Finance Stoch (2018) 22:297–326
https://doi.org/10.1007/s00780-017-0353-3

An expansion in the model space in the context of utility maximization

Kasper Larsen[1] · **Oleksii Mostovyi**[2] · **Gordan Žitković**[3]

Received: 10 August 2016 / Accepted: 26 September 2017 / Published online: 14 December 2017

Abstract In the framework of an incomplete financial market where the stock price dynamics are modeled by a continuous semimartingale (not necessarily Markovian), an explicit second-order expansion formula for the power investor's value function—seen as a function of the underlying market price of risk process—is provided. This allows us to provide first-order approximations of the optimal primal and dual controls. Two specific calibrated numerical examples illustrating the accuracy of the method are also given.

Keywords Continuous semimartingales · Second-order expansion · Incomplete markets · Power utility · Convex duality · Optimal investment

Mathematics Subject Classification (2010) 91G10 · 91G80 · 60K35

JEL Classification C61 · G11

1 Introduction

In an incomplete financial setting with noise governed by a continuous martingale and in which the investor's preferences are modeled by a negative power utility function,

✉ K. Larsen
kl756@math.rutgers.edu

O. Mostovyi
oleksii.mostovyi@uconn.edu

G. Žitković
gordanz@math.utexas.edu

[1] Department of Mathematics, Rutgers University, Piscataway, NJ 08854, USA

[2] Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA

[3] Department of Mathematics, University of Texas at Austin, Austin, TX 78712, USA

we provide a second-order Taylor expansion of the investor's value function with respect to perturbations of the underlying market price of risk process. We show that tractable models can be used to approximate highly intractable ones as long as the latter can be interpreted as perturbations of the former. As a by-product of our analysis, we explicitly construct first-order approximations of both the primal and the dual optimizers. Finally, we apply our approximation in two numerical examples.

There are two different ways of looking at our contribution: as a tool to approximate the value function and perform numerical computations, or as a stability result with applications to statistical estimation. Let us elaborate on these, and the related work, in order.

An approximation interpretation. The conditions for existence and uniqueness of the investor's utility optimizers are well established (see [21, 27]). However, in general settings, the numerical computation of the investor's value function and corresponding optimal trading strategy remain a challenging problem. Various existing approaches include the following:

1. In Markovian settings, the value function can typically be characterized by an HJB equation. Its numerical implementation through a finite-grid approximation is naturally subject to the curse of dimensionality. Many authors (see [23, 40, 7, 25, 33]) opt for affine and quadratic models for which closed-form solutions exist. Going beyond these specifications in high-dimensional settings by using PDE techniques seems to be very hard computationally. We refer to [26] and the references therein for recent advances on numerically solving the PDE stemming from the HJB equation.

2. In general (i.e., not necessarily Markovian) complete models, [10] and [12] provide efficient Monte Carlo simulation techniques based on the martingale method for complete markets developed in [9] and [20].

3. Other approximation methods are based on various Taylor-type expansions. The authors of [4] and [5] log-linearize the investor's budget constraint as well as the investor's first-order condition for optimality. Kogan and Uppal [24] expand in the investor's risk-aversion coefficient around the log-investor (the myopic investor's problem is known to be tractable even in incomplete settings). When solving the HJB equation numerically (using a Longstaff–Schwartz type of technique), Brandt et al. [3] expand the value function in the wealth variable to a fourth-degree Taylor approximation.

4. Based on the duality results in [21], Haugh et al. [14] provide an upper bound on the error stemming from using suboptimal strategies. Bick et al. [2] propose a method based on minimizing over a subset of dual elements. This subset is chosen such that the corresponding dual utility can be computed explicitly and transformed into a feasible primal strategy.

5. It is also important to mention the recent explosion in research in asymptotic methods in a variety of different areas in mathematical finance (transaction costs, pricing, ambiguity aversion, etc.). Since we focus on model expansion in utility maximization in this paper, we simply point the reader to some of the most recent papers, namely [1, 18, 16], and the references therein, for further information.

In our work, no Markovian assumption is imposed and we deal with general, possibly incomplete markets with continuous price processes. We consider the utility functions $U(x) := x^p/p$ for $x > 0$. We note that while our results apply only to

$p < 0$, it is possible to extend them to $p \in (0, 1)$ at the cost of imposing additional integrability requirements. We do not pursue such an extension because the parameter range $p \in (0, 1)$ which we leave out seems to lie outside the typical range of risk-aversion parameters observed in practice (see e.g. [39]).

A stability interpretation. As mentioned above, our contribution can also be seen as a stability result. It is well known (see e.g. [38]) that even in Samuelson's model, estimating the drift is far more challenging than estimating the volatility. Larsen and Žitković [31] identify the kinds of perturbations of the market price of risk process under which the value function behaves continuously. In the present paper, we take the stability analysis one step further and provide a second-order Taylor expansion in an infinite-dimensional space of market price of risk processes. In this way, we not only identify the "continuous" directions, but also those features of the market price of risk process that affect the solution of the utility maximization problem the most (at least locally). Any statistical procedure which is performed with utility maximization in mind should therefore focus on those salient features in order to use the scarce data most efficiently.

Similar perturbations have been considered by [34], but in a somewhat different setting. [34] is based on Malliavin calculus and produces a first-order expansion for the utility indifference price of an exponential investor in an Itô-process driven market; some of the ideas used can be traced to the related work [11].

Mathematical challenges. From a mathematical point of view, our approach is founded on two ideas. One of them is to extend the techniques and results of [31]; indeed, the basic fact that the dual minimizers converge when the market price of risk process does is heavily exploited. It does not, however, suffice to get the full picture. For that, one needs to work on the primal and the dual problems simultaneously and use a pair of bounds. The ideas used there are related to, and can be interpreted as, a nonlinear version of the primal–dual second-order error estimation techniques first used in [15] in the context of mathematical finance. The first-order expansion in the quantity of the unspanned contingent claim developed in [15] was generalized in [29] (see also [28]). The arguments in these papers rely on convexity and concavity properties in the expansion parameter (wealth and number of unspanned claims). This is not the case in the present paper; indeed, when seen as a function of the underlying market price of risk process, the investor's value function is neither convex nor concave and a more delicate, local analysis needs to be performed.

Numerical examples. In Sect. 5, we use two examples to illustrate how our approximation performs under realistic conditions. First, we consider the Kim–Omberg model (see [23]) which is widely used in the financial literature. Under a calibrated set of parameters, we find that our approximation is indeed very accurate when compared to the exact values.

Our second set of examples belongs to the class of extended affine models introduced in [8]. The authors show that this class of models has superior empirical properties when compared to popular affine and quadratic specifications (such as those used e.g. in [33]). The resulting optimal investment problem for the extended affine models unfortunately does not seem to be explicitly solvable. Our approximation technique turns out to be easily applicable, and our error bounds are quite tight in the relevant parameter ranges. Furthermore, unlike numerical methods based on

PDEs, our method's computation time grows linearly in the number of underlying factors. Therefore, we can and do apply our theory to a high-dimensional extended affine model.

2 A family of utility maximization problems

2.1 The setup

We work on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ with a finite time horizon $T > 0$. We assume that the filtration $\mathbb{F}$ is right-continuous and that the σ-algebra $\mathcal{F}_0$ consists of all $\mathbb{P}$-trivial subsets of $\mathcal{F}$.

Let M be a continuous local martingale, and let $R^{(\varepsilon)}$, $\varepsilon \geq 0$, be a family of continuous $\mathbb{F}$-semimartingales given by

$$R_s^{(\varepsilon)} := M_s + \int_0^s \lambda_t^{(\varepsilon)}\, d\langle M\rangle_t, \quad s \in [0,T], \text{ where } \lambda^{(\varepsilon)} := \lambda + \varepsilon\lambda', \tag{2.1}$$

for a pair $\lambda, \lambda' \in \mathcal{P}_M^2$, where $\mathcal{P}_M^2$ denotes the collection of all progressively measurable processes π with $\int_0^T \pi_t^2\, d\langle M\rangle_t < \infty$. Since $S^{(\varepsilon)} := \mathcal{E}(R^{(\varepsilon)})$ (where $\mathcal{E}$ denotes the stochastic exponential) will be interpreted as the price process of a financial asset, the assumption that $\lambda^{(\varepsilon)} \in \mathcal{P}_M^2$ can be taken as a minimal no-arbitrage-type condition. We remark right away that further integrability conditions on λ and λ' need to be imposed below for our main results to hold.

2.2 The utility maximization problem

Given $x > 0$ and $\varepsilon \geq 0$, let $\mathcal{X}^{(\varepsilon)}(x)$ denote the set of all nonnegative wealth processes starting from initial wealth x in the financial market consisting of $S^{(\varepsilon)} := \mathcal{E}(R^{(\varepsilon)})$ and a zero-interest bond, i.e.,

$$\mathcal{X}^{(\varepsilon)}(x) := \{x\mathcal{E}(\pi \cdot R^{(\varepsilon)})_t, t \in [0,T] : \pi \in \mathcal{P}_M^2\}.$$

Here, π is interpreted as the fraction of wealth invested in the risky asset $S^{(\varepsilon)}$. The investor's preferences are modeled by a CRRA (power) utility function with the risk-aversion parameter $p < 0$, i.e.,

$$U(x) := \frac{x^p}{p}, \quad x > 0.$$

The value function of the corresponding optimal investment problem is defined by

$$u^{(\varepsilon)}(x) := \sup_{X\in\mathcal{X}^{(\varepsilon)}(x)} \mathbb{E}[U(X_T)], \quad x > 0. \tag{2.2}$$

2.3 The dual of the utility maximization problem

As is usual in the utility maximization literature, a fuller picture is obtained if one also considers an appropriate version of the optimization problem dual to (2.2). For that, we need to examine the no-arbitrage properties of the set of models introduced in Sect. 2.1 above.

We observe first that the assumptions we placed on the market price of risk processes $\lambda^{(\varepsilon)}$ above are not sufficient to guarantee the existence of an equivalent martingale measure (and hence NFLVR). They do preclude so-called "arbitrages of the first kind" and imply the related NUBPR condition. In particular, for all $x, y > 0$ and $\varepsilon \geq 0$, there exists a (strictly) positive càdlàg supermartingale Y with the property that $Y_0 = y$ and YX is a supermartingale for each $X \in \mathcal{X}^{(\varepsilon)}(x)$; we denote the set of all such processes by $\mathcal{Y}^{(\varepsilon)}(y)$. While this is a consequence of the NUBPR condition in general, in our case, an example of a process in $\mathcal{Y}^{(\varepsilon)}(y)$ is given explicitly as $yZ^{(\varepsilon)}$, where $Z^{(\varepsilon)}$ is the *minimal* local martingale density, that is,

$$Z_t^{(\varepsilon)} = \mathcal{E}(-\lambda^{(\varepsilon)} \cdot M)_t.$$

Having described the dual domain, we remind the reader that the *conjugate* utility function $V : (0, \infty) \to \mathbb{R}$ is defined by

$$V(y) := \sup_{x>0} \big(U(x) - xy\big) = \frac{y^{-q}}{q}, \quad \text{where } q := \frac{p}{1-p} \in (-1, 0).$$

We define the *dual value function* $v^{(\varepsilon)} : (0, \infty) \to \mathbb{R}$ by

$$v^{(\varepsilon)}(y) := \inf_{Y \in \mathcal{Y}^{(\varepsilon)}(y)} \mathbb{E}[V(Y_T)], \quad y > 0, \varepsilon \geq 0. \tag{2.3}$$

Due to negativity (and a fortiori finiteness) of the primal value function $u^{(\varepsilon)}$, the (abstract) Theorem 3.2 of [27] can now be applied (see also [35]). Its main assumption, namely the bipolar relationship between the primal and dual domains, holds due to the existence of the numéraire process given explicitly by $1/Z^{(\varepsilon)}$ (see Theorem 4.12 in [19]). One can also use a simpler argument (see [30]), which applies only to the case of a CRRA utility with $p < 0$, to obtain the following conclusions for all $\varepsilon \geq 0$:

1. Both $u^{(\varepsilon)}$ and $v^{(\varepsilon)}$ are finite, and we have the conjugacy relationships

$$v^{(\varepsilon)}(y) = \sup_{x>0} \big(u^{(\varepsilon)}(x) - xy\big), \qquad u^{(\varepsilon)}(x) = \inf_{y>0} \big(v^{(\varepsilon)}(y) + xy\big). \tag{2.4}$$

2. For all $x, y > 0$, there exist solutions $\hat{X}^{(\varepsilon)}(x) \in \mathcal{X}^{(\varepsilon)}(x)$ and $\hat{Y}^{(\varepsilon)}(y) \in \mathcal{Y}^{(\varepsilon)}(y)$ of (2.2) and (2.3), respectively, and they are related by

$$U'\big(\hat{X}_T^{(\varepsilon)}(x)\big) = \hat{Y}_T^{(\varepsilon)}\big(y^{(\varepsilon)}(x)\big), \quad \text{where } y^{(\varepsilon)}(x) := \frac{d}{dx} u^{(\varepsilon)}(x) = px^{p-1}u^{(\varepsilon)}(1).$$

3. The product $\hat{X}^{(\varepsilon)}\hat{Y}^{(\varepsilon)}$ is a uniformly integrable martingale. In particular,

$$\mathbb{E}[\hat{X}_T^{(\varepsilon)}\hat{Y}_T^{(\varepsilon)}] = xy.$$

The homogeneity of the utility function U and its conjugate V transfers to the value functions $u^{(\varepsilon)}$ and $v^{(\varepsilon)}$ and the solutions $\hat{X}^{(\varepsilon)}$ and $\hat{Y}^{(\varepsilon)}$, i.e.,

$$\begin{aligned} u^{(\varepsilon)}(x) &= x^p u^{(\varepsilon)}, & v^{(\varepsilon)}(y) &= y^{-q} v^{(\varepsilon)}, \\ \hat{X}^{(\varepsilon)}(x) &= x\hat{X}^{(\varepsilon)}, & \hat{Y}^{(\varepsilon)}(y) &= y\hat{Y}^{(\varepsilon)}, \end{aligned} \tag{2.5}$$

where, to simplify the notation, we write $u^{(\varepsilon)}$, $v^{(\varepsilon)}$, $\hat{X}^{(\varepsilon)}$, and $\hat{Y}^{(\varepsilon)}$ for $u^{(\varepsilon)}(1)$, $v^{(\varepsilon)}(1)$, $\hat{X}^{(\varepsilon)}(1)$, and $\hat{Y}^{(\varepsilon)}(1)$, respectively.

2.4 A change of measure

For $\varepsilon = 0$, we denote by $\hat{\pi}^{(0)}$ the primal optimizer, that is, the process in $\mathcal{P}^2_M$ such that

$$\hat{X}^{(0)}_t = \mathcal{E}(\hat{\pi}^{(0)} \cdot R^{(0)})_t.$$

We define the probability measure $\tilde{\mathbb{P}}^{(0)}$ by

$$\frac{d\tilde{\mathbb{P}}^{(0)}}{d\mathbb{P}} := \hat{X}^{(0)}_T \hat{Y}^{(0)}_T = \frac{1}{v^{(0)}} V(\hat{Y}^{(0)}_T) = \frac{1}{u^{(0)}} U(\hat{X}^{(0)}_T), \tag{2.6}$$

where the equalities follow from the identities $xU'(x) = pU(x)$, $yV'(y) = -qV(y)$ and the relations between the value functions outlined above.

The measure $\tilde{\mathbb{P}}^{(0)}$ has been in the mathematical finance literature for a while (see e.g. [27, Sect. 2]). The explicit form of $\tilde{\mathbb{P}}^{(0)}$ is not generally available, but we note that by Girsanov's theorem, the process

$$\tilde{M}^p_s := M_s + \int_0^s (\lambda_t - \hat{\pi}^{(0)}_t)\, d\langle M\rangle_t \tag{2.7}$$

is a $\tilde{\mathbb{P}}^{(0)}$-local martingale. This fact will be used below in the proof of Proposition 4.3.

3 The problem and the main results

We first provide first-order expansions and error estimates of the primal and dual value functions. Secondly, we provide an expansion of the optimal controls in the Brownian setting.

3.1 Value functions

At the basic level, we are interested in the first-order properties of the convergence, as $\varepsilon \searrow 0$, of the value functions $u^{(\varepsilon)}$ and $v^{(\varepsilon)}$ to the value functions $u^{(0)}$ and $v^{(0)}$ of the "base" model (corresponding to $\varepsilon = 0$). To familiarize ourselves with the flavor of the results we can expect in the general case, we start by analyzing a similar problem for logarithmic utility. It has the advantage that it admits a simple explicit solution. Let $u^{(\varepsilon)}_{\log}(x)$ and $v^{(\varepsilon)}_{\log}(y)$ denote the value functions of the problems as in (2.2) and

(2.3) above, but with $U(x) := \log x$ and $V(y) := \sup_x (U(x) - xy) = -\log y - 1$. It is a classical result that as long as $\mathbb{E}[\int_0^T (\lambda_t^2 + (\lambda_t')^2)\, d\langle M\rangle_t] < \infty$, we have

$$u_{\log}^{(\varepsilon)}(x) = \log x + \frac{1}{2}\mathbb{E}\bigg[\int_0^T (\lambda_t^{(\varepsilon)})^2\, d\langle M\rangle_t\bigg], \qquad v_{\log}^{(\varepsilon)} = u_{\log}^{(\varepsilon)} - 1.$$

The (exact) second-order expansion in ε of $u_{\log}^{(\varepsilon)}(x)$ is thus given by

$$\begin{aligned} u_{\log}^{(\varepsilon)}(x) &= u_{\log}^{(0)}(x) + \varepsilon\mathbb{E}\bigg[\int_0^T \lambda_t\lambda_t'\, d\langle M\rangle_t\bigg] + \frac{1}{2}\varepsilon^2\mathbb{E}\bigg[\int_0^T (\lambda_t')^2\, d\langle M\rangle_t\bigg] \\ &= u_{\log}^{(0)}(x) + \varepsilon\mathbb{E}\bigg[\int_0^T \lambda_t'\, dR_t^{(0)}\bigg] + \frac{1}{2}\varepsilon^2\mathbb{E}\bigg[\int_0^T (\lambda_t')^2\, d\langle M\rangle_t\bigg], \end{aligned}$$

where $R^{(0)}$ is defined in (2.1). We cannot expect the value function to be a second-order polynomial in ε in the case of a general power utility. We do obtain a formally similar first-order expansion in Theorem 3.1 below and an analogous error estimate in Theorem 3.2. Section 4 is devoted to their proofs. We remind the reader of the homogeneity relationships in (2.5); they allow us to assume from now on that $x := y := 1$.

Theorem 3.1 *In the setting of Sect. 2, we assume that*

$$\int_0^T (\lambda_t')^2\, d\langle M\rangle_t \in \mathbb{L}^{1-p}(\mathbb{P}) \quad \textit{and} \quad \int_0^T \lambda_t'\, dR_t^{(0)} \in \bigcup_{s>(1-p)} \mathbb{L}^s(\mathbb{P}).$$

Then with $\Delta^{(0)} := \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\int_0^T \lambda_t'\, dR_t^{(0)}]$, *where* $\tilde{\mathbb{P}}^{(0)}$ *is defined by* (2.6), *we have*

$$\frac{d}{d\varepsilon}u^{(\varepsilon)}\bigg|_{\varepsilon=0+} := \lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}(u^{(\varepsilon)} - u^{(0)}) = pu^{(0)}\Delta^{(0)}, \tag{3.1}$$

$$\frac{d}{d\varepsilon}v^{(\varepsilon)}\bigg|_{\varepsilon=0+} := \lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}(v^{(\varepsilon)} - v^{(0)}) = qv^{(0)}\Delta^{(0)}. \tag{3.2}$$

Theorem 3.2 *In the setting of Sect. 2, we assume that*

$$\int_0^T (\lambda_t')^2\, d\langle M\rangle_t, \int_0^T \lambda_t'\, dR_t^{(0)} \in \mathbb{L}^{2(1-p)}(\mathbb{P}) \quad \textit{and} \quad \Phi^2 e^{\varepsilon_0|p|\Phi^-} \in \mathbb{L}^1(\tilde{\mathbb{P}}^{(0)}) \tag{3.3}$$

for some $\varepsilon_0 > 0$, *where* $\Phi := \int_0^T \hat{\pi}_t^{(0)}\lambda_t'\, d\langle M\rangle_t$ *and* $\Phi^- := -\min\{\Phi, 0\}$. *Then there exist constants* $C > 0$ *and* $\varepsilon_0' \in (0, \varepsilon_0]$ *such that for all* $\varepsilon \in [0, \varepsilon_0']$, *we have*

$$|u^{(\varepsilon)} - u^{(0)} - \varepsilon pu^{(0)}\Delta^{(0)}| \le C\varepsilon^2, \tag{3.4}$$

$$|v^{(\varepsilon)} - v^{(0)} - \varepsilon qv^{(0)}\Delta^{(0)}| \le C\varepsilon^2. \tag{3.5}$$

Remark 3.3 1. It is perhaps more informative to think of the results in Theorems 3.1 and 3.2 on the logarithmic scale. As is evident from (3.1) and (3.2), the functions

$u^{(\varepsilon)}$ and $v^{(\varepsilon)}$ admit the right *logarithmic* derivatives $p\Delta^{(0)}$ and $q\Delta^{(0)}$, respectively, at $\varepsilon = 0$. Moreover, we have the small-ε asymptotics

$$u^{(\varepsilon)} = u^{(0)} e^{\varepsilon p \Delta^{(0)} + O(\varepsilon^2)}, \qquad v^{(\varepsilon)} = v^{(0)} e^{\varepsilon q \Delta^{(0)} + O(\varepsilon^2)}.$$

If one takes one step further and uses the *certainty equivalent* $\mathrm{CE}^{(\varepsilon)}$ given by

$$U(\mathrm{CE}^{(\varepsilon)}) = u^{(\varepsilon)},$$

we note that $\Delta^{(0)}$ is precisely the infinitesimal growth rate of $\mathrm{CE}^{(\varepsilon)}$ at $\varepsilon = 0$—an ε-change of the market price of risk in the direction λ' leads to an $e^{\varepsilon\Delta^{(0)}}$-fold increase in the certainty equivalent of the initial wealth.

2. A careful analysis of the proof of Theorem 3.2 below yields the following additional information:

(a) The proof of Proposition 4.3 reveals that $\Delta^{(0)} = \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\Phi]$.

(b) The condition involving Φ in (3.3) is needed only for the upper bound in (3.4) and the lower bound in (3.5). The other two bounds hold for all $\varepsilon \geq 0$, even if (3.3) holds with $\varepsilon_0 = 0$.

(c) The constants C and ε_0' depend (in a simple way) on ε_0, p and the $\mathbb{L}^{2(1-p)}(\tilde{\mathbb{P}}^{(0)})$- and $\mathbb{L}^1(\tilde{\mathbb{P}}^{(0)})$-bounds of the random variables in (3.3). For two one-sided bounds, explicit formulas are given in Propositions 4.2 and 4.3. The other two bounds are somewhat less informative; so we do not compute them explicitly.

(d) Even though we cannot claim that the functions $u^{(\varepsilon)}$ and $v^{(\varepsilon)}$ are convex or concave in ε, it is possible to show their local *semiconcavity* (see [6, Definition 1.1.1]). This can be done via the techniques from the proof of Theorem 3.2.

3. The assumption of constant risk-aversion (power utility) allows us to incorporate many stochastic interest rate models into our setting. Indeed, provided that $c := \mathbb{E}[e^{p\int_0^T r_t dt}] < \infty$, we can introduce the probability measure $\mathbb{P}^r$ defined by

$$\frac{d\mathbb{P}^r}{d\mathbb{P}} := c e^{p\int_0^T r_t\, dt}$$

on $\mathcal{F}_T$. For any admissible wealth process X, we then have

$$\mathbb{E}[U(X_T)] = c\,\mathbb{E}^{\mathbb{P}^r}[U(X_T e^{-\int_0^T r_u\, du})].$$

In this way, the utility maximization under $\mathbb{P}^r$ with a zero interest rate becomes equivalent to the utility maximization problem under $\mathbb{P}$ with the interest rate process $(r_t)_{t\in[0,T]}$.

A practical implementation of the above idea depends on how explicit one can be about the Girsanov transformation associated with $\mathbb{P}^r$. It turns out, fortunately, that many of the widely used interest rate models such as Vasiček, CIR or the quadratic normal models (see e.g. [36, Chap. 7] for a textbook discussion of these models) allow a fully explicit description (often due to their affine structure). For example, in the Vasiček model, the Girsanov drift under $\mathbb{P}^r$ can be computed quite explicitly, due to the underlying affine structure. Indeed, suppose that r has Ornstein–Uhlenbeck dynamics of the form

$$dr_t = \kappa(\theta - r_t)\, dt + \beta\, dB_t, \quad r_0 \in \mathbb{R},$$

where B is a Brownian motion and $\kappa > 0$, $\theta, \beta \in \mathbb{R}$. Then the process

$$B^{(p)} := B - \int_0^{\cdot} b(T-t)\,dt, \quad \text{where } b(t) := \frac{\beta p}{\kappa}(1-e^{-\kappa t}),$$

is a $\mathbb{P}^r$-Brownian motion.

3.2 Optimal controls

We use Landau notation in the following sense. The notation $f \in O(\varepsilon^r)$ for $r \in \mathbb{N}$ means that there exists a constant $C \geq 0$ such that $|f(\varepsilon)| \leq C\varepsilon^r$ for all ε small. Therefore, the estimates (3.4) and (3.5) are of type $O(\varepsilon^2)$. To see this, we note that a slight adjustment to the proof below of Proposition 4.3 shows that the wealth process $\tilde{X} := \mathcal{E}(\int \hat{\pi}^{(0)} dR^{(\varepsilon)})$ satisfies (see (4.12))

$$|\mathbb{E}[U(\tilde{X}_T)] - u^{(0)}(1+\varepsilon p \Delta^{(0)})| \leq \frac{1}{2} p^2 \varepsilon^2 |u^{(0)}| \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\Phi^2 e^{\varepsilon |p| \Phi^-}].$$

Therefore, under the conditions of Theorem 3.2, $\hat{\pi}^{(0)}$ is an $O(\varepsilon^2)$-optimal control for the ε-model because the triangle inequality produces a constant $C > 0$ such that

$$|\mathbb{E}[U(\tilde{X}_T)] - u^{(\varepsilon)}| \leq C\varepsilon^2,$$

for all $\varepsilon > 0$ small enough. In this section, we provide a correction term to $\hat{\pi}^{(0)}$ such that the resulting wealth process upgrades the convergence to $o(\varepsilon^2)$.

For simplicity, we consider the (augmented) filtration generated by (B, W), where B and W are two independent Brownian motions with values in $\mathbb{R}$ and $\mathbb{R}^d$, $d \in \mathbb{N}$, respectively. The scalar-valued Brownian motion B drives the stock returns, while the presence of the multidimensional Brownian motion W allows model incompleteness. More specifically, in (2.1) we take

$$dM_t := \sigma_t dB_t, \quad M_0 := 0,$$

for a process $\sigma \in \mathcal{P}_B^2$ with $\sigma \neq 0$. We define $\tilde{\mathbb{P}}^{(0)}$ by (2.6) and denote by $(B^{\tilde{\mathbb{P}}^{(0)}}, W^{\tilde{\mathbb{P}}^{(0)}})$ the corresponding $\tilde{\mathbb{P}}^{(0)}$-Brownian motions. Provided that $\Phi := \int_0^T \hat{\pi}_t^{(0)} \lambda_t' \sigma_t^2 dt$ is in $\mathbb{L}^2(\tilde{\mathbb{P}}^{(0)})$, Φ has under $\tilde{\mathbb{P}}^{(0)}$ the unique representation

$$\Phi = \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\Phi] + \int_0^T \gamma_t^B \sigma_t dB_t^{\tilde{\mathbb{P}}^{(0)}} + \int_0^T \gamma_t^W dW_t^{\tilde{\mathbb{P}}^{(0)}}, \tag{3.6}$$

where we have used $\sigma \neq 0$. Because $\Phi \in \mathbb{L}^2(\tilde{\mathbb{P}}^{(0)})$, the two processes γ^W and γ^B in (3.6) satisfy the integrability conditions

$$\mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}\left[\int_0^T ((\gamma_t^B \sigma_t)^2 + (\gamma_t^W)^2) dt\right] < \infty.$$

These square-integrability properties are used in the proof of the next theorem.

Theorem 3.4 *In the above Brownian setting, we assume*

$$\int_0^T (\lambda_t')^2 \sigma_t^2 dt \in \mathbb{L}^{1-p}(\mathbb{P}) \cap \mathbb{L}^1(\tilde{\mathbb{P}}^{(0)}) \quad \textit{and} \quad \int_0^T \hat{\pi}_t^{(0)} \lambda_t' \sigma_t^2 dt \in \mathbb{L}^2(\tilde{\mathbb{P}}^{(0)}), \tag{3.7}$$

as well as the existence of a constant $\varepsilon_0 > 0$ *such that*

$$\mathbb{E}\Big[e^{\int_0^T (p(1-p)^2 \varepsilon \hat{\pi}^{(0)} \lambda' + \frac{1}{2}\varepsilon^2 p^2 \gamma^B ((p-2) p \gamma^B - 2\lambda'))\sigma^2 dt}\Big] < \infty \tag{3.8}$$

for all $\varepsilon \in (0, \varepsilon_0)$. *Then we have*

$$u^{(\varepsilon)} - u^{(0)} - \varepsilon p u^{(0)} \Delta^{(0)} - \frac{1}{2}\varepsilon^2 p u^{(0)} \big(\Delta^{(00)} + p(\Delta^{(0)})^2\big) = O(\varepsilon^3), \tag{3.9}$$

$$v^{(\varepsilon)} - v^{(0)} - \varepsilon q v^{(0)} \Delta^{(0)} - \frac{1}{2}\varepsilon^2 q v^{(0)} \big(\Delta^{(00)} + q(\Delta^{(0)})^2\big) = O(\varepsilon^3), \tag{3.10}$$

as $\varepsilon \searrow 0$. *In* (3.9) *and* (3.10), *we have defined*

$$\Delta^{(00)} := \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}} \left[\int_0^T \left(p|\gamma_t^W|^2 + \frac{(\lambda_t')^2 + p\gamma_t^B(\gamma_t^B + 2\lambda_t')}{1-p} \sigma_t^2 \right) dt \right], \tag{3.11}$$

where the processes γ^B *and* γ^W *are given by the representation* (3.6).

Remark 3.5 1. The proof of Theorem 3.4 below shows that the process

$$\tilde{\pi} := \hat{\pi}^{(0)} + \varepsilon \frac{\lambda' + p\gamma^B}{1-p} \tag{3.12}$$

is an $O(\varepsilon^3)$-optimal control for the ε-model in the sense that the wealth process $\tilde{X} := \mathcal{E}(\int \tilde{\pi}\, dR^{(\varepsilon)})$ satisfies

$$\mathbb{E}[U(\tilde{X}_T)] - u^{(\varepsilon)} = O(\varepsilon^3) \quad \text{as } \varepsilon \searrow 0.$$

2. Because the filtration is generated by (B, W), the orthogonal component $\hat{H}^{(0)}$ in the dual optimizer $Z^{(0)} \hat{H}^{(0)}$ for the dual problem (2.3) can be written in the form $\hat{H}^{(0)} = \mathcal{E}(-\hat{\nu}^{(0)} \cdot W)$ for a d-dimensional process $\hat{\nu}^{(0)}$ in $\mathcal{P}_W^2$. The proof of Theorem 3.4 below also shows that the process

$$\tilde{\nu} := \hat{\nu}^{(0)} - \varepsilon p \gamma^W \tag{3.13}$$

is an $O(\varepsilon^3)$-optimal dual control in the ε-model.

3. Throughout the paper, we have considered $\varepsilon = 0$ as the base model. Because we can write

$$\lambda + (\bar{\varepsilon} + \varepsilon)\lambda' = \lambda + \bar{\varepsilon}\lambda' + \varepsilon\lambda'$$

for any $\bar{\varepsilon} \in [\varepsilon_L, \varepsilon_U]$ with $\varepsilon_L < \varepsilon_U$, we can use Theorem 3.4 for the base model $\lambda + \bar{\varepsilon}\lambda'$ to provide a second-order Taylor expansion around any point $\bar{\varepsilon}$. Therefore, whenever

$\Delta^{(0)}$ and $\Delta^{(00)}$ are bounded uniformly in $\bar{\varepsilon} \in [\varepsilon_L, \varepsilon_U]$, Theorem 3 in [37] ensures that $u^{(\varepsilon)}$ is twice differentiable in ε.

4. As illustrated in the examples in Sect. 5.3 below, the exponential moment condition (3.8) can often be made to hold by imposing some smallness condition on either $T > 0$ or $\varepsilon_0 > 0$.

4 Proofs of the main theorems

We start the proof with a short discussion of the special structure that the dual domain $\mathcal{Y}^{(\varepsilon)}$ has when the stock price process $S^{(\varepsilon)} = \mathcal{E}(R^{(\varepsilon)})$ is continuous. Indeed, it has been shown in [31, Proposition 3.2] that in that case, the maximal elements in $\mathcal{Y}^{(\varepsilon)}$ (for the pointwise order) are precisely local martingales of the form

$$Y = Z^{(\varepsilon)} H, \quad H \in \mathcal{H},$$

where $\mathcal{H}$ denotes the set of all M-orthogonal strictly positive local martingales H with $H_0 = 1$. We remark that even though the results in [31] were written under the NFLVR assumption, the proof of [31, Proposition 3.2] does not use this condition. Hence we can write

$$v^{(\varepsilon)} = \inf_{H \in \mathcal{H}} \mathbb{E}[V(Z_T^{(\varepsilon)} H_T)],$$

and the minimizer $\hat{Y}^{(\varepsilon)}$ always has the form

$$\hat{Y}^{(\varepsilon)} = Z^{(\varepsilon)} \hat{H}^{(\varepsilon)} \quad \text{for some } \hat{H}^{(\varepsilon)} \in \mathcal{H}. \tag{4.1}$$

Finally, we introduce two shortcuts for expressions that appear frequently in the proof, namely

$$\eta := \int_0^T \lambda_t' \, dR_t^{(0)}, \qquad \Lambda := \int_0^T (\lambda_t')^2 \, d\langle M \rangle_t, \tag{4.2}$$

and remind the reader that $\Phi := \int_0^T \hat{\pi}_t^{(0)} \lambda_t' \, d\langle M \rangle_t$ and $\Delta^{(0)} := \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\eta]$. It will be useful to keep in mind that $p < 0$, $q \in (-1, 0)$, $(1-p)(1+q) = 1$, and that $-1/q$ and $1-p$ are conjugate exponents.

4.1 Proof of Theorem 3.1

The proof of Theorem 3.1 is based on the stability results of [31] and the following lemma.

Lemma 4.1 *Let $(K^{(\varepsilon)})_{\varepsilon \ge 0}$ be a family of positive random variables such that*

1. $\mathbb{E}[Z_T^{(\delta)} K^{(\varepsilon)}] \le 1$ *for all* $\varepsilon, \delta \ge 0$;
2. $K^{(\varepsilon)} \to K^{(0)}$ *in probability as* $\varepsilon \searrow 0$.

Then under the conditions of Theorem 3.1, we have

$$\lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\mathbb{E}\big[V(Z_T^{(\varepsilon)}K^{(\varepsilon)})-V(Z_T^{(0)}K^{(\varepsilon)})\big]=q\mathbb{E}[V(Z_T^{(0)}K^{(0)})\eta].$$

Proof The map $\varepsilon\mapsto Z_T^{(\varepsilon)}$ is almost surely continuously differentiable; indeed, we have

$$\log Z_T^{(\varepsilon)}=\log Z_T^{(0)}-\varepsilon\int_0^T\lambda'_t\,dR_t^{(0)}-\frac{1}{2}\varepsilon^2\int_0^T(\lambda'_t)^2\,d\langle M\rangle_t,$$

and so

$$\frac{d}{d\varepsilon}Z_T^{(\varepsilon)}=-Z_T^{(\varepsilon)}(\eta+\varepsilon\Lambda)\quad\text{a.s.}$$

Therefore, we have

$$V(Z_T^{(\varepsilon)}K)-V(Z_T^{(0)}K)=\int_0^\varepsilon qV(Z_T^{(\delta)}K)(\eta+\delta\Lambda)d\delta \tag{4.3}$$

for each ε and each positive random variable K. Thus,

$$V(Z_T^{(\varepsilon)}K^{(\varepsilon)})-V(Z_T^{(0)}K^{(\varepsilon)})-\varepsilon qV(Z_T^{(0)}K^{(0)})\eta=A_\varepsilon+B_\varepsilon,$$

where

$$\begin{aligned}A_\varepsilon&:=\int_0^\varepsilon q\big(V(Z_T^{(\delta)}K^{(\varepsilon)})-V(Z_T^{(0)}K^{(0)})\big)\eta\,d\delta,\\ B_\varepsilon&:=\int_0^\varepsilon qV(Z_T^{(\delta)}K^{(\varepsilon)})\Lambda\,\delta\,d\delta.\end{aligned} \tag{4.4}$$

Hölder's inequality implies that (recall that $q<0$)

$$\mathbb{E}[B_\varepsilon]\le\frac{1}{2}\varepsilon^2\sup_{\delta\in[0,\varepsilon]}(\mathbb{E}[Z_T^{(\delta)}K^{(\varepsilon)}]^{-q}\mathbb{E}[\Lambda^{1-p}]^{1+q})\le\frac{1}{2}\varepsilon^2\mathbb{E}[\Lambda^{1-p}]^{1+q}. \tag{4.5}$$

Thus we have $\frac{1}{\varepsilon}\mathbb{E}[B_\varepsilon]\to 0$ as $\varepsilon\searrow 0$. To show that $\frac{1}{\varepsilon}\mathbb{E}[A_\varepsilon]\to 0$, we claim that $\mathbb{E}[A_\varepsilon]=\int_0^\varepsilon f(\varepsilon,\delta)\,d\delta$, where the function $f:[0,\infty)^2\to\mathbb{R}$ is given by

$$f(\varepsilon,\delta):=q\mathbb{E}\big[\big(V(Z_T^{(\delta)}K^{(\varepsilon)})-V(Z_T^{(0)}K^{(0)})\big)\eta\big]. \tag{4.6}$$

This claim follows from Fubini's theorem which can be applied because Tonelli's theorem and Hölder's inequality produce

$$\mathbb{E}\left[\int_0^\varepsilon(Z_T^{(\delta)}K^{(\varepsilon)})^{-q}|\eta|d\delta\right]=\int_0^\varepsilon\mathbb{E}[(Z_T^{(\delta)}K^{(\varepsilon)})^{-q}|\eta|]d\delta\le\varepsilon\mathbb{E}[|\eta|^{1-p}]^{\frac{1}{1-p}}.$$

The expectation in the last term above is finite because $\eta \in \mathbb{L}^s$ for some $s > 1 - p$. Because $f(0,0) = 0$, it will be enough to show that f is continuous at $(0,0)$. By the assumptions of the lemma and the definition of $Z^{(\delta)}$, we have

$$V(Z_T^{(\delta_n)} K^{(\varepsilon_n)}) \longrightarrow V(Z_T^{(0)} K^{(0)}) \quad \text{in probability}$$

for each sequence $(\varepsilon_n, \delta_n) \in [0,\infty)^2$ such $(\varepsilon_n, \delta_n) \to (0,0)$. Therefore, it suffices to establish uniform integrability of the expression inside the expectation in (4.6). For that, we can use the theorem of de la Vallée-Poussin with exponent $r > 1$: For conjugate exponents (α, α') with $\alpha > 1$ and $\alpha' := \frac{\alpha}{\alpha - 1}$, Hölder's inequality produces

$$\mathbb{E}\big[(Z_T^{(\delta)} K^{(\varepsilon)})^{-qr} |\eta|^r\big] \le \mathbb{E}\big[(Z_T^{(\delta)} K^{(\varepsilon)})^{-qr\alpha}\big]^{\frac{1}{\alpha}} \mathbb{E}[|\eta|^{r\alpha'}]^{\frac{1}{\alpha'}}.$$

Because $s > 1 - p$, we can find $r > 1$ and $\alpha > 1$ such that

$$r\alpha = -\frac{1}{q}, \quad r\alpha' = s.$$

The conclusion then follows from the assumption that $\eta \in \mathbb{L}^s$. □

Proof of Theorem 3.1 Thanks to the optimality of $Z_T^{(\varepsilon)} \hat{H}_T^{(\varepsilon)}$, we have the upper estimate

$$\frac{1}{\varepsilon}\mathbb{E}\big[V(Z_T^{(\varepsilon)} \hat{H}_T^{(\varepsilon)}) - V(Z_T^{(0)} \hat{H}_T^{(0)})\big] \le \frac{1}{\varepsilon}\mathbb{E}\big[V(Z_T^{(\varepsilon)} \hat{H}_T^{(0)}) - V(Z_T^{(0)} \hat{H}_T^{(0)})\big]. \tag{4.7}$$

Similarly, by using the optimality of $Z_T^{(0)} \hat{H}_T^{(0)}$, we get the lower estimate

$$\frac{1}{\varepsilon}\mathbb{E}\big[V(Z_T^{(\varepsilon)} \hat{H}_T^{(\varepsilon)}) - V(Z_T^{(0)} \hat{H}_T^{(0)})\big] \ge \frac{1}{\varepsilon}\mathbb{E}\big[V(Z_T^{(\varepsilon)} \hat{H}_T^{(\varepsilon)}) - V(Z_T^{(0)} \hat{H}_T^{(\varepsilon)})\big]. \tag{4.8}$$

Our next task is to prove that the limits of the right-hand sides of (4.7) and (4.8) exist and both coincide with the right-hand side of (3.2). In each case, Lemma 4.1 can be applied; in the first with $K^{(\varepsilon)} = \hat{H}_T^{(0)}$, and in the second with $K^{(\varepsilon)} = \hat{H}_T^{(\varepsilon)}$. In both cases, assumption 1 of Lemma 4.1 follows directly from the fact that $Z^{(\varepsilon)} \hat{H}^{(0)} \in \mathcal{Y}^{(\varepsilon)}$ and $Z^{(\varepsilon)} \hat{H}^{(\varepsilon)} \in \mathcal{Y}^{(\varepsilon)}$. As for assumption 2, it trivially holds in the first case. In the second case, we need to argue that $\hat{H}_T^{(\varepsilon)} \to \hat{H}_T^{(0)}$ in probability as $\varepsilon \searrow 0$. That in turn follows easily from [31, Lemma 3.10]; as mentioned above, the stronger NFLVR assumption made in [31] is not necessary, and its results hold under the weaker NUBPR condition.

Having proved (3.2), we turn to (3.1). Thanks to (2.5), the conjugacy relationship (2.4) takes in our setting the simple form

$$pu^{(\varepsilon)} = (qv^{(\varepsilon)})^{1-p}. \tag{4.9}$$

Therefore, $u^{(\varepsilon)}$ is right-differentiable at $\varepsilon = 0$, and we have

$$p\frac{d}{d\varepsilon}u^{(\varepsilon)}\bigg|_{\varepsilon=0+} = (1-p)(qv^{(0)})^{-p} q^2 v^{(0)} \Delta^{(0)} = p^2 u^{(0)} \Delta^{(0)}.$$

□

4.2 Remaining proofs

Proposition 4.2 *Suppose that* $\eta \in \mathbb{L}^{2(1-p)}$ *and* $\Lambda, \Lambda\eta \in \mathbb{L}^{1-p}$. *Then for all* $\varepsilon \geq 0$, *we have*

$$v^{(\varepsilon)} - v^{(0)} - \varepsilon q v^{(0)} \Delta^{(0)} \leq \frac{1}{2} C_v \varepsilon^2 + \frac{1}{2} C'_v \varepsilon^3, \tag{4.10}$$

where $C_v = |q| \|\eta\|^{1/2}_{\mathbb{L}^{2(1-p)}} + \|\Lambda\|_{\mathbb{L}^{1-p}}$ *and* $C'_v = |q| \|\eta\Lambda\|_{\mathbb{L}^{1-p}}$.

Proof The upper estimate (4.7) and the representation (4.3) imply that

$$\mathbb{E}\big[V(Z_T^{(\varepsilon)} \hat{H}_T^{(\varepsilon)}) - V(Z_T^{(0)} \hat{H}_T^{(0)}) - \varepsilon q V(Z_T^{(0)} \hat{H}_T^{(0)})\eta\big] \leq \mathbb{E}[A_\varepsilon] + \mathbb{E}[B_\varepsilon],$$

where A_ε and B_ε are defined by (4.4), with $K^{(\varepsilon)} = K^{(0)} = \hat{H}_T^{(0)}$. As in (4.5), we have

$$\mathbb{E}[B_\varepsilon] \leq \frac{1}{2} \varepsilon^2 \|\Lambda\|_{\mathbb{L}^{1-p}}.$$

To deal with A_ε we note that its structure allows us to apply the representation from (4.3) once again to see that

$$\frac{1}{q^2} A_\varepsilon = \int_0^\varepsilon \int_0^\delta V(Z_T^{(\beta)} \hat{H}_T^{(0)}) \eta(\eta + \beta\Lambda)\, d\beta\, d\delta.$$

This in turn can be estimated via Hölder's inequality, as in (4.5), by

$$\mathbb{E}[A_\varepsilon] \leq \frac{1}{2} |q| \varepsilon^2 \sup_{\beta \in [0,\varepsilon]} \mathbb{E}\big[\big(\eta(\eta + \beta\Lambda)\big)^{1-p}\big]^{1+q} \leq \frac{1}{2} |q| \varepsilon^2 (\|\eta^2\|_{\mathbb{L}^{1-p}} + \varepsilon \|\eta\Lambda\|_{\mathbb{L}^{1-p}}),$$

producing the bound in (4.10). □

Unfortunately, the same idea cannot be applied to obtain a similar lower bound. Instead, we turn to the primal problem and establish a lower bound for it.

Proposition 4.3 *Let* $\Phi := \int_0^T \hat{\pi}_t^{(0)} \lambda'_t \, d\langle M \rangle_t$, *let* Λ *be defined by* (4.2) *and* $\tilde{\mathbb{P}}^{(0)}$ *by* (2.6). *Given* $\varepsilon_0 > 0$, *assume that* $\Lambda \in \mathbb{L}^{1-p}$ *and* $\Phi^2 e^{\varepsilon_0 |p| \Phi^-} \in \mathbb{L}^1(\tilde{\mathbb{P}}^{(0)})$. *Then*

$$u^{(\varepsilon)} - u^{(0)} - \varepsilon p u^{(0)} \Delta^{(0)} \geq -C_u(\varepsilon)\varepsilon^2 \quad \text{for } \varepsilon \in [0, \varepsilon_0],$$

where $C_u(\varepsilon) := \frac{1}{2} p^2 |u^{(0)}| \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\Phi^2 e^{\varepsilon |p| \Phi^-}]$.

Proof For $\tilde{X} := \mathcal{E}(\hat{\pi}^{(0)} \cdot R^{(\varepsilon)})$, we have $\tilde{X} \in \mathcal{X}^{(\varepsilon)}$ so that by optimality,

$$u^{(\varepsilon)} - u^{(0)} - p\varepsilon \mathbb{E}[U(\hat{X}_T^{(0)})\Phi] \geq \mathbb{E}[U(\tilde{X}_T) - U(\hat{X}_T^{(0)}) - p\varepsilon U(\hat{X}_T^{(0)})\Phi]. \tag{4.11}$$

Thanks to the form of $\tilde{X}$, the right-hand side of (4.11) can be written as $\mathbb{E}[U(\hat{X}_T^{(0)})D_\varepsilon]$, where $D_\varepsilon := \exp(p\varepsilon\Phi) - 1 - p\varepsilon\Phi = \int_0^\varepsilon \int_0^\delta p^2\Phi^2 e^{p\beta\Phi}\, d\beta\, d\delta$. Thus,

$$\begin{aligned}\mathbb{E}[U(\hat{X}_T^{(0)})D_\varepsilon] &= p^2 \int_0^\varepsilon \int_0^\delta \mathbb{E}[U(\hat{X}_T^{(0)})\Phi^2 e^{p\beta\Phi}]\, d\beta\, d\delta \\ &\geq \frac{1}{2} p^2\varepsilon^2 \mathbb{E}[U(\hat{X}_T^{(0)})\Phi^2 e^{\varepsilon|p|\Phi^-}]. \end{aligned} \tag{4.12}$$

Therefore $u^{(\varepsilon)} - u^{(0)} - \varepsilon p\mathbb{E}[U(\hat{X}_T^{(0)})\Phi] \geq -C_u(\varepsilon)\varepsilon^2$ for $\varepsilon \in [0, \varepsilon_0]$ with C_u as in the statement (here we use the change of measure in (2.6)).

It remains to show that $\mathbb{E}[U(\hat{X}_T^{(0)})\Phi] = \mathbb{E}[U(\hat{X}_T^{(0)})\eta]$, which is equivalent to showing $\mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\Phi] = \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\eta]$ by the definition of $\tilde{\mathbb{P}}^{(0)}$. We define the local $\tilde{\mathbb{P}}^{(0)}$-martingale $\tilde{M}^p$ by (2.7). Therefore, $N = \int_0^\cdot \lambda_t'\, d\tilde{M}_t^p$ is also a local martingale. The desired equality is therefore equivalent to the equality $\mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[N_T] = 0$ by the definition of η and Φ. In turn, it is sufficient to show that N is an $\mathcal{H}^2$-martingale under $\tilde{\mathbb{P}}^{(0)}$. Since $\langle N\rangle_T = \int_0^T (\lambda_t')^2\, d\langle M\rangle_t = \Lambda$, Hölder's inequality implies that

$$\mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\langle N\rangle_T] = (qv^{(0)})^{-1}\mathbb{E}[(\hat{Y}_T^{(0)})^{-q}\Lambda] \leq (qv^{(0)})^{-1}\mathbb{E}[\Lambda^{1-p}]^{1+q} < \infty. \qquad \square$$

Remark 4.4 If one is interested in an error estimate which does not feature the optimal portfolio $\hat{\pi}^{(0)}$ (through Φ), one can adopt an alternative approach in the proof (and the statement) of Proposition 4.3. More specifically, by using $\tilde{X} = \hat{X}^{(0)}\mathcal{E}(\varepsilon\lambda' \cdot R^{(\varepsilon)})$ as a test process (instead of $\mathcal{E}(\hat{\pi}^{(0)} \cdot R^{(\varepsilon)})$), one obtains a constant $C_u(\varepsilon)$ which depends only on the primal and dual optimizers $\hat{X}^{(0)}$ and $\hat{Y}^{(0)}$, in addition to λ', η and Λ.

Proof of Theorem 3.2 Two of the four inequalities in Theorem 3.2 have been established in Propositions 4.2 and 4.3. For the remaining two, we use the special form (4.9) of the conjugacy relationship between $u^{(\varepsilon)}$ and $v^{(\varepsilon)}$. Thanks to Proposition 4.3 and the positivity of $pu^{(\varepsilon)}$, $qv^{(\varepsilon)}$ and $1+q$, we have

$$q\big(v^{(\varepsilon)} - v^{(0)} - \varepsilon qv^{(0)}\Delta^{(0)}\big) = (pu^{(\varepsilon)})^{1+q} - (pu^{(0)})^{1+q} - \varepsilon q(pu^{(0)})^{1+q}\Delta^{(0)}.$$

The right-hand side above is further bounded from above, for ε in a (right) neighborhood of 0, by

$$F(\varepsilon) := (pu^{(0)} + \varepsilon pu^{(0)}\Delta^{(0)} - pC\varepsilon^2)^{1+q} - (pu^{(0)})^{1+q} - \varepsilon q(pu^{(0)})^{1+q}\Delta^{(0)},$$

where $C := C_u(\varepsilon_0)$ is from Proposition 4.3. Now F is a C^2-function in some neighborhood of 0 with $F(0) = F'(0) = 0$; hence on each compact subset of that neighborhood, it is bounded by a constant multiple of ε^2. In particular, we have

$$v^{(\varepsilon)} - v^{(0)} - \varepsilon qv^{(0)}\Delta^{(0)} \geq -C\varepsilon^2$$

for some $C > 0$ and for ε in some (right) neighborhood of 0. A similar argument, but based on Proposition 4.2, shows that (3.4) holds as well. $\square$

Proof of Theorem 3.4 The first part of (3.7) means that $\Lambda \in \mathbb{L}^{1-p}(\mathbb{P})$; hence the second half of the proof of Proposition 4.3 shows that $\mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\Phi] = \Delta^{(0)}$. Therefore the representation (3.6) can be written as

$$\Phi = \Delta^{(0)} + \int_0^T \gamma_t^B \sigma_t dB_t^{\tilde{\mathbb{P}}^{(0)}} + \int_0^T \gamma_t^W dW_t^{\tilde{\mathbb{P}}^{(0)}}.$$

Because the filtration is generated by the Brownian motions (B, W), we can find $\hat{\nu}^{(0)} \in \mathcal{P}_W^2$ such that the dual optimizer $\hat{H}^{(0)}$ can be represented as

$$\hat{H}^{(0)} = \mathcal{E}(-\hat{\nu}^{(0)} \cdot W).$$

Therefore, Girsanov's theorem ensures that under $\tilde{\mathbb{P}}^{(0)}$, the processes

$$dB_t^{\tilde{\mathbb{P}}^{(0)}} := dB_t + (\lambda_t - \hat{\pi}_t^{(0)})\sigma_t dt, \quad dW_t^{\tilde{\mathbb{P}}^{(0)}} := dW_t + \hat{\nu}_t^{(0)} dt \tag{4.13}$$

are independent Brownian motions.

We start with the primal problem and define $\tilde{\pi} := \hat{\pi}^{(0)} + \varepsilon\delta$ for $\varepsilon \in (0, \varepsilon_0)$ and $\delta := q\gamma^B + \frac{\lambda'}{1-p} \in \mathcal{P}_B^2$. Then we have

$$(\tilde{X})^p := \mathcal{E}(\tilde{\pi} \cdot R^{(\varepsilon)})^p = (\hat{X}^{(0)})^p e^{p\int(\varepsilon\hat{\pi}^{(0)}\lambda' + \varepsilon^2(\delta\lambda' - \frac{1}{2}\delta^2))\sigma^2 dt + p\varepsilon\int \delta\sigma dB^{\tilde{\mathbb{P}}^{(0)}}}.$$

To see that $\mathbb{E}[(\tilde{X}_T)^p] < \infty$, we apply Hölder's inequality twice with the exponents $-1/q$ and $1 - p$ to see that

$$\begin{aligned}
&\mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}\left[e^{p\int_0^T(\varepsilon\hat{\pi}^{(0)}\lambda' + \varepsilon^2(\delta\lambda' - \frac{1}{2}\delta^2))\sigma^2 dt + p\varepsilon\int_0^T \delta\sigma dB_t^{\tilde{\mathbb{P}}^{(0)}}}\right] \\
&= \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}\left[e^{p\int_0^T(\varepsilon\hat{\pi}^{(0)}\lambda' + \varepsilon^2(\delta\lambda' - \frac{1}{2}\delta^2))\sigma^2 dt + q\varepsilon\int_0^T(\lambda' + p\gamma^B)\sigma dB_t^{\tilde{\mathbb{P}}^{(0)}}}\right] \\
&\le \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}\left[e^{p(1-p)\int_0^T(\varepsilon\hat{\pi}^{(0)}\lambda' + \varepsilon^2(\delta\lambda' - \frac{1}{2}\delta^2))\sigma^2 dt - \frac{1}{2}q(1-p)\varepsilon^2\int_0^T(\lambda' + p\gamma^B)^2\sigma^2 dt}\right]^{\frac{1}{1-p}} \\
&= \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}\left[e^{p\int_0^T((1-p)(\varepsilon\hat{\pi}^{(0)}\lambda' + \varepsilon^2(\delta\lambda' - \frac{1}{2}\delta^2)) - \frac{1}{2}\varepsilon^2(\lambda' + p\gamma^B)^2)\sigma^2 dt}\right]^{\frac{1}{1-p}} \\
&\le \mathbb{E}\left[e^{p(1-p)\int_0^T((1-p)(\varepsilon\hat{\pi}^{(0)}\lambda' + \varepsilon^2(\delta\lambda' - \frac{1}{2}\delta^2)) - \frac{1}{2}\varepsilon^2(\lambda' + p\gamma^B)^2)\sigma^2 dt}\right]^{\frac{1}{(1-p)^2}}.
\end{aligned}$$

By inserting $\delta := q\gamma_B + \frac{\lambda'}{1-p}$, we see that the coefficient in front of $(\lambda')^2$ inside the exponent is $-\varepsilon p^2\sigma^2/2 < 0$. Consequently, by deleting the $(\lambda')^2$-term, we see that (3.8) is an upper bound for $\mathbb{E}[\tilde{X}_T^p]$.

For any random variable A with $\mathbb{E}[e^{\varepsilon_0 A}] < \infty$, the representation

$$e^x = 1 + x + \frac{1}{2}x^2 + \int\int\int_{[0,1]^3} tsx^3 e^{rstx} dr ds dt, \quad x \in \mathbb{R},$$

allows us to find a constant $C > 0$ such that

$$\mathbb{E}\left[\left|e^{\varepsilon A} - 1 - \varepsilon A - \frac{1}{2}\varepsilon^2 A^2\right|\right] \le \varepsilon^3 C \int\int_{[0,1]^2} ts\mathbb{E}[e^{\varepsilon_0 A}] ds dt$$

for all $\varepsilon \in (0, \varepsilon_0)$ sufficiently small. Consequently, we can find a function $C_u(\varepsilon)$ which is in $O(\varepsilon^3)$ such that

$$\begin{aligned}\mathbb{E}[U(\tilde{X}_T)] &= u^{(0)}\mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}\Big[e^{p\int_0^T(\varepsilon\hat{\pi}^{(0)}\lambda'+\varepsilon^2(\delta\lambda'-\frac{1}{2}\delta^2))\sigma^2dt+p\varepsilon\int_0^T\delta\sigma dB^{\tilde{\mathbb{P}}^{(0)}}}\Big]\\ &= u^{(0)}\bigg(1+p\varepsilon\Delta^{(0)}+\frac{1}{2}p\varepsilon^2\big(p(\Delta^{(0)})^2+\Delta^{(00)}\big)\bigg)+C_u(\varepsilon). \qquad (4.14)\end{aligned}$$

We then turn to the dual problem. For the perturbed dual control $\tilde{\nu} := \hat{\nu}^{(0)} - \varepsilon p\gamma^W$ in $\mathcal{P}^2_W$, we have

$$\begin{aligned}&\big(Z^{(\varepsilon)}\mathcal{E}(-\tilde{\nu}\cdot W)\big)^{-q}\\ &= e^{q\int(\lambda+\varepsilon\lambda')\sigma dB+q\int(\hat{\nu}^{(0)}-\varepsilon p\gamma^W)dW+q\frac{1}{2}\int((\lambda+\varepsilon\lambda')\sigma^2+|\hat{\nu}^{(0)}-\varepsilon q\gamma^W|^2)dt}\\ &= (Z^{(0)}\hat{H}^{(0)})^{-q}e^{\varepsilon q\int\lambda'\sigma dB^{\tilde{\mathbb{P}}^{(0)}}-\varepsilon qp\int\gamma^W dW^{\tilde{\mathbb{P}}^{(0)}}+q\frac{1}{2}\int(\varepsilon^2(\lambda')^2\sigma^2+\varepsilon^2p^2|\gamma^W|^2+2\varepsilon\lambda'\pi^{(0)}\sigma^2)dt}.\end{aligned}$$

Since $\tilde{\nu}$ is admissible in the ε-problem, we find

$$\begin{aligned}&v^{(\varepsilon)}\\ &\le \frac{1}{q}\mathbb{E}\big[\big(Z_T^{(\varepsilon)}\mathcal{E}(-\tilde{\nu}\cdot W)_T\big)^{-q}\big]\\ &= v^{(0)}\mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}\Big[e^{\varepsilon q\int_0^T\lambda'\sigma dB^{\tilde{\mathbb{P}}^{(0)}}-\varepsilon qp\int_0^T\gamma^W dW^{\tilde{\mathbb{P}}^{(0)}}+q\frac{1}{2}\int_0^T(\varepsilon^2(\lambda')^2\sigma^2+\varepsilon^2p^2|\gamma^W|^2+2\varepsilon\lambda'\pi^{(0)}\sigma^2)dt}\Big].\end{aligned}$$

Finiteness of $v^{(\varepsilon)}$ ensures that the $\tilde{\mathbb{P}}^{(0)}$-expectation appearing in the last line above is also finite. Indeed, because $q<0$, we have $v^{(\varepsilon)}, v^{(0)} \in (-\infty, 0)$, and so the $\tilde{\mathbb{P}}^{(0)}$-expectation is bounded from above by $\frac{v^{(\varepsilon)}}{v^{(0)}}$. As in the primal problem, this allows us to replace e^x with its Taylor series and in turn implies that we can find a function $C_v(\varepsilon) \in O(\varepsilon^3)$ such that

$$\begin{aligned}&v^{(0)}\mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}\Big[e^{\varepsilon q\int_0^T\lambda'\sigma dB^{\tilde{\mathbb{P}}^{(0)}}-\varepsilon qp\int_0^T\gamma^W dW^{\tilde{\mathbb{P}}^{(0)}}+q\frac{1}{2}\int_0^T(\varepsilon^2(\lambda')^2\sigma^2+\varepsilon^2p^2|\gamma^W|^2+2\varepsilon\lambda'\pi^{(0)}\sigma^2)dt}\Big]\\ &= v^{(0)}\bigg(1+q\varepsilon\Delta^{(0)}+\frac{1}{2}q\varepsilon^2\big(q(\Delta^{(0)})^2+\Delta^{(00)}\big)\bigg)+C_v(\varepsilon).\end{aligned}$$

By combining this estimate and (4.14) with the primal–dual relation (4.9), we find

$$\begin{aligned}&u^{(0)}\bigg(1+p\varepsilon\Delta^{(0)}+\frac{1}{2}p\varepsilon^2\big(p(\Delta^{(0)})^2+\Delta^{(00)}\big)\bigg)+C_u(\varepsilon)\\ &\le u^{(\epsilon)}\\ &= \frac{1}{p}(qv^{(\epsilon)})^{1-p}\\ &\le \frac{1}{p}\bigg(qv^{(0)}\Big(1+q\epsilon\Delta^{(0)}+\frac{1}{2}q\epsilon^2\big(q(\Delta^{(0)})^2+\Delta^{(00)}\big)+C_v(\epsilon)\Big)\bigg)^{1-p}. \qquad (4.15)\end{aligned}$$

The function $x \mapsto x^{1-p}$ is real analytic on $(0, \infty)$. Therefore, the fact that $C_v \in O(\varepsilon^3)$ ensures that the last line of (4.15) agrees with the first line of (4.15) up to $O(\varepsilon^3)$-terms. This establishes (3.9). A similar argument produces (3.10). □

5 Examples

5.1 First examples

We start this section with a short list of trivial and extreme cases. They are not here to illustrate the power of our main results, but simply to help the reader understand them better. They also tell a similar qualitative story: loosely speaking, the improvement in the utility (on a log-scale) is proportional both to the base market price of risk process and to the size of the deviation. Locally around λ, the value function of the utility maximization problem—parametrized by the market price of risk process $\tilde{\lambda}$—is well approximated by an exponential function of the form

$$u(\tilde{\lambda}) \approx u(\lambda) e^{\langle \tilde{\lambda}-\lambda, \hat{\pi}^{(0)} \rangle_0}, \quad \text{with } \langle \rho, \pi \rangle_0 = \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}} \left[\int_0^T \rho_t \pi_t \, dt \right], \tag{5.1}$$

where $u(\tilde{\lambda})$ and $u(\lambda)$ denote the values of the utility maximization problems with market price of risk processes $\tilde{\lambda}$ and λ, respectively.

Example 5.1 (Small market price of risk) Suppose that $\lambda \equiv 0$ so that we can think of $S^{(\varepsilon)}$ as the stock price in a market with a "small" market price of risk. Since $Z^{(0)} \equiv 1$, it is clearly the dual optimizer at $\varepsilon = 0$ and we have $\hat{\pi}^{(0)} \equiv 0$. Consequently, under the assumptions of Theorem 3.2, we have $\tilde{\mathbb{P}}^{(0)} = \mathbb{P}$ and

$$\Delta^{(0)} = \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}} \left[\int_0^T \lambda'_t \, dM_t \right] = 0.$$

It follows that

$$u^{(\varepsilon)} = u^{(0)} + O(\varepsilon^2), \qquad v^{(\varepsilon)} = v^{(0)} + O(\varepsilon^2),$$

and the effects of $\varepsilon\lambda'$ are felt only to the second order, regardless of the risk-aversion coefficient $p < 0$.

Example 5.2 (Deviations from the Black–Scholes model) Suppose that $M = B$ is an $\mathbb{F}$-Brownian motion and $\lambda \neq 0$ is a constant process (we also use λ for the value of the constant). In that case, it is classical that the dual minimizer in the base market is $Z^{(0)} = \mathcal{E}(-\lambda B)$, and consequently that $\frac{d\tilde{\mathbb{P}}^{(0)}}{d\mathbb{P}} = \mathcal{E}(q\lambda B)$. It follows that

$$\Delta^{(0)} = \frac{\lambda}{1-p} \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}} \left[\int_0^T \lambda'_t \, dt \right].$$

As we shall see below, this form is especially convenient for computations.

Example 5.3 (Uniform deviations) Another special case where it is particularly easy to compute the (logarithmic) derivative $\Delta^{(0)}$ is when the perturbation λ' is a constant process (whose value is also denoted by λ'). Indeed, in that case,

$$\Delta^{(0)} = \lambda' \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}} \left[\int_0^T \hat{\pi}_t^{(0)} \, dt \right]. \tag{5.2}$$

It is especially instructive to consider the case where the base model is the Black–Scholes model since everything becomes explicit: the optimal portfolio is given by the Merton proportion $\hat{\pi}_t^{(0)} = \lambda/(1-p)$, and the values $u^{(0)}$ and $v^{(0)}$ are given by

$$pu^{(0)} = \exp\left(\frac{1}{2} q \lambda^2 T\right), \qquad qv^{(0)} = \exp\left(\frac{1}{2} \frac{q}{1-p} \lambda^2 T\right).$$

Using (5.2) or by performing a straightforward direct computation, we easily get

$$p\Delta^{(0)} = q\lambda'\lambda T,$$

making the approximation in (5.1) exact.

5.2 The Kim–Omberg model

The Kim–Omberg model (see [23]) is one of the most widely used models for the market price of risk process. Because it allows explicit expressions for all quantities involved in CRRA utility maximization, it serves as an excellent test case for the practical implementation of our main results. Appendix A contains all technical details.

We assume that $\mathbb{F}$ is the augmentation of the filtration generated by two independent one-dimensional Brownian motions B and W and define λ^{KO} to be the Ornstein–Uhlenbeck process given by

$$d\lambda_t^{\mathrm{KO}} = \kappa(\theta - \lambda_t^{\mathrm{KO}})dt + \beta dB_t + \gamma dW_t, \quad \lambda_0^{\mathrm{KO}} \in \mathbb{R}, \tag{5.3}$$

where κ, θ, β and γ are constants. We define the volatility $M_t := B_t$ in what follows. The closed-form expressions for the primal value function and the primal optimizer can be found in Theorem A.1 in Appendix A (the result is proved in [23]).

To illustrate our approximation, we think of the Kim–Omberg model as a perturbation of a base model. As base model, we consider a model with “totally unhedgeable coefficients” (see [22, Example 6.7.4]), namely

$$d\lambda_t = \kappa(\theta - \lambda_t)dt + \gamma \, dW_t, \quad \lambda_0 = \lambda_0^{\mathrm{KO}} \in \mathbb{R}. \tag{5.4}$$

In this way, $\lambda^{\mathrm{KO}} = \lambda + \varepsilon\lambda'$, where $\varepsilon := \beta$ and

$$d\lambda'_t = -\kappa\lambda'_t \, dt + dB_t, \quad \lambda'_0 = 0. \tag{5.5}$$

Based on Theorem A.1, Lemma A.2 in Appendix A provides a coupled system of ODEs for smooth deterministic functions $C_j : [0, T] \to \mathbb{R}$, $C_j(T) = 0$, such that the integrands appearing in the representation (3.6) of $\Phi := \int_0^T \pi_t^{(0)} \lambda_t' dt$ are given by

$$\gamma_t^B = \frac{1}{p-1}\big(C_2(t) + C_6(t)\lambda_t\big), \tag{5.6}$$

$$\gamma_t^W = \frac{\gamma}{p-1}\big(C_4(t) + 2C_5(t)\lambda_t + C_6(t)\lambda_t'\big). \tag{5.7}$$

5.2.1 *Exact computations*

The proof of Lemma A.2 in Appendix A shows that

$$\Delta^{(0)} := \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}\left[\int_0^T \lambda_s' \hat{\pi}_s^{(0)} ds\right] = \frac{1}{p-1}\big(C_1(0) + C_4(0)\lambda_0 + C_5(0)\lambda_0^2\big).$$

This relation, a similar one (whose exact form and derivation we omit) for the second-order term $\Delta^{(00)}$ of (3.11), and the availability of an exact expression for the value function allow an efficient numerical computation of the zeroth-, first- and second-order approximation, and their comparison with the exact values. The model parameters used in Table 1 below are the calibrated model parameters for the market portfolio reported in Sect. 4.2 in [32] (we ignore the constant interest rate and constant volatility used in Sect. 4 in [32]). Moreover, we use negative values of ε because the empirical covariation between the excess return and the stock's return is typically negative (see e.g. the discussion in [32, Sect. 4.2]).

Instead of hard-to-interpret expected utility values, we report their certainty equivalents (i.e., their compositions with the function CE $:= U^{-1}$; see Remark 3.3, 1). We set

$$\delta^{(0)} := u^{(0)} + \varepsilon p u^{(0)} \Delta^{(0)},$$

$$\delta^{(00)} := u^{(0)} + \varepsilon p u^{(0)} \Delta^{(0)} + \frac{\varepsilon^2}{2} p u^{(0)} \big(\Delta^{(00)} + p(\Delta^{(0)})^2\big).$$

This gives the following table:

Table 1 Certainty equivalents for the zeroth-, first- and second-order approximations and the exact values in the Kim–Omberg model with $\beta := \varepsilon$ and unit initial wealth. The model parameters used are $\gamma = 0.04395$, $\kappa = 0.0404$, $\theta = 0.117$, $p = -1$ and $T = 10$

ε	λ_0	CE($u^{(0)}$)	CE($\delta^{(0)}$)	CE($\delta^{(00)}$)	CE($u^{(\varepsilon)}$)
−0.01	0.1	1.046	1.047	1.048	1.048
−0.05	0.1	1.046	1.054	1.081	1.084
−0.10	0.1	1.046	1.063	1.181	1.206
−0.01	0.5	1.614	1.647	1.648	1.649
−0.05	0.5	1.614	1.794	1.850	1.846
−0.10	0.5	1.614	2.020	2.339	2.272

5.2.2 Monte Carlo-based computations

One of the advantages of our approach is that it lends itself easily to computational methods based on Monte Carlo (MC) simulation. For the Kim–Omberg model, we use the standard explicit Euler scheme from MC simulation to compute the involved quantities of interest. In other words, we do not rely on the availability of exact expressions for the value functions or the correction terms $\Delta^{(0)}$ and $\Delta^{(00)}$.

For a portfolio π and the model perturbation parameter ε, we define the constant $\mathrm{CE}^{(\varepsilon)}(\pi) \in (0, \infty)$ uniquely by

$$U\big(\mathrm{CE}^{(\varepsilon)}(\pi)\big) = \mathbb{E}\big[U\big(\mathcal{E}(\pi \cdot R^{(\varepsilon)})_T\big)\big].$$

In other words, $\mathrm{CE}^{(\varepsilon)}(\pi)$ is the dollar amount whose utility value matches that of the expected utility an investor would obtain in the ε-model when using the strategy π. We remind the reader that $\hat{\pi}^{(0)}$ denotes the optimizer in the base ($\varepsilon = 0$) model, $\tilde{\pi}^{(\varepsilon)}$ is the second-order improvement (as in (3.12) above) of $\hat{\pi}^{(0)}$, and $\hat{\pi}^{(\varepsilon)}$ is the exact optimizer in the ε-model. Both quantities $\mathrm{CE}^{(\varepsilon)}(\hat{\pi}^{(0)})$ and $\mathrm{CE}^{(\varepsilon)}(\tilde{\pi}^{(\varepsilon)})$ serve as lower bounds for the exact value $\mathrm{CE}(u^{(\varepsilon)})$. The second one, which we also denote by

$$\mathrm{LB} := \mathrm{CE}^{(\varepsilon)}(\tilde{\pi}^{(\varepsilon)}),$$

is second-order optimal and appears in our simulations. To obtain a corresponding upper bound, we simulate the dynamics of the dual process, based on (4.9) and the second-order optimal dual control $\tilde{\nu}$ defined by (3.13). We define

$$\mathrm{UB} := U^{-1}\bigg(\frac{1}{p}\mathbb{E}\big[\big(Z_T^{(\varepsilon)}\mathcal{E}(-\tilde{\nu} \cdot W)_T\big)^{-q}\big]^{1-p}\bigg).$$

To quantify the simulation errors, we report the 95%-confidence intervals based on MC-simulated values of $\mathrm{CE}^{(\varepsilon)}(\hat{\pi}^{(0)})$, LB and UB in Table 2 below. The value $\mathrm{CE}(u^{(\varepsilon)})$, computed without MC simulation and included for comparison only, is exact to 3 decimal places.

Table 2 95%-confidence intervals for certainty equivalents for the upper and lower bounds as well as the base model optimizer $\hat{\pi}^{(0)}$ for the Kim–Omberg model. The true exact values for the ε-model are included in the last column for comparison. Except for the last column, the numbers are based on MC simulation using an Euler scheme with one million paths each with time-step size 0.001. The model parameters are the same as in Table 1

ε	λ_0	$\mathrm{CE}^{(\varepsilon)}(\hat{\pi}^{(0)})$	LB	UB	$\mathrm{CE}(u^{(\varepsilon)})$
−0.01	0.10	[1.047, 1.048]	[1.048, 1.049]	[1.048, 1.049]	1.048
−0.05	0.10	[1.052, 1.053]	[1.083, 1.084]	[1.083, 1.085]	1.084
−0.10	0.10	[1.057, 1.058]	[1.200, 1.201]	[1.204, 1.208]	1.206
−0.01	0.50	[1.644, 1.649]	[1.647, 1.653]	[1.646, 1.657]	1.649
−0.05	0.50	[1.760, 1.764]	[1.844, 1.850]	[1.843, 1.857]	1.846
−0.10	0.50	[1.868, 1.871]	[2.248, 2.256]	[2.266, 2.286]	2.272

In Table 2, we note the significant difference between the performance of the base-model optimizer $\tilde{\pi}^{(0)}$ and its second-order improvement $\tilde{\pi}^{(\varepsilon)}$, especially for larger values of ε. Furthermore, the lower and upper bounds appear to be quite tight.

5.3 Extended affine models

The two models considered in this section do not have closed-form expressions for the value functions u and v. We consider models belonging to the class of extended affine specifications of the market price of risk process developed in [8]. The second example is based on a high-dimensional underlying Markov process which prevents PDE methods from being applicable. Appendix B contains all technical details. In particular, as we explain in Appendix B, in these affine models, the integrability conditions (3.7) and (3.8) hold provided that $T > 0$ is sufficiently small.

5.3.1 One-dimensional extended affine model

As in the Kim–Omberg model above, we let the augmented filtration be generated by two independent Brownian motions B and W. A central role is played by the Feller process F given by

$$dF_t = \kappa(\theta - F_t)dt + \sqrt{F_t}\,(\beta dB_t + \gamma dW_t), \quad F_0 > 0,$$

where κ, θ, β, and γ are strictly positive constants such that the (strict) Feller condition $2\kappa\theta > \beta^2 + \gamma^2$ holds. This ensures in particular that F is strictly positive on $[0, T]$ (in Appendix B, we show that the strict Feller condition also ensures that (3.7) holds whenever $T > 0$ is small). Unlike in the Kim–Omberg model, the appropriate volatility normalization turns out to be $\sqrt{F_t}$; that is, we define

$$M_s := \int_0^s \sqrt{F_t}\, dB_t.$$

A particular extended affine specification for the market price of risk process which was considered in [8] is given by

$$\lambda_t^{\mathrm{CFK}} := \frac{\varepsilon}{F_t} + 1,$$

where ε is a (positive or negative) constant. Unless $\varepsilon = 0$, there is currently no known closed-form solution to the corresponding optimal investment problem.[1] However, for $\varepsilon = 0$, the resulting model is covered by the analysis in [25] (see Theorem B.1 in Appendix B which is from [25]). Therefore, we choose the constant market price of risk process

$$\lambda_t := 1 \tag{5.8}$$

[1] Theorem 4.5 in [13] expresses the corresponding value function as an infinite sum of weighted generalized Laguerre polynomials.

Table 3 95%-confidence intervals for certainty equivalents for the upper and lower bounds as well as the base model optimizer $\hat{\pi}^{(0)}$ for the extended affine model. The parameter values are $\kappa = 5$, $\theta = 0.0169$, $\beta = -0.1$, $\gamma = 0.1744$, $p = -1$ and $T = 10$. The numbers are based on MC simulation using an Euler scheme with one million paths each with time-step size 0.001

ε	F_0	$\mathrm{CE}^{(\varepsilon)}(\hat{\pi}^{(0)})$	LB	UB	$\mathrm{CE}^{(0)}(\hat{\pi}^{(0)})$
0.10	0.01	[1.724, 1.726]	[10.159, 10.399]	[10.226, 10.481]	1.043
0.05	0.01	[1.342, 1.343]	[2.141, 2.151]	[2.131, 2.149]	1.043
0.01	0.01	[1.097, 1.098]	[1.118, 1.119]	[1.117, 1.120]	1.043
0.10	0.05	[1.728, 1.729]	[9.660, 9.877]	[9.766, 10.000]	1.045
0.05	0.05	[1.344, 1.345]	[2.105, 2.115]	[2.102, 2.120]	1.045
0.01	0.05	[1.099, 1.100]	[1.119, 1.121]	[1.117, 1.121]	1.045

for the base model, whereas we define the perturbation process λ' by

$$\lambda'_t := \frac{1}{F_t}.$$

Table 3 is the analogue of Table 2 for the extended affine model with parameters taken from Fig. 4 in [32, Sect. 3.3]. The methodology and the simulated quantities are the same as for Table 2.

Perhaps even more than in the Kim–Omberg model, the numbers in Table 3 above illustrate the superiority of the second-order approximations (columns 4 and 5) over its first-order version (column 3) as well as the zeroth order values (column 6). Again, the bounds in Table 3 appear quite tight when compared to the first-order approximations for moderate values of ε.

5.3.2 High-dimensional extended affine model

We now let the augmented filtration be generated by three independent Brownian motions $(W^{(1)}, W^{(2)}, W^{(3)})$ (in the notation of Sect. 3.2, we have $B := W^{(2)}$ and $W := (W^{(1)}, W^{(3)})$). We consider the class $A_1(3)$ from Appendix B of [8], where the central role is played by the three processes

$$dF_t^{(1)} = (a_1 + b_{11}F_t^{(1)})dt + \sqrt{F_t^{(1)}}\,dW_t^{(1)}, \qquad F_0^{(1)} > 0,$$

$$dF_t^{(2)} = (a_2 + b_{21}F_t^{(1)} + b_{22}F_t^{(2)} + b_{23}F_t^{(3)})dt + \sqrt{1 + F_t^{(1)}}\,dW_t^{(2)}, \qquad F_0^{(2)} \in \mathbb{R},$$

$$dF_t^{(3)} = (a_3 + b_{31}F_t^{(1)} + b_{32}F_t^{(2)} + b_{33}F_t^{(3)})dt + \sqrt{1 + F_t^{(1)}}\,dW_t^{(3)}, \qquad F_0^{(3)} \in \mathbb{R}.$$

In these dynamics, the constants a_i and b_{ij} satisfy the Feller condition $2a_1 \geq 1$, but are otherwise arbitrary. We use from Appendix B of [8] the model specification

$$\lambda_t^{\mathrm{CFK}} := \frac{\lambda_{20} + \lambda_{21}F_t^{(1)} + \lambda_{22}F_t^{(2)} + \lambda_{23}F_t^{(3)}}{1 + F_t^{(1)}}, \qquad M := \int_0^\cdot \sqrt{1 + F_t^{(1)}}\,dW_t^{(2)}, \tag{5.9}$$

Table 4 95%-confidence intervals for certainty equivalents for the upper and lower bounds as well as the base model optimizer $\hat{\pi}^{(0)}$ for the model (5.10), (5.11). The factor process parameter values are taken from Table 8 in Sect. 5 in [8], $p = -1$ and $T = 10$. The numbers are based on MC simulation using an Euler scheme with one million paths each with time-step size 0.001

ε	$(F_0^{(i)})_{i=1}^3$	$\mathrm{CE}^{(\varepsilon)}(\hat{\pi}^{(0)})$	LB	UB	$\mathrm{CE}^{(0)}(\hat{\pi}^{(0)})$
0.05	0.01	[2.420, 2.427]	[6.600, 6.646]	[6, 682, 6.802]	2.313
0.01	0.01	[2.382, 2.392]	[2.474, 2.485]	[2.471, 2.495]	2.313
0.005	0.01	[2.347, 2.356]	[2.371, 2.382]	[2.361, 2.383]	2.313
0.05	0.05	[2.423, 2.429]	[6.618, 6.664]	[6.776, 6.896]	2.315
0.01	0.05	[2.382, 2.391]	[2.479, 2.491]	[2.474, 2.489]	2.315
0.005	0.05	[2.354, 2.364]	[2.370, 2.381]	[2.376, 2.398]	2.315

where $\lambda_{20}, \ldots, \lambda_{23}$ are some constants (and $\sigma := \sqrt{1 + F^{(1)}}$ in the notation of Sect. 3.2). In general, there is no closed-form solution available for the power investor's optimal investment problem for this model. However, the following submodel of (5.9) is covered by the closed-form expressions provided in [25] (constant market price of risk process):

$$\lambda_t := \lambda_{21}, \tag{5.10}$$

with the corresponding volatility M as in (5.9). Therefore, we choose the model (5.10) for the base model, whereas we define the perturbation process by

$$\lambda'_t := \frac{\lambda_{20} - \lambda_{21} + \lambda_{22} F_t^{(2)} + \lambda_{23} F_t^{(3)}}{1 + F_t^{(1)}}. \tag{5.11}$$

Theorem B.1 in Appendix B is from [25] and provides the primal and dual optimizers for the base model (5.10) in closed form (we note that (5.10) is another example of a model with completely unhedgeable coefficients).

Table 4 is the analogue of Tables 2 and 3 for the model (5.10) and (5.11) with parameter coefficients taken from Table 8 in [8, Sect. 5].

Acknowledgements The authors would like to thank the anonymous referees, the anonymous Associate Editor, the Editor Martin Schweizer, Milica Čudina, Claus Munk, Mihai Sîrbu, and Kim Weston for useful comments.

During the preparation of this work, the first author has been supported by the National Science Foundation under Grant No. DMS-1411809 (2014–2017). The second author has been supported by the National Science Foundation under grant No. DMS-1600307 (2015–2018). The third author has been supported by the National Science Foundation under Grant No. DMS-1107465 (2012–2017) and Grant No. DMS-1516165 (2015–2018). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

Appendix A: Details on the Kim–Omberg model

The following result summarizes the main properties in [23].

Theorem A.1 (Kim and Omberg 1996) *Let the market price of risk process be defined by* (5.3), $M := B$, *and let* $p < 0$. *Then there exist continuously differentiable functions* $b, c : [0, \infty) \to \mathbb{R}$ *such that for* $t \in [0, T)$, *we have*

$$\begin{aligned} b'(t) &= -\alpha_4 b(t) - \alpha_1 c(t) + \alpha_2 b(t)c(t), \qquad & b(T) &= 0, \\ c'(t) &= q - 2\alpha_4 c(t) + \alpha_2 c^2(t), \qquad & c(T) &= 0, \end{aligned}$$

where $\alpha_1 := \theta\kappa$, $\alpha_2 := (1+q)\beta^2 + \gamma^2$, $\alpha_3 := \beta^2 + \gamma^2$ *and* $\alpha_4 := q\beta - \kappa$. *Furthermore, the primal optimizer is given by*

$$\hat{\pi}_t^{KO} = \frac{b(t)\beta + (c(t)\beta - 1)\lambda_t^{KO}}{p-1}, \qquad t \in [0, T].$$

For $p < 0$, the above Riccati equation describing c has a "normal non-exploding solution" as defined in the Appendix of [23]. Therefore, all three functions a, b and c are bounded on any finite subinterval $[0, T]$ of $[0, \infty)$.

Lemma A.2 *Let* (λ, λ') *be defined by* (5.4) *and* (5.5) *and let* $p < 0$. *For the basis model* λ, *the primal and dual optimizers are given by*

$$\hat{\pi}_t^{(0)} = \frac{\lambda_t}{1-p}, \qquad \hat{\nu}_t^{(0)} = \gamma\big(b(t) + c(t)\lambda_t\big), \qquad t \in [0, T].$$

Furthermore, the processes (γ^B, γ^W) *appearing in the representation* (3.6) *of* Φ *are given by* (5.6) *and* (5.7), *where the functions* C_j *satisfy the ODEs*

$$\begin{aligned} -C_1'(t) &= \tilde{b}(t)\, C_4(t) + \gamma^2 C_5(t), & C_1(T) &= 0, \\ -C_2'(t) &= \tilde{b}(t)\, C_6(t) - \kappa\, C_2(t), & C_2(T) &= 0, \\ -C_4'(t) &= q\, C_2(t) - \tilde{c}(t)\, C_4(t) + 2\tilde{b}(t)\, C_5(t), & C_4(T) &= 0, \\ -C_5'(t) &= q\, C_6(t) - 2\tilde{c}(t)\, C_5(t), & C_5(T) &= 0, \\ -C_6'(t) &= -\big(\kappa + \tilde{c}(t)\big)\, C_6(t) - 1, & C_6(T) &= 0 \end{aligned}$$

on $[0, T)$, *with* (a, b, c) *as in Theorem* A.1 *with* $\beta := 0$, $\tilde{b}(t) := \kappa\theta - \gamma^2 b(t)$ *and* $\tilde{c}(t) := \kappa + \gamma^2 c(t)$. *Furthermore, for the measure* $\tilde{\mathbb{P}}^{(0)}$ *defined by* (2.6) *and for all* $T > 0$, *we have*

$$\Delta^{(0)} := \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}\left[\int_0^T \lambda_s' \hat{\pi}_s^{(0)} ds\right] = -\frac{1}{1-p}\big(C_1(0) + C_4(0)\lambda_0 + C_5(0)\lambda_0^2\big).$$

Proof The first part follows from Theorem A.1 applied to the case $\beta := 0$. To find the representation (3.6), we define the function

$$f(t, x, \lambda) := \frac{x^p}{p} e^{-a(t) - b(t)\lambda - \frac{1}{2}c(t)\lambda^2}, \qquad t \in [0, T], x > 0, \lambda \in \mathbb{R},$$

where the functions (b, c) are as in Theorem A.1 and a is given by

$$a'(t) = -\alpha_1 b(t) - \frac{1}{2}\alpha_3 c(t) + \frac{1}{2}\alpha_2 b^2(t), \qquad a(T) = 0.$$

The martingale properties of $(f(t, \hat{X}_t^{(0)}, \lambda_t))$ and $(\hat{X}_t^{(0)}\hat{Y}_t^{(0)})$ as well as the proportionality property $(\hat{X}_T^{(0)})^p \propto \hat{X}_T^{(0)}\hat{Y}_T^{(0)}$ produce

$$pf(t, \hat{X}_t^{(0)}, \lambda_t) = p\mathbb{E}[f(T, \hat{X}_T^{(0)}, \lambda_T)|\mathcal{F}_t] \propto \mathbb{E}[\hat{Y}_T^{(0)}\hat{X}_T^{(0)}|\mathcal{F}_t] = \hat{X}_t^{(0)}\hat{Y}_t^{(0)}.$$

By computing the dynamics of the left-hand side, we see from Girsanov's theorem that the two processes (see (4.13))

$$\begin{aligned} dB_t^{\tilde{\mathbb{P}}^{(0)}} &= -q\lambda_t dt + dB_t, \\ dW_t^{\tilde{\mathbb{P}}^{(0)}} &= \big(b(t) + c(t)\lambda_t\big)\gamma dt + dW_t, \end{aligned}$$

are independent Brownian motions under $\tilde{\mathbb{P}}^{(0)}$. These dynamics and Itô's lemma ensure that

$$N_t := \int_0^t \lambda_s' \lambda_s ds - C_1(t) - C_2(t)\lambda_t' - C_4(t)\lambda_t - C_5(t)\lambda_t^2 - C_6(t)\lambda_t\lambda_t'$$

is a $\tilde{\mathbb{P}}^{(0)}$-local martingale. Because the processes (λ, λ') remain Ornstein–Uhlenbeck processes under $\tilde{\mathbb{P}}^{(0)}$ and the functions $C_1, \ldots, C_6$ are bounded, N is indeed a $\tilde{\mathbb{P}}^{(0)}$-martingale. Furthermore, thanks to the zero terminal conditions imposed on $C_1, \ldots, C_6$, we see that

$$\Phi = \frac{1}{1-p}\int_0^T \lambda_t \lambda_t' dt = \frac{1}{1-p} N_T = \frac{1}{1-p} N_0 + \int_0^T \gamma_t^B dB_t^{\tilde{\mathbb{P}}^{(0)}} + \int_0^T \gamma_t^W dW_t^{\tilde{\mathbb{P}}^{(0)}}$$

for (γ^B, γ^W) defined by (5.6) and (5.7). □

Appendix B: Details on the extended affine models

The following result is from [25].

Theorem B.1 (Kraft 2005) *Let $p < 0$ and consider the model (5.8) with $\lambda_t := 1$. The primal optimal control is $\hat{\pi}_t^{(0)} = \frac{f(t)-1}{p-1}$, the dual optimal control corresponding to W is $\hat{\nu}_t^{(0)} = f(t)\sqrt{F_t}$, and the deterministic function f is given by*

$$\begin{aligned} f'(t) &= \frac{-f(t)(f(t)(\beta^2 + \gamma^2) + 2\kappa) + p(f(t)^2\gamma^2 - 1 + 2f(t)(\beta + \kappa))}{2(p-1)}, \\ f(T) &= 0. \end{aligned}$$

For the model (5.10) *with* $\lambda_t := \lambda_{21}$, *the primal optimal control is* $\hat{\pi}_t^{(0)} = \frac{\lambda_{21}}{1-p}$, *the dual optimal controls corresponding to* $(W^{(1)}, W^{(3)})$ *are*

$$\hat{\nu}_t^{(0)} = \Big(f(t)\sqrt{F_t^{(1)}}, 0\Big),$$

and the deterministic function f *is given by*

$$f'(t) = \frac{1}{2}\bigg(-2b_{11}f(t) + f(t)^2 + \frac{\lambda_{21}^2 p}{1-p}\bigg), \qquad f(T) = 0.$$

We then turn to the representation (3.6) of $\Phi := \int_0^T \hat{\pi}_t^{(0)}\lambda'_t\sigma_t^2 dt$. For the model (5.8), this is trivial because $\int_0^T \hat{\pi}_t^{(0)} dt$ is deterministic (implying that $\gamma^B = \gamma^W = 0$). The next lemma provides the representation for the model in (5.11).

Lemma B.2 *Let* $p < 0$, *let* (λ, λ') *be defined by* (5.10) *and* (5.11), *and define* $\hat{\pi}_t^{(0)} := \frac{\lambda_{21}}{1-p}$. *The integrands* (γ^B, γ^W) *appearing in the representation* (3.6) *of*

$$\Phi := \int_0^T \hat{\pi}_t^{(0)}(\lambda_{20} - \lambda_{21} + \lambda_{22}F_t^{(2)} + \lambda_{23}F_t^{(3)})dt$$

are given by

$$\gamma_t^B = -C_2(t), \quad \gamma_t^W = -\Big(C_1(t)\sqrt{F_t^{(1)}}, C_3(t)\sqrt{1 + F_t^{(1)}}\Big),$$

where the functions C_j *satisfy on* $[0, T]$ *the ODEs*

$$\begin{aligned}
&C_1'(t) = -b_{11}C_1 - b_{31}C_3(t) + C_1(t)f(t) - C_2(t)(b_{21} - \lambda_{21} + \hat{\pi}_t^{(0)}),\\
&C_1(T) = 0,\\
&C_2'(t) = -b_{22}C_2(t) - b_{32}C_3(t) + \lambda_{22}\hat{\pi}_t^{(0)},\\
&C_2(T) = 0,\\
&C_3'(t) = -b_{23}C_2(t) - b_{33}C_3(t) + \lambda_{23}\hat{\pi}_t^{(0)},\\
&C_3(T) = 0.
\end{aligned}$$

Proof This is similar to the proof of Lemma A.2. Because the functions C_j satisfy the above ODEs, we have that

$$\begin{aligned}
N_t := &\int_0^t \hat{\pi}_u^{(0)}(\lambda_{20} - \lambda_{21} + \lambda_{22}F_u^{(2)} + \lambda_{23}F_u^{(3)})du\\
&- C_0(t) - C_1(t)F_t^{(1)} - C_2(t)F_t^{(2)} - C_3(t)F_t^{(3)}
\end{aligned}$$

is a $\tilde{\mathbb{P}}^{(0)}$-martingale (here we use the $\tilde{\mathbb{P}}^{(0)}$-Brownian motions defined in (4.13)). The function C_0 is given by

$$C_0'(t) = -a_1 C_1(t) - a_2 C_2(t) - a_3 C_3(t) + C_2(t)\lambda_{21} - \big(C_2(t) - \lambda_{20} + \lambda_{21}\big)\hat{\pi}_t^{(0)}$$

with $C_0(T) = 0$. Then we have

$$\mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[\Phi|\mathcal{F}_t] = \mathbb{E}^{\tilde{\mathbb{P}}^{(0)}}[N_T|\mathcal{F}_t] = N_t.$$

Hence, we find the integrands appearing in the representation (3.6) of Φ to be as in the statement. □

Finally, we verify the integrability conditions (3.7) and (3.8) in Theorem 3.4, provided that $T > 0$ is small. We can define the second-order optimal controls $(\tilde{\pi}, \tilde{\nu})$ by (3.12) and (3.13). We start by considering the model (5.8). Theorem 3.1 in [17] states that the joint Laplace transform of

$$\Lambda := \int_0^T \frac{1}{F_s}\,ds, \qquad Q := \int_0^T F_s\,ds$$

is finite for all $T > 0$ in some neighborhood of 0 as soon as the strict Feller condition $2\kappa\theta > 1$ holds. This implies that both Λ and Q have some finite exponential positive moments (and consequently, Λ and Q have all moments). Because the integrand γ^B appearing in the representation (3.6) of Φ is zero in this model, the integrand appearing inside the exponential in (3.8) is deterministic. Consequently, (3.8) holds for all $\varepsilon > 0$.

For the model (5.11), we have the estimate

$$\int_0^T (\lambda_t')^2(1 + F_t^{(1)})dt \le \int_0^T (\lambda_{20} - \lambda_{21} + \lambda_{22}F_t^{(2)} + \lambda_{23}F_t^{(3)})^2 dt,$$

which shows that the moment requirements in (3.7) hold. Furthermore, by inserting λ' defined in (5.11) and $\sigma_t^2 = 1 + F_t^{(1)}$ into the exponent in (3.8), we see that this exponent is affine in $F^{(1)}$, $F^{(2)}$ and $F^{(3)}$ (with time-dependent coefficients). Consequently, the expectation in (3.8) is finite for small positive values of ε.

References

1. Altarovici, A., Muhle-Karbe, J., Soner, H.M.: Asymptotics for fixed transaction costs. Finance Stoch. **19**, 363–414 (2015)
2. Bick, B., Kraft, H., Munk, C.: Solving constrained consumption-investment problems by simulation of artificial market strategies. Manag. Sci. **59**, 485–503 (2013)
3. Brandt, M.W., Goyal, A., Santa-Clara, P., Stroud, J.R.: A simulation approach to dynamic portfolio choice with an application to learning about return predictability. Rev. Financ. Stud. **18**, 831–873 (2005)
4. Campbell, J.Y.: Intertemporal asset pricing without consumption data. Am. Econ. Rev. **LXXXIII**, 487–512 (1993)
5. Campbell, J.Y., Viceira, L.M.: Consumption and portfolio decisions when expected returns are time varying. Q. J. Econ. **114**, 433–495 (1999)

6. Cannarsa, P., Sinestrari, C.: Semiconcave Functions, Hamilton–Jacobi Equations, and Optimal Control. Progress in Nonlinear Differential Equations and their Applications, vol. 58. Birkhäuser Boston, Boston (2004)
7. Chacko, G., Viceira, L.M.: Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. Rev. Financ. Stud. **18**, 1369–1402 (2005)
8. Cheridito, P., Filipović, D., Kimmel, R.L.: Market price of risk specifications for affine models: theory and evidence. J. Financ. Econ. **83**, 123–170 (2007)
9. Cox, J.C., Huang, C.F.: Optimal consumption and portfolio policies when asset prices follow a diffusion process. J. Econ. Theory **49**, 33–83 (1989)
10. Cvitanić, J., Goukasian, L., Zapatero, F.: Monte Carlo computation of optimal portfolios in complete markets. J. Econ. Dyn. Control **27**, 971–986 (2003)
11. Davis, M.H.A.: Optimal hedging with basis risk. In: Kabanov, Yu., et al. (eds.) From Stochastic Calculus to Mathematical Finance. The Shiryaev Festschrift, pp. 169–187. Springer, Berlin (2006)
12. Detemple, J.B., Garcia, R., Rindisbacher, M.: A Monte Carlo method for optimal portfolio. J. Finance **LVIII**, 401–446 (2003)
13. Guasoni, P., Robertson, S.: Static fund separation of long-term investments. Math. Finance **25**, 789–826 (2015)
14. Haugh, M.B., Kogan, L., Wang, J.: Evaluating portfolio policies: a duality approach. Oper. Res. **54**, 405–418 (2006)
15. Henderson, V.: Valuation of claims on nontraded assets using utility maximization. Math. Finance **12**, 351–373 (2002)
16. Herrmann, S., Muhle-Karbe, J., Seifried, F.T.: Hedging with small uncertainty aversion. Finance Stoch. **21**, 1–64 (2017)
17. Hurd, T., Kuznetsov, A.: Explicit formulas for Laplace transforms of stochastic integrals. Markov Process. Relat. Fields **14**, 277–290 (2008)
18. Kallsen, J., Muhle-Karbe, J.: Option pricing and hedging with small transaction costs. Math. Finance **25**, 702–723 (2015)
19. Karatzas, I., Kardaras, C.: The numéraire portfolio in semimartingale financial models. Finance Stoch. **11**, 447–493 (2007)
20. Karatzas, I., Lehoczky, J.P., Shreve, S.E.: Optimal portfolio and consumption decisions for a "small" investor on a finite horizon. SIAM J. Control Optim. **25**, 1557–1586 (1987)
21. Karatzas, I., Lehoczky, J.P., Shreve, S.E., Xu, G.L.: Martingale and duality methods for utility maximization in an incomplete market. SIAM J. Control Optim. **29**, 702–730 (1991)
22. Karatzas, I., Shreve, S.E.: Methods of Mathematical Finance, 1st edn. Applications of Mathematics (New York), vol. 39. Springer, New York (1998)
23. Kim, T.S., Omberg, E.: Dynamic nonmyopic portfolio behavior. Rev. Financ. Stud. **9**, 141–161 (1996)
24. Kogan, L., Uppal, R.: Risk aversion and optimal policies in partial and general equilibrium economies. Working paper (2001). Available online at http://www.nber.org/papers/w8609
25. Kraft, H.: Optimal portfolios and Heston's stochastic volatility model: an explicit solution for power utility. Quant. Finance **5**, 303–313 (2005)
26. Kraft, H., Seiferling, T., Seifried, F.T.: Optimal consumption and investment with Epstein–Zin recursive utility. Finance Stoch. **21**, 187–226 (2017)
27. Kramkov, D., Schachermayer, W.: The asymptotic elasticity of utility functions and optimal investment in incomplete markets. Ann. Appl. Probab. **9**, 904–950 (1999)
28. Kramkov, D., Sîrbu, M.: On the two-times differentiability of the value functions in the problem of optimal investment in incomplete markets. Ann. Appl. Probab. **16**, 1352–1384 (2006)
29. Kramkov, D., Sîrbu, M.: Sensitivity analysis of utility-based prices and risk-tolerance wealth processes. Ann. Appl. Probab. **16**, 2140–2194 (2006)
30. Larsen, K.: A note on the existence of the power investor's optimizer. Finance Stoch. **15**, 183–190 (2011)
31. Larsen, K., Žitković, G.: Stability of utility-maximization in incomplete markets. Stoch. Process. Appl. **117**, 1642–1662 (2007)
32. Larsen, L., Munk, C.: The costs of suboptimal dynamic asset allocation: general results and applications to interest rate risk, stock volatility risk, and growth/value tilts. J. Econ. Dyn. Control **36**, 266–293 (2012)
33. Liu, J.: Portfolio selection in stochastic environments. Rev. Financ. Stud. **20**, 1–39 (2007)
34. Monoyios, M.: Malliavin calculus method for asymptotic expansion of dual control problems. SIAM J. Financ. Math. **4**, 884–915 (2013)

35. Mostovyi, O.: Necessary and sufficient conditions in the problem of optimal investment with intermediate consumption. Finance Stoch. **19**, 135–159 (2015)
36. Munk, C.: Financial Asset Pricing Theory, vol. 6. Oxford University Press, London (2013)
37. Oliver, H.W.: The exact Peano derivative. Trans. Am. Math. Soc. **76**, 444–456 (1954)
38. Rogers, L.C.G.: The relaxed investor and parameter uncertainty. Finance Stoch. **5**, 131–154 (2001)
39. Szpiro, G.G.: Measuring risk aversion: an alternative approach. Rev. Econ. Stat. **68**, 156–159 (1986)
40. Wachter, J.: Portfolio and consumption decisions under mean-reverting returns: an exact solution for complete markets. J. Financ. Quant. Anal. **37**, 63–91 (2002)