Mathematics and Financial Economics (2026) 20:171–202
https://doi.org/10.1007/s11579-025-00411-4

Check for updates

Adaptive-Robust Portfolio Optimisation

Theerawat Bhudisaksang[1] · Álvaro Cartea[2,3] · Leandro Sánchez-Betancourt[2,3]

Received: 23 January 2025 / Accepted: 4 December 2025 / Published online: 16 December 2025
© The Author(s) 2025

Abstract
An agent solves an exponential utility maximisation problem that is robust to parameter misspecification and where the optimal strategy continuously adapts to new information. The agent invests in a risk-free asset and in risky stocks whose prices follow geometric diffusion processes. The agent does not know the drift parameters of the stock price dynamics, so she considers a set of alternative measures to make the investment problem robust to model misspecification and employs a continuous-time estimator to learn the value of the drift parameters as new information arrives during the investment horizon. For the two risky asset case, the agent's value function is characterised as the solution to a non-linear PDE. We show that the value function has a stochastic representation and use it to analyse the optimal adaptive-robust strategy and to compare it with various benchmarks.

Keywords Adaptive-robust control · Model uncertainty · Stochastic control · Time-consistency · Dynamic programming · Online learning · Algorithmic trading

1 Introduction

A classical problem in finance consists in finding the optimal investment strategy for an agent who maximises the expected utility of terminal wealth, where the agent's wealth dynamics are determined by the evolution of the value of her investment portfolio, see, e.g., [14]. Generally, the dynamics of the assets are modelled by a set of stochastic differential equations (SDEs) and one assumes that the agent knows the dynamics of the assets, which includes knowing the parameters of the SDEs. For example, when the price dynamics are described by a geometric Brownian motion, the agent knows both the volatility and the drift of the returns of the assets.

In the classical setup, the agent uses the price dynamics to derive the optimal investment strategy. However, in practice, one does not know the true dynamics of the asset nor the value of the model parameters, i.e., the value of the parameters of the SDEs that describe the evolution of prices are unkown. In this paper, the risky assets follow a geometric Brownian

TB acknowledges support of the Second Century Fund (C2F), Chulalongkorn University and the Mathematical Institute, University of Oxford.

✉ Leandro Sánchez-Betancourt
sanchezbetan@maths.ox.ac.uk

1 Department of Mathematics and Computer Science, Chulalongkorn University, Bangkok, Thailand

2 Mathematical Institute,University of Oxford, Oxford, England

3 Oxford-Man Institute of Quantitative Finance, Oxford, England

motion, of which the agent knows the volatility of returns but does not know the drift of the returns of the asset. We employ the continuous-time adaptive-robust framework in [2] — based on the discrete-time framework first proposed in [4] — to derive the agent's optimal strategy when the drift parameters of the stock processes are unknown.[1] Here, we apply the theory of adaptive robust control developed in [2, 3] to the Merton optimal investment problem and derive a tractable characterisation of the value function with which we provide insights about the role of the adaptive and robust components of the strategy. In the adaptive component of this approach, the agent employs a continuous-time estimator to learn the value of the drift parameters so that the control problem adapts to the arrival of new information throughout the agent's trading horizon. On the other hand, in the robust component of this approach, the agent considers a set of alternative measures to make the control problem robust to model misspecification that arise from not knowing the drift of the risky assets.

For the two risky asset case, the agent's value function is characterised by a non-linear partial differential equation (PDE), and we show that the value function has a stochastic representation which one evaluates by numerically computing a deterministic integral. This calculation bypasses the use of numerical methods to solve the PDE directly, which is usually not feasible in high-dimensional problems such as the one we study. With this representation, we analyse the agent's optimal adaptive-robust investment strategy. The optimal adaptive-robust investment strategy depends on the values $(\hat{\theta}_1, \hat{\theta}_2)$ of the adaptive estimates of the two unknown drifts. We find that there are nine regions in the $(\hat{\theta}_1, \hat{\theta}_2)$ space that modulate the optimal strategy. We show how these investment regions change as the correlation between assets change. More precisely, when the correlation is positive (resp. negative), we show how the area corresponding to long-short/short-long (resp. long-long/short-short) positions in the first-second assets expands. We build our understanding of the optimal strategy under a different set of assumptions on the volatility of the underlying assets and we study, numerically, the performance of the adaptive-robust strategy.

Our work is related to two strands of the literature that focus on either parameter uncertainty or learning model parameters, while our paper includes both robustness and learning simultaneously to solve the agent's investment problem. In [6] the agent maximises expected exponential utility of wealth without perfect knowledge of the drift parameter in the stock price dynamics. [15] develop general results on utility maximisation under parameter uncertainty in the drift, volatility, and jump terms in the dynamics of stock prices. [12] study utility maximisation problems where parameters lie in a time-varying interval. These articles focus on parameter uncertainty in the agent's utility maximisation problem and, in particular, the agent cannot estimate the relevant parameters as time evolves during the trading horizon. In work that focuses on learning, [5] consider the agent's utility maximisation problem with both Bayesian online updates of model parameters and derive a PDE that characterises the agent's value function. [9] study the Markowitz portfolio selection problem where the agent does not know the drift vector of stock prices. In this paper we study both, continuous-time adaptivity and robustness in the Merton problem.

The remainder of the paper is organised as follows. Section 2 presents the utility maximisation problem in the adaptive-robust framework and we derive the main theoretical results of the paper. In particular, we show that the value function of the problem has a stochastic representation. Section 4 discusses the optimal strategy as a function of the correlation between the returns of the assets. Section 5 conducts asymptotic analysis of the adaptive-robust strategy when stock prices are independent and Section 6 showcases the performance of the adaptive-robust framework and compares it with various benchmarks.

[1] For recent developments see [1, 8].

2 Portfolio investment strategy

The agent solves a dynamic asset allocation problem where shortselling is allowed. The agent trades in two risky assets and in a risk-free asset to maximise the expected utility of terminal wealth. Let r denote the risk-free rate and let $S_t = (S_t^1,\ S_t^2)$ denote the prices of the risky assets at time t, which satisfy the SDEs

$$\begin{aligned} dS_t^1 &= \theta^{*,1}\, S_t^1\, dt + \sigma_{11}\, S_t^1\, dW_t^1 + \sigma_{12}\, S_t^1\, dW_t^m\,, \\ dS_t^2 &= \theta^{*,2}\, S_t^2\, dt + \sigma_{21}\, S_t^2\, dW_t^2 + \rho\, \sigma_{22}\, S_t^2\, dW_t^m\,, \end{aligned} \tag{2.1}$$

under the probability measure $\mathbb{P}^*$. Here, W^1 and W^2 are independent Brownian motions that denote idiosyncratic shocks to the prices of the assets, and $\rho \in [-1, 1]$ is a correlation parameter. The Brownian motion W^m, independent of W^1 and W^2, represents the market risk in the price innovations of both assets. We work with two risky assets because one can derive explicit characterisations of the trading regions and study the investment strategies. Although the analysis can be extended to n assets, the expressions would become much more convoluted.

The agent knows that the dynamics of the prices are given by (2.1), knows the value of the volatility parameters $\sigma_{ij} > 0$ for $i, j \in \{1, 2\}$, and knows the value of the correlation parameter ρ. However, the agent does not know the values of the drift parameters $\theta^{*,1}$ and $\theta^{*,2}$ which we denote by $\theta^* = [\theta^{*,1} \quad \theta^{*,2}]^\intercal$, where $\intercal$ is the transpose operator. If one uses quadratic variation arguments, the volatility parameters are known because asset prices are observed continuously. Thus, in this paper we focus on the continuous estimation of the drift parameters.

The agent's controlled wealth process is denoted by X_t^α and satisfies the SDE

$$dX_t^\alpha = (\alpha_t^\intercal\, (\theta^* - \mathbf{1}\, r) + r\, X_t^\alpha)\, dt + \alpha_t^1\, \sigma_{11}\, dW_t^1 + \alpha_t^1\, \sigma_{12}\, dW_t^m + \alpha_t^2\, \sigma_{21}\, dW_t^2 + \alpha_t^2\, \rho\, \sigma_{22}\, dW_t^m\,, \tag{2.2}$$

where the control α_t^i is the amount of money invested in asset i, $\alpha_t = [\alpha_t^1,\ \alpha_t^2]^\intercal$, and $\mathbf{1} = [1,\ 1]^\intercal$.

In what follows, we focus on the optimal investment problem without transaction costs and study the case of exponential utility. We expect that the framework and results could be extended to account for transaction costs, in which cases we expect similar insights; see [10, 17] for works that incorporate small transaction costs to the Merton problem.

The adaptive-robust framework we develop benefits from efficient computation times because the joint process of the state variables and the adaptive estimator remains Markovian, with which we characterise the value function via a Hamilton–Jacobi–Bellman (HJB) equation. This representation enables tractable numerical schemes, even in higher dimensions, subject to the usual trade-offs in discretization. In contrast, approaches based on offline estimators (i.e., estimators computed from fixed historical samples and not updated continuously) typically result in non-Markovian dynamics when coupled with the controlled state process. As a result, the associated stochastic control problems lack an HJB characterisation, making them computationally more challenging and limiting the applicability of standard dynamic programming techniques. For this specific problem with a higher number of assets (more than two), the defined subspace becomes more complicated and harder to state explicitly.

The agent acknowledges that she does not know the true value of the drift parameters. Hence, the agent employs the adaptive-robust methodology — proposed by [4] for the discrete-time setting and extended to continuous-time in [3] — to derive an investment strategy. The key advantage of the adaptive-robust approach is that the agent uses the evolution of the underlying stochastic price processes to update the estimates of the unknown parameters continuously while ensuring that the decisions are robust to model misspecification.

The agent is risk-averse and derives utility from wealth. Her preferences are described by the exponential utility function $U(x) = -\exp(-\gamma\, x)$, with $\gamma > 0$, which is increasing and concave, and bounded from above. The agent's adaptive-robust problem is given by the value function

$$v(t, x, S, \hat{\theta}) = \sup_{\alpha \in \mathcal{A}} \inf_{\mathbb{P} \in \mathcal{P}(t,x,S,\hat{\theta},G)} \mathbb{E}^{\mathbb{P}} \left[U(X_T^{\alpha}) \right] , \tag{2.3}$$

where $X_t = x$, $S_t = S$, $\mathcal{A}$ is the set of admissible control processes (progressively measurable and square integrable), and $T > 0$ denotes the terminal date of the agent's investment horizon. Here, $\mathbb{E}^{\mathbb{P}}[\,\cdot\,]$ is the expectation operator under the measure $\mathbb{P}$, which is a probability measure in the set of alternative measures $\mathcal{P}(t, x, S, \hat{\theta}, G)$ considered by the investor. In particular, the deterministic function $G(t, \hat{\theta})$ specifies the investor's model uncertainty that stems from the estimation process of the unknown drift parameters.

Next, we describe the drift estimator process for stock $i \in \{1, 2\}$. First, over a time-step $\Delta t > 0$ the expected return of the stock under the true measure $\mathbb{P}^*$ is given by

$$\mathbb{E}^{\mathbb{P}^*} \left[\frac{S^i_{t+\Delta t} - S^i_t}{S^i_t} \right] \approx \theta^{*,i}\, \Delta t\, . \tag{2.4}$$

Indeed, use (2.1) to see that for an integrable deterministic function $g(t)$ one has

$$\mathbb{E}^{\mathbb{P}^*} \left[\int_0^t \frac{g(u)}{S^i_u}\, dS^i_u \right] = \theta^{*,i} \int_0^t g(u)\, du\, , \tag{2.5}$$

and from the equation above, solve for $\theta^{*,i}$ to see that

$$\frac{1}{\int_0^t g(u)\, du} \int_0^t \frac{g(u)}{S^i_u}\, dS^i_u \tag{2.6}$$

is an unbiased estimator of the drift parameter $\theta^{*,i}$. Here, we choose $g(u) = (1 + u)^{L-1}$, where $L > 0$ is a learning parameter, and write

$$\hat{\theta}^i_t = \frac{L}{(1+t)^L} \int_0^t \frac{(1+u)^{L-1}}{S^i_u}\, dS^i_u \tag{2.7}$$

for $i = 1, 2$, where $\hat{\theta}^i$ denotes the estimate of the drift of asset i. The choice of the function $g(u)$ ensures that the estimator in (2.7) converges as $t \to \infty$; see Proposition 2.1 below. Then, by Itô's lemma, the estimators of the drift parameters in (2.7) are

$$\begin{aligned} d\hat{\theta}^1_t &= \beta_t \left((\theta^{*,1} - \hat{\theta}^1_t)\, dt + \sigma_{11}\, dW^1_t + \sigma_{12}\, dW^m_t \right) , \\ d\hat{\theta}^2_t &= \beta_t \left((\theta^{*,2} - \hat{\theta}^2_t)\, dt + \sigma_{21}\, dW^2_t + \rho\, \sigma_{22}\, dW^m_t \right) , \end{aligned} \tag{2.8}$$

under the true probability $\mathbb{P}^*$, where the learning rate is

$$\beta_t = \frac{L}{1+t}\, . \tag{2.9}$$

The choice of the function g is not arbitrary. Note that for a general deterministic function g, one can rearrange (2.6) to obtain

$$\begin{aligned} d\hat{\theta}_t^1 &= \frac{g(t)}{G(t)} \left((\theta^{*,1} - \hat{\theta}_t^1)\, dt + \sigma_{11}\, dW_t^1 + \sigma_{12}\, dW_t^m \right), \\ d\hat{\theta}_t^2 &= \frac{g(t)}{G(t)} \left((\theta^{*,2} - \hat{\theta}_t^2)\, dt + \sigma_{21}\, dW_t^2 + \rho\, \sigma_{22}\, dW_t^m \right), \end{aligned} \tag{2.10}$$

where $G(t) = \int_0^t g(s)\, ds$. From the result in [2], the learning rate must satisfy

$$\int_0^\infty \frac{g(t)}{G(t)} dt = \infty \qquad \text{and} \qquad \int_0^\infty \frac{g^2(t)}{G^2(t)} dt < \infty$$

to allow for both the coverage of the domain space and the convergence of the estimator. The choice of $g(t) = (1+t)^{L-1}$ is a tractable example that satisfies both conditions.

Proposition 2.1 *Let $i \in \{1, 2\}$ and let $\hat{\theta}_t^i$ be given in (2.7), then, as $t \to \infty$, the estimator $\hat{\theta}_t^i$ converges to $\theta^{*,i}$ in probability. The distribution of the random variable $\hat{\theta}_t^i$ for $i = 1, 2$ is normal with mean value*

$$\hat{\theta}_0^i \left(\frac{1}{1+t} \right)^L + \theta^{*,i} \left(1 - \frac{1}{(1+t)^L} \right)$$

and when $L \neq 1/2$, the variance is

$$(\sigma_{i\,1}^2 + I_i^2\, \sigma_{i\,2}^2)\, \frac{L^2}{2\,L-1} \left(\frac{(1+t)^{2\,L-1}}{(1+t)^{2\,L}} - \frac{(1)^{2\,L-1}}{(1+t)^{2\,L}} \right),$$

and when $L = 1/2$, the variance is

$$\frac{(\sigma_{i\,1}^2 + I_i^2\, \sigma_{i\,2}^2)}{4} \left(\frac{\log(1+t)}{1+t} \right).$$

For a proof see Appendix A.

We observe that as t becomes large, the variance of the estimator $\hat{\theta}_t^i$ behaves like $c/1+t$ where c is a scaling factor. Therefore, the uncertainty set should scale as $\hat{\theta}_t^i \pm c/\sqrt{1+t}$, where c corresponds to the desired confidence level.

The larger the value of L, the quicker the estimator converges to the true value as $t \to \infty$. But the estimator would fluctuate more when t is small. Thus, it should be up to the investor to decide what constitutes a good choice for L.

Next, to specify the set of alternative measures $\mathcal{P}(t, x, S, \hat{\theta}, G)$, we first introduce the set-valued function

$$G(t, \hat{\theta}) = \left\{ (\tilde{\theta}_1, \tilde{\theta}_2) \in \mathbb{R}^2 \mid \hat{\theta}_i - c/\sqrt{1+t} \le \tilde{\theta}_i \le \hat{\theta}_i + c/\sqrt{1+t} \quad \text{for } i = 1, 2 \right\}, \tag{2.11}$$

which represents the model uncertainty that stems from not knowing the true value of the drift parameter; here, the uncertainty parameter $c \ge 0$ specifies the width of the uncertainty set.

Remark 2.2 The agent chooses the parameter $c \ge 0$ to control the width of the uncertainty set. Heuristically, one uses the variance of the estimators $\hat{\theta}^i$ for $i = 1, 2$ and for a given L to choose c such that the estimators are within a ball around the true values of the parameters with a given confidence (e.g., 90%). See Appendix A for explicit formulae of the variances of the estimators which we can use to establish the link between c and L.

To solve the investment problem, the agent derives a Hamilton–Jacobi–Bellman–Isaacs (HJBI) equation to characterise the value function v. The characterisation is under the measure $\tilde{\mathbb{P}} \in \mathcal{P}(t, x, S, \hat{\theta}, G)$ because the agent searches over all strategies and outcomes that depend on all the models determined by the choice of $\tilde{\theta}$ in (2.11). This approach requires to write the estimator process in (2.8) under the measure $\tilde{\mathbb{P}}$, so we proceed as follows.

The set $\mathcal{P}(t, x, S, \hat{\theta}, G)$ is defined as follows. Each alternative measure $\tilde{\mathbb{P}} \in \mathcal{P}(t, x, S, \hat{\theta}, G)$ considered by the agent is associated with a progressively measurable process $\tilde{\theta}$ such that $\tilde{\theta}_u \in G(t, \hat{\theta}_u)$ for all $u \in [t, T]$ and

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}^*} = \exp\left(-\sum_{i=1}^{2}\int_0^t \frac{\theta^{*,i} - \tilde{\theta}^i_s}{\sigma_{i\,1}}\, dW^i - \frac{1}{2}\sum_{i=1}^{2}\int_0^t \left(\frac{\theta^{*,i} - \tilde{\theta}^i_s}{\sigma_{i\,1}}\right)^2 ds\right). \tag{2.12}$$

Then, by Girsanov's theorem, the process

$$d\tilde{W}^i_t = dW^i_t + \frac{\theta^{*,i}_t - \tilde{\theta}^i_t}{\sigma_{i\,1}}\, dt\,, \qquad \text{for} \quad i = 1,\, 2\,,$$

is a Brownian motion under the probability measure $\tilde{\mathbb{P}}$; see [16]. Therefore, under the measure $\tilde{\mathbb{P}}$, the estimators in (2.8) are given by

$$\begin{aligned} d\hat{\theta}^1_t &= \beta_t \left((\tilde{\theta}^1_t - \hat{\theta}^1_t)\, dt + \sigma_{11}\, d\tilde{W}^1_t + \sigma_{12}\, dW^m_t\right), \\ d\hat{\theta}^2_t &= \beta_t \left((\tilde{\theta}^2_t - \hat{\theta}^2_t)\, dt + \sigma_{21}\, d\tilde{W}^2_t + \rho\, \sigma_{22}\, dW^m_t\right), \end{aligned} \tag{2.13}$$

where $\tilde{W}^1$, $\tilde{W}^2$, and W^m are independent Brownian motions. Thus, equipped with (2.13), the value function in (2.3) is the solution of an HJBI for the alternative models $\tilde{\mathbb{P}} \in \mathcal{P}(t, x, S, \hat{\theta}, G)$, which include the continuous-time updates of the estimated drift parameters.

Note that for a fixed time t, when the value of the uncertainty parameter c in the function $G(t, \theta)$ is high (resp. low), the agent is less (resp. more) certain about the estimate of the drift. On the other hand, for a fixed value of $c > 0$, as time evolves, the intervals in (2.11) shrink because the agent is more confident about the estimate of the drift of prices as more data are used to compute each update of $\hat{\theta}_t$.

In contrast, when the agent fully trusts the estimator of the drift parameter, she fixes the value of the uncertainty parameter c to zero, so the set-valued function $G(t, \hat{\theta})$ is the set $\{(\hat{\theta}_1, \hat{\theta}_2)\}$. In this case, the agent fully commits to the value of each update of the estimate $\hat{\theta}^i_t$ in the model of price dynamics (2.1). Consequently, the agent's optimisation problem only considers the 'adaptive' part of (2.3) (there is no robustness actions because the agent fully trusts the estimates of the drift parameters), so the optimal investment strategy is the solution to the adaptive control problem

$$v(t, x, S, \hat{\theta}) = \sup_{\alpha \in \mathcal{A}_0} \mathbb{E}^{\hat{\mathbb{P}}}\left[U(X^\alpha_T)\right], \tag{2.14}$$

where the probability measure $\hat{\mathbb{P}}$ is defined by

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}^*} = \exp\left(\sum_{i=1}^{2} -\int_0^t \frac{\theta^{*,i} - \hat{\theta}^i_s}{\sigma_{i\,1}}\, dW^i - \frac{1}{2}\sum_{i=1}^{2}\int_0^t \left(\frac{\theta^{*,i} - \hat{\theta}^i_s}{\sigma_{i\,1}}\right)^2 ds\right). \tag{2.15}$$

The investment problem in (2.14) is only 'adaptive' because at every point in time, the agent searches for optimal strategies that use the continuous updates of the estimates of the drift parameter of the stock prices. In the ideal case in which the agent knows the true value of the drift parameter in (2.1), the investment problem in (2.3) reduces to the classical Merton problem

$$v(t, x, S, \hat{\theta}) = \sup_{\alpha \in \mathcal{A}_0} \mathbb{E}^{\mathbb{P}^*}\left[U(X_T^{\alpha})\right], \tag{2.16}$$

where $\mathbb{P}^*$ is the true probability measure. One can show (see e.g., Chapter 5 in [7]) that the Merton optimal investment strategy for the exponential utility case $U(x) = -e^{-\gamma x}$ is given by

$$\alpha_t^{*,M} = \frac{\Sigma^{-1}\,(\theta^* - \mathbf{1}\,r)}{\gamma\,\exp(r\,(T-t))}, \tag{2.17}$$

where Σ is the variance-covariance matrix returns.

In the general case, the uncertainty parameter c is greater than zero, and the agent solves the adaptive-robust problem in (2.3). In Appendix B we prove that the value function v associated with the above adaptive-robust control problem is finite, and by Proposition 2.19 in [2], one can show that the value function satisfies the HJBI equation

$$\begin{aligned} 0 = {} & \partial_t v + r\,x\,\partial_x v + \tfrac{1}{2}\,\beta_t^2\,\sigma_1^2\,\partial_{\hat{\theta}_1\hat{\theta}_1} v + \tfrac{1}{2}\,\beta_t^2\,\sigma_2^2\,\partial_{\hat{\theta}_2\hat{\theta}_2} v + \beta_t^2\,\rho\,\sigma_c^2\,\partial_{\hat{\theta}_1\hat{\theta}_2} v \\ & + \sup_{\alpha} \inf_{\tilde{\theta} \in G(t,\hat{\theta})} \left\{ \beta_t\,(\tilde{\theta} - \hat{\theta})^{\intercal}\,\partial_{\hat{\theta}} v + \alpha^{\intercal}\,(\tilde{\theta} - \mathbf{1}\,r)\,\partial_x v + \tfrac{1}{2}\,\alpha^{\intercal}\,\Sigma\,\alpha\,\partial_{xx} v + \beta_t\,\alpha^{\intercal}\,\Sigma\,\partial_{x\hat{\theta}}\,v \right\}, \end{aligned} \tag{2.18}$$

where

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\,\sigma_c^2 \\ \rho\,\sigma_c^2 & \sigma_2^2 \end{bmatrix}, \quad \text{with} \quad \sigma_1^2 = \sigma_{11}^2 + \sigma_{12}^2, \quad \sigma_2^2 = \sigma_{21}^2 + \rho^2\,\sigma_{22}^2, \quad \text{and} \quad \sigma_c^2 = \sigma_{12}\,\sigma_{22}. \tag{2.19}$$

Here, Σ is the variance-covariance matrix of the returns of assets S^1 and S^2 and we recall that $\rho \in [-1, 1]$. Note that the correlation between the prices of the two risky assets is $\rho\,\sigma_c^2/\sqrt{\sigma_1\,\sigma_2}$ and that the HJBI does not depend on the true value of the drift parameter θ^*. To simplify notation, we define

$$\mathbf{1}^{+,+} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{1}^{-,-} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \mathbf{1}^{+,-} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{1}^{-,+} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \mathbf{1}^{+}=1, \quad \mathbf{1}^{-}=-1.$$

The investment strategy is a function of various quantities, including the value of model parameters; e.g., the correlation parameter ρ, the value of the uncertainty parameter c, and the estimate of the drift processes. Thus, to streamline the discussion, we delineate nine non-overlapping investment regions in $\mathbb{R}^2$ which we use below to derive the agent's investment strategy.

The first four regions, denoted by $A^{+,+}$, $A^{+,-}$, $A^{-,+}$, and $A^{-,-}$, are given by

$$A^{j,k} = \left\{ \hat{\theta} \in \mathbb{R}^2 \mid \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{-j,-k}\,c/\sqrt{1+t}\right) \in \mathbb{R}^{j,k} \right\} \qquad \text{for} \quad j,k \in \{+,-\},$$

where for $j,k \in \{+,-\}$ we use the notation $\mathbb{R}^{j,k} = \mathbb{R}^j \times \mathbb{R}^k$, for which $\mathbb{R}^+ = (0,\infty)$ and $\mathbb{R}^- = (-\infty, 0)$.

The remaining five regions are $A^{0,-}$, $A^{0,+}$, $A^{-,0}$, $A^{+,0}$, and $A^{0,0}$. In region $A^{0,-}$, $\hat{\theta}_2 - r + c/\sqrt{1+t} \le 0$ and

$$\frac{\rho\,\sigma_c^2}{\sigma_2^2}\,(\hat{\theta}_2 - r + c/\sqrt{1+t}) + r - c/\sqrt{1+t} \le \hat{\theta}_1 \le \frac{\rho\,\sigma_c^2}{\sigma_2^2}\,(\hat{\theta}_2 - r + c/\sqrt{1+t}) + r + c/\sqrt{1+t};$$

in region $A^{0,+}$, $\hat\theta_2 - r - c/\sqrt{1+t} \ge 0$ and

$$\frac{\rho\,\sigma_c^2}{\sigma_2^2}\,(\hat\theta_2 - r - c/\sqrt{1+t}) + r - c/\sqrt{1+t} \le \hat\theta_1 \le \frac{\rho\,\sigma_c^2}{\sigma_2^2}\,(\hat\theta_2 - r - c/\sqrt{1+t}) + r + c/\sqrt{1+t}\,;$$

in region $A^{-,0}$, $\hat\theta_1 - r + c/\sqrt{1+t} \le 0$ and

$$\frac{\rho\,\sigma_c^2}{\sigma_1^2}\,(\hat\theta_1 - r + c/\sqrt{1+t}) + r - c/\sqrt{1+t} \le \hat\theta_2 \le \frac{\rho\,\sigma_c^2}{\sigma_1^2}\,(\hat\theta_1 - r + c/\sqrt{1+t}) + r + c/\sqrt{1+t}\,;$$

in region $A^{+,0}$, $\hat\theta_1 - r - c/\sqrt{1+t} \ge 0$ and

$$\frac{\rho\,\sigma_c^2}{\sigma_1^2}\,(\hat\theta_1 - r - c/\sqrt{1+t}) + r - c/\sqrt{1+t} \le \hat\theta_2 \le \frac{\rho\,\sigma_c^2}{\sigma_1^2}\,(\hat\theta_1 - r - c/\sqrt{1+t}) + r + c/\sqrt{1+t}\,;$$

and in region $A^{0,0}$,

$$r - c/\sqrt{1+t} \le \hat\theta_1 \le r + c/\sqrt{1+t} \quad \text{and} \quad r - c/\sqrt{1+t} \le \hat\theta_2 \le r + c/\sqrt{1+t}\,.$$

Finally, define the function

$$F(t,\hat\theta) = \begin{cases} \frac{1}{2}\left(\hat\theta - \mathbf{1}r + \mathbf{1}^{-j,-k}\,\frac{c}{\sqrt{1+t}}\right)^{\intercal}\Sigma^{-1}\left(\hat\theta - \mathbf{1}r + \mathbf{1}^{-j,-k}\,\frac{c}{\sqrt{1+t}}\right), & \hat\theta \in A^{j,k}\,,\; j,k \in [+,-]\,, \\ \frac{1}{2}\left(\hat\theta_1 - r + \mathbf{1}^{-j}\,\frac{c}{\sqrt{1+t}}\right)\frac{1}{\sigma_1^2}\left(\hat\theta_1 - r + \mathbf{1}^{-j}\,\frac{c}{\sqrt{1+t}}\right), & \hat\theta \in A^{j,0}\,,\; j \in [+,-]\,, \\ \frac{1}{2}\left(\hat\theta_2 - r + \mathbf{1}^{-j}\,\frac{c}{\sqrt{1+t}}\right)\frac{1}{\sigma_2^2}\left(\hat\theta_2 - r + \mathbf{1}^{-j}\,\frac{c}{\sqrt{1+t}}\right), & \hat\theta \in A^{0,j}\,,\; j \in [+,-]\,, \\ 0\,, & \hat\theta \in A^{0,0}\,, \end{cases} \tag{2.20}$$

which we use below to write the optimal investment strategy of the agent.

Proposition 2.3 ***Adaptive-robust strategy.*** *Let $U(x) = -e^{-\gamma x}$ be the agent's utility function, where $\gamma > 0$ is the risk-aversion parameter. Let the stock prices satisfy the SDEs in (2.1), the learning parameter $L > 0$, and investment horizon $T > 0$. Then, the adaptive-robust investment strategy, in feedback form, for the investor's problem in (2.3) is*

$$\alpha_t^* = \begin{cases} \dfrac{\Sigma^{-1}\left(\hat\theta - \mathbf{1}r + \mathbf{1}^{-j,-k}\,c/\sqrt{1+t}\right) - \beta_t\;\partial_{\hat\theta}\bar H(t,\hat\theta)}{\gamma\;\exp(r\,(T-t))}\,, & \hat\theta \in A^{j,k}\,,\; j,k \in [+,-]\,, \\ \dfrac{\left[\hat\theta_1 - r + \mathbf{1}^{-j}\,c/\sqrt{1+t}\quad 0\right]^{\intercal} - \sigma_1^2\,\beta_t\;\partial_{\hat\theta}\bar H(t,\hat\theta)}{\sigma_1^2\,\gamma\;\exp(r\,(T-t))}\,, & \hat\theta \in A^{j,0}\,,\; j \in [+,-]\,, \\ \dfrac{\left[0\quad \hat\theta_2 - r + \mathbf{1}^{-j}\,c/\sqrt{1+t}\right]^{\intercal} - \sigma_2^2\,\beta_t\;\partial_{\hat\theta}\bar H(t,\hat\theta)}{\sigma_2^2\,\gamma\;\exp(r\,(T-t))}\,, & \hat\theta \in A^{0,j}\,,\; j \in [+,-]\,, \\ \dfrac{-\beta_t\;\partial_{\hat\theta}\bar H(t,\hat\theta)}{\gamma\;\exp(r\,(T-t))}\,, & \hat\theta \in A^{0,0}\,, \end{cases} \tag{2.21}$$

where $\hat\theta = [\hat\theta_1 \quad \hat\theta_2]^{\intercal}$ is the estimate of the drift of the stock returns at time t. The stochastic representation of the function $\bar H$ is

$$\bar H(t,\hat\theta) = \mathbb{E}^{\mathbb{P}^*}\left[\int_t^T F(u, Z_u)\,du \;\middle|\; Z_t = \hat\theta\right], \tag{2.22}$$

with F in (2.20), where the stochastic process $Z = [Z^1 \quad Z^2]^{\intercal}$, with $Z_t = \hat\theta$, follows

$$\begin{aligned} dZ_u^1 &= \beta_t\left((r - Z_u^1)\,du + \sigma_{11}\;dW_u^1 + \sigma_{12}\;dW_u^m\right), \\ dZ_u^2 &= \beta_t\;\left((r - Z_u^2)\,du + \sigma_{21}\;dW_u^2 + \rho\,\sigma_{22}\;dW_u^m\right)\,. \end{aligned} \tag{2.23}$$

Here, W^1, W^2, and W^m are mutually independent Brownian motions under the probability measure $\mathbb{P}^*$ (see price dynamics in (2.1)), and β_t is the learning rate in (2.9). Conditional on $Z_t = \hat{\theta}$, the process $Z_u \in \mathbb{R}^2$ is normally distributed under the probability measure $\mathbb{P}^*$. For all values of the learning parameter L, the mean of Z_u^i is

$$\hat{\theta}_i \left(\frac{1+t}{1+u}\right)^L + r\left(1 - \frac{(1+t)^L}{(1+u)^L}\right). \tag{2.24}$$

When $L \neq 1/2$, the covariance matrix of Z_u is

$$\Sigma \int_t^u L^2 \, \frac{(1+v)^{2L-2}}{(1+u)^{2L}} \, dv = \Sigma \, \frac{L^2}{2L-1} \left(\frac{(1+u)^{2L-1}}{(1+u)^{2L}} - \frac{(1+t)^{2L-1}}{(1+u)^{2L}}\right), \tag{2.25}$$

and when $L = 1/2$, the covariance matrix of Z_u is

$$\Sigma \int_t^u L^2 \, \frac{(1+v)^{2L-2}}{(1+u)^{2L}} \, dv = \frac{\Sigma}{4} \left(\frac{\log(1+u) - \log(1+t)}{1+u}\right), \tag{2.26}$$

where Σ is in (2.19).

Proof Recall that the value function v satisfies the HJBI in (2.18), subject to the terminal condition $v(T, x, \hat{\theta}) = -e^{-\gamma x}$. Substitute the ansatz

$$v(t, x, \hat{\theta}) = -\,U(t, \hat{\theta}) \, \exp\left(-\gamma \, x \, \exp\left(r\,(T - t)\right)\right)$$

into the HJBI in (2.18) to write

$$\begin{aligned} 0 = {} & \partial_t U + \tfrac{1}{2}\,\beta_t^2\,\sigma_1^2\,\partial_{\hat{\theta}_1\hat{\theta}_1} U + \tfrac{1}{2}\,\beta_t^2\,\sigma_2^2\,\partial_{\hat{\theta}_2\hat{\theta}_2} U + \beta_t^2\,\rho\,\sigma_c^2\,\partial_{\hat{\theta}_1\hat{\theta}_2} U \\ & + \inf_{\alpha} \sup_{\tilde{\theta}\in G(t,\hat{\theta})} \left\{\beta_t\,(\tilde{\theta} - \hat{\theta})^\intercal\,\partial_{\hat{\theta}} U - \alpha^\intercal\,(\tilde{\theta} - \mathbf{1}\,r)\,U + \tfrac{1}{2}\,\alpha^\intercal\,\Sigma\,\alpha\,U - \beta_t\,\alpha^\intercal\,\Sigma\,\partial_{\hat{\theta}}\,U\right\} \end{aligned} \tag{2.27}$$

with terminal condition $U(T, \hat{\theta}) = 1$.

Next, we discuss the inf-sup solutions to determine the optimal investment strategy. The term inside the sup sign depends on $\tilde{\theta}^\intercal\,(\beta_t\,\partial_{\hat{\theta}} U - \alpha\,U)$, where $\tilde{\theta} \in G(t, \hat{\theta})$; therefore, we consider the value of $\beta_t\,\partial_{\hat{\theta}} U - \alpha\,U$ in different investment regions of $\mathbb{R}^2$, see Appendix C.1 for the detailed calculation.

Now, we solve the inf-sup problem in (2.27) for each $\hat{\theta}$ in the nine non-overlapping investment regions of $\mathbb{R}^2$ discussed above. Let $U(t, \hat{\theta}) = \exp(-H(t, \hat{\theta}))$ and let $C(t, \hat{\theta})$ denote the inf-sup term in (2.27). Then, the optimal investment strategy depends on nine regions of $\mathbb{R}^2$ — see Appendix C.2 and Appendix C.3 for more details. Moreover, the optimal investment strategy satisfies (2.21) and

$$C(t, \hat{\theta}) = -U\left(\beta_t\,(\mathbf{1}\,r - \hat{\theta})^\intercal\,\partial_{\hat{\theta}} H + F(t, \hat{\theta})\right), \tag{2.28}$$

for all t and $\hat{\theta}$. Use (2.27) for $\hat{\theta} \in \mathbb{R}^2$ and recall the terminal condition $H(T, \hat{\theta}) = 0$ to write

$$\partial_t H + \tfrac{1}{2}\,\beta_t^2\,\sigma_1^2\,\partial_{\hat{\theta}_1\hat{\theta}_1} H + \tfrac{1}{2}\,\beta_t^2\,\sigma_2^2\,\partial_{\hat{\theta}_2\hat{\theta}_2} H + \beta_t^2\,\rho\,\sigma_c^2\,\partial_{\hat{\theta}_1\hat{\theta}_2} H + \beta_t\,(\mathbf{1}\,r - \hat{\theta})^\intercal\,\partial_{\hat{\theta}} H + F(t, \hat{\theta}) = 0, \tag{2.29}$$

where F is in (2.21).

To show that the functions H and $\bar{H}$ in (2.22) are equal, we define the stochastic process $Z = [Z^1 \quad Z^2]^\intercal$, which follows

$$dZ_u^1 = \beta_t\left((r - Z_u^1)\,du + \sigma_{11}\,dW_u^1 + \sigma_{12}\,dW_u^m\right),$$

$$dZ_u^2 = \beta_t \left((r - Z_u^2)\, du + \sigma_{21}\, dW_u^2 + \rho\, \sigma_{22}\, dW_u^m \right),$$

with $Z_t = \hat{\theta}$, $u \in [t, T]$, where W^1, W^2, and W^m are independent Brownian motions under the probability measure $\mathbb{P}^*$. From the PDE in (2.29) and the terminal condition $H(T, \hat{\theta}) = 0$, the function H has the stochastic representation in (2.22) because the process

$$H(t, Z_t) + \int_t^T F(u, Z_u)\, du$$

is a martingale, by Itô's lemma. Next, we show that the random variable Z_u for $t \le u \le T$ is normally distributed. We apply Itô's lemma to $\exp(\int_t^u \beta_s\, ds)\, Z_u$ and write

$$\begin{aligned}
\exp\left(\int_t^u \beta_s\, ds\right) Z_u^i &= Z_t^i + \int_t^u \exp\left(\int_t^v \beta_s\, ds\right) dZ_v^i + \int_t^u \beta_v \exp\left(\int_t^v \beta_s\, ds\right) Z_v^i\, dv \\
&= Z_t^i + \int_t^u r\, \beta_v \exp\left(\int_t^v \beta_s\, ds\right) dv + \int_t^u \sigma_{i\,1}\, \beta_v \exp\left(\int_t^v \beta_s\, ds\right) dW_v^i \\
&\quad + \int_t^u I_i\, \sigma_{i\,2}\, \beta_v \exp\left(\int_t^v \beta_s\, ds\right) dW_v^m,
\end{aligned}$$

where, to simplify notation, $I_1 = 1$ and $I_2 = \rho$. Therefore,

$$\begin{aligned}
Z_u^i &= Z_t^i \exp\left(-\int_t^u \beta_s\, ds\right) + \int_t^u r\, \beta_v \exp\left(-\int_v^u \beta_s\, ds\right) dv \\
&\quad + \int_t^u \sigma_{i\,1}\, \beta_v \exp\left(-\int_v^u \beta_s\, ds\right) dW_v^i + \int_t^u I_i\, \sigma_{i\,2}\, \beta_v \exp\left(-\int_v^u \beta_s\, ds\right) dW_v^m \\
&= \hat{\theta}_i \left(\frac{1+t}{1+u}\right)^L \\
&\quad + L \left(\int_t^u r\, \frac{(1+v)^{L-1}}{(1+u)^L}\, dv + \int_t^u \sigma_{i\,1}\, \frac{(1+v)^{L-1}}{(1+u)^L}\, dW_v^i + \int_t^u I_i\, \sigma_{i\,2}\, \frac{(1+v)^{L-1}}{(1+u)^L}\, dW_v^m \right).
\end{aligned}$$

Thus, it is easy to see that the mean of Z_u^i is given by (2.24), and the covariance matrix of Z_u is given in (2.25) when $L \neq 1/2$ and in (2.26) when $L = 1/2$. □

In the next proposition, we let $c = 0$ and derive the closed-form optimal adaptive strategy α^{ad} for the investment problem in (2.14).

Proposition 2.4 ***Adaptive strategy.*** *Let the prices of the asset satisfy the SDEs in* (2.1)*,* $L > 0$ *the learning parameter,* $T > 0$ *the investment horizon, the uncertainty parameter* $c = 0$*, and* $\rho \in [-1, 1]$. *The adaptive investment strategy for the investor's problem* (2.14) *is*

$$\alpha_t^{ad} = \frac{\Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\, r\right)}{\gamma\ \exp(r\,(T-t))} - \frac{\beta_t\ \partial_{\hat{\theta}} \bar{H}^0(t, \hat{\theta})}{\gamma\ \exp(r\,(T-t))}, \tag{2.30}$$

where $\bar{H}^0$ *has the stochastic representation*

$$\bar{H}^0(t, \hat{\theta}) = \mathbb{E}^{\mathbb{P}^*}\left[\int_t^T \frac{1}{2}\, (Z_u - \mathbf{1}\, r)^\intercal\ \Sigma^{-1}\ (Z_u - \mathbf{1}\, r)\ du\right], \tag{2.31}$$

with Z *defined in* (2.23)*, and*

$$\partial_{\hat{\theta}} \bar{H}^0(t, \hat{\theta}) = \Sigma^{-1}\, (\hat{\theta} - \mathbf{1}\, r) \int_t^T \left(\frac{1+t}{1+u}\right)^{2L} du. \tag{2.32}$$

Proof Let $c = 0$ in Proposition 2.3 to obtain (2.30) and (2.31). Next, we show (2.32). Denote the variance of the random variable X and the covariance between the random variables X and Y by $\mathrm{Var}(X)$ and $\mathrm{Cov}(X, Y)$, respectively. From (2.31), we have that

$$\bar{H}^0(t,\hat{\theta}) = \frac{1}{2\,(\sigma_1^2\,\sigma_2^2 - \rho^2\,\sigma_c^4)} \times \int_t^T \left(\sigma_2^2\,\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^1 - r)^2\right] + \sigma_1^2\,\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^2 - r)^2\right] - 2\,\rho\,\sigma_c^2\,\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^1 - r)\,(Z_u^2 - r)\right]\right) du\,.$$

The expectations that appear above are given by

$$\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^1 - r)^2\right] = \mathbb{E}^{\mathbb{P}^*}\left[(Z_u^1 - r)\right]^2 + \mathrm{Var}(Z_u^1) = (\hat{\theta}_1 - r)^2 \left(\frac{1+t}{1+u}\right)^{2L} + \mathrm{Var}(Z_u^1)\,, \tag{2.33}$$

$$\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^2 - r)^2\right] = (\hat{\theta}_2 - r)^2 \left(\frac{1+t}{1+u}\right)^{2L} + \mathrm{Var}(Z_u^2)\,, \tag{2.34}$$

and

$$\mathbb{E}^{\mathbb{P}^*}\left[(Z_u^1 - r)\,(Z_u^2 - r)\right] = (\hat{\theta}_1 - r)\,(\hat{\theta}_2 - r) \left(\frac{1+t}{1+u}\right)^{2L} + \mathrm{Cov}(Z_u^1, Z_u^2)\,. \tag{2.35}$$

The terms $\mathrm{Var}(Z_u^1)$, $\mathrm{Var}(Z_u^2)$, and $\mathrm{Cov}(Z_u^1, Z_u^2)$ do not depend on $\hat{\theta}$, thus, we have that

$$\partial_{\hat{\theta}_1} \bar{H}^0(t,\hat{\theta}) = \frac{\sigma_2^2\,(\hat{\theta}_1 - r) - \rho\,\sigma_c^2\,(\hat{\theta}_2 - r)}{\sigma_1^2\,\sigma_2^2 - \rho^2\,\sigma_c^4} \int_t^T \left(\frac{1+t}{1+u}\right)^{2L} du \tag{2.36}$$

and

$$\partial_{\hat{\theta}_2} \bar{H}^0(t,\hat{\theta}) = \frac{-\rho\,\sigma_c^2\,(\hat{\theta}_1 - r) + \sigma_1^2\,(\hat{\theta}_2 - r)}{\sigma_1^2\,\sigma_2^2 - \rho^2\,\sigma_c^4} \int_t^T \left(\frac{1+t}{1+u}\right)^{2L} du\,. \tag{2.37}$$

Finally, the expression in (2.32) follows from the two equations above. □

From Proposition 2.3, the optimal strategy for the adaptive-robust model in (2.21) is the sum of two components. The first term on the right-hand side of (2.21) is a Merton-type investment strategy with a truncation that depends on the sign of $\hat{\theta} - r$. The second term on the right-hand side of (2.21) is an adjustment that results from the integral of the estimate of the risk-premium with a truncation ($\bar{H}$). Similarly, from Proposition 2.4, the optimal strategy of the adaptive model in (2.30) is the sum of two components. The first term on the right-hand side of (2.30) is a Merton-type investment strategy. The second term on the right-hand side of (2.30) is an adjustment that results from the integral of the estimate of the risk-premium ($\bar{H}^0$) over the remaining trading horizon.

The adaptive strategy is compensated by the partial derivative (with respect to $\hat{\theta}$) of the integral of the variance of the estimated risk-premium with a truncation for the adaptive-robust strategy and without truncation for the adaptive strategy. We refer to the term $\beta_t\,\partial_{\hat{\theta}}\bar{H}^0$ as "compensated parameter estimation term" because the term only appears when the agent simultaneously estimates the value of the unknown parameter and solves the optimal control problem. When the agent accounts for parameter uncertainty, the term $\beta_t\,\partial_{\hat{\theta}}\bar{H} - \beta_t\,\partial_{\hat{\theta}}\bar{H}^0$ depends on c to account for parameter uncertainty — we refer to this term as the "compensated parameter uncertainty term". We return to these points below when we study the optimal investment strategies when the value of the correlation parameter ρ is zero.

In the literature, the second term on the right-hand side of (2.30) is generally referred to as "intertemporal hedging term". [11] study a HARA utility maximisation problem where the risk premium of the stock follows an Ornstein–Uhlenbeck (OU) process with known coefficients. The optimal investment strategy consists of two components: the usual optimal strategy for the Merton problem and a term to hedge risk-premium uncertainty. Moreover, they show that the optimal strategy can short the stock when the risk-premium of the stock is positive. Similar to their result, we show that when the value of the learning parameter L is less than one, there are instances where the agent shorts the risky asset when the risk premium is positive. In an extension to Kim and Omberg, [18] assumes that the agent learns the coefficients of the OU process, which leads to and additional term in the optimal investment strategy. Similarly, [13] show that the intertemporal hedging term of the predictable return accounts for the sensitivity of the risk-premium to the stock process.

3 The adaptive strategy

In this section, we discuss the behaviour of the adaptive strategy when the correlation between the stock prices is zero (i.e., $\rho = 0$). We write the adaptive strategy in (2.30) as

$$\alpha_t^{ad,i} = \frac{\hat{\theta}_i - r}{\gamma\, \sigma_i^2 \exp(r\,(T-t))} \times A(t, T, L)\,, \tag{3.1}$$

where

$$A(t, T, L) = \begin{cases} 1 - \frac{L}{2L-1}\left(1 - \left(\frac{1+t}{1+T}\right)^{2L-1}\right) & \text{if} \quad L \neq 1/2\,, \\ 1 - \frac{1}{2}\,(\log(1+T) - \log(1+t)) & \text{if} \quad L = 1/2\,. \end{cases} \tag{3.2}$$

Recall that the first term on the right-hand side of the adaptive strategy in (3.1) is the Merton-type investment strategy. The term $A(t, T, L)$ is an adjustment that results from the updates in the estimate of the drift parameters. To understand the intuition of the adaptive strategy we study the adjustment term $A(t, T, L)$ as a function of $T - t$ and as a function of the value of the learning parameter L. The proposition below shows the upper and lower bounds of the adjustment term for $L > 0$ and $t \leq T$.

Proposition 3.1 *Let $L > 0$ be the learning parameter, $T > 0$ the horizon of the investment, and let the correlation parameter $\rho = 0$. Then, the adjustment term $A(t, T, L)$ is increasing in t, and obeys the bounds*

$$A(0, T, L) \leq A(t, T, L) \leq A(T, T, L) = 1\,, \tag{3.3}$$

where

$$A(0, T, L) = \begin{cases} \frac{L-1}{2L-1} + \frac{L}{2L-1}\left(\frac{1}{1+T}\right)^{2L-1} & \textit{if} \quad L \neq 1/2\,, \\ 1 - \frac{1}{2}\log(1+T) & \textit{if} \quad L = 1/2\,. \end{cases} \tag{3.4}$$

For a proof see Appendix D.

The sign of the lower bound (3.4) could be positive or negative. The next proposition shows that when $L \geq 1$, $A(t, T, L)$ is always positive and when $L < 1$, there exists a time $t^L < T$ such that $A(t^L, T, L) = 0$.

Proposition 3.2 *Assume that the horizon of the investment is $T > 0$ and that the correlation parameter is $\rho = 0$. (1) If $L \geq 1$, then the adjustment term $A(t, T, L)$ is always positive. (2)*

If $L < 1$, then there exists

$$t^L = \begin{cases} (1+T)\left(\frac{1-L}{L}\right)^{1/(2L-1)} - 1 & \text{if} \quad L \neq 1/2\,, \\ (1+T)^{1/2} - 1 & \text{if} \quad L = 1/2\,, \end{cases} \tag{3.5}$$

such that $A(t^L, T, L) = 0$. Thus, if $t > t^L$, then $A(t, T, L) > 0$, and if $t < t^L$, then $A(t, T, L) < 0$.

For a proof see Appendix E.

When the learning parameter $L \geq 1$, the sign of the holdings in the stock for the adaptive and the Merton-type strategies is the same. For example, if $L \geq 1$ and $\hat{\theta}_i < r$ both the Merton-type strategy and the adaptive strategy will hold a short position in stock i. On the other hand, when $L < 1$, the sign of the holdings in the stock held by the adaptive strategy is not necessarily the same as the sign of the holdings of the Merton-type strategy.

Next, we study $A(t, T, L)$ as a function of L. For $L \neq 1/2$, we have that

$$\frac{\partial A(t,T,L)}{\partial L} = \frac{1}{(2L-1)^2}\left(1 - \left(\frac{1+t}{1+T}\right)^{2L-1}\right) + \frac{L}{(2L-1)}(2L-1)\left(\frac{(1+t)^{2L-2}}{(1+T)^{2L-1}}\right), \tag{3.6}$$

and the second term on the right-hand side of (3.6) is always greater than zero for all values of L. When $L > 1/2$, the first term on the right-hand side of (3.6) is greater than zero, and if $L < 1/2$, the first term on the right-hand side of (3.6) can be less than zero.

The sensitivity to the integral of the variance of the estimated risk-premium in (2.31) has an important effect on the adaptive strategy (2.31). When $L > 1/2$, as the value of L increases, the sensitivity of the integral of the variance of the estimated risk-premium decreases because the investor acknowledges that learning the risk-premium happens quickly because the value of L is large, and therefore, the integral of the variance of the estimated risk-premium in (2.31) and its sensitivity in (2.32) are smaller. For example, assume that $\hat{\theta} > r$, then, there are two forces at play in the adaptive strategy. First, the investor goes long the stock because the expected returns exceed the risk free rate, and second, the investor makes adjustments to the amount to be invested to account for the arrival of new information and the updates of the drift parameter. Recall that the investor's utility function is increasing and concave in wealth. Thus, all else being equal, if the sensitivity of the integral of the variance of the estimated risk-premium in (2.32) increases, the quantity held in the stock decreases. Similarly, an increase in the variance of the estimator of the drift process reduces the holdings in the stock.

The choice of the value of the learning parameter L is crucial. From Proposition 3.2, when $L \geq 1$, the adjustment term is always greater than zero regardless of the values of t and T. Therefore, when $L \geq 1$, the sign of the adaptive strategy is the same as that of the Merton-type strategy because the adjustment term is always greater than zero. In contrast, when $L < 1$, the adjustment term can be greater or less than zero depending on the value of t. Therefore, when $L < 1$ the sign of the adaptive strategy is not necessarily the same as that of the Merton investment strategy in (2.17).

4 The adaptive-robust strategy

In this section, we study the optimal adaptive-robust investment strategy in (2.3) when the correlation between the innovations in the prices of the two risky assets in (2.1) is (i) $\rho = 0$,

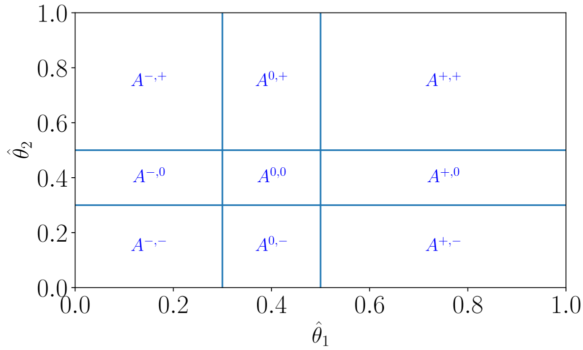

Fig. 1 Investment regions when $\rho = 0$. Model parameters are $r = 0.4$, $c = 0.2$, and $t = 3$. The x-axis represents the drift $\hat{\theta}_1$ and the y-axis represents the drift $\hat{\theta}_2$

(ii) $\rho > 0$, (iii) $\rho < 0$, and we also study the effect of the uncertainty coefficient c on the optimal strategy.

4.1 Uncorrelated asset prices

In this subsection, we assume that the correlation between S^1 and S^2 is zero, i.e., $\rho = 0$. Figure 1 plots the nine investment regions in $\mathbb{R}^2$, where the centre region is $A^{0,0}$, the upper right region is $A^{+,+}$, and the remaining regions are $A^{+,0}$, $A^{+,-}$, $A^{0,-}$, $A^{-,-}$, $A^{-,0}$, $A^{-,+}$, $A^{0,+}$ in clockwise order, and other model parameters are $r = 0.4$, $c = 0.2$, and $t = 3$.

The investment decisions are centred around $r = 0.4$. The width of the regions $A^{0,0}$, $A^{\pm,0}$, and $A^{0,\pm}$ shrinks as time increases: $A^{+,+} = [r + c/\sqrt{1+t}, \infty) \times [r + c/\sqrt{1+t}, \infty)$ and $A^{-,-} = (-\infty,\ r - c/\sqrt{1+t}] \times (-\infty,\ r - c/\sqrt{1+t}]$. When $\hat{\theta} \in A^{+,+}$ the agent acquires a long position in both stocks because the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ of the drift parameters are greater than $r + c/\sqrt{1+t}$, where c is the uncertainty coefficient. Similarly, when $\hat{\theta} \in A^{-,-}$ the agent acquires a short position in both stocks because both estimates of the drift are less than $r - c/\sqrt{1+t}$.

When $\hat{\theta} \in A^{+,0}$, the agent acquires a long position in the first stock, because $\hat{\theta}_1$ is greater than $r + (c/\sqrt{1+t})$. For the other asset, one cannot determine whether the position is long or short because $\hat{\theta}_2$ is in $[r - (c/\sqrt{1+t}),\ r + (c/\sqrt{1+t})]$, so the value of $\hat{\theta}_2$ neither exceeds nor is less than the risk-free rate r enough to account for the effect of the uncertainty coefficient c. Similarly, when $\hat{\theta} \in A^{-,0}$, the agent acquires a short position in the first stock because $\hat{\theta}_1$ is less than $r - (c/\sqrt{1+t})$, and the position in the other asset could be long or short because $\hat{\theta}_2$ is in between $[r - (c/\sqrt{1+t}),\ r + (c/\sqrt{1+t})]$.

Finally, when $\hat{\theta} \in A^{-,+}$, the agent takes a short position in the first stock and takes a long position in the second stock because the estimate $\hat{\theta}_1$ is less than $r - c/\sqrt{1+t}$ and the estimate $\hat{\theta}_2$ is greater than $r + c/\sqrt{1+t}$. The investment region $A^{+,-}$ has a similar interpretation.

4.2 Positive correlation between asset prices

In this subsection, the correlation between the prices of the risky assets is positive i.e., $\rho > 0$. Figure 2 plots the nine regions in $\mathbb{R}^2$, where the centre region is $A^{0,0}$, the upper right region is $A^{+,+}$, and the remaining regions are $A^{+,0}$, $A^{+,-}$, $A^{0,-}$, $A^{-,-}$, $A^{-,0}$, $A^{-,+}$, $A^{0,+}$ in clockwise order.

When $\rho > 0$, the regions $A^{+,-}$ and $A^{-,+}$ take a larger share of the $(\hat{\theta}_1, \hat{\theta}_2)$ space. This is expected because with a positive correlation, the agent favours opposite positions in the

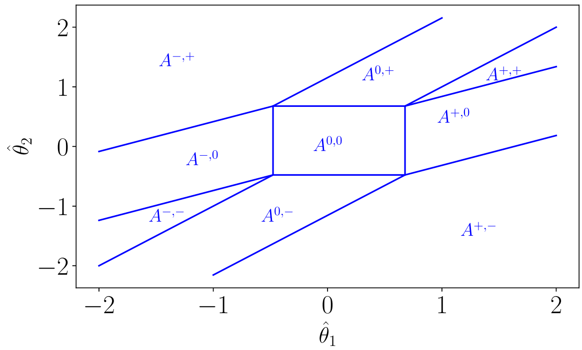

Fig. 2 Investment regions when $\rho > 0$. Model parameters are $r = 0.4$, $c = 0.2$, and $t = 3$. The x-axis represents the drift $\hat{\theta}_1$ and the y-axis represents the drift $\hat{\theta}_2$

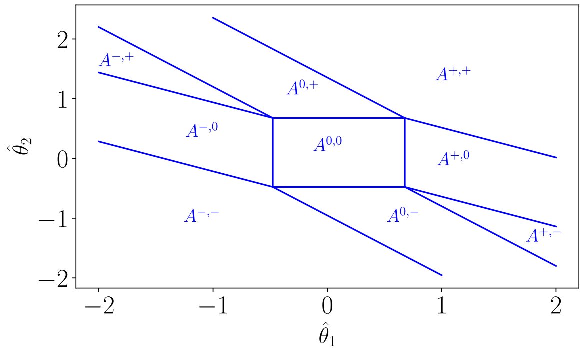

Fig. 3 Investment regions when $\rho < 0$. Model parameters are $r = 0.4$, $c = 0.2$, and $t = 3$. The x-axis represents the drift $\hat{\theta}_1$ and the y-axis represents the drift $\hat{\theta}_2$

assets more due to risk aversion. Figure 2 also shows that $A^{+,+} \subset [r + c/\sqrt{1+t}, \infty) \times [r + c/\sqrt{1+t}, \infty)$, so there exist $\hat{\theta} \in [r + c/\sqrt{1+t}, \infty) \times [r + c/\sqrt{1+t}, \infty)$ which is not in $A^{+,+}$. This is due to the positive correlation between the two risky assets, which makes the diversification of a long-long position less efficient under the risk-adjusted returns. Therefore, it is possible for the agent to acquire a long position in only one asset even though both of $\hat{\theta}_1$ and $\hat{\theta}_2$ are greater than $r + (c/\sqrt{1+t})$. Similarly, for the region $A^{-,-} \subset (-\infty,\ r - c/\sqrt{1+t}] \times (-\infty,\ r - c/\sqrt{1+t}]$, it is possible for the agent to acquire a short position in only one asset, even if both $\hat{\theta}_1$ and $\hat{\theta}_2$ are less than $r - c/\sqrt{1+t}$.

Similarly, Figure 3 shows the regions when the the correlation between the prices of the risky assets is negative, i.e., $\rho < 0$. The centre region is $A^{0,0}$, the upper right most region is $A^{+,+}$, and the remaining regions are $A^{+,0}$, $A^{+,-}$, $A^{0,-}$, $A^{-,-}$, $A^{-,0}$, $A^{-,+}$, $A^{0,+}$ in clockwise order. Here, the regions $A^{+,+}$ and $A^{-,-}$ take a larger share of the $(\hat{\theta}_1, \hat{\theta}_2)$ space. Again, this is expected because the negative correlation makes similar positions (e.g., long-long or short-short) favour diversification of risk. We omit the interpretations of the strategies in each region because they are similar to the those when $\rho > 0$.

5 Asymptotic analysis: $c \to 0$ and $\rho = 0$

The optimal strategy in (2.21) is not available in closed-form because we cannot find an explicit solution for the stochastic representation of $\bar{H}$ in (2.22). Thus, we consider the optimal strategy in Proposition 2.3 when the uncertainty parameter c is close to zero. We also explore the effect of the learning parameter L on the optimal strategy for both the adaptive strategy and the adaptive-robust strategy. We assume that the correlation between the prices of the stocks is zero (i.e, $\rho = 0$) — one can follow a similar approach when stock prices are correlated.

Next, we state a corollary from Proposition 2.3 that provides the optimal strategy of the adaptive-robust control problem in (2.3) when $\rho = 0$ and $c > 0$ — we call this optimal strategy "the adaptive-robust strategy".

Corollary 5.1 *Let the uncertainty parameter $c > 0$ and $\rho = 0$. The optimal adaptive-robust investment strategy is*

$$\alpha_t^{adr,i} = \begin{cases} \frac{(\hat{\theta}_i - r + c/\sqrt{1+t}) - \beta_t\,\sigma_i^2\,\partial_{\hat{\theta}_i}\bar{H}(t,\hat{\theta})}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}\,, & \hat{\theta}_i - r \le -c/\sqrt{1+t}\,, \\ \frac{(\hat{\theta}_i - r - c/\sqrt{1+t}) - \beta_t\,\sigma_i^2\,\partial_{\hat{\theta}_i}\bar{H}(t,\hat{\theta})}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}\,, & \hat{\theta}_i - r \ge c/\sqrt{1+t}\,, \\ \frac{-\beta_t\;\partial_{\hat{\theta}_i}\bar{H}(t,\hat{\theta})}{\gamma\;\exp(r\,(T-t))}\,, & \mid \hat{\theta}_i - r \mid < c/\sqrt{1+t}\,, \end{cases} \tag{5.1}$$

where $i = 1, 2$ and the function $\bar{H}$ in (2.22) is given by

$$\begin{aligned}\bar{H}(t,\hat{\theta}) = \sum_{i=1}^{2} \frac{1}{2\,\sigma_i^2}\,\mathbb{E}^{\mathbb{P}}\left[\int_t^T \left(Z_u^i - r - c/\sqrt{1+u}\right)^2 \mathbf{1}_{\{Z_u^i - r >\; c/\sqrt{1+u}\}}\,du \;\mid\; Z_t = \hat{\theta}\right] \\ + \frac{1}{2\,\sigma_i^2}\,\mathbb{E}^{\mathbb{P}}\left[\int_t^T \left(Z_u^i - r + c/\sqrt{1+u}\right)^2 \mathbf{1}_{\{Z_u^i - r <\; -c/\sqrt{1+u}\}}\,du \;\mid\; Z_t = \hat{\theta}\right],\end{aligned} \tag{5.2}$$

which is an integral of the estimated variance with a truncation when $c > 0$ and it is an integral of the estimated variance when $c = 0$; also, it is easy to see that $\bar{H}(t,\hat{\theta}; c = 0) > \bar{H}(t,\hat{\theta}; c < 0)$.

Before we proceed to the main results of this section, we provide some preliminary identities to compute (5.2). We denote by $\phi(x)$ and $\Phi(x)$ the density function and cumulative distribution of a standard normal variable, respectively.

Lemma 5.2 *Let the random variable Y be normally distributed with mean μ and variance σ^2, i.e., $Y \sim \mathcal{N}(\mu, \sigma^2)$, then*

$$\mathbb{E}^{\mathbb{P}}\left[Y^2\,\mathbf{1}_{\{Y>0\}}\right] = \left(\mu^2 + \sigma^2\right)\left(1 - \Phi(-\mu/\sigma)\right) + \frac{\mu\,\sigma}{\sqrt{2\,\pi}}\,\exp\left(-\mu^2/2\,\sigma^2\right)\,. \tag{5.3}$$

For a proof see Appendix F.

From (5.3), we also have that

$$\mathbb{E}^{\mathbb{P}}\left[Y^2\,\mathbf{1}_{\{Y<0\}}\right] = \left(\mu^2 + \sigma^2\right)\Phi(-\mu/\sigma) - \frac{\mu\,\sigma}{\sqrt{2\,\pi}}\,\exp\left(-\mu^2/2\,\sigma^2\right)\,. \tag{5.4}$$

Now, define the function

$$h(u,\hat{\theta}_i) = \sum_{i=1}^{2}\mathbb{E}^{\mathbb{P}}\left[\left(Z_u^i - r - \frac{c}{\sqrt{1+u}}\right)^2 \mathbf{1}_{\{Z_u^i - r > \frac{c}{\sqrt{1+u}}\}} + \left(Z_u^i - r + \frac{c}{\sqrt{1+u}}\right)^2 \mathbf{1}_{\{Z_u^i - r < \frac{-c}{\sqrt{1+u}}\}}\right], \tag{5.5}$$

so from (5.2), we have that

$$\bar{H}(t,\hat{\theta}) = \sum_{i=1}^{2} \frac{1}{2\,\sigma_i^2} \int_t^T h(u,\hat{\theta}_i)\,du\,. \tag{5.6}$$

The next lemma characterises the gradient $\partial_{\hat{\theta}_i} h$.

Lemma 5.3 *Let the function h be as in (5.5). Then, h is $C^{1,2}([0,T]\times\mathbb{R})$ and the gradient $\partial_{\hat{\theta}_i} h$ is given by*

$$\partial_{\hat{\theta}_i} h = \Big[2\, f\,(1-\Phi(-f/\boldsymbol{\sigma}_i)) + 2\,\boldsymbol{\sigma}_i\,\phi(-f/\boldsymbol{\sigma}_i) + 2\,\tilde{f}\,\Phi\left(-\tilde{f}/\boldsymbol{\sigma}_i\right) - 2\,\Sigma_i\,\phi\left(-\tilde{f}/\boldsymbol{\sigma}_i\right)\Big]\,\partial_{\hat{\theta}_i} f\,. \quad (5.7)$$

Here

$$f(u,\hat{\theta}_i) = (\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - \frac{c}{\sqrt{1+u}}\,, \qquad \tilde{f}(u,\hat{\theta}_i) = (\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + \frac{c}{\sqrt{1+u}}\,,$$

where

$$\boldsymbol{\sigma}_i(u) = \sqrt{\frac{L^2\,\sigma_i^2}{2\,L-1}\left(\frac{(1+u)^{2\,L-1}}{(1+u)^{2\,L}} - \frac{(1+t)^{2\,L-1}}{(1+u)^{2\,L}}\right)} \quad \text{if} \quad L \neq 1/2$$

and

$$\boldsymbol{\sigma}_i(u) = \frac{\sigma_i^2}{4}\left(\frac{\log(1+u)-\log(1+t)}{1+u}\right) \quad \text{if} \quad L = 1/2\,.$$

For a proof see Appendix G.

Note that the function $\boldsymbol{\sigma}_i$ depends on the value of t. The next lemma shows the asymptotic formula of the function h when the value of the uncertainty parameter c is small.

Lemma 5.4 *Let the function h be as in (5.5). Then, there exists a constant $\delta_{t,\hat{\theta}_i}$ and an integrable function $R(u,\hat{\theta}_i)$ with respect to u such that*

$$\left|\partial_{\hat{\theta}_i} h(u,\hat{\theta}_i) - 2\,(\hat{\theta}_i - r) - \tfrac{2\,c}{\sqrt{1+u}}\Big[\Phi\left(-(\hat{\theta}_i - r)/\boldsymbol{\sigma}_i(u)\right) - \Phi\left((\hat{\theta}_i - r)/\boldsymbol{\sigma}_i(u)\right)\Big]\right| \leq c^2\,R(u,\hat{\theta}_i)\,, \quad (5.8)$$

for all $u \in [t,T]$ and for all $c < \delta_{t,\hat{\theta}_i}$, where $\delta_{t,\hat{\theta}_i}$ depends on t and $\hat{\theta}_i$.

For a proof see Appendix H.

In the lemma above, it is necessary to show that the function R is integrable with respect to u because $\partial_{\hat{\theta}_i}\bar{H}$ is the integral of $\partial_{\hat{\theta}_i} h$; see (5.6). Next, we show the asymptotic formula for the partial derivative of $\bar{H}$ with respect to $\hat{\theta}_i$, which is a direct result from Lemma 5.4.

Lemma 5.5 *Let the function $\bar{H}$ be as in (2.22). Then, $\bar{H}$ is $C^{1,2}([0,T]\times\mathbb{R}^n)$ and the gradient $\partial_{\hat{\theta}_i}\bar{H}$ is given by*

$$\partial_{\hat{\theta}_i}\bar{H} = \frac{1}{\sigma_i^2}\int_t^T (\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^{2\,L} du - \frac{c}{\sigma_i^2}\left(\int_t^T \boldsymbol{H}^L(t,u,\hat{\theta}_i)\,du\right) + \mathcal{O}(c^2)\,, \quad (5.9)$$

where

$$\boldsymbol{H}^L(t,u,\hat{\theta}_i) = \frac{1}{\sqrt{1+u}}\left(\frac{1+t}{1+u}\right)^{2\,L}\Bigg[\Phi\left(-(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L/\,\boldsymbol{\sigma}_i(u)\right) - \Phi\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L/\,\boldsymbol{\sigma}_i(u)\right)\Bigg]. \quad (5.10)$$

Proof Recall the definition of the functions $\bar{H}$ and h and write

$$\partial_{\hat{\theta}_i} \bar{H}(t, \hat{\theta}_i) = \frac{1}{2\,\sigma_i^2} \int_t^T \partial_{\hat{\theta}_i} h(u, \hat{\theta}_i)\, du\,.$$

Therefore, from Lemma 5.4, equation (5.9) holds because the function $\tilde{R}$ is integrable. □

Here, the term

$$\Phi\left(-(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L / \boldsymbol{\sigma}_i(u)\right) - \Phi\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L / \boldsymbol{\sigma}_i(u)\right) \tag{5.11}$$

is positive when $\hat{\theta}_i - r < 0$ and negative when $\hat{\theta}_i - r > 0$. Therefore, $\boldsymbol{H}^L(t, u, \hat{\theta}_i)$ is positive when $\hat{\theta}_i - r < 0$ and negative when $\hat{\theta}_i - r > 0$.

Next, we study the properties of the function $\boldsymbol{H}^L$ in (5.10) to develop an intuition for the adaptive-robust strategy. First, we show the convergence of the function $L \int_t^T \boldsymbol{H}^L(t, u, \hat{\theta}_i)\, du$ with respect to the parameter L.

Proposition 5.6 *For each $\hat{\theta}_i$ and t, the function $L \int_t^T \boldsymbol{H}^L(t, u, \hat{\theta}_i)\, du$ converges to 0 as $L \to \infty$.*

Proof It suffices to consider the case $\hat{\theta}_i > 0$. When $\hat{\theta}_i > 0$, the function $\boldsymbol{H}^L(t, u, \hat{\theta})$ is negative, so $L \int_t^T \boldsymbol{H}^L(t, u, \hat{\theta}_i)\, du$ is negative. Next,

$$\begin{aligned}\lim_{L\to\infty} L\, \boldsymbol{H}^L(t, u, \hat{\theta}_i) &= \lim_{L\to\infty} \frac{L}{\sqrt{1+u}} \left(\frac{1+t}{1+u}\right)^{2L} \\ &\quad \times \lim_{L\to\infty} \left[1 - 2\,\Phi\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L / \boldsymbol{\sigma}_i(u)\right)\right] = 0\end{aligned}$$

because both limits converge to 0. Next, note that

$$\left|\frac{L}{\sqrt{1+u}}\left(\frac{1+t}{1+u}\right)^{2L} \times \left[1 - 2\,\Phi\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L / \boldsymbol{\sigma}_i(u)\right)\right]\right| \le \frac{3L}{\sqrt{1+u}}\left(\frac{1+t}{1+u}\right)^{2L}$$

and that the function

$$u \to \frac{3L}{\sqrt{1+u}}\left(\frac{1+t}{1+u}\right)^{2L}$$

is integrable from t to T. Thus,

$$L \int_t^T \boldsymbol{H}^L(t, u, \hat{\theta}_i)\, du \to 0 \qquad \text{as} \qquad L \to \infty\,,$$

by the dominated convergence theorem. □

Now, given the asymptotic formula for $\partial_{\hat{\theta}_i} \bar{H}$ when the value of c is close to zero in (5.9), we study three cases of the optimal adaptive-robust investment strategy in (2.3) for a sufficient large L.

(1) $\hat{\theta}_i > r + c/\sqrt{1+t}$:

$$\alpha_t^{adr,i} = \alpha_t^{ad,i} + \frac{-\frac{c}{\sqrt{1+t}} + L\,\frac{c}{1+t} \int_t^T \boldsymbol{H}^L(t, u, \hat{\theta}_i)\, du + \mathcal{O}(c^2)}{\gamma\, \sigma_i^2\, \exp(r\,(T-t))}\,. \tag{5.12}$$

The second term on the right-hand side is negative because $\boldsymbol{H}^L(t, u, \hat{\theta}_i)$ is negative when $\hat{\theta}_i > r + c/\sqrt{1+t}$. The long position of the adaptive-robust strategy is lower than that of the adaptive strategy when $\hat{\theta}_i > r + c/\sqrt{1+t}$. Intuitively, the agent adjusts the adaptive-robust strategy because the estimate of the drift is positive.

(2) $\hat{\theta}_i < r - c/\sqrt{1+t}$:

$$\alpha_t^{adr,i} = \alpha_t^{ad,i} + \frac{\frac{c}{\sqrt{1+t}} + L\,\frac{c}{1+t}\int_t^T \boldsymbol{H}^L(t, u, \hat{\theta}_i)\,du + \mathcal{O}(c^2)}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}\,. \tag{5.13}$$

The second term on the right-hand side is positive because $\boldsymbol{H}^L(t, u, \hat{\theta}_i)$ is positive; thus, the amount of the risk stock held short in adaptive-robust strategy is smaller than that of the adaptive strategy when $\hat{\theta}_i < r - c/\sqrt{1+t}$.

(3) $|\,\hat{\theta}_i - r\,| < c/\sqrt{1+t}$:

$$\alpha_t^{adr,i} = \alpha_t^{ad,i} + \frac{L\,\frac{c}{1+t}\int_t^T \boldsymbol{H}^L(t, u, \hat{\theta}_i)\,du + \mathcal{O}(c^2)}{\gamma\,\sigma_i^2\,\exp(r\,(T-t))}\,. \tag{5.14}$$

The second term on the right-hand side is negative when $0 < \hat{\theta}_i - r < c/\sqrt{1+t}$. Hence, the agent of the adaptive-robust strategy builds a smaller long position than that of the adaptive strategy. Similarly, when $0 > \hat{\theta}_i - r > -c/\sqrt{1+t}$, the difference $\Delta\alpha_t^i$ is positive. Therefore, the agent with the adaptive-robust strategy builds a smaller short position than that given by the adaptive strategy. In both cases, the adaptive-robust strategy is more conservative than the adaptive strategy when the estimate of the drift is close to the interest rate r.

6 Numerical results

The value function v can be characterised by the equation in (2.18), see [2]. In general, the value function v can be solved numerically using standard numerical PDE techniques. However, in the multidimensional cases (as it would be the case in practice), most numerical schemes become infeasible. In this paper, instead of solving the PDE in (2.18) numerically, we evaluate the integral (5.7) to obtain a numerical solution of the value function v.

In this section, we study two aspects of the solution we found. First, we study the behaviour of the optimal strategy of the adaptive-robust problem for various values of the uncertainty parameter c. Second, we compare the performance of the adaptive-robust strategies to three benchmarks where the agent (i) knows the true value of the drift parameter, (ii) misspecifies the value of the drift parameter, or (iii) employs a robust strategy. In the robust strategy, the agent uses the framework derived above, but does not learn the value of the unknown parameters.

6.1 The effect of the uncertainty parameter

Here, we explore the agent's adaptive-robust strategy for various choices of uncertainty parameter c. We assume that $T = 5$, $\sigma = 0.3$, $r = 0.01$, $L = 1$, and we calculate (numerically) the optimal strategy in (2.21) for one asset. The following figures show the optimal strategy when $t \in \{0, 1, 2, 4\}$. Figure 4 shows the cases $t = 0$ and $t = 1$, and Figure 5 shows the cases $t = 2$ and $t = 4$.

As expected, when $c = 0$ the optimal strategy is linearly dependent on the estimate of the drift parameter. When $c > 0$, the agent truncates the optimal strategy to almost zero (there is a

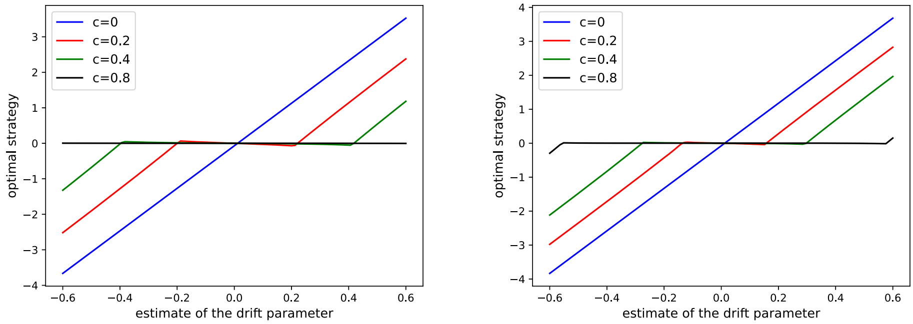

Fig. 4 Optimal adaptive-robust strategy when $t = 0$ (left) and $t = 1$ (right) as a function of the estimate of the drift parameter and for various values of the uncertainty parameter c

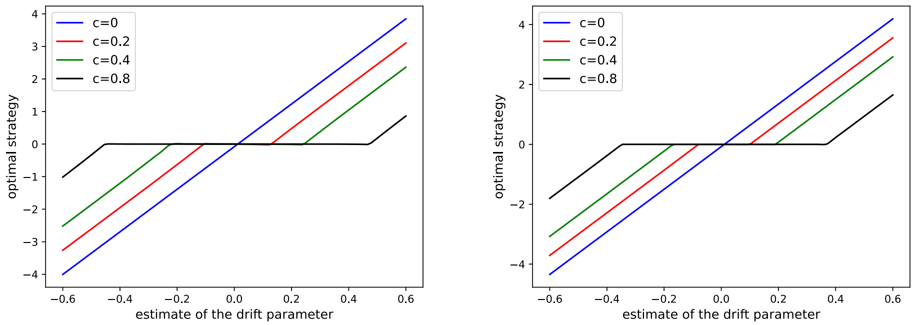

Fig. 5 Optimal adaptive-robust strategy when $t = 2$ (left) and $t = 4$ (right) as a function of the estimate of the drift parameter and for various values of the uncertainty parameter c

small term left) when the estimator is not too far from the risk-free rate r. Also, everything else being equal, as the value of c increases, the truncation area expands. Moreover, everything else being equal, as t increases, the truncation area shrinks; this is encoded in the definition of G. Furthermore, the choice of the value of c, which dictates the speed at which shrinking happens, can be linked to the confidence intervals of the estimators as discussed in Remark 2.2.

6.2 Performance of the adaptive-robust strategy

In this subsection, we compare the performance of the adaptive-robust strategy ($\star$) with that of three benchmarks where the agent: (i) knows the true value of the drift parameter, (ii) employs a wrong drift parameter, or (iii) employs a robust strategy. In the robust strategy, the agent uses the framework derived above but does not learn the value of the unknown parameter. Instead, the agent assumes that the true parameter $\theta^* \in [\underline{\theta}, \overline{\theta}]$ where $\underline{\theta}$ is the lowest possible value for the the true parameter and $\overline{\theta}$ is the highest possible value for the the true parameter. To obtain the robust strategy, the agent solves the problem

$$v(t, x, y) = \sup_{\alpha \in \mathcal{A}_0} \inf_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}} \left[U(X_T^{\alpha}) \right] , \tag{6.1}$$

where the set $\mathcal{P}$ contains all probability measure $\mathbb{P}_{\tilde{\theta}}$, such that $\tilde{\theta}_u \in [\underline{\theta}, \overline{\theta}]$ for all $u \in [t, T]$.

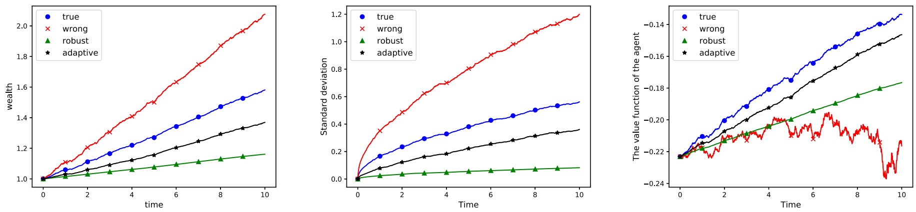

Fig. 6 Mean wealth process, standard deviation of the wealth process, and value function of agent

The set of possible values of the parameter θ^{*1} is in [0.03, 0.20], θ^{*2} is in [0, 0.20], the terminal time is $T = 10$ minutes, $L = 1$, and $c = 0.1$. The other model parameters are: $X_0 = 1$, $S_0^1 = 10$, $S_0^2 = 5$, $\gamma = 0.5$, $\theta^{*,1} = 0.09$, $\theta^{*,2} = 0.03$, $r = 0.01$, $\hat{\theta}_0^1 = 0.14$, $\hat{\theta}_0^2 = 0$, $\sigma_1 = 0.1$, and $\sigma_2 = 0.3$.

Recall that the perfect-knowledge strategy, i.e., the investor knows the true value of the drift parameter θ^*, is given by

$$\alpha_t^{*,1} = \frac{\theta^{*,1} - r}{\gamma\, \sigma_1^2\, \exp(r\,(T-t))} \quad \text{and} \quad \alpha_t^{*,2} = \frac{\theta^{*,2} - r}{\gamma\, \sigma_2^2\, \exp(r\,(T-t))}\,. \tag{6.2}$$

Specifically, we compare the ($\star$) adaptive-robust strategy in (2.21) to the following benchmarks: (i) the agent employs (6.2) with the true drift $\theta = \theta^*$, (ii) the agent employs the robust strategy without learning in (6.1), in which case the investment strategy is as in (6.2) with $\theta^1 = 0.03$ and $\theta^2 = 0$, and (iii) the agent employs (6.2) with the wrong drift parameters $\theta^1 = 0.15$ and $\theta^2 = 0.17$. In the figures below, we label strategy ($\star$) as "adaptive", strategy (i) as "true", strategy (ii) as "robust", and strategy (iii) as "wrong".

We discretise the time space into 8,000 time steps and employ 1,000 simulations to study the performance of the four strategies. The left panel of Figure ?? shows the wealth process of the agent. The strategy that overestimates the expected growth of the risky asset overinvests on the stock, and the mean value of the portfolio is higher than that of the other three strategies. On the other hand, the right panel of the figure shows that the standard deviation of the wealth for the wrong strategy is also the highest. Finally, the third panel shows the value function of the agent for the four strategies, where the highest value function is that of the perfect-knowledge strategy followed by that of the adaptive-robust strategy.

In this example, Figure 6 shows that using a wrong (predetermined) drift leads the agent to be over-aggressive, because the correct drift value is less than the wrong drift value. Thus, although it achieves high wealth, it does so at the expense of a high standard deviation and, as expected, the value function is lowest. Similarly, as expected, the adaptive strategy is somewhere between the oracle strategy (knowing the true drift parameters) and the robust strategy. The hyperparameters in the adaptive and robust strategies give the agent additional levers to adjust the performance in these plots.

Appendix A Proof of the convergence of the estimator in (2.8)

Let $I_1 = 1$ and $I_2 = \rho$, apply Itô's lemma to $\exp(\int_0^t \beta_s\, ds)\, \hat{\theta}_t^i$ to obtain

$$\exp\left(\int_0^t \beta_s\, ds\right)\, \hat{\theta}_t^i = \hat{\theta}_0^i + \int_0^t \exp\left(\int_0^s \beta_u\, du\right)\, d\hat{\theta}_s^i + \int_0^t \beta_s\, \exp\left(\int_0^s \beta_u\, du\right)\, \hat{\theta}_s^i\, ds$$

$$
\begin{aligned}
&= \hat{\theta}^i_t + \int_0^t \theta^{*,i}\, \beta_s\, \exp\left(\int_0^s \beta_u\, du\right) ds + \int_0^t \sigma_{i\,1}\, \beta_s\, \exp\left(\int_0^s \beta_u\, du\right) dW^i_s \\
&\quad + \int_0^t I_i\, \sigma_{i\,2}\, \beta_s\, \exp\left(\int_0^s \beta_u\, du\right) dW^m_s .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\hat{\theta}^i_t =& \hat{\theta}^i_0 \exp\left(-\int_0^t \beta_s\, ds\right) + \int_0^t \theta^{*,i}\, \beta_s\, \exp\left(-\int_s^t \beta_u\, du\right) ds + \int_0^t \sigma_{i\,1}\, \beta_s\, \exp\left(-\int_s^t \beta_u\, du\right) dW^i_s \\
&+ \int_0^t I_i\, \sigma_{i\,2}\, \beta_s\, \exp\left(-\int_s^t \beta_u\, du\right) dW^m_s \\
=& \hat{\theta}^i_0 \left(\frac{1}{1+t}\right)^L \\
&+ L\left(\int_0^t \theta^{*,i}\, \frac{(1+s)^{L-1}}{(1+t)^L}\, ds + \int_0^t \sigma_{i\,1}\, \frac{(1+s)^{L-1}}{(1+t)^L}\, dW^i_s + \int_0^t I_i\, \sigma_{i\,2}\, \frac{(1+s)^{L-1}}{(1+t)^L}\, dW^m_s\right) .
\end{aligned}
$$

On the right-hand side of the above equation the stochastic integrals are with respect to deterministic functions, therefore the distribution of the estimator $\hat{\theta}^i$ is normal. Then, for all values of the learning parameter L, the mean of $\hat{\theta}^i_t$ is

$$
\hat{\theta}^i_0 \left(\frac{1}{1+t}\right)^L + \theta^{*,i}\left(1 - \frac{1}{(1+t)^L}\right) ,
$$

and when $L \neq 1/2$, the variance of $\hat{\theta}^i_t$ is

$$
(\sigma^2_{i\,1} + I^2_i\, \sigma^2_{i\,2}) \int_0^t L^2\, \frac{1^{2\,L-2}}{(1+s)^{2\,L}}\, ds = (\sigma^2_{i\,1} + I^2_i\, \sigma^2_{i\,2})\, \frac{L^2}{2\,L-1}\left(\frac{(1+t)^{2\,L-1}}{(1+t)^{2\,L}} - \frac{(1)^{2\,L-1}}{(1+t)^{2\,L}}\right) ,
$$

and when $L = 1/2$, the variance of $\hat{\theta}^i_t$ is

$$
(\sigma^2_{i\,1} + I^2_i\, \sigma^2_{i\,2}) \int_0^t L^2\, \frac{(1+s)^{2\,L-2}}{(1+t)^{2\,L}}\, ds = \frac{(\sigma^2_{i\,1} + I^2_i\, \sigma^2_{i\,2})}{4}\left(\frac{\log(1+t)}{1+t}\right) .
$$

Thus, for each $\epsilon > 0$

$$
\mathbb{P}^*\left(\left|\hat{\theta}^i_t - \theta^{*,i}\right| > \epsilon\right) \leq \frac{\mathbb{E}^{\mathbb{P}^*}\left[(\hat{\theta}^i_t - \theta^{*,i})^2\right]}{\epsilon^2} \leq C_\epsilon \max\left\{\frac{\log(1+t)}{1+t}, \frac{1}{1+t}\right\} , \tag{A.1}
$$

where C_ϵ is a constant that depends on ϵ. Then, $\hat{\theta}^i_t$ converges to $\theta^{*,i}$ in probability as t goes to infinity.

Appendix B Proof that the value function is finite

First, there exists a constant M such that $M \geq U(X)$ for all $X \in \mathbb{R}$, so $v \leq M$. The constant process $\alpha = 0$ is an admissible control satisfying $X^\alpha_T = X\, e^{r\,(T-t)}$ because with this control process the agent only invests in the risk-free asset. Then, substitute X^α_T to obtain

$$
v(t, X, S, \hat{\theta}) \geq \inf_{\mathbb{P} \in \mathcal{P}(t,x,G)} \mathbb{E}^{\mathbb{P}}\left[-e^{-\gamma\, X\, e^{r\,(T-t)}}\right] = -e^{-\gamma\, X\, e^{r\,(T-t)}} . \tag{B.1}
$$

Therefore, $-\infty < v \leq M$, which implies that v is finite and concludes the proof.

Appendix C Additional details in the proof of Proposition 2.3

Appendix C.1 First part

(1) If $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{+,+}$ (first quadrant of $\mathbb{R}^2$), then

$$\begin{aligned}\sup_{\tilde\theta\in G(t,\hat\theta)} &\left\{\beta_t\,(\tilde\theta-\hat\theta)\,\partial_{\hat\theta}U-\alpha^\intercal\left(\tilde\theta-\mathbf{1}\,r\right)U+\tfrac12\,\alpha^\intercal\,\Sigma\,\alpha\,U-\beta_t\,\alpha^\intercal\,\Sigma\,\partial_{\hat\theta}\,U\right\}\\ &=\beta_t\,c/\sqrt{1+t}\,\mathbf{1}^\intercal\,\partial_{\hat\theta}U-\alpha^\intercal\left(\hat\theta-\mathbf{1}\,(r-c/\sqrt{1+t})\right)U\\ &\quad+\tfrac12\,\alpha^\intercal\,\Sigma\,\alpha\,U-\beta_t\,\alpha^\intercal\,\Sigma\,\partial_{\hat\theta}\,U\,,\end{aligned}\tag{C.1}$$

and the infimum is attained at either

$$\alpha^*=\beta_t\,\frac{\partial_{\hat\theta}U}{U}\qquad\text{or}\qquad\alpha^*=\Sigma^{-1}\left(\hat\theta-\mathbf{1}\,r+\mathbf{1}^{+,+}\,c/\sqrt{1+t})\right)+\beta_t\,\frac{\partial_{\hat\theta}\,U}{U}\,,\tag{C.2}$$

or

$$\alpha^{*,1}=\beta_t\,\frac{\partial_{\hat\theta_1}U}{U}\qquad\text{and}\qquad\alpha^{*,2}=\frac{\hat\theta_2-r+c/\sqrt{1+t}}{\sigma_2^2}+\beta_t\,\frac{\partial_{\hat\theta_2}\,U}{U}\,,\tag{C.3}$$

or

$$\alpha^{*,2}=\beta_t\,\frac{\partial_{\hat\theta_2}U}{U}\qquad\text{and}\qquad\alpha^{*,1}=\frac{\hat\theta_1-r+c/\sqrt{1+t}}{\sigma_1^2}+\beta_t\,\frac{\partial_{\hat\theta_1}\,U}{U}\,.\tag{C.4}$$

(2) If $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{-,+}$ (second quadrant of $\mathbb{R}^2$), the infimum is attained at either

$$\alpha^*=\beta_t\,\frac{\partial_{\hat\theta}U}{U}\qquad\text{or}\qquad\alpha^*=\Sigma^{-1}\left(\hat\theta-\mathbf{1}\,r+\mathbf{1}^{-,+}\,c/\sqrt{1+t}\right)+\beta_t\,\frac{\partial_{\hat\theta}\,U}{U}\,,\tag{C.5}$$

or

$$\alpha^{*,1}=\beta_t\,\frac{\partial_{\hat\theta_1}U}{U}\qquad\text{and}\qquad\alpha^{*,2}=\frac{\hat\theta_2-r+c/\sqrt{1+t}}{\sigma_2^2}+\beta_t\,\frac{\partial_{\hat\theta_2}\,U}{U}\,,\tag{C.6}$$

or

$$\alpha^{*,2}=\beta_t\,\frac{\partial_{\hat\theta_2}U}{U}\qquad\text{and}\qquad\alpha^{*,1}=\frac{\hat\theta_1-r-c/\sqrt{1+t}}{\sigma_1^2}+\beta_t\,\frac{\partial_{\hat\theta_1}\,U}{U}\,.\tag{C.7}$$

(3) If $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{-,-}$ (third quadrant of $\mathbb{R}^2$), the infimum is attained at either

$$\alpha^*=\beta_t\,\frac{\partial_{\hat\theta}U}{U}\qquad\text{or}\qquad\alpha^*=\Sigma^{-1}\left(\hat\theta-\mathbf{1}\,r+\mathbf{1}^{-,-}\,c/\sqrt{1+t})\right)+\beta_t\,\frac{\partial_{\hat\theta}\,U}{U}\,,\tag{C.8}$$

or

$$\alpha^{*,1}=\beta_t\,\frac{\partial_{\hat\theta_1}U}{U}\qquad\text{and}\qquad\alpha^{*,2}=\frac{\hat\theta_2-r-c/\sqrt{1+t}}{\sigma_2^2}+\beta_t\,\frac{\partial_{\hat\theta_2}\,U}{U}\,,\tag{C.9}$$

or

$$\alpha^{*,2}=\beta_t\,\frac{\partial_{\hat\theta_2}U}{U}\qquad\text{and}\qquad\alpha^{*,1}=\frac{\hat\theta_1-r-c/\sqrt{1+t}}{\sigma_1^2}+\beta_t\,\frac{\partial_{\hat\theta_1}\,U}{U}\,.\tag{C.10}$$

(4) If $\beta_t\, \partial_{\hat\theta} U - \alpha\, U \in \mathbb{R}^{+,-}$ (fourth quadrant of $\mathbb{R}^2$), the infimum is attained at either

$$\alpha^*=\beta_t\,\frac{\partial_{\hat\theta}U}{U}\qquad\text{or}\qquad\alpha^*=\Sigma^{-1}\left(\hat\theta-\mathbf{1}\,r+\mathbf{1}^{+,-}\,c/\sqrt{1+t})\right)+\beta_t\,\frac{\partial_{\hat\theta}\,U}{U}\,,\tag{C.11}$$

or

$$\alpha^{*,1} = \beta_t \frac{\partial_{\hat{\theta}_1} U}{U} \quad \text{and} \quad \alpha^{*,2} = \frac{\hat{\theta}_2 - r - c/\sqrt{1+t}}{\sigma_2^2} + \beta_t \frac{\partial_{\hat{\theta}_2} U}{U}, \tag{C.12}$$

or

$$\alpha^{*,2} = \beta_t \frac{\partial_{\hat{\theta}_1} U}{U} \quad \text{and} \quad \alpha^{*,1} = \frac{\hat{\theta}_1 - r + c/\sqrt{1+t}}{\sigma_1^2} + \beta_t \frac{\partial_{\hat{\theta}_1} U}{U}. \tag{C.13}$$

Appendix C.2 Second part

Consider $\hat{\theta} \in A^{-,-}$ and study the value α in the following four cases.
Case 1: $\beta_t \, \partial_{\hat{\theta}} U - \alpha\, U \in \mathbb{R}^{+,+}$, then the infimum is attained at

$$\alpha^* = \Sigma^{-1} \left(\hat{\theta} - \mathbf{1} r + \mathbf{1}^{+,+} c/\sqrt{1+t}) \right) + \beta_t \frac{\partial_{\hat{\theta}} U}{U},$$

because $\hat{\theta} \in A^{-,-}$.
Case 2: $\beta_t \, \partial_{\hat{\theta}} U - \alpha\, U \in \mathbb{R}^{-,+}$. Then

$$\theta_1 - r + \frac{c}{\sqrt{1+t}} \leq 0 \quad \text{and} \quad \sigma_2^2 \left(\hat{\theta}_1 - r + \frac{c}{\sqrt{1+t}} \right) \leq \rho\, \sigma_c^2 \left(\hat{\theta}_2 - r + \frac{c}{\sqrt{1+t}} \right)$$

because $\hat{\theta} \in A^{-,-}$. Therefore, the value $\beta_t \, \partial_{\hat{\theta}} U - \alpha\, U \notin \mathbb{R}^{-,+}$ when

$$\alpha^2 = \beta_t \frac{\partial_{\hat{\theta}_2} U}{U} \quad \text{and} \quad \alpha^1 = (\sigma_1^2)^{-1} \left(\hat{\theta}_1 - r - c/\sqrt{1+t}) \right) + \beta_t \frac{\partial_{\hat{\theta}_1} U}{U}, \tag{C.14}$$

or

$$\alpha = \Sigma^{-1} \left(\hat{\theta} - \mathbf{1} r + \mathbf{1}^{-,+} c/\sqrt{1+t}) \right) + \beta_t \frac{\partial_{\hat{\theta}} U}{U}.$$

Therefore, the infimum is either attained at

$$\alpha^1 = \beta_t \frac{\partial_{\hat{\theta}_1} U}{U} \quad \text{and} \quad \alpha^2 = (\sigma_2^2)^{-1} \left(\hat{\theta}_2 - r + c/\sqrt{1+t}) \right) + \beta_t \frac{\partial_{\hat{\theta}_2} U}{U}, \tag{C.15}$$

or at $\alpha = \beta_t \partial_{\hat{\theta}} U / U$.
Case 3: $\beta_t \, \partial_{\hat{\theta}} U - \alpha\, U \in \mathbb{R}^{+,-}$. Similar to Case 2, the infimum is attained at

$$\alpha^2 = \beta_t \frac{\partial_{\hat{\theta}_2} U}{U} \quad \text{and} \quad \alpha^1 = (\sigma_1^2)^{-1} \left(\hat{\theta}_1 - r + c/\sqrt{1+t}) \right) + \beta_t \frac{\partial_{\hat{\theta}_1} U}{U}, \tag{C.16}$$

or at $\alpha = \beta_t \partial_{\hat{\theta}} U / U$.
Case 4: $\beta_t \, \partial_{\hat{\theta}} U - \alpha\, U \in \mathbb{R}^{+,-}$. It is easy to check that the infimum is only attained at $\alpha = \beta_t \partial_{\hat{\theta}} U / U$.

Note that all possible optimal strategies are of the form $K + \beta_t \, \partial_{\hat{\theta}} U / U$ for some K. Therefore, by substituting into (C.1), we have that

$$\sup_{\tilde{\theta} \in G(t,x)} \left\{ \beta_t \, (\tilde{\theta} - \hat{\theta})\, \partial_{\hat{\theta}} U - (K + \beta_t \partial_{\hat{\theta}} U/U)^\intercal \left(\tilde{\theta} - r \right) U \right\} + \frac{1}{2} U\, K^\intercal\, \Sigma\, K - \frac{\beta_t^2}{2U}\, \partial_{\hat{\theta}} U^\intercal\, \Sigma\, \partial_{\hat{\theta}} U. \tag{C.17}$$

For Case 1: $K = \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right)$ and $\beta_t\,\partial_{\hat{\theta}} U - \alpha\,U \in \mathbb{R}^{+,+}$, then (C.17) is

$$\begin{aligned}
&\frac{\beta_t\,c}{\sqrt{1+t}}(\partial_{\theta_1} U + \partial_{\theta_2} U) - \frac{\beta_t\,(\partial_\theta U)^\intercal}{U}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right)U \\
&\quad - \frac{1}{2}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right)^\intercal \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right) \\
&= \beta_t\,(\partial_\theta U)^\intercal\left(\hat{\theta} - \mathbf{1}\,r\right) - \frac{1}{2}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right)^\intercal \\
&\quad \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t})\right).
\end{aligned} \tag{C.18}$$

Similarly, for Case 2, we have that (C.17) is

$$\beta_t\,(\partial_\theta U)^\intercal\left(\hat{\theta} - \mathbf{1}\,r\right) - \frac{1}{2}\left(\hat{\theta}_2 - r + c/\sqrt{1+t})\right)^\intercal (\sigma_2^2)^{-1}\left(\hat{\theta}_2 - r + c/\sqrt{1+t})\right), \tag{C.19}$$

for Case 3, we have that (C.17) is

$$\beta_t\,(\partial_\theta U)^\intercal\left(\hat{\theta} - \mathbf{1}\,r\right) - \frac{1}{2}\left(\hat{\theta}_1 - r + c/\sqrt{1+t})\right)^\intercal (\sigma_1^2)^{-1}\left(\hat{\theta}_1 - r + c/\sqrt{1+t})\right), \tag{C.20}$$

and for Case 4, we have that (C.17) is

$$\beta_t\,(\partial_\theta U)^\intercal\left(\hat{\theta} - \mathbf{1}\,r\right). \tag{C.21}$$

Then, it is straightforward to check Case 1 includes the minimal value. For the other regions, we follow the same argument and this completes the proof.

Appendix C.3 Third part

Next, we solve the inf-sup problem in (2.27) for each $\hat{\theta}$ in the nine non-overlapping investment regions of $\mathbb{R}^2$ discussed above. If $\hat{\theta} \in A^{-,-} \subset \mathbb{R}^2$, the optimal investment strategy is

$$\alpha^* = \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,+}\,c/\sqrt{1+t}\right) - \beta_t\,\partial_{\hat{\theta}} H, \tag{C.22}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t\,(\mathbf{1}\,r - \hat{\theta})^\intercal\,\partial_{\hat{\theta}} H + F(t,\hat{\theta})\right). \tag{C.23}$$

If $\hat{\theta} \in A^{+,-} \subset \mathbb{R}^2$, the optimal strategy is

$$\alpha^* = \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{-,+}\,c/\sqrt{1+t})\right) - \beta_t\,\partial_{\hat{\theta}} H, \tag{C.24}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t\,(\mathbf{1}\,r - \hat{\theta})^\intercal\,\partial_{\hat{\theta}} H + F(t,\hat{\theta})\right). \tag{C.25}$$

If $\hat{\theta} \in A^{-,+} \subset \mathbb{R}^2$, the optimal strategy is

$$\alpha^* = \Sigma^{-1}\left(\hat{\theta} - \mathbf{1}\,r + \mathbf{1}^{+,-}\,c/\sqrt{1+t})\right) - \beta_t\,\partial_{\hat{\theta}} H, \tag{C.26}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t\,(\mathbf{1}\,r - \hat{\theta})^\intercal\,\partial_{\hat{\theta}} H + F(t,\hat{\theta})\right). \tag{C.27}$$

If $\hat{\theta} \in A^{+,+} \subset \mathbb{R}^2$, the optimal strategy is

$$\alpha^* = \Sigma^{-1} \left(\hat{\theta} - \mathbf{1} r + \mathbf{1}^{-,-} c/\sqrt{1+t})\right) - \beta_t \; \partial_{\hat{\theta}} H \, , \tag{C.28}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t \, (\mathbf{1} r - \hat{\theta})^\intercal \, \partial_{\hat{\theta}} H + F(t,\hat{\theta})\right) . \tag{C.29}$$

If $\hat{\theta} \in A^{0,-} \subset \mathbb{R}^2$, then

$$\alpha^{*,1} = -\beta_t \; \partial_{\hat{\theta}_1} H \quad \text{and} \quad \alpha^{*,2} = \frac{\hat{\theta}_2 - r + c/\sqrt{1+t}}{\sigma_2^2} - \beta_t \; \partial_{\hat{\theta}_2} H \, , \tag{C.30}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t \, (\mathbf{1} r - \hat{\theta})^\intercal \, \partial_{\hat{\theta}} H + \frac{1}{2}\left(\hat{\theta}_2 - r + c/\sqrt{1+t}\right) \frac{1}{\sigma_2^2} \left(\hat{\theta}_2 - r + c/\sqrt{1+t}\right)\right) . \tag{C.31}$$

If $\hat{\theta} \in A^{+,0} \subset \mathbb{R}^2$, then

$$\alpha^{*,1} = \frac{\hat{\theta}_1 - r - c/\sqrt{1+t}}{\sigma_1^2} - \beta_t \; \partial_{\hat{\theta}_1} H \quad \text{and} \quad \alpha^{*,2} = -\beta_t \; \partial_{\hat{\theta}_2} H \, , \tag{C.32}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t \, (\mathbf{1} r - \hat{\theta})^\intercal \, \partial_{\hat{\theta}} H + \frac{1}{2}\left(\hat{\theta}_1 - r - c/\sqrt{1+t}\right) \frac{1}{\sigma_1^2} \left(\hat{\theta}_1 - r - c/\sqrt{1+t}\right)\right) . \tag{C.33}$$

If $\hat{\theta} \in A^{0,+} \subset \mathbb{R}^2$, then

$$\alpha^{*,1} = -\beta_t \; \partial_{\hat{\theta}_1} H \quad \text{and} \quad \alpha^{*,2} = \frac{\hat{\theta}_2 - r - c/\sqrt{1+t}}{\sigma_2^2} - \beta_t \; \partial_{\hat{\theta}_2} H \, , \tag{C.34}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t \, (\mathbf{1} r - \hat{\theta})^\intercal \, \partial_{\hat{\theta}} H + \frac{1}{2}\left(\hat{\theta}_2 - r - c/\sqrt{1+t}\right) \frac{1}{\sigma_2^2} \left(\hat{\theta}_2 - r - c/\sqrt{1+t}\right)\right) . \tag{C.35}$$

If $\hat{\theta} \in A^{-,0} \subset \mathbb{R}^2$, then

$$\alpha^{*,1} = \frac{\hat{\theta}_1 - r + c/\sqrt{1+t}}{\sigma_1^2} - \beta_t \; \partial_{\hat{\theta}_1} H \quad \text{and} \quad \alpha^{*,2} = -\beta_t \; \partial_{\hat{\theta}_2} H \, , \tag{C.36}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t \, (\mathbf{1} r - \hat{\theta})^\intercal \, \partial_{\hat{\theta}} H + \frac{1}{2}\left(\hat{\theta}_1 - r + c/\sqrt{1+t}\right) \frac{1}{\sigma_1^2} \left(\hat{\theta}_1 - r + c/\sqrt{1+t}\right)\right) . \tag{C.37}$$

If $\hat{\theta} \in A^{0,0} \subset \mathbb{R}^2$, then

$$\alpha^* = -\beta_t \; \partial_{\hat{\theta}} H \, , \tag{C.38}$$

and

$$C(t,\hat{\theta}) = -U\left(\beta_t \, (\mathbf{1} r - \hat{\theta})^\intercal \, \partial_{\hat{\theta}} H\right) . \tag{C.39}$$

Appendix D Proof of Proposition 3.1

First, we show that $A(t, T, L)$ is an increasing function on t. Recall that

$$A(t,T,L) = \begin{cases} 1 - \frac{L}{2L-1}\left(1 - \left(\frac{1+t}{1+T}\right)^{2L-1}\right) & \text{if} \quad L \neq 1/2\,, \\ 1 - \frac{1}{2}\,(\log(1+T) - \log(1+t)) & \text{if} \quad L = 1/2\,. \end{cases} \tag{D.1}$$

Then, by direct computation, the partial derivative of A with respect to t is

$$\partial_t A(t,T,L) = \begin{cases} L\,\frac{(1+t)^{2L-2}}{(1+T)^{2L-1}} & \text{if} \quad L \neq 1/2\,, \\ \frac{1}{2\,(1+t)} & \text{if} \quad L = 1/2\,. \end{cases} \tag{D.2}$$

Therefore, the partial derivative is always positive and the function $A(t, T, L)$ is increasing in t. Thus, for all $t \leq T$, the adjustment term satisfies $A(0, T, L) \leq A(t, T, L) \leq A(T, T, L) = 1$. Then, by substitution, we obtain the upper and lower bounds.

Appendix E Proof of Proposition 3.2

Next, we discuss the adaptive investment strategy for four ranges of the value of the learning parameter L. If $L \geq 1$, it is easy to see that the term $A(0, T, L)$

$$A(0,T,L) = 1 - \frac{L}{2\,L-1}\left(1 - \left(\frac{1}{1+T}\right)^{2\,L-1}\right) > 0\,,$$

for $0 \leq t \leq T$ because $L/(2\,L-1) \leq 1$ and

$$\left(1 - \left(\frac{1+t}{1+T}\right)^{2\,L-1}\right) \leq 1\,.$$

Therefore, $A(t, T, L)$ is always positive. When $L \neq 1/2$ and $L < 1$, we have that

$$\left(\frac{1-L}{L}\right)^{1/(2\,L-1)} < 1$$

because either $(1-L)/L < 1$ and $1/(2\,L-1) > 0$ or $(1-L)/L > 1$ and $1/(2\,L-1) < 0$. Consider

$$t^L = (1+T)\left(\frac{1-L}{L}\right)^{1/(2\,L-1)} - 1\,, \tag{E.1}$$

then, it is easy to check that $A(t^L, T, L) = 0$. When $L = 1/2$, we consider

$$t^L = (1+T)^{1/2} - 1\,,$$

and it is easy to check that $A(t^L, T, L) = 0$. From Proposition 3.1, the function $A(t, T, L)$ is an increasing function on t. Then, if $t > t^L$, then $A(t, T, L) > 0$ and if $t < t^L$, then $A(t, T, L) < 0$.

Appendix F Proof of lemma 5.2

The second equation follows directly from the first equation. Let $Z \sim \mathcal{N}(0,1)$ and $Y = \mu + \sigma Z$. Write

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}\left[Y^2\,\mathbf{1}_{\{Y>0\}}\right] &= \mathbb{E}^{\mathbb{P}}\left[(\mu+\sigma\,Z)^2\,\mathbf{1}_{\{Z>-\mu/\sigma\}}\right] \\ &= \mu^2\,(1-\Phi(-\mu/\sigma)) + 2\,\mu\,\sigma\,\mathbb{E}^{\mathbb{P}}\left[Z\,\mathbf{1}_{\{Z>-\mu/\sigma\}}\right] \\ &\quad + \sigma^2\,\mathbb{E}^{\mathbb{P}}\left[Z^2\,\mathbf{1}_{\{Z>-\mu/\sigma\}}\right]. \end{aligned} \tag{F.1}$$

From the density of the standard normal distribution we obtain

$$\mathbb{E}^{\mathbb{P}}\left[Z\,\mathbf{1}_{\{Z>-\mu/\sigma\}}\right] = \frac{1}{\sqrt{2\,\pi}}\int_{-\mu/\sigma}^{\infty} z\,e^{-z^2/2}\,dz = \frac{1}{\sqrt{2\,\pi}}\,\exp\left(-\mu^2/2\,\sigma^2\right) \tag{F.2}$$

and

$$\mathbb{E}^{\mathbb{P}}\left[Z^2\,\mathbf{1}_{\{Z>-\mu/\sigma\}}\right] = \int_{-\mu/\sigma}^{\infty} z^2\,e^{-z^2/2}\,dz = -\frac{1}{\sqrt{2\,\pi}}\,\mu/\sigma\,\exp\left(-\mu^2/2\,\sigma^2\right)+1-\Phi(-\mu/\sigma)\,. \tag{F.3}$$

Therefore, (F.1), (F.2), and (F.3) imply (5.3).

Appendix G Proof of lemma 5.3

To simplify notation, let $\Sigma_i = \Sigma_i(u)$. From Lemma 5.2, we write

$$h_1(u,\hat{\theta}_i) = (f^2+\Sigma_i^2)\,(1-\Phi(-f/\Sigma_i)) + f\,\Sigma_i\,\phi(-f/\Sigma_i)$$

and compute its gradient with respect to $\hat{\theta}$ to obtain

$$\begin{aligned} \partial_{\hat{\theta}_i} h_1 =&\; 2\,f\,\partial_{\hat{\theta}_i} f\,(1-\Phi(-f/\Sigma_i)) + (f^2+\Sigma_i^2)\,\partial_{\hat{\theta}_i}\Phi(f/\Sigma_i) + \Sigma_i\,\phi(-f/\Sigma_i) \\ &+ f\,\Sigma_i\,\partial_{\hat{\theta}_i}\phi(-f/\Sigma_i) \\ =&\; 2\,f\,\partial_{\hat{\theta}_i} f\,(1-\Phi(-f/\Sigma_i)) + (f^2+\Sigma_i^2)\,\partial_{\hat{\theta}_i} f/\Sigma_i\,\phi(f/\Sigma_i) \\ &+ \partial_{\hat{\theta}_i} f\,\Sigma_i\,\phi(-f/\Sigma_i) - (f^2/\Sigma_i)\,\partial_{\hat{\theta}} f\,\phi(-f/\Sigma_i) \\ =&\; 2\,f\,\partial_{\hat{\theta}_i} f\,(1-\Phi(-f/\Sigma_i)) + 2\,\partial_{\hat{\theta}_i} f\,\Sigma_i\,\phi(-f/\Sigma_i)\,, \end{aligned} \tag{G.1}$$

where we use

$$\partial_{\hat{\theta}_i}\Phi(f/\Sigma_i) = 1/\Sigma_i\,\partial_{\hat{\theta}_i} f\,\phi(f/\Sigma_i) \quad \text{and} \quad \partial_{\hat{\theta}_i}\phi(-f/\Sigma_i) = -f/\Sigma_i^2\,\partial_{\hat{\theta}_i} f\,\phi(-f/\Sigma_i)\,.$$

Similarly, for

$$h_2(u,\hat{\theta}_i) = (\tilde{f}^2+\Sigma_i^2)\,\Phi\left(-\tilde{f}/\Sigma_i\right) - \tilde{f}\,\Sigma_i\,\Phi\left(-\tilde{f}/\Sigma_i\right),$$

its gradient with respect to $\hat{\theta}$ is given by

$$\partial_{\hat{\theta}} h_2 = 2\,\tilde{f}\,\partial_{\hat{\theta}_i}\tilde{f}\,\Phi\left(-\tilde{f}/\Sigma_i\right) - 2\,\Sigma_i\,\partial_{\hat{\theta}_i}\tilde{f}\,\phi\left(-\tilde{f}/\Sigma_i\right). \tag{G.2}$$

Let $h(u,\hat{\theta}_i) = h_1(u,\hat{\theta}_i) + h_2(u,\hat{\theta}_i)$; from Lemma 5.2 we have that

$$h(u,\hat{\theta}_i) = \mathbb{E}^{\mathbb{P}}\left[\left(Z_u^i - r - c/\sqrt{1+u}\right)^2\,\mathbf{1}_{\{Z_u^i > c/\sqrt{1+u}\}}\right]$$

$$+\,\mathbb{E}^{\mathbb{P}}\left[\left(Z_u^i - r + c/\sqrt{1+u}\right)^2 \mathbf{1}_{\{Z_u^i < -c/\sqrt{1+u}\}}\right].$$

Then,

$$\partial_{\hat{\theta}_i} h = \partial_{\hat{\theta}_i} f \left[2\, f\,(1-\Phi(-f/\Sigma_i)) + 2\,\Sigma_i\,\phi(-f/\Sigma_i) + 2\,\tilde{f}\,\Phi\left(-\tilde{f}/\Sigma_i\right) - 2\,\Sigma_i\,\phi\left(-\tilde{f}/\Sigma_i\right)\right]. \tag{G.3}$$

Appendix H Proof of lemma 5.4

Proof Recall that

$$f(u,\hat{\theta}_i) = (\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}\,,$$
$$\tilde{f}(u,\hat{\theta}_i) = (\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c/\sqrt{1+u}\,,$$

and $\Sigma_i = \Sigma_i(u)$. We write

$$\partial_{\hat{\theta}_i} h = \partial_{\hat{\theta}_i} f \left[2\, f\,\Phi(f/\Sigma_i) + 2\,\Sigma_i\,\phi(-f/\Sigma_i) + 2\,\tilde{f}\,\Phi\left(-\tilde{f}/\Sigma_i\right) - 2\,\Sigma_i\,\phi\left(-\tilde{f}/\Sigma_i\right)\right]. \tag{H.1}$$

By Taylor expansion, the following equality holds

$$\begin{aligned}
&2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}\right)\Phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}}{\Sigma}\right)\\
&= 2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}\right)\\
&\quad\left[\Phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right) - \frac{c}{\sqrt{1+u}\,\Sigma_i(u)}\,\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right.\\
&\quad\left.+\frac{c^2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u}\right)}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\,\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u}}{\Sigma_i(u)}\right)\right].
\end{aligned} \tag{H.2}$$

Next, collect the terms with at least an order of c^2 on the right-hand side of (H.2), and we see that these terms are dominated by

$$\begin{aligned}
\tilde{R}_1(u,\hat{\theta}_i) = &\frac{2\,\phi\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L/\Sigma_i(u)\right)}{(1+u)\,\Sigma_i(u)}\\
&+\max_{0\le c'\le c}\left[\frac{\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u}\right)\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}\right)}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\times\right.
\end{aligned}$$

$$\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u}}{\Sigma_i(u)}\right)\Bigg]. \tag{H.3}$$

Now, we need to choose c such that the function $\tilde{R}_1(u, \hat{\theta}_i)$ is integrable. The first term on the right-hand side of (H.3) is integrable by the property of the normal density. For the second term, the maximum is attained at either 0 or c, and the local minimum satisfies

$$(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u} = \Sigma_i(u).$$

Note that the following sum is an integrable function

$$\Bigg[\frac{\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u}\right)\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}\right)}{2\,\Sigma(u)^3\,(1+u)^{3/2}} \times$$
$$\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}}{\Sigma(u)}\right)\Bigg]$$
$$+\left[\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u}\right)}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\,\phi\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L/\Sigma_i(u)\right)\right],$$

and it dominates the second term on the right-hand side of (H.3) when the maximum is attained at either 0 or c. For the local minimum, we choose $\delta^1_{t,\hat{\theta}_i}$ small enough such that for all $c < \delta^1_{t,\hat{\theta}_i}$, the local minimum lies in $[0, c]$ when $u > t$. Therefore, $\tilde{R}_1$ is integrable. Next, we apply Taylor expansion around $c = 0$ to the second term on the right-hand side of (H.1) to obtain

$$2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c/\sqrt{1+u}\right)\Phi\left(-\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c/\sqrt{1+u}}{\Sigma_i}\right)$$
$$= 2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c/\sqrt{1+u}\right)$$
$$\left[\Phi\left(-\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right) - \frac{c}{\sqrt{1+u}\,\Sigma_i(u)}\,\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right.$$
$$\left. + \frac{c^2\left((\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c'/\sqrt{1+u}\right)}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\,\phi\left(\frac{(\hat{\theta}_i - r)\left(\frac{1+t}{1+u}\right)^L + c'/\sqrt{1+u}}{\Sigma_i(u)}\right)\right]. \tag{H.4}$$

Then, we define an integrable function $\tilde{R}_2$ and $\delta^2_{t,\hat{\theta}_i}$ similar to the case in (H.3). For the the third term and the fourth term on the right-hand side of (H.1), we write

$$
\begin{aligned}
&2\,\Sigma_i(u)\,\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c/\sqrt{1+u}}{\Sigma_i}\right)\\
&= 2\,\Sigma_i(u)\left[\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\Sigma(u)}\right) + \frac{c\,(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\sqrt{1+u}\,\Sigma_i(u)^2}\,\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right.\\
&\left.+\frac{c^2\,((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u})}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right],
\end{aligned}
\tag{H.5}
$$

and

$$
\begin{aligned}
&-2\,\Sigma_i(u)\,\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L + c/\sqrt{1+u}}{\Sigma_i}\right)\\
&= -2\,\Sigma_i(u)\left[\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right) - \frac{c\,(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\sqrt{1+u}\,\Sigma_i(u)}\,\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right.\\
&\left.+\frac{c^2\,((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L - c'/\sqrt{1+u})}{2\,\Sigma_i(u)^3\,(1+u)^{3/2}}\phi\left(\frac{(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L}{\Sigma_i(u)}\right)\right],
\end{aligned}
\tag{H.6}
$$

and there are integrable functions $\tilde{R}_3$ and $\tilde{R}_4$ and $\delta^3_{t,\hat{\theta}_i}$ and $\delta^4_{t,\hat{\theta}_i}$. By combining (H.2), (H.4), (H.5), and (H.6), we obtain

$$
\begin{aligned}
&\left|\partial_{\hat{\theta}_i} h(u,\hat{\theta}_i) - 2\,(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^{2L} - \frac{2\,c}{\sqrt{1+u}}\,\left(\frac{1+t}{1+u}\right)^{2L}\right.\\
&\qquad\left.\left[\Phi\left(-(\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L/\Sigma_i(u)\right) - \Phi\left((\hat{\theta}_i - r)\,\left(\frac{1+t}{1+u}\right)^L/\Sigma_i(u)\right)\right]\right|\\
&\le c^2\,R(u,\hat{\theta}_i)\,,
\end{aligned}
\tag{H.7}
$$

where $\tilde{R} = \tilde{R}_1 + \tilde{R}_2 + \tilde{R}_3 + \tilde{R}_4$ and $\delta_{t,\hat{\theta}} = \min\left\{\delta^1_{t,\hat{\theta}_i}, \delta^2_{t,\hat{\theta}_i}, \delta^3_{t,\hat{\theta}_i}, \delta^4_{t,\hat{\theta}_i}\right\}$. □

Data Availability Data sets of simulations generated are available from the corresponding author.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Bayraktar, E., Chen, T.: Nonparametric adaptive robust control under model uncertainty. SIAM J. Control. Optim. **61**(5), 2737–2760 (2023)
2. Bhudisaksang, T., Cartea, Á.: Adaptive robust control in continuous time. SIAM J. Control. Optim. **59**(5), 3912–3945 (2021)
3. Bhudisaksang, T., Cartea, Á.: Online drift estimation for jump-diffusion processes. Bernoulli **27**(4), 2494–2518 (2021)
4. Bielecki, T.R., Chen, T., Cialenco, I., Cousin, A., Jeanblanc, M.: Adaptive robust control under model uncertainty. SIAM J. Control. Optim. **57**(2), 925–946 (2019)
5. Bismuth, A., Guéant, O., Pu, J.: Portfolio choice, portfolio liquidation, and portfolio transition under drift uncertainty. Math. Financ. Econ. **13**(4), 661–719 (2019)
6. Bordigoni, G., Matoussi, A., and Schweizer, M.: A stochastic control approach to a robust utility maximization problem. In Stochastic analysis and applications, pages 125–151. Springer (2007)
7. Cartea, Á., Jaimungal, S., Penalva, J.: Algorithmic and high-frequency trading. Cambridge University Press (2015)
8. Cialenco, I. and Kováčová, G.: Vector-valued robust stochastic control. arXiv preprint arXiv:2407.00266 (2024)
9. De Franco, C., Nicolle, J., Pham, H.: Bayesian learning for the Markowitz portfolio selection problem. International Journal of Theoretical and Applied Finance **22**(07), 1950037 (2019)
10. Kallsen, J. and Muhle-Karbe, J.: On using shadow prices in portfolio optimization with transaction costs. The Annals of Applied Probability, pages 1341–1358 (2010)
11. Kim, T.S., Omberg, E.: Dynamic nonmyopic portfolio behavior. The Review of Financial Studies **9**(1), 141–161 (1996)
12. Liang, Z., Ma, M.: Robust consumption-investment problem under CRRA and CARA utilities with time-varying confidence sets. Mathematical Finance (2018)
13. Liu, R., Muhle-Karbe, J.: Portfolio choice with stochastic investment opportunities: A user's guide. arXiv preprint arXiv:1311.1715 (2013)
14. Merton, R.: Optimum consumption and portfolio rules in a continuous-time model. Journal of Economic Theory **3**(4), 373–413 (1971)
15. Neufeld, A., Nutz, M.: Robust utility maximization with Lévy processes. Math. Financ. **28**(1), 82–105 (2018)
16. Papaioannou, D.: Applied multidimensional Girsanov theorem. Available at SSRN 1805984, (2012)
17. Shreve, S.E., Soner, H. M.: Optimal investment and consumption with transaction costs. The Annals of Applied Probability, pages 609–692 (1994)
18. Xia, Y.: Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation. J. Financ. **56**(1), 205–246 (2001)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.