Journal of
Risk and Financial Management

MDPI

Article

Almost Perfect Shadow Prices

Eberhard Mayerhofer

Department of Mathematics & Statistics, University of Limerick, V94 T9PX Limerick, Ireland;
eberhard.mayerhofer@ul.ie

Abstract: Shadow prices simplify the derivation of optimal trading strategies in markets with transaction costs by transferring optimization into a more tractable, frictionless market. This paper establishes that a naïve shadow price ansatz for maximizing long-term returns, given average volatility yields a strategy that is, for small bid–ask spreads, asymptotically optimal at the third order. Considering the second-order impact of transaction costs, such a strategy is essentially optimal. However, for risk aversion different from one, we devise alternative strategies that outperform the shadow market at the fourth order. Finally, it is shown that the risk-neutral objective rules out the existence of shadow prices.

Keywords: transaction costs; portfolio choice; shadow prices; reflected diffusions

MSC: 91G10; 91G80

1. Introduction

With a little help of my friends.[1]

—The Beatles

Shadow prices (or *consistent price systems*) are a tool for characterizing the absence of arbitrage opportunities in markets with proportional transaction costs (see, for example, Czichowsky and Schachermayer (2016); Guasoni et al. (2010); Kabanov et al. (2002)), or for deriving optimal strategies for various objectives (see, for example, Guasoni and Muhle-Karbe (2013); Gerhold et al. (2013, 2014); Herdegen et al. (2023); Kallsen and Muhle-Karbe (2010)). This paper investigates the applicability of shadow prices to the optimization of long-term returns given average volatility.

Strategies that are optimal in frictionless markets[2] such as the delta-hedging of European-type options, or constant proportion strategies, lead to immediate bankruptcy under proportional costs.[3] To ensure solvency, trading frequency needs to be modulated to finite variation, trading as little as necessary to stay close to the target exposures. This paper relates to the objectives of long-run investors (Gerhold et al. (2014); Guasoni and Mayerhofer (2019, 2023); Taksar et al. (1988))[4], who consider it optimal to keep the fraction of wealth π invested in the risky asset within an interval around a target exposure by engaging only in trading whenever this fraction hits the boundaries of the interval.[5] For example, for constant investment opportunities and a sufficiently small relative bid–ask spread ε, the trading boundaries $\pi_- < \pi_+$ of an investor with risk-aversion γ are approximately

$$\pi_\pm = \pi_* \pm \left(\frac{3}{4\gamma}\pi_*^2(\pi_* - 1)^2\right)^{1/3} \varepsilon^{1/3}, \tag{1}$$

where $\pi_* = \frac{\mu}{\gamma\sigma^2}$ is the well-known Merton fraction, μ being the annualized average return of the risky asset, and σ its volatility.

Absurdly, such finite variation strategies are mathematically more challenging than the infinite variation strategies typically encountered in frictionless markets. Shadow prices

check for updates

Citation: Mayerhofer, Eberhard. 2024. Almost Perfect Shadow Prices. *Journal of Risk and Financial Management* 17: 70. https://doi.org/10.3390/jrfm17020070

Academic Editors: Svetlozar (Zari) Rachev and W. Brent Lindquist

Received: 13 December 2023
Revised: 3 February 2024
Accepted: 8 February 2024
Published: 10 February 2024

allow to transfer optimization into a more tractable, frictionless (but fictitious) market. A shadow price is a frictionless asset that evolves in the bid–ask spread of the risky asset and for which the optimal strategy buys (respectively sells) whenever its price agrees with the ask (respectively bid) of this risky asset. For objectives which are monotone functions of wealth, such as power utility, the strategy in the shadow market is also optimal in the original market because by trading in the shadow market, the investor is generally better off. Furthermore, shadow markets provide an elegant, intuitive derivation of optimal trading policies for different objectives and market models. It therefore may come as surprise that Guasoni and Mayerhofer (2019, 2023) use the more traditional Hamilton–Jacobi–Bellman equations both for the heuristic derivation of the candidate optimal control limit policy (with asymptotics (1)) *and* the verification of optimality. This is even more surprising, as the respective objectives lend themselves to very tractable candidate shadow prices and trading strategies (see Section 3.1 below). However, the local mean variance criterion is, in general, not monotone in wealth.[6] Therefore, the verification of optimality fails, leaving open the question of whether trading strategies derived in the shadow market are also optimal in the original market with transaction costs.

Guasoni and Mayerhofer (2019) show that, in the presence of transaction costs, maximizing returns is well posed, even without controlling for volatility—transaction costs act as a penalty in the objective. As a consequence, the efficient frontier is not a straight line as in the classical Merton problem but reaches a maximum for finite volatility, after which taking on even further risk may result in negative alpha. However, in frictionless markets, such an objective gives the incentive to seek arbitrary leverage, unless the asset has zero expected excess return. Thus, shadow prices are destined to fail as an optimization tool.

Nevertheless, a candidate shadow price can be found for a risk-averse investor. A construction similar to Gerhold et al. (2013) yields trading policies of the form (1), and thus, they are indistinguishable at the first order from the optimal one. Moreover, at the second order, they are distinguished by a mere change of sign in the second-order coefficient. Even more surprisingly, the equivalent safe rate of the shadow price trading strategy agrees at the third order with the maximum. In view of the second-order impact of transaction costs, it is essentially optimal. However, we devise trading policies that strictly outperform the shadow price trading strategy at fourth order.

Program of Paper

The paper is structured as follows: Section 2 presents the market model, encompassing a risky Black–Scholes asset with transaction costs, the mean–variance objective, and a recap of the optimal trading policies established in Guasoni and Mayerhofer (2019). Section 2.1 introduces control limit policies, evaluating their long-run performance along with small-transaction cost asymptotics (Lemma 3). In Section 3.1, a naïve ansatz for a shadow price is proposed, and asymptotic expansions of the trading boundaries are provided. Theorem 2 demonstrates their third-order asymptotic optimality and Theorem 3 establishes their strict sub-optimality. Section 3.3 provides a rigorous proof that for maximizing expected returns without controlling for volatility, no shadow price exists (Theorem 4). The final Section 4 summarizes our findings and points out directions for future research. The appendix computes a high-order approximation of the candidate shadow price to support the proof of Theorem 3.

2. Materials and Methods

The market model is comprised of two assets: a safe safe asset that is continuously compounded at a constant rate of $r \geq 0$ and a risky asset S purchased at its ask price S_t and satisfying the dynamics

$$\frac{dS_t}{S_t} = (\mu + r)dt + \sigma dB_t, \quad S_0, \sigma, \mu > 0, \tag{2}$$

where B is a standard Brownian motion. The risky asset's bid (selling) price is $(1-\varepsilon)S_t$, which implies a constant relative bid–ask spread of $\varepsilon > 0$, or, equivalently, constant proportional transaction costs.

Let w be the wealth associated with a self-financing trading strategy[7]. The mean–variance trade-off is captured by maximizing the equivalent safe rate,

$$\mathrm{ESR} := \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T \frac{dw_t}{w_t} - \frac{\gamma}{2}\left\langle \int_0^\cdot \frac{dw_t}{w_t}\right\rangle_T\right]. \tag{3}$$

With π, the proportion of wealth invested in the risky asset, and with φ_t, the number of shares $\varphi_t = \varphi_t^\uparrow - \varphi_t^\downarrow$ being the difference of purchases $\varphi_t^\uparrow$ minus sales $\varphi_t^\downarrow$, one can rewrite the objective[8] as follows

$$\mathrm{ESR} := r + \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T \left(\mu\pi_t - \frac{\gamma\sigma^2}{2}\pi_t^2\right)dt - \varepsilon\int_0^T \pi_t \frac{d\varphi_t^\downarrow}{\varphi_t}\right]. \tag{4}$$

In the absence of transaction costs ($\varepsilon = 0$), the objective is maximized by the constant proportion portfolio $\pi_* := \frac{\mu}{\gamma\sigma^2}$ dating back to Markowitz and Merton. The risk-neutral objective $\gamma = 0$ reduces to the average annualized return over a long horizon, which is well posed for transaction costs (Guasoni and Mayerhofer 2019, Theorem 3.2) but meaningless in the traditional framework with zero bid–ask spread, where a strategy can be arbitrarily levered. The case $\gamma = 1$ reduces to logarithmic utility, which is solved by the Taksar et al. (1988) for the unlevered case $\frac{\mu}{\gamma\sigma^2} < 1$.

An optimal strategy maximizing the equivalent safe rate exists. The following is a shortened version of (Guasoni and Mayerhofer 2019, Theorem 3.1), characterizing optimality.

Theorem 1. *Let* $\frac{\mu}{\gamma\sigma^2} \neq 1$.

1. *For any* $\gamma > 0$, *there exists* $\varepsilon_0 > 0$ *such that for all* $\varepsilon < \varepsilon_0$, *there is a unique solution* (W, ζ_-, ζ_+), *with* $\zeta_- < \zeta_+$ *for the free boundary problem*

$$\tfrac{1}{2}\sigma^2\zeta^2 W''(\zeta) + (\sigma^2+\mu)\zeta W'(\zeta) + \mu W(\zeta) - \tfrac{1}{(1+\zeta)^2}\left(\mu - \gamma\sigma^2\tfrac{\zeta}{1+\zeta}\right) = 0, \tag{5}$$

$$W(\zeta_-) = 0 \tag{6}$$

$$W'(\zeta_-) = 0, \tag{7}$$

$$W(\zeta_+) = \tfrac{\varepsilon}{(1+\zeta_+)(1+(1-\varepsilon)\zeta_+)}, \tag{8}$$

$$W'(\zeta_+) = \tfrac{\varepsilon(\varepsilon-2(1-\varepsilon)\zeta_+-2)}{(1+\zeta_+)^2(1+(1-\varepsilon)\zeta_+)^2} \tag{9}$$

2. *The trading strategy that buys at* $\pi_- := \zeta_-/(1+\zeta_-)$ *and sells at* $\pi_+ := \zeta_+/(1+\zeta_+)$ *as little as possible to keep the risky weight* π_t *within the interval* $[\pi_-, \pi_+]$ *is optimal.*
3. *The maximum performance is*

$$\max_{\varphi\in\Phi}\lim_{T\to\infty}\tfrac{1}{T}\mathbb{E}\left[\int_0^T\left(\mu\pi_t - \tfrac{\gamma\sigma^2}{2}\pi_t^2\right)dt - \varepsilon\int_0^T \pi_t\tfrac{d\varphi_t^\downarrow}{\varphi_t}\right] = \mu\pi_- - \tfrac{\gamma\sigma^2}{2}\pi_-^2, \tag{10}$$

where Φ *is the set of admissible strategies in Definition 1 below,* $\varphi_t = \pi_t w_t/S_t$ *is the number of shares held at time* t, *and* $\varphi_t^\downarrow$ *is the cumulative number of shares sold up to time* t.

4. *The trading boundaries* π_- *and* π_+ *have the asymptotic expansions*

$$\pi_\pm = \pi_* \pm \left(\tfrac{3}{4\gamma}\pi_*^2(\pi_*-1)^2\right)^{1/3}\varepsilon^{1/3} - \tfrac{(1-\gamma)\pi_*}{\gamma}\left(\tfrac{\gamma\pi_*(\pi_*-1)}{6}\right)^{1/3}\varepsilon^{2/3} + O(\varepsilon). \tag{11}$$

5. *The equivalent safe rate (*ESR*) has the expansion*

$$\mathrm{ESR} = r + \frac{\gamma\sigma^2}{2}\pi_*^2 - \frac{\gamma\sigma^2}{2}\left(\frac{3}{4\gamma}\pi_*^2(\pi_* - 1)^2\right)^{2/3}\varepsilon^{2/3} + O(\varepsilon). \tag{12}$$

2.1. Admissible Strategies and Their Long-Run Performance

In view of transaction costs, only finite-variation trading strategies are consistent with solvency. This is illustrated by the following example.

Example 1. *Consider the dynamic hedging part of* $1/\varepsilon$ *variance swaps*[9] *on the asset* S *with maturity* $T = 2$*, that requires to hold*

$$\varphi_t = \frac{1}{\varepsilon S_t}$$

units of the underlying at each time $t \geq 0$*. Trading discretely, along a mesh of size* Δ*, one needs to sell at* $t + \Delta$ *if and only if* $S_{t+\Delta} > S_t$*, which incurs a cost of*

$$\varepsilon \times 1/\varepsilon \times S_{t+\Delta}(1/S_t - 1/S_{t+\Delta}) = (S_{t+\Delta}/S_t - 1).$$

Let $x_+ = \max(0, x)$ *and* $N = T/\Delta$*, then the total transaction cost amounts to*

$$C_N = \sum_{i=0}^{N-1}(S_{(i+1)\Delta}/S_{i\Delta} - 1)_+.$$

Note that this sum counts all positive simple returns of the asset, which can be approximated by logarithmic returns. Thus, as $N \to \infty$*,* $C_N \to C$*, the semivariation of a Brownian motion* B *with drift,*

$$C = \lim_{\Delta t \to 0}\sum_{i=0}^{T/\Delta t - 1}(B_{(i+1)\Delta t} - B_{i\Delta t})_+ = \infty,$$

almost surely. This shows that, under proportional transaction costs, such a dynamic trading strategy results in immediate bankruptcy.

Denote by X_t and Y_t the wealth in the safe and risky positions, respectively, and by $(\varphi_t^\uparrow)_{t\geq 0}$ and $(\varphi_t^\downarrow)_{t\geq 0}$, the cumulative number of shares bought and sold, respectively. The self-financing condition prescribes that (X, Y) satisfies the dynamics

$$dX_t = rX_t dt - S_t d\varphi_t^\uparrow + (1-\varepsilon)S_t d\varphi_t^\downarrow, \quad dY_t = S_t d\varphi_t^\uparrow - S_t d\varphi_t^\downarrow + \varphi_t dS_t. \tag{13}$$

A strategy is admissible if it is non-anticipative and solvent, up to a small increase in the spread.

Definition 1. *Let* $x > 0$ *(the initial capital) and let* $(\varphi_t^\uparrow)_{t\geq 0}$ *and* $(\varphi_t^\downarrow)_{t\geq 0}$ *be continuous, increasing processes, adapted to the augmented natural filtration of* B*. Then,* $(x, \ \varphi_t = \varphi_t^\uparrow - \varphi_t^\downarrow)$ *is an* admissible trading strategy *if the following apply:*

1. *Its liquidation value is strictly positive at all times: there exists* $\varepsilon' > \varepsilon$ *such that the discounted asset* $\widetilde{S}_t := e^{-rt}S_t$ *satisfies*

$$x - \int_0^t \widetilde{S}_s d\varphi_s + \widetilde{S}_t\varphi_t - \varepsilon'\int_0^t \widetilde{S}_s d\varphi_s^\downarrow - \varepsilon'\varphi_t^+\widetilde{S}_t > 0 \qquad \textit{a.s. for all } t \geq 0. \tag{14}$$

2. *The following integrability conditions hold:*

$$\mathbb{E}\left[\int_0^t |\pi_u|^2 du\right] < \infty, \quad \mathbb{E}\left[\int_0^t \pi_u \frac{d\|\varphi_u\|}{\varphi_u}\right] < \infty \quad \textit{for all } t \geq 0, \tag{15}$$

where $\|\varphi_t\|$ denotes the total variation of φ on $[0, t]$.
The family of admissible trading strategies is denoted by Φ.

The following lemma describes the dynamics of the wealth process w_t, the risky weight π_t, and the risky-safe ratio ζ_t.

Lemma 1. *For any admissible trading strategy φ:*

$$\frac{d\zeta_t}{\zeta_t} = \mu dt + \sigma dB_t + (1+\zeta_t)\frac{d\varphi_t^{\uparrow}}{\varphi_t} - (1+(1-\varepsilon)\zeta_t)\frac{d\varphi_t^{\downarrow}}{\varphi_t}, \tag{16}$$

$$\frac{dw_t}{w_t} = rdt + \pi_t(\mu dt + \sigma dB_t - \varepsilon\frac{d\varphi_t^{\downarrow}}{\varphi_t}), \tag{17}$$

$$\frac{d\pi_t}{\pi_t} = (1-\pi_t)(\mu dt + \sigma dB_t) - \pi_t(1-\pi_t)\sigma^2 dt + \frac{d\varphi_t^{\uparrow}}{\varphi_t} - (1-\varepsilon\pi_t)\frac{d\varphi_t^{\downarrow}}{\varphi_t}. \tag{18}$$

For any such strategy, the functional

$$F_T(\varphi) := \frac{1}{T}\mathbb{E}\left[\int_0^T \frac{dw_t}{w_t} - \frac{\gamma}{2}\left\langle\int_0^T \frac{dw_t}{w_t}\right\rangle_T\right] \tag{19}$$

can be rewritten as

$$F_T(\varphi) = r + \frac{1}{T}\mathbb{E}\left[\int_0^T \left(\mu\pi_t - \frac{\gamma\sigma^2}{2}\pi_t^2\right)dt - \varepsilon\int_0^T \pi_t\frac{d\varphi_t^{\downarrow}}{\varphi_t}\right]. \tag{20}$$

Proof. See (Guasoni and Mayerhofer 2019, Lemma A.2). □

Lemma 2. *Let $\eta_- < \eta_+$ be such that either $\eta_+ < -1/(1-\varepsilon)$ or $\eta_- > 0$. Then, there exists an admissible trading strategy $\hat{\varphi}$ such that the risky-safe ratio η_t satisfies SDE (16). Moreover, $(\eta_t, \hat{\varphi}_t^{\uparrow}, \hat{\varphi}_t^{\downarrow})$ is a reflected diffusion on the interval $[\eta_-, \eta_+]$. In particular, η_t has stationary density equal to*

$$\nu(\eta) := \frac{\frac{2\mu}{\sigma^2} - 1}{\eta_+^{\frac{2\mu}{\sigma^2}-1} - \eta_-^{\frac{2\mu}{\sigma^2}-1}}\eta^{\frac{2\mu}{\sigma^2}-2}, \quad \eta \in [\eta_-, \eta_+], \tag{21}$$

when $\eta_- > 0$, and otherwise equals

$$\nu(\eta) := \frac{\frac{2\mu}{\sigma^2} - 1}{|\eta_-|^{\frac{2\mu}{\sigma^2}-1} - |\eta_+|^{\frac{2\mu}{\sigma^2}-1}}|\eta|^{\frac{2\mu}{\sigma^2}-2}, \quad \eta \in [\eta_-, \eta_+]. \tag{22}$$

Proof. See (Guasoni and Mayerhofer 2019, Lemma B.5). □

Definition 2. *For the rest of the paper, the strategy in Lemma 2 is called "control limit policy for $\eta_\pm$", an adaption of the name of similar policies in Taksar et al. (1988), where "limit" actually relates to the boundaries of the interval $[\eta_-, \eta_+]$. Note that the strategy in Theorem 1 (2) is exactly of this kind: it entails no trading as long as $\zeta_t \in (\zeta_-, \zeta_+)$ and trades as little as necessary at $\zeta_\pm$ to keep the risky-safe ratio in the interval $[\zeta_-, \zeta_+]$. Alternative strategies, such as trading into the middle of the no-trade region, incur significantly larger transaction costs.*[10]

The following computes the statistics contributing to the ESR of any trading strategy as in Lemma 2 (not just the optimal one) in terms of the risky-safe ratio.

Lemma 3. *Consider a control limit policy for $\eta_\pm$. Long-run mean $\hat{m}$, long-run standard deviation $\hat{\sigma}$ and average transaction costs ATC are given by the almost sure limits,*

$$\hat{m} = \lim_{T\to\infty} \frac{1}{T}\int_0^T \frac{dw_t}{w_t}dt = r + \mu \int_{\eta_-}^{\eta_+} \left(\frac{\zeta}{1+\zeta}\right)\nu(d\eta), \tag{23}$$

$$\hat{\sigma}^2 = \lim_{T\to\infty} \frac{1}{T}\left\langle \int_0^\cdot \frac{dw_t}{w_t}\right\rangle_T = \sigma^2 \int_{\eta_-}^{\eta_+} \left(\frac{\zeta}{1+\zeta}\right)^2 \nu(d\eta), \tag{24}$$

$$\mathrm{ATC} = \varepsilon \lim_{T\to\infty} \frac{1}{T}\int_0^T \pi_t \frac{d\varphi_t^\downarrow}{\varphi_t} = \frac{\sigma^2(2\mu/\sigma^2-1)}{2}\left(\frac{\frac{\varepsilon\zeta_+}{(1+\zeta_+)(1+(1-\varepsilon)\zeta_+)}}{1-\left(\frac{\zeta_-}{\zeta_+}\right)^{2\mu/\sigma^2-1}}\right), \tag{25}$$

where ν is the stationary density of Lemma 2.

Proof. All the formulae use the ergodic theorem and thus can be obtained with the methods of Guasoni and Mayerhofer (2019). In particular, identity (25) holds in the view of (Gerhold et al. 2014, Lemma C.1). □

Using the analytic expressions of (23)–(25) with MATHEMATICA, we obtain explicit asymptotics, precise at the third order in $\varepsilon^{1/3}$:

Lemma 4. *For the optimal strategy of Theorem 1, the statistics of Lemma 3 satisfy the following asymptotics:*

$$\hat{m} = r + \frac{\mu^2}{\gamma\sigma^2} - \frac{\mu(2\pi_*-1)}{\gamma}\left(\frac{\gamma\pi_*(\pi_*-1)}{6}\right)^{1/3}\varepsilon^{2/3} + O(\varepsilon^{4/3}), \tag{26}$$

$$\hat{\sigma}^2 = \frac{\mu^2}{\gamma^2\sigma^2} - \frac{\sigma^2\pi_*(7\pi_*-3)}{2\gamma}\left(\frac{\gamma\pi_*(\pi_*-1)}{6}\right)^{1/3}\varepsilon^{2/3} + O(\varepsilon^{4/3}), \tag{27}$$

$$\mathrm{ATC} = \frac{3\sigma^2}{\gamma}\left(\frac{\gamma\pi_*(\pi_*-1)}{6}\right)^{4/3}\varepsilon^{2/3} - \frac{\mu(\gamma-1)}{2\gamma}\pi_*(\pi_*-1)\varepsilon + O(\varepsilon^{4/3}). \tag{28}$$

The maximum equivalent safe rate satisfies

$$\mathrm{ESR} = r + \tfrac{\gamma\sigma^2}{2}\pi_*^2 - \tfrac{\gamma\sigma^2}{2}\left(\tfrac{3}{4\gamma}\pi_*^2(\pi_*-1)^2\right)^{2/3}\varepsilon^{2/3} + \tfrac{\mu(\gamma-1)}{2\gamma}\pi_*(\pi_*-1)\varepsilon + O(\varepsilon^{4/3}). \tag{29}$$

Remark 1.

1. *All asymptotics of Lemma 4 improve those of (Guasoni and Mayerhofer 2019, Theorem 3.1) in precision by one order. Note that (26) is a corrected version of (Guasoni and Mayerhofer 2019, Theorem 3.1, eq. (3.8)), where the bracket $(5\pi_*-3)$ is given, instead of the correct term $(2\pi_*-1)$ in (26).*
2. *One can run a consistency check that compares the asymptotics (29) of the maximum ESR (computed, by developing $r+\hat{m}-\frac{\gamma}{2}\hat{\sigma}^2 - ATC$ into a formal power series in $\varepsilon^{1/3}$) with the asymptotic expansion of the shorter formula $r+\mu\pi_- \frac{\gamma\sigma^2}{2}\pi_-^2$ from Theorems 1 and 3.*

3. Results

3.1. Asymptotically Optimal Shadow Policies

In this section, a shadow price for the mean–variance objective (3) is constructed, and asymptotic formulas for the implied strategy that is optimal in the shadow market, are derived. The exposition is motivated by the shadow price construction for log-utility investors, cf. (Gerhold et al. 2013, Chapter 3), and see also Guasoni and Muhle-Karbe (2013); Gerhold et al. (2014). Assume the following functional form of the shadow price $\widetilde{S}_t$,

$$\widetilde{S}_t = g(\pi_t)S_t, \tag{30}$$

where g satisfies the boundary conditions

$$g(\pi_-) = 1, \quad g(\pi_+) = (1-\varepsilon), \tag{31}$$

reflecting that an optimal strategy (such as of Theorem 1) is a control limit policy for $\pi_\pm$, which buys (respectively sells) the frictionless asset $\widetilde{S}$ precisely when its price equals the ask price S, and sells precisely when it equals the bid price $(1-\varepsilon)S$.

If $\widetilde{S}$ satisfies (30) with twice differentiable g, and if g satisfies (34), then Itô's formula yields the dynamics of instantaneous returns[11]

$$\frac{d\widetilde{S}_t}{\widetilde{S}_t} = rdt + d\widetilde{\mu}_t + \widetilde{\sigma}_t dB_t,$$

with

$$d\widetilde{\mu}_t = \mu dt + \frac{g'(\pi_t)\big(\pi_t(1-\pi_t)\mu dt + \pi_t(1-\pi_t)^2\sigma^2\big)}{g(\pi_t)} + \frac{\frac{1}{2}g''(\pi_t)\pi_t^2(1-\pi_t)^2\sigma^2 dt + g'(\pi_t)\left(\pi_t\frac{d\varphi_t^\uparrow}{\varphi_t} - (1-\varepsilon\pi_t)\frac{d\varphi_t^\downarrow}{\varphi_t}\right)}{g(\pi_t)} \tag{32}$$

and diffusion coefficient

$$\widetilde{\sigma}_t = (\sigma g(\pi_t) + g'(\pi_t)\pi_t(1-\pi_t)\sigma)/g(\pi_t). \tag{33}$$

The smooth pasting condition

$$g'(\pi_-) = g'(\pi_+) = 0 \tag{34}$$

is imposed such that the instantaneous drift of the shadow price becomes absolutely continuous (the condition removes the local time terms $\frac{d\varphi_t^\uparrow}{\varphi_t}$ and $\frac{d\varphi_t^\downarrow}{\varphi_t}$), and thus $d\widetilde{\mu}_t = \widetilde{\mu}_t dt$, with

$$\widetilde{\mu}_t = \mu + \frac{g'(\pi_t)(\pi_t(1-\pi_t)\mu + \pi_t(1-\pi_t)^2\sigma^2) + \frac{1}{2}g''(\pi_t)\pi_t^2(1-\pi_t)^2\sigma^2}{g(\pi_t)}. \tag{35}$$

The fraction of wealth $\widetilde{\pi}$ invested in the risky asset, evaluated at the shadow price, satisfies

$$\widetilde{\pi}_t = \frac{Y_t g(\pi_t)}{X_t + Y g(\pi_t)} = \frac{\pi_t g(\pi_t)}{(1-\pi_t) + \pi_t g(\pi_t)}. \tag{36}$$

The mean–variance optimality in the shadow market holds when the proportion of wealth in the shadow market's risky asset $\widetilde{S}$ equals the Merton fraction, that is,

$$\widetilde{\pi}_t = \frac{\widetilde{\mu}_t}{\gamma\widetilde{\sigma}_t^2}.$$

Equating this solution with (36), and using (35), (33) entails that g satisfies the ODE

$$\frac{1}{2}g''(\pi)\pi^2(1-\pi)^2\sigma^2 = \frac{\gamma\pi\sigma^2(g + g'(\pi)\pi(1-\pi))^2}{1-\pi+\pi g(\pi)} - \mu g(\pi) - g'(\pi)(\pi(1-\pi)\mu + \pi(1-\pi)^2\sigma^2). \tag{37}$$

Define Ψ implicitly as

$$g(\pi) = \frac{\Psi(Y/X)}{Y/X} =: \frac{\Psi(\zeta)}{\zeta},$$

and set
$$\zeta_\pm = \frac{\pi_\pm}{1-\pi_\pm}. \tag{38}$$

Then, (Ψ, ζ_-, ζ_+) satisfy the problem

$$\Psi''(\zeta) = \frac{2\gamma\Psi'^2(\zeta)}{(1+\Psi(\zeta))} - \frac{2\mu}{\sigma^2}\frac{\Psi'(\zeta)}{\zeta}, \tag{39}$$
$$\Psi'(\zeta_-) = \Psi(\zeta_-)/\zeta_- = 1, \tag{40}$$
$$\Psi'(\zeta_+) = \Psi(\zeta_+)/\zeta_+ = (1-\varepsilon). \tag{41}$$

This is a free boundary problem, because both Ψ and the trading boundaries $\zeta_\pm$ for the control limit policy are unknown.

Using the explicit solution Ψ of the corresponding initial value problem (39) and (40), and respecting terminal conditions (41), one obtains a non-linear system of equations for (ζ_-, ζ_+). For small ε, this very system allows a unique solution with asymptotic expansion[12]

$$\tilde{\zeta}_\pm = \frac{\pi_*}{1-\pi_*} \pm \left(\frac{3}{4\gamma}\right)^{1/3}\left(\frac{(\pi_*)^2}{1-\pi_*}\right)^{2/3}\varepsilon^{1/3} - \frac{(1+2\gamma)\pi_*}{2\gamma(1-\pi_*)^2}\left(\frac{\gamma\pi_*(\pi_*-1)}{6}\right)^{1/3}\varepsilon^{2/3} + O(\varepsilon). \tag{42}$$

In comparison, the optimal strategy of Theorem 1 is a control limit policy whose limits $\zeta_\pm$, in terms of the risky-safe ratio, have the expansion[13]

$$\zeta_\pm = \frac{\pi_*}{1-\pi_*} \pm \left(\frac{3}{4\gamma}\right)^{1/3}\left(\frac{\pi_*}{(\pi_*-1)^2}\right)^{2/3}\varepsilon^{1/3} - \frac{(5-2\gamma)\pi_*}{2\gamma(\pi_*-1)^2}\left(\frac{\gamma\pi_*(\pi_*-1)}{6}\right)^{1/3}\varepsilon^{2/3} + O(\varepsilon). \tag{43}$$

Note the factor $(1+2\gamma)$ in the $\varepsilon^{2/3}$ term in (42), which differs from the factor $(5-2\gamma)$ in (43). Accordingly, the associated trading boundaries have an asymptotic expansion,

$$\tilde{\pi}_\pm = \pi_* \pm \left(\frac{3}{4\gamma}(\pi_*)^2(1-\pi_*)^2\right)^{1/3}\varepsilon^{1/3} + \frac{(1-\gamma)\pi_*}{\gamma}\left(\frac{\gamma\pi_*(\pi_*-1)}{6}\right)^{1/3}\varepsilon^{2/3} + O(\varepsilon). \tag{44}$$

The expansion (42), respectively, (44), agrees with the above expansions (43), respectively, (11), up to the first order (they agree in constant and in $\varepsilon^{1/3}$ terms). But they disagree in a quite subtle way for any $\gamma \neq 1$ at the second order: the absolute values, but not the signs of the second-order term of $\pi_\pm$ (see (44)) and $\tilde{\pi}_\pm$ (see (11)), are identical.

The following establishes asymptotic optimality of the third order of the strategy obtained from the shadow market (the proof exclusively uses MATHEMATICA and higher-order expansions of (44)).

Theorem 2. *The asymptotic expansion of the equivalent safe rate and average transaction costs of the control limit policy for $\tilde{\pi}_\pm$ are of the exact same form as (29), respectively, (25). Thus, the strategy is asymptotically optimal at the third order. The long-run mean and variance defined by Lemma 3 satisfy the following asymptotics:*

$$\tilde{m} = r + \frac{\mu^2}{\gamma\sigma^2} - \frac{\mu(2\gamma\pi_*-1)}{\gamma}\left(\frac{\gamma\pi_*(\pi_*-1)}{6}\right)^{1/3}\varepsilon^{2/3} + O(\varepsilon^{4/3}), \tag{45}$$
$$\tilde{\sigma}^2 = \frac{\mu^2}{\gamma^2\sigma^2} - \frac{\sigma^2\pi_*(\pi_*(8\gamma+1)-3)}{2\gamma}\left(\frac{\gamma\pi_*(\pi_*-1)}{6}\right)^{1/3}\varepsilon^{2/3} + O(\varepsilon^{4/3}). \tag{46}$$

Remark 2. *Note that, almost miraculously,*

$$\hat{m} - \frac{\gamma}{2}\hat{\sigma}^2 = \tilde{m} - \frac{\gamma}{2}\tilde{\sigma}^2 = O(\varepsilon^{4/3})$$

because the average transaction costs as well as the equivalent safe rate agree for both strategies up to the third order, and the mean and variance's third-order terms vanish (compare (45) and (46) with the mean and variance of the optimal strategy in Lemma 4).

3.2. Outperforming the Shadow Market

In Theorem 2, it is shown that $\widetilde{S}$ is an asymptotic shadow price, as the strategy that is optimal in the frictionless market is even optimal at the third order in the original market. For the proof of this statement, it is crucial to have precise asymptotic expansions of the trading boundaries $\widetilde{\pi}_{\pm}$.

The objective of this section is to prove that this strategy is not optimal. To this end, it would be useful to have higher (fourth and fifth) order terms in the expansion of the optimal trading boundaries $\zeta_{\pm}$ and thus, the maximum performance (29). However, the free boundary problem of (A1)–(A5) associated with the optimal solution of Theorem 1 is notoriously difficult to deal with, even with MATHEMATICA, while the free boundary problem (A9) and (A10) arising from the shadow price ansatz is much more tractable. Therefore, instead of developing the maximum performance to even higher precision, a strategy is found that merely outperforms the shadow market.

Theorem 3. *Suppose $\gamma \notin \{0, 1\}$. For any $\theta \in \mathbb{R}$, the family of control limit policies for*

$$\widetilde{\pi}^{\theta}_{\pm} := \widetilde{\pi}_{\pm} + (\theta - 1) \times \frac{(\gamma-1)(\pi_*)^2(1-\pi_*)}{6}\left(\frac{6}{\gamma\pi_*(1-\pi_*)}\right)^{2/3}\varepsilon^{2/3} \quad (47)$$

has an equivalent safe rate

$$\begin{aligned}\mathrm{ESR} = r &+ \frac{\gamma\sigma^2}{2}\pi_*^2 - \frac{\gamma\sigma^2}{2}\left(\frac{3}{4\gamma}\pi_*^2(\pi_*-1)^2\right)^{2/3}\varepsilon^{2/3} + \frac{\mu(\gamma-1)}{2\gamma}\pi_*(\pi_*-1)\varepsilon \\ &- \frac{\sigma^2 \times k(\theta)}{20\gamma}\left(\frac{\gamma\pi_*(1-\pi_*)}{6}\right)^{2/3}\varepsilon^{4/3} + O(\varepsilon^{5/3}),\end{aligned} \quad (48)$$

where

$$k(\theta) := -9 + 2\pi_*\Big(9 + \pi_*\Big(3 + 12\gamma(\gamma-2) + (10\theta + 5\theta^2)(\gamma-1)^2\Big)\Big)$$

and, thus, is asymptotically optimal at the third order. For sufficiently small ε, the best performance at the fourth order is achieved for $\theta = -1$, strictly outperforming the shadow performance ($\theta = 1$).

Proof. Using the method in Appendix A, derive asymptotic expansions of c and s (whence $\widetilde{\zeta}_{\pm}$ up to the sixth order), satisfying the free boundary problem (A1)–(A5) at the same order. Modifying the second-order term by including a factor θ as in (47), one arrives at (48). The fourth-order coefficient $k(\theta)$ is a polynomial of the second order in θ, with a global minimum at $\theta = -1$. The comparison sign and magnitude of this factor are straightforward and reveal that $\theta = -1$ outperforms any other control limit policy for $\theta \neq -1$. □

Remark 3. *Note that for $\theta = -1$, the control limit policy for (47) is, up to order two, equal to the optimal strategy (11) of Theorem 1. This does not mean that it is optimal at order four or beyond, as higher-order coefficients of $\widetilde{\pi}^{\theta}_{\pm}$ may not agree with those of the optimal boundaries $\pi_{\pm}$.*

3.3. The Limits of Shadow Prices

Recall that a shadow price $\widetilde{S}$ is a frictionless process evolving in the bid–ask spread

$$(1-\varepsilon)S_t \le \widetilde{S}_t \le S_t, \quad t \ge 0 \quad (49)$$

such that the optimal strategy φ is also optimal in the original market, and buys (respectively sells) precisely when $\widetilde{S}_t = S_t$ (respectively $\widetilde{S}_t = (1-\varepsilon)S_t$).

To start with, the dynamics of the risky-safe ratio, wealth and proportion of wealth in the shadow market, for any finite variation strategy, is stated.

Lemma 5. *Suppose the shadow price satisfies the dynamics*

$$\frac{d\widetilde{S}_t}{\widetilde{S}_t} = (r + \widetilde{\mu}_t)dt + \widetilde{\sigma}_t dB_t. \tag{50}$$

For any finite variation trading strategy φ*,*

$$\frac{d\widetilde{\zeta}_t}{\widetilde{\zeta}_t} = \widetilde{\mu}_t dt + \widetilde{\sigma}_t dB_t + (1+\widetilde{\zeta}_t)\frac{d\varphi_t^{\uparrow}}{\varphi_t} - (1+\widetilde{\zeta}_t)\frac{d\varphi_t^{\downarrow}}{\varphi_t}, \tag{51}$$

$$\frac{d\widetilde{w}_t}{\widetilde{w}_t} = rdt + \widetilde{\pi}_t(\widetilde{\mu}_t dt + \widetilde{\sigma}_t dB_t), \tag{52}$$

$$\frac{d\widetilde{\pi}_t}{\widetilde{\pi}_t} = (1-\widetilde{\pi}_t)(\widetilde{\mu}_t dt + \widetilde{\sigma}_t dB_t) - \widetilde{\pi}_t(1-\widetilde{\pi}_t)\widetilde{\sigma}_t^2 dt + \frac{d\varphi_t^{\uparrow}}{\varphi_t} - \frac{d\varphi_t^{\downarrow}}{\varphi_t}. \tag{53}$$

Proof. A similar proof as that of Lemma 1 applies. □

Lemma 6. *If a shadow price exists, then for the optimal strategy, the cash positions in the original and shadow markets agree (*$\widetilde{X} = X$*), and the fraction of wealth invested in the shadow price satisfies*

$$\pi_t \le \widetilde{\pi}_t \le \frac{(1-\varepsilon)\pi_t}{1-\varepsilon\pi_t}.$$

In particular, if the optimal strategy satisfies $\pi_t \in [\pi_-, \pi_+]$*, then*

$$\pi_- \le \widetilde{\pi}_t \le \frac{(1-\varepsilon)\pi_+}{1-\varepsilon\pi_+} \le (1-\varepsilon)\pi_+(1+\varepsilon\pi_+). \tag{54}$$

Proof. The optimal strategy trades the risky asset at the same price in both markets; therefore, the cash positions agree.

The lower bound is proved by observing that for $a, b > 0$, the function

$$\frac{a\xi}{-b + a\xi}$$

is strictly decreasing for any $\xi > b/a$ (which corresponds to positive wealth), and since $\widetilde{S} \le S$,

$$\widetilde{\pi}_t = \frac{\varphi_t \widetilde{S}_t}{X_t + \varphi_t \widetilde{S}_t} \ge \frac{\varphi_t S_t}{X_t + \varphi_t S_t} = \frac{\varphi_t S_t}{w_t} = \pi_t.$$

Similarly, the upper bound follows from

$$\widetilde{\pi}_t = \frac{\varphi_t \widetilde{S}_t}{X_t + \varphi_t \widetilde{S}_t} \le \frac{(1-\varepsilon)\varphi_t S_t}{X_t + \varphi_t(1-\varepsilon)S_t} = \frac{(1-\varepsilon)\pi_t}{1-\varepsilon\pi_t}.$$

The constant bounds in terms of the trading boundaries $\pi_\pm$ are an obvious conclusion. The last inequality in (54) follows from the summation formula of the geometric series, knowing that solvency implies $\varepsilon\pi_+ < 1$. □

For proportional transaction costs, maximizing the expected excess returns

$$\lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T \frac{dw_t}{w_t}\right]$$

over all admissible strategies $\varphi \in \Phi$ is well posed. By (Guasoni and Mayerhofer 2023, Theorem 3.2), for sufficiently small ε, there exists $0 < \pi_- < \pi_+ < \infty$ such that the trading strategy $\hat{\varphi}$ that buys at π_- and sells at π_+ to keep the risky weight π_t within the interval $[\pi_-, \pi_+]$ is optimal. The maximum expected return of this optimal strategy is given by the almost sure limit

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{dw_t}{w_t} = r + \mu\pi_- \tag{55}$$

and the trading boundaries have the series expansions

$$\pi_- = (1-\kappa)\kappa^{1/2}\left(\frac{\mu}{\sigma^2}\right)^{1/2} \varepsilon^{-1/2} + 1 + O(\varepsilon^{1/2}), \tag{56}$$

$$\pi_+ = \kappa^{1/2}\left(\frac{\mu}{\sigma^2}\right)^{1/2} \varepsilon^{-1/2} + 1 + O(\varepsilon^{1/2}), \tag{57}$$

where $\kappa \approx 0.5828$ is the unique solution of

$$\frac{3}{2}\xi + \log(1-\xi) = 0, \quad \xi \in (0,1).$$

The remainder of this section is dedicated to showing that a shadow market does not exist.

For technical reasons, it is assumed in this section that any shadow price satisfies the following.

Assumption 1. *A shadow price $\widetilde{S}$ is a continuous process satisfying the dynamics* (50) *with drift and diffusion coefficients being ergodic in the sense that, almost surely, for some $\bar{\mu}, \bar{\sigma}^2 \in \mathbb{R}$,*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \widetilde{\mu}_t dt = \bar{\mu}, \quad \lim_{T\to\infty} \frac{1}{T} \int_0^T \widetilde{\sigma}_t^2 dt = \bar{\sigma}^2. \tag{58}$$

Remark 4. *The fairly general Assumption 1 is natural in that it applies to all known constructions of shadow prices in continuous-time models. In fact, typically, the ratio $\frac{\widetilde{S}_t}{S_t}$ is equal to $g(\pi_t)$, where g is a real analytic function and π_t is a stationary process: the optimal proportion of wealth in the risky asset, evolving within an interval $[\pi_-, \pi_+]$, where one buys (respectively sells) precisely at the trading boundary π_i (respectively π_+) and satisfying $g(\pi_-) = 1$ and $g(\pi_+) = 1 - \varepsilon$, reflecting the very definition of shadow price, agreeing with the ask (respectively bid) whenever shares are purchased (respectively sold). As these functions in the literature are all analytic, one can use Itô's formula to derive the dynamics* (50) *in a more explicit form.*[14] *There exist continuous functions h and H such that*

$$\widetilde{\mu}_t = h(\pi_t), \quad \widetilde{\sigma}_t^2 = H(\pi_t)$$

By the ergodic theorem (Borodin and Salminen 2002, II.35 and II.36), one obtains the finite limits in (58).

Lemma 7. *If $\mu > \sigma^2/2$, then $\bar{\mu} > \bar{\sigma}^2/2$. In particular, $\bar{\mu} \neq 0$.*

Proof. Write the fraction $\widetilde{S}_t/S_t$ as explicit solutions, where the accrual factor e^{rt} factors out. As $\mu > \sigma^2/2$, by the law of iterated logarithms, $e^{-rt}S_t$ almost surely tends to ∞ as $t \to \infty$. If $\bar{\mu} \leq \bar{\sigma}^2/2$, then for sufficiently large t, $\widetilde{S}_t$ behaves like a geometric Brownian motion with drift $\bar{\mu}$ and volatility $\bar{\sigma}$, whence by a similar argument, $\liminf_{t\to\infty} \widetilde{S}_t = 0$, almost surely. Thus, $\liminf_{t\to\infty} \widetilde{S}_t/S_t = 0 < (1-\varepsilon)$, a contradiction to (49). □

Theorem 4. *If $\mu > \sigma^2/2$, then a shadow price satisfying the dynamics of* (50) *and Assumption 1 does not exist.*

Proof. Assume, for a contradiction, there exists a shadow price $\widetilde{S}$. By Lemma 5, the shadow wealth $\widetilde{w}_t = \varphi_t \widetilde{S}_t + \widetilde{X}_t$ satisfies the SDE (52). Furthermore, by Lemma 6

$$\int_0^t \widetilde{\pi}_t^2 \widetilde{\sigma}_u^2 du \leq (1-\varepsilon)^2 (\pi_+^2 (1+\varepsilon \pi_+)^2 \int_0^t \widetilde{\sigma}_u^2 du < \infty$$

and thus, the integral of the Brownian term is a martingale by Assumption 1. Thus, the strategy φ with associated wealth $\widetilde{w}$ achieves its optimum at

$$\lambda := \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int_0^T \frac{d\widetilde{w}_t}{\widetilde{w}_t}\right] = r + \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int_0^T \widetilde{\mu}_t \widetilde{\pi}_t dt\right].$$

Note that, by Assumption 1,

$$\bar{\mu} = \lim_{T\to\infty} \frac{1}{T} \int_0^T \widetilde{\mu}_t dt = \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int_0^T \widetilde{\mu}_t dt\right]$$

and by Lemma 7, $\bar{\mu} > 0$. Furthermore, by Lemma 6, there exist $0 < L < U < \infty$ such that

$$L\bar{\mu} < \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int_0^T \widetilde{\mu}_t \widetilde{\pi}_t dt\right] < U\bar{\mu}. \tag{59}$$

Any alternative strategy $\varphi^\star$, whose proportion of wealth in the shadow price satisfies

$$\widetilde{\pi}^* \geq U \tag{60}$$

outperforms φ because

$$\lambda^* \geq r + U\bar{\mu} > \lambda.$$

Trading strategies that keep the exposure in the shadow asset constant to U exist, but they are of infinite variation. To obtain finite variation strategies satisfying (60), recall that by Lemma 5, the fraction of wealth in the shadow asset $\widetilde{w}$ associated with a finite variation strategy satisfies (53). One can modify this strategy by allowing bulk trades: let φ^* be the finite variations strategy that does refrain from trading whenever $\widetilde{\pi}^\star \in (U, 2U)$ but buys (respectively sells) the shadow asset in bulk whenever $\widetilde{\pi}^\star$ hits U (respectively $2U$) so as to reset $\widetilde{\pi}^\star$ to the midpoint $3U/2$. Such a strategy can be constructed pathwise and satisfies

$$\frac{d\widetilde{\pi}_{t-}^\star}{\widetilde{\pi}_{t-}^\star} = (1-\widetilde{\pi}_{t-}^\star)(\widetilde{\mu}_t dt + \widetilde{\sigma}_t dB_t) - \widetilde{\pi}_{t-}^\star (1-\widetilde{\pi}_{t-}^\star)\widetilde{\sigma}_t^2 dt + \frac{\Delta\varphi_t^{\star,\uparrow}}{\varphi_t^\star} - \frac{\Delta\varphi_t^{\star,\downarrow}}{\varphi_t^\star}.$$

The existence of such a strategy contradicts the optimality of φ, and thus a shadow price does not exist. □

Remark 5. *The finite variation strategy in the end of the proof cannot be replaced by a (standard) reflected diffusion with two reflecting boundaries because for the existence of strong solutions to the associated SDE on convex domains (Tanaka 1979, Theorem 4.1), one would need $\widetilde{\mu}_t$ and $\widetilde{\sigma}_t$ to be regular enough functions of $\widetilde{\pi}_t^\star$, an assumption that is too strong in this context. Also, it is unknown whether such a strategy is solvent in the original market with transaction costs.*

4. Discussion

Optimizing portfolios in continuous-time markets with proportional costs presents mathematically challenging problems. Strategies that are optimal in frictionless markets must be adjusted to prevent immediate bankruptcy as exemplified by the dynamic hedging component of a variance swap (see Example 1). The strategies considered in this paper are stationary[15] and, thus, ergodic theorems are used to determine their long-run performance. To gain insights into trading frequency, transaction costs, and long-run performance, we derive asymptotic expansions of the trading boundaries for small bid–ask spread.

The paper explores the (candidate) shadow prices for local mean–variance investors, with a threefold contribution.

First, we discover that the optimal strategy in the (candidate) shadow market differs from the optimal one in the original market but only in the second-order terms of the asymptotic expansion of the trading boundaries.[16] Theorem 2 demonstrates that, for risk aversion $\gamma > 0$, the equivalent safe rate of the shadow market strategy agrees at the third order with the maximum. As transaction costs are of the second order, we conclude that the performance of the shadow market strategy is essentially optimal. It is worth noting that the same is true[17] for a long-run power–utility investor (cf. Gerhold et al. (2014)), as their trading boundaries also agree at the first order with (1). Second, Theorem 3 establishes that for $\gamma \neq 1$, the (potential) shadow market strategy $\widetilde{\pi}$ is not optimal, as it can be outperformed. The alternative strategy is not necessarily optimal, even though it agrees up to the second order with the optimal one. In summary, the (candidate) shadow price is an asymptotic shadow price. Third, Theorem 4 demonstrates that for risk-neutral investors ($\gamma = 0$), no such shadow market exists.

The findings of this paper prompt the following research problems. First, we conjecture that a minor modification of the objective will render the shadow price candidate of Section 3.1 optimal in the original market with transaction costs. Motivated by Martin (2012, 2016), we propose to replace the equivalent safe rate in (3) by an infinite horizon, the local mean–variance utility function[18]

$$\mathrm{ESR} := \mathbb{E}\left[\int_0^\infty \delta e^{-\delta t}\frac{dw_t}{w_t} - \frac{\gamma}{2}\int_0^\infty \delta e^{-\delta t}\left\langle \frac{dw_t}{w_t}\right\rangle_t\right] \tag{61}$$

for some discount rate $\delta > 0$. In the absence of transaction costs ($\varepsilon = 0$), the maximum equivalent safe rate agrees with that of the old objective (3). More importantly, this objective leads to the exact same shadow market construction as in Section 3.1. The question remains if our shadow market policy maximizes also (61) in the original market, surpassing its third-order optimality (Theorem 2).

Second, the mathematical treatment of optimization problems involving transaction costs is always uniquely tailored to a specific objective. This results in free boundary problems that vary significantly, encompassing scenarios from Riccati Differential Equations Gerhold et al. (2012) and linear equations (Guasoni and Mayerhofer 2019, Theorem 3.3) to the nonlinear problem (37) addressed in this paper, and even singular problems for zero risk-aversion (Guasoni and Mayerhofer 2019, Theorem 3.2). The question persists: can a unified approach be devised that accommodates a diverse range of objectives? To explore this possibility, one might aim for conformity to a common format—a second-order free boundary problem stated as follows:

$$F(g, g', g'') = 0, \tag{62}$$
$$g(\pi_-) = 1, \quad g(\pi_+) = 1 - \varepsilon, \tag{63}$$
$$g'(\pi_-) = 0, \quad g'(\pi_+) = 0. \tag{64}$$

This problem involves an unknown scalar function $g = g(\pi)$ that must satisfy a second-order nonlinear ODE (62), along with buy and sell boundaries π_- and π_+, respectively. The latter boundaries must adhere to zeroth-order boundary conditions (63) and first-order conditions (64).[19] In practical trading applications, a second-order approximation of the trading boundaries would suffice. Such an approximation might be achieved through a general polynomial ansatz for an approximation of (62).

Third, most of the literature[20] regarding the existence of an optimal strategy and its asymptotic expansion depends on the assumption of a "sufficiently small" bid–ask spread ε, without providing a minimum ε_0, for which these statements hold. Are they applicable to actual bid–ask spreads observed in markets (for liquid assets ranging in the basis points)? Addressing this question involves either demonstrating optimality for all $\varepsilon \in (0, 1)$ or identifying counterexamples where optimality breaks down for larger transaction costs, along with determining the explicit lower bound ε_0 at which control limit policies remain optimal. Such a lower bound would be contingent on model parameters γ, μ/σ^2, and risk aversion. Most likely, it will depend on the chosen objective.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A. The Free Boundary Problem for the Shadow Price Candidate

Let us introduce two new parameters $c = 1/\zeta_-$ and $s = \zeta_+/\zeta_-$. By defining the new function ϕ implicitly via

$$\Psi(\zeta) := \phi(c\zeta)/c,$$

the free boundary problem (39)–(41) produces a similar one for $\phi(z)$,

$$\phi''(z) = \frac{2\gamma\phi'^2(z)}{(c+\phi(z))} - 2\gamma\pi^* \frac{\phi'(z)}{z}, \tag{A1}$$

$$\phi(1) = 1, \tag{A2}$$

$$\phi'(1) = 1, \tag{A3}$$

$$\phi(s) = (1-\varepsilon)s, \tag{A4}$$

$$\phi'(s) = (1-\varepsilon). \tag{A5}$$

Remark A1. *Since for small transaction costs, trading strategies will be control limit policies on sufficiently small intervals, only the following cases need to be distinguished:*

- $\zeta_- < \zeta_+ < -1$ *(levered case): Then,* $c < 0$*, and therefore* $z > 0$*, so* $s < 1$*. Conversely,* $s < 1$ *implies* $\zeta_- < -1$*.*
- $0 < \zeta_- < \zeta_+$ *(unlevered case): Then,* $c > 0$*, and therefore the argument* $z < 0$*, so* $s > 1$*. Conversely,* $s > 1$ *implies* $\zeta_- > 0$*.*

For the sake of brevity, let us only consider the levered case, that is, $\zeta_- < \zeta_+ < -1$ and $\phi(\zeta) < -1$. Since also $c < 0$, one obtains $\phi(z) > -c$ for all z. Also, $c + 1 > 0$. Dividing (A1) by ϕ' and integrating once, one thus obtains

$$\log(\phi'(z)) = 2\gamma \log(c + \phi(z)) - 2\gamma\pi^* \log z - 2\gamma \log(c+1),$$

where the initial condition (A3) is respected. Taking antilogarithms, one thus obtains

$$\frac{\phi'(z)}{(c+\phi(z))^{2\gamma}} = \frac{z^{-2\gamma\pi^*}}{(c+1)^{2\gamma}}. \tag{A6}$$

Exclude in the following the singular cases $\gamma \neq 1/2$ and $\mu/\sigma^2 \neq 1/2$ (those cases can be dealt individually, leading to simpler solutions of the ODE (A6).). Integrating once again, one obtains

$$\frac{(c+\phi(z))^{1-2\gamma}}{1-2\gamma} = \frac{(z^{1-2\gamma\pi^*}-1)}{(1-2\gamma\pi^*)(c+1)^{2\gamma}} + \frac{(c+1)^{1-2\gamma}}{1-2\gamma}, \tag{A7}$$

where the initial condition (A2) is respected. Thus,

$$\phi(z) = -c + \left(\frac{\frac{(1-2\gamma)}{(1-2\gamma\pi^*)}(z^{1-2\gamma\pi^*}-1)}{(c+1)^{2\gamma}} + (c+1)^{1-2\gamma} \right)^{\frac{1}{1-2\gamma}}. \tag{A8}$$

Until this stage, the terminal boundary conditions (A4) and (A5) have not been involved. Those allow to reformulate the free boundary problem in terms of a system of non-linear equations for s and c.

Lemma A1. *Let $\gamma \neq 1/2$ and $\mu/\sigma^2 \neq 1/2$. ϕ, c, s is a solution to the free boundary problem (A1)–(A5) if and only if s and c satisfy the following system of non-linear equations:*

$$\left(\frac{c+(1-\varepsilon)s}{c+1}\right)^{1-2\gamma} - 1 = \frac{1-2\gamma}{1-2\gamma\pi^*}\frac{s^{1-2\gamma\pi^*}-1}{c+1}, \tag{A9}$$

$$(1-\varepsilon)^{\frac{1}{2\gamma}} s^{\pi^*} = \frac{c+(1-\varepsilon)s}{c+1}. \tag{A10}$$

Proof. The initial value problem (A1)–(A3), parameterized in c, has the explicit solution (A8). What remains is to involve the boundary conditions (A4) and (A5). Starting from (A7) and using (A4) yields

$$\frac{(c+(1-\varepsilon)s)^{1-2\gamma}}{1-2\gamma} = \frac{(s^{1-2\gamma\pi^*}-1)}{(1-2\gamma\pi^*)(c+1)^{2\gamma}} + \frac{(c+1)^{1-2\gamma}}{1-2\gamma}, \tag{A11}$$

from which (A9) follows. Using (A4)–(A6), one obtains

$$\frac{(1-\varepsilon)}{(c+(1-\varepsilon)s)^{2\gamma}} = \frac{s^{-2\gamma\pi^*}}{(c+1)^{2\gamma}}.$$

Taking the 2γths root, one obtains (A10). The proof of the converse implication is similar. □

Appendix B. Asymptotics of the Free Boundaries

Recall that $\pi^* = \frac{\mu}{\gamma\sigma^2}$ and note that

$$\zeta_- = \frac{1}{c}, \quad \zeta_+ = \frac{s}{c} \tag{A12}$$

and the associated trading boundaries $\pi_\pm$ satisfy

$$\pi_\pm := \frac{\zeta_\pm}{1+\zeta_\pm}. \tag{A13}$$

We introduce the abbreviations

$$\bar{c} := \frac{1-\pi^*}{\pi^*}, \quad \Delta := \left(\frac{6}{\gamma\pi^*(1-\pi^*)}\right)^{1/3}\varepsilon^{1/3}.$$

Proposition A1. *For sufficiently small $\varepsilon > 0$, the free boundary problem (A1)–(A5) has a unique solution $(h(\zeta), c, s)$. Moreover, the following asymptotics hold as $\varepsilon \to 0$:*

$$\begin{aligned} c = \bar{c} &+ \frac{1-\pi^*}{2\pi^*}\Delta + \frac{(1-\pi^*)(3-\pi^*(2\gamma+1))}{12\pi^*}\Delta^2 \\ &- \frac{(\pi^*-1)\left((4\gamma^2+22\gamma+1)(\pi^*)^2 - 24(2\gamma+1)\pi^* + 36\right)}{360\pi^*}\Delta^3 + O(\varepsilon^{4/3}), \end{aligned} \tag{A14}$$

$$\begin{aligned} s = 1 + \Delta &+ \frac{\Delta^2}{2} + \frac{1}{180}\left(\left(4\gamma^2 - 8\gamma + 1\right)(\pi^*)^2 + 3(4\gamma-3)\pi^* + 36\right)\Delta^3 \\ &+ \frac{(8\gamma^2 - 26\gamma + 2)(\pi^*)^2 + 2(17\gamma-9)\pi^* + 27}{360}\Delta^4 + O(\varepsilon^{5/3}). \end{aligned} \tag{A15}$$

Proof. The proof is inspired by (Gerhold et al. 2013, Proposition 6.1), where a similar result is developed for log utility from consumption and for unlevered strategies.[21] Having already solved the initial value problem (A1)–(A3), parameterized in c, which has the explicit solution (A8), it remains to involve the boundary conditions (A4) and (A5). A naïve approach would be to define for sufficiently small δ the map $F := (F_1, F_2)^\top$, where

$$F_1(\delta, c, s) = \left(\frac{c + (1-\delta^3)s}{c+1}\right)^{1-2\gamma} - 1 - \frac{1-2\gamma}{1-2\gamma\pi^*}\frac{s^{1-2\gamma\pi^*}-1}{c+1}, \tag{A16}$$

$$F_2(\delta, c, s) = (1-\delta^3)^{\frac{1}{2\gamma}} s^{\pi^*} - \frac{c+(1-\delta^3)s}{c+1}, \tag{A17}$$

and to show, by means of the implicit function theorem, that F has a unique zero $(s(\delta), c(\delta))$ at $(c = \bar{c}, s = 1)$, which is analytic in δ. Note, however, that the implicit function theorem cannot be applied in this case: even though $F(\delta_0 = 0, c_0 = \bar{c}, s_0 = s) = 0$, the Jacobian vanishes at the critical point $(0, \bar{c}, 1)$.

Consider the levered case only, as the other case can be proved quite similarly. In this case, $s < 1$. Having a look at Equation (A6), one sees that for $z < 1$, z is sufficiently close to $z = 1$, $\phi'(z) > 0$, and since $\phi(1) = 1$, this implies $\phi(z) < 1$. Since $\phi = \phi(z, c)$ is an analytic function in (z, c) near $c = \bar{c}$ and $z = 1$, it satisfies an expansion of the form

$$\phi(z,c) = 1 + (z-1) + \sum_{i\geq 2}\sum_{j\geq 0} a_{ij}(z-1)^i(c-\bar{c})^j$$

with coefficients a_{ij}, which can be calculated recursively. Furthermore, $a_{0j} = a_{1j} = \delta_{0j}$ for $j \geq 0$ due to the initial conditions (A2) and (A3). One now solves for c, s, invoking the terminal conditions (A4) and (A5). The latter imply that

$$\varepsilon s = s - \phi(s,c), \quad \text{and} \quad \phi(s,c) - s\phi'(z = s, c) = 0.$$

Dividing by $s - 1$, reflecting that the solution $s = 1$ is not interesting, a Taylor expansion yields

$$\frac{\phi(s,c) - s\phi'(s,c)}{s-1} = \sum_{i\geq 0}\sum_{j\geq 0} b_{ij}(s-1)^i(c-\bar{c})^j \tag{A18}$$

for certain coefficients b_{ij}. By using (A1)–(A3) and L'Hospital's rule, one obtains

$$b_{0,0} = \lim_{z\to 1, c\to\bar{c}} \frac{\phi(z,c) - z\phi'(z,c)}{z-1} = -\phi''(1,\bar{c}) = 0,$$

and, further by a twofold application of L'Hospital's rule,

$$b_{1,0} = \lim_{z\to 1, c\to\bar{c}} \frac{\phi(z,c) - z\phi'(z,c)}{(z-1)^2} = -\lim_{z\to 1, c\to\bar{c}} \phi^{(3)}(z,c) = 2\gamma\pi^*(1-\pi^*) \neq 0.$$

Hence, the implicit function theorem is applicable and yields $s(c) = H(c)$ as a function of c such that

$$H(\bar{c}) = 1, \quad H'(\bar{c}) = \frac{2\pi^*}{1-\pi^*}.$$

Inserting this function into (A10), one obtains the problem

$$g(c,\delta) := (1-\delta^3)^{\frac{1}{2\gamma}} H^{\pi^*}(c) - \frac{c + (1-\varepsilon)H(c)}{c+1} = 0.$$

Since $g(c = \bar{c}, \delta = 0) = 0$ and $\partial_c g(c, \delta) = \frac{1}{\pi^*} \neq 0$, one can apply the implicit function theorem which asserts that for sufficiently small δ, a unique and analytic solution $c = c(\delta)$ exists to $g(c, \delta) = 0$ and $c(0) = \bar{c}$. Therefore, $c(\delta), s(\delta) = H(c(\delta))$ is the unique solution of our problem for small δ.

Finally, one derives the asymptotic formulas (A14) and (A15): let (ϕ, c, s) be the unique solution of (A1)–(A5). Due to Lemma A1, s and c satisfy the system (A9) and (A10).

Substitute $c = c(s)$ from (A10) into (A9), and replace ε by δ^3 in all equations. Then, one plugs into the modified Equation (A9) a power series ansatz for s, namely, $s = 1 + s_1\delta + \dots s_6\delta^6$. Developing both sides as power series in δ and comparing coefficients leads to (A15). This result is then plugged into (A10), yielding, quite similarly, (A14). □

Remark A2. *Using the formulae (A12), the asymptotics (42) for the trading boundaries $\widetilde{\zeta}_\pm$ in terms of the risky-safe ratio follow from the asymptotics of Proposition A1. The asymptotics (44) for $\widetilde{\pi}_\pm$ then follows from the relationship (A13).*

Notes

1 Some tedious computations in this paper where performed by MATHEMATICA. For motivating this research topic and providing feedback, I am indebted to Professor Paolo Guasoni.

2 More generally, the term *market frictions* encompasses, for example, price impact, short-selling constraints, and margin requirements (see Guasoni and Muhle-Karbe (2013), Guasoni and Weber (2020) and Guasoni et al. (2023) and the references therein).

3 Example 1 below shows this failure for a variance swap hedge.

4 These references deal with particularly tractable, long-run problems of local or global utility maximization; however, the first papers in this field, starting with Magill and Constantinides (1976), where optimal investment and consumption problems on an infinite horizon, which exhibit similar strategies and asymptotics. For an overview of this research field, see (Guasoni and Mayerhofer 2019, Chapter 1) and Guasoni and Muhle-Karbe (2013).

5 For these strategies, the name "control limit policy" from Taksar et al. (1988) is adopted, see Definition 2 below.

6 When $\gamma = 1$, the local-mean variance objective agrees with logarithmic utility, for which monotonicity holds and the shadow market strategy is the optimal one cf. Gerhold et al. (2013).

7 For the dynamics of the wealth process, see Lemma 1 below.

8 This follows from the respective finite-horizon objective (20), expressed in terms of π_t and φ_t, see Lemma 1.

9 It is well-known that a variance swap with maturity T on a continuous semimartingale S can be perfectly hedged by holding $2/(TS_t)$ units of the underlying at time $t \leq T$ (the dynamic hedging term), and a static portfolio of European puts and calls with expiry T, Bossu et al. (2005).

10 By ergodicity, the strategy that makes bulk trades into the middle of the optimal no-trade region incurs average transaction costs of higher order, namely proportional to $\varepsilon^{1/3}$. (Compare the ATC (28) which is of second order.)

11 The product rule gives

$$\frac{d\widetilde{S}_t}{\widetilde{S}_t} = \frac{dS_t}{S_t} + \frac{dg}{g} + \frac{d\langle S_t, g\rangle}{S_t g} =: (\widetilde{\mu}_t + r)dt + \widetilde{\sigma}_t dB_t,$$

from which the particular form of drift and diffusion coefficients (32), (33) can be computed.

12 For the details leading to this and other asymptotics, see Appendix A, Proposition A1 and Remark A2.

13 This expression is readily obtained from (11) by expanding (38) into formal power series in $\varepsilon^{1/3}$.

14 The general form of drift and diffusion coefficients follows from the typical smooth pasting conditions $g'(\pi_\pm) = 0$, along the same arguments as in Section 3.1 that turn (32) into (35), by removing local-time terms.

15 More precisely, certain portfolio statistics, such as π_t or ζ_t, exhibit stationarity.

16 That this second order discrepancy is not essential, can be seen also by a numerical robustness check, with trades at daily frequency and with a finite time horizon of, say five years. Numerical examples are already elaborated for a similar objective in great detail in (Guasoni and Mayerhofer 2023, Section 6 (Figures 4 and 5)).

17 This assertion can be proven using the same method as in Theorem 2.

18 Note that we use portfolio returns, as opposed to changes of wealth in Martin (2012, 2016). Besides, Martin's work cares about asymptotic optimality at lowest order, similar to Kallsen and Muhle-Karbe (2017).

19 Such a general representation bears the advantage that the stochastic process $\widetilde{S}_t := g(\pi_t)S_t$ could be interpreted as a (candidate) shadow price.

20 (Taksar et al. 1988, Theorem 6.16) appears to be an exception, which does not refer to te smallness of transaction costs.

21 Similar methods to derive asymptotic expansions in small transaction costs are found in the papers Gerhold et al. (2012, 2014); Guasoni and Mayerhofer (2019, 2023).

References

Borodin, Andrei N., and Paavo Salminen. 2002. *Handbook of Brownian Motion: Facts and Formulae*. Berlin and Heidelberg: Springer.

Bossu, Sebastien, Eva Strasser, and Regis Guichard. 2005. Just what you need to know about variance swaps. *JPMorgan Equity Derivatives Report* 4: 1–29.

Czichowsky, Christoph, and Walter Schachermayer. 2016. Duality theory for portfolio optimisation under transaction costs. *Annals of Applied Probability* 26: 1888–941. [CrossRef]

Gerhold, Stefan, Johannes Muhle-Karbe, and Walter Schachermayer. 2012. Asymptotics and duality for the Davis and Norman problem. *Stochastics An International Journal of Probability and Stochastic Processes* 84: 625–41. [CrossRef]

Gerhold, Stefan, Johannes Muhle-Karbe, and Walter Schachermayer. 2013. The dual optimizer for the growth-optimal portfolio under transaction costs. *Finance and Stochastics* 17: 325–54. [CrossRef]

Gerhold, Stefan, Paolo Guasoni, Johannes Muhle-Karbe, and Walter Schachermayer. 2014. Transaction costs, trading volume, and the liquidity premium. *Finance and Stochastics* 18: 1–37. [CrossRef]

Guasoni, Paolo, and Eberhard Mayerhofer. 2019. The limits of leverage. *Mathematical Finance* 29: 249–84. [CrossRef]

Guasoni, Paolo, and Eberhard Mayerhofer. 2023. Leveraged funds: Robust replication and performance evaluation. *Quantitative Finance* 23: 1155–76. [CrossRef]

Guasoni, Paolo, and Johannes Muhle-Karbe. 2013. Portfolio choice with transaction costs: A user's guide. In *Paris-Princeton Lectures on Mathematical Finance 2013*. Cham: Springer, vol. 2081, pp. 169–201.

Guasoni, Paolo, and Marko Hans Weber. 2020. Nonlinear price impact and portfolio choice. *Mathematical Finance* 30: 341–76. [CrossRef]

Guasoni, Paolo, Eberhard Mayerhofer, and Mingchuan Zhao. 2023. Options Portfolio Selection with Position Limits. *Michael J. Brennan Irish Finance Working Paper Series Research Paper* 22: 1–33. [CrossRef]

Guasoni, Paolo, Miklós Rásonyi, and Walter Schachermayer. 2010. The fundamental theorem of asset pricing for continuous processes under small transaction costs. *Annals of Finance* 6: 157–91. [CrossRef]

Herdegen, Martin, David Hobson, and Alex S. L. Tse. 2023. *Optimal Investment and Consumption with Epstein-Zin Stochastic Differential Utility and Proportional Transaction Costs*. Private communication.

Kabanov, Yuri, Miklós Rásonyi, and Christophe Stricker. 2002. No-arbitrage criteria for financial markets with efficient friction. *Finance and Stochastics* 6: 371–82. [CrossRef]

Kallsen, Jan, and Johannes Muhle-Karbe. 2010. On using shadow prices in portfolio optimization with transaction costs. *Annals of Applied Probability* 20: 1341–58. [CrossRef]

Kallsen, Jan, and Johannes Muhle-Karbe. 2017. The general structure of optimal investment and consumption with small transaction costs. *Mathematical Finance* 27: 695–703. [CrossRef]

Magill, Michael J. P., and George M. Constantinides. 1976. Portfolio selection with transactions costs. *Journal of Economic Theory* 13: 245–63. [CrossRef]

Martin, Richard J. 2012. Optimal multifactor trading under proportional transaction costs. *arXiv* arXiv:1204.6488.

Martin, Richard J. 2016. Universal trading under proportional transaction costs. *arXiv* arXiv:1603.06558.

Taksar, Michael, Michael J. Klass, and David Assaf. 1988. A diffusion model for optimal portfolio selection in the presence of brokerage fees. *Mathematics of Operations Research* 13: 277–94. [CrossRef]

Tanaka, Hiroshi. 1979. Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Mathematical Journal* 9: 163–77. [CrossRef]