Math Finan Econ (2009) 2:189–210
DOI 10.1007/s11579-009-0019-9

Representation results for law invariant time consistent functions

Michael Kupper · Walter Schachermayer

Received: 22 August 2008 / Accepted: 30 June 2009 / Published online: 1 September 2009

Abstract We show that the only dynamic risk measure which is law invariant, time consistent and relevant is the entropic one. Moreover, a real valued function c on $L^\infty(a,b)$ is normalized, strictly monotone, continuous, law invariant, time consistent and has the Fatou property if and only if it is of the form $c(X) = u^{-1} \circ \mathbb{E}[u(X)]$, where $u : (a,b) \to \mathbb{R}$ is a strictly increasing, continuous function. The proofs rely on a discrete version of the Skorohod embedding theorem.

Keywords Law invariance · Time consistency · Certainty equivalent · Dynamic risk measures · Skorohod embedding theorem

Mathematics Subject Classification (2000) 91B30 · 91B16 · 91B55

1 Introduction and main results

The theory of preferences and their numerical representations goes back to Bernoulli [4]. Axiomatic foundations have been given, among others, by Alt [1], von Neumann and

We thank Freddy Delbaen, Damir Filipović, Hans Föllmer, Fabio Maccheroni, Eberhard Mayerhofer, Alexander Mürmann, Ludger Rüschendorf, Alexander Schied, Mihai Sirbu, Josef Teichmann, Nicolas Vogelpoth and especially Patrick Cheridito as well as two anonymous referees for helpful comments and suggestions. Financial support from the Vienna Science and Technology Fund (WWTF) is gratefully acknowledged. Financial support from the Austrian Science Fund (FWF) under grant P19456, from Vienna Science and Technology Fund (WWTF) under grant MA13 and by the Christian Doppler Research Association (CDG) is gratefully acknowledged.

M. Kupper (✉)
Vienna Institute of Finance, University of Vienna and Vienna University of Economics and Business Administration, Heiligenstädter Strasse 46-48, 1190 Vienna, Austria
e-mail: mkupper@vif.ac.at

W. Schachermayer
University Vienna, Nordbergstr. 15, 1090 Vienna, Austria
e-mail: walter.schachermayer@univie.ac.at

Morgenstern [32], Savage [34], Ellsberg [15], Gilboa and Schmeidler [20] and Maccheroni et al. [30]. If $\succeq$ is a preference order (c.f. Föllmer and Schied [18, Sect. 2.1]) on the set of all probability distributions with bounded support in the interval (a, b), where $-\infty \le a < b \le \infty$, such that $\succeq$ satisfies the *independence axiom* and the *Archimedean axiom* (c.f. Föllmer and Schied [18, Sect. 2.1]) then it has an *affine numerical representation* U, which under additional continuity assumptions is a *von Neumann-Morgenstern representation* (c.f. Föllmer and Schied [18, Theorems 2.21 and 2.28]). The independence axiom is crucial for a von Neumann-Morgenstern representation. It is demonstrated by Machina [29] and others (see [29] for the references) that preferences which do not satisfy the independence axiom lead to dynamic inconsistencies, at least if the preferences are not updated in an adequate way. In this paper, we show that any numerical representation of a preference order which is defined on the space of bounded random variables and which is strictly monotone, normalized on constants, law invariant and time consistent, necessarily is (under some technical continuity conditions, see Theorem 1.4 below) a certainty equivalent of an expected utility.

Here is the formal setting for our results. Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{N}_0}, \mathbb{P}\right)$ be a standard filtered probability space, i.e., $(\Omega, \mathcal{F}, \mathbb{P})$ is isomorphic to $[0, 1]^{\mathbb{N}_0}$ equipped with its Borel sigma-algebra $\mathcal{B}([0, 1]^{\mathbb{N}_0})$ and the product of Borel measures $\lambda^{\mathbb{N}_0}$. The filtration $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$ is generated by the coordinate functions. We identify random variables which are a.s. identical. All equalities and inequalities between random variables are understood in the a.s. sense. For $-\infty \le a < b \le \infty$, we denote by $L^\infty(a, b) = L^\infty(\Omega, \mathcal{F}, \mathbb{P}; (a, b))$ the set of all random variables which are essentially bounded and take values in the open interval (a, b). $L^\infty_t(a, b)$ consists of all random variables in $L^\infty(a, b)$ which are $\mathcal{F}_t$-measurable. Throughout we will simply write L^∞ for $L^\infty(-\infty, \infty)$ and L^∞_t for $L^\infty_t(-\infty, \infty)$. For $X, Y \in L^\infty(a, b)$ we write $X \preceq_1 Y$ if Y (first order) stochastically dominates X, i.e., $\mathbb{P}[X \le m] \ge \mathbb{P}[Y \le m]$ for all $m \in \mathbb{R}$.

1.1 A representation result for law invariant, time consistent certainty equivalents

On $L^\infty(a, b)$ we consider a preference order $\succeq$ with representation $U : L^\infty(a, b) \to \mathbb{R}$ (c.f. Föllmer and Schied [18, Sect. 2.1]) such that X is preferred to Y if $U(X) \ge U(Y)$. We assume that U is *law invariant*, that is $U(X) = U(Y)$ if X and Y have the same distribution. In the literature, law invariant functions are also referred to as distribution based functions. Thus, the representation U can also be viewed as a function acting on the probability distributions with bounded support in (a, b). We further assume that there exists a *certainty equivalent* $c_0 : L^\infty(a, b) \to \mathbb{R}$ for the numerical representation U, which is implicitly defined through $U(X) = U(c_0(X))$, $X \in L^\infty(a, b)$. If in addition $U(m_1) > U(m_2)$ for all $m_1, m_2 \in (a, b)$ with $m_1 > m_2$, then the certainty equivalent is *normalized on constants*, i.e., $c_0(m) = m$ for all $m \in (a, b)$. Note that c_0 is a numerical representation for the preference order $\succeq$ and c_0 is strictly monotone exactly when U is strictly monotone.

Definition 1.1 A function $c_0 : L^\infty(a, b) \to \mathbb{R}$ is said to be

- normalized on constants if $c_0(m) = m$ for all $m \in (a, b)$;
- strictly monotone if $X \ge Y$ and $\mathbb{P}[X > Y] > 0$ imply $c_0(X) > c_0(Y)$;
- law invariant if $c_0(X) = c_0(Y)$ whenever $\text{law}(X) = \text{law}(Y)$;
- time consistent if for each $t \in \mathbb{N}_0$ there exists a mapping $c_t : L^\infty(a, b) \to L^\infty_t(a, b)$ which satisfies the local property

$$1_A X = 1_A Y \quad \text{implies} \quad 1_A c_t(X) = 1_A c_t(Y) \quad \text{for all } A \in \mathcal{F}_t,\ X, Y \in L^\infty(a, b) \quad (1)$$

and

$$c_0(X) = c_0(c_t(X)) \quad \text{for all } X \in L^\infty(a,b). \tag{2}$$

For any function $c_0 : L^\infty(a,b) \to \mathbb{R}$ which is normalized on constants and strictly monotone, there exists at most one function $c_t : L^\infty(a,b) \to L_t^\infty(a,b)$, which satisfies Eqs. 1 and 2. If there exists such c_t, one verifies that it is normalized on constants ($c_t(m) = m$ for all $m \in L_t^\infty(a,b)$) and monotone ($X \geq Y$ implies $c_t(X) \geq c_t(Y)$). In view of Eq. 2, every time consistent function $c_0 : L^\infty(a,b) \to L_t^\infty(a,b)$ restricted to $\cup_{t\in\mathbb{N}} L_t^\infty(a,b)$ can be computed by backward recursion

$$\begin{cases} c_t(X) = X & \text{if } t \geq T \\ c_t(X) = c_t\left(c_{t+1}(X)\right) & \text{if } t < T \end{cases}, \tag{3}$$

where $X \in L_T^\infty(a,b)$ for some $T \in \mathbb{N}$.

Definition 1.2 [12,26] Let E be a subset of L^∞. A function $f : E \to \mathbb{R}$ has the *Fatou property* if $f(X) \geq \limsup_{n\to\infty} f(X_n)$ for any $||\cdot||_\infty$-bounded sequence $X_n \in E$ converging to $X \in E$ in probability.

A function $f : E \to \mathbb{R}$ has the *Lebesgue property* if $f(X) = \lim_{n\to\infty} f(X_n)$, for every $||\cdot||$-bounded sequence $X_n \in E$ converging a.s. to $X \in E$.

Obviously, a function which has the Lebesgue property also has the Fatou property.

Let $b^{\varepsilon,p}$ denote a Bernoulli random variable taking the values $+\varepsilon$ and $-\varepsilon$ with probabilities p and $1-p$.

Definition 1.3 A function $f : L^\infty(a,b) \to \mathbb{R}$ satisfies the continuity *condition (C)* if for any $\varepsilon > 0$ and all $x \in (a,b)$ with $a+\varepsilon < x < b-\varepsilon$ there is $p = p(x) \in (0,1)$ such that

$$f(x + b^{\varepsilon,p}) < x. \tag{4}$$

Note that condition (C) is satisfied if f is strictly monotone and has the Fatou property. Indeed, suppose that f has the Fatou property and let $(p_n)_{n\in\mathbb{N}} \subset (0,1)$ be a sequence with $p_n \searrow 0$. Then, $x + b^{\varepsilon,p_n} \to x - \varepsilon$ in probability and the Fatou property and the strict monotonicity of f imply

$$\limsup_{n\to\infty} f(x + b^{\varepsilon,p_n}) \leq f(x-\varepsilon) < x.$$

Hence, there is $p_{n_0} \in (0,1)$ with $f(x + b^{\varepsilon,p_{n_0}}) < x$.

Our first main result can now be formulated as follows.

Theorem 1.4 *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{N}_0}, \mathbb{P})$ be a standard filtered probability space and fix numbers $-\infty \leq a < b \leq +\infty$. A function $c_0 : L^\infty(a,b) \to \mathbb{R}$ is normalized on constants, strictly monotone, $||\cdot||_\infty$-continuous, law invariant, time consistent and satisfies condition (C) if and only if*

$$c_0(X) = u^{-1} \circ \mathbb{E}\left[u(X)\right] \tag{5}$$

for a strictly increasing, continuous function $u : (a,b) \to \mathbb{R}$. In this case, the function u is uniquely defined up to affine transformations of the form $u \mapsto \alpha u + \beta$, $\alpha > 0$, $\beta \in \mathbb{R}$, and

$$c_t(X) = u^{-1} \circ \mathbb{E}\left[u(X) \mid \mathcal{F}_t\right] \quad \textit{for all } t \in \mathbb{N}. \tag{6}$$

The proof of Theorem 1.4 is postponed to the Sect. 3.

Remark 1.5 Theorem 1.4 extends to a dynamic setting the representation results on means by Nagumo [31], Kolmogorov [27] and de Finetti [16]. These representation results give necessary and sufficient conditions for a function $M(x_1, \dots, x_n)$ being a mean, i.e., $M(x_1, \dots, x_n) = \phi^{-1}(\sum_{i=1}^n \phi(x_i))$ for a continuous, strictly increasing function ϕ and any $n \geq 1$ and all values $x_1, \dots, x_n$ in an interval $[a, b]$. Hardy et al. [25] give a similar representation result in terms of distribution functions. For further discussions on means, we refer to Hardy et al. [25] and Bullen [5].

Building on the Nagumo-Kolmogorov-de Finetti Theorem, Cerreia-Vioglio et al. [7] recently proved that a function $c_0 : L^\infty \to \mathbb{R}$ is of the form 5 if and only if c_0 is normalized on constants, law invariant, monotone, has the Lebesgue property and satisfies

$$c_0(X1_A) > c_0(Y1_A) \quad \Leftrightarrow \quad c_0(X1_A + Z1_{A^c}) > c_0(Y1_A + Z1_{A^c}) \tag{7}$$

for all $A \in \mathcal{F}$ and all $X, Y, Z \in L^\infty$. Lemma 2 in Carreia-Vioglio et al. [7] further shows that for the case $a = -\infty$ and $b = +\infty$, the function c_0 is quasi-concave if and only if u is concave.

Remark 1.6 For any continuous and strictly increasing function $u : \mathbb{R} \to \mathbb{R}$, the functional $\pi(X) := u^{-1} \circ \mathbb{E}[u(X)]$ defines an insurance premium principle, which is called the *mean value principle* (c.f. Gerber [21, Chapter 5, Sect. 4]). Gerber shows in [22] that any law invariant premium principle π which is iterative (i.e. $\pi(X) = \pi(\pi(X \mid Y))$ for all $X, Y \in L^\infty$, where $\pi(X \mid Y)$ denotes the premium for X given the random variable Y) and for which $[0, 1] \ni p \mapsto \pi(b^{\varepsilon,p})$ is continuous and strictly increasing has to be the mean value principle. Iterativity means that π is time consistent for every sub σ-agebra $\mathcal{G} \subset \mathcal{F}$, i.e., for all $\mathcal{G} \subset \mathcal{F}$ there is $\pi_{\mathcal{G}} : L^\infty \to L^\infty(\mathcal{G})$ such that $\pi(X) = \pi(\pi_{\mathcal{G}}(X))$ for all $X \in L^\infty$ and $\pi_{\mathcal{G}}$ has the local property Eq. 1. A premium principle π which is time consistent for all sub σ-agebras $\mathcal{G} \subset \mathcal{F}$ satisfies Eq. 7 as $\pi(X) = \pi(\pi_{\sigma(1_A)}(X))$ for all $A \in \mathcal{F}$ and all $X \in L^\infty$.

It is shown in [21] that π is cash invariant (i.e. $\pi(X + m) = \pi(X) + m$, $m \in \mathbb{R}$) if and only if u is an exponential or linear function, see also Nagumo [31] and de Finetti [16]. Cash invariant, time consistent and law invariant functions are discussed in the Sect. 1.2.

Remark 1.7 Suppose that, additionally to the assumptions of Theorem 1.4, we have $c_0(X) < \mathbb{E}[X]$ for all $X \in L^\infty(a, b)$ unless X is constant. Then u is strictly concave (see for instance Proposition 2.35 in Föllmer and Schied [18]).

1.2 A representation result for law invariant, time consistent dynamic risk measures

In this subsection, we consider functionals which are *cash invariant* (in the literature this property is also referred to as translation invariance [2,3]). This extra condition allows us to prove the main Theorem 1.4 for functions which are not assumed to satisfy condition (C) and are relevant instead of strictly increasing.

Definition 1.8 A *dynamic risk measure* is a family $(\rho_t)_{t \in \mathbb{N}_0}$ of mappings $\rho_t : L^\infty \to L_t^\infty$ such that, for all $X, Y \in L^\infty$, the following properties are satisfied

(i) normalization: $\rho_t(0) = 0$;
(ii) cash invariance: $\rho_t(X + m) = \rho_t(X) - m$ for all $m \in L_t^\infty$;
(iii) monotonicity: $X \geq Y$ implies $\rho_t(X) \leq \rho_t(Y)$.

A dynamic risk measure is

- convex if $\rho_t(\lambda X + (1 - \lambda)Y) \leq \lambda \rho_t(X) + (1 - \lambda)\rho_t(Y)$ for all $\lambda \in L_t^\infty$ with $0 \leq \lambda \leq 1$ and all $t \in \mathbb{N}_0$;
- law invariant if $\rho_0(X) = \rho_0(Y)$ whenever $\text{law}(X) = \text{law}(Y)$;

- time consistent if $\rho_0(X) = \rho_0(-\rho_t(X))$ for all $t \in \mathbb{N}_0$;
- relevant if $\rho_0(-\varepsilon 1_A) > 0$ for all $A \in \mathcal{F}$ and all $\varepsilon > 0$.

The theory of risk measures has been initiated by the influential paper by Artzner et al. [2]. Since then, risk measures have been generalized in several directions. For an overview of static *convex risk measures* (mappings $\rho : L^\infty \to \mathbb{R}$ which are normalized, cash invariant, monotone and convex) we refer to Föllmer and Schied [18]. We here are mainly interested in *law invariant* risk measures which are studied for instance in Kusuoka [28], Frittelli et al. [19], Jouini et al. [26] and Cheridito and Li [10]. For *dynamic risk measures* their representations and related concepts such as time consistency, we refer to Artzner et al. [3], Cheridito et al. [8], Cheridito and Kupper [9], Föllmer and Penner [17] and the references therein.

Remark 1.9 It is shown in Jouini et al. [26] that any law invariant convex risk measure ρ_0 automatically has the *Fatou property*.

Here is our second main result.

Theorem 1.10 *Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{N}_0}, \mathbb{P}\right)$ be a standard filtered probability space. The family $(\rho_t)_{t\in\mathbb{N}_0}$ is a law invariant, time consistent, relevant dynamic risk measure if and only if there is $\gamma \in (-\infty, \infty]$ such that*

$$\rho_t(X) = \frac{1}{\gamma} \ln \mathbb{E}\left[\exp(-\gamma X) \mid \mathcal{F}_t\right] \quad \textit{for all } t \in \mathbb{N}_0. \tag{8}$$

The limiting cases $\gamma = 0$ and $\gamma = \infty$ are defined as

$$\rho_t(X) = \begin{cases} \mathbb{E}\left[-X \mid \mathcal{F}_t\right] & \gamma = 0 \\ ess\,sup_{Z\in\mathcal{P}_t} \mathbb{E}\left[Z(-X) \mid \mathcal{F}_t\right] & \gamma = \infty \end{cases}, \tag{9}$$

where $\mathcal{P}_t$ denotes the set of all positive integrable random variables Z with $\mathbb{E}\left[Z \mid \mathcal{F}_t\right] = 1$.

In addition, the dynamic risk measure $(\rho_t)_{t\in\mathbb{N}_0}$ is convex (resp. coherent) iff $\gamma \in [0, \infty]$ (resp. $\gamma \in \{0, \infty\}$).

Note that $\operatorname{ess\,sup}_{Z\in\mathcal{P}_t} \mathbb{E}\left[Z(-X) \mid \mathcal{F}_t\right]$ is the time t conditional *worst case risk measure*. Let us give some remarks and compare Theorem 1.10 with the existing literature.

Remark 1.11 Due to Corollary 4.59 in Föllmer and Schied [18], any law invariant, convex risk measure is relevant. As a corollary of Theorem 1.10 we deduce that any law invariant, time consistent, convex dynamic risk measure is of the form Eq. 8 for some $\gamma \in [0, \infty]$.

Remark 1.12 Closely related to Theorem 1.10 is a result by Delbaen [13]. In a continuous time framework, under a filtration for which all martingales are continuous, it is shown that the only law invariant, time consistent, dynamic coherent risk measure (a dynamic convex risk measure, which additionally satisfies $\rho_t(\lambda X) = \lambda\rho_t(X)$ for all $\lambda \in (L_t^\infty)_+$) is either the negative of the expected value or the worst case risk measure. Independently of the present paper, Delbaen presented at the Oberwolfach meeting (2008) on “Stochastic Analysis in Finance and Insurance” a version of Theorem 1.10 in continuous time, under a filtration generated by a Brownian motion. The result is based on a representation result for dynamic penalty functions by Delbaen et al. [14].

Remark 1.13 A continuous time dynamic risk measure can be embedded in our discrete time framework. Indeed, if $(\rho_t)_{t\in[0,T]}$ is a time consistent, dynamic risk measure in continuous time, then for any strictly increasing sequence $0 = t_0 < t_1 < \cdots, t_i \in [0, T]$, the family

$\tilde{\rho}_n = \rho_{t_n} : L^\infty \to L^\infty_{t_n}$ is a time consistent dynamic risk measure in discrete time. If $\rho_0 = \tilde{\rho}_0$ is law invariant and relevant then Theorem 1.10 states that ρ_0 has to be the entropic risk measure.

Remark 1.14 Every law invariant, time consistent, dynamic risk measure $(\rho_t)_{t\in\mathbb{N}_0}$ is additive for independent random variables, i.e.,

$$\rho_0(X+Y) = \rho_0(X) + \rho_0(Y) \quad \text{for all } X, Y \in L^\infty \text{ being independent.} \tag{10}$$

In Goovaerts et al. [24] it is shown that any risk measure satisfying Eq. 10 is a weighted average of entropic risk measures:

$$\begin{aligned} -\rho_0(X) = G(-\infty)\text{ess.inf}X + \int\limits_{(-\infty,\infty)} -\frac{1}{\gamma}\ln \mathbb{E}\left[\exp(-\gamma X)\right] dG(\gamma) \\ + (1-(G(\infty)))\text{ess.sup}X, \end{aligned}$$

for an increasing function $G : [-\infty, \infty] \to [0, 1]$. Related results for premium principles satisfying a different monotonicity assumption have axiomatically been characterized by Gerber and Goovaerts [23].

We finally sketch why Eq. 10 follows from the above assumptions. Indeed, there exist $X' \in L^\infty_1$ and $Y' \in L^\infty_2$ such that Y' is independent of $\mathcal{F}_1$, $\text{law}(X') = \text{law}(X)$ and $\text{law}(Y') = \text{law}(Y)$. By Lemma 2.2 below, it follows

$$\rho_0(X+Y) = \rho_0(X'+Y') = \rho_0(X' - \rho_1(Y')) = \rho_0(X) + \rho_0(Y).$$

Remark 1.15 Weber [35] studies law invariant dynamic convex risk measures which satisfy a weaker time consistency property. More precisely, he shows that if ρ_t is weakly acceptance and rejection consistent then ρ_0 has to be a *shortfall risk measure*.

Remark 1.16 Artzner et al. [3] and Cheridito and Stadje [11] provide explicit counter-examples which demonstrate that the average value at risk $AV@R$ and the value at risk $V@R$ are not time consistent. Note that both risk measures are law invariant, but $V@R$ even fails to be convex.

2 Proof of Theorem 1.10

Throughout this section, $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{N}_0}, \mathbb{P}\right)$ is a standard filtered probability space.

Proof of the "if"-part of Theorem 1.4 It is straightforward to check that any dynamic risk measure $(\rho_t)_{t\in\mathbb{N}_0}$ of the form 8 defines a law invariant, time consistent, relevant dynamic risk measure. □

Proof of the "only if"-part of Theorem 1.4 Let $(\rho_t)_{t\in\mathbb{N}_0}$ be a law invariant, time consistent, relevant dynamic risk measure. Let us define the collection of utility functions

$$u_\gamma(x) = \begin{cases} \frac{1-\exp(-\gamma x)}{1-\exp(-\gamma)} & \text{if } \gamma \in \mathbb{R}\setminus\{0\} \\ x & \text{if } \gamma = 0 \end{cases},$$

satisfying $u_\gamma(0) = 0$, $u_\gamma(1) = 1$ and $u_\gamma : \mathbb{R} \to \mathbb{R}$ is increasing, for all $\gamma \in \mathbb{R}$. For every sequence $(\gamma_k)_{k\in\mathbb{N}} \subset \mathbb{R}$ with $\gamma_k \to \gamma \in \mathbb{R}$, the sequence u_{γ_k} converges uniformly on compacts to u_γ. The entropic risk measure with risk aversion parameter $\gamma \in \mathbb{R}$ is defined by

$$\rho^\gamma(X) = \begin{cases} \frac{1}{\gamma} \ln \mathbb{E}\left[\exp(-\gamma X)\right] & \text{if } \gamma \in \mathbb{R} \setminus \{0\} \\ \mathbb{E}[-X] & \text{if } \gamma = 0 \end{cases}, \quad X \in L^\infty.$$

Note that $\rho^\gamma(X) = -u_\gamma^{-1} \circ \mathbb{E}\left[u_\gamma(X)\right]$ for all $X \in L^\infty$. □

Lemma 2.1 *The collection of entropic risk measures ρ^γ, $\gamma \in \mathbb{R}$, satisfies for all $X \in L^\infty$*

(i) $\lim_{\gamma\to\infty} \rho^\gamma(X) = \sup_{Z\in\mathcal{P}_0} \mathbb{E}[Z(-X)] = ess.sup(-X)$,
(ii) $\lim_{\gamma\to 0} \rho^\gamma(X) = \mathbb{E}[-X]$,
(iii) $\lim_{\gamma\to -\infty} \rho^\gamma(X) = \inf_{Z\in\mathcal{P}_0} \mathbb{E}[Z(-X)] = ess.inf(-X)$ *and*
(iv) *the function* $\mathbb{R} \ni \gamma \mapsto \rho^\gamma(X)$ *is increasing.*

The set $\mathcal{P}_0$ consists of all probability densities, i.e., all positive integrable random variables Z with $\mathbb{E}[Z] = 1$.

Proof (i) follows from the well-known dual representation for the entropic risk measure (see for instance in [8,17,18])

$$\rho^\gamma(X) = \sup_{Z\in\mathcal{P}_0} \left\{ \mathbb{E}[Z(-X)] - \frac{1}{\gamma}\mathbb{E}[Z \ln Z] \right\}, \quad X \in L^\infty, \ \gamma > 0. \tag{11}$$

Equation 11 further yields

$$\lim_{\gamma \searrow 0} \rho^\gamma(X) = \mathbb{E}[-X] \quad \text{and} \quad \mathbb{R}_+ \ni \gamma \mapsto \rho^\gamma(X) \text{ is increasing.} \tag{12}$$

(ii)–(iv) then follow from (i), 12 and the equality $\rho^\gamma(X) = -\rho^{-\gamma}(-X)$ valid for all $X \in L^\infty$ and all $\gamma \in \mathbb{R}$. □

Lemma 2.1 justifies the Definition 9. Let ρ^∞ ($\rho^{-\infty}$) be defined as the worst (best) case risk measure. Let $b_1, b_2, b_3, \dots$ denote a sequence of independent Bernoulli random variables, such that b_t is independent of $\mathcal{F}_{t-1}$ and assumes the values 1 and -1 with probabilities $1/2$. Define

$$\eta_\varepsilon := \rho_0(\varepsilon b_1) \quad \text{implying} \quad \rho_0(\varepsilon b_1 + \eta_\varepsilon) = 0. \tag{13}$$

Clearly, $-\varepsilon \le \eta_\varepsilon \le \varepsilon$. For instance, if ρ_0 is the worst case risk measure, then $\eta_\varepsilon = \varepsilon$. Define $\gamma_\varepsilon \in [-\infty, \infty]$ implicitly through $\rho^{\gamma_\varepsilon}(\varepsilon b_1 + \eta_\varepsilon) = 0$. In particular, if $\gamma_\varepsilon \in \mathbb{R}$, then

$$u_{\gamma_\varepsilon}(x) = \mathbb{E}\left[u_{\gamma_\varepsilon}(x + \varepsilon b_1 + \eta_\varepsilon)\right], \quad \text{for all } x \in \mathbb{R}. \tag{14}$$

The goal is to show that there exists $\gamma \in (-\infty, \infty]$ such that

$$\rho_0(X) = \rho^\gamma(X) \quad \text{for all } X \in L^\infty, \tag{15}$$

which in turn implies Eq. 8.

Lemma 2.2 *Let $t \in \mathbb{N}$ and $X, Y \in L^\infty$ with $law(X) = law(Y)$ and Y is independent of $\mathcal{F}_t$. Then, we have $\rho_0(X) = \rho_t(Y)$.*

Proof Suppose that $\rho_t(Y)$ is not constant. Then, there exist $m \in \mathbb{R}$ and $A, A' \in \mathcal{F}_t$ with $\mathbb{P}[A] = \mathbb{P}[A'] > 0$ such that $\rho_t(Y) - m > 0$ on A and $\rho_t(Y) - m \le 0$ on A'. By time consistency, local property and cash invariance of ρ_t, we deduce

$$\rho_0\left(1_A(m - \rho_t(Y))\right) = \rho_0\left(1_A(-\rho_t(Y + m))\right) = \rho_0\left(-\rho_t(1_A(Y + m))\right) = \rho_0(1_A(Y + m)) \tag{16}$$

and analogously

$$\rho_0\left(1_{A'}(m-\rho_t(Y))\right)=\rho_0(1_{A'}(Y+m)). \tag{17}$$

On the one hand, since ρ_0 is relevant, it follows

$$\rho_0\left(1_A(m-\rho_t(Y))\right)>0\geq\rho_0\left(1_{A'}(m-\rho_t(Y))\right). \tag{18}$$

On the other hand, $\text{law}(1_A(Y+m))=\text{law}(1_{A'}(Y+m))$ implying that $\rho_0(1_A(Y+m))=\rho_0(1_{A'}(Y+m))$ in contradiction to Eqs. 16–18. Hence, $\rho_t(Y)$ has to be constant. Then, since $\text{law}(X)=\text{law}(Y)$, we get

$$\rho_0(X)=\rho_0(Y)=\rho_0(-\rho_t(Y))=\rho_t(Y)$$

by time consistency and cash invariance of ρ_0. This completes the proof. □

For any $\varepsilon>0$, we define the random walk with drift

$$R_t^\varepsilon:=R_0^\varepsilon+\sum_{j=1}^t\left(\varepsilon b_j+\eta_\varepsilon\right),\quad t\in\mathbb{N}_0, \tag{19}$$

starting at $R_0^\varepsilon\in\mathbb{R}$. The following Lemma shows that any R^ε of the form Eq. 19 satisfies $\rho_t(R_s^\varepsilon)=-R_t^\varepsilon$ for all $s\geq t$, which can be viewed as a generalized martingale property with respect to the non-linear conditional expectation ρ_t.

Lemma 2.3 *Let R^ε be a stochastic process which follows the dynamics Eq. 19 and let τ be a bounded stopping time. Then, we have $\rho_0(R_\tau^\varepsilon)=-R_0^\varepsilon$.*

Proof We first show that

$$\rho_s(R_{s+1}^\varepsilon)=-R_s^\varepsilon\quad\text{for all } s\in\mathbb{N}_0. \tag{20}$$

Indeed, suppose that R_s^ε assumes the values $\{x_1,\dots,x_N\}$, i.e., $\mathbb{P}[R_s^\varepsilon\in\{x_1,\dots,x_N\}]=1$. In view of Eq. 19 we have

$$R_{s+1}^\varepsilon=\sum_{n=1}^N 1_{\{R_s^\varepsilon=x_n\}}\left(x_n+\varepsilon b_{s+1}+\eta_\varepsilon\right).$$

Hence, the local property of ρ_s yields

$$\rho_s(R_{s+1}^\varepsilon)=\sum_{n=1}^N 1_{\{R_s^\varepsilon=x_n\}}\rho_s\left(x_n+\varepsilon b_{s+1}+\eta_\varepsilon\right). \tag{21}$$

Since $\text{law}(x_n+\varepsilon b_{s+1}+\eta_\varepsilon)=\text{law}(x_n+\varepsilon b_1+\eta_\varepsilon)$ a.s., and $x_n+\varepsilon b_{s+1}+\eta_\varepsilon$ is independent of $\mathcal{F}_s$, Lemma 2.2 and Eq. 13 imply

$$\rho_s(x_n+\varepsilon b_{s+1}+\eta_\varepsilon)=\rho_0(x_n+\varepsilon b_1+\eta_\varepsilon)=-x_n. \tag{22}$$

We then derive Eq. 20 from Eqs. 21 and 22.

We next show by backward induction that

$$\rho_t(R_\tau^\varepsilon)=-R_{t\wedge\tau}^\varepsilon\quad\text{for all } t\in\mathbb{N}_0. \tag{23}$$

Indeed, since τ is a bounded stopping time there is $T \in \mathbb{N}$ with $\tau \le T$. By cash invariance of ρ_T we have $\rho_T(R^\varepsilon_\tau) = -R^\varepsilon_\tau = -R^\varepsilon_{T\wedge\tau}$. For the induction step, we assume that Eq. 23 holds for all $t \ge s+1$. In view of Eq. 20, we deduce on $A := \{\tau \ge s+1\} \in \mathcal{F}_s$

$$1_A\rho_s(R^\varepsilon_\tau) = 1_A\rho_s(-\rho_{s+1}(R^\varepsilon_\tau)) = 1_A\rho_s(R^\varepsilon_{(s+1)\wedge\tau}) = 1_A\rho_s(R^\varepsilon_{s+1}) = -1_A R^\varepsilon_{s\wedge\tau}. \tag{24}$$

By cash invariance of ρ_s, we deduce on $A^c = \{\tau \le s\}$

$$1_{A^c}\rho_s(R^\varepsilon_\tau) = 1_{A^c}\rho_s(-\rho_{s+1}(R^\varepsilon_\tau)) = 1_{A^c}\rho_s(R^\varepsilon_{(s+1)\wedge\tau}) = 1_{A^c}\rho_s(R^\varepsilon_{s\wedge\tau}) = -1_{A^c} R^\varepsilon_{s\wedge\tau}. \tag{25}$$

Combining Eq. 24 with Eq. 25 implies $\rho_s(R^\varepsilon_\tau) = -R^\varepsilon_{s\wedge\tau}$ and the induction step is completed. □

Remark 2.4 For any R^ε of the form Eq. 19 with respective $\gamma_\varepsilon \in \mathbb{R}$, the stochastic process $u_{\gamma_\varepsilon}(R^\varepsilon_t)$ is a martingale. Indeed, for all $t \in \mathbb{N}_0$ we deduce from Eq. 14 that

$$\mathbb{E}\left[u_{\gamma_\varepsilon}\left(R^\varepsilon_{t+1}\right) \mid \mathcal{F}_t\right] = \mathbb{E}\left[u_{\gamma_\varepsilon}\left(R^\varepsilon_t + \varepsilon b_{t+1} + \eta_\varepsilon\right) \mid \mathcal{F}_t\right] = u_{\gamma_\varepsilon}\left(R^\varepsilon_t\right). \tag{26}$$

The proof is based on the following discrete version of the Skorohod embedding theorem (see for instance Revuz and Yor [33, Chapter VI, Sect. 5], and the references therein).

Lemma 2.5 *Let $X \in L^\infty$ and $(\varepsilon_k)_{k\in\mathbb{N}}$ be a sequence tending to zero such that $\gamma_k := \gamma_{\varepsilon_k} \in \mathbb{R}$ and $(\gamma_k)_{k\in\mathbb{N}}$ converges to some $\gamma \in \mathbb{R}$ as k tends to infinity. Then, there exists a subsequence of (ε_k) (still denoted by (ε_k)), such that for any $k \in \mathbb{N}$ we may find stochastic processes $R^{\varepsilon_k,+}$, $R^{\varepsilon_k,-}$ of the form Eq. (19) as well as bounded stopping times σ_k^+ and σ_k^-, such that X satisfies*

$$R^{\varepsilon_k,+}_{\sigma_k^+} \succeq_1 X \succeq_1 R^{\varepsilon_k,-}_{\sigma_k^-}, \tag{27}$$

and

$$\lim_{k\to\infty}\left|u_{\gamma_k}(R^{\varepsilon_k,+}_0) - \mathbb{E}\left[u_{\gamma_k}(X)\right]\right| = \lim_{k\to\infty}\left|u_{\gamma_k}(R^{\varepsilon_k,-}_0) - \mathbb{E}\left[u_{\gamma_k}(X)\right]\right| = 0. \tag{28}$$

Proof A discrete version of the Skorohod embedding theorem is the heart of the construction. The technical details are a little messy but the basic idea is straightforward. For the convenience of the reader we first informally sketch the idea. The process R^{ε_k} is a random walk with drift. Further, $u_{\gamma_k}(R^{\varepsilon_k})$ is a martingale. We approximate a given $X \in L^\infty$ in law by the terminal value $R^{\varepsilon_k}_\sigma$, where σ is a bounded stopping time. The goal is to have $u_{\varepsilon_k}(R^{\varepsilon_k}_0) = \mathbb{E}\left[u_{\varepsilon_k}(R^{\varepsilon_k}_\sigma)\right] \approx \mathbb{E}\left[u_{\varepsilon_k}(X)\right]$. We first assume that X only assumes two values $x_1 < x_2$. We start the random walk R^{ε_k} at $R^{\varepsilon_k}_0 = u^{-1}_{\gamma_k} \circ \mathbb{E}\left[u_{\gamma_k}(X)\right] \in (x_1, x_2)$ and define the stopping time $\sigma = \inf\left\{t \in \mathbb{N}_0 \mid R^{\varepsilon_k}_t \approx x_1 \text{ or } R^{\varepsilon_k}_t \approx x_2\right\}$. By martingale convergence, σ is a.s. finite and $\mathbb{P}[R^{\varepsilon_k}_\sigma \approx x_1] \approx \mathbb{P}[X = x_1]$ and $\mathbb{P}[R^{\varepsilon_k}_\sigma \approx x_2] \approx \mathbb{P}[X = x_2]$. We then need some technicalities to make the meaning of "$\approx$" precise and to replace σ by a bounded stopping time. We then repeat the above argument along a binomial tree.

We now state the proof in full detail. It is enough to prove the lemma for random variables X which assume 2^N different values for some $N \in \mathbb{N}$. Indeed, approximate $X \in L^\infty$ by some $X^N \in L^\infty$ which takes 2^N different values such that $||X - X^N||_\infty \le 1/N$ and

$$X^N + 1/N \succeq_1 X \succeq_1 X^N - 1/N.$$

Applying the lemma (valid for random variables X which assume 2^N values) on $X^N + 1/N$ and $X^N - 1/N$ yields stochastic processes $R^{\varepsilon_k,N,+}$, $R^{\varepsilon_k,N,-}$ and bounded stopping times $\sigma^+_{k,N}$ and $\sigma^-_{k,N}$ such that

$$R^{\varepsilon_k,N,+}_{\sigma^+_{k,N}} \succeq_1 X^N + 1/N \succeq_1 X \succeq_1 X^N - 1/N \succeq_1 R^{\varepsilon_k,N,-}_{\sigma^-_{k,N}}$$

$$\lim_{k\to\infty} \left| u_{\gamma_k}(R^{\varepsilon_k,N,+}_0) - \mathbb{E}\left[u_{\gamma_k}(X^N + 1/N)\right]\right|$$
$$= \lim_{k\to\infty} \left| u_{\gamma_k}(R^{\varepsilon_k,N,-}_0) - \mathbb{E}\left[u_{\gamma_k}(X^N - 1/N)\right]\right| = 0.$$

The claim then follows since for any $k \in \mathbb{N}$ there is $N = N(k)$ such that

$$\left|\mathbb{E}\left[u_{\gamma_k}(X^N + 1/N)\right] - \mathbb{E}\left[u_{\gamma_k}(X)\right]\right| \le 1/k.$$

Suppose now that X takes the values $x_1 < \cdots < x_{2^N}$ with respective probabilities $p_1, \ldots, p_{2^N}$. Let us introduce the finite probability space

$$\left(\hat{\Omega} = \{\omega_1, \ldots, \omega_{2^N}\}, \hat{\mathcal{F}} = 2^{\hat{\Omega}}, \hat{\mathbb{P}} = \{p_1, \ldots, p_{2^N}\}\right).$$

$\hat{\mathbb{E}}$ denotes the (conditional) expectation with respect to $\hat{\mathbb{P}}$. The filtration $\hat{\mathcal{F}}_n = \sigma(\hat{A}^1_n, \ldots, \hat{A}^{2^n}_n)$, $n = 0, \ldots, N$, is generated by the time n atoms

$$\hat{A}^j_n = \left\{\omega_k \mid k \in \left\{1 + (j-1)\frac{2^N}{2^n}, 2 + (j-1)\frac{2^N}{2^n}, \ldots, j\frac{2^N}{2^n}\right\}\right\}, \quad j = 1, \ldots, 2^n.$$

The random variable $\hat{X} = (x_1, \ldots, x_{2^N})$ on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ has the same distribution as X on $(\Omega, \mathcal{F}, \mathbb{P})$. Fix $k \in \mathbb{N}$ and define the stochastic process $\hat{Y}^{\varepsilon_k}$ inductively by $\hat{Y}^{\varepsilon_k}_N = \hat{X}$ and

$$\hat{Y}^{\varepsilon_k}_n = u^{-1}_{\gamma_k} \circ \hat{\mathbb{E}}\left[u_{\gamma_k}(\hat{Y}^{\varepsilon_k}_{n+1} + 2\varepsilon_k) \mid \hat{\mathcal{F}}_n\right], \quad n = 0, \ldots, N-1. \tag{29}$$

Let $\hat{y}^1_n < \cdots < \hat{y}^{2^n}_n$ denote the 2^n different values of the random variable $\hat{Y}^{\varepsilon_k}_n$, so that, by construction $\hat{A}^j_n = \{\hat{Y}^{\varepsilon_k}_n = \hat{y}^j_n\}$. Further, let $R^{\varepsilon_k,+}$ be the stochastic process which follows the dynamics 19 and starts at $R^{\varepsilon_k,+}_0 = \hat{Y}^{\varepsilon_k}_0$.

Step 1. There exists an increasing sequence of a.s. finite stopping times $0 = \tilde{\sigma}_0 \le \tilde{\sigma}_1 \le \cdots \le \tilde{\sigma}_N$ and for each $n = 0, \ldots, N$ there is an almost sure partition $(A^j_n)^{2^n}_{j=1}$ of Ω such that

$$R^{\varepsilon_k,+}_{\tilde{\sigma}_n} \ge \hat{y}^j_n \quad \text{on } A^j_n, \quad n = 0, \ldots, N, \quad j = 1, \ldots, 2^n, \tag{30}$$

$$\sum_{i=j}^{2^n} \mathbb{P}[A^i_n] > \sum_{i=j}^{2^n} \hat{\mathbb{P}}[\hat{A}^i_n], \quad n = 1, \ldots, N, \quad j = 2, \ldots, 2^n. \tag{31}$$

In particular, $\mathbb{P}[R^{\varepsilon_k,+}_{\tilde{\sigma}_n} \ge \hat{y}^j_n] > \hat{\mathbb{P}}[\hat{Y}^{\varepsilon_k}_n \ge \hat{y}^j_n]$ for all $n = 1, \ldots, N$, $j = 2, \ldots, 2^n$ and therefore

$$R^{\varepsilon_k,+}_{\tilde{\sigma}_n} \succeq_1 \hat{Y}^{\varepsilon_k}_n, \quad n = 0, \ldots, N. \tag{32}$$

The proof of Eqs. 30 and 31 is by induction on $n = 0, \ldots, N$. By construction $R^{\varepsilon_k,+}_{\tilde{\sigma}_0} \ge \hat{Y}^{\varepsilon_k}_0$. Fix $1 \le m \le N-1$. We assume that Eqs. 30 and 31 hold for all $n \le m$ and we will show that there exists a finite stopping time $\tilde{\sigma}_{m+1} \ge \tilde{\sigma}_m$ and an almost sure partition $(A^j_{m+1})^{2^{m+1}}_{j=1}$

of Ω such that Eqs. 30 and 31 also hold for $m+1$. Fix $1 \leq j_0 \leq 2^m$. On $\hat{A}_m^{j_0}$, the random variable $\hat{Y}_m^{\varepsilon_k}$ equals $\hat{y}_m := \hat{y}_m^{j_0}$, whereas $\hat{Y}_{m+1}^{\varepsilon_k}$ takes the two values $\hat{y}_{m+1}^d := \hat{y}_{m+1}^{2j_0-1}$ and $\hat{y}_{m+1}^u := \hat{y}_{m+1}^{2j_0}$. Then

$$u_{\gamma_k}(\hat{y}_m) = \hat{p}^d u_{\gamma_k}(\hat{y}_{m+1}^d + 2\varepsilon_k) + \hat{p}^u u_{\gamma_k}(\hat{y}_{m+1}^u + 2\varepsilon_k), \tag{33}$$

for the conditional probabilities $0 < \hat{p}^d, \hat{p}^u < 1$. On $A_m^{j_0}$, the stopping time $\tilde{\sigma}_{m+1}$ is defined by

$$\tilde{\sigma}_{m+1} = \inf\left\{t \geq \tilde{\sigma}_m \mid R_t^{\varepsilon_k,+} \in [\hat{y}_{m+1}^d, \hat{y}_{m+1}^d + 2\varepsilon_k) \text{ or } R_t^{\varepsilon_k,+} \geq \hat{y}_{m+1}^u\right\}. \tag{34}$$

In view of Eq. 33 we deduce $\hat{y}_{m+1}^d + 2\varepsilon_k < \hat{y}_m < \hat{y}_{m+1}^u + 2\varepsilon_k$, which together with Eq. 30 yields $\hat{y}_{m+1}^d + \varepsilon_k < R_{\tilde{\sigma}_m}^{\varepsilon_k,+}$. Due to Remark 3.4 $u_{\varepsilon_k}(R^{\varepsilon_k,+})$ is a martingale. Hence the martingale convergence theorem implies that $\tilde{\sigma}_{m+1}$ is a.s. finite. Thus, the sets

$$A_{m+1}^{2j-1} = \left\{R_{\tilde{\sigma}_{m+1}}^{\varepsilon_k,+} \in [\hat{y}_{m+1}^{2j-1}, \hat{y}_{m+1}^{2j-1} + 2\varepsilon_k)\right\} \quad \text{and}$$
$$A_{m+1}^{2j} = \left\{R_{\tilde{\sigma}_{m+1}}^{\varepsilon_k,+} \geq \hat{y}_{m+1}^{2j}\right\}, \quad j = 1, \ldots 2^m,$$

form an almost sure partition of Ω and Eq. 30 holds for $n = m+1$. Since $u_{\gamma_k}(R^{\varepsilon_k,+})$ is a martingale, Eq. 30 yields $\mathbb{E}[u_{\gamma_k}(R_{\tilde{\sigma}_{m+1}}^{\varepsilon_k,+}) \mid A_m^{j_0}] \geq u_{\varepsilon_k}(\hat{y}_m)$. If $R_{\tilde{\sigma}_m}^{\varepsilon_k,+} > \hat{y}_{m+1}^u$, then we have $\mathbb{P}[R_{\tilde{\sigma}_{m+1}}^{\varepsilon_k,+} \geq \hat{y}_{m+1}^u \mid A_m^{j_0}] = 1$. If $R_{\tilde{\sigma}_m}^{\varepsilon_k,+} \leq \hat{y}_{m+1}^u$, then $R_{\tilde{\sigma}_{m+1}}^{\varepsilon_k,+} \leq \hat{y}_{m+1}^u + 2\varepsilon_k$, which in view of Eqs. 33 and 34 leads to $\mathbb{P}[R_{\tilde{\sigma}_{m+1}}^{\varepsilon_k,+} \geq \hat{y}_{m+1}^u \mid A_m^{j_0}] > \hat{p}^u$. This shows

$$\mathbb{P}[R_{\tilde{\sigma}_{m+1}}^{\varepsilon_k,+} \geq \hat{y}_{m+1}^u \mid A_m^{j_0}] > \hat{p}^u \quad \text{and} \quad \mathbb{P}[R_{\tilde{\sigma}_{m+1}}^{\varepsilon_k,+} \geq \hat{y}_{m+1}^d \mid A_m^{j_0}] \geq 1. \tag{35}$$

We next prove Eq. 31 for $n = m+1$. By construction,

$$\hat{\mathbb{P}}[\hat{A}_m^{j_0}]\hat{p}^u = \hat{\mathbb{P}}[\hat{A}_{m+1}^{2j_0}] \quad \text{and} \quad \hat{\mathbb{P}}[\hat{A}_m^{j_0}](1-\hat{p}^u) = \hat{\mathbb{P}}[\hat{A}_{m+1}^{2j_0-1}].$$

On the one hand, if $\mathbb{P}[A_m^{j_0}] \geq \hat{\mathbb{P}}[\hat{A}_m^{j_0}]$, the induction hypothesis Eqs. 31 and 35 imply

$$\sum_{i=2j_0}^{2^{m+1}} \mathbb{P}[A_{m+1}^i] = \mathbb{P}[A_m^{j_0}]\mathbb{P}[R_{\tilde{\sigma}_{m+1}}^{\varepsilon_k,+} \geq \hat{y}_{m+1}^{2j_0} \mid A_m^{j_0}] + \sum_{i=j_0+1}^{2^m} \mathbb{P}[A_m^i]$$
$$> \hat{\mathbb{P}}[\hat{A}_m^{j_0}]\hat{p}^u + \sum_{i=j_0+1}^{2^m} \hat{\mathbb{P}}[\hat{A}_m^i] = \sum_{i=2j_0}^{2^{m+1}} \hat{\mathbb{P}}[\hat{A}_m^i].$$

On the other hand, if $\mathbb{P}[A_m^{j_0}] < \hat{\mathbb{P}}[\hat{A}_m^{j_0}]$, we deduce

$$\sum_{i=1}^{2j_0-1} \mathbb{P}[A_{m+1}^i] = \mathbb{P}[A_m^{j_0}]\mathbb{P}[R_{\tilde{\sigma}_{m+1}}^{\varepsilon_k,+} < \hat{y}_{m+1}^{2j_0} \mid A_m^{j_0}] + \sum_{i=1}^{j_0-1} \mathbb{P}[A_m^i]$$
$$< \hat{\mathbb{P}}[\hat{A}_m^{j_0}](1-\hat{p}^u) + \sum_{i=1}^{j_0-1} \hat{\mathbb{P}}[\hat{A}_m^i] = \sum_{i=1}^{2j_0-1} \hat{\mathbb{P}}[\hat{A}_m^i].$$

This shows $\sum_{i=2j_0}^{2^{m+1}} \mathbb{P}[A^i_{m+1}] > \sum_{i=2j_0}^{2^{m+1}} \hat{\mathbb{P}}[\hat{A}^i_m]$. Moreover, if $j_0 \geq 2$ (only in this case we have to check the induction hypothesis Eq. 31), we deduce

$$\sum_{i=2j_0-1}^{2^{m+1}} \mathbb{P}[A^i_{m+1}] = \sum_{i=j_0}^{2^m} \mathbb{P}[A^i_m] > \sum_{i=j_0}^{2^m} \hat{\mathbb{P}}[\hat{A}^i_m] = \sum_{i=2j_0-1}^{2^{m+1}} \hat{\mathbb{P}}[\hat{A}^i_m].$$

This completes the induction step.
Step 2. There is an increasing sequence of bounded stopping times $0 = \sigma_0 \leq \cdots \leq \sigma_N$ such that

$$R^{\varepsilon_k,+}_{\sigma_n} \succeq_1 \hat{Y}^{\varepsilon_k}_n, \quad n = 0, \ldots, N. \tag{36}$$

Indeed, since (c.f. 31) $\mathbb{P}[R^{\varepsilon_k,+}_{\tilde{\sigma}_n} \geq \hat{y}^j_n] > \hat{\mathbb{P}}[\hat{Y}^{\varepsilon_k}_n \geq \hat{y}^j_n]$ for all $n = 1, \ldots, N$ and all $j = 2, \ldots, 2^n$, the dominated convergence theorem implies $T \in \mathbb{N}$, such that

$$\mathbb{P}[R^{\varepsilon_k,+}_{\tilde{\sigma}_n \wedge T} \geq \hat{y}^j_n] \geq \hat{\mathbb{P}}[\hat{Y}^{\varepsilon_k}_n \geq \hat{y}^j_n] \quad \text{for all } n = 1, \ldots, N \text{ and all } j = 2, \ldots, 2^n.$$

Hence, for the bounded stopping times $\sigma_n = \tilde{\sigma}_n \wedge T$ we get Eq. 36.

Step 3. There is a subsequence of (ε_k), which we still denote by (ε_k), such that $\hat{Y}^{\varepsilon_k}_n$ converges to $\hat{Y}_n$ for all $n = 0, \ldots, N-1$. Then, by continuity of u_γ it follows $u_{\gamma_k}(\hat{Y}^{\varepsilon_k}_{N-1}) \to u_\gamma(\hat{Y}_{N-1})$ and Eq. 29 yields

$$u_{\gamma_k}(\hat{Y}^{\varepsilon_k}_{N-1}) = \hat{\mathbb{E}}\left[u_{\gamma_k}(\hat{X} + 2\varepsilon_k) \mid \hat{\mathcal{F}}_{N-1}\right] \to \hat{\mathbb{E}}\left[u_\gamma(\hat{X}) \mid \hat{\mathcal{F}}_{N-1}\right],$$

showing that $u_\gamma(\hat{Y}_{N-1}) = \hat{\mathbb{E}}[u_\gamma(\hat{X}) \mid \hat{\mathcal{F}}_{N-1}]$. Backward induction yields

$$u_\gamma(\hat{Y}_n) = \hat{\mathbb{E}}[u_\gamma(\hat{X}) \mid \hat{\mathcal{F}}_n] \quad \text{and} \quad \left|u_{\gamma_k}(\hat{Y}^{\varepsilon_k}_n) - \hat{\mathbb{E}}[u_{\gamma_k}(\hat{X}) \mid \hat{\mathcal{F}}_n]\right| \to 0$$
$$\text{for all } n = 0, \ldots, N-1.$$

In particular, since $\hat{\mathbb{E}}[u_{\gamma_k}(\hat{X})] = \mathbb{E}[u_{\gamma_k}(X)]$, it follows

$$\left|u_{\gamma_k}(R^{\varepsilon_k,+}_0) - \mathbb{E}\left[u_{\gamma_k}(X)\right]\right| \to 0. \tag{37}$$

Using similar argument as before, there exist a stochastic processes $R^{\varepsilon_k,-}$ of the form Eq. 19 and bounded stopping times σ^-_k such that $X \succeq R^{\varepsilon_k,-}_{\sigma^-_k}$ and $\left|u_{\gamma_k}(R^{\varepsilon_k,-}_0) - \mathbb{E}\left[u_{\gamma_k}(X)\right]\right| \to 0$. □

We now finish the proof of Theorem 1.4. We distinguish between three different cases.

Case 1 There is a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ tending to zero such that $\gamma_k := \gamma_{\varepsilon_k} \in \mathbb{R}$ and $(\gamma_k)_{k\in\mathbb{N}}$ converges to some $\gamma \in \mathbb{R}$ as k tends to infinity. The sequence $(u_{\gamma_k})_{k\in\mathbb{N}}$ then converges uniformly on compacts to the function u_γ. Let X be a random variable in L^∞. ρ_0 is monotone with respect to $\succeq_1$. Indeed, let U be a random variable that is uniformly distributed on $(0, 1)$. If Y_1, Y_2 are random variables such that $Y_1 \succeq_1 Y_2$, then $F^{-1}_{Y_1}(U)$ has the same distribution as Y_1, $F^{-1}_{Y_2}(U)$ has the same distribution as Y_2, and $F^{-1}_{Y_1}(U) \geq F^{-1}_{Y_2}(U)$. Since ρ_0 is law invariant, one obtains that $\geq$-monotonicity implies $\succeq_1$-monotonicity. Hence, Lemmas 2.3 and 2.5 imply

$$\rho_0(X) \geq \rho_0(R^{\varepsilon_k,+}_{\sigma^+_k}) = -R^{\varepsilon_k,+}_0 = -\lim_{k\to\infty} u^{-1}_{\gamma_k} \circ \mathbb{E}\left[u_{\gamma_k}(X)\right] = -u^{-1}_\gamma \circ \mathbb{E}\left[u_\gamma(X)\right].$$

Analogously, it follows $\rho_0(X) \leq -u^{-1}_\gamma \circ \mathbb{E}\left[u_\gamma(X)\right]$, showing that $\rho_0(X) = \rho^\gamma(X)$.

Case 2 There is a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ tending to zero such that $\liminf_{k\to\infty}\gamma_k = +\infty$. Let $X \in L^\infty$, fix $M \in \mathbb{N}$ and define $\tilde{\gamma}_k := \gamma_k \wedge M$ with respective $\tilde{\eta}_{\varepsilon_k} := \rho^{\tilde{\gamma}_k}(\varepsilon_k b_1)$. By Lemma 2.5 there exist $\tilde{R}^{\varepsilon_k}$ of the form Eq. 19 and a bounded stopping times $\tilde{\sigma}_k$ such that

$$\tilde{R}^{\varepsilon_k}_{\tilde{\sigma}_k} \succeq_1 X \quad \text{and} \quad \lim_{k\to\infty}\left|u_{\tilde{\gamma}_k}(\tilde{R}^{\varepsilon_k}_0) - \mathbb{E}\left[u_{\tilde{\gamma}_k}(X)\right]\right| = 0. \tag{38}$$

Define $R^{\varepsilon_k}_t := \tilde{R}^{\varepsilon_k}_0 + \sum_{j=1}^t(\varepsilon_k b_j + \eta_{\varepsilon_k})$. Since $\tilde{\gamma}_k \le \gamma_k$, Lemma 2.1(iv) implies $\tilde{\eta}_{\varepsilon_k} \le \eta_{\varepsilon_k}$ and whence $R^{\varepsilon_k}_{\tilde{\sigma}_k} \succeq_1 \tilde{R}^{\varepsilon_k}_{\tilde{\sigma}_k}$. Lemma 2.3 in combination with Eq. 38 yields

$$\rho_0(X) \ge \limsup_{k\to\infty}\rho_0(R^{\varepsilon_k}_{\tilde{\sigma}_k}) = -\liminf_{k\to\infty} R^{\varepsilon_k}_0 = -\liminf_{k\to\infty}\tilde{R}^{\varepsilon_k}_0 = -u_M^{-1}\circ\mathbb{E}\left[u_M(X)\right] = \rho^M(X).$$

Taking the limes $M \to \infty$ we deduce $\rho_0(X) \ge \sup_{Z\in\mathcal{P}_0}\mathbb{E}\left[Z(-X)\right]$ from Lemma 2.1(i). Hence, ρ_0 is dominated below by the worst case risk measure, that is, ρ_0 has to be the worst case risk measure itself.

Case 3 If there is a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ tending to zero such that $\limsup_{k\to\infty}\gamma_k = -\infty$, then similar arguments as given in Case 2 imply that $\rho_0(X) = \inf_{Z\in\mathcal{P}_0}\mathbb{E}\left[Z(-X)\right]$. However, the best case risk measure $\inf_{Z\in\mathcal{P}_0}\mathbb{E}\left[Z(-X)\right]$ is not relevant. Hence, the Case 3 is excluded and we are left with either Cases 1 or 2. The proof is completed. □

3 Proof of Theorem 1.4

Throughout this section, $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{N}_0}, \mathbb{P}\right)$ is a standard filtered probability space.

Proof of the "if"-part of Theorem 1.4 Let us assume that $c_0 : L^\infty(a,b) \to \mathbb{R}$ is of the form $u^{-1}\circ\mathbb{E}\left[u(X)\right]$, where $u : (a,b) \to \mathbb{R}$ is a strictly increasing, continuous function. Obviously, c_0 is normalized on constants, strictly monotone, $||\cdot||_\infty$-continuous, law invariant and satisfies condition (C). Moreover, for $t \in \mathbb{N}$ we define $c_t(X) := u^{-1}\circ\mathbb{E}\left[u(X) \mid \mathcal{F}_t\right]$ and deduce on $A \in \mathcal{F}_t$

$$1_A c_t(X) = 1_A u^{-1}\circ\mathbb{E}\left[u(X) \mid \mathcal{F}_t\right] = 1_A u^{-1}\circ\mathbb{E}\left[u(Y) \mid \mathcal{F}_t\right] = 1_A c_t(Y),$$

for all $X, Y \in L^\infty(a,b)$ with $X1_A = Y1_A$. Moreover,

$$c_0(c_t(X)) = u^{-1}\circ\mathbb{E}\left[u\left(u^{-1}\circ\mathbb{E}\left[u(X) \mid \mathcal{F}_t\right]\right)\right] = c_0(X),$$

showing that c_0 is time consistent. □

Preparations for the "only if"-part of Theorem 1.4 Suppose that $c_0 : L^\infty(a,b) \to \mathbb{R}$ is normalized on constants, strictly monotone, $||\cdot||_\infty$-continuous, law invariant, time consistent and satisfies condition (C). Lemma 2.2 specializes in the present context as follows.

Lemma 3.1 *Let $t \in \mathbb{N}$ and $X, Y \in L^\infty(a,b)$ with law$(X) =$ law(Y) a.s. and Y is independent of $\mathcal{F}_t$. Then, we have $c_0(X) = c_t(Y)$.*

Proof If $c_t(Y)$ is constant, then, since c_0 is normalized on constants, time consistent, law invariant and law$(X) =$ law(Y) a.s., it follows

$$c_t(Y) = c_0(c_t(Y)) = c_0(Y) = c_0(X).$$

We therefore assume that $c_t(Y)$ is not constant. If $c_t(Y) \le c_0(X)$ and $\mathbb{P}\left[c_t(Y) < c_0(X)\right] > 0$ then strict monotonicity and time consistency of c_0 yield $c_0(Y) = c_0(c_t(Y)) < c_0(X)$

which is a contradiction. Analogously, $c_t(Y) \geq c_0(X)$ and $\mathbb{P}[c_t(Y) > c_0(X)] > 0$ is absurd. Thus, there exist $A, A' \in \mathcal{F}_t$ with $\mathbb{P}[A] = \mathbb{P}[A'] > 0$ such that $c_t(Y) < c_0(X)$ on A and $c_t(Y) > c_0(X)$ on A'. In view of the local property of c_t, time consistency and strict monotonicity of c_0 we deduce for $m \in (a, b)$

$$\begin{aligned} c_0(1_A Y + 1_{A^c} m) &= c_0(c_t(1_A Y + 1_{A^c} m)) = c_0(1_A c_t(Y) + 1_{A^c} m) \\ &< c_0(1_A c_0(X) + 1_{A^c} m) \end{aligned} \tag{39}$$

$$\begin{aligned} c_0(1_{A'} Y + 1_{A'^c} m) &= c_0(c_t(1_{A'} Y + 1_{A'^c} m)) = c_0(1_{A'} c_t(Y) + 1_{A'^c} m) \\ &> c_0(1_{A'} c_0(X) + 1_{A'^c} m). \end{aligned} \tag{40}$$

On the other hand, $\mathrm{law}(1_A Y + 1_{A^c} m) = \mathrm{law}(1_{A'} Y + 1_{A'^c} m)$ a.s. as well as $\mathrm{law}(1_A c_0(X) + 1_{A^c} m) = \mathrm{law}(1_{A'} c_0(X) + 1_{A'^c} m)$ a.s., which in view of Eqs. 39 and 40 is a contradiction to the law invariance of c_0. This shows that $c_t(Y)$ is a constant and whence $c_0(X) = c_t(Y)$. □

Let us fix a compact interval $[A, B] \subset (a, b)$ for some $A, B \in \mathbb{R}$ with $A < B$. For any $\varepsilon_n := (B - A)/n$, $n \in \mathbb{N}$, we define

$$\mathcal{I}_{\varepsilon_n} = \{A + \varepsilon_n, A + 2\varepsilon_n, \ldots, A + (n-2)\varepsilon_n, A + (n-1)\varepsilon_n\}.$$

Lemma 3.2 *For all ε_n and all $x \in [A + \varepsilon_n, B - \varepsilon_n]$ there exist Bernoulli random variables $b_+^{\varepsilon_n}(x)$ and $b_-^{\varepsilon_n}(x)$ taking the values $+\varepsilon_n$ and $-\varepsilon_n$ with probabilities*

$$\begin{aligned} \mathbb{P}\left[b_+^{\varepsilon_n}(x) = \varepsilon_n\right] &= p_+^{\varepsilon_n}(x) \quad \text{and} \quad \mathbb{P}\left[b_+^{\varepsilon_n}(x) = -\varepsilon_n\right] = 1 - p_+^{\varepsilon_n}(x), \\ \mathbb{P}\left[b_-^{\varepsilon_n}(x) = \varepsilon_n\right] &= p_-^{\varepsilon_n}(x) \quad \text{and} \quad \mathbb{P}\left[b_-^{\varepsilon_n}(x) = -\varepsilon_n\right] = 1 - p_-^{\varepsilon_n}(x), \end{aligned}$$

as well as increasing, continuous functions $u_{\varepsilon_n}^+ : [A, B] \to [0, 1]$ and $u_{\varepsilon_n}^- : [A, B] \to [0, 1]$ such that

$$c_0(x + b_+^{\varepsilon_n}(x)) \leq x, \quad c_0(x + b_-^{\varepsilon_n}(x)) \geq x, \quad \text{for all } x \in [A + \varepsilon_n, B - \varepsilon_n], \tag{41}$$

$$u_{\varepsilon_n}^+(x) = \mathbb{E}\left[u_{\varepsilon_n}^+(x + b_+^{\varepsilon_n}(x))\right], \quad u_{\varepsilon_n}^-(x) = \mathbb{E}\left[u_{\varepsilon_n}^-(x + b_-^{\varepsilon_n}(x))\right], \quad \text{for all } x \in \mathcal{I}_{\varepsilon_n}, \tag{42}$$

and

$$||u_{\varepsilon_n}^+ - u_{\varepsilon_n}^-||_\infty \to 0. \tag{43}$$

Proof Fix ε_n. For any $x \in [A + \varepsilon_n, B - \varepsilon_n]$ we define

$$p(x) := \sup\left\{p \in [0, 1] \mid c_0(x + b^{\varepsilon, p}) \leq x\right\}. \tag{44}$$

Recall that $b^{\varepsilon, p}$ denotes a Bernoulli random variable taking the values ε and $-\varepsilon$ with probabilities p and $1 - p$. Condition (C) implies that $p(x) \in (0, 1]$.

Step 1. There is an increasing, continuous function $u_{\varepsilon_n} : [A, B] \to [0, 1]$ such that $u_{\varepsilon_n}(A) = 0$, $u_{\varepsilon_n}(B) = 1$ and

$$u_{\varepsilon_n}(x) = \mathbb{E}\left[u_{\varepsilon_n}\left(x + b^{\varepsilon_n, p(x)}\right)\right], \quad \text{for all } x \in \mathcal{I}_{\varepsilon_n}. \tag{45}$$

Indeed, define $\tilde{u}_{\varepsilon_n}(A) = 0$, $\tilde{u}_{\varepsilon_n}(A + \varepsilon_n) = 1$ and inductively

$$\tilde{u}_{\varepsilon_n}(A + (k+1)\varepsilon_n) := \frac{1}{p(k\varepsilon_n)}\left[\tilde{u}_{\varepsilon_n}(A + k\varepsilon_n) - (1 - p(k\varepsilon_n))\tilde{u}_{\varepsilon_n}(A + (k-1)\varepsilon_n)\right]$$

for all $k = 1, \ldots, n-1$. Then $\tilde{u}_{\varepsilon_n}(x) \geq \tilde{u}_{\varepsilon_n}(y)$ for all $x, y \in \mathcal{I}_{\varepsilon_k}$ with $x \geq y$. The normalized function $u_{\varepsilon_n}(x) := \tilde{u}_{\varepsilon_n}(x)/\tilde{u}_{\varepsilon_n}(B)$, $x \in \mathcal{I}_{\varepsilon_n} \cup \{A, B\}$, then satisfies $u_{\varepsilon_n}(A) = 0$, $u_{\varepsilon_n}(B) = 1$

and Eq. 45. By linear interpolation u_{ε_n} extends to an increasing function on $[A, B]$, i.e., for $A + k\varepsilon_n < x < A + (k+1)\varepsilon_n$ we define

$$u_{\varepsilon_n}(x) := u_{\varepsilon_n}(A + k\varepsilon_n) + \frac{u_{\varepsilon_n}(A + (k+1)\varepsilon_n) - u_{\varepsilon_n}(A + k\varepsilon_n)}{\varepsilon_n}(x - (A + k\varepsilon_n)).$$

Step 2. Let $(p_k^+)_{k\in\mathbb{N}}$ be a sequence satisfying $0 < p_k^+(x) < p(x)$ and $p_k^+(x) \nearrow p(x)$ for all $x \in [A+\varepsilon_n, B-\varepsilon_n]$. Due to Eq. 44 we have $c_0(x+b^{\varepsilon_n, p_k^+(x)}) \leq x$ for all $x \in [A+\varepsilon_n, B-\varepsilon_n]$. Let $u_{\varepsilon_k}^{k,+}$ denote the increasing, continuous function constructed in Step 1 associated to $p_k^+(x)$. Then

$$u_{\varepsilon_n}^{k,+}(x) = \mathbb{E}\left[u_{\varepsilon_n}^{k,+}\left(x + b^{\varepsilon_n, p_k^+(x)}\right)\right], \quad \text{for all } x \in \mathcal{I}_{\varepsilon_n}.$$

Further, since $p_k^+(x) \nearrow p(x)$ for all $x \in \mathcal{I}_{\varepsilon_n}$, the sequence $(u_{\varepsilon_n}^{k,+})_{k\in\mathbb{N}}$ converges uniformly on $[A, B]$ to u_{ε_n}. Hence, there is $k_0 \in \mathbb{N}$ such that

$$b_+^{\varepsilon_n}(x) := b^{\varepsilon_n, p_{k_0}^+(x)} \quad \text{and} \quad u_{\varepsilon_n}^+(x) := u_{\varepsilon_k}^{k_0,+}(x)$$

satisfy Eqs. 41, 42 and $\|u_{\varepsilon_n}^+ - u_{\varepsilon_n}\|_\infty \leq \varepsilon_n$. Analogously, there exists a sequence $p(x) \leq p_k^-(x) \leq 1$ with $p_k^-(x) \to p(x)$ for all $x \in [A+\varepsilon_n, B-\varepsilon_n]$ and $c_0(x + b^{\varepsilon_n, p_k^-(x)}) \geq x$ for all $x \in [A+\varepsilon_n, B-\varepsilon_n]$. Note that strict monotonicity of c_0 yields $c_0(x + b^{\varepsilon_n, 1}) > x$. Let $u_{\varepsilon_k}^{k,-}$ denote the increasing, continuous function constructed in Step 1 associated to $p_k^-(x)$. The sequence $(u_{\varepsilon_n}^{k,-})_{k\in\mathbb{N}}$ converges uniformly on $[A, B]$ to u_{ε_n}. Again, there is $k_0 \in \mathbb{N}$ such that

$$b_-^{\varepsilon_n}(x) := b^{\varepsilon_n, p_{k_0}^-(x)} \quad \text{and} \quad u_{\varepsilon_n}^-(x) := u_{\varepsilon_k}^{k_0,-}(x)$$

satisfy Eqs. 41, 42 and $\|u_{\varepsilon_n}^- - u_{\varepsilon_n}\|_\infty \leq \varepsilon_n$. Together, $\|u_{\varepsilon_n}^+ - u_{\varepsilon_n}^-\|_\infty \leq 2\varepsilon_n \to 0$ as $n \to \infty$. □

In view of Lemma A.1 there exists a subsequence of (ε_n) which we denote by (ε_k) such that $(u_{\varepsilon_k}^+)$ and $(u_{\varepsilon_k}^-)$ converge pointwise to an increasing function $u : [A, B] \to [0, 1]$. For fixed k and $t \in \mathbb{N}_0$ and any $\mathcal{F}_t$-measurable random variable X with values in $\{x_1, \ldots, x_N\}$, $x_i \in [A+\varepsilon_k, B-\varepsilon_k]$, let $b_{t+1}^{\varepsilon_k,\pm}(X)$ denote the $\mathcal{F}_{t+1}$–measurable random variables $b_{t+1}^{\varepsilon_k,+}(X)$ and $b_{t+1}^{\varepsilon_k,-}(X)$ with distribution

$$\text{law}(b_{t+1}^{\varepsilon_k,\pm}(X) \mid X = x) = \text{law}(b_\pm^{\varepsilon_k}(x)) \quad \text{for all } x \in \{x_1, \ldots, x_N\}.$$

Let $(R_t^{\varepsilon_k,\pm})_{t\in\mathbb{N}_0}$ denote the stopped random walks which start at $R_0^{\varepsilon_k,\pm} \in [A, B]$ and follows the dynamics

$$R_{t+1}^{\varepsilon_k,\pm} = \begin{cases} R_t^{\varepsilon_k,\pm} + b_{t+1}^{\varepsilon_k,\pm}(R_t^{\varepsilon_k,\pm}) & \text{if } t < \tau_0^\pm \\ R_{\tau_0}^{\varepsilon_k,\pm} & \text{if } t \geq \tau_0^\pm \end{cases}, \tag{46}$$

for the stopping times $\tau_0^\pm = \inf\{t \in \mathbb{N}_0 \mid R_t^{\varepsilon_k,\pm} \in [A, A+\varepsilon_k] \text{ or } R_t^{\varepsilon_k,\pm} \in [B-\varepsilon_k, B]\}$, where the infimum over the empty set is defined as $+\infty$. By construction, $R^{\varepsilon_k,\pm}$ are Markov processes with values in $(R_0^{\varepsilon_k,\pm} + \varepsilon_k\mathbb{N}) \cap [A, B]$ and transition probabilities

$$\mathbb{P}\left[R_{t+1}^{\varepsilon_k,\pm} = x + \varepsilon_k \mid R_t^{\varepsilon_k,\pm} = x\right] = p_\pm^{\varepsilon_k}(x),$$
$$\mathbb{P}\left[R_{t+1}^{\varepsilon_k,\pm} = x - \varepsilon_k \mid R_t^{\varepsilon_k,\pm} = x\right] = 1 - p_\pm^{\varepsilon_k}(x).$$

Lemma 2.3 specializes to the present context as follows. The proof is a straightforward modification.

Lemma 3.3 *Let $R^{\varepsilon_k,\pm}$ denote the stochastic processes which follow the dynamics Eq. 46 and let τ be a bounded stopping time. Then, we have*

$$c_0(R_\tau^{\varepsilon_k,+}) \leq R_0^{\varepsilon_k,+} \quad \textit{and} \quad c_0(R_\tau^{\varepsilon_k,-}) \geq R_0^{\varepsilon_k,-}.$$

Proof Suppose that $R_s^{\varepsilon_k,+}$ assumes the values $\{x_1,\dots,x_N\}$. Due to Eq. 46 we have

$$R_{s+1}^{\varepsilon_k,+} = 1_{\{s\geq\tau_0^+\}} R_s^{\varepsilon_k,+} + 1_{\{s+1\leq\tau_0^+\}} \sum_{n=1}^N 1_{\{R_s^{\varepsilon_k,+}=x_n\}} \left(x_n + b_{s+1}^{\varepsilon_k,+}(x_n)\right).$$

The local property of c_s yields

$$c_s(R_{s+1}^{\varepsilon_k,+}) = 1_{\{s\geq\tau_0^+\}} c_s(R_s^{\varepsilon_k,+}) + 1_{\{s+1\leq\tau_0^+\}} \sum_{n=1}^N 1_{\{R_s^{\varepsilon_k,+}=x_n\}} c_s\left(x_n + b_{s+1}^{\varepsilon_k,+}(x_n)\right). \tag{47}$$

Since $\operatorname{law}(x_n + b_{s+1}^{\varepsilon_k,+}(x_n) \mid \mathcal{F}_s) = \operatorname{law}(x_n + b_+^{\varepsilon_k}(x_n))$ a.s., Lemmas 3.1 and 3.2 yield

$$c_s(x_n + b_{s+1}^{\varepsilon_k,+}(x_n)) = c_0(x_n + b_+^{\varepsilon_k}(x_n)) \leq x_n. \tag{48}$$

We get $c_s(R_{s+1}^{\varepsilon_k,+}) \leq R_s^{\varepsilon_k,+}$ from Eqs. 47 and 48. We then proceed by backward induction as in the proof of Lemma 2.3. □

Remark 3.4 For any $R^{\varepsilon_k,\pm}$ of the form Eq. 46 starting at $R_0^{\varepsilon_k,\pm} \in \mathcal{I}_{\varepsilon_k}$, the stochastic processes $u_{\varepsilon_k}^\pm(R_t^{\varepsilon_k,\pm})$ are martingales. Indeed, since $\mathbb{P}[R_t^{\varepsilon_k,\pm} \in \mathcal{I}_{\varepsilon_k}] = 1$ for all $t \in \mathbb{N}_0$ we deduce from Eq. 42 that

$$\begin{aligned}\mathbb{E}\left[u_{\varepsilon_k}^\pm\left(R_{t+1}^{\varepsilon_k,\pm}\right) \mid \mathcal{F}_t\right] &= 1_{\{t<\tau_0^\pm\}} \mathbb{E}\left[u_{\varepsilon_k}^\pm\left(R_t^{\varepsilon_k,\pm} + b_{t+1}^{\varepsilon_k,\pm}(R_t^{\varepsilon_k,\pm})\right) \mid \mathcal{F}_t\right] \\ &\quad + 1_{\{t\geq\tau_0^\pm\}} \mathbb{E}\left[u_{\varepsilon_k}^\pm(R_t^{\varepsilon_k,\pm}) \mid \mathcal{F}_t\right] = u_{\varepsilon_k}^\pm\left(R_t^{\varepsilon_k,\pm}\right).\end{aligned} \tag{49}$$

Lemma 3.5 *The function $u : [A, B] \to [0, 1]$ is continuous and strictly increasing. For any random variable X taking at most two values $x_1, x_2 \in (A, B)$, we have*

$$c_0(X) = u^{-1} \circ \mathbb{E}[u(X)].$$

Proof Step 1. Suppose that X assumes the values $A < x_1 < x_2 < B$ with probabilities $0 < p_1, p_2 < 1$. Let

$$f_{\varepsilon_k}^+(x) := (u_{\varepsilon_k}^+)^{-1}\left\{p_1 u_{\varepsilon_k}^+(x_1 + x) + p_2 u_{\varepsilon_k}^+(x_2 + x)\right\}$$

be a sequence of functions defined on $[-\kappa, \kappa]$ for some $\kappa > 0$. The functions $f_{\varepsilon_k}^+$ are continuous, increasing and bounded by $x_1 - \kappa \leq f_{\varepsilon_k}^+(x) \leq x_2 + \kappa$ for all $x \in [-\kappa, \kappa]$. In view of Lemma A.1, there exists a subsequence of (ε_k), which we still denote by (ε_k), such that $f_{\varepsilon_k}^+$ converges pointwise to an increasing function $f^+ : [-\kappa, \kappa] \to [x_1 - \kappa, x_2 + \kappa]$. In this first step, we assume in addition that 0 is a continuity point of f^+. Define

$$Y_0^{\varepsilon_k} := (u_{\varepsilon_k}^+)^{-1}\left\{p_1 u_{\varepsilon_k}^+(x_1 + \varepsilon_k) + p_2 u_{\varepsilon_k}^+(x_2 + \varepsilon_k)\right\} \in [x_1 + \varepsilon_k, x_2 + \varepsilon_k]. \tag{50}$$

Here, we assume that k is large enough such that $x_2 + 2\varepsilon_k < B$. Let $R^{\varepsilon_k,+}$ be defined as the stochastic process which follows the dynamics Eq. 46 and starts at

$$R_0^{\varepsilon_k,+} := \min\left\{x \in \mathcal{I}_{\varepsilon_k} \mid x \geq Y_0^{\varepsilon_k}\right\} \in [x_1 + \varepsilon_k, x_2 + 2\varepsilon_k].$$

Define the stopping time

$$\sigma := \inf\left\{t \geq 0 \mid R_t^{\varepsilon_k,+} \in [x_1, x_1+\varepsilon_k) \text{ or } R_t^{\varepsilon_k,+} \geq x_2\right\}.$$

Due to Remark 3.4, the process $u^+_{\varepsilon_k}(R_t^{\varepsilon_k,+})$ is a martingale. Hence, the martingale convergence theorem yields that σ is a.s. finite. If $R_0^{\varepsilon_k,+} \geq x_2$ then $\mathbb{P}[R_\sigma^{\varepsilon_k,+} \geq x_2] = 1$. If $R_0^{\varepsilon_k,+} < x_2$ then $R_\sigma^{\varepsilon_k,+} < x_2 + \varepsilon_k$ and the martingale stopping theorem yields

$$p_1 u^+_{\varepsilon_k}(x_1+\varepsilon_k) + p_2 u^+_{\varepsilon_k}(x_2+\varepsilon_k) \leq u^+_{\varepsilon_k}(R_0^{\varepsilon_k,+}) = \mathbb{E}\left[u^+_{\varepsilon_k}(Y_\sigma^{\varepsilon_k,+})\right].$$

Hence, $\mathbb{P}[R_\sigma^{\varepsilon_k,+} \geq x_2] > p_2$. The dominating convergence theorem implies $T \in \mathbb{N}$ such that $\mathbb{P}[R_{\sigma\wedge T}^{\varepsilon_k,+} \geq x_2] \geq p_2$. Hence, for the bounded stopping time $\sigma_+^k := \sigma \wedge T$ we end up with

$$R_{\sigma_+^k}^{\varepsilon_k,+} \succeq_1 X. \tag{51}$$

Since 0 is a continuity point of f^+ we deduce from Lemma A.1

$$\left|(u^+_{\varepsilon_k})^{-1}\left\{p_1 u^+_{\varepsilon_k}(x_1+\varepsilon_k) + p_2 u^+_{\varepsilon_k}(x_2+\varepsilon_k)\right\} - (u^+_{\varepsilon_k})^{-1}\left\{p_1 u^+_{\varepsilon_k}(x_1) + p_2 u^+_{\varepsilon_k}(x_2)\right\}\right| \to 0$$

as $k \to \infty$, and therefore in combination with Eq. 50 and $|R_0^{\varepsilon_k,+} - Y_0^{\varepsilon_k}| \leq \varepsilon_k$, we deduce

$$\left|R_0^{\varepsilon_k,+} - (u^+_{\varepsilon_k})^{-1} \circ \mathbb{E}\left[u^+_{\varepsilon_k}(X)\right]\right| \to 0 \quad \text{as } k \to \infty. \tag{52}$$

Analogously, if 0 is a continuity point of f^- defined as a converging subsequence of

$$f^-_{\varepsilon_k}(x) := (u^-_{\varepsilon_k})^{-1}\left\{p_1 u^-_{\varepsilon_k}(x_1+x) + p_2 u^-_{\varepsilon_k}(x_2+x)\right\},$$

then there exist $R^{\varepsilon_k,-}$ of the form Eq. 46 starting at $R_0^{\varepsilon_k,-} \in \mathcal{I}_{\varepsilon_k}$ and a bounded stopping times σ_k^- such that $X \succeq_1 R_{\sigma_k^-}^{\varepsilon_k,-}$ and $|R_0^{\varepsilon_k,-} - (u^-_{\varepsilon_k})^{-1} \circ \mathbb{E}[u^-_{\varepsilon_k}(X)]| \to 0$.

*Step 2.*Suppose that X takes the values $A < x_1 < x_2 < B$ with probabilities $0 < p_1, p_2 < 1$, such that $c_0(X)$ is a continuity point of the function u. Fix $\delta > 0$. There exists $x \in (-\kappa, 0)$ such that X^δ taking the values $x_1^\delta = x_1 + x$ and $x_2^\delta = x_2 + x$ with probabilities p_1 and p_2 satisfies

(i) $\left|c_0(X) - c_0(X^\delta)\right| \leq \delta$,
(ii) $c_0(X^\delta)$ is a continuity point of u,
(iii) x is a continuity point of f^+,
(iv) $X \geq X^\delta$ and $||X^\delta - X||_\infty \leq \delta$.

Due to Step 1, there are $(R^{\varepsilon_k,+})$ and (σ_k^+) such that Eqs. 51 and 52 hold, i.e.,

$$R_{\sigma_k^+}^{\varepsilon_k,+} \succeq_1 X^\delta \quad \text{and} \quad \xi_k := \left|R_0^{\varepsilon_k,+} - (u^+_{\varepsilon_k})^{-1} \circ \mathbb{E}\left[u^+_{\varepsilon_k}(X^\delta)\right]\right| \to 0.$$

c_0 is monotone with respect to $\succeq_1$. Hence, Lemma 3.3 implies $R_0^{\varepsilon_k,+} \geq c_0(R_{\sigma_k^+}^{\varepsilon_k,+}) \geq c_0(X^\delta)$, from which we deduce

$$(u^+_{\varepsilon_k})^{-1} \circ \mathbb{E}\left[u^+_{\varepsilon_k}(X^\delta)\right] \geq c_0(X^\delta) - \xi_k.$$

Since $X \geq X^\delta$, X^δ assumes only two values, and $c_0(X^\delta)$ is a continuity point of u, we derive

$$\mathbb{E}\left[u(X)\right] \geq \mathbb{E}\left[u(X^\delta)\right] \geq u(c_0(X^\delta)), \quad \text{as } k \to \infty.$$

Finally, if we let δ tending to zero, it follows $||X^\delta - X||_\infty \to 0$ and $\left|c_0(X^\delta) - c_0(X)\right| \to 0$. Since $c_0(X)$ is a continuity point of u, we conclude $\mathbb{E}[u(X)] \geq u(c_0(X))$. Analogously, if we approximate X by X^δ from above and bound it by $R^{\varepsilon_k,-}$ from below, we get $\mathbb{E}[u(X)] \leq u(c_0(X))$. Together, we conclude $\mathbb{E}[u(X)] = u(c_0(X))$ for any random variable X taking at most two values in (A, B) and $c_0(X)$ is a continuity point of u.

Step 3. The function $u : [A, B] \to [0, 1]$ is continuous. Indeed, suppose by way of contradiction that there is $x \in (A, B]$ such that $u(x-) < u(x)$ (the case $u(x) < u(x+)$ works analogously). Since c_0 is $||\cdot||_\infty$-continuous and strictly increasing, there exist sequences (x_1^n) and (x_2^n) in (A, B) such that

(i) $x_1^n < x_2^n < x$, $x_1^n \nearrow x_1$ and $x_2^n \nearrow x_2 = x$,
(ii) $c_0(X^n)$ and $c_0(X)$ are continuity points of u, and
(iii) $||X^n - X||_\infty \to 0$,

where X and X^n assume the values $\{x_1, x_2\}$ and $\left\{x_1^n, x_2^n\right\}$ with probabilities $1/2$. Indeed, the function $x_1 \mapsto c_0(X)$ maps from (A, x_2) in (A, x_2). Since the set of discontinuity points of u in (A, x_2) is countable and c_0 is strictly increasing, there are at most countably many $x_1 \in (A, x_2)$ for which $c_0(X)$ is a discontinuity point of u. Hence, there is $x_1 \in (A, x_2)$ for which $c_0(X)$ is a continuity point of u. We then approximate X by X^n from below in $L^\infty(A, B)$ such that $c_0(X^n)$ are continuity points of u. Then, $c_0(X^n) \nearrow c_0(X)$ and Step 2 yields

$$\mathbb{E}[u(X^n)] = u(c_0(X^n)) \nearrow u(c_0(X)) = \mathbb{E}\left[u(X)\right],$$

in contradiction to $u(x_2^n) \nearrow u(x_2-) < u(x_2)$. Whence, u has to be continuous. In particular, Step 2 yields $\mathbb{E}[u(X)] = u(c_0(X))$ for all random variables X taking at most two values in (A, B).

Step 4. The function $u : [A, B] \to [0, 1]$ is strictly increasing. Indeed, by way of contradiction assume there exist $A < r < s < B$ such that $u(r) = u(s)$. Since u is not constant, there exists $t < r$ such that $u(t) < u(r)$ or $t > s$ such that $u(t) > u(s)$. In the second case, denote $v := \inf\{x > s \mid u(x) > u(s)\}$. By continuity of u one has $s \leq v < t$ and $u(r) = u(s) = u(v)$. Since c_0 is $||\cdot||_\infty$-continuous, there exists $w \in (v, t]$ such that $c_0(X) \leq v$, where X is a random variable taking the values r and w with probability $1/2$ each. If Y is a random variable that is equal to v and w with probabilities $1/2$, then $c_0(Y) > v$ and therefore, $u(c_0(Y)) > u(c_0(X))$. On the other hand, $u(c_0(Y)) = \mathbb{E}[u(X)] = \mathbb{E}[u(Y)] = u(c_0(X))$, which is a contradiction. The proof is completed. □

Lemma 2.5 specializes to the present context as follows.

Lemma 3.6 *Let $N \in \mathbb{N}$ and $X \in L^\infty(A, B)$ be a random variable taking 2^N different values with strictly positive probabilities. There exists a subsequence of (ε_k) (still denoted by (ε_k)), such that for any $k \in \mathbb{N}$ large enough, we may find stochastic processes $R^{\varepsilon_k,+}$, $R^{\varepsilon_k,-}$ of the form Eq. 46 as well as bounded stopping times σ_k^+ and σ_k^-, such that X satisfies*

$$R^{\varepsilon_k,+}_{\sigma_k^+} \succeq_1 X \succeq_1 R^{\varepsilon_k,-}_{\sigma_k^-}, \tag{53}$$

and

$$\lim_{k\to\infty} \left|u^+_{\varepsilon_k}(R_0^{\varepsilon_k,+}) - \mathbb{E}\left[u^+_{\varepsilon_k}(X)\right]\right| = \lim_{k\to\infty} \left|u^-_{\varepsilon_k}(R_0^{\varepsilon_k,-}) - \mathbb{E}\left[u^-_{\varepsilon_k}(X)\right]\right| = 0. \tag{54}$$

Sketch of the proof The proof is a straightforward modification of the proof of Lemma 2.5. In a first step, we define $\hat{Y}_N^{\varepsilon_k} = \hat{X}$ and

$$\hat{Y}_n^{\varepsilon_k} = (u_{\varepsilon_k}^+)^{-1} \circ \hat{\mathbb{E}}\left[u_{\varepsilon_k}^+(\hat{Y}_{n+1}^{\varepsilon_k} + \varepsilon_k) \mid \hat{\mathcal{F}}_n\right], \quad n = 0, \ldots, N-1, \tag{55}$$

on the filtered probability space $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_n)_{n=0}^N, \hat{\mathbb{P}})$. Here we assume that ε_k is small enough such that $\hat{Y}_n^{\varepsilon_k} \in (A+\varepsilon_k, B-\varepsilon_k)$ for all $n = 0, \ldots, N$. We then define $R^{\varepsilon_k,+}$ as the stochastic process which follows the dynamics Eq. 46 and starts at

$$R_0^{\varepsilon_k,+} = \min\left\{x \in \mathcal{I}_{\varepsilon_k} \mid x \geq \hat{Y}_0^{\varepsilon_k}\right\} \in (A, B).$$

The martingale stopping arguments given in the proof of Lemma 2.5 imply the existence of a bounded stopping time σ_k^+ such that

$$R_{\sigma_k^+}^{\varepsilon_k,+} \succeq_1 X.$$

Finally, there is a subsequence of (ε_k), which we still denote by (ε_k), such that $\hat{Y}_n^{\varepsilon_k}$ converges to $\hat{Y}_n$ for all $n = 0, \ldots, N-1$. By continuity of u, the arguments given in Step 3 of the proof of Lemma 2.5 imply $|u_{\varepsilon_k}^+(\hat{Y}_0^{\varepsilon_k}) - \mathbb{E}[u_{\varepsilon_k}^+(X)]| \to 0$. Since $|\hat{Y}_0^{\varepsilon_k} - R_0^{\varepsilon_k,+}| \leq \varepsilon_k$, we conclude Eq. 54. The proof for the lower bound works analogously.

Proof of the "only if"-part of Theorem 1.4 Let $[A, B]$ denote a compact interval in (a, b). Fix $X \in L^\infty(A, B)$ and $\delta > 0$. By $||\cdot||_\infty$-continuity of c_0, there exists X^δ taking 2^N different values, such that $X^\delta \in L^\infty(A, B)$, $||X^\delta - X||_\infty \leq \delta$ and $\left|c_0(X^\delta) - c_0(X)\right| \leq \delta$. Since c_0 is monotone with respect to $\succeq_1$ we deduce from the Lemmas 3.3 and 3.6 that

$$R_0^{\varepsilon_k,+} \geq c_0(R_{\sigma_k^+}^{\varepsilon_k,+}) \geq c_0(X^\delta) \geq c_0(R_{\sigma_k^-}^{\varepsilon_k,-}) \geq R_0^{\varepsilon_k,-}$$

and therefore

$$u_{\varepsilon_k}^+(R_0^{\varepsilon_k,+}) - \mathbb{E}\left[u_{\varepsilon_k}^+(X^\delta)\right] \geq u_{\varepsilon_k}^+(c_0(X^\delta)) - \mathbb{E}\left[u_{\varepsilon_k}^+(X^\delta)\right],$$
$$u_{\varepsilon_k}^-(c_0(X^\delta)) - \mathbb{E}\left[u_{\varepsilon_k}^-(X^\delta)\right] \geq u_{\varepsilon_k}^-(R_0^{\varepsilon_k,-}) - \mathbb{E}\left[u_{\varepsilon_k}^-(X^\delta)\right].$$

Letting k tending to infinity yields $u^{-1} \circ \mathbb{E}[u(X^\delta)] \geq c_0(X^\delta)$ and $c_0(X^\delta) \geq u^{-1} \circ \mathbb{E}[u(X^\delta)]$ and therefore $c_0(X^\delta) = u^{-1} \circ \mathbb{E}[u(X^\delta)]$. Letting δ converging to zero, it follows

$$c_0(X) = u^{-1} \circ \mathbb{E}[u(X)] \quad \text{for all } X \in L^\infty(A, B),$$

as u is uniformly continuous on $[A, B]$.

Hence, for any compact interval $[A, B] \subset (a, b)$ there is $u_{A,B} : [A, B] \to [0, 1]$ strictly increasing and continuous such that $c_0(X) = u_{A,B}^{-1} \circ \mathbb{E}\left[u_{A,B}(X)\right]$ for all $X \in L^\infty(A, B)$. We next give some standard arguments showing that $u_{A,B}$ on $[A, B]$ is uniquely determined up to affine transformations. Suppose there exist two strictly increasing, continuous functions $u, \tilde{u} : [A, B] \to \mathbb{R}$ such that

$$c_0(X) = u^{-1} \circ \mathbb{E}[u(X)] = \tilde{u}^{-1} \circ \mathbb{E}\left[\tilde{u}(X)\right] \quad \text{for all } X \in L^\infty(A, B). \tag{56}$$

Define the strictly increasing and continuous function $\psi(x) = \tilde{u} \circ u^{-1}(x)$. Then $\tilde{u}(x) = \psi \circ u(x)$ and Eq. 56 yields $\tilde{u} \circ u^{-1} \circ \mathbb{E}[u(X)] = \mathbb{E}\left[\tilde{u}(X)\right]$. This shows

$$\psi \circ \mathbb{E}[u(X)] = \mathbb{E}[\psi \circ u(X)] \quad \text{for all } X \in L^\infty(A, B), \tag{57}$$

and consequently $\psi(x) = \alpha x + \beta$ for some $\alpha > 0$ and $\beta \in \mathbb{R}$ (approximate ψ uniformly on compacts by polynomials). We therefore can extend $u_{A,B} : [A, B] \to \mathbb{R}$ with $c_0(X) =$

$u_{A,B}^{-1} \circ \mathbb{E}[u_{A,B}(X)]$ for all $X \in L^\infty(A, B)$ to $u_{A',B'} : [A', B'] \to \mathbb{R}$ such that $[A, B] \subset [A', B']$, $u_{A',B'}$ restricted to $[A, B]$ coincides with $u_{A,B}$ and $c_0(X) = u_{A',B'}^{-1} \circ \mathbb{E}[u_{A',B'}(X)]$ for all $X \in L^\infty(A', B')$. By exhausting (a, b) with compact intervals $[A^n, B^n] \subset (a, b)$, there is a continuous and strictly increasing function $u : (a, b) \to \mathbb{R}$ which is unique up to affine transformations and satisfies $c_0(X) = u^{-1} \circ \mathbb{E}[u(X)]$ for all $X \in L^\infty(a, b)$ with values in some compact interval $[A, B] \subset (a, b)$. By $|| \cdot ||_\infty$-continuity of c_0 it follows that $c_0(X) = u^{-1} \circ \mathbb{E}[u(X)]$ for all $X \in L^\infty(a, b)$ and the proof of Theorem 1.4 is completed. □

Appendix A: Helly's theorem

The following lemma is well-known. For the sake of completeness we give a proof.

Lemma A.1 *Let $f_n : [A, B] \to [0, 1]$ be a sequence of increasing, continuous functions. Then, there is a subsequence (f_{n_k}) and an increasing function $f : [A, B] \to [0, 1]$ such that*

$$f_{n_k}(x) \to f(x) \quad \text{for all } x \in [A, B] \text{ as } k \to \infty. \tag{58}$$

The function f has at most countably many discontinuity points. Moreover, $f_{n_k}(x_k) \to f(x)$ for any sequence $x_k \in [A, B]$ which converges to some continuity point $x \in [A, B]$ of f.

Remark A.2 Helly's theorem is usually stated as a convergence result only for the continuity points of the limiting function f. It was observed in Campi and Schachermayer [6] that one may also obtain convergence on the discontinuity points of f.

Proof Let $(z_j)_{j\in\mathbb{N}}$ be a sequence running through $\mathcal{I} := [A, B] \cap \mathbb{Q}$. Let $(f_{n_k^1})$ be a subsequence of (f_n) such that $f_{n_k^1}(z_1)$ converges to $f(z_1) \in [0, 1]$. Let $(f_{n_k^2})$ be a subsequence of $(f_{n_k^1})$ such that $f_{n_k^2}(z_2) \to f(z_2) \in [0, 1]$ and so on. Then, $f_{n_k^k}(x) \to f(x)$ for all $x \in \mathcal{I}$. The function f is increasing on $\mathcal{I}$, i.e., for any $x, y \in \mathcal{I}$ with $x \le y$ it follows $f(x) \le f(y)$. Therefore, f has at most countably many discontinuities $(z_j)_{j\ge1}$. Let $(f_{\xi_k^1})$ be a subsequence of $(f_{n_k^k})$ such that $f_{\xi_k^1}(x) \to f(x)$ for all $x \in \mathcal{I} \cup \{z_1\}$, $(f_{\xi_k^2})$ a subsequence of $(f_{\xi_k^1})$ such that $f_{\xi_k^2}(z_2) \to f(z_2)$ and so on. Define $f_{n_k} := f_{\xi_k^k}$ which is a subsequence of (f_n). Then, $f_{n_k}(x) \to f(x)$ for all $x \in \cup_{j\ge1}\{z_j\} \cup \mathcal{I}$. By construction, any $x \in [A, B] \setminus (\cup_{j\ge1}\{z_j\})$ is a continuity point of $f : \mathcal{I} \to [0, 1]$, whence we define $f(x) = \lim_{n\to\infty} f(x_n)$ for an arbitrary sequence $x_n \in \mathcal{I}$ converging to x. The function $f : [A, B] \to [0, 1]$ is increasing.

Let $x_k \in [A, B]$ be a sequence with limit x, and assume that f is continuous at x. Fix $\delta > 0$. There exist $y_1, y_2 \in \mathcal{I}$ with $y_1 < x < y_2$ such that $|f(y) - f(x)| \le \delta/4$ for all $y \in [y_1, y_2]$. Furthermore, there is $k_0 \in \mathbb{N}$ such that

$$y_1 \le x_k \le y_2, \quad |f_{n_k}(y_1) - f(y_1)| \le \frac{\delta}{4}, \quad |f_{n_k}(y_2) - f(y_2)| \le \frac{\delta}{4} \quad \text{for all } k \ge k_0.$$

Then

$$\begin{aligned} \left|f_{n_k}(x_k) - f(x)\right| &\le \left|f_{n_k}(y_2) - f(x)\right| + \left|f(x) - f_{n_k}(y_1)\right| \\ &\le \left|f_{n_k}(y_2) - f(y_2)\right| + \left|f(y_1) - f_{n_k}(y_1)\right| + \frac{\delta}{2} \le \delta \quad \text{for all } k \ge k_0. \end{aligned}$$

In particular, $f_{n_k}(x) \to f(x)$ for all $x \in [A, B]$ and f is increasing. The proof is completed. □

References

1. Alt, F.: Über die Messbarkeit des Nutzens. Zeitschrift für Nationalökonomie. Verlag von Julius Springer, Wien (1935)
2. Artzner, Ph., Delbaen, F., Eber, J.M., Heath, D.: Coherent risk measures. Math. Finance **9**(3), 203–228 (1999)
3. Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., Ku, H.: Coherent multiperiod risk adjusted values and Bellman's principle. Ann. Oper. Res. **152**, 5–22 (2007)
4. Bernoulli, D.: Specimen theoriae novae de mensura sortis. Commentarii *Academiae Scientarum Imperalis Petropolitanae*, vol. 5, pp. 175–192 (1738) [Translated by L. Sommer: *Econometrica* **22** (1954), pp. 23–36]
5. Bullen, P.S.: Handbook of Means and Their Inequalities. Kluwer, Dordrecht (2003)
6. Campi, L., Schachermayer, W.: A super-replication theorem in Kabanov's Model of Transaction Costs. Finance Stoch. **10**(4), 579–596 (2006)
7. Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., Montrucchio, L.: Risk measures: rationality and diversification (2009, preprint)
8. Cheridito, P., Delbaen, F., Kupper, M.: Dynamic monetary risk measures for bounded discrete-time processes. Electron. J. Probab. **11**, 57–106 (2006)
9. Cheridito, P., Kupper, M.: Composition of time-consistent dynamic monetary risk measures in discrete time (2007, preprint)
10. Cheridito, P., Li, T.: Dual characterization of properties of risk measures on Orlicz hearts. Math. Financial Econ. **2**(1), 29–55 (2008)
11. Cheridito, P., Stadje, M.: Time-inconsistency of VaR and time-consistent alternatives. Finance Res. Lett. **6**(1), 40–46 (2009)
12. Delbaen, F.: Coherent risk measures on general probability spaces. Advances in Finance and Stochastics. pp. 1–37. Springer, Berlin (2002)
13. Delbaen, F.: The structure of m-stable sets and in particular of the set of riskneutral measures. In: Yor, M., Emery, M. (eds.) In Memoriam Paul-André Meyer—Seminaire de Probabilités XXXIX, pp. 215–258. Springer, Berlin (2006)
14. Delbaen, F., Peng, S., Rosazza Gianin, E.: Representation of the penalty term of dynamic concave utilities (2008, preprint)
15. Ellsberg, D.: Risk, ambiguity and Savage axioms. Quart. J. Econ. **75**, 643–669 (1961)
16. de Finetti, B.: Sul concetto di media. Giornale Dell'Instituto Italiano Degli Attuari **2**, 369–396 (1931)
17. Föllmer, H., Penner, I.: Convex risk measures and the dynamics of their penalty functions. Stat. Decis. **24**(1), 61–96 (2006)
18. Föllmer, H., Schied, A.: Stochastic Finance. An Introduction in Discrete Time. de Gruyter Studies in Mathematics 27 (2002c)
19. Frittelli, M., Rosazza Gianin, E.: Law-invariant convex risk measures. Adv. Math. Econ. **7**, 33–46 (2005)
20. Gilboa, I., Schmeidler, D.: Maxmin expected utility with non-unique prior. J. Math. Econ. **18**, 141–153 (1989)
21. Gerber, H.U.: An Introdution to Mathematical Risk Theory. S.S. Huebner Foundation for Insurance Education, Wharton School, University of Pennsylvania, Philadelphia (1979)
22. Gerber, H.U.: On iterative Premium Calculation Principles. Sonderabdruck aus den Mitteilungen der Vereinigung schweizerischer Versicherungsmathematiker, vol. 74. Band, Heft 2 (1974)
23. Gerber, H.U., Goovaerts, M.: On the representation of additive principles of premium calculation. Scand. Actuarial J. **4**, 221–227 (1981)
24. Goovaerts, M., Kaas, R., Laeven, R., Tang, Q.: A comonotonic image of independence for additive risk measures. Insur. Math. Econ. **35**, 581–594 (2004)
25. Hardy, G., Littlewood, J.E., Pólya, G.: Inequalities, 2nd edn. Cambridge University press, Cambridge (1952)
26. Jouini, E., Schachermayer, W., Touzi, N.: Law invariant risk measures have the Fatou property. Adv. Math. Econ. **9**, 49–71 (2006)
27. Kolmogorov, A.N.: Sur la notion de la moyenne. Atti Della R. Accademia Nazionale Dei Lincei **12**, 388–391 (1930)
28. Kusuoka, S.: On law-invariant coherent risk measures. Adv. Math. Econ. **3**, 83–95 (2001)
29. Machina, M.: Dynamic consistency and non-expected utility models of choice under uncertainty. J. Econ. Lit. **27**, 1622–1668 (1989)
30. Maccheroni, F., Marinacci, M., Rustichini, A.: Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica **74**, 1447–1498 (2006)
31. Nagumo, M.: Über eine Klasse der Mittelwerte. Jpn. J. Math. **7**, 71–79 (1930)

32. Von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press, Princeton (1944)
33. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion, 2nd edn. Springer, Berlin (1994)
34. Savage, L.J.: The Foundations of Statistics. Wiley Publ. Stat., Wiley, New York (1954)
35. Weber, S.: Distribution-invariant risk measures, information, and dynamic consistency. Math. Finance **16**(2), 419–442 (2006)