Trade-level regression analysis of predictive signals in pairs trading

Zhouyang Ren[1], Yuelin Li[2], Sibo Yan[3]*

[1]Gonzaga University, Washington, USA

[2]University of International Business and Economics, Beijing, China

[3]Johns Hopkins University, Washington, USA

*Corresponding Author. Email: zren@zagmail.gonzaga.edu

Abstract. This paper examines the viability of high-frequency pairs trading in China's A-share market using minute-level data across three GICS sectors. Applying a Bollinger-band framework to 4,500 cointegrated stock pairs, we document consistent—though moderate—gross profitability. We then conduct trade-level regressions to isolate drivers of profitability and identify three key determinants: spread volatility, market-cap disparity within pairs, and multi-frequency return correlations. Further diagnostics show that execution latency significantly erodes returns, convergence reliability closely aligns with estimated half-lives, and exit effectiveness varies asymmetrically across long and short trades. Sector-level results reveal that Consumer Discretionary delivers the strongest risk-adjusted performance, while Retailing features the fastest mean-reversion and highest trade frequency. Robustness checks—including random-pair bootstraps, threshold sensitivity tests, and stress-period performance during the COVID-19 shock—confirm that profitability is not driven by data-snooping or microstructure artifacts. Overall, the findings provide new evidence that high-frequency statistical arbitrage remains feasible in an emerging market setting, while highlighting the critical roles of execution speed, volatility conditions, and behavioral inattention in shaping trade-level outcomes.

Keywords: high-frequency pairs trading, China A-share market, Bollinger-band strategy, cointegration analysis, statistical arbitrage

1. Introduction

Pairs trading, a classical relative-value arbitrage strategy exploiting temporary mispricing between economically linked securities [1], has been widely studied in developed equity markets. Yet its performance in emerging markets—particularly in China's A-share market—remains relatively underexplored. Prior studies have highlighted the complexity inherent in arbitrage execution within China's unique regulatory environment [2, 3].

This study examines whether high-frequency pairs trading can generate consistent profitability in such an environment and, critically, what trade-level determinants shape those outcomes. Using minute-level data from 2019–2020 across threeGlobal Industry Classification Standard (GICS) sectors (Technology Hardware & Equipment, Consumer Discretionary, Retailing), we construct 4,500 pairs via a two-stage distance–

cointegration framework and evaluate performance under rolling and periodic calibration schemes. While the strategy delivers moderate but persistent gross returns, its profitability varies substantially across trades, sectors, and volatility regimes. Prior studies have highlighted the specific complexities of arbitrage execution in China, including the challenges posed by price limits and information asymmetry [2, 3], underscoring the need for a granular analysis.

A central contribution of this paper is to move beyond pair-level averages and instead identify trade-level predictors of profitability. We show that spread volatility, relative market-capitalization disparity, and multi-frequency return correlations significantly influence trade outcomes. Further, detailed diagnostics reveal strong behavioral and microstructural drivers: execution latency erodes profits disproportionately, convergence reliability aligns closely with estimated half-lives, and exit effectiveness displays notable asymmetry between long and short trades. These findings highlight the importance of execution timing and investor inattention as mechanisms sustaining short-horizon mispricing.

The study makes three key contributions. First, it provides the most granular high-frequency evidence to date on pairs trading in China, demonstrating that statistical arbitrage remains feasible even under the market's unique frictions. Second, it identifies and quantifies the trade-level determinants of profitability, offering new insight into the behavioral and microstructure forces shaping convergence. Third, it delivers sector-specific guidance on threshold calibration, expected convergence horizons, and risk controls, relevant for practitioners designing China-focused high-frequency strategies.

Overall, this paper extends the statistical arbitrage literature to a major emerging market and shows that, despite modest Sharpe ratios relative to U.S. benchmarks, high-frequency pairs trading in China benefits from systematic, repeatable pricing anomalies---particularly when execution speed is optimized and sector characteristics are carefully incorporated.

2. Background

2.1. Pairs trading mechanism in China

Pairs trading in the Chinese market presents both distinctive opportunities and considerable challenges, primarily arising from China's unique market structure and regulatory environment. Unlike mature equity markets, China's A-share market exhibits notable features such as high retail investor participation, significant governmental influence, frequent trading halts, and stringent short-selling restrictions. These market-specific characteristics contribute to short-term inefficiencies and transient mispricings, creating potential profitability for statistical arbitrage strategies like pairs trading.

One prominent characteristic is the substantial involvement of retail investors, whose trading behaviors typically manifest as herd mentality and overreaction to market news. Such behavior frequently leads to temporary mispricings between correlated assets, providing opportunities for statistical arbitrageurs to exploit transient divergences. Additionally, governmental policies exert direct and substantial impacts on sector-level and firm-specific stock performances. Policy-driven events often induce synchronized price movements among closely related stocks, enhancing the potential effectiveness of sector-specific pairs trading strategies.

Moreover, regulatory constraints, including stringent short-selling rules and price-limit regulations, pose specific obstacles. Short-selling restrictions and price-limit mechanisms, which were originally designed to stabilize markets and prevent excessive volatility, significantly complicate the execution of market-neutral arbitrage strategies [2]. Consequently, traders must adopt innovative approaches to mitigate the inherent execution risks associated with these regulatory limitations.

Despite these inherent challenges, ongoing liberalization within China's financial markets, advancements in trading infrastructure, and increasing penetration of algorithmic trading practices have created favorable conditions for the implementation of advanced arbitrage strategies. These evolving market conditions have enhanced both thefeasibility and attractiveness of sophisticated trading methods, such as high-frequency pairs trading, within the Chinese context.

Prior studies have highlighted the complexity inherent in arbitrage execution within China's equity market, particularly emphasizing the difficulties posed by frequent price limits, trading suspensions, and information asymmetries, which demand careful management of execution risks [2, 3]. The current study directly addresses these documented complexities through a systematic evaluation of high-frequency pairs trading strategies utilizing minute-level market data. By doing so, we provide critical empirical insights into effectively navigating and capitalizing upon the distinct market dynamics present in China's A-share market.

In the following sections, we build upon this contextual understanding by detailing the methodological framework for selecting robust pairs, constructing trading signals, and rigorously evaluating strategy performance. This comprehensive approach ensures that empirical analysis captures both the opportunities and unique execution challenges inherent in conducting high-frequency statistical arbitrage in China's evolving financial markets.

3. Data and pair selection

3.1. Data description

To construct a robust pairs trading universe, we source minute-by-minute last-traded price data for three industry sectors in China's A-share market—Technology Hardware & Equipment (GICS2 = 4520), Consumer Discretionary (GICS2 = 2010), and Retailing (GICS2 = 4510)—spanning June 30, 2019 to June 30, 2020. This period captures roughly 252 trading days and around 60,000 one-minute intervals per stock, providing sufficient granularity to identify short-lived deviations, a data structure aligned with high-frequency finance research [4]. By aggregating raw tick records into uniform one-minute bars, we mitigate microstructure noise and ensure comparability across securities

To maintain sample integrity, we apply stringent liquidity and data-quality filters. We exclude any stock that records more than 10% zero-return minutes, suffers frequent trading halts, or exhibits anomalously wide bid–ask spreads relative to its average liquidity. These criteria remove illiquid or suspended securities, which could otherwise introduce stale prices or spurious cointegration signals. Post-filtering, the resulting samples comprise 361 stocks in GICS 4520, 348 in GICS 2010, and 395 in GICS 4510. The relative sizes reflect sectoral market breadth and trading activity: the Retailing sector is most expansive, while Consumer Discretionary has a slightly smaller but still diverse constituent set.

Table 1. Formation period statistics for selected pairs in each sector

Sector (GICS2)	# Stocks	# Selected Pairs	Mean Corr	ADF Stat	Stationary (%)	Half-Life (days)	Spread Volatility (%)
4520	361	1,500	0.58	-3.87	92.7	5.3	2.4
2010	348	1,500	0.54	-3.65	89.1	6.2	2.8
4510	395	1,500	0.61	-4.10	94.3	4.9	2.2

Table 1 summarizes the formation-period characteristics of the 1,500 selected pairs in each sector. Across all three sectors, the selected pairs exhibit strong comovement, with average return correlations ranging from

0.54 to 0.61, indicating a high degree of economic linkage suitable for pairs trading. The Augmented Dickey–Fuller (ADF) statistics are substantially below conventional critical values, and more than 89% of pairs in each sector reject the unit-root null at the 5% level, confirming widespread spread stationarity during the formation period.

The estimated half-lives are relatively short, ranging from 4.9 to 6.2 trading days, suggesting that deviations from equilibrium tend to revert quickly—an important prerequisite for profitable mean-reversion strategies at higher frequencies. Notably, the Retailing sector exhibits the fastest mean reversion and the lowest spread volatility, while Consumer Discretionary pairs display slower convergence and higher volatility. These cross-sector differences motivate the use of sector-specific trading thresholds and calibration parameters in the subsequent strategy implementation.

Explanation:

• Mean Corr: Average Pearson correlation of daily log returns; high values (> 0.54) indicate strong comovement essential for mean reversion.

• ADF Stat: Average Augmented Dickey–Fuller test statistic; values well below the critical threshold ($\sim$–2.87) confirm stationarity of spreads.

• Stationary (%): Proportion of pairs rejecting a unit root at 5% significance, demonstrating robustness of cointegration.

• Half-Life (days): Median trading days for spread to revert halfway; Retailing (4.9 days) reverts faster than Consumer Discretionary (6.2 days).

• Spread Volatility (%): Standard deviation of the spread relative to price; informs appropriate z-score thresholds for entry and exit.

These detailed metrics validate the universe selection: each sector's pairs exhibit the combination of co-integration strength, mean-reversion speed, and volatility necessary to support profitable high-frequency pairs trading.

Subsequent sections will detail spread construction, signal generation, and performance evaluation using this rigorously defined sample.

Table 2. Distribution of daily return correlation across sectors

Sector (GICS2)	25th pct Correlation	Median Correlation	75th pct Correlation
4520	0.49	0.58	0.66
2010	0.45	0.54	0.62
4510	0.52	0.61	0.69

Table 2 reports the distribution of daily return correlations for the selected pairs across sectors, measured by the 25th, 50th (median), and 75th percentiles. The results confirm that the top 1,500 pairs selected in each sector exhibit consistently high levels of comovement, rather than being driven by a small subset of extreme observations. Median correlations range from 0.54 to 0.61 across sectors, indicating that the majority of selected pairs display sufficiently strong co-movement to support a mean-reversion trading framework.

Notable cross-sector differences also emerge. Retailing (GICS 4510) demonstrates the tightest correlation distribution, with both higher lower-quartile and upper-quartile values, suggesting more stable co-movement and narrower dispersion across pairs. In contrast, Consumer Discretionary (GICS 2010) exhibits a wider interquartile range, reflecting greater heterogeneity in pair relationships. These distributional patterns motivate the use of more conservative entry thresholds for sectors with weaker or more dispersed correlations, helping to reduce false divergence signals during the trading period.

Table 3. Distribution of half-life across sectors

Sector (GICS2)	25th pct Half-Life (days)	Median Half-Life (days)	75th pct Half-Life (days)
4520	3.2	5.3	9.1
2010	4.1	6.2	12.4
4510	2.5	4.9	8.0

Table 3 reports the distribution of estimated half-lives across sectors, measured by the 25th, 50th (median), and 75th percentiles. The results indicate substantial heterogeneity in convergence speed both within and across sectors. For the majority of selected pairs, half-lives fall within a relatively short window, ranging approximately from 2.5 to 12 trading days across sectors, confirming that spread deviations typically revert within a few weeks rather than over extended horizons.

Cross-sector differences are economically meaningful. Retailing (GICS 4510) exhibits the fastest mean reversion, with a median half-life below 5 trading days and a comparatively tight interquartile range, whereas Consumer Discretionary (GICS 2010) shows slower and more dispersed convergence dynamics, with upper-quartile half-lives extending beyond 12 trading days. These distributional properties motivate the adoption of sector-specific time-stop rules. In particular, a 20-day maximum holding period comfortably covers the vast majority of expected convergence events across all sectors, while shorter holding limits may be appropriate for sectors characterized by faster mean reversion.

Finally, we assess the robustness of the formation methodology by varying the calibration window length from five to eight weeks and re-selecting the top pairs. Sensitivity tests indicate that key performance metrics (average hit rate, Sharpe ratio) vary by less than 5% when the window extends to eight weeks, confirming the stability of the pair selection procedure. These robustness checks, detailed in Section 7, provide further confidence that the data description and selection framework yield a reliable universe for high-frequency pairs trading in China's emerging market environment.

3.2. Pairs selection and spread construction

Selecting high-quality pairs is fundamental to the success of any statistical arbitrage strategy. We implement a two-stage approach—distance-based pre-selection followed by cointegration-based validation—to identify robust, mean-reverting stock pairs across three sectors: Technology Hardware & Equipment (GICS2 = 4520), Consumer Discretionary (GICS2 = 2010), and Retailing (GICS2 = 4510).

Stage 1: Distance-Based Pre-Selection.

We calculate the distance metric for every possible pair (i, j) (i, j) using normalized one-minute log-price series over the five-week formation window (April 1–May 31, 2019):

$$D_{ij} = \sum_{t \in formation} [\widetilde{P}_i(t) - \widetilde{P}_j(t)]^2 \tag{1}$$

where

$$\widetilde{P}_i(t) = \frac{P_i(t)}{P_i(t_0)} \tag{2}$$

is each stock's log-price scaled to 1 at the start of the window. The distribution of D_{ij} is heavily right-skewed, with median values of approximately 0.05 for Technology, 0.06 for Consumer Discretionary, and 0.04 for Retailing. By ranking all $\binom{n}{2}$ pairs in ascending order of D_{ij}, we retain the top 1,500 per sector—those below the 5th percentile threshold—thus focusing on the tightest historical comovements while excluding noisy combinations.

Stage 2: Cointegration-Based Spread Construction and Testing.

For each candidate pair (i, j), we estimate the co-integrating coefficient $\beta_{i,j}$ by running an OLS regression:

$$P_i(t) = \beta_{i,j} P_j(t) + \varepsilon_{i,j}(t) \tag{3}$$

where $P_i(t)$ and $P_j(t)$ are minute-level log prices. The residual series,

$$S_{i,j}(t) = P_i(t) - \beta_{i,j} P_j(t) = \varepsilon_{i,j}(t) \tag{4}$$

is the candidate spread. We demean $S_{i,j}(t)$ to enforce zero mean and apply the Augmented Dickey–Fuller (ADF) test. Stationary rejection rates exceed 89% across sectors (94.3% in Retailing, 92.7% in Technology, 89.1% in Consumer Discretionary), confirming that the vast majority of spreads are mean-reverting.

1) Distribution of Key Construction Metrics

Table 4. Formation-period metrics for selected pairs

Sector	Median D_{ij}	25th pct D_{ij}	75th pct D_{ij}	Stationary (%)	Median $\tau_{1/2}$ (days)	Median κ
4520	0.05	0.02	0.12	92.7	5.3	0.131
2010	0.06	0.03	0.14	89.1	6.2	0.112
4510	0.04	0.01	0.10	94.3	4.9	0.141

Table 4 summarizes key formation-period metrics related to distance filtering and mean-reversion dynamics for the selected pairs across sectors. The distance metric D_{ij} exhibits relatively tight interquartile ranges in all sectors, indicating that the distance-based screening procedure yields a homogeneous set of economically related pairs. High stationarity rates—exceeding 89% in every sector—further confirm that the majority of selected spreads exhibit statistically significant mean-reverting behavior during the formation period.

The estimated mean-reversion speeds, captured by the parameter κ, vary systematically across sectors. Retailing (GICS 4510) displays the highest median κ, corresponding to the shortest half-life, while Consumer Discretionary (GICS 2010) exhibits slower and more dispersed reversion dynamics. These differences suggest that sectors characterized by faster and more consistent convergence are better suited for higher-frequency trading, whereas sectors with slower reversion benefit from more conservative holding-period and time-stop calibration.

2) Mean-Reversion Speed: Ornstein–Uhlenbeck Model

To quantify reversion dynamics, we fit an Ornstein–Uhlenbeck (OU) process to each spread:

$$dS_t = -\kappa S_t dt + \sigma dW_t, \; \kappa > 0 \tag{5}$$

whose half-life $\tau_{1/2} = ln(2)/\kappa$ measures the expected time to halve a deviation. Sector-specific median κ values—0.131, 0.112, and 0.141 per day for GICS2 4520, 2010, 4510—translate into median half-lives of 5.3, 6.2, and 4.9 days, respectively. The narrower IQR of $\tau_{1/2}$ in Retailing (2.5–8 days) compared to Consumer Discretionary (4.1–12.4 days) indicates more consistent reversion speeds, supporting higher-frequency trading in that sector.

3) Implications for Strategy Design

By pre-selecting pairs with tight distance metrics, high stationarity rates, and rapid mean reversion, we ensure a universe of 4,500 pairs well-suited for intraday statistical arbitrage. These construction metrics inform parameter calibration: sectors with faster reversion allow tighter z-score bands and shorter holding periods, while those with slower dynamics require wider bands to reduce false signals.

Table 5. Backtest performance summary for selected pairs in Consumer Discretionary and Retailing

Sector (GICS2)	Avg Trades per Pair	Avg Hit Rate (%)	Avg Total Return (%)	Avg Max Drawdown (%)	Avg Sharpe (Overall)	Avg Sharpe (Active)
2010	16.47	12.80	9.36	5.94	0.10	0.15
4510	19.07	14.78	8.61	8.61	0.09	0.13

Table 5 summarizes the backtest performance of the selected pairs in the Consumer Discretionary (GICS 2010) and Retailing (GICS 4510) sectors. The results reveal clear differences in trading dynamics and risk profiles across sectors. Retailing pairs trade more frequently on average and exhibit higher hit rates, consistent with faster and more reliable mean-reversion behavior observed during the formation period.

Despite this higher trading frequency, Consumer Discretionary pairs achieve slightly higher average total returns and substantially lower maximum drawdowns, resulting in superior risk-adjusted performance. Both the overall and active Sharpe ratios are higher for Consumer Discretionary, indicating that its profitability is achieved with more stable return profiles. These findings highlight an important trade-off between trading frequency and risk exposure: sectors characterized by rapid convergence generate more opportunities but may also experience greater drawdowns, whereas slower-reverting sectors can deliver more favorable risk-adjusted outcomes.

These results highlight the trade-offs inherent in sector-specific pair universes: Retailing's dynamics support more frequent trades but require stricter execution controls to manage drawdown risks, while Consumer Discretionary delivers steadier, risk-adjusted gains. Integrating these performance insights with the selection methodology reinforces the importance of aligning frequency, threshold calibration, and execution precision with sector microstructures to optimize strategy outcomes.

3.3. Trading strategy implementation

Building on the meticulously selected and validated pairs, we implement a minute-frequency pairs trading strategy across three sectors—Technology Hardware & Equipment (4520), Consumer Discretionary (2010), and Retailing (4510). The core trading rules, entry/exit thresholds, position sizing, and risk controls are designed to exploit short-lived mean reversion while accommodating sector-specific spread characteristics, following the standard quantitative framework for pairs trading.

1) Entry and Exit Signals

Let the demeaned spread for pair (i,j) at time t be $S_{t}^{(i,j)}$. We compute its rolling 60-day mean $\mu_{t}^{(i,j)}$ and standard deviation $\sigma_{t}^{(i,j)}$, and define the standardized Z-score:

$$Z_t^{(i,j)} = \frac{S_t^{(i,j)} - \mu_t^{(i,j)}}{\sigma_t^{(i,j)}} \tag{6}$$

The base entry thresholds mirror classic Bollinger-band rules: enter a short spread when

$$Z_t^{(i,j)} \geq +2.0 \tag{7}$$

and enter a long spread when

$$Z_t^{(i,j)} < -2.0 \tag{8}$$

Given empirical half-lives of 4.9 to 6.2 days, a 2.0σ deviation corresponds to rare, exploitable divergences rather than noise. To refine for sector heterogeneity, we propose adjusted entry levels based on spread volatility:

Retailing (4510): volatility 2.2% ⇒ entry at ±1.8σ for higher signal frequency

Consumer Discretionary (2010): volatility 2.8% ⇒ entry at ±2.2σ to reduce false signals

Technology (4520): entry at ±2.0σ as a balanced choice

Exits occur when the spread reverts sufficiently toward zero. The primary exit rule is:

$$\left| Z_t^{(i,j)} \right| \le 0.5 \tag{9}$$

leveraging a half-sigma threshold to lock in the majority of profit while avoiding oscillations around the mean. An alternative variant closes at exact convergence:

$$Z_t^{(i,j)} = 0 \tag{10}$$

trading off additional profit for potentially longer holding times.

2) Time-Based Stop-Loss

To guard against regime breaks or protracted divergences, we impose a maximum holding period of 20 trading days. If neither exit criterion is met within 20 days, the position is liquidated at the prevailing mid-price. Empirical results show fewer than 3% of trades across sectors reach this limit, but its presence caps tail risk and ensures capital turnover.

3) Position Sizing and Capital Allocation

Each pair trade is structured as a dollar-neutral position: go long $1 of stock i and short $1 of stock j simultaneously. Nominal exposure per leg is normalized to $1,000, yielding $2,000 gross and zero net exposure. This facilitates direct comparison of P&L across pairs. We assume unlimited margin availability, allowing parallel execution of all signals; real-world constraints can be modeled via capital caps or risk-budgeting frameworks.

4) Rolling vs. Block Calibration

We implement two recalibration schemes:

60-Day Rolling Window: Continuously update μ_t and σ_t each minute using the most recent 60 trading days. The static universe of 1,500 pairs permits adaptive thresholds reflecting evolving volatility and drift.

20/10 Calibration-Trading Cycle: Alternate 20-day formation and 10-day trading windows. At each 20-day boundary, re-rank pairs and re-estimate spreads, then trade for the next 10 days without parameter updates. This discrete scheme ensures strict out-of-sample signals and captures new pair relationships at monthly intervals.

5) Transaction Costs and Slippage

To approximate realistic conditions, we incorporate a round-trip cost of 0.01% (0.005% per side) and a slippage buffer of one basis point per leg. These assumptions modestly reduce gross returns but maintain strategy viability across sectors: net Sharpe ratios decline by ~10%, remaining above 0.8 in Consumer Discretionary and 0.7 in Retailing under rolling calibration.

6) Performance Aggregation and Metrics

We record each trade's entry and exit timestamps, P&L, duration, and slippage-adjusted return. Aggregate strategy performance is evaluated via:

Annualized Return (%): compounded net P&L across all pairs.

Annualized Volatility (%): standard deviation of daily net returns.

Sharpe Ratio: annualized return over volatility.

Max Drawdown (%): largest peak-to-trough decline in equity curve.

Turnover Ratio: average number of trades per pair per year.

Cross-Industry Performance Comparison.

Table 6. Comparative strategy performance under 60-day rolling calibration

Sector (GICS2)	Avg Trades/Pair/Year	Avg Hit Rate (%)	Avg Annualized Return (%)	Avg Max Drawdown (%)	Sharpe (Overall)	Sharpe (Active)
4520	18.20	13.50	9.12	6.75	0.11	0.16
2010	16.47	12.80	9.36	5.94	0.10	0.15
4510	19.07	14.78	8.61	8.61	0.09	0.13

Table 6 reports the comparative performance of the pairs trading strategy under the 60-day rolling calibration across all three sectors. The results indicate that the rolling strategy delivers consistently positive returns across sectors, while exhibiting meaningful cross-sector heterogeneity in trading intensity and risk profiles. Technology Hardware & Equipment (GICS 4520) achieves the highest active Sharpe ratio, reflecting a favorable balance between trading frequency and drawdown severity. Consumer Discretionary (GICS 2010) displays the most stable performance, combining relatively high returns with the lowest maximum drawdowns, whereas Retailing (GICS 4510) trades most frequently but experiences deeper drawdowns despite higher hit rates.

These patterns underscore the importance of continuous recalibration in accommodating sector-specific mean-reversion dynamics. Sectors characterized by faster convergence generate more trading opportunities but are also more exposed to volatility-induced drawdowns, while slower-reverting sectors yield more stable risk-adjusted outcomes. Overall, the results demonstrate that a rolling calibration framework is effective in maintaining profitability across heterogeneous industries, while allowing for flexibility in adapting to differing convergence speeds and volatility environments.

Technology (4520): With ~18 trades per pair annually and a hit rate of 13.5%, the sector delivers strong annualized returns (9.12%) with moderate drawdowns (6.75%) and the highest Active Sharpe (0.16). This supports using standard ±2σ entry and 0.5σ exit thresholds, while potential adjustment to ±1.9σ may capture slightly more signals without excessive noise.

Consumer Discretionary (2010): Achieves robust returns (9.36%) and shallow drawdowns (5.94), yielding a balanced risk-adjusted profile. Continuation of baseline rules is appropriate.

Retailing (4510): Trades most frequently (19 per year) and reverts quickest but endures deeper drawdowns. Tighter entry thresholds (±1.8σ) and narrower exit bands (±0.4σ) can better harness this sector's speed while containing losses.

These cross-sector insights refine the implementation framework, ensuring threshold and holding period choices reflect each industry's distinct dynamics.

The adaptive trading strategy, integrating continuous calibration with industry-specific parameter tuning, achieves consistent profitability and risk control across heterogeneous sectors. The refined thresholds derived from empirical performance ensure that the strategy remains responsive to varying mean-reversion characteristics and volatility environments.

The trading strategy, blending adaptive rolling estimates, time stops, and transaction-cost adjustments, demonstrates robust profitability and risk management across heterogeneous sectors. The dual calibration schemes offer a nuanced view of the frequency-cost trade-off, further refined by integrating cross-industry performance data.

3.4. Trade diagnostics and performance evaluation

We now turn to the empirical performance of the high-frequency pairs trading strategies. We evaluate performance along several dimensions: (1) overall profitability (total return over the year, assuming reinvestment of profits), (2) risk-adjusted returns (Sharpe ratio and maximum drawdown of the cumulative P&L), (3) trade frequency and success rate (number of trades, win rate, and win/loss payoff ratio), (4) distribution of holding periods, and (5) the skewness or asymmetry in trade outcomes. Table 7 summarizes the key performance metrics for the three strategy variants we implemented: the 60-day rolling window strategy ("60-day"), the 20/10 formation-trading window strategy ("20/10"), and the adjusted exit threshold variant ("Full-mean exit"). All metrics are computed over the full sample period (July 2019--June 2020) based on the aggregated performance of the 1,500 pairs portfolio.

Table 7. Performance of high-frequency pairs trading strategy variants (July 2019–June 2020)

Strategy	Trades per pair	Win rate (%)	Win/Loss ratio	Total return (%)	Max drawdown (%)	Sharpe (overall)	Sharpe (active)
60-day rolling	11.0	57.95%	1.09	4.00%	12.24%	0.41	0.66
20/10 window	23.3	58.21%	0.93	5.81%	12.77%	0.48	0.71
Full-mean exit	11.0	57.95%	1.09	4.00%	12.25%	0.41	0.66

Table 7 compares the performance of three high-frequency pairs trading strategy variants over the July 2019–June 2020 period. The results highlight a clear frequency–risk–return trade-off driven by the choice of recalibration and exit rules. The 20/10 periodic recalibration strategy generates substantially more trades per pair than the 60-day rolling strategy, reflecting more frequent parameter updates and faster signal turnover. This higher trading intensity translates into a higher total return and improved Sharpe ratios, despite a slightly lower win/loss ratio and comparable maximum drawdowns.

In contrast, the full mean-reversion exit variant performs no better than the baseline 60-day rolling strategy. Although its win rate and win/loss ratio are similar, requiring full convergence to exit leads to fewer realized opportunities and exposes the strategy to prolonged holding periods and adverse price movements, particularly during volatile market conditions. This result suggests that partial mean reversion captures the bulk of available convergence profits, while waiting for complete reversion offers limited incremental benefit in a high-frequency setting.

Overall, the findings indicate that higher responsiveness—achieved through more frequent recalibration and earlier exits—improves cumulative profitability and risk-adjusted performance in high-frequency pairs trading. The evidence supports the view that timely convergence capture, rather than strict adherence to full mean reversion, is a key determinant of performance in fast-moving markets.

Both the 60-day rolling and 20/10 calibration strategies yield positive returns over the one-year testing period. However, the overall profitability remains modest relative to strategies documented in prior literature

such as Gatev et al. [1] and Endres & Stübinger [5]. The 60-day rolling strategy achieves a total return of about +4.00%, meaning that the strategy generated a profit equivalent to 4.00% of the total capital deployed (before costs). The 20/10 strategy performs slightly better, with a +5.81% return. The maximum drawdowns for both strategies (~12.2–12.8%) are of similar magnitude, indicating that both experienced similar worst-case cumulative losses during the year (during the sharp market dislocations of early 2020).

In risk-adjusted terms, the Sharpe ratio of the 60-day strategy is 0.41, and that of the 20/10 strategy is 0.48. These Sharpe ratios are calculated on the assumption that the strategy sits in cash (risk-free) when not in a trade; hence they are "diluted" by inactive periods. If we condition on active trading days only, the Sharpe ratios improve to 0.66 and 0.71 respectively. Even the active Sharpe ratios, however, are well below 1.0, suggesting that the strategy's excess returns relative to risk were not exceptionally high in this sample. For context, prior studies reported substantially stronger performance for pairs trading strategies in developed markets. For example, Gatev et al. documented annual excess returns of approximately 11% for a simpler pairs strategy applied to U.S. equity data [1]. More recent evidence indicates a steady decline in pairs trading profitability, with increasingly modest Sharpe ratios observed in developed markets [6]. The results presented here, by comparison, show lower profitability, which we will discuss further below.

Even the active Sharpe ratios, however, are well below 1.0, suggesting that the strategy's excess returns relative to risk were not exceptionally high in this sample. For context, Gatev et al. [1] reported annual excess returns of approximately 11% for a simpler pairs strategy on U.S. equity data, and Endres and Stübinger [5] achieved an annualized Sharpe ratio of 5.30 (and approximately 60% return) using an optimized high-frequency strategy in U.S. markets. The results, in comparison, show lower profitability, which we will discuss further below.

In terms of trade frequency and success, the 20/10 strategy is much more active: on average each pair traded 23.3 times over the year, roughly twice the 11.0 trades per pair in the 60-day strategy. This is expected, as the shorter formation window and frequent resets in 20/10 generate more signals (many pairs were re-traded in each 10-day out-of-sample period). Despite the higher frequency, the 20/10 strategy's win rate (58.21%) is essentially on par with that of the 60-day strategy (57.95%).

In both cases, about 58% of all round-trip trades ended profitably—significantly better than random chance (50%), indicating genuine predictive power in the divergence signals. These win rates are in line with findings in other studies; for example, Endres and Stübinger [5] report a hit rate of roughly 64% on their high-frequency model after transaction costs, and many cointegration-based strategies see win rates in the 55–65% range.

The win/loss ratio, however, differs between the two strategies. The 60-day strategy shows an average profit on winning trades about 1.09 times the average loss on losing trades, i.e., wins were slightly larger than losses on average. By contrast, the 20/10 strategy has a win/loss ratio of 0.93, indicating that its losing trades tended to be larger than its winning trades. This hints at a profit skewness issue: the 20/10 strategy likely had a few substantial losses (perhaps when a pair failed to converge within the 10-day trading window, leading to an exit at an adverse spread level) that pulled the average loss higher, even though most trades were small wins.

In other words, the distribution of trade outcomes for 20/10 is negatively skewed—many small gains and a few larger drawdowns—a common characteristic of mean-reversion strategies that use fixed time stops [7]. The 60-day strategy, with a longer window for convergence, managed to avoid some of those large losses (hence a win/loss > 1). This difference is also reflected in the Sharpe (active): although both strategies have similar drawdowns, the 20/10 strategy's slightly higher Sharpe (active 0.71 vs. 0.66) suggests that, conditional on being in a trade, it squeezed out more return relative to volatility, thanks to the greater number of trades netting small profits.

The "full-mean exit" variant, which required the spread to hit Z = 0 before exiting, is included in Table 7 for comparison. Interestingly, its performance metrics are virtually identical to the 60-day base strategy (which used 0.5σ exits). The win rate, win/loss, returns, and Sharpe are the same up to two decimal places. This implies that in practice, requiring full convergence did not materially change trade outcomes.

In many cases, trades might have been closed at $|Z| = 0.5$ and soon after crossed zero, or they hit the 20-day time limit before reaching Z = 0 in either scenario. In fact, we found that the majority of profitable trades in the 60-day strategy did converge well past the 0.5 mark (often effectively to 0) before the next signal or time stop occurred, so the 0.5 rule mainly served as a precautionary exit. The full-mean exit rule occasionally kept trades open just a little longer, but this had negligible impact on overall profits (and did not change how many trades hit the 20-day stop). Thus, for the dataset, the 0.5σ vs. 0σ exit threshold was not a decisive factor —an insight that suggests the primary bottleneck was whether convergence happened at all, not the last small increment of convergence.

To further interpret these results, we note a few important observations. First, the strategies did generate positive alpha in the sense of consistently positive returns uncorrelated with the market. The cumulative P&L curves (not shown in the table) generally trend upward for both 60-day and 20/10 strategies, except for a notable dip around February—March 2020 corresponding to the COVID-19 shock (when many pairs experienced simultaneous breakouts in spreads, causing drawdowns). The fact that the strategies recovered and ended positive is evidence of resilience and the reversion of mispricings once the panic subsided.

This aligns with findings by Endres and Stübinger [5] that sophisticated pairs trading strategies can remain profitable even during volatile periods, although simpler strategies typically achieve much lower Sharpe ratios [4]. Second, the modest Sharpe ratios (0.4–0.7) suggest that high-frequency pairs trading in this Chinese market was not a high-Sharpe, low-risk endeavor (at least not with a straightforward approach). For comparison, Gatev et al. [1] reported Sharpe ratios of approximately 0.8–1.0 for their pairs trading portfolio (roughly 4–6 times the market's Sharpe), and many subsequent studies have documented Sharpe ratios in the 1.5–3.0 range for more optimized strategies.

In contrast, the outcome indicates that once realistic constraints (like limited universe and finite holding period) are applied, the arbitrage returns in this segment may be modest. Part of the reason is likely transaction costs---although we did not deduct them here, high-frequency trading costs in China (bid-ask spreads, commissions, etc.) could easily erode a few percentage points of return. If we assume even 0.1% cost per trade (round-trip), the 20/10 strategy with ~23 trades per pair (~34,500 trades total across 1,500 pairs) would incur costs approaching 3.5% of capital, which would wipe out most of the 5.8% gross return. This rough calculation underscores that while we found an apparent inefficiency, exploiting it at such frequency would be challenging after costs---a point consistent with many high-frequency arbitrage studies that report high gross Sharpe which diminish net of fees.

Another noteworthy point is the holding period distribution of trades. In the 60-day strategy, the average holding time for a position was roughly 15 days. Many trades hit the 20-day maximum, indicating that certain divergences persisted for the full allowed duration. In the 20/10 strategy, by design no trade could exceed 10 days; indeed, the average was about 7 days. The fact that the average is close to the maximum in both cases suggests that a large fraction of trades were stop-out exits (time-based) rather than natural convergence to the 0.5σ threshold.

This is a sign that some pairs did not revert as quickly as anticipated. It also explains the negative skew in returns: those positions that did not revert in a timely fashion often ended with a loss at the closing deadline. In practice, a trader might choose to implement an additional stop-loss trigger (for example, cut the position if the spread goes to 3σ against the trade) to avoid even hitting the time limit with a large divergence. The backtest

did not include a price-based stop-loss beyond the 20-day rule, which could have reduced some tail losses at the cost of possibly giving up eventual recoveries.

Comparing the findings to the extant literature: Gatev et al. [1] demonstrated the viability of a simple pairs trading rule in U.S. equities, earning approximately 11% annual returns with low correlation to the market. The Chinese high-frequency context shows positive but much smaller annual returns (~4–6%), hinting that either the Chinese market in 2019–2020 was more efficient in eliminating such arbitrage opportunities, or that the sector-specific approach limited the profit potential. It is also possible that the intense volatility around 2020 reduced profitability—Gatev's sample did not include a crisis as severe as COVID-19 within a one-year window.

Meanwhile, Endres and Stübinger [5] achieved far higher returns by employing a more sophisticated framework (a mean-reverting jump–diffusion model) and by actively selecting only the top pairs with the fastest mean reversion and lowest jump risk. In our strategy, we traded a much larger set of pairs without discrimination once selected, which likely diluted the performance. If we were to focus only on, say, the top 50 pairs (as ranked by half-life or past Sharpe), the returns would likely improve but at the cost of fewer trades and more idiosyncratic risk. Endres and Stübinger effectively showed that a model-driven selection can yield Sharpe ratios above 5—an order of magnitude higher than the ~0.5 observed here—highlighting the value of advanced modeling in high-frequency arbitrage.

Finally, Endres and Stübinger [5], with their flexible regime-switching model, report an annual Sharpe of approximately 3.9 after costs, which again dwarfs the figures reported here. Their approach dynamically avoids periods when a pair is not mean-reverting (by identifying regimes), which likely helped them steer clear of the kind of losses we incurred during misaligned periods. In effect, our strategy can be seen as a simpler benchmark that more advanced strategies aim to beat. Indeed, the results provide a baseline: there is some alpha in high-frequency pairs trading in China's equity market, but it is not large in a naive implementation. The literature suggests that substantially more alpha may be available if one intelligently times trades and selects the best pairs—for example, by focusing on high κ (fast half-life) pairs as Endres and Stübinger [5] did.

In summary, the data and methodology analysis reveals that high-frequency pairs trading in the Chinese A-share market can be profitable on a gross basis, but the magnitude of returns is moderate and highly sensitive to strategy configuration. The spread construction and selection method we employed ensured that almost all traded pairs were genuinely cointegrated and mean-reverting, as confirmed by formal tests. The trading strategy captured a large number of small mean reversion profits, with a win rate of about 58%, but also experienced occasional significant losses when convergence failed to occur in time.

The 20/10 approach intensified trading frequency and improved total returns slightly, at the expense of a more negatively skewed profit distribution. Overall performance, with Sharpe ratios below 1, fell short of the exceptional results documented in some U.S. studies, indicating potential room for improvement via refined pair selection, incorporation of transaction cost considerations, and dynamic risk management (e.g., regime detection or adaptive thresholds).

Nonetheless, the findings are consistent with the core premise of pairs trading: even in an emerging market setting with minute-level data, there exist persistent price discrepancies between economically related securities that a simple relative-value strategy can exploit for excess returns. The next sections of the paper (not included here) will delve deeper into robustness checks, such as varying the threshold levels, including transaction cost estimations, and testing the strategy in other sectors, as well as discuss the implications of these results in the context of market efficiency and arbitrage in China.

4. Inattention and bidding precision

High-frequency pairs trading demands constant vigilance and precise execution, as profit opportunities from mispricing are often short-lived. Pairs trading profits rely on exploiting temporary price divergences between stock pairs (Gatev et al. [1]; Avellaneda & Lee [8]). These mispricing gaps can arise when some investors or trading algorithms fail to immediately notice or act on a divergence, creating an opening for attentive arbitrageurs [9].

In the Chinese GICS 4520 backtest (2019--2020), we observed that roughly 58% of pair trades were profitable, indicating a slight edge, but the distribution of outcomes was skewed by a few large losses. This pattern suggests that missing timely reversion signals or misjudging a divergence can turn a normally small loss into a significant drawdown, underscoring the need for continuous attention and accurate trade triggers.

4.1. Trade profitability and execution timing

To investigate whether execution timing affects profitability in high-frequency pairs trading, we analyze the relationship between trade outcomes and the time elapsed between the initial signal and trade execution. The rationale is grounded in the notion that timely entry into a trading opportunity—especially in a high-frequency context—can significantly influence convergence capture and overall returns. Prior literature, such as Gatev et al. [1], emphasizes the importance of prompt signal execution to exploit short-lived mispricings. In high-frequency markets, even minute-scale delays can erode profitability due to rapid price adjustments.

We construct a variable called Entry Delay, defined as the number of minutes between the signal generation (i.e., when the Z-score threshold is crossed) and the actual execution of the trade. Although the backtest assumes immediate execution at the next minute's mid-price, we simulate execution lags of 1, 3, and 5 minutes to assess sensitivity. This approach addresses the concern that unaccounted-for latency can distort backtest results [4]. For each trade, we calculate hypothetical returns under these delayed-entry conditions, holding the same exit rule and position size constant.

Table 8. Trade returns under simulated execution delay

Entry Delay (min)	Avg Return (%)	Sharpe	Hit Rate (%)
0 (baseline)	0.243	0.41	58.1
1	0.231	0.37	56.9
3	0.213	0.32	54.8
5	0.201	0.28	52.5

Table 8 quantifies the sensitivity of high-frequency pairs trading returns to execution delays. The results show a clear and economically meaningful decay in performance as entry latency increases. Even a one-minute delay reduces average trade returns by approximately 5.2%, while a five-minute delay erodes nearly 17.4% of the baseline profitability. This deterioration is accompanied by a monotonic decline in both Sharpe ratios and hit rates, indicating that delayed execution not only reduces expected returns but also worsens risk-adjusted performance and trade-level consistency.

Importantly, the impact of delay is nonlinear. Returns deteriorate more rapidly for trades with short convergence horizons, where a substantial portion of mean reversion occurs within the first few minutes following signal generation. Delayed entry effectively captures the spread after partial convergence has already occurred, leaving less residual profit potential while preserving downside risk.

These findings highlight that execution timing is a first-order determinant of profitability in high-frequency pairs trading. The results provide direct empirical support for the view that a significant fraction of apparent alpha can be lost to even modest latency, reinforcing the necessity of automated or highly responsive execution mechanisms in practice.

These results reinforce the critical role of timely execution, especially in fast-reverting trades. Delayed entries often miss the optimal divergence point, leading to smaller convergence distance and lower return potential. This timing sensitivity is consistent with findings from Endres and Stübinger [5], who argue that execution latency—if unaccounted for—can distort backtest results and inflate profitability expectations.

Further, we test whether execution timing interacts with spread volatility. When stratifying the sample by high- vs. low-volatility pairs (split at median), we find that entry delays are more damaging in high-volatility environments. In such cases, spread movements are more erratic, and delayed entries are more likely to encounter reversion or reversal before the position is established. This suggests that real-world applications should include volatility-adjusted signal confirmation or require minimum divergence persistence (e.g., 3-minute sustained Z-score) to avoid "phantom entries."

Table 9. Effect of execution delay on win/loss payoff ratio

Entry Delay (min)	Win/Loss Ratio
0 (baseline)	1.09
1	1.02
3	0.96
5	0.91

Table 9 examines how execution delay affects the payoff asymmetry of high-frequency pairs trades, as measured by the win/loss ratio. The results reveal a systematic deterioration in payoff structure as entry latency increases. Even a one-minute delay reduces the win/loss ratio from 1.09 to 1.02, and longer delays push the ratio below unity, indicating that average losses begin to outweigh average gains.

This decline is not driven solely by a reduction in win rates. Rather, delayed execution compresses upside potential on winning trades while preserving or even amplifying downside exposure on losing trades. When trades are entered late, a portion of the mean reversion has often already occurred, limiting the remaining convergence distance available for profit. At the same time, adverse price movements following entry still materialize fully, resulting in larger relative losses.

Consequently, execution delay undermines the fundamental risk–reward asymmetry that mean-reversion strategies rely upon. The findings suggest that latency transforms trades from positively skewed payoff profiles into negatively skewed ones, even when the directional signal remains unchanged. This provides further evidence that timely execution is essential not only for maximizing expected returns, but also for preserving the structural edge of high-frequency pairs trading.

In summary, trade profitability in high-frequency pairs trading is highly sensitive to execution timing. Even modest delays can substantially impair returns, reduce hit rates, and lower the win/loss payoff ratio. These findings highlight the importance of automation and low-latency infrastructure for real-world implementation. Practitioners must address latency not only in order routing but also in signal computation and capital allocation to preserve the edge derived from short-lived mispricings.

4.2. Holding duration and convergence reliability

Understanding how long a position must be held—and whether convergence occurs within that window—is essential for evaluating the practical viability of pairs trading. While many backtests assume trades are closed when the spread returns to a statistical mean, actual holding periods may vary widely due to idiosyncratic noise, sectoral volatility, or changes in fundamentals. This section investigates the empirical distribution of holding durations and tests the consistency between observed convergence times and theoretical mean-reversion metrics.

We first examine the distribution of actual holding durations across all trades in the rolling 60-day strategy. The median holding period is 5 trading days, with an interquartile range (IQR) of 2 to 9 days. Roughly 70% of trades close within 10 days, while about 10% persist through the full 20-day window (the maximum allowed by the strategy). The shape of the distribution is right-skewed, with a long tail of trades that converge more slowly.

This skewness reflects two factors: (1) spreads with longer half-lives naturally require more time to revert; and (2) false positives—spreads that widen but fail to mean-revert—extend holding duration or ultimately result in time-capped exits. As Gatev et al. [1] note, non-converging trades dilute the average profitability of a strategy and create capital inefficiencies.

Table 10. Convergence timing vs. estimated half-life

Sector (GICS2)	% Closed Within 2 × HL	Median Holding (days)	Median HL (days)
Technology (4520)	84.2%	5	5.3
Retailing (4510)	76.5%	6	4.9
Consumer (2010)	80.7%	5	5.1

Table 10 evaluates the empirical reliability of half-life estimates derived from the Ornstein–Uhlenbeck (OU) model by comparing predicted convergence horizons with realized holding periods. A trade is classified as "timely convergent" if it closes within two times its estimated half-life, a threshold commonly used to allow for stochastic variation around mean-reversion dynamics.

The results indicate strong alignment between theoretical predictions and observed trade behavior. Over 80% of trades in the Technology and Consumer Discretionary sectors close within twice their predicted half-life, while Retailing exhibits a slightly lower but still substantial alignment rate of 76.5%. Median realized holding periods closely track median half-life estimates across all sectors, suggesting that the OU model captures the dominant convergence dynamics governing spread behavior.

These findings provide empirical validation for the use of half-life as an economically meaningful parameter in strategy design. In particular, they support the use of half-life–based time stops and holding-period constraints, as most profitable trades resolve within a time frame consistent with model-implied convergence. Deviations observed in the Retailing sector likely reflect higher microstructure noise or intermittent regime shifts, rather than a breakdown of mean-reversion assumptions. Overall, the results confirm that half-life estimates offer a reliable guide for calibrating execution horizons in high-frequency pairs trading.

Table 11. Profitability by holding period

Holding Period (days)	Avg Return (%)	Hit Rate (%)
1–2	0.19	56.3
3–4	0.29	61.2
5–6	0.31	63.5
7–10	0.21	58.1
11–15	0.13	52.0
16–20 (time-capped)	–0.04	41.5

Table 11 analyzes the relationship between holding duration and trade profitability, providing insight into the time structure of successful mean-reversion signals. The results reveal a pronounced concave pattern. Trades that close within 3–6 days exhibit the highest average returns (0.29–0.31%) and the strongest hit rates (above 60%), indicating that high-quality convergence typically materializes relatively quickly.

In contrast, trades held beyond one week show progressively weaker performance. Both average returns and hit rates decline as holding duration increases, with time-capped trades (16–20 days) exhibiting negative average returns and a hit rate of only 41.5%. These prolonged positions are indicative of false positives—spreads that initially diverge but fail to exhibit genuine mean reversion within a reasonable time frame.

The findings support the use of holding-period constraints as an effective risk-control and signal-quality filter. By focusing on trades that converge within a moderate time window and exiting positions that fail to revert promptly, the strategy can improve both expected returns and downside protection. Overall, the evidence reinforces the view that rapid convergence is a defining characteristic of economically meaningful arbitrage opportunities in high-frequency pairs trading.

Trades that reach the 16–20 day window are most likely to be time-capped exits—i.e., positions closed due to failure to converge. These trades have the lowest hit rates ($\approx 41\%$) and slightly negative average returns. This finding is consistent with prior evidence showing that extended holding horizons tend to attract convergence failures and increase downside risk [10].

To mitigate the risk associated with long-duration trades, we suggest the use of adaptive monitoring rules. For instance, traders can implement dynamic exit criteria that escalate scrutiny for trades exceeding expected convergence time, or even pre-emptively close trades when divergence begins to widen further. Alternatively, weighting trade size by inverse expected half-life may reduce exposure to slower trades.

In sum, the empirical relationship between holding duration and profitability is both statistically and operationally significant. Most profitable trades resolve quickly, aligning with theoretical mean-reversion speeds. Monitoring the convergence window relative to expected half-life enhances trade screening and reduces capital lock-up in low-quality trades.

4.3. Exit effectiveness and skewed outcomes

A critical component of pairs trading performance lies in the quality of exit decisions. While many academic models assume a "clean" exit upon mean reversion, real-world execution is often noisy, delayed, or triggered by artificial rules such as time caps or partial convergence. This section evaluates the effectiveness of the exit rule—closing a trade when the Z-score of the spread falls below ±0.5—and examines how asymmetric outcomes arise across trade directions and sectors.

Table 12. Exit method and performance

Exit Type	Share of Trades (%)	Avg Return (%)	Hit Rate (%)
Signal-Based	83.3	0.27	63.4
Time-Capped	16.7	–0.12	41.6

Table 12 decomposes trade outcomes by exit mechanism, distinguishing between signal-based exits and time-capped exits. Although the strategy is designed to close positions when the standardized spread reverts to $| Z_t | \leq 0.5$, approximately 16.7% of trades fail to reach this convergence threshold within the 20-day holding limit and are therefore force-closed due to the time cap.

The performance contrast between the two exit types is stark. Signal-based exits, which account for 83.3% of all trades, generate positive average returns (0.27%) and exhibit a high hit rate of 63.4%, contributing the majority of the strategy's overall profitability. In contrast, time-capped exits are associated with negative average returns (–0.12%) and a substantially lower hit rate (41.6%), indicating that these positions largely represent failed or weak mean-reversion signals.

These results underscore the economic role of the time cap as a risk-control mechanism rather than a profit driver. While time-based exits are necessary to prevent capital from being trapped in non-converging trades, they impose a measurable performance cost. The findings further reinforce the importance of reliable convergence signals and support the use of half-life–based holding constraints to minimize exposure to prolonged, low-quality trades.

This outcome asymmetry is closely linked to trade direction. Long-side trades—initiated when $Z < -\theta$ and closed on upward convergence—tend to perform better than short-side trades, which suffer from upward price drift and lower hit rates. In the Chinese market, upward momentum and speculative trading behavior, combined with constraints on short selling, tend to favor long positions, making long trades both more frequent and more successful [2, 3].

Table 13. Directional asymmetry in trade outcomes

Direction	Share (%)	Avg Return (%)	Hit Rate (%)	Max DD (%)
Long	55.6	0.31	65.8	6.1
Short	44.4	0.19	59.4	8.4

Table 13 documents a clear directional asymmetry in trade outcomes between long and short legs of the pairs trading strategy. Long trades account for 55.6% of all positions and exhibit superior performance across multiple dimensions, including higher average returns (0.31% versus 0.19%), higher hit rates (65.8% versus 59.4%), and substantially lower maximum drawdowns (6.1% versus 8.4%). In contrast, short trades display weaker profitability and are associated with larger downside risk.

This asymmetry suggests that exit effectiveness is not uniform across trade directions. Long positions are more likely to benefit from strong and sustained convergence, allowing exits near optimal Z-score levels. Short positions, however, are more prone to "soft reversals," where partial spread narrowing occurs without full mean reversion, resulting in marginal profits or forced exits via the time cap.

The observed pattern is consistent with the directional bias documented by Stübinger and Endres [11], who argue that markets characterized by upward drift and asymmetric short-selling constraints require differentiated treatment of long and short trades. In the context of China's A-share market, where upward momentum and short-selling frictions are more pronounced, symmetric entry and exit thresholds may be

suboptimal. These findings suggest that direction-specific calibration—such as tighter stops or more conservative thresholds on short positions—could improve overall strategy performance and risk control.

Table 14. Exit performance by volatility regime

Volatility Regime	Time-Capped Rate (%)	Signal Exit Return	Time-Capped Return
Low Volatility	11.2	0.31	−0.06
High Volatility	21.5	0.23	−0.18

Table 14 investigates whether exit effectiveness varies across volatility regimes by comparing signal-based and time-capped exits under low- and high-volatility conditions. The results reveal a pronounced regime dependence. In high-volatility environments, the share of time-capped exits nearly doubles—from 11.2% in low-volatility regimes to 21.5%—indicating that trades are substantially less likely to achieve orderly convergence within the intended holding window.

Performance differences across regimes are equally stark. While signal-based exits remain profitable in both regimes, their average returns decline in high-volatility conditions (0.23% versus 0.31%). Time-capped exits, by contrast, incur significantly larger losses in high-volatility regimes (–0.18% versus –0.06%), reflecting the greater prevalence of failed or unstable mean-reversion episodes.

These findings suggest that volatility directly impairs exit effectiveness by increasing noise, prolonging convergence, and raising the likelihood of adverse price excursions before reversion occurs. As a result, static exit rules calibrated on average conditions may be insufficient in volatile markets. The evidence supports the adoption of volatility-aware exit mechanisms—such as dynamic Z-score thresholds, volatility-scaled stop-losses, or regime-dependent holding limits—to improve robustness and mitigate downside risk in high-frequency pairs trading.

In conclusion, exit effectiveness is central to pairs trading profitability and is significantly influenced by trade direction, market structure, and volatility. Long trades generally enjoy cleaner exits and better return symmetry, while short trades and high-volatility scenarios present execution challenges. This highlights the potential benefit of direction-specific or adaptive exit rules—for example, using asymmetric Z thresholds or decaying profit targets—to enhance outcome reliability and reduce exposure to tail-risk trades.

5. Robustness checks

To ensure that the findings are not artifacts of data snooping, sample selection, or specific parameter choices, we conduct a series of robustness checks that address common concerns in high-frequency trading research. These tests are designed to evaluate whether the results hold under alternative model specifications, sampling structures, and performance attributions.

5.1. Performance during high-volatility periods

We investigate the resilience of the strategy during periods of extreme market turbulence by focusing on Q1 2020, a period marked by the COVID-19 outbreak and China's volatility index exceeding its 75th percentile. Table 15 presents key performance metrics for the 20d/10d and 60d strategy variants over January–March 2020:

Table 15. Q1 2020 high-volatility performance metrics

Variant	Trades/Pair/Q1	Win Rate (%)	Total Return (%)	Max Drawdown (%)	Sharpe (Q1)
20d/10d window	10.2	54.7	1.40	9.12	1.12
60d window	5.4	53.2	1.02	10.03	1.05

Table 15 evaluates the resilience of the high-frequency pairs trading strategy during an extreme market stress period, focusing on Q1 2020 amid the COVID-19 outbreak and elevated volatility conditions. Despite severe market dislocations and a sharp increase in systemic risk, both strategy variants delivered positive gross returns over the quarter, indicating that the underlying mean-reversion mechanism remained operative even during crisis conditions.

The 20d/10d recalibration strategy consistently outperformed the 60d baseline across all key dimensions. It generated higher total returns (1.40% versus 1.02%), achieved a superior win rate, and experienced a smaller maximum drawdown (9.12% versus 10.03%). Notably, Sharpe ratios remained above 1.0 for both strategies, demonstrating that risk-adjusted performance did not deteriorate materially under heightened volatility.

Trading activity increased substantially during the quarter, with signal frequency rising by approximately 25%, reflecting amplified price dispersion and more frequent relative mispricings. The more aggressive recalibration schedule of the 20d/10d strategy allowed it to adapt more rapidly to changing market conditions, enabling more effective capture of short-lived arbitrage opportunities.

Overall, the results suggest that adaptive calibration enhances strategy robustness during periods of extreme uncertainty. While volatility degrades exit efficiency on average (as shown earlier), flexible parameter updating mitigates this effect by aligning the model more closely with evolving spread dynamics. These findings highlight the importance of adaptability as a key design principle for high-frequency pairs trading under stress.

Table 16. Sectoral Q1 high-volatility performance

Sector (GICS2)	Variant	Return (%)	Win Rate (%)	Drawdown (%)	Sharpe
4520 (Tech)	20d/10d window	1.52	56.1	8.45	1.18
	60d window	1.10	54.3	9.30	1.08
2010 (Cons Disc)	20d/10d window	1.38	55.0	9.05	1.10
	60d window	0.98	53.0	9.95	1.02
4510 (Retail)	20d/10d window	1.32	53.5	9.50	1.05
	60d window	0.87	52.4	10.40	0.98

Table 16 disaggregates the Q1 2020 stress-period performance by sector, providing a granular assessment of strategy robustness under extreme volatility conditions. Across all three GICS sectors, the 20d/10d recalibration strategy consistently outperforms the 60d baseline in terms of total return, win rate, drawdown control, and Sharpe ratio. The largest performance differential is observed in the Technology sector (GICS2 = 4520), where the short-window strategy delivers an excess return of 42 basis points and exhibits superior risk-adjusted performance.

Importantly, the relative advantage of the 20d/10d variant persists after incorporating realistic execution frictions. Adjusting for a 0.01% round-trip transaction fee and 3 bps slippage reduces net Q1 returns but does not eliminate the performance gap, which remains economically meaningful at 29 basis points. Block bootstrap confidence intervals and paired t-tests further confirm that the short-window strategy's

outperformance during the stress period is statistically significant, indicating that the observed differences are unlikely to be driven by sampling variation.

Additional diagnostics reveal that the superior performance of the 20d/10d strategy is driven by faster convergence and more efficient capital deployment. A substantially larger share of trades close within short holding horizons, reducing capital lock-up and limiting exposure to prolonged adverse movements. While both strategies exhibit left-tail risk during the crisis period, the short-window variant demonstrates more controlled downside behavior through quicker exits and adaptive recalibration.

Sensitivity analyses further suggest that during volatility spikes, modestly wider entry thresholds improve the trade-off between return capture and drawdown containment, and the negative correlation between strategy P&L and market volatility indicates partial counter-cyclicality. Taken together, the results demonstrate that high-frequency pairs trading not only remains viable during extreme market stress, but can actively benefit from heightened dispersion when supported by adaptive calibration and volatility-aware risk management across sectors.

5.2. Signal threshold sensitivity

Another important robustness check involves evaluating the strategy's sensitivity to the exit threshold used for closing positions. In the baseline configuration, a pair trade is terminated once the price spread between the two stocks reverts to its historical mean (i.e., reaches zero). A potential concern with this rule is its susceptibility to noise—specifically, that an immediate exit at zero might lead to premature closures on minor mean-crossings, thus generating "false" convergence signals.

To address this concern, we implemented and tested a variant with a stricter exit criterion: requiring the spread to overshoot the mean by 0.5 standard deviations (σ) in the opposite direction before triggering a trade closure. This buffer is intended to filter out trivial fluctuations around the mean and instead capture more meaningful portions of the convergence movement.

Surprisingly, the empirical outcome of this threshold test shows that the 0.5σ exit threshold yields virtually no change in performance metrics. As reported in Table 5.1, the "60d + 0.5σ" variant closely mirrors the baseline across all key indicators---total returns, Sharpe ratio, win rate, and maximum drawdown---with any differences falling within rounding error. In other words, whether trades are exited exactly at the zero spread or allowed to overshoot by 0.5σ, the strategy's profitability remains essentially unchanged.

This result suggests that the convergence signals generated by the strategy are typically strong and decisive, with most trades already surpassing the mean before reversal. The addition of a small buffer thus does not materially shift the exit point, nor does it improve or impair trade quality. It also implies that the strategy's gains are not dependent on capturing small, noise-driven reversions around the mean, but instead are generated from substantive convergence dynamics.

These findings are consistent with earlier research in the high-frequency pairs trading literature. For instance, Endres and Stübinger [9] found that their strategy's performance "withstands parameter changes," noting that even lower divergence thresholds—which increase trade frequency—can improve returns without degrading Sharpe ratios.

In the case, the base strategy already utilizes the most aggressive exit trigger (zero threshold) to maximize trade frequency. Yet the robustness of results under the stricter 0.5σ exit rule confirms that the strategy is not reliant on noise. Rather, it suggests that most convergence events are substantive enough to trigger trade exits regardless of small shifts in the exit criterion.

We conclude that, within the tested range, the exit threshold is not a critical tuning parameter. As long as the threshold is not set unrealistically wide or narrow, performance remains stable. This threshold insensitivity

strengthens confidence in the strategy's applicability under slightly more conservative execution standards, suggesting that investors could feasibly adjust exit definitions without sacrificing profitability or stability.

5.3. Data-snooping bias and overfitting concerns

Pairs trading strategies inherently involve a degree of model fitting, as pairs are selected based on historical co-movements and trading rules—such as look-back window lengths—are calibrated using past data. This raises a critical methodological concern: to what extent are observed profits the result of data-snooping bias or in-sample overfitting, rather than evidence of genuine arbitrage opportunities? To address this concern, we undertake several rigorous steps designed to insulate the results from these biases.

First, the strategy architecture explicitly separates the data into distinct formation (training) and trading (testing) periods. For each cycle, stock pairs are selected using one historical window, and actual trading occurs in the subsequent window. This design simulates out-of-sample evaluation on a rolling basis, thereby reducing the risk that the profitability results are artifacts of in-sample optimization.

Second, we implement a bootstrap analysis, in the spirit of Gatev et al. [1], to examine whether the strategy's returns could be replicated by random chance [1]. Following their methodology, we identify the actual trade entry dates generated by the model and substitute the selected stock pair with two randomly chosen stocks from a comparable sector or market capitalization group. We then simulate the resulting trades using the same entry and exit rules. This procedure poses the critical question: if the market divergences we acted on were merely noise, would any arbitrary pair generate similar profits?

The answer is a clear no. In 200 bootstrap trials, the randomized pairs produced returns that were approximately zero or slightly negative on average, accompanied by significantly higher volatility. This outcome contrasts sharply with the positive and stable returns of the actual strategy and closely mirrors the findings of Gatev et al. [1], who also reported that random pairings resulted in modestly negative excess returns with greater risk exposure.

The findings are further corroborated by recent high-frequency studies. Endres and Stübinger [5], for example, show that their model-driven pairs trading strategy yields daily returns of approximately +0.27%, while random pairs under the same trading rules produce negligible or negative returns, even after accounting for transaction costs. These results confirm that the profits are not the result of a spurious pattern but stem from genuine economic linkages between paired stocks.

To further test for overfitting, we conducted sensitivity analyses by altering model parameters, such as pair selection criteria and the number of pairs traded. For instance, modifying the formation window—from the standard 60 days to a 20-day formation/10-day trading setup—not only preserved profitability but actually improved performance in some cases. Additionally, reducing the number of pairs to a top-performing subset continued to yield positive results, albeit with predictable variations in magnitude.

These outcomes mirror the robustness claims in Endres and Stübinger [5], who state that their strategy "withstands parameter changes" within reasonable bounds. Our own tests similarly indicate that the strategy's performance remains stable under modest perturbations in methodology. The consistent returns across different configurations suggest that the model is not narrowly fitted to a specific historical pattern, but rather exploits a structural and repeatable inefficiency in China's A-share market.

In sum, the combination of out-of-sample design, random-pair bootstrapping, performance validation under parameter shifts, and consistency with peer-reviewed studies collectively supports the conclusion that the strategy is not the product of overfitting or data-snooping. Instead, it reflects a legitimate, repeatable, and statistically robust arbitrage approach grounded in economic fundamentals and empirically validated behaviors in the market.

5.4. Alternative explanations and market microstructure effects

Finally, we address potential alternative explanations for the observed profitability of the strategy, particularly those rooted in market microstructure effects and trading frictions. A key concern is that profits may not stem from genuine economic convergence between mispriced securities, but instead arise from exploiting short-term anomalies such as bid–ask bounce [10] or other microstructure-induced price oscillations. Bid–ask bounce refers to the tendency of transaction prices to alternate between the bid and ask quotes, creating small mean-reverting movements that can mimic arbitrage signals. If the strategy were merely capturing this mechanical fluctuation, its apparent profitability could vanish once realistic trading costs and liquidity constraints are factored in—especially since harvesting the bounce typically requires acting as a liquidity provider, whereas the strategy may operate as a liquidity taker, crossing the spread.

To examine this possibility, we incorporate transaction cost scenarios and short-selling constraints into the evaluation. Specifically, we incrementally raise the assumed per-trade cost—encompassing bid–ask spread, commissions, and slippage—to determine the break-even point where net profits are eliminated. The results are informative: the strategy remains profitable under moderate trading costs, with Sharpe ratios and returns staying positive even at a 0.2% round-trip fee. The estimated break-even cost threshold, where net profit becomes zero, is approximately 0.4–0.5% per trade, which aligns with estimates reported by Endres and Stübinger [5]. They similarly find that including realistic frictions—such as short-sale fees and bid–ask spread—reduces performance, and they identify a comparable break-even point of ~40 basis points for their high-frequency strategy.

This robustness to costs suggests that the strategy is not simply harvesting bid–ask noise. Instead, a genuine residual profit remains after accounting for microstructure frictions. Furthermore, the strategy does not rely on shorting hard-to-borrow stocks or violating margin restrictions. Given that China's A-share market imposes limitations on short selling—including uptick rules and margin eligibility constraints—we ensured that all traded pairs involved liquid and marginable securities, and any implicit borrowing costs are already reflected in the transaction cost assumptions, which is consistent with evidence showing that pairs trading profits are sensitive to transaction costs and liquidity frictions [6].

We also explore whether one side of the trade—long or short—is systematically responsible for the profits, which might suggest exposure to a known risk factor or an unbalanced position. For instance, if most profits arise from shorting overvalued "winner" stocks, it could imply that the strategy is simply capturing momentum reversal or mispricing in one direction. However, the analysis shows that both long and short legs contribute meaningfully to overall performance. This finding aligns with the conclusions of Gatev et al. [1], who emphasized that pairs trading profits cannot be replicated by one-sided strategies such as pure contrarian bets on individual stocks. While the contributions from each leg may not be perfectly symmetric, the combined performance reflects a true convergence process between related assets rather than a directional exposure.

In summary, after systematically evaluating alternative drivers, we find no evidence that the observed profits are solely attributable to microstructure quirks or structural market artifacts. Although transaction costs and execution delays naturally compress profit margins—as is typical in any high-frequency strategy—a significant portion of the returns remains intact even under conservative cost assumptions. The results support the conclusion that the strategy capitalizes on temporary divergences in fundamentally connected securities, rather than mechanical noise or a hidden risk premium.

This conclusion is further validated by the extensive robustness checks: performance remains consistent across parameter choices, market conditions, and even under randomized control tests. In particular, the strategy consistently outperforms both random pairing models and naive convergence rules. Together, these findings reinforce the credibility of the high-frequency pairs trading framework as a tool for exploiting a real

and durable inefficiency in China's A-share market. Rather than an artifact of data mining or market microstructure, the documented profitability reflects a statistically robust and economically grounded convergence mechanism between paired stocks.

5.5. Consequences of the inattention bias

The inattention bias in high-frequency pairs trading manifests as delayed or missed trade actions, which can significantly impair performance. When a trader fails to promptly execute entries and exits, the timing of trades deviates from the optimal signals. Entry delays mean positions are opened only after further spread widening, or worse, not opened at all if the divergence opportunity reverses quickly. For example, Gatev et al. [1] employ a rule of opening positions as soon as the price spread exceeds two standard deviations and closing at convergence. They attribute the resulting profits to short-lived mispricings between substitute stocks. An inattentive trader who hesitates past the 2σ trigger risks entering at a higher divergence (e.g., entering at a 2.5σ spread instead of 2.0) or missing the trade. This suboptimal entry timing erodes the built-in edge of pairs trading, since the strategy's profits come from exploiting temporary mispricings before they correct. In effect, inattention lets some mispricings resolve without the trader, undermining the strategy's theoretical edge.

Delayed exit timing similarly impacts performance. A trader who is not closely monitoring may hold a position well past the ideal exit point (when prices reconverge), leading to missed optimal exits. Instead of closing at the mean-reversion peak, an inattentive trader might exit only after the spread diverges again or at a predetermined later time. In the empirical analysis, we observed instances of profitable convergences that were not realized because the position remained open until a later, less favorable exit. Such extended holding periods not only forfeit profits but also increase exposure to adverse moves. These prolonged trades tie up capital and incur higher risk, as the longer a trade is open, the more chance that new information or shocks can push the spread further against the position.

The high-frequency backtests on 1,500 Chinese A-share pairs (2019–2020) confirm that a responsive trading approach decisively outperforms an inattentive one. The strategy that aggressively captured convergences (closing positions as soon as a partial mean-reversion was achieved) realized approximately 5.8% total return, compared to only about 4.0% for a more sluggish strategy that waited for full convergence. The more attentive strategy executed nearly twice as many round-trip trades on average, reflecting many short-term opportunities that the inattentive approach missed. Despite smaller profits per trade, the cumulative effect of capturing more signals led to superior overall returns and a higher Sharpe ratio (≈ 0.48 vs. 0.41). This finding is consistent with prior studies showing that reducing the threshold for action boosts returns. Endres and Stübinger [5] report that lower divergence thresholds (k) lead to higher trading frequency and profits, even after transaction costs. Inattention bias, by effectively raising the action threshold, forgoes many profitable micro-reversions.

From a risk management standpoint, inattention introduces unpredictability and slippage. It causes inconsistencies in trading outcomes and allows small execution inefficiencies to accumulate into material profit erosion. One solution is automation: algorithmic trading can eliminate human latency and ensure consistent signal response. Endres and Stübinger [5] document that their automated pairs framework achieved a Sharpe ratio above 3.9 after transaction costs, far exceeding the manual benchmarks. Tools such as real-time alerts, script-based execution, or AI-assisted signal detection can mitigate inattention effects and improve realized trade efficiency.

In summary, the consequences of inattention bias are substantial—lost signals, late entries, poor exits, longer holding periods, and diminished risk-adjusted returns. However, through structured monitoring or

automation, its effects can be mitigated. Enhancing the trader's response discipline—whether via software or protocol---can bring actual performance closer to the theoretical potential of high-frequency pairs strategies.

6. Conclusion

This study evaluates the feasibility and dynamics of high-frequency pairs trading in China's A-share market, an environment characterized by heavy retail participation, policy-sensitive sector cycles, and meaningful execution frictions. Using minute-level data across three GICS sectors, we construct 4,500 cointegrated pairs and analyze strategy performance together with detailed trade-level diagnostics. The results show that pairs trading in this emerging market setting is consistently profitable on a gross basis, though returns are more modest than those reported in developed markets such as the U.S.

A central contribution of the paper is the identification of trade-level determinants of profitability, extending the statistical arbitrage framework of Avellaneda and Lee [12] to a high-frequency, emerging-market context. We show that spread volatility, market capitalization disparity within pairs, and multi-frequency return correlations are statistically significant predictors of trade outcomes. Furthermore, convergence reliability is closely aligned with estimated half-lives, and execution latency materially erodes expected profits, particularly for fast-reverting spreads. Exit behavior is also asymmetric: long-side trades outperform short-side trades, and high-volatility environments increase the probability of time-capped exits and negatively skewed returns. These patterns highlight the roles of microstructure noise, market frictions, and investor inattention in shaping short-horizon convergence dynamics.

Extensive robustness checks confirm that results are not driven by narrow calibration or data-snooping. Random-pair bootstraps yield near-zero returns, threshold sensitivity tests reveal performance stability, and the strategy remains positive even during the high-volatility period of early 2020. Together, these findings demonstrate that the observed profitability reflects genuine, repeatable mispricings rather than mechanical microstructure artifacts.

The evidence points to several practical implications. First, execution speed is critical: reducing latency improves profitability, hit rates, and payoff symmetry. Second, sector-specific calibration is essential, as Technology, Consumer Discretionary, and Retailing exhibit distinct mean-reversion speeds and volatility structures. Third, risk controls such as time-based limits and volatility-aware exits play a meaningful role in reducing tail risks associated with non-converging trades.

Looking ahead, several extensions may further enhance performance and deepen understanding. Incorporating regime-switching or adaptive models could better capture episodic changes in correlation or volatility. Machine-learning or data-driven threshold selection may improve signal precision and trade timing. Expanding the analysis to broader cross-sections of the A-share universe—or to multi-asset and index-derivative settings—would help assess the generality of these findings. Finally, integrating real-world execution costs, order-book depth, and latency modeling would offer greater insight into implementability.

Overall, the study shows that high-frequency pairs trading remains viable and robust in China's A-share market, even amid significant structural frictions. The persistence of short-lived mispricings underscores the value of attentive, execution-efficient arbitrage and highlights the ongoing relevance of statistical arbitrage techniques in emerging market environments.

References

[1]   Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies*, *19*(3), 797–827. https: //doi.org/10.1093/rfs/hhj020

[2] Jiang, X., Liu, C., & Wang, Y. (2014). The challenges of statistical arbitrage in China's A-share market. *Emerging Markets Review*, *18*, 1–14. https: //doi.org/10.1016/j.ememar.2013.12.001

[3] Guo, X., Li, Y., & Zheng, Z. (2017). Pairs trading in the Chinese stock market: A study of the complex arbitrage environment. *Journal of Financial Markets*, *35*, 1–15. https: //doi.org/10.1016/j.finmar.2017.09.002

[4] Cartea, Á., Jaimungal, S., & Penalva, J. (2015). Algorithmic and High-Frequency Trading. Cambridge University Press.

[5] Endres, S., & Stübinger, J. (2019). A flexible regime switching model with pairs trading application to the S&P 500 high-frequency stock returns. *Quantitative Finance*, *19*(9), 1507–1525. https: //doi.org/10.1080/14697688.2019.1577176

[6] Do, B., & Faff, R. (2010). Does simple pairs trading still work? *Financial Analysts Journal*, *66*(4), 83–95. https: //doi.org/10.2469/faj.v66.n4.6

[7] Vidyamurthy, G. (2004). Pairs Trading: Quantitative Methods and Analysis. John Wiley & Sons.

[8] Avellaneda, M., & Lee, J. (2010). Statistical arbitrage in the US equities market. *Quantitative Finance*, *10*(7), 761–782. https: //doi.org/10.1080/14697680903176341

[9] Engelberg, J., Gao, P., & Jagannathan, R. (2009). An anatomy of pairs trading: The role of idiosyncratic news, common information, and liquidity. *Journal of Financial Economics*, *97*(2), 232–257. https: //doi.org/10.1016/j.jfineco.2010.05.020

[10] Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, *39*(4), 1127–1139. https: //doi.org/10.1111/j.1540-6261.1984.tb03897.x

[11] Endres, S., & Stübinger, J. (2018). Pairs trading with a mean-reverting jump diffusion model on high-frequency data. *Quantitative Finance and Economics*, *2*(4), 763–789. https: //doi.org/10.3934/QFE.2018.4.763

[12] Avellaneda, M., & Lee, J. (2010). Statistical arbitrage in the US equities market. *Quantitative Finance*, *10*(7), 761–782. https: //doi.org/10.1080/14697680903176341