mathematics

Article

Optimal Strategy of the Dynamic Mean-Variance Problem for Pairs Trading under a Fast Mean-Reverting Stochastic Volatility Model

Yaoyuan Zhang and Dewen Xiong *

School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai 200240, China; s.yaoyuan.zhang@gmail.com
* Correspondence: xiongdewen@sjtu.edu.cn

Abstract: We discuss the dynamic mean-variance (MV) problem for pairs trading with the assumptions that one of the security prices satisfies a stochastic volatility model (SVM) and the corresponding price spread follows an Ornstein–Uhlenbeck (OU) process. We provide a semi-closed-form of the optimal strategy based on the solution of a PDE, which is difficult to solve explicitly. Thus, we assume that one of the security prices satisfies the Scott model, a fast-mean-reverting volatility model, and give a closed-form approximation for the optimal strategy. Empirical studies, by using historical data from Chinese security markets, show that the Scott model produces a more stable strategy by better capturing mean-reverting volatility.

Keywords: stochastic volatility; Ornstein–Uhlenbeck process; asymptotic analysis; dynamic mean-variance problem; pairs trading

MSC: 91G10; 91G80; 93E20; 35C20

check for updates

Citation: Zhang, Y.; Xiong, D. Optimal Strategy of the Dynamic Mean-Variance Problem for Pairs Trading under a Fast Mean-Reverting Stochastic Volatility Model. *Mathematics* **2023**, *11*, 2191. https://doi.org/10.3390/math11092191

Academic Editors: Jingtao Shi and Zhibin Liang

Received: 8 March 2023
Revised: 13 April 2023
Accepted: 4 May 2023
Published: 6 May 2023

1. Introduction

Pairs trading is a statistical arbitrage strategy that emerged from a Morgan Stanley quantitative group from in the 1980s. Investors choose a pair of highly-correlated securities, buy the relatively under-priced security, and sell the relatively over-priced security simultaneously, with the expectation of making a profit from the price spread regression. In this paper, we discuss the dynamic mean-variance (MV) problem for pairs trading in the Scott model, a fast-mean-reverting volatility model.

The cointegration approach is popular in pairs trading and was proposed by Vidyamurthy [1]. Vidyamurthy assessed the co-movement of securities through conintegration testing and designed a trading rule based on a simple nonparametric threshold. Lin et al. [2] studied the optimal trading threshold problem by introducing a minimum profit condition for a conintegrated pair of stocks. In a recent work by Yan et al. [3], they discussed pairs trading under a delayed cointegration model. In the cointegration framework, it is very important to model price spread. Elliott et al. [4] described spread using a mean-reverting Gaussian Markov chain model and developed an analytical framework for pairs trading strategies. Bertram [5] developed a statistical arbitrage model for the spread of two log price series under the assumption of the OU process. Many authors have viewed the optimal pairs trading problem as a stochastic control problem and have considereded it by maximizing a variety of utility functions. Jurek and Yang [6] discussed asset allocation strategies between a mean reverting arbitrage opportunity described by an OU process and a risk-free asset and assigned a closed-form optimal allocation for CRRA utility over a finite time horizon. Suzuki [7,8] and Endres and Stübinger [9] solved an optimal regime switching problem with the constraints of finite transaction times and transaction fees. Liu and Timmermann [10] derived a closed-form optimal strategy based on the cointegration assumption with the power utility over the terminal wealth. Chiu and Wong [11] assumed that the log prices of the risky assets satisfy the linear stochastic differential equation with a constant matrix of cointegration coefficients. They considered the time-consistent dynamic mean-variance problem (according to dynamic programming)and provided a closed-form optimal strategy. Recently, Zhu et al. [12] assumed that price spread follows an OU process and that one of the corresponding securities satisfies the geometric Brownian motion(GBM) model with a constant volatility. They considered the time-consistent mean-variance problem (see Björk et al. [13]) for pairs trading and provided a closed-form optimal strategy.

However, in all of the models mentioned above, the drift and the volatility of the (log) price processes are all deterministic functions or constants that are unable to capture some of the stochastic volatility characteristics that can be observed in many markets, such as mean reversion effects, see Cont [14,15], Teräsvitra and Zhao [16], and Gatheral et al. [17] for empirical evidence related to this. Therefore, in this paper, we assume that price spread follows an OU process and that one of the log prices of the securities satisfies a stochastic volatility model (SVM). Additionally, we discuss the dynamic mean-variance problem in the means of Björk et al. [13]. Because an SVM can describe price dynamics better, it is more likely to develop strategies that can be used to control trading risk more precisely and bring about greater utility. In this paper, we first discuss the optimal strategy for pairs trading under a general SVM by a PDE. Then, we specify a fast mean-reverting stochastic volatility model, the Scott model [18], and discuss the approximate optimal strategy.

The main contributions of this paper are as follows: First, we provide a semi-closed-form of the optimal strategy under a general SVM based on the solution of a PDE. Second, we provide an closed-form approximation of the optimal strategy under a fast mean-reverting volatility model that captures the mean-reverting property of volatilities by using the asymptotic analysis technique. Our approximate formula can be proven to have sufficient precision and extremely high computational efficiency compared with the traditional finite difference method(FDM), which is of great practical value. Finally, we calibrate the model parameters using securities data from the Chinese stock markets, and demonstrate the effect of our approximated optimal strategy by comparing it with the optimal strategy described in Zhu et al. [12]. Empirical studies show that the Scott model can produce a more stable strategy by better capturing mean-reverting volatility.

The remainder of this paper is organized as follows: In Section 2, we describe the model specifications as well as the optimal dynamic MV problem and provide the semi-closed-form optimal strategy based on a PDE. In Section 3, we provide a closed-form approximation of the optimal strategy for the Scott volatility model. In Section 4, we validate the trading strategy empirically using Chinese securities market data. Section 5 concludes the paper.

2. The Dynamic Mean-Variance Problem for a General Stochastic Volatility Model

In this section, we set up the dynamic mean-variance(MV) problem for pairs trading under a general SVM. Since the MV problem is time inconsistent, we discuss the optimal strategy, according to the definition of the equilibrium strategy introduced by [13] by transforming the dynamic MV problem to a non-cooperative Nash equilibrium game.

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space and that $W_1 = \{W_1(t); t \in [0,T]\}$, $W_2 = \{W_2(t); t \in [0,T]\}$ and $W_3 = \{W_3(t); t \in [0,T]\}$ are Brownian motions with

$$dW_1(t)dW_2(t) = \rho_{12}dt, \quad dW_1(t)dW_3(t) = \rho_{13}dt, \quad dW_2(t)dW_3(t) = \rho_{23}dt.$$

We assume that $\mathbb{F} = \{\mathcal{F}_t; t \in [0,T]\}$ is the filtration generated by W_1, W_2 and W_3. The conditional expectation and conditional variance with respect to $\mathcal{F}_t$ are denoted as $\mathbf{E}_t(\cdot)$ and $\mathbf{Var}_t(\cdot)$.

We assume that there is a pair of conintegrated securities denoted as P and Q. The price processes of P and Q are denoted by P_t and Q_t, and there is a tradable risk-free asset

Π whose price process is denoted by Π_t. Furthermore, we also assume that the market is frictionless, i.e., there are no transaction costs and taxes, and that short selling is allowed.

Assume that the dynamic of the price Q_t satisfies the following stochastic volatility model:

$$\begin{aligned} dQ_t &= \xi Q_t dt + \gamma(y_t) Q_t dW_1(t), \\ dy_t &= \alpha(y_t)dt + \beta(y_t)dW_2(t), \end{aligned} \tag{1}$$

where ξ is a constant. We assume that the spread of the log-prices of P and Q satisfies an OU process. Let $X_t = \ln(P_t) - \ln(Q_t)$ be the spread of the log-prices; then X_t satisfies the following SDE:

$$dX_t = \kappa(\theta - X_t)dt + \eta dW_3(t), \tag{2}$$

where κ, θ, and η are all constants. The dynamic of the risk-free asset Π_t is given by

$$d\Pi_t = r\Pi_t dt.$$

Remark 1. *Since $P_t = Q_t e^{X_t}$, according to Itô's formula, one can see that P_t satisfies the following SDE:*

$$\begin{aligned} dP_t = \ & P_t\left[\kappa(\theta - X_t) + \xi + \frac{1}{2}\eta^2 + \rho_{13}\eta\gamma(y_t)\right]dt \\ & + P_t\gamma(y_t)dW_1(t) + P_t\eta dW_3(t). \end{aligned} \tag{3}$$

We denote h_t as the weights invested in the securities P and Q at time t in a symmetric pairs trading strategy, and the corresponding wealth process V_t^h is given by

$$dV_t^h = V_t^h\left(h_t\frac{dP_t}{P_t} - h_t\frac{dQ_t}{Q_t} + \frac{d\Pi_t}{\Pi_t}\right). \tag{4}$$

Substituting (1) and (3) into (4), one can see that

$$dV_t^h = V_t^h h_t\left[(\kappa(\theta - X_t) + \frac{1}{2}\eta^2 + \rho_{13}\eta\gamma(y_t))dt + \eta dW_3(t)\right] + V_t^h r dt. \tag{5}$$

Let $\pi_t := e^{-rt}V_t^h h_t$ be the discounted money invested in the security P, which can be viewed as a strategy; then, the discounted wealth process $\overline{V}_t(\pi) := V_t^h e^{-rt}$ is given by

$$d\overline{V}_t(\pi) = \pi_t\left[(\kappa(\theta - X_t) + \frac{1}{2}\eta^2 + \rho_{13}\eta\gamma(y_t))\right]dt + \pi_t\eta dW_3(t).$$

Assume that the discounted wealth at time $t \in [0, T)$ is $\overline{V}_t$, then

$$\overline{V}_T(\pi) = \overline{V}_t + \int_t^T \pi_u\left(\kappa(\theta - X_u) + \frac{1}{2}\eta^2 + \rho_{13}\eta\gamma(y_u)\right)du + \int_t^T \pi_u\eta dW_3(u). \tag{6}$$

Let

$$J(t, \overline{V}_t; \pi) := \mathbf{E}_t\left(\overline{V}_T(\pi)\right) - \lambda\mathbf{Var}_t\left(\overline{V}_T(\pi)\right), \tag{7}$$

where $\lambda > 0$, we consider the following dynamic MV problem:

$$J(t, \overline{V}_t) = \sup_{\pi_u; u\in[t,T]} J(t, \overline{V}_t; \pi). \tag{8}$$

Because of the time-inconsistency of the MV problem (8), we introduce the optimal strategy according to the definition of equilibrium strategy provided in Björk et al. [13].

Definition 1. *The strategy $\pi^* = \{\pi^*_u; u \in [0, T]\}$ is called an **optimal strategy** if for any permutation $\pi^{\hat{\pi},\varepsilon}_u = \hat{\pi}_u I_{u\in[t,t+\varepsilon)} + \pi^*_u I_{u\in[t+\varepsilon,T]}$,*

$$\limsup_{\varepsilon\to 0} \frac{1}{\varepsilon}\Big\{J\big(t, \overline{V}_t(\pi^*); \pi^{\hat{\pi},\varepsilon}\big) - J\big(t, \overline{V}_t(\pi^*); \pi^*\big)\Big\} \le 0$$

holds for any $t \in [0, T]$.

We have the following theorem

Theorem 1 (Main result I). *Define $M(x, y) := \kappa(\theta - x) + \frac{1}{2}\eta^2 + \rho_{13}\eta\gamma(y)$. The optimal strategy for the dynamic MV problem (8) is given by*

$$\pi^*_t = \frac{1}{2\lambda\eta^2} M(X_t, y_t) - f_x(t, X_t, y_t) - \frac{\rho_{23}\beta(y_t)}{\eta} f_y(t, X_t, y_t), \tag{9}$$

where $f(t, x, y)$ is a solution to the following equation:

$$\begin{aligned} 0 = \; & \frac{1}{2\lambda\eta^2} M^2(x, y) + f_t(t, x, y) - f_x(t, x, y)\left(\frac{1}{2}\eta^2 + \rho_{13}\eta\gamma(y)\right) \\ & - \frac{\rho_{23}\beta(y_t)}{\eta} M(x, y) f_y(t, x, y) + \alpha(y) f_y(t, x, y) \\ & + \frac{1}{2}\eta^2 f_{xx}(t, x, y) + \frac{1}{2}\beta^2(y) f_{yy}(t, x, y) + \rho_{23}\eta\beta(y) f_{xy}(t, x, y), \end{aligned} \tag{10}$$

where $(x, y, t) \in \mathbf{R} \times \mathbf{R} \times (0, T]$ with the terminal condition $f(T, \cdot, \cdot) = 0$.

Lemma 1. *Let π^* be the strategy given in Theorem 1 and $f(t, x)$ be a solution of the Equation (10); then,*

$$f(t, X_t, y_t) = \mathbf{E}_t\left[\int_t^T \pi^*_u\left(\kappa(\theta - X_u) + \frac{1}{2}\eta^2 + \rho_{13}\eta\gamma(y_u)\right) du\right]. \tag{11}$$

Proof. Let $F_t : = f(t, X_t, y_t)$. It follows from Itô's formula that

$$\begin{aligned} dF_t = \; & \left[f_t + f_x\kappa(\theta - X_t) + f_y\alpha(y_t) + \frac{1}{2} f_{xx}\eta^2 + \frac{1}{2} f_{yy}\beta^2(y_t) + f_{xy}\eta\beta(y_t)\rho_{23}\right] dt \\ & + f_x\eta dW_3(t) + f_y\beta(y_t) dW_2(t). \end{aligned}$$

Let

$$\mathcal{M}_t : = F_t + \int_0^t \pi^*_u M(X_t, y_t) du,$$

then

$$\begin{aligned} d\mathcal{M}_t = \; & \pi^*_t M(X_t, y_t) dt + dF_t \\ = \; & \left[\frac{1}{2\lambda\eta^2} M(X_t, y_t)^2 - f_x(t, X_t, y_t)\Big(M(X_t, y_t) - \kappa(\theta - X_t)\Big)\right. \\ & - \frac{\rho_{23}\beta(y_t)}{\eta} f_y(t, X_t, y_t) M(X_t, y_t) + f_t(t, X_t, y_t) + f_y(t, X_t, y_t)\alpha(y_t) \\ & \left. + \frac{1}{2} f_{xx}(t, X_t, y_t)\eta^2 + \frac{1}{2} f_{yy}(t, X_t, y_t)\beta^2(y_t) + f_{xy}(t, X_t, y_t)\eta\beta(y_t)\rho_{23}\right] dt \\ & + f_x(t, X_t, y_t)\eta dW_3(t) + f_y(t, X_t, y_t)\beta(y_t) dW_2(t) \\ = \; & f_x(t, X_t, y_t)\eta dW_3(t) + f_y(t, X_t, y_t)\beta(y_t) dW_2(t). \end{aligned}$$

One can see that $\mathcal{M}_t$ is a martingale; thus,

$$F_t = f(t, X_t, y_t) = \mathbf{E}_t\left[\int_t^T \pi_u^* M(X_t, y_t)du\right],$$

which implies (11). □

Proof of Theorem 1. For the given $\varepsilon > 0$, let $\pi_u^{\hat{\pi},\varepsilon} : \ = \ \hat{\pi}_u I_{u\in[t,t+\varepsilon)} + \pi_u^* I_{u\in[t+\varepsilon,T]}$ be any permutation of π^*. For any strategy π, introduce

$$\Delta_\varepsilon \overline{V}_t(\pi) = \overline{V}_{t+\varepsilon}(\pi) - \overline{V}_t(\pi), \quad \Delta_\varepsilon f_t = f(t+\varepsilon, X_{t+\varepsilon}, y_{t+\varepsilon}) - f(t, X_t, y_t).$$

Since π^* is not dependent on the corresponding discounted wealth process $\overline{V}(\pi^*)$, it follows from (6) that

$$\begin{aligned}\overline{V}_T(\pi^{\hat{\pi},\varepsilon}) &= \overline{V}_t(\pi^*) + \{\overline{V}_{t+\varepsilon}(\hat{\pi}) - \overline{V}_t(\hat{\pi})\} + \{\overline{V}_T(\pi^*) - \overline{V}_{t+\varepsilon}(\pi^*)\} \\ &= \overline{V}_T(\pi^*) + \Delta_\varepsilon \overline{V}_t(\hat{\pi} - \pi^*).\end{aligned}$$

Because

$$\begin{aligned}&\mathbf{Var}_t\Big(\overline{V}_T(\pi^{\hat{\pi},\varepsilon})\Big) = \mathbf{E}_t\Big(\mathbf{Var}_{t+\varepsilon}\Big(\overline{V}_T(\pi^{\hat{\pi},\varepsilon})\Big)\Big) + \mathbf{Var}_t\Big(\mathbf{E}_{t+\varepsilon}\Big(\overline{V}_T(\pi^{\hat{\pi},\varepsilon})\Big)\Big) \\ &= \mathbf{E}_t\Big(\mathbf{Var}_{t+\varepsilon}\Big(\overline{V}_T(\pi^*)\Big)\Big) + \mathbf{Var}_t\Big(\overline{V}_{t+\varepsilon}(\pi^*) + \Delta_\varepsilon f_t + \Delta_\varepsilon \overline{V}_t(\hat{\pi} - \pi^*)\Big) \\ &= \mathbf{E}_t\Big(\mathbf{Var}_{t+\varepsilon}\Big(\overline{V}_T(\pi^*)\Big)\Big) + \mathbf{Var}_t\Big(\Delta_\varepsilon f_t + \Delta_\varepsilon \overline{V}_t(\hat{\pi})\Big),\end{aligned}$$

one can see that

$$\begin{aligned}&J\big(t, \overline{V}_t(\pi^*); \pi^{\hat{\pi},\varepsilon}\big) = \mathbf{E}_t\Big(\overline{V}_T(\pi^{\hat{\pi},\varepsilon})\Big) - \lambda \mathbf{Var}_t\Big(\overline{V}_T(\pi^{\hat{\pi},\varepsilon})\Big) \\ &= \mathbf{E}_t\Big(\overline{V}_T(\pi^*)\Big) + \mathbf{E}_t\Big(\Delta_\varepsilon \overline{V}_t(\hat{\pi} - \pi^*)\Big) - \lambda \mathbf{E}_t\Big(\mathbf{Var}_{t+\varepsilon}\Big(\overline{V}_T(\pi^*)\Big)\Big) \\ &\quad - \lambda \mathbf{Var}_t\Big(\Delta_\varepsilon f_t + \Delta_\varepsilon \overline{V}_t(\hat{\pi})\Big) \\ &= \mathbf{E}_t\Big(J(t+\varepsilon, \overline{V}_{t+\varepsilon}(\pi^*); \pi^*)\Big) + \mathbf{E}_t\Big(\Delta_\varepsilon \overline{V}_t(\hat{\pi} - \pi^*)\Big) - \lambda \mathbf{Var}_t\Big(\Delta_\varepsilon f_t + \Delta_\varepsilon \overline{V}_t(\hat{\pi})\Big).\end{aligned}$$

Furthermore, it follows from the proof of Lemma 1 and (6) that

$$\begin{aligned}\Delta_\varepsilon f_t + \Delta_\varepsilon \overline{V}_t(\hat{\pi}) &= \int_t^{t+\varepsilon} (\hat{\pi}_u - \pi_u^*) M(X_u, y_u) du \\ &\quad + \int_t^{t+\varepsilon} \Big\{\hat{\pi}_u + f_x(u, X_u, y_u)\Big\} \eta dW_3(u) + \int_t^{t+\varepsilon} f_y(u, X_u, y_u)\beta(y_u) dW_2(u),\end{aligned}$$

thus

$$\begin{aligned}&\lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\Big(J\big(t, \overline{V}_t(\pi^*); \pi^{\hat{\pi},\varepsilon}\big) - J\big(t, \overline{V}_t(\pi^*); \pi^*\big)\Big) \\ &= \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\Big[\mathbf{E}_t\Big(\Delta_\varepsilon \overline{V}_t(\hat{\pi} - \pi^*)\Big) - \lambda \mathbf{Var}_t\Big(\Delta_\varepsilon f_t + \Delta_\varepsilon \overline{V}_t(\hat{\pi})\Big) + \lambda \mathbf{Var}_t(\Delta_\varepsilon f_t + \Delta_\varepsilon \overline{V}_t(\pi^*))\Big] \\ &= (\hat{\pi}_t - \pi_t^*) M(X_t, y_t) - \lambda\Big(\hat{\pi}_t + f_x(t, X_t, y_t)\Big)^2 \eta^2 \\ &\quad - 2\lambda\Big(\hat{\pi}_t + f_x(t, X_t, y_t)\Big) f_y(t, X_t, y_t)\beta(y_t)\eta\rho_{23} - \lambda f_y(t, X_t, y_t)^2 \beta(y_t)^2 \\ &\quad + \lambda\Big(\pi_t^* + f_x(t, X_t, y_t)\Big)^2 \eta^2 \\ &\quad + 2\lambda\Big(\pi_t^* + f_x(t, X_t, y_t)\Big) f_y(t, X_t, y_u)\beta(y_u)\eta\rho_{23} + \lambda f_y(u, X_u, y_u)^2 \beta(y_u)^2 \\ &= -\lambda\eta^2(\hat{\pi}_t - \pi_t^*)^2 \leq 0,\end{aligned}$$

which completes the proof. □

Remark 2. *From the proof of Theorem 1, we can see that π_t^* is the solution of the following HJB equation:*

$$0 = \sup_{\pi_t} \left\{ \mathbf{E}_t(dJ(t, X_t, y_t)) - \lambda \mathbf{Var}_t\Big(df(t, X_t, y_t) + d\bar{V}_t(\pi_t)\Big) \right\}.$$

Remark 3. *It is unlikely that a closed-form solution of the PDE (10) can be achieved without specifying $\alpha, \beta,$ and γ. Therefore, we will consider the Scott model, one of the most widely known mean-reverting stochastic volatility models, and will discuss the approximate solution.*

3. Closed-Form Approximation under the Scott Model

In order to capture the fast mean-reverting characteristics of volatility (see Fouque [19,20] for empirical studies), we introduce the Scott model, initially proposed by Scott (1987) [18], which is a well-known mean-reverting volatility model. Under the Scott model, the underlying security price is modeled by:

$$\begin{aligned} dQ_t &= \xi Q_t dt + e^{y_t} Q_t dW_1(t), \\ dy_t &= a(b - y_t)dt + \sigma dW_2(t), \end{aligned}$$

where $a > 0$ and $\sigma > 0$. Since $\alpha(y) = a(b-y)$, $\beta(y) = \sigma$, $\gamma(y) = e^y$. Clearly, the volatility is modeled by an OU process. We assume $a \gg 1$ to ensure the fast mean-reverting property and set $\rho = 0$ for convenience. According to Theorem 1, the corresponding PDE for the optimal strategy is given by:

$$\begin{aligned} 0 = &\ \frac{1}{2\lambda\eta^2} M^2(x, y) + f_t(t, x, y) - f_x(t, x, y)\left(\frac{1}{2}\eta^2 + \rho_{13}\eta e^y\right) \\ &+ a(b - y) f_y(t, x, y) + \frac{1}{2}\eta^2 f_{xx}(t, x, y) + \frac{1}{2}\sigma^2 f_{yy}(t, x, y), \end{aligned} \tag{12}$$

where $(x, y, t) \in \mathbf{R} \times \mathbf{R} \times (0, T]$ with the terminal condition $f(T, \cdot, \cdot) = 0$. It is difficult to solve the PDE (12) explicitly and to achieve a closed-form optimal strategy. Therefore, in this section, we will use the asymptotic analysis technique to find a closed-form approximation of the PDE (12).

Let $g(t, x, y) := f_x(t, x, y)$; one can see from (12) that $g(t, x, y)$ satisfies the following PDE:

$$\begin{aligned} 0 = &\ g_t(t, x, y) + \frac{1}{2}\eta^2 g_{xx}(t, x, y) + \frac{1}{2}\sigma^2 g_{yy}(t, x, y) \\ &- g_x(t, x, y)\left(\frac{1}{2}\eta^2 + \rho_{13}\eta e^y\right) + a(b - y) g_y(t, x, y) - \frac{\kappa}{\lambda\eta^2} M(x, y). \end{aligned} \tag{13}$$

Let $\epsilon = a^{-1}$ and $\nu^2 = \frac{\epsilon}{2}\sigma^2$. Since $a \gg 1$, it is reasonable to assume that $0 < \nu < 1$. We introduce the following operator

$$\mathcal{L}^\epsilon = \frac{1}{\epsilon}\mathcal{L}_0 + \mathcal{L}_1,$$

where

$$\begin{aligned} \mathcal{L}_0 &= \nu^2 \frac{\partial^2}{\partial y^2} + (b - y)\frac{\partial}{\partial y}, \\ \mathcal{L}_1 &= \frac{\partial}{\partial t} + \frac{1}{2}\eta^2 \frac{\partial^2}{\partial x^2} - \left(\frac{1}{2}\eta^2 + \rho_{13}\eta e^y\right)\frac{\partial}{\partial x}. \end{aligned}$$

Then the PDE (13) can be written in the following form

$$\mathcal{L}^\epsilon(g) = \frac{\kappa}{\lambda\eta^2} M(x, y), \qquad g(T, \cdot, \cdot) = 0. \tag{14}$$

To provide an approximation for $g(t,x,y)$, we need to introduce $\phi(y) : = \frac{1}{\sqrt{2\pi v}} e^{-\frac{(y-b)^2}{2v^2}}$ and define the functional $\Psi(\cdot)$ as

$$\Psi(h) := \int_{-\infty}^{+\infty} \phi(y)h(y)dy$$

for all h such that $\Psi(|h|) < +\infty$. The following lemma can be found in Fouque [19].

Lemma 2. *The following equation*

$$\mathcal{L}_0 u(t,x,y) = h(t,x,y)$$

only has a solution if $h(t,x,y)$ satisfies

$$\Psi_{t,x}(h) : = \int_{-\infty}^{+\infty} \phi(y)h(t,x,y)dy = 0,$$

for each $t \in [0,T]$, $x \in (-\infty,+\infty)$.

3.1. An Approximation for $g(t,x,y)$

Assume that g is a solution of (14). We can construct an asymptotic expansion with respect to $\sqrt{\epsilon}$ for g as the following

$$g(t,x,y) = g_0(t,x,y) + \sqrt{\epsilon} g_1(t,x,y) + \epsilon g_2(t,x,y) + \cdots .$$

Substituting it into (14), we obtain

$$\begin{aligned} &\frac{\kappa}{\lambda\eta^2} M(x,y) = \mathcal{L}^\epsilon g(t,x,y) \\ = &\frac{1}{\epsilon}\mathcal{L}_0 g_0(t,x,y) + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_0 g_1(t,x,y) + (\mathcal{L}_0 g_2(t,x,y) + \mathcal{L}_1 g_0(t,x,y)) \\ &+ \sqrt{\epsilon}(\mathcal{L}_0 g_3(t,x,y) + \mathcal{L}_1 g_1(t,x,y)) + \cdots . \end{aligned} \tag{15}$$

For simplicity, we assume that $g_0(t,x,y) \equiv g_0(t,x)$, $g_1(t,x,y) \equiv g_1(t,x)$. Since the operator $\mathcal{L}_0$ only involves partial derivatives with respect to y, one can see that

$$\begin{aligned} \mathcal{L}_0 g_0(t,x,y) = 0, \\ \mathcal{L}_0 g_1(t,x,y) = 0. \end{aligned}$$

Then, Equation (15) can be simplified into the following form

$$\begin{aligned} \frac{\kappa}{\lambda\eta^2} M(x,y) = &\ (\mathcal{L}_0 g_2(t,x,y) + \mathcal{L}_1 g_0(t,x)) \\ &+ \sqrt{\epsilon}(\mathcal{L}_0 g_3(t,x,y) + \mathcal{L}_1 g_1(t,x)) + \cdots . \end{aligned} \tag{16}$$

It is natural to consider the following equations

$$\mathcal{L}_0 g_2(t,x,y) = -\mathcal{L}_1 g_0(t,x) + \frac{\kappa}{\lambda\eta^2} M(x,y), \tag{17}$$

$$\mathcal{L}_0 g_3(t,x,y) = -\mathcal{L}_1 g_1(t,x). \tag{18}$$

To make sure the Equation (17) has a solution, it follows from Lemma 2 that for each (t,x), the right part of (17) should satisfy the following equation

$$\Psi_{t,x}\left(\mathcal{L}_1 g_0(t,x) - \frac{\kappa}{\lambda\eta^2} M(x,y)\right) = 0. \tag{19}$$

Since $\mathcal{L}_1 g_0(t,x) = \frac{\partial}{\partial t} g_0(t,x) + \frac{1}{2}\eta^2 \frac{\partial^2}{\partial x^2} g_0(t,x) - (\frac{1}{2}\eta^2 + \rho_{13}\eta e^y)\frac{\partial}{\partial x} g_0(t,x)$, if we let $c: = \frac{1}{2}\eta^2 + \eta\rho_{13}\Psi(e^y)$, one can see that

$$\Psi_{t,x}(\mathcal{L}_1 g_0(t,x)) = \frac{\partial}{\partial t} g_0(t,x) + \frac{1}{2}\eta^2 \frac{\partial^2}{\partial x^2} g_0(t,x) - c g_0(t,x).$$

Furthermore, from the definition of $M(x,y)$, one can see that

$$\begin{aligned} \Psi_{t,x}(M(x,y)) =& \ \Psi_{t,x}\left(\kappa(\theta - x) + \frac{1}{2}\eta^2 + \rho_{13}\eta e^y\right) \\ =& \ \kappa(\theta - x) + c, \end{aligned}$$

which is a linear function of x. Assigning the boundary value of $g_0(t,x)$ as $g_0(T,\cdot) = 0$, the solution can be directly given by

$$g_0(t,x) = \frac{\kappa^2}{\lambda\eta^2}(t-T)(\theta - x) - \frac{c\kappa^2}{2\lambda\eta^2}(t-T)^2 + \frac{c\kappa}{\lambda\eta^2}(t-T). \tag{20}$$

Similarly, to ensure that Equation (18) has a solution, we need

$$\Psi_{t,x}(\mathcal{L}_1 g_1(t,x)) = \frac{\partial}{\partial t} g_1(t,x) + \frac{1}{2}\eta^2 \frac{\partial^2}{\partial x^2} g_1(t,x) - c g_1(t,x) = 0.$$

With the bound condition $g_1(T,\cdot) = 0$, one can see that

$$g_1(t,x) = 0$$

is a solution. Therefore, a natural approximation for $g(t,x,y)$ is given by

$$g^*(t,x,y) = \ g_0(t,x) + \sqrt{\epsilon} g_1(t,x).$$

Thus, we have the following theorem:

Theorem 2 (Main result II). *Denote* $c: = \frac{1}{2}\eta^2 + \rho_{13}\eta\Psi(e^y)$. *Let*

$$g^*(t,x,y) = \frac{\kappa^2}{\lambda\eta^2}(t-T)(\theta - x) - \frac{c\kappa^2}{2\lambda\eta^2}(t-T)^2 + \frac{c\kappa}{\lambda\eta^2}(t-T). \tag{21}$$

Then, there exists a constant C that is independent of ϵ *such that*

$$|g(t,x,y) - g^*(t,x,y)| \le \sqrt{\epsilon} C(1 + e^{|y|}). \tag{22}$$

An approximate optimal strategy is given by

$$\hat{\pi}_t^* = \ \frac{1}{2\lambda\eta^2}\kappa(\theta - X_t) + \frac{1}{4\lambda} + \frac{1}{2\lambda\eta}\rho_{13} e^{y_t} - g^*(t, X_t, y_t).$$

Remark 4. *Compared with the traditional finite difference method(FDM), we can obtain an explicit solution with a significant advantage in computational complexity using the asymptotic analysis technique. In practice, at each decision moment, we only need to calculate the optimal strategy at one specific point* (x,y,t)*, which describes the market state at that time, rather than all values on a series of grid points. In addition, since our problem is defined on an unbounded domain, in order to apply the FDM, the solution area must be cutoff from infinity and additional artificial boundary conditions must be added. This will introduce additional boundary errors and increase the complexity of the theoretical analysis and numerical calculations.*

Remark 5. *From the error estimation given in Theorem 2, we can see that the accuracy of our closed-form approximation of the optimal strategy is mainly controlled by volatility mean-reversion speed parameter a. Faster mean-reversion speed implies a more accurate strategy.*

3.2. Some Auxiliary Approximations

We established the following lemmas to helping prove Theorem 2.

Lemma 3. (i) *If $y < \nu^2 + b$, then*

$$\frac{1}{\phi(y)}\int_{-\infty}^{y}\phi(z)dz \le \sqrt{2\pi}\nu e^{\frac{1}{2}}, \tag{23}$$

$$\frac{1}{\phi(y)}\int_{-\infty}^{y}e^{z}\phi(z)dz \le \sqrt{2\pi}\nu e^{y}. \tag{24}$$

(ii) *If $y \ge \nu^2 + b$, then*

$$\frac{1}{\phi(y)}\int_{y}^{+\infty}e^{z}\phi(z)dz \le \sqrt{2\pi}\nu e^{y}. \tag{25}$$

Lemma 4. *Given $h(y)$ that satisfies $\Psi(h) = 0$ and $|h(y)| \le C_1(1+e^{y})$, for some positive constants C_1, let $\mathcal{X}(y)$ satisfy the following equation*

$$\mathcal{L}_0\mathcal{X}(y) = h(y).$$

Then, $|\mathcal{X}(y)| < C_2(1+e^{|y|})\epsilon^{-\frac{1}{2}}$ for some positive constant C_2 independent of ϵ.

This lemma provides an estimation for the linear operator $\mathcal{L}_0$. If we choose $C_1 = 0$ in Lemma 4, by using (A1), we obtained the following corollary:

Corollary 1. *Let $\mathcal{X}(y)$ be the solution of the equation $\mathcal{L}_0\mathcal{X}(y) = 0$; then, there exists a positive constant C_2 independent of ϵ such that*

$$|\mathcal{X}(y)| \le C_2.$$

Lemma 5. *Let $\{y_t; t \in [0,T]\}$ be the OU process in the Scott model; then, for all $\tau \ge 0$, there exists a positive constant $\hat{C}$ independent of ϵ such that*

$$\mathbf{E}\Big[\, e^{|y_{t+\tau}|} \,\Big|\, y_t = y \,\Big] \le \hat{C}e^{|y|}. \tag{26}$$

Proofs of Lemmas 3–5 are given in the Appendices A–C.

3.3. The Proof of Theorem 2

We only need to consider error estimation. The residue portion can be defined as

$$R_\epsilon(t,x,y) := g_0(t,x) + \sqrt{\epsilon}g_1(t,x) + \epsilon g_2(t,x,y) + \epsilon\sqrt{\epsilon}g_3(t,x,y) - g(t,x,y).$$

Recalling $\mathcal{L}^{\epsilon}g(t,x,y) = \frac{\kappa}{\lambda\eta^2}M(x,y)$ and

$$\begin{aligned}
\mathcal{L}_0 g_0(t,x) =&\ 0,\\
\mathcal{L}_0 g_1(t,x) =&\ 0,\\
\mathcal{L}_0 g_2(t,x,y) =&\ -\mathcal{L}_1 g_0(t,x) + \frac{\kappa}{\lambda\eta^2}M(x,y),\\
\mathcal{L}_0 g_3(t,x,y) =&\ -\mathcal{L}_1 g_1(t,x),
\end{aligned} \tag{27}$$

we have

$$\begin{aligned}\mathcal{L}^{\epsilon}R_{\epsilon}(t,x,y) = &\ \frac{1}{\epsilon}\mathcal{L}_0 g_0(t,x) + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_0 g_1(t,x) + \{\mathcal{L}_0 g_2(t,x,y) + \mathcal{L}_1 g_0(t,x)\} \\ &+ \sqrt{\epsilon}\{\mathcal{L}_0 g_3(t,x,y) + \mathcal{L}_1 g_1(t,x)\} \\ &+ \epsilon\mathcal{L}_1 g_2(t,x,y) + \epsilon\sqrt{\epsilon}\mathcal{L}_1 g_3(t,x,y) - \mathcal{L}^{\epsilon} g(t,x,y) \\ = &\ \epsilon(\mathcal{L}_1 g_2(t,x,y) + \sqrt{\epsilon}\mathcal{L}_1 g_3(t,x,y)).\end{aligned}$$

Defining

$$\begin{aligned}G_{\epsilon}(t,x,y) = &\ \mathcal{L}_1 g_2(t,x,y) + \sqrt{\epsilon}\mathcal{L}_1 g_3(t,x,y), \\ H_{\epsilon}(x,y) = &\ g_2(T,x,y) + \sqrt{\epsilon} g_3(T,x,y).\end{aligned}$$

One can see that $R_{\epsilon}(t,x,y)$ solves following equation:

$$\begin{cases}\mathcal{L}^{\epsilon}R_{\epsilon}(t,x,y) = \epsilon G_{\epsilon}(t,x,y), \\ R_{\epsilon}(T,\cdot,\cdot) = \epsilon H_{\epsilon}(\cdot,\cdot).\end{cases}$$

Applying Feymann–Kac formula, R_{ϵ} demonstrates probabilistic representation as follows:

$$R_{\epsilon}(t,x,y) = \epsilon\mathbf{E}_t\left[H_{\epsilon}(\hat{X}_T, y_T) + \int_t^T G_{\epsilon}(s,\hat{X}_s,y_s)ds \;\middle|\; \hat{X}_t = x, y_t = y\right], \tag{28}$$

where $\hat{X}_t$ is driven by

$$d\hat{X}_t = -\left(\frac{1}{2}\eta^2 + \rho_{13}\eta e^{y_t}\right)dt + \eta dW_3(t).$$

From (28), one can see that the boundary of R_{ϵ} is controlled by both G_{ϵ} and H_{ϵ}. In the following sections, we provide estimations for these two parts.

Let $\psi(y)$ solves

$$\mathcal{L}_0\psi(y) = e^y - \Psi(e^y).$$

According to Lemma 4,

$$|\psi(y)| \le C_2(e^{|y|}+1)\epsilon^{-\frac{1}{2}} \tag{29}$$

holds for some positive constants C_2. We choose

$$g_2(t,x,y) := \frac{\kappa\rho_{13}}{\lambda\eta}[\kappa(T-t)+1]\psi(y), \tag{30}$$

$$g_3(t,x,y) := 0. \tag{31}$$

One can see from (19) that

$$\begin{aligned}&\mathcal{L}_0 g_2(t,x,y) = \frac{\kappa\rho_{13}}{\lambda\eta}[\kappa(T-t)+1](e^y - \Psi(e^y))] \\ &= -\{\mathcal{L}_1 g_0(t,x) - \Psi_{t,x}(\mathcal{L}_1 g_0(t,x))\} + \frac{\kappa}{\lambda\eta^2}\{M(x,y) - \Psi_{t,x}(M(x,y))\} \\ &= -\mathcal{L}_1 g_0(t,x) + \frac{\kappa}{\lambda\eta^2}M(x,y), \\ &\mathcal{L}_0 g_3(t,x,y) = 0.\end{aligned}$$

Thus, $g_0(t,x)$, $g_1(t,x)$, $g_2(t,x,y)$, and $g_3(t,x,y)$ satisfy Equation (27). Furthermore, one can easily see that there exists a constant C^* such that

$$\begin{aligned}&|g_2(t,x,y)| \le C^*(e^{|y|}+1)\epsilon^{-\frac{1}{2}}, \\ &|G_{\epsilon}(t,x,y)| = \left|\mathcal{L}_1 g_2(t,x,y) + \sqrt{\epsilon}\mathcal{L}_1 g_3(t,x,y)\right|\end{aligned}$$

$$\leq C^*\left(e^{|y|}+1\right)\epsilon^{-\frac{1}{2}},$$
$$|H_\epsilon(x,y)| \leq C^*(e^{|y|}+1)\epsilon^{-\frac{1}{2}}.$$

Using Lemma 5, one can see that

$$\begin{aligned}
|R_\epsilon(t,x,y)| =\ & \epsilon\left|\mathbf{E}\left[H_\epsilon(\hat{X}_T,y_T)+\int_t^T G_\epsilon(s,\hat{X}_s,y_s)ds \;\middle|\; \hat{X}_t=x,y_t=y\right]\right| \\
\leq & C^*\epsilon^{\frac{1}{2}}\mathbf{E}\left[e^{|y_T|}+1+\int_t^T\left(e^{|y_s|}+1\right)ds \;\middle|\; y_t=y\right] \\
\leq & C^*\epsilon^{\frac{1}{2}}\left(\mathbf{E}_t\left[e^{|y_{t+(T-t)}|}\right]+T+1\right)+C^*\epsilon^{\frac{1}{2}}\int_0^{T-t}\mathbf{E}_t\left[e^{|y_{t+\tau}|}\right]d\tau \\
\leq & C^*\epsilon^{\frac{1}{2}}\left(\hat{C}e^{|y|}+T+1\right)+C^*\epsilon^{\frac{1}{2}}T\hat{C}e^{|y|} \\
\leq & \left[C^*(\hat{C}+1)(T+1)\right]\left(e^{|y|}+1\right)\epsilon^{\frac{1}{2}}.
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
|g_0(t,x)-g(t,x,y)| =\ & \left|R_\epsilon(t,x,y)-\epsilon g_2(t,x,y)-\epsilon\sqrt{\epsilon}g_3(t,x,y)\right| \\
\leq & |R_\epsilon(t,x,y)|+\epsilon|g_2(t,x,y)|+\epsilon\sqrt{\epsilon}|g_3(t,x,y)| \\
\leq & \left[C^*(\hat{C}+1)(T+1)+C^*\right]\left(e^{|y|}+1\right)\epsilon^{\frac{1}{2}}+C_3\epsilon^{\frac{3}{2}}.
\end{aligned}$$

Denote

$$C = C^*\left[(\hat{C}+1)(T+1)+1\right],$$

then (22) follows. The approximate optimal strategy is given by directly applying Theorem 1.

4. Empirical Experiments

In this section, we compare the effect of our strategy (the optimal approximate strategy in the Scott model) and the strategy proposed by Zhu et al. (the optimal strategy in the constant volatility model (see Zhu et al. [12])) on both real scenarios and simulated scenarios. We select three stock pairs (listed in Table 1) traded on the Chinese security markets SSE and SZSE to illustrate our results using the standard cointegration testing method mentioned by Chambers [21]. For the estimation of the Scott model, we combine the maximum likelihood estimation(MLE) with the extended Kalman filter to produce an on-line updated estimation (see Wang et al. [22] and Simon [23] for details). We also recommended Aihara [24] for an alternative robust filtering estimation. Then, we empirically validate the strategies given in Section 3 based on the real market data from the Chinese security markets SSE and SZSE.

Table 1. Selected stock pairs.

Label	Stock P	Stock Q	Industry
1	Cecep Wind Power	Cecep Solar Energy	Clean Energy Industry
2	Shanxi Lu'An Environmental Energy	Shanxi Coking Coal Energy	Coal Industry
3	Haitong Securities	Citic Securities	Security Industry

The sample period is from 1 June 2019 to 1 December 2022 across different industries, but the stocks in each pair are in the same industry and are highly correlated in terms of both fundamentals and price series. The data were obtained from TDX software, and we only used the daily closing prices. Typically, we used the forward-adjusted prices to avoid the dividend effect. Figure 1 presents the forward-adjusted stock prices and the dynamics of the corresponding price spread for the first pair as an example.

Figure 1. Stock prices and price spread dynamics for Cecep Wind Power and Cecep Solar Energy.

Starting from 1 December 2021, we performed out-of-sample testing for all cases using the moving-window method. Parameters were updated everyday by using data from last 375 trading days (on and a half years). We chose $\lambda = 0.5$ and equally allocate the initial endowment of 100 units among the selected pairs. The paths of the wealth processes obtained from our strategy under the Scott model and from the strategy proposed by Zhu et al. are shown in Figure 2, respectively. Table 2 presents some of the commonly used statistics denoting strategy performance.

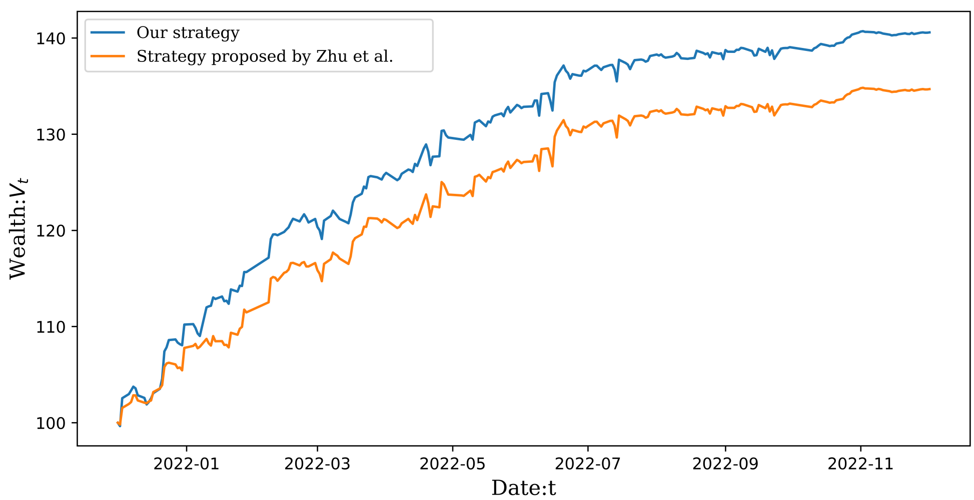

Figure 2. The wealth dynamics of the out-of-sample testing. Strategy proposed by Zhu et al. [12].

Table 2. Statistics for the out-of-sample testing.

Model	Win Rate	Profit-Loss Ratio	Average Profit	Max Drawdown	Sharp Ratio
Our strategy	56.967%	1.762	0.140%	−2.136%	3.582
Strategy proposed by Zhu et al. [12]	55.328%	1.670	0.122%	−1.909%	3.209

Figure 2 and Table 2 indicate that our strategy under the Scott model outperforms the strategy proposed by Zhu et al. under the constant volatility model with respect to important indicators such as the Sharp ratio, the profit-loss ratio and the win rate in the out-of-sample testing.

To further compare the value of the mean-variance objective function J for these two strategies, we implemented simulations with respect to the parameters estimated using real market data for the stock pairs listed in Table 1. Parameters for the simulation are given in Table 3. As for the strategy proposed by Zhu et al., we used $e^{|b|}$ as the constant volatility, which is the long-term average volatility in the Scott model.

Table 3. Parameters for simulation.

Label	κ	θ	η	ξ	a	b	σ	ρ_{13}	y_0
1	11.633	−0.233	0.377	1.206	5.324	−0.695	2.826	−0.338	−0.554
2	5.183	−0.173	0.366	0.792	4.109	−0.796	2.986	−0.348	−0.876
3	7.961	−0.244	0.205	−0.022	9.987	−1.246	3.235	−0.398	−1.314

We chose $T = 1.5, r = 0.03, dt = 1/250, X_0 = 0, V_0 = 100$, and simulated each pair 1000 times with risk aversion λ values ranging from 0.5 to 1.5. The statistics of the discounted terminal wealth for each pair are shown in Tables 4–6. J is the value of the objective function defined in (7).

Table 4. Cecep Wind Power and Cecep Solar Energy.

λ	Strategy Proposed by Zhu et al. [12]			Our Strategy		
	Mean	**S.D.**	**J**	**Mean**	**S.D.**	**J**
0.25	767.267	43.255	299.519	548.729	27.084	365.340
0.5	435.940	21.627	202.065	326.670	13.542	234.976
0.75	325.497	14.418	169.581	252.651	9.028	191.521
1.0	270.276	10.814	153.339	215.641	6.771	169.794
1.25	237.143	8.651	143.593	193.435	5.417	156.758
1.5	215.055	7.209	137.096	178.631	4.514	148.067

Table 5. Shanxi Lu'An Environmental Energy and Shanxi Coking Coal Energy.

λ	Strategy Proposed by Zhu et al. [12]			Our Strategy		
	Mean	**S.D.**	**J**	**Mean**	**S.D.**	**J**
0.25	220.566	18.173	138.004	181.554	8.569	163.198
0.5	162.589	9.086	121.308	143.083	4.284	133.905
0.75	143.263	6.058	115.743	130.259	2.856	124.141
1.0	133.600	4.543	112.960	123.847	2.142	119.258
1.25	127.803	3.635	111.290	120.000	1.714	116.329
1.5	123.938	3.029	110.177	117.436	1.428	114.376

Table 6. Haitong Securities and Citic Securities.

λ	Strategy Proposed by Zhu et al. [12]			Our Strategy		
	Mean	**S.D.**	**J**	**Mean**	**S.D.**	**J**
0.25	981.336	50.993	331.271	518.334	17.535	441.463
0.5	542.969	25.496	217.937	311.468	8.768	273.033
0.75	396.847	16.998	180.158	242.513	5.845	216.889
1.0	323.786	12.748	161.269	208.035	4.384	188.818
1.25	279.949	10.199	149.936	187.349	3.507	171.975
1.5	250.725	8.499	142.380	173.558	2.923	160.746

Tables 4–6 indicate the effectiveness of both strategies by comparing the average terminal wealth with the initial asset $V_0 = 100$. Both strategies yield a discounted final wealth greater than V_0 in every case, which means that profit is always higher than the risk-free return. Furthermore, comparing the strategy statistics of the strategy proposed by Zhu et al. and our strategy in each case, we found that the standard deviation of the terminal wealth of our strategy is always smaller than that of the strategy proposed by Zhu et al. Although the mean may be lower, ultimately, the J value of our strategy is always greater than that of the strategy proposed by Zhu et al. This phenomenon suggests that the approximate optimal strategy under the Scott model outperforms the optimal strategy under the constant volatility strategy, by producing more stable profits. It is noteworthy that λ plays a critical role in controlling the uncertainty of the outcome result of both strategies. The mean and the standard deviation of the terminal wealth decrease as λ increases. Intuitively, a larger λ indicates more risk aversion, which leads to smaller allocation on the risky assets and thus lower uncertainty.

Clearly, our strategies show effectiveness for both simulated and real out-of-sample data. The comparison of the strategy proposed by Zhu et al. and our approximate strategy shows that the Scott model can better capture the mean-reverting characteristic of volatility, resulting in a more stable trading strategy.

5. Conclusions

In this paper, we provide a semi-closed-form optimal strategy of the mean-variance problem for pairs trading by assuming that one of the security prices satisfies a general stochastic volatility model and that the corresponding price spread follows the Ornstein–Uhlenbeck process. Then, we provide a closed-form approximate formula for the optimal strategy in the Scott model using the asymptotic analysis technique. Our approximate formula has extremely high computational efficiency and has been proven to be accurate. We implemented our approximate optimal strategy on the real historical data selected from Chinese security markets and compared it with the optimal strategy under the constant volatility model proposed by Zhu et al. [12]. The numerical results show that both strategies are effective and that the Scott model produces a more stable strategy by better capturing mean-reverting volatility.

Author Contributions: Conceptualization, Y.Z.; methodology, Y.Z. and D.X.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.Z. and D.X.; visualization, Y.Z.; supervision, D.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China under Grant No. 11671257.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proof of Lemma 3

(i) Let
$$c_0(y) := \frac{y-b}{\sqrt{2}v}.$$
One can see that $c_0(y) \le \frac{\nu}{\sqrt{2}}$. Through direct computation, we have
$$\frac{1}{\phi(y)}\int_{-\infty}^{y}\phi(z)dz = e^{\frac{(y-b)^2}{2v^2}}\int_{-\infty}^{y} e^{-\frac{(z-b)^2}{2v^2}}dz = \sqrt{2}\nu e^{c_0^2(y)}\int_{-\infty}^{c_0(y)} e^{-x^2}dx.$$
If $c_0(y) \le 0$, then
$$\begin{aligned}\frac{1}{\phi(y)}\int_{-\infty}^{y}\phi(z)dz = &\ \sqrt{2}\nu e^{c_0^2(y)}\int_{-\infty}^{0} e^{-(u+c_0^2(y))}du\\ \le &\sqrt{2}\nu\int_{-\infty}^{0} e^{-u^2}du = \frac{\sqrt{2\pi}}{2}\nu;\end{aligned}$$
if $c_0(y) > 0$, then
$$\frac{1}{\phi(y)}\int_{-\infty}^{y}\phi(z)dz \le \sqrt{2}\nu e^{c_0(y)^2}\int_{-\infty}^{+\infty} e^{-x^2}dx \le \sqrt{2\pi}\nu e^{\frac{\nu^2}{2}}.$$
Recalling that $0 < \nu < 1$, (23) follows. Furthermore, if we introduce
$$c_1(y) = \frac{y-b-\nu^2}{\sqrt{2}\nu},$$
then
$$\begin{aligned}\frac{1}{\phi(y)}\int_{-\infty}^{y} e^z\phi(z)dz = &\ e^{\frac{(y-b)^2}{2v^2}}\int_{-\infty}^{y} e^{z-\frac{(z-b)^2}{2v^2}}dz\\ = &\ e^{\frac{(y-b)^2}{2v^2}+b+\frac{1}{2}\nu^2}\int_{-\infty}^{y} e^{-\frac{(z-b-\nu^2)^2}{2v^2}}dz\\ = &\ \frac{\sqrt{2}}{\nu}e^{\frac{(y-b)^2}{2v^2}+b+\frac{1}{2}\nu^2}\int_{-\infty}^{c(y)} e^{-x^2}dx\\ = &\ \sqrt{2}\nu e^{y+c_1(y)^2}\int_{0}^{+\infty} e^{-(x-c_1(y))^2}dx\\ \le &\sqrt{2}\nu e^{y+c_1(y)^2}\int_{0}^{+\infty} e^{-(x^2+c_1(y)^2)}dx\\ = &\ \sqrt{2\pi}\nu e^{y},\end{aligned}$$
and (24) follows.

(ii) If $y \ge b+\nu^2$, then $c_1(y) \ge 0$, one can see that
$$\begin{aligned}\frac{1}{\phi(y)}\int_{y}^{+\infty} e^z\phi(z)dz = &\ \sqrt{2}\nu e^{y+c_1(y)^2}\int_{0}^{+\infty} e^{-(x+c_1(y))^2}dx\\ \le &\sqrt{2}\nu e^{y+c_1(y)^2}\int_{0}^{\infty} e^{-(x^2+c_1(y)^2)}dx\\ = &\ \sqrt{2\pi}\nu e^{y},\end{aligned}$$
and (25) follows.

Appendix B. Proof of Lemma 4

One can easily see that

$$\mathcal{L}_0\mathcal{X}(y) = \frac{\nu^2}{\phi(y)}\frac{\partial}{\partial y}\left(\phi(y)\frac{\partial \mathcal{X}}{\partial y}\right) = h(y),$$

thus

$$\mathcal{X}_y(y) = \frac{1}{\nu^2\phi(y)}\int_{-\infty}^{y} h(z)\phi(z)dz = -\frac{1}{\nu^2\phi(y)}\int_{y}^{+\infty} h(z)\phi(z)dz. \tag{A1}$$

(i) If $y \geq \nu^2 + b$, we can see from Lemma 3 (ii) that

$$\begin{aligned}
\left|\mathcal{X}_y(y)\right| = & \frac{1}{\nu^2\phi(y)}\left|\int_{y}^{+\infty} h(z)\phi(z)dz\right| \leq \frac{C_1}{\nu^2\phi(y)}\int_{y}^{+\infty}(1+e^z)\phi(z)dz \\
\leq & \frac{C_1}{\nu^2\phi(y)}e^{-b-\nu^2}\int_{y}^{+\infty}(e^z + e^{z+b+\nu^2})\phi(z)dz \\
\leq & \frac{C_1\left(1+e^{-b}\right)}{\nu^2\phi(y)}\int_{y}^{\infty} e^z\phi(z)dz \\
\leq & \frac{C_1}{\nu}\sqrt{2\pi}\left(1+e^{-b}\right)e^y.
\end{aligned}$$

Since

$$\mathcal{X}(y) = \int_{\nu^2+b}^{y}\mathcal{X}_y(z)dz + \mathcal{X}(\nu^2+b),$$

one can see that:

$$\begin{aligned}
|\mathcal{X}(y)| \leq & \int_{\nu^2+b}^{y}\left|\mathcal{X}_y(s)\right|ds + \left|\mathcal{X}(\nu^2+b)\right| \\
\leq & \frac{C_1}{\nu}\sqrt{2\pi}\left(1+e^{-b}\right)\int_{\nu^2+b}^{y} e^z dz + \left|\mathcal{X}(\nu^2+b)\right| \\
= & \frac{C_1}{\nu}\sqrt{2\pi}\left(1+e^{-b}\right)\left(e^y - e^{\nu^2+b}\right) + \left|\mathcal{X}(\nu^2+b)\right| \\
\leq & \sqrt{2}\sigma^{-1}\epsilon^{-\frac{1}{2}}\left[C_1\sqrt{2\pi}(1+e^{-b}) + |\mathcal{X}(\nu^2+b)|\right](e^y+1) \\
\leq & \hat{C}_0\epsilon^{-\frac{1}{2}}(e^{|y|}+1),
\end{aligned} \tag{A2}$$

where $\hat{C}_0 = \sigma^{-1}\left[C_1\sqrt{2\pi}(1+e^{-b}) + |\mathcal{X}(\nu^2+b)|\right]$.

(ii) If $y < \nu^2 + b$, from Lemma 3 (i), we can see that

$$\begin{aligned}
\left|\mathcal{X}_y(y)\right| \leq & \frac{C_1}{\nu^2\phi(y)}\int_{-\infty}^{y}(1+e^z)\phi(z)dz \\
= & \frac{C_1}{\nu^2\phi(y)}\int_{-\infty}^{y} e^z\phi(z)dz + \frac{C_1}{\nu^2\phi(y)}\int_{-\infty}^{y}\phi(z)dz \\
\leq & \frac{C_1}{\nu}\sqrt{2\pi}e^y + \frac{C_1}{\nu}\sqrt{2\pi}e^{\frac{1}{2}} \\
\leq & \frac{C_1}{\nu}\sqrt{2\pi}\left(e^{|y|} + e^{\frac{1}{2}+|y|}\right) \\
\leq & \frac{C_1}{\nu}\left(1+e^{\frac{1}{2}}\right)e^{|y|}.
\end{aligned}$$

Since

$$\mathcal{X}(y) = -\int_{\nu^2+b}^{y}\mathcal{X}_y(z)dz + \mathcal{X}(\nu^2+b),$$

one can see that:

$$
\begin{aligned}
|\mathcal{X}(y)| \leq & \int_y^{\nu^2+b} \left|\mathcal{X}_y(s)\right| ds + \left|\mathcal{X}(\nu^2+b)\right| \\
\leq & \frac{C_1}{\nu}\left(1+e^{\frac{1}{2}}\right) \int_y^{\nu^2+b} e^{|z|} dz + \left|\mathcal{X}(\nu^2+b)\right|.
\end{aligned}
$$

Furthermore, if $\nu^2+b<0$, we have

$$
\begin{aligned}
|\mathcal{X}(y)| \leq & \frac{C_1}{\nu}\left(1+e^{\frac{1}{2}}\right) \int_y^{\nu^2+b} e^{-z} dz + \left|\mathcal{X}(\nu^2+b)\right| \\
\leq & \frac{C_1}{\nu}\left(1+e^{\frac{1}{2}}\right)\left(e^{-y}-e^{-(\nu^2+b)}\right) + \left|\mathcal{X}(\nu^2+b)\right| \\
\leq & \frac{C_1}{\nu}\left(1+e^{\frac{1}{2}}\right) e^{|y|} + \left|\mathcal{X}(\nu^2+b)\right| \\
\leq & \hat{C}_1 \epsilon^{-\frac{1}{2}}\left(e^{|y|}+1\right),
\end{aligned} \tag{A3}
$$

where $\hat{C}_1=\sigma^{-1}\left[C_1\sqrt{2\pi}\left(1+e^{\frac{1}{2}}\right)+\left|\mathcal{X}(\nu^2+b)\right|\right]$. If $\nu^2+b\geq 0$, we have

$$
\begin{aligned}
|\mathcal{X}(y)| \leq & \frac{C_1}{\nu}\left(1+e^{\frac{1}{2}}\right)\left(\int_{-|y|}^{0} e^{-z} dz + \int_0^{\nu^2+b} e^{z} dz\right) + \left|\mathcal{X}(\nu^2+b)\right| \\
\leq & \frac{C_1}{\nu}\left(1+e^{\frac{1}{2}}\right)\left(e^{|y|}+e^{\nu^2+b}\right) + \left|\mathcal{X}(\nu^2+b)\right| \\
\leq & \frac{C_1}{\nu}\left(1+e^{\frac{1}{2}}\right)\left(1+e^{\nu^2+b}\right) e^{|y|} + \left|\mathcal{X}(\nu^2+b)\right| \\
\leq & \hat{C}_2 \epsilon^{-\frac{1}{2}}\left(e^{|y|}+1\right),
\end{aligned} \tag{A4}
$$

where $\hat{C}_2=\sigma^{-1}\left[C_1\sqrt{2\pi}\left(1+e^{\frac{1}{2}}\right)\left(1+e^{1+b}\right)+\left|\mathcal{X}(\nu^2+b)\right|\right]$

Using (A2)–(A4) to define

$$
C_2=\sigma^{-1}\left[C_1\sqrt{2\pi}\left(1+e^{\frac{1}{2}}\right)\left(1+e^{1+|b|}\right)+\sup_{y\in(b,1+b)}|\mathcal{X}(y)|\right],
$$

we know that C_2 is greater than $\hat{C}_0, \hat{C}_1, \hat{C}_2$ and is independent of ϵ. Naturally, we have

$$
|\mathcal{X}(y)| \leq C_2\left(1+e^{|y|}\right)\epsilon^{-\frac{1}{2}}.
$$

Appendix C. Proof of Lemma 5

We denote $\mathbf{E}_t[\cdot]=\mathbf{E}[\cdot|y_t=y]$, $\mathbf{Var}_t[\cdot]=\mathbf{Var}[\cdot|y_t=y]$ for convenience. Since $y_{t+\tau}$ can be written as

$$
y_{t+\tau}=(1-e^{-a\tau})b+y_t e^{-a\tau}+\sigma\int_t^{t+\tau} e^{-a(\tau-s)} dW_3(s),
$$

one can see that

$$
\begin{aligned}
\mu_\tau := & \ \mathbf{E}_t[y_{t+\tau}] = b+(y-b)e^{-a\tau}, \\
\sigma_\tau^2 := & \ \mathbf{Var}_t[y_{t+\tau}] = \sigma^2\frac{1-e^{-2a\tau}}{2a},
\end{aligned}
$$

and $y_{t+\tau}$ follows the Gaussian law $\mathcal{N}(\mu_\tau, \sigma_\tau^2)$; thus,

$$\begin{aligned}
&\mathbf{E}_t\left[e^{|y_{t+\tau}|}\right] \\
=\ &\frac{1}{\sigma_\tau\sqrt{2\pi}}\left[\int_0^{+\infty} e^x e^{-\frac{(x-\mu_\tau)^2}{2\sigma_\tau^2}}\,dx + \int_{-\infty}^0 e^{-x} e^{-\frac{(x-\mu_\tau)^2}{2\sigma_\tau^2}}\,dx\right] \\
=\ &\frac{1}{\sqrt{\pi}} e^{\mu_\tau+\frac{1}{2}\sigma_\tau^2} \int_{-\left(\frac{\mu_\tau}{\sqrt{2}\sigma_\tau}+\frac{\sqrt{2}}{2}\sigma_\tau\right)}^{+\infty} e^{-z^2}\,dz + \frac{1}{\sqrt{\pi}} e^{-\mu_\tau+\frac{1}{2}\sigma_\tau^2} \int_{\frac{\mu_\tau}{\sqrt{2}\sigma_\tau}-\frac{\sqrt{2}}{2}\sigma_\tau}^{+\infty} e^{-z^2}\,dz \\
\leq\ &2e^{|\mu_\tau|+\frac{1}{2}\sigma_\tau^2} \\
=\ &2e^{b(1-e^{-a\tau})+\frac{\sigma^2}{4a}(1-e^{-2a\tau})+|y|e^{-a\tau}} \\
\leq\ &2e^{|b|+\frac{\sigma^2}{4a}} e^{|y|}.
\end{aligned}$$

Since $a > 1$, let $\hat{C} = 2e^{|b|+\frac{\sigma^2}{4}}$, then (26) follows.

References

1. Vidyamurthy, G. *Pairs Trading: Quantitative Methods and Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2004; Volume 217.
2. Lin, Y.X.; McCRAE, M.; Gulati, C. Loss protection in pairs trading through minimum profit bounds: A cointegration approach. *Adv. Decis. Sci.* **2006**, *2006*, 73803. [CrossRef]
3. Yan, T.; Chiu, M.C.; Wong, H.Y. Pairs trading under delayed cointegration. *Quant. Financ.* **2022**, *22*, 1627–1648. [CrossRef]
4. Elliott, R.J.; Van Der Hoek, J.; Malcolm, W.P. Pairs trading. *Quant. Financ.* **2005**, *5*, 271–276. [CrossRef]
5. Bertram, W.K. Analytic solutions for optimal statistical arbitrage trading. *Phys. A Stat. Mech. Its Appl.* **2010**, *389*, 2234–2243. [CrossRef]
6. Jurek, J.W.; Yang, H. Dynamic portfolio selection in arbitrage. In Proceedings of the EFA 2006 Meetings Paper, Zurich, Switzerland, 23–26 August 2006.
7. Suzuki, K. Optimal switching strategy of a mean-reverting asset over multiple regimes. *Automatica* **2016**, *67*, 33–45. [CrossRef]
8. Suzuki, K. Optimal pair-trading strategy over long/short/square positions—Empirical study. *Quant. Financ.* **2018**, *18*, 97–119. [CrossRef]
9. Endres, S.; Stübinger, J. A flexible regime switching model with pairs trading application to the S&P 500 high-frequency stock returns. *Quant. Financ.* **2019**, *19*, 1727–1740.
10. Liu, J.; Timmermann, A. Optimal convergence trade strategies. *Rev. Financ. Stud.* **2013**, *26*, 1048–1086. [CrossRef]
11. Chiu, M.C.; Wong, H.Y. Dynamic cointegrated pairs trading: Mean–variance time-consistent strategies. *J. Comput. Appl. Math.* **2015**, *290*, 516–534. [CrossRef]
12. Zhu, D.M.; Gu, J.W.; Yu, F.H.; Siu, T.K.; Ching, W.K. Optimal pairs trading with dynamic mean-variance objective. *Math. Methods Oper. Res.* **2021**, *94*, 145–168. [CrossRef]
13. Björk, T.; Murgoci, A.; Zhou, X.Y. Mean–variance portfolio optimization with state-dependent risk aversion. *Math. Financ. Int. J. Math. Stat. Financ. Econ.* **2014**, *24*, 1–24. [CrossRef]
14. Cont, R. Empirical properties of asset returns: Stylized facts and statistical issues. *Quant. Financ.* **2001**, *1*, 223–236. [CrossRef]
15. Cont, R. Volatility clustering in financial markets: Empirical facts and agent-based models. In *Long Memory in Economics*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 289–309.
16. Teräsvirta, T.; Zhao, Z. Stylized facts of return series, robust estimates and three popular models of volatility. *Appl. Financ. Econ.* **2011**, *21*, 67–94. [CrossRef]
17. Gatheral, J.; Jaisson, T.; Rosenbaum, M. Volatility is rough. *Quant. Financ.* **2018**, *18*, 933–949. [CrossRef]
18. Scott, L.O. Option pricing when the variance changes randomly: Theory, estimation, and an application. *J. Financ. Quant. Anal.* **1987**, *22*, 419–438. [CrossRef]
19. Fouque, J.P.; Papanicolaou, G.; Sircar, K.R. Mean-reverting stochastic volatility. *Int. J. Theor. Appl. Financ.* **2000**, *3*, 101–142. [CrossRef]
20. Fouque, J.P.; Lorig, M.J. A fast mean-reverting correction to Heston's stochastic volatility model. *SIAM J. Financ. Math.* **2011**, *2*, 221–254. [CrossRef]
21. Chambers, M.J. The estimation of continuous time models with mixed frequency data. *J. Econom.* **2016**, *193*, 390–404. [CrossRef]
22. Wang, X.; He, X.; Bao, Y.; Zhao, Y. Parameter estimates of Heston stochastic volatility model with MLE and consistent EKF algorithm. *Sci. China Inf. Sci.* **2018**, *61*, 042202. [CrossRef]

23. Simon, D. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
24. Aihara, S. Estimation of stochastic volatility in the Hull-White model. *Appl. Math. Financ.* **2000**, *7*, 153–181. [CrossRef]