Research Article

Journal of Financial Intermediation, Sustainable Banking and Marketing

Optimal Pairs Trading Strategies with Flexible Position Sizing and Stop-Loss Boundaries

Yitong Niu, & Huayue Zhang*

School of Finance, Nankai University, Tianjin, 300071, p. R. China

*Corresponding author: Huayue Zhang, School of Finance, Nankai University, Tianjin, 300071, p. R. China.

Submitted: 12 January 2026 **Accepted:** 19 January 2026 **Published:** 27 January 2026

Citation: *Zhang, H, & Niu, Y. (2026). Optimal Pairs Trading Strategies with Flexible Position Sizing and Stop-Loss Boundaries. J of Fin Int Sus Ban Mar, 2(1), 01-15.*

Abstract

Our study aims to investigate the profitability of an optimized mean reversion trading strategy in the context of the US equity market. In contrast to conventional pair trading strategies, a thorough method that combines time series, stochastic control techniques, and cointegration is utilized to build the best static pair trading portfolio for the Ornstein-Uhlenbeck process, with parameters that are precisely calculated. Initial assessment in- volves a comprehensive evaluation of multiple metrics to ensure the selected pairs exhibit pre-trade mean reversion characteristics. Subsequently, Ornstein Uhlen beck process parameters are finely tuned to address the varying degrees of stationarity observed in different spread scenarios. Dynamic contrarian trading signals are then derived from model parameters, with thresholds and in-sample period lengths optimized through iterative testing. Analysis of historical data pertaining to five pairs demonstrates that the proposed pair trading strategy outperforms traditional cointegrated pairs, yielding higher returns both within and beyond sample periods, with an average excess annualized return exceeding 8%. Notably, the strategy's adaptability is highlighted by the dynamic adjustment of model parameters and trading strategies over time, including position sizing, directional bias, and stop- loss thresholds, thereby enhancing robustness and adaptability. Furthermore, validation of the model's ability to swiftly adjust portfolios in response to high-risk events validates its effective- ness in mitigating risks while maximizing returns.

Keywords: Pair Trading, Cointegration, Stop-Loss Aversion, Time Series.

Introduction

The first practice of statistical pairs trading is attributed to Wall Street quant Nunzio Tartaglia, who was at Morgan Stanley in the mid-1980s. Pairs trading is a common hedge fund strategy in which one trades based on a model of the relative value between a pair of stocks; e.g. see the books by Vidyamurthy (2004) and Whistler (2004). In pairs trading, it is important to determine when to initiate a pairs trade (i.e., how much divergence is sufficient to trigger a trade) and when to close the position (when to lock in profits if the stocks perform as expected or when to cut losses if the trade goes sour) [1].

It is the purpose of this paper to focus on the mathematics of pairs trading. In particular, we consider the case when a difference of a pair satisfies a mean reversion model with transactions cost to determine these key thresholds, and find the optimality.

Mean-reversion models are often used in financial markets to capture price movements that have the tendency to move towards an "equilibrium" level. There are many studies in connection with mean reversion stock returns; For example, a common approach is to define the spread between two stocks as the price of the first stock minus a multiple of the price of the second stock. Typically, this spread will be mean-reverting around some value. When the spread deviates from its average value, one then makes trades based on the assumption of a return to the average. Using quantitative models that capture this basic idea, a number of researchers have proposed methods for profitably trading pairs; e.g. see Elliott et al. (2005), Gatev et al. (2006), Do et al. (2006), Mudchanatongsuk et al. (2008), Tourin and Yan (2013), Song and Zhang (2013), Deshpande and Barmish (2016) and references therein. Mathematical trading rules have been studied for many years. For example, Suleyman Basak (2010) used dynamic programming to derive the Hamilton-Jacobi-Behrman (HIB) equation, and the model solved the time inconsistency problem in the dynamic mean-variance portfolio problem.

Moreover, from the viewpoint of transaction costs, the practical for business and the trans- action cost with respect to the amount of rebalancing affects the optimal trading strategy and trading rule. The issue of trading under transaction costs has a long history in finance, and the literature is extensive. In the context of portfolio selection, early significant work includes Magill and Constantinides (1976), Constantinides (1979), and Davis and Norman (1990), Davis et al. (1993), Chan and Lakonishok (1995),Keim and Madhavan (1997), and Jones and Lipson (2001) . In this paper, we try to find an optimal pairs trading rule in which a pairs (long-short) position consists of a long position of one stock and a short position of the other. A fixed (commission or slippage) cost will be imposed to each transaction to optimally trade pairs while explicitly incorpo- rating the effects of transaction costs, and to explore its characteristics by numerical simulations.

Estimation of Spread Process Parameters

Although the prices of two highly correlated stocks may be unstable individually, their linear combination will fluctuate around a fixed level. This section presents two ways to estimate stock portfolio spreads.We begin by describing the prices, spread, and wealth dynamics.

Price, Spread and Wealth Dynamics

Consider a pair of two stocks P and Q, and Pt and Q_t denote their prices at time t,respectively.We adopted the cointegration method to construct stock pairs, assuming that the log- arithmic prices of two stocks satisfy a cointegration relationship.Expanding on we modified the logarithmic spread to a non-proportional form. Addi- tionally, the trading unit proportion dynamically changes over time based on the linear fit results of a rolling window period. Denote Xt the difference of the logarithms of the two stock prices, i.e.,

$$X_t = \ln P_t - \beta \ln Q_t, \tag{2.1}$$

where β is the cointegration coefficient which is a time varying parameter obtained through linear fitting of the data over a rolling window period of 200 days and given later.We assumed that the spread follows an Ornstein-Uhlenbeck (OU)process:

$$dX_t = k(\mu - X_t)dt + \sigma dB_t, \tag{2.2}$$

where k ∈ R+ is the speed of mean reversion, μ R is the long-term equilibrium level to which the spread reverts, σ R+ measures the strength of noise interference,and Bt is a standard Brownian motion defined on a filtered probability space. Let Wt be the value of a self-financing pairs-trading portfolio and let πt be the number held for stocks P and Q at time t. Then, the wealth dynamics of the portfolio value is given by: $dW_t = \pi_t dX_t = k\pi_t(\mu - X_t)dt + \sigma\pi_t dB_t$It is easy to see that changes in wealth function W is a function of the control process πt and k, μ, σ. Therefore, a model that accurately estimates the OU process parameters is crucial for solving the subsequent optimization problem to get the optimal *pi**. Next, we first give two ways to estimate parameters for the two stationary classification situations of the sequence.

Estimation of Parameters for Stationary Mean-Reverting Process

Due to the contamination of high-frequency data by microstructure noise in the market, ne- glecting biased parameter estimation may lead to significant estimation bias. It is crucial to employ estimation models robust to noise E_i N (0, σ²). The research findings of [2].Indicate the necessity of considering this aspect in the con- text of OU processes observed at discrete equidistant time intervals, contaminated by independent Gaussian white noise (additive noise model) $X_i = X_i - E_i, i = 0, \ldots, n$ It follows an ARMA(1,1) process:

$X_t = \alpha + \varphi X_{i-1} + \theta S_{i-1} + S_i, S_i \square N(0, \sigma^2).$

We only observe the process Xi at a finite number of discrete times $0 = T_0 < T_1 < \cdots T_n = 1$.Without loss of generality,we let t ∈ [0, 1], and $\Delta = T_i - T_{i-1} = {}^{1}$ is the time period between observations. Then we provide its maximum likelihood estimation results. After verifying the stationarity of the spread in the testing window, our model initially assumes that daily closing price data at a frequency of one day was an OU process contaminated by independent white noise. We fitted the parameters of the ARMA(1,1) model to the known time series data of the spread and then estimated the parameters of the OU process. However, we observed that the estimation results for our standard spread parameters using the ARMA(1,1) model were particularly unsatisfactory. The model fitting report indicates poor parameter fitting for the MA(1) process, with a significant estimation bias.

```
                              ARMA Model Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  100
Model:                     ARMA(1, 1)   Log Likelihood                 379.862
Method:                       css-mle   S.D. of innovations              0.005
Date:                Thu, 23 Nov 2023   AIC                           -751.724
Time:                        15:04:32   BIC                           -741.303
Sample:                             0   HQIC                          -747.507

==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.7094      0.006    120.384      0.000       0.698       0.721
ar.L1.y        0.9065      0.049     18.388      0.000       0.810       1.003
ma.L1.y        0.1145      0.084      1.357      0.178      -0.051       0.280
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.1031           +0.0000j            1.1031            0.0000
MA.1           -8.7373           +0.0000j            8.7373            0.5000
```

Figure 1: Arima Model Results

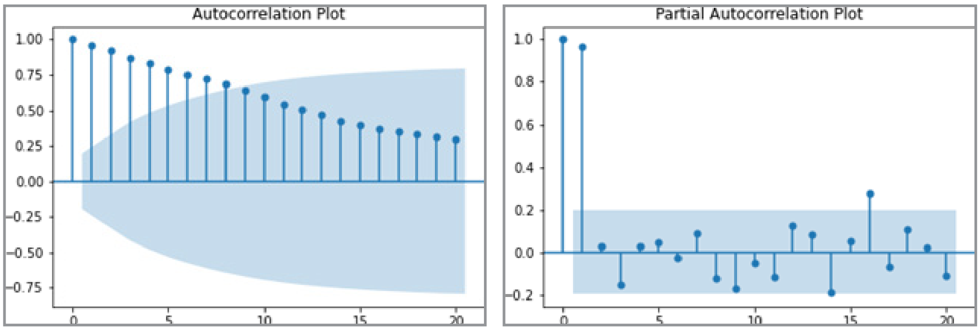

Figure 2: ACF and PACF Plot

From the Figure 2 of the autocorrelation function (ACF) and partial autocorrelation function (PACF), it is evident that the ACF of daily closing price data exhibits a tail, while the PACF is truncated at lag one. This suggests that, in practical terms, the standardized spread, after taking logarithms, approximately follows a discrete AR(1) model. Therefore, the logarithmically standardized daily spread data is relatively stationary, with negligible influence from white noise.

Therefore, when noise is negligible, it is evident that the discretized OU process corresponds to an AR(1) process. Here we predicted the stationary sequence using the AR(1) model:

$X_t = \alpha + \varphi X_{i-1} + L_i,\ V_i \sim N(0, \sigma^2)$(2.6)

Thus, the parameters of the Ornstein-Uhlenbeck process are obtained through the maximum like- lihood method

$$\hat{k} = \frac{1}{\Delta}\log\hat{\varphi}$$

$$\hat{\mu} = \frac{\hat{\alpha}}{1-\hat{\varphi}}$$

$$\hat{\sigma}^2 = -2\frac{1}{\Delta}\frac{\hat{\sigma}_l^2}{1-\hat{\varphi}^2}\log\hat{\varphi}$$

The hat of the symbol represents an estimate derived directly or indirectly from an observation.

Estimation of Parameters for Non-Stationary Mean-Reverting Process

In the previous section we proved that the mean reversion process of stock spreads in the real world is negligibly affected by noise. Therefore, when our price difference series fails the stability test, we can also approximately regard it as an OU process. However, since the time series AR (1) model is used to fit the parameters, the price difference needs to satisfy strict stationarity. In this case not applicable. Therefore, for price difference sequences that fail the strict stationarity test, we provide a way to approximately estimate the parameters of the OU process using numerical simulation methods [3]. demonstrated that the parameters of the discrete Vasicek model belonging to the OU process could be obtained through the maximum likelihood method:

$$\hat{k} = -\delta^{-1}\log(\hat{\beta}_1)$$

$$\hat{\mu} = \hat{\beta}_2 \quad \hat{\sigma}^2 = 2\hat{k}\hat{\beta}_3\left(1-\hat{\beta}_1^{2}\right)^{-1}$$

where:

$$\hat{\beta}_1 = \frac{n^{-1}\sum_{i=1}^{n} X_i X_{i-1} - n^{-2}\sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_{i-1}}{n^{-1}\sum_{i=1}^{n} X^2{}_{i-1} - n^{-2}\left(\sum_{i=1}^{n} X_{i-1}\right)^2}$$

$$\hat{\beta}_2 = \frac{n^{-1}\sum_{i=1}^{n}(X_i - \hat{\beta}_1 X_{i-1})}{1-\hat{\beta}_1}$$

$$\hat{\beta}_3 = n^{-1}\sum_{i=1}^{n}\{X_i - \hat{\beta}_1 X_{i-1} - \hat{\beta}_2(1-\hat{\beta}_1)\}^2$$

Therefore, we first examined the stationarity of the spread through the ADF method. The stationary spread obtained from the window period of the stationarity test was used as the training sample for fitting the AR model. The parameters of the AR(1) process were then reverse-calculated to obtain the parameters of the OU process. For the non-stationary spread, we used the mean and variance of the window period, as well as the spread on the first day of the window period, as initial values. We iteratively seeked the numerical solution for the model, and we checked that the deviation at the end of each window period was relatively small. Below, we calculated the residuals of the estimated daily spread and presented the residual plot and boxplot of residuals.

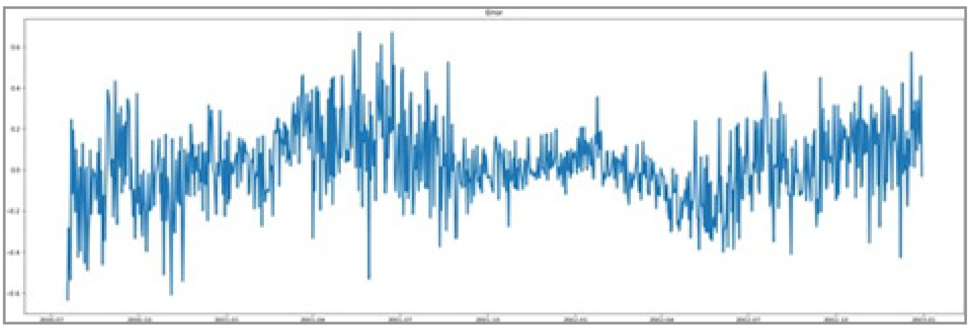

Figure 3: Residual Plot

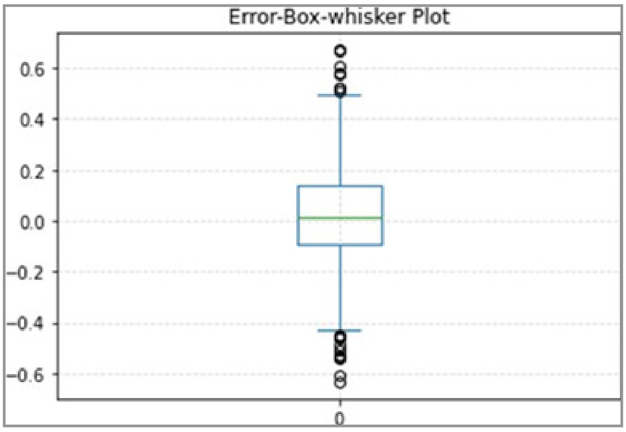

Figure 4: Boxplot of Residuals

By examining the trend plot of the model residuals, it is evident that the residuals fluctuate around 0, indicating relative stability. According to the boxplot, the median and mean of the residuals are approximately 0, suggesting that the model does not exhibit significant overestimation or underestimation. The quartiles of the residuals are relatively small, indicating that the model's predictions are stable with minimal fluctuation. While there are occasional larger outliers on a few days, they do not exceed the range of the y-axis. More than 50% of the residuals are within 0.2,indicating a high concentration and a relatively symmetrical distribution of residuals. Furthermore, we calculated the Residual Sum of Squares (RSS) for the model as 32.94, and the Mean Squared Error (MSE) as 0.03672. This suggests that our model has a small estimation error compared to the true data, indicating a relatively accurate estimation result.

Optimal Pair Trading Strategy

Optimal Problem and the Solution of HJB Equation

Assumption

Assuming it is the position of our trading portfolio at time and it can take any real number, in short combination in the opposite direction).

According to the above Assumptions, we formulate the portfolio optimization pair-trading prob- lem as a stochastic optimal control problem. We assume that an investor's preference can be represented by the utility function U (x) = ^{1} x$^{\gamma}$, with x 0 and $\gamma < 1$. Therefore, our objective is to maximize expected utility at the final time T .Thus, we seek to solve

$$\max_{\pi\in\Pi} \quad \sup_{\pi_t} \quad E[\tfrac{1}{\gamma}W_T{}^{\gamma}]$$
$$\text{subject to} \quad dW_t = k\pi_t(\mu - X_t)dt + \sigma\pi_t dB_t \tag{3.1}$$

Remark

To solve the optimization problem, we applied transformations and time dilation using Itoˆ's lemma to convert the equation of the OU process (Equation (2)) into a dimensionless system[4].

$$X \rightarrow \frac{(X-\mu)}{\sigma}\sqrt{\text{k}}$$
$$\pi \rightarrow \frac{\pi}{\sqrt{k}}\sigma$$
$$t \rightarrow kt$$

Note that in the standardized case, terminal wealth does not change. Remark 3.1.2: According to the equation of theorem 3.1, we have:

$$dX_t = -X_t dt + dW_t \tag{3.2}$$

We refered to Equation (8) as the dimensionless system because Xt does not depend on the model parameters. As the above transformation is linear, each Xt value in the dimensionless system corre- sponds to a unique original system Xt value. Let J(t, w, x) denote the value function. By standard arguments, one may show that the Hamilton-Jacobi-Bellman (HJB) equation coresponding to our stochastic control problem is

$$\sup_{\pi_t}(J_t - xJ_x - \pi xJ_w + \frac{1}{2}J_{xx} + \frac{1}{2}\pi^2 J_{ww} + \pi J_{xw}) = 0 \tag{3.3}$$

subject to the terminal condition

$$J(T,w,x) = \frac{1}{\gamma}w^{\gamma} \tag{3.4}$$

where the subscripts on J denote partial derivative. The first-order condition for the optimal position in the dimensionless system is obtained as:

$$\pi^*(w,x,t) = x\frac{J_w}{J_{ww}} - \frac{J_{xw}}{J_{ww}} \tag{3.5}$$

Remark

According to Remark and Theorem 3.2,substituting this first-order condi- tion(28) into the Hamilton-Jacobi-Bellman equation (27), we obtained the nonlinear partial differ- ential equation:

$$J_t + \frac{1}{2}J_{xx} - xJ_x - \frac{1}{2}J_{ww}(\frac{J_{xw}}{J_{ww}} - x\frac{J_w}{J_{ww}})^2 = 0 \tag{3.6}$$

Next, we solved this nonlinear differential equation.

Determination of the Optimal Trading

extended the Andrew Morton model under simple as- assumptions and provided a well-defined explicit solution to the Hamilton-Jacobi-Bellman (HJB) equation. Let $\tau = T - t$, the following is defined [5]:

$$\nu = \frac{1}{\sqrt{1-\gamma}} \tag{3.7}$$

$$C(\tau) = \cosh\nu\tau + \nu\sinh\nu\tau \tag{3.8}$$

$$C'(\tau) = \frac{\mathrm{dC}(\tau)}{\mathrm{d}\tau} = \nu\sinh\nu\tau + \nu^2\cosh\nu\tau(3.9) \tag{3.9}$$

$$D(\tau) = \frac{C'(\tau)}{C(\tau)} \tag{3.10}$$

Theorem 3.3: For $\gamma<0$ or $0<\gamma<1$, the optimal strategy is:

$$\pi_t^* = -wxD(\tau) \tag{3.11}$$

Under the conclusion ,the value function is given by the following expression:

$$J(w,x,t) = \frac{1}{\gamma}w^{\gamma}\sqrt{e^{\gamma}C(\tau)^{\gamma-1}}\exp(\frac{x^2}{2}(1+(\gamma-1)D(\tau))) \tag{3.12}$$

Where τ, $C(\tau)$ and $D(\tau)$ are defined by equations (12)-(14), and $X_t = x, W_t = w$.

Determination and Discussion of Trading Rules
Building upon the optimal trading positions obtained from the optimization process described as above, we extended the model to incorporate trading rules considering transaction costs. These rules include opening and closing conditions (timing for entry and exit), stop-loss thresholds, and stop-loss conditions. We also discussed factors that generally affect investor value functions and trading positions, providing practical interpretations. Discussion on Entry and Exit Timing

Suggests that in the U.S. stock market, the performance of distance-based currency pair trading slightly improves when the VIX in the trading system increases or is high. However, this improvement is not sufficient to have statistical and economic significance after deducting transaction costs. This indicates that the profits of pair trading quickly appear after the deviation between the two assets, and incorporating VIX timing into our trading strategy adds little economic value to currency trading.

Therefore, our model immediately enters when there is a trading opportunity (ie. profit after deducting all costs is greater than 0), and no opening threshold is set. The profitability of pair trading is closely related to the mean-reverting process of the spread.

However, due to the influence of external market conditions, the spread may deviate significantly, and there is a risk of being difficult to recover to the original mean level. In such cases, it is necessary to set an effective stop-loss threshold that does not affect profitability [6].

As we know, the direction of constructing positions in pair trading is opposite to the direction of deviation of the spread from the mean, and we waited for an opportunity for the spread to revert to the mean. Theorem 4.1: According to Ito's lemma, the diffusion term of $d\pi t$ can be obtained from Equation Therefore, the correlation coefficient between A and B is given by:

$$\mathrm{cov}(d\pi, dX) = -D(\tau)(W_t + \pi_t X_t) = W_t D(\tau)(-1 + X^2 D(\tau)) \tag{4.1}$$

Similar to the model , we obtained that the expression in Equation (36) was negative. This implies that the range where the spread and the direction of position changes are opposite is:

$$|X| \leq \sqrt{1/D(\tau)} \tag{4.2}$$

Remark
According to the first equation of Equation (36), another way to understand this threshold is obtained: once the losses brought by the position exceed the existing wealth, the loss positions start to be reduced.

Remark
We transform it back to the dimensionless system, the stop-loss threshold is:

$$|X - \mu| \leq \sigma \sqrt{\frac{1}{kD(\frac{\tau}{k})}} \tag{4.3}$$

Remark
According to the Inequality (38) of Remark 4.1.2, It can be seen that this dynamic stop-loss threshold has a good interpretability:

1. The original system's stop-loss condition includes σ, so the triggering condition for stop-loss is also related to volatility. When the volatility of the spread is relatively large, the threshold will correspondingly increase, avoiding frequent triggering of the threshold and affecting the profit level.
2. The stop-loss condition includes the function ν of γ, so this stop-loss condition is also related to consumer risk aversion.
3. The stop-loss condition includes the regression speed k, indicating that the regression speed of the spread over the time period will affect the stop-loss threshold.
4. The stop-loss condition is related to the remaining time τ.

Demonstrated through simulations how $D(\tau)$ depends on the remaining time τ for different values of γ. The results show that for investors with relatively low risk aversion γ, as τ increases, $D(\tau)$ gradually decreases and stabilizes after reaching a certain threshold.

Below we consider other stop loss thresholds: the Average True Range (ATR) indicator was first introduced by J. Welles Wilder. This indicator is commonly used in stock and commodity markets and can effectively assist traders in anticipating possible future price volatility, providing valuable help in setting stop-loss or take-profit targets. The ATR represents the concept of volatility, illustrating traders' expectations and enthusiasm, and reflecting the level of market trading activity. Therefore, we applied it to our spread system as a stop-loss threshold.

This indicator represents the average trading range of price movements over a period of N days, typically with a time period of 14 trading days. The calculation method is as follows.
First, calculate the True Range:

$$TR_t = \max(hight - lowt, |closet-1 - hight|, |closet-1 - lowt|)$$

Where high, low , close represent the highest price, lowest price, and closing price, with subscripts denoting time. Evidently, TR_t is the maximum of the following three volatility measures: the distance between the highest and lowest points of the day, the distance between the previous day's closing price and the current day's highest price, and the distance between the previous day's closing price and the current day's lowest price.

On the basis of TR_{it}, $i = A, B$, taking a certain time period (usually the average of 14 periods by default), the Average True Range (ATRit, i = A, B) can be obtained. Therefore, the ATRit of two stocks is added according to the pairing ratio and average is calculated over a window period of 200 days, resulting in the rolling threshold ATR_t for the stock pair's spread. In addition, traditional pair trading often uses measures such as the volatility of spread devi- ation from the mean, multiples of standard deviation, historical price percentage, and so on, as stop-loss thresholds.

Therefore, we compared six stop-loss thresholds, each corresponding to two stop-loss strate- gies: maintaining positions without further accumulation and liquidation. We compared these six strategies against the stock pairs: COCA COLA and PEPSICO, during bull and bear markets (January 1, 2000, to December 31, 2002), examining their respective returns.

Table 1: The Results of Six Stop-Loss Thresholds with Two Stop-Loss strategies

Negative correlation	coefficient	Standard deviation	Historical rolling volatility	Historical percentage method (30%)	Cointegration test	ATR method
liquidation	1.2080	1.2004	1.2057	1.0211	1.0124	1.1925
Maintaining positions	1.2088	1.2006	1.2057	1.061	1.0106	1.2080

Negative correlation coefficient is defined as the boundary threshold in Equation (38), beyond which stop-loss occurs; Standard deviation means stop-loss when the volatility of the spread exceeds three times the standard deviation within the window period; Historical rolling volatility means stop-loss when the spread exceeds three times the average value of the historical spread in the window period; Historical percentage method (30%) means stop-loss when the spread exceeds 30% of its value at the time of the window period; Cointegration test means stop-loss by checking whether the cointegration is stable at that time, if not stable, stop-loss; ATR method means stop-loss when exceeding the ATR threshold in the window period.

Comparison reveals that the method with a negative correlation coefficient has the highest profitability when maintaining the position unchanged. Besides, both the standard deviation and historical rolling volatility methods show good profitability under both stop-loss measures. With the exception of the cointegration test, almost all stop-loss methods result in higher profits when maintaining the position unchanged. Therefore, we can draw a conclusion from behavioral finance regarding the impact of investor psychology on returns: in the trading process, a chasing and selling strategy is not advisable. For investors, maintaining a stable mindset and the courage to hold positions are crucial. When profits are slightly compromised, having the patience to wait and the courage to hold positions, especially in the presence of transaction costs, often leads to better results than liquidating positions.

Considering that the method with a negative correlation coefficient is related to consumer risk aversion, it has stronger interpretability and generality. In the subsequent analysis, we will adopt this stop-loss threshold.

Discussion on Transaction Costs

Due to the frequent adjustment of positions required in pair trading, transaction cost is a crucial factor that must be considered. Accordingly, we defined transaction costs as 2% of the trading amount per transaction.

Changes in position will incur transaction costs, and the consideration of the remaining trading time by investors during the current trade will also generate opportunity costs related to time. Theorem 4.2: Similar to the approach taken by, the value function is decomposed into the following three multiplicative terms (U1,V1 and V2) and it is observed how it depends on the remaining trading time.

$$U_1(w) = \frac{1}{\gamma} w^{\gamma} \tag{4.4}$$

$$V_1(\tau) = \sqrt{e^{\tau} C(\tau)^{\gamma - 1}} \tag{4.5}$$

$$V_2(\tau, x) = \exp(\frac{x^2}{2}(1 + (\gamma - 1) D(\tau)) \tag{4.6}$$

Where τ , C(τ) and D(τ) are defined by equations (31)-(33), and D(τ) .

Remark 4.2.1: Term U1 represents the utility value obtained from the current investor's held wealth, term V1 is the utility value of remaining time, and term V2 is the utility value of immediate investment demonstrated through numerical simulations that assuming no immediate opportunities (i.e., when X = 0), the investor's strategy with logarithmic utility was independent of time, and the value function J linearly increased with remaining time τ . Extending the trading period beyond a certain minimum length does not significantly increase the value function for risk-averse investors. When there are immediate investment opportunities, the value function generally exhibits exponential growth. Furthermore, for investors with any level of risk aversion, an increase in the spread will always lead to a decrease in their utility level.

Discussion on Trading Positions

Remark

According to Theorem 3.3, The optimal trading position in the dimensionless system is restored to the original system as:

$$\pi_t^* = -\frac{k}{\sigma^2} wx D(\frac{\tau}{k}) \tag{4.7}$$

Clearly, the position increases with the growth of investor's wealth. It can also be seen that an increase in the regression speed of the spread and a decrease in volatility will lead to an increase in the optimal trading position, indicating more aggressive trading behavior. Therefore, in uncertain market conditions, overestimating volatility and underestimating the regression speed of the spread is a safer approach. Demonstrated through simulation that as the time horizon approaches, traders with lower risk aversion became more aggressive. It was also noted that this might be related to the completion of performance indicators for traders towards the end of trading.

Empirical Results

In the first section, we presented an optimal stock selection strategy based on [7].

In the second section, we compared the return results of our stock selection strategy with two other common strategies, demonstrating the effectiveness and superiority of our approach. In the third section, we searched for the optimal stock pairs within each industry stock pool and gave the returns of each industry.

In the fourth section, we lifted the restriction on the industry for stock selection and searched for the optimal stock pairs across all industries.

- In the fifth section, we constrained the industry to be the same and compared the returns of our stock selection strategy during the bull and bear markets of 2000 and 2020.
- In the sixth section, we presented the model's performance in extreme scenarios, providing excess returns.
- In the seventh section, we considered trading in 3 pairs and 5 pairs simultaneously, reducing the opportunity cost of waiting for a single trading pair.

- In the eighth section, we discussed the returns for different risk aversion profiles of investors targeting the same investment portfolio and derived a balance theory between the opportunity cost of waiting time and the aggressiveness of the trading process.

We assumed an initial wealth W0=1 , transaction costs 2% of the transaction amount, and neglected impact costs and market frictions due to the small trading volume. For each stock pair, we provided price trends, standardized spread and stop-loss thresholds, position change charts, and wealth change charts. Evaluation metrics for return performance included annualized return rate, relative return rate compared to the corresponding industry index, Sharpe ratio, industry-specific Sharpe ratio, the percentage of days with returns less than or equal to 0 as a proportion of the total number of days among others. The research of indicates the profitability of fixed pair trading is related to the time trend, showing a continuous declining trend over time. Therefore, our trading cycle for the same stock pair should not exceed two years.

Selection of Pairs

According to the findings of, We conducted simulations with different stock selection and window periods. Following the principle of profit maximization, we determined the optimal lookback period for stock selection: 3-year stock observation period with 200-day trading window. Many studies have indicated that a low-volatility stock pool can affect the returns of pair trading. However, higher volatility poses a greater challenge for our stock selection. Considering that volatility should be at a moderate level in the market, we need to strike a balance in our stock- picking strategy, So the components of the Standard & Poor (S&P) industry index as of October 20, 2023 make up the stock pool used for our selection.

Specifically, we chose the constituents of the S&P daily consumer, energy, materials, financial, and industrial indices as our stock pool and excluded stocks with missing closing prices during the retrospective period (financial and industrial sectors). Our model calculates the NZC and SSD of the stocks in the chosen industry pool for the three years prior to the trading period (without any trading cycle overlap) before initiating any trading activity. Since our trading units are integers and the pairs are linear, we only need to consider the linear correlation coefficient, i.e., the Pearson correlation coefficient, between stocks.

$$\rho(price_A, price_B) = \frac{COV(price_A, price_B)}{\sqrt{Var(price_A)Var(price_B)}}$$

The selection criteria were: choosing the top 15 based on correlation, then selecting the top 8 from SSD, and finally picking the top 5 stock from NZC. We assumed our trading dates were from January 1, 2000, to December 31, 2002 (spanning both bear market (January 14, 2000, to October 9, 2002) and bull market (October 10, 2002, to December 31, 2002) periods), with a stock retrospective period from January 1, 1997, to December 31, 1999, i.e., a three-year stock selection period. Assuming our trading period is from January 1, 2000, to December 31, 2002, with a stock lookback period from January 1, 1997, to December 31, 1999, totaling a three-year stock selection period.Below we give an example of the stock picking process in the Consumer Discretionary industry.

I. Begin by plotting the correlation matrix of the industry stock pool, resulting in the correlation heat map shown as below:

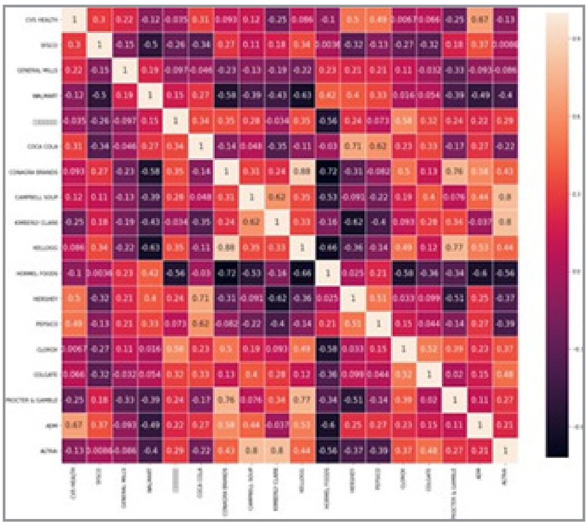

Figure 5: Correlation Heat Map

II. Identify the top fifteen stocks with the highest correlation:

Table 2: The Correlation Coefficient of Pairs

	Stock A	Stock B	correlation coefficient
1	KELLOGG	CONAGRA BRANDS	0.87663
2	ALTRIA	KIMBERLY CLARK	0.802376
3	WBA	CAMPBELL SOUP	0.79551
4	PROCTER&GAMBLE	KELLOGG	0.771291
5	WBA	CONAGRA BRANDS	0.764481

6	HERSHEY	COCA COLA	0.705664
7	ADM	CVS HEALTH	0.670761
8	PEPSICO	COCA COLA	0.620463
9	KIMBERLY CLARK	CAMPBELL SOUP	0.618508
10	ADM	CONAGRA BRANDS	0.582673
11	CLOROX	WBA	0.577459
12	ADM	KELLOGG	0.530131
13	COLGATE	CLOROX	0.524405
14	PEPSICO	HERSHEY	0.507376
15	HERSHEY	CVS HEALTH	0.502026

Identify The Top Eight Stocks with the Minimum SSD

Table 3: The SSD Of Pairs

	Stock A	Stock B	SSD
1	HERSHEY	COCA COLA	1177512
2	PROCTER & GAMBLE	CONAGRA BRANDS	1217281
3	CLOROX	WBA	1283241
4	ALTRIA	KIMBERLY CLARK	1388386
5	PEPSICO	COCA COLA	1397406
6	HERSHEY	CVS HEALTH	1441898
7	PROCTER & GAMBLE	KELLOGG	1532596
8	ADM	CONAGRA BRANDS	1533604

Identify the top five stocks with the maximum NZC

Table 4: The NZC Of Pairs

	Stock A	Stock B	NZC
1	HERSHEY	COCA COLA	487
2	PROCTER & GAMBLE	KELLOGG	485
3	PEPSICO	COCA COLA	412
4	CLOROX	WBA	412
5	HERSHEY	CVS HEALTH	345

In our Empirical study, we selected the pairs with the maximum NZC: 'HERSHEY' and 'COCA COLA'.

Effectiveness of Stock Selection Strategy

Taking the daily consumer industry as an example, we compared the returns of the stock port- folios obtained through our stock selection strategy with those obtained only based on correlation and the distance method (based solely on the magnitude of the price trends of two stocks).

It can be observed that the optimal stock portfolio returns obtained through our strategy for 'ALTRIA' and 'KIMBERLY CLARK' are better than the stock portfolio selected based on high correlation and small distance: 'PEPSICO' and 'COCA COLA.' This indicates the effectiveness and robustness of our stock selection strategy.

1. Hershey, Coca Cola

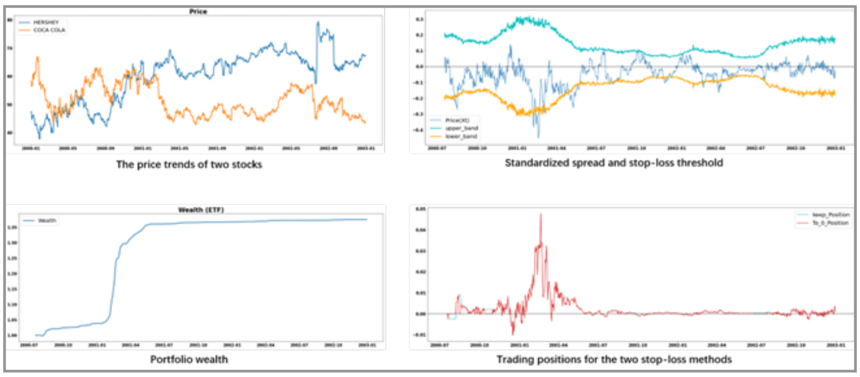

Figure 6: Hershey and Coca Cola (2) Pepsico, Coca Cola

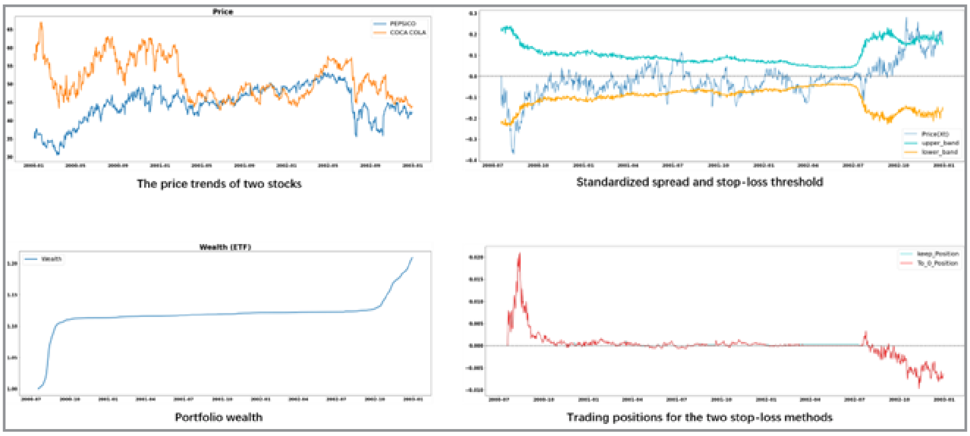

Figure 7: Pepsico and Coca Cola

The comparative results of the returns for the three aforementioned strategies are presented in the following table.

Table 5: The Comparative Results of Three Strategies

Strategy	Annualized Return	Annualized Return of the Industry S&P Index	Relative Return	Percentage of ¡U¨ 0 Days	Annualized Sharpe Ratio	Sharpe Ratio of the Industry Index
Our strategy	18.77%	-0.56%	19.33%	13.38%	0.08382	-0.1606
Strategy1	10.49%	-0.56%	11.05%	12.04%	0.0329	-0.1606
Strategy2	10.49%	-0.56%	11.05%	12.04%	0.0329	-0.1606

Stock pairs within the Same Industry

According to the stock selection strategy, we obtained the optimal stock pairs for each industry as follows:

Table 6: The Optimal Stock Pairs for Each Industry

S&P Industry	Optimal Stock Pair
Consumer Goods Industry Energy Industry Materials Industry Financial Industry Industrial Industry	HERSHEY,COCA COLA DVN.N¡, CTRA.N NUE.N, IP.N KEY.N, CB.N SNA.N, NOC.N

We can see that in the later period of the energy industry, the correlation between the two stocks decreased, and the volatility of the spread increased. Our trading threshold also expanded accordingly, making it less likely to trigger stop-loss and ensuring a stable increase in wealth.

Considering that the denominator of the Sharpe ratio represents the standard deviation of the portfolio's annualized return, it focuses on the volatility of the entire asset, including both upward and downward movements, i.e., the standard deviation. However, from the principles of pair trading, we know that the fluctuation of the spread within a certain range is favorable for us. Therefore, we introduced a new performance measure, the monthly Sortino ratio. The Sortino ratio only considers the downside volatility of the asset, i.e., the drawdown standard deviation. In simple terms, both the Sharpe ratio and the Sortino ratio have the excess return as the numerator, but the Sortino ratio distinguishes between good and bad volatility. Only when the monthly return of the pair trading portfolio is less than the risk-free rate for that period, the volatility of that period, i.e., the standard deviation of the pair trading portfolio's return, will be recorded. Therefore, we summarized the performance of the individual optimal pair trading portfolios selected from the aforementioned same-industry pairs as follows:

Table 7: The Results of Individual Optimal Pair from the Same Industry

S&P Industry Annualized Return Annualized Return of the Industry S&P			Index Relative Return	Percentage of ¡U¨ 0 Days	Percentage of ¡U¨ 0 Days	harpe Ratio of the Industry Index Monthly Sortino Ratio	
Consumer Goods Industry	18.77%	-0.56%	19.33%	13.38%	0.08382	-0.1606	8.5968
Energy Industry	21.63%	-4.86%	26.49%	11.73%	0.09797	-0.4059	23.8193
Materials Industry	17.03%	-1.08%	18.11%	15.87%	0.07501	-1.0801	9.1583

Industrial Industry	9.21%	-0.74%	9.95%	23.02%	0.03344	-0.7374	4.1939

1. Energy Industry(‘DVN.N’, ’CTRA.N’)

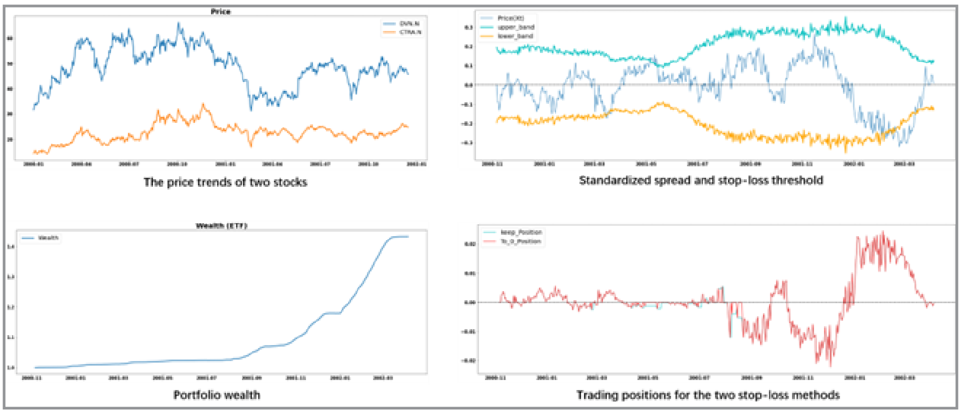

Figure 8: Energy Industry (2)Materials Industry(’NUE.N’, ’IP.N’)

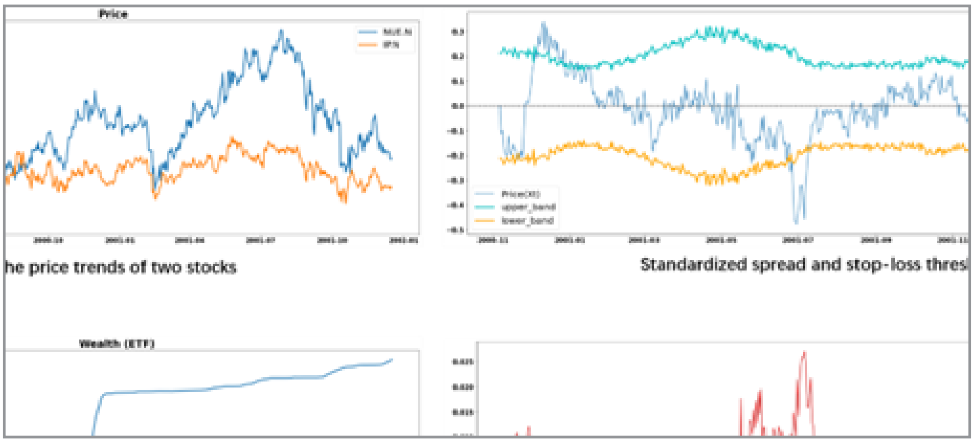

Figure 9: Materials Industry (3)Financial Industry(’KEY.N’, ’CB.N’)

Figure 10: Financial Industry (4)Industrial Industry(’SNA.N’, ’NOC.N’)

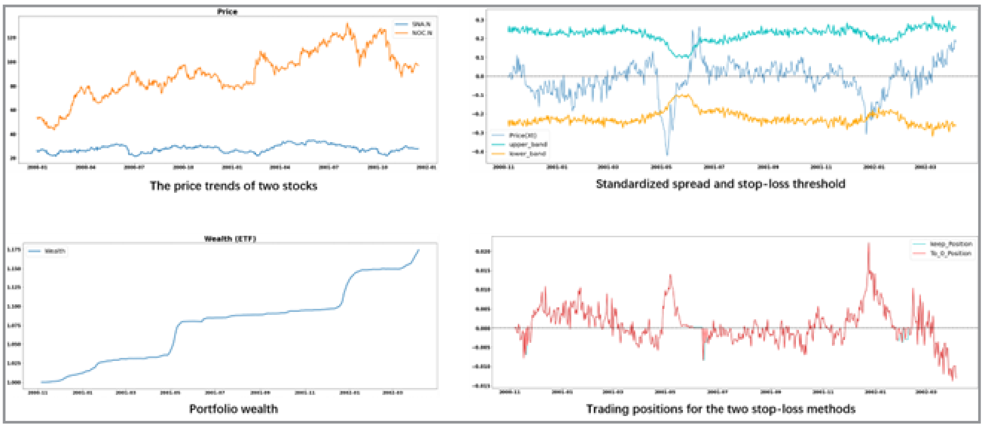

Figure 11: Industrial Industry

Stock Pairs Within Different Industry

Found that the trading results from the unrestricted industry criterion showed a very high percentage of loss making trades resulting from a high arbitrage or divergence risk and tested that industry homogeneity had a significant impact in reducing the divergence risk of the strategy on the basis and did not have a particularly significant impact on market efficiency.

To examine whether our model is subject to the limitations of industry homogeneity, in this section, we no longer restrict stock pairs to come from the same S&P industry but instead employ a cross-industry stock selection approach. We selected stocks based on the correlation, SSD, and NZC criteria for the entire industry during the three-year backtesting period from January 1, 1997, to December 31, 1999. The top three outstanding stock pairs selected are ('PROCTER & GAMBLE', 'AJG.N'), ('CLOROX', 'AJG.N'), and ('KEY.N', 'CB.N'). Notably, the first two pairs are from entirely different industries, and their pair trading situations are as follows:

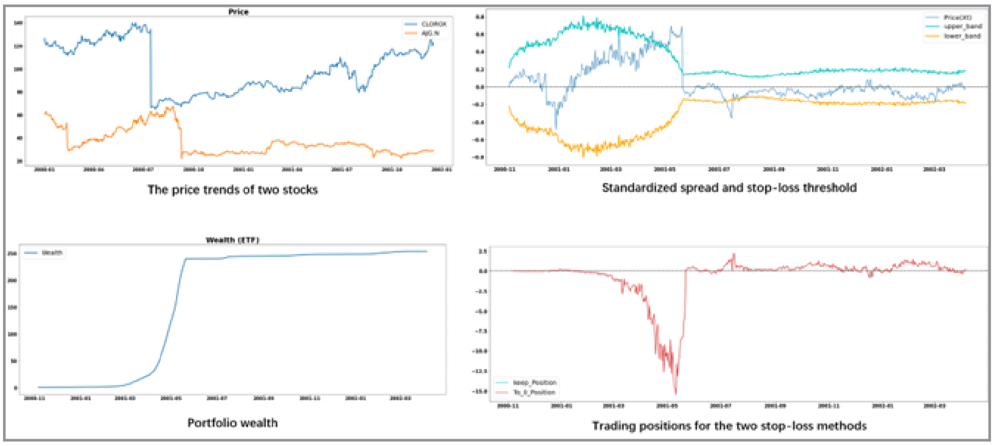

Figure 12: PROCTER & GAMBLE and AJG.N

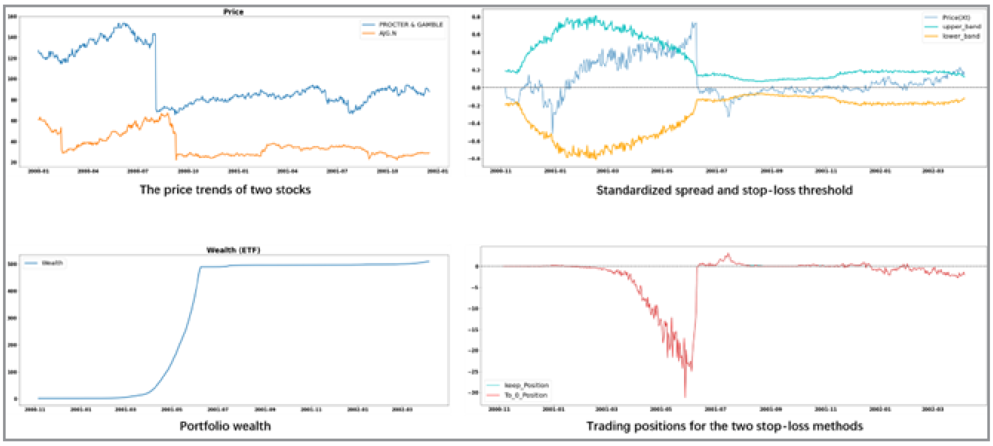

Figure 13: CLOROX and AJG.N

It can be observed that the two pairs selected across industries have achieved more excess returns compared to stock selections within the same industry. This indicates that our stock selection strategy is not only effective for stocks within the same industry but also applicable to the entire stock pool. It also provides insight to investors: when engaging in pair trading, it is advisable not to restrict stock pairs to the same industry; instead, one should conduct screening across all industries, as cross-industry stock selection can yield surprisingly positive results. Found the reason behind the weak performance of the different industry strategy was a very high percentage of stop losses triggered implying a higher divergence risk. However, by observing the trading process of the two pairs of trades above (top right corner of the chart), we found that the frequency of triggering thresholds was not very high, even lower than some pairs of trades in the same industry. This indicates that the real-time dynamic stop- loss thresholds of our strategy are effective and reasonable. Compared to traditional pair trading strategies, it can avoid the problem of frequent triggering of stop-loss thresholds mentioned by, which leads to a decline in returns, and maintain a robust upward trend in returns.

According to, there are two factors: arbitrage risk effect and market efficiency which significantly affect pair trading returns, our strategy eliminates significant arbitrage risks among pairs in different industries and enhances market efficiency at a higher level.

Results of Strategies in Different year

In this section, we select the industry as the daily consumer industry and compare our strategy in two time periods: January 1, 2000 (experiencing a bear market from January 14, 2000, to October 9, 2002, and a bull market from October 10, 2002, to December 31, 2002) and January 1, 2020, to December 31, 2022 (experiencing a bear market from January 21, 2020, to March 23, 2020, and a bull market from March 24, 2020, to December 31, 2022). We use the stock selection strategy based on the backtesting period from January 1, 2017, to December 31, 2019, and the optimal investment pairs obtained is 'WALMART' and 'COCA COLA.' Below is the trading situation of this pairs starting from January 1, 2020:

Figure 14: Pair Starting from January 1, 2020

We can see that this pairs went through the bull and bear markets of 2020, and while the trading returns were not as high as in 2000, they still maintained stable positive returns. Observing the price trends of 'WALMART' and 'COCA COLA' stocks in the two time periods of 2000 and 2020, we can find that in 2000 , the price volatility of both stocks was significantly higher than in 2020. We can conclude that in more volatile bull and bear markets, a higher degree of oscillation in the two stocks of the pair will lead to higher returns. This also confirms the conclusion of.

Excess Returns in Extreme Scenarios

We explored a unique stock pair in the energy industry: OKE.N and MRO.N. During the stock selection period, these two stocks exhibited a high correlation and ranked third in the excellent stock selection. Below are the trading details of this pair:

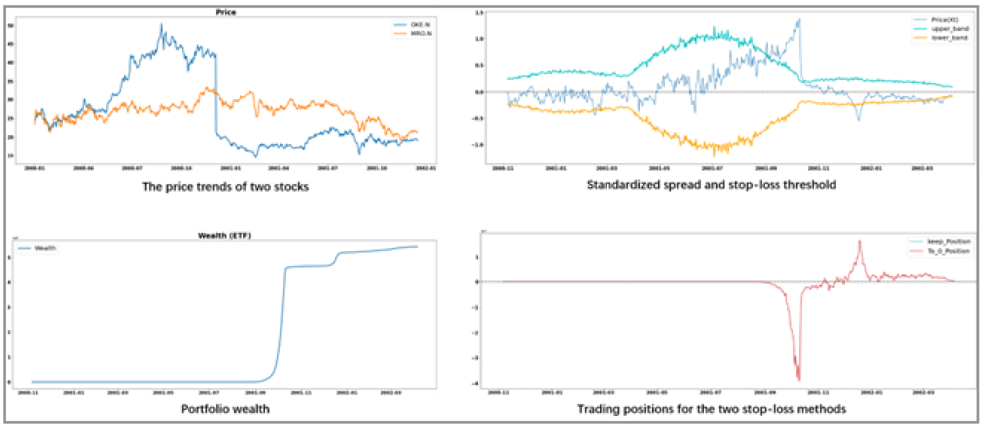

Figure 15: OKE.N and MRO.N

We can see that the sharp bear market decline of OKE.N at the end of 2000 caused the spread between the two to deviate from the average level. Due to the drop in OKE.N's price, the prices of the two stocks were at similar levels, resulting in a low cost to construct a long-short portfolio. We were able to increase positions to a large extent as suggested by the strategy, ultimately achieving unexpected excess returns! Subsequently, when OKE.N fell to a level below its original price and remained stable, our portfolio quickly adjusted positions, continued to increase positions in the opposite direction, and sustained profits. As shown in the upper right graph, when stock prices experienced significant changes, our strategy appropriately expanded the stop-loss threshold to ensure that trades were not closed easily. Our strategy also effectively identified moments when the prices of the two stocks reversed, allowing us to build positions in the opposite direction and continue to gain profits. This provides insight for investors: during the stock selection phase, choosing stocks with higher correlation, even when stock prices unexpectedly change during the trading period, maintaining confidence in one's strategy can lead to excess returns.

Results of Multiple Pairs

According to Table 1 , it can be seen that during the trading period, choosing to trade a single stock pair may result in days with zero returns due to the set stop-loss threshold. On average, the percentage of days with zero returns is 15%. To avoid the opportunity cost loss caused by this waiting time, one method is to choose stocks with higher correlation to avoid excessive occurrences beyond the threshold. However, this might lead to significant deviations from expected outcomes. Another method is to trade multiple pairs during the trading period and allocate the initial wealth based on the excellence of the selected pairs. For example, let's observe the returns of multiple trading pairs in the daily consumer and energy industries. We will examine the cases of 3 pairs and 5 pairs. We still assume the initial wealth is $1. If we have 5 pairs, we allocate weights based on the selected NZC from largest to smallest, giving weights of 0.3, 0.2, 0.2, 0.2, and 0.1 to the five pairs, respectively. This means the initial amounts for trading the five pairs are $0.3, $0.2, $0.2, $0.2, and $0.1. For 3 pairs, we assign weights of 0.5, 0.3 and 0.2. Here are the stocks selected by
our model:

Daily Consumer Industry Stock Selection (Top five pairs):

1. Hershey,Coca Cola
2. Procter & Gamble, Kellogg
3. Coca Cola, Pepsico
4. Clorox,WBA
5. Hershey,CVS Health

Energy Industry Stock Selection (Top five pairs):

1. Eog.N,DVN.N
2. DVN.N ,Ctra.N
3. HES.N ,Apa.O
4. HES .N,DVN .N
5. MRO.N ,CTRA.N

The returns from trading according to the initial weights are as follows (January 1, 2000, to December 31, 2002):

Table 8: The Results of Trading with Initial Weights

Industry	Annualize-dReturn	Annualize-dReturnof the Indus-tryS&PIn-dex	Relative Return	Percenta-geof≤0Days	Annualized SharpeR-atio	Sharpe Ratio of the Indus-tryIndex	Monthly SortinoR-atio
Consumer Goods Industry (5 pairs)	105.48%	-0.56%	106.04%	2.00%	0.3364	-0.1606	77.2439
Consumer Goods Industry (3 pairs)	105.04%	-0.56%	105.60%	3.01%	0.3209	-0.1606	46.2561
Consumer Goods Industry (1 pairs)	18.77%	-0.56%	19.33%	13.38%	0.08382	-0.1606	8.5968
Energy Industry (5 pairs)	43.84%	-4.86%	48.70%	0.58%	0.1099	-0.4059	3[2]*inf
Energy Industry (3 pairs)	19.11%	-4.86%	23.97%	2.32%	0.0185	-0.4059	32.06334
Energy Industry (1 pairs)	21.63%	-4.86%	26.49%	11.73%	0.09797	-0.4059	23.8193

As seen in the daily consumer industry, compared to using a single optimal investment pair, multiple pairs achieved a higher annualized return, a higher monthly Sortino ratio, and a significant reduction in the percentage of days with zero returns. This substantial reduction in zero-return days not only greatly lowers our opportunity cost but also brings about new trading opportunities. Additionally, the results of the 5 pairs outperformed those from the 3 pairs, highlighting the investment philosophy of ”not putting all your eggs in one basket.” Broadening our stock pool often leads to greater returns. In the energy industry, compared to a single optimal investment pair, the annualized return of the 3 pairs is not as high, but the percentage of days with zero returns significantly decreases, and the monthly Sortino ratio also increases. The annualized return of the 5 pairs significantly improves, indicating that the down-market risk of multiple pairs brings higher returns while also leading to an increase in the overall return, which remains far above the industry index's return.

Results for Different Levels of Risk Aversion

From the analysis of the results in the fifth section, we can see that the investor's risk aversion coefficient will affect our trading positions and stop-loss thresholds, thereby influencing the ultimate returns of our trades. Therefore, we adjusted the investor's risk aversion level for the daily consumption, energy, and materials industries respectively, and observed the final results of

Table 9: The Results of Different Industry After Adjusting γ

ConsumerGood-sIndustry	γ	AnnualizedReturn	Annualized SharpeRatio	Monthly Sor-tinoRatio	Percentage of≤-0Days
	0.9	310.00%	6.1846	425.9797	11.26%
	0.5	71.56%	0.49776	36.7544	11.37%
	-0.5	32.90%	0.16022	25.284	12.04%
	-1	27.74%	0.12771	13.6036	12.93%
	-1.5	24.17%	0.10599	11.2559	12.60%
EnergyIndustry	γ	AnnualizedReturn	Annualized Sharp-eRatio	Monthly SortinoR-atio	Percentageof≤-0Days
	0.9	412.96%	8.98214	724.2562	9.48%
	0.5	85.20%	0.60635	61.6076	11.03%
	-0.5	38.59%	0.20149	27.3916	10.03%
	-1	28.35%	0.13547	24.1415	10.44%
	-1.5	24.55%	0.11287	22.5645	9.09%
MaterialsIndustry	γ	AnnualizedReturn	Annualized Sharp-eRatio	Monthly SortinoR-atio	Percentageof≤-0Days
	0.9	198.24%	3.3253	366.8659	15.86%
	0.5	55.34%	0.41612	59.0984	14.31%
	-0.5	26.56%	0.14592	17.9804	14.51%
	-1	21.40%	0.10844	11.8281	15.86%
	-1.5	18.52%	0.08938	10.1732	14.89%
	-2	17.03%	0.07501	9.1583	15.87%

We can see that investors with higher risk aversion levels 1 γ typically choose low-risk in- vestment portfolios, resulting in lower annualized returns from the investment portfolios. The monthly Sortino ratio is generally used to measure the returns generated by unit downside risk, with higher Sortino ratios considered favorable as they indicate higher returns per unit of risk. It can be seen that as the investor's risk aversion level increases, the monthly Sortino ratio decreases. This suggests that a higher degree of risk aversion makes investors more inclined to choose low- risk investment portfolios, leading to missing out on higher-yield investment opportunities and adecline in overall annualized returns. Particularly in extreme cases $\gamma = 0.9$, our model achieved excess annualized returns, Sharpe ratios, and Sortino ratios. Further observation of the percentage of days with returns less than or equal to 0 reveals that, in the daily consumption and energy in- dustries, except for the reversal of the monotonic relationship between risk aversion coefficients -1 and -1.5, the percentage of days with returns equal to 0 increases with an increase in risk aversion level. This indicates that a risk-averse mindset leads to an opportunity cost of waiting. In the materials industry, under the extreme case of $\gamma = 0.9$, the percentage of days with returns less than or equal to 0 is slightly higher than the case of γ being lower. Further analysis shows that the percentage of days with negative returns is 15.85%, and the percentage of days with returns equal to 0 is 0%. This suggests that at an extremely low level of risk aversion, there is almost no waiting time, and the opportunity cost of waiting is 0. However, due to a more aggressive trading strategy, 15.86% of the days have negative returns. Although the final annualized return is still relatively high, negative returns during the trading process may affect investor psychology, leading to the possibility of early termination of trades. Therefore, different risk aversion coefficients for investors reflect a balance between the opportunity cost of waiting and the robustness of returns during the trading process.

Conclusion and Recommendations

Highlights

A robust pairs trading model must possess four fundamental attributes: reliable criteria for pair selection, an effective stochastic model for simulating currency pair movements and robust parameter estimation techniques, a low-risk and high-return trading strategy for a given currency pair, and a comprehensive back-testing method (with non-overlapping test and training sets) to determine the optimal pairs and window lengths for trading. Therefore, we utilized various criteria such as industry, SSD, and NZC for stock selection in the stock pool, and demonstrated the effec- tiveness of our stock selection strategy in empirical studies. For cointegrated and non-cointegrated spread processes, we employed two models to predict the regression model of the spread. By back-testing the residuals of real and predicted values, we found high accuracy in our prediction process. Our strategy aims to maximize investor utility. We solved the optimal position through the HJB equation, determined the stop-loss threshold based on the correlation coefficient between the optimal position and the spread's sign change, and proposed two stop-loss measures. Our model uses a 200-day window for data fitting, strictly adhering to the principle of non-overlapping test and training sets [8].

According to empirical results, in terms of absolute returns, our strategy yielded positive returns in both bull and bear markets. In terms of relative returns, our model achieved excess returns compared to both S&P industry indices and other pair selection strategies. In the empirical study, we introduced metrics such as

Sharpe ratio, percentage of days with zero returns, and monthly Sortino ratio to comprehensively evaluate the performance of the strategy. The results showed that our model achieved favorable returns. Additionally, based on the threshold setting principle of the model, once the wealth reaches zero, trading stops, and short-selling or leveraged trading ceases. This limits the maximum loss. For example, we illustrated the extreme case of the energy sector's price trend to demonstrate that our model can quickly respond when the spread reverses, adjust the position size, and reverse the position to continue gaining profits, highlighting the flexibility of our strategy. It is worth mentioning that the time periods of our stock selection model and strategy model do not overlap. The estimation of strategy model parameters is based on a 200-day sliding window before the trading day, which does not overlap with the trading period. The test results confirm the robustness and consistency of the model [9].

Future Research

Every model has its limitations, and investors need to make appropriate adjustments and optimizations based on their own situations in practical applications. In our model, market frictions and the impact cost of bulk trading on prices have not been considered. When the trading order volume is large or different investors compete for the same pairs trading opportunity, market microstructure feedback may affect the optimal strategy. To address this, we can refer to the method proposed by and suggest incorporating linear impact costs of trading volume on prices into the model, where the coefficient can be fitted through historical data and is usually expressed as a percentage.

In real trading, stock prices may also undergo structural changes and jumps, so more factor testing and pattern recognition of stock price processes before trading are necessary. In addition, investing all capital in a single pairs trading portfolio carries a certain risk, so we can also consider choosing portfolios involving multiple risky assets ETFs.

References

1. Barzykin, A. V., & Lillo, F. (2019). Optimal VWAP execution under transient price impact. ERN: Optimization Techniques; Programming Models; Dynamic Analysis. https://api.semanticscholar.org/CorpusID:159326986
2. Bertram, W. K. (2010). Analytic solutions for optimal statistical arbitrage trading. Physica A: Statistical Mechanics and Its Applications, 389(11), 2234–2243. https://doi.org/10.1016/j.physa.2010.01.045
3. Boguslavsky, M., & Boguslavskaya, E. (2004). Arbitrage under power. Unpublished manuscript. https://api.semanticscholar.org/CorpusID:59935978
4. Do, B. H., & Faff, R. W. (2010). Do simple pairs trading still work? Financial Analysts Journal, 66(4), 83–95. https://api.semanticscholar.org/CorpusID:154852025
5. Gautam, A., & Goyal, R. (2023). Analysis of performance of the pairs trading strategy using public sector enterprises stocks. Recent Advances in Sciences, Engineering, Information Technology & Management. https://api.semanticscholar.org/CorpusID:259476874
6. Holý, V., & Tomanová, P. (2018). Estimation of Ornstein–Uhlenbeck process using ultra-high-frequency data with application to intraday pairs trading strategy. arXiv: Statistical Finance. https://api.semanticscholar.org/CorpusID:158274096
7. Huck, N. (2015). Pairs trading: Does volatility timing matter? Applied Economics, 47(57), 6239–6256. https://doi.org/10.1080/00036846.2015.1068923
8. Liu, J., & Timmermann, A. (2013). Optimal convergence trade strategies. Review of Financial Studies, 26(4), 1048–1086. https://api.semanticscholar.org/CorpusID:5320852
9. Tang, C. Y., & Chen, S. X. (2009). Parameter estimation and bias correction for diffusion processes. Journal of Econometrics, 149(1), 65–81. https://api.semanticscholar.org/CorpusID:4896632