Original Research Article

Evaluating Dynamic Investment Scaling in Pairs Trading Across U.S. Equity Sectors

Arnav Rajadhyaksha

Glenelg High School, 14025 Burntwoods Road, Glenelg, MD 21737, United States

ABSTRACT

This study evaluates whether dynamic investment scaling based on spread magnitude enhances the performance of pairs trading strategies. Using daily closing prices of U.S. equities across eleven sectors from 2021 to March 2026, pairs are identified through a two-stage selection process combining zero-crossing frequency and the Augmented Dickey–Fuller (ADF) tests applied to rolling z-score–normalized spreads. Pairs are formed over a 12-month formation period and evaluated over the subsequent 3-month trading windows. Trading rules incorporate threshold-based entry, exit, and stop-loss conditions, while position sizes are adjusted dynamically using a parameterized scaling function (k-values) as spreads change. Sensitivity analysis suggests that intermediate entry thresholds ($z \approx 1.25$) balance trade frequency and signal quality, while wider stop-loss thresholds ($z \approx 3.5$) help mitigate extreme losses. Results show that no single scaling parameter consistently maximizes returns across sectors or time periods. Regression analysis indicates that scaling parameters are not statistically significant predictors of returns, whereas ADF statistics and selected sector indicators exhibit significance. However, the explanatory power remains limited (adjusted $R^2 \approx 0.03$), consistent with the noisy nature of financial returns. In contrast, scaling parameters are significantly associated with maximum drawdown, suggesting an effect on downside risk rather than expected returns. Overall, dynamic investment scaling does not materially improve returns but can reduce drawdowns. Strategy performance is more strongly driven by pair selection characteristics, particularly spread stationarity, and sector-specific factors. These results should be interpreted with caution given simplifying assumptions such as excluding transaction costs and relatively limited out-of-sample evaluation data.

Keywords: Pairs Trading; Mean Reversion; Investment Scaling; Statistical Arbitrage; Augmented Dickey–Fuller test; Zero-Crossings; Z-score normalization

INTRODUCTION

Pairs trading is a statistical arbitrage strategy that seeks to exploit temporary deviations in the relative prices of historically related securities. It aims to capture mean-reverting behavior between asset pairs while minimizing exposure to broader market movements (1). Despite its widespread use, several implementation decisions remain critical to strategy performance, including pair selection, trade timing, and capital allocation.

While prior research has extensively examined pair selection techniques and signal generation, relatively less attention has been given to how much capital should be allocated dynamically as spreads evolve over time. Most implementations rely on fixed position sizing, regardless

Corresponding author: Arnav Rajadhyaksha, E-mail: arnav.rajadhyaksha09@gmail.com.

Accepted May 11, 2026
https://doi.org/10.70251/HYJR2348.4398106

of the magnitude of mispricing. This raises an important question: can adjusting investment size as a function of spread magnitude improve performance or risk characteristics?

This study addresses this gap by evaluating a dynamic investment scaling framework within a sector-based pairs trading strategy. Using U.S. equity data across eleven sectors, the analysis examines whether scaling positions based on spread magnitude enhances returns or improves downside risk measures such as maximum drawdown. The contribution of this research is twofold. First, it integrates a two-stage pair selection methodology based on zero-crossing frequency and Augmented Dickey–Fuller (ADF) statistics. Second, it systematically evaluates the role of dynamic position sizing across sectors and time periods. The findings provide insights into whether investment scaling meaningfully contributes to performance, or whether outcomes are primarily driven by pair selection and market characteristics.

LITERATURE REVIEW

The theoretical foundation of pairs trading is closely linked to market efficiency and mean reversion. Fama discusses how markets are informationally efficient with prices reflecting all available information (2). However, Malkiel argues that while markets are largely efficient, anomalies do exist, implying that statistical arbitrage opportunities may arise if such mispricing is identified prior to reversion (3). Similarly, Lo suggests that market efficiency is not static but evolves over time (3).

Gatev *et al*. provide a comprehensive analysis of pairs trading, demonstrating that distance-based strategies can generate persistent excess returns by exploiting mean reversion in relative prices (1). Using a 40-year sample, they find that the top-ranked pairs produce average returns of approximately 1.4% per month, corresponding to about 18% annually, even after accounting for transaction costs. They further implement a formation period of twelve months followed by a six-month trading period as their preferred design for strategy execution. Additionally, they also determined that pairs trading is not affected by certain known risks, including size, value, market beta, momentum, and short-term reversal (1). Furthering the point that pairs trading is a robust procedure to use, they found that their methods work when the pairs are actually similar, as when compared to stocks that were randomly selected. Their paper discovered that pairs trading works for pairs across all industries and can be used as long as there are two stocks whose movement mirrors each other. Finally, they noticed that once proper pairs were found, losses were statistically insignificant.

Subsequent research has explored variations in implementation and robustness. Krauss examines statistical arbitrage strategies in high-frequency settings and highlights the sensitivity of profitability to transaction costs underscoring the importance of realistic assumptions when evaluating trading strategies (5). A critical component of pairs trading is the identification of suitable pairs. Chan assesses multiple statistical techniques (Augmented Dickey-Fuller (ADF) which is used here, Hurst exponent, Variance ratio, and Johansen test) to detect mean reversion or stationarity (6). He also emphasizes the importance of robust backtesting and iterative model refinement. However, Bailey *et al*. caution that extensive parameter tuning can lead to overfitting, potentially overstating strategy performance (7).

While pair selection and signal generation have been extensively studied, entry and exit rules have received comparatively less attention. Gatev *et al*. propose entering trades when the distance between normalized prices exceeds approximately two standard deviations and exiting trades when there is mean reversion (spread crosses zero) (1). Another dimension that has received limited focus is position sizing. Kelly highlights the importance of optimal bet sizing for long-term growth (8), while Frazzini and Pedersen demonstrate that risk-adjusted strategies can outperform naive allocations (9). Despite this, most pairs trading implementations assume fixed investment levels, regardless of spread magnitude. Finally, competing evidence suggests potential risks to mean-reversion strategies. Momentum effects documented by Jegadeesh and Titman (10) and Moskowitz *et al*. (11) indicate that price trends may persist, potentially leading to prolonged deviations and losses in pairs trading strategies.

Taken together, the literature highlights that while pairs trading is well established, the role of dynamic investment scaling remains underexplored. This study builds on existing research by systematically examining how position sizing, as a function of spread magnitude, influences both returns and downside risk.

METHODS AND MATERIALS

Pairs are defined as two assets whose prices move in tandem over a period of time. The overall approach to evaluating a pairs trading strategy entails identifying pairs using statistical techniques, generating trading signals, executing the proposed strategy (longing the undervalued asset and shorting the overvalued asset), and assessing performance. This study introduces an enhanced pair selection methodology along with dynamic investment scaling based on spread magnitude.

Data Collection and Processing

Daily closing prices for selected U.S. equities across eleven sectors were collected from October 2021 to March 2026. Observations with missing values were removed to ensure alignment of trading days. Data from 2021 were used solely to initialize rolling calculations.

Within each sector, the top 9–10 companies were selected as shown in Table 1, and all possible pairwise combinations were generated. For example, ten InfoTech stocks resulted in 45 pairs. Similarly, all pairs were assessed in each of the eleven sectors.

Table 1. *List of various tickers within eleven different sectors to then be used to identify top pairs.*

Sectors	Companies
InfoTech	AAPL, META, NVDA, MSFT, GOOG, AMZN, ORCL, IBM, CSCO, INTC
Communication	T, VZ, TMUS, CMCSA, DIS, NFLX, CHTR, DISH, FOXA, IPG
Consumer Cyclical	HD, LOW, MCD, NKE, SBUX, TJX, ROST, KMX, BBY, DG
Financial	JPM, BAC, WFC, C, GS, MS, AXP, PNC, USB, BRK.B, V, MA
Healthcare	JNJ, PFE, MRK, ABBV, TMO, ABT, DHR, BMY, LLY, MDT
Utilities	DUK, SO, NEE, AEP, EXC, SRE, PEG, ED, XEL, ES
Energy	XOM, CVX, COP, SLB, EOG, PSX, VLO, MPC, KMI, OXY
Consumer Staples	PG, KO, PEP, WMT, COST, MDLZ, CL, TGT, CVS, GIS
Industrials	UNP, BA, CAT, LMT, MMM, GE, HON, DE, UPS, FDX
Materials	LIN, APD, SHW, ECL, NEM, DD, FCX, MLM, VMC, PPG
Real estate	AMT, PLD, CCI, EQIX, PSA, SPG, DLR, AVB, EQR, VTR

Prices were normalized for each stock i using rolling z-scores computed over a 60-day window:

$$Z_{i,t} = \frac{P_{i,t} - \mu_{i,t}}{\sigma_{i,t}}$$

where $\mu_{i,t}$ and $\sigma_{i,t}$ represent the rolling mean and standard deviation. The spread between two assets was defined as the difference between their normalized prices and represents relative mispricing between the assets.

Figure 1 shows the closing prices for MSFT and ORCL, while Figure 2 shows the normalized prices for MSFT and ORCL and Figure 3 shows the spread between the normalized prices.

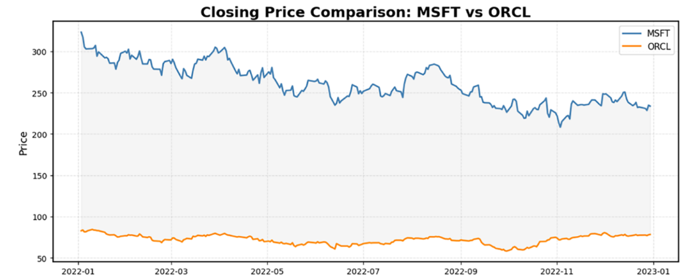

Figure 1. *Closing Prices of MSFT and ORCL from Jan 2022 through Dec 2022.*

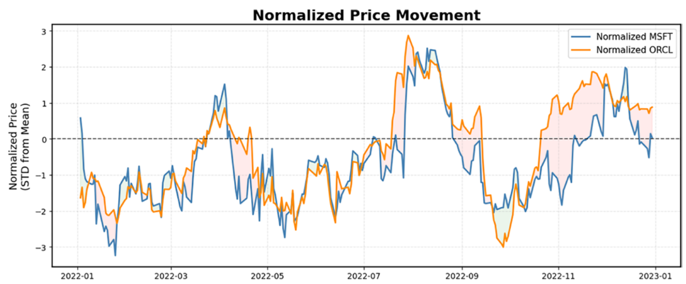

Figure 2. *Normalized Prices of MSFT and ORCL using rolling mean and standard deviation of prior 60 day calculated like a z-score.*

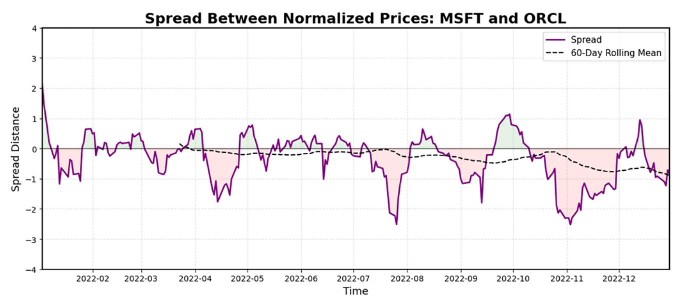

Figure 3. *Spread between Normalized Prices of MSFT and ORCL. Zero-crossing is computed by how many times the spread crosses zero during the evaluation window and calculated to be 30+ times in 2022 for this pair.*

Pair Selection Criteria

The pair selection methodology employs a two-stage filtering approach combining zero-crossing frequency and ADF statistics. This reflects a common practice in the pairs trading literature, where heuristic measures of mean reversion are used alongside formal statistical tests to identify candidate pairs (4, 6). The use of zero-crossings captures the frequency of mean-reverting behavior, while the ADF test assesses statistical stationarity.

First, zero-crossing analysis was conducted by computing the number of times the spread crossed zero over a 12-month period. A higher number of zero-crossings indicates stronger mean-reverting behavior. For example, as illustrated in Figure 3, the spread between MSFT and ORCL crossed zero more than 30 times in 2022.

Second, the ADF test was applied to the spread series to assess stationarity. More negative ADF statistics indicate stronger evidence of stationarity and, by extension, mean reversion. For instance, the ADF statistic for the MSFT–ORCL pair in 2022 was −4.327, suggesting a relatively strong mean-reverting relationship.

Within each sector, the top ten pairs based on zero-crossings were selected, and from these, the top seven pairs with the most negative ADF statistics were retained. These pairs were then evaluated in the subsequent 3-month trading period. InfoTech selections of top seven pairs over two years (2022 and 2023) are presented in Table 2. Similar assessment was done across all the eleven sectors of the economy.

To assess relevance of threshold selection of ten for zero-crossings and seven for ADF statistics, sensitivity analysis was conducted by varying the number of selected pairs at each stage. The results remain qualitatively consistent across alternative specifications, suggesting that the findings are robust to the choice of selection parameters. Picking a higher number for threshold implies higher computing and not missing out on pairs with lower statistics that may have better returns.

While the thresholds used in this study are empirically motivated, alternative selection frameworks such as continuous ranking or single-stage optimization approaches (e.g., power-statistic methods) may provide a more systematic approach and represent a direction for future research.

Trading Strategy

Once the pairs are identified, the trading strategy consists of taking positions when the spread between the normalized prices of the two stocks deviates from zero: shorting the relatively overvalued stock and taking a long position in the relatively undervalued stock.

However, the spread between normalized prices fluctuates over time, giving rise to multiple considerations regarding trade execution, including the timing of entry and exit, as well as the allocation of capital for each position. If a trade is executed every time the normalized spread crosses zero momentarily, it can result in too many trades. Similarly, holding onto the pairs if the normalized spread widens and staying for an extended period can result in high losses and sizable drawdowns. To address these issues, threshold-based trading rules are introduced. Specifically, trades are initiated only when the normalized spread exceeds a predefined entry threshold, and positions are closed either upon mean reversion or when the spread reaches an extreme level. The latter serves as a stop-loss mechanism, designed to limit losses in the presence of structural breaks or deviations from the assumed mean-reverting behavior.

This approach is consistent with the methodology of Gatev *et al.* who propose entering trades when the distance between normalized prices exceeds approximately two standard deviations (1). While their

Table 2. *For InfoTech sector in 2022 and 2023, the top pairs that were in the top ten Zero-crossings and top seven highest absolute ADF statistics.*

Year	Best Pair	# of Zero-crossings	Absolute ADF Stat
2022	(‘MSFT’, ‘ORCL’)	37	-4.33
2022	(‘META’, ‘GOOG’)	30	-3.45
2022	(‘MSFT’, ‘GOOG’)	32	-3.44
2022	(‘NVDA’, ‘MSFT’)	38	-3.31
2022	(‘GOOG’, ‘AMZN’)	35	-2.96
2022	(‘AAPL’, ‘AMZN’)	29	-2.89
2022	(‘NVDA’, ‘GOOG’)	36	-2.69
2023	(‘NVDA’, ‘ORCL’)	33	-4.60
2023	(‘META’, ‘NVDA’)	36	-4.20
2023	(‘NVDA’, ‘AMZN’)	31	-4.17
2023	(‘AMZN’, ‘ORCL’)	37	-3.95
2023	(‘META’, ‘MSFT’)	30	-3.59
2023	(‘META’, ‘ORCL’)	31	-3.55
2023	(‘AAPL’, ‘MSFT’)	34	-3.43

framework is based on price distance rather than a z-score formulation, the underlying intuition is similar: large deviations are more likely to represent temporary mispricing and thus exhibit higher probabilities of mean reversion. They further observe that higher thresholds (e.g., three standard deviations) reduce trading frequency but increase average profitability per trade, whereas lower thresholds increase trading activity and potential for total profit at the cost of greater noise and transaction expenses. Their primary exit rule is based on convergence rather than an explicit stop-loss.

Building on this framework and given limited literature identified around this, a grid-based sensitivity analysis was implemented using z-score–normalized spreads within the Information Technology sector. Table 3 shows the combinations of entry and stop-loss thresholds that were evaluated based on risk-adjusted performance, measured as return divided by standard deviation. Lower entry thresholds (z = 0.5) had worse results. The results also indicate that entry thresholds above 1.25 lead to a substantial reduction in trading opportunities, limiting overall profitability. Consequently, an entry threshold of z = 1.25 was identified as a balanced choice. A stop-loss threshold in the range of z = 3 to 4 is found to effectively limit downside risk without prematurely terminating positions during normal spread fluctuations. Positions are exit upon mean reversion (z ≈ 0), consistent with standard practice.

Table 3. *Return-to-standard-deviation ratio (risk-adjusted return) for the Information Technology sector under varying entry (z) and stop-loss (z) thresholds, based on z-score–normalized spreads.*

Stop-Loss (z)	Entry = 0.5	Entry = 1.0	Entry = 1.25
3.0	-36%	-8%	8%
3.5	-34%	-6%	9%
4.0	-33%	-6%	9%

The results highlight the importance of threshold selection in balancing trade frequency and signal quality. Lower entry thresholds generate excessive trades with poor risk-adjusted returns, while excessively high thresholds reduce trading opportunities. The chosen parameters reflect a compromise between these competing effects and are supported by robustness across adjacent parameter values. Other entry and stop-loss thresholds can be evaluated in future works of this strategy.

Dynamic Investment Scaling

The study introduces a parameter k to control how investment size changes with spread magnitude, as illustrated in Figure 4. Investment is scaled from 1 to 0 as the spread increases from 1.25 to 3.5. When k = 1, the reduction in investment is linear. For k > 1, the reduction follows a nonlinear (convex) pattern, allowing the strategy to maintain higher exposure at lower spread levels before decreasing more sharply as the spread approaches the stop-loss threshold.

Positions were rebalanced as spreads evolved. Transaction costs were assumed to be zero to isolate the impact of the strategy. Transaction costs can be considered for an evolved approach of this strategy. The assumption of zero transaction costs is a simplifying assumption and represents a limitation, particularly for strategies involving frequent position adjustments.

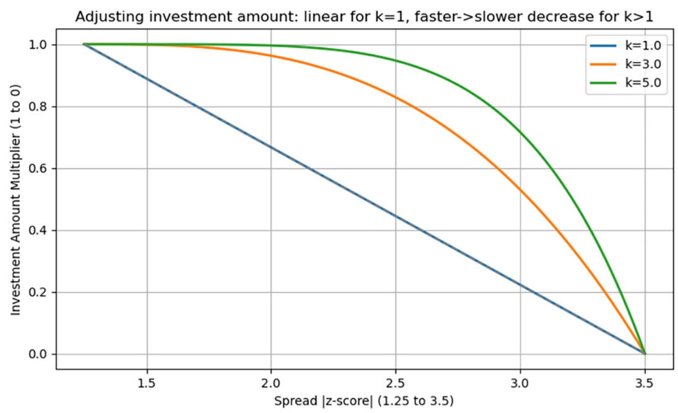

Figure 4. *Varying investment amount factor calculated for different k-values as normalized spread increases from 1.25 (entry) to 3.5 (stop-loss). k =1 corresponds to a linear drop in investment while k = 5 represents a gradual scaling of investment as spread increases from 1.25 to 3.5.*

Trading strategies are assessed for all the top pairs identified in the subsequent quarter for varying values of k (values of 1, 3, 5 were evaluated). For example, MSFT and ORCL are the top pairs from 2022 and that pair is assessed for Quarter 1 of 2023. Figure 5 demonstrates how the strategy enters and exits the trades for MSFT and ORCL trades in Quarter 1 of 2023.

Performance Metrics

Each Strategy is evaluated for key outcomes such as Returns (profit over initial capital), Sharpe Ratio, Maximum Drawdown, Win Rate, Number of trades. Sharpe Ratio is calculated as excess return divided

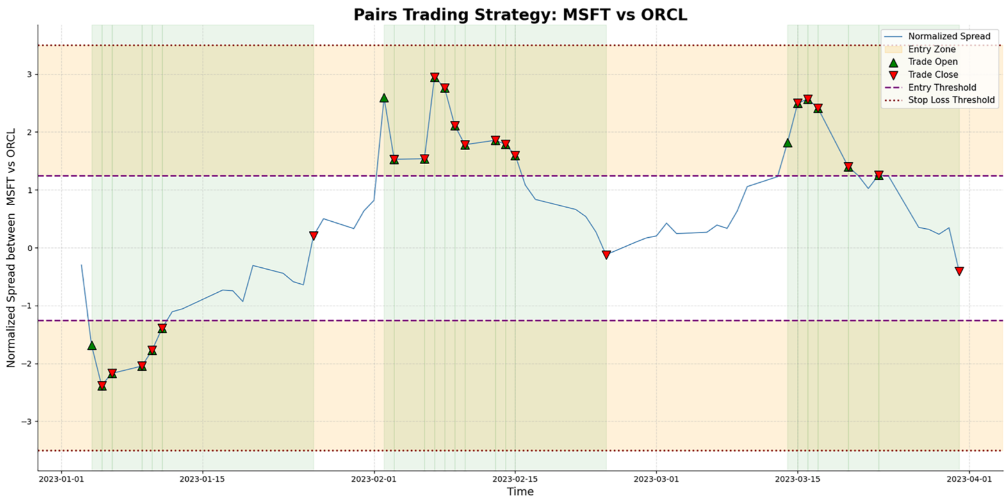

Figure 5. *Executing trading strategy in Q1 2023 for top pair (MSFT & ORCL) identified from 2022 spreads. Trade is entered when absolute value of spread is higher than 1.25 and scaled, while trade is exit when spread crosses 0.*

by standard deviation, while Win Rate is defined as percentage of winning trades over all executed trades and Maximum Drawdown measures the largest drop in the portfolio from its peak to trough. Additionally, linear regression was performed to assess the relationship between returns and explanatory variables including ADF statistic, number of zero-crossings, k-value, and sector.

RESULTS

Performance Across k-values

Across all sectors and testing periods, no single k-value consistently produced the highest returns. Both low and high k-values achieved the highest performance depending on the specific pair and time period.

Extreme k-values (k = 1 and k = 5) most frequently corresponded to the higher returns, while intermediate values were less frequently optimal. For example, k-value of 1 indicates higher portfolio return for MSFT-ORCL in Q1-2023 while lower negative returns (i.e. higher returns) are seen for NVDA-GOOG for k-value of 1 as seen in Table 4. Similar inconsistencies are seen across other sectors as well.

Using an integrated dataset across the various sectors for the top seven pairs for 5 years of data, initial linear regression was performed with return (Net change in balance over starting balance) as the dependent variable and k-values, ADF statistic, number of zero-crossings, various sectors as the independent variable. To avoid

Table 4. *InfoTech Sector performance in Q1-2023 of two pairs (MSFT-ORCL and NVDA-GOOG) identified from 2022 for different k-values indicates that certain k-values are not indicative of higher returns. k=5 has higher returns for MSFT-ORCL while k=1 has higher returns for NVDA-GOOG.*

Sector	*Trading Period*	*Ticker 1*	*Ticker 2*	*ADF in prior 12 months*	*# 0-crossing (prior 12)*	*k-value*	*Portfolio Return*	*Sharpe Ratio*	*Max Drawdown*	*Win Rate*
infotech	*Q1-23*	*MSFT*	*ORCL*	*(4.33)*	*37*	*1*	*3.2%*	*7.0%*	*-5.2%*	*40.9%*
infotech	*Q1-23*	*MSFT*	*ORCL*	*(4.33)*	*37*	*3*	*4.2%*	*7.6%*	*-6.1%*	*40.9%*
infotech	*Q1-23*	*MSFT*	*ORCL*	*(4.33)*	*37*	*5*	*4.7%*	*8.2%*	*-6.3%*	*40.9%*
infotech	*Q1-23*	*NVDA*	*GOOG*	*(2.69)*	*36*	*1*	*-7.4%*	*-11.0%*	*-16.3%*	*39.1%*
infotech	*Q1-23*	*NVDA*	*GOOG*	*(2.69)*	*36*	3	*-11.8%*	*-15.0%*	*-21.5%*	*39.1%*
infotech	*Q1-23*	*NVDA*	*GOOG*	*(2.69)*	*36*	5	*-12.3%*	*-15.1%*	*-22.1%*	*39.1%*

overfitting, variables with consistently weak explanatory power across specifications were excluded. Variables that indicated limited effect were k-value, number of zero-crossings and sectors such as Real Estate, Utilities, Financial, Consumer Cyclical, Consumer Staples, Energy, Materials.

Excluding the variables with limited effect, a revised linear regression analysis was performed with a 75% train and 25% test (out of sample) dataset.

Table 5 reports the results of an Ordinary Least Squares regression examining the relationship between returns and selected explanatory variables, including the ADF statistic and selective sector indicators.

Table 5. *OLS Regression Results (Dependent Variable: Returns) indicates ADF statistic, Healthcare and Information Technology sectors are significant at <5% level. Notes: Standard errors in parentheses. ***, ** denote significance at 5% levels, respectively.*

Variables	Coefficient	(Std. Error)
Intercept	-0.0165	(0.0240)
ADF Statistic	-0.0053**	(0.0023)
Healthcare Sector	-0.0406**	(0.0132)
Information Technology	0.0339**	(0.0144)
R^2	0.035	
Adjusted R^2	0.029	
F-statistic	6.078	
Prob (F-statistic)	0.0005	

The regression results indicate that the model is jointly significant (F-statistic $p < 0.01$), although the explanatory power is relatively modest, with an adjusted R^2 of 0.029. Low adjusted R-squares are common in financial return data, where outcomes are influenced by a large number of unobserved or noisy factors such as liquidity conditions, short-term market microstructure effects, news events, and execution timing. As a result, even economically meaningful relationships can produce low explanatory power in a statistical sense. Fama asserts that returns are largely unpredictable from publicly available variables and low explanatory power of regression is expected in efficient markets (2).

The ADF statistics, which are negative in value and are intended to capture mean-reversion characteristics, has a negative coefficient indicating that higher absolute value of ADF statistic indicates more returns and is significant at 5% level. Sectoral effects are also statistically significant. The coefficient for the healthcare sector is negative and significant at the 5% level, indicating that, holding other factors constant, pairs associated with this sector tend to generate lower returns. In contrast, the information technology sector exhibits a positive and statistically significant coefficient at the 5% level, suggesting relatively higher returns compared to the omitted baseline sector.

The intercept term is not statistically significant, indicating no baseline return once explanatory variables are accounted for. The out of sample test population had a positive but slightly lower R^2 (0.02) than the train population R^2 (0.035) indicating the model still works but has a slight amount of over-fitting. The lower sample size represents an opportunity to build a better model.

Maximum drawdown is a negative outcome and represents how much the portfolio went negative at its worst point. Linear Regression on the maximum drawdown shows that the maximum drawdown has a negative k-value coefficient that is significant as seen in Table 6. The k-value representing the trading intensity or position scaling parameter, is negative and highly statistically significant. This suggests that higher values of k are associated with reduced maximum drawdown, indicating that gradually reducing position reduces maximum drawdown but not expected returns as seen from the initial regression.

Table 6. *OLS Regression Results (Dependent Variable: Maximum Drawdown) indicates k-value is significant at 1% level along with ADF Statistic and Healthcare & Information Technology sectors. Notes: Standard errors in parentheses. ***, ** denote significance at the 1% and 5% levels, respectively.*

Variables	*Coefficient*	*(Std. Error)*
Intercept	*-7.1557****	*(1.401)*
ADF Statistic	*-0.1258***	*(0.061)*
Healthcare Sector	*-1.6462***	*(0.804)*
Information Technology	*-2.1292***	*(0.842)*
K-value	*-0.4597****	*(0.142)*
R^2	*0.029*	
Adjusted R^2	*0.024*	
F-statistic	*5.178*	
Prob (F-statistic)	*0.0004*	

DISCUSSION

The findings indicate that dynamic investment sizing does not produce a universally optimal strategy across all market conditions. A key observation is the tradeoff between risk and return. Lower k-values (faster drop in investment), tended to have higher max drawdowns. In contrast, higher k-values (gradual drop in investment) maintain greater exposure, allowing for lower maximum drawdowns.

The lack of significance of the k-value in regression results suggests that while position sizing influences the distribution of outcomes, it is not the primary driver of returns. Instead, the quality of pair selection plays a more central role.

The results also highlight the importance of pair selection. The relationship between ADF statistics and performance suggests that stationarity is a critical factor in determining profitability. In contrast, the number of zero-crossings appears to be a weaker indicator of performance, despite being useful in initial filtering.

Sector-level differences indicate that certain industries may exhibit more stable mean-reverting relationships, contributing to more predictable outcomes. This aligns with the view that market efficiency varies across sectors and over time.

These findings are consistent with the adaptive markets hypothesis by Lo, which suggests that the effectiveness of trading strategies varies across market environments (3). They also highlight potential limitations related to overfitting, as noted by Bailey *et al*., given the evaluation of multiple parameter configurations (7).

The study has certain limitations that can be further enhanced in future research on this topic. The approach assumes no transaction costs as the position is modified every day to understand the impact of scaling investment. This is not realistic in a real-world scenario where it's needed to have relatively fewer investment adjustments. Zero transaction cost assumption was adopted to isolate the impact of the proposed investment scaling mechanism and pair selection methodology, independent of market frictions. Given the high turnover associated with dynamic position adjustments, incorporating realistic transaction costs would likely reduce profitability and may render certain strategies infeasible. This limitation is particularly relevant in high-frequency trading contexts, where costs can significantly erode returns (5). The results should therefore be interpreted as an upper bound on achievable performance under frictionless conditions. In practice, the viability of the strategy would depend on execution efficiency, market liquidity, and transaction cost structures. The low R^2 indicates weak explanatory power of the model, though these are somewhat expected in efficient markets. Targeting more features can potentially improve R^2, however, that can result in overfitting and Bailey *et al*. (7) caution against overfitting. Expanding the backtest to multiple years and increasing sample size for the model build can potentially improve the model's explanatory power as well as reduce risk of overfitting. Moreover, the backtesting trading time window of next three months is relatively short and may drive more variability in outcomes. This time window was selected because the predictive power of mean reversion in pairs trading may diminish when the same pair is tested over extended horizons; however, it remains worthwhile to evaluate the strategy across longer testing windows.

CONCLUSION

This study develops and evaluates a sector-based pairs trading framework that combines zero-crossing frequency and Augmented Dickey–Fuller (ADF) statistics for pair selection, alongside a dynamic investment scaling mechanism based on spread magnitude. Using rolling formation and trading windows, the analysis examines whether adjusting position size as a function of spread deviation improves strategy performance across multiple U.S. equity sectors.

The empirical findings indicate that no single investment scaling specification consistently maximizes returns across sectors or time periods. While certain configurations perform well in specific contexts, the absence of a stable relationship between scaling parameters and returns suggests that dynamic position sizing is not a primary driver of return generation. Regression results support this observation, showing that scaling parameters are not statistically significant predictors of returns, whereas measures of spread stationarity (ADF statistics) and sectoral characteristics exhibit greater explanatory relevance. However, the overall explanatory power of the return model remains modest, reflecting the inherently noisy and multifactor nature of financial return processes.

In contrast, dynamic scaling appears to have a more consistent relationship with downside risk. The negative and statistically significant association between the scaling parameter and maximum drawdown suggests that more gradual position adjustment (higher k-values) can mitigate extreme losses. This indicates that while dynamic scaling may not enhance expected returns, it can influence the distribution of outcomes by improving risk characteristics. Taken together, the results suggest that the effectiveness of pairs trading strategies is more strongly determined by the quality of pair selection - particularly the degree of spread stationarity - than by the choice of investment scaling rule. Sectoral differences further contribute to variation in performance, indicating that mean-reverting relationships may not be uniformly present across industries.

The analysis also suggests that entry, exit and stop-loss thresholds are also drivers of return and represent a future topic of research. These findings should be interpreted considering several limitations. The analysis assumes zero transaction costs, which likely overstates achievable performance given the relatively high turnover implied by dynamic scaling. The evaluation of multiple parameter configurations raises the possibility of overfitting. Future research could extend this framework by incorporating realistic transaction costs, testing longer evaluation horizons, and exploring alternative approaches to both pair selection and position sizing, including optimization-based or model-driven methods.

Overall, this study contributes to the literature by providing evidence that, within the context of spread-based pairs trading, dynamic investment scaling plays a secondary role relative to pair selection and sector characteristics. While scaling mechanisms can improve downside risk management, their impact on return generation appears limited under the conditions examined.

FUNDING SOURCES

The author has no funding sources that supported the research and preparation of this article.

CONFLICT OF INTEREST

The author declares that there are no conflicts of interest regarding the publication of this article.

REFERENCES

1. Gatev E, Goetzmann WN & Rouwenhorst KG. Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*. 2006; 19 (3): 797-827. https://doi.org/10.1093/rfs/hhj020
2. Fama EF. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*. 1970; 25 (2): 383-417. https://doi.org/10.2307/2325486, https://doi.org/10.1111/j.1540-6261.1970.tb00518.x
3. Malkiel BG. The efficient market hypothesis and its critics. *Journal of Economic Perspectives*. 2003; 17 (1): 59-82. https://doi.org/10.1257/089533003321164958
4. Lo AW. The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management*. 2004; 30 (5): 15-29. https://doi.org/10.3905/jpm.2004.442611
5. Krauss C. Statistical arbitrage strategies and high-frequency data. *Quantitative Finance*. 2017; 17 (11): 1613-1628.
6. Chan EP. Algorithmic trading: Winning strategies and their rationale. Wiley. 2013. https://doi.org/10.1002/9781118676998
7. Bailey D, Borwein J, López de Prado M & Zhu Q. Probability of backtest overfitting. *Journal of Computational Finance*. 2017; 20 (4): 39-69. https://doi.org/10.21314/JCF.2016.322
8. Kelly JL. A new interpretation of information rate. *Bell System Technical Journal*. 1956; 35 (4): 917-926. https://doi.org/10.1002/j.1538-7305.1956.tb03809.x
9. Frazzini A & Pedersen LH. Betting against beta. *Journal of Financial Economics*. 2014; 111 (1): 1-25. https://doi.org/10.1016/j.jfineco.2013.10.005
10. Jegadeesh N & Titman S. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*. 1993; 48 (1): 65-91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x
11. Moskowitz TJ, Ooi YH & Pedersen LH. Time series momentum. *Journal of Financial Economics*. 2012; 104 (2): 228-250. https://doi.org/10.1016/j.jfineco.2011.11.003