Pairs Trading with Stock Borrowing Fee

Ziyi Chen *

Jia-Wen Gu †

Harry Zheng ‡

Abstract

Pairs trading is a strategy that involves simultaneously longing one asset and shorting another related asset, aiming to profit from the price difference between them. In this paper, we discuss a pairs trading problem with stock borrowing fee under a mean-variance (MV) framework. We assume that the difference in the logarithm of two stock prices follows an OU process and focus on the trading strategy that always shorts one stock and longs the other in equal dollar amount. When borrowing stocks for short selling, an interest fee is incurred. By combining dynamic programming and BSDE methods, we establish the existence of solutions to the BSDE and, based on this result, derive a semi-closed-form equilibrium strategy that can be decomposed into three parts: the first part reflects myopic demand, which is expressed as a linear function of the price spread. The second part, characterized by the solution of a BSDE, represents hedging demand. The third part arises from

*Department of Mathematics, Southern University of Science and Technology, Shenzhen, China. E-mail: 12231268@mail.sustech.edu.cn.

†Department of Mathematics, Southern University of Science and Technology, Shenzhen, China. E-mail: gujw@sustech.edu.cn or jwgu.hku@outlook.com. Supported in part by the Key Project of National Natural Science Foundation of China 72432005, Guangdong Basic and Applied Basic Research Foundation 2023A1515030197.

‡Department of Mathematics, Imperial College London, UK. E-mail: h.zheng@imperial.ac.uk. Supported in part by Engineering and Physical Sciences Research Council of UK (Grant No. EP/V008331/1).

the stock borrowing fee. We also consider the impact of trading constraints on the equilibrium strategy. Finally, we adopt a deep learning-based approach to numerically solve the problem and present simulation results to show the performance of the equilibrium strategy.

Key words: Pairs trading, Mean variance framework, Time inconsistency, Equilibrium strategy, Stock borrowing fee, Deep learning, OU process
AMS MSC2010: 49L20, 60H30

1 Introduction

Pairs trading is a market-neutral trading strategy aimed at exploiting the price differentials between two highly correlated financial assets. It revolves around the concept of statistical arbitrage, where investors simultaneously buy the undervalued asset and sell the overvalued asset, anticipating the convergence of their prices back to their long-term relationship. The study by Gatev et al. [?] indicates that pairs trading can maintain stable returns in both bear and bull markets.

The strategy for pairs trading typically relies on quantitative analysis and model building to identify and capitalize on price divergence (see Vidyamurthy [?], Cartea [?]). Existing research papers on pairs trading using mathematical methods can be broadly categorized into two main types: those considering continuous scenarios and those focusing on discrete scenarios. In continuous scenarios, pairs trading is modeled as a portfolio optimization problem where the goal is to determine the optimal amount of money (or number of shares) to invest in paired stocks at each moment in time. Mudchanatongsuk [?] consider a self-financing portfolio composed of paired stocks and bonds. They solve the optimal strategy in closed form at each moment in time under the utility model. Tourin and Yan [?] extend the work of Mudchanatongsuk [?], studying pairs trading strategies with arbitrary amounts in each stock. Zhu et al. [?] and Yu et al. [?] consider a similar model under the mean-variance criterion. Due to time inconsistency in mean-variance problems, Zhu et al. [?] seek an equilibrium strategy, while Yu et al. [?] have a pre-commitment strategy. It is worth noting that these papers do not have trading frictions. Li and Tourin [?] model the trading rates that affect the execution prices of the two stocks. In discrete scenarios, pairs trading is formulated as an optimal switching problem. Zhang and Zhang [?] and Tie et al. [?] solve optimal two-regime (buy and sell) switching problems with transaction costs. The former considers an exponential OU process, while the latter examines a 2-dimensional geometric Brownian motion. Based on their work, Song and Zhang [?] and Liu et al. [?] introduce the stop-loss mechanism. Furthermore, Suzuki [?], Ngo and Pham [?], and Suzuki [?] study the switching problem between three regimes: flat position (no holding stocks), one long and the other short and one short and the other long. These papers obtain optimal strategies by solving the HJB equations, which are quasi-variational inequalities. Leung and Li [?] and Leung et al. [?] employ a probabilistic approach to rigorously derive optimal price intervals for entering and exiting the market.

Trading costs play a crucial role in the design and performance of quantitative strategies (see, e.g., Almgren and Chriss [?], Bismuth et al.[?], Schied et al. [?], Shreve and Soner [?]). One of the salient features of pairs trading is the requirement to short sell one of the stocks in the pair, which incurs a stock borrowing fee. The existing literature has investigated the influence of short selling costs and short constraints. Miller [?] theoretically proposes that when there is a wide divergence in investor opinions regarding a certain risky asset, short-selling costs and constraints force pessimistic investors out of the market. As a result, stock prices fail to incorporate the information held by these pessimistic investors, leading to upward bias and potential overvaluation of the stock. Duffie et al. [?] propose an asset valuation model in which short-selling is achieved by finding security lenders and negotiating the lending fee and show that if the security is difficult to borrow, its initial price would be higher and is expected to decline over time. Jone and Lamont [?], Billingsley and Kovacs [?], and Chang et al. [?] conduct empirical analysis by collecting stock market data, and their results consistently support Miller [?]'s theory. These papers highlight the importance of short selling costs and short constraints. Pairs trading is based on buying undervalued stocks and short selling overvalued stocks. Since the stock borrowing fee rate when short selling is often much higher than the risk-free rate (Duffie, et al. [?]), this greatly affects the performance of pairs trading strategies. The past literature has not thoroughly explored the impact of stock borrowing fees on pairs trading, which is the focus of this research.

We consider the dynamic trading problem in a continuous setting and assume that our wealth consists of paired stocks and a bond. To capture the mean-reversion property of paired stocks, we model the spread between the logarithms of their prices as an OU process, which is a standard assumption in the literature on trading co-integrated assets (see, e.g., Bergault [?], Cartea [?], Emschwiller [?], Lehalle [?]). We engage in short selling of one stock and allocate the proceeds to another stock. This is referred to as a zero-cost strategy, which achieves trading without the need for initial capital investment through such a combination. Additionally, our model incorporates trading frictions, where borrowing stocks for short-selling incurs a specific borrowing fee that is modeled to be proportional to the dollar value of the stock being shorted. Moreover, we introduce short-selling constraints. The mean-variance criterion is adopted to select the optimal strategy. In existing literature, pre-commitment and equilibrium strategies are widely used for time-inconsistent mean-variance problems. The pre-commitment strategy is established at the initial time and is strictly adhered to throughout the investment horizon, without adjustments due to intervening events or new information (see, Strotz [?], Dang and Forsyth [?] and Pedersen and Peskir [?]). The equilibrium strategy refers to a strategy that is optimal at any point in time, given that the same strategy is to be followed in the future. (see, Basak and Chabakauri [?], Björk and Murgoci [?], Björk et al. [? ?], and Dai et al. [?]). In this paper, we adopt the equilibrium strategy. We first derive the extended HJB equations and then express the equilibrium strategy in the form of a BSDE utilizing the connection between semi-linear PDE and BSDE. To address the technical existence problem of BSDE with quadratic generators, we adopt the method developed by Briand and Hu [?]. Finally, we employ deep learning methods to solve the BSDE and provide numerical simulation results. To the best of our knowledge, only Xu and Yang [?] have studied the optimal pairs trading model with costly short-selling under the utility maximization criterion. There are significant differences between their study and ours. We adopt the mean-variance criterion, whereas they employ a utility-based framework. The mean-variance criterion captures risk explicitly by penalizing the variance of terminal wealth, thereby directly reflecting the risk of losses caused by non-convergence of price spreads and excessively large risk positions. Besides, we derive semi-closed-form strategies that facilitate theoretical analysis, while Xu and Yang [?] rely exclusively on numerical analysis in the case of a power utility function.

In summary, our innovations are three-fold. Firstly, we incorporate the stock borrowing fee into our model to better capture real-world trading dynamics and consider the impact of portfolio constraints (or short selling constraints). Secondly, we establish the existence of solutions to the FBSDE, which plays a crucial role in deriving the semi-closed-form equilibrium strategy. The semi-closed-form equilibrium strategy allows us to theoretically analyze similarities and differences between pairs trading strategies with and without short-selling costs, providing a solid foundation for the analysis of optimal decision-making. Thirdly, we employ deep learning techniques to visualize the equilibrium strategies, which help us gain an intuitive understanding of properties of strategies and how parameter values influence the trading strategy. We also conduct several simulation experiments to illustrate the impact of stock borrowing fees and short-selling constraints.

The paper is structured as follows. Section 2 presents the model setup and MV problem. Section 3 introduces the methodology for finding the equilibrium strategy. Section 4 shows a deep learning method for solving FBSDE. Section 5 presents the simulation result. Section 6 concludes. Some technical proofs are given in the appendix.

2 Problem Formulation

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a filtered complete probability space with $\{\mathcal{F}_t\}_{t\geq 0}$ being a natural filtration generated by two standard Brownian motions $\hat{W}$ and W with correlation $\rho \in [-1, 1]$, i.e., $\mathrm{d}\langle W, \hat{W}\rangle_t = \rho \mathrm{d}t$. We consider a self-financing portfolio that contains a risk-free asset M_t and two stocks S_t^1 and S_t^2. The price of a risk-free asset satisfies the following equation:

$$dM_t = rM_t dt, \tag{1}$$

where r is the risk-free interest rate, a positive constant. The stock price S_t^2 evolves according to

$$dS_t^2 = bS_t^2 dt + \sigma S_t^2 d\hat{W}_t \tag{2}$$

where the growth rate b and volatility σ are positive constants. Let X_t denote the (logarithmic) price spread between stocks S_t^1 and S_t^2, i.e., $X_t = lnS_t^1 - lnS_t^2$. We assume it follows an OU process:

$$dX_t = k(\theta - X_t)dt + \eta dW_t, \tag{3}$$

where $k > 0$ is the mean reversion rate, θ the long-term mean, $\eta > 0$ the volatility. Applying Itô's formula, we see that the dynamics of S_t^1 satisfy the following equation:

$$dS_t^1 = [k(\theta - X_t) + b + \frac{1}{2}\eta^2 + \rho\sigma\eta]S_t^1 dt + \sigma S_t^1 d\hat{W}_t + \eta S_t^1 dW_t. \tag{4}$$

We consider a zero-cost strategy, that is, we short sell one stock and use the proceeds to buy another stock. Specifically, a dollar amount of u_t is allocated to S_t^1, and a dollar amount of $-u_t$ is allocated to S_t^2. Assume that the stock borrowing fee rate is δ. Let V_t denote the self-financing wealth process. It can be described by

$$dV_t = u_t \frac{dS_t^1}{S_t^1} - u_t \frac{dS_t^2}{S_t^2} + V_t \frac{dM_t}{M_t} - \delta |u_t| dt. \tag{5}$$

Substituting (1), (2) and (4) into (5), we get

$$dV_t = u_t \mu(X_t)dt - \delta |u_t| dt + rV_t dt + \eta u_t dW_t. \tag{6}$$

where $\mu(x) = k(\theta - x) + \frac{1}{2}\eta^2 + \rho\sigma\eta$.

Remark 1. If we switch the labels of two assets, the only difference lies in the sign of the control u_t of (6). It would not affect the optimal strategy since we consider a symmetric strategy.

Let T be the investment horizon. We define the admissible feedback strategy as follows:

Definition 1. $u = \{u_s, t \le s \le T\}$ is called an admissible feedback strategy if

1. $u_s = u(s, X_s, V_s)$, where $u(\cdot, \cdot, \cdot)$ is a deterministic mapping from $[t, T] \times \mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$.

2. For each initial point (t, x, v), the SDE (6) has a unique strong solution, denoted by V^u.

3. $E_t^{x,v} \int_t^T u_s^2 ds < \infty$, where $E_t^{x,v}[\cdot]$ denotes $E[\cdot | X_t = x, V_t = v]$.

The collection of all admissible feedback strategies starting at t is denoted by $\mathcal{U}_t$. We consider the following reward function, which balances risk and return:

$$J(t, x, v; u) = E_t^{x,v}[V_T^u] - \frac{\gamma}{2} Var_t^{x,v}[V_T^u], \tag{7}$$

where $Var_t^{x,v}[\cdot]$ denotes $Var[\cdot|X_t = x, V_t = v]$, $\gamma > 0$ is the risk aversion parameter. In what follows, we will omit the superscripts (x, v) in $E_t^{x,v}$ and $Var_t^{x,v}$ when no confusion arises. Our goal is to find an admissible control $\hat{u}$ that maximizes the reward function:

$$J(t, x, v; \hat{u}) = \sup_{u \in \mathcal{U}_t} J(t, x, v; u). \tag{8}$$

Note that due to the quadratic term of the conditional variance in the reward function, the MV problem (8) exhibits time inconsistency. To address this issue, we adopt the equilibrium strategy method proposed by Björk et al. [?], which is locally optimal at any given time, assuming the same strategy is to be followed in the future. We now introduce the definition of equilibrium strategy as follows:

Definition 2. Consider an admissible feedback strategy $\hat{u} \in \mathcal{U}_t$. It is an equilibrium strategy if for any $(t, x, v) \in [0, T] \times R \times R$, the following inequality holds:

$$\liminf_{h \to 0^+} \frac{J(t, x, v; \hat{u}) - J(t, x, v; u^h)}{h} \geq 0,$$

where u^h is an admissible perturbation strategy for $\hat{u}$, defined by

$$u^h(s, x, v) = \begin{cases} w, & \text{if } t \leq s \leq t + h, \\ \hat{u}(s, x, v), & \text{if } t + h \leq s \leq T, \end{cases}$$

with $0 < h < T - t$ and w is an arbitrarily given constant.

Remark 2. It is worth noting that the existing literature on terminology issues is limited and there is no clear consensus on the naming conventions for classes of control processes. Here we adopt the definition of feedback strategy provided by Björk et al. [?] and Dai et al. [?].

If such an equilibrium strategy $\hat{u}$ exists, we denote the equilibrium value function as G, i.e., $G(t, x, v) = J(t, x, v; \hat{u})$.

3 Equilibrium Strategy

In this section, we demonstrate how to combine the PDE and BSDE methods to solve the equilibrium strategy for the MV problem.

We first introduce the notation $F(t,x,v) = E_t^{x,v} V_T^{\hat{u}}$, which is the condition expectation of terminal wealth under the equilibrium strategy. Similar to Björk et al. [?] and Gu et al. [?], we can derive the extended HJB equation system:

$$\begin{cases} \sup_u \{\mathcal{A}^u G - \frac{\gamma}{2}(\eta^2 F_x^2 + \eta^2 u^2 F_v^2 + 2\eta^2 u F_x F_v)\} = 0, \\ \mathcal{A}^{\hat{u}} F = 0, \end{cases} \tag{9}$$

with terminal conditions $G(T,x,v) = F(T,x,v) = v$, where $\mathcal{A}^u$ is an infinitesimal generator defined by

$$\begin{aligned} \mathcal{A}^u :=& \frac{\partial}{\partial t} + k(\theta - x)\frac{\partial}{\partial x} + [\mu(x)u - \delta|u| + rv]\frac{\partial}{\partial v} \\ &+ \frac{1}{2}\eta^2\frac{\partial^2}{\partial x^2} + \frac{1}{2}\eta^2 u^2 \frac{\partial^2}{\partial v^2} + \eta^2 u \frac{\partial^2}{\partial x \partial v}. \end{aligned}$$

Note that (9) is a coupled system of nonlinear equations. Directly solving it is challenging, we first consider dimension reduction. We conjecture that the solution has the following form:

$$\begin{cases} G(t,x,v) = e^{r(T-t)}v + g(t,x), \quad g(T,x) = 0, \\ F(t,x,v) = e^{r(T-t)}v + f(t,x), \quad f(T,x) = 0. \end{cases} \tag{10}$$

Substituting (10) into (9) and simplifying the expression, we have

$$\begin{cases} g_t + k(\theta - x)g_x + \frac{1}{2}\eta^2 g_{xx} - \frac{\gamma}{2}\eta^2 f_x^2 + \sup_u \{R(u;t,x,f_x)\} = 0, \\ f_t + k(\theta - x)f_x + \frac{1}{2}\eta^2 f_{xx} + e^{r(T-t)}[\mu(x)\hat{u} - \delta|\hat{u}|] = 0. \end{cases} \tag{11}$$

where $R(u;t,x,f_x) := -\gamma\eta^2 e^{2r(T-t)}u^2/2 + \mu(x)e^{r(T-t)}u - \gamma\eta^2 e^{r(T-t)}u f_x - \delta e^{r(T-t)}|u|$.

Lemma 3.1. *There exists a unique $\hat{u}$ at which $R(u;t,x,f_x)$ achieves its maximum* *value and $\hat{u}$ is given by*

$$\hat{u} = \begin{cases} \dfrac{\mu(x) - \gamma\eta^2 f_x - \delta}{\gamma\eta^2 e^{r(T-t)}}, & \text{if } \mu(x) - \gamma\eta^2 f_x \geq \delta, \\ \dfrac{\mu(x) - \gamma\eta^2 f_x + \delta}{\gamma\eta^2 e^{r(T-t)}}, & \text{if } \mu(x) - \gamma\eta^2 f_x \leq -\delta, \\ 0, \quad \text{if } -\delta < \mu(x) - \gamma\eta^2 f_x < \delta. \end{cases} \tag{12}$$

Proof. It is not difficult to check that $\hat{u} = \underset{u}{\operatorname{argmax}}\, R(u; t, x, f_x)$. Fixing (t, x, f_x), the term $-\gamma\eta^2 e^{2r(T-t)}u^2/2 + \mu(x)e^{r(T-t)}u - \gamma\eta^2 e^{r(T-t)}uf_x$ is a quadratic function in u with a negative leading coefficient, hence it is strictly concave. the term $-\delta e^{r(T-t)}|u|$ is clearly concave in u. Since the sum of a strictly concave function and a concave function remains strictly concave, it follows that R is strictly concave. Consequently, the maximizer $\hat{u}$ is unique. □

Substituting (12) into the second equation of (11), we obtain

$$f_t + k(\theta - x)f_x + \frac{1}{2}\eta^2 f_{xx} + m(x, \eta f_x) = 0, \ \ f(T, x) = 0, \tag{13}$$

where $m(x, z)$ is a continuous function given by

$$m(x, z) = \begin{cases} \dfrac{(\mu(x) - \delta)(\mu(x) - \gamma\eta z - \delta)}{\gamma\eta^2}, & \text{if } \mu(x) - \gamma\eta z \geq \delta, \\ \dfrac{(\mu(x) + \delta)(\mu(x) - \gamma\eta z + \delta)}{\gamma\eta^2}, & \text{if } \mu(x) - \gamma\eta z \leq -\delta, \\ 0, \quad \text{if } -\delta < \mu(x) - \gamma\eta z < \delta. \end{cases}$$

Remark 3. The continuity of m is due to the fact that it is a piecewise continuous function and remains continuous at the breakpoints. Besides, m is quadratically dominated. Indeed, by the Cauchy inequality, it easy to see that

$$|m(x, z)| \leq C_1 + C_2 x^2 + C_3 z^2,$$

where C_1, C_2, C_3 are constants.

Due to the low regularity of m, solving the equation (13) is challenging and it may not yield a classical solution, which makes it difficult to apply the verification theorem. We use the nonlinear Feynman-Kac formula (see section 6.3 in Pham [**?**

]) to transform the semi-linear PDE into a BSDE which requires weaker regularity conditions. Next, we would focus on the following FBSDE:

$$\begin{cases} -dY_s = m(X_s, Z_s)ds - Z_s dW_s, \quad Y_T = 0, \\ dX_s = k(\theta - X_s)ds + \eta dW_s, \quad X_0 = x. \end{cases} \tag{14}$$

According to the nonlinear Feynman-Kac formula, we have $f(t, X_t) = Y_t$, which will be verified later. Note that (14) is a decoupled FBSDE. The forward process X is a well defined OU process. The generator m is unbounded and exhibits quadratic growth with respect to (x, z), addressing the existence problem requires a technical method. We first establish some preliminary results, focusing in particular on estimating exponential moments of $\int_0^T X_t^2 dt$, namely, $E\exp\{p\int_0^T X_t^2 dt\}$ for $p > 0$. For simplicity, we assume $\theta = 0$ and define $\tilde{u}(t, x) = E[\exp\{p\int_t^T X_t^2 dt\}|X_t = x]$. By Feynman-Kac formula, $\tilde{u}(t, x)$ satisfies the following PDE:

$$\begin{cases} \tilde{u}_t - kx\tilde{u}_x + \frac{1}{2}\eta^2\tilde{u}_{xx} + px^2\tilde{u} = 0, \\ \tilde{u}(T, x) = 1. \end{cases} \tag{15}$$

We assume that the solution takes the form of :

$$\tilde{u}(t, x) = \exp(A(t)x^2 + B(t)), \tag{16}$$

where $A(t)$ and $B(t)$ are undetermined functions. Substituting (16) into (15) and comparing the coefficients of corresponding terms, we obtain

$$\begin{cases} A'(t) - 2kA(t) + 2\eta^2 A^2(t) + p = 0, \quad A(T) = 0, \\ B'(t) + \eta^2 A(t) = 0, \quad B(T) = 0. \end{cases} \tag{17}$$

In the following, we focus on the discussion of the existence of $A(t)$, since once a solution for $A(t)$ is established, $B(t)$ can subsequently be determined. Let $y(t) = A(T - t)$. Then $y(t)$ satisfies

$$y' = h(y), \quad y(0) = 0, \tag{18}$$

where we use the simple notation $h(y) = -2ky + 2\eta^2 y^2 + p$. Next, we turn to study the existence and uniqueness of solution to ODE (18).

Proposition 3.1. *Consider the case where* $p > 0$.

1. *If* $p \leq \frac{k^2}{2\eta^2}$, *there exists an unique solution* $y(\cdot)$ *defined on* $[0, \infty)$. *In particular,* $y(t)$ *increases from* 0 *to a constant value, which corresponds to a zero of the function* h.

2. *If* $p > \frac{k^2}{2\eta^2}$, *there exists an unique solution* $y(\cdot)$ *defined on* $[0, \xi)$. ξ *is a positive constant satisfying*

$$\frac{\pi}{2\sqrt{2\eta^2 p}} < \xi < \frac{\pi}{2\sqrt{ac}},$$

where

$$a = -\frac{2\eta^2 k^2}{k^2 + p\eta^2} + 2\eta^2 > 0, \quad c = \frac{p}{2} - \frac{k^2}{4\eta^2} > 0. \tag{19}$$

In particular, $y(t)$ *is also increasing but* $\lim_{t \to \xi} y(t) = +\infty$.

Proof. The proof is given in Appendix 7.1. □

Returning to the ODE system (17), we have

Corollary 3.1. *Consider ODE system* (17) .

1. *If* $0 < p \leq \frac{k^2}{2\eta^2}$, *then there exists a unique solution* $(A(t), B(t))$ *defined on* $[0, T]$.

2. *If* $p > \frac{k^2}{2\eta^2}$, *then there exists a unique solution* $(A(t), B(t))$ *defined on* $(T - \xi, T]$.

In the second case of the Corollary 3.1, $(A(\cdot), B(\cdot))$ would blow up as $t \downarrow T - \xi$, which implies that $(A(0), B(0))$ is not defined when $T > \xi$. Recall that $\tilde{u}(t, x) = E[\exp\{p \int_t^T X_t^2 dt\} | X_t = x]$, along with the expression for $\tilde{u}$ in (16). According to Corollary 3.1, we can get the desired result for the estimate for the exponential moment of $\int_0^T X_t^2 dt$.

Corollary 3.2. *If* $0 < p \leq \frac{k^2}{2\eta^2}$, *then* $\int_0^T X_t^2 dt$ *have finite exponential moments of such* p. *If* $p > \frac{k^2}{2\eta^2}$, $\int_0^T X_t^2 dt$ *have finite exponential moments of such* p *only when* $T < \xi$.

Combining the results discussed above, we can derived the existence of FBSDE (14).

Lemma 3.2. *Suppose $kT < \frac{\pi}{2}$, then the FBSDE (14) admits a solution (Y, Z) satisfying*

$$\begin{cases} |Y_t| \leq \dfrac{1}{\varepsilon+\tilde{\varepsilon}} \log E_t \left(\exp \left((\varepsilon+\tilde{\varepsilon}) \displaystyle\int_t^T \alpha(r)dr \right) \right), & \forall t \in [0,T], \\ E\left(\displaystyle\int_0^T |Z_s|^2 ds \right) < +\infty, \end{cases} \tag{20}$$

where $\varepsilon, \tilde{\varepsilon}$ are positive constants satisfying

$$\begin{cases} 0 < \varepsilon < \gamma(\dfrac{\pi^2}{8k^2T^2} - \dfrac{1}{2}), \\ 0 < \tilde{\varepsilon} < -\Big(\dfrac{k^2}{2\gamma\eta^2} + \dfrac{k^2}{4\eta^2\varepsilon} + \dfrac{\varepsilon}{2}\Big) + \dfrac{1}{2}\sqrt{\Big(\dfrac{k^2}{\gamma\eta^2} + \dfrac{k^2}{2\eta^2\varepsilon} + \varepsilon\Big)^2 + 4\Big(\dfrac{\pi^2}{8\eta^2T^2} - \dfrac{k^2\varepsilon}{\gamma\eta^2} - \dfrac{k^2}{2\eta^2}\Big)}, \end{cases} \tag{21}$$

and $\alpha(t) = (\frac{k^2}{\gamma\eta^2} + \frac{k^2}{2\eta^2\varepsilon} + \tilde{\epsilon})X_t^2 + K$, K is a positive constant.

Proof. The proof is given in Appendix 7.2. □

It's worth mentioning that the condition of the Lemma are not the most relaxed, as the inequality estimation in the proof is not the sharpest. One can attempt to relax the assumption, but more technical handing is required. Thanks to Lemma 3.2, we can derive the equilibrium strategy.

Theorem 3.1. (***Verification Theorem***) *Let (X_t, Y_t, Z_t) be the solution of (14). Then an equilibrium strategy of the MV problem (8) is given by*

$$\hat{u}(t, X_t) = \begin{cases} \dfrac{\mu(X_t) - \gamma\eta Z_t - \delta}{\gamma\eta^2 e^{r(T-t)}}, & \text{if } \mu(X_t) - \gamma\eta Z_t \geq \delta, \\ \dfrac{\mu(X_t) - \gamma\eta Z_t + \delta}{\gamma\eta^2 e^{r(T-t)}}, & \text{if } \mu(X_t) - \gamma\eta Z_t \leq -\delta, \\ 0, \quad \text{if } -\delta < \mu(X_t) - \gamma\eta Z_t < \delta, \end{cases} \tag{22}$$

and the expected growth of wealth under the equilibrium strategy is given by $E_t[V_T^{\hat{u}} - e^{r(T-t)}V_t] = Y_t$. Furthermore, if the solution (Y, Z) of (14) is unique, then there exists a deterministic function f such that $Y_t = f(t, X_t)$. If $f \in C^{1,2}$, then f is a classical solution of (13) with $Z_t = \eta f_x(t, X_t)$.

Proof. The proof is given in Appendix 7.3. □

From the expression of equilibrium strategy (22), we can see that due to the existence of the stock borrowing fee, holding a non-zero position in the paired stocks at all time is not optimal. Instead, there exists a region where, when the price spread falls within this range, we do not hold any risky assets. This is similar to the situation where transaction costs are present (e.g., Dai and Yi [?] and Davis and Norman [?]). This result differs from that in Zhu et al. [?], where they also consider the pairs trading problem but the stock borrowing fee is ignored.

Consider the expression of $e^{r(T-t)}\hat{u}_t$, we can divide it into three parts. The first term, $\mu(X_t)/\gamma\eta^2$, is determined by the price spread X_t at the current moment, reflecting myopic demand, which implies that the strategy's immediate response to the current price differential illustrates a tendency to focus predominantly on short-term fluctuations rather than potential long-term implications or benefits. The second term, $-Z_t/\eta$, is commonly referred to as the intertemporal hedging demand, which represents the strategy to mitigate the risk associated with unpredictable changes in investment opportunities that are external to investors and beyond their control. The third term, $\pm\delta/\gamma\eta^2$, is included to offset the effects of stock borrowing fee. When $\delta = 0$ this term vanishes. Moreover, note that

$$\lim_{T\to+\infty} \hat{u}(t, X_t) = 0,$$

which implies that as the investment horizon T extends, the optimal position size in stock pairs decreases significantly. This observation underscores the notion that, to mitigate volatility risk, traders are likely to adopt smaller positions when the trading duration is prolonged.

Corollary 3.3. *If there is no stock borrowing fee, i.e., $\delta = 0$, the equilibrium strategy is given by*

$$\hat{u}(t, x) = \frac{[k + 2k^2(T-t)](\theta - x)}{\gamma\eta^2 e^{r(T-t)}} + \frac{[k(T-t)+1]^2(\eta^2 + 2\rho\sigma\eta)}{2\gamma\eta^2 e^{r(T-t)}}. \tag{23}$$

Proof. The proof is given in Appendix 7.4. □

When other parameters are given, the equilibrium strategy (23) is linearly and negatively correlated with the price spread, reflecting the investment insight that pairs trading seeks to exploit the convergence of price spreads to their long-term mean for arbitrage opportunities. When the volatility of the price spread η is high, investors' expectations for short-term profits (i.e., myopic demand) are suppressed. This is because high volatility implies larger price fluctuations and higher risk, leading investors to reduce their short-term positions to mitigate potential losses. When the risk aversion coefficient γ is high, it suppresses both myopic demand and hedging demand. Overall, a high risk aversion coefficient makes investors more conservative, leading them to reduce their risk exposure in both the short and long term to mitigate potential losses. It is interesting to observe that, although $x > \theta$ indicates that stock S^1 is relatively overpriced and mean-reversion favors a short position in S^1, the optimal strategy (23) may still prescribe a long position of S^1 and a short position of S^2. This occurs because the second term of (23), which arises from the drift of S^1 in (4), captures the risk premium of S^1 relative of S^2 and remains positive whenever $\rho > -\frac{\eta}{2\sigma}$. The equilibrium strategy takes a trade-off between a mean-reversion signal (the first term, advising a short position in S^1 when $x > \theta$) and the attraction of a risk premium of S^1 (the second term, advising a long position of S^1 when $\rho > -\frac{\eta}{2\sigma}$). For moderate deviations above θ, the allure of longing S^1 to capture its risk premium can dominate the weak mean-reversion signal, justifying the establishment of a long position of S^1 and a short position of S^2 ($\hat{u} > 0$). The strategy only dictates a short position of S^1 once the expected profit from a strong mean-reversion (a large $x \gg \theta$) outweighs the risk premium.

Furthermore, since the equilibrium strategy (23) is linearly negatively correlated with the spread, this implies that when the spread deviates significantly from the mean without converging, investors would increase investment in risk assets, leading to greater risk exposure. We therefore consider imposing constraints on the positions in stock pairs in the next result, whose proof is similar to that of Theorem 3.1 and is omitted.

Proposition 3.2. *Consider the following constraints on the strategy:*

$$\underline{u}_t \leq u_t \leq \overline{u}_t,$$

where $\underline{u}_t, \overline{u}_t$ *satisfy* $-\infty \leq \underline{u}_t \leq 0$ *and* $0 \leq \overline{u}_t \leq +\infty$. *Suppose the following FBSDE* *has a solution* (Y, Z)*:*

$$\begin{cases} -dY_s = n(s, X_s, Z_s)ds - Z_s dW_s, \quad Y_T = 0, \\ dX_s = k(\theta - X_s)ds + \eta dW_s, \quad X_0 = x, \end{cases}$$

where $n(t, x, z)$ *is defined by*

$$n(t,x,z) = \begin{cases} (\mu(x) - \delta)\Big[\dfrac{(\mu(x) - \gamma\eta z - \delta)}{\gamma\eta^2} \wedge e^{r(T-t)}\overline{u}_t\Big], & \text{if } \mu(x) - \gamma\eta z \geq \delta, \\ (\mu(x) + \delta)\Big[\dfrac{(\mu(x) - \gamma\eta z + \delta)}{\gamma\eta^2} \vee e^{r(T-t)}\underline{u}_t\Big], & \text{if } \mu(x) - \gamma\eta z \leq -\delta, \\ 0, \quad \text{if} -\delta < \mu(x) - \gamma\eta z < \delta. \end{cases}$$

Then the equilibrium strategy is given by

$$\hat{u}(t, X_t) = \begin{cases} \dfrac{\mu(X_t) - \gamma\eta Z_t - \delta}{\gamma\eta^2 e^{r(T-t)}} \wedge \overline{u}_t, & \text{if } \mu(X_t) - \gamma\eta Z_t \geq \delta, \\ \dfrac{\mu(X_t) - \gamma\eta Z_t + \delta}{\gamma\eta^2 e^{r(T-t)}} \vee \underline{u}_t, & \text{if } \mu(X_t) - \gamma\eta Z_t \leq -\delta, \\ 0, \quad \text{if } -\delta < \mu(X_t) - \gamma\eta Z_t < \delta. \end{cases} \tag{24}$$

Note that if $\underline{u}_t = 0$, it means that one can only take a long position on stock S^1 and a short position on stock S^2. Even $\delta = 0$, the constrained equilibrium strategy (24) does not have an explicit analytical expression.

4 Numerical Method for FBSDE

Due to the lack of explicit solutions for FBSDE (14) in general cases, we employ deep learning-based numerical methods to solve FBSDE (14) in this section. The method of using deep learning to solve FBSDEs is first proposed by E et al. [?] and Han et al. [?]. The convergence of the algorithm is provided by Han and Long [?]. Various variants have also emerged (see Chan-Wai-Nam [?] and Raissi [?]).

We divide the time horizon $[0, T]$ into a sequence of time grids $0 = t_0 < t_1 < \cdots < t_N = T$, where $t_i = i\Delta t$, $i = 0, 1, \cdots, N$, $\Delta t = T/N$. For notational simplicity, we denote $X_n := X_{t_n}, Y_n := Y_{t_n}, Z_n := Z_{t_n}$. We first employ the Euler scheme to approximate the OU process $\{X_t\}_{t\in[0,T]}$:

$$X_{n+1} = X_n + k(\theta - X_n)(t_{n+1} - t_n) + \eta(W_{t_{n+1}} - W_{t_n}) \tag{25}$$

for $n = 0, \cdots, N-1$ with $X_0 = x$. Analogously, we can obtain a discretization of BSDE as

$$Y_{n+1} = Y_n - m(X_n, Z_n)(t_{n+1} - t_n) + Z_n(W_{t_{n+1}} - W_{t_n}). \tag{26}$$

Note that $\{X_n\}_{n=0}^{N}$ can be obtained directly from the discretization, whereas $\{Y_n\}_{n=0}^{N-1}$ and $\{Z_n\}_{n=0}^{N-1}$ are the quantities that we need to determine.

We employ a deep learning approximation for the initial condition Y_0 and $\{Z_n\}_{n=0}^{N-1}$ but not for $\{Y_n\}_{n=1}^{N-1}$ while taking $\{X_n\}_{n=0}^{N}$ as input. Specifically, given $X_0 = x$, we consider $Y_0^\nu \in R$ and $Z_0^\nu \in R$ to be suitable approximations of Y_0 and Z_0, respectively. Let $Z_n^\nu : R \to R$ be networks functions that serve as suitable approximations Z_n for $n = 1, \cdots, N-1$. Then Y_n^ν can be obtained by (26), that is

$$Y_{n+1}^\nu = Y_n^\nu - m(X_n, Z_n^\nu(X_n))(t_{n+1} - t_n) + Z_n^\nu(X_n)(W_{t_{n+1}} - W_{t_n})$$

for $n = 0, \cdots, N-1$. Inspired by the terminal condition for BSDE, i.e., $Y_T = 0$, we choose the following loss function to update the parameters by stochastic gradient decent (SGD) method:

$$\nu \mapsto E[|Y_N^\nu|^2].$$

In actual training processes, we use $N-1$ fully-connected feedforward neural networks without bias term to represent $\{Z_n^\nu\}_{n=1}^{N-1}$. Each neural network has the same structure: one input layer, two hidden layers, and one output layer. Both the input and output layers are 1 dimensional, while each hidden layer has 11 dimensions. We adopt batch normalization immediately after each matrix multiplication and before applying the ReLU activation function. At each epoch, use the Adam optimizer to update the parameters with mini-batches of 64 samples, that is, at each iteration step, we simulate 64 sample trajectories of OU process X by (25) and use them as inputs. Before training, all training parameters are initialized with a normal or a uniform distribution without pre-training. We present the architecture of the network in Fig. 1, where $h_i^j, i \in \{1, \cdots, N-1\}, j = 1, 2$ are hidden layers.

Remark 4. Fixing the initial point is more convenient for further experiments and can also help the neural network converge faster. However, when X_0 is changed, we need to re-train the neural network. Another method is to replace Z_0^ν and Y_0^ν with neural networks mapping R to R and sampling X_0 uniformly in a given set D as the initial value at each epoch.

Remark 5. Under conditions of higher regularity, Raissi [?] uses a single neural network $Y^\nu(t,x) : [0,T] \times R^d \mapsto R$ to approximate $Y_t, t \in [0,T]$ and considers the derivative of $Y^\nu(t,x)$ as suitable approximations of $Z_t, t \in [0,T]$ due to the non-linear Feynman-Kac formula. However, we do not adopt this approach because of the low regularity of our equation.

We illustrate in Figure 2 the mean loss function during an experiment and present several sample trajectories of the learned solution Y_t based on the newly generated X_t (not seen during training) in Figure 3. The average loss function decreased to around 10^{-3} after 5000 iterations.

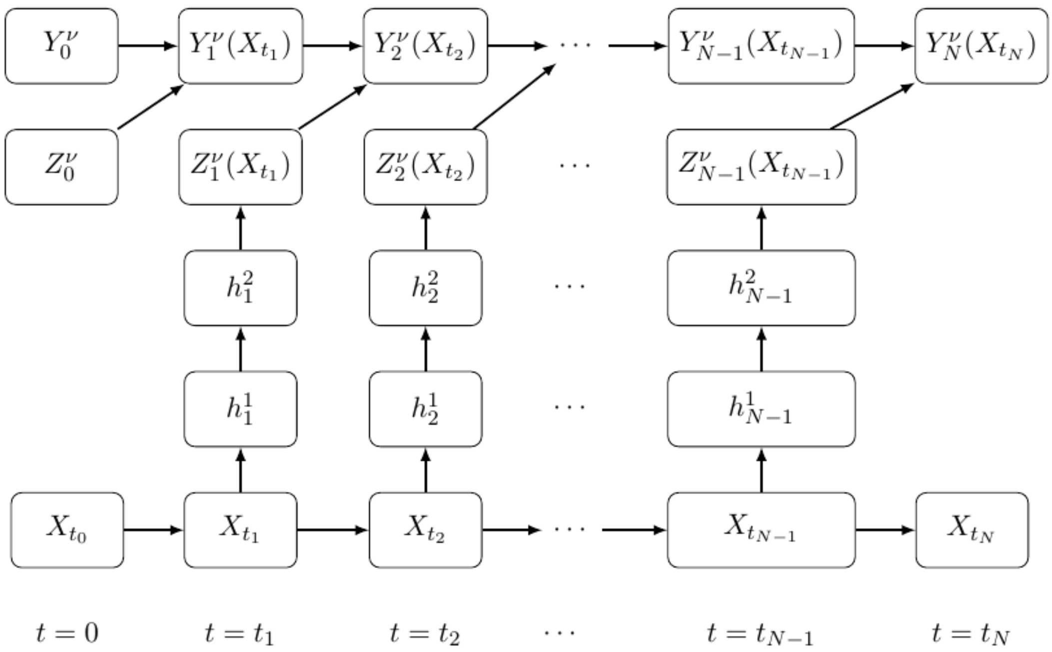

Figure 1: Architecture of Neural Network

5 Simulation Study

In this section, we conduct simulations to illustrate equilibrium strategies and the evolution of wealth under these strategies. Since the purpose of the numerical analysis is to demonstrate the equilibrium strategy under various parameter configurations

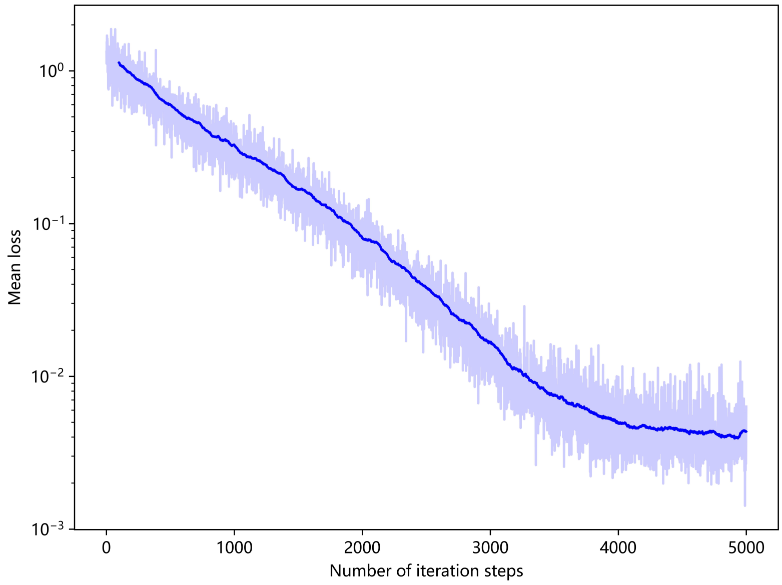

Figure 2: Mean of the loss function of an experiment with parameters $k = 0.5, \theta = 0, \eta = 0.8, \delta = 0.2, \sigma = 0.1, \rho = 0.1, \gamma = 0.1, r = 0, T = 0.5, N = 100, X_0 = 0$

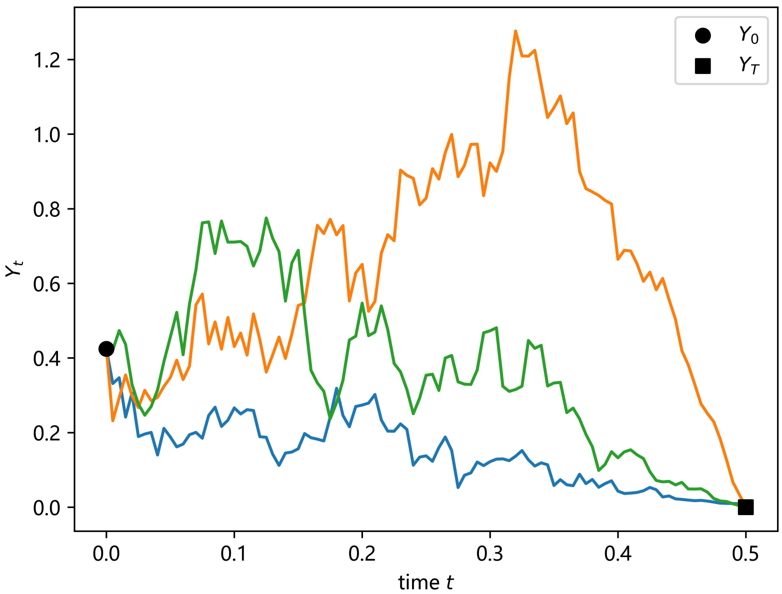

Figure 3: Learning solution Y_t

and to evaluate its performance, rather than to estimate parameters, the parameter values are fixed in advance. Unless otherwise stated, the parameter values that we use are as follows: $k = 0.5$, $\theta = 0$, $\eta = 0.8$, $\sigma = 0.1$, $\rho = 0.1$, $\gamma = 0.1$, $r = 0$, $T = 0.5$, $v_0 = 0$, $X_0 = 0$. For approaches to parameter calibration, we refer the interested readers to Mudchanatongsuk [**?**].

In Figure 4, we illustrate the equilibrium strategies under different stock borrowing fee rates without portfolio constraints. In the absence of a stock borrowing fee, the optimal investment amount is linearly negatively correlated with the price spread. This can also be discerned from the analytical solution (23). In the presence of a stock borrowing fee, there exists a region where holding a flat position in paired stocks is optimal when the price spread falls within that range. Moreover, a larger value of δ will result in a larger region of flat position. This is because only when the price spread deviates from the long-term mean by a certain extent can the profits from the strategy offset the impact of trading frictions. Besides, we can also observe that the strategy suggests a long spread position ($\hat{u} > 0$) when the spread is slightly above its long-term mean in the presence of a stock borrowing fee. This, as we explained earlier, occurs when risk premium of S^1 relative to S^2 dominates mean-reversion signal. In Figure 5, we examine the impact of time on equilibrium strategies. We observe the strategies are time-varying under the mean-variance criterion, which differs from that in the logarithmic utility case presented by Xu and Yang [**?**]. As the investment horizon approaches, the optimal amount of money invested in risky assets tends to decrease. Figure 6 illustrates the impact of spread volatility η on equilibrium strategy. Higher spread volatility implies more conservative trading behavior, as the same spread movement leads to smaller position adjustments. Moreover, at higher η, when spread is slightly above its long-term mean, the equilibrium strategy suggests taking a long position in S^1 and a short position in S^2, reflecting that increased volatility amplifies effect of risk premium of S^1. Figure 7 shows how the mean reversion rate k affects the equilibrium strategy. As expected, a lower mean reversion rate implies that the spread reverts to its long-term mean more slowly, leading to more conservative behavior with relatively smaller positions in spread. Moreover, a lower mean reversion rate k enlarges the zero-position region, as the effect of the borrowing fee becomes relatively more significant.

In Figure 8, we investigate the influence of trading constraints and observe,

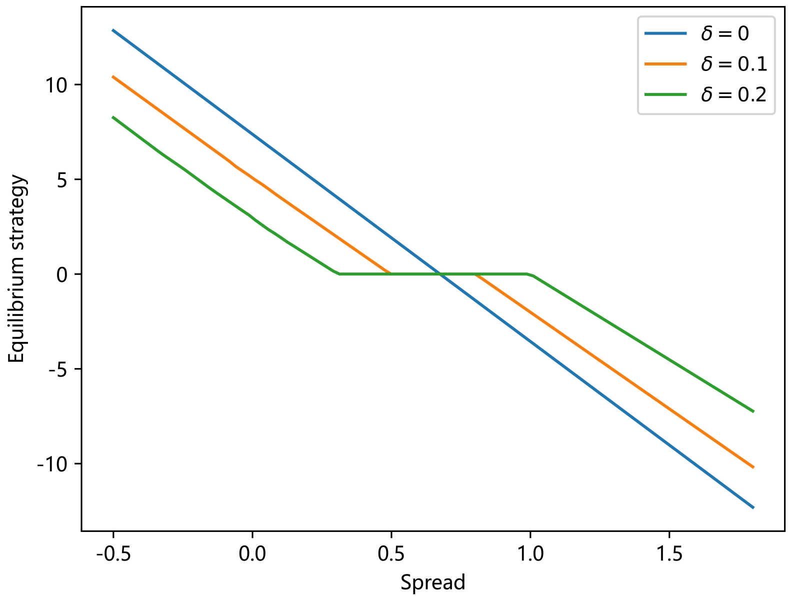

Figure 4: Equilibrium strategy across different δ at $t = 0.1$. Other parameters are taken as $k = 0.5$, $\theta = 0$, $\eta = 0.8$, $\sigma = 0.1$, $\rho = 0.1$, $\gamma = 0.1$, $r = 0$, $T = 0.5$.

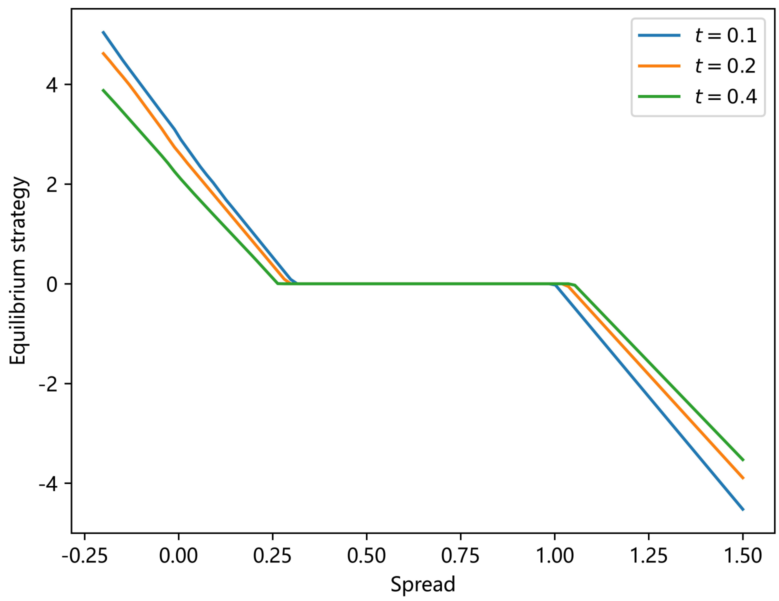

Figure 5: Equilibrium strategy across different t with $\delta = 0.2$.Other parameters are taken as $k = 0.5$, $\theta = 0$, $\eta = 0.8$, $\sigma = 0.1$, $\rho = 0.1$, $\gamma = 0.1$, $r = 0$, $T = 0.5$.

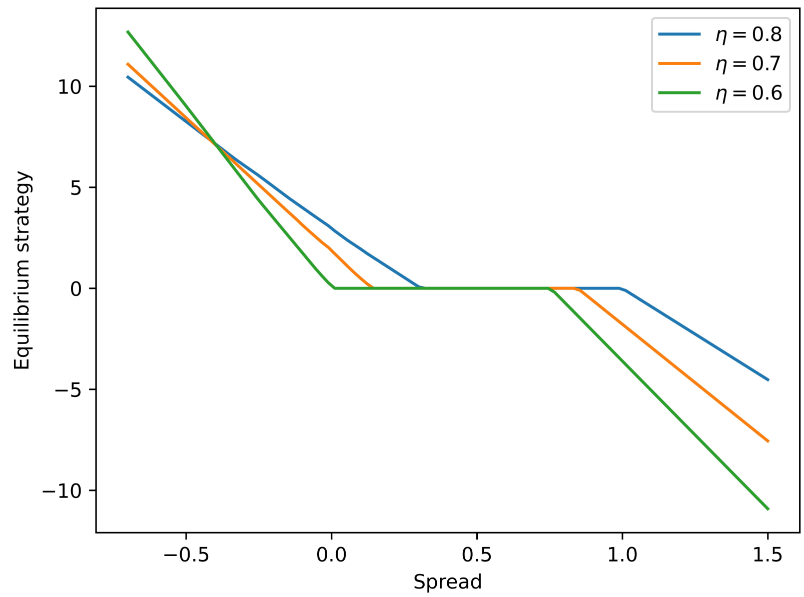

Figure 6: Equilibrium strategy across different η at $t = 0.1$. Other parameters are taken as $k = 0.5$, $\theta = 0$, $\sigma = 0.1$, $\rho = 0.1$, $\gamma = 0.1$, $r = 0$, $T = 0.5$, $\delta = 0.2$.

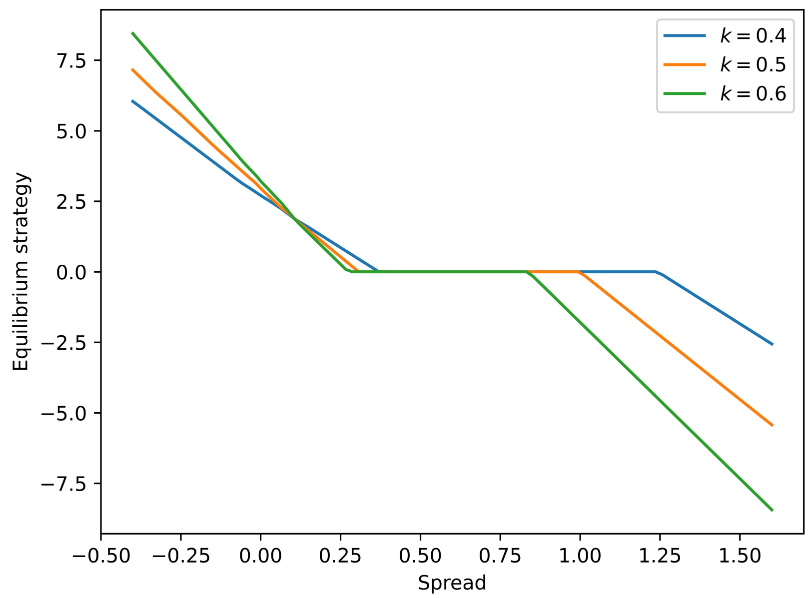

Figure 7: Equilibrium strategy across different k at $t = 0.1$. Other parameters are taken as $\theta = 0$, $\eta = 0.8$, $\sigma = 0.1$, $\rho = 0.1$, $\gamma = 0.1$, $r = 0$, $T = 0.5$, $\delta = 0.2$.

under portfolio constraints, a wider spread range where maintaining a flat position is optimal, compared to the unconstrained scenario. The optimal positions in risk assets under portfolio constraints are consistently lower compared to those without such constraints. Furthermore, we have identified an intriguing phenomenon: anticipating the likely binding of constraints in the future, investors react proactively, which implies a non-myopic effect. These findings collectively indicate that under trading constraints, our investment strategy tends to be more conservative.

To assess the effectiveness of the equilibrium strategy on terminal wealth, we

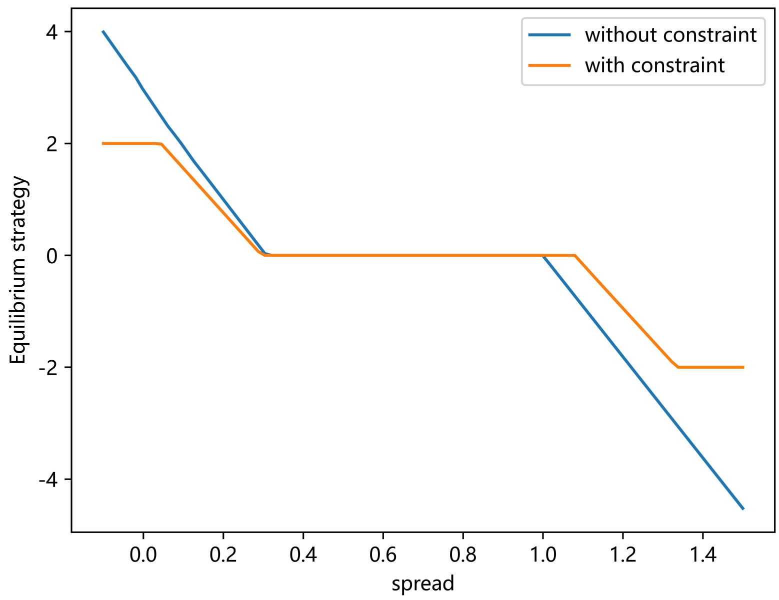

Figure 8: Equilibrium strategy with and without constraints at $t = 0.1$ and $\delta = 0.2, \overline{u} = 2, \underline{u} = -2$.

simulate a total of 10,000 sample trajectories, presenting the statistical results in Table 1 and illustrating the dynamics of wealth, strategy, and interest for a sampled path in Figure 11. Figure 11 shows the evolution of wealth along this sample path can be decomposed into three phases. First, the strategy initially takes a long position in S^1 ($\hat{u} > 0$) while the spread rises, which increases wealth, and the position size gradually decreases as the spread climbs. Second, as the spread moves into the no-trade region, $\hat{u} \approx 0$ and wealth remains essentially flat. Third, as the spread gradually reaches its peak and begins to fall, the strategy takes a short position ($\hat{u} < 0$), and this directional exposure generates further gains, resulting in a positive terminal wealth. In Table 1, as anticipated, the rise in the stock borrowing fee rate leads to a decrease in average wealth growth. However, the standard deviation of the terminal wealth and the proportion of positive terminal wealth will correspondingly increase. This suggests that transaction costs will make our investment strategy more conservative. It is intriguing to observe that, while the average terminal wealth values are consistently positive, the average loss during periods of negative terminal wealth actually exceeds the average gain when terminal wealth is positive. To explore the reasons behind this phenomenon, we plot the histogram and box plot of the wealth distribution in Figure 9 and Figure 10. The wealth distribution exhibits characteristics of heavy tails and left skewness. We can also observe that as the stock borrowing fee rate increases, the range of wealth fluctuations becomes smaller and more concentrated. Besides, when the terminal wealth ends in negative value, the average loss is substantial, which implies the necessity of contemplating the potential for significant losses with low probabilities. It enlightens us to implement efficient methods to mitigate tail risks such as posing portfolio constraints or implementing a stop-loss order. Thus, we choose an extreme sample

Table 1: Simulation result

δ	0	0.02	0.1	0.2
Average terminal wealth	1.3915	1.2539	0.8137	0.4336
Average positive terminal wealth	2.9647	2.7433	1.9915	1.2112
Average negative terminal wealth	-3.2659	-3.1436	-2.7119	-2.1996
Average interest	None	0.0626	0.2380	0.3081
Standard deviation of terminal wealth	3.3373	3.1586	2.5434	1.8608
Percentage of paths ending with positive wealth(%)	74.70	74.75	74.96	77.20

path with the greatest loss due to non-convergent spread in Figure 12, to illustrate how trading constraints can protect investors. The bounds $\overline{u}$ and $\underline{u}$ are set to 7 and -7. It can be observed that when the spread does not converge and continues to decline, the constraint mechanism is triggered, thereby reducing the losses. Finally, in the presence of borrowing fees, we compare the performance of the optimal strategy (22) with cost-naive strategy (23), which disregards the impact of borrowing fees. We use the same data for simulation and the results are reported in Table 2. It is unsurprising that the average terminal wealth under the cost-naive strategy exceeds that of the optimal strategy, as the cost-naive strategy is, as previously noted, more aggressive. In addition, the cost-naive strategy leads to higher volatility in terminal wealth. To better compare the two strategies, we also calculate the values of the re-

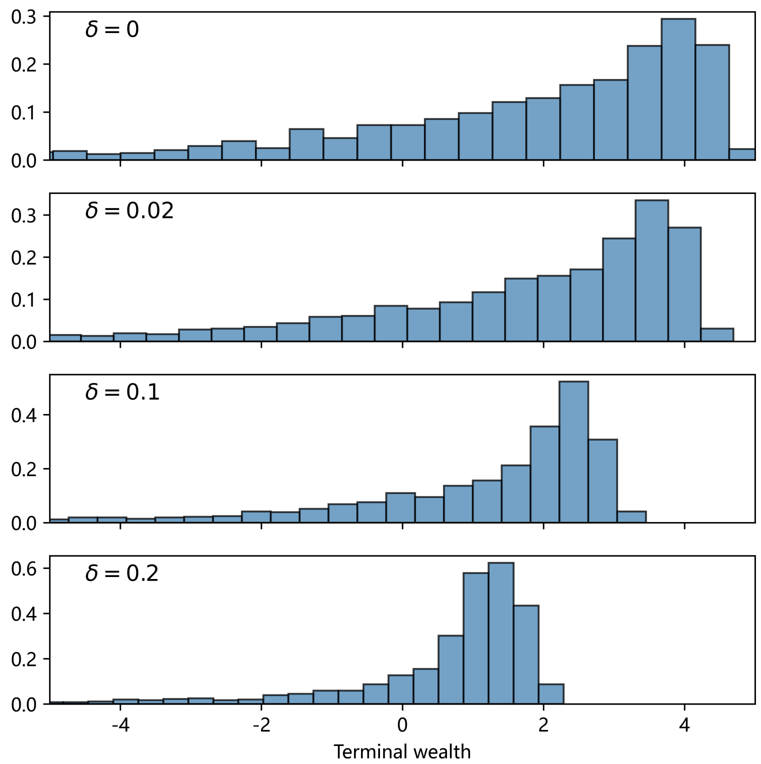

Figure 9: Histogram of terminal wealth distribution

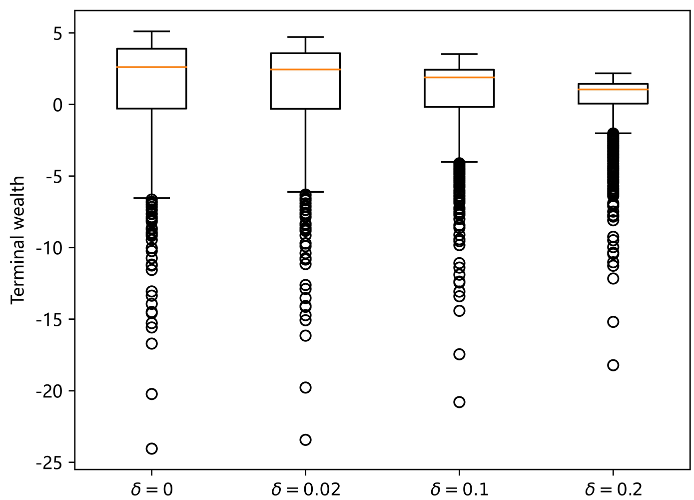

Figure 10: Box plot of wealth distribution

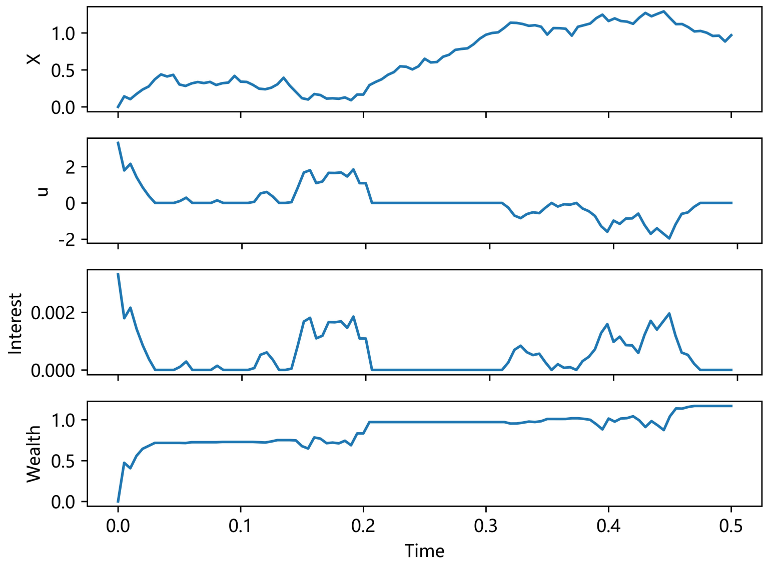

Figure 11: Dynamics of equilibrium strategy, interest and wealth for a simulated sample path

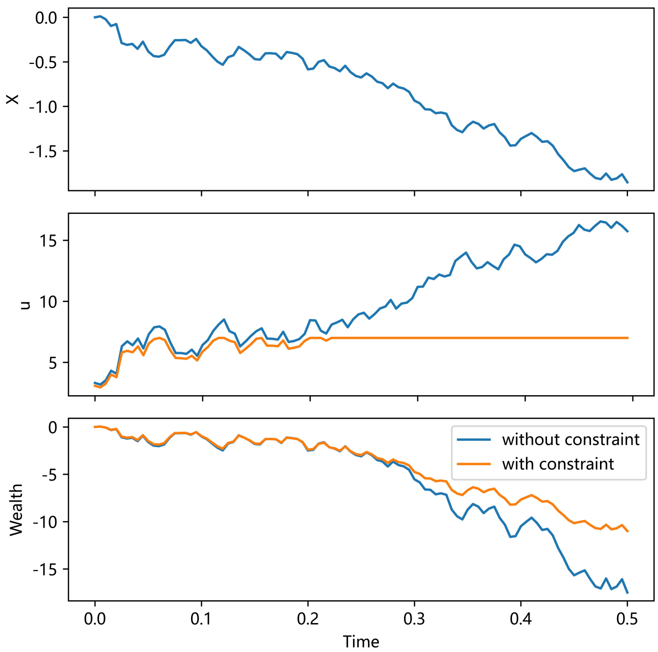

Figure 12: Dynamics of wealth under non-convergent spread

ward function (7). When the stock borrowing fee rate is very low, the performance of both strategies is similar. However, when the fee rate is sufficiently high, the optimal strategy performs significantly better. This suggests that when the stock borrowing fee rate is high, we cannot ignore the impact it has.

Table 2: The performance of optimal strategy and cost-naive strategy

	$\delta = 0.02$		$\delta = 0.1$		$\delta = 0.2$	
	Optimal	Cost-naive	Optimal	Cost-naive	Optimal	Cost-naive
Average terminal wealth	1.2539	1.3246	0.8137	1.0571	0.4336	0.7227
Average positive terminal wealth	2.7433	2.9254	1.9915	2.7902	1.2112	2.6224
Average negative terminal wealth	-3.1436	-3.3303	-2.7119	-3.3927	-2.1996	-3.5217
Average interest	0.0626	0.0669	0.2380	0.3344	0.3081	0.6688
Standard deviation of terminal wealth	3.1586	3.3571	2.5434	3.4372	1.8608	3.5389
Value of reward function	0.7551	0.7611	0.4903	0.4664	0.2604	0.0965

6 Conclusion

In this paper, we examine the pairs trading problem with stock borrowing fee under MV criterion. We investigate the equilibrium strategy from both theoretical and numerical perspectives. From a theoretical perspective, we employ a technical method to prove the existence of the solution to the FBSDE and use this solution to construct the equilibrium strategy in a semi-closed form by integrating the PDE approach with the BSDE technique. We employ the deep learning method to numerically solve the FBSDE and visualize the equilibrium strategies. We find that the presence of short selling cost makes the equilibrium strategy more conservative and that portfolio constraints mitigate myopic behavior. We also present some simulation results to test the performance of the equilibrium strategy and compare such strategy and cost-naive strategy in the presence of stock borrowing fee. The results indicate that we cannot simply ignore the costs of short-selling.

This paper sheds some light on the introduction of trading costs to pairs trading in a continuous scenario, which still requires further in-depth study in the existing literature. For future research, we plan to incorporate proportional and fixed transaction costs into our model.

Acknowledgments. The authors are very grateful to three anonymous reviewers and the AE for their constructive comments and suggestions that have helped to improve the paper of the previous version.

Disclosure of interest. The authors have no relevant financial or non-financial interests to disclose.

7 Appendix

7.1 Proof of Proposition 3.1

Proof. We regard h as function of (t, y) defined on $[0, \infty) \times (-\infty, +\infty)$. Note that h is locally Lipschitz continuous. Therefore, according to the extension theorem for solution of ODE, there exists an unique solution through the initial point $(t, y) = (0, 0)$ and it can be continuously extended up to the boundary of $[0, \infty) \times (-\infty, +\infty)$. The maximal existence interval of such solution can be analyzed by studying the properties of h.

1. If $0 < p < \frac{k^2}{2\eta^2}$, it is straightforward to that the function h has two distinct zeros, denoted by y_1, y_2, satisfying $0 < y_1 < y_2$. On the interval $(0, y_1)$, we have $h(y) > 0$. Therefore, the unique solution $y(t)$ is increasing from 0 and converges to y_1 without ever attaining y_1. It implies that the maximal existence interval is $[0, +\infty)$.

2. If $p = \frac{k^2}{2\eta^2}$, there exists an unique zero of h denoted by y_0 and $h(y) \geq 0$. By similar argument as above, we know that $y(t)$ is increasing and converges to y_0. The maximal existence interval is also $[0, \infty)$.

3. If $p > \frac{k^2}{2\eta^2}$, then $h(y) > 0$. We consider the following two ODEs:

$$\bar{y}' = a\bar{y}^2 + c, \quad \bar{y}(0) = 0, \tag{27}$$

where

$$a = -\frac{2\eta^2 k^2}{k^2 + p\eta^2} + 2\eta^2 > 0, \quad c = \frac{p}{2} - \frac{k^2}{4\eta^2} > 0.$$

It is easy to check that $h(y) \geq ay^2 + c$. Solving (27), we get

$$\bar{y}(t) = \sqrt{\frac{c}{a}} \tan(\sqrt{ac}t).$$

By the comparison theorem of ODE, we have $y(t) \geq \bar{y}(t)$. Since $\lim_{t \to \frac{\pi}{2\sqrt{ac}}} \bar{y}(t) = \infty$, $y(t)$ would blow up. The maximal existence interval of $y(t)$ is $[0, \xi)$, where $\xi < \frac{\pi}{2\sqrt{ac}}$. Similarly, to estimate the lower bound of ξ, we consider the following ODE:

$$\hat{y}' = 2\eta^2 \hat{y}^2 + p, \quad \hat{y}(0) = 0. \tag{28}$$

Obviously, $h(y) \leq 2\eta^2 y^2 + p$ and the solution of (28) is given by

$$\hat{y}(t) = \sqrt{\frac{p}{2\eta^2}} \tan(\sqrt{2\eta^2 p} t).$$

By the comparison theorem again, we have $y(t) \leq \hat{y}(t)$. Since $\lim_{t \to \frac{\pi}{2\sqrt{2\eta^2 p}}} \hat{y}(t) = +\infty$, $\xi > \frac{\pi}{2\sqrt{2\eta^2 p}}$.

□

7.2 Proof of Lemma 3.2

Proof. We apply the existence theorem of quadratic BSDEs in Briand and Hu [?] to prove the the existence of solution (Y, Z) of the FBSDE (14). For the reader's convenience, we restate their theorem as follows.

Theorem 7.1. *Consider the following quadratic BSDE:*

$$dY_t = -f(s, Y_s, Z_s)ds + Z_s dW_s, \quad Y_T = \zeta, \tag{29}$$

where the generator function f is a mapping from $\Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$ satisfying:

1. *for all $t \in [0,T], (y,z) \mapsto f(t,y,z)$ is continuous;*

2. *for each $(t,z) \in [0,T] \times \mathbb{R}$,*

$$\forall y \in \mathbb{R}, \; y(f(t,y,z) - f(t,0,z)) \leq \lambda |y|^2,$$

   where $\lambda \geq 0$ is a constant.

3. *growth condition:*

$$\forall (t,y,z) \in [0,T] \times \mathbb{R} \times \mathbb{R}, \; |f(t,y,z)| \leq \alpha(t) + \varphi(|y|) + \frac{c}{2}|z|^2,$$

where $c > 0$ is constant, $\alpha(t)$ is a progressively measurable nonnegative stochastic process and φ is a deterministic continuous nondecreasing function from $\mathbb{R}_+$ to $\mathbb{R}_+$ with $\varphi(0) = 0$.

Let $|\alpha|_1 := \int_0^T \alpha(s)ds$. If $|\zeta| + |\alpha|_1$ has exponential moment of order $p > ce^{\lambda T}$, i.e., $E\exp\{p(|\zeta| + |\alpha|_1)\} < \infty$, then (29) has a solution (Y, Z) such that

$$\begin{cases} |Y_t| \leq \frac{1}{c}\log E_t\left(\exp\left(ce^{\lambda(T-t)}|\xi| + c\int_t^T \alpha(r)e^{\lambda(r-t)}dr\right)\right), \quad \forall t \in [0,T], \\ E\left(\int_0^T |Z_s|^2 ds\right) < +\infty. \end{cases} \tag{30}$$

Specifically, in our case, it is easy to check that the generator function m in BSDE (14) satisfies

$$|m(x,z)| \leq K + \left(\frac{k^2}{\gamma\eta^2} + \frac{k^2}{2\eta^2\varepsilon} + \tilde{\varepsilon}\right)x^2 + \frac{\varepsilon + \tilde{\varepsilon}}{2}z^2,$$

where $\epsilon, \tilde{\epsilon}$ are fixed constants satisfying condition (21) and K is a constant. For simple notation, we let $A_\varepsilon = \frac{k^2}{\gamma\eta^2} + \frac{k^2}{2\eta^2\varepsilon} + \tilde{\varepsilon}$ and $\alpha(t) = A_\varepsilon X_t^2 + K$. We proceed to examine condition on p that ensure $E\exp\left\{p\int_0^T \alpha(t)dt\right\} < \infty$ which is equivalent to $E\exp\left\{pA_\epsilon \int_0^T X_t^2 dt\right\} < \infty$. Since $kT < \frac{\pi}{2}$, it follows that

$$\frac{k^2}{2\eta^2 A_\epsilon} < \frac{\pi^2}{8\eta^2T^2A_\epsilon}. \tag{31}$$

Next, we show that

$$\varepsilon + \tilde{\varepsilon} < \frac{\pi^2}{8\eta^2T^2A_\varepsilon} \tag{32}$$

holds under condition (21). Note that $\varepsilon\left(\frac{k^2}{\gamma\eta^2} + \frac{k^2}{2\eta^2\varepsilon}\right) < \frac{\pi^2}{8\eta^2T^2}$ since $0 < \varepsilon < \gamma(\frac{\pi^2}{8k^2T^2} - \frac{1}{2})$. Consider the quadratic function $l(\tilde{\varepsilon}) = (\varepsilon + \tilde{\varepsilon})\left(\frac{k^2}{\gamma\eta^2} + \frac{k^2}{2\eta^2\varepsilon} + \tilde{\varepsilon}\right)$. It easy to check that when

$$0 < \tilde{\varepsilon} < -\left(\frac{k^2}{2\gamma\eta^2} + \frac{k^2}{4\eta^2\varepsilon} + \frac{\varepsilon}{2}\right) + \frac{1}{2}\sqrt{\left(\frac{k^2}{\gamma\eta^2} + \frac{k^2}{2\eta^2\varepsilon} + \varepsilon\right)^2 + 4\left(\frac{\pi^2}{8\eta^2T^2} - \frac{k^2\varepsilon}{\gamma\eta^2} - \frac{k^2}{2\eta^2}\right)},$$

we have $l(\tilde{\varepsilon}) < \frac{\pi^2}{8\eta^2T^2}$. Thus inequality (32) holds. Combing (31) and (32), we can derive that there exists p satisfying

$$\begin{cases} \dfrac{k^2}{2\eta^2} < pA_\epsilon < \dfrac{\pi^2}{8\eta^2 T^2}, \\ \varepsilon + \tilde{\varepsilon} < p. \end{cases} \tag{33}$$

The first condition in (33) can ensure $E \exp\left\{pA_\epsilon \int_0^T X_t^2 dt\right\} < \infty$ according to Corollary 3.2. Combing the second condition in (33) and the sufficient condition $p > ce^{\lambda T}$ stated in Theorem 7.1 (where, in our case, $c = \varepsilon + \tilde{\varepsilon}, \lambda = 0$), we conclude that FBSDE (14) admits a solution (Y, Z) satisfying (30). □

7.3 Proof of Theorem 3.1

Proof. The proof is divided into three parts.

1. $\hat{u}$ is an admissible control.

According to the expression of $\hat{u}$ in (22), by the Cauchy inequality we can derive

$$\hat{u}^2(t, X_t) \le C_1 X_t^2 + C_2 Z_t^2 + C_3.$$

Since X_t is an OU process and Z_t satisfies (20), we can know $E_t \int_t^T X_t^2 dt < +\infty$ and $E_t \int_t^T Z_t^2 dt < +\infty$. It follows that $E_t \int_t^T \hat{u}_t^2 dt < +\infty$. Then $\hat{u}$ is an admissible control.

2. $\hat{u}$ is an equilibrium strategy.

Let $\tilde{V}_t = e^{r(T-t)} V_t$. Then $\tilde{V}_t$ satisfies the following equation:

$$d\tilde{V}_t = [k(\theta - X_t) + \frac{1}{2}\eta^2 + \rho\sigma\eta] e^{r(T-t)} u_t dt - \delta |e^{r(T-t)} u_t| dt + \eta e^{r(T-t)} u_t dW_t \tag{34}$$

Substituting (22) into the dynamics of wealth process (34), integrating from t to T and taking the expectation, we can get

$$E_t \tilde{V}_T^{\hat{u}} - \tilde{V}_t^{\hat{u}} = E_t \int_t^T m(X_s, Z_s) ds = Y_t. \tag{35}$$

We define the perturbation strategy u_s as follows:

$$u_s = \begin{cases} w, & \text{if } t \le s \le t + h, \\ \hat{u}_s, & \text{if } t + h \le s \le T, \end{cases}$$

It is easy to check that u is also an admissible strategy. Since these two strategies are identical from $t+h$ to T, we have $\tilde{V}_T^{\hat{u}} - \tilde{V}_{t+h}^{\hat{u}} = \tilde{V}_T^u - \tilde{V}_{t+h}^u$. It follows that $E_{t+h}\tilde{V}_T^u = Y_{t+h} + \tilde{V}_{t+h}^u$ due to (35). In addition, noting that $\tilde{V}_T = V_T$, we have

$$
\begin{aligned}
& J(t, X_t, V_t; u) \\
=& E_t V_T^u - \frac{\gamma}{2} Var_t V_T^u \\
=& E_t \tilde{V}_T^u - \frac{\gamma}{2} Var_t \tilde{V}_T^u \\
=& E_t\left[E_{t+h}\tilde{V}_T^u\right] - \frac{\gamma}{2}\left\{E_t[Var_{t+h}\tilde{V}_T^u] + Var_t[E_{t+h}\tilde{V}_T^u]\right\} \\
=& E_t[Y_{t+h} + \tilde{V}_{t+h}^u] - \frac{\gamma}{2}\left\{E_t[Var_{t+h}(\tilde{V}_T^u - \tilde{V}_t^u)] + Var_t(Y_{t+h} + \tilde{V}_{t+h}^u)\right\} \\
=& E_t\left[Y_{t+h} + \tilde{V}_{t+h}^{\hat{u}}\right] + E_t[\tilde{V}_{t+h}^u - \tilde{V}_{t+h}^{\hat{u}}] - \frac{\gamma}{2}\left\{E_t[Var_{t+h}(\tilde{V}_T^{\hat{u}} - \tilde{V}_t^{\hat{u}}) + Var_t(Y_{t+h} + \tilde{V}_{t+h}^u)\right\} \\
=& E_t\left[Y_{t+h} + \tilde{V}_{t+h}^{\hat{u}}\right] - \frac{\gamma}{2}\left\{E_t[Var_{t+h}\tilde{V}_T^{\hat{u}}] + Var_t[E_{t+h}\tilde{V}_T^{\hat{u}}]\right\} \\
& + \frac{\gamma}{2} Var_t[E_{t+h}\tilde{V}_T^{\hat{u}}] - \frac{\gamma}{2} Var_t(Y_{t+h} + \tilde{V}_{t+h}^u) + E_t[\tilde{V}_{t+h}^u - \tilde{V}_{t+h}^{\hat{u}}] \\
=& E_t V_T^{\hat{u}} - \frac{\gamma}{2} Var_t V_T^{\hat{u}} + \frac{\gamma}{2} Var_t[Y_{t+h} + \tilde{V}_{t+h}^{\hat{u}}] - \frac{\gamma}{2} Var_t(Y_{t+h} + \tilde{V}_{t+h}^u) + E_t[\tilde{V}_{t+h}^u - \tilde{V}_{t+h}^{\hat{u}}] \\
=& J(t, X_t, V_t; \hat{u}) + \frac{\gamma}{2} Var_t[Y_{t+h} + \tilde{V}_{t+h}^{\hat{u}}] - \frac{\gamma}{2} Var_t(Y_{t+h} + \tilde{V}_{t+h}^u) + E_t[\tilde{V}_{t+h}^u - \tilde{V}_{t+h}^{\hat{u}}].
\end{aligned}
\tag{36}
$$

For simple notation, we introduce $a(t, x, u) = \mu(x)e^{r(T-t)}u - \delta|e^{r(T-t)}u|$. Therefore,

$$
\begin{cases}
\tilde{V}_{t+h}^u - \tilde{V}_t = \int_t^{t+h} a(s, X_s, u_s)ds + \eta e^{r(T-t)} u_s dW_s, \\
\tilde{V}_{t+h}^{\hat{u}} - \tilde{V}_t = \int_t^{t+h} a(s, X_s, \hat{u}_s)ds + \eta e^{r(T-t)} \hat{u}_s dW_s.
\end{cases}
\tag{37}
$$

It follows that

$$
E_t\left[\tilde{V}_{t+h}^u - \tilde{V}_{t+h}^{\hat{u}}\right] = E_t \int_t^{t+h} \left[a(s, X_s, u_s) - a(s, X_s, \hat{u}_s)\right] ds. \tag{38}
$$

On the other hand, since $a(s, X_s, \hat{u}_s) = m(X_s, Z_s)$, we have

$$
Y_{t+h} - Y_t = -\int_t^{t+h} a(s, X_s, \hat{u}_s)ds + \int_t^{t+h} Z_s dW_s. \tag{39}
$$

Combining (37) and (39), we can get

$$\begin{aligned}\tilde{V}^{\hat{u}}_{t+h}+Y_{t+h}=&\tilde{V}_t+Y_t+\int_t^{t+h}\Big[a(s,X_s,u_s)-a(s,X_s,\hat{u}_s)\Big]ds\\&+\int_t^{t+h}\eta e^{r(T-t)}\hat{u}_s dW_s+\int_t^{t+h}Z_s dW_s\end{aligned}\tag{40}$$

Taking the conditional variance of both sides of (40), we have

$$\begin{aligned}&Var_t(\tilde{V}^{\hat{u}}_{t+h}+Y_{t+h})\\=&E_t\int_t^{t+h}\Big[\eta e^{r(T-t)}\hat{u}_s+Z_s\Big]^2ds\\&+Var_t\left\{\int_t^{t+h}\Big[a(s,X_s,u_s)-a(s,X_s,\hat{u}_s)\Big]ds\right\}\\&+Cov\Big(\int_t^{t+h}\Big[a(s,X_s,u_s)-a(s,X_s,\hat{u}_s)\Big]ds,\int_t^{t+h}\Big[\eta e^{r(T-t)}u_s+Z_s\Big]dW_s\Big)\\=&E_t\int_t^{t+h}\Big[\eta e^{r(T-t)}\hat{u}_s+Z_s\Big]^2ds+o(h).\end{aligned}\tag{41}$$

Similarly,

$$Var_t(\tilde{V}^{u}_{t+h}+Y_{t+h})=E_t\int_t^{t+h}\Big[\eta e^{r(T-t)}u_s+Z_s\Big]^2ds+o(h).\tag{42}$$

Substituting (38), (41) and (42) into (36), we obtain

$$J(t,X_t,V_t;\hat{u})-J(t,X_t,V_t;u)=E_t\int_t^{t+h}\Big[R(s,X_s,\hat{u}_s)-R(s,X_s,u_s)\Big]ds+o(h),\tag{43}$$

where

$$R(t,x,u)=-\gamma\eta^2e^{2r(T-t)}u^2/2+\mu(x)e^{r(T-t)}u-\gamma\eta e^{r(T-t)}uZ_t-\delta|e^{r(T-t)}u|.$$

Dividing both sides of (43) by h and let $h\to 0+$, by Dominate Convergence Theorem, we have

$$\liminf_{h\to 0+}\frac{J(t,X_t,V_t;\hat{u})-J(t,X_t,V_t;u)}{h}=R(t,X_t,\hat{u}_t)-R(t,X_t,u_t).$$

Recalling Lemma 3.1, $R(t, X_t, \cdot)$ is maximized at $\hat{u}_t$, which implies

$$\liminf_{h\to 0+} \frac{J(t, X_t, V_t; \hat{u}) - J(t, X_t, V_t; u)}{h} \geq 0.$$

Consequently, $\hat{u}$ is an equilibrium strategy.

3. f is a classical solution of (13).
According to the Markov property of the FBSDE (14), there exists a deterministic function f such that $Y_t = f(t, X_t)$. If $f \in C^{1,2}$, applying Itô's formula to f, we have

$$df(t, X_t) = \Big[f_t + k(\theta - X_t)f_x(t, X_t) + \frac{1}{2}\eta^2 f_{xx}(t, X_t)\Big]dt + \eta f_x(t, X_t)dW_t.$$

Comparing it with

$$-dY_s = m(X_s, Z_s)ds - Z_s dW_s,$$

we can derive that $Z_t = \eta f_x(t, X_t)$ and f satisfies (13).

□

7.4 Proof of Corollary 3.3

Proof. According to Theorem 3.1, we just need to solve the following PDE:

$$f_t - \Big(\frac{1}{2}\eta^2 + \rho\sigma\eta\Big)f_x + \frac{1}{2}\eta^2 f_{xx} = -\frac{\mu^2(x)}{\gamma\eta^2}, \tag{44}$$

with terminal condition $f(T, x) = 0$. We make an ansatz for f:

$$f(t, x) = a_2(t)\mu^2(x) + a_1(t)\mu(x) + a_0(t). \tag{45}$$

Substituting (45) into (44) and comparing the coefficients, we obtain a system of ODEs:

$$\begin{cases} a_2'(t) = -\dfrac{1}{\gamma\eta^2}, \quad a_2(T) = 0, \\ a_1'(t) + (k\eta^2 + 2k\rho\sigma\eta)a_2(t) = 0, \quad a_1(T) = 0, \\ a_0'(t) + \Big(\dfrac{1}{2}k\eta^2 + k\rho\sigma\eta\Big)a_1(t) + k^2\eta^2 a_2(t) = 0, \quad a_0(T) = 0. \end{cases}$$

Solving the system yields:

$$\begin{cases} a_2(t) = \dfrac{T-t}{\gamma\eta^2}, \\ a_1(t) = \dfrac{k\eta^2 + 2k\rho\sigma\eta}{2\gamma\eta^2}(T-t)^2, \\ a_0(t) = \dfrac{(k\eta^2 + 2k\rho\sigma\eta)^2}{12\gamma\eta^2}(T-t)^3 + \dfrac{k^2}{2\gamma}(T-t)^2. \end{cases}$$

Substituting f into (22), we obtain (23).

□