Is Pairs Trading a thing of the past?

Jesus Villota [1]

CEMFI, Casado del Alisal 5, E-28014 Madrid, Spain

< jesus.villota@cemfi.edu.es >

Abstract

JEL Codes:

Keywords:

[1]I acknowledge

1. Introduction

Revisiting pairs trading 20 years after Gatev et al. (2006) Pairs trading first took shape in the mid-1980s when Nunzio Tartaglia's quantitative desk began systematically scanning U.S. equities for temporary deviations in the relative prices of stocks that had historically moved in tandem. The recipe is intuitive: locate two securities whose past price paths exhibit a tight long-run relationship, short the one that has recently run ahead, go long the laggard, and close the position as soon as the spread mean-reverts. Gatev et al. (2006) famously showed that this seemingly naïve contrarian rule delivered annualised abnormal returns of roughly 11 percent between 1962 and 2002–even after conservative trading-cost assumptions–prompting an explosion of interest among hedge funds and proprietary "market-neutral" desks. Yet the very popularity that followed appears to have eroded the opportunity: these same authors already noted shrinking profits in the late 1990s, and the follow-up study by Do and Faff (2010) documented a further decline. Extending the evidence through 2024, Figure 2 shows that the canonical implementation of the strategy has failed to generate significant positive excess returns since 2010, a decay reminiscent of what befell momentum once it, too, became widely traded.

[Insert Figure 2 about here]

Generalizing Pairs Trading: From Pairs to Sparse Synthetic Replicas The decay in profitability of the classic pairs trade, however, does not necessarily invalidate the underlying principle of relative-value arbitrage. Rather, it suggests that the traditional one-to-one pairing constraint may be too restrictive for modern, complex markets where simple, stable relationships between two individual stocks have become scarce. The core insight of pairs trading –that a security's value can be assessed relative to a close substitute– remains powerful. The challenge lies in identifying or creating a more robust substitute.

Traditional pairs trading imposes a severe cardinality constraint by limiting the replication of a target asset to a single substitute security. This restriction, while conceptually elegant, may be unnecessarily limiting in practice. Consider General Motors: rather than searching for a single stock whose normalized price series closely tracks GM's movements, we could construct a synthetic replica using a linear combination of multiple securities. This approach maintains the fundamental economic intuition of pairs trading–exploiting temporary deviations between economically related assets–while providing greater flexibility in replica construction.

The theoretical foundation remains rooted in relative pricing theory, where securities serving as close economic substitutes should exhibit similar price dynamics. However, our framework relaxes the stringent requirement that such substitution be achieved through a single asset. Instead, we allow for the construction of synthetic substitutes through sparse linear combinations, potentially improving the replication of the target asset.

This paper proposes a generalization of the pairs trading framework that relaxes this rigid cardinality constraint. Instead of searching for a single, naturally occurring substitute for a target asset, we propose to construct a "synthetic replica" of it. Methodologically, this is achieved by regressing the normalized price series of the target asset against a broad "donor pool" of potential substitutes. The fitted values from this regression form the price series of the synthetic replica –a portfolio of assets weighted to best track the target. The spread in this generalized framework is therefore the regression error: the price difference between the target asset and its synthetic replica.

Making the replicating portfolio sparse A crucial challenge, however, is that an unconstrained Ordinary Least Squares (OLS) regression using a large donor pool would yield a dense synthetic replica, composed of hundreds of assets. Such a portfolio would be prohibitively expensive to trade due to transaction costs and slippage, rendering the strategy impractical and suffering from the curse of dimensionality. To overcome this, we introduce a penalized regression approach using the Least Absolute Shrinkage and Selection Operator (LASSO). The LASSO penalty is ideally suited for this context as it promotes sparsity by forcing the coefficients of less relevant assets in the donor pool to exactly zero. This technique simultaneously performs automated stock selection and estimates the weights for the synthetic replica, effectively identifying the most important constituents for creating a parsimonious, and therefore tradable, replicating portfolio.

Law of One Price By the Law of One Price, the target stock should have the same (normalized) price as its replicating portfolio, hence, any deviation in their difference (i.e.:, in the regression error or spread) should eventually close. Our trading strategy capitalizes on this by betting on the mean reversion of the spread, shorting the target asset when the residual is positive (i.e., the target is overpriced relative to its replica) and going long when it is negative. In this context, the cointegrating vector is

Cointegration Gatev et al argue that pairs trading is related to cointegration (in the sense of Engle and Granger), as the spread of pairs is expected to mean revert (i.e: it is stationary). Formally, for a particular pair of stocks i and j, their *assumed* cointegrating vector is the difference between their canonical vectors $\boldsymbol{\alpha} = \mathbf{e}_i - \mathbf{e}_j$. Figure 2 shows that finding cointegrating vectors of that shape used to be an easy task in the past, as pair spreads where highly mean-reverting, and therefore, it was highly profitable to bet on this characteristic. Howevever, this property has recently become harder to trade, as markets are now more efficient. The methodology proposed in this paper relaxes the structure imposed on the cointegrating vector to $\boldsymbol{\alpha} = \mathbf{e}_i - \boldsymbol{\beta}$, where i denotes the target stock and $\boldsymbol{\beta}$ contains a 0 in the target stock's position and the regression coefficients of the target stock on a donor pool of assets, and a

In this context, there is a branch of literature in statistical arbitrage called "mean reverting portfolios", whose main concern is to build portfolios of stocks that are cointegrated. Both, pairs trading and our generalized approach based on pairs trading a replicating portfolio can be embedded withing this umbrella. Note that the weights of these mean-reverting portfolios are precisely the cointegrating vectors specified in the previous paragraph.

Our methodology adapts the core formation and trading period structure of Gatev et al. (2006) to this sparse, synthetic framework. We re-examine whether this more flexible and robust form of pairs trading–trading an asset against its sparse synthetic counterpart–can restore profitability in the modern era.

2. Methodology

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.

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

where $p_{i,t}$ represents the normalized price of the target asset[2], $\mathcal{D}$ denotes the donor pool of potential replicating assets, and $\varepsilon_{i,t}$ captures the spread between the target and its synthetic replica.

Traders should then be interested in trading the spread between the target stock and the replicating portfolio, which is simply the residual error $\varepsilon_{i,t}$, by betting on its mean-reversion.

While **??** describes the unrestricted case, where no structure is imposed in the replicating portfolio, we can show that a restricted version from this general framework delivers the classical

[2] The normalized price time series of a stock i is built from a position where 1 dollar is invested at the beginning of the estimation window. Formally, letting $r_{i,0} = 0$, the normalized price time series at time t is:

$$p_{i,t} = \prod_{\tau=0}^{t} (1 + r_{i,\tau})$$

pairs trading methodology.

Recasting Gatev et al. (2006)'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. In this case, the only nonzero element of the coefficient vector will correspond to that of the stock whose normalized price time series minimizes the euclidean distance with respect to the target stock's normalized price time series.

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 one-to-one relationship by imposing the restriction $\boldsymbol{\beta} \in \mathcal{B}$.

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

which is equivalent to selecting the stock whose price series minimizes the least-squares distance (or equivalently, the euclidean distance) with respect to that of the target stock. [3]

$$j^* = \arg\min_{j \in \mathcal{D}} \sum_{t=0}^{T} \Big(p_{i,t} - p_{j,t} \Big)^2 \tag{3}$$

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. To promote sparsity while avoiding hard cardinality constraints, we employ LASSO regularization.

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^{|\mathcal{D}|}} \left\{ \sum_{t=0}^{T} \left(p_{i,t} - \sum_{j \in \mathcal{D}} \beta_j p_{j,t} \right)^2 + \lambda \sum_{j \in \mathcal{D}} |\beta_j| \right\} \tag{4}$$

Note that, for λ close to 0, the lasso penalization is very small, so our replicating portfolio will be composed of a large amount of assets. In the limit, $\lambda = 0$ and we recover eq. (1); in this case, the average number of stocks in the replicating portfolio averages 200 stocks, which becomes too

[3] Note that pairs trading with a hedge ratio is equivalent to restricting only the cardinality of the coefficient vector. Or equivalently, to restricting the coefficient vector to lie in $\mathcal{B}^+ := \{\alpha \mathbf{e}_j : j \in \mathcal{D}, \alpha \in \mathbb{R}\}$. In this case, the replicating portfolio program is:

$$\begin{bmatrix} \min\limits_{\boldsymbol{\beta} \in \mathbb{R}^{|\mathcal{D}|}} & \sum_{t=0}^{T} \Big(p_{i,t} - \sum_{j \in \mathcal{D}} \beta_j p_{j,t} \Big)^2 \\ \text{s.t.} & \|\boldsymbol{\beta}\|_0 = 1 \end{bmatrix} \equiv \min_{\boldsymbol{\beta} \in \mathcal{B}^+} \sum_{t=0}^{T} \Big(p_{i,t} - \sum_{j \in \mathcal{D}} \beta_j p_{j,t} \Big)^2$$

unwieldy and costly to trade (transaction costs rapidly explode). On the other hand, for large λ we recover sparser solutions. In our applications we experiment with a grid of λ values and report all of them to show that results are robust to the choice of this hyperparameter. The best level of penalization is $\lambda = 0.001$, which delivers replicating portfolios composed of 20 stocks on average.

The following section details our empirical implementation and assess whether this generalized approach can restore profitability to spread-based pairs trading strategies.

[Insert Figure 2 about here]

3. Empirical Application

Data

- We work with CRSP daily files from January 1st 1962 to December 31rd 2024.
- Exchanges NYSE, AMEX, and NASDAQ (EXCHCD 1, 2, and 3)
- Share codes 10 and 11

Our empirical design retains the two-stage rolling-window approach of Gatev et al. (2006). For every month in the sample we first estimate the strategy over a 12-month *formation window* and then trade the resulting spreads during the subsequent 6-month *trading window*.

Formation period 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.

Then, for every stock in this universe we construct its associated cumulative-return index, which that tracks the value of one dollar invested in that stock at the start of the window and continuously reinvests cash dividends.

For each stock in our selected universe, we treat it as a potential target and seek to construct its sparse synthetic replica. As outlined previously, this is achieved by regressing the target's normalized price series against the series of all other liquid stocks in the donor pool, using a LASSO penalty to enforce sparsity.

For our main analysis, we select a penalization parameter of $\lambda = 0.001$, which effectively shrinks the coefficients of irrelevant assets to zero and yields a parsimonious replicating portfolio, typically composed of around 20 constituent securities. The spread is the time series of residuals from this regression–the difference between the target stock's price and the price of its synthetic replica. Following the formation period, we carry forward the top 20 target-replica strategies with the lowest sum of squared residuals for trading.

Trading period The trading phase commences on the first day following the formation period. Our trading rules are designed to capitalize on the expected mean reversion of the spread. We initiate a dollar-neutral position when a target stock's price diverges significantly from that of its synthetic replica. Specifically, a position is opened when the spread widens beyond two of its historical standard deviations, as estimated during the 12-month formation period. This involves shorting one dollar of the relatively overvalued asset and simultaneously investing one dollar in the undervalued asset. The position is held until the spread reverts and crosses its mean (i.e., the prices cross), at which point it is closed. If a position is still open at the end of the 6-month trading window, it is liquidated at the prevailing market prices. Similarly, if a constituent stock is delisted, the position is closed out using the delisting return or the last available price. Positions may open, close, and reopen multiple times within the six-month window.

Payoffs The payoffs from this strategy are equivalent to excess returns, as they are derived from self-financing, dollar-neutral positions. The total excess return for each target-replica strategy over the trading period is the sum of its reinvested payoffs. Positions are marked-to-market daily, and the daily returns on the long and short portfolios are compounded to calculate monthly returns, which mirrors a buy-and-hold approach for the underlying assets.

The realized payoffs from this strategy are characterized by a sequence of cash flows: positive flows accrue each time a spread opens and subsequently converges within the trading window, while unresolved spreads generate a final cash flow at the end of the period. If no divergence occurs, no trade is initiated. 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

$$r_{P,t} = \frac{\sum_{i \in P} w_{i,t} r_{i,t}}{\sum_{i \in P} w_{i,t}}$$

$$w_{i,t} = w_{i,t-1}(1 + r_{i,t-1}) = \prod_{s=1}^{t-1}(1 + r_{i,s})$$

where $r_{i,t}$ denotes the return and $w_{i,t}$ 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: the return on committed capital (which assumes a dollar is allocated to every pair selected for trading, regardless of whether a trade is triggered). That is, it divides aggregate pay-offs by the number of spreads selected for trading, charging the strategy for capital tied up even if a spread never opens.The fully invested return scales payoffs by the actual capital deployed in open trades. The former is a conservative metric, reflecting the opportunity cost of capital commitment, while the latter provides a more realistic gauge of trading profitability for flexible investors, reflecting 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.

References

B. Do and R. Faff. Does simple pairs trading still work? *Financial Analysts Journal*, 66(4):83–95, July 2010. ISSN 1938-3312. doi: 10.2469/faj.v66.n4.1. URL http://dx.doi.org/10.2469/faj.v66.n4.1.

E. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst. Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*, 19(3):797–827, 2006. ISSN 1465-7368. doi: 10.1093/rfs/hhj020. URL http://dx.doi.org/10.1093/rfs/hhj020.

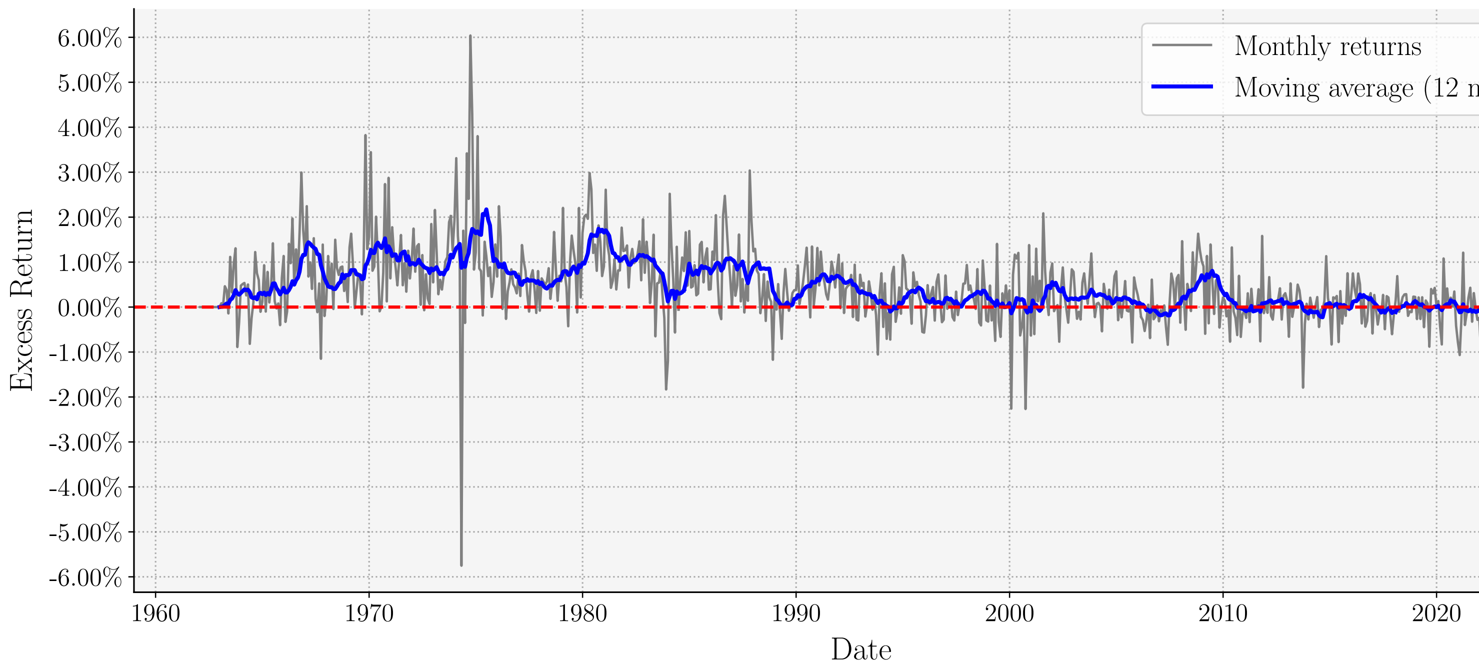
Table 1: Excess returns of unrestricted pairs trading strategies

Pairs portfolio	Top 5	Top 20	Pairs 101-120	All Pairs
A. Excess return distribution (no waiting)				
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
B. Excess return distribution (one day waiting)				
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

Table 2: Systematic risk of pairs trading strategies

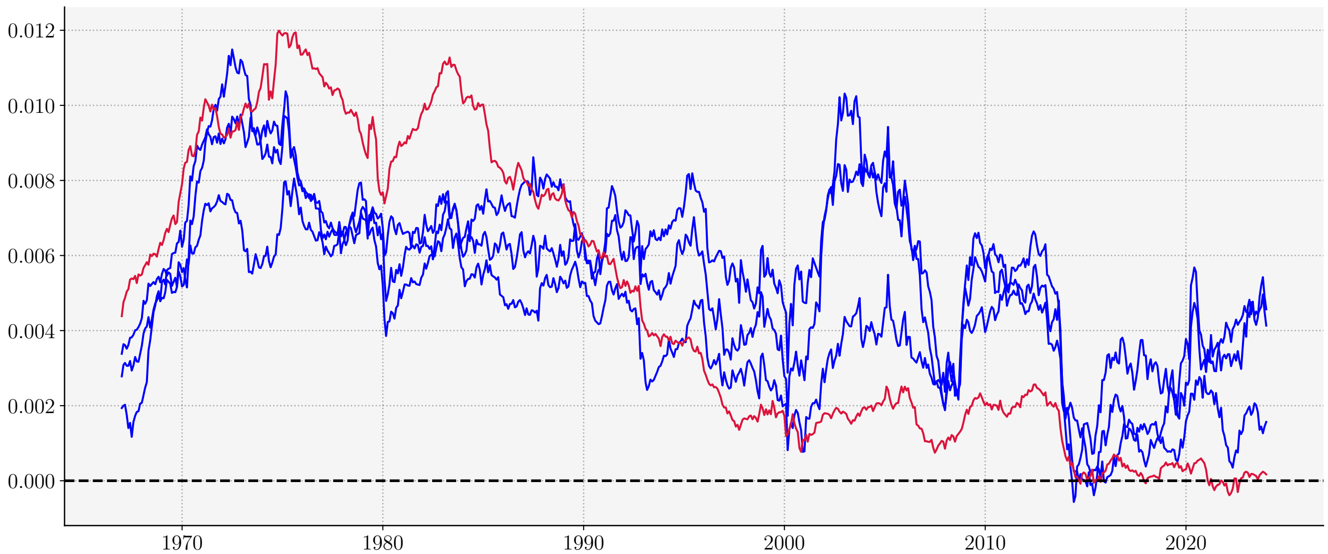
	Top 5	Top 20	20 after top 100	All
"Wait one day" portfolio performance				
Mean excess return	0.00853	0.00883	0.00854	0.00745
Standard deviation	0.03714	0.02494	0.02726	0.01828
Sharpe Ratio	0.80	1.23	1.08	1.41
Monthly serial correlation	-0.03	0.02	0.02	0.02
FF5 + Reversals				
Intercept	0.00618	0.00551	0.00525	0.00409
	(4.31)	(5.55)	(4.46)	(6.19)
Market	-0.15365	-0.07676	-0.06021	-0.01491
	(-3.68)	(-3.12)	(-2.21)	(-0.76)
SMB	-0.06277	-0.13786	-0.21888	-0.08163
	(-1.27)	(-2.97)	(-3.13)	(-1.51)
HML	-0.13321	0.03773	0.00239	-0.01717
	(-1.75)	(0.77)	(0.04)	(-0.41)
Momentum	-0.09301	-0.06905	-0.14525	-0.17550
	(-2.43)	(-2.19)	(-4.82)	(-8.36)
Short-Term Reversal	0.15247	0.18033	0.23888	0.23198
	(2.67)	(4.37)	(4.83)	(6.26)
Long-Term Reversal	-0.01133	-0.07164	0.05008	-0.01632
	(-0.16)	(-1.26)	(0.84)	(-0.42)
R^2	0.05	0.11	0.19	0.41

Figure 1: Decay of pairs trading excess returns



Note: This figure plots the monthly returns to pairs trading along with its 12-month moving average. The strategy is implemented as in Gatev et al. (2006) and the figure is displayed as in Do and Faff (2010), but extending the sample to the end of 2024. We thank Binh Do and Alexander Rubesam for facilitating the code to replicate this figure.

Figure 2: Decay of pairs trading excess returns



Note: 40-month moving average of monthly excess returns of traditional pairs trading versus generalized pairs trading. As we can see, restricting the coefficient of the replicating portfolio to the canonical basis added economic until 1990, there on, as 1-to-1 spreads became more heavily traded, the profitability of this type of spread trading decreaased substantially. Hoewever, the generalized approach that allows more flexibility in the coefficient vector of the replicating portfolio maintains profitability even in the last decade, where markets have become heavily efficient (note that excess returns to traditional pairs trading are 0 in the last decade).