Probability, Uncertainty and Quantitative Risk
Vol. 9, No. 4, 2024, 529–552

https://www.aimsciences.org/PUQR
DOI:10.3934/puqr.2024022

The optimal strategy of the dynamic mean–variance problem for pairs trading with a common stochastic factor

Yaoyuan Zhang, Dewen Xiong*

School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai 200240, *China*

Email: volwc@sjtu.edu.cn, xiongdewen@sjtu.edu.cn

Abstract This paper studies the optimal pairs trading strategy of the mean−variance (MV) objective function under a continuous-time cointegration model with a common stochastic factor. Although this common stochastic factor is not directly tradable, it significantly impacts asset prices. We first provide a semiclosed-form solution under a general model. We then specify the common factor model to be a mean-reverting process with time-varying parameters and provide closed-form optimal strategies for pairs trading with fixed and flexible ratios, respectively. Empirical analysis based on historical data from Chinese securities markets shows the effectiveness of both optimal strategies. The optimal flexible-ratio strategy outperforms the optimal fixed-ratio strategy in terms of both profit and risk.

Keywords Continuous-time cointegration model, Dynamic mean–variance problem, Pairs trading, Mean-reverting process

2020 Mathematics Subject Classification 91G10, 93E20, 35Q62

1. Introduction

Pairs trading is a type of statistical arbitrage strategy that emerged in the 1980s, which profits by capturing short-term deviations between stock pairs from some long-term equilibrium relationship. In this paper, we consider the dynamic mean–variance (MV) problem for pairs trading with a common stochastic factor.

An important tool in pairs trading is cointegration, which was developed by Engle and Granger [8]. Vidyamurthy [15] proposed a cointegration framework for pairs trading problems, in which the comovement of assets is identified by cointegration testing. They designed a simple trading strategy based on a nonparametric threshold. However, this strategy is not optimal. Researchers often employ the stochastic control method with different utility functions to find the optimal strategy for pairs trading. Jurek and Yang [10] provided a closed-form solution of the optimal asset allocation strategy between a mean-reverting trading pair and a risk-free asset with the CRRA utility. Liu and Timmerman [11] studied the optimal strategy on cointegrated asset pairs with the power utility. Chiu and Wong [4–6] explored the optimal dynamic trading

Received 20 November 2023; Accepted 13 June 2024; Early access 2 September 2024
*Corresponding author

strategy of cointegrated assets with the MV criterion, and extended it to asset–liability management problems with insurance liability. However, the optimal strategies in [4–6] are time-inconsistent. Thus, Chiu and Wong [7] explored the time-consistent dynamic MV problem for pairs trading with a constant cointegration matrix and provided a closed-form optimal strategy by solving the corresponding Hamilton–Jacobi–Bellman (HJB) equation directly. Shen and Sui [14] discussed the optimal investment and consumption problem for the power and logarithmic utility, while Ma and Zhu [12] explored the exponential utility. More recently, Zhu et al. [22] assumed that the price spread of two correlated assets follows a mean-reverting Ornstein-Uhlenbeck (OU) process and explored the dynamic MV problem, obtaining a closed-form optimal trading strategy. In a model similar to that of Zhu et al. [22], Yu et al. [19] explored the constrained optimal control problem in the MV criterion; they transformed the problem into a family of linear-quadratic optimal control problems and derived the optimal strategy for both static and dynamic problems. Zhang and Xiong [20] examined the optimal MV problem in a fast mean-reverting stochastic volatility model and provided an approximate solution. Notably, the models in [19, 20, 22] are not in the cointegration framework.

This paper discusses the MV problem for pairs trading in a continuous-time cointegration model with a common stochastic factor. It is inspired by the recent work of Chen et al. [3], who introduced a regime-switching cointegration model to capture discrete structural changes in the stock market. However, in reality, stock returns are often influenced by common stochastic factors not directly tradable, such as stock market indices. These indices serve as prevalent descriptors for various market indicators, including market risk levels and investor sentiment. Consequently, we incorporate a common stochastic factor into the continuous-time cointegration model as outlined in [4–7, 12], and assume that there are n risky assets, whose drift and volatility are influenced by this common stochastic factor. We discuss the optimal pairs trading strategies for the dynamic MV problem in terms of the optimal equilibrium strategy introduced by Björk et al. [1]. Empirical evidence, highlighted in [16, 18], confirms the widespread presence of the mean-reverting effect in Chinese stock markets. Thus, our empirical experiments presume the common stochastic factor to be the Shanghai Securities Composite Index (SSCI), characterized by a mean-reverting stochastic process. This assumption allows us to explore the interrelations between stocks across different industries and broader market context. Then, we compare the performance of the optimal strategies across different models.

The main contributions of this paper are outlined as follows. Firstly, we investigate the optimal MV problem in a continuous-time cointegration model with a common stochastic factor. We provide a semiclosed-form optimal strategy using the solutions of PDEs. Second, we derive a closed-form optimal strategy termed **the optimal flexible-ratio strategy** with a specifically defined mean-reverting common stochastic factor. Furthermore, by adopting a fixed ratio congruent with an eigenvector of the cointegrated matrix, we also derive a corresponding closed-form optimal strategy, termed **the optimal fixed-ratio strategy**. The optimal flexible-ratio strategy allows a variable fund ratio investment across different assets over time. Finally, through empirical experiments, utilizing both real data from Chinese stock markets and simulated data, we demonstrate the effectiveness of the two strategies and compare their performance with the optimal strategy provided by Chiu and Wong [7].

The remainder of this paper is organized as follows: Section 2 elaborates on the dynamic MV problem by establishing a general model for the continuous-time cointegration model with a common stochastic factor. We introduce a semiclosed-form optimal strategy for this general model. Section 3 specifies the general model into a Gaussian model and provides a closed-form optimal strategy. In Section 4, with a predetermined fixed ratio based on an eigenvector of the cointegrated matrix $\boldsymbol{A}$, we explore the optimal fixed-ratio strategy within this Gaussian model. The outcomes of the empirical experiments are presented in Section 5. Section 6 concludes the paper.

2. The optimal strategy in a general model

This section extends the continuous-time cointegration model in [4–7, 12] to account for cases where both the drift and volatility of the log-price of each selected risky asset are influenced by a common stochastic factor. *Throughout this paper, vectors or matrices are denoted by bold symbols, while nonbold symbols denote scalars. Unless noted otherwise, the subscript t signifies time in stochastic processes.*

Assume that there are n risky assets in the market, and the log-price vector of these assets at time t is expressed as $\boldsymbol{X}_t := (X_t^{(1)}, ..., X_t^{(n)})^T$. The drift and volatility of $\boldsymbol{X} = \{\boldsymbol{X}_t\}_{t\in[0,T]}$ are influenced by common stochastic factor $Y = \{Y_t\}_{t\in[0,T]}$, such as the Shanghai Securities Composite Index (SSCI). The dynamics of $\boldsymbol{X}$ satisfies the following stochastic differential equation (SDE):

$$\mathrm{d}\boldsymbol{X}_t = [\boldsymbol{\theta}(t, Y_t) - \boldsymbol{A}(t)\boldsymbol{X}_t]\,\mathrm{d}t + \boldsymbol{\sigma}(t, Y_t)\mathrm{d}\boldsymbol{W}_t, \tag{1}$$

where $\boldsymbol{W} = \{\boldsymbol{W}_t\}_{t\in[0,T]}$, $\boldsymbol{W}_t := (W_t^{(1)}, ..., W_t^{(n)})^T$, is a standard n-dimensional Brownian motion. $\boldsymbol{A}(t)$ is an $n \times n$ coefficient matrix of cointegration (called the **cointegration matrix**), and the cointegration matrices in [4–7, 12] are all constant matrices. The stochastic factor $Y = \{Y_t\}_{t\in[0,T]}$ is assumed to satisfy the following SDE:

$$\mathrm{d}Y_t = a(t, Y_t)\mathrm{d}t + b(t, Y_t)\mathrm{d}B_t, \tag{2}$$

where $B = \{B_t\}_{t\in[0,T]}$ is a {standard} Brownian motion with:

$$\mathrm{d}\langle B, \boldsymbol{W}\rangle_t = \boldsymbol{\rho}(t)\mathrm{d}t$$

and $\boldsymbol{\rho}(t) := (\rho^{(1)}(t), ..., \rho^{(n)}(t))^T$. We assume that $\left(\Omega, \mathbb{P}, \mathbb{F} := \{\mathcal{F}_t\}_{t\in[0,T]}\right)$ is a complete probability space, where $\mathbb{F}$ is the filtration generated by $\boldsymbol{W}$ and B. Furthermore, given that $t \in [0, T]$ and $y \in \mathbb{R}$, we assume:

(i) $\boldsymbol{A}(t), \boldsymbol{\theta}(t, y), \boldsymbol{\sigma}(t, y), \boldsymbol{\rho}(t), a(t, y), b(t, y)$ and $\boldsymbol{\Sigma}(t, y) := \boldsymbol{\sigma}(t, y)\boldsymbol{\sigma}(t, y)^T$ are all deterministic continuous functions, where $\boldsymbol{A}(t), \boldsymbol{\theta}(t, y), \boldsymbol{\sigma}(t, y)$ and $b(t, y)$ are bounded and $|b(t, y)| > 0$;

(ii) $a(t, y)$ and $b(t, y)$ satisfy the Lipschitz condition and a linear growth condition such that SDE (2) has a unique strong solution, denoted as $\{Y_t\}_{t\in[0,T]}$;

(iii) the inverse matrix $\boldsymbol{\Sigma}^{-1}(t, y)$ exists and is bounded;

(iv) $|\rho^{(i)}(t)| < 1$ for all $i = 1, 2, \cdots, n$.

Given $\boldsymbol{X}_0 = \boldsymbol{x}_0, Y_0 = y_0$, it follows from Theorem 5.2.1 of Øksendal [13] that the SDEs (1–2) have a unique solution with

$$E\left[\int_0^T \left\{|\boldsymbol{X}_u|^2 + |Y_u|^2\right\}\mathrm{d}u\right] < +\infty. \tag{3}$$

Let $\boldsymbol{S} = \left\{\boldsymbol{S}_t := (S_t^{(1)}, ..., S_t^{(n)})^T\right\}_{t\in[0,T]}$ be the price vector corresponding to $\boldsymbol{X}$. Under the continuous-time cointegration model (1), it follows from Itô's formula that the following SDE gives the dynamics of asset prices:

$$\mathrm{d}\boldsymbol{S}_t = \mathrm{diag}(\boldsymbol{S}_t)\left\{\left[\boldsymbol{\theta}(t,Y_t) - \boldsymbol{A}(t)X_t + \frac{1}{2}\boldsymbol{\mathcal{D}}(t,Y_t)\right]\mathrm{d}t + \boldsymbol{\sigma}(t,Y_t)\mathrm{d}\boldsymbol{W}_t\right\}, \tag{4}$$

where

$$\boldsymbol{\mathcal{D}}(t,Y_t) := \left(\sum_{j=1}^{n}\sigma_{1j}^2(t,Y_t),...,\sum_{j=1}^{n}\sigma_{nj}^2(t,Y_t)\right)^T.$$

In addition to the risky assets, we assume that there is a risk-free asset in the market whose price $\Pi(t)$ is given by

$$\mathrm{d}\Pi(t) = r\Pi(t)\mathrm{d}t,$$

where r denotes the constant risk-free rate.

We now discuss the self-financing strategy. Consider an investor with an initial wealth $V_0 > 0$. At any time t prior to T, the investor decides on the allocation of investment across n risky assets $S_t^{(1)},...,S_t^{(n)}$ and the risk-free asset $\Pi(t)$. Denoting $\boldsymbol{h}(t) := (h^{(1)}(t),...,h^{(n)}(t))^T$ as the weights invested in $S_t^{(1)},...,S_t^{(n)}$ at time t, the wealth process $V_t^{\boldsymbol{h}}$ is given by:

$$\begin{aligned}\mathrm{d}V_t^{\boldsymbol{h}} &= \sum_{i=1}^{n} V_t^h h^{(i)}(t)\frac{\mathrm{d}S_t^{(i)}}{S_t^{(i)}} + V_t^{\boldsymbol{h}}\left(1-\sum_{i=1}^{n}h^{(i)}(t)\right)\frac{\mathrm{d}\Pi(t)}{\Pi(t)}\\ &= V_t^{\boldsymbol{h}}\left[\boldsymbol{h}(t)^T\boldsymbol{\alpha}(t,\boldsymbol{X}_t,Y_t)\mathrm{d}t + \boldsymbol{h}(t)^T\boldsymbol{\sigma}(t,Y_t)\mathrm{d}\boldsymbol{W}_t + r\mathrm{d}t\right],\end{aligned} \tag{5}$$

with

$$\boldsymbol{\alpha}(t,\boldsymbol{X}_t,Y_t) := \boldsymbol{\theta}(t,Y_t) - \boldsymbol{A}(t)\boldsymbol{X}_t + \frac{1}{2}\boldsymbol{\mathcal{D}}(t,Y_t) - r(1,...,1)^T. \tag{6}$$

Let $\boldsymbol{\pi}(t) := e^{-rt}V_t^{\boldsymbol{h}}\boldsymbol{h}(t)$ represent the discounted money invested in $S^{(1)},...,S^{(n)}$, which is considered the strategy. Similar to Zhou and Li [21], we define the admissible strategy set $\mathcal{A}$ as:

$$\mathcal{A} := \left\{\boldsymbol{\pi} = \{\boldsymbol{\pi}(u,\boldsymbol{X_u},Y_u)\}_{u\in[0,T]}\middle|\begin{array}{c}\boldsymbol{\pi}(u,\boldsymbol{x},y) \text{ is a function such that}\\ \mathbf{E}\left[\int_0^T|\boldsymbol{\pi}(u,\boldsymbol{X_u},Y_u)|^2\mathrm{d}u\right] < +\infty.\end{array}\right\}$$

The discounted wealth process $\overline{V}_t(\boldsymbol{\pi}) := e^{-rt}V_t^{\boldsymbol{h}}$ is then given by

$$\mathrm{d}\overline{V}_t(\boldsymbol{\pi}) = \boldsymbol{\pi}_t^T\boldsymbol{\alpha}(t,\boldsymbol{X}_t,Y_t)\mathrm{d}t + \boldsymbol{\pi}_t^T\boldsymbol{\sigma}(t,Y_t)\mathrm{d}\boldsymbol{W}_t; \tag{7}$$

thus,

$$\overline{V}_T(\boldsymbol{\pi}) = \overline{V}_t(\boldsymbol{\pi}) + \int_t^T\boldsymbol{\pi}_u^T\boldsymbol{\alpha}(u,\boldsymbol{X}_u,Y_u)\mathrm{d}u + \int_t^T\boldsymbol{\pi}_u^T\boldsymbol{\sigma}(u,Y_u)\mathrm{d}\boldsymbol{W}_u. \tag{8}$$

For any $\boldsymbol{\pi}\in\mathcal{A}$, one can see that $\overline{V}_t = \{\overline{V}_t(\boldsymbol{\pi})\}_{t\in[0,T]}$ is an $\mathbb{F}$-adapted process with $\boldsymbol{E}\left[\left|\overline{V}_t(\boldsymbol{\pi})\right|\right] < +\infty$. For simplicity, we denote

$$\overline{V}_{t,T}(\boldsymbol{\pi}) := \overline{V}_T(\boldsymbol{\pi}) - \overline{V}_t(\boldsymbol{\pi}).$$

Hereinafter, $\mathbf{E}[\cdot|\mathcal{F}_t]$ is denoted as $\mathbf{E}_t[\cdot]$, while $\mathbf{Var}[\cdot|\mathcal{F}_t]$ is denoted as $\mathbf{Var}_t[\cdot]$. Given a strategy $\boldsymbol{\pi}\in\mathcal{A}$ with $\overline{V}_t(\boldsymbol{\pi}) = \overline{V}_t$, let

$$
\begin{aligned}
J(t, \boldsymbol{X}_t, Y_t, \overline{V}_t; \boldsymbol{\pi}) &:= \mathbf{E}_t\left[\overline{V}_T(\boldsymbol{\pi})\right] - \lambda \mathbf{Var}_t\left[\overline{V}_T(\boldsymbol{\pi})\right] \\
&= \overline{V}_t + \mathbf{E}_t\left[\overline{V}_{t,T}(\boldsymbol{\pi})\right] - \lambda \mathbf{Var}_t\left[\overline{V}_{t,T}(\boldsymbol{\pi})\right]. \qquad (9)
\end{aligned}
$$

We consider the following dynamic mean–variance problem

$$
\max_{\pi \in \mathcal{A}} J(t, \boldsymbol{X}_t, Y_t, \overline{V}_t; \boldsymbol{\pi}). \qquad (10)
$$

Given that problem (10) is not time-consistent, we adopt the optimal equilibrium strategy, introduced by Björk et al.[1].

Definition 2.1 *For given strategy $\boldsymbol{\pi}^* \in \mathcal{A}$, choose an arbitrary admissible strategy $\hat{\boldsymbol{\pi}}$ and $\varepsilon > 0$. We define a* ***perturbation strategy*** *of $\boldsymbol{\pi}^*$:*

$$
\boldsymbol{\pi}^{\hat{\boldsymbol{\pi}},\varepsilon}(s, \boldsymbol{x}, y) = \begin{cases} \hat{\boldsymbol{\pi}}(s, \boldsymbol{x}, y), & t \leqslant s < t + \varepsilon, \\ \boldsymbol{\pi}^*(s, \boldsymbol{x}, y), & t + \varepsilon \leqslant s \leqslant T. \end{cases} \qquad (11)
$$

If the following limit holds for all $\hat{\boldsymbol{\pi}} \in \mathcal{A}$ and $t \in [0, T]$:

$$
\limsup_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \left\{ J\left(t, \boldsymbol{X}_t, Y_t, \overline{V}_t(\boldsymbol{\pi}^*); \boldsymbol{\pi}^{\hat{\pi},\varepsilon}\right) - J\left(t, \boldsymbol{X}_t, Y_t, \overline{V}_t(\boldsymbol{\pi}^*); \boldsymbol{\pi}^*\right) \right\} \leqslant 0,
$$

the strategy $\boldsymbol{\pi}^*$ is called an ***optimal equilibrium strategy****.*

Remark 2.2 *In subsequent discussions, any reference to "the optimal strategy" implies "the optimal equilibrium strategy" as defined in Definition 2.1.*

Definition 2.3 *For any domain $\mathcal{D} \subset [0, T] \times \mathbb{R}^n \times \mathbb{R}$, we define the following function classes:*

(i) *$C^{1,2,2}(\mathcal{D})$ denotes the class of real-valued functions $f(t, \boldsymbol{x}, y) : \mathcal{D} \mapsto \mathbb{R}$ that are continuously differentiable with respect to t and twice continuously differentiable with respect to both $\boldsymbol{x}$ and y;*

(ii) $\mathcal{H}(\mathcal{D}) = \{f : \exists B, \beta > 0 s.t. |f(t, \boldsymbol{x}, y)| \leqslant Be^{\beta(|\boldsymbol{x}|^2 + y^2)}\}$;

(iii) $C^{1,2,2}_{\mathcal{H}}(\mathcal{D}) = C^{1,2,2}(\mathcal{D}) \cap \mathcal{H}(\mathcal{D})$.

Remark 2.4 *Definition 2.3 (ii) represents a classical growth condition in PDE theory. According to Fleming [9], this growth condition is crucial for ensuring the uniqueness of the classical solution to the Cauchy problem associated with parabolic equations.*

For $f \in \mathcal{C}^{1,2,2}_{\mathcal{H}}([0, T] \times \mathbb{R}^n \times \mathbb{R})$, we introduce the following operator:

$$
\begin{aligned}
\mathcal{L}f(t, \boldsymbol{x}, y) = {} & \partial_t f(t, \boldsymbol{x}, y) + \partial_{\boldsymbol{x}}^T f(t, \boldsymbol{x}, y) \left[\boldsymbol{\theta}(t, y) - \boldsymbol{A}(t)\boldsymbol{x}\right] + \partial_y f(t, \boldsymbol{x}, y) a(t, y) \\
& + \frac{1}{2} tr\left(\partial_{\boldsymbol{x}\boldsymbol{x}} f(t, \boldsymbol{x}, y) \boldsymbol{\Sigma}(t, y)\right\} + \frac{1}{2} \partial_{yy} f(t, \boldsymbol{x}, y) b(t, y)^2 \\
& + b(t, y) \partial_{y\boldsymbol{x}}^T f(t, \boldsymbol{x}, y) \boldsymbol{\sigma}(t, y) \boldsymbol{\rho}(t),
\end{aligned}
$$

where $tr(\cdot)$ denotes the trace of a matrix, $\partial_x f := \left(\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}\right\}^T$, $\partial_{yx} f := \left(\frac{\partial^2 f}{\partial y \partial x_1}, ..., \frac{\partial^2 f}{\partial y \partial x_n}\right\}^T$, and $\partial_{\boldsymbol{x}\boldsymbol{x}} f := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1,\cdots,n}$. Thus, we have the following theorem:

Theorem 2.5 (Main result I) *Assume that functions $f(t, \boldsymbol{x}, y), g(t, \boldsymbol{x}, y) \in \mathcal{C}^{1,2,2}_{\mathcal{H}}([0, T] \times \mathbb{R}^{n \times n} \times \mathbb{R})$ are solutions to the following PDEs:*

$$
\begin{aligned}
& \mathcal{L}f(t, \boldsymbol{x}, y) + \frac{1}{2\lambda} \boldsymbol{\alpha}^T(t, \boldsymbol{x}, y) \boldsymbol{\Sigma}(t, y)^{-1} \boldsymbol{\alpha}(u, \boldsymbol{x}, y) - \boldsymbol{\alpha}^T(t, \boldsymbol{x}, y) \partial_{\boldsymbol{x}} f(t, \boldsymbol{x}, y) \\
& - \partial_y f(t, \boldsymbol{x}, y) b(t, y) \boldsymbol{\alpha}^T(t, \boldsymbol{x}, y) \boldsymbol{\sigma}^T(t, y)^{-1} \boldsymbol{\rho}(t) = 0, \qquad (12)
\end{aligned}
$$

$$\begin{aligned}&\mathcal{L}g(t,\boldsymbol{x},y)+\frac{1}{4\lambda}\boldsymbol{\alpha}^T(t,\boldsymbol{x},y)\boldsymbol{\Sigma}(t,y)^{-1}\boldsymbol{\alpha}(t,\boldsymbol{x},y)\\&-\boldsymbol{\alpha}^T(t,\boldsymbol{x},y)\partial_{\boldsymbol{x}}f(t,\boldsymbol{x},y)-\partial_y f(t,\boldsymbol{x},y)b(t,y)\boldsymbol{\alpha}^T(t,\boldsymbol{x},y)\boldsymbol{\sigma}^T(t,y)^{-1}\boldsymbol{\rho}(t)\\&+\lambda\left(\partial_y f(t,\boldsymbol{x},y)\right)^2 b(t,y)^2\left(\boldsymbol{\rho}(t)^T\boldsymbol{\rho}(t)-1\right)=0,\end{aligned}\tag{13}$$

with the terminal conditions $f(T,\boldsymbol{x},y)=0$ and $g(T,\boldsymbol{x},y)=0$, respectively. Assume

$$\boldsymbol{E}\left[\int_0^T\left\{|\partial_{\boldsymbol{x}}f(u,\boldsymbol{X}_u,Y_u)|^2+|\partial_y f(u,\boldsymbol{X}_u,Y_u)|^2\right\}\mathrm{d}u\right]<+\infty,$$

let

$$\begin{aligned}\boldsymbol{\pi}^*(u,\boldsymbol{X}_u,Y_u):=\frac{1}{2\lambda}\boldsymbol{\Sigma}(u,Y_u)^{-1}\boldsymbol{\alpha}(u,\boldsymbol{X}_u,Y_u)-\partial_{\boldsymbol{x}}f(u,\boldsymbol{X}_u,Y_u)\\-\partial_y f(u,\boldsymbol{X}_u,Y_u)b(u,Y_u)\boldsymbol{\sigma}^T(u,Y_u)^{-1}\boldsymbol{\rho}(u).\end{aligned}\tag{14}$$

Then, $\boldsymbol{\pi}^*=\{\boldsymbol{\pi}^*(u,\boldsymbol{X}_{\boldsymbol{u}},Y_u)\}_{u\in[0,T]}\in\mathcal{A}$ is an optimal strategy for problem (10), fulfilling:

$$\mathbf{E}_t\left[\overline{V}_{t,T}(\boldsymbol{\pi}^*)\right]=f(t,\boldsymbol{X}_t,Y_t),\tag{15}$$

$$\mathbf{E}_t\left[\overline{V}_{t,T}(\boldsymbol{\pi}^*)\right]-\lambda\mathbf{Var}_t[\overline{V}_{t,T}(\boldsymbol{\pi}^*)\big]=g(t,\boldsymbol{X}_t,Y_t).\tag{16}$$

Remark 2.6 *The optimal strategy defined in (14) depends on $f(t,\boldsymbol{x},y)$ rather than $g(t,\boldsymbol{x},y)$. However, $g(t,\boldsymbol{x},y)$ describes the optimal MV utility with $\boldsymbol{X}_t=\boldsymbol{x}$ and $Y_t=y$. PDE (13) depends on the solution of PDE (12).*

Proof Based on Fleming [9] (Theorem 6.4.1 and Corollary 6.4.2), the classical solutions of PDEs (12) and (13) are unique. Let $\boldsymbol{\pi}^*$ be the strategy as defined in (14). The Markov property ensures the existence of deterministic functions $f^*(t,\boldsymbol{x},y)$ and $g^*(t,\boldsymbol{x},y)$ such that

$$\begin{aligned}f^*(t,\boldsymbol{X}_t,Y_t)&:=\mathbf{E}_t\left[\overline{V}_{t,T}(\boldsymbol{\pi}^*)\right],\\g^*(t,\boldsymbol{X}_t,Y_t)&:=\mathbf{E}_t\left[\overline{V}_{t,T}(\boldsymbol{\pi}^*)\right]-\lambda\mathbf{Var}_t[\overline{V}_{t,T}(\boldsymbol{\pi}^*)\big].\end{aligned}\tag{17}$$

We note that $f^*(t,\boldsymbol{X}_t,Y_t)+\overline{V}_t(\boldsymbol{\pi}^*)$ constitutes a square-integrable martingale, leading to:

$$\mathcal{L}f^*(u,\boldsymbol{X}_u,Y_u)+\boldsymbol{\pi}^*(u,\boldsymbol{X}_u,Y_u)^T\boldsymbol{\alpha}(u,\boldsymbol{X}_u,Y_u)=0.$$

Substituting $\boldsymbol{\pi}^*$ from (14) into the above equation, it is evident that $f^*(t,\boldsymbol{x},y)$ satisfies PDE (12) with the terminal condition $f^*(T,\boldsymbol{x},y)=0$. By the uniqueness of the classical solution of PDE (12), we conclude $f^*(t,\boldsymbol{x},y)=f(t,\boldsymbol{x},y)$. Moreover,

$$\begin{aligned}\mathrm{d}\left\{\overline{V}_t(\boldsymbol{\pi}^*)+f\left(t,\boldsymbol{X}_t,Y_t\right)\right\}=\Big\{\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)+\boldsymbol{\pi}^*(t,\boldsymbol{X}_t,Y_t)\Big\}^T\boldsymbol{\sigma}(t,Y_t)\mathrm{d}\boldsymbol{W}_t\\+\partial_y f(t,\boldsymbol{X}_t,Y_t)b(t,Y_t)\mathrm{d}B_t.\end{aligned}\tag{18}$$

To finish the proof, we must establish that the strategy defined in (14) is optimal and that its utility matches (16).

For any given $\varepsilon>0$ and any admissible strategy $\hat{\boldsymbol{\pi}}$, we construct a perturbation strategy $\boldsymbol{\pi}^{\hat{\pi},\varepsilon}$ for $\boldsymbol{\pi}^*$. For ease of subsequent discussions, we denote $\boldsymbol{\pi}_t^*:=\boldsymbol{\pi}^*(t,\boldsymbol{X}_t,Y_t)$ and $\hat{\boldsymbol{\pi}}_t:=\hat{\boldsymbol{\pi}}(t,\boldsymbol{X}_t,Y_t)$. Then, we have:

$$\overline{V}_T(\boldsymbol{\pi}^{\hat{\pi},\varepsilon})=\overline{V}_t(\boldsymbol{\pi}^*)+\overline{V}_{t,t+\varepsilon}(\hat{\boldsymbol{\pi}})+\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*).$$

It follows from (17) that:

$$
\begin{aligned}
&\mathbf{E}_t\left[\overline{V}_{t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + g^*(t+\varepsilon, \boldsymbol{X}_{t+\varepsilon}, Y_{t+\varepsilon})\right] \\
&= \mathbf{E}_t\left\{\overline{V}_{t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + \mathbf{E}_{t+\varepsilon}\left[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)\right] - \lambda \mathbf{Var}_{t+\varepsilon}\left[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)\right]\right\} \\
&= \mathbf{E}_t\left\{\mathbf{E}_{t+\varepsilon}[\overline{V}_T(\boldsymbol{\pi}^{\hat{\pi},\varepsilon})] - \lambda \mathbf{Var}_{t+\varepsilon}\left[\overline{V}_T(\boldsymbol{\pi}^{\hat{\pi},\varepsilon})\right]\right\} \\
&= \mathbf{E}_t\left[\overline{V}_T(\boldsymbol{\pi}^{\hat{\pi},\varepsilon})\right] - \lambda \mathbf{Var}_t\left[\overline{V}_T(\boldsymbol{\pi}^{\hat{\pi},\varepsilon})\right] + \lambda \mathbf{Var}_t\left[E_{t+\varepsilon}\left[\overline{V}_T(\boldsymbol{\pi}^{\hat{\pi},\varepsilon})\right]\right] \\
&= J\left(t, \boldsymbol{X}_t, Y_t, \overline{V}_t(\boldsymbol{\pi}^*); \boldsymbol{\pi}^{\hat{\pi},\varepsilon}\right) + \lambda \mathbf{Var}_t\left[\overline{V}_{t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + f\left(t+\varepsilon, \boldsymbol{X}_{t+\varepsilon}, Y_{t+\varepsilon}\right\}\right],
\end{aligned}
$$

which implies

$$
\begin{aligned}
J\left(t, \boldsymbol{X}_t, Y_t, \overline{V}_t(\boldsymbol{\pi}^*); \boldsymbol{\pi}^{\hat{\pi},\varepsilon}\right) = {}& \mathbf{E}_t\left\{\overline{V}_{t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + \mathbf{E}_{t+\varepsilon}\left[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)\right] - \lambda \mathbf{Var}_{t+\varepsilon}\left[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)\right]\right\} \\
& - \lambda \mathbf{Var}_t\left[\overline{V}_{t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + f\left(t+\varepsilon, \boldsymbol{X}_{t+\varepsilon}, Y_{t+\varepsilon}\right\}\right]. \qquad (19)
\end{aligned}
$$

Therefore, for convenience, we denote $\Delta_\varepsilon f(t, \boldsymbol{X}_t, Y_t) := f(t+\varepsilon, \boldsymbol{X}_{t+\varepsilon}, Y_{t+\varepsilon}) - f(t, \boldsymbol{X}_t, Y_t)$. Hence,

$$
\begin{aligned}
& J(t, \boldsymbol{X}_t, Y_t, \overline{V}_t(\boldsymbol{\pi}^*); \boldsymbol{\pi}^{\hat{\pi},\varepsilon}) - J(t, \boldsymbol{X}_t, Y_t, \overline{V}_t(\boldsymbol{\pi}^*); \boldsymbol{\pi}^*) \\
&= \mathbf{E}_t\left\{\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + \mathbf{E}_{t+\varepsilon}[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)] - \mathbf{E}_t[\overline{V}_{t,T}(\boldsymbol{\pi}^*)]\right. \\
&\quad \left. - \lambda\left[\mathbf{Var}_{t+\varepsilon}[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)] - \mathbf{Var}_t[\overline{V}_{t,T}(\boldsymbol{\pi}^*)]\right]\right\} - \lambda \mathbf{Var}_t\left[\overline{V}_{t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + f\left(t+\varepsilon, \boldsymbol{X}_{t+\varepsilon}, Y_{t+\varepsilon}\right\}\right] \\
&= \mathbf{E}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + \Delta_\varepsilon f(t, \boldsymbol{X}_t, Y_t)\right] \\
&\quad - \lambda \mathbf{E}_t\left[\mathbf{Var}_{t+\varepsilon}[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)] - \mathbf{Var}_t[\overline{V}_{t,T}(\boldsymbol{\pi}^*)]\right] - \lambda \mathbf{Var}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + \Delta_\epsilon f\left(t, \boldsymbol{X}_t, Y_t\right\}\right].
\end{aligned}
$$

Since

$$
\begin{aligned}
& \mathbf{Var}_t[\overline{V}_{t,T}(\boldsymbol{\pi}^*)] = \mathbf{Var}_t[\overline{V}_T(\boldsymbol{\pi}^*)] \\
&= \mathbf{E}_t\left[\mathbf{Var}_{t+\varepsilon}[\overline{V}_T(\boldsymbol{\pi}^*)]\right] + \mathbf{Var}_t\left[\mathbf{E}_{t+\varepsilon}[\overline{V}_T(\boldsymbol{\pi}^*)]\right] \\
&= \mathbf{E}_t\left[\mathbf{Var}_{t+\varepsilon}[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)]\right] + \mathbf{Var}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^*) + \mathbf{E}_{t+\varepsilon}[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)] - \mathbf{E}_t[\overline{V}_{t,T}(\boldsymbol{\pi}^*)]\right] \\
&= \mathbf{E}_t\left[\mathbf{Var}_{t+\varepsilon}[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)]\right] + \mathbf{Var}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^*) + \Delta_\varepsilon f(t, \boldsymbol{X}_t, Y_t)\right],
\end{aligned}
$$

we have

$$
\mathbf{E}_t\left[\mathbf{Var}_{t+\varepsilon}[\overline{V}_{t+\varepsilon,T}(\boldsymbol{\pi}^*)] - \mathbf{Var}_t[\overline{V}_{t,T}(\boldsymbol{\pi}^*)]\right] = -\mathbf{Var}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^*) + \Delta_\varepsilon f(t, \boldsymbol{X}_t, Y_t)\right].
$$

We then obtain

$$
\begin{aligned}
& J(t, \boldsymbol{X}_t, Y_t, \overline{V}_t(\boldsymbol{\pi}^*); \boldsymbol{\pi}^{\hat{\pi},\varepsilon}) - J(t, \boldsymbol{X}_t, Y_t, \overline{V}_t(\boldsymbol{\pi}^*); \boldsymbol{\pi}^*) \\
&= \mathbf{E}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + \Delta_\varepsilon f(t, \boldsymbol{X}_t, Y_t)\right] \\
&\quad + \lambda \mathbf{Var}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^*) + \Delta_\varepsilon f(t, \boldsymbol{X}_t, Y_t)\right] - \lambda \mathbf{Var}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + \Delta_\varepsilon f\left(t, \boldsymbol{X}_t, Y_t\right\}\right].
\end{aligned}
$$

Since

$$
\begin{aligned}
& \overline{V}_{t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon}) + f\left(t+\varepsilon, \boldsymbol{X}_{t+\varepsilon}, Y_{t+\varepsilon}\right) \\
&= \overline{V}_t(\boldsymbol{\pi}^*) + f\left(t, \boldsymbol{X}_t, Y_t\right) + \int_t^{t+\varepsilon}\left\{\mathcal{L}f(u, \boldsymbol{X}_u, Y_u) + \hat{\boldsymbol{\pi}}_u^T \boldsymbol{\alpha}(u, \boldsymbol{X}_u, Y_u)\right\}\mathrm{d}u \\
&\quad + \int_t^{t+\varepsilon}\left\{\partial_{\boldsymbol{x}} f(u, \boldsymbol{X}_u, Y_u) + \hat{\boldsymbol{\pi}}_u\right\}^T \boldsymbol{\sigma}(u, Y_u)\mathrm{d}\boldsymbol{W}_u + \int_t^{t+\varepsilon} \partial_y f(u, \boldsymbol{X}_u, Y_u) b(t, Y_t)\mathrm{d}B_u,
\end{aligned}
$$

it can be observed that

$$\begin{aligned}
&\mathbf{E}_t\left[\left\{\overline{V}_{t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon})+f\left(t+\varepsilon,\boldsymbol{X}_{t+\varepsilon},Y_{t+\varepsilon}\right)\right\}^2\right]\\
=&\left\{\overline{V}_t(\boldsymbol{\pi}^*)+f\left(t,\boldsymbol{X}_t,Y_t\right)\right\}^2\\
&+\mathbf{E}_t\left[2\left\{\overline{V}_t(\boldsymbol{\pi}^*)+f\left(t,\boldsymbol{X}_t,Y_t\right)\right\}\int_t^{t+\varepsilon}\left(\hat{\boldsymbol{\pi}}_u-\boldsymbol{\pi}_u^*\right)^T\boldsymbol{\alpha}(u,\boldsymbol{X}_u,Y_u)\mathrm{d}u\right]\\
&+\mathbf{E}_t\left[\int_t^{t+\varepsilon}\left\{\left[\partial_{\boldsymbol{x}}f(u,\boldsymbol{X}_u,Y_u)+\hat{\boldsymbol{\pi}}_u\right]^T\boldsymbol{\Sigma}(u,Y_u)\left[\partial_{\boldsymbol{x}}f(u,\boldsymbol{X}_u,Y_u)+\hat{\boldsymbol{\pi}}_u\right]\right.\right.\\
&+\left(\partial_y f(u,\boldsymbol{X}_u,Y_u)\right)^2 b(t,Y_t)^2\\
&\left.\left.+\ 2\partial_y f(u,\boldsymbol{X}_u,Y_u)b(u,Y_u)\left[\partial_{\boldsymbol{x}}f(u,\boldsymbol{X}_u,Y_u)+\hat{\boldsymbol{\pi}}_u\right]^T\boldsymbol{\sigma}(u,Y_u)\boldsymbol{\rho}(u)\right\}\mathrm{d}u\right].
\end{aligned}$$

Hence,

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathbf{Var}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon})+\Delta_\varepsilon f\left(t,\boldsymbol{X}_t,Y_t\right\}\right]\\
&=\left[\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)+\hat{\boldsymbol{\pi}}_t\right]^T\boldsymbol{\Sigma}(t,Y_t)\left[\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)+\hat{\boldsymbol{\pi}}_t\right]\\
&\quad+\left(\partial_y f(t,\boldsymbol{X}_t,Y_t)\right)^2 b(t,Y_t)^2+2\partial_y f(t,\boldsymbol{X}_t,Y_t)b(t,Y_t)\left[\partial_{\boldsymbol{x}}f(u,\boldsymbol{X}_t,Y_t+\hat{\boldsymbol{\pi}}_t\right]^T\boldsymbol{\sigma}(t,Y_t)\boldsymbol{\rho}(t).
\end{aligned}\tag{20}$$

Similarly, one can see from (18) that

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\mathbf{Var}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^*)+\Delta_\varepsilon f\left(t,\boldsymbol{X}_t,Y_t\right\}\right]\\
&=\left[\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)+\boldsymbol{\pi}_t^*\right]^T\boldsymbol{\Sigma}(t,Y_t)\left[\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)+\boldsymbol{\pi}_t^*\right]\\
&\quad+\left(\partial_y f(t,\boldsymbol{X}_t,Y_t)\right)^2 b(t,Y_t)^2+2\partial_y f(t,\boldsymbol{X}_t,Y_t)b(t,Y_t)\left[\partial_{\boldsymbol{x}}f(u,\boldsymbol{X}_t,Y_t+\boldsymbol{\pi}_t^*\right]^T\boldsymbol{\sigma}(t,Y_t)\boldsymbol{\rho}(t).
\end{aligned}\tag{21}$$

From (20) and (21), we obtain that

$$\begin{aligned}
&\lim_{\varepsilon\to 0}\frac{J(t,\boldsymbol{X}_t,Y_t,\overline{V}_t(\boldsymbol{\pi}^*);\boldsymbol{\pi}^{\hat{\pi},\varepsilon})-J(t,\boldsymbol{X}_t,Y_t,\overline{V}_t(\boldsymbol{\pi}^*);\boldsymbol{\pi}^*)}{\varepsilon}\\
&=\hat{\boldsymbol{\pi}}_t^T\boldsymbol{\alpha}(t,\boldsymbol{X}_t,Y_t)+\mathcal{L}f(t,\boldsymbol{X}_t,Y_t)\\
&\quad+\lambda\left(\boldsymbol{\pi}_t^*+\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)\right)^T\boldsymbol{\Sigma}(t,Y_t)\left(\boldsymbol{\pi}_t^*+\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)\right)\\
&\quad-\lambda\left(\hat{\boldsymbol{\pi}}_t+\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)\right)^T\boldsymbol{\Sigma}(t,Y_t)\left(\hat{\boldsymbol{\pi}}_t+\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)\right)\\
&\quad+2\lambda\partial_y f(t,\boldsymbol{X}_t,Y_t)b(t,Y_t)(\boldsymbol{\pi}_t^*-\hat{\boldsymbol{\pi}}_t)^T\boldsymbol{\rho}(t)\\
&:=\Psi(t,\hat{\boldsymbol{\pi}}_t).
\end{aligned}$$

It follows from (12) and (14) that

$$\left(\boldsymbol{\pi}_t^*\right)^T\boldsymbol{\alpha}(t,\boldsymbol{X}_t,Y_t)+\mathcal{L}f(t,\boldsymbol{X}_t,Y_t)=0,$$

which implies

$$\begin{aligned}
\Psi(t,\hat{\boldsymbol{\pi}}_t)=&-\lambda(\hat{\boldsymbol{\pi}}_t-\boldsymbol{\pi}_t^*)^T\boldsymbol{\Sigma}(t,Y_t)(\hat{\boldsymbol{\pi}}_t-\boldsymbol{\pi}_t^*)\\
&-2\lambda(\hat{\boldsymbol{\pi}}_t-\boldsymbol{\pi}_t^*)^T\left[-\frac{1}{2\lambda}\boldsymbol{\alpha}(t,\boldsymbol{X}_t,Y_t)\right.\\
&\left.+\boldsymbol{\Sigma}(t,Y_t)(\boldsymbol{\pi}_t^*+\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t))+\partial_y f(t,\boldsymbol{X}_t,Y_t)b(t,Y_t)\boldsymbol{\rho}(t)\right]\\
=&-\lambda(\hat{\boldsymbol{\pi}}_t-\boldsymbol{\pi}_t^*)^T\boldsymbol{\Sigma}(t,Y_t)(\hat{\boldsymbol{\pi}}_t-\boldsymbol{\pi}_t^*).
\end{aligned}$$

One can see that $\Psi(t,\hat{\boldsymbol{\pi}}_t)$ is maximized at $\boldsymbol{\pi}_t^*$ with $\Psi(t,\boldsymbol{\pi}_t^*)=0$. As a result, $\boldsymbol{\pi}_t^*$ defined in (14) is an optimal strategy defined by Definition 2.1.

We now prove (16), i.e., $g^*(t,\boldsymbol{x},y)=g(t,\boldsymbol{x},y)$. In fact, one can see from (19) and (20) that

$$\begin{aligned}&J(t,\boldsymbol{X}_t,Y_t,\overline{V}_t(\boldsymbol{\pi}^*);\boldsymbol{\pi}^{\hat{\pi},\varepsilon})-J(t,\boldsymbol{X}_t,Y_t,\overline{V}_t(\boldsymbol{\pi}^*);\boldsymbol{\pi}^*)\\ &=\mathbf{E}_t\left[\overline{V}_{t,t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon})+g^*(t+\varepsilon,\boldsymbol{X}_{t+\varepsilon},Y_{t+\varepsilon})-g^*(t,\boldsymbol{X}_t,Y_t)\right]\\ &\quad-\lambda\mathbf{Var}_t\left[\overline{V}_{t+\varepsilon}(\boldsymbol{\pi}^{\hat{\pi},\varepsilon})+f\left(t+\varepsilon,\boldsymbol{X}_{t+\varepsilon},Y_{t+\varepsilon}\right\}\right],\end{aligned}$$

thus $\Psi(t,\hat{\boldsymbol{\pi}}_t)$ can also be represented by the following:

$$\begin{aligned}\Psi(t,\hat{\boldsymbol{\pi}}_t)&=\lim_{\varepsilon\to0}\frac{J(t,\boldsymbol{X}_t,Y_t,\overline{V}_t(\boldsymbol{\pi}^*);\boldsymbol{\pi}^{\hat{\pi},\varepsilon})-J(t,\boldsymbol{X}_t,Y_t,\overline{V}_t(\boldsymbol{\pi}^*);\boldsymbol{\pi}^*)}{\varepsilon}\\ &=\hat{\boldsymbol{\pi}}_t^T\boldsymbol{\alpha}(t,\boldsymbol{X}_t,Y_t)+\mathcal{L}g^*(t,\boldsymbol{X}_t,Y_t)\\ &\quad-\lambda\left[\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)+\hat{\boldsymbol{\pi}}_t\right]^T\boldsymbol{\Sigma}(t,Y_t)\left[\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)+\hat{\boldsymbol{\pi}}_t\right]\\ &\quad-\lambda\left(\partial_yf(t,\boldsymbol{X}_t,Y_t)\right\}^2b(t,Y_t)^2-2\lambda\partial_yf(t,\boldsymbol{X}_t,Y_t)b(t,Y_t)[\partial_{\boldsymbol{x}}f(t,\boldsymbol{X}_t,Y_t)+\hat{\boldsymbol{\pi}}_t]^T\boldsymbol{\sigma}(t,Y_t)\boldsymbol{\rho}(t).\end{aligned}$$

Using $\Psi(t,\boldsymbol{\pi}_t^*)=0$ and substituting $\boldsymbol{\pi}_t^*$ in (14) into the above equation, we can see that $g^*(t,\boldsymbol{x},y)$ is the solution of PDE (13) with the terminal condition $g^*(T,\boldsymbol{x},y)=0$. It follows from the uniqueness of the classical solution of PDE (13) that $g^*(t,\boldsymbol{x},y)=g(t,\boldsymbol{x},y)$. □

3. The optimal strategy for the Gaussian model

Generally, solving the PDE (12) explicitly is challenging. Hence, it is difficult to apply Theorem 2.5 in practice. In this section, we specify the model outlined in (1) and (2) into a Gaussian model and assume that the dynamics of the log-price vector of the risky assets is given by

$$\mathrm{d}\boldsymbol{X}_t=\left[\boldsymbol{c}(t)+\boldsymbol{d}(t)Y_t-\boldsymbol{A}(t)\boldsymbol{X}_t\right]\mathrm{d}t+\boldsymbol{\sigma}(t)\mathrm{d}\boldsymbol{W}_t,\tag{22}$$

and the dynamics of the common stochastic factor is given by

$$\mathrm{d}Y_t=\kappa(t)(\theta(t)-Y_t)\mathrm{d}t+b(t)\mathrm{d}B_t,\tag{23}$$

where $\kappa(t),\theta(t),b(t),\boldsymbol{c}(t),\boldsymbol{d}(t),\boldsymbol{\sigma}(t),\boldsymbol{\rho}(t),\boldsymbol{A}(t)$ and $\boldsymbol{\Sigma}(t):=\boldsymbol{\sigma}(t)\boldsymbol{\sigma}(t)^T$ are all deterministic continuous-bounded functions with respect to $t\in[0,T]$ and $|b(t)|>0$. The inverse matrix $\boldsymbol{\Sigma}^{-1}(t)$ exists and is bounded. Our objective is to derive a closed-form optimal strategy of the dynamic mean–variance (MV) problem (10) in this Gaussian model.

In this case, we introduce:

$$\boldsymbol{\mathcal{D}}(t):=\left(\sum_{j=1}^n\sigma_{1j}^2(t),...,\sum_{j=1}^n\sigma_{nj}^2(t)\right)^T,$$

where it can be seen that $\boldsymbol{\alpha}(t,\boldsymbol{X}_t,Y_t)$ in (6) can be simplified into the following form

$$\boldsymbol{\alpha}(t,\boldsymbol{X}_t,Y_t)=\overline{\boldsymbol{c}}(t)+\boldsymbol{d}(t)Y_t-\boldsymbol{A}(t)\boldsymbol{X}_t,$$

where

$$\overline{\boldsymbol{c}}(t)=\boldsymbol{c}(t)+\frac{1}{2}\boldsymbol{\mathcal{D}}(t)-r(1,...,1)^T.$$

Furthermore, to find the explicit form of the optimal strategy in this Gaussian model, we must introduce

$$\boldsymbol{p}(t) := \boldsymbol{c}(t) - \overline{\boldsymbol{c}}(t), \quad \boldsymbol{q}(t) := b(t)\left(\boldsymbol{\sigma(t)}^T\right)^{-1}\boldsymbol{\rho}(t), \quad \overline{q}(t) := -\kappa(t) - \boldsymbol{q}(t)^T\boldsymbol{d}(t).$$

Theorem 3.1 (Main result II) *Let* $Q(t) := \int_0^t \overline{q}(s)\mathrm{d}s$, *and introduce*

$$\begin{cases} C_1(t;\lambda) := e^{-2Q(t)}\displaystyle\int_t^T e^{2Q(s)}\frac{\boldsymbol{d}(s)^T\boldsymbol{\Sigma}(s)^{-1}\boldsymbol{d}(s)}{2\lambda}\mathrm{d}s, \\ \boldsymbol{C}_2(t;\lambda) := e^{-Q(t)}\displaystyle\int_t^T e^{Q(s)}\left[2\boldsymbol{A}(s)^T\boldsymbol{q}(s)C_1(s;\lambda) - \frac{\boldsymbol{A}(s)^T\boldsymbol{\Sigma}(s)^{-1}\boldsymbol{d}(s)}{\lambda}\right]\mathrm{d}s, \\ \boldsymbol{C}_3(t;\lambda) := \displaystyle\int_t^T\left[\boldsymbol{C}_2(s;\lambda)\boldsymbol{q}(s)^T\boldsymbol{A}(s) + \frac{\boldsymbol{A}(s)^T\boldsymbol{\Sigma}(s)^{-1}\boldsymbol{A}(s)}{2\lambda}\right]\mathrm{d}s, \\ C_4(t;\lambda) := e^{-Q(t)}\displaystyle\int_t^T e^{Q(s)}\left[\boldsymbol{C}_2(s;\lambda)^T\boldsymbol{p}(s) + 2C_1(s;\lambda)(\kappa(s)\theta(s) - \boldsymbol{q}(s)^T\overline{\boldsymbol{c}}(s)) + \frac{\overline{\boldsymbol{c}}(s)^T\boldsymbol{\Sigma}(s)^{-1}\boldsymbol{d}(s)}{\lambda}\right]\mathrm{d}s, \\ \boldsymbol{C}_5(t;\lambda) := \displaystyle\int_t^T\Big\{[\boldsymbol{C_3}(s;\lambda) + \boldsymbol{C}_3(s;\lambda)^T]\boldsymbol{p}(s) + \boldsymbol{C}_2(s;\lambda)(\kappa(s)\theta(s) - \boldsymbol{q}(s)^T\overline{\boldsymbol{c}}(s)) \\ \qquad + \boldsymbol{A}(s)\boldsymbol{q}(s)C_4(s;\lambda) - \dfrac{\boldsymbol{A}(s)^T\boldsymbol{\Sigma}(s)^{-1}\overline{\boldsymbol{c}}(s)}{\lambda}\Big\}\mathrm{d}s, \\ C_6(t;\lambda) := \displaystyle\int_t^T\Big\{\boldsymbol{C}_2(s;\lambda)^T b(s)\boldsymbol{\sigma}(s)\boldsymbol{\rho}(s) + b(s)^2C_1(s;\lambda) + \boldsymbol{C}_5(s;\lambda)^T\boldsymbol{p}(s) \\ \qquad + (\kappa(s)\theta(s)! - \boldsymbol{q}(s)^T\overline{\boldsymbol{c}}(s))C_4(s;\lambda) \\ \qquad + \dfrac{1}{2}tr\left(\left[\boldsymbol{C}_3(s;\lambda) + \boldsymbol{C}_3(s;\lambda)^T\right]\boldsymbol{\Sigma}(s)\right) + \dfrac{\overline{\boldsymbol{c}}(s)^T\boldsymbol{\Sigma}(s)^{-1}\overline{\boldsymbol{c}}(s)}{2\lambda}\Big\}\mathrm{d}s. \end{cases} \tag{24}$$

Then, the strategy defined as

$$\begin{aligned}\boldsymbol{\pi}^*(t,\boldsymbol{X}_t,Y_t) =& \frac{1}{2\lambda}\boldsymbol{\Sigma}(t)^{-1}\{\overline{\boldsymbol{c}}(t) + \boldsymbol{d}(t)Y_t - \boldsymbol{A}(t)\boldsymbol{X}_t\} - \{[\boldsymbol{C}_3(t;\lambda) + \boldsymbol{C}_3(t;\lambda)^T]\boldsymbol{X}_t + \boldsymbol{C_2(t;\lambda)}Y_t + \boldsymbol{C}_5(t;\lambda)\} \\ & - [\boldsymbol{X}_t^T\boldsymbol{C}_2(t;\lambda) + 2C_1(t;\lambda)Y_t + C_4(t;\lambda)]\boldsymbol{q}(t)\end{aligned} \tag{25}$$

is an optimal strategy of the dynamic MV problem (10), and

$$\begin{aligned}\mathbf{E}_t[\overline{V}_{t,T}(\boldsymbol{\pi}^*)] = & C_1(t;\lambda)Y_t^2 + \boldsymbol{X}_t^T\boldsymbol{C}_2(t;\lambda)Y_t + \boldsymbol{X}_t^T\boldsymbol{C}_3(t;\lambda)\boldsymbol{X}_t \\ & + C_4(t;\lambda)y + \boldsymbol{X}_t^T\boldsymbol{C}_5(t;\lambda) + C_6(t;\lambda).\end{aligned} \tag{26}$$

Remark 3.2 *The optimal strategy in (25) is referred to as* ***the optimal flexible-ratio strategy****. This strategy is linearly dependent on* $\boldsymbol{X}_t$ *and* Y_t. *Regarding the coefficients,* $\boldsymbol{C}_3(t;\lambda)$ *is an* $n\times n$ *matrix, whereas both* $\boldsymbol{C}_2(t;\lambda)$ *and* $\boldsymbol{C}_5(t;\lambda)$ *are* $n\times 1$ *vectors.*

Proof of Theorem 3.1 It follows from (3) that $\boldsymbol{\pi}^*$ given by (25) is admissible. According to Theorem 2.5, we just need prove that

$$\begin{aligned}f(t,\boldsymbol{x},y) = & C_1(t;\lambda)y^2 + \boldsymbol{x}^T\boldsymbol{C}_2(t;\lambda)y + \boldsymbol{x}^T\boldsymbol{C}_3(t;\lambda)\boldsymbol{x} \\ & + C_4(t;\lambda)y + \boldsymbol{x}^T\boldsymbol{C}_5(t;\lambda) + C_6(t;\lambda)\end{aligned} \tag{27}$$

is a solution of PDE (12) with the terminal condition $f(T,\boldsymbol{x},y)=0$. First, it can be seen from (24) that $\{C_1(t;\lambda), \boldsymbol{C}_2(t;\lambda), \boldsymbol{C}_3(t;\lambda), C_4(t;\lambda), \boldsymbol{C}_5(t;\lambda), C_6(t;\lambda)\}$ is a solution of the following ODE system:

$$\begin{cases} C_1'(t;\lambda) + 2\overline{q}(t)C_1(t;\lambda) + \dfrac{\boldsymbol{d}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{d}(t)}{2\lambda} = 0, \\ \boldsymbol{C}_2'(t;\lambda) + \overline{q}(t)\boldsymbol{C}_2(t;\lambda) + 2C_1(t;\lambda)\boldsymbol{A}(t)^T\boldsymbol{q}(t) - \dfrac{\boldsymbol{A}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{d}(t)}{\lambda} = 0, \\ \boldsymbol{C}_3'(t;\lambda) + \boldsymbol{C}_2(t;\lambda)\boldsymbol{q}(t)^T\boldsymbol{A}(t) + \dfrac{\boldsymbol{A}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{A}(t)}{2\lambda} = 0, \\ C_4'(t;\lambda) + \overline{q}(t)C_4(t;\lambda) + \boldsymbol{C}_2(t;\lambda)^T\boldsymbol{p}(t) \\ \qquad +2C_1(t;\lambda)(\kappa(t)\theta(t) - \boldsymbol{q}(t)^T\overline{\boldsymbol{c}}(t)) + \dfrac{\overline{\boldsymbol{c}}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{d}(t)}{\lambda} = 0, \\ \boldsymbol{C}_5'(t;\lambda) + \left[\boldsymbol{C}_3(t;\lambda) + \boldsymbol{C}_3(t;\lambda)^T\right]\boldsymbol{p}(t) + \boldsymbol{C}_2(t;\lambda)(\kappa(t)\theta(t) - \boldsymbol{q}(t)^T\overline{\boldsymbol{c}}(t)) \\ \qquad +\boldsymbol{A}(t)\boldsymbol{q}(t)^T C_4(t;\lambda) - \dfrac{\boldsymbol{A}(t)^T\boldsymbol{\Sigma}(t)^{-1}\overline{\boldsymbol{c}}(t)}{\lambda} = 0, \\ C_6'(t;\lambda) + \boldsymbol{C}_2(t;\lambda)^T b(t)\boldsymbol{\sigma}(t)\boldsymbol{\rho}(t) + b(t)^2C_1(t;\lambda) + \boldsymbol{C}_5(t;\lambda)^T\boldsymbol{p}(t) \\ \qquad +(\kappa(t)\theta(t) - \boldsymbol{q}(t)^T\overline{\boldsymbol{c}}(t))C_4(t;\lambda) \\ \qquad +\dfrac{1}{2}tr\left(\left[\boldsymbol{C}_3(t;\lambda) + \boldsymbol{C}_3(t;\lambda)^T\right]\boldsymbol{\Sigma}(t)\right\} + \dfrac{\overline{\boldsymbol{c}}(t)^T\boldsymbol{\Sigma}(t)^{-1}\overline{\boldsymbol{c}}(t)}{2\lambda} = 0. \end{cases} \tag{28}$$

Taking the partial derivative with respect to $t, \boldsymbol{x}, y$ in (27), it can be seen that

$$\begin{cases} \partial_t f(t,\boldsymbol{x},y) = C_1'(t;\lambda)y^2 + \boldsymbol{x}^T\boldsymbol{C}_2'(t;\lambda)y + \boldsymbol{x}^T\boldsymbol{C}_3'(t;\lambda)\boldsymbol{x} \\ \qquad\qquad + C_4'(t;\lambda)y + \boldsymbol{x}^T\boldsymbol{C}_5'(t;\lambda) + C_6'(t;\lambda), \\ \partial_{\boldsymbol{x}} f(t,\boldsymbol{x},y) = [\boldsymbol{C}_3(t;\lambda) + \boldsymbol{C}_3(t;\lambda)^T]\boldsymbol{x} + \boldsymbol{C}_2(t;\lambda)y + \boldsymbol{C}_5(t;\lambda), \\ \partial_y f(t,\boldsymbol{x},y) = \boldsymbol{x}^T\boldsymbol{C}_2(t;\lambda) + 2C_1(t;\lambda)y + C_4(t;\lambda), \\ \partial_{\boldsymbol{xx}} f(t,\boldsymbol{x},y) = \boldsymbol{C}_3(t;\lambda) + \boldsymbol{C}_3(t;\lambda)^T, \\ \partial_{yy} f(t,\boldsymbol{x},y) = 2C_1(t;\lambda), \\ \partial_{\boldsymbol{x}y} f(t,\boldsymbol{x},y) = \boldsymbol{C}_2(t;\lambda). \end{cases} \tag{29}$$

Substituting (29) into PDE (12), it follows from (28) that

$$\begin{aligned} &\mathcal{L}f(t,\boldsymbol{x},y) + \frac{1}{2\lambda}\boldsymbol{\alpha}^T(t,\boldsymbol{x},y)\boldsymbol{\Sigma}(t,y)^{-1}\boldsymbol{\alpha}(u,\boldsymbol{x},y) \\ &\quad - \boldsymbol{\alpha}^T(t,\boldsymbol{x},y)\partial_{\boldsymbol{x}} f(t,\boldsymbol{x},y) - \partial_y f(t,\boldsymbol{x},y)\boldsymbol{\alpha}^T(t,\boldsymbol{x},y)\boldsymbol{q}(t) \\ &= \partial_t f(t,\boldsymbol{x},y) + \partial_{\boldsymbol{x}} f^T(t,\boldsymbol{x},y)\boldsymbol{p}(t) + \partial_y f(t,\boldsymbol{x},y)[\kappa(t)(\theta(t) - y) - \boldsymbol{\alpha}^T(t,\boldsymbol{x},y)\boldsymbol{q}(t)] \\ &\quad + \frac{1}{2}tr(\partial_{\boldsymbol{xx}} f(t,\boldsymbol{x},y)\boldsymbol{\Sigma}(t)) + \frac{1}{2}\partial_{yy} f(t,\boldsymbol{x},y)b(t)^2 + \partial_{\boldsymbol{x}y} f^T(t,\boldsymbol{x},y)b(t)\boldsymbol{\sigma}(t)\boldsymbol{\rho}(t) \\ &= \boldsymbol{x}^T\left[\boldsymbol{C}_3'(t;\lambda) + \boldsymbol{C}_2(t;\lambda)\boldsymbol{q}(t)^T\boldsymbol{A}(t) + \frac{\boldsymbol{A}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{A}(t)}{2\lambda}\right]\boldsymbol{x} \\ &\quad + \left[C_1'(t;\lambda) + 2\overline{q}(t)C_1(t;\lambda) + \frac{\boldsymbol{d}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{d}(t)}{2\lambda}\right]y^2 \\ &\quad + \boldsymbol{x}^T\left[\boldsymbol{C}_2'(t;\lambda) + \overline{q}(t)\boldsymbol{C}_2(t;\lambda) + 2C_1(t;\lambda)\boldsymbol{A}(t)^T\boldsymbol{q}(t) - \frac{\boldsymbol{A}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{d}(t)}{\lambda}\right]y \\ &\quad + \boldsymbol{x}^T\left\{\boldsymbol{C}_5'(t;\lambda) + \left[\boldsymbol{C}_3(t;\lambda) + \boldsymbol{C}_3(t;\lambda)^T\right]\boldsymbol{p}(t) + \boldsymbol{C}_2(t;\lambda)(\kappa(t)\theta(t) - \boldsymbol{q}(t)^T\overline{\boldsymbol{c}}(t))\right\} \\ &\quad + y\left[C_4'(t;\lambda) + \overline{q}(t)C_4(t;\lambda) + \boldsymbol{C}_2(t;\lambda)^T\boldsymbol{p}(t) + 2C_1(t;\lambda)(\kappa(t)\theta(t) - \boldsymbol{q}(t)^T\overline{\boldsymbol{c}}(t)) + \frac{\overline{\boldsymbol{c}}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{d}(t)}{\lambda}\right] \\ &\quad + C_6'(t;\lambda) + \boldsymbol{C}_2(t;\lambda)^T b(t)\boldsymbol{\sigma}(t)\boldsymbol{\rho} + b(t)^2C_1(t;\lambda) + \boldsymbol{C}_5(t;\lambda)^T\boldsymbol{p}(t) + (\kappa(t)\theta(t) - \boldsymbol{q}(t)^T\overline{\boldsymbol{c}}(t))C_4(t;\lambda) \\ &\quad + \frac{1}{2}tr\left(\left[\boldsymbol{C}_3(t;\lambda) + \boldsymbol{C}_3(t;\lambda)^T\right]\boldsymbol{\Sigma}(t)\right\} + \frac{\overline{\boldsymbol{c}}(t)^T\boldsymbol{\Sigma}(t)^{-1}\overline{\boldsymbol{c}}(t)}{2\lambda} \\ &= 0. \end{aligned}$$

Thus, $f(t,\boldsymbol{x},y)$ given by (27) is a solution of PDE (12) with the terminal condition $f(T,\boldsymbol{x},y)=0$, and according to Theorem 2.5, $\boldsymbol{\pi}^*$ given by (25) is an optimal strategy. □

Corollary 3.3 *Assume that* $\boldsymbol{\pi}^*=\{\boldsymbol{\pi}^*(u,\boldsymbol{X}_u,Y_u)\}_{u\in[0,T]}$ *is the optimal flexible-ratio strategy, and* $f(t,\boldsymbol{x},y)$ *is given by (27). Let*

$$z(t,\boldsymbol{x},y)=\overline{C}_1(t;\lambda)y^2+\boldsymbol{x}^T\overline{\boldsymbol{C}}_2(t;\lambda)y+\boldsymbol{x}^T\overline{\boldsymbol{C}}_3(t;\lambda)\boldsymbol{x}+\overline{C}_4(t;\lambda)y+\boldsymbol{x}^T\overline{\boldsymbol{D}}_5(t;\lambda)+\overline{C}_6(t;\lambda),$$

where $\{\overline{C}_1(t;\lambda),\overline{\boldsymbol{C}}_2(t;\lambda),\overline{\boldsymbol{C}}_3(t;\lambda),\overline{C}_4(t,\lambda),\overline{\boldsymbol{C}}_5(t;\lambda),\overline{C}_6(t;\lambda)\}$ *is the solution of the following ODE system*

$$\begin{aligned}
&\overline{C}_1'(t;\lambda)+(\overline{\boldsymbol{C}}_2(t;\lambda)^T\boldsymbol{d}(t)-2\kappa(t)\overline{C}_1(t;\lambda))-\frac{1}{4\lambda}\boldsymbol{d}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{d}(t)\\
&\quad+4\lambda C_1(t;\lambda)^2b(t)^2(\boldsymbol{\rho}(t)^T\boldsymbol{\rho}(t)-1)=0,\\
&\overline{\boldsymbol{C}}_2'(t;\lambda)+(\overline{\boldsymbol{C}}_3(t;\lambda)+\overline{\boldsymbol{C}}_3(t;\lambda)^T)\boldsymbol{d}(t)-(\boldsymbol{A}(t)^T+\kappa(t))\overline{\boldsymbol{C}}_2(t;\lambda)+\frac{1}{2\lambda}\boldsymbol{A}(t)\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{d}(t)\\
&\quad+4\lambda b(t)^2(\boldsymbol{\rho}(t)^T\boldsymbol{\rho}(t)-1)\boldsymbol{C}_2(t;\lambda)C_1(t;\lambda)=\boldsymbol{0},\\
&\overline{\boldsymbol{C}}_3'(t;\lambda)-(\overline{\boldsymbol{C}}_3(t;\lambda)+\overline{\boldsymbol{C}}_3(t;\lambda)^T)\boldsymbol{A}(t)-\frac{1}{4\lambda}\boldsymbol{A}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{A}(t)\\
&\quad+\lambda b(t)^2(\boldsymbol{\rho}(t)^T\boldsymbol{\rho}(t)-1)\boldsymbol{C}_2(t;\lambda)\boldsymbol{C}_2(t;\lambda)^T=\boldsymbol{0},\\
&\overline{C}_4'(t;\lambda)+\overline{\boldsymbol{C}}_2(t;\lambda)^T\boldsymbol{c}(t)+\overline{\boldsymbol{C}}_5(t;\lambda)^T\boldsymbol{d}(t)+2\kappa(t)\theta(t)\overline{C}_1(t;\lambda)-\kappa(t)\overline{C}_4(t;\lambda)\\
&\quad-\frac{1}{2\lambda}\overline{c}(t)^T\boldsymbol{\Sigma}(t)^{-1}\boldsymbol{d}(t)+4\lambda b(t)^2(\boldsymbol{\rho}(t)^T\boldsymbol{\rho}(t)-1)C_1(t;\lambda)C_4(t;\lambda)=0,\\
&\overline{\boldsymbol{C}}_5'(t;\lambda)+(\overline{\boldsymbol{C}}_3(t;\lambda)+\overline{\boldsymbol{C}}_3(t;\lambda)^T)\boldsymbol{c}(t)-\overline{\boldsymbol{A}}(t)^T\overline{\boldsymbol{C}}_5(t;\lambda)+\kappa(t)\theta(t)\overline{\boldsymbol{C}}_2(t;\lambda)\\
&\quad+\frac{1}{2\lambda}\boldsymbol{A}(t)^T\boldsymbol{\Sigma}(t)^{-1}\overline{\boldsymbol{c}}(t)+2\lambda b(t)^2(\boldsymbol{\rho}(t)^T\boldsymbol{\rho}(t)-1)\boldsymbol{C}_2(t;\lambda)C_4(t;\lambda)=\boldsymbol{0},\\
&\overline{C}_6'(t;\lambda)+\overline{\boldsymbol{C}}_5(t;\lambda)^T\boldsymbol{c}(t)+\kappa(t)\theta(t)\overline{C}_4(t;\lambda)+\frac{1}{2}tr([\overline{\boldsymbol{C}}_3(t;\lambda)+\overline{\boldsymbol{C}}_3(t;\lambda)^T]\boldsymbol{\Sigma}(t))\\
&\quad+\overline{C}_1(t;\lambda)b(t)^2+\overline{\boldsymbol{C}}_2(t;\lambda)^Tb(t)\boldsymbol{\sigma}(t)\boldsymbol{\rho}(t)-\frac{1}{4\lambda}\overline{\boldsymbol{c}}(t)\boldsymbol{\Sigma}(t)^{-1}\overline{\boldsymbol{c}}(t)\\
&\quad+\lambda b(t)^2(\boldsymbol{\rho}(t)^T\boldsymbol{\rho}(t)-1)C_4(t;\lambda)^2=0.
\end{aligned}$$

Then, the optimal utility is given by

$$\mathbf{E}_t[\overline{V}_{t,T}(\boldsymbol{\pi}^*)]-\lambda\mathbf{Var}_t[\overline{V}_{t,T}(\pi^*)]=f(t,\boldsymbol{X}_t,Y_t)+z(t,\boldsymbol{X}_t,Y_t).$$

4. The optimal fixed-ratio strategy

In traditional pairs trading, agents select two cointegrated assets, and examine the mean-reverting spread process: they buy one asset and sell the other asset simultaneously. This approach represents a type of fixed-ratio strategy, where the fixed proportion vector is $(1,-1)^T$. This section examines the fixed-ratio pairs trading strategy for n risky assets in a continuous-time cointegration model.

We still assume that the dynamics of the log-price vector $\boldsymbol{X}_t$ is given by (22) and the common stochastic factor Y_t is given by (23). Furthermore, we assume that $\kappa(t),\theta(t),b(t),\boldsymbol{c}(t),\boldsymbol{d}(t)$, $\boldsymbol{\sigma}(t),\boldsymbol{\rho}(t)$ and $\boldsymbol{\Sigma}(t):=\boldsymbol{\sigma}(t)\boldsymbol{\sigma}(t)^T$ satisfy the same conditions as specified in Section 3. Additionally, we assume that the cointegrated matrix $\boldsymbol{A}$ is a constant positive definite matrix. We now aim to investigate the optimal pairs trading strategy using the fixed proportion vector $\boldsymbol{\xi}$, which is an eigenvector of $\boldsymbol{A}$.

More precisely, the proportion in the strategy utilizes a unit eigenvector $\boldsymbol{\xi} = (\xi^{(1)}, ..., \xi^{(n)})^T$ of $\boldsymbol{A}^T = \boldsymbol{A}$, with k being the corresponding real eigenvalue, that is,

$$\boldsymbol{A}^T\boldsymbol{\xi} = k\boldsymbol{\xi}.$$

Accordingly, the log spread process is denoted by

$$D_t^{\xi} := \boldsymbol{\xi}^T\boldsymbol{X}_t = \xi_1 X_1(t) + \cdots + \xi_n X_n(t),$$

and the dynamics of spread D_t^{ξ} are characterized by

$$\begin{aligned} \mathrm{d}D_t^{\xi} &= \left[\boldsymbol{\xi}^T\boldsymbol{c}(t) + \boldsymbol{\xi}^T\boldsymbol{d}(t)Y_t - \boldsymbol{\xi}^T\boldsymbol{A}\boldsymbol{X}_t\right]\mathrm{d}t + \boldsymbol{\xi}^T\boldsymbol{\sigma}(t)\mathrm{d}\boldsymbol{W}_t \\ &= \left[\boldsymbol{\xi}^T\boldsymbol{c}(t) + \boldsymbol{\xi}^T\boldsymbol{d}(t)Y_t - kD_t^{\xi}\right]\mathrm{d}t + \boldsymbol{\xi}^T\boldsymbol{\sigma}(t)\mathrm{d}\boldsymbol{W}_t. \end{aligned} \tag{30}$$

Notably, D_t^{ξ} is a mean-reverting process, which is crucial in pairs trading. The eigenvalue k governs the speed of mean reversion. Financial intuition suggests that a larger k results in a faster mean reversion, potentially offering more trading opportunities and, consequently, higher potential profits. We define:

$$\sigma^{\xi}(t) := \sqrt{\boldsymbol{\xi}^T\boldsymbol{\Sigma}(t)\boldsymbol{\xi}}, \quad \rho^{\xi}(t) := \boldsymbol{\xi}^T\boldsymbol{\sigma}(t)\boldsymbol{\rho}(t), \quad \mathrm{d}\overline{W}_t := \frac{1}{\boldsymbol{\sigma}^{\xi}(t)}\boldsymbol{\xi}^T\boldsymbol{\sigma}(t)\mathrm{d}\boldsymbol{W}_t, \tag{31}$$

where $\overline{W}_t$ is a standard Brownian motion such that $\mathrm{d}\overline{W}_t\mathrm{d}B_t = \rho^{\xi}(t)\mathrm{d}t$ and

$$\mathrm{d}D_t^{\xi} = \left[c^{\xi}(t) + d^{\xi}(t)Y_t - kD_t^{\xi}\right]\mathrm{d}t + \sigma^{\xi}(t)\mathrm{d}\overline{W}_t,$$

where $c^{\xi}(t) := \boldsymbol{\xi}^T\boldsymbol{c}(t), d^{\xi}(t) := \boldsymbol{\xi}^T\boldsymbol{d}(t)$. We use a scalar value $\hbar(t)$ to control the amount of money invested in the asset pair with the fixed proportion $\boldsymbol{\xi}$, and the corresponding wealth process $V_t^{\hbar}$ is given by

$$\mathrm{d}V_t^{\hbar} = V_t^{\hbar}\left[\hbar(t)M(t, D_t^{\xi}, Y_t)\mathrm{d}t + \hbar(t)\sigma^{\xi}(t)\mathrm{d}\overline{W}(t) + r\mathrm{d}t\right], \tag{32}$$

where

$$M(t, x, y) = c^{\xi}(t) + d^{\xi}(t)y - kx - r\sum_{i=1}^{n}\xi_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\xi_i\sigma_{ij}(t)^2.$$

Similar to before, we set $\pi(t) := e^{-rt}V_t^{\hbar}\hbar(t)$ as the strategy and consider the discounted wealth process: $\overline{V}_t(\pi) := e^{-rt}V_t^{\hbar}$. We consider the strategy depending solely on D^{ξ} and Y, i.e., $\pi = \pi(u, D_u^{\xi}, Y_u)$, and investigate the following dynamic MV problem:

$$\max_{\pi\in\mathcal{A}} \mathbf{E}_t\left[\overline{V}_T(\pi)\right] - \lambda\mathbf{Var}_t\left[\overline{V}_T(\pi)\right]. \tag{33}$$

In fact, this fixed-ratio strategy can be regarded as a special case of the flexible-ratio strategy when $n = 1$, although the assumptions may differ slightly. Employing a similar approach as in the proof of Theorem 3.1 allows us to derive the fixed-ratio optimal strategy. Nonetheless, to facilitate the comparison analysis in Section 5, it is important to obtain the closed-form expression of this strategy. Thus, similar to Section 3, we introduce:

$$\begin{aligned} p(t) &:= r\sum_{i=1}^{n}\xi_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\xi_i\sigma_{ij}(t)^2, \quad q(t) := \frac{b(t)\rho^{\xi}(t)}{\sigma^{\xi}(t)}, \\ \overline{c}(t) &:= c^{\xi}(t) - p(t), \quad \overline{q}(t) := -\kappa(t) - q(t)d^{\xi}(t), \end{aligned} \tag{34}$$

then,

$$M(t,x,y) = \overline{c}(t) - kx + d^{\xi}(t)y.$$

For $f \in \mathcal{C}_{\mathcal{H}}^{1,2,2}([0,T]\times\mathbb{R}\times\mathbb{R})$, we introduce the following operator

$$\begin{aligned} \mathcal{L}^{\xi}f(t,x,y) &= \partial_t f(t,x,y) + \partial_x f(t,x,y)\left[c^{\xi}(t) - kx + d^{\xi}(t)y\right] + \kappa(t)(\theta(t)-y)\partial_y f(t,x,y) \\ &+ \frac{1}{2}\left(\sigma^{\xi}(t)\right)^2\partial_{xx}f(t,x,y) + \frac{1}{2}b(t)^2\partial_{yy}f(t,x,y) + b(t)\sigma^{\xi}(t)\rho^{\xi}(t)\partial_{xy}f(t,x,y). \end{aligned}$$

Similar to Theorem 2.5, we have the following lemma.

Lemma 4.1 *Assume that* $\overline{f}(t,x,y), \overline{g}(t,x,y) \in \mathcal{C}_{\mathcal{H}}^{1,2,2}([0,T]\times\mathbb{R}\times\mathbb{R})$ *are the solutions of the following PDEs*

$$\begin{aligned} &\mathcal{L}^{\xi}\overline{f}(t,x,y) + \frac{1}{2\lambda\left(\sigma^{\xi}(t)\right)^2}M(t,x,y)^2 \\ &- M(t,x,y)\partial_x\overline{f}(t,x,y) - q(t)M(t,x,y)\partial_y\overline{f}(t,x,y) = 0, \end{aligned} \tag{35}$$

$$\begin{aligned} &\mathcal{L}^{\xi}\overline{g}(t,x,y) + \frac{1}{4\lambda\left(\sigma^{\xi}(t)\right)^2}M(t,x,y)^2 - M(t,x,y)\partial_x\overline{f}(t,x,y) \\ &- q(t)M(t,x,y)\partial_y\overline{f}(t,x,y) - \lambda\left(b(t)\partial_y\overline{f}(t,x,y)\right\}^2\left(1-\left(\rho^{\xi}(t)\right)^2\right) = 0, \end{aligned} \tag{36}$$

with the terminal conditions $\overline{f}(T,\boldsymbol{x},y) = 0$ *and* $\overline{g}(T,\boldsymbol{x},y) = 0$*, respectively. We also assume the same condition for* $\overline{f}$ *and* $\overline{g}$ *as in Theorem 2.5. Let*

$$\pi^*(t,D_t^{\xi},Y_t) := \frac{1}{2\lambda\left(\sigma^{\xi}(t)\right)^2}M(t,D_t^{\xi},Y_t) - \partial_x\overline{f}(t,D_t^{\xi},Y_t) - q(t)\partial_y\overline{f}(t,D_t^{\xi},Y_t). \tag{37}$$

Then, $\pi^* = \{\pi^*(t,D_t^{\xi},Y_t)\}_{t\in[0,T]} \in \mathcal{A}$ *is an optimal strategy for problem (33), fulfilling*

$$\begin{aligned} \mathbf{E}_t[\overline{V}_{t,T}(\pi)] &= \overline{f}(t,D_t^{\xi},Y_t), \\ \mathbf{E}_t[\overline{V}_{t,T}(\pi)] - \lambda\mathbf{Var}_t[\overline{V}_{t,T}(\pi)] &= \overline{g}(t,D_t^{\xi},Y_t). \end{aligned}$$

Remark 4.2 *The PDE (35) can be simplified into the following form*

$$\begin{aligned} &\partial_t\overline{f}(t,x,y) + p(t)\partial_x\overline{f}(t,x,y) + \left[\kappa(t)\theta(t) - q(t)\overline{c}(t) + kq(t)x + \overline{q}(t)y\right]\partial_y\overline{f}(t,x,y) \\ &+ \frac{1}{2}\left(\sigma^{\xi}(t)\right)^2\partial_{xx}\overline{f}(t,x,y) + \frac{1}{2}b(t)^2\partial_{yy}\overline{f}(t,x,y) + b(t)\sigma^{\xi}(t)\rho^{\xi}(t)\partial_{xy}\overline{f}(t,x,y) \\ &+ \frac{1}{2\lambda\left(\sigma^{\xi}(t)\right)^2}M(t,x,y)^2 = 0. \end{aligned}$$

Similar to Theorem 3.1, using Lemma 4.1, the optimal fixed-ratio strategy is obtained in the following corollary.

Corollary 4.3 *Let* $Q(t) := \int_0^t \overline{q}(t)\mathrm{d}s$*, and introduce*

$$\begin{cases} C_1(t;\lambda) := e^{-2Q(t)} \displaystyle\int_t^T e^{2Q(s)} \frac{\left(d^\xi(t)\right)^2}{2\lambda\left(\sigma^\xi(s)\right)^2} \mathrm{d}s, \\ C_2(t;\lambda) := e^{-Q(t)} \displaystyle\int_t^T e^{Q(s)} \left[2kq(s)C_1(s;\lambda) - \frac{kd^\xi(s)}{\lambda\left(\sigma^\xi(s)\right)^2}\right] \mathrm{d}s, \\ C_3(t;\lambda) := \displaystyle\int_t^T \left[kq(s)C_2(s;\lambda) + \frac{k^2}{2\lambda\left(\sigma^\xi(s)\right)^2}\right] \mathrm{d}s, \\ C_4(t;\lambda) := e^{-Q(t)} \displaystyle\int_t^T e^{Q(s)} \left[2\left(\kappa(s)\theta(s) - q(s)\overline{c}(s)\right) C_1(s;\lambda) + p(s)C_2(s;\lambda) + \frac{\overline{c}(s)d^\xi(s)}{\lambda\left(\sigma^\xi(s)\right)^2}\right] \mathrm{d}s, \\ C_5(t;\lambda) := \displaystyle\int_t^T \left[\left(\kappa(s)\theta(s) - q(s)\overline{c}(s)\right) C_2(s;\lambda) + 2p(s)C_3(s;\lambda) + kq(s)C_4(s;\lambda) - \frac{k\overline{c}(s)}{\lambda\left(\sigma^\xi(s)\right)^2}\right] \mathrm{d}s, \\ C_6(t;\lambda) := \displaystyle\int_t^T \Big[b(s)^2C_1(s;\lambda) + b(s)\sigma^\xi(s)\rho(s)C_2(s;\lambda) + (\sigma^\xi(s))^2C_3(s;\lambda) \\ \qquad + \left(\kappa(s)\theta(s) - q(s)\overline{c}(s)\right) C_4(s;\lambda) + p(s)C_5(s;\lambda) + \dfrac{\overline{c}(s)^2}{\lambda\left(\sigma^\xi(s)\right)^2}\Big] \mathrm{d}s. \end{cases} \tag{38}$$

Then, the strategy given by

$$\begin{aligned} \pi^*(t, D_t^\xi, Y_t) = &\left[\frac{\overline{c}(t)}{2\lambda\left(\sigma^\xi(t)\right)^2} - q(t)C_4(t;\lambda) - C_5(t;\lambda)\right] - \left[\frac{k}{2\lambda\left(\sigma^\xi(t)\right)^2} + q(t)C_2(t;\lambda) + 2C_3(t;\lambda)\right] D_t^\xi \\ &+ \left[\frac{d^\xi(t)}{2\lambda\left(\sigma^\xi(t)\right)^2} - 2q(t)C_1(t;\lambda) - C_2(t;\lambda)\right] Y_t \end{aligned} \tag{39}$$

is an optimal strategy of problem (33), and

$$\begin{aligned} \mathbf{E}_t\left[\overline{V}_{t,T}(\pi^*)\right] = &C_3(t;\lambda)\left(D_t^\xi\right)^2 + C_2(t;\lambda)D_t^\xi Y_t + C_1(t;\lambda)Y_t \\ &+ C_5(t;\lambda)D_t^\xi + C_4(t;\lambda)Y_t + C_6(t;\lambda). \end{aligned} \tag{40}$$

Remark 4.4 *The strategy given by (39) is termed* ***the optimal fixed-ratio strategy*** *for multiple assets pairs trading. It is observed that the optimal fixed-ratio strategy, corresponding to an eigenvector $\boldsymbol{\xi}$, exhibits linear dependence on the spread D_t^ξ and the common factor Y_t.*

Remark 4.5 *According to Lemma 4.1, it suffices to show that*

$$\begin{aligned} \overline{f}(t,x,y) := &C_1(t;\lambda)y^2 + C_2(t;\lambda)xy + C_3(t;\lambda)x^2 \\ &+ C_4(t;\lambda)y + C_5(t;\lambda)x + C_6(t;\lambda) \end{aligned} \tag{41}$$

satisfies the PDE (35) under the terminal condition $\overline{f}(T,x,y) = 0$. The proof is similar to the proof of Theorem 3.1, the details of which are omitted here.

Remark 4.6 *Similar to Corollary 3.3, we can define*

$$z(t,x,y) = \overline{C}_1(t;\lambda)y^2 + \overline{C}_2(t;\lambda)xy + \overline{C}_3(t;\lambda)y^2 + \overline{C}_4(t;\lambda)y + \overline{C}_5(t;\lambda)x + \overline{C}_6(t;\lambda),$$

where $\left\{\overline{C}_1(t), \overline{C}_2(t), \overline{C}_3(t), \overline{C}_4(t), \overline{C}_5(t), \overline{C}_6(t)\right\}$ is the solution of the following ODE system

$$
\begin{aligned}
&\overline{C}_1(t;\lambda)' + d^{\xi}(t)\overline{C}_2(t;\lambda) - 2\kappa(t)\overline{C}_1(t;\lambda) - \frac{(d^{\xi}(t))^2}{4\lambda(\sigma^{\xi}(t))^2} + 4\lambda b(t)^2((\rho^{\xi}(t))^2-1)C_1(t;\lambda)^2 = 0,\\
&\overline{C}_2(t;\lambda)' + 2d^{\xi}(t)\overline{C}_3(t;\lambda) - (k+\kappa(t))\overline{C}_2(t;\lambda) + \frac{kd^{\xi}(t)}{2\lambda(\sigma^{\xi}(t))^2}\\
&\qquad +4\lambda b(t)^2((\rho^{\xi}(t))^2-1)C_1(t;\lambda)C_2(t;\lambda) = 0,\\
&\overline{C}_3(t;\lambda)' - 2k\overline{C}_3(t;\lambda) - \frac{k^2}{4\lambda(\sigma^{\xi}(t))^2} + \lambda b(t)^2((\rho^{\xi}(t))^2-1)C_2(t;\lambda)^2 = 0,\\
&\overline{C}_4(t;\lambda)' + c^{\xi}(t)\overline{C}_2(t;\lambda) + 2\kappa(t)\theta(t)\overline{C}_1(t;\lambda) + d(t)^{\xi}\overline{C}_5(t;\lambda) - \kappa(t)\overline{C}_4(t;\lambda) - \frac{\overline{c}^{\xi}(t)d^{\xi}(t)}{2\lambda(\sigma^{\xi}(t))^2}\\
&\qquad +4\lambda b(t)^2((\rho^{\xi}(t))^2-1)C_1(t;\lambda)C_4(t;\lambda) = 0,\\
&\overline{C}_5(t;\lambda)' + 2c(t)^{\xi}\overline{C}_3(t;\lambda) + \kappa(t)\theta(t)\overline{C}_2(t;\lambda) - k\overline{C}_5(t;\lambda) + \frac{k\overline{c}^{\xi}(t)}{2\lambda(\sigma^{\xi}(t))^2}\\
&\qquad +2\lambda b(t)^2((\rho^{\xi}(t))^2-1)C_2(t;\lambda)C_4(t;\lambda) = 0,\\
&\overline{C}_6(t;\lambda)' + (\sigma^{\xi}(t))^2\overline{C}_3(t;\lambda) + b(t)\sigma^{\xi}(t)\rho^{\xi}(t)\overline{C}_2(t;\lambda) + b(t)^2\overline{C}_1(t;\lambda) + c^{\xi}(t)\overline{C}_5(t;\lambda) + \kappa(t)\theta(t)\overline{C}_4(t;\lambda)\\
&\qquad -\frac{\overline{(}c^{\xi}(t))^2}{4\lambda(\sigma^{\xi}(t))^2} + \lambda b(t)^2((\rho^{\xi}(t))^2-1)C_4(t;\lambda)^2 = 0,
\end{aligned}
$$

with the condition $\overline{C}_1(T;\lambda)=0, \overline{C}_2(T;\lambda)=0, \overline{C}_3(T;\lambda)=0, \overline{C}_4(T;\lambda)=0, \overline{C}_5(T;\lambda)=0$, $C_6(T;\lambda)=0$. *Then, the utility for the optimal fixed-ratio strategy is given by*

$$
\mathbf{E}_t\left[\overline{V}_{t,T}(\pi^*)\right] - \lambda \mathbf{Var}_t\left[\overline{V}_{t,T}(\pi^*)\right] = f\left(t, D_t^{\xi}, Y_t\right) + z\left(t, D_t^{\xi}, Y_t\right).
$$

5. Empirical experiments

This section conducts several experiments in which groups of three stocks are selected from the Chinese security markets SSE and SZSE. We compare the performance of the optimal flexible-ratio strategy, the optimal fixed-ratio strategy, and the optimal strategy proposed by Chiu and Wong [7] in the classical cointegration model, where there is no common stochastic factor. Here, the common stochastic factor is set as the log price of the Shanghai Securities Composite Index, one of the most important benchmarks for the Chinese stock market, albeit not directly tradable. Traditionally, pairs trading involves selecting stocks from the same industry with similar fundamentals to ensure cointegration. However, introducing a common stochastic factor makes it possible to select stocks from different industries, thus broadening the trading opportunities.

Without loss of generality, we assume $\boldsymbol{\sigma}$ to be a lower triangular matrix. In practice, parameters $b, \boldsymbol{c}, \boldsymbol{\rho}, \boldsymbol{d}, \boldsymbol{\sigma}, \kappa$ and θ can be estimated using an online updated procedure combined with the maximum likelihood estimation (which is commonly used in calibrating stochastic volatility models, see Wang et al.[17] for more detail) rolling on the historical data set. We also refer to Chambers [2] for details about testing and estimating continuous-time cointegration models.

Regarding the optimal fixed-ratio strategy, we employ eigenvectors corresponding to the largest and smallest eigenvalues of the cointegrated matrix $\boldsymbol{A}$ as fixed proportion vectors to assess performance in both cases. Empirical experiments indicate that larger eigenvalues are associated with faster mean reversion, leading to better strategies, which aligns with financial intuition.

We next discuss the effectiveness of both the optimal flexible-ratio and the optimal fixed-ratio strategies associated with the largest and smallest eigenvalues. Additionally, we empirically illustrate the optimal strategy proposed by Chiu and Wong [7]. We compare the performance of their strategy with our optimal flexible-ratio strategy on the same selection of stocks. The numerical experiment is an out-of-sample test employing real historical data from the Chinese stock markets SSE and SZSE.

We select six groups of stocks for the period from January 23, 2020 to February 28, 2023. Each group comprises three stocks from different industries. The selected stocks in each group are listed in Table 1. We use the forward-adjusted daily closing price to avoid the impact of dividends.

Table 1 Selected stocks in each group

Label	Stock code	Stock name	Industry
	000554	Sinopec Shandong Taishan Petroleum	Petroleum
1	600386	Beijing Bashi Media	Automobile
	600202	Harbin Air Conditioning	Equipment
	000677	Chtc Helon	Chemical
2	002495	Guangdong Jialong Food	Food
	000520	Chang Jiang Shipping Group Phoenix	Transportation
	002308	Vtron Group	Computer
3	603499	Shanghai Sunglow Packaging Technology	Manufacturing
	002303	Mys Group	Manufacturing
	002836	Guangdong New Grand Long Packing	Manufacturing
4	002917	Shenzhen King Explorer Science And Technology	Chemical
	002769	Shenzhen Prolto Supply Chain Management	Transportation
	601512	China-Singapore Suzhou Industrial Park Development Group	Real Estate
5	600624	Shanghai Fudan Forward S&t	Medical
	600604	Shanghai Shibei Hi-Tech	Real Estate
	603790	Argus (Shanghai) Textile Chemicals	Chemical
6	002786	Shenzhen Silver Basis Technology	Mechanical
	002308	Vtron Group	Computer

Figure 1 illustrates the adjusted log prices of stocks from the last group listed in Table 1, along with the corresponding log price of the Shanghai Securities Composite Index.

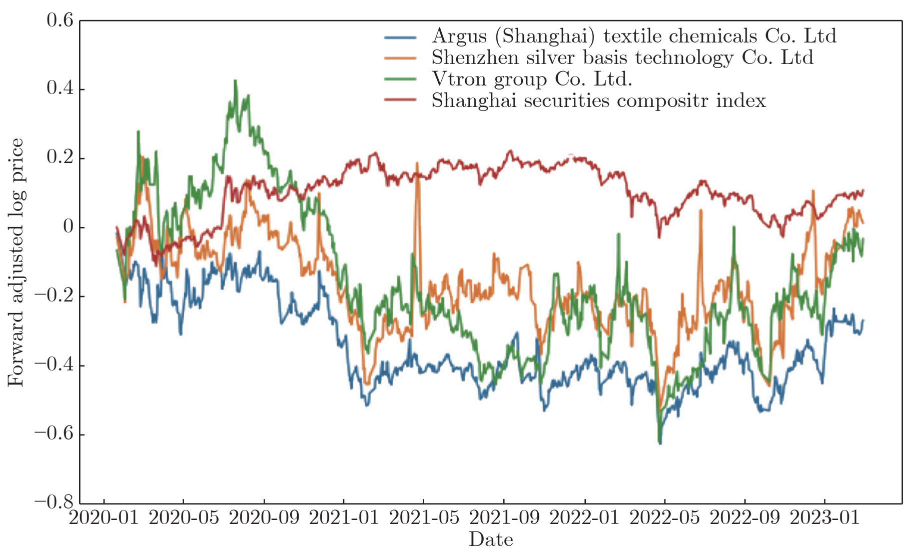

Figure 1 Forward-adjusted log prices of stocks and the Shanghai Securities Composite Index

We perform the out-of-sample test for each group starting from February 28, 2022. All parameters were estimated daily using data from the previous two years (which is approximately 500 trading days). We allocate the initial endowment of 1000 for each group of stocks and set $T = 1, \lambda = 1$ and $r = 0.02$. Figures 2–4 present the comparison of the strategy profit for groups 1,2 and 5 given by Table 1.

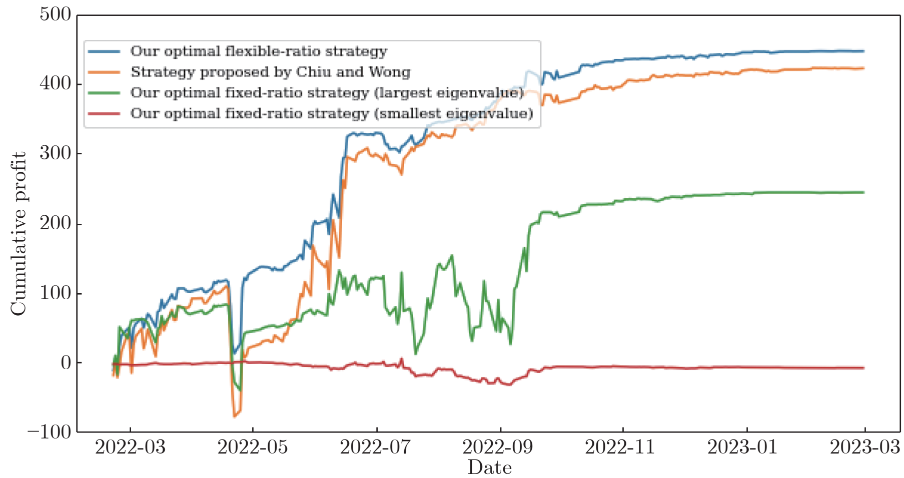

Figure 2 Cumulative profit of group 1

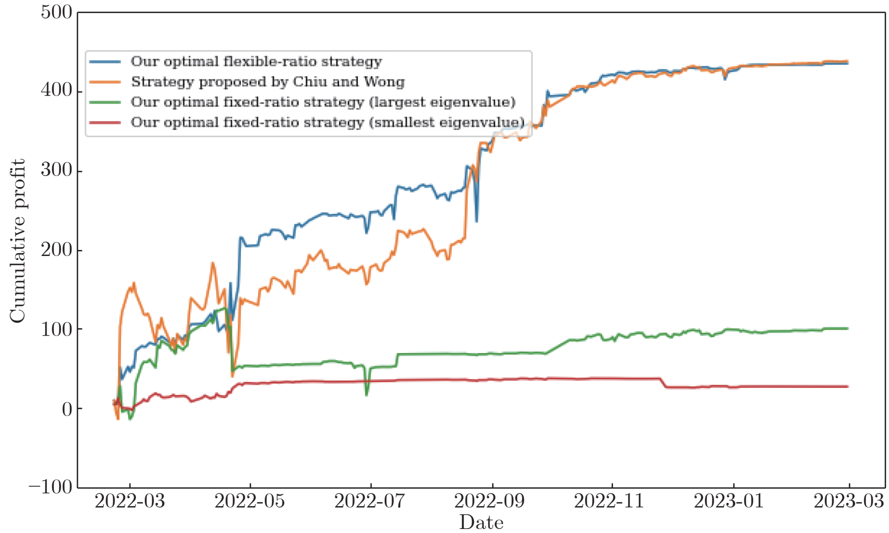

Figure 3 Cumulative profit of group 2

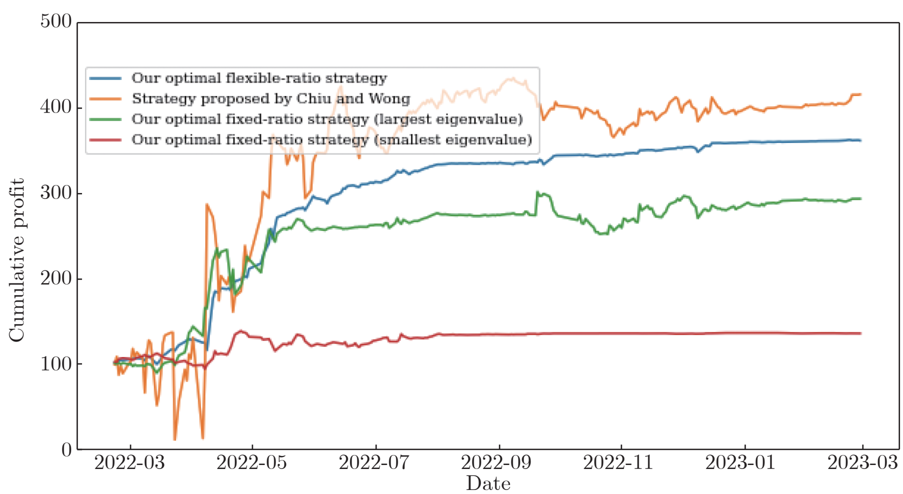

Figure 4 Cumulative profit of group 5

The results reveal that our optimal flexible-ratio strategy, the optimal fixed-ratio strategy for the largest eigenvalue and the strategy proposed by Chiu and Wong[7] all yield significant profitability in out-of-sample testing. However, the fixed-ratio strategy for the smallest eigenvalue did not. Moreover, while the optimal strategy proposed by Chiu and Wong[7] may occasionally produce higher profits, it exhibits significantly greater volatility than our optimal flexible-ratio strategy. This result suggests that our optimal flexible-ratio strategy is better than the optimal strategy proposed by Chiu and Wong[7] in terms of controlling risk. Tables 2–5 provide some important statistics of the out-of-sample trading results for all strategies.

Table 2 Performance of our optimal flexible-ratio strategy

Label	Winning rate/%	Profit-loss ratio	Average profit/%	Max drawdown/%	Sharp ratio
1	59.274	1.381	0.153	−9.88	2.358
2	55.242	1.662	0.142	−5.522	2.289
3	58.35	1.507	0.099	−3.917	2.142
4	58.753	1.635	0.114	−3.602	3.302
5	64.185	1.975	0.144	−2.634	2.975
6	59.557	1.711	0.121	−2.642	3.855

Table 3 Performance of the strategy proposed by Chiu and Wong [7]

Label	Winning rate/%	Profit-loss ratio	Average profit/%	Max drawdown/%	Sharp ratio
1	57.258	1.16	0.149	−18.459	1.548
2	54.435	1.308	0.143	−12.981	1.609
3	49.899	1.206	0.104	−19.888	0.722
4	61.972	1.267	0.101	−3.083	3.042
5	53.924	1.076	0.165	−26.941	0.729
6	53.119	1.11	0.119	−11.546	1.156

Table 4 Performance of our optimal fixed-ratio strategy for the largest eigenvalue

Label	Winning rate/%	Profit-loss ratio	Average profit/%	Max drawdown/%	Sharp ratio
1	54.032	1.065	0.089	−12.059	0.931
2	57.258	1.008	0.037	−10.29	0.894
3	50.302	1.504	0.095	−8.837	1.313
4	51.509	1.219	0.05	−7.98	1.058
5	51.509	1.466	0.115	−8.91	1.72
6	55.131	1.325	0.114	−8.197	1.853

Table 5 Performance of our optimal fixed-ratio strategy for the smallest eigenvalue

Label	Winning rate/%	Profit-loss ratio	Average profit/%	Max drawdown/%	Sharpe ratio
1	45.565	1.143	−0.002	−3.789	−0.18
2	52.016	1.224	0.009	−1.525	0.905
3	49.296	1.214	0.009	−3.073	0.734
4	45.473	1.419	0.005	−1.143	0.863
5	51.71	1.239	0.026	−4.493	1.105
6	48.29	0.967	−0.006	−9.7	−0.224

Examining the data in Tables 2 and 3 indicates that both strategies are effective, as the average profits are all positive. However, our optimal flexible-ratio strategy demonstrates a lower drawdown risk and a consistently higher Sharpe ratio across nearly all cases, indicating less uncertainty in returns. The optimal fixed-ratio strategy for the largest eigenvalue always outperforms the corresponding strategy for the smallest eigenvalue regarding the average profit and Sharpe ratio. It is notable that the optimal fixed-ratio strategy for the largest eigenvalue outperforms the optimal strategy proposed by Chiu and Wong [7] in terms of the Sharpe ratio in half the cases; however, it consistently generates lower profits.

In practice, we can allocate initial funds of 1000 units equally across all six groups to compute the average cumulative profit of each strategy. Figure 5 compares the average cumulative profit for these four optimal strategies.

Compared to the optimal strategy proposed by Chiu and Wong [7], Figure 5 reveals that the cumulative profits for our optimal flexible-ratio and optimal fixed-ratio strategies for the largest eigenvalue exhibit smoother trajectories, indicating better risk control. Furthermore, while the optimal fixed-ratio strategy for the largest eigenvalue is inferior to the optimal flexible-ratio strategy in terms of both profitability and risk control, it significantly outperforms the optimal fixed-ratio strategy for the smallest eigenvalue in profitability. Figure 5 illustrates that the volatility of the cumulative profit for each optimal strategy diminishes over time. This phenomena is further illustrated in Figure 6. Figure 6 shows the trends in the average amount of money invested in stocks under our optimal flexible-ratio strategy. As the terminal time T draws closer, the money

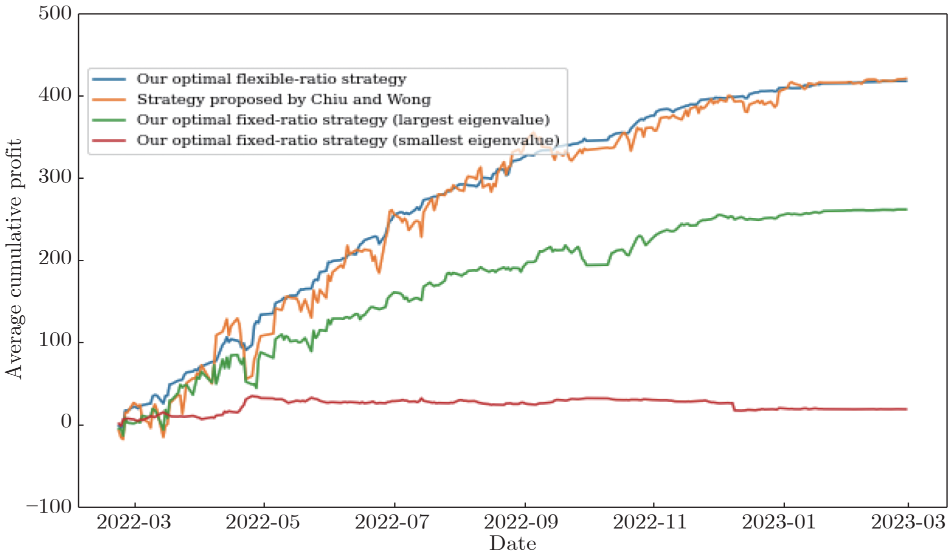

Figure 5 The average cumulative profit in out-of-sample testing

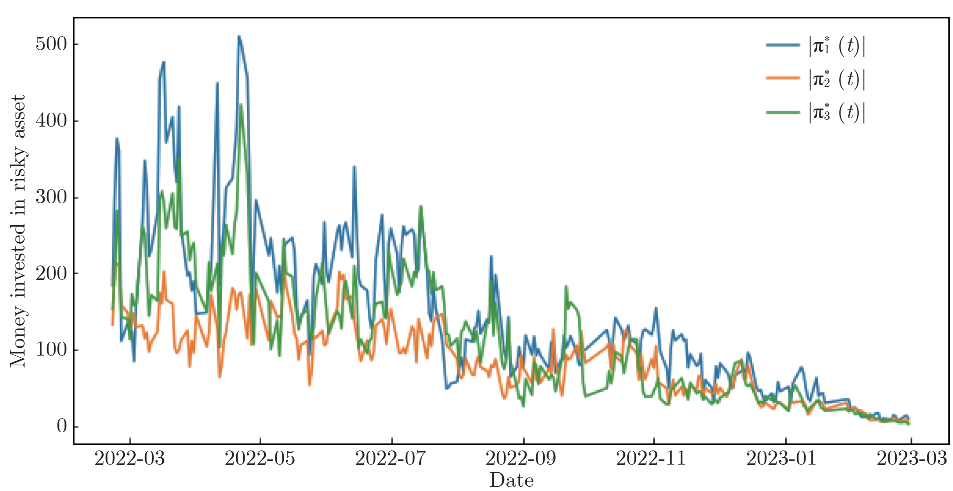

Figure 6 The average amount of money invested on each stocks

invested in risky assets progressively declines, reducing the uncertainty of the strategy's profit over time.

To further compare the value of the mean–variance utility J of our optimal flexible-ratio and optimal fixed-ratio strategies for the largest eigenvalue, simulations are performed using parameters estimated from real market data for the stocks in the groups listed Table 1. The parameters for these simulations are given in Tables 6–8.

Table 6 Parameters for simulation (Part I)

Label	A_{11}	A_{12}	A_{13}	A_{21}	A_{22}	A_{23}	A_{31}	A_{32}	A_{33}
1	27.149	−8.321	−8.195	−7.767	13.125	−0.842	−14.506	−7.1	20.258
2	23.987	−10.259	−8.171	−8.92	15.186	2.726	−16.319	−1.559	29.822
3	10.82	−13.804	17.59	−15.466	15.72	17.702	−6.273	−10.814	37.786
4	18.95	−9.67	−1.133	−9.918	14.299	−4.653	−10.989	−4.058	16.122
5	34.271	−15.111	0.941	−12.141	24.34	−7.264	8.592	−17.473	27.786
6	34.564	−9.982	−6.899	−11.949	20.802	−8.374	−6.951	3.174	9.369

Table 7 Parameters for simulation (Part II)

Label	c_1	c_2	c_3	d_1	d_2	d_3	ρ_1	ρ_2	ρ_3
1	1.727	0.333	0.247	−5.8	0.591	−8.986	0.368	0.315	0.099
2	−1.621	−0.178	−2.699	1.392	−1.812	5.125	0.341	0.317	0.137
3	−3.961	−7.224	−7.848	−6.579	2.706	−3.201	0.339	0.237	0.262
4	1.174	−0.015	0.136	−19.265	8.448	−7.201	0.279	0.251	0.225
5	−6.628	−1.725	−9.332	4.504	4.95	4.208	0.404	0.234	0.211
6	−10.288	3.303	0.242	−0.52	−1.707	−4.062	0.483	0.219	0.115

Table 8 Parameters for simulation (Part III)

Label	σ_{11}	σ_{12}	σ_{13}	σ_{21}	σ_{22}	σ_{23}	σ_{31}	σ_{32}	σ_{33}
1	0.423	0.0	0.0	0.113	0.35	0.0	0.167	0.128	0.438
2	0.376	0.0	0.0	0.13	0.356	0.0	0.135	0.108	0.356
3	0.471	0.0	0.0	0.132	0.368	0.0	0.106	0.098	0.233
4	0.428	0.0	0.0	0.108	0.471	0.0	0.144	0.085	0.436
5	0.252	0.0	0.0	0.062	0.341	0.0	0.146	0.118	0.295
6	0.306	0.0	0.0	0.134	0.505	0.0	0.216	0.048	0.418

We chose $T = 1/12, r = 0.02, \mathrm{d}t = 1/250, \kappa = 3.583, \theta = 0.111, b = 0.010$, and $V_0 = 100$. Setting the initial values provided by Table 9, we perform simulations for each group 1000 times, varying the risk-aversion parameter λ from 0.25 to 1.5. The statistics of the discounted terminal wealth for each group are shown in Table 10, where J represents the value of the mean–variance utility as defined in (9).

Table 9 Initial values for simulation

Label	Y	X_1	X_2	X_3
1	0.097	0.21	0.155	0.163
2	0.097	−0.135	0.011	−0.128
3	0.097	−0.065	−0.093	−0.254
4	0.097	0.155	0.133	0.161
5	0.097	−0.21	−0.353	−0.443
6	0.097	−0.277	0.02	−0.065

Table 10 Simulation results for each group

Label	λ	The optimal flexible-ratio strategy			The optimal fixed-ratio strategy		
		Mean	S.D.	J	Mean	S.D.	J
1	0.25	117.104	6.447	106.715	107.806	6.439	97.442
	0.5	108.552	3.223	103.357	103.901	3.219	98.721
	1.0	104.276	1.612	101.679	101.946	1.608	99.361
	1.5	102.851	1.074	101.119	101.292	1.07	99.574
2	0.25	120.853	7.447	106.988	111.114	7.175	98.244
	0.5	110.427	3.724	103.494	105.557	3.587	99.122
	1.0	105.213	1.862	101.747	102.778	1.794	99.561
	1.5	103.476	1.241	101.165	101.852	1.196	99.708
3	0.25	120.829	6.514	110.221	114.079	12.983	71.939
	0.5	110.415	3.257	105.11	107.041	6.493	85.962
	1.0	105.207	1.629	102.555	103.522	3.249	92.967
	1.5	103.472	1.086	101.703	102.35	2.169	95.296
4	0.25	115.299	5.524	107.671	102.855	5.735	94.632
	0.5	107.65	2.762	103.835	101.427	2.868	97.316
	1.0	103.825	1.381	101.918	100.714	1.434	98.658
	1.5	102.55	0.921	101.278	100.476	0.956	99.105
5	0.25	166.812	17.873	86.956	79.765	26.535	−96.266
	0.5	133.406	8.936	93.478	89.883	13.268	1.869
	1.0	116.703	4.468	96.739	94.942	6.633	50.939
	1.5	111.135	2.979	97.826	96.628	4.422	67.298
6	0.25	120.606	7.183	107.706	113.41	7.237	100.317
	0.5	110.303	3.592	103.853	106.705	3.619	100.157
	1.0	105.152	1.796	101.927	103.351	1.81	100.075
	1.5	103.434	1.197	101.284	102.233	1.207	100.047

It can be concluded that both strategies are effective, since the mean of the discounted terminal wealth is greater than the initial asset value of $V_0 = 100$ in all cases. Beyond achieving a higher mean terminal wealth, the optimal flexible-ratio strategy consistently exhibits a lower standard deviation of terminal wealth than the optimal fixed-ratio strategy for the largest eigenvalue. This indicates that the optimal flexible-ratio strategy outperforms the optimal fixed-ratio strategy in generating higher returns and ensuring the stability of those returns. Consequently, the optimal flexible-ratio strategy's mean–variance utility J is always larger.

Notably, terminal wealth's mean and standard deviation decreases as λ increases. This phenomenon, as discussed by Zhu et al. [22] and Chiu and Wong [7], suggests that λ is crucial in managing the risks associated with these strategies. A higher λ indicates a greater risk aversion, reducing the amount of money invested in risky assets, thereby mitigating uncertainty.

Finally, our empirical experiments have validated the effectiveness of both our optimal flexible-ratio and optimal fixed-ratio strategies for the largest eigenvalue and the optimal strategy proposed by Chiu and Wong [7]. Additionally, comparing trading results shows that our optimal flexible-ratio strategy, under a cointegration model with a common stochastic factor yields more stable returns than the optimal strategy proposed by Chiu and Wong [7] in the traditional cointegration model. Our model captures not only the interrelations between the selected stocks but also their connection to broader market dynamics via the common stochastic factor. Although not directly tradable, this factor significantly influences stock returns. Furthermore, empirical experiments show that our optimal flexible-ratio strategy always outperforms the optimal fixed-ratio strategy for the largest eigenvalue and significantly outperforms the strategy for the smallest eigenvalue, consistent with financial intuition.

6. Conclusion

This paper investigates optimal pairs trading strategies for the mean–variance problem under a continuous-time cointegration model with a common stochastic factor. This model is adept at capturing the interrelations between the selected stocks and their relationship with broader market dynamics. We first provide a semiclosed-form optimal strategy for a general model, derived from the PDE solutions. Next, we specify the model's assumptions into a Gaussian model with a common mean-reverting factor, providing a closed-form solution for the optimal flexible-ratio strategy. Furthermore, we assume that the fixed proportion vector is an eigenvector of the cointegrated matrix and give an closed-form expression of the optimal fixed-ratio strategy based on the spread process and the common factor. Finally, we conduct numerical experiments on real data from the Chinese market and evaluate the performance of the optimal flexible-ratio and fixed-ratio strategies and the optimal strategy proposed by Chiu and Wong [7] in the traditional cointegration model. The empirical experiments reveal that our optimal flexible-ratio strategy outperforms that proposed by Chiu and Wong [7] and the optimal fixed-ratio strategy in delivering higher and more stable profits.

References

[1] Björk, T., Murgoci, A. and Zhou, X. Y., Mean–variance portfolio optimization with state-dependent risk aversion, Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics, 2014, 24(1): 1–24.

[2] Chambers, M. J., The estimation of continuous time models with mixed frequency data, Journal of Econometrics, 2016, 193(2): 390–404.

[3] Chen, K., Chiu, M. C. and Wong, H. Y., Time-consistent mean-variance pairs-trading under regime-switching cointegration, SIAM Journal on Financial Mathematics, 2019, 10(2): 632–665.

[4] Chiu, M. C. and Wong, H. Y., Mean–variance portfolio selection of cointegrated assets, Journal of Economic Dynamics and Control, 2011, 35(8): 1369–1385.

[5] Chiu, M. C. and Wong, H. Y., Mean–variance asset–liability management: Cointegrated assets and insurance liability, European Journal of Operational Research, 2012, 223(3): 785–793.

[6] Chiu, M. C. and Wong, H. Y., Optimal investment for an insurer with cointegrated assets: Crra utility, Insurance: Mathematics and Economics, 2013, 52(1): 52–64.

[7] Chiu, M. C. and Wong, H. Y., Dynamic cointegrated pairs trading: Mean–variance time-consistent strategies, Journal of Computational and Applied Mathematics, 2015, 290: 516–534.

[8] Engle, R. F. and Granger, C. W. J., Co-integration and error correction: Representation, estimation, and testing, Econometrica, Journal of the Econometric Society, 1987, 55(2): 251–276.

[9] Fleming, W. H., Stochastic differential equations and applications, Vol. 1 (Avner Friedman), SIAM Review, 1977, 19(2): 366–367.

[10] Jurek, J. W. and Yang, H., Dynamic portfolio selection in arbitrage, In EFA 2006 Meetings Paper, 2007.

[11] Liu, J. and Timmermann, A., Optimal convergence trade strategies, The Review of Financial Studies, 2013, 26(4): 1048–1086.

[12] Ma, G. and Zhu, S.-P., Optimal investment and consumption under a continuous-time cointegration model with

exponential utility, Quantitative Finance, 2019, 19(7): 1135–1149.

[13] Øksendal, B., Stochastic Differential Equations: An Introduction with Applications, 6th edn., Springer, Berlin, Heidelberg, 2014.

[14] Shen, Y. and Siu, T. K., Optimal investment and consumption in a continuous-time co-integration model, IMA Journal of Management Mathematics, 2017, 28(4): 501–530.

[15] Vidyamurthy, G., Pairs Trading: Quantitative Methods and Analysis, Vol. 217, John Wiley & Sons, 2004.

[16] Wang, J., Zhang, D. and Zhang, J., Mean reversion in stock prices of seven Asian stock markets: Unit root test and stationary test with fourier functions, International Review of Economics & Finance, 2015, 37: 157–164.

[17] Wang, X., He, X., Bao, Y. and Zhao, Y., Parameter estimates of heston stochastic volatility model with MLE and consistent EKF algorithm, Science China Information Sciences, 2018, 61(4): 1–17.

[18] Wu, Y., Momentum trading, mean reversal and overreaction in chinese stock market, Review of Quantitative Finance and Accounting, 2011, 37: 301–323.

[19] Yu, F., Ching, W.-K., Wu, C. and Gu, J.-W., Optimal pairs trading strategies: A stochastic mean–variance approach, Journal of Optimization Theory and Applications, 2023, 196(1): 36–55.

[20] Zhang, Y. and Xiong, D., Optimal strategy of the dynamic mean-variance problem for pairs trading under a fast mean-reverting stochastic volatility model, Mathematics, 2023, 11(9): 2191.

[21] Zhou, X. Y. and Li, D., Continuous-time mean-variance portfolio selection: A stochastic LQ framework, Applied Mathematics & Optimization, 2000, 42(1): 19–33.

[22] Zhu, D.-M., Gu, J.-W., Yu, F.-H., Siu, T.-K. and Ching, W.-K., Optimal pairs trading with dynamic mean-variance objective, Mathematical Methods of Operations Research, 2021, 94(1): 145–168.