Math Finan Econ (2010) 3:115–138
DOI 10.1007/s11579-010-0030-1

Foundations of continuous-time recursive utility: differentiability and normalization of certainty equivalents

Holger Kraft · Frank Thomas Seifried

Received: 27 January 2010 / Accepted: 27 May 2010 / Published online: 9 June 2010

Abstract This paper relates recursive utility in continuous time to its discrete-time origins and provides an alternative to the approach presented in Duffie and Epstein (Econometrica 60:353–394, 1992), who define recursive utility in continuous time via backward stochastic differential equations (stochastic differential utility). We show that the notion of Gâteaux differentiability of certainty equivalents used in their paper has to be replaced by a different concept. Our approach allows us to address the important issue of normalization of aggregators in non-Brownian settings. We show that normalization is always feasible if the certainty equivalent of the aggregator is of the expected utility type. Conversely, we prove that in general Lévy frameworks this is essentially also necessary, i.e. aggregators that are not of the expected utility type cannot be normalized in general. Besides, for these settings we clarify the relationship of our approach to stochastic differential utility and, finally, establish dynamic programming results.

Keywords Recursive utility · Stochastic differential utility · Lévy framework · Certainty equivalents · Normalization · Dynamic programming

JEL Classification D81 · D91 · C61

1 Introduction and motivation

In this paper, we study the existence and characterization of the continuous-time limit of a canonical discrete-time recursive utility model that has been popular in various applications,

H. Kraft (✉)
Department of Finance, Goethe-University, Grüneburgplatz 1, 60323 Frankfurt am Main, Germany
e-mail: holgerkraft@finance.uni-frankfurt.de

F. T. Seifried
Department of Mathematics, University of Kaiserslautern, Erwin-Schrödinger-Straße, 67663 Kaiserslautern, Germany
e-mail: seifried@mathematik.uni-kl.de

especially in asset pricing and certain macroeconomic settings. Its time-t_k continuation value is defined via the backward recursion

$$V_{t_k} = W\left(t_{k+1} - t_k, c_{t_k}, \mathfrak{m}\left(\mathcal{L}\left(V_{t_{k+1}}|\mathfrak{F}_{t_k}\right)\right)\right)$$

where W is called an aggregator, since it combines current consumption, c_{t_k}, and the certainty equivalent of future utility, $\mathfrak{m}(\mathcal{L}(V_{t_{k+1}}|\mathfrak{F}_{t_k}))$, to determine current utility, V_{t_k}.[1] We show that the continuous-time limit has the form that was conjectured by Duffie and Epstein [10], who instead showed the existence of a direct continuous-time formulation of the model in a Brownian setting. In particular, we prove that, in a Lévy framework, the continuous-time recursive utility index is given by

$$V_t = \mathbb{E}\left[\int_t^T \left\{ f(c_s, V_s) + \tfrac{1}{2}\sigma_s^2 A(V_s) + \int_{\mathbb{R}_*^\ell} J(V_s, \Psi_s(x))\vartheta(\mathrm{d}x) \right\} \mathrm{d}s \,\middle|\, \mathfrak{F}_t\right]$$

for an aggregator $(f, \mathfrak{m})$ with associated variance multiplier A and jump term J. Furthermore, we study whether the previous representation of V_t can be normalized, i.e. whether A and J can be transformed away at the same time. For instance, this is a relevant issue if decision problems with recursive utility are studied via Bellman equations.

To be able to address these points, we firstly provide an alternative approach to recursive utility in continuous time that directly relates the continuous-time formulation to its discrete-time counterpart via the condition

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \mathfrak{m}(\mathcal{L}(V_{t+s}|\mathfrak{F}_t)) = -f(c_t, V_t),$$

where $\mathcal{L}(V_{t+s}|\mathfrak{F}_t)$ denotes the conditional distribution of V_{t+s} given time-t information. To distinguish this concept from that of stochastic differential utility (SDU) as defined in Duffie and Epstein [10], we refer to it as continuous-time recursive utility (CRU). This alternative concept is also mentioned in Duffie and Epstein [10], but only to heuristically motivate SDU. In this motivation, they use an inappropriate concept of differentiability of certainty equivalents, namely Gâteaux differentiability.[2] We will introduce a suitable notion of differentiability that forms the basis for our formulation of CRU. We also clarify the connection to SDU, thereby also providing a natural discrete-time foundation for SDU. It is then shown that CRU is exactly the right approach to study the above-mentioned issue of normalization. In particular, we demonstrate that normalization is feasible if and only if the certainty equivalent of the aggregator is of the expected utility type. Therefore, aggregators that allow for normalization are behaviorally indistinguishable from aggregators with expected utility certainty equivalents. We wish to point out that our results also imply that aggregators which are not of the expected utility type cannot be normalized in general (e.g. Chew–Dekel preferences as in Proposition 5.10 and Example 8.2). Finally, using extended versions of the so-called stochastic Gronwall–Bellman inequalities, we prove that the results of Duffie and Epstein [10] on Hamilton–Jacobi–Bellman equations generalize to Lévy settings.

Recursive utility plays an increasingly important rôle in the literature on optimal consumption and portfolio choice. As mentioned above, normalization is necessary to obtain tractable Hamilton–Jacobi–Bellman equations in the dynamic programming approach of

[1] See, e.g., [15].

[2] We wish to point out that their definition of SDU does not rely on this motivation. Therefore, none of their formal results is affected.

Bhamra et al. [3], Benzoni [2], Fisher and Gilles [17] and Bhamra et al. [4], among others. Moreover, normalization is crucial for addressing questions such as the existence of recursive utility indices and the representation of recursive utility in certainty equivalent form, which can be regarded as a different type of normalization. Consequently, normalization is also relevant for the utility gradient approach pioneered by Duffie and Skiadas [12] and extended in Schroder and Skiadas [23–25] when this method is applied to SDU.[3] In a different context, Ma [21] provides an existence result for SDU in a finite-intensity Lévy framework assuming a normalized variance multiplier. However, there are only few papers studying the transition of recursive utility from discrete to continuous time. To the best of our knowledge, apart from Duffie and Epstein [10] only Svensson [29] looks at a related issue by presenting a heuristic dynamic programming approach based on a continuous-time limit. Skiadas [28] provides an intuitive interpretation of the impact of jumps on recursive utility and studies their effects in the presence of ambiguity. Finally, for axiomatic foundations of recursive utility we refer to Kreps and Porteus [19] and Skiadas [26], and to Skiadas [27] for a general overview.

Since subsequent work adopted the notion of Gâteaux differentiability, it is important to point out that this notion has to be replaced by another concept such as the one proposed in our paper. This point is highlighted by Example 5.1 and resolved in Sect. 5. Interestingly, the local gradient representations computed by Duffie and Epstein [10] can be interpreted as U-derivatives in the sense of our paper.

To summarize, the main contributions of this paper are the following: Firstly, we present an alternative approach to recursive utility in continuous time which is directly related to its discrete-time foundations. For this purpose, we secondly introduce a novel notion of differentiability of certainty equivalents. Thirdly, we use our approach to clarify the crucial issue of normalization and show that normalization is essentially feasible if and only if the certainty equivalent is of the expected utility type. Fourthly, we establish a dynamic programming result for the maximization of recursive utility in a Lévy framework.

The remainder of this paper is organized as follows: In Sect. 2, we fix our mathematical framework and introduce some terminology. Section 3 briefly summarizes the fundamental concepts of recursive utility in discrete time, setting the basis for the transition to continuous time in Sect. 4, which contains our definition of CRU. Additionally, we show why differentiability is relevant. Section 5 then thoroughly analyzes this important point. In Sect. 6, we study CRU in a Lévy framework and clarify its relationship to SDU. Section 7 presents our results on normalization. In Sect. 8, a verification theorem is derived, and Sect. 9 concludes. Appendix A contains some of the proofs, and Appendix B collects stochastic Gronwall–Bellman results.

2 Mathematical setting and notation

We let $(\Omega, \mathfrak{F}, \mathbb{P})$ be a probability space endowed with a filtration $\{\mathfrak{F}_t\}_{t\in[0,T]}$ satisfying the usual conditions of completeness and right-continuity such that $\mathfrak{F}_0$ is $\mathbb{P}$-trivial. Moreover, we fix a set $\mathcal{C} \subseteq \mathbb{R}^k$ of feasible consumption rates and a subinterval $\mathcal{V} \subseteq \mathbb{R}$ of the real line. We denote by C a class of predictable $\mathcal{C}$-valued processes with time horizon $[0, T]$, which we take as a model for the consumption processes to be ranked.

Sometimes we restrict attention to a *Lévy setting* by which we mean the following: We assume as given a standard Wiener process $B = \{B_t\}_{t\in[0,T]}$ and a Poisson random measure ν on $(\mathbb{R}^\ell_*, \mathfrak{B}(\mathbb{R}^\ell_*))$ with intensity ϑ where $\mathbb{R}^\ell_* \triangleq \mathbb{R}^\ell \setminus \{0\}$. The associated compensated random

[3] Normalization transforms can also be used to simplify the analysis of translation-invariant and homothetic aggregators.

measure is denoted by $\tilde{\nu}$, i.e. $\tilde{\nu}(\mathrm{d}t, \mathrm{d}x) = \nu(\mathrm{d}t, \mathrm{d}x) - \mathrm{d}t\vartheta(\mathrm{d}x)$. We suppose that the underlying filtration $\{\mathfrak{F}_t\}_{t\in[0,T]}$ is generated by W, ν and the class of $\mathbb{P}$-negligible sets. Further, we say that a càdlàg process $V = \{V_t\}_{t\in[0,T]}$ is *regular* if

$$\mathrm{d}V_t = \xi_t \mathrm{d}t + \sigma_t \mathrm{d}W_t + \int_{\mathbb{R}^\ell_*} \Psi_t(x)\tilde{\nu}(\mathrm{d}t, \mathrm{d}x) \tag{1}$$

where $\xi = \{\xi_t\}_{t\in[0,T]}$ and $\sigma = \{\sigma_t\}_{t\in[0,T]}$ are progressive processes and $\{\Psi_t(\,\cdot\,)\}_{t\in[0,T]}$ is a predictable[4] process that satisfy the integrability conditions

$$\mathbb{E}\left[\int_0^T |\xi_t|^p \mathrm{d}t\right] < \infty,\ \mathbb{E}\left[\int_0^T |\sigma_t|^p \mathrm{d}t\right] < \infty,\ \mathbb{E}\left[\int_{[0,T]\times\mathbb{R}^\ell_*} |\Psi_t(x)|^p \mathrm{d}t\vartheta(\mathrm{d}x)\right] < \infty \tag{2}$$

for any $p \in [1, \infty)$. Finally, we take C to be a class of predictable $\mathcal{C}$-valued processes $c = \{c_t\}_{t\in[0,T]}$ with

$$\mathbb{E}\left[\int_0^T |c_t|^p \mathrm{d}t\right] < \infty \quad \text{for all } p \in [1, \infty).$$

3 Recursive utility in discrete time: a brief review

We are interested in a mapping $\mathfrak{u} : \mathsf{C} \to \mathbb{R}, c \mapsto \mathfrak{u}(c)$ that ranks consumption streams in such a way that $\mathfrak{u}(c) \geq \mathfrak{u}(c')$ if and only if c is weakly preferred to c'. The notion of recursive utility provides a paradigm to construct such a functional via a so-called continuation value process V^c associated to c by setting

$$\mathfrak{u}(c) \triangleq V^c_0 \quad \text{for every consumption stream } c.$$

The mapping $\mathfrak{u}$ is referred to as a *recursive utility function*, and $\mathfrak{u}(c)$ is the *utility index* of c. In the following, we recapitulate the exact definition of continuation value processes. Let $[t_0, t_1, \ldots, t_N]$ be a partition of $[0, T]$ with $t_0 = 0, t_N = T$, and suppose that $c = \{c_{t_k}\}_{k=0,1,\ldots,N}$ is a discrete-time *deterministic* consumption stream. Then $V^c = \{V^c_{t_k}\}_{k=0,1,\ldots,N}$ is defined by means of the backward recursion

$$V^c_{t_k} \triangleq W\left(t_{k+1} - t_k, c_{t_k}, V^c_{t_{k+1}}\right) \quad \text{for } k = N-1, \ldots, 0 \text{ with } V^c_{t_N} = 0. \tag{3}$$

Here, the mapping $W : [0, \infty) \times \mathcal{C} \times \mathcal{V} \to \mathcal{V}, (\Delta, c, v) \mapsto W(\Delta, c, v)$ is of class C^0 with $W(0, c, v) = v$ for $c \in \mathcal{C}, v \in \mathcal{V}$, and describes the *temporal aggregation* of present consumption c_{t_k} and the value $V^c_{t_{k+1}}$ of future consumption outstanding. In the presence of *randomness*, the quantity $V^c_{t_{k+1}}$ is not known as of time t_k. As a substitute, the agent may resort to its conditional distribution $\mathcal{L}(V^c_{t_{k+1}}|\mathfrak{F}_{t_k})$ given the information available to her at t_k, which is a *lottery* on future utility. Thus as a further ingredient a certainty equivalent $\mathfrak{m}$ is required, and (3) canonically generalizes to

$$V^c_{t_k} \triangleq W\left(t_{k+1} - t_k, c_{t_k}, \mathfrak{m}(\mathcal{L}(V^c_{t_{k+1}}|\mathfrak{F}_{t_k}))\right) \quad \text{for } k = N-1, \ldots, 0 \text{ with } V^c_{t_N} = 0, \tag{4}$$

where $\mathfrak{u}(c) \triangleq V^c_0$ is a deterministic quantity. Formally we state

[4] This means that $\Psi : [0, T] \times \Omega \times \mathbb{R}^\ell_* \to \mathbb{R}$ is $\mathfrak{P} \otimes \mathfrak{B}(\mathbb{R}^\ell_*)$-measurable, see Jacod and Shiryaev [18].

Definition 3.1 (*Certainty equivalent*) Let $\mathcal{M}_1(\mathcal{V})$ denote the set of probability measures on the Borel σ-field $\mathfrak{B}(\mathcal{V})$ of $\mathcal{V}$ with moments of all orders. Then a functional

$$\mathfrak{m} : \mathcal{M}_1(\mathcal{V}) \to \mathbb{R}, \quad \mu \mapsto \mathfrak{m}(\mu),$$

is called a *certainty equivalent* on $\mathcal{V}$ if[5] $\mathfrak{m}(\delta_v) = v$ for all $v \in \mathcal{V}$. If there exists a strictly increasing polynomially bounded C^2 function $h : \mathcal{V} \to \mathbb{R}$ such that

$$\mathfrak{m}(\mu) = h^{-1}\left(\int_{\mathcal{V}} h \mathrm{d}\mu\right) \text{ for all } \mu \in \mathcal{M}_1(\mathcal{V}),$$

then $\mathfrak{m}$ is said to be an *expected utility*, or *EU*, certainty equivalent. If h is the identity mapping, then $\mathfrak{m}$ is said to be *risk-neutral*.

The pair $(W, \mathfrak{m})$ completely describes an agent's preferences over discrete-time stochastic consumption streams via the associated recursive utility function $\mathfrak{u}$.

Definition 3.2 (*Discrete-time aggregator*) Suppose that $W : [0, \infty) \times \mathcal{C} \times \mathcal{V} \to \mathcal{V}$ is of class C^0 with $W(0, c, v) = v$ for all $c \in \mathcal{C}$ and $v \in \mathcal{V}$ and let $\mathfrak{m}$ be a certainty equivalent on $\mathcal{V}$. Then the pair $(W, \mathfrak{m})$ is said to be a *discrete-time aggregator* on $\mathcal{V}$.

If $V^c = \{V^c_{t_k}\}_{k=0,1,\ldots,N}$ is such that $\mathbb{E}[|V^c_{t_k}|^p] < \infty$ for all $k = 0, 1, \ldots, N$, $p \in [1, \infty)$ and V^c satisfies (4), then we refer to V^c as the *continuation value process* of c under $(W, \mathfrak{m})$.

We are now going to address the issue of normalization in discrete time. Recall that the interpretation of recursive utility indices is ordinal rather than cardinal. Hence, if $\bar{\mathcal{V}} \subseteq \mathbb{R}$ is an interval, the function $\Phi : \mathcal{V} \to \bar{\mathcal{V}}$ is strictly increasing, and we set

$$\bar{\mathfrak{u}} : \mathsf{C} \to \mathbb{R}, \quad \bar{\mathfrak{u}}(c) \triangleq \Phi(\mathfrak{u}(c)),$$

then $\bar{\mathfrak{u}}$ is a recursive utility function describing the same preference structure as $\mathfrak{u}$. In this situation, we say that $\mathfrak{u}$ and $\bar{\mathfrak{u}}$ are *equivalent*. Two discrete-time aggregators $(W, \mathfrak{m})$ and $(\bar{W}, \bar{\mathfrak{m}})$ with associated recursive utility functions $\mathfrak{u}$ and $\bar{\mathfrak{u}}$ are said to be *ordinally equivalent* if $\mathfrak{u}$ and $\bar{\mathfrak{u}}$ are equivalent.

Proposition 3.3 (Transformation of aggregators) *Let* $(W, \mathfrak{m})$ *and* $(\bar{W}, \bar{\mathfrak{m}})$ *be discrete-time aggregators on* $\mathcal{V}$ *and* $\bar{\mathcal{V}}$ *and suppose that*[6]

$$\mathfrak{m}(\mu) = \Phi^{-1}\left(\bar{\mathfrak{m}}(\mu^{\Phi})\right) \quad \textit{for all } \mu \in \mathcal{M}_1(\mathcal{V})$$

and

$$W(\Delta, c, v) = \Phi^{-1}\left(\bar{W}(\Delta, c, \Phi(v))\right) \quad \textit{for } \Delta \geq 0,\ c \in \mathcal{C},\ v \in \mathcal{V}$$

for some strictly increasing polynomially bounded function $\Phi : \mathcal{V} \to \bar{\mathcal{V}}$ *with* $\Phi(0) = 0$. *Then* $(W, \mathfrak{m})$ *and* $(\bar{W}, \bar{\mathfrak{m}})$ *are ordinally equivalent.*

Proof Let $c = \{c_{t_k}\}_{k=0,1,\ldots,N}$ be a consumption process, let $V^c = \{V^c_{t_k}\}_{k=0,1,\ldots,N}$ denote the corresponding continuation value process, and let $\bar{V}^c$ be given by $\bar{V}^c_{t_k} \triangleq \Phi(V^c_{t_k})$ for

5 δ_v denotes the Dirac point measure located at v.

6 $\mu^{\Phi} \triangleq \mu \circ \Phi^{-1} \in \mathcal{M}_1(\bar{\mathcal{V}})$ denotes the image of $\mu \in \mathcal{M}_1(\mathcal{V})$ under the Borel mapping $\Phi : \mathcal{V} \to \bar{\mathcal{V}}$.

$k = 0, 1, \ldots, N$. By the L^p-contraction property of conditional expectations and the polynomial boundedness assumption, we have $\mathbb{E}[|\bar{V}^c_{t_k}|^p] < \infty, k = 0, 1, \ldots, N, p \in [1, \infty)$. Substituting into the recursion (4), we obtain

$$\begin{aligned}\bar{V}^c_{t_k} &= \Phi\left(V^c_{t_k}\right) = \Phi\left(W\left(t_{k+1} - t_k, c_{t_k}, \mathfrak{m}\left(\mathcal{L}\left(V^c_{t_{k+1}}|\mathfrak{F}_{t_k}\right)\right)\right)\right)\\ &= \bar{W}\left(t_{k+1} - t_k, c_{t_k}, \bar{\mathfrak{m}}\left(\mathcal{L}\left(V^c_{t_{k+1}}|\mathfrak{F}_{t_k}\right)^{\Phi}\right)\right) = \bar{W}\left(t_{k+1} - t_k, c_{t_k}, \bar{\mathfrak{m}}\left(\mathcal{L}\left(\bar{V}^c_{t_{k+1}}|\mathfrak{F}_{t_k}\right)\right)\right)\end{aligned}$$

for $k = N - 1, \ldots, 0$ and $\bar{V}^c_{t_N} = 0$ since $\Phi(0) = 0$. Hence, $\bar{V}^c$ is the continuation value process of c under $(\bar{W}, \bar{\mathfrak{m}})$, and the corresponding recursive utility functions satisfy

$$\bar{\mathfrak{u}}(c) = \bar{V}^c_0 = \Phi(V^c_0) = \Phi(\mathfrak{u}(c)).$$

Since the consumption stream c is arbitrary, the claim follows. □

Let $(W, \mathfrak{m}_h)$ be an aggregator with an EU certainty equivalent induced by a function h. Then, by taking $\Phi \triangleq h - h(0)$, we obtain

Corollary 3.4 (Normalization of EU certainty equivalents) *Every discrete-time aggregator with an EU certainty equivalent is ordinally equivalent to an aggregator whose certainty equivalent is risk-neutral.*

4 From discrete to continuous time

Duffie and Epstein [10] use a heuristic limiting argument to motivate their SDU approach to recursive utility in continuous time.[7] Recall that SDU is rigorously defined in continuous time and does not rely on this discrete-time motivation. In this section, we provide an alternative definition of recursive utility in continuous time via CRU. Our approach directly relates recursive utility in continuous time to its discrete-time counterpart. We fix a discrete-time aggregator $(W, \mathfrak{m})$ and a left-continuous consumption process $c = \{c_t\}_{t\in[0,T]}$. For a partition $\pi = [t_0, \ldots, t_N]$ with $t_0 = 0, t_N = T$ write $|\pi| \triangleq \max_{k=0,\ldots,N-1}(t_{k+1} - t_k)$. Then Eq. 4 leads to the requirement that the continuation value process $V^c = \{V^c_t\}_{t\in[0,T]}$ satisfies

$$V^c_{t_k} = W\left(t_{k+1} - t_k, c_{t_k}, \mathfrak{m}\left(\mathcal{L}\left(V^c_{t_{k+1}}|\mathfrak{F}_{t_k}\right)\right)\right) + o(|\pi|) \quad \text{for } k = N - 1, \ldots, 0,\ \ V^c_T = 0.$$

On the other hand, we have $V^c_{t_k} = W(0, c_{t_k}, V^c_{t_k})$ for $k = N - 1, \ldots, 0$, so with $t = t_k$ and $\Delta = t_{k+1} - t_k$ we obtain from a Taylor expansion

$$\begin{aligned}0 &= W\left(\Delta, c_t, \mathfrak{m}\left(\mathcal{L}\left(V^c_{t+\Delta}|\mathfrak{F}_t\right)\right)\right) - W\left(0, c_t, V^c_t\right) + o(|\pi|)\\ &= \frac{\partial W}{\partial \Delta}\left(0, c_t, V^c_t\right)\Delta + \frac{\partial W}{\partial v}\left(0, c_t, V^c_t\right)\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \mathfrak{m}\left(\mathcal{L}\left(V^c_{t+s}|\mathfrak{F}_t\right)\right)\Delta + o(|\pi|),\end{aligned}$$

provided W is of class C^1 with $\frac{\partial W}{\partial v}(0, c, v) > 0$ for $c \in \mathcal{C}$, $v \in \mathcal{V}$, and the real function

$$s \mapsto \mathfrak{m}\left(\mathcal{L}\left(V^c_{t+s}|\mathfrak{F}_t\right)\right) \text{ is a.s. differentiable at } s = 0. \tag{5}$$

In this case, the continuation value process V^c must satisfy[8]

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \mathfrak{m}\left(\mathcal{L}\left(V^c_{t+s}|\mathfrak{F}_t\right)\right) = -f\left(c_t, V^c_t\right) \text{ a.s., } V^c_T = 0, \tag{6}$$

[7] See (16) in Duffie and Epstein [10].

[8] This relation can also be found in Epstein [14] and Duffie and Epstein [10].

where $f : \mathcal{C} \times \mathcal{V} \to \mathbb{R}$ is given by

$$f(c, v) \triangleq \frac{\frac{\partial W}{\partial \Delta}(0, c, v)}{\frac{\partial W}{\partial v}(0, c, v)}. \tag{7}$$

The above derivation parallels the one given by Duffie and Epstein [10] and is not completely rigorous. However, since our focus is on normalization and an appropriate notion of differentiability, we are mainly interested in the commonly accepted Eqs. 6 and 7. A rigorous proof of convergence is beyond the scope of this paper and is left for future research. For instance, it would be interesting to relate the limit of the discrete-time backward recursion to recent results on the numerical simulation of backward stochastic differential equations (see, e.g. [5]).[9]

Observe that (5) requires differentiability in the sense of ordinary calculus. However, to study this condition in detail, one needs a chain rule involving two different types of non-standard derivatives. This is addressed in Sect. 5. At this point we give

Definition 4.1 (*Continuous-time aggregator, continuation value process*) Let $f : \mathcal{C} \times \mathcal{V} \to \mathbb{R}$ be Borel measurable and let $\mathfrak{m}$ be a certainty equivalent on $\mathcal{V}$. Then the pair $(f, \mathfrak{m})$ is called a *continuous-time aggregator* on $\mathcal{V}$. If $c \in \mathsf{C}$ is a consumption process and $V^c = \{V_t^c\}_{t\in[0,T]}$ is a $\mathcal{V}$-valued semimartingale with $\mathbb{E}[\sup_{t\in[0,T]} |V_t^c|^p] < \infty$ for all $p \in [1, \infty)$ such that the differentiability condition (5) holds and that (6) is satisfied for a.e. $t \in [0, T]$, then V^c is called a *continuation value process* of c.[10]

Definition 4.1 is formulated for an abstract aggregator $(f, \mathfrak{m})$ and is valid even without referring to a discrete-time model. However, if f is derived from a discrete-time aggregator W via (7), then the above discussion justifies that we consider $(f, \mathfrak{m})$ as the continuous-time analog of $(W, \mathfrak{m})$. Thus Definition 4.1 has the advantage that it preserves the intuitive interpretation of aggregators. We now return to the construction of a continuous-time aggregator from its discrete-time analog. Equation 7 yields a general method to determine f from a discrete-time aggregator $(W, \mathfrak{m})$, as illustrated in the following

Example 4.2 (Epstein–Zin preferences) Let $h, u : (0, \infty) \to \mathbb{R}$, $h(v) \triangleq \frac{1}{\rho} v^\rho$ and $u(v) \triangleq \frac{1}{\gamma} v^\gamma$ for $\rho, \gamma < 1$, and define

$$W(\Delta, c, v) \triangleq h\left(\mathrm{e}^{-\alpha\Delta} u(h^{-1}(v)) + u(c)\Delta\right) \quad \text{for } \Delta \geq 0,\ c, v > 0.$$

Then straightforward computations using Eq. 7 show that

$$f(c, v) = \frac{h'(h^{-1}(v))}{u'(h^{-1}(v))}\left[u(c) - \alpha u(h^{-1}(v))\right] = -\beta_1 v^{1-\frac{\gamma}{\rho}}\left[c^\gamma - \beta_2 v^{\frac{\gamma}{\rho}}\right], \quad c, v > 0,$$

where $\beta_1 \triangleq \frac{1}{\gamma}\rho^{1-\frac{\gamma}{\rho}}$ and $\beta_2 \triangleq \alpha\rho^{\frac{\gamma}{\rho}}$.

If for each consumption process c there exists a unique corresponding continuation value process V^c, then we can define recursive utility in continuous time as follows.

Definition 4.3 (*Continuous-time recursive utility, CRU function*) Let $(f, \mathfrak{m})$ be a continuous-time aggregator on $\mathcal{V}$ and suppose that for each consumption process $c \in \mathsf{C}$ there exists an a.s. uniquely determined continuation value process $V^c = \{V_t^c\}_{t\in[0,T]}$. Then

[9] We thank an anonymous referee for pointing this out.

[10] The space $[0, T]$ is endowed with the Lebesgue measure.

we define the corresponding *continuous-time recursive utility function*, or *CRU function*, by setting

$$\mathfrak{u} : \mathsf{C} \to \mathbb{R}, \quad \mathfrak{u}(c) \triangleq V_0^c$$

and say that $(f, \mathfrak{m})$ *generates* the CRU function $\mathfrak{u}$.

5 Differentiability of certainty equivalents

This section addresses the question of when the differentiability condition (5) is satisfied. To give (5) a precise meaning, it is clear that we need a suitable chain rule. Therefore, we have to clarify what it means for

$$\mathfrak{m} : \mathcal{M}_1(\mathcal{V}) \to \mathcal{V} \text{ to be differentiable,} \tag{8}$$

and we have to clarify what it means for the function

$$h \mapsto \mathfrak{L}(V_{t+h}^c | \mathfrak{F}_t) \text{ to be differentiable.} \tag{9}$$

Recall that in Duffie and Epstein [10] differentiability in (8) is taken in the sense of Gâteaux derivatives[11] on the convex set $\mathcal{M}_1(\mathcal{V})$ of probability distributions on $\mathcal{V}$. We however wish to stress that this is inappropriate. Indeed, consider the following

Example 5.1 (Gâteaux differentiability is inappropriate) Let for an arbitrary $\alpha \in (0, 1)$ the certainty equivalent $\mathfrak{m}_\alpha$ on $\mathbb{R}$ be given by

$$\mathfrak{m}_\alpha : \mathcal{M}_1(\mathbb{R}) \to \mathbb{R}, \quad \mathfrak{m}_\alpha(\mu) \triangleq \sup\{x \in \mathbb{R} : \mu((-\infty, x]) \leq \alpha\},$$

which assigns to each probability distribution its α-quantile. For $v \in \mathbb{R}$ and a finite signed measure ρ with $\rho(\mathbb{R}) = 0$ observe that

$$(\delta_v + h\rho)((-\infty, x]) = h\rho((-\infty, x]) \leq |h| \, \|\rho\| \leq \alpha \quad \text{for all } x \in (-\infty, v)$$

and

$$(\delta_v + h\rho)((-\infty, v]) = 1 + h\rho((-\infty, v]) \geq 1 - |h| \, \|\rho\| > \alpha \quad \text{if } 0 < h < \delta,$$

where $\|\rho\|$ denotes the total variation of ρ and $\delta \triangleq \frac{1}{\|\rho\|} \min(\alpha, 1 - \alpha)$. Hence, we have

$$\mathfrak{m}_\alpha(\delta_v + h\rho) = v \quad \text{for sufficiently small } h > 0.$$

Thus $\mathfrak{m}$ is smooth at certainty in the sense of Duffie and Epstein [10], and its Gâteaux derivative vanishes identically.[12] In particular, the local gradient representation in the sense of Duffie and Epstein [10] violates Eq. 13 below.[13]

Intuitively, the failure of Gâteaux differentiability for an α-quantile certainty equivalent is due to the fact that it is insensitive at deterministic payoffs. In the light of this example, it is apparent that Eq. 13 below is *not* valid for Gâteaux derivatives. Somewhat oversimplified, the problem is due to the fact that for two point masses δ_v and δ_w we have

$$\delta_v + \delta_w \neq \delta_{v+w}.$$

[11] Duffie and Epstein [10] require $\mathfrak{m}$ to be 'smooth at certainty', i.e. Gâteaux differentiable at point masses.

[12] As pointed out by an anonymous referee, the same argument can be used to show that the Fréchet derivative is also zero.

[13] See also the discussion following Eq. 14 in that paper.

More precisely, our notion of U-differentiability in Definition 5.4 is based on the linear structure of the *underlying* space, whereas the notion of Gâteaux differentiability is based on the linear structure of the space of *probability measures*.

Therefore, we now provide novel definitions of differentiability for both (8) and (9); we first address (9). Let us fix a class U of polynomially bounded C^2 test functions defined on $\mathcal{V}$. The rôle of U is technical and will become clear from Definitions 5.2 and 5.4. The set U may depend on the aggregator $(f, \mathfrak{m})$ under consideration.

Definition 5.2 (U-*differentiability*) A family $\{\mu_s\}_{s\geq 0} \subseteq \mathcal{M}_1(\mathcal{V})$ of probability measures on $\mathcal{V}$ is said to be U-*differentiable* at $s = 0$ if $\mu_0 = \delta_v$ for some $v \in \mathcal{V}$ and

$$\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \int_{\mathcal{V}} u \mathrm{d}\mu_s = \lim_{s\downarrow 0} \frac{1}{s}\left[\int_{\mathcal{V}} u\mathrm{d}\mu_s - u(v)\right] \text{ exists for any } u \in \mathrm{U}. \tag{10}$$

In this case, we refer to the operator

$$\dot{\mu}_0 : \mathrm{U} \to \mathbb{R}, \quad u \mapsto \dot{\mu}_0[u] \triangleq \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \int_{\mathcal{V}} u\mathrm{d}\mu_s$$

as the U-*derivative* of $\{\mu_s\}_{s\geq 0}$ at $s = 0$.

As an important application, let us consider a Lévy setting. Assume that each $u \in \mathrm{U}$ is such that u' and u'' are bounded, and $\mu_s = \mathcal{L}(V_{t+s}|\mathfrak{F}_t)$, $s \geq 0$, where $V = \{V_s\}_{s\geq 0}$ is regular and given by (1), i.e. $\mathrm{d}V_s = \xi_s\mathrm{d}s + \sigma_s\mathrm{d}W_s + \int_{\mathbb{R}^\ell_*} \Psi_s(x)\tilde{\nu}(\mathrm{d}s, \mathrm{d}x)$. Then Itô's formula yields

$$\mathrm{d}u(V_s) = \left\{\xi_t u'(V_t) + \tfrac{1}{2}\sigma_t^2 u''(V_t) + \int_{\mathbb{R}^\ell_*} [u(V_t + \Psi_t(x)) - u(V_t) - u'(V_t)\Psi_t(x)]\, \vartheta(\mathrm{d}x)\right\} \mathrm{d}t + \mathrm{d}Z_t,$$

where $Z = \{Z_t\}_{t\in[0,T]}$ is a martingale due to the integrability conditions stated in (2). By Fubini's theorem for conditional expectations as stated in Proposition B.2 and the fundamental theorem of calculus, which is justified by (2), we obtain for a.e. $t \in [0, T]$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \mathbb{E}[u(V_{t+s})|\mathfrak{F}_t] &= \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \mathbb{E}\left[\int_t^{t+s} \left\{\xi_r u'(V_r) + \tfrac{1}{2}\sigma_r^2 u''(V_r)\right.\right.\\ &\quad \left.\left.+ \int_{\mathbb{R}^\ell_*} [u(V_r + \Psi_r(x)) - u(V_r) - u'(V_r)\Psi_r(x)]\, \vartheta(\mathrm{d}x)\right\} \mathrm{d}r \,\middle|\, \mathfrak{F}_t\right]\\ &= \lim_{s\downarrow 0} \frac{1}{s} \int_t^{t+s} \mathbb{E}\left[\xi_r u'(V_r) + \tfrac{1}{2}\sigma_r^2 u''(V_r)\right.\end{aligned}$$

$$+ \int_{\mathbb{R}^\ell_*} \left[u(V_r + \Psi_r(x)) - u(V_r) - u'(V_r)\Psi_r(x) \right] \vartheta(\mathrm{d}x) \middle| \mathfrak{F}_t \right] \mathrm{d}r$$

$$= \xi_t u'(V_t) + \tfrac{1}{2}\sigma_t^2 u''(V_t) + \int_{\mathbb{R}^\ell_*} [u(V_t + \Psi_t(x))$$

$$- u(V_t) - u'(V_t)\Psi_t(x) \big] \vartheta(\mathrm{d}x) \text{ a.s.}$$

Hence, the family $\{\mathcal{L}(V_{t+s}|\mathfrak{F}_t)\}_{s\geq 0}$ is a.s. U-differentiable with derivative

$$\dot{\mathcal{L}}(V_t|\mathfrak{F}_t)[u] = \xi_t u'(V_t) + \tfrac{1}{2}\sigma_t^2 u''(V_t) + \int_{\mathbb{R}^\ell_*} \left[u(V_t + \Psi_t(x)) - u(V_t) - u'(V_t)\Psi_t(x) \right] \vartheta(\mathrm{d}x). \tag{11}$$

Remark 5.3 In the terminology of Markov semigroups, if $\mu_s = \mathcal{L}(X_s)$, $s \geq 0$, for some stochastic process X with a corresponding semigroup $\{K_t\}$ on the function space $C_b(\mathcal{V})$, then the U-derivative of Definition 5.2 resembles the notion of the (weak*) generator of the associated dual semigroup $\{K_t^*\}$ on $\mathcal{M}_1(\mathcal{V}) \subseteq C_b(\mathcal{V})^*$. Note, however, that Definition 5.2 is not restricted to Markovian settings.

Let us now introduce a corresponding notion of differentiability for certainty equivalents.

Definition 5.4 (*Differentiability of certainty equivalents*) A certainty equivalent $\mathfrak{m}$ on $\mathcal{V}$ is said to be U-*differentiable* if there exists a continuous function $M : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$, $(v, w) \mapsto M(v, w)$ with $M(v, \cdot\,) \in \mathrm{U}$ for each $v \in \mathcal{V}$ such that whenever $\{\mu_s\}_{s\geq 0}$ is U-differentiable at $s = 0$ with $\mu_0 = \delta_v$, it follows that the function $s \mapsto \mathfrak{m}(\mu_s)$ is differentiable at $s = 0$ and

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \mathfrak{m}(\mu_s) = \dot{\mu}_0[M(v, \cdot\,)]. \tag{12}$$

In this case, M is called a *local gradient representation* of $\mathfrak{m}$. A continuous-time aggregator $(f, \mathfrak{m})$ is U-*differentiable* if $\mathfrak{m}$ is U-differentiable.

Equation 12 is to be understood as a *chain rule*. Moreover, the rôle of U becomes clear at this point: U must be chosen large enough so that differentiability in (8) holds and small enough so that differentiability in (9) holds. In particular, by Proposition 5.8, one may take $\mathrm{U} = \{h\}$ for any EU certainty equivalent induced by a function h. Interestingly, Definition 5.4 is the natural concept of differentiability corresponding to the notion of continuity in Epstein and Zin [15]. Finally, note that local gradient representations need not be unique.

Example 5.5 (Example 5.1 continued) It is clear that the family $\{\delta_{v+s}\}_{s\geq 0}$ is U-differentiable, and since $\mathfrak{m}(\delta_{v+s}) = v + s$ for any $s \geq 0$, we find $\dot{\mu}_0[u] = u'(v)$ for any $u \in \mathrm{U}$. Thus (12) implies

$$\frac{\partial M}{\partial w}(v, v) = 1 \quad \text{for all } v \in \mathcal{V}. \tag{13}$$

On the other hand, consider the family $\{\mu_s\}_{s\geq 0}$ given by $\mu_s \triangleq (1-s)\delta_v + s\delta_{\bar{v}}$ with $v, \bar{v} \in \mathcal{V}$. This family is U-differentiable with $\dot{\mu}_0[u] = u(\bar{v}) - u(v)$, $u \in \mathrm{U}$. Let $\mathfrak{m}_\alpha$ be the certainty equivalent from Example 5.1. Then we have $\mathfrak{m}_\alpha(\mu_s) = v$ for sufficiently small $s \geq 0$. Now, if $\mathfrak{m}_\alpha$ were U-differentiable, we would obtain

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \mathfrak{m}(\mu_s) = M(v, \bar{v}) - M(v, v).$$

Since $v, \bar{v} \in \mathcal{V}$ are arbitrary, this would imply that $\frac{\partial M}{\partial w}(v, v) = 0$, contradicting (13). Thus, $\mathfrak{m}_\alpha$ is not U-differentiable.

As emphasized in Sect. 1, the fundamental difference between U-differentiability and Gâteaux differentiability is that the former is based on the linear structure of $\mathcal{V}$, whereas the latter is based on that of $\mathcal{M}_1(\mathcal{V})$. In this context, it is in order to clarify

Remark 5.6 (Relationship to Machina [22]) Machina [22] and subsequent work study preference structures over probability measures, i.e. on $\mathcal{M}_1(\mathcal{V})$, and investigate when preferences are 'locally linear in probabilities'. In contrast to our purposes here, the appropriate concept of differentiability to address this issue is indeed based on the linear structure of $\mathcal{M}_1(\mathcal{V})$. From what has been said above, it is apparent that this notion of differentiability is conceptually different from that of Definition 5.4.

Example 5.7 Similarly as in Example 5.5, we can take μ_s to be a distribution with mean v and variance s that is compactly supported in $\mathcal{V}$ and deduce that the U-derivative of $\{\mu_s\}_{s\geq 0}$ is given by $\dot{\mu}_0[u] = \frac{1}{2}u''(v)$ for $u \in \mathrm{U}$. We conclude that if $\mathfrak{m}$ is *risk-averse* in the sense that $\mathfrak{m}(\mu) \leq \mathfrak{m}(\delta_{E(\mu)})$ for all $\mu \in \mathcal{M}_1(\mathcal{V})$ where $E(\mu) \triangleq \int_{\mathcal{V}} x\mu(\mathrm{d}x)$ is the mean of μ, then it follows that $\frac{\partial^2 M}{\partial w^2}(v, v) \leq 0$ for all $v \in \mathcal{V}$.

Proposition 5.8 (Differentiation of EU certainty equivalents) *Let $\mathfrak{m}_h$ denote the EU certainty equivalent on $\mathcal{V}$ induced by the function $h \in \mathrm{U}$. Then $\mathfrak{m}_h$ is* U*-differentiable, and its local gradient representation M_h is given by*

$$M_h(v, w) = \frac{h(w)}{h'(v)} \quad \textit{for } v, w \in \mathcal{V}. \tag{14}$$

In particular, the local gradient representation of a risk-neutral certainty equivalent is the identity mapping $\mathrm{Id} : \mathcal{V} \times \mathcal{V} \to \mathcal{V}, (v, w) \mapsto w$.

Proof Let $\{\mu_s\}_{s\geq 0}$ be a U-differentiable family with $\mu_0 = \delta_v$. Then we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \mathfrak{m}(\mu_s) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} h^{-1}\left(\int_{\mathcal{V}} h \mathrm{d}\mu_s\right) = \frac{1}{h'(v)} \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \int_{\mathcal{V}} h\mathrm{d}\mu_s = \dot{\mu}_0\left[\frac{h}{h'(v)}\right],$$

and therefore the local gradient representation is given by (14). □

Remark 5.9 In the special case of Proposition 5.8, the local gradient representation in the sense of Duffie and Epstein [10], which refers to Gâteaux derivatives, happens to be given by the same formula since we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \mathfrak{m}(\delta_v + s\rho) = \frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} h^{-1}\left(h(v) + s\int_{\mathcal{V}} h\mathrm{d}\rho\right) = \int_{\mathcal{V}} \frac{h(w)}{h'(v)}\,\rho(\mathrm{d}w),$$

whenever ρ is a signed measure on $\mathfrak{B}(\mathcal{V})$ such that $\rho(\mathcal{V}) = 0$.

Finally, we calculate the local gradient representation of the *Chew–Dekel certainty equivalent*[14] $\mathfrak{m}_H$ induced by the C^2 function $H : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$. Here it is assumed that $\frac{\partial H}{\partial m}$ is strictly positive and $H(w, w) = 0$ for $w \in \mathcal{V}$. Then $\mathfrak{m}_H : \mathcal{M}_1(\mathcal{V}) \to \mathbb{R}$ is determined by the condition that

$$m = \mathfrak{m}_H(\mu) \text{ is the unique solution to } \int_{\mathcal{V}} H(m, w)\mu(\mathrm{d}w) = 0 \quad \text{for every } \mu \in \mathcal{M}_1(\mathcal{V}).$$

Note that for $H(m, w) = h(m) - h(w)$ the Chew–Dekel certainty equivalent collapses to an EU certainty equivalent.

Proposition 5.10 (Differentiation of Chew–Dekel certainty equivalents) *Let $\mathfrak{m}_H$ denote the Chew–Dekel certainty equivalent induced by the function H. Assume that $H(m, \cdot)$, $\frac{\partial H}{\partial m}(m, \cdot) \in \mathrm{U}$ for $m \in \mathcal{V}$, and suppose that $\frac{\partial H}{\partial m}$ and $\frac{\partial^2 H}{\partial m^2}$ are uniformly bounded. Then $\mathfrak{m}_H$ is* U*-differentiable with local gradient representation*

$$M_H : \mathcal{V} \times \mathcal{V} \to \mathbb{R}, \quad M_H(v, w) = -\frac{H(v, w)}{\frac{\partial H}{\partial m}(v, v)}. \tag{15}$$

Proof See Appendix A. □

If we choose $H(m, w) = 1_{(-\infty,m]}(w) - \alpha$ for some $\alpha \in (0, 1)$, the corresponding Chew–Dekel certainty equivalent coincides with $\mathfrak{m}_\alpha$ from Example 5.1. Since $\mathfrak{m}_\alpha$ is not U-differentiable, this shows that the differentiability assumptions on H in Proposition 5.10 cannot be dropped in general.

In special cases of Chew–Dekel certainty equivalents, one can establish U-differentiability under weaker assumptions. For weighted utility this is shown in the following proposition.[15]

Proposition 5.11 (Differentiation of weighted utility certainty equivalents) *Let $h : \mathcal{V} \to \mathbb{R}$ be a strictly increasing function of class* C^2 *and let $g : \mathcal{V} \to (0, \infty)$. Assume that g and h are polynomially bounded and $g, gh \in \mathrm{U}$. Then the weighted utility certainty equivalent*

$$\mathfrak{m} : \mathcal{M}_1(\mathcal{V}) \to \mathbb{R}, \quad \mathfrak{m}(\mu) \triangleq h^{-1}\left(\frac{\int_{\mathcal{V}} gh\,\mathrm{d}\mu}{\int_{\mathcal{V}} g\,\mathrm{d}\mu}\right)$$

is U*-differentiable and its local gradient representation is given by*

$$M : \mathcal{V} \times \mathcal{V} \to \mathbb{R}, \quad M(v, w) = \frac{g(w)[h(w) - h(v)]}{g(v)h'(v)}. \tag{16}$$

Proof If $\{\mu_s\}_{s\geq 0}$ is a U-differentiable family with $\mu_0 = \delta_v$, we have

$$\begin{aligned} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \mathfrak{m}(\mu_s) &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} h^{-1}\left(\frac{\int_{\mathcal{V}} gh\,\mathrm{d}\mu_s}{\int_{\mathcal{V}} g\,\mathrm{d}\mu_s}\right) \\ &= \frac{1}{h'(v)}\left[\frac{\dot{\mu}_0[gh]}{g(v)} - \frac{g(v)h(v)\dot{\mu}_0[g]}{g(v)^2}\right] = \dot{\mu}_0\left[\frac{g[h - h(v)]}{g(v)h'(v)}\right]. \end{aligned}$$

Hence, $\mathfrak{m}$ is U-differentiable with local gradient representation given by the stated formula. □

14 See Dekel [9] and Chew [8].

15 See Chew [7,8]. Note that for $g = 1$ the weighted utility certainty equivalent collapses to an EU certainty equivalent.

Remark 5.12 The local gradient representation of a weighted utility certainty equivalent can also be calculated via Proposition 5.10 if its assumptions are satisfied. Of course in this case (15) reduces to (16).

To summarize, let us return to the problem formulated at the beginning of this section concerning the derivative in Eq. 6 of Definition 4.1. We suppose that the continuation value process V^c is such that $\{\mathfrak{L}(V^c_{t+s}|\mathfrak{F}_t)\}_{s\geq 0}$ is a.s. U-differentiable at $s=0$ for a.e. $t\in[0,T]$ and assume that the certainty equivalent $\mathfrak{m}$ is U-differentiable in the sense of Definition 5.4. Then condition (5) is satisfied and

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0}\mathfrak{m}\left(\mathfrak{L}\left(V^c_{t+s}|\mathfrak{F}_t\right)\right)=\dot{\mathfrak{L}}\left(V^c_t|\mathfrak{F}_t\right)\left[M\left(V^c_t,\ \cdot\right)\right]\ \text{ exists a.s. for a.e. } t\in[0,T]$$

where $\dot{\mathfrak{L}}(V^c_t|\mathfrak{F}_t)$ denotes the U-derivative of $\{\mathfrak{L}(V^c_{t+s}|\mathfrak{F}_t)\}_{s\geq 0}$. Substituting into (6), we obtain

Theorem 5.13 (U-differentiability and CRU) *Let $(f,\mathfrak{m})$ be a* U*-differentiable continuous-time aggregator on $\mathcal{V}$ and let M denote the local gradient representation of $\mathfrak{m}$. Moreover, suppose that for each consumption process $c=\{c_t\}_{t\in[0,T]}$ there exists an a.s. unique $\mathcal{V}$-valued semimartingale $V^c=\{V^c_t\}_{t\in[0,T]}$ with $\mathbb{E}[\sup_{t\in[0,T]}|V^c_t|^p]<\infty$ for all $p\in[1,\infty)$ such that $\{\mathfrak{L}(V^c_{t+s}|\mathfrak{F}_t)\}_{s\geq 0}$ is a.s.* U*-differentiable for each $t\in[0,T]$, $V^c_T=0$, and*

$$\dot{\mathfrak{L}}\left(V^c_t|\mathfrak{F}_t\right)[M(V_t,\ \cdot)]=-f(c_t,V^c_t)\ \textit{a.s. for a.e. } t\in[0,T].$$

Then $(f,\mathfrak{m})$ generates a CRU function.

6 Lévy settings and stochastic differential utility

Throughout this section, we assume that we are in a Lévy setting and that u',u'' are bounded for each $u\in$ U. For a U-differentiable aggregator $(f,\mathfrak{m})$ and a regular process $V^c=\{V^c_t\}_{t\in[0,T]}$ with $\mathrm{d}V^c_t=\xi_t\mathrm{d}t+\sigma_t\mathrm{d}W_t+\int_{\mathbb{R}^\ell_*}\Psi_t(x)\tilde{\nu}(\mathrm{d}t,\mathrm{d}x)$ as in (1), the discussion preceding Theorem 5.13 leads to the condition

$$-\xi_t=\tfrac{1}{2}\sigma_t^2A(V^c_t)+\int\limits_{\mathbb{R}^\ell_*}J(V^c_t,\Psi_t(x))\vartheta(\mathrm{d}x)+f(c_t,V^c_t)\ \text{ a.s. for a.e. } t\in[0,T]$$

where the *variance multiplier* A and the *jump term* J associated to $\mathfrak{m}$ are defined by

$$A(v)\triangleq\frac{\partial^2M}{\partial w^2}(v,v)\quad\text{and}\quad J(v,\psi)\triangleq M(v,v+\psi)-M(v,v)-\psi\quad\text{for } v,v+\psi\in\mathcal{V}. \tag{17}$$

Intuitively, A represents the investor's aversion towards diffusion risk, whereas J essentially captures aversion towards jump risk, see Skiadas [28]. Hence, we obtain for a.e. $t\in[0,T]$

$$V^c_t=\mathbb{E}\left[\int\limits_t^T\left\{f\left(c_s,V^c_s\right)+\tfrac{1}{2}\sigma_s^2A\left(V^c_s\right)+\int\limits_{\mathbb{R}^\ell_*}J\left(V^c_s,\Psi_s(x)\right)\vartheta(\mathrm{d}x)\right\}\mathrm{d}s\,\middle|\,\mathfrak{F}_t\right]\ \text{a.s.} \tag{18}$$

Following Duffie and Epstein [10], one may now take (18) as a formal definition of recursive utility in continuous time.

Definition 6.1 (*Stochastic differential utility, SDU function*) In a Lévy framework, suppose $f : \mathcal{C} \times \mathcal{V} \to \mathbb{R}$ is Borel measurable and $M : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is such that $M(v, \cdot)$ is of class C^2 for every $v \in \mathcal{V}$. Let A and J be given by (17). If for each $c \in \mathsf{C}$ there is an a.s. unique $\mathcal{V}$-valued semimartingale V^c such that $\mathrm{d}V_t^c = \xi_t \mathrm{d}t + \sigma_t \mathrm{d}W_t + \int_{\mathbb{R}_*^\ell} \Psi_t(x)\tilde{\nu}(\mathrm{d}t, \mathrm{d}x)$ and the backward stochastic differential equation 18 is satisfied, then the function

$$\mathfrak{u} : \mathsf{C} \to \mathbb{R}, \quad \mathfrak{u}(c) \triangleq V_0^c$$

is said to be the *stochastic differential utility function*, or *SDU function*, associated to (f, M).

Note that Definition 6.1 also captures generalized stochastic differential utility in the sense of El Karoui et al. [13] and Lazrak and Quenez [20]. The following result shows that the notions of CRU and SDU are essentially equivalent in Lévy settings.

Theorem 6.2 (CRU versus SDU) *In a Lévy framework, suppose that $(f, \mathfrak{m})$ is a* U*-differentiable continuous-time aggregator with local gradient representation M. If $(f, \mathfrak{m})$ generates a CRU function and each continuation value process is regular, then (f, M) generates an SDU function. Conversely, if (f, M) generates an SDU function, then $(f, \mathfrak{m})$ generates a CRU function. In both cases, the corresponding CRU and SDU functions coincide.*

Proof It is clear from the derivation of (18) that if $(f, \mathfrak{m})$ generates a CRU function $\mathfrak{u}$ and each continuation value process is regular, then (f, M) generates the SDU function $\mathfrak{u}$. The converse follows from Theorem 5.13 if we recall from Sect. 5 that the conditional distributions of any regular process are a.s. U-differentiable. □

The remainder of this section is concerned with existence and uniqueness results for CRU and SDU in Lévy settings with $\mathcal{V} = \mathbb{R}$. For SDU, the relevant Eq. 18 is in general hard to deal with. Nevertheless, for Poisson random measures with finite intensities, Ma [21] establishes the existence of SDU indices using the fact (see Corollary 7.3 below) that the variance multiplier can be transformed away. In the special case when both the variance multiplier A and the jump term J are zero, Eq. 18 simplifies to

$$V_t^c = \mathbb{E}\left[\int_t^T f(c_s, V_s^c)\mathrm{d}s \,\middle|\, \mathfrak{F}_t \right] \quad \text{a.s.} \quad \text{for a.e. } t \in [0, T]. \tag{19}$$

In the terminology of Sect. 7, this means that $(f, \mathfrak{m})$ is *normalized*. Existence (in the class of regular processes) and uniqueness (modulo indistinguishability in the class of L^p-bounded càdlàg processes, for arbitrary $p \in [1, \infty)$) of solutions to (19) follow[16] under the assumption that f satisfies Lipschitz and linear growth conditions, i.e.

$$|f(c, v) - f(c, w)| \leq \alpha|v - w| \quad \text{for all } c \in \mathcal{C}, \; v, w \in \mathbb{R} \tag{20}$$

for some $\alpha > 0$ and $|f(c, 0)| \leq \beta_0 + \beta_1|c|$, $c \in \mathcal{C}$, for some $\beta_1, \beta_2 > 0$. However, it is not at all clear at this point under which conditions the normalization (19) is feasible; this will be addressed in the next section. The SDU existence result just discussed yields the following existence result for CRU.

[16] See Lemma 2.4 of Tang and Li [30] and Theorem 2.1 of Barles et al. [1]. Note that this result is stated in Tang and Li [30] and Barles et al. [1] for $p = 2$, but it can be extended to the case $p > 2$, see the arguments in the proof of Proposition 3.1 in Buckdahn and Pardoux [6].

Corollary 6.3 (Existence of CRU) *Let $\mathfrak{m}_n$ denote a risk-neutral certainty equivalent. In a Lévy setting with $\mathcal{V} = \mathbb{R}$, consider an aggregator $(f, \mathfrak{m}_n)$ where f satisfies Lipschitz and linear growth conditions. Then $(f, \mathfrak{m}_n)$ generates a CRU function $\mathfrak{u}$, and $\mathfrak{u}$ coincides with the SDU function generated by (f, Id) where $\mathrm{Id}(v, w) = w$ for $v, w \in \mathbb{R}$.*

Proof The result immediately follows from Theorem 6.2 since risk-neutral certainty equivalents are U-differentiable with local gradient representation Id. □

Remark 6.4 The boundedness condition on u' and u'' for $u \in \mathrm{U}$ and the Lipschitz condition on f are not satisfied for some relevant classes of certainty equivalents. This is however a generic technical problem that already occurs for stochastic differential utility in Brownian settings. We refer the reader to Duffie and Lions [11] for an approach via partial differential equations and to Schroder and Skiadas [23] for an approach via backward stochastic differential equations.

We also remark that with the help of Theorem 6.2 and the Gronwall–Bellman results established in Appendix B, the desirable properties of stochastic differential utility functions in Sect. 5 of Duffie and Epstein [10] can be extended to Lévy frameworks.

7 Ordinal equivalence and normalization

Returning to the discussion of ordinal equivalence at the end of Sect. 3, we say that two aggregators $(f, \mathfrak{m})$ and $(\bar{f}, \bar{\mathfrak{m}})$ generating recursive utility functions $\mathfrak{u}$ and $\bar{\mathfrak{u}}$ are *ordinally equivalent* if $\mathfrak{u}$ and $\bar{\mathfrak{u}}$ are equivalent. The degree of freedom implicit in this notion can be used to perform a change of scale analogously to the discrete-time normalization of Proposition 3.3. A certainty equivalent $\mathfrak{m}$ is said to be *normalized* if it is U-differentiable and the associated variance multiplier and jump term as defined in (17) vanish. For instance, this is the case for risk-neutral certainty equivalents. An aggregator $(f, \mathfrak{m})$ is said to be normalized if $\mathfrak{m}$ is normalized. It turns out that the concept of CRU is especially suitable to study normalization as Theorem 7.1 shows. We wish to point that this theorem is not restricted to Lévy settings.

Theorem 7.1 (Transformation of aggregators) *Let $(f, \mathfrak{m})$ and $(\bar{f}, \bar{\mathfrak{m}})$ be aggregators on $\mathcal{V}$ and $\bar{\mathcal{V}}$ and suppose that $(f, \mathfrak{m})$ generates a CRU function. Furthermore, let $\Phi : \mathcal{V} \to \bar{\mathcal{V}}$ be a C^2 function with $\Phi'(v) > 0$ for $v \in \mathcal{V}$ and $\Phi(0) = 0$ such that Φ and Φ^{-1} are polynomially bounded.*[17] *Given that*[18]

$$\mathfrak{m}(\mu) = \Phi^{-1}\left(\bar{\mathfrak{m}}(\mu^{\Phi})\right) \quad \text{for all } \mu \in \mathcal{M}_1(\mathcal{V}) \tag{21}$$

and

$$f(c, v) = \frac{\bar{f}(c, \Phi(v))}{\Phi'(v)} \quad \text{for } c \in \mathcal{C},\ v \in \mathcal{V}, \tag{22}$$

then $(\bar{f}, \bar{\mathfrak{m}})$ generates a CRU function, and $(f, \mathfrak{m})$ and $(\bar{f}, \bar{\mathfrak{m}})$ are ordinally equivalent. Moreover, if $\bar{\mathfrak{m}}$ is $\bar{\mathrm{U}}$-differentiable and $\bar{u} \circ \Phi \in \mathrm{U}$ whenever $\bar{u} \in \bar{\mathrm{U}}$, then $\mathfrak{m}$ is U-differentiable with local gradient representation

[17] Here and in the following, polynomial boundedness assumptions are required because of the generality of our framework: We are working within the class of L^p-bounded semimartingales, see Definition 4.1. For a given aggregator, one might be able to relax these conditions by using a setting exactly tailored to this aggregator.

[18] $\mu^{\Phi} = \mu \circ \Phi^{-1}$ continues to denote the image of $\mu \in \mathcal{M}_1(\mathcal{V})$ under $\Phi : \mathcal{V} \to \bar{\mathcal{V}}$.

$$M(v, w) = \frac{\bar{M}(\Phi(v), \Phi(w))}{\Phi'(v)} \quad \textit{for } v, w \in \mathcal{V}. \tag{23}$$

Proof Let $c \in \mathsf{C}$ and denote by V^c the associated continuation value process. Then

$$\begin{aligned} -f\left(c_t, V_t^c\right) &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \mathfrak{m}\left(\mathfrak{L}\left(V_{t+s}^c|\mathfrak{F}_t\right)\right) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \Phi^{-1}\left(\bar{\mathfrak{m}}\left(\mathfrak{L}\left(V_{t+s}^c|\mathfrak{F}_t\right)^{\Phi}\right)\right) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \Phi^{-1}\left(\bar{\mathfrak{m}}\left(\mathfrak{L}\left(\bar{V}_{t+s}^c|\mathfrak{F}_t\right)\right)\right) = \frac{1}{\Phi'(V_t)} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \bar{\mathfrak{m}}\left(\mathfrak{L}\left(\bar{V}_{t+s}^c|\mathfrak{F}_t\right)\right) \text{ a.s.,} \end{aligned}$$

where $\bar{V}^c = \{\bar{V}_t^c\}_{t\in[0,T]}$ is given by $\bar{V}_t^c \triangleq \Phi(V_t^c)$, $t \in [0, T]$. Note that $\mathbb{E}[\sup_{t\in[0,T]} |\bar{V}_t^c|^p] < \infty$ for $p \in [1, \infty)$. Thus if conditions (22) and (21) hold, then $\bar{V}^c$ is a continuation value process of c under $(\bar{f}, \bar{\mathfrak{m}})$. Applying the preceding argument to $\bar{V}^c$ and the mapping Φ^{-1}, we see that $\bar{V}^c$ is uniquely determined. Consequently, $(\bar{f}, \bar{\mathfrak{m}})$ generates a CRU function $\bar{\mathfrak{u}}$ with

$$\bar{\mathfrak{u}}(c) = \bar{V}_0^c = \Phi(V_0^c) = \Phi(\mathfrak{u}(c)) \quad \text{for any } c \in \mathsf{C},$$

and thus $\bar{\mathfrak{u}}$ and $\mathfrak{u}$ are equivalent. Next let $\{\mu_s\}_{s\geq 0}$ be a U-differentiable family on $\mathcal{V}$ with $\mu_0 = \delta_v$. The identity

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \int_{\bar{\mathcal{V}}} \bar{u}\, \mathrm{d}\mu_s^{\Phi} = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \int_{\mathcal{V}} \bar{u} \circ \Phi\, \mathrm{d}\mu_s = \dot{\mu}_0[\bar{u} \circ \Phi]$$

shows that $\{\mu_s^{\Phi}\}_{s\geq 0}$ is $\bar{\mathrm{U}}$-differentiable with $\dot{\mu}_0^{\Phi}[\bar{u}] = \dot{\mu}_0[\bar{u} \circ \Phi]$ for $\bar{u} \in \bar{\mathrm{U}}$. Now, if $\bar{\mathfrak{m}}$ is $\bar{\mathrm{U}}$-differentiable with local gradient representation $\bar{M}$, then it follows that

$$\begin{aligned} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \mathfrak{m}(\mu_s) &= \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \Phi^{-1}\left(\bar{\mathfrak{m}}\left(\mu_s^{\Phi}\right)\right) \\ &= \frac{1}{\Phi'(v)} \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \bar{\mathfrak{m}}\left(\mu_s^{\Phi}\right) = \frac{1}{\Phi'(v)} \dot{\mu}_0^{\Phi}\left[\bar{M}(\Phi(v), \Phi(\,\cdot\,))\right]. \end{aligned}$$

Hence, $\mathfrak{m}$ is U-differentiable, and its local gradient representation M is given by (23). □

If $\mathfrak{m}$ is of the EU type, then setting $\Phi \triangleq h - h(0)$ leads to

Corollary 7.2 (Normalization of EU certainty equivalents) *Suppose $(f, \mathfrak{m}_h)$ is an aggregator with an EU certainty equivalent induced by a function h where h^{-1} is polynomially bounded. Furthermore, assume that $(f, \mathfrak{m}_h)$ generates a CRU function. Then there exists an ordinally equivalent normalized aggregator $(\bar{f}, \bar{\mathfrak{m}})$ where $\bar{\mathfrak{m}}$ can even be chosen to be risk-neutral.*

In particular, it follows that, in a general semimartingale framework, existence and uniqueness results for the normalized backward stochastic differential equation (19) can be applied to establish existence of CRU functions associated to aggregators with EU certainty equivalents. U-differentiability has to be checked separately.

In the remainder of this section, we restrict ourselves to a Lévy setting. In the situation of Theorem 7.1, the respective variance multipliers A and $\bar{A}$ satisfy

$$A(v) = \bar{A}(v)\Phi'(v) + \frac{\Phi''(v)}{\Phi'(v)} \quad \text{for } v \in \mathcal{V},$$

provided Φ is of class C^2. This can be interpreted as an ordinary differential equation of second order for Φ that allows us to transform away $\bar{A}$.

Corollary 7.3 (Transformation of variance multipliers) *In a Lévy framework, suppose that* $(f, \mathfrak{m})$ *is a* U*-differentiable aggregator that generates a CRU function. Let* A *denote the variance multiplier of* $\mathfrak{m}$*. Moreover suppose that there is a solution to* $\Phi'' = A\Phi'$, $\Phi(0) = 0$, *such that* Φ *and* Φ^{-1} *are polynomially bounded. Then there exists an ordinally equivalent aggregator with vanishing variance multiplier.*

If the jump term is to disappear as well, then EU form is essentially necessary in the following sense.

Theorem 7.4 (Necessity of EU form) *In a Lévy setting, suppose that the aggregator* $(f, \mathfrak{m})$ *on* $\mathcal{V}$ *is ordinally equivalent to an aggregator* $(\bar{f}, \bar{\mathfrak{m}})$ *on* $\bar{\mathcal{V}}$ *with normalized jumps, i.e.* $\bar{J}(\bar{v}, \bar{\psi}) = 0$ *for* $\bar{v}, \bar{v} + \bar{\psi} \in \bar{\mathcal{V}}$*. Then the local gradient representation* M *of* $\mathfrak{m}$ *can be taken to be of EU form (14), i.e.*

$$M(v, w) = \frac{\Phi(w)}{\Phi'(v)} \quad \textit{for } v, w \in \mathcal{V}.$$

Proof By the definition of $\bar{J}$, the function $\bar{M}(\bar{v}, \cdot)$ must be affine-linear for each $\bar{v} \in \bar{\mathcal{V}}$. Hence, by (13) the local gradient representation $\bar{M}$ must satisfy $\bar{M}(\bar{v}, \bar{w}) = \bar{\alpha}(\bar{v}) + \bar{w}$ for all $\bar{v}, \bar{w} \in \bar{\mathcal{V}}$ with some function $\bar{\alpha} : \bar{\mathcal{V}} \to \mathbb{R}$. Then (23) yields

$$M(v, w) = \alpha(v) + \frac{\Phi(w)}{\Phi'(v)} \quad \text{for } v, w \in \mathcal{V}$$

where $\alpha(v) \triangleq \frac{\bar{\alpha}(\Phi(v))}{\Phi'(v)}$. However, it is obvious from Eq. 11 that the latter term cancels out, and the claim thus follows. □

In a discontinuous Lévy setting ($\nu \neq 0$), we arrive at the following important conclusion: If $\mathfrak{u}$ is a recursive utility function that arises from a continuous-time aggregator $(f, \mathfrak{m})$ after normalization, then the local gradient representation of $\mathfrak{m}$ must be of EU form. Thus from a behavioral point of view, the restriction to aggregators which can be normalized is equivalent to a restriction to EU certainty equivalents. Note that this is in line with the normalization result of Duffie and Epstein [10] because in their Brownian setting 'normalization' refers to the variance multiplier only.

8 Dynamic programming with recursive utility

Generalizing Duffie and Epstein [10], we study a stochastic control problem whose criterion is defined by a recursive utility function. Assume that we are in a Lévy framework and that the recursive utility function $\mathfrak{u}$ coincides with the normalized stochastic differential utility functional induced by $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ where f satisfies the Lipschitz condition (20). The state process X has the dynamics

$$dX_t = b(t, X_t, c_t)dt + a(t, X_t, c_t)dW_t + \int_{\mathbb{R}^\ell_*} e(t, X_{t-}, c_t, y)\tilde{\nu}(dt, dy), \quad X_0 = x, \quad (24)$$

with b, a and e being suitable coefficients and $x \in \mathbb{R}^d$. The process $c = \{c_t\}_{t\in[0,T]}$ is chosen from the class of *admissible controls*

$$\mathsf{C}(x) \triangleq \left\{c \in \mathsf{C} : (24) \text{ has a unique solution } X^{x,c} \text{ and } c_t \in \Gamma\left(t, X^{x,c}_t\right) \text{ for } t \in [0, T]\right\}$$

where the function[19] $\Gamma : [0,T]\times\mathbb{R}^d \to 2^{\mathcal{C}}$ models a possibly state-dependent constraint. Given an initial value $x\in\mathbb{R}^d$, the *optimization problem* is to maximize utility over the class $\mathsf{C}(x)$ of admissible strategies, i.e. to

$$\text{find } c^\star \in \mathsf{C}(x) \text{ such that } \mathfrak{u}(c^\star) = \max_{c\in\mathsf{C}(x)} \mathfrak{u}(c). \tag{25}$$

Problem (25) is invariant with respect to ordinally equivalent transformations. Therefore, the results of Sect. 7 provide sufficient conditions such that the assumption that $\mathfrak{u}$ is a normalized stochastic differential utility functional is satisfied. We now formulate a dynamic programming equation for problem (25). Therefore, define the controlled *generator* L^c for $c\in\mathcal{C}$ as

$$\begin{aligned} L^c[u](t,x) &\triangleq \tfrac{\partial u}{\partial t}(t,x) + b(t,x,c)\tfrac{\partial u}{\partial x}(t,x) + \tfrac{1}{2}a(t,x,c)^2\tfrac{\partial^2 u}{\partial x^2}(t,x) \\ &\quad + \int_{\mathbb{R}^\ell_*} \left[u(t,x+e(t,x,c,y)) - u(t,x) - \tfrac{\partial u}{\partial x}(t,x)e(t,x,c,y)\right]\vartheta(\mathrm{d}y) \end{aligned}$$

for $(t,x)\in[0,T]\times\mathbb{R}^d$ and $u\in\mathrm{C}^{1,2}([0,T]\times\mathbb{R}^d)$. Moreover, a Borel function $\gamma : [0,T]\times\mathbb{R}^d \to \mathcal{C}$ is said to be an *admissible feedback control* if the equation

$$\begin{aligned} \mathrm{d}X_t &= b(t,X_t,\gamma(t,X_t))\mathrm{d}t + a(t,X_t,\gamma(t,X_t))\mathrm{d}W_t \\ &\quad + \int_{\mathbb{R}^\ell_*} e(t,X_{t-},\gamma(t,X_{t-}),y)\tilde{\nu}(\mathrm{d}t,\mathrm{d}y), \quad X_0 = x, \end{aligned}$$

has a unique solution $X^{x,\gamma}$ such that $c=\{c_t\}_{t\in[0,T]} \triangleq \{\gamma(t,X^{x,\gamma}_t)\}_{t\in[0,T]} \in \mathsf{C}(x)$. We now establish a verification result for problem (25).

Theorem 8.1 (Verification theorem) *Let $w\in\mathrm{C}^{1,2}([0,T]\times\mathbb{R}^d)$ be a solution of the dynamic programming equation*

$$\sup_{c\in\Gamma(t,x)} L^c[w](t,x) + f(c,w(t,x)) = 0, \quad w(T,\,\cdot\,) = 0, \tag{26}$$

and assume that the local martingales

$$\int_0^{\cdot} \tfrac{\partial w}{\partial x}\left(s,X^{x,c}_s\right) a\left(s,X^{x,c}_s,c_s\right)\mathrm{d}W_s \;\; \textit{and} \;\; \int_{[0,\,\cdot\,]\times\mathbb{R}^\ell_*} J[w]\left(s,X^{x,c}_{s-},e\left(s,X^{x,c}_{s-},c_s,y\right)\right)\tilde{\nu}(\mathrm{d}s,\mathrm{d}y),$$

where $J[w](s,x,y)\triangleq w(s,x+y)-w(s,x)$ for $s\in[0,T]$, $x,y\in\mathbb{R}$, are martingales for every $c\in\mathsf{C}(x)$. Further, suppose there is a measurable function $\gamma^\star : [0,T]\times\mathbb{R}^d\to\mathcal{C}$ with

$$L^{\gamma^\star(t,x)}[w](t,x) + f(\gamma^\star(t,x),w(t,x)) = 0 \;\; \textit{for all } (t,x)\in[0,T]\times\mathbb{R}^d.$$

Then the feedback control $\gamma^\star$ is optimal, and w is the value function of problem (25). In particular,

$$w(0,x) = \max_{c\in\mathsf{C}(x)} \mathfrak{u}(c) \;\; \textit{for all } x\in\mathbb{R}^d.$$

Proof See Appendix A. □

[19] $2^{\mathcal{C}}$ denotes the power set of $\mathcal{C}$.

Finally, the following example shows that Hamilton–Jacobi–Bellman equations can involve non-standard terms if normalization is not feasible.

Example 8.2 (Chew–Dekel preferences) Consider an aggregator $(f, \mathfrak{m}_H)$ where $\mathfrak{m}_H$ is the Chew–Dekel certainty equivalent induced by H. If the assumptions of Proposition 5.10 are satisfied, $\mathfrak{m}_H$ is U-differentiable with local gradient representation (15). Even if we assume that the variance multiplier is 0, an additional term appears in the Hamilton–Jacobi–Bellman equation (26) that cannot be transformed away:

$$\sup_{c\in\Gamma(t,x)} L^c[w](t,x) + f(c, w(t,x)) - \int_{\mathbb{R}^\ell_*} \left\{ w(t, x + e(t,x,c,y)) - w(t,x) \right.$$
$$\left. + \frac{H(w(t,x), w(t, x + e(t,x,c,y)))}{\frac{\partial H}{\partial m}(w(t,x), w(t,x))} \right\} \vartheta(\mathrm{d}y) = 0.$$

To understand the meaning of the non-standard terms, denote the generator without the jump terms by L^c_{cont}. We then obtain

$$\sup_{c\in\Gamma(t,x)} L^c_{\text{cont}}[w](t,x) + f(c, w(t,x)) - \int_{\mathbb{R}^\ell_*} \left\{ \frac{H(w(t,x), w(t, x + e(t,x,c,y)))}{\frac{\partial H}{\partial m}(w(t,x), w(t,x))} \right.$$
$$\left. + \frac{\partial w}{\partial x}(t,x) e(t,x,c,y) \right\} \vartheta(\mathrm{d}y) = 0.$$

since the remaining terms cancel out with the corresponding term in the generator. Setting $w \triangleq w(t,x)$, $e \triangleq e(t,x,c,y)$ and $\hat{w} \triangleq w(t, x+e)$ the integrand $\frac{H(w,\hat{w})}{\frac{\partial H}{\partial m}(w,w)} + \frac{\partial w}{\partial x} e$ thus captures the non-linearity of the certainty equivalent. Note that in the special case where H is normalized, i.e. $H(m, w) = m - w$, the non-standard term collapses into the usual jump component of the generator:

$$-\left\{ \frac{H(w,\hat{w})}{\frac{\partial H}{\partial m}(w,w)} + \frac{\partial w}{\partial x} e \right\} = \hat{w} - w - \frac{\partial w}{\partial x} e.$$

To summarize, the non-standard terms in the Bellman equation are a consequence of the non-linearity of the certainty equivalent and the presence of jumps.

9 Conclusion

This paper provides an alternative derivation of recursive utility in continuous time. In contrast to stochastic differential utility, we directly link continuous-time recursive utility to its discrete-time counterpart by applying a novel notion for differentiating certainty equivalents. We have shown that this approach is useful to study the question of when aggregators can be normalized. It turns out that in the presence of jumps normalization is essentially feasible if and only if certainty equivalents are of the expected utility type (i.e. linear homogenous). Consequently, related Hamilton–Jacobi–Bellman equations have the well-known form. For instance, this is so for Epstein–Zin preferences. However, if certainty equivalents are not linear homogenous, which is for instance the case for Chew–Dekel preferences, then aggregators cannot be normalized in general. This implies that Bellman equations involve non-standard terms. This is a crucial result and has to be taken into account in future research on portfolio optimization and asset pricing using non-standard preferences such as preferences modeled by weighted utility theory. Furthermore, normalization is also relevant if the existence of stochastic differential utility is studied. For this reason, our paper also contributes to this strand of research and shows that in the presence of jumps it is not sufficient to study stochastic differential utility for normalized aggregators only.

Acknowledgements Holger Kraft gratefully acknowledges financial support by the Deutsche Forschungsgemeinschaft. Frank Seifried gratefully acknowledges financial support by the Rheinland-Pfalz Cluster of Excellence 'Dependable Adaptive Systems and Mathematical Modeling'. We thank Harjoat Bhamra, Darrell Duffie, Elyès Jouini, Costis Skiadas, and anonymous referees for very helpful comments. All remaining errors are of course our own.

Appendix A: Proofs of Proposition 5.10 and Theorem 8.1

Proof of Proposition 5.10 Let $\{\mu_s\}_{s\geq 0}$ be U-differentiable with $\mu_0 = \delta_v$, and set $g(s) \triangleq \mathfrak{m}(\mu_s)$ for $s \geq 0$. Observe that, by construction,

$$\frac{1}{s}\int_{\mathcal{V}} H(g(s), w)\mu_s(\mathrm{d}w) = 0 \quad \text{for all } s \geq 0. \tag{27}$$

Applying the mean value theorem to $H(\,\cdot\,, w)$, we obtain for $w \in \mathcal{V}$ and $s \geq 0$ that

$$\begin{aligned} H(g(s), w) &= H(g(0), w) + [g(s) - g(0)]\frac{\partial H}{\partial m}(\theta(s, w)g(0) + [1-\theta(s, w)]g(s), w) \\ &= H(g(0), w) + sg'(0)\frac{\partial H}{\partial m}(g(0), w) + \varphi(s, w) \end{aligned} \tag{28}$$

where $\theta(s, w) \in [0, 1]$ and $\varphi:\ [0,\infty)\times\mathcal{V}\to\mathbb{R}$ is given by

$$\begin{aligned} \varphi(s, w) \triangleq\ & [g(s) - g(0) - sg'(0)]\frac{\partial H}{\partial m}(\theta(s, w)g(0) + [1-\theta(s, w)]g(s), w) \\ & + sg'(0)\left[\frac{\partial H}{\partial m}(\theta(s, w)g(0) + [1-\theta(s, w)]g(s), w) - \frac{\partial H}{\partial m}(g(0), w)\right]. \end{aligned}$$

Note that clearly $\frac{1}{s}|[g(s) - g(0) - sg'(0)]\frac{\partial H}{\partial m}(\theta(s, w)g(0) + [1-\theta(s, w)]g(s), w)| \leq |\frac{g(s)-g(0)}{s} - g'(0)|K \to 0$ as $s \downarrow 0$ where $K \in (0,\infty)$ is a uniform bound for $\frac{\partial H}{\partial m}$ and $\frac{\partial^2 H}{\partial m^2}$. Furthermore, we have $|g'(0)[\frac{\partial H}{\partial m}(\theta(s, w)g(0) + [1-\theta(s, w)]g(s), w) - \frac{\partial H}{\partial m}(g(0), w)]| \leq |g'(0)|K|g(s) - g(0)| \to 0$ as $s \downarrow 0$. Hence, it follows that

$$\sup_{w\in\mathcal{V}} \frac{|\varphi(s, w)|}{s} \to 0 \quad \text{as } s \downarrow 0.$$

Substituting (28) into (27) yields

$$0 = \frac{1}{s}\int_{\mathcal{V}} H(g(0), w)\mu_s(\mathrm{d}w) + g'(0)\int_{\mathcal{V}} \frac{\partial H}{\partial m}(g(0), w)\mu_s(\mathrm{d}w) + \int_{\mathcal{V}} \frac{\varphi(s, w)}{s}\mu_s(\mathrm{d}w),\ \ s \geq 0.$$

Here the last summand tends to 0 as $s \downarrow 0$, and since $g(0) = v$ and $H(v,\ \cdot),\ \frac{\partial H}{\partial m}(v,\ \cdot) \in \mathrm{U}$

$$0 = \dot{\mu}_0[H(g(0),\ \cdot)] + g'(0)\int_{\mathcal{V}} \frac{\partial H}{\partial m}(g(0), w)\mu_0(\mathrm{d}w) = \dot{\mu}_0[H(v,\ \cdot)]$$
$$+ \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \mathfrak{m}(\mu_s)\frac{\partial H}{\partial m}(v, v).$$

After rearranging we obtain $\frac{\mathrm{d}}{\mathrm{d}s}|_{s=0}\mathfrak{m}(\mu_s) = -\frac{\dot{\mu}_0[H(v,\ \cdot)]}{\frac{\partial H}{\partial m}(v,v)} = \dot{\mu}_0[M_H(v,\ \cdot)]$ where M_H is defined as in the assertion. This completes the proof. □

Proof of Theorem 8.1 We adapt the line of argument given in the proof of Proposition 9 in Duffie and Epstein [10]. Let $x \in \mathbb{R}^d$ and $c \in \mathsf{C}(x)$ be an arbitrary admissible control. To shorten notation, $X = X^{x,c}$ and $V = V^c$ denote the controlled process and the continuation value process associated to c, respectively. Itô's formula implies that

$$w(t, X_t) = w(t, X_t) - w(T, X_T) = -\int_t^T L^{c_s}[w](s, X_s)\mathrm{d}s - M_T + M_t \text{ a.s.,}$$

where M is a martingale. Taking conditional expectation yields $w(t, X_t) = -\mathbb{E}[\int_t^T L^{c_s}[w](s, X_s)\mathrm{d}s|\mathfrak{F}_t]$ a.s. On the other hand, we have $V_t = \mathbb{E}[\int_t^T f(c_s, V_s)\mathrm{d}s|\mathfrak{F}_t]$ a.s. due to the definition of V. Hence,

$$w(t, X_t) - V_t = -\mathbb{E}\left[\int_t^T \{L^{c_s}[w](s, X_s) + f(c_s, V_s)\}\mathrm{d}s \,\middle|\, \mathfrak{F}_t\right] \text{ a.s.} \quad \text{for all } t \in [0, T]. \tag{29}$$

The dynamic programming equation and the Lipschitz property of f imply

$$\begin{aligned} L^{c_s}[w](s, X_s) + f(c_s, V_s) &= L^{c_s}[w](s, X_s) + f(c_s, w(s, X_s)) + f(c_s, V_s) \\ &\quad - f(c_s, w(s, X_s)) \leq f(c_s, V_s) - f(c_s, w(s, X_s)) \\ &\leq \alpha\,|w(s, X_s) - V_s| \quad \text{for } s \in [0, T]. \end{aligned} \tag{30}$$

Combining Eq. 29 with 30, it follows that $Y = \{Y_t\}_{t\in[0,T]} \triangleq \{w(t, X_t) - V_t\}_{t\in[0,T]}$ satisfies $Y_t = \mathbb{E}[\int_t^T H_s\mathrm{d}s|\mathfrak{F}_t]$ with $H_t \geq -\alpha|Y_t|$ for all $t \in [0, T]$. Thus we can apply the generalized version of Skiadas' Lemma B.4 and obtain $\mathfrak{u}(c) = V_0^c \leq w(0, X_0^{x,c}) = w(0, x)$. Since $c \in \mathsf{C}(x)$ is arbitrary, we have $\max_{c\in\mathsf{C}(x)} \mathfrak{u}(c) \leq w(0, x)$.

Conversely, under the assumptions of the theorem, the feedback control $\gamma^\star$ is admissible and (30) is satisfied as equality for $\gamma^\star$. Therefore, the above argument applies to both $\{w(t, X_t^{x,c^\star}) - V_t^{c^\star}\}_{t\in[0,T]}$ and $\{V_t^{c^\star} - w(t, X_t^{x,c^\star})\}_{t\in[0,T]}$. Consequently, $\mathfrak{u}(c^\star) = w(0, x)$ where $c^\star = \{c_t^\star\}_{t\in[0,T]} \triangleq \{\gamma^\star(t, X_t^{x,\gamma^\star})\}_{t\in[0,T]}$. Hence, $\gamma^\star$ is an optimal feedback control. □

Appendix B: Stochastic Gronwall–Bellman inequalities

We extend the results of Appendix B in Duffie and Epstein [10] to discontinuous processes; the appendix of Ma [21] contains a related result. Since in the absence of continuity the line of argument has to be refined, we give complete proofs. Throughout this section, we assume as given a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ endowed with an arbitrary filtration $\{\mathfrak{F}_t\}_{t\in[0,T]}$ satisfying the usual conditions. Recall the general version of

Lemma B.1 (Gronwall–Bellman inequality) *Let $h : [0, T] \to \mathbb{R}$ be a bounded measurable function such that for some $\alpha \in (0, \infty)$*

$$h(t) \geq \alpha \int_t^T h(s)\mathrm{d}s \quad \text{for a.e. } t \in [0, T].$$

Then it follows that $h(t) \geq 0$ for a.e. $t \in [0, T]$.

Below we need a Fubini-type theorem for conditional expectations that we state explicitly for ease of reference.

Proposition B.2 (Conditional Fubini theorem) *Let $Y = \{Y_t\}_{t\in[0,T]}$ be a measurable[20] process on the probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ with $\int_0^T \mathbb{E}[|Y_t|]\mathrm{d}t < \infty$ and let $\mathfrak{G} \subseteq \mathfrak{F}$ be a sub-σ-field. Then there exists a measurable process $H = \{H_t\}_{t\in[0,T]}$ with*

$$H_t = \mathbb{E}[Y_t|\mathfrak{G}] \text{ a.s.} \quad \text{for all } t \in [0, T].$$

Moreover, whenever $H = \{H_t\}_{t\in[0,T]}$ is a measurable process with the stated property,

$$\mathbb{E}\left[\int_0^t Y_s\mathrm{d}s \,\middle|\, \mathfrak{G}\right] = \int_0^t H_s\mathrm{d}s \text{ a.s.} \quad \text{for all } t \in [0, T].$$

A proof can be found in Ethier and Kurtz [16]. Note that it is not trivial that (some version of) the process $\{\mathbb{E}[Y_t|\mathfrak{G}]\}_{t\in[0,T]}$ is 'well-behaved'. In particular, H is *not* unique modulo indistinguishability; in Duffie and Epstein [10] it is assumed to be continuous.

Theorem B.3 (Stochastic Gronwall–Bellman inequality) *Let $Y = \{Y_t\}_{t\in[0,T]}$ be a right-continuous adapted process such that $\mathbb{E}[\sup_{s\in[0,T]} |Y_s|] < \infty$ and suppose that for some $\alpha \in (0, \infty)$ we have*

$$Y_t \geq \alpha\mathbb{E}\left[\int_t^T Y_s\mathrm{d}s \,\middle|\, \mathfrak{F}_t\right] \text{ a.s.} \quad \text{for every } t \in [0, T].$$

Then $Y_t \geq 0$ for all $t \in [0, T]$ a.s.

Proof Since Y is right-continuous, it suffices to prove $Y(t_0) \geq 0$ a.s. for each $t_0 \in [0, T]$. If the result is established for $t_0 = 0$, then applying it to each of the processes $\{Y_{t_0+t}\}_{t\in[0,T-t_0]}$ for $t_0 \in [0, T]$ yields the claim.[21] Hence, it is sufficient to show that $Y_0 \geq 0$ a.s. We choose a measurable modification H of the conditional expectations process $\{\mathbb{E}[Y_t|\mathfrak{F}_0]\}_{t\in[0,T]}$ and observe that

$$H_t \geq \alpha\mathbb{E}\left[\mathbb{E}\left[\int_t^T Y_s\mathrm{d}s \,\middle|\, \mathfrak{F}_t\right] \,\middle|\, \mathfrak{F}_0\right] = \alpha\mathbb{E}\left[\int_t^T Y_s\mathrm{d}s \,\middle|\, \mathfrak{F}_0\right] = \alpha\int_t^T H_s\mathrm{d}s \text{ a.s.}, \ t \in [0, T],$$

[20] A process $Y = \{Y_t\}_{t\in[0,T]}$ is measurable if the mapping $Y : [0, T]\times\Omega \to \mathbb{R}$ is $\mathfrak{B}([0, T])\otimes\mathfrak{F}$-measurable.

[21] Recall that we are not assuming that $\mathfrak{F}_0$ is trivial.

by iterated conditioning and Proposition B.2. It follows that $H \geq \alpha \int_{\cdot}^{T} H_s \mathrm{d}s$ a.e. on $[0, T] \times \Omega$, i.e. it is a.s. true that

$$H_t \geq \alpha \int_t^T H_s \mathrm{d}s \quad \text{for a.e. } t \in [0, T]. \tag{31}$$

Conditional dominated convergence yields $H_t = \mathbb{E}[Y_t | \mathfrak{F}_0] \to Y_0$ as $t \downarrow 0$ in L^1. Lemma B.1 and (31) imply that $H_t \geq 0$ for a.e. $t \in [0, T]$, a.s., i.e. $H_t \geq 0$ a.s. for a.e. $t \in [0, T]$. Choosing $\{t_n\}_{n \in \mathbb{N}}$ such that $t_n \downarrow 0$ and $H_{t_n} \geq 0$ a.s. for each $n \in \mathbb{N}$, and extracting a subsequence if necessary, we obtain $H_{t_n} \to Y_0$ a.s., whence $Y_0 \geq 0$ a.s. □

Finally, we generalize a crucial lemma by Costis Skiadas.

Corollary B.4 (Skiadas' Lemma) *Let $Y = \{Y_t\}_{t \in [0,T]}$ be a right-continuous adapted process with $Y_T = 0$ and $\mathbb{E}[\sup_{s \in [0,T]} |Y_s|] < \infty$. Moreover, assume that there exist a progressive process H and a constant $\alpha \in (0, \infty)$ such that*

$$Y_t = \mathbb{E}\left[\int_t^T H_s \mathrm{d}s \middle| \mathfrak{F}_t \right] \text{ a.s.} \quad \text{and} \quad H_t \geq -\alpha |Y_t| \text{ a.s.} \quad \text{for all } t \in [0, T]. \tag{32}$$

Then $Y_t \geq 0$ for all $t \in [0, T]$ a.s.

Proof By a similar argument as in the proof of Theorem B.3, it suffices to show that $Y_0 \geq 0$ a.s. We define the stopping time

$$\tau \triangleq \inf \{t \in [0, \infty) \,:\, Y_t > 0\} \wedge T.$$

Note that $Y_\tau \geq 0$ since Y is right-continuous and $Y_T = 0$. By (32) it follows[22] that $Y_t + \int_0^t H_s \mathrm{d}s = \mathbb{E}[\int_0^T H_s \mathrm{d}s | \mathfrak{F}_t]$ a.s., $t \in [0, T]$. So $\{Y_t + \int_0^t H_s \mathrm{d}s\}_{t \in [0,T]}$ is a martingale. Thus, $\mathbb{E}[1_{\{\tau > t\}}(Y_\tau + \int_0^\tau H_s \mathrm{d}s) | \mathfrak{F}_t] = 1_{\{\tau > t\}}(Y_t + \int_0^t H_s \mathrm{d}s)$ a.s. by optional stopping. Hence,

$$1_{\{\tau > t\}} Y_t = \mathbb{E}\left[\int_t^\tau 1_{\{\tau > t\}} H_s \mathrm{d}s + 1_{\{\tau > t\}} Y_\tau \middle| \mathfrak{F}_t \right] \text{ a.s.} \quad \text{for } t \in [0, T]. \tag{33}$$

The assumption on H yields $H_t(\omega) \geq -\alpha |Y_t(\omega)| = \alpha Y_t(\omega)$ for all $(t, \omega) \in [0, T] \times \Omega$ with $0 \leq t < \tau(\omega)$, and substituting this into (33), we get

$$1_{\{\tau > t\}} Y_t \geq \mathbb{E}\left[\int_t^\tau 1_{\{\tau > t\}} H_s \mathrm{d}s \middle| \mathfrak{F}_t \right] \geq \alpha \mathbb{E}\left[\int_t^T 1_{\{\tau > s\}} Y_s \mathrm{d}s \middle| \mathfrak{F}_t \right] \text{ a.s.}$$

Applying the Stochastic Gronwall–Bellman Inequality B.3 to $\{1_{\{\tau > t\}} Y_t\}_{t \in [0,T]}$, we find that $1_{\{\tau > 0\}} Y_0 \geq 0$ a.s. By definition of τ, we have $1_{\{\tau = 0\}} Y_0 \geq 0$. Hence, $Y_0 \geq 0$ a.s. □

References

1. Barles, G., Buckdahn, R., Pardoux, E.: Backward stochastic differential equations and integral-partial differential equations. Stoch. Stoch. Rep. **60**, 57–83 (1997)

[22] Here the assumption is needed that H is progressive.

2. Benzoni, L., Collin-Dufresne, P., Goldstein, R.: Can standard preferences explain the prices of out-of-the-money S&P 500 put options? NBER Working Paper W11861 (2005)
3. Bhamra, H.S., Kuehn, L.-A., Strebulaev, I.A.: The levered equity risk premium and credit spreads: a unified framework. Rev. Financ. Stud. **23**, 645–703 (2010)
4. Bhamra, H.S., Kuehn, L.-A., Strebulaev, I.A.: The aggregate dynamics of capital structure and macroeconomic risk. Rev. Financ. Stud. (2010) (forthcoming)
5. Bouchard, B., Touzi, N.: Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations. Stoch. Process. Appl. **111**, 175–206 (2004)
6. Buckdahn, R., Pardoux, E.: BSDE's with jumps and associated integro-partial differential equations, preprint (1994)
7. Chew, S.: A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the allais paradox. Econometrica **51**, 1065–1092 (1983)
8. Chew, S.: Axiomatic utility theories with the betweenness property. Ann. Oper. Res. **19**, 273–298 (1989)
9. Dekel, E.: An axiomatic characterization of preferences under uncertainty. J. Econ. Theory **40**, 304–318 (1986)
10. Duffie, D., Epstein, L.: Stochastic differential utility. Econometrica **60**, 353–394 (1992)
11. Duffie, D., Lions, P.-L.: PDE solutions of stochastic differential utility. J. Math. Econ. **21**, 577–606 (1992)
12. Duffie, D., Skiadas, C.: Continuous-time security pricing: a utility gradient approach. J. Math. Econ. **23**, 107–131 (1994)
13. El Karoui, N., Peng, S., Quenez, M.C.: Backward stochastic differential equations in finance. Math. Financ. **7**, 1–71 (1997)
14. Epstein, L.: the global stability of efficient intertemporal allocations. Econometrica **55**, 329–358 (1987)
15. Epstein, L., Zin, S.: Substitution, risk aversion and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica **57**, 937–969 (1989)
16. Ethier, S., Kurtz, T.: Markov Processes: Characterization and Convergence. 2nd edn. Wiley, New York (1986)
17. Fisher, M., Gilles, C.: Consumption and asset prices with recursive preferences. FEDS Working Paper 98–40 (1998)
18. Jacod, J., Shiryaev, A.: Limit Theorems for Stochastic Processes. 2nd edn. Springer, Berlin (1987)
19. Kreps, D., Porteus, E.: Temporal resolution of uncertainty and dynamic choice theory. Econometrica **46**, 185–200 (1978)
20. Lazrak, A., Quenez, M.C.: A generalized stochastic differential utility. Math. Oper. Res. **28**, 154–180 (2003)
21. Ma, C.: An existence theorem of intertemporal recursive utility in the presence of Lévy jumps. J. Math. Econ. **34**, 509–526 (2000)
22. Machina, M.: 'Expected utility' analysis without the independence axiom. Econometrica **50**, 277–323 (1982)
23. Schroder, M., Skiadas, C.: Optimal consumption and portfolio selection with stochastic differential utility. J. Econ. Theory **89**, 68–126 (1999)
24. Schroder, M., Skiadas, C.: Optimal lifetime consumption-portfolio strategies under trading constraints and generalized recursive preferences. Stoch. Process. Appl. **108**, 155–202 (2003)
25. Schroder, M., Skiadas, C.: Optimality and state pricing in constrained financial markets with recursive utility under continuous and discontinuous information. Math. Financ. **18**, 199–238 (2008)
26. Skiadas, C.: Recursive utility and preferences for information. Econ. Theory **12**, 293–312 (1998)
27. Skiadas, C.: Dynamic portfolio choice and risk aversion. In: Birge, J.R., Linetsky, V. (eds.) Handbooks in Operation Research and Management Science, vol. 15, pp. 789–843. Elsevier, New York (2008)
28. Skiadas, C.: Smooth ambiguity aversion toward small risks and continuous-time recursive utility, preprint. http://ssrn.com/abstract=1238156 (2008)
29. Svensson, L.E.O.: Portfolio choice with non-expected utility in continuous time. Econ. Lett. **30**, 313–317 (1989)
30. Tang, S., Li, X.: Necessary conditions for optimal control of stochastic systems with random jumps. SIAM J. Control Optim. **32**, 1447–1475 (1994)