Sensitivity of Optimal Consumption Streams*

Martin Herdegen† Johannes Muhle-Karbe‡

June 25, 2018

Abstract

We study the sensitivity of optimal consumption streams with respect to perturbations of the random endowment. At the leading order, the consumption adjustment does not matter: any choice that matches the budget constraint simply shifts the original utility by the marginal value of the perturbation. Nontrivial results can be obtained by considering the next-to-leading order. Here, one first solves the problem for a deterministic perturbation, which leads to a "prognosis measure". The desired consumption adjustment for a general endowment perturbation is in turn given by the conditional expectation of the latter, computed under this measure and appropriately weighted with the conditional expectations of the remaining risk-tolerance.

Keywords: optimal consumption, random endowment, asymptotic analysis.

AMS MSC 2010: Primary, 91B16; Secondary, 49N15, 60G44, 90C59

JEL Classification: G11, D91, E21

1 Introduction

Consumption-savings problems are ubiquitous in economics (cf., e.g., [27, 21, 3, 8] and the references therein). For example, households need to smooth the consumption financed by their labor income, and governments have to decide how to spend their tax revenue.[1]

Existence, uniqueness, and duality for problems of this kind are well understood, even in more general settings that also allow for investment in a financial market. See, e.g., [18, 19, 31] and the references therein. In contrast, beyond standard utilities and very particular endowment streams (see, e.g. [24, 5]), little is known about the qualitative and quantitative properties of the solution.

In this paper, we shed new light on this issue by means of asymptotic techniques. To wit, we start from a tractable benchmark endowment $(Y_t)_{t\in[0,T]}$ for which the solution is well understood – the canonical example is the case of a deterministic endowment. Then, we perturb the latter by

*We thank Christoph Czichowsky, Sebastian Herrmann, Semyon Malamud, Yaroslav Melnyk, Kasper Larsen, Rémy Praz, Frank Seifried, and, in particular, Hao Xing for fruitful discussions. Moreover, we thank two anonymous referees for their careful reading and pertinent remarks.

†University of Warwick, Department of Statistics, Coventry, CV4 7AL, UK, email `m.herdegen@warwick.ac.uk`.
Supported by the Swiss National Science Foundation (SNF) under grant 150101.

‡Carnegie Mellon University, Department of Mathematical Sciences, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA, email `johannesmk@cmu.edu`.

[1]A problem of this type is studied in [11] using the results of the present study.

an additional small endowment stream $(\varepsilon \Delta Y_t)_{t\in[0,T]}$, which can be completely general. For this perturbation, we in turn perform a sensitivity analysis of the consumption problem.

As suggested by the envelope theorem (see, e.g., [29, Theorem M.L.1]), at the leading order $O(\varepsilon)$, *any* adjustment for which the budget constraint is binding, i.e., which consumes all of the additional endowment, has the same asymptotic effect. To wit, expected utility is simply shifted by the marginal value of the endowment, evaluated using the state price density corresponding to the benchmark problem.

To understand the sensitivity of the optimal consumption stream with respect to changes in the endowment, we therefore pass to the next-to-leading order $O(\varepsilon^2)$ and exhibit a consumption adjustment that is asymptotically optimal. To this end, we proceed as follows.

First, we show that the perturbation of the general consumption problem at hand can be approximated by a problem for *quadratic* utility (with time- and state-dependent risk aversion) at the next-to-leading order $O(\varepsilon^2)$.[2] To prove this result, we first obtain a lower bound by analysing our concrete candidate strategy. An upper bound valid for all competitors is in turn derived by considering a suitable dual element. This methodology for asymptotic verification first seems to have appeared in the work of Henderson [10]. Different variants have since been used in a number of contexts by [22, 14, 25].

Next, we turn our attention to the approximating quadratic problem. Here, the basic idea is to decompose the analysis into two steps. The first is generic, in that it only depends on the baseline problem. The second depends on the specific perturbation at hand.

To wit, one first considers a *deterministic* perturbation of the baseline endowment, i.e., the question of how to optimally consume one extra dollar.[3] The solution of the corresponding quadratic problem yields a martingale, that can be used to define an auxiliary "prognosis measure". By switching to the latter, the case of general, random risk tolerances and interest rates can be reduced to the case where these quantities are deterministic, and explicit solution are readily available.[4] To wit, if discounted risk tolerances are deterministic, then the optimal consumption adjustment is given by the agent's prognosis of future endowment shocks, computed under her marginal pricing measure and suitably weighted by her risk-tolerance. In the general case, an analogous representation still obtains, *if* the estimates of future endowments and risk tolerances are computed under the prognosis measure determined in the first step. As a byproduct, we also obtain a formula for the second-order effect of the additional random endowment. Here, the key ingredients are the fluctuations of the endowment prognosis, weighted by future risk tolerances.

To carry out this program, we first consider a discretized version of the problem. Here, the steps outlined above can be carried through by means of explicit (backward) constructions. The continuous-time analogues can in turn be obtained by passing to the limit in an appropriate manner. Alternatively, under additional regularity conditions, they can be characterized directly by means of a BSDE driven by the investor's direct risk-tolerance process.[5]

[2] This parallels results for asymptotic utility-based pricing and hedging without intermediate consumption, where a quadratic hedging criterion also approximates its general counterpart [28, 1, 23, 16].

[3] For problems with a financial market but without intermediate consumption, an analogous notion plays a key role in the work of Kramkov and Sirbu [22].

[4] This is reminiscent of mean-variance hedging problems. Indeed, these admit an explicit solution for deterministic mean-variance tradeoffs [32]. In general, one first solves a "pure investment problem" independent of the specific random endowment. This in turn leads to a change of measure that neutralizes the effect of the random mean-variance tradeoff, see [4] and the references therein.

[5] This is in analogy to the risk-tolerance wealth process of Kramkov and Sirbu [22], which also has a backward

The results of the present study play a key role in [11], where they are used study how transaction costs are optimally consumed by the entity receiving them, e.g., a government reinvesting taxes or the operator of an exchange spending the fees it receives. Such consumption problems are difficult due to the singular nature of the transaction cost payments. However, for small costs, they can be made tractable using the asymptotic approach proposed here.

In our analysis, we focus on a pure consumption-savings problem without trading in risky assets. This is in some sense orthogonal to the work of Kramkov and Sirbu [22, 23], who consider consumption at the terminal time only and study the leading-order effect of a small random endowment on the investment in risky assets. It is an intriguing question whether the two approaches can be combined to obtain a full picture of the sensitivity of optimal investment and consumption with respect to small random endowments. However, this is bound to compound the substantial technical difficulties inherent in each part of the analysis. We therefore defer this challenging problem to future research.

The remainder of this article is organized as follows. In Section 2, we describe our setting for a general consumption-savings problem. Afterwards, we recall the well-known "first-order condition" for optimality. Section 4 contains the leading-order analysis of the problem. The main results of the paper, concerning the analysis at the next-to-leading order, are collected in Section 5. To provide some intuition, we first derive these results on an informal level, and then state and prove them in precise mathematical terms. The most technical aspects of these proofs are delegated to Appendices A and B.

2 Setting

This section describes our setup for a general consumption-savings problem. Throughout, we fix a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}, P)$ with finite time horizon $T > 0$; the filtration $\mathbb{F}$ is right-continuous and the initial σ-field $\mathcal{F}_0$ is P-trivial.

We consider an investor who receives a cumulative monetary random endowment $(Y_t)_{t\in[0,T]}$, which is adapted.[6] She uses her endowment to purchase a perishable consumption good. Her *consumption clock* μ, a finite (deterministic) measure on $[0, T]$, describes how she values consumption over time. We assume that

$$\mu([0,t]) < \mu([0,T]) \quad \text{for } t < T, \tag{2.1}$$

so that the effective time horizon of the investor is indeed T. Other than that, the clock μ can be completely general. Standard examples include $\mu(\mathrm{d}t) = \sum_{k=0}^{N} \delta_{\frac{kT}{N}}(\mathrm{d}t)$ (discrete consumption), $\mu(\mathrm{d}t) = \mathbf{1}_{(0,T)}(t)\,\mathrm{d}t$ (continuous consumption), or $\mu(\mathrm{d}t) = \mathbf{1}_{(0,T)}(t)\,\mathrm{d}t + \delta_T(\mathrm{d}t)$ (continuous consumption and terminal lump sum consumption).

Denote by $(B_t)_{t\in[0,T]}$ the exogenous exchange rate of consumption good against money, where $B_0 = 1$. We assume that B is absolutely continuous and positive, so that there exists an adapted

representation. The solution of this BSDE formally corresponds to the "indirect risk-tolerance process", that plays a pivotal role in the asymptotic analysis of optimization problems with small trading costs, compare [33, 15, 30].

[6] The endowment process Y need not be nondecreasing, of finite variation, or even a semimartingale, even though this will typically be the case in applications.

interest rate process $(r_t)_{t\in[0,T]}$ such that

$$\mathrm{d}B_t = r_t B_t\,\mathrm{d}t, \quad B_0 = 1.$$

By means of the bank account B, the investor can save and thereby smooth consumption over time. More specifically, she may purchase adapted *consumption processes* $(c_t)_{t\in[0,T]}$ satisfying $\int_{[0,T]} \frac{|c_t|}{B_t}\,\mu(\mathrm{d}u) < \infty$ P-a.s. and the *budget constraint*

$$\int_{[0,T]} \frac{c_t}{B_t}\,\mu(\mathrm{d}t) \leq Y_T \;\; P\text{-a.s.} \tag{2.2}$$

This means that no debt is allowed at the terminal time T; in contrast, the investor may borrow against future endowment at earlier times.

For $k \in \mathbb{N}_0$, we say that a consumption stream c is *kth-moment feasible* for Y and write $c \in \mathcal{A}_k(Y)$, if

$$E\left[\left(\int_{[0,T]} \frac{|c_t|}{B_t}\,\mu(\mathrm{d}u)\right)^k\right] < \infty.$$

We denote by $\mathcal{A}_k^*(Y)$ the subset of all consumption streams $c \in \mathcal{A}_k(Y)$ for which the budget constraint (2.2) is binding; with increasing marginal utilities and if μ has an atom at T, one can restrict to these policies without loss of generality. $\mathcal{A}_0(Y)$ is the largest class for which consumption-savings problems can be formulated. For technical reasons, we sometimes have to restrict ourselves to some subclass $\mathcal{A}_k(Y)$, $k > 0$. To wit, for given $k \in \mathbb{N}_0$, the investor maximises

$$\hat{c} = \operatorname*{argmax}_{c\in\mathcal{A}_k(Y)} E\left[\int_{[0,T]} U_t(\omega, c_t(\omega))\mu(\mathrm{d}t)\right]. \tag{2.3}$$

Here, her preferences are described by a *utility random field* $U : [0,T]\times\Omega\times\mathbb{R} \to \mathbb{R}\cup\{-\infty\}$ satisfying the following standard properties:

(a) for fixed x, the process $U_\cdot(\cdot, x)$ is progressively measurable,

(b) there exists $x_U \in \{-\infty, 0\}$ such that $U_\cdot(\cdot, x) \equiv -\infty$ for all $x \in (-\infty, x_U]$,

(c) for fixed t and ω, $U_t(\omega,\cdot)$ is increasing, strictly concave, and C^3 on $(x_U, +\infty)$.

As is customary, we usually drop the dependence on ω in the notation. Moreover, we write U' for $\frac{\partial U}{\partial x}$, etc., which is of course only defined on (x_u, ∞). Finally, we set $E\left[\int_{[0,T]} U_t(c_t)\,\mu(\mathrm{d}t)\right] := -\infty$ for $c \in \mathcal{A}_0(Y)$ with $E\left[\int_{[0,T]} U_t^-(c_t)\,\mu(\mathrm{d}t)\right] = -\infty$.

Example 2.1. U is called *standard deterministic utility field* if it does not depend on ω and is of the form

$$U_t(x) = \exp\left(-\int_0^t \beta_u\,\mathrm{d}u\right)u(x),$$

where $(\beta_t)_{t\in[0,T]}$ is a deterministic and bounded *impatience rate* and $u : \mathbb{R}\to\mathbb{R}\cup\{-\infty\}$ is a utility *function* satisfying

(a) there is $x_u \in \{-\infty, 0\}$ such that $u(x) = -\infty$ for all $x \in (-\infty, x_u]$,

(b) u is increasing, strictly concave and C^3 on (x_u, ∞),

(c) $\lim_{x \searrow x_u} u'(x) = +\infty$, $\lim_{x\to\infty} u'(x) = 0$. (2.4)

Typical choices for u are the exponential utility $u(x) = -\exp(-\gamma x)$ with $\gamma > 0$, the power utility $u(x) = \frac{x^{1-\gamma}}{1-\gamma}\mathbf{1}_{\{x>0\}} - \infty\mathbf{1}_{\{x\le 0\}}$ with $\gamma \in (0,\infty)\setminus\{1\}$, or the logarithmic utility $u(x) = \log(x)\mathbf{1}_{\{x>0\}} - \infty\mathbf{1}_{\{x\le 0\}}$.

For any utility random field U, we denote by $A^U, P^U : [0,T]\times\Omega\times(x_U,\infty) \to \mathbb{R}\cup\{-\infty\}$ its *absolute risk aversion* and *absolute prudence* (cf. [8]):

$$A_t^U(\omega,x) := -\frac{U_t''(\omega,x)}{U_t'(\omega,x)}, \qquad P_t^U(\omega,x) := -\frac{U_t'''(\omega,x)}{U_t''(\omega,x)}.$$

We say that U has *decreasing absolute risk aversion (DARA)* if $A_t^U(\omega,\cdot)$ is nonincreasing for fixed t and ω, and *decreasing absolute prudence (DAP)* if $P_t^U(\omega,\cdot)$ is nonincreasing for fixed t and ω, compare [8, p.25]. Note that DARA implies $U''' > 0$ and $P^U > 0$, whereas DAP implies that U''' is decreasing. If U is a standard deterministic utility field and the corresponding utility function u is exponential, power or logarithmic, then U satisfies DARA and DAP.

Finally, $\tilde U : [0,T]\times\Omega\times(0,\infty)\to\mathbb{R}$ denotes the *conjugate* of U:

$$\tilde U_t(\omega,y) := \sup_{x\in\mathbb{R}}(U_t(\omega,x) - xy). \tag{2.5}$$

By a standard result in convex analysis, for fixed t and ω, $\tilde U_t(\omega,\cdot)$ is decreasing, strictly convex and C^3. Moreover:

$$\tilde U_t'(y) = -(U_t')^{-1}(y), \tag{2.6}$$
$$\tilde U_t''(y) = -\frac{1}{U_t''((U_t')^{-1}(y))}, \tag{2.7}$$
$$\tilde U_t'''(y) = \frac{U_t'''((U_t')^{-1}(y))}{(U_t''((U_t')^{-1}(y)))^3} = \frac{-P_t^U((U_t')^{-1}(y))}{(U_t''((U_t')^{-1}(y)))^2},$$

where $\tilde U' = \frac{\partial \tilde U}{\partial y}$, etc. Note that under DARA and DAP, $\tilde U'''$ is negative and increasing.

3 A sufficient condition for optimality

Before turning to our perturbation analysis, we recall the well-known "first-order condition": a consumption stream $\hat c$ is optimal if the budget constraint (2.2) binds and the marginal utility, evaluated along $\hat c$, is a state price density.

Lemma 3.1. *Let $k \ge 2$. If there are $\hat c \in \mathcal{A}_k^*(Y)$ with $\hat c > x_U$ and $E\left[\int_{[0,T]} |U_t(\hat c_t)|\,\mu(\mathrm{d}t)\right] < \infty$*

and a positive square-integrable martingale $(\hat{Z}_t)_{t\in[0,T]}$ *such that*[7]

$$U_t'(\hat{c}_t) = \frac{\hat{Z}_t}{B_t}, \quad t \in [0,T], \tag{3.1}$$

then $\hat{c}$ *is optimal in* $\mathcal{A}_k(Y)$.

Proof. Let $c \in \mathcal{A}_k(Y)$ be a competing consumption stream with $E\left[\int_{[0,T]} U_t^-(c_t)\mu(\mathrm{d}t)\right] < \infty$. Define the measure $\hat{Q} \approx P$ on $\mathcal{F}_T$ by $\mathrm{d}\hat{Q} = \frac{\hat{Z}_T}{\hat{Z}_0}\,\mathrm{d}P$. Then concavity of U, the first-order condition (3.1), Corollary A.2(a), the budget constraint (2.2) for c and $\hat{c}$ (which is binding for $\hat{c}$) give

$$\begin{aligned} E\left[\int_{[0,T]} (U_t(c_t) - U_t(\hat{c}_t))\,\mu(\mathrm{d}t)\right] &\leq E\left[\int_{[0,T]} U_t'(\hat{c}_t)(c_t - \hat{c}_t)\,\mu(\mathrm{d}t)\right] \\ &= E\left[\int_{[0,T]} \hat{Z}_t\left(\frac{c_t}{B_t} - \frac{\hat{c}_t}{B_t}\right)\mu(\mathrm{d}t)\right] \\ &= \hat{Z}_0 E^{\hat{Q}}\left[\int_{[0,T]} \left(\frac{c_t}{B_t} - \frac{\hat{c}_t}{B_t}\right)\mu(\mathrm{d}t)\right] \leq \hat{Z}_0 E^{\hat{Q}}\left[Y_T - Y_T\right] = 0, \end{aligned}$$

as claimed. □

Standing assumption Henceforth, we assume that there exist a consumption stream $\hat{c} \in \mathcal{A}_2^*(Y)$ and a positive square-integrable martingale $\hat{Z}$ satisfying the conditions of Lemma 3.1. We denote by $\hat{Q} \approx P$ on $\mathcal{F}_T$ the *marginal pricing measure* given by $\mathrm{d}\hat{Q} = \frac{\hat{Z}_T}{\hat{Z}_0}\,\mathrm{d}P$ (cf. [6, 17]).

In general, the first-order condition is only guaranteed to hold for a dual variable from a larger class of supermartingale densities, compare [18, 19, 31]. However, it is often satisfied in concrete examples, cf., e.g., [5, 11]. In particular, it holds if all primitives of the model are deterministic. In this case, which serves as the expansion point for the perturbation analysis of *small* endowments,[8] the optimal consumption stream $\hat{c}$ and the martingale $\hat{Z}$ in Lemma 3.1 are deterministic and can be computed explicitly up to the solution of a scalar equation:

Lemma 3.2. *Suppose the total endowmentx* $Y_T > x_u$ *is constant, the interest rate* $(r_t)_{t\in[0,T]}$ *is deterministic and bounded, and* U *is a standard deterministic utility field (cf. Example 2.1). Then the process*

$$\hat{c}_t = (u')^{-1}\left(\hat{z}\exp\left(\int_0^t \beta_u - r_u\,\mathrm{d}u\right)\right) \tag{3.2}$$

is optimal in $\mathcal{A}_k(Y)$ *for all* $k \geq 2$. *Here,* $\hat{z} > 0$ *is the unique solution of*

$$\int_0^T (u')^{-1}\left(\hat{z}\exp\left(\int_0^t (\beta_u - r_u)\,\mathrm{d}u\right)\right)\exp\left(-\int_0^t r_u\,\mathrm{d}u\right)\mu(\mathrm{d}t) = Y_T; \tag{3.3}$$

the martingale $\hat{Z}$ *from Lemma 3.1 is given by* $\hat{Z} \equiv \hat{z}$.

Proof. By the Inada conditions (2.4), the function $(u')^{-1}$ is defined on $(0,\infty)$, is continuous, strictly decreasing, and satisfies $\lim_{y\searrow 0}(u')^{-1}(y) = +\infty$ and $\lim_{y\to\infty}(u')^{-1}(y) = x_u$. (Recall that

[7] Since consumptions streams are typically not nonnegative here, we need to assume that $\hat{Z}$ is square-integrable in order to apply Bayes' theorem in the form of Corollary A.2(a).

[8] Note, however, that our results allow to expand around arbitrary, not necessarily deterministic, endowments.

x_u is either $-\infty$ or 0.) Hence, as r and β are deterministic and bounded and μ is a finite measure on $[0,T]$, the function

$$f(z) := \int_0^T (u')^{-1} \left(\hat{z} \exp \left(\int_0^t (\beta_u - r_u)\,\mathrm{d}u \right) \right) \exp \left(- \int_0^t r_u \,\mathrm{d}u \right) \mu(\mathrm{d}t)$$

is well defined on $(0,\infty)$, continuous, strictly decreasing, and satisfies $\lim_{z\searrow 0} f(z) = +\infty$ as well as $\lim_{z\to\infty} f(z) = x_u$. Thus, by the intermediate value theorem, there exists a unique $\hat{z} > 0$ satisfying (3.3). If we define $\hat{c}$ by (3.2) and set $\hat{Z} \equiv \hat{z}$, then $t \mapsto \hat{c}_t$ is continuous, deterministic and bounded, and belongs to $\mathcal{A}_k^*(Y)$ for any $k \geq 2$. The claim now follows from Lemma 3.1. $\square$

For exponential, power, or logarithmic utilities, the constant $\hat{z}$ and the corresponding consumption stream $\hat{c}$ from Lemma 3.2 can be readily calculated explicitly.

4 First-order optimality

We now turn to the sensitivity analysis of the optimal consumption problem. Suppose that the investor receives the perturbed cumulative endowment stream

$$Y_t^\varepsilon = Y_t + \varepsilon \Delta Y_t, \quad t \in [0,T],$$

for $\varepsilon > 0$ and an adapted process $(\Delta Y_t)_{t\in[0,T]}$, null at 0. The leading-order effect of the perturbation $\varepsilon\Delta Y$ can in turn be described as follows:

Theorem 4.1. *Suppose that $E^{\hat{Q}}[|\Delta Y_T|] < \infty$ and U satisfies DARA. Moreover, assume there exist $\Delta c \in \mathcal{A}_2^*(\Delta Y)$ and $\varepsilon_0 > 0$ such that $\hat{c} - \varepsilon_0|\Delta c| > x_U$ P-a.s. and*

$$E\left[\int_{[0,T]} -U_t''(\hat{c}_t - \varepsilon_0|\Delta c_t|)(\Delta c_t)^2 \,\mu(\mathrm{d}t)\right] < \infty. \tag{4.1}$$

Then the process $(\hat{c}^\varepsilon)_{t\in[0,T]}$, defined by

$$\hat{c}^\varepsilon = \hat{c} + \varepsilon \Delta c, \quad \varepsilon \in [0, \varepsilon_0),$$

belongs to $\mathcal{A}_2(Y^\varepsilon)$ and is optimal in $\mathcal{A}_2(Y^\varepsilon)$ at the leading order $O(\varepsilon)$, i.e.:

$$E\left[\int_{[0,T]} U_t(\hat{c}_t^\varepsilon)\,\mu(\mathrm{d}t)\right] \geq \sup_{c^\varepsilon\in\mathcal{A}_2(Y^\varepsilon)} E\left[\int_{[0,T]} U_t(c_t^\varepsilon)\,\mu(\mathrm{d}t)\right] - o(\varepsilon).$$

The corresponding leading-order maximal utility is given by

$$E\left[\int_{[0,T]} U_t(\hat{c}_t^\varepsilon)\,\mu(\mathrm{d}t)\right] = E\left[\int_{[0,T]} U_t(\hat{c}_t)\,\mu(\mathrm{d}t)\right] + \varepsilon \hat{Z}_0 E^{\hat{Q}}[\Delta Y_T] - O(\varepsilon^2). \tag{4.2}$$

Theorem 4.1 states that – modulo integrability conditions[9] – *any* consumption correction $\varepsilon\Delta c$

[9]The second-order moments are needed are needed to control the second-order remainder terms appearing in the Taylor expansion that we use to approximate general utilities by quadratic ones in the proof of Theorem 4.1.

is optimal at the leading order $O(\varepsilon)$ as long as the budget constraint $\int_{[0,T]} \frac{\varepsilon \Delta c_t}{B_t}\, \mu(\mathrm{d}t) = \varepsilon \Delta Y_T$ is binding, i.e., all extra endowment is consumed eventually. Whence, the timing of the extra consumption only has a second-order effect. The corresponding leading-order welfare correction $\varepsilon \hat{Z}_0 E^{\hat{Q}}[\Delta Y_T]$ is simply the marginal-utility based price of the perturbation [6, 17].

A sufficient set of conditions for Theorem 4.1 is that the endowment correction ΔY_T is uniformly bounded and μ has an atom at T. (The unperturbed endowment Y_T can be general.) In this case, the consumption correction Δc can be chosen uniformly bounded (regardless whether the unperturbed optimizer $\hat{c}$ is also bounded or not) and (4.1) is then easily seen to be satisfied for standard utility functions – assuming also that the unperturbed optimizer $\hat{c}$ is uniformly bounded away from 0 for utilities on the positive real line.

Remark 4.2. Theorem 4.1 can be generalized as follows: Assume that $Y^\varepsilon = Y + \varepsilon \Delta Y^\varepsilon$, where ΔY^ε converges to ΔY in $L^1(\hat{Q})$. If $\Delta c^\varepsilon \in \mathcal{A}_2^*(\Delta Y^\varepsilon)$ with $\hat{c} - \varepsilon|\Delta c^\varepsilon| > x_U$ P-a.s. and $\limsup_{\varepsilon \downarrow 0} E\left[\int_{[0,T]} -U_t''(\hat{c}_t - \varepsilon|\Delta c_t^\varepsilon|)(\Delta c_t^\varepsilon)^2\, \mu(\mathrm{d}t)\right] < \infty$, then an inspection of the proof of Theorem 4.1 shows that it remains valid (with $\hat{c}^\varepsilon = \hat{c} + \varepsilon \Delta c^\varepsilon$) if $O(\varepsilon^2)$ is replaced by $o(\varepsilon)$ in (4.2).

Proof of Theorem 4.1. We first establish a primal lower bound for the candidate $\hat{c}^\varepsilon$, then derive a dual upper bound for any competitor $c^\varepsilon \in \mathcal{A}_2(Y^\varepsilon)$, and finally compare the two.

Primal lower bound. For $\varepsilon \in [0, \varepsilon_0)$ and for fixed t and ω, a Taylor expansion of order one with Lagrange remainder term gives

$$U_t(\hat{c}_t^\varepsilon) = U_t(\hat{c}_t) + U_t'(\hat{c}_t)(\hat{c}_t^\varepsilon - \hat{c}_t) + \frac{1}{2} U_t''(\tilde{c}(t,\omega))(\hat{c}_t^\varepsilon - \hat{c}_t)^2,$$

where $\tilde{c}(t,\omega)$ takes values in the interval with endpoints $\hat{c}_t(\omega)$ and $\hat{c}_t^\varepsilon(\omega)$. By definition of $\hat{c}^\varepsilon$, the fact that U'' is increasing by DARA, and since $\varepsilon < \varepsilon_0$, we obtain the pointwise inequality

$$U_t(\hat{c}_t^\varepsilon) \geq U_t(\hat{c}_t) + \varepsilon U_t'(\hat{c}_t)\Delta c_t + \frac{1}{2}\varepsilon^2 U_t''(\hat{c}_t - \varepsilon_0|\Delta c_t|)\Delta c_t^2, \quad t \in [0,T]. \tag{4.3}$$

Take the expectation of the integral of the first-order term on the right-hand side of (4.3) and use the first-order condition (3.1), Corollary A.2(a), and $\Delta c \in \mathcal{A}_2^*(\Delta Y)$ to obtain

$$\begin{aligned} \varepsilon E\left[\int_{[0,T]} U_t'(\hat{c}_t)\Delta c_t\, \mu(\mathrm{d}t)\right] &= \varepsilon E\left[\int_{[0,T]} B_t U_t'(\hat{c}_t)\frac{\Delta c_t}{B_t}\, \mu(\mathrm{d}t)\right] = \varepsilon E\left[\int_{[0,T]} \hat{Z}_t \frac{\Delta c_t}{B_t}\, \mu(\mathrm{d}t)\right] \\ &= \varepsilon \hat{Z}_0 E^{\hat{Q}}\left[\int_{[0,T]} \frac{\Delta c_t}{B_t}\, \mu(\mathrm{d}t)\right] = \varepsilon \hat{Z}_0 E^{\hat{Q}}[\Delta Y_T]. \end{aligned}$$

Now taking the expectations of the integral on both sides of (4.3) and using (4.1) yields (4.2).

Dual upper bound. We proceed to show that

$$\sup_{c^\varepsilon \in \mathcal{A}_2(Y^\varepsilon)} E\left[\int_{[0,T]} U_t(c_t)\, \mu(\mathrm{d}t)\right] \leq E\left[\int_{[0,T]} U_t(\hat{c}_t)\, \mu(\mathrm{d}t)\right] + \varepsilon \hat{Z}_0 E^{\hat{Q}}[\Delta Y_T]. \tag{4.4}$$

To this end, let $c^\varepsilon \in \mathcal{A}_2(Y^\varepsilon)$ be any consumption stream with $E\left[\int_{[0,T]} U_t^-(c_t^\varepsilon)\, \mu(\mathrm{d}t)\right] < \infty$. The definition of the conjugate $\tilde{U}$ (cf. 2.5) and the first-order condition (3.1) give the pointwise estimate

$$U_t(c_t^\varepsilon) \leq \tilde{U}_t\left(\frac{\hat{Z}_t}{B_t}\right) + \frac{\hat{Z}_t}{B_t}c_t^\varepsilon = U_t(\hat{c}_t) - \frac{\hat{Z}_t}{B_t}\hat{c}_t + \frac{\hat{Z}_t}{B_t}c_t^\varepsilon = U_t(\hat{c}_t) + \frac{\hat{Z}_t}{B_t}(c_t^\varepsilon - \hat{c}_t). \tag{4.5}$$

Now the same argument as in the proof of Lemma 3.1 gives

$$E\left[\int_{[0,T]} (U_t(c_t^\varepsilon) - U_t(\hat{c}_t))\, \mu(\mathrm{d}t)\right] \leq \hat{Z}_0 E^{\hat{Q}}\left[Y_T^\varepsilon - Y_T\right] \leq \varepsilon \hat{Z}_0 E^{\hat{Q}}\left[\Delta Y_T\right].$$

Comparison of lower and upper bound. Comparing (4.2) to (4.4) shows that $\hat{c}^\varepsilon$ is indeed optimal for $\mathcal{A}_2(Y^\varepsilon)$ at the leading order $O(\varepsilon)$. □

Remark 4.3. The dual considerations in the proof of Theorem 4.1 also show that $\hat{Z}$ is first-order optimal for the minimization problem dual to 2.3 (cf. [18, 19, 31, 14] for more details.)

5 Second-order optimality

Theorem 4.1 shows that, at the leading order $O(\varepsilon)$, *any* consumption correction $\Delta c^\varepsilon := c^\varepsilon - \hat{c}$ for the perturbed endowment $Y^\varepsilon = Y + \varepsilon \Delta Y$ is optimal as long as the corresponding budget constraint is satisfied with equality. To understand the sensitivity of optimal consumption streams with respect to a perturbations of the endowment, it is therefore necessary to perform a second-order expansion including terms of order $O(\varepsilon^2)$. This is done in the present section. To provide some intuition, we start with a heuristic derivation. Afterwards, we state and prove our main results in precise mathematical terms.

5.1 Heuristics

A key ingredient for our perturbation analysis is the *direct risk tolerance process* $(\rho_t)_{t\in[0,T]}$ with respect to current consumption:[10]

$$\rho_t = -\frac{U_t'(\hat{c}_t)}{U_t''(\hat{c}_t)}. \tag{5.1}$$

In view of Theorem 4.1, c^ε is *first-order* optimal for the perturbed endowment $Y^\varepsilon = Y + \varepsilon\Delta Y$ if $c^\varepsilon = \hat{c} + \varepsilon \Delta c$ and $\Delta c \in \mathcal{A}_2^*(\Delta Y)$. For each such Δc, a formal *second-order* Taylor expansion gives

$$U_t(c_t^\varepsilon) = U_t(\hat{c}_t) + \varepsilon U_t'(\hat{c}_t)\Delta c_t + \frac{1}{2}\varepsilon^2 U_t''(\hat{c}_t)(\Delta c_t)^2 + o(\varepsilon^2).$$

As $U'' < 0$, we therefore have to solve the following quadratic minimisation problem to determine a second-order optimal policy:

$$\Delta\hat{c} = \underset{\Delta c \in \mathcal{A}_2^*(\Delta Y)}{\operatorname{argmin}}\ E\left[\int_{[0,T]} -U_t''(\hat{c}_t)(\Delta c_t)^2\, \mu(\mathrm{d}t)\right].$$

[10] Likewise, without intermediate consumption, the terminal risk-tolerance ρ_T is the crucial object in the analysis of utility-based prices and hedging strategies [23]. The direct risk-tolerance process also plays a pivotal role in models with small trading costs [15].

Set $\Delta c = \rho \Delta Z$ for some process $(\Delta Z_t)_{t\in[0,T]}$ to be determined. Together with the first-order condition (3.1) for $\hat{c}$ and the definition of ρ, this leads to the following minimisation problem for ΔZ:

$$\text{Minimise } E^{\hat{Q}}\left[\int_{[0,T]} \frac{\rho_t}{B_t}(\Delta Z_t)^2\,\mu(\mathrm{d}t)\right], \text{ subject to } \int_{[0,T]} \frac{\rho_t}{B_t}\Delta Z_t\,\mu(\mathrm{d}t) = \Delta Y_T. \tag{5.2}$$

If $\Delta\hat{Z}$ is a $\hat{Q}$-martingale satisfying the constraint $\int_{[0,T]} \frac{\rho_t}{B_t}\Delta\hat{Z}_t\,\mu(\mathrm{d}t) = \Delta Y_T$, then $\Delta\hat{Z}$ is optimal for (5.2).[11] Indeed, let ΔZ be any competitor satisfying $\int_{[0,T]} \frac{\rho_t}{B_t}\Delta Z_t\,\mu(\mathrm{d}t) = \Delta Y_T$. We have to show that

$$E^{\hat{Q}}\left[\int_{[0,T]} \frac{\rho_t}{B_t}((\Delta Z_t)^2 - (\Delta\hat{Z}_t)^2)\,\mu(\mathrm{d}t)\right] \geq 0.$$

In view of the pointwise inequality $(\Delta Z_t)^2 - (\Delta\hat{Z}_t)^2 \geq 2(\Delta Z_t - \Delta\hat{Z}_t)\Delta\hat{Z}_t$, it suffices to establish

$$E^{\hat{Q}}\left[\int_{[0,T]} \Delta\hat{Z}_t\frac{\rho_t}{B_t}(\Delta Z_t - \Delta\hat{Z}_t)\,\mu(\mathrm{d}t)\right] = 0.$$

To ease notation, define $K_t = \int_{[0,t]} \frac{\rho_u}{B_u}(\Delta Z_u - \Delta\hat{Z}_u)\,\mu(\mathrm{d}u)$. Integration by parts, $K_T = 0$, and the $\hat{Q}$-martingale property of $\Delta\hat{Z}$ yield

$$\begin{aligned} E^{\hat{Q}}\left[\int_{[0,T]} \Delta\hat{Z}_t\frac{\rho_t}{B_t}(\Delta Z_t - \Delta\hat{Z}_t)\,\mu(\mathrm{d}t)\right] &= E^{\hat{Q}}\left[\Delta\hat{Z}_0 K_0 + \int_{(0,T]} \Delta\hat{Z}_t\,\mathrm{d}K_t\right] \\ &= E^{\hat{Q}}\left[\Delta\hat{Z}_T K_T - \int_{(0,T]} K_{t-}\,\mathrm{d}\Delta\hat{Z}_t\right] = 0. \end{aligned}$$

In summary, second-order optimality boils down to finding a $\hat{Q}$-martingale $\Delta\hat{Z}$ such that

$$\int_{[0,T]} \frac{\rho_t}{B_t}\Delta\hat{Z}_t\,\mu(\mathrm{d}t) = \Delta Y_T. \tag{5.3}$$

To solve (5.3), first consider the case where the discounted risk tolerance $(\frac{\rho_t}{B_t})_{t\in[0,T]}$ is *deterministic*. Define the *remaining (discounted) risk tolerance*

$$R_t = \int_{[t,T]} \frac{\rho_u}{B_u}\,\mu(\mathrm{d}u), \quad t \in [0,T],$$

which measures the investor's sensitivity with respect to changes in future consumption. Integra-

[11] To motivate this, denote by $\hat{Z}^\varepsilon$ the dual martingale corresponding to $\hat{c}^\varepsilon$. Then, a Taylor expansion of the first-order condition $\frac{\hat{Z}^\varepsilon_t}{B_t} = U_t'(\hat{c}^\varepsilon_t)$ for $\hat{c}^\varepsilon$ yields $\frac{\hat{Z}^\varepsilon_t}{B_t} + O(\varepsilon^2) = U_t'(\hat{c}_t) + \varepsilon U_t''(\hat{c}_t)\Delta\hat{c}_t = U_t'(\hat{c}_t)(1 - \varepsilon\frac{\Delta\hat{c}_t}{\rho_t})$. The first-order condition (3.1) for $\hat{c}$ in turn gives $\hat{Z}^\varepsilon_t = \hat{Z}_t(1 - \varepsilon\frac{\Delta\hat{c}_t}{\rho_t}) + O(\varepsilon^2)$, which suggests that $\frac{\Delta\hat{c}}{\rho}$ should be a $\hat{Q}$-martingale.

tion by parts yields

$$\begin{aligned}
\Delta Y_T &= \int_{[0,T]} \Delta\hat{Z}_t \frac{\rho_t}{B_t}\mu(\mathrm{d}t) = \Delta\hat{Z}_0 \frac{\rho_0}{B_0}\mu(\{0\}) - \int_{(0,T]} \Delta\hat{Z}_t \,\mathrm{d}R_{t+} \\
&= \Delta\hat{Z}_0 \frac{\rho_0}{B_0}\mu(\{0\}) + \Delta\hat{Z}_0 R_{0+} - \Delta\hat{Z}_T R_{T+} + \int_{(0,T]} R_t \,\mathrm{d}\Delta\hat{Z}_t \\
&= \Delta\hat{Z}_0 R_0 + \int_{(0,T]} R_t \,\mathrm{d}\Delta\hat{Z}_t. \qquad (5.4)
\end{aligned}$$

Denote by $M_t^{\Delta Y} = E^{\hat{Q}}[\Delta Y_T \,|\, \mathcal{F}_t]$ the $\hat{Q}$-martingale generated by the total endowment, i.e., the agent's *best prognosis* of ΔY_T under the marginal pricing measure. Taking conditional $\hat{Q}$-expectations in (5.4) and using that $\Delta\hat{Z}$ needs to be a $\hat{Q}$-martingale gives

$$M_t^{\Delta Y} = \Delta\hat{Z}_0 R_0 + \int_{(0,t]} R_u \,\mathrm{d}\Delta\hat{Z}_u,$$

and in turn

$$\Delta\hat{Z}_t = \frac{M_0^{\Delta Y}}{R_0} + \int_{(0,t]} \frac{1}{R_u} \,\mathrm{d}M_u^{\Delta Y}. \qquad (5.5)$$

In summary, for deterministic discounted direct risk tolerances ρ/B, the optimal consumption correction $\Delta\hat{c}$ is the product of the direct risk-tolerance ρ and a $\hat{Q}$-martingale $\Delta\hat{Z}$. The latter is given explicitly by Formula (5.5) as the agent's best prognosis of the future endowment ΔY_T at time t, appropriately weighted by her remaining risk-tolerance R.

Let us now pass to the general case with possibly stochastic interest rates and risk tolerances. Here, the idea is to reduce to the case studied above by a suitable change of measure. The latter is determined by the solution for the problem with a unit perturbation. Indeed, suppose that we can find a positive $\hat{Q}$-martingale $(Z_t^\rho)_{t\in[0,T]}$ satisfying

$$\int_{[0,T]} \frac{\rho_t}{B_t} Z_t^\rho \,\mu(\mathrm{d}t) = 1, \qquad (5.6)$$

so that $(Z_t^\rho)_{t\in[0,T]}$ is the optimal consumption correction for $\Delta Y_T = 1$.[12] This in turn allows to reduce the case of random risk-tolerances to the deterministic one by passing from the marginal pricing measure $\hat{Q}$ to the "prognosis measure" $Q^\rho \approx \hat{Q} \approx P$ defined by $\mathrm{d}Q^\rho = \frac{Z_T^\rho}{Z_0^\rho}\,\mathrm{d}P$.[13] Note that if the the total discounted risk-tolerance $(\frac{\rho_t}{B_t})_{t\in[0,T]}$ is deterministic, then no change of measure is necessary because Z^ρ is given by the constant $Z^\rho \equiv (\int_{[0,T]} \frac{\rho_t}{B_t}\,\mu(\mathrm{d}t))^{-1}$ in this case.

To carry out this program, write the $\hat{Q}$-martingale $\Delta\hat{Z}$ as $Z^\rho Z^{\Delta Y}$, where $(Z_t^{\Delta Y})_{t\in[0,T]}$ is a Q^ρ-martingale to be determined and denote by $(R_t)_{t\in[0,T]}$ the Q^ρ-expected remaining (discounted) risk tolerance:

$$R_t = E^{Q^\rho}\left[\int_{[t,T]} \frac{\rho_u}{B_u}\,\mu(\mathrm{d}u) \,\middle|\, \mathcal{F}_t\right], \quad t\in[0,T]. \qquad (5.7)$$

[12] The process $(Z_t^\rho)_{t\in[0,T]}$ depends on the interest rate, the direct risk-tolerance, and the marginal pricing measure, but not on the particular endowment perturbation at hand. In Markovian baseline settings, it can be readily determined from the benchmark policy.

[13] This parallels the situation for mean-variance hedging [9, 4] and asymptotic utility-based pricing and hedging [28, 1, 16, 22, 23], where one first solves a "pure investment problem" independent of the particular random endowment, and in turn simplifies the original problem by a suitable change of measure (and, sometimes, numeraire).

This is investor's best estimate of her discounted future risk tolerances, computed under the prognosis measure Q^ρ. By Bayes' theorem and the definition of Z^ρ in (5.6),

$$Z^\rho_t R_t = E^{\hat{Q}}\left[\int_{[t,T]} Z^\rho_u \frac{\rho_u}{B_u}\,\mu(\mathrm{d}u)\,\middle|\,\mathcal{F}_t\right] = 1 - \int_{[0,t)} Z^\rho_u \frac{\rho_u}{B_u}\,\mu(\mathrm{d}u). \tag{5.8}$$

In particular, $Z^\rho_0 R_0 = 1$ and an integration by parts yields

$$\begin{aligned}
\Delta Y_T = \int_{[0,T]} Z^{\Delta Y}_t Z^\rho_t \frac{\rho_t}{B_t}\,\mu(\mathrm{d}t) &= Z^{\Delta Y}_0 Z^\rho_0 \frac{\rho_0}{B_0}\mu(\{0\}) - \int_{(0,T]} Z^{\Delta Y}_t \,\mathrm{d}(Z^\rho_t R_{t+}) \\
&= Z^{\Delta Y}_0 Z^\rho_0 \frac{\rho_0}{B_0}\mu(\{0\}) + Z^{\Delta Y}_0 Z^\rho_0 R_{0+} - Z^{\Delta Y}_T Z^\rho_T R_{T+} + \int_{(0,T]} Z^\rho_t R_t \,\mathrm{d}Z^{\Delta Y}_t \\
&= Z^{\Delta Y}_0 Z^\rho_0 R_0 + \int_{(0,T]} Z^\rho_t R_t \,\mathrm{d}Z^{\Delta Y}_t \\
&= Z^{\Delta Y}_0 + \int_{(0,T]} Z^\rho_t R_t \,\mathrm{d}Z^{\Delta Y}_t.
\end{aligned} \tag{5.9}$$

As the final ingredient, define the investor's best estimate of her random endowment, also computed under the prognosis measure:

$$M^{\Delta Y}_t = E^{Q^\rho}\left[\Delta Y_T \,|\, \mathcal{F}_t\right], \quad t \in [0,T].$$

Taking conditional Q^ρ-expectations in (5.9) and using that $Z^{\Delta Y}$ needs to be a Q^ρ-martingale gives

$$M^{\Delta Y}_t = Z^{\Delta Y}_0 + \int_{(0,t]} Z^\rho_u R_u \,\mathrm{d}Z^{\Delta Y}_u,$$

or, equivalently:

$$Z^{\Delta Y}_t = M^{\Delta Y}_0 + \int_{(0,t]} \frac{1}{Z^\rho_u R_u}\,\mathrm{d}M^{\Delta Y}_u. \tag{5.10}$$

For general ρ and B, the optimal (normalised) consumption correction $\Delta\hat{c}$ therefore is the product of three terms. Like for deterministic ρ/B, the first one is the direct risk-tolerance ρ. The new ingredient is the martingale Z^ρ, which governs the change of measure to the prognosis measure. The last factor, $Z^{\Delta Y}$, is given by Formula (5.10) as the investor's best prognosis of her random endowment, weighted by her expected risk tolerances. The only difference to the case of deterministic ρ/B is that these estimates are now computed under the prognosis measure Q^ρ instead of the marginal pricing measure $\hat{Q}$.

Remark 5.1. Suppose the underlying filtration is continuous and $\mu(\mathrm{d}t) = \mathbf{1}_{(0,T)}(t)\,\mathrm{d}t + \delta_T(\mathrm{d}t)$. Then, the above heuristic arguments suggest a characterization of the processes R and Z^ρ by means of a quadratic BSDE. Indeed, (5.8) and the product formula yield

$$Z^\rho_t\,\mathrm{d}R_t + R_t \mathrm{d}Z^\rho_t + \mathrm{d}\langle Z^\rho, R\rangle_t = -Z^\rho_t \rho_t\,\mathrm{d}t, \quad Z^\rho_T R_T = Z^\rho_T \rho_T,$$

or, equivalently:

$$\mathrm{d}R_t = -\frac{R_t}{Z_t^\rho}\mathrm{d}Z_t^\rho - \frac{1}{Z_t^\rho}\,\mathrm{d}\langle Z^\rho, R\rangle_t - \rho_t\,\mathrm{d}t, \quad R_T = \rho_T. \tag{5.11}$$

Now, define the martingale

$$\mathrm{d}M_t^\rho = -\frac{R_t}{Z_t^\rho}\mathrm{d}Z_t^\rho, \quad M_0^\rho = Z_0^\rho. \tag{5.12}$$

Plugging this into (5.11) in turn leads to a quadratic BSDE for the process R:

$$\mathrm{d}R_t = \mathrm{d}M_t^\rho + \frac{1}{R_t}\,\mathrm{d}\langle M^\rho\rangle_t - \rho_t\,\mathrm{d}t, \quad R_T = \rho_T. \tag{5.13}$$

Given a solution (R, M) of (5.13), the process Z^ρ is in turn determined by (5.12) via the linear SDE $\mathrm{d}Z_t^\rho = -\frac{Z_t^\rho}{R_t}\,\mathrm{d}M_t^\rho$ with initial condition $Z_0^\rho = M_0^\rho$. Together with the informal arguments in [15, Appendix B.1], this suggests that the Q^ρ-expected remaining risk tolerance R should in fact coincide with the risk tolerance of the *indirect* utility process in the investor's unperturbed problem. The latter is readily obtained for Markovian benchmark problems by differentiating the corresponding value *function*.

We note in passing that to prove existence and uniqueness of a solution to (5.13) (for "sufficiently nice" ρ and B; cf. also Remark 5.7), it is simpler[14] to consider the process R^{-1}, which solves the BSDE

$$\mathrm{d}R_t^{-1} = \mathrm{d}\tilde{M}_t^\rho + \rho_t (R_t^{-1})^2\,\mathrm{d}t, \quad R_T^{-1} = \rho_T^{-1}. \tag{5.14}$$

Here, $\tilde{M}_t^\rho$ solves the SDE $\mathrm{d}\tilde{M}_t^\rho = (R_t^{-1})^2\,\mathrm{d}M_t^\rho = \frac{1}{R_t Z_t^\rho}\,\mathrm{d}Z_t^\rho$ with initial condition $\tilde{M}_0^\rho = M_0^\rho = Z_0^\rho$.

5.2 Uniqueness, existence, and structure of $\Delta\hat{Z}$

In this section we study the uniquess, the existence, and the structure of a $\hat{Q}$-martingale $\Delta\hat{Z}$ satisfying

$$\int_{[0,T]} \frac{\rho_t}{B_t}\Delta\hat{Z}_t\,\mu(\mathrm{d}t) = \Delta Y_T,$$

as required for (5.3). First, we establish uniqueness of $\Delta\hat{Z}$ under quite general conditions:

Lemma 5.2. *Suppose that $E^{\hat{Q}}\left[(\int_{[0,T]} \frac{\rho_t}{B_t}\mu(\mathrm{d}t))^3\right] < \infty$. Then there exists at most one $\hat{Q}$-martingale $(\Delta\hat{Z}_t)_{t\in[0,T]}$ with $E^{\hat{Q}}\left[|\Delta\hat{Z}_T|^3\right] < \infty$ such that*

$$\int_{[0,T]} \frac{\rho_t}{B_t}\Delta\hat{Z}_t\,\mu(\mathrm{d}t) = \Delta Y_T. \tag{5.15}$$

Proof. First, note that if ΔZ is any $\hat{Q}$-martingale satisfying $E^{\hat{Q}}\left[|\Delta Z|^3\right] < \infty$, then Hölder's

[14]This has been kindly pointed out to us by Hao Xing.

inequality and Doob's maximal inequality (with the universal constant $C_3 > 0$) give

$$\begin{aligned}
E^{\hat{Q}}\left[|\Delta Z_T| \int_{[0,T]} \frac{\rho_t}{B_t} |\Delta Z_t| \,\mu(\mathrm{d}t)\right] &\leq E^{\hat{Q}}\left[\Big(\sup_{t\in[0,T]} |\Delta Z_t|\Big)^2 \int_{[0,T]} \frac{\rho_t}{B_t}\,\mu(\mathrm{d}t)\right] \\
&\leq E^{\hat{Q}}\left[\Big(\sup_{t\in[0,T]} |\Delta Z_t|\Big)^3\right]^{\frac{2}{3}} E^{\hat{Q}}\left[\Big(\int_{[0,T]} \frac{\rho_t}{B_t}\,\mu(\mathrm{d}t)\Big)^3\right]^{\frac{1}{3}} \\
&\leq C_3^{\frac{2}{3}} E^{\hat{Q}}\left[|\Delta Z_T|^3\right]^{\frac{2}{3}} E^{\hat{Q}}\left[\Big(\int_{[0,T]} \frac{\rho_t}{B_t}\,\mu(\mathrm{d}t)\Big)^3\right]^{\frac{1}{3}} < \infty.
\end{aligned}$$

Next, assume that there are two $\hat{Q}$-martingales $\Delta Z^{(1)}$ and $\Delta Z^{(2)}$ with $E^{\hat{Q}}\left[|\Delta Z_T^{(i)}|^3\right] < \infty$ satisfying (5.15) with $\Delta\hat{Z}$ replaced by $\Delta Z^{(i)}$, $i \in \{1,2\}$. Set $\tilde{Z} := \Delta Z^{(1)} - \Delta Z^{(2)}$. Then $\tilde{Z}$ is also a $\hat{Q}$-martingale with $E^{\hat{Q}}\left[|\tilde{Z}_T|^3\right] < \infty$ and satisfies

$$\int_{[0,T]} \frac{\rho_t}{B_t} \tilde{Z}_t \,\mu(\mathrm{d}t) = 0. \tag{5.16}$$

Now, multiply both sides of (5.16) by $\tilde{Z}_T$, take $\hat{Q}$-expectations, and use Fubini's theorem, the tower property of conditional expectations and the $\hat{Q}$-martingale property of $\tilde{Z}$. This gives

$$\begin{aligned}
0 = E^{\hat{Q}}\left[\tilde{Z}_T \int_{[0,T]} \frac{\rho_t}{B_t} \tilde{Z}_t \,\mu(\mathrm{d}t)\right] &= \int_{[0,T]} E^{\hat{Q}}\left[\tilde{Z}_T \frac{\rho_t}{B_t} \tilde{Z}_t\right] \mu(\mathrm{d}t) \\
= \int_{[0,T]} E^{\hat{Q}}\left[\frac{\rho_t}{B_t} \tilde{Z}_t E^{\hat{Q}}\left[\tilde{Z}_T \,\middle|\, \mathcal{F}_t\right]\right] \mu(\mathrm{d}t) &= \int_{[0,T]} E^{\hat{Q}}\left[\frac{\rho_t}{B_t} \tilde{Z}_t^2\right] \mu(\mathrm{d}t).
\end{aligned}$$

As a consequence, $E\left[\frac{\rho_t}{B_t}\tilde{Z}_t^2\right] = 0$ for μ-a.e. $t \in [0,T]$. Since $\frac{\rho_t}{B_t}\tilde{Z}_t^2 \geq 0$ P-a.s. for each $t \in [0,T]$, it follows that $\frac{\rho_t}{B_t}\tilde{Z}_t^2 = 0$ P-a.s. for μ-a.e. $t \in [0,T]$. As $\frac{\rho_t}{B_t} > 0$ P-a.s. for each $t \in [0,T]$, this yields $\tilde{Z}_t = 0$ P-a.s. for μ-a.e. $t \in [0,T]$. The $\hat{Q}$-martingale property of $\tilde{Z}$ then gives $\tilde{Z} \equiv 0$ P-a.s. on $[0,t]$ for μ-a.e. $t \in [0,T]$. Finally, Condition (2.1) implies that $\tilde{Z} \equiv 0$ P-a.s. on $[0,T)$; by the martingale convergence theorem, we may therefore conclude that $\tilde{Z} \equiv 0$ P-a.s. on $[0,T]$. Thus, $\Delta Z^{(1)} = \Delta Z^{(2)}$ P-a.s., establishing the claimed uniqueness. □

Next, we establish existence of $\Delta\hat{Z}$ under the key additional assumption that μ has an atom at T. Otherwise, we cannot expect existence to hold, see Example 5.4.

Theorem 5.3. *Suppose that $\mu(\{T\}) > 0$. Moreover, assume that $E^{\hat{Q}}\left[(\sup_{t\in[0,T]} \frac{\rho_t}{B_t})^3\right] < \infty$, $E^{\hat{Q}}\left[(\frac{\rho_T}{B_T})^{-3}\right] < \infty$, and ΔY_T is bounded. Then, there exists a (unique) $\hat{Q}$-martingale $\Delta\hat{Z}$ with $E^{\hat{Q}}\left[|\Delta\hat{Z}_T|^3\right] < \infty$ satisfying*

$$\int_{[0,T]} \frac{\rho_t}{B_t} \Delta\hat{Z}_t \,\mu(\mathrm{d}t) = \Delta Y_T \quad P\text{-a.s.} \tag{5.17}$$

Moreover:

(a) *If $\Delta Y_T = 1$, then $\Delta\hat{Z}$ is nonnegative.*

(b) *If $\frac{\rho_T}{B_T}$ is uniformly bounded from below, then $\Delta\hat{Z}$ is bounded.*

The proof of Theorem 5.3 is rather lengthy and technical, and therefore delegated to Appendix B. Here we just sketch the main ideas. In a first step, we construct a solution to (5.17) in finite discrete time, i.e., on a finite time grid $0 = t_0, \dots, t_N = T$. This is done exactly as outlined in Section 5.1. To wit, we first construct a positive $\hat{Q}$-martingale $Z^{\rho,N}$ satisfying the discretized analogue of (5.6). Next, we construct the "prognosis measure" $Q^{\rho,N} \approx \hat{Q} \approx P$ by $\mathrm{d}Q^{\rho,N} = \frac{Z_T^{\rho,N}}{Z_0^{\rho,N}}\,\mathrm{d}P$ and define $Z^{\Delta Y,N}$ as the discretized analogue of the SDE (5.10). Finally, we set $\Delta Z^N := Z^{\rho,N} Z^{\Delta Y,N}$. In a second step, we make the grid finer and finer and check that the resulting limit ΔZ (in the "Komlós sense") satisfies (5.6).

The following example shows that if the consumption clock μ does not have an atom at T, we cannot expect ΔZ to exist in general, even if the discounted risk-tolerances ρ_t/B_t are deterministic. Indeed, if the quadratic variation of the prognosis $M^\rho = E^{\hat{Q}}\left[\Delta Y_T \,|\, \mathbb{F}\right]$ of the endowment does not vanish quickly enough as the horizon T nears, it may not be possible to satisfy the budget constraint (5.15) with equality. This is because there might be an unexpected move of the endowment close to the time horizon which cannot be "consumed away".[15]

Example 5.4. Assume that $T = 1$, $B \equiv 1$, $\rho \equiv 1$, $\mu(\mathrm{d}t) = \mathbf{1}_{(0,T)}(t)\,\mathrm{d}t$, so that $\hat{Q} = P$ (cf. Lemma 3.2 with $r \equiv 0$ and $U(t,x) = \exp(-x)$). Let $\Delta Y_1 = \exp(-W_1^2)$, where $(W_t)_{t\in[0,1]}$ is Brownian motion. Then there does not exist a P-martingale $(\Delta\hat{Z}_t)_{t\in[0,1]}$ with $E\left[|\Delta\hat{Z}_T|^3\right] < \infty$ satisfying (5.17). Indeed, we trivially have $E\left[(\int_{[0,T]} \frac{\rho_t}{B_t}\mu(\mathrm{d}t))^3\right] < \infty$. Thus, in the notation of Corollary 5.6 below, $R_t = (1-t)$ and $M_t^{\Delta Y} = E\left[\exp(-W_1^2) \,|\, \mathcal{F}_t\right] = \frac{1}{\sqrt{3-2t}} \exp\left(-\frac{W_t^2}{3-2t}\right)$. If there were a P-martingale $\Delta\hat{Z}$ with $E\left[|\Delta\hat{Z}_T|^3\right] < \infty$ satisfying (5.17), then by Corollary 5.6 below, it would satisfy the SDE

$$\mathrm{d}\Delta Z_t = \frac{1}{1-t}\,\mathrm{d}M_t^{\Delta Y} = -\frac{2}{(1-t)(3-2t)^{\frac{3}{2}}} \exp\left(-\frac{W_t^2}{3-2t}\right) \mathrm{d}W_t, \quad \Delta Z_0 = \frac{M_0^{\Delta Y}}{R_0} = \frac{1}{\sqrt{3}}. \tag{5.18}$$

It is an easy exercise to verify that (5.18) does not have a solution on $[0,1]$.[16]

Finally, suppose that the process Z^ρ from Equation (5.6) exists (e.g., under the conditions of Theorem 5.3) and is *positive*. Then, if the process $\Delta\hat{Z}$ exists as well (e.g., under the conditions of Theorem 5.3) it is of the form as derived in the heuristics:

Lemma 5.5. *Suppose that $E^{\hat{Q}}\left[(\int_{[0,T]} \frac{\rho_t}{B_t}\mu(\mathrm{d}t))^3\right] < \infty$ and that there exists a* positive *$\hat{Q}$ martingale Z^ρ satisfying*

$$\int_{[0,T]} \frac{\rho_t}{B_t} Z_t^\rho\,\mu(\mathrm{d}t) = 1. \tag{5.19}$$

Define the measure $Q^\rho \approx \hat{Q} \approx P$ on $\mathcal{F}_T$ by $\mathrm{d}Q^\rho = \frac{Z_T^\rho}{Z_0^\rho}\,\mathrm{d}\hat{Q}$. If the $\hat{Q}$-martingale $\Delta\hat{Z}$ from Lemma 5.2 exists, it is of the form

$$\Delta\hat{Z} = Z^\rho Z^{\Delta Y}, \tag{5.20}$$

[15] A related phenomenon is the "facelift" observed in [26]. For absolutely continuous endowments, however, these issues typically do not arise, cf. [5].

[16] Of course, (5.18) does have a unique (strong) solution on $[0,1)$, but this is not enough for our purposes.

where $Z^{\Delta Y}$ is the Q^ρ-martingale with dynamics

$$\mathrm{d}Z_t^{\Delta Y} = \frac{1}{Z_t^\rho R_t}\,\mathrm{d}M_t^{\Delta Y}, \quad Z_0^{\Delta Y} = M_0^{\Delta Y}. \tag{5.21}$$

Here, $M_t^{\Delta Y} = E^{Q^\rho}[\Delta Y_T \,|\, \mathcal{F}_t]$ is the (càdlàg version) of the Q^ρ-martingale generated by the total endowment, and the (làdlàg) process $R_t = E^{Q^\rho}\left[\int_{[t,T]} \frac{\rho_u}{B_u}\,\mu(\mathrm{d}u) \,\middle|\, \mathcal{F}_t\right]$ measures the Q^ρ-expected remaining (discounted) risk tolerance.

The following corollary covers the important special case of deterministic discounted risk tolerances ρ_t/B_t.[17] It follows immediately from Lemma 5.5 using that $Z^\rho \equiv (\int_{[0,T]} \frac{\rho_t}{B_t}\mu(\mathrm{d}t))^{-1}$ and $R_0 = \frac{1}{Z_0^\rho}$ in this case.

Corollary 5.6. *Suppose that ρ/B is deterministic and $\int_{[0,T]} \frac{\rho_t}{B_t}\mu(\mathrm{d}t) < \infty$. If the $\hat{Q}$-martingale $\Delta\hat{Z}$ from Lemma 5.2 exists, it satisfies the SDE*

$$\mathrm{d}\Delta\hat{Z}_t = \frac{1}{R_t}\,\mathrm{d}M_t^{\Delta Y}, \quad \Delta Z_0 = \frac{M_0^{\Delta Y}}{R_0}. \tag{5.22}$$

Here, $M_t^{\Delta Y} = E^{\hat{Q}}[\Delta Y_T \,|\, \mathcal{F}_t]$ is the (càdlàg version) of the $\hat{Q}$-martingale generated by the total endowment, and the (càglàd) function $R_t = \int_{[t,T]} \frac{\rho_u}{B_u}\,\mu(\mathrm{d}u)$ measures the remaining (discounted) risk tolerance.

Proof of Lemma 5.5. Set $Z_t^{\Delta Y} = \frac{\Delta\hat{Z}}{Z^\rho}$. Then $Z^{\Delta Y}$ is a Q^ρ-martingale by Bayes' theorem. Define the process $(R_t)_{t\in[0,T]}$ by

$$R_t = \frac{1}{Z_t^\rho}\left(1 - \int_{[0,t)} Z_u^\rho \frac{\rho_u}{B_u}\,\mu(\mathrm{d}u)\right).$$

Then R is positive on $[0,T)$ by Assumption (2.1) on μ. (It is positive on $[0,T]$ if and only if $\mu(\{T\}) > 0$.) Note that R has làdlàg paths and ZR is nonincreasing and left-continuous; in particular, it is predictable. By (a conditional extension of) Corollary A.2(b),

$$R_t = E^{Q^\rho}\left[\int_{[t,T]} \frac{\rho_u}{B_u}\,\mu(\mathrm{d}u) \,\middle|\, \mathcal{F}_t\right].$$

Arguing as in (5.9), it follows that

$$\Delta Y_T = Z^{\Delta Y_0} + \int_{(0,T]} Z_t^\rho R_T\,\mathrm{d}Z_t^{\Delta Y}. \tag{5.23}$$

As $Z^\rho R$ is nonincreasing from 1 to $Z_T^\rho R_T \geq 0$ and $E^{Q^\rho}\left[|Z_T^{\Delta Y}|\right] < \infty$ because $Z^{\Delta Y}$ is a Q^ρ-martingale, it follows from Lemma A.1 that $\int Z^\rho R\,\mathrm{d}Z^{\Delta Y}$ is a (true) Q^ρ-martingale. So in particular the right-hand side of (5.23) is Q^ρ-integrable. Then so is the left-hand side ΔY_T, which shows that the Q^ρ-martingale $(M_t^{\Delta Y})_{t\in[0,T]}$ given by $M_t^{\Delta Y} = E^{Q^\rho}[\Delta Y_T \,|\, \mathcal{F}_t]$ is well defined. Now, taking conditional Q^ρ-expectations in (5.23) shows that $M^{\Delta Y}$ satisfies

$$\mathrm{d}M_t^{\Delta Y} = Z_t^\rho R_T\,\mathrm{d}Z_t^{\Delta Y}, \quad M_0^{\Delta Y} = Z_0^{\Delta Y}. \tag{5.24}$$

[17]In essence, this is the case of exponential utility and deterministic interest rates.

Note that (5.24) implies in particular that $M^{\Delta Y}_T - M^{\Delta Y}_{T-} = 0$ if $R_T = 0$. Using again that $Z^\rho R$ is nonincreasing and strictly positive on $[0,T)$, we may deduce that $Z^{\Delta Y}$ satisfies (5.21). $\square$

Remark 5.7. Even if Z^ρ exists, it need not be positive (but only nonnegative) in general, in which case the decomposition (5.20) fails. However, if ρ and B are "sufficiently nice", Z^ρ is indeed positive. For example, suppose the filtration is continuous and ρ/B is uniformly bounded from above and away from zero. Then, as has been kindly pointed out to us by Hao Xing, standard BSDE arguments show that the BSDE (5.14) has a unique solution $(R^{-1}, \tilde{M}^\rho)$, where R^{-1} is uniformly bounded from above and away from zero, and $\tilde{M}^\rho$ is a BMO $\hat{Q}$-martingale. $R := 1/R^{-1}$ is in turn well defined and uniformly bounded, whence $\int R\,\mathrm{d}\tilde{M}^\rho$ is also a BMO $\hat{Q}$-martingale and $Z^\rho := \mathcal{E}\left(\int R\,\mathrm{d}\tilde{M}^\rho\right)$ is a positive (true) $\hat{Q}$-martingale (cf. [20]). It is now easy to check that Z^ρ solves (5.19).

5.3 The main result

After the preparations of the previous section, we can now formulate our main result about the second-order sensitivities of the consumption-savings problem:

Theorem 5.8. *Suppose that $\hat{c} \in \mathcal{A}^*_3(Y)$, $E\left[|\hat{Z}^3_T|\right] < \infty$, $E^{\hat{Q}}\left[\int_{[0,T]}(\frac{\rho_t}{B_t})^3\mu(\mathrm{d}t)\right] < \infty$, as well as $E^{\hat{Q}}\left[\Delta Y_T^2\right] < \infty$. Assume that U satisfies DARA and DAP, and suppose that there exists a $\hat{Q}$-martingale $\Delta\hat{Z}$ with $E^{\hat{Q}}\left[|\Delta Z_T|^3\right] < \infty$ satisfying*

$$\int_{[0,T]} \frac{\rho_t}{B_t}\Delta\hat{Z}_t\,\mu(\mathrm{d}t) = \Delta Y_T \quad P\text{-a.s.} \tag{5.25}$$

Moreover, assume that $\Delta\hat{Z}$ is locally bounded, there is $\varepsilon_0 > 0$ such that $\hat{c} - \varepsilon_0\rho_t|\Delta\hat{Z}_t| > x_U$ P-a.s., and we have the estimates

$$E\left[\left(\int_{[0,T]}\frac{\rho_t}{B_t}|\Delta\hat{Z}_t|\,\mu(\mathrm{d}t)\right)^3\right] < \infty, \tag{5.26}$$

$$E\left[\int_{[0,T]} U'''_t(\hat{c}_t - \varepsilon_0\rho_t|\Delta\hat{Z}_t|)\rho_t^3|\Delta\hat{Z}_t|^3\,\mu(\mathrm{d}t)\right] < \infty, \tag{5.27}$$

as well as

$$E\left[\left(\sup_{t\in[0,T]}|\Delta\hat{Z}_t|\right)^3\int_{[0,T]} -\tilde{U}'''_t\left(\frac{\hat{Z}_t}{2B_t}\right)\frac{\hat{Z}^3_t}{B^3_t}\,\mu(\mathrm{d}t)\right] < \infty. \tag{5.28}$$

Then, the process

$$\hat{c}^\varepsilon = \hat{c} + \varepsilon\rho\Delta\hat{Z}, \quad \varepsilon \in [0,\varepsilon_0)$$

is second-order optimal for $\mathcal{A}_3(Y^\varepsilon)$. It satisfies

$$E\left[\int_{[0,T]} U_t(\hat{c}^\varepsilon_t)\,\mu(\mathrm{d}t)\right] = E\left[\int_{[0,T]} U_t(\hat{c}_t)\,\mu(\mathrm{d}t)\right] + \varepsilon \hat{Z}_0 E^{\hat{Q}}\left[\Delta Y_T\right] - \frac{1}{2}\varepsilon^2 \hat{Z}_0 E^{\hat{Q}}\left[\Delta \hat{Z}_T \Delta Y_T\right] + o(\varepsilon^2). \tag{5.29}$$

Moreover, if there exists a positive *$\hat{Q}$-martingale Z^ρ with $\int_{[0,T]} \frac{\rho_t}{B_t} Z^\rho_t\,\mu(\mathrm{d}t) = 1$, then*

$$\hat{c}^\varepsilon = \hat{c} + \varepsilon \rho Z^\rho Z^{\Delta Y}, \quad \varepsilon \in [0, \varepsilon_0),$$

where $Z^{\Delta Y}$ is as in Lemma 5.5, and the welfare expansion can be written as

$$\varepsilon \hat{Z}_0 E^{\hat{Q}}\left[\Delta Y_T\right] - \frac{1}{2}\varepsilon^2 \hat{Z}_0 E^{\hat{Q}}\left[\frac{(M^{\Delta Y}_0)^2}{R_0} + \int_{(0,T]} \frac{1}{R_t}\,\mathrm{d}\left[M^{\Delta Y}\right]_t\right] + o(\varepsilon^2),$$

with $(M^{\Delta Y}_t)_{t\in[0,T]}$ and $(R_t)_{t\in[0,T]}$ as in Lemma 5.5.

Theorem 5.8 states that – under the stated integrability assumptions[18] – the (normalized) second-order consumption correction $\Delta\hat{c} = (\hat{c}^\varepsilon - \hat{c})/\varepsilon$ is the product of the direct risk-tolerance ρ, the martingale Z^ρ, and the prognosis martingale $Z^{\Delta Y}$. The first-order welfare correction $\varepsilon Z_0 E^{\hat{Q}}\left[\Delta Y_T\right]$, is the marginal-utility-based price of the terminal perturbation $\varepsilon \Delta Y_T$. The second-order welfare correction $-\frac{1}{2}\varepsilon^2 \hat{Z}_0 E^{\hat{Q}}\left[\frac{(M^{\Delta Y}_0)^2}{R_0} + \int_{(0,T]} \frac{1}{R_t}\,\mathrm{d}\left[M^{\Delta Y}\right]_t\right]$ is the expected fluctuation of the terminal perturbation $\varepsilon\Delta Y_T$ weighted by the indirect risk-tolerance R. In the special case that ΔY_T is deterministic, it simplifies to $-\frac{1}{2}\varepsilon^2 \hat{Z}_0 \frac{1}{R_0}(\Delta Y_T)^2$.

A sufficient set of conditions for the first part of Theorem 5.8 is that μ has an atom at T, the endowment correction ΔY is bounded (the unperturbed endowment Y_T may be general), and ρ and B are uniformly bounded from above and from zero.[19] Then ΔZ exists and is bounded by Theorem 5.3, and for standard utility functions (5.26) – (5.28) are easily seen to be satisfied.

Proof of Theorem 5.8. It follows from (5.26) that $\rho\Delta\hat{Z} \in \mathcal{A}^*_3(\Delta Y)$ and $\hat{c}^\varepsilon \in \mathcal{A}^*_3(Y^\varepsilon)$.

The basic idea for the remaining proof is similar to the one of Theorem 4.1: a lower bound is established by considering our concrete strategy; a universal upper bound is in turn obtained from a duality argument.

Primal lower bound. For $\varepsilon \in [0, \varepsilon_0)$ and fixed t and ω, a Taylor expansion of order two with Lagrange remainder term gives

$$U_t(\hat{c}^\varepsilon_t) = U_t(\hat{c}_t) + U'_t(\hat{c}_t)(\hat{c}^\varepsilon_t - \hat{c}_t) + \frac{1}{2}U''_t(\hat{c}_t)(\hat{c}^\varepsilon_t - \hat{c}_t)^2 + \frac{1}{6}U'''_t(\tilde{c}(t,\omega))(\hat{c}^\varepsilon_t - \hat{c}_t)^3,$$

where $\tilde{c}(t,\omega)$ takes values in the interval with endpoints $\hat{c}_t(\omega)$ and $\hat{c}^\varepsilon_t(\omega)$. By the definition of $\hat{c}^\varepsilon$,

[18] The third-order moments are needed to control the third-order remainder terms appearing in the Taylor expansion that we use to approximate general utilities by quadratic ones in the proof of Theorem 5.8.

[19] For utilities on the positive real line, we also have to assume that the benchmark optimizer $\hat{c}$ is uniformly bounded away from 0.

the fact that U''' is positive and nonincreasing by DARA and DAP, and since $\varepsilon < \varepsilon_0$, we have

$$U_t(\hat{c}_t^\varepsilon) \geq U_t(\hat{c}_t) + \varepsilon U_t'(\hat{c}_t)\rho_t \Delta\hat{Z}_t + \frac{1}{2}\varepsilon^2 U_t''(\hat{c}_t)\hat{\rho}_t^2\Delta\hat{Z}_t^2 - \frac{1}{6}\varepsilon^3 U_t'''(\hat{c}_t - \varepsilon_0\rho_t|\Delta\hat{Z}_t|)\rho_t^3|\Delta\hat{Z}_t|^3. \quad (5.30)$$

Now, take the expectation of the integral of the first-order term on the right-hand side of (5.30). Then, it follows as in the proof of Theorem 4.1 that

$$\varepsilon E\left[\int_{[0,T]} U_t'(\hat{c}_t)\rho_t\Delta\hat{Z}_t\,\mu(\mathrm{d}t)\right] = \varepsilon\hat{Z}_0 E^{\hat{Q}}\left[\Delta Y_T\right]. \quad (5.31)$$

Next, take the expectation of the integral of the second-order term on the right hand side of (5.30). Using the definition of ρ in (5.1), the first-order condition (3.1), Corollary A.2(b), Lemma A.1 and the budget constraint (5.25), we obtain

$$\begin{aligned}
\frac{1}{2}\varepsilon^2 E\left[\int_{[0,T]} U_t''(\hat{c}_t)\rho_t^2\Delta\hat{Z}_t^2\,\mu(\mathrm{d}t)\right] &= -\frac{1}{2}\varepsilon^2 E\left[\int_{[0,T]} \hat{Z}_t\Delta\hat{Z}_t\frac{\rho_t}{B_t}\Delta\hat{Z}_t\,\mu(\mathrm{d}t)\right] \\
&= -\frac{1}{2}\varepsilon^2\hat{Z}_0 E^{\hat{Q}}\left[\int_{[0,T]} \Delta\hat{Z}_t\frac{\rho}{B_t}\Delta\hat{Z}_t\,\mu(\mathrm{d}t)\right] \\
&= -\frac{1}{2}\varepsilon^2\hat{Z}_0 E^{\hat{Q}}\left[\Delta\hat{Z}_T\Delta Y_T\right]. \quad (5.32)
\end{aligned}$$

Finally, taking the expectation of the integral on both sides of (5.30) and using (5.31), (5.32), and (5.27) yields (5.29).

Dual upper bound. We proceed to show that

$$\begin{aligned}
\sup_{c^\varepsilon\in\mathcal{A}_3(Y^\varepsilon)} E\left[\int_{[0,T]} U_t(c_t^\varepsilon)\,\mu(\mathrm{d}t)\right] \leq E\left[\int_{[0,T]} U_t(\hat{c}_t)\,\mu(\mathrm{d}t)\right] + \varepsilon\hat{Z}_0 E^{\hat{Q}}\left[\Delta Y_T\right] \\
- \frac{1}{2}\varepsilon^2\hat{Z}_0 E^{\hat{Q}}\left[\Delta\hat{Z}_T\Delta Y_T\right] + o(\varepsilon^2). \quad (5.33)
\end{aligned}$$

As $\Delta\hat{Z}$ is locally bounded by assumption, there exists a nondecreasing sequence of $\mathbb{F}$-stopping times $(\tau_n)_{n\in\mathbb{N}}$ with values in $[0,T]$ such that $\lim_{n\to\infty} P[\tau_n = T] = 1$ and, for each $n\in\mathbb{N}$, the stopped martingale $\Delta\hat{Z}^{\tau_n}$ is bounded. Set $\tau_0 := 0$. For $\varepsilon\in(0,1/(2\Delta\hat{Z}_0))$, let

$$n(\varepsilon) = \max\left\{n\in\{0,\dots,\lfloor 1/\varepsilon\rfloor\} : \sup_{t\in[0,T]}|\Delta\hat{Z}_t^{\tau_n}| \leq \frac{1}{2\varepsilon}\ P\text{-a.s.}\right\}.$$

Define $\Delta\hat{Z}^\varepsilon := \Delta\hat{Z}^{\tau_{n(\varepsilon)}}$. Then $\hat{\sup}_{t\in[0,T]}|\Delta\hat{Z}_t^\varepsilon| \leq \frac{1}{2\varepsilon}$ P-a.s. for each $\varepsilon\in(0,1/(2|\Delta\hat{Z}_0|))$ and $\lim_{\varepsilon\searrow 0}\Delta\hat{Z}^\varepsilon = \Delta\hat{Z}$ P-a.s. as $\lim_{\varepsilon\searrow 0} n(\varepsilon) = +\infty$. Set $\hat{Z}^\varepsilon := \hat{Z}(1-\varepsilon\Delta\hat{Z}^\varepsilon)$ and note that $\hat{Z}^\varepsilon \geq \frac{1}{2}\hat{Z}$ P-a.s.

Fix $\varepsilon\in(0,1/(2\Delta\hat{Z}_0))$ and let $c^\varepsilon\in\mathcal{A}_3(Y^\varepsilon)$ with $E\left[\int_{[0,T]} U_t^-(c_t^\varepsilon)\,\mu(\mathrm{d}t)\right] < \infty$. By definition of the conjugate $\tilde{U}$, we have $U_t(c_t^\varepsilon) \leq \tilde{U}_t\left(\frac{\hat{Z}_t^\varepsilon}{B_t}\right) + \frac{\hat{Z}_t^\varepsilon}{B_t}c_t^\varepsilon$ for each fixed t and ω. Whence, a Taylor expansion of order two with Lagrange remainder term gives

$$\begin{aligned}\tilde{U}_t\left(\frac{\hat{Z}_t^\varepsilon}{B_t}\right) = \tilde{U}_t\left(\frac{\hat{Z}_t}{B_t}\right) - \varepsilon\tilde{U}_t'\left(\frac{\hat{Z}_t}{B_t}\right)\frac{\hat{Z}_t}{B_t}\Delta\hat{Z}_t^\varepsilon + \frac{1}{2}\varepsilon^2\tilde{U}_t''\left(\frac{\hat{Z}_t}{B_t}\right)\frac{\hat{Z}_t^2}{B_t^2}(\Delta\hat{Z}_t^\varepsilon)^2 \\ - \frac{1}{6}\varepsilon^3\tilde{U}_t'''(\zeta(t,\omega))\frac{\hat{Z}_t^3}{B_t^3}(\Delta\hat{Z}_t^\varepsilon)^3,\end{aligned}$$

where $\zeta(t,\omega)$ lies in the interval with endpoints $\frac{\hat{Z}_t(\omega)}{B_t(\omega)}$ and $\frac{\hat{Z}_t^\varepsilon(\omega)}{B_t(\omega)}$. Now, use that $\tilde{U}_t'\left(\frac{\hat{Z}_t}{B_t}\right) = -\hat{c}_t$ (which follows from (3.1) and (2.6)), $\tilde{U}_t''\left(\frac{\hat{Z}_t}{B_t}\right) = -\frac{1}{U_t''(\hat{c}_t)} = \rho_t\frac{B_t}{\hat{Z}_t}$ (which follows from (3.1) and (2.7)), the fact that $\tilde{U}'''$ is negative and nondecreasing by DARA and DAP, and $\hat{Z}_t^\varepsilon \geq \frac{\hat{Z}_t}{2}$. This yields the following estimate:

$$\tilde{U}_t\left(\frac{\hat{Z}_t^\varepsilon}{B_t}\right) \leq \tilde{U}_t\left(\frac{\hat{Z}_t}{B_t}\right) + \varepsilon\hat{Z}_t\Delta\hat{Z}_t^\varepsilon\frac{\hat{c}_t}{B_t} + \frac{1}{2}\varepsilon^2\hat{Z}_t\frac{\rho_t}{B_t}(\Delta\hat{Z}_t^\varepsilon)^2 - \frac{1}{6}\varepsilon^3\tilde{U}_t'''\left(\frac{\hat{Z}_t}{2B_t}\right)\frac{\hat{Z}_t^3}{B_t^3}|\Delta\hat{Z}_t^\varepsilon|^3.$$

Combining this with the identity $\tilde{U}_t(\frac{\hat{Z}_t}{B_t}) = U_t(\hat{c}_t) - \frac{\hat{Z}_t}{B_t}\hat{c}_t$ (which follows from (3.1)) and (4.5):

$$U_t(c_t^\varepsilon) \leq U_t(\hat{c}_t) + \frac{\hat{Z}_t^\varepsilon}{B_t}(c_t^\varepsilon - \hat{c}_t) + \frac{1}{2}\varepsilon^2\hat{Z}_t\frac{\rho_t}{B_t}(\Delta\hat{Z}_t^\varepsilon)^2 - \frac{1}{6}\varepsilon^3\tilde{U}_t'''\left(\frac{\hat{Z}_t}{2B_t}\right)\frac{\hat{Z}_t^3}{B_t^3}|\Delta\hat{Z}_t^\varepsilon|^3. \tag{5.34}$$

Now, consider separately each of the last three terms on the right-hand side of (5.34). First, calculate the expectation of the integral of the second one. Corollary A.2(a), positivity of $\hat{Z}^\varepsilon$, the budget constraints for c^ε and $\hat{c}$, the definition of $\hat{Z}^\varepsilon$, Bayes' theorem and dominated convergence give

$$\begin{aligned}E\left[\int_{[0,T]}\hat{Z}_t^\varepsilon\frac{c_t^\varepsilon - \hat{c}_t}{B_t}\,\mu(\mathrm{d}t)\right] &= E\left[\hat{Z}_T^\varepsilon\int_{[0,T]}\frac{c_t^\varepsilon - \hat{c}_t}{B_t}\,\mu(\mathrm{d}t)\right] \leq E\left[\hat{Z}_T^\varepsilon\varepsilon\Delta Y_T\right] \\ &= \varepsilon\hat{Z}_0E^{\hat{Q}}\left[(1-\varepsilon\Delta\hat{Z}^\varepsilon)\Delta Y_T\right] = \varepsilon\hat{Z}_0E^{\hat{Q}}\left[\Delta Y_T\right] - \varepsilon^2\hat{Z}_0E^{\hat{Q}}\left[\Delta\hat{Z}_T^\varepsilon\Delta Y_T\right] \\ &= \varepsilon\hat{Z}_0E^{\hat{Q}}\left[\Delta Y_T\right] - \varepsilon^2\hat{Z}_0E^{\hat{Q}}\left[\Delta\hat{Z}_T\Delta Y_T\right] + o(\varepsilon^2).\end{aligned}$$

Next, calculate the expectation of the integral of the third term on the right-hand side of (5.34). Corollary A.2(b), dominated convergence, (5.25), and Lemma A.1 give

$$\begin{aligned}\frac{1}{2}\varepsilon^2E\left[\int_{[0,T]}\hat{Z}_t\frac{\rho_t}{B_t}(\Delta\hat{Z}_t^\varepsilon)^2\,\mu(\mathrm{d}t)\right] &= \frac{1}{2}\varepsilon^2\hat{Z}_0E^{\hat{Q}}\left[\int_{[0,T]}\frac{\rho_t}{B_t}(\Delta\hat{Z}_t^\varepsilon)^2\,\mu(\mathrm{d}t)\right] \\ &= \frac{1}{2}\varepsilon^2\hat{Z}_0E^{\hat{Q}}\left[\int_{[0,T]}\frac{\rho_t}{B_t}\Delta\hat{Z}_t^2\,\mu(\mathrm{d}t)\right] + o(\varepsilon^2) \\ &= \frac{1}{2}\varepsilon^2E^{\hat{Q}}\left[\int_{[0,T]}\Delta\hat{Z}_t\frac{\rho_t}{B_t}\Delta\hat{Z}_t\,\mu(\mathrm{d}t)\right] + o(\varepsilon^2) \\ &= \frac{1}{2}\varepsilon^2E^{\hat{Q}}\left[\Delta\hat{Z}_T\Delta Y_T\right] + o(\varepsilon^2).\end{aligned}$$

Finally, compute the expectation of the integral of the last term on the right-hand side of (5.34).

By (5.28), we have

$$\begin{aligned}
&\frac{1}{6}\varepsilon^3 E\left[\int_{[0,T]} -\tilde{U}_t'''\left(\frac{\hat{Z}_t}{2}\right)\frac{\hat{Z}_t^3}{B_t^3}|\Delta\hat{Z}_t^\varepsilon|^3\,\mu(\mathrm{d}t)\right] \\
&\qquad\le \frac{1}{6}\varepsilon^3 E\left[\left(\sup_{t\in[0,T]}|\Delta\hat{Z}_t^\varepsilon|\right)^3\int_{[0,T]} -\tilde{U}_t'''\left(\frac{\hat{Z}_t}{2}\right)\frac{\hat{Z}_t^3}{B_t^3}\,\mu(\mathrm{d}t)\right] \\
&\qquad\le \frac{1}{6}\varepsilon^3 E\left[\left(\sup_{t\in[0,T]}|\Delta\hat{Z}_t|\right)^3\int_{[0,T]} -\tilde{U}_t'''\left(\frac{\hat{Z}_t}{2}\right)\frac{\hat{Z}_t^3}{B_t^3}\,\mu(\mathrm{d}t)\right] = O(\varepsilon^3) = o(\varepsilon^2).
\end{aligned}$$

Now taking the expectation of the integral on both sides of (5.34) and then the supremum over all $c^\varepsilon \in \mathcal{A}_3(Y^\varepsilon)$ with $E\left[\int_{[0,T]} U_t^-(c_t^\varepsilon)\,\mu(\mathrm{d}t)\right] < \infty$ gives (5.33).

Comparison of lower and upper bound. (5.29) and (5.33) show that $\hat{c}^\varepsilon$ is indeed second-order optimal for $\mathcal{A}_3(Y^\varepsilon)$.

Additional Claim. Finally suppose that the martingale Z^ρ from Lemma 5.2 exists and is positive. Define $Q^\rho \approx \hat{Q}$, $(Z_t^{\Delta Y})_{t\in[0,T]}$, $(M_t^{\Delta Y})_{t\in[0,T]}$, and $(R_t)_{t\in[0,T]}$ as in Lemma 5.5. Then it suffices to show that

$$E^{\hat{Q}}\left[\Delta\hat{Z}_T\Delta Y_T\right] = E^{\hat{Q}}\left[\int_{[0,T]}\frac{1}{R_t}\,\mathrm{d}\left[M^{\Delta Y}\right]_t\right].$$

Bayes' theorem, the dynamics of $Z^{\Delta Y}$ (cf. (5.21)), the product formula, and Lemma A.1 yield

$$\begin{aligned}
E^{\hat{Q}}\left[\Delta\hat{Z}_T\Delta Y_T\right] &= Z_0^\rho E^{Q^\rho}\left[Z_T^{\Delta Y}M^{\Delta Y_t}\right] = Z_0^\rho E^{Q^\rho}\left[\int_{[0,T]}\frac{1}{R_tZ_t^\rho}\,\mathrm{d}\left[M^{\Delta Y}\right]_t\right] \\
&= E^{\hat{Q}}\left[Z_T^\rho\int_{[0,T]}\frac{1}{R_tZ_t^\rho}\,\mathrm{d}\left[M^{\Delta Y}\right]_t\right] = E^{\hat{Q}}\left[\int_{[0,T]}\frac{Z_t^\rho}{R_tZ_t^\rho}\,\mathrm{d}\left[M^{\Delta Y}\right]_t\right] \\
&= E^{\hat{Q}}\left[\int_{[0,T]}\frac{1}{R_t}\,\mathrm{d}\left[M^{\Delta Y}\right]_t\right].
\end{aligned}$$

This establishes the last assertion and thereby completes the proof of the theorem. □

Remark 5.9. The dual considerations in the proof of Theorem 4.1 also show that the martingale Z^ε is second-order optimal for the minimization problem dual to (2.3).

A On Bayes' theorem

In this appendix, we recall a simple – but apparently not so well-known – sufficient condition for the stochastic integral of a finite variation integrand with respect to a martingale integrator to be a true martingale. Moreover, we deduce two versions of Bayes' theorem which are used in several of the proofs.

We denote by $\|A\|$ the total variation of a finite variation process A, and use the convention that $A_{0-} := 0$.

Lemma A.1. *Consider an adapted càdlàg process of finite variation* $(A_t)_{t\in[0,T]}$ *and a martingale* $(M_t)_{t\in[0,T]}$ *such that*

$$E\left[\|A\|_T|M_T|\right] < \infty.$$

Then, the stochastic integral $\int A_- \, \mathrm{d}M$ *is a martingale and, for any stopping time* τ *taking values in* $[0,T]$*, we have*

$$E\left[\int_{[0,\tau]} M_t \, \mathrm{d}A_t\right] = E\left[M_\tau A_\tau\right]. \tag{A.1}$$

Proof. Let τ be an arbitrary stopping time with values in $[0,T]$. Then, by the submartingale property of $|M|$ and the tower property of conditional expectations:

$$E\left[\|A\|_\tau|M_\tau|\right] \le E\left[\|A\|_\tau E\left[|M_T| \,|\, \mathcal{F}_\tau\right]\right] = E\left[\|A\|_\tau|M_T|\right] \le E\left[\|A\|_T|M_T|\right] < \infty.$$

Hence, by (the optional version of) [7, Theorem VI.57 and Remark VI.58d)] and using that the optional projection of M_τ is the stopped martingale M^τ, we obtain

$$E\left[M_\tau A_\tau\right] = E\left[\int_{[0,\tau]} M_\tau \, \mathrm{d}A_t\right] = E\left[\int_{[0,\tau]} M_t^\tau \, \mathrm{d}A_t\right] = E\left[M_0A_0 + \int_{(0,\tau]} M_t \, \mathrm{d}A_t\right]. \tag{A.2}$$

This gives (A.1). Moreover, an integration by parts yields

$$E\left[M_0A_0 + \int_{(0,\tau]} M_t \, \mathrm{d}A_t\right] = E\left[M_\tau A_\tau - \int_{(0,\tau]} A_{t-} \, \mathrm{d}M_t\right]. \tag{A.3}$$

Together, (A.2) and (A.3) show

$$E\left[\int_{(0,\tau]} A_{t-} \, \mathrm{d}M_t\right] = 0.$$

As τ was arbitrary, it follows that the process $\int A_- \, \mathrm{d}M$ has constant expectation 0 over stopping times and therefore is a martingale. □

Corollary A.2. *Let* $(Z_t)_{t\in[0,T]}$ *be a positive martingale and* $Q \approx P$ *defined by* $\mathrm{d}Q = \frac{Z_T}{Z_0}\mathrm{d}P$ *be the corresponding equivalent measure. Let* $(A_t)_{t\in[0,T]}$ *be an adapted càdlàg process of finite variation, and assume that at least one of the following conditions is satisfied:*

(a) $\|A\|_T$ *and* Z_T *are square-integrable under* P.

(b) $E\left[\int_{[0,T]} Z_t \, \mathrm{d}\|A\|_t\right] < \infty$.

Then the process $(\int_{[0,t]} Z_s \, \mathrm{d}A_s)_{t\in[0,T]}$ *is* P*-integrable, and*

$$E\left[\int_{[0,T]} Z_s \, \mathrm{d}A_s\right] = E\left[Z_TA_T\right] = Z_0E^Q\left[A_T\right].$$

Proof. By Lemma A.1, it suffices to show that $E\left[\|A\|_TZ_T\right] < \infty$. For (a), this follows from the Cauchy–Schwarz inequality. For (b), this is a consequence of (the optional version of) [7, Theorem VI.57]. □

B Proof of Theorem 5.3

This appendix contains the proof of Theorem 5.3, which is broken up into several auxiliary results for better readability. Throughout, notation is eased by replacing $\hat{Q}$ and ρ/B in Theorem 5.3 with P and ρ, respectively.

The first step is to establish existence of the martingale Z^ρ in finite discrete time, by means of an explicit backward construction inspired by the heuristics from Section 5.1:

Lemma B.1. *Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_k)_{k\in\{0,\dots,N\}}, P)$ be a filtered probability space, and assume that $\mathcal{F}_0$ is P-trivial. Let $(\rho_k)_{k\in\{0,\dots,N\}}$ be an nonnegative adapted process such that $E\left[\rho_k^3\right] < \infty$ for $k \in \{0,\dots,N\}$, $\rho_N > 0$ P-a.s., and $E\left[\rho_N^{-3}\right] < \infty$. Then, there exists a unique martingale $(Z_k^\rho)_{k\in\{0,\dots,N\}}$ with $E\left[|Z_N^\rho|^3\right] < \infty$ such that*

$$\sum_{k=0}^{N} \rho_k Z_k^\rho = 1 \;\; P\text{-a.s.} \tag{B.1}$$

Moreover, Z^ρ is positive and satisfies

$$Z_N^\rho \le \rho_N^{-1} \;\; P\text{-a.s.} \tag{B.2}$$

Proof. Uniqueness of Z^ρ follows by the same argument as in Lemma 5.2.

To establish existence, define the process $(R_k)_{k\in\{0,\dots,N\}}$ recursively[20] by $R_N := \rho_N$ and

$$R_{k-1} := \rho_{k-1} + E\left[R_k^{-1} \,\middle|\, \mathcal{F}_{k-1}\right]^{-1}, \quad k \in \{1,\dots,N\}. \tag{B.3}$$

This process is well defined and satisfies $R_k > 0$ P-a.s. as well as $E\left[R_k^{-1}\right] < \infty$ for $k \in \{1,\dots,N\}$. Indeed, for $k = N$, this follows from the assumptions that $\rho_N > 0$ P-a.s. and $E\left[\rho_N^{-1}\right] < \infty$. By backward induction, if we know that $R_k > 0$ and $E\left[R_k^{-1}\right] < \infty$ for some $k \in \{1,\dots,N\}$, then R_{k-1} is finite and satisfies $R_{k-1} \ge E\left[R_k^{-1} \,\middle|\, \mathcal{F}_{k-1}\right]^{-1} > 0$. Rearranging, taking expectations and using the tower property in turn yields

$$E\left[R_{k-1}^{-1}\right] \le E\left[E\left[R_k^{-1} \,\middle|\, \mathcal{F}_{k-1}\right]\right] = E\left[R_k^{-1}\right] < \infty.$$

Now, define the positive martingale $(Z_k^\rho)_{k\in\{1,\dots,N\}}$ recursively[21] by $Z_0^\rho := (R_0)^{-1}$ and

$$Z_{k+1}^\rho := Z_k^\rho \frac{R_{k+1}^{-1}}{E\left[R_{k+1}^{-1} \,\middle|\, \mathcal{F}_k\right]}, \quad k \in \{0,\dots N-1\}. \tag{B.4}$$

This process is well defined by the properties of R. We proceed to show by backward induction

[20] This definition is motivated by the BSDE (5.14) for R^{-1}. If we discretize the latter – in a slightly nonstandard way – as $\Delta R_k^{-1} = \rho_{k-1} E\left[R_{k-1}^{-1} R_k^{-1} \,\middle|\, \mathcal{F}_{k-1}\right] + \Delta\tilde{M}_k^\rho$, take conditional $\mathcal{F}_{k-1}$-expectations yielding $E\left[R_k^{-1} \,\middle|\, \mathcal{F}_{k-1}\right] - R_{k-1}^{-1} = \rho_{k-1} R_{k-1}^{-1} E\left[R_k^{-1} \,\middle|\, \mathcal{F}_{k-1}\right]$, and solve for R_{k-1}, we get (B.3).

[21] This is again motivated by the BSDE (5.14) for R^{-1} and the SDE $\mathrm{d}\tilde{M}_t^\rho = \frac{R^{-1}}{Z_t^\rho}\,\mathrm{d}\tilde{Z}_t^\rho$. Discretizing the latter as $\Delta\tilde{M}_{k+1}^\rho = \frac{E\left[R_{k+1}^{-1} \,\middle|\, \mathcal{F}_k\right]}{Z_k^\rho} \Delta Z_{k+1}^\rho$, using $\Delta\tilde{M}_{k+1}^\rho = R_{k+1}^{-1} - E\left[R_{k+1}^{-1} \,\middle|\, \mathcal{F}_k\right]$, and solving for Z_k^ρ, we get (B.4).

that

$$Z_n^\rho R_n = \sum_{k=n}^{N} Z_k^\rho \rho_k, \quad n \in \{0, \dots, N\}. \tag{B.5}$$

The induction basis $n = N$ is trivial. So assume that (B.5) holds for all $k \in \{n, \dots, N\}$ for some $n \in \{1, \dots, N\}$. Then by the induction hypothesis, (B.4), and (B.3):

$$\sum_{k=n-1}^{N} Z_k^\rho \rho_k = Z_{n-1}^\rho \rho_{n-1} + Z_n^\rho R_n = Z_{n-1}^\rho \rho_{n-1} + Z_{n-1}^\rho E\left[R_n^{-1} \,\middle|\, \mathcal{F}_{n-1}\right]^{-1} = Z_{n-1}^\rho R_{n-1}.$$

Now (B.5) for $n = 0$ yields

$$\sum_{k=0}^{N} \rho_k Z_k^\rho = R_0 Z_0^\rho = 1. \tag{B.6}$$

in accordance with (B.1). The estimate (B.2) follows immediately from the fact that $0 < Z_N^\rho \rho_N < 1$ P-a.s. by nonnegativity of ρ, positivity of Z^ρ and ρ_N, and (B.1). As $E\left[\rho_N^{-3}\right] < \infty$, (B.2) immediately implies that Z^ρ has finite third moments. □

Next, we establish existence of $\Delta \hat{Z}$ in finite discrete time, again by working with the discrete analogues of the relations in Section 5.1.

Lemma B.2. *Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_k)_{k \in \{0,\dots,N\}}, P)$ be a filtered probability space, where $\mathcal{F}_0$ is P-trivial. Let ΔY_N be a bounded $\mathcal{F}_N$-measurable random variable and let $(\rho_k)_{k \in \{0,\dots,N\}}$ be a non-negative adapted process. Assume that $E\left[\rho_k^3\right] < \infty$ for $k \in \{0, \dots, N\}$, $\rho_N > 0$ P-a.s., and $E\left[\rho_N^{-3}\right] < \infty$. Then, there exists a unique martingale $(\Delta \hat{Z}_k)_{k \in \{0,\dots,N\}}$ with $E\left[|\Delta \hat{Z}_N|^3\right] < \infty$ such that*

$$\sum_{k=0}^{N} \rho_k \Delta \hat{Z}_k = \Delta Y_N \ \ P\text{-a.s.} \tag{B.7}$$

Moreover, we have

$$|\Delta \hat{Z}_N| \le 2 \rho_N^{-1} \|\Delta Y_N\|_\infty \ \ P\text{-a.s.} \tag{B.8}$$

Proof. Uniqueness of $\Delta \hat{Z}$ follows by the same argument as in Lemma 5.2.

To establish existence, let $(Z_k^\rho)_{k \in \{0,\dots,N\}}$ and $(R_k)_{k \in \{1,\dots,N\}}$ be as in the proof of Lemma B.1. It follows from (B.5) and (B.1) that

$$Z_n^\rho R_n = 1 - \sum_{k=0}^{n-1} Z_k^\rho \rho_k, \quad n \in \{0, \dots, N\},$$

in accordance with (5.7). In particular, $Z^\rho R$ is predictable. By nonnegativity of ρ, positivity of Z^ρ and ρ_N, and the definition of R, we may also deduce that $\frac{1}{Z^\rho R}$ is well defined, nondecreasing and satisfies

$$\frac{1}{Z_n^\rho R_n} \le \frac{1}{Z_N^\rho R_N} = \frac{1}{Z_N^\rho \rho_N}, \quad n \in \{1, \dots, N\}. \tag{B.9}$$

As motivated by the discussion in Section 5.1, define $Q^\rho \approx P$ on $\mathcal{F}_N$ by $\mathrm{d}Q^\rho = \frac{Z_N^\rho}{Z_0^\rho} \,\mathrm{d}P$ and the Q^ρ-martingale $(M_k^{\Delta Y})_{k \in \{0,\dots N\}}$ by

$$M_k^{\Delta Y} = E^{Q^\rho}\left[\Delta Y_N \,\middle|\, \mathcal{F}_k\right].$$

Clearly, $M_N^{\Delta Y} = \Delta Y_N$. In analogy to the continuous-time SDE (5.10), now define the process $(Z_k^{\Delta Y})_{k\in\{0,\dots,N\}}$ by

$$Z_k^{\Delta Y} = M_0^{\Delta Y} + \sum_{j=1}^{k} \frac{1}{Z_j R_j}(M_j^{\Delta Y} - M_{j-1}^{\Delta Y}) = \frac{M_0^{\Delta Y}}{Z_0^\rho R_0} + \sum_{j=1}^{k} \frac{1}{Z_j^\rho R_j}(M_j^{\Delta Y} - M_{j-1}^{\Delta Y}),$$

where the second equality follows from the fact that $Z_0^\rho R_0 = 1$ by (B.6). Then $Z^{\Delta Y}$ is a martingale transform of the Q^ρ-martingale $M^{\Delta Y}$. As

$$E^{Q^\rho}\left[|M_N^{\Delta Y}|\frac{1}{\rho_N Z_N^\rho}\right] = Z_0^\rho E\left[\frac{|\Delta Y_N|}{\rho_N}\right] \leq \|\Delta Y_N\|_\infty E\left[\rho_N^{-1}\right] < \infty,$$

it follows from Lemma A.1 that $M^{\Delta Y}$ is a Q^ρ-martingale. Moreover, integration by parts yields

$$Z_N^{\Delta Y} = \frac{M_T^{\Delta Y}}{Z_N^\rho R_N} - \sum_{j=1}^{N} M_{j-1}^{\Delta Y}\left(\frac{1}{Z_j^\rho R_j} - \frac{1}{Z_{j-1}^\rho R_{j-1}}\right).$$

Thus, by a telescopic sum and (B.9),

$$|Z_N^{\Delta Y}| \leq \|\Delta Y_N\|_\infty \left(\frac{1}{Z_N^\rho R_N} + \sum_{j=1}^{N}\left(\frac{1}{Z_j R_j} - \frac{1}{Z_{j-1}^\rho R_{j-1}}\right)\right) \leq 2\|\Delta Y_N\|_\infty \frac{1}{Z_N^\rho \rho_N}. \tag{B.10}$$

Finally, define the process $(\Delta\hat{Z}_k)_{k\in\{0,\dots,N\}}$ by

$$\Delta\hat{Z}_k = Z_k^\rho Z_k^{\Delta Y}.$$

Then $\Delta\hat{Z}$ is a P-martingale by Bayes theorem and satisfies (B.8) because of (B.10). Note that as $E\left[\rho_N^{-3}\right] < \infty$, (B.8) implies in particular that $\Delta\hat{Z}$ has finite third moments. It remains to show that $\Delta\hat{Z}$ satisfies (B.7). By the definitions of $\Delta\hat{Z}$, $Z^{\Delta Y}$, and R, (B.5) and an integration by parts, we have

$$\begin{aligned}
\sum_{k=0}^{N} \rho_k \Delta\hat{Z}_k &= \sum_{k=0}^{N} Z_k^{\Delta Y}\rho_k Z_k^\rho = Z_N^{\Delta Y}\rho_N Z_N^\rho - \sum_{k=1}^{N} Z_{k-1}^{\Delta Y}(R_k Z_k^\rho - R_{k-1}Z_{k-1}^\rho) \\
&= Z_0^{\Delta Y} R_0 Z_0^\rho + \sum_{k=1}^{N} R_k Z_k^\rho (Z_k^{\Delta Y} - Z_{k-1}^{\Delta Y}) \\
&= M_0^{\Delta Y} + \sum_{k=1}^{N}(M_k^{\Delta Y} - M_{k-1}^{\Delta Y}) = M_N^{\Delta Y} = \Delta Y_N.
\end{aligned}$$

This completes the proof. □

The next step is to extend the two results above to continuous time by passing to the limit in an appropriate manner. This requires a rather technical lemma on the existence of special partitions of $[0, T]$, whose mesh sizes do not only go to 0 with respect to the Lebesgue-measure but also relative to the consumption clock μ.

Lemma B.3. *Let μ be a finite measure on $([0, T], \mathcal{B}_{[0,T]})$. Then:*

(a) *For each $N \in \mathbb{N}$, there exists a partition $0 = t_0^N < \cdots < t_N^N = T$ of $[0,T]$ such that*

$$\max_{k=1,\dots,N}(t_k^N - t_{k-1}^N) \le \frac{2T}{N} \quad \text{and} \quad \max_{k=1,\dots,N} \mu((t_{k-1}^N, t_k^N)) \le \frac{2\mu((0,T))}{N}. \tag{B.11}$$

(b) *Let $\{t_0^N, \dots, t_N^N\}_{N\in\mathbb{N}}$ be any sequence of partitions of $[0,T]$ satisfying (B.11) for each N. Then for each càdlàg function $f : [0,T] \to \mathbb{R}$, we have*

$$\lim_{N\to\infty} \sum_{k=1}^{N} \int_{(t_{k-1}^N, t_k^N)} |f(t) - f(t_{k-1})| \, \mu(\mathrm{d}t) = 0.$$

Proof. We may assume without loss of generality that $\mu((0,T)) > 0$; otherwise simply take $t_k^N := \frac{k}{N}T$, $k \in \{0, \dots, N\}$ for (a), and (b) is trivial.

(a) Fix $N \in \mathbb{N}$. Set $s_0^N := 0$ and define recursively, for $k \in \{0, \dots, N-1\}$,

$$s_{k+1}^N = \begin{cases} \inf\left\{t > s_k^N : \mu((s_k^N, t]) > \frac{2\mu((0,T))}{N}\right\}, & \text{if } \mu((s_k^N, T]) > \frac{2\mu((0,T))}{N}, \\ T, & \text{if } \mu((s_k^N, T]) \le \frac{2\mu((0,T))}{N}. \end{cases}$$

Then $\mu((s_{k+1}^N, s_k^N)) \le \frac{2\mu((0,T))}{N} \le \mu((s_{k+1}^N, s_k^N])$ for all $k \in \{0, \dots, N-1\}$ with $s_k^N < T$ and hence

$$\mu((0,T)) \ge \mu((s_0^N, s_k^N]) = \sum_{j=1}^{k} \mu((s_{j-1}^N, s_j^N]) \ge k\frac{2\mu((0,T))}{N}, \tag{B.12}$$

for all $k \in \{0, \dots, N-1\}$ with $s_k^N < T$. Set $k_N = \min\{k \in \{1, \dots, N\} : s_k^N = T\}$. Then $k_N \le \frac{N}{2} + 1$ by (B.12). Now for $t \in (0,T)$, set $j^N(t) := \min\{k \in \{1, \dots, k^N\} : s_k^N > t\}$ and define $t_0^N, \dots, t_N^N$ recursively by $t_0^N = s_0^N = 0$ and, for $k \in 0, \dots, N-1$:

$$t_{k+1}^N = \begin{cases} t_k^N + \frac{2N}{T}, & \text{if } j^N(t_k^N) - t_k^N > \frac{2N}{T}, \\ s_{j^N(t_k^N)}^N, & \text{if } j^N(t_k^N) - t_k^N \le \frac{2N}{T} \quad \text{and} \quad j^N(t_k^N) < k_N, \\ t_k^N + \frac{T - t_k^N}{2}, & \text{if } j^N(t_k^N) - t_k^N \le \frac{2N}{T} \quad \text{and} \quad j^N(t_k^N) = k_N \quad \text{and} \quad k < N-1, \\ T, & \text{if } j^N(t_k^N) - t_k^N \le \frac{2N}{T} \quad \text{and} \quad j^N(t_k^N) = k_N \quad \text{and} \quad k = N-1. \end{cases}$$

In this recursion, we have $j^N(t_k^N) - t_k^N > \frac{2N}{T}$ for at most $\frac{N-1}{2}$ many $k \in \{0, \dots, N-1\}$. Moreover, for at most $k_N - 1 \le \frac{N}{2}$ many $k \in \{0, \dots, N-1\}$ we have $j^N(t_k^N) - t_k^N \le \frac{2N}{T}$ and $j^N(t_k^N) < k^N$. Thus, $j^N(t_{N-1}^N) - t_{N-1}^N \le \frac{2N}{T}$ and $j^N(t_{N-1}^N) = k_N$. This shows that $t_N^N = T$ and the partition $0 = t_0^N < \cdots < t_N^N = T$ satisfies (B.11).

(b) Let $\{t_0^N, \dots, t_N^N\}_{N\in\mathbb{N}}$ be any sequence of partitions of $[0,T]$ satisfying (B.11) for each N. Fix a càdlàg function $f : [0,T] \to \mathbb{R}$, and let $\varepsilon > 0$ be given. By [2, Lemma 1 in Chapter 3], there exists a partition $0 = u_0 < \cdots < u^M = T$ of $[0,T]$ such that

$$\max_{k\in\{1,\dots,M\}} \sup_{u\in[u_{k-1}, u_k)} |f(u) - f(u_{k-1})| < \frac{\varepsilon}{2\mu((0,T))}.$$

Choose N large enough that $\frac{8M\mu((0,T))}{N}\|f\|_\infty < \frac{\varepsilon}{2}$ and define

$$K := \{\ell \in \{1,\dots,N\} : \text{there is } k \in \{1,\dots,M\} \text{ with } u_k \in (t^N_{\ell-1}, t^N_\ell)\}.$$

Clearly $|K| \le M$. Then:

$$\begin{aligned}\sum_{k=1}^N \int_{(t_{k-1},t_k)} |f(t) - f(t_{k-1})|\,\mu(\mathrm{d}t) &= \sum_{k\in K} \int_{(t_{k-1},t_k)} |f(t) - f(t_{k-1})|\,\mu(\mathrm{d}t) \\ &\quad + \sum_{k\in\{1,\dots,N\}\setminus K} \int_{(t_{k-1},t_k)} |f(t) - f(t_{k-1})|\,\mu(\mathrm{d}t) \\ &< |K| \times 2\|f\|_\infty \frac{2\mu((0,T))}{N} + \frac{\varepsilon}{2\mu((0,T))}\mu((0,T)) \\ &= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

This completes the proof. □

With Lemma B.3 at hand, we proceed to extend Lemma B.2 to continuous time by means of Komlós' Lemma:

Lemma B.4. *Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, P)$ be a filtered probability space, where $\mathcal{F}_0$ is P-trivial. Let μ be a finite measure on $([0,T], \mathcal{B}_{[0,T]})$ with $\mu(\{T\}) > 0$. Let ΔY_T be a bounded $\mathcal{F}_T$-measurable random variable and $(\rho_t)_{t\in[0,t]}$ a nonnegative, adapted, càdlàg process. Assume that $\rho_T > 0$ P-a.s., $E\left[\rho_T^{-3}\right] < \infty$, and $E\left[\left(\sup_{t\in[0,T]}\rho_t\right)^3\right] < \infty$. Then there exists a unique martingale $(\Delta\hat{Z}_t)_{t\in[0,T]}$ with $E\left[|\Delta\hat{Z}_T|^3\right] < \infty$ such that*

$$\int_{[0,T]} \rho_t \Delta\hat{Z}_t\,\mu(\mathrm{d}t) = \Delta Y_T \quad P\text{-a.s.},$$

and we have

$$|\Delta\hat{Z}_T| \le 2\rho_T^{-1}\mu(\{T\})^{-1}\|\Delta Y_T\|_\infty \quad P\text{-a.s.}$$

Moreover:

(a) *If $\Delta Y_T = 1$, then $\Delta\hat{Z}$ is nonnegative.*

(b) *If $\frac{\rho_T}{B_T}$ is uniformly bounded from below, then $\Delta\hat{Z}$ is bounded.*

Proof. Uniqess follows from Lemma 5.2.

To establish existence, we use an approximation argument. For $N \in \mathbb{N}$, let $(t^N_i)_{i\in\{0,\dots,N\}}$ be a partition of $[0,T]$ as in Lemma B.3 and let $(\rho^N_k)_{k\in\{0,\dots,N\}}$ be defined by

$$\rho^N_k = \begin{cases} \rho_{t^N_k}\mu([t^N_k, t^N_{k+1})) & \text{if } k \in \{0,\dots,N-1\}, \\ \rho_T\mu(\{T\}) & \text{if } k = N. \end{cases}$$

By Lemma B.2, for each $N \in \mathbb{N}$, there exists a martingale $(\Delta\hat{Z}^N_t)_{t\in[0,T]}$ with finite third moments such that

$$\sum_{k=0}^{N} \rho_k^N \Delta \hat{Z}_{t_k^N} = \Delta Y_T \;\; P\text{-a.s.} \tag{B.13}$$

Moreover, if $\Delta = 1$, it follows from Lemma B.1 that each $\Delta\hat{Z}^N$ is positive. Define the martingale $(N^\rho)_{t\in[0,T]}$ by

$$N_t^\rho = E\left[\frac{1}{\rho_T \mu(\{T\})} \,\middle|\, \mathcal{F}_t \right].$$

Then, by (B.7):

$$|\Delta\hat{Z}_t^N| \leq E\left[|\Delta\hat{Z}_T^N| \,\middle|\, \mathcal{F}_t \right] \leq 2\|\Delta Y_N\|_\infty N_t^\rho \;\; P\text{-a.s.} \tag{B.14}$$

We proceed to show that $\int_{[0,T]} \rho_t \Delta\hat{Z}_t^N \,\mu(\mathrm{d}t)$ converges to ΔY_T in L^1. By (B.13) and (B.14),

$$\begin{aligned}
\left| \int_{[0,T]} \rho_u^N \Delta\hat{Z}_u^N \,\mu(\mathrm{d}u) - \Delta Y_T \right| &= \left| \int_{[0,T]} \rho_u \Delta\hat{Z}_u^N \,\mu(\mathrm{d}u) - \sum_{k=0}^{N} \rho_k^N \Delta\hat{Z}_{t_k^N} \right| \\
&= \left| \sum_{k=1}^{N} \int_{(t_{k-1}^N, t_k^N)} \left(\rho_u \Delta\hat{Z}_u^N - \rho_{t_{k-1}^N} \Delta\hat{Z}_{t_{k-1}^N} \right) \mu(\mathrm{d}u) \right| \\
&\leq \sum_{k=1}^{N} \int_{(t_{k-1}^N, t_k^N)} \rho_u \left| \Delta\hat{Z}_u^N - \Delta\hat{Z}_{t_{k-1}^N} \right| \mu(\mathrm{d}u) \\
&\quad + \sum_{k=1}^{N} \int_{(t_{k-1}^N, t_k^N)} |\Delta\hat{Z}_{t_{k-1}^N}^N| \left| \rho_u - \rho_{t_{k-1}^N} \right| \mu(\mathrm{d}u) \\
&\leq \left(\sup_{t\in[0,T]} \rho_t \right) \sum_{k=1}^{N} \int_{(t_{k-1}^N, t_k^N)} \left| \Delta\hat{Z}_u^N - \Delta\hat{Z}_{t_{k-1}^N}^N \right| \mu(\mathrm{d}u) \\
&\quad + 2\|\Delta Y_N\|_\infty \left(\sup_{t\in[0,T]} |N_t^\rho| \right) \sum_{k=1}^{N} \int_{(t_{k-1}^N, t_k^N)} \left| \rho_u - \rho_{t_{k-1}^N} \right| \mu(\mathrm{d}u).
\end{aligned} \tag{B.15}$$

Both summands on the right-hand side of (B.15) converge P-a.s. to 0 by Lemma B.3. Moreover, the right hand side of (B.15) is P-a.s. bounded from above by

$$(2 + 4\|\Delta Y_N\|_\infty) \left(\sup_{t\in[0,T]} \rho_t \right) \left(\sup_{t\in[0,T]} |N_t^\rho| \right) \mu((0,T)).$$

Hölder's inequality and Doob's maximal inequality yield

$$E\left[\left(\left(\sup_{t\in[0,T]} \rho_t \right) \left(\sup_{t\in[0,T]} |\Delta\hat{Z}_t^N| \right) \right)^{3/2} \right] \leq cE\left[\left(\sup_{t\in[0,T]} \rho_t \right)^3 \right]^{\frac{1}{2}} E\left[(N_T^\rho)^3 \right]^{\frac{1}{2}} < \infty,$$

for some constant $c > 0$. Thus, the right-hand side of (B.15) is uniformly integrable. As a result:

$$\lim_{N\to\infty} E\left[\left| \int_{[0,T]} \rho_u \Delta\hat{Z}_u^N \,\mu(\mathrm{d}u) - \Delta Y_T \right| \right] = 0. \tag{B.16}$$

Finally, we construct $\Delta\hat{Z}$ as the "Komlós limit" of the $\Delta\hat{Z}^N$. Since the sequence $(\Delta\hat{Z}_T^N)_{N\in\mathbb{N}}$ is bounded in L^3 (and a fortiori in L^1) by (B.14), Komlós' Lemma [12, Theorem 5.2.1] shows that there exist a sequence $(\Delta\tilde{Z}^N_T)_{N\in\mathbb{N}}$ with $\Delta\tilde{Z}^N_T \in \mathrm{conv}(\Delta\hat{Z}^N_T, \Delta\hat{Z}^{N+1}_T, \Delta\hat{Z}^{N+2}_T, \ldots)$ and a random variable $\Delta\hat{Z}_T \in L^1(\mathcal{F}_T)$ such that $\Delta\tilde{Z}^N_T$ converges P-a.s. to $\Delta\hat{Z}_T$. Note that if $\Delta Y_T = 1$, then each $\Delta\tilde{Z}^N_T$ is positive, and so $\Delta\hat{Z}_T$ is nonnegative. As each $\Delta\tilde{Z}^N_T$ satisfies the estimate (B.14), the sequence $(\Delta\tilde{Z}^N_T)_{N\in\mathbb{N}}$ is bounded in L^3 and therefore converges to $\Delta\hat{Z}_T$ also in L^2 by the de la Vallée-Poussin's criterion for uniform integrability (cf., e.g., the remark before [13, Lemma 4.10]). Denote by $\Delta\hat{Z} = (\Delta\hat{Z}_t)_{t\in[0,T]}$ the càdlàg version of the martingale generated by $\Delta\hat{Z}_T$. Clearly, $\Delta\hat{Z}$ has finite third moments. We claim that

$$\int_{[0,T]} \rho_u \Delta\hat{Z}_u \,\mu(\mathrm{d}u) = \Delta Y \quad P\text{-a.s.}$$

Since $\int_{[0,T]} \rho_u \Delta\tilde{Z}^N_u \,\mu(\mathrm{d}u)$ converges to ΔY_T in L^1 by (B.16) and linearity of the integral, it suffices to show that $\int_{[0,T]} \rho_u \Delta\tilde{Z}^N_u \,\mu(\mathrm{d}u)$ converges to $\int_{[0,T]} \rho_u \Delta\hat{Z}_u \,\mu(\mathrm{d}u)$ in L^1. By Hölder's inequality and Doob's maximal inequality, we have

$$\begin{aligned} E\left[\left|\int_{[0,T]} \rho_u(\Delta\tilde{Z}^N_u - \Delta\hat{Z}_u)\,\mu(\mathrm{d}u)\right|\right] &\le \mu([0,T]) E\left[\left(\sup_{t\in[0,T]} \rho_t\right)\left(\sup_{t\in[0,T]} |\Delta\tilde{Z}^N_u - \Delta\hat{Z}_u|\right)\right] \\ &\le 2\mu([0,T])) E\left[\left(\sup_{t\in[0,T]} \rho_t\right)^2\right]^{1/2} E\left[(\Delta\tilde{Z}^N_T - \Delta\hat{Z}_T)^2\right]^{1/2}. \end{aligned}$$

The claim now follows from the fact that $\Delta\tilde{Z}^N_T$ converges to $\Delta\hat{Z}_T$ in L^2 as $N \to \infty$. □

References

[1] D. Becherer, *Bounded solutions to backward SDEs with jumps for utility optimization and indifference hedging*, Ann. Appl. Probab. **16** (2006), no. 4, 2027–2054.

[2] P. Billingsley, *Convergence of Probability Measures*, second ed., John Wiley & Sons, Inc., New York, 1999.

[3] C. D. Carroll and M. S. Kimball, *On the concavity of the consumption function*, Econometrica **64** (1996), no. 4, 981–992.

[4] A. Černý and J. Kallsen, *On the structure of general mean-variance hedging strategies*, Ann. Probab. **35** (2007), no. 4, 1479–1531.

[5] P. O. Christensen and K. Larsen, *Incomplete continuous-time securities markets with stochastic income volatility*, Rev. Asset Pricing Stud. **4** (2014), no. 2, 247–285.

[6] M. H. A. Davis, *Option pricing in incomplete markets*, Mathematics of Derivative Securities, Cambridge University Press, Cambridge, 1997, p. 216–226.

[7] C. Dellacherie and P.-A. Meyer, *Probabilities and Potential. B*, North-Holland Publishing Co., Amsterdam, 1982.

[8] C. Gollier, *The Economics of Risk and Time*, MIT Press, Cambridge, MA, 2001.

[9] C. Gourieroux, J. P. Laurent, and H. Pham, *Mean-variance hedging and numéraire*, Math. Finance **8** (1998), no. 3, 179–200.

[10] V. Henderson, *Valuation of claims on nontraded assets using utility maximization*, Math. Finance **12** (2002), no. 4, 351–373.

[11] M. Herdegen and J. Muhle-Karbe, *Stability of Radner equilibria with respect to small frictions*, Finance Stoch. **22** (2018), no. 2, 443–502.

[12] Y. Kabanov and M. Safarian, *Markets with Transaction Costs: Mathematical Theory*, Springer, Berlin, 2009.

[13] O. Kallenberg, *Foundations of modern probability.*, second ed., Springer, Berlin, 2002.

[14] J. Kallsen and S. Li, *Portfolio optimization under small transaction costs: a convex duality approach*, Preprint, 2013.

[15] J. Kallsen and J. Muhle-Karbe, *The general structure of optimal investment and consumption with small transaction costs*, Math. Finance **27** (2017), no. 2, 659–703.

[16] J. Kallsen and T. Rheinländer, *Asymptotic utility-based pricing and hedging for exponential utility*, Statist. Decisions **28** (2011), no. 1, 17–36.

[17] I. Karatzas and S. G. Kou, *On the pricing of contingent claims under constraints*, Ann. Appl. Probab. **6** (1996), no. 2, 321–369.

[18] I. Karatzas, J. P. Lehoczky, S. E. Shreve, and G.-L. Xu, *Martingale and duality methods for utility maximization in an incomplete market*, SIAM J. Control Optim. **29** (1991), no. 3, 702–730.

[19] I. Karatzas and G. Žitković, *Optimal consumption from investment and random endowment in incomplete semimartingale markets*, Ann. Probab. **31** (2003), no. 4, 1821–1858.

[20] N. Kazamaki, *Continuous Exponential Martingales and BMO*, Springer, Berlin, 1994.

[21] M. S. Kimball, *Precautionary saving in the small and in the large*, Econometrica **58** (1990), no. 1, 53–73.

[22] D. Kramkov and M. Sirbu, *Sensitivity analysis of utility-based prices and risk-tolerance wealth processes*, Ann. Appl. Probab. **16** (2006), no. 4, 2140–2194.

[23] ______, *Asymptotic analysis of utility-based hedging strategies for small number of contingent claims*, Stochastic Process. Appl. **117** (2007), no. 11, 1606–1620.

[24] P. Lakner and E. Slud, *Optimal consumption by a bond investor: the case of random interest rate adapted to a point process*, SIAM J. Control Optim. **29** (1991), no. 3, 638–655.

[25] K. Larsen, O. Mostovyi, and G. Žitković, *An expansion in the model space in the context of utility maximization*, Finance Stoch. **22** (2018), no. 2, 297–326.

[26] K. Larsen, G. Zitkovic, and H. M. Soner, *Facelifting in Utility Maximization*, Finance Stoch. **20** (2014), no. 1, 99–121.

[27] H. E. Leland, *Saving and uncertainty: the precautionary demand for saving*, Q. J. Econ. **82** (1968), no. 3, 465–473.

[28] M. Mania and M. Schweizer, *Dynamic exponential utility indifference valuation*, Ann. Appl. Probab. **15** (2005), no. 3, 2113–2143.

[29] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic theory*, Oxford University Press, New Yorwk, N.Y., 1995.

[30] L. Moreau, J. Muhle-Karbe, and H. M. Soner, *Trading with small price impact*, Math. Finance **27** (2017), no. 2, 350–400.

[31] O. Mostovyi, *Optimal investment with intermediate consumption and random endowment*, Math. Finance **27** (2017), no. 1, 96–114.

[32] M. Schweizer, *Approximating random variables by stochastic integrals*, Ann. Probab. **22** (1994), no. 3, 1536–1575.

[33] H. M. Soner and N. Touzi, *Homogenization and asymptotics for small transaction costs*, SIAM J. Control Optim. **51** (2013), no. 4, 2893–2921.