal{P}_{t+1}^T(\omega, \cdot)\big) < +\infty \qquad Q_t\text{-almost surely}.$$

Similarly, we conclude that $E_{Q'(\cdot)}[X(\omega, \cdot)^- + Y(\omega, \cdot) + |\Delta S_s(\omega, \cdot)|] < +\infty$ Q_t-almost surely for all $t + 2 \le s \le T$. Thus it holds

$$Q'(\cdot) \in \mathcal{M}_{t+1}^T(\omega, \cdot) \qquad Q_t\text{-almost surely}.$$

But then it follows from the definition of $\mathcal{D}_{t+1}^n$ that

$$E_{Q'(\cdot)}\big[X(\omega, \cdot) \wedge n\big] - H\big(Q'(\cdot), \mathcal{P}_{t+1}^T(\omega, \cdot)\big) \le \mathcal{D}_{t+1}^n(\omega \otimes_t \cdot)$$

Q_t-almost surely, so that

$$\begin{aligned} E_{Q_t}\big[\mathcal{D}_{t+1}^n(\omega \otimes_t \cdot)^-\big] &\le E_{Q_t}\big[\big(E_{Q'(\cdot)}\big[X(\omega, \cdot) \wedge n\big] - H\big(Q'(\cdot), \mathcal{P}_{t+1}^T(\omega, \cdot)\big)\big)^-\big] \\ &\le E_Q\big[X(\omega, \cdot)^-\big] + H\big(Q, \mathcal{P}_t^T(\omega)\big) < +\infty \end{aligned}$$

by Jensen's inequality and Lemma 4.4. Therefore, one has $Q_t \in \mathcal{M}_t(0)$ and it follows that

$$\begin{aligned} &E_Q\big[X(\omega,\cdot)\wedge n\big] - H\big(Q,\mathcal{P}_t^T(\omega)\big) \\ &\quad = E_{Q_t}\big[E_{Q'(\cdot)}\big[X(\omega,\cdot)\wedge n\big] - H\big(Q'(\cdot),\mathcal{P}_{t+1}^T(\omega,\cdot)\big)\big] - H\big(Q_t,\mathcal{P}_t(\omega)\big) \\ &\quad \le \sup_{R\in\mathcal{M}_t(0)} \big(E_R\big[\mathcal{D}_{t+1}^n(\omega\otimes_t\cdot)\big] - H\big(R,\mathcal{P}_t(\omega)\big)\big) = \mathcal{E}_t^n(\omega,x) - x \end{aligned}$$

and as $Q \in \mathcal{M}_t^T(\omega)$ was arbitrary, that indeed $\mathcal{D}_t^n(\omega)+x \le \mathcal{E}_t(\omega,x)$. Coupled with the other inequality which was shown before, it holds $\mathcal{E}_t(\omega,x) = \mathcal{D}_t^n(\omega)+x$ and the proof is complete. □

The following lemma will be important in the proof of the dynamic programming principle. Since it was already shown that $\mathcal{E}_t(\omega,x) = \mathcal{D}_t(\omega)+x$, the proof is almost one to one to the one for [38], Lemma 3.7. For the sake of completeness, a proof is given in the Appendix.

LEMMA 4.6 ([38], Lemma 3.7). *For every $0 \le t \le T-1$ and $x \in \mathbb{R}$, there exists a process $\vartheta^* \in \Theta$ such that*

$$\begin{aligned} &\mathcal{E}_s\big(\omega, x + (\vartheta^*\cdot S)_t^s(\omega)\big) \\ &\quad = \sup_{P\in\mathcal{P}_s(\omega)} \log E_P\big[\exp\big(\mathcal{E}_{s+1}\big(\omega\otimes_s\cdot, x + (\vartheta^*\cdot S)_t^{s+1}(\omega,\cdot)\big)\big)\big] \end{aligned}$$

for all $t \le s \le T-1$ and $\omega \in \Omega_s$.

PROOF OF THEOREM 4.1—DYNAMIC PROGRAMMING. We turn to the proof of the dynamic programming principle, that is, we show that

$$C := \inf_{\vartheta\in\Theta} \sup_{P\in\mathcal{P}_t^T(\omega)} \log E_P\big[\exp\big(X(\omega,\cdot)+x+(\vartheta\cdot S)_t^T(\omega,\cdot)\big)\big] = \mathcal{E}_t(\omega,x) \tag{4.8}$$

for all x, $\omega \in \Omega_t$, and $0 \le t \le T-1$ and that the infimum over $\vartheta \in \Theta$ is attained. Again, fix some x, t, and $\omega \in \Omega_t$. By the first part of the proof of Theorem 4.1, that is, the part which focuses on duality, it holds $\mathcal{E}_t(\omega,x) = \mathcal{D}_t(\omega)+x$. Therefore, x can be subtracted on both sides of (4.8) and there is no loss of generality in assuming that $x = 0$. This will lighten notation.

First we focus on the inequality $C \ge \mathcal{E}_t(\omega,x)$. Fix some $\vartheta \in \Theta$, $P \in \mathcal{P}_t^T(\omega)$, and $Q \in \mathcal{M}_t^T(\omega)$. If

$$C' := \log E_P\big[\exp\big(X(\omega,\cdot)+(\vartheta\cdot S)_t^T(\omega,\cdot)\big)\big] \ge E_Q\big[X(\omega,\cdot)\big] - H(Q,P),$$

then the claim follows by taking the supremum over all those Q and P, and in a second step the infimum over all $\vartheta \in \Theta$. To show this, one may assume that C' and $H(Q,P)$ are finite, otherwise there is nothing to prove. Define

$$Z := X(\omega,\cdot) + (\vartheta\cdot S)_t^T(\omega,\cdot).$$

Applying the elementary inequality $ab \le \exp(a) + b \log b$ to "$a = Z^+$" and "$b = dQ/dP$" yields

$$E_Q[Z^+] \le E_P[\exp(Z^+)] + H(Q, P) \le \exp(C') + 1 + H(Q, P) < +\infty.$$

Therefore, it holds

$$E_Q[(\vartheta \cdot S)_t^T(\omega, \cdot)^+] \le E_Q[X(\omega, \cdot)^-] + E_Q[Z^+] < +\infty$$

by the definition of $\mathcal{M}_t^T(\omega)$. But then it follows from a result on local martingales (see [27], Theorem 1 and 2) that $(\vartheta \cdot S)_t^T(\omega, \cdot)$ is actually integrable with respect to Q and has expectation 0. Hence $E_Q[Z^-] < +\infty$ and, therefore, Lemma A.1 yields

$$C' = \log E_P[\exp(Z)] \ge E_Q[Z] - H(Q, P) = E_Q[X(\omega, \cdot)] - H(Q, P),$$

which is what we wanted to show.

We complete the proof by showing that $C \le \mathcal{E}_t(\omega, 0)$ and that an optimal strategy $\vartheta^* \in \Theta$ exists. Let ϑ^* be the as in Lemma 4.6, that is, such that

$$\mathcal{E}_s(\omega, (\vartheta^* \cdot S)_t^s(\omega)) = \sup_{P \in \mathcal{P}_s(\omega)} \log E_P[\exp(\mathcal{E}_{s+1}(\omega \otimes_s \cdot, (\vartheta^* \cdot S)_t^{s+1}(\omega, \cdot))] \tag{4.9}$$

for all $t \le s \le T - 1$. Then ϑ^* is optimal and $C \le \mathcal{E}_t(\omega, 0)$. Indeed, let $P = P_t \otimes \cdots \otimes P_{T-1} \in \mathcal{P}_t^T(\omega)$ and fix some $t \le s \le T - 1$. Then it follows from (4.9) that

$$\begin{aligned}
&\log E_P[\exp(\mathcal{E}_s(\omega \otimes_t \cdot, (\vartheta^* \cdot S)_t^s(\omega, \cdot))] \\
&\quad = \log E_{P_t \otimes \cdots \otimes P_{s-1}}[\exp(\mathcal{E}_s(\omega \otimes_t \cdot, (\vartheta^* \cdot S)_t^s(\omega, \cdot)))] \\
&\quad \ge \log E_{P_t \otimes \cdots \otimes P_{s-1}}[\exp(\log E_{P_s(\cdot)}[\exp(\mathcal{E}_{s+1}(\omega \otimes_t \cdot, (\vartheta^* \cdot S)_t^{s+1}(\omega, \cdot)))])] \\
&\quad = \log E_P[\exp(\mathcal{E}_{s+1}(\omega \otimes_t \cdot, (\vartheta^* \cdot S)_t^{s+1}(\omega, \cdot)))],
\end{aligned}$$

and an iteration yields

$$\begin{aligned}
\mathcal{E}_t(\omega, 0) &= \log E_P[\exp(\mathcal{E}_t(\omega, (\vartheta^* \cdot S)_t^t(\omega)))] \\
&\ge \log E_P[\exp(\mathcal{E}_T(\omega \otimes_t \cdot, (\vartheta^* \cdot S)_t^T(\omega, \cdot)))] \\
&= \log E_P[\exp(X(\omega, \cdot) + (\vartheta^* \cdot S)_t^T(\omega, \cdot))].
\end{aligned}$$

As $P \in \mathcal{P}_t^T(\omega)$ was arbitrary, it holds $C \le \mathcal{E}_t(\omega, x)$ and since $\vartheta^* \in \Theta$, it follows from the previously shown inequality that ϑ^* is optimal. $\square$

4.2. *The case with options.* Fix some function $Y\colon \Omega \to [0, +\infty)$ such that $-Y$ is upper semianalytic and recall that $\mathcal{M}(Y) := \{Q \in \mathcal{M} : E_Q[Y] < +\infty\}$ and $\mathcal{M}_g(Y) := \{Q \in \mathcal{M}_g : E_Q[Y] < +\infty\}$, where $\mathcal{M}$ and $\mathcal{M}_g$ where defined in Section 2. Moreover, fix some Borel function $Z\colon \Omega \to \mathbb{R}$.

We first claim that for every upper semianalytic function $X\colon \Omega \to \mathbb{R}$ bounded by Z, that is, $|X| \le Z$, one has

$$\begin{aligned}\inf\{m \in \mathbb{R} : m + (\vartheta \cdot S)_0^T \ge X\mathcal{P}\text{-q.s. for some } \vartheta \in \Theta\} \\ = \sup_{Q \in \mathcal{M}(Y)} E_Q[X]\end{aligned} \tag{4.10}$$

in case of no options and, if $|g^i| \le Z$ for $1 \le i \le e$, then

$$0 \in \operatorname{ri}\{E_Q[g^e] : Q \in \mathcal{M}_{\hat{g}}(Y)\}, \tag{4.11}$$

where $\hat{g} := (g^1, \dots, g^{e-1})$ and also

$$\inf\left\{m \in \mathbb{R} : \begin{array}{c} m + (\vartheta \cdot S)_0^T + \alpha g \ge X\mathcal{P}\text{-q.s.} \\ \text{for some } (\vartheta, \alpha) \in \Theta \times \mathbb{R}^e \end{array}\right\} = \sup_{Q \in \mathcal{M}_g(Y)} E_Q[X]. \tag{4.12}$$

All these claims are proven in [10] if one relaxes $\mathcal{M}$ in the sense that the relative entropy does not need to be finite. In fact, Bouchard and Nutz deduce (4.12) from (4.11), and (4.11) from (4.10); see Theorem 4.9 as well as equation (5.1) and Theorem 5.1 in [10], respectively. The same can be done here (with the exact same arguments as in [10]), so we shall only give a (sketch of a) proof for (4.10). Consider first the one-period case and define

$$\mathcal{C}' := \{Q \in \mathfrak{P}(\Omega) : E_Q[|\Delta S| + Y] < +\infty \text{ and } Q \ll P \text{ for some } P \in \mathcal{P}\},$$

and $\mathcal{M}' := \{Q \in \mathcal{C} : E_Q[\Delta S] = 0\}$. Then the following superhedging duality

$$\inf\{m \in \mathbb{R} : m + h\Delta S \ge X\mathcal{P}\text{-q.s. for some } h \in \mathbb{R}^d\} = \sup_{Q \in \mathcal{M}'} E_Q[X]$$

(see [10], Theorem 3.4) is a consequence of the fact that $0 \in \operatorname{ri}\{E_Q[\Delta S] : Q \in \mathcal{C}'\}$; see Lemma 3.5 and Lemma 3.6 in [10]. However, since

$$0 \in \operatorname{ri}\{E_Q[g] : Q \in \mathcal{C}\} \qquad \text{for } \mathcal{C} = \{Q \in \mathcal{C} : H(Q, \mathcal{P}) < +\infty\}$$

by Lemma 3.5, the same arguments as for [10], Theorem 3.4, show that

$$\inf\{m \in \mathbb{R} : m + h\Delta S \ge X\mathcal{P}\text{-q.s. for some } h \in \mathbb{R}^d\} = \sup_{Q \in \mathcal{M}(Y)} E_Q[X],$$

in particular $\sup_{Q \in \mathcal{M}'} E_Q[X] = \sup_{Q \in \mathcal{M}(Y)} E_Q[X]$. For the transition to the multiperiod case define recursively $m_T := m'_T := X$ and

$$m'_t(\omega) := \sup_{Q \in \mathcal{M}'_t(\omega)} E_Q[m'_{t+1}(\omega, \cdot)] \quad \text{and} \quad m_t(\omega) = \sup_{Q \in \mathcal{M}_t^Z(\omega)} E_Q[m_{t+1}(\omega, \cdot)],$$

for $0 \le t \le T - 1$ and $\omega \in \Omega_t$, where

$$\mathcal{M}'_t(\omega) := \{Q \in \mathfrak{P}(\Omega_1) : Q \ll P \text{ for some } P \in \mathcal{P}_t(\omega) \text{ and } E_Q[\Delta S_{t+1}(\omega, \cdot)] = 0\},$$

$$\mathcal{M}_t^Z(\omega) := \{Q \in \mathcal{M}'_t(\omega) : E_Q[Z] + E_Q[m_{t+1}(\omega, \cdot)^-] + H(Q, \mathcal{P}_t(\omega)) < +\infty\}$$

and $Z\colon \Omega_1 \to [0,+\infty)$ an arbitrary universally measurable function. A backward induction shows that $m_t = m'_t$ for each t. Moreover, following the exact same arguments as in the part of the proof for Theorem 4.1 which focuses on duality, one can show that $m_t(\omega) = \sup_{Q\in\mathcal{M}_t^T(\omega)} E_Q[X(\omega,\cdot)]$ where we recall

$$\mathcal{M}_t^T(\omega) = \left\{ Q \in \mathfrak{P}(\Omega_{T-t}) : \begin{array}{l} \big(S_s(\omega,\cdot)\big)_{t\le s\le T} \text{ is a } Q\text{-martingale and} \\ E_Q\big[X(\omega,\cdot)^- + Y(\omega,\cdot)\big] + H\big(Q,\mathcal{P}_t^T(\omega)\big) < +\infty \end{array} \right\}$$

so that $\mathcal{M}_0^T = \mathcal{M}(Y)$. Since it is shown in (or rather within the proof of) [10], Lemma 4.13, that

$$\inf\big\{m\in\mathbb{R} : m + (\vartheta\cdot S)_0^T \ge X\mathcal{P}\text{-q.s. for some } \vartheta\in\Theta\big\} = m'_0,$$

the claim follows from $m'_0 = m_0 = \sup_{Q\in\mathcal{M}(Y)} E_Q[X]$.

PROOF OF THEOREM 2.2. The proof is an induction over e. For $e=0$, the statement is a special case of Theorem 4.1, so assume that both claims (duality and existence) are true for $e-1\ge 0$. By assumption, there is a Borel function Z such that $|X| + |g^i| \le Z$ for every $1\le i\le e$. Using the induction hypothesis, it follows that

$$\inf_{(\vartheta,\alpha)\in\Theta\times\mathbb{R}^e} \sup_{P\in\mathcal{P}} \log E_P\big[\exp\big(X + (\vartheta\cdot S)_0^T + \alpha g\big)\big] \tag{4.13}$$

$$\begin{aligned} &= \inf_{\beta\in\mathbb{R}} \min_{(\vartheta,\hat\alpha)\in\Theta\times\mathbb{R}^{e-1}} \sup_{P\in\mathcal{P}} \log E_P\big[\exp\big(X + (\vartheta\cdot S)_0^T + \hat\alpha\hat g + \beta g^e\big)\big] \\ &= \inf_{\beta\in\mathbb{R}} \sup_{Q\in\mathcal{M}_{\hat g}} \big(E_Q[X] + \beta E_Q\big[g^e\big] - H(Q,\mathcal{P})\big) = \inf_{\beta\in\mathbb{R}} \sup_{Q\in\mathcal{M}_{\hat g}} J(Q,\beta), \end{aligned} \tag{4.14}$$

where $\hat g = (g^1,\dots,g^{e-1})$ and

$$J\colon \mathcal{M}_{\hat g}\times\mathbb{R}\to\mathbb{R}, \qquad (Q,\beta)\mapsto E_Q[X] + \beta E_Q\big[g^e\big] - H(Q,\mathcal{P}).$$

It is already shown that $0\in \operatorname{ri}\{E_Q[g^e] : Q\in\mathcal{M}_{\hat g}\}$ (see (4.11)), which can be used exactly as in the proof of Theorem 3.1 to prove that

$$\inf_{|\beta|\le n} \sup_{Q\in\mathcal{M}_{\hat g}} J(Q,\beta) = \inf_{\beta\in\mathbb{R}} \sup_{Q\in\mathcal{M}_{\hat g}} J(Q,\beta) = \sup_{Q\in\mathcal{M}_{\hat g}} \inf_{\beta\in\mathbb{R}} J(Q,\beta) \tag{4.15}$$

for some $n\in\mathbb{N}$; see (3.5) for the first, and the text below (3.4) for the second equality. Hence

$$\begin{aligned} &\inf_{(\vartheta,\alpha)\in\Theta\times\mathbb{R}^e} \sup_{P\in\mathcal{P}} \log E_P\big[\exp\big(X + (\vartheta\cdot S)_0^T + \alpha g\big)\big] \\ &\quad = \inf_{\beta\in\mathbb{R}} \sup_{Q\in\mathcal{M}_{\hat g}} J(Q,\beta) = \sup_{Q\in\mathcal{M}_{\hat g}} \inf_{\beta\in\mathbb{R}} J(Q,\beta) = \sup_{Q\in\mathcal{M}_g} \big(E_Q[X] - H(Q,\mathcal{P})\big) \end{aligned}$$

showing that duality holds. The first equality in (4.15) together with the lower semicontinuity of $\sup_{Q\in\mathcal{M}_{\hat{g}}} J(Q,\cdot)$ imply that there is some $\beta^*\in\mathbb{R}$ such that

$$\sup_{Q\in\mathcal{M}_{\hat{g}}} J(Q,\beta^*) = \inf_{\beta\in\mathbb{R}} \sup_{Q\in\mathcal{M}_{\hat{g}}} J(Q,\beta).$$

For this β^*, the induction hypotheses (4.14) guarantees the existence of an optional strategy $(\vartheta^*,\hat{\alpha}^*)\in\Theta\times\mathbb{R}^{e-1}$ showing that $(\vartheta^*,\alpha^*)\in\Theta\times\mathbb{R}^e$ is optimal for (4.13), where $\alpha^* := (\hat{\alpha}^*,\beta^*)$. This completes the proof. □

PROOF OF THEOREM 2.4. Since Θ and $\mathbb{R}^e$ are vector spaces, it follows from Theorem 2.2 that

$$\begin{aligned}\pi_\gamma(X) &= \inf_{(\vartheta,\alpha)\in\Theta\times\mathbb{R}^e} \sup_{P\in\mathcal{P}} \frac{1}{\gamma}\log E_P\big[\exp\big(\gamma X + (\vartheta\cdot S)_0^T + \alpha g\big)\big] \\ &= \frac{1}{\gamma}\sup_{Q\in\mathcal{M}_g}\big(E_Q[\gamma X] - H(Q,\mathcal{P})\big) = \sup_{Q\in\mathcal{M}_g}\Big(E_Q[X] - \frac{1}{\gamma}H(Q,\mathcal{P})\Big).\end{aligned}$$

This formula implies both that π_γ is increasing in γ and, by interchanging the suprema over γ and Q, that $\sup_\gamma \pi_\gamma(X) = \sup_{Q\in\mathcal{M}_g} E_Q[X]$. The latter term coincides by (4.12) with $\pi(X)$, hence the proof is complete. □

APPENDIX A: TECHNICAL PROOFS

We start by proving Remark 2.3, Remark 2.5, and the statements of Section 2.3.

PROOF OF REMARK 2.3. 1) Let $T=d=1$, $\Omega=\mathbb{R}$, $S_0=0$, $S_1(\omega)=\omega$, and define $\mathcal{P}=\operatorname{conv}\{\delta_x : x\in[0,1]\}$ so that NA($\mathcal{P}$) fails. Then in (2.3) the left-hand side is always larger or equal than $X(0)$, and the right-hand side equals $X(0)$ since $\mathcal{M}=\{\delta_0\}$. For the choice $X=-1_{\{0\}}$, a short computation yields that the left-hand side actually equals 0, showing that there is a gap.

2) Let again $T=d=1$, $\Omega=\mathbb{R}$, $S_0=0$, $S_1(\omega)=\omega$ and define $\mathcal{P}=\operatorname{conv}\{\delta_{-1},\delta_x : x\in(0,1]\}$. Then NA($\mathcal{P}$) holds true and every martingale measure Q with $H(Q,\mathcal{P})<+\infty$ satisfies $Q(\{-1\})>0$. In particular, for $X:=-\infty 1_{\{-1\}}$ the right-hand side of (2.3) equals $-\infty$ while the right-hand side satisfies

$$\inf_{h\in\mathbb{R}}\sup_{P\in\mathcal{P}}\log E_P\big[\exp(X+h\Delta S)\big] \geq \inf_{h\in\mathbb{R}}\lim_{x\downarrow 0}\log\frac{\exp(-\infty)+\exp(hx)}{2} = \log\frac{1}{2}$$

as $(\delta_{-1}+\delta_x)/2\in\mathcal{P}$ for every $x\in(0,1]$. To see that an optimal strategy $h\in\mathbb{R}$ needs not to exists, take the same X but let $\mathcal{P}=\{(\delta_{-1}+\delta_1)/2\}$. □

PROOF OF REMARK 2.5. We claim that for any probability P satisfying the classical no-arbitrage, it is possible to construct $\mathcal{P}_t$ such that $\mathcal{P}=\{P\}$ and

$\mathrm{NA}(\mathcal{P}_t(\omega))$ holds for every t and $\omega \in \Omega_t$ and only sketch the proof. Write $P = P_0 \otimes \cdots \otimes P_{T-1}$ for the kernels P_t from Remark B.2 and define $N_t := \{\omega \in \Omega_t : \mathrm{NA}(P_t(\omega)) \text{ fails}\}$. Then it holds

$$N_t = \pi \left\{ (\omega, h) \in \Omega_t \times \mathbb{R}^d : \begin{array}{l} P_t(\omega)\big(h \Delta S_{t+1}(\omega, \cdot) \geq 0\big) = 1 \text{ and} \\ P_t(\omega)\big(h \Delta S_{t+1}(\omega, \cdot) > 0\big) > 0 \end{array} \right\}$$

and by the classical fundamental theorem of asset pricing

$$N_t^c = \pi\big\{(\omega, Q) \in \Omega_t \times \mathfrak{P}(\Omega_1) : E_Q\big[\Delta S_{t+1}(\omega, \cdot)\big] = 0 \text{ and } Q \sim P_t(\omega)\big\}.$$

In both cases, π denotes the projection onto Ω_t. It can be shown that both sets, which the projection acts on, are Borel. Thus N_t and N_t^c are analytic sets. Now define $\mathcal{P}_t(\omega) = \{P_t(\omega)\}$ if $\omega \in N_t^c$ and $\mathcal{P}_t(\omega) = \{\delta_{S_t(\omega)}\}$ else. Then $\mathcal{P}_t$ has analytic graph and since N_t is a zero set under P, it follows that $\mathcal{P} = \{P\}$. $\square$

PROOFS FOR SECTION 2.3. (a) The graphs of (2.5), (2.9) and (2.10) are Borel: For $\mathcal{P}_t$ defined by (2.5), notice that

$$g : \Omega_t \times \mathfrak{P}(\Omega_1) \to [0, +\infty], \qquad (\omega, R) \mapsto \mathrm{dist}\big(R, P_t(\omega)\big)/\varepsilon_t(\omega)$$

is Borel, hence $\operatorname{graph} \mathcal{P}_t = \{g \leq 1\}$ is Borel and, therefore, analytic. The proofs for (2.9) and (2.10) are analogue.

(b) If $\mathcal{R}_t$ has analytic graph, then so do $\hat{\mathcal{P}}_t$ and $\mathcal{P}_t$: Define

$$g : \Omega_t \times \mathfrak{P}(\Omega_1) \to \Omega_t \times \mathfrak{P}(\Omega_1), \qquad (\omega, P) \mapsto \big(\omega, P \circ f_t(\omega, \cdot)^{-1}\big)$$

and notice that g is Borel by Lemma B.1 and [7], Proposition 7.26. Therefore, $\operatorname{graph} \hat{\mathcal{P}}_t = g(\operatorname{graph} \mathcal{R}_t)$ is an analytic set, as the image of such set under a Borel function. As for $\mathcal{P}_t$, define the Borel function

$$\begin{aligned} g_n : \big((\Omega_t \times \mathfrak{P}(\Omega_1))^n \cap \Delta_n\big) \times C_n &\to \Omega_t \times \mathfrak{P}(\Omega_1), \\ \big((\omega^i, P^i)_i, \lambda\big) &\mapsto \big(\omega^1, \lambda_1 P^1 + \cdots + \lambda_n P^n\big) \end{aligned}$$

for every $n \in \mathbb{N}$, where

$$\begin{aligned} \Delta_n &:= \big\{(\omega^i, P^i)_i \in (\Omega_t \times \mathfrak{P}(\Omega_1))^n : \omega^1 = \omega^i \text{ for } 1 \leq i \leq n\big\}, \\ C_n &:= \big\{\lambda \in [0, +\infty) : \lambda_1 + \cdots + \lambda_n = 1\big\}. \end{aligned}$$

Therefore, as the countable union of the images under Borel functions of analytic sets,

$$\operatorname{graph} \mathcal{P}_t = \bigcup_n g_n\big((\operatorname{graph} \hat{\mathcal{P}}_t)^n \cap \Delta_n\big) \times C_n))$$

is again an analytic set.

(c) On the no-arbitrage condition. We only prove the claim for the Wasserstein distance of order p, that is, $\mathcal{R}_t$ given by (2.9), the proof for $\mathcal{R}_t$ given by (2.10)

works similarly. Fix $\omega \in \Omega_t$ and let $h \in \mathbb{R}$ such that $h\Delta S_{t+1}(\omega, \cdot) \geq 0$ $\mathcal{P}_t(\omega)$-q.s. If $f_t(\omega, \mathbb{R}) = \{S_t(\omega)\}$, then trivially $h\Delta S_{t+1}(\omega, \cdot) = 0$ $\mathcal{P}_t(\omega)$-q.s. Otherwise, there are $y^{\pm} \in \mathbb{R}$ such that $\pm f(\omega, y^{\pm}) > 0$ by assumption. Now define $R^{\pm} := \lambda^{\pm}\delta_{y^{\pm}} + (1 - \lambda^{\pm}) \operatorname{law} X$, where $\lambda^{\pm} := 1 \wedge 1/(\operatorname{dist}(\delta_{y^{\pm}}, \operatorname{law} X)\varepsilon_t(\omega))$ is strictly positive since X has finite pth moment. By convexity,

$$\operatorname{dist}\bigl(R^{\pm}, \operatorname{law} X\bigr) \leq \lambda^{\pm} \operatorname{dist}(\delta_{\pm x}, \operatorname{law} X) + \bigl(1 - \lambda^{\pm}\bigr) \operatorname{dist}(\operatorname{law} X, \operatorname{law} X) \leq \varepsilon_t(\omega)$$

so that $R^{\pm} \in \mathcal{R}_t(\omega)$. Hence $hf(\omega, y^{\pm}) \geq 0$, which in turn implies $h = 0$.

(d) The binomial and Black–Scholes model. A computation as in a) shows that the graph of Φ_t defined by

$$\Phi_t(\omega) := \bigl\{(q, a, b) : p \in \bigl[\underline{p}_t(\omega), \overline{p}_t(\omega)\bigr], a \in \bigl[\underline{a}_t(\omega), \overline{a}_t(\omega)\bigr], b \in \bigl[\underline{b}_t(\omega), \overline{b}_t(\omega)\bigr]\bigr\}$$

is an analytic set. Since

$$g\colon \Omega_t \times \mathbb{R}^3 \to \Omega_t \times \mathfrak{P}(\Omega_1), \qquad (\omega, p, a, b) \mapsto \bigl(\omega, p\delta_a + (1 - p)\delta_b\bigr)$$

is continuous, it follows that $\operatorname{graph} \mathcal{R}_t = g(\operatorname{graph} \Phi_t)$ is an analytic set. The proof for the Black–Scholes model works similarly. $\square$

The following lemma is related to [21], Lemma 3.29, where X is assumed to be bounded.

LEMMA A.1. *Let $X\colon \Omega \to \mathbb{R}$ be measurable and let $P \in \mathfrak{P}(\Omega)$. Then one has*

$$\log E_P\bigl[\exp(X)\bigr] = \sup_{Q \in \mathcal{A}} \bigl(E_Q[X] - H(Q, P)\bigr),$$

where $\mathcal{A} := \{Q \in \mathfrak{P}(\Omega) : H(Q, P) + E_Q[X^-] < +\infty\}$.

PROOF. For each natural number n, define Q_n by

$$\frac{dQ_n}{dP} := \frac{\exp(X \wedge n)}{E_P[\exp(X \wedge n)]}.$$

Then Q_n is equivalent to P and since $\exp(X \wedge n)X^- \leq 1$, it follows that X^- is integrable with respect to Q_n. By equivalence of P and Q_n, one can write

$$\frac{dQ}{dP} = \frac{dQ}{dQ_n}\frac{dQ_n}{dP} \qquad \text{for any } Q \in \mathcal{A}.$$

Applying Jensen's inequality to the convex function $[0, \infty) \to [-1, \infty), x \mapsto x \log x$ with "$x = dQ/dQ_n$" yields

$$\begin{aligned} H(Q, P) &= E_{Q_n}\left[\frac{dQ}{dQ_n} \log \frac{dQ}{dQ_n}\right] + E_Q\left[\log \frac{dQ_n}{dP}\right] \\ &\geq E_Q\left[\log \frac{dQ_n}{dP}\right] = E_Q[X \wedge n] - \log E_P\bigl[\exp(X \wedge n)\bigr] \end{aligned}$$

with equality if (and only if) $Q = Q_n$. Since the right-hand side is finite for $Q = Q_n$, it follows that $H(Q_n, P) < +\infty$ and, therefore, $Q_n \in \mathcal{A}$. Rearranging the terms which appear in the inequality above yields

$$\log E_P\big[\exp(X \wedge n)\big] \geq E_Q[X \wedge n] - H(Q, P)$$

for all $Q \in \mathcal{A}$ with equality for $Q = Q_n \in \mathcal{A}$. This shows the claim if X were bounded. The general case follows by letting n tend to infinity. Indeed, since the set $\mathcal{A}$ does not depend on n, we can interchange two suprema and conclude that

$$\log E_P\big[\exp(X)\big] = \sup_n \sup_{Q \in \mathcal{A}} \big(E_Q[X \wedge n] - H(Q, P)\big) = \sup_{Q \in \mathcal{A}} \big(E_Q[X] - H(Q, P)\big).$$

The use of the monotone convergence theorem in the last step was justified because $E_Q[X^-] < +\infty$ for every $Q \in \mathcal{A}$. $\square$

LEMMA A.2. *Let V and W be two Polish spaces and $P, Q \in \mathfrak{P}(V \times W)$ with representation $P = \mu \otimes K$, $Q = \mu' \otimes K'$ for measures $\mu, \mu' \in \mathfrak{P}(V)$ and universally measurable kernels $K, K' \colon V \to \mathfrak{P}(W)$. Then one has*

$$Q \ll P \quad \textit{if and only if} \quad \mu' \ll \mu \quad \textit{and} \quad K'(v) \ll K(v)$$

for μ'-almost every v.

PROOF. If $\mu' \ll \mu$ and $K'(v) \ll K(v)$ for μ'-almost every v, it follows from the definition that $Q \ll P$. Indeed, for any Borel set $A \subset V \times W$ such that $0 = P(A) = E_{\mu(dv)}[K(v)(A_v)]$, it holds $Q(A) = E_{\mu'(dv)}[K'(v)(A_v)] = 0$. Here, $A_v := \{w \in W : (v, w) \in A\}$.

The other direction needs more work. The idea is to show that the generalized Radon–Nikodym derivative (see, e.g., [21], Theorem A.13) is measurable with respect to the kernels. Assume that $Q \ll P$ and first notice that $\mu' \ll \mu$. If this were not the case, then $\mu'(A) > 0$ while $\mu(A) = 0$ for some Borel set $A \subset V$ which implies $Q(A \times W) = \mu'(A) > 0$ but $P(A \times W) = 0$. We proceed to show the absolute continuity of the kernels. Notice that the mapping

$$\mathfrak{P}(W) \times \mathfrak{P}(W) \times W, \qquad \big(R', R, w\big) \mapsto \frac{dR'}{dR}(w)$$

can be shown to be Borel, where dR'/dR denotes the Radon–Nikodym derivative of the absolutely continuous part of R' with respect to R. This result, due to Doob, can be found, for example, in [16], Theorem V.58, and the subsequent remark. Hence

$$V \times W \to \mathfrak{P}(W) \times \mathfrak{P}(W) \times W, \qquad (v, w) \mapsto \big(K(v), K'(v), w\big)$$

is universally measurable. Thus, since $(R, R') \mapsto (R + R')/2$ is Borel, it follows that $Z \colon V \times W \to [0, +\infty]$,

$$Z(v, w) := \frac{dK'(v)}{d(K(v) + K'(v))/2}(w)\bigg(\frac{dK(v)}{d(K(v) + K'(v))/2}(w)\bigg)^{-1}$$

is universally measurable, with the convention $x/0 := +\infty$ for all $x \geq 0$. A straightforward computation as in [21], Theorem A.13, yields $K(v)(Z(v,\cdot) = +\infty) = 0$,

$$K'(v)(B) = K'(v)\big(B \cap \{Z(v,\cdot) = +\infty\}\big) + E_{K(v)}\big[1_B Z(v,\cdot)\big]$$

for any universally measurable set $B \subset W$, and as a consequence that

$$K'(v) \ll K(v) \quad \text{if and only if} \quad K'(v)\big(Z(v,\cdot) = +\infty\big) = 0.$$

Heading for a contradiction, assume that the set of all such v has not full μ' measure and define the universally measurable set

$$A := \{(v,w) : Z(v,w) = +\infty\}.$$

Then

$$Q(A) = E_{\mu'(dv)}\big[K'(v)\big(Z(v,\cdot) = +\infty\big)\big] > 0,$$

while on the other hand $P(A) = E_{\mu(dv)}[K(v)(Z(v,\cdot) = +\infty)] = 0$. This contradicts the absolute continuity of Q with respect to P. $\square$

PROOF OF LEMMA 4.3. The goal is to show that

$$H(Q,P) = \sum_{s=t}^{T-1} E_Q\big[H\big(Q_s(\cdot), P_s(\cdot)\big)\big]. \tag{A.1}$$

(a) We first comment on the measurability of terms appearing later. Fix some $t \leq s \leq T-1$ and notice as in the proof of Lemma A.2 that

$$\Omega_{s-t} \to \mathfrak{P}(\Omega_1) \times \mathfrak{P}(\Omega_1), \qquad \bar{\omega} \mapsto \big(Q_s(\bar{\omega}), P_s(\bar{\omega})\big)$$

is universally measurable. Since the entropy H is Borel by Lemma 4.2, one can check that

$$\Omega_{s-t} \to [0,+\infty], \qquad \bar{\omega} \mapsto H\big(Q_s(\bar{\omega}), P_s(\bar{\omega})\big)$$

is universally measurable. Similarly, using Doob's result on the measurability of the Radon–Nikodym derivative as in the proof of Lemma A.2, it follows that

$$\Omega_{s-t} \times \Omega_1 \to [0,+\infty], \qquad (\bar{\omega}, \omega') \mapsto \frac{dQ_s(\bar{\omega})}{dP_s(\bar{\omega})}(\omega')$$

is universally measurable. Moreover, by [7], Lemma 7.29, the same holds true if $\bar{\omega} \in \Omega_{s-t}$ is fixed in the above mapping, and the latter is considered as a function of ω'.

(b) A direct application of Lemma A.2 shows that

$$Q \ll P \quad \text{if and only if} \quad Q_s \ll P_s, \qquad Q_t \otimes \cdots \otimes Q_{s-1}\text{-almost surely}$$

for all $t \le s \le T-1$, where in case $s = t$ the above should be understood as $Q_t \ll P_t$. This implies that whenever Q is not absolutely continuous with respect to P, then both sides in (A.1) are equal to $+\infty$. Hence we may assume that $Q \ll P$. Then dQ/dP can be expressed as the product of $dQ_s(\cdot)/dP_s(\cdot)$, where s ranges from t to $T-1$. Therefore, for any $t \le s \le T-1$, it follows that

$$E_Q\left[\left(\log \frac{dQ_s(\cdot)}{dP_s(\cdot)}\right)^-\right] = E_{Q_t \otimes \cdots \otimes Q_{s-1}}\left[E_{P_s(\cdot)}\left[\frac{dQ_s(\cdot)}{dP_s(\cdot)}\left(\log \frac{dQ_s(\cdot)}{dP_s(\cdot)}\right)^-\right]\right] \le 1,$$

where the last inequality holds since $x(\log x)^- \le 1$ for all $x \ge 0$. By integrability, the same steps may be repeated without the negative parts so that

$$H(Q, P) = E_Q\left[\log \frac{dQ}{dP}\right] = \sum_{s=t}^{T-1} E_Q\left[\log \frac{dQ_s(\cdot)}{dP_s(\cdot)}\right] = \sum_{s=t}^{T-1} E_Q\big[H\big(Q_s(\cdot), P_s(\cdot)\big)\big]$$

as claimed. $\square$

PROOF OF LEMMA 4.6. First we claim that for any $\vartheta \in \Theta$ and $0 \le t \le T-1$, there exists a universally measurable mapping $\hat{h}_t \colon \Omega_t \to \mathbb{R}^d$ such that

$$\begin{aligned} &\mathcal{E}_t\big(\omega, x + (\vartheta \cdot S)_0^t(\omega)\big) \\ &\quad = \sup_{P \in \mathcal{P}_t(\omega)} \log E_P\big[\exp\big(\mathcal{E}_{t+1}\big(\omega \otimes_t \cdot, x + (\vartheta \cdot S)_0^t(\omega) \\ &\qquad + \hat{h}_t(\omega)\Delta S_{t+1}(\omega, \cdot)\big)\big)\big] \end{aligned} \tag{A.2}$$

for all $\omega \in \Omega_t$. To that end, fix some $\vartheta \in \Theta$, $0 \le t \le T-1$, and recall that $\mathcal{F}_t$ was defined as the universal completion of the Borel σ-field on Ω_t. From the first part of the proof of Theorem 4.1, we already know that $\mathcal{E}_t(\omega, x) = \mathcal{D}_t(\omega) + x$ for all $\omega \in \Omega_t$ and $x \in \mathbb{R}$ and that $\mathcal{D}_t$ is upper semianalytic, in particular $\mathcal{F}_t$-measurable. This implies that $\mathcal{E}_t$ is $\mathcal{F}_t \otimes \mathcal{B}(\mathbb{R})$-measurable. Define the function

$$\phi(\omega, x, h) := \sup_{P \in \mathcal{P}_t(\omega)} \log E_P\big[\exp\big(\mathcal{D}_{t+1}(\omega \otimes_t \cdot) + x + h\Delta S_{t+1}(\omega, \cdot)\big)\big].$$

For fixed x and h, it follows from [7], Proposition 7.47 (as in the first first part of the proof of Theorem 4.1) that $\phi(\cdot, x, h)$ is upper semianalytic. Moreover, for fixed ω, an application of Fatou's lemma (as in part (b) of the proof of Theorem 3.1) shows that $\phi(\omega, \cdot, \cdot)$ is lower semicontinuous. Therefore, we can conclude by [10], Lemma 4.12, that ϕ is $\mathcal{F}_t \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^d)$-measurable. Now fix $x \in \mathbb{R}$ and define the set-valued mapping

$$\Phi(\omega) := \big\{h \in \mathbb{R}^d : \phi\big(\omega, x + (\vartheta \cdot S)_0^t(\omega), h\big) = \mathcal{E}_t\big(\omega, x + (\vartheta \cdot S)_0^t(\omega)\big)\big\}.$$

By Theorem 3.1, it holds $\Phi(\omega) \neq \varnothing$ and by the above its graph is in $\mathcal{F}_t \otimes \mathcal{B}(\mathbb{R}^d)$. Hence it follows by Theorem 5.5 in [30] or rather the corollary and scholim after, that Φ admits an $\mathcal{F}_t$-measurable selector $\hat{h}_t$.

To conclude the proof of the lemma, define $\vartheta_s^* := 0$ for $s \le t$, let $\hat{h}_t$ be an optimal strategy for time t and define $\vartheta_{t+1}^* := \hat{h}_t$. By the above, there is a universally measurable mapping $\hat{h}_{t+1}\colon \Omega_{t+1} \to \mathbb{R}$ such that (A.2) holds for $t+1$. Define $\vartheta_{t+2}^* := \hat{h}_{t+1}$. Proceeding in a recursive matter until $t = T$, we construct $\vartheta^* \in \Theta$ which fulfills the requirements of the lemma. $\square$

APPENDIX B: ANALYTIC SETS

We briefly recall the used terminology and give a short overview on the theory of analytic sets; for more details see, for example, Chapter 7 in the book of Bertsekas and Shreve [7]. Throughout, fix two Polish spaces V and W. A subset of a Polish space is called analytic, if it is the image of a Borel set of another Polish space under a Borel function. Similarly, a function $f\colon V \to [-\infty, +\infty]$ is upper semianalytic, if $\{f \ge c\} \subset V$ is an analytic set for every real number c. Further define $\mathcal{B}(V)$ to be the Borel σ-field on V and $\mathfrak{P}(V)$ to be the set of all probability measures on $\mathcal{B}(V)$. The set $\mathfrak{P}(V)$ is endowed with the weak topology induced by all continuous bounded functions, that is, $\sigma(\mathfrak{P}(V), C_b(V))$. Then $\mathfrak{P}(V)$ becomes a Polish space itself. The set of universally measurable subsets of V is defined as $\bigcap\{\mathcal{B}(V)^P : P \in \mathfrak{P}(V)\}$, where $\mathcal{B}(V)^P$ is the completion of $\mathcal{B}(V)$ with respect to the probability P. A function $f\colon V \to W$ is said to be universally measurable, if $\{f \in B\}$ is universally measurable for every $B \in \mathcal{B}(W)$. It follows from the definition that every Borel set is analytic, and from Lusin's theorem (see [7], Proposition 7.42) that every analytic set is universally measurable. The same of course holds true if we replace sets by functions in the previous sentence. A set-valued function $\Psi\colon V \to W$ is said to have analytic graph, if

$$\operatorname{graph}\Psi := \big\{(v, w) : v \in V, w \in \Psi(v)\big\} \subset V \times W$$

is an analytic set. Finally, given a set $\mathcal{P} \subset \mathfrak{P}(V)$, a set $N \subset V$ is said to be $\mathcal{P}$-polar if $P(N) = 0$ for all $P \in \mathcal{P}$. Similarly, a property is said to hold $\mathcal{P}$-quasi surely (q.s. for short), if it holds outside a $\mathcal{P}$-polar set.

One can readily verify that $(v, P(dw)) \mapsto E_P[X(v, \cdot)]$ is continuous, whenever $X\colon V \times W \to \mathbb{R}$ is uniformly continuous and bounded. The following lemma generalizes this.

LEMMA B.1 ([7], Propositions 7.29/7.46/7.48). *Let $X\colon V \times \mathfrak{P}(W) \times W \to [-\infty, +\infty]$ be Borel/upper semianalytic/universally measurable. Then the mapping $V \times \mathfrak{P}(W) \to [-\infty, +\infty]$, $(v, P) \mapsto E_P[X(v, P, \cdot)]$ is Borel/upper semianalytic/universally measurable.*

PROOF. The proof is an application of Proposition 7.29/Proposition 7.46/Proposition 7.48 in [7], depending on the given measurability. Indeed, using the notation of [7], define the Borel spaces $\mathcal{X} := V \times \mathfrak{P}(W)$ and $\mathcal{Y} = W$ as well

as the Borel/upper semianalytic/universally measurable mapping $f\colon \mathcal{X} \times \mathcal{Y} \to [-\infty, +\infty]$ $f(x, y) = f(c, P, w) := X(v, P, w)$ and Borel kernel $q(dy, x) = q(dw, (v, P)) := P(dw)$. By the mentioned proposition the mapping $V \times \mathfrak{P}(W) = \mathcal{X} \to [-\infty, +\infty]$, $(v, P) = x \mapsto \int f(x, y) q(dy, x) = E_P[X(v, P, \cdot)]$ is Borel/upper semianalytic/universally measurable. $\square$

REMARK B.2. By means of the disintegration theorem, every probability $P \in \mathfrak{P}(V \times W)$ can be written as $P = \mu \otimes K$, where $\mu \in \mathfrak{P}(V)$ and $K\colon V \to \mathfrak{P}(W)$ is Borel. In fact, it is possible to construct the kernel K in a way such that the mapping

$$V \times \mathfrak{P}(V \times W) \to \mathfrak{P}(W), \qquad (v, P) \mapsto K(v)$$

is Borel; see [7], Proposition 7.27.

Acknowledgments. The author would like to thank Michael Kupper and two anonymous referees for valuable comments.

REFERENCES

[1] ACCIAIO, B., BEIGLBÖCK, M., PENKNER, F. and SCHACHERMAYER, W. (2016). A model-free version of the fundamental theorem of asset pricing and the super-replication theorem. *Math. Finance* **26** 233–251. MR3481303

[2] BACKHOFF VERAGUAS, J. D. and FONTBONA, J. (2016). Robust utility maximization without model compactness. *SIAM J. Financial Math.* **7** 70–103. MR3466196

[3] BARTL, D. (2016). Conditional nonlinear expectations. Preprint. Available at arXiv:1612.09103v2.

[4] BARTL, D., DRAPEAU, S. and TANGPI, L. (2017). Computational aspects of robust optimized cer-tainty equivalents and option pricing. *Math. Finance*. To appear.

[5] BEIGLBÖCK, M., HENRY-LABORDÈRE, P. and PENKNER, F. (2013). Model-independent bounds for option prices—A mass transport approach. *Finance Stoch.* **17** 477–501. MR3066985

[6] BEIGLBÖCK, M., NUTZ, M. and TOUZI, N. (2017). Complete duality for martingale optimal transport on the line. *Ann. Probab.* **45** 3038–3074. MR3706738

[7] BERTSEKAS, D. P. and SHREVE, S. E. (1978). *Stochastic Optimal Control: The Discrete Time Case. Mathematics in Science and Engineering* **139**. Academic Press, New York. MR0511544

[8] BLANCHARD, R. and CARASSUS, L. (2017). Convergence of utility indifference prices to the superreplication price in a multiple-priors framework. Preprint. Available at arXiv:1709.09465.

[9] BLANCHARD, R. and CARASSUS, L. (2018). Multiple-priors optimal investment in discrete time for unbounded utility function. *Ann. Appl. Probab.* **28** 1856–1892. MR3809479

[10] BOUCHARD, B. and NUTZ, M. (2015). Arbitrage and duality in nondominated discrete-time models. *Ann. Appl. Probab.* **25** 823–859. MR3313756

[11] BURZONI, M., FRITTELLI, M., HOU, Z., MAGGIS, M. and OBŁÓJ, J. (2016). Pointwise arbitrage pricing theory in discrete time. *Math. Oper. Res.* To appear.

[12] CHERIDITO, P. and KUPPER, M. (2011). Composition of time-consistent dynamic monetary risk measures in discrete time. *Int. J. Theor. Appl. Finance* **14** 137–162. MR2780781

[13] CHERIDITO, P., KUPPER, M. and TANGPI, L. (2015). Representation of increasing convex functionals with countably additive measures. Preprint. Available at arXiv:1502.05763.

[14] CHERIDITO, P., KUPPER, M. and TANGPI, L. (2017). Duality formulas for robust pricing and hedging in discrete time. *SIAM J. Financial Math*. **8** 738–765. MR3705784

[15] DELBAEN, F., GRANDITS, P., RHEINLÄNDER, T., SAMPERI, D., SCHWEIZER, M. and STRICKER, C. (2002). Exponential hedging and entropic penalties. *Math. Finance* **12** 99–123. MR1891730

[16] DELLACHERIE, C. and MEYER, P.-A. (1988). *Probabilities and Potential. C*: *Potential Theory for Discrete and Continuous Semigroups*. North-Holland, Amsterdam. MR0939365

[17] DENG, S., TAN, X. and YU, X. (2018). Utility maximization with proportional transaction costs under model uncertainty. Preprint. Available at arXiv:1805.06498.

[18] DENIS, L. and KERVAREC, M. (2013). Optimal investment under model uncertainty in non-dominated models. *SIAM J. Control Optim*. **51** 1803–1822. MR3048198

[19] DUPUIS, P. and ELLIS, R. S. (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, New York. MR1431744

[20] FÖLLMER, H. and LEUKERT, P. (1999). Quantile hedging. *Finance Stoch*. **3** 251–273. MR1842286

[21] FÖLLMER, H. and SCHIED, A. (2011). *Stochastic Finance*: *An Introduction in Discrete Time*, 3rd revised and extended ed. de Gruyter, Berlin. MR2779313

[22] FRITTELLI, M. (2000). The minimal entropy martingale measure and the valuation problem in incomplete markets. *Math. Finance* **10** 39–52. MR1743972

[23] GILBOA, I. and SCHMEIDLER, D. (1989). Maxmin expected utility with nonunique prior. *J. Math. Econom.* **18** 141–153. MR1000102

[24] GUNDEL, A. (2005). Robust utility maximization for complete and incomplete market models. *Finance Stoch*. **9** 151–176. MR2211122

[25] HOBSON, D. G. (1998). Robust hedging of the lookback option. *Finance Stoch*. **2** 329–347.

[26] HU, Y., IMKELLER, P. and MÜLLER, M. (2005). Utility maximization in incomplete markets. *Ann. Appl. Probab.* **15** 1691–1712. MR2152241

[27] JACOD, J. and SHIRYAEV, A. N. (1998). Local martingales and the fundamental asset pricing theorems in the discrete-time case. *Finance Stoch*. **2** 259–273. MR1809522

[28] KRAMKOV, D. and SCHACHERMAYER, W. (1999). The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Ann. Appl. Probab.* **9** 904–950. MR1722287

[29] LACKER, D. (2018). Liquidity, risk measures, and concentration of measure. *Mathematics of Operations* **43**.

[30] LEESE, S. J. (1978). Measurable selections and the uniformization of Souslin sets. *Amer. J. Math*. **100** 19–41. MR0507445

[31] MACCHERONI, F., MARINACCI, M. and RUSTICHINI, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica* **74** 1447–1498. MR2268407

[32] MANIA, M. and SCHWEIZER, M. (2005). Dynamic exponential utility indifference valuation. *Ann. Appl. Probab.* **15** 2113–2143. MR2152255

[33] MATOUSSI, A., POSSAMAÏ, D. and ZHOU, C. (2015). Robust utility maximization in nondominated models with 2BSDE: The uncertain volatility model. *Math. Finance* **25** 258–287. MR3321250

[34] MERTON, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *J. Econom. Theory* **3** 373–413. MR0456373

[35] NEUFELD, A. and NUTZ, M. (2018). Robust utility maximization with Lévy processes. *Math. Finance* **28** 82–105. MR3758917

[36] NEUFELD, A. and SIKIC, M. (2017). Nonconcave robust optimization under knightian uncertainty. Preprint. Available at arXiv:1711.03875.

[37] NEUFELD, A. and ŠIKIĆ, M. (2018). Robust utility maximization in discrete-time markets with friction. *SIAM J. Control Optim.* **56** 1912–1937. MR3807937

[38] NUTZ, M. (2016). Utility maximization under model uncertainty in discrete time. *Math. Finance* **26** 252–268. MR3481304

[39] NUTZ, M. and VAN HANDEL, R. (2013). Constructing sublinear expectations on path space. *Stochastic Process. Appl.* **123** 3100–3121. MR3062438

[40] OWARI, K. (2011). Robust utility maximization with unbounded random endowment. In *Advances in Mathematical Economics*. *Adv. Math. Econ.* **14** 147–181. Springer, Tokyo. MR3220855

[41] PENG, S. (2007). G-expectation, G-Brownian motion and related stochastic calculus of Itô type. In *Stochastic Analysis and Applications*. *Abel Symp.* **2** 541–567. Springer, Berlin. MR2397805

[42] QUENEZ, M.-C. (2004). Optimal portfolio in a multiple-priors model. In *Seminar on Stochastic Analysis*, *Random Fields and Applications IV*. *Progress in Probability* **58** 291–321. Birkhäuser, Basel. MR2096294

[43] SCHIED, A. (2006). Risk measures and robust optimization problems. *Stoch. Models* **22** 753–831. MR2263865

[44] SION, M. (1958). On general minimax theorems. *Pacific J. Math.* **8** 171–176. MR0097026

[45] VILLANI, C. (2009). *Optimal Transport*: *Old and New*. *Grundlehren der Mathematischen Wissenschaften* **338**. Springer, Berlin. MR2459454

UNIVERSITY OF KONSTANZ
UNIVERSITÄTSSTRASSE 10
78464 KONSTANZ
GERMANY
E-MAIL: daniel.bartl@uni-konstanz.de