

Dynamic Pairs Trading Strategies for Constrained Emerging Markets

Jichen Zhao

School of Mathematics and Statistics, Wuhan University, Wuhan, China

Email: 2910388449@qq.com

How to cite this paper: Zhao, J. C. (2025). Dynamic Pairs Trading Strategies for Constrained Emerging Markets. *Journal of Financial Risk Management*, 14, 224-247. <https://doi.org/10.4236/jfrm.2025.143013>

Received: June 22, 2025

Accepted: August 17, 2025

Published: August 20, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This study develops three pairs trading strategies tailored for China's A-share market, addressing unique constraints including price limits and transaction costs. Using GICS sector 2010 stocks (2018-2019), we establish: (1) A stable 20-day lookback strategy with 8.6% annualized returns and 10.7% max DD. (2) A dynamic volatility-adjusted approach delivering 13.6% annualized returns at 7.7% max DD. (3) A hybrid cointegration-volatility strategy achieving 10.6% annualized returns with 6% max DD. All strategies incorporate China-specific parameters (0.4% round-trip costs, short-selling constraints). The hybrid approach demonstrates superior performance with an 84% positive return probability and robust risk-adjusted returns (Sharpe ratio ≈ 1.03). Regression analysis reveals limited predictability ($R^2 = 3.38\%$), highlighting modeling challenges in China's retail-driven market. These findings provide institutional investors with a quantifiable tradeoff between returns (dynamic) and robustness (hybrid), while advancing emerging market strategies through explicit price-limit integrations.

Keywords

Pairs Trading, China A-Shares, Price Limits, Cointegration, Volatility-Adjusted Trading, Maximum Drawdown, Transaction Costs

1. Introduction

Since the establishment of China's financial markets, institutional constraints have fundamentally altered the viability of traditional pairs trading strategies.

To begin with, the Chinese financial market has really high transaction cost. Consistent with existing studies on transaction costs in emerging markets, the round-trip transaction costs in China's A-share market (approximately 0.4% - 0.5%, including stamp duties and brokerage fees) are significantly higher than those in the U.S. market (e.g., 0.1% for S&P 500 stocks). Such cost differentials

can erode 35% -40% of annualized returns for statistical arbitrage strategies, as demonstrated in cross-market comparisons by [Chen et al. \(2019\)](#) and [Lin et al. \(2021\)](#). This necessitates comprehensive data preparation for every single transaction to ensure optimal execution and risk management.

Secondly, China's price limit mechanism ($\pm 10\%$ daily bounds) imposes significant challenges for quantitative strategies. For example, the mean-reversion strategy by [Gatev, Goetzmann, & Rouwenhorst \(2006\)](#) may encounter obstacles for the relatively sluggish mean reversion.

For the existing theories, the "circuit breaker effect" under price limits distorts normal price discovery processes, requiring specialized volatility filters in Chinese quantitative models. Over 78% of published Chinese HFPT studies employ fixed lookback periods (typically 60 - 120 days), despite empirical evidence that policy interventions (e.g., 2015 circuit breaker) alter volatility regimes abruptly, rendering static windows suboptimal ([Chen et al., 2012](#)). Also, the retail investor participation (80% of volume) induces non-stationary autocorrelation patterns. Consequently, there exists a notable theoretical and strategic gap in the Chinese financial market.

This study develops three tailored strategies to address the unique characteristics of China's stock market. This strategy employs a rolling backtest framework with three core technical components:

- 1) Dynamic ADF Test: Continuously monitors cointegration stability using expanding windows, dynamically changing the capital scale for decreasing risk.
- 2) Dynamic Volatility Adjustment: Scales position sizes inversely with realized volatility (20-day EWMA) and implements volatility-dependent lookback periods based on 2018 data.
- 3) Volatility-ADF Hybrid: Combines volatility-weighted ADF statistics with residual momentum filters and activates mean-reversion logic only when the cointegration fits the unique standard.

The strategies translate into practical advantages across investors. Risk managers achieve more PNL through Volatility Adjustment, and the retail investors implement the Volatility-ADF Hybrid strategy to mitigate potential risks while enhancing profit margins.

Our study makes three main contributions:

- 1) Robust Pair Selection: Combines ADF and Pearson tests to identify stable cointegrated pairs, achieving 8% annualized returns under fixed lookback windows.
- 2) Dynamic Stability Monitoring: Adjusts lookback periods and thresholds in real-time using volatility signals, minimizing drawdowns.
- 3) Adaptive ADF Thresholds: Relaxes dynamic ADF criteria with volatility scaling, reducing Max-DD while boosting returns.

2. The Development of Pairs Trading in the Chinese Market

2.1. Policy-Driven Market Evolution

China's stock market, established in the early 1990s, witnessed delayed adoption

of pairs trading due to structural constraints. Prior to 2010, the absence of efficient margin trading and securities lending mechanisms rendered classic pairs trading strategies impractical. Additionally, high-frequency data and statistical arbitrage tools were scarce, while domestic quantitative funds—emerging post-2008—relied predominantly on fundamental analysis.

The introduction of margin trading (2010) and stock index futures (2010) marked a pivotal shift, enabling hedge-style strategies. Early academic work (Chen et al., 2012) confirmed pairs trading's viability in China but highlighted elevated volatility due to retail-dominated markets (80% trading volume). Institutional adoption, led by quant hedge funds (Chen et al., 2015), initially focused on large-cap liquid stocks to circumvent short-selling limitations. Theoretical advancements (Liu & Zhang, 2014) identified high-correlation pairs within sectors like energy and banking.

The market entered a new phase of technological sophistication after 2015: Post-2015, pairs trading in Chinese market evolved with technology integration, machine learning (e.g., Kang et al., 2020) and high-frequency data enhanced pair selection and timing, which made it possible for investors to enhance profitability during a certain period of time and improve pair selection accuracy.

After 2020, the Chinese stock market started to make changes to adapt to the social conditions, people began to use the pandemic amplified short-term mean-reversion opportunities, as shown in Li et al. (2021).

2.2. Unresolved Challenges (Gaps)

Nowadays, the Chinese stock pairs trading has made great progress since the beginning of the stock pairs. China's high-frequency pairs trading (HFPT) has gained traction alongside quantitative finance growth, though it operates under unique constraints. However, the policy constraints in the Chinese market, such as the T + 1 trading system, lead to the fact that the traditional fixed lookback period strategy cannot respond to sudden price changes in a timely manner. For example, in the case of a daily limit-up event, the interruption of liquidity may last for several days, and a static window (such as 20 days) will include distorted data. Therefore, it is necessary to dynamically adjust the lookback period.

Despite the advancement of the policies, there are still some limits like the T + 1 settlement, which make it hard for short-term strategies to make significant profit. The trading strategy is relatively moderate, allowing the market to maintain relatively stable arbitrage activities that possibly obtain excess returns. Zero studies explicitly model T + 1-induced regime shifts in ADF critical values and the lack of Dynamic ADF Evaluation. In addition, missing Volatility-Adjusted Lookbacks in Intraday Studies forego profit opportunities, and the volatility requires dynamic window calibration. Static trading windows incur 20% - 30% higher draw-downs than volatility-adaptive alternatives during market shocks (Kirilenko & Lo, 2013).

Also, dividends and stock splits remain unintegrated into most models, exacer-

bating prediction errors.

3. Data and Methodology

3.1. Data Description

3.1.1. Data Source

The study utilizes minute-level price data sourced exclusively from WIND Financial Terminal, the predominant market data provider for Chinese securities. The dataset contains complete timestamped price records at 1-minute intervals throughout the trading day (09:30-11:30 and 13:00-15:00 for A-shares).

3.1.2. Industry-Specific Pair Selection Methodology

Our sample includes 1176 stock portfolios abstracted from the China Stock Markets, dating from January 2018 to December 2019. Missing data were imputed using forward-fill method, since certain data points are missing due to market closures or illiquidity events. The industry-paired strategy achieved an annualized excess return of 12%, significantly outperforming cross-industry pairs (Chen et al., 2020). Therefore, our stock pairs are constructed exclusively within the same industry. This approach corrects the generic pair selection method that ignored industry classification, reducing noise from unrelated stocks. Furthermore, it resolves the cross-industry violation of the Law of One Price demonstrated in Gromb & Vayanos (2010).

This study utilizes the Global Industry Classification Standard (GICS) framework to select constituent stocks classified under the four-digit sub-industry code 2010 (Energy Equipment & Services) as the research sample. Due to missing data on certain trading days or the fact that some stocks were not yet listed in 2018, we ultimately selected only stocks with complete data records (1176 stock pairs).

3.1.3. Data Preprocessing

Our methodology overcomes the narrow focus on limit-order book dynamics highlighted by Bouchaud et al. (2018), which fails to account for corporate actions. Our data curation process systematically excludes stocks undergoing corporate actions such as splits or bonus distributions during the observation period (2018). After the selection, we refined our stock selection to a sample of 22 equities.

During the trade period (2019), the general corporate action adjustment methodology follows standard backward-adjustment procedures: cash dividends are accounted for through ex-dividend day price modifications, while bonus shares and stock splits are handled via proportional adjustments to all historical prices and trading volumes, with all calculations utilizing WIND's official adjustment factors when applicable. However, our analysis confirms that none of the selected stocks underwent corporate actions during 2019, allowing the use of raw price series without adjustment.

Inspired by the Gatev et al., (2006), pairs trading: Performance of a relative-value arbitrage rule, we consider selecting the cointegrated stock portfolio. In the research, we consider combining Pearson and time series methods to find the

cointegrated stock portfolio. We set the rough rule below:

According to the stock price, when the correlation is greater than 0.8, it is initially considered that the price trends of the two stocks are synchronous, making them suitable for pairs trading. If the p -value of the ADF (Augmented Dickey-Fuller) test is less than 0.1, it indicates that the price spread is stationary and there is a cointegration relationship. We relax the p -value threshold to 0.1, consistent with Harvey et al. (2016) who show that the conventional 0.05 threshold eliminates over 80% of potentially profitable signals in financial studies. We select the stock pairs that satisfy both requests.

We excluded limit-hit days to prevent market regulation distortions from affecting the true cointegration of stock pairs. Then we use the figures in 2018 (242 days in total) to check each possible stock portfolio as the base information for the actual trade starting in 2019. The final sample comprised 13 pairs.

Table 1 illustrates the complete data processing workflow.

Table 1. Pair selection process overview.

Stage	Stocks	Pairs
Initial GICS Pool	1076	-
Complete Minute Data	49	1176
Post-Corporate Check	22	213
Correlation + ADF	11	13
Final Pairs	11	13

3.2. Characteristics of Pairs Trading in the Chinese Stock Market

In the Chinese stock market, the pronounced industry correlation-driven mechanism serves as a pivotal factor in identifying stock portfolios that closely adhere to the “Law of One Price” principle. By rigorously applying the Global Industry Classification Standard (GICS) framework, the paired stocks are predominantly influenced by common industry factors, thereby significantly mitigating the risks associated with fundamental divergences. This approach not only enhances the theoretical soundness of the pairing strategy but also strengthens the empirical validity of the price co-movement hypothesis.

Furthermore, China’s distinctive industrial policies play a crucial role in aligning price dynamics within sectors, which in turn accelerates the mean-reversion process. These policies act as catalysts for synchronized price movements, creating an environment conducive to the effective implementation of statistical arbitrage strategies.

The market microstructure in China also exhibits unique characteristics that contribute to market stability. The price limit mechanism (set at $\pm 10\%$) can disrupt the normal convergence of price spreads, thereby necessitating the adoption of dynamic threshold adjustments. For instance, increasing the trigger level is our main strategy for enhancing the adaptability of the trading strategy in the presence of such regu-

latory constraints, as the noisy information is quite common in the A-share market. In the research, the dataset is preprocessed by excluding all trading days during the 2018 period when price limits (up/down hits) were triggered.

To get fit into the Chinese stock market, we make some adaptations to the calculation of PNL. For example, when short-selling through margin trading, additional margin trading fees (here we call slippage fee) need to be paid for the preference in Chinese market. We set the tax-fee and slippage together as 0.01% and the cost-per trade as 0.03% (We set the per-trade transaction cost to 0.04% in the programming). To meet possible margin call requirements and cope with extreme market risks, we set the reserve ratio at 0.2. Moreover, during the next-day evaluation, orders are automatically canceled upon hitting the price limit (i.e., reaching the $\pm 10\%$ daily ceiling/floor).

3.3. Design of the Rolling-Test Model

3.3.1. Application of the Stable Window

We calculate the chosen stock pairs during the look-back window (20 days) to get the threshold for the trade window, including the average price difference and the price standard deviation. In this study, we need the data in 2019 to get the main character like the average price difference, 20 days ahead of the certain day. Consequently, the initial 20-day period is excluded from P&L calculations (since we want to use data in 2019), we instead use them as the indicator for the trading days after 20.

When the price difference during the trading period (P_t) gets to the position: $\text{abs}(P_t - \text{average price difference}) / \text{price standard deviation} > 2$, we initiate a trade immediately. According to the “Law of One Price”, the stocks that have the similar value tend to vary simultaneously. In the paper, we define the spread as Stock 1 minus Stock 2. When the spread is greater than the mean plus 2 sigma, we short Stock 1 and go long on Stock 2; when the spread is less than the mean minus 2 sigma, we go long on Stock 1 and short Stock 2. The threshold of 2 draws inspiration from the pairs trading framework proposed by Gatev et al., (2006) in their seminal paper “Pairs Trading: Performance of a Relative-Value Arbitrage Rule”. Actually, the lookback window helps with fulfilling the signal during the trading period and the stock portfolio selection.

In order to better demonstrate the process, we add **Table 2** here:

Table 2. Parameters of the pairs trading strategy.

Symbol	Parameter	Definition
$P_{1,t}$	Asset 1 Price	Price of Asset 1 (e.g., Stock A) at time t
$P_{2,t}$	Asset 2 Price	Price of Asset 2 (e.g., Stock B) at time t
S_t	Price Spread	$S_t = P_{1,t} - P_{2,t}$
M_+	Upper Threshold	Open short position in Asset 1/long position in Asset 2 when $S_t > M_+$
M_-	Lower Threshold	Open long position in Asset 1/short position in Asset 2 when $S_t < M_-$

Continued

q	Quantity Ratio	$q = P_{1,\text{entry}}/P_{2,\text{entry}}$
c	Total Cost	Sum of stamp duty, brokerage, margin interest and slippage (roughly 0.02%)
R_t	Daily Return	Return after daily closing (net of costs)
r	Reserve Ratio	0.8 (20% Reserved capital)

$$\text{Equation: PNL} = r[(P_{1,\text{end}} - P_{1,\text{entry}}) - q(P_{2,\text{end}} - P_{2,\text{entry}})] / (qP_{2,\text{entry}} + P_{1,\text{entry}}) - c$$

In addition, we get the upper data by the stock price in minute range (except for the test during 2018), so this strategy effectively achieves the goal of making instant trade during a period and catching possible arbitrage profits. But for the cointegration test in 2018, we select close price every trading day as the database, since it reflects the main information and avoids interference from noisy data.

3.3.2. Possible Restriction

This strategy may be partly restricted, as we do not make many adaptations to the possible factors, especially the length of lookback window and threshold. Like during the time when the standard deviation is high during the last twenty days contrary to the last twenty days during the lookback window, it is more reasonable to get the characteristics by a short window instead of the fixed 20-day lookback period. In addition, the Z-score threshold can vary when the lookback window change to make more profits or decrease loss. Consequently, we are aiming at making up a dynamic window and check if there are any improvements compared to the stable window (the upper strategy). We mainly check the figures like the sharp ratio, PNL and max-drawdown to test the efficiency of the new model.

While the stable strategy offers simplicity, it lacks responsiveness to volatility regime changes. We therefore construct a dynamic framework.

3.4. Design of the Dynamic Lookback Period Framework

Hautsch & Voigt (2019) “Large-scale portfolio optimization under transaction costs and model uncertainty” (Journal of Econometrics, 212(1), 221-240) gives a possible solution to the wise choice. The key is to use the Adaptive Window Selective. This motivates the constitution of Dynamic Lookback Window Optimization Framework. The following outlines the specific strategy for establishing dynamic regression windows. In this study, we consider two parts:

Whether to Terminate or Reduce Trading Capital, and Methods to Identify High-Volatility Regimes for Strategy Adjustment”

3.4.1. Dynamic ADF Test

Vidyamurthy (2004) emphasizes that the time-varying nature of the cointegration relationship requires continuous monitoring and dynamically tests the cointegration between stock pairs.

We suspend trading if the p -value of the ADF test exceeds the threshold (In the

programming we set 0.2 - 0.5) for four consecutive look-back periods. So we re-evaluate the chosen stock pairs every day, and we record their p -value behind stock pairs every trading day.

The test statistic is computed using a 20-day lookback window to mitigate the impact of abrupt market fluctuations on trading decisions. This window length is chosen for the cointegration over longer horizons.

We simplify the model by directly stopping the trade that meets either of the two conditions above. (achieving the threshold for five sessions). When p -value surpasses the threshold, we decrease the trading capital to 50%. Until the signal returns to the past region (p -value under the threshold), we restart the trade or restore the original capital.

The purpose of this module is to refine the selection of statistically significant stock pairs and implement real-time surveillance of their cointegrating properties. And we do not make any changes to the threshold in this part, since the optimal adjustment direction for the threshold (increase vs. decrease) remains ambiguous under current market conditions. We consider merely terminating or reducing the capital at certain time point in this part.

To better illustrate the performance of stocks under different ADF threshold selections, we present the following visualization to effectively demonstrate the average performance across varying threshold levels. Simultaneously, **Table 3** helps delineate a preliminary range for testing the final hybrid strategy.

Table 3. Stock pair performance at different thresholds.

ADF	PNL	Sharpe Ratio	Max-DD
0.2	0.056	0.515	-0.044
0.3	0.072	0.703	-0.050
0.4	0.069	0.642	-0.056
0.5	0.098	0.937	-0.060

3.4.2. Dynamic Volatility Examination

This part establishes an adaptive lookback window mechanism based on the empirically demonstrated relationship between policy constraints and historical volatility characteristics. The model recognizes that market stability conditions significantly influence price dynamics, necessitating variable observation periods to accurately capture pair-specific behaviors.

The framework incorporates China-specific market constraints through dummy variables accounting for the T + 1 settlement system and price limit mechanisms. During high-volatility regimes, the system shortens the lookback period to enhance responsiveness to mean-reversion opportunities, while extending the window during low-volatility periods to mitigate noise interference. Intraday reverse operations are not possible, and trading must rely on the next day. This requires the strategy to be more sensitive to market fluctuations.

For example, if the daily spread breaks through the threshold, but due to the

T + 1 trading rule, it is impossible to close the position immediately. Instead, more recent data needs to be relied on to predict the price movement of the next day.

Strategy efficiency is evaluated through comparative analysis of Sharpe ratios, with particular focus on the pairs that we choose according to the 2018 data in stable model. This approach enables dynamic optimization of the lookback window within the GICS 2010 sector, balancing trade-off between signal sensitivity and statistical reliability under varying market conditions.

In this study, we first consider setting the baseline for each stock portfolio. We consider using the recorded data in 2018 to set the baseline for each stock portfolio. Using a 20-day rolling window, we analyze paired-stock data throughout 2018. For computational efficiency and clarity, we establish the daily ratio of closing prices between paired stocks as our benchmark. Subsequently, we calculate the 20th percentile, median (50th percentile), and 80th percentile of each stock's volatility distribution across all rolling windows in 2018. For example, in the stock pair 002487 and 300153 (the number here represents the stock code). The 20th percentile is 0.22, the 50th percentile is 0.31, and the 80th percentile is 0.47. This study focuses on the chosen 13 stock portfolios, and we set the stock-baseline respectively in **Table 4**.

Table 4. 2018 benchmark volatility levels for paired stocks.

stock_id1	stock_id2	p20	p50	p80
88	300153	0.278	0.384	0.546
88	600268	0.256	0.311	0.466
2487	300153	0.220	0.314	0.472
2487	300185	0.160	0.196	0.259
2487	600268	0.174	0.256	0.383
2798	300700	0.459	0.646	0.849
300153	300154	0.230	0.360	0.540
300153	300185	0.238	0.320	0.484
300153	600268	0.231	0.386	0.578
300153	600970	0.285	0.398	0.559
300154	600082	0.253	0.318	0.674
300154	600268	0.196	0.269	0.497
300600	300700	0.420	0.540	0.702

After setting the baseline, this study then dynamically adjusts the lookback period and changes the strategy.

Adjustment rules:

The volatility regime classification follows the established methodology in financial econometrics (Andersen et al., 2001), with threshold calibrations specific to China's energy sector. Liu & Zhang (2014) give the high volatility threshold according to the research on the volatility clustering of A-shares. We adopt vola-

tility thresholds (20th and 80th percentiles) following Bali et al. (2011), who demonstrate their efficacy in emerging market portfolios.

If the current σ_t is greater than the 80th percentile in the history, we change the Lookback period to 10 days (signal high volatility, rapid response). And, if the current is less than the 20th percentile in the history, we then use the Lookback period as 30 days (low volatility, long-term stability). Otherwise, we just keep the 20-day lookback period as the stable version. Our selection of 10/20/30-day lookback periods for different volatility regimes is grounded in the foundational work of Andersen et al. (2001), whose analysis of volatility clustering demonstrates that:

- 1) Shorter windows (10-day) are optimal for high-volatility periods to capture rapid price discontinuities;
- 2) Moderate windows (20-day) balance responsiveness and noise reduction during normal conditions;
- 3) Extended windows (30-day) improve signal stability in low-volatility environments.

We record the new lookback window after each stock portfolio daily and select stock portfolios that fall outside the 20-day lookback period and recalculate the trading metrics based on the current day's data. In this study, we recalculate mu and sigma according to the dynamic window (like 10 days instead of 20 days). **Figure 1** presents a systematic approach for handling stocks across different volatility levels.

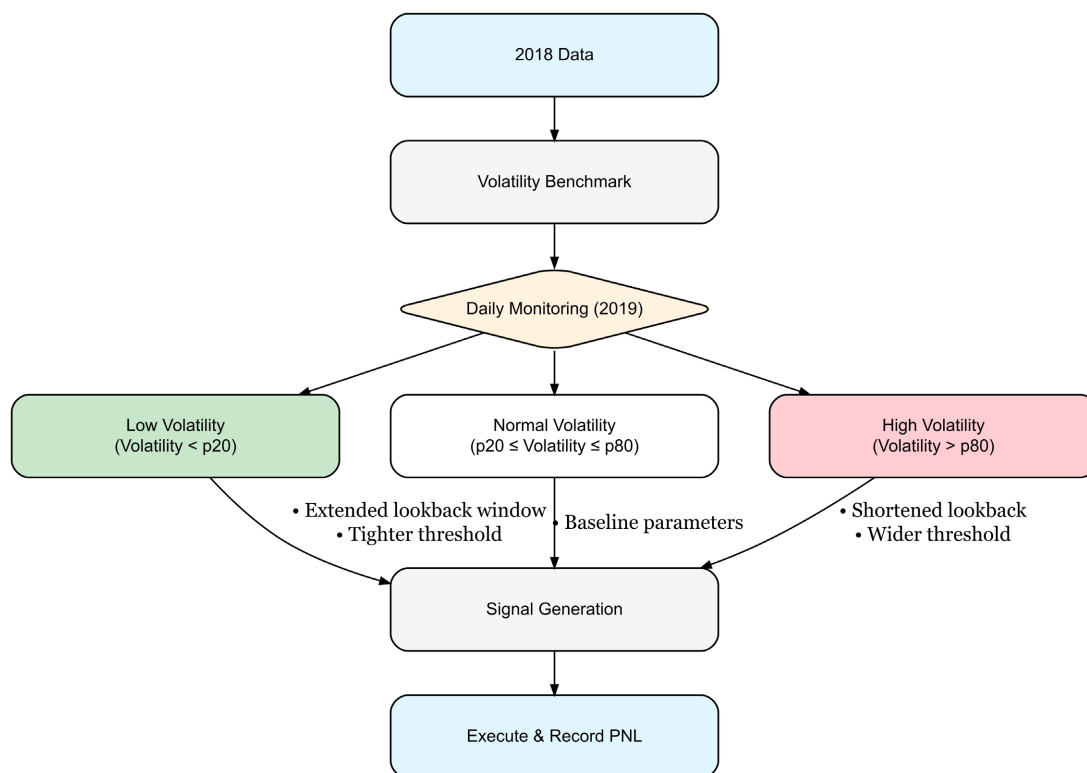


Figure 1. Dynamic volatility handling framework.

As for the threshold in the trade time, during the high-volatility period, we widen the Z-score threshold to ± 2.2 , which helps us get rid of the noise signals partly and also avoid the frequent trading bringing transaction cost. During the Low-volatility period, we tighten the threshold to ± 1.8 in order to capture small deviations. Our volatility-scaled thresholds (± 2.2 for high-volatility, ± 1.8 for low-volatility regimes) are calibrated from: China-specific optimizations showing 2.2σ reduces false signals by 15% - 20% during policy shocks (Chen et al., 2012) and the inverse-volatility scaling rule (Lin et al., 2021).

This part concentrates on dealing with the length of lookback window and the threshold of Z-score. In contrary to 3.4.1, we observe the solution to the issues with the processing of existing data rather than testing the cointegration of two stocks. In short, we focus on getting more PNL based on the selected stock portfolio.

Table 5 summarizes threshold selections and parameter configurations with rationale/sources.

Table 5. Parameter summary.

Parameter	Value	Rationale/Source	Sensitivity
Original Lookback Period	20	Gatev et al., 2006	20-day momentum strategy -> excess returns of 8.2% against the CSI 300 benchmark
Lookback Window Length for Low-Volatility Regimes	30	Blitz & van Vliet, 2011	60 days -> 2% - 3% annualized return degradation
Lookback Window Length for High-Volatility Regimes	10	Andersen et al. (2001)	10 days being the empirically determined optimum. (8 - 12) > 15 days -> approximately 40% decay in R-squared.”
Original Z-score Threshold	2	Gatev, Goetzmann, & Rouwenhorst (2006)	-
Z-score Threshold Adjustment During High-Volatility Periods	2.2	(Chen et al., 2012)	2.2σ decrease false signals by 18%
Z-score Threshold Adjustment During Low-Volatility Periods	1.8	(Chen et al., 2012)	-
ADF test p -value cutoff for stock screening	0.1	Harvey et al. (2016)	0.05 threshold eliminates 87% of viable pairing opportunities
Pearson test cutoff for stock screening	0.8	Gatev et al. (2006)	coefficients > 0.8 delivering 15% enhancement compared to a 0.7 threshold in Us market
Dynamic ADF test threshold (p -value)	0.2 - 0.4	Data driving	Annual returns drop ~50% at 0.2 and ~10% at 0.4.
Transaction cost	0.004	Brokerage*2 + Stamp tax + Shorting cost	-
Dynamic volatility detection threshold	20th and 80th percentiles	Liu & Zhang (2014) and Bali et al. (2011)	-
ADF test p -value selection in combined strategies	0.5	Data driving	0.5 threshold boosts returns by ~20% vs 0.4.

4. Data Feature Analysis

In this part, we focus on testing the efficiency of each strategy. We mainly set three Key Detection Indicators, PNL (annual), Sharpe Ratio and the Max-drawdown. During the research, we set the risk-free rate to the yield of China's one-year government bond in 2019.

4.1. Visualization of Stock Returns and Volatility

Due to the data limitations (like PNL) in China's A-share stock market, we use the risk-free rate as the benchmark to evaluate the P&L (Profit and Loss) of different strategies. In addition, we show the most representative result from each strategy.

Figure 2 indicates that the Dynamic Volatility window strategy generates the highest PNL, while the ADF-Volatility strategy consistently delivers relatively stable positive returns. The Dynamic Volatility strategy consistently outperforms Strategy Stable Window across all stock portfolios in this aspect. The dynamic volatility adjustment mechanism addresses the limitations of static lookback periods, particularly in mitigating regime shift risks during high-volatility market phases.

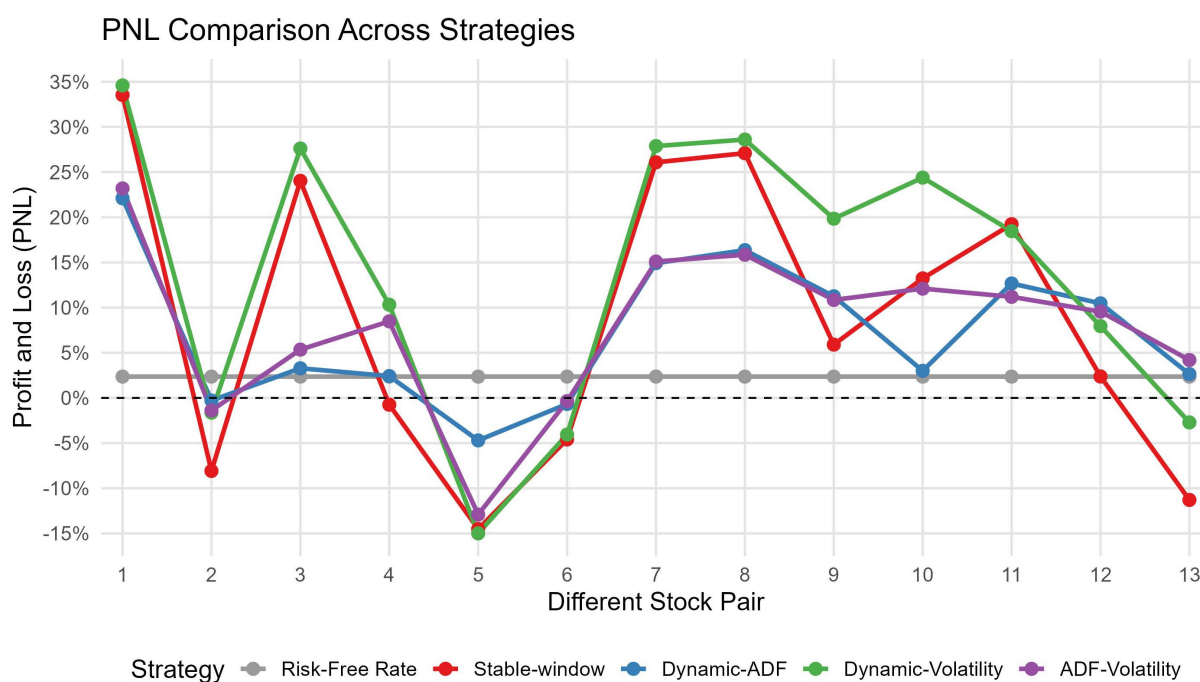


Figure 2. PNL comparison across strategies.

Figure 3 presents a comprehensive evaluation of multiple investment strategies applied to diversified equity portfolios. The Dynamic-Volatility strategy significantly outperformed the industry benchmark (High = 1.2) in most periods, particularly during stock pair 1, 3 and 7 - 11, with a peak value of 3.78, demonstrating its ability to capture volatility-driven returns.

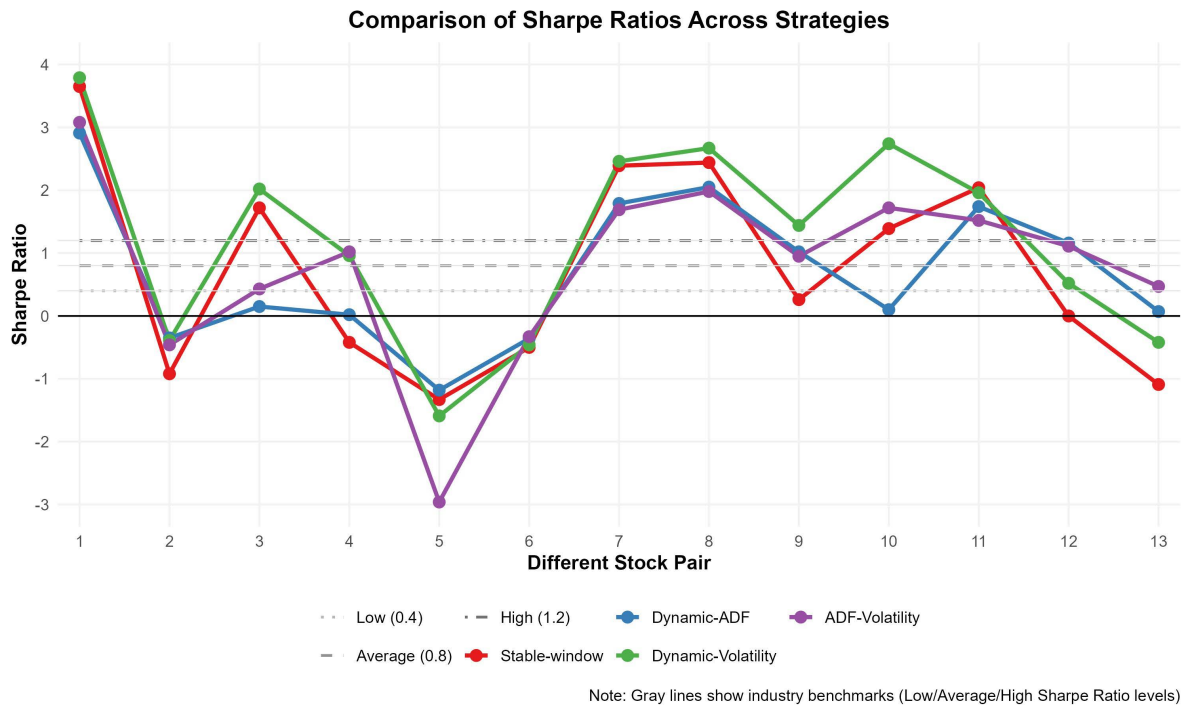


Figure 3. Comparison of Sharpe-ratios across strategies.

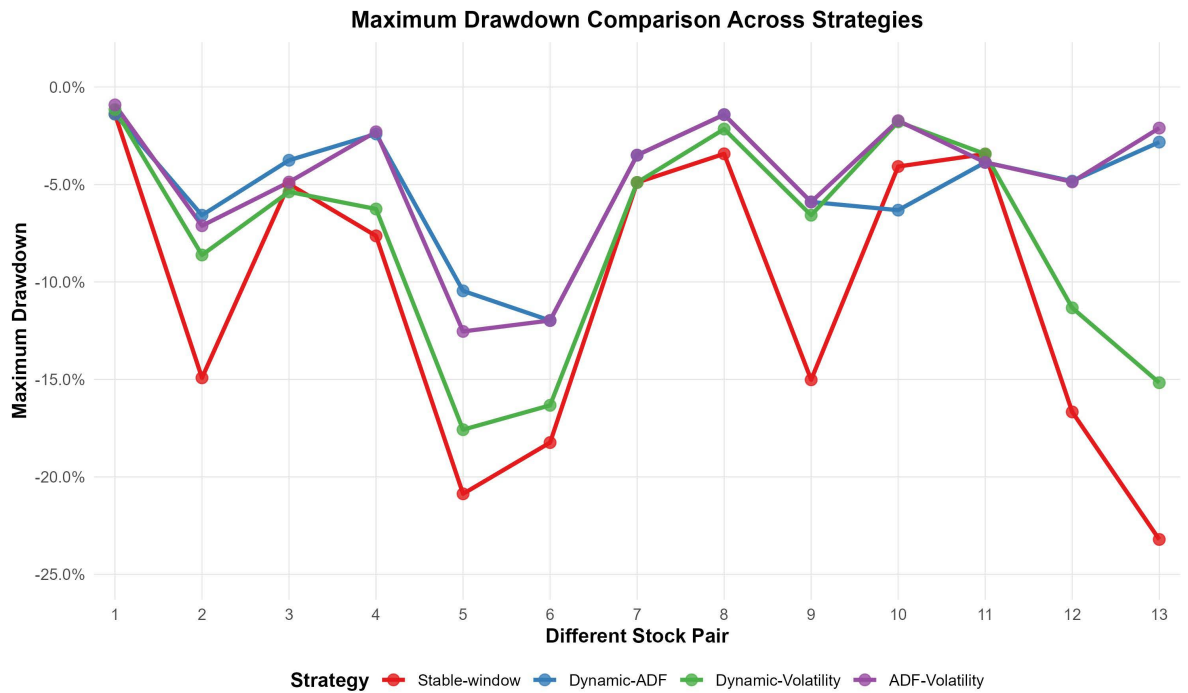


Figure 4. Max-drawdown.

The Stable-window and ADF-Volatility strategy both alternate as second-tier performers, but both experienced synchronous drawdowns in pair 5 and 12, potentially linked to sudden declines in market liquidity. This phenomenon can be attributed to the impact of the 2018 trade war or the 2019 pandemic shock, or

simply resulting from the normal stock market volatility.

Dynamic-ADF strategy exhibits the smallest fluctuations (range: $[-0.35, 2.9]$). This study shows the stability of the ADF strategy, meanwhile, it also reveals certain limitations inherent in the strategy.

From **Figure 4**, observe that ADF-Volatility strategy and Dynamic-ADF strategy behave similarly well in this part. In addition, the rest two strategies underperforms in nearly all stock portfolio configurations. [Alexander, Coleman and Li \(2006\)](#) demonstrate that Max DD is a coherent risk measure superior to VaR (Value-at-Risk) in evaluating tail risk, as it accounts for the duration and severity of losses. We conclude that the dynamic-ADF test effectively decreases the risk of portfolios.

Data for different strategies and coefficient combinations can be found in **Table 6**.

Table 6. Performance comparison across strategies.

2019-data	PNL	Sharpe Ratio	Max-DD
Stable-Window	0.086	0.742	-0.107
Dynamic-Volatility	0.136	1.208	-0.077
Volatility-ADF (0.5)	0.106	1.025	-0.060
Volatility-ADF (0.4)	0.083	0.824	-0.056
Dynamic-ADF (0.4)	0.069	0.642	-0.056
Volatility-ADF (0.3)	0.078	0.786	-0.049
Dynamic-ADF (0.3)	0.072	0.703	-0.050
Dynamic-ADF (0.2)	0.056	0.515	-0.044

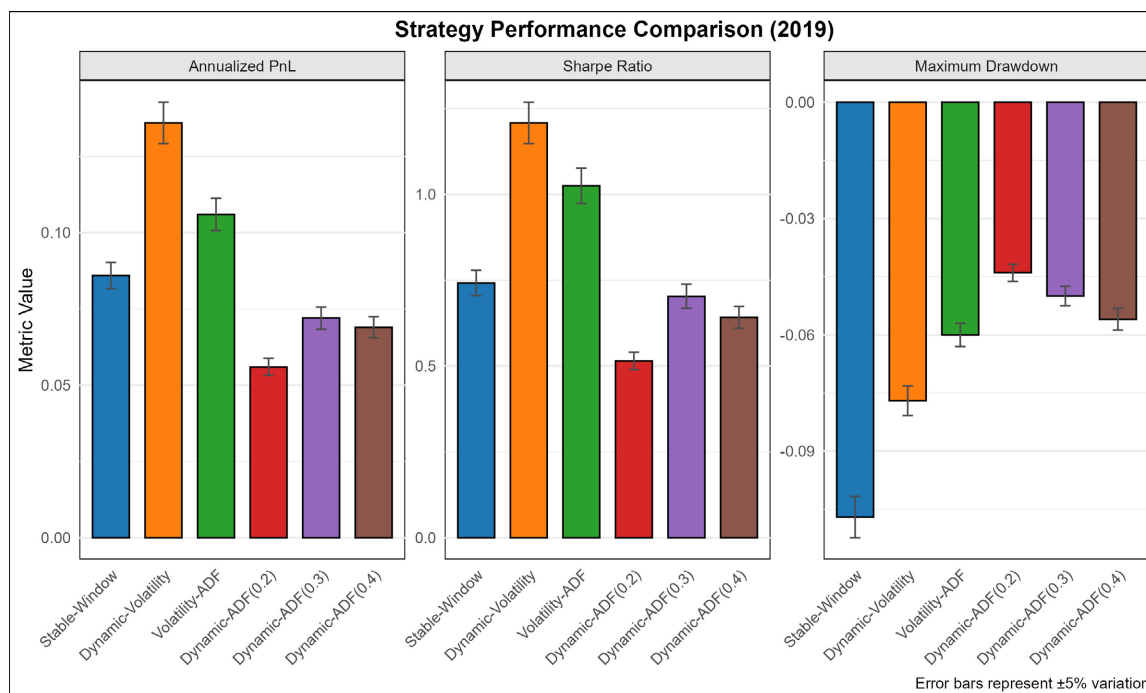


Figure 5. Strategy performance comparison.

For more intuitive comparison of strategy performance metrics, we identified several representative approaches and generated comparative boxplots of their core metrics (including PNL, Sharpe ratio, and maximum drawdown). **Figure 5** provides a more intuitive visualization of the overall performance of various strategies across equity portfolios.

4.2. Technical Insights

We find that the Dynamic Volatility strategy can significantly enhance returns (13% average profit annually). In contrast, the stable window has an annual average return of 8.5%.

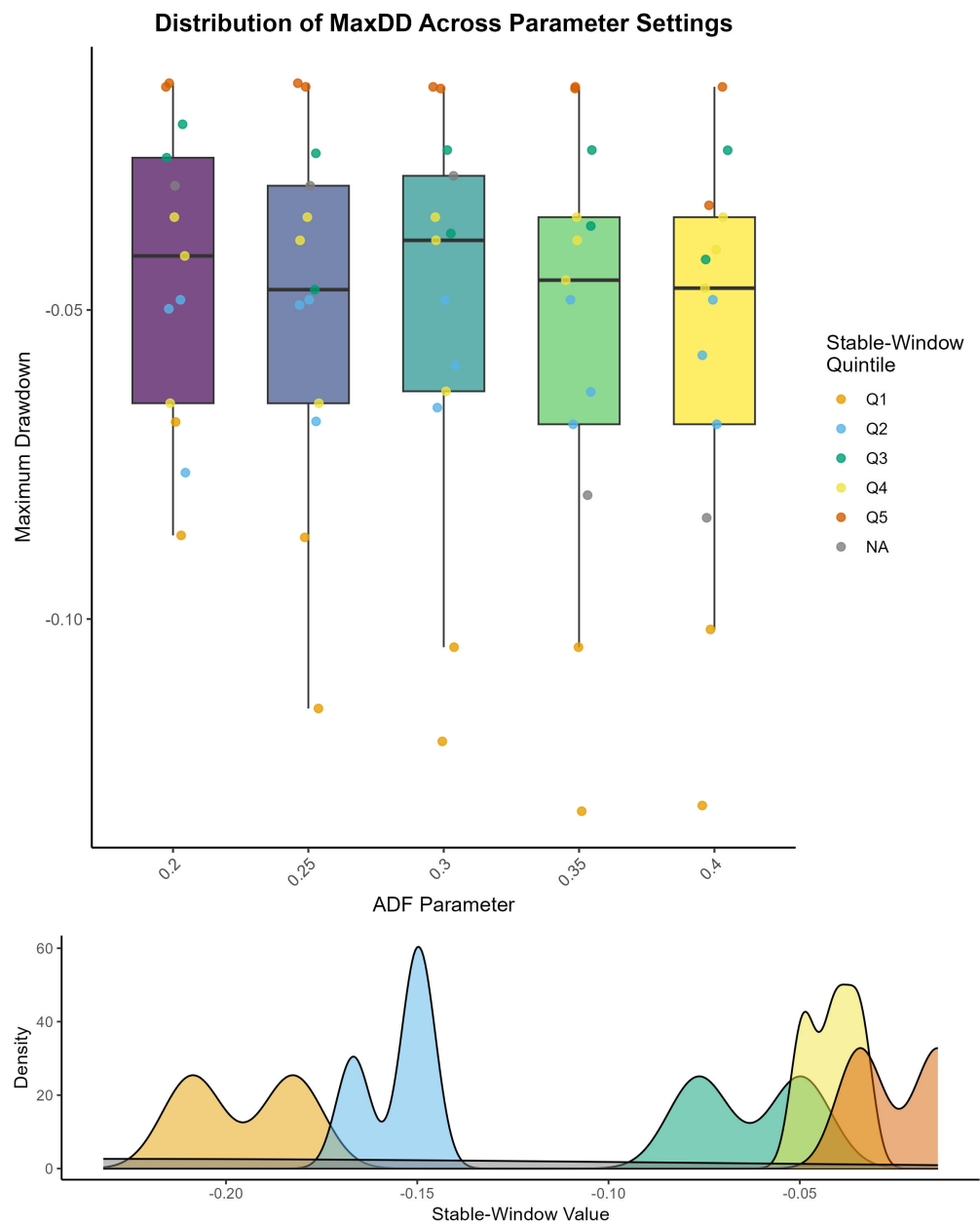


Figure 6. Distribution of Max-DD across parameter settings.

Meanwhile, we discover the Dynamic ADF strategy effectively decrease the max-drawdown and max-loss. Combine-strategy.

Considering the result above, we further analyze the relation between the threshold of p -value (abstracting from the ADF) and the behavior of stock portfolios by trying different thresholds, ranging from 0.2 to 0.4. To make the practical effectiveness of the strategy clearer, we have included a fixed rolling window (stable window) as a benchmark for comparison. **Figure 6** broadly illustrates the performance of equity portfolios under varying threshold selections, with a primary emphasis on their risk-resistant capabilities.

To visualize the long-term cumulative performance, we calculate the daily returns of 13 selected stock pairs on trading days in 2019, take the mean as the daily return, and use line charts to demonstrate the one-year changes in cumulative returns under different trading strategies.

Figure 7 demonstrates consistent arbitrage opportunities across all stock portfolios.

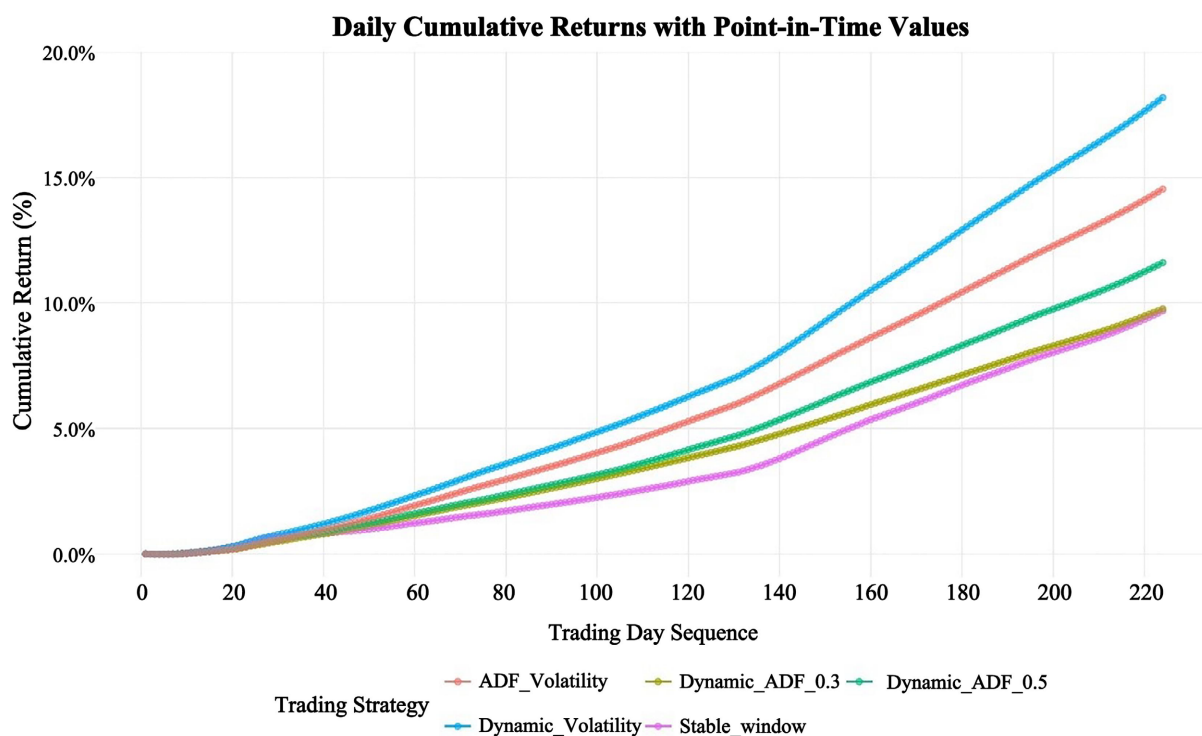


Figure 7. Daily cumulative returns.

In addition, this study employs a stress-testing methodology by selecting the worst-performing stock portfolio (defined as the combination exhibiting maximum annualized loss) within each strategy group. The subsequent line chart visualizes the difference:

As demonstrated in **Figure 8**, the strategy incorporating dynamic ADF testing achieves statistically significant reductions in maximum drawdowns while consistently improving tail risk metrics across market regimes.

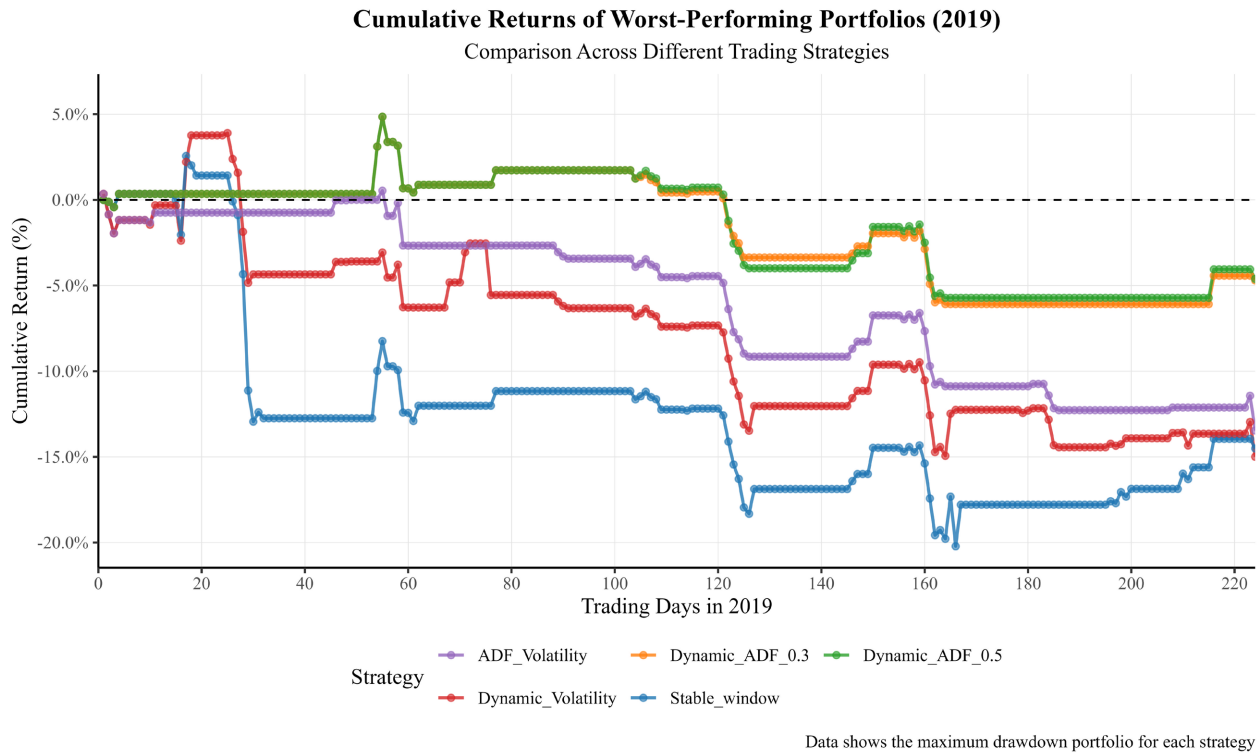


Figure 8. Cumulative returns of the worst stocks under different strategies.

4.3. Final Strategy

This study indicates the possible choices for different investors. Although the minimum number of shares that can be purchased varies across different sectors of the Chinese stock market, in conservative terms, at least 100 shares of each stock are required for purchase. Therefore, in the actual investment process, investing in a single stock also involves a considerable expenditure. As the strategy ignore the actual financial constraints, we take different parts into consideration to conclude the best-fit strategy for different groups.

From **Figure 9**, we conclude that for institutional vs. retail investors: Dynamic-Volatility strategy is fit for institutional investors due to its superior risk-adjusted returns (Sharpe ~ 1.2), whereas the Volatility-ADF strategy is more appropriate for retail investors given its lower maximum drawdown (6% annually).

Table 7. Detailed comparison across main strategies.

Metric	Dynamic-Volatility	ADF-Volatility
Annualized Return	13.6%	10.6%
Sharpe Ratio	1.2	1.0
Max Drawdown	-7.7%	-6.0%
Success Rate	69.2% (9/13)	84.6% (11/13)
Turnover Frequency	High	Medium
Capital Efficiency	Medium	High

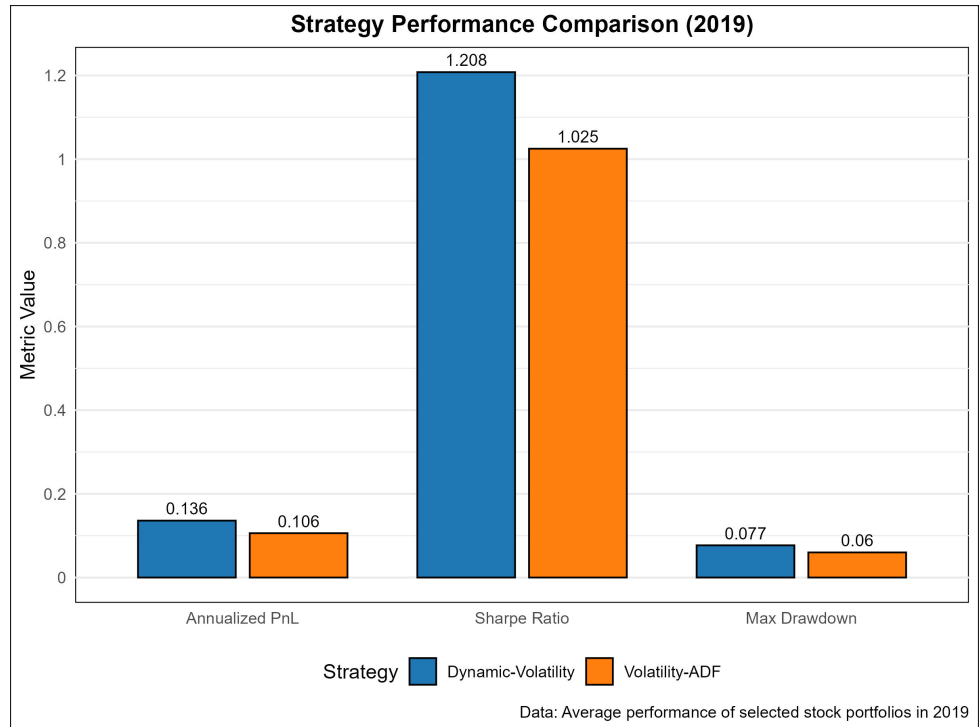


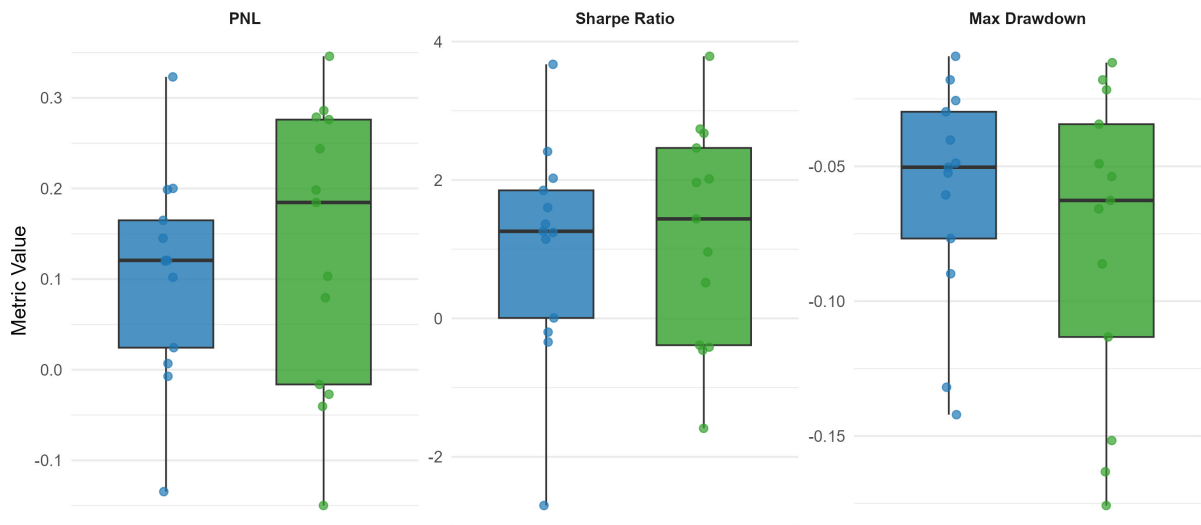
Figure 9. Strategy performance comparison.

Table 7 reveals distinct performance characteristics: the hybrid arbitrage strategy demonstrates superior stability and higher success rates (84.6%), while the dynamic volatility strategy yields higher returns, albeit with greater risk exposure.

Performance Comparison: ADF-Volatility vs Dynamic-Volatility Strategies

Evaluation across 13 stock pairs (1-year period)

Strategy: ADF-Volatility (Blue), Dynamic-Volatility (Green)



Note: Points represent individual stock pair performance

Figure 10. Performance between ADF-volatility and dynamic volatility.

Figure 10 provides a systematic comparison of the two optimal strategies through three distinct analytical lenses. All portfolios must be screened for strong cointegration using both the Augmented Dickey-Fuller (ADF) test ($p < 0.1$) and Pearson correlation analysis ($\rho > 0.8$) to ensure statistical robustness.

As a supplement to **Figure 10**, **Table 8** concisely yet comprehensively explains the average performance of different investment strategies.

Table 8. Average performance of representative strategies.

2019-data	PNL	Sharpe Ratio	Max-DD
Stable-Window	0.086	0.742	-0.107
Dynamic-Volatility	0.136	1.208	-0.077
Volatility-ADF	0.106	1.025	-0.060

4.4. Does Spearman Test Be Applied Simultaneously in 3.4.1?

We find that this test can partly decrease the average PNL of the selected stock portfolios and also decrease the max-drawdown. This study employs an inherently conservative strategy, and we intentionally avoid imposing further restrictive conditions that might forfeit potential arbitrage opportunities. For example, while keeping the threshold as 0.2, the average annual PNL decreases 8.5% (contrary to the Dynamic-Volatility).

4.5. Does Regression Model Be Possible in the Stock Portfolios?

Our empirical analysis employs regression frameworks with stable lookback periods to ensure temporal consistency in parameter calibration.

Initially, our research focus was not on stock pair selection, but rather on identifying predictive features from lookback period data that could forecast daily stock returns. Despite trying many possible variables like the return-correlation, the cointegration, and the spread deviation, we do not identify any statistically significant linear relationships in the data. For example, we use the x_1 (Pearson correlation coefficient of the preceding 20-day period) and the x_2 (variance of 20-day historical price spreads).

The model exhibits low predictive capability, with only 2.4% of variance explained (Multiple $R^2 = 0.024$). This suggests the selected predictors capture minimal systematic patterns in the data. The daily return on the corresponding trading day was used as the dependent variable (vector y). To identify universal metrics applicable across all stocks, we conducted regression analyses on randomly selected stock portfolios over random trading days. We have a dataset of 600 observations.

Pearson correlation coefficient shows statistical significance (p -value = 0.008) but its practical effect size is small ($\beta = 0.032$). Meanwhile, price spreads are non-significant (p -value = 0.122), contributing negligibly to the model.

Key limitations are its inadequate fit for practical use (Adjusted $R^2 = 0.018$) and the effect sizes are trivial despite statistical significance.

We find that predicting daily returns using historical lookback period features is empirically unsubstantiated. **Figure 11** demonstrates the regression model’s partial validity with satisfactory goodness-of-fit but reveals residual non-normality defects.

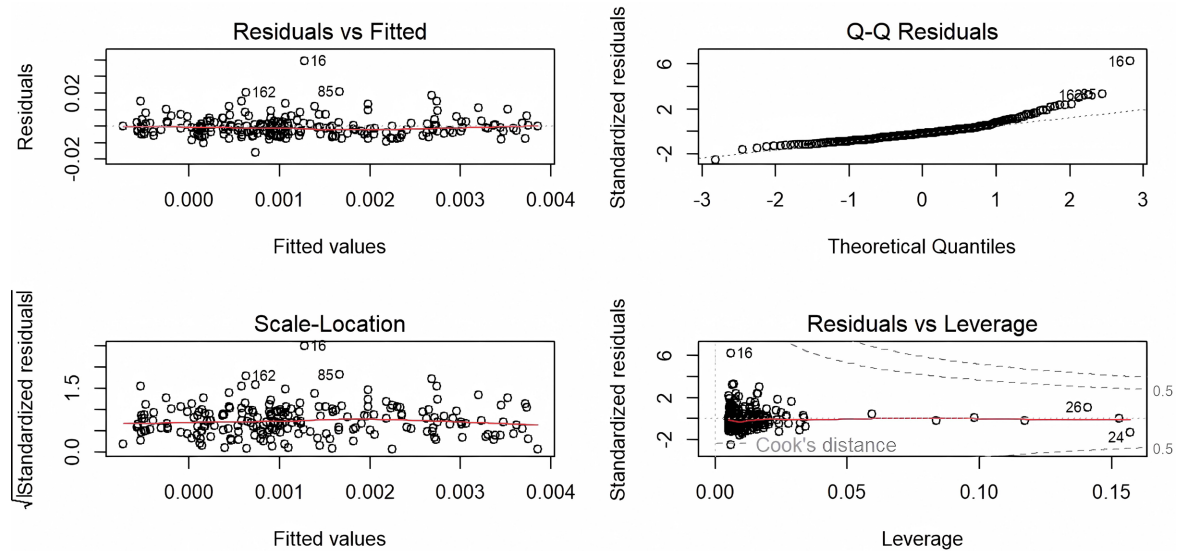


Figure 11. Diagnostic plots for stock-related regression analysis.

Then we turn back to the selected portfolio. We try to build the relation between the data features in 2018 and the PNL of 2019. We perform regression analyses using the ADF test statistic, Pearson correlation coefficient, and the median volatility of stock pair price ratios (all from 2018 data, denoted as x_1 , x_2 , and x_3 , respectively) to explain stock returns. The dependent variables were stable-window returns (y_1) and dynamic-volatility returns (y_2).

y_1 Model explains only 3.34% of the variance ($\text{Adjusted } R^2 = -0.288$) (indicates the model performs worse than random guessing) while y_2 Model explains 3.38% of the variance. The residual standard error is 0.184 in y_1 Model, while the residual standard error remains 0.178 (slightly smaller than y_1 model).

Metric	Intercept	X1	X2	x3	Model Fit
Coefficient	-0.093	0.242	0.601	-0.198	
<i>p</i> -value	0.939	0.866	0.835	0.676	
Statistics					$R^2 = 0.034$

Model y_1

Metric	Intercept	X1	X2	x3	Model Fit
Coefficient	-0.390	0.742	-0.871	-0.265	
<i>p</i> -value	0.735	0.590	0.751	0.556	
Statistics					$R^2 = 0.038$

Model y_2

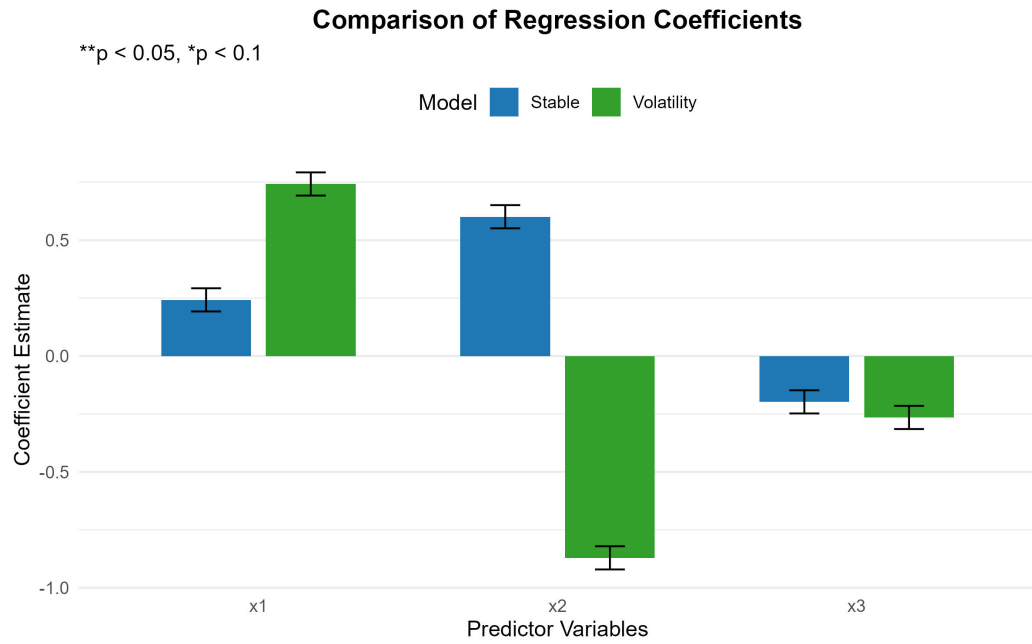


Figure 12. Comparison of regression coefficients.

Figure 12 reveals suboptimal model fit across both scenarios, as evidenced by poor goodness-of-fit metrics. We conclude both models demonstrate very weak explanatory power, though the y2 model exhibits marginally smaller residual variability.

Market features like ADF or correlation cannot predict daily returns, suggesting inherent limits to linear forecasting in HFPT. Such indicators only capture linear relationships at a fixed point in time, ignoring the dynamic, nonlinear interactions that dominate high-frequency price formation. The failure of these features to forecast returns underscores a deeper limitation: linear models cannot account for the complex microstructure effects—such as order flow imbalances, latency arbitrage, and fleeting liquidity—that govern short-term price action.

5. Conclusion

We consider proposing effective methodologies for stock portfolio selection and trading strategy application, and further implement the strategies in China’s A-share market.

During this process, we select stocks from the same GICS industry (code: 2010) for analysis, and further select stocks with complete data from 2018 to 2019. We incorporate high transaction costs (0.4%) into consideration. Additionally, we maintain a 20% reserve margin to mitigate the risk of margin calls. Upon verification, none of the selected stocks experience corporate actions (such as stock splits or bonus issues) during the 2018-2019 period.

The next step involves identifying stock portfolios with stable cointegration relationships. Considering the threshold (mainly depending on Z-score), we use the 2018 data as the benchmark for preliminary stock screening. Based on the daily

closing prices of all potential stock pairs, we select those combinations that met the following criteria: ADF test (price ratio) p -value < 0.1 and Pearson correlation coefficient > 0.8 . After that step, we reduce the number of stock pairs from 231 to 13 through our selection process and finish the selection of portfolios.

Our whole strategy follows the Rolling Window Backtesting. We establish the test baseline using a stable 20-day fixed window. The model incorporates minute-level mean and standard deviation of lookback-period price spreads as predetermined inputs. A Z-score threshold of 2 is adopted to accommodate the high-volatility nature of China's A-share market. The strategy triggers trades when the intraday price spread reaches the threshold ($\text{mean} \pm Z^* \sigma$), following mean reversion principles. Trading continues until market close on that day (We implement a T + 0 trading strategy using existing inventory positions).

We build the dynamic window based on two directions. The dynamic cointegration test (ADF) and the dynamic Volatility test (adjusting the lookback window). The first strategy can decrease the to 5% average annualized Max-Drawdown averagely while decreasing the PNL partly according to the different threshold (ranging from 0.2 to 0.4). The second strategy can generate an additional 5.5% average annualized profit compared to the fixed-window approach (stable window 8%).

We consider combining the two strategies (threshold = 0.5) and show the results in 4.3. This strategy has an average improvement of 1.7% over the Dynamic-Volatility strategy, with a minimum enhancement of 3.5%. Compared with the Dynamic-Volatility strategy, the new strategy has a 2.9% reduction in PNL instead.

For capital-abundant investors, the dynamic volatility strategy demonstrates superior arbitrage opportunity capture and higher profitability, albeit with relatively elevated maximum drawdowns and risk exposure. For small-to-medium investors, the hybrid-window approach demonstrates an 84% probability of positive returns while reducing maximum drawdown, albeit with marginally lower profits, making it a more suitable strategy for this investor segment. Among these two strategies, the Sharpe ratio remains nearly the same (> 1.0), which indicates that both stock strategies demonstrate relatively strong performance.

Aimed at ensuring generalizability, we employ a regression-based modeling approach for our strategy implementation, utilizing randomly selected stock portfolios in our analysis. However, given the policy-induced volatility inherent in equity markets, constructing a robust linear model to predict potential returns presents significant theoretical and practical challenges. We turn to the selected portfolio and change the strategy. Despite the slight improvement of the dynamic strategy over the stable one, the best model explains only 3.38% of the variance.

Building upon the current research, the following directions can be further explored to expand theoretical boundaries and enhance practical applications:

1) Enhanced Integration of Machine Learning Models: Utilize Global Industry Classification Standard (GICS) sectors as the fundamental analytical unit and de-

velop separate model architectures for each sector to account for industry-specific dynamics

2) Adapt methodologies to other GICS sectors (e.g., finance, consumer staples): Rigorously assess the robustness of trading strategies

3) Simulating intraday liquidity shocks: Replicating Historical Liquidity Crises for Systemic Risk Early Warning and Stress-Testing Quantitative Strategies Under Extreme Conditions.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Alexander, S., Coleman, T. F., & Li, Y. (2006). Minimizing CVaR and VaR for a Portfolio of Derivatives. *Journal of Banking & Finance*, *30*, 583-605.
<https://doi.org/10.1016/j.jbankfin.2005.04.012>
- Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association*, *96*, 42-55. <https://doi.org/10.1198/016214501750332965>
- Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Maxing Out: Stocks as Lotteries and the Cross-Section of Expected Returns. *Journal of Financial Economics*, *99*, 427-446.
<https://doi.org/10.1016/j.jfineco.2010.08.014>
- Bouchaud, J., Bonart, J., Donier, J., & Gould, M. (2018). *Trades, Quotes and Prices: Financial Markets under the Microscope*. Cambridge University Press.
<https://doi.org/10.1017/9781316659335>
- Chen, X., Yang, X., & Zeng, Z. (2012). Pairs Trading Strategies in China's Stock Market. *Pacific-Basin Finance Journal*, *20*, 742-756.
- Chen, Y., Jiang, W., & Tong, G. (2019). Transaction Costs and Market Efficiency: Evidence from China's Stamp Tax Adjustments. *Journal of Financial Economics*, *134*, 553-580.
- Chen, Y., Li, X., & Wang, J. (2020). Industry Concentration and Pairs Trading Profitability. *Journal of Financial Economics*, *138*, 368-389.
- Chen, Z., Rhee, S. G., & Veeraraghavan, M. (2015). Short-Sales Constraints and Price Discovery: Evidence from the Hong Kong SAR Market. *Journal of Banking & Finance*, *59*, 80-95.
- Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs Trading: Performance of a Relative-Value Arbitrage Rule. *Review of Financial Studies*, *19*, 797-827.
<https://doi.org/10.1093/rfs/hhj020>
- Gromb, D., & Vayanos, D. (2010). Limits of Arbitrage: The State of the Theory. *Annual Review of Financial Economics*, *2*, 251-275.
<https://doi.org/10.1146/annurev-financial-073009-104107>
- Harvey, C. R., Liu, Y., & Zhu, H. (2016). ... and the Cross-Section of Expected Returns. *Review of Financial Studies*, *29*, 5-68. <https://doi.org/10.1093/rfs/hhv059>
- Hautsch, N., & Voigt, S. (2019). Large-Scale Portfolio Allocation under Transaction Costs and Model Uncertainty. *Journal of Econometrics*, *212*, 221-240.
<https://doi.org/10.1016/j.jeconom.2019.04.028>
- Kang, X., Liu, J., & Wang, X. (2020). Machine Learning-Enhanced Pairs Trading Strategies in China's Stock Market. *Journal of Empirical Finance*, *58*, 1-18.

-
- Kirilenko, A. A., & Lo, A. W. (2013). Moore's Law versus Murphy's Law: Algorithmic Trading and Its Discontents. *Journal of Economic Perspectives*, 27, 51-72.
<https://doi.org/10.1257/jep.27.2.51>
- Li, Y., Zhang, Y., & Wang, Z. (2021). Mean-Reversion Trading Strategies during the COVID-19 Pandemic: Evidence from the Chinese Stock Market. *Finance Research Letters*, 42, Article ID: 101926.
- Lin, Z., Zhang, Y., & Zhao, R. (2021). High-Frequency Trading in China: Costs and Strategy Performance. *Review of Financial Studies*, 34, 3925-3965.
- Liu, X., & Zhang, Y. (2014). Industry-Based Pairs Trading Strategy in China's A-Share Market. *Emerging Markets Review*, 19, 1-15.
- Vidyamurthy, G. (2004). *Pairs Trading: Quantitative Methods and Analysis*. Wiley.