Is Pairs Trading a Thing of the Past?

Jesus Villota Miranda

CEMFI

32nd Finance Forum (July 2, 2025)

Outline

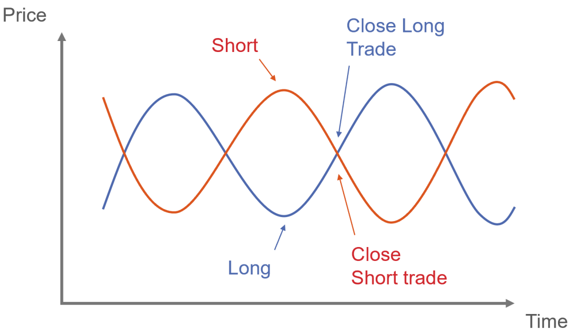
1. Introduction
2. Methodology
3. Empirical Application
4. Results
5. Conclusions

History of Pairs Trading

- Developed by a group of "*quants*" at Morgan Stanley in the 1980s.
  - Pioneered by Gerry Bamberger
  - Advanced by a team of mathematicians and physicists led by Nunzio Tartaglia
- In 1987, Morgan Stanley made $50 mill. on Tartaglia's team's automated systems
- Although the Morgan Stanley group disbanded in 1989 after a period of performance, **pairs trading has since become an increasingly popular "market-neutral" investment strategy** used by hedge funds, individual and institutional traders,.

How the strategy works

1. Search for **two securities** whose **prices** tend to move together.
2. Hypothesize that this relationship will hold over time.
3. When a price **divergence** occurs, simultaneously:
   - LONG the UNDERPERFORMING security
   - SHORT the OUTPERFORMING security
4. When asset prices **converge**, close both positions to realize a profit.



Theoretical Underpinnings: Asset Pricing

Asset Pricing can be viewed in **absolute** and **relative** terms:

- **Absolute pricing** $\Rightarrow$ securities are valued from **fundamentals**, such as disconted future cash flows
- **Relative pricing** $\Rightarrow$ two securities that are **close substitutes** for each other should sell for the **same price** (*it does not say what that price should be!*)
    - The **L**aw of **O**ne **P**rice is applicable to relative pricing (*even if that price is wrong!*)
    - According to Ingersoll (1987), the LOP is the proposition that "*two investments with the same payoff in every state of nature must have the same current value*".
    - In other words, two securities with the **same payoffs** in all states of the world should sell for the **same price**

Theoretical Underpinnings: Co-Integration

- Pairs trading can be justified within an **Equilibrium Asset-Pricing framework with nonstationary common factors**
- If the long and short components fluctuate with common nonstationary factors, then the **prices of the component portfolios** (*aka the spread*) would be **co-integrated** and the pairs trading strategy would be expected to work.
- To interpret a pair (i, j) as co-integrated prices, we need to assume that there exists a **co-integrating vector that has only two nonzero coordinates**:

$$\boldsymbol{\alpha} = \mathbf{e}_i - \mathbf{e}_j$$

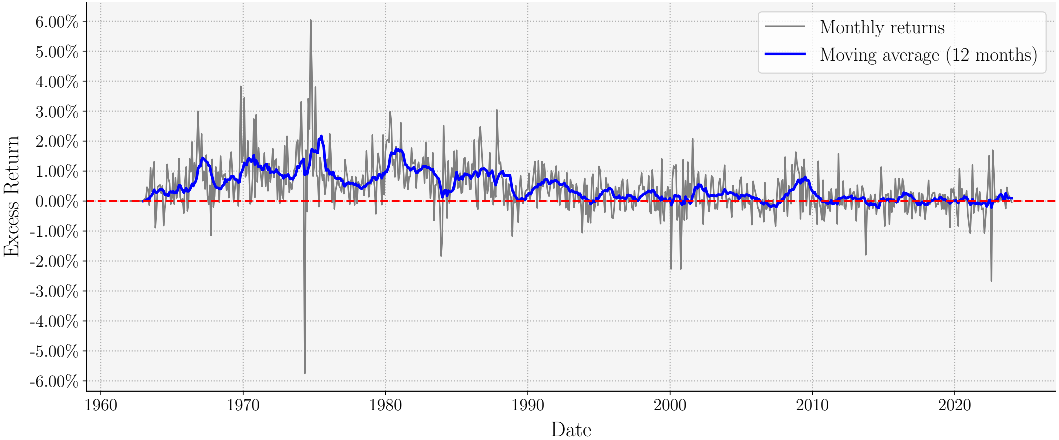
- In that case, the sum or difference of scaled prices will be **reverting to zero** and a trading rule could be constructed to exploit the expected temporary deviations.

Literature Review

- **[Gatev, Goetzmann, & Rouwenhorst (GGR), 2006]**
  - Their paper, *"Pairs Trading: Performance of a Relative-Value Arbitrage Rule"* is the most-cited and seminal work in the literature.
  - First circulated as a working paper in 1999 and later published in the *Review of Financial Studies* (2006).
  - It provided the **first large-scale, rigorous empirical evidence of the strategy's historical profitability**, establishing it as a legitimate field of study.
- **[Do and Faff, 2010]**
  - In their *Financial Analysts Journal* article, they re-examined the GGR strategy with an extended dataset running through 2009.
  - They confirmed that profitability was steadily **decaying**, likely due to the strategy's popularity and increasing market efficiency.
  - However, they also discovered that profits temporarily **resurged** during periods of market turmoil, such as the 2008 Global Financial Crisis (GFC).

The Decay of Pairs Trading

Figure: Monthly excess returns to Pairs Trading



The Decay of Pairs Trading

- The **popularization** of PT has led to a significant **decline in profitability**
- However, this decay **does not invalidate** the underlying principle of **relative-value arbitrage**
- The issue may lie in the **restrictive nature** of traditional **1-to-1 pairing**:
  - Modern markets are **increasingly efficient**
  - Simple, stable relationships between individual stocks have become **scarce**
  - Traditional pairwise constraints may be **too limiting**

Challenge: Identifying more **robust substitutes** in an increasingly efficient market

Recent Literature on Pairs Trading

- The pairs trading literature has **stagnated conceptually** since GGR
- Research focus has shifted toward **methodological refinements** rather than fundamental innovation:
  - Advanced statistical techniques for pair identification
  - Sophisticated modeling of joint stock behavior
  - Complex signal generation and position sizing methods
- This has produced a **fragmented literature** with several concerning patterns:
  - **Incremental improvements** without clear theoretical justification
  - **Method selection bias** - unclear why specific techniques are chosen
  - **Publication bias** - potential cherry-picking of favorable results

Need for: principled evolution of the traditional PT framework

Pairs Trading

Structural Break-Aware

Pairs Trading with Markov Regime-Switching Model

Trading Based on Quantile Forecasting of Smooth Transition GARCH Models

Dynamic Copula Framework for Pairs Trading

Cointegration Approach for the Pair Trading Based on the Kalman Filter

An Intelligent Model for Pairs Trading Using Genetic Algorithms

Pairs Trading with a Mean-Reverting Jump-Diffusion Model on High-Frequency Data

Walk-Forward Optimized Forex Cointegration Strategy

ing: Distance, Cointegration & Copula Methods

Hidden Markov Model for Statistical Arbitrage in

A Pairs Trading Strategy Based on Mixed Copulas

Research on High-Frequency Stock Pair Trading Strategy Based on MS-GARCH Model

Review of Stochastic Differential Equations in Statistical Arbitrage Pairs Trading

Black-Litterman Bayesian n in Statistical Arbitrage

Can we reframe the idea of Pairs Trading?

The Limitation of the Traditional Approach

- Traditional pairs trading imposes a **severe cardinality constraint**:
  - Limited to finding **one single substitute** security for any target asset
  - While conceptually elegant, this restriction may be **unnecessarily limiting** in practice

A More Flexible Approach

- Instead of searching for one perfect substitute, why not construct a **synthetic replica** using multiple securities?
- **Example - General Motors**: Rather than finding a single stock that tracks GM's movements:
  - Use a **linear combination** of multiple securities
  - Create a more accurate synthetic substitute
  - Capture complex relationships that no single stock can replicate

Key Idea: We can generalize the search for substitutes from **1-to-1** to **1-to-portfolio**

Generalizing Pairs Trading: a Replicating Portfolio Approach

Traditional Approach

- **Structure:** Rigid 1-to-1 pairing.
- **Substitute:** A single, similar stock.
- **Spread:** Price(A) - Price(B).
- **Limitation:** Stable, reliable pairs are now scarce.

Replicating Portfolio Approach

- **Structure:** Flexible 1-to-many.
- **Substitute:** A **synthetic replica** (a portfolio).
- **Spread:** Price(Target) - Price(Replica).
- **Advantage:** More robust and flexible replication.

Our approach preserves the same Asset pricing and Cointegration underpinnings as traditional PT

Outline

1. Introduction
2. Methodology
3. Empirical Application
4. Results
5. Conclusions

Pairs Trading as a Replicating Portfolio problem

- The economic intuition behind pairs trading –exploit temporary mispricing between two close substitutes– can be cast more generally as a problem of *relative valuation*.
- Rather than forcing the substitute for a target stock i to be a *single* partner, we allow it to be a *replicating portfolio*: a linear combination of securities whose joint price path mimics that of i.

$$p_{i,t} = \sum_{j \in \mathcal{D}} \beta_j \, p_{j,t} + \varepsilon_{i,t}, \qquad t = 0, \ldots, T, \tag{1}$$

where:

- $p_{i,t}$ represents the *normalized price* (i.e.: cumulative return) of the target asset
- $\mathcal{D}$ denotes the donor pool of potential replicating assets, and
- $\varepsilon_{i,t}$ denotes the spread between the target and its synthetic replica.

Recasting Gatev's Pairs Trading as a Replicating Portfolio problem

Note that traditional pairs trading, as described in Gatev et al, is a special case where the regression coefficient vector is restricted to the canonical basis.

- Let $\mathcal{B} := \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{|\mathcal{D}|}\}$, denote the canonical basis of $\mathbb{R}^{|\mathcal{D}|}$.
- "Traditional" pairs trading fixes a target stock i and selects a single-asset replicating portfolio in a 1-to-1 relationship by imposing the restriction $\beta \in \mathcal{B}$.

$$\min_{\beta \in \mathcal{B}} \sum_{t=0}^{T} \left(p_{i,t} - \sum_{j \in \mathcal{D}} \beta_j p_{j,t} \right)^2$$

- "Traditional" pairs trading fixes a target stock i and selects a single-asset replicating portfolio in a 1-to-1 relationship by imposing the restriction $\beta \in \mathcal{B}$.

$$\min_{\beta \in \mathcal{B}} \sum_{t=0}^{T} \left(p_{i,t} - \sum_{j \in \mathcal{D}} \beta_j p_{j,t} \right)^2$$

- Under this formulation, the only nonzero element of the coefficient vector will correspond to that of the **stock** whose **price** series **minimizes** the **least-squares (or euclidean) distance with respect to that of the target stock**.

$$j^* = \arg\min_{j \in \mathcal{D}} \sum_{t=0}^{T} \left(p_{i,t} - p_{j,t} \right)^2 \tag{2}$$

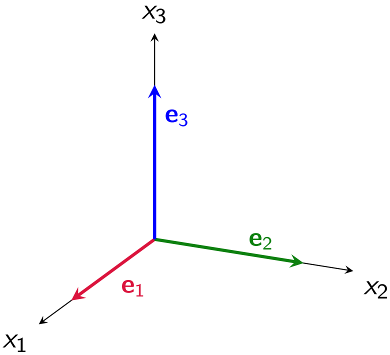
- Note that pairs trading with a hedge ratio is equivalent to restricting $\boldsymbol{\beta}$ to lie in $\mathcal{B}^{+} := \{\alpha \mathbf{e}_j : j \in \mathcal{D}, \alpha \in \mathbb{R}\}$. Hence:

$$\min_{\boldsymbol{\beta} \in \mathcal{B}^{+}} \sum_{t=0}^{T} \left(p_{i,t} - \sum_{j \in \mathcal{D}} \beta_j p_{j,t} \right)^2$$

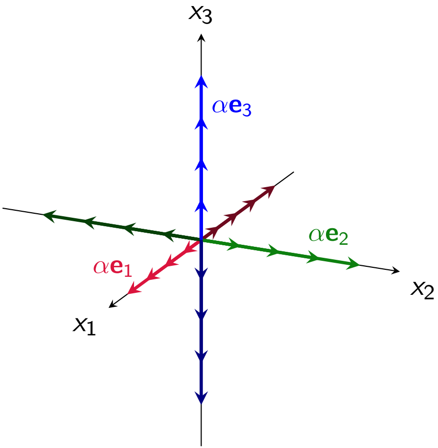
- Which is equivalent to choosing $\boldsymbol{\beta}$ freely but imposing unit cardinality in the coefficient vector

$$\begin{aligned} \min_{\boldsymbol{\beta} \in \mathbb{R}^{|\mathcal{D}|}} \quad & \sum_{t=0}^{T} \left(p_{i,t} - \sum_{j \in \mathcal{D}} \beta_j p_{j,t} \right)^2 \\ \text{s.t.} \quad & \|\boldsymbol{\beta}\|_0 = 1 \end{aligned}$$

Canonical vs. Augmented Canonical Space



$$\mathcal{B} = \{\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3\}$$



$$\mathcal{B}^+ = \{\alpha\, \mathbf{e}_i : \alpha \in \mathbb{R},\ i = 1, 2, 3\}$$

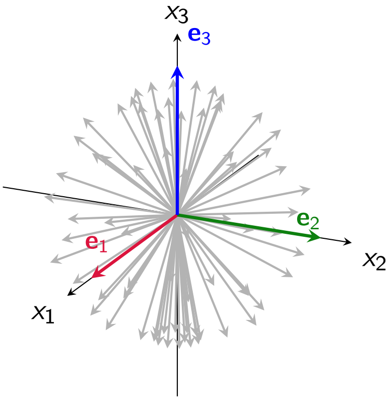
Pairs Trading a Sparse Replicating Portfolio

- We avoid imposing such restrictive cardinality restrictions, and allow for a denser regression coefficient vector. However, we deliberately promote sparsity to avoid excessive transaction costs.

- For this purpose, we employ the LASSO regularization [Tibshirani, 2006]

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^{|\mathcal{D}|}} \left\{ \sum_{t=0}^{T} \left(p_{i,t} - \sum_{j \in \mathcal{D}} \beta_j p_{j,t} \right)^2 + \lambda \sum_{j \in \mathcal{D}} |\beta_j| \right\} \tag{3}$$

Replicating Portfolio Approach: Any Vector in $\mathbb{R}^3$



$$\mathbb{R}^3 = \mathrm{span}(\mathcal{B})$$

Trade-off Between Sparsity and Replication Quality

- **Low λ ($\to$ 0):** Minimal penalization leads to dense portfolios
  - Recovers the unrestricted regression from equation (1)
  - Average portfolio size: $\sim$200 stocks
  - **Problem:** Excessive transaction costs make strategy impractical
- **High λ (> 1):** Strong penalization promotes sparsity
  - Fewer assets in replicating portfolio
  - Lower transaction costs
  - **Risk:** Degradation in replication quality

Our Empirical Choice: $\lambda = 0.01$

- **Resulting portfolio size:** $\sim$15 stocks on average
- **Rationale:** Balances replication accuracy with practical tradability
- Robustness to other choices of λ

Outline

1. Introduction
2. Methodology
3. Empirical Application
4. Results
5. Conclusions

Data

- We work with CRSP daily files from January 1st 1962 to December 31st 2024.
- Exchanges: NYSE, AMEX, and NASDAQ (exchange codes 1, 2, and 3)
- Common stocks (share codes 10 and 11)

Sample splits

- Our empirical design retains the two-stage rolling-window approach of [Gatev et al., 2006].
- For every month in the sample we:
  1. Estimate the strategy over a 12-month *formation window* and then,
  2. Trade the resulting spreads during the subsequent 6-month *trading window*.

Formation window (12 months)

1. During each formation period, we define our universe of stocks by filtering out from CRSP any security that exhibits at least one non-trading day in that period.

$$\mathcal{U} = \{j : \exists\ r_{j,t}\ \forall t \in \mathcal{T}_{for}\}$$

2. Then, for every stock in this universe we construct its "*normalized price*" as the cumulative return. This tracks the value of investing 1$ in the stock at the start of the window and continuously reinvest cash dividends. Set $r_{j,0} = 0$, then:

$$p_{j,t} = \prod_{\tau=0}^{t} (1 + r_{j,\tau}), \quad t \in \mathcal{T}_{for}$$

Formation window (12 months)

3. We treat each stock $i \in \mathcal{U}$ as a potential target and construct its replicating portfolio from the donor pool $\mathcal{D}_i = \mathcal{U} \setminus \{i\}$

$$\widehat{\boldsymbol{\beta}}_i := \underset{\boldsymbol{\beta} \in \mathbb{R}^{|\mathcal{D}_i|}}{\arg\min} \left\{ \sum_{t \in \mathcal{T}_{for}} \left(p_{i,t} - \sum_{j \in \mathcal{D}_i} \beta_j p_{j,t} \right)^2 + \lambda \sum_{j \in \mathcal{D}_i} |\beta_j| \right\}$$

4. We rank "pairs" in ascending order of MSE

$$\text{rank}(i) = \sum_{j \in \mathcal{U}} \mathbf{1}\{MSE_i \geq MSE_j\}$$

where $MSE_i := \sum_{t \in \mathcal{T}_{for}} (p_{i,t} - \sum_{j \in \mathcal{D}_i} \beta_j p_{j,t})^2$

Trading window (6 months)

- The trading window commences on the first day following the formation period.
- The trading object is the spread between the target and its replicating portfolio

$$\varepsilon_{i,t} := p_{i,t} - \sum_{j \in \mathcal{D}_i} \beta_j p_{j,t}$$

- The trading rule capitalizes on the expected mean reversion of the spread: $\varepsilon_{i,t} \sim I(0)$
- We initiate a dollar-neutral position whenever it widens beyond two historical standard deviations (estimated in the formation window)

$$TR(\varepsilon_{i,t}) = \begin{cases} \text{LONG} & \text{if} \quad \epsilon_{i,t} < -2\hat{\sigma}_i \\ \text{SHORT} & \text{if} \quad \epsilon_{i,t} > +2\hat{\sigma}_i \\ \text{NEUTRAL} & \text{if} \quad \text{else} \end{cases}$$

Trading window (6 months)

- This trading strategy involves shorting 1$ of the *relatively overvalued* asset and simultaneously longing 1$ in the *relatively undervalued* asset.
- The position is **closed** when the **spread mean reverts** (i.e., prices cross)
- If a **position** is **still open at the end of the 6-month** trading window, it is **liquidated at the prevailing market prices**.
- If a constituent stock is **delisted**, the position is closed out using the **delisting return** or the last available price.

Payoffs

- Returns are computed as **excess returns**, with all positions **marked to market** daily and returns **compounded to the monthly level**.
- Specifically, the daily return on a portfolio P is calculated as

$$r_{P,t} = \frac{\sum_{i \in P} w_{i,t} r_{i,t}}{\sum_{i \in P} w_{i,t}} \qquad \text{where} \qquad w_{i,t} = w_{i,t-1}(1 + r_{i,t-1}) = \prod_{s=1}^{t-1}(1 + r_{i,s})$$

  - $r_{i,t}$ denotes the return
  - $w_{i,t}$ is the evolving weight of asset i at time t.
- This approach is equivalent to a **buy-and-hold strategy**, with daily returns compounded to obtain monthly performance.

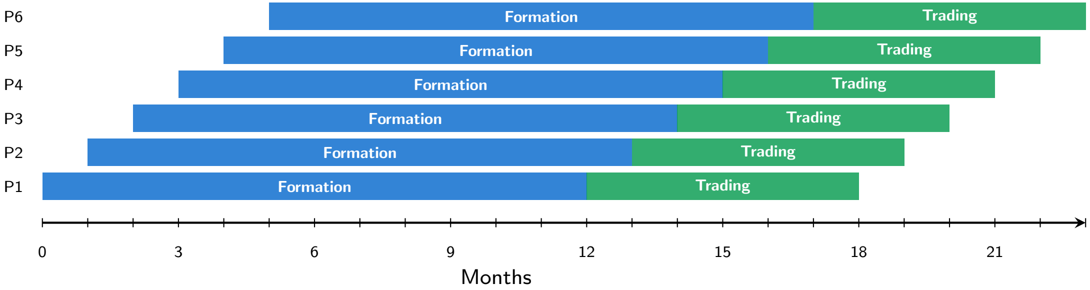
Excess-return metrics

Following [Gatev et al., 2006], we report two measures of excess return:

- **Return on committed capital**.
  - It assumes 1$ is allocated to every pair selected for trading (regardless of whether a trade was triggered)
  - Conservative metric: reflects the opportunity cost of capital commitment
- **Return on fully invested capital**.
  - It assumes 1$ is allocated only to pairs that trigger a trade
  - Reflects the viewpoint of a flexible trading desk that is able to redeploy idle funds

Overlapping portfolios

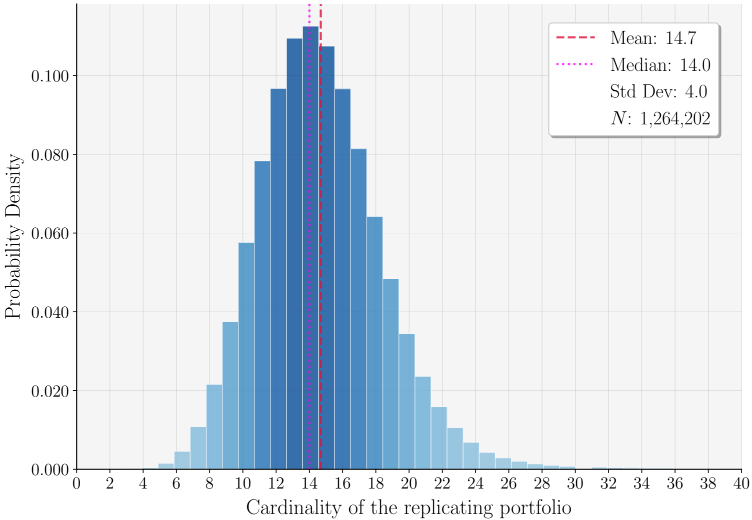
- To generate a continuous time series of returns, we initiate this entire process at the beginning of every month in our sample (excluding the initial 12 months required for the first formation period). This creates a series of overlapping 6-month trading periods.
- To correct for the serial correlation induced by this overlap, we follow the standard procedure of averaging the monthly returns across all strategies that are concurrently active, as in Jegadeesh (1993).
- The resulting series can be interpreted as the net payoff to a trading desk managing six distinct portfolios, with each portfolio staggered by one month.


P6
P5
P4
P3
P2
P1
Formation
Trading
Formation
Trading
Formation
Trading
Formation
Trading
Formation
Trading
Formation
Trading
0
3
6
9
12
15
18
21
Months

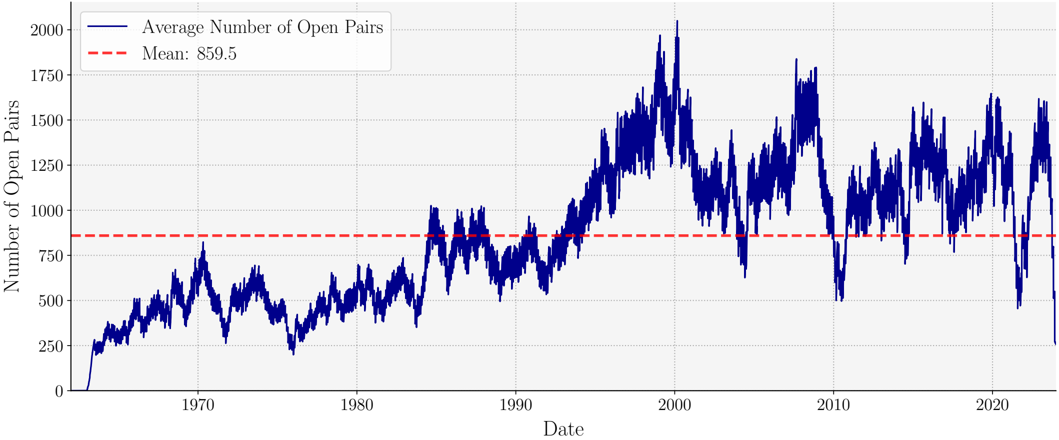
Outline

1. Introduction
2. Methodology
3. Empirical Application
4. Results
5. Conclusions

Number of constituents in the replicating portfolio



Open pairs over time



Monthly Excess Return Distribution [no watiting]

Pairs portfolio	Top 5	Top 20	Pairs 101-120	All Pairs
Average excess return (fully invested)	0.01013	0.00971	0.01013	0.00859
Standard error (Newey-West)	0.00135	0.00103	0.00112	0.00074
t-Statistic	7.50	9.39	9.01	11.59
Excess return distribution				
Median	0.00891	0.00871	0.01003	0.00826
Standard deviation	0.03661	0.02511	0.02726	0.01873
Skewness	0.49	-0.10	-0.32	-0.25
Kurtosis	4.47	4.90	7.96	17.44
Minimum	-0.09500	-0.14352	-0.18062	-0.16669
Maximum	0.18704	0.09846	0.17219	0.12407
Observations with excess return < 0	39%	35%	32%	27%
Average excess return on committed capital	0.00644	0.00513	0.00499	0.00294

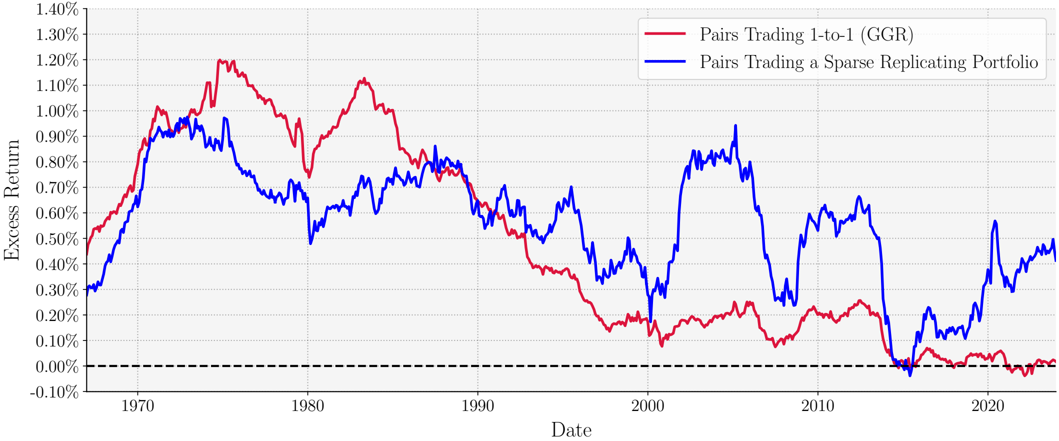
Period: January 1963 to December 2024

Monthly Excess Return Distribution [1 day watiting]

Pairs portfolio	Top 5	Top 20	Pairs 101-120	All Pairs
Average monthly return (fully invested)	0.00853	0.00883	0.00854	0.00745
Standard error (Newey-West)	0.00129	0.00099	0.00113	0.00068
t-Statistic	6.60	8.95	7.58	10.94
Excess return distribution				
Median	0.00650	0.00895	0.00863	0.00711
Standard deviation	0.03714	0.02494	0.02726	0.01828
Skewness	0.72	-0.03	-0.71	-0.18
Kurtosis	6.23	5.57	12.28	18.20
Minimum	-0.10801	-0.15005	-0.19062	-0.16497
Maximum	0.22563	0.10843	0.18429	0.12462
Observations with excess return < 0	41%	36%	35%	30%
Average excess return on committed capital	0.00533	0.00440	0.00426	0.00250

Period: January 1963 to December 2024

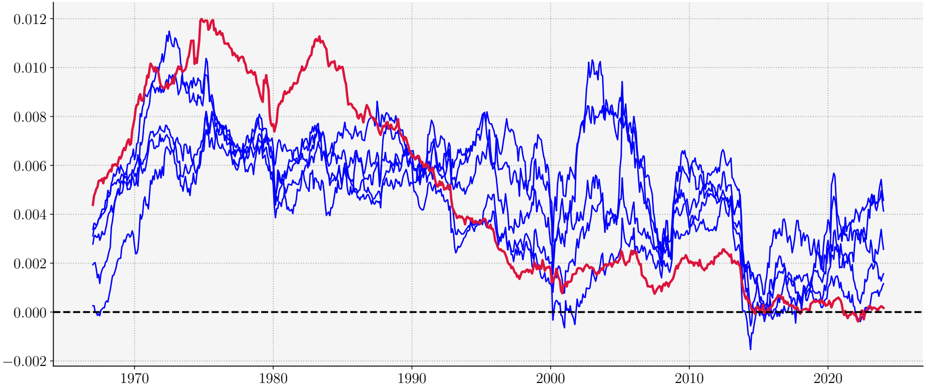
Monthly excess returns [1 day waiting rule] (60 month rolling average)



Robustness to λ

Blue: Replicating Portfolios for different λ values

Red: Traditional Pairs Trading

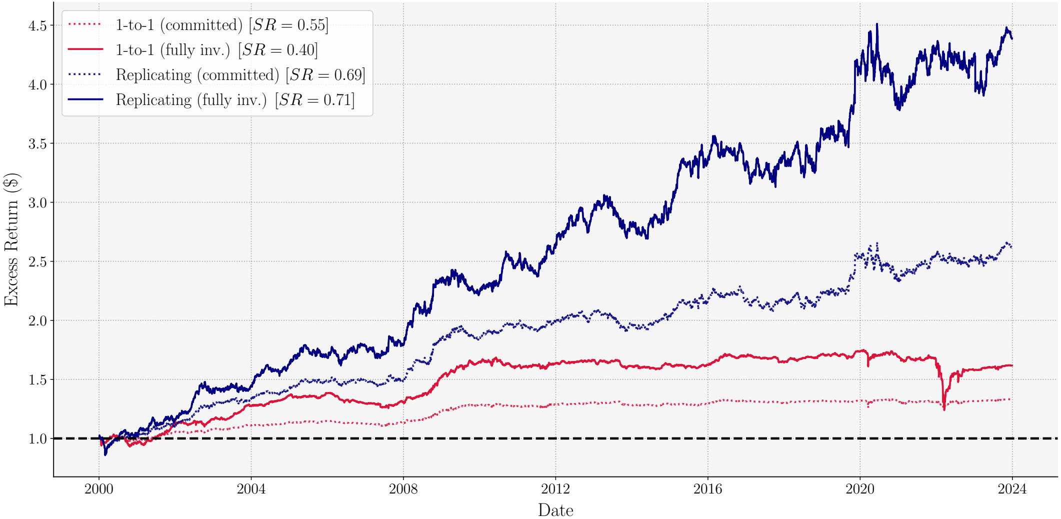


Systematic Risk decomposition: FF5 + Reversals [1 day waiting]

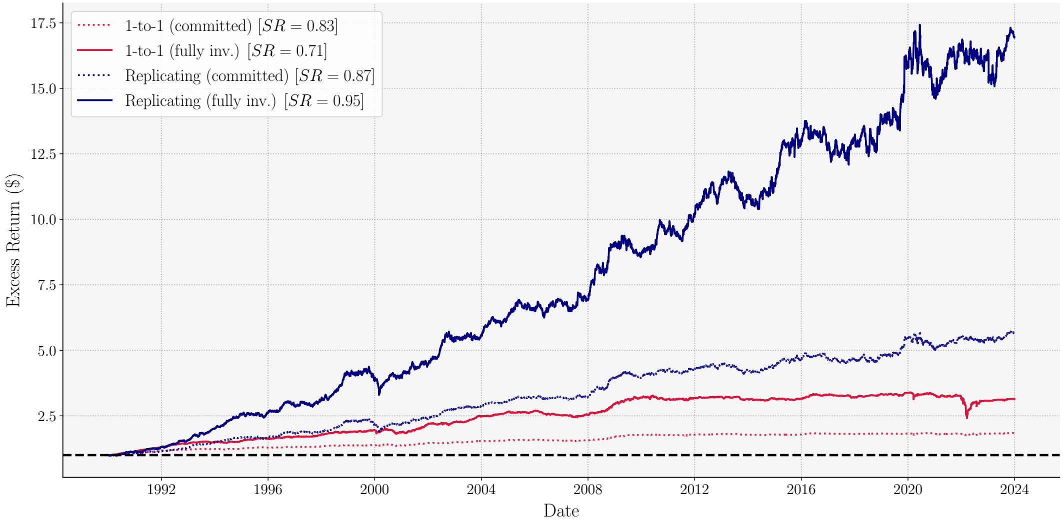
	Top 5	Top 20	20 after top 100	All
Mean excess return	0.0085	0.0088	0.0085	0.0074
Standard deviation	0.0371	0.0249	0.0273	0.0183
Sharpe Ratio	0.80	1.23	1.08	1.41
FF5+Reversals				
Intercept	0.0062 (4.31)	0.0055 (5.55)	0.0052 (4.46)	0.0041 (6.19)
Market	-0.1537 (-3.68)	-0.0768 (-3.12)	-0.0602 (-2.21)	-0.0149 (-0.76)
SMB	-0.0628 (-1.27)	-0.1379 (-2.97)	-0.2189 (-3.13)	-0.0816 (-1.51)
HML	-0.1332 (-1.75)	0.0377 (0.77)	0.0024 (0.04)	-0.0172 (-0.41)
Momentum	-0.0930 (-2.43)	-0.0691 (-2.19)	-0.1452 (-4.82)	-0.1755 (-8.36)
Short-Term Reversal	0.1525 (2.67)	0.1803 (4.37)	0.2389 (4.83)	0.2320 (6.26)
Long-Term Reversal	-0.0113 (-0.16)	-0.0716 (-1.26)	0.0501 (0.84)	-0.0163 (-0.42)
R^2	0.05	0.11	0.19	0.41

Period: January 1963 to December 2024

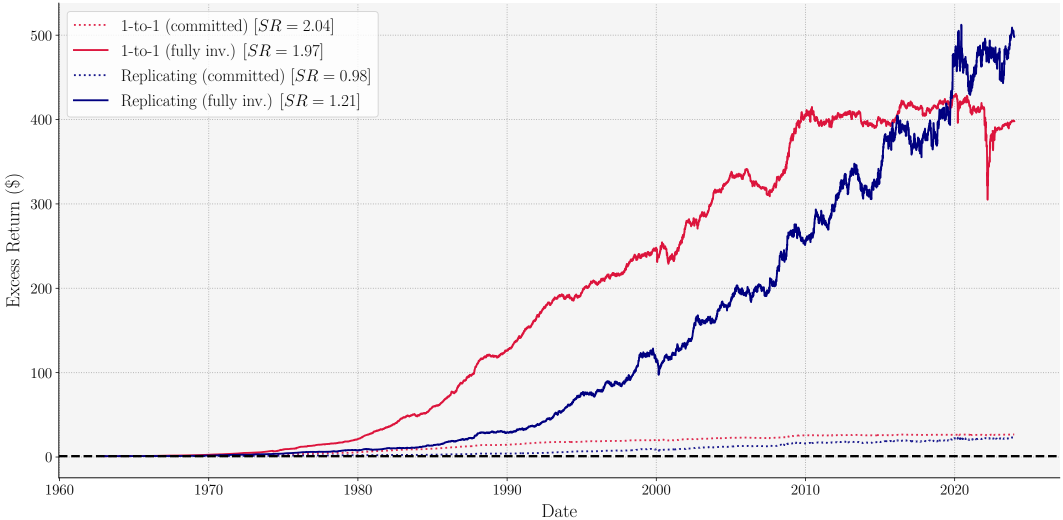
Equity curve to Pairs Trading since 2000



Equity curve since 1990



Equity curve since 1963



Outline

1. Introduction
2. Methodology
3. Empirical Application
4. Results
5. Conclusions

Methodological Contributions

- **Conceptual advancement**: We reframe pairs trading from a rigid cardinality-constrained problem to a flexible replicating portfolio framework
- **Practical implementation**: LASSO regularization creates manageable portfolios ($\approx$15 constituents) while maintaining superior replication quality
- **Theoretical consistency**: The approach preserves the fundamental economic intuition of relative-value arbitrage while expanding the opportunity set

Key Takeaway

Relative-value arbitrage is not obsolete (yet)

- **Traditional pairs trading's profitability has steadily declined since the 1990s**, confirming the decay documented in prior literature.
- The issue is **methodological limitation**: flexibility in substitute construction restores profitability
- **The replicating portfolio approach successfully revitalizes the strategy** by relaxing the restrictive 1-to-1 pairing constraint
- Our approach offers a **principled path forward** for pairs trading in efficient markets

References I

Do, B. and Faff, R. (2010).
Does simple pairs trading still work?
Financial Analysts Journal, 66(4):83–95.

Gatev, E., Goetzmann, W. N., and Rouwenhorst, K. G. (2006).
Pairs trading: Performance of a relative-value arbitrage rule.
Review of Financial Studies, 19(3):797–827.