JOURNAL OF ECONOMIC SURVEYS

doi: 10.1111/joes.12153

STATISTICAL ARBITRAGE PAIRS TRADING STRATEGIES: REVIEW AND OUTLOOK

Christopher Krauss*

Department of Statistics and Econometrics
University of Erlangen-Nürnberg

Abstract. This survey reviews the growing literature on pairs trading frameworks, i.e., relative-value arbitrage strategies involving two or more securities. Research is categorized into five groups: The distance approach uses nonparametric distance metrics to identify pairs trading opportunities. The cointegration approach relies on formal cointegration testing to unveil stationary spread time series. The time-series approach focuses on finding optimal trading rules for mean-reverting spreads. The stochastic control approach aims at identifying optimal portfolio holdings in the legs of a pairs trade relative to other available securities. The category "other approaches" contains further relevant pairs trading frameworks with only a limited set of supporting literature. Finally, pairs trading profitability is reviewed in the light of market frictions. Drawing from a large set of research consisting of over 100 references, an in-depth assessment of each approach is performed, ultimately revealing strengths and weaknesses relevant for further research and for implementation.

Keywords. Mean-reversion; Pairs Trading; Spread Trading; Relative-value Arbitrage

1. Introduction

According to Gatev *et al.* (2006), the concept of pairs trading is surprisingly simple and follows a two-step process. First, find two securities whose prices have moved together historically in a formation period. Second, monitor the spread between them in a subsequent trading period. If the prices diverge and the spread widens, short the winner and buy the loser. In case the two securities follow an equilibrium relationship, the spread will revert to its historical mean. Then, the positions are reversed and a profit can be made. The concept of univariate pairs trading can be extended: In quasi-multivariate frameworks, one security is traded against a weighted portfolio of comoving securities. In multivariate frameworks, groups of stocks are traded against other groups of stocks. Terms of reference for such strategies are (quasi-)multivariate pairs trading, generalized pairs trading, or statistical arbitrage. We further consider all these strategies under the umbrella term of "pairs trading," since it is the ancestor of more complex approaches (Vidyamurthy, 2004; Avellaneda and Lee, 2010).

The most cited paper in this domain is published by Gatev *et al.* (2006), hereafter GGR. A simple yet compelling algorithm is tested on a large sample of U.S. equities, while rigorously controlling for data snooping bias. The strategy yields annualized excess returns of up to 11% at low exposure to systematic sources of risk. More importantly, profitability cannot be explained by previously documented reversal profits as in Jegadeesh (1990) and Lehmann (1990) or momentum profits as in Jegadeesh and Titman (1993). These unexplained excess returns elevate GGR's pairs trading to one of the few capital market phenomena that have stood the test of time[1] as well as independent scrutiny by later authors, most notably Do and Faff (2010, 2012).

*Corresponding author contact email: christopher.krauss@fau.de; Tel: +49 (0) 911 5302 290.

Journal of Economic Surveys (2017) Vol. 31, No. 2, pp. 513–545

Despite these findings, academic research about pairs trading is still small compared to contrarian and momentum strategies.[2] However, interest has recently surged, and pairs trading research is gaining momentum along five streams of literature:

- **Distance approach:** In the formation period, distance metrics are leveraged to identify comoving securities. In the trading period, simple nonparametric threshold rules are used to trigger trading signals. The strategy is simple and transparent, allowing for large-scale empirical applications. With these studies, pairs trading is established as profitable across different markets, asset classes, and time frames.
- **Cointegration approach:** In the formation period, cointegration tests are applied to identify comoving securities. In the trading period, most authors use simple algorithms to generate trading signals; the majority of them based on GGR's threshold rule. The key benefit of these strategies is the econometrically more reliable equilibrium relationship of identified pairs.
- **Time-series approach:** All authors in this domain assume that a set of comoving securities has been established by prior analyses, so the formation period is generally ignored. Instead, the focus lies on the trading period and on generating optimized trading signals by different methods of time-series analysis.
- **Stochastic control approach:** As in the time-series approach, the formation period is ignored. Instead, the focus lies on identifying optimal portfolio holdings in the legs of a pairs trade compared to other assets. Stochastic control theory is used to determine value and optimal policy functions for this portfolio problem.
- **Other approaches:** This bucket contains further pairs trading frameworks with only a limited set of supporting literature and limited relation to previously mentioned approaches. Included in this category are the machine learning (ML) and combined forecasts approach, the copula approach, and the principal components analysis (PCA) approach.

Table 1 provides an overview of representative studies per approach, the data sample, and returns per annum.[3] Tables A.1–A.5 in the Appendix give a more detailed picture.

Considering the diversity of these categories, the contribution of this survey is threefold: First, a comprehensive review of pairs trading literature is conducted along the five approaches. Second, the most relevant contributions per category are discussed in detail. Third, pairs trading profitability is assessed in light of market frictions. Drawing from a large set of literature consisting of over 100 references, an in-depth assessment of each approach is possible, ultimately revealing strengths and weaknesses relevant for further research and for implementation. The latter aspect makes this survey relevant for researchers and practitioners alike. The remainder of this paper is organized as follows: Section 2 covers the distance approach and its various empirical applications. Section 3 reviews uni- and multivariate frameworks for the cointegration approach. Section 4 covers the time-series approach and discusses different models aiming at the identification of optimal trading thresholds. Section 5 reviews the stochastic control approach and how to determine optimal portfolio holdings. Section 6 covers the remaining approaches and Section 7 discusses pairs trading and market frictions. Finally, Section 8 concludes and summarizes directions for further research.

2. Distance Approach

2.1 *The Baseline Approach—Gatev, Goetzmann, and Rouwenhorst*

The distance approach is introduced with the seminal paper of Gatev *et al.* (2006). Their study is performed on all liquid U.S. CRSP stocks from 1962 to 2002. First, a cumulative total return index is constructed for each stock and normalized to the first day of a 12-month formation period. Second, with n stocks

Table 1. Overview Pairs Trading Approaches

Approach	Representative Studies	Sample	Return p.a.
Distance	Gatev *et al.* (2006)	U.S. CRSP 1962–2002	0.11
	Do and Faff (2010)	U.S. CRSP 1962–2009	0.07
Cointegration	Vidyamurthy (2004)	–	–
	Rad *et al.* (2015)	U.S. CRSP 1962–2014	0.10
Time series	Elliott *et al.* (2005)	–	–
	Cummins and Bucca (2012)	Energy futures 2003–2010	≥ 0.18
Stochastic	Jurek and Yang (2007)	Selected stocks 1962–2004	0.28–0.43
control	Liu and Timmermann (2013)	Selected stocks 2006–2012	0.06–0.23
Others: ML, combined	Huck (2009)	U.S. S&P 100 1992–2006	0.13–0.57
forecasts	Huck (2010)	U.S. S&P 100 1993–2006	0.16–0.38
Others:	Krauss and Stübinger (2015)	U.S. S&P 100 1990–2014	0.07–0.08
Copula	Rad *et al.* (2015)	U.S. CRSP 1962–2014	0.05
Others: PCA	Avellaneda and Lee (2010)	U.S. subset 1997–2007	–

under consideration, the sum of Euclidean squared distance (SSD) for the price time series[4] of $n(n-1)/2$ possible combinations of pairs is calculated. The top 20 pairs with minimum historic SSD are considered in a subsequent six-month trading period. Trades are opened when the spread diverges by more than two historical standard deviations and closed upon mean-reversion, at the end of the trading period, or upon delisting. The advantages of this methodology are relatively clear: As Do *et al.* (2006) point out, GGR's approach is economic model-free, and as such not subject to model misspecifications and misestimations. It is easy to implement, robust to data snooping, and results in statistically significant risk-adjusted excess returns. However, the choice of Euclidean squared distance as selection metric is analytically suboptimal. Let us assume that a rational pairs trader has the objective of maximizing excess returns per pair. With constant initial invest, this amounts to maximizing profits per pair. The latter are the product of number of trades per pair and profit per trade. As such, a pairs trader aims for spreads exhibiting frequent and strong divergences from and subsequent convergences to equilibrium. In other words, the profit-maximizing rational investor seeks out pairs with high spread variance and strong mean-reversion properties. These two attributes generate a high number of round-trip trades with high profits per trade. Let us now examine how GGR's ranking logic relates to these requirements.

Spread variance. $p_{i,t}$ and $p_{j,t}$ denote realizations of the normalized price processes $P_i = (P_{i,t})_{t\in T}$ and $P_j = (P_{j,t})_{t\in T}$ of the securities i and j of a pair and $s^2(\cdot)$ sample variance. Empirical spread variance $s^2_{P_i-P_j}$ can be expressed as

$$s^2_{P_i-P_j} = \frac{1}{T}\sum_{t=1}^{T}\left(p_{i,t}-p_{j,t}\right)^2 - \left(\frac{1}{T}\sum_{t=1}^{T}\left(p_{i,t}-p_{j,t}\right)\right)^2 \tag{1}$$

We can solve for the average sum of squared distances $\overline{ssd}_{P_i,P_j}$ in the formation period:

$$\overline{ssd}_{P_i,P_j} = \frac{1}{T}\sum_{t=1}^{T}\left(p_{i,t}-p_{j,t}\right)^2 = s^2_{P_i-P_j} + \left(\frac{1}{T}\sum_{t=1}^{T}\left(p_{i,t}-p_{j,t}\right)\right)^2 \tag{2}$$

First of all, it is trivial to see that an "ideal pair" in the sense of GGR with zero SSD has a spread of zero and thus produces no profits. The latter is indicative for a suboptimal selection metric, since we would expect the number one pair of the ranking to generate the highest profits. Next, let us consider pairs with low average SSD at the top of GGR's ranking. Equation (2) shows that constraining for low SSD is the same as minimizing the sum of (i) spread variance and (ii) squared spread mean. We see that summand (ii) grows with spread mean drifting away from zero. Conversely, summand (i) grows with increasing deviations from this mean. It is hard to say which effect dominates the minimization problem in an empirical application. However, table 2 of GGR's results clearly shows decreasing spread volatility with decreasing SSD. Thus, GGR's selection metric is prone to form pairs with low spread variance and limited profit potential.

Mean-reversion. GGR interpret the pair's price time series as cointegrated in the sense of Bossaerts (1988). However, the author develops a rigorous cointegration test based on canonical correlation analysis. Conversely, GGR perform no cointegration testing whatsoever (Galenko *et al.*, 2012). As such, the high correlation[5] may well be spurious, since high correlation is not related to a cointegration relationship (Alexander, 2001). Omitting cointegration testing is contradictory to the requirements of a rational investor. Spurious relationships based on an assumption of return parity are not mean-reverting. The potential lack of an equilibrium relationship leads to higher divergence risks and thus to potential losses. Do and Faff (2010) confirm that 32% of all identified distance pairs do not converge. Huck and Afawubo (2015) show that pairs selected based on cointegration more frequently exhibit mean-reverting behavior compared to distance pairs, even if they do not necessarily converge until the end of the trading period (see share of nonconvergent profitable trades in Huck and Afawubo (2015, table 3, p. 606)).

Given these deficiencies, cointegration testing constitutes an advanced selection metric. Mean-reverting spreads are more prone to be selected and restrictions on spread variance from OLS estimation less rigid. These assertions are confirmed in a recent comparison study of Huck and Afawubo (2015), showing that cointegrated pairs are more profitable and their spread volatility twice as high compared to distance pairs.

2.2 *Expanding on the GGR Sample*

Do and Faff (2010, 2012) replicate GGR's methodology on the same stock universe with extended sample period until 2009. They confirm declining profitability, mainly due to an increasing share of nonconverging pairs. With the inclusion of trading costs, pairs trading according to GGR's baseline methodology becomes largely unprofitable. The authors then use refined selection criteria to improve pairs identification. First, they only allow for matching securities within the 48 Fama–French industries. This restriction potentially leads to more meaningful pairs and fewer spurious correlations. However, there is also potential to miss out on interindustry opportunities. For example, Cohen and Frazzini (2008) find substantial customer–supplier links in the U.S. stock market allowing for return predictability. Second, pairs with a high number of zero-crossings in the formation period are favored. This heuristic is used as a proxy for mean-reversion strength, in lieu of cointegration testing. The top portfolios incorporating industry restrictions, the number of zero-crossings as well as SSD in the selection algorithm are still slightly profitable, even after trading costs. However, the methodology of Do and Faff (2012) is more susceptible to data snooping, since they test a total of 29 different combinations of selection algorithms. Nevertheless, their studies significantly corroborate GGR's findings and help establish pairs trading as a capital market anomaly.

2.3 *From SSD to Pearson Correlation and Quasi-Multivariate Pairs Trading*

Chen *et al.* (2012) use the same data set and time frame as GGR, but opt for Pearson correlation on return level for identifying pairs. In a five-year formation period, monthly pairwise return correlations are calculated for all stocks. Then, the authors construct an empirical metric to quantify return divergence d_{ij} between the returns of stock i and comover j,

$$d_{ij,t} = \beta(r_{i,t} - r_f) - (r_{j,t} - r_f) \qquad t \in T \tag{3}$$

where β denotes the regression coefficient of stock i's return realizations r_i on its comover's return realizations r_j and r_f is the risk-free rate. Chen *et al.* (2012) consider two cases for the comover return r_j: In the univariate case, it is the return of the most highly correlated partner for stock i. In the quasi-multivariate case, it is the return of a comover portfolio, consisting of the equal-weighted returns of the 50 most highly correlated partners for stock i. Subsequent to the formation period follows a one-month trading period. All stocks are sorted in descending order based on their previous month's return divergence and split into deciles. A dollar-neutral portfolio is constructed by longing decile 10 and shorting decile 1, and held for one month.

The key question is how the correlation selection metric on return level differs from the SSD selection metric on price level. First, let us consider sample variance of spread returns, defined as return on buy minus return on sell (Pole, 2008):

$$s^2_{R_i - R_j} = s^2_{R_i} + s^2_{R_j} - 2\hat{\rho}_{R_i,R_j}\sqrt{s^2_{R_i}}\sqrt{s^2_{R_j}} \tag{4}$$

We immediately see from equation (4) that constraining for high return correlation $\hat{\rho}_{R_i,R_j}$ between stocks i and j leads to lower variance of spread returns. However, the return time series of the individual securities may still exhibit vastly different variances. Clearly, this selection metric is more flexible than minimizing SSD. Following an example of Chen *et al.* (2012), consider two securities with perfect return correlation, but one stock return is always twice that of the other. Such return divergences can successfully be captured in the correlation-based framework. In case the selection metric is meaningful and divergences are reversed in the following month, a profit can be made. Conversely, the SSD metric misses out on this opportunity, since the price spread between these two stocks is clearly divergent. The second difference to GGR's study stems from the higher information level contained in a diversified comover portfolio as opposed to single stocks. A third difference lies in the trading method, which is executed mechanically once a month. As such, it is clear in advance how many pairs are traded and which amount of capital has to be allocated. For an equal-weighted portfolio, Chen *et al.* (2012) report average monthly raw returns of 1.70%, almost twice as high as those of GGR. The majority of this increase can be explained by the advantages of the comover portfolio. Reducing the number of stocks in the comover portfolio to one leads to a drop in returns by almost one third.[6] The rest of the edge versus the GGR methodology most likely stems from the higher flexibility of return correlation as pairs selection metric. Nonetheless, it needs to be pointed out that return correlation may be favorable to SSD from this empirical point of view, but it is also far from optimal. Two securities correlated on return level do not necessarily share an equilibrium relationship and there is no theoretical foundation that divergences are reversed. Many correlations in the study may well be spurious, so cointegration testing is strongly recommended.

Perlin (2007, 2009) also tests the advantages of quasi-multivariate versus univariate pairs trading on the 57 most liquid stocks in the Brazilian market from 2000 to 2006. Pairs selection is performed by maximizing Pearson correlation between standardized price time series, which is equivalent to minimizing SSD.[7] In the univariate context, one comover stock is identified and in the quasi-multivariate case, a weighted comover portfolio of five partner stocks. Trades are triggered with simple threshold rules. In the univariate case, the author goes long the undervalued and short the overvalued security. In the quasi-multivariate case, only the reference component of each pair is traded to avoid high transaction costs. Similar to Chen *et al.* (2012), quasi-multivariate pairs trading is found to outperform its univariate counterpart.

2.4 *Explaining Pairs Trading Profitability*

Gatev *et al.* (2006) show that risk-adjusted excess returns of disjoint pairs portfolios exhibit high correlation and surmise that these returns are a compensation for a yet undiscovered latent risk factor. Subsequent studies focus on the sources of pairs trading profitability.

Andrade *et al.* (2005) replicate GGR's approach in the Taiwanese stock market from 1994 to 2004. First, they confirm GGR's findings out-of-sample. Then, they link uninformed demand shocks with pairs trading risk and return characteristics. Andrade *et al.* (2005) find that the dominant factor behind spread divergence is uninformed buying and conclude that pairs trading profits are a compensation for liquidity provision to uninformed buyers.

Papadakis and Wysocki (2007) use GGR's pairs trading rule on a subset of the U.S. equity market to analyze the impact of accounting events on pairs trading profitability between 1981 and 2006. They find that trades are often opened around earnings announcements and analyst forecasts. Such trades are significantly less profitable than those in nonevent periods, which can be explained by investor underreaction. Incremental excess returns are earned by delaying the closures up to three weeks after accounting events. This research suggests that drift in stock prices after such events is a significant factor affecting pairs trading profitability. However, Do and Faff (2010) could not replicate these results on the extended GGR sample, casting doubt on the findings.

Engelberg *et al.* (2009) test a variant of the GGR algorithm on the CRSP U.S. stock universe from 1993 to 2006. They find that pairs trading profitability exponentially decreases over time and that this profitability is strongly related to events at the time of spread divergence. Idiosyncratic information and idiosyncratic liquidity shocks are unfavorable, since they have no impact on the paired firm and render spread divergences permanent. The combination of common information to both stocks with market frictions such as illiquidity is advantageous. It leads to faster absorption of information in the price of one stock of the pair and not the other. A lead-lag relationship is the result, which can be profitably exploited.

Chen *et al.* (2012) confirm that pairs trading profitability is partly driven by delays in information diffusion across the two legs of a pair. Also, pairs trading profitability is highest in poorer information environments. Yet, contrary to Engelberg *et al.* (2009), they find no evidence that short-term liquidity provision is related to pairs trading profitability. Additionally, their strategy performs poorly during the financial crisis in 2008, a low liquidity environment with potential rewards for liquidity providing strategies. Note that Chen *et al.* (2012) are the only prominent exception in this respect, indicating that pairs selection with return correlations results in a vastly different trading strategy compared to GGR's approach.

Jacobs and Weber (2013) test a variant of the GGR algorithm on a subset of the U.S. market from 1960 to 2008 and on several international markets. They confirm that pairs trading returns are linked to different diffusion speeds of common information across the two securities forming a pair. In particular, pairs are more likely to open on so-called high-divergence days, where investor attention is primarily focused on the market level instead of individual stocks, due to a high quantity of unexpected new information per day. High distraction leads to slower diffusion of common information, creating profitable lead-lag relationships. Hence, pairs opened on such days are more likely to converge and thus more profitable. Jacobs and Weber (2015) expand on this study on a comprehensive U.S. data set and 34 international markets. They find that pairs trading returns are a persistent phenomenon. The U.S. sample reveals that profitability is positively affected by limits to arbitrage and negatively affected by increasing pair visibility. Jacobs (2015) tests 20 groups of long-short anomalies. The author finds pairs trading to be the top five anomaly on a large sample of U.S. stocks, with abnormal returns exceeding 1% per month. The strategy barely loads on investor sentiment proxies, but seems to be related to limits to arbitrage.

Huck (2013) finds on an S&P 500 sample that GGR's pairs trading returns are highly sensitive to the length of the formation period. Strong positive results are achieved with durations of 6, 18, and 24 months. Surprisingly, there is a slump in abnormal returns for the 12-month formation period of GGR. In a later study, Huck (2015) examines the impact of volatility timing on pairs trading. On an international sample of S&P 500 and Nikkei 225 constituents, he finds that pairs trading returns cannot be further improved by timing volatility.

2.5 *Further Out-of-Sample Testing of GGR's Strategy*

Nath (2003) applies pairs trading to the entire secondary market of U.S. government debt. The data stem from GovPX and range from 1994 to 2000. The author uses SSD between standardized dirty prices as selection metric in a 40-day formation period. In the subsequent 40-day trading period, positions are entered when SSD reaches certain trigger levels around the median, defined as percentiles of the empirical distribution of historical SSD. Trades are closed upon reversion to the median, at the end of the trading period, or if the stop loss percentiles are hit. The strategies outperform their benchmarks in terms of Sharpe and gain–loss ratio. Bianchi *et al.* (2009) examine GGR's strategy in commodity market futures from 1990 to 2008. They find statistically and economically significant excess returns with low exposure to systematic sources of risk. Bowen *et al.* (2010) examine GGR's strategy on the FTSE 100 constituents from January to December 2007, in a true high-frequency setting with 60-minute-binned returns. The authors use a 264-hour formation period and a 132-hour trading period. At first, they find intraday pairs trading to be profitable with low exposure to systematic risk factors and excess returns of approximately 20% p.a. However, these results are extremely sensitive to transaction costs and speed of execution. Factoring in transaction costs of 15 basis points and delaying execution by a 60-minute interval leads to full elimination of excess returns. Mori and Ziobrowski (2011) test GGR's trading rule for the U.S. stock market compared to the subset of the U.S. Real Estate Investment Trust (REIT) market 1987–2008. Over the entire sample period, REIT pairs produce higher profits at lower risk compared to common stocks. The superiority of REITs is mainly due to the high industry homogeneity, leading to more stable pairs with clear, long-term relationships. However, the effect disappears after the year 2000, either due to structural changes or due to investor recognition of pairs trading opportunities in this market. Broussard and Vaihekoski (2012) replicate GGR's algorithm in the Finnish stock market from 1987 to 2008. They confirm GGR's results for their sample while highlighting potential implementation hurdles. Bowen and Hutchinson (2014) apply GGR's strategy to the U.K. equity market from 1979 to 2012. They find statistically significant risk-adjusted excess returns that do not load on systematic risk factors.

3. Cointegration approach

3.1 *Univariate Pairs Trading*

3.1.1 *Development of a Theoretical Framework*

Vidyamurthy (2004) provides the most cited work for cointegration-based pairs trading. His design is generally ad-hoc and for practitioners, yet with many relevant insights. The framework relies on three key steps: First, pairs are preselected based on statistical or fundamental similarity measures. Second, tradability is assessed, following an adapted version of the Engle–Granger cointegration test. Third, optimal entry/exit thresholds are designed with nonparametric methods. Despite lacking empirical applications, his concept lays the groundwork for all further cointegration-based pairs trading studies.

3.1.2 *A Large-Scale Empirical Application*

Rad *et al.* (2015) provide the first large-scale empirical application of the cointegration approach on U.S. CRSP data from 1962 until 2014, following GGR and Vidyamurthy (2004). The spread $\varepsilon_{ij,t}$ between two stocks can be defined as

$$\varepsilon_{ij,t} = P_{i,t} + \gamma P_{j,t} \tag{5}$$

with P_i and P_j denoting the $I(1)$-nonstationary price processes of stocks i and j. An intercept is neglected. The cointegration coefficient γ is a nonzero real number, so that the spread $\varepsilon_{ij,t}$ as linear combination of P_i and P_j is $I(0)$-stationary, hence mean-reverting. We say the two price processes are cointegrated or assume an error correction form.

Following GGR, Rad *et al.* (2015) identify the stocks with minimum SSD in a 12-month formation period. Next, they test these stocks for cointegration with the Engle–Granger approach and only retain the top 20 stocks of the SSD ranking that are also cointegrated. Trading signals are generated with a variant of GGR's threshold rule; one USD is invested in the long leg of each pair and the dollar amount in the short leg is determined in line with the cointegration relationship in (5). Rad *et al.* (2015) find monthly excess returns of 0.83% prior to transaction costs for the cointegration approach—very similar to the 0.88% return of the distance method they run as benchmark. This lack of outperformance of the cointegration approach is possibly driven by the two-step pairs selection metric. Limiting cointegration to the subset of pairs with minimum SSD introduces a clear selection bias. In particular, following the ideas of subsection 2.1, the distance constraint reduces spread variance and the subsequent cointegration test most likely only picks pairs that would have been part of the top k distance pairs anyways. As such, the similarity in return behavior between distance and cointegration approach comes as no surprise. Going forward, we recommend an end-to-end cointegration-based selection and trading algorithm, as introduced by Caldeira and Moura (2013) or by Huck (2015) - discussed in further detail in subsection 3.1.4.

3.1.3 *A Deep-Dive on the Development of Optimal Trading Thresholds*

Lin *et al.* (2006) develop a minimum profit condition for a pair of securities that is cointegrated over the relevant time horizon as in equation (5). The cointegration coefficient γ is assumed to be less than zero on all occasions. Stock i is used for short positions and stock j for long positions. Naturally, the price of j at the opening of the trade is always lower than the price of i. For each n shares long of stock j, $n/|\gamma|$ shares of stock i are held short in one pair, i.e., the proportion of shares held is determined by the cointegration relationship. When t_o denotes opening time, t_c closing time, and we use the relation from (5), the total profit per trade TP_{ij,t_c} amounts to:

$$TP_{ij,t_c} = -\frac{n}{|\gamma|}\left(\left(\varepsilon_{ij,t_c} - P_{i,t_c}\right) - \left(\varepsilon_{ij,t_o} - P_{i,t_o}\right)\right) + \frac{n}{|\gamma|}\left(P_{i,t_o} - P_{i,t_c}\right) = \frac{n\left(\varepsilon_{ij,t_o} - \varepsilon_{ij,t_c}\right)}{|\gamma|} > K \tag{6}$$

If a trader sets the minimum required profit per trade to K and chooses the entry threshold ε_{ij,t_o} and exit threshold ε_{ij,t_c}, the number of shares n can be calculated according to (6). Lin *et al.* (2006) test this approach for different thresholds in a simulation study and on one exemplary stock pair. The concept has several weaknesses. First, the minimum profit per trade is set in absolute terms, so the profitability scaled by initial investment can become quite low, which Lin *et al.* (2006) confirm in their application. Second, the simulation study lacks diversity with only one cointegration model under review. It would be interesting to see how the trading rule performs in different cointegration settings, potentially allowing for the equilibrium relationship to shift or rupture. Third, the empirical application is limited to two stocks and a time frame of less than two years. Fourth, as Vidyamurthy (2004) points out, total profit over a trading period is a function of the number of trades and the profit per trade. Optimizing profit per trade usually does not optimize total profit over the trading horizon, since a higher minimum profit per trade leads to a lower number of trades and vice versa. The latter point is addressed in a subsequent paper by Puspaningrum *et al.* (2010). They fit an AR(1)-process to the spread $\varepsilon_{ij,t}$ of two cointegrated stocks and use an integral equation approach to numerically evaluate the estimated number of trades for any given trading threshold or minimum profit per trade. This approach allows for the optimization of total profit per trading period, but it still lacks larger empirical applications.

3.1.4 *A Review of Further Empirical Applications*

The first empirical applications are found in the domain of futures markets under the keyword "spread trading." A representative study is by Girma and Paulson (1999). They focus on the "crack spread," i.e., the price difference between petroleum futures and futures on its refined end products from 1983 to 1994. Different variants of this spread are stationary according to the augmented Dickey–Fuller (ADF) and the Phillips–Perron test. Trades are entered when the spread deviates a multiple k of its cross-sectional standard deviation from its cross-sectional moving average, both calculated over n-days and all available contract months. Positions are closed when the spread returns to its own n-day moving average. Girma and Paulson (1999) test both 5- and 10-day moving averages and five different entry thresholds. The results are promising: After consideration of USD 100 transaction costs per full turn, average return still exceeds 15% p.a.[8] Dunis *et al.* (2006c) and Cummins and Bucca (2012) confirm the profit potential for the crack spread in later years with different models. Following Girma and Paulson (1999), Simon (1999) successfully trades the crush spread, i.e., the difference between soybean futures prices and its end products. Similarly, Emery and Liu (2002) analyze the spark spread, i.e., the difference between natural gas and electricity futures prices with positive results. However, for the gold–silver spread, Wahab and Cohn (1994) find trading to be unsuccessful. This direction of research is promising, since there are fundamental reasons for the cointegration relationship between raw materials and refined end products.

Hong and Susmel (2003) are the first authors to implement a rudimentary version of the cointegration approach to stocks. They choose 64 American Depositary Receipts (ADRs) and the corresponding shares in the local markets from 1991 to 2000. Hong and Susmel (2003) assume these pairs to be cointegrated, but provide no test results. Also, they do not calculate the spread according to the cointegration relationship as in equation (5), but on a 1:1 basis. Pairs trades are entered when the spread diverges more than a fixed threshold and reversed upon return to equilibrium. Only ADRs may be shorted due to potential short-selling restrictions in local markets. Despite the methodological weaknesses in terms of cointegration testing, the results are impressive with annualized returns of 33%. However, Broumandi and Reuber (2012) note that these large returns may be driven by an appreciation of local currencies, casting doubt on the findings.

Dunis *et al.* (2010) test the univariate cointegration approach in a daily and a high-frequency setting on the constituents of the EuroStoxx 50 index. They restrict pairs formation to 10 industry groups and come up with 176 possible pairs, which may or may not be cointegrated. The spread for each pair is defined as

$$\varepsilon_{ij,t} = P_{i,t} - \gamma_t P_{j,t} \tag{7}$$

where the time-varying parameter γ_t is estimated with the Kalman Filter. Next, the spreads of all pairs are calculated with equation (7), standardized, and then traded according to a simple standard deviation logic similar to GGR. However, if pairs formation simply relies on industry classification, the out-of-sample results are not very convincing. Therefore, Dunis *et al.* (2010) test the relationship between several in-sample indicators and out-of-sample information ratios. They find that in-sample *t*-statistics of the ADF test as part of the Engle–Granger cointegration test and in-sample information ratios seem to have certain predictive power for out-of-sample information ratios. Much better out-of-sample information ratios are realized when trading only the top five pairs preselected by one or both of these indicators. Miao (2014) confirms these findings when applying a similar strategy to high-frequency data.

Caldeira and Moura (2013) apply the univariate cointegration approach to the 50 most liquid stocks of the Brazilian stock index IBovespa. They use the Engle–Granger approach as well as the Johansen method at the 5% significance level to test for cointegration relationships for all 1225 combinations of pairs over a 12-month formation period. On average, they find 90 cointegrated pairs. Following Dunis *et al.* (2010), these pairs are ranked according to the in-sample Sharpe ratio. The top 20 pairs are selected for a four-month trading period. Positions are opened and closed based on a modified standard deviation rule similar to GGR. Caldeira and Moura (2013) show statistically significant excess returns after consideration of transaction costs. Their findings are robust to data snooping based on the reality check of White (2000) and the SPA test of Hansen (2005). These results are definitely convincing, but one improvement is key for future studies. Currently, no control for the multiple comparisons problem is implemented. Clearly, the Engle–Granger and the Johansen tests are not statistically independent when used on the same data set. In contrast, in most cases, they probably yield the same results. For simplicity's sake, let us assume the latter were the case. Cointegration testing on 1225 pairs thus produces 61 cointegrated pairs as false positives in expectation, at the 5% significance level the authors used. Even though this estimate is aggressive, Caldeira and Moura (2013) most likely have a significant share of false positives in their rankings. However, the subsequent heuristic of filtering the pairs by in-sample Sharpe ratio most likely improves tradability. Nevertheless, it would be an improvement to actively control for familywise error rate as in Cummins and Bucca (2012).

Gutierrez and Tse (2011) provide an appealing conceptual framework that is applied to a set of three cointegrated stocks. For each pair, the Granger-leaders and Granger-followers are identified. When a classical pairs trading strategy is applied, the majority of profitability stems from the Granger-follower, whereas the Granger-leader barely contributes. Clearly, the results are not representative, but the concept deserves rigorous testing on a larger sample.

Finally, there is a set of further applications: Li *et al.* (2014) show that cointegration-based pairs trading is profitable in the Chinese AH-share markets. Baronyan *et al.* (2010) evaluate a set of 14 market-neutral trading strategies on the constituents of the Dow Jones Industrial Average. Similar to Gutierrez and Tse (2011), they find improved performance in case Granger causality testing is taken into account. Huck and Afawubo (2015) also run a comparison study. They analyze the cointegration approach and the distance approach for the S&P 500 constituents and under varying parameterizations. After consideration of risk loadings and transaction costs, they find that the cointegration approach significantly outperforms the distance method. The latter corroborates the hypothesis that the cointegration approach identifies econometrically more sound equilibrium relationships. Bogomolov (2011) arrives at a similar conclusion for the Australian stock market.

3.2 *Multivariate Cointegration Approach*

3.2.1 *Passive Index Tracking and Enhanced Indexation Strategies*

Dunis and Ho (2005) pursue two objectives. First, the authors use cointegration relationships to construct index tracking portfolios for the EuroStoxx 50 index. More specifically, Dunis and Ho (2005) take different subsets (5, 10, 15, or 20 stocks) of the index constituents and estimate the joint cointegration vector for these constituents and their index. Then, they measure the tracking error return of this basket versus the index for different rebalancing frequencies. They find that the tracking baskets produce a positive tracking error, resulting in an outperformance versus the benchmark. Second, the authors develop an advanced indexation strategy. Such an approach is characterized by creating tracking baskets for artificial benchmarks. A synthetic “plus” benchmark is constructed by adding uniformly distributed returns of z percent p.a. to the daily returns of the EuroStoxx 50. Analogously, a “minus” benchmark can be created. Next, the Johansen procedure is used to find adequate securities among the 50 index constituents to track these benchmarks. According to Dunis and Ho (2005), going long the “plus” benchmark and short the “minus” benchmark allows for a market neutral investment strategy with potential “double alpha.” The authors find significant outperformance compared to the EuroStoxx 50 index. Alexander (1999), Alexander (2001), and Alexander and Dimitriu (2005) develop very similar strategies. However, all of the above exhibit one key issue. Whereas the index tracking strategy relies on a “natural” cointegration relationship between the index and its constituents, it is troublesome to make such an assumption for the artificial benchmark. Alexander and Dimitriu (2005) also follow this approach, even though Alexander (1999) herself shows in a powerful example, that adding a miniscule daily incremental return to one of two cointegrated time series may break up the entire cointegration relationship.

3.2.2 *Active Statistical Arbitrage Strategies*

In the previous section, tracking and enhanced indexation strategies have been discussed. These are passive strategies, since they are closely tied to an underlying index as benchmark (Galenko *et al.*, 2012). In contrast, Galenko *et al.* (2012) develop an active statistical arbitrage strategy based on a multivariate cointegration framework. Whereas correlation reflects short-term linear dependence in returns, cointegration models long-term dependencies in prices (Alexander, 2001). As such, compared to the distance methods from Section 2, the cointegration approach has a higher potential of identifying true long-term equilibrium relationships between several assets. Their framework heavily relies on properties they develop for the return process of cointegrated assets. The latter is defined as weighted sum of asset returns, where the weighting scheme is designed according to the components of the cointegration vector. The authors are able to show that this weighted return process is mean-reverting under certain conditions. A trading strategy capitalizing on this mean-reversion effect is shown to have positive expected profits. Empirical applications on several index exchange traded funds (ETFs) result in an outperformance versus the benchmark. However, Galenko *et al.* (2012) perform extensive data mining. First, they apply their strategy to daily and weekly data. Second, for each of these setups, they take a different duration of the formation period to estimate the cointegration vector. Finally, they test nine different lag parameters, over which the return process is cumulated to a price time series. The latter is troublesome, considering that this parameter should be infinity based on their theoretical model (Galenko *et al.*, 2012, p. 94). It is thus unclear, why they also experiment with short-term values, such as five days. Also, excess returns are not tested for statistical significance.

3.3 *Adjacent Developments*

Burgess (1999) develops a holistic statistical arbitrage framework relying on a combination of cointegration and emerging techniques such as neural networks and genetic algorithms. Burgess (2003) presents a simplified variant of this approach, solely relying on cointegration testing. D’Aspremont (2011) uses canonical correlation analysis to construct mean-reverting portfolios with a limited number of assets. Karakas (2009) applies fractional cointegration to dual-class firms and Liu and Chou (2003) to gold and silver markets. Peters *et al.* (2011), Gatarek *et al.* (2011), and Gatarek *et al.* (2014) use Bayesian procedures for cointegration testing and apply them to a very limited set of securities. Finally, Cheng *et al.* (2011) develop a statistical arbitrage strategy with a cointegration model based on logistic mixture autoregressive equilibrium errors.

4. Time-Series Approach

4.1 *Modeling the Spread in State Space*

Elliott *et al.* (2005) are the most cited authors in this domain. They explicitly describe the spread with a mean-reverting Gaussian Markov chain, observed in Gaussian noise. The latter can be achieved with a state space model, consisting of a state and a measurement equation. We will briefly present their approach, starting with the *state equation*: It is assumed that the latent state variable x_k follows a mean-reverting process:

$$x_{k+1} - x_k = (a - bx_k)\,\tau + \sigma\sqrt{\tau}\varepsilon_{k+1} \tag{8}$$

Thereby, $a, \sigma \in \mathcal{R}_0^+, b \in \mathcal{R}^+$, and $\varepsilon_k \overset{iid}{\sim} \mathcal{N}(0, 1)$. Time $t_k = k\tau$ for $k = 0, 1, 2, ...$ is discrete. This process reverts to its mean $\mu = a/b$ with mean-reversion strength b. It can also be written as:

$$x_{k+1} = A + Bx_k + C\varepsilon_{k+1} \tag{9}$$

where $A = a\tau$, $B = 1 - b\tau$, and $C = \sigma\sqrt{\tau}$. In continuous time, it is possible to describe the state process with the well-known Ornstein–Uhlenbeck (OU) process:

$$dX_t = \rho\,(\mu - X_t)\,dt + \sigma dW_t \tag{10}$$

where dW_t is a standard Brownian motion defined on some probability space. The parameter $\mu = a/b$ denotes the mean and $\rho = b$ describes the speed of mean-reversion. The second component to a state space model is the *measurement equation*: Here, the observed spread is defined as the sum of the state variable x_k and some Gaussian noise $\omega_k \overset{iid}{\sim} \mathcal{N}(0, 1)$:

$$y_k = x_k + D\omega_k, \quad D > 0 \tag{11}$$

According to this model, a pairs trade is entered when $y_k \geq \mu + c(\sigma/\sqrt{2\rho})$, or when $y_k \leq \mu - c(\sigma/\sqrt{2\rho})$. Thereby, c denotes a fixed parameter, for which Elliott *et al.* (2005) give no guidance on how to determine it. The position is reversed at time $\mathcal{T}$, denoting the first passage time result for the Ornstein–Uhlenbeck process.

Following Do *et al.* (2006), this approach has three advantages: First, the model is fully tractable, meaning that its parameters can be estimated using the Kalman Filter and the state space model. The estimator is based on maximum likelihood and optimal in terms of minimum mean squared error. There are well-known implementations at hand; Elliott *et al.* (2005) use the expectation maximization (EM) algorithm. Second, the continuous time model can be exploited for forecasting purposes. Critical questions about pairs trading, such as expected holding times and expected returns, can be answered explicitly, provided that the spread really follows this rigid model. Third, the approach is fundamentally based on mean-reversion.

However, Do *et al.* (2006) also criticize the model of Elliott *et al.* (2005). Essentially, they point out that this rigid model is only applicable to securities in return parity, which is rarely observed in practice. Exceptions are dual-listed companies or cross-listings, limiting the applicability to a small subset of securities. This assertion is basically valid, but variants of the concept are successfully applied to other securities, as in Avellaneda and Lee (2010). Further critique is provided by Cummins and Bucca (2012): A major limitation lies in the Gaussian nature of the OU-process, which is in conflict with the stylized facts of financial data. However, this disadvantage is largely compensated by analytic simplicity. Thus, the concept constitutes a valuable asset to pairs trading research and potentially a true improvement compared to nonparametric trading rules.

Following Elliott *et al.* (2005), Do *et al.* (2006) develop a pairs trading methodology that models mispricing at the return instead of the price level. Their proposed state space model can be summarized as follows (Puspaningrum, 2012, p. 27):

$$x_{k+1} = A + Bx_k + C\varepsilon_{k+1} \tag{12}$$

$$y_k = x_k + \Gamma U_k + D\omega_k \tag{13}$$

This representation is very similar to equations (9) and (11) of Elliott *et al.* (2005). The first difference is that the observed spread y_k is defined as the difference of asset returns of the two stocks of a pair. The second adjustment consists of the loading matrix Γ and the variable U_k, which are exogenous inputs stemming from arbitrage pricing theory (APT). Essentially, Do *et al.* (2006) use APT to generate a fundamental justification for their pairs trading framework.[9] Similar to Elliott *et al.* (2005), the authors discuss different estimation procedures and also opt for the EM algorithm. After successful model calibration, a long-short position is taken whenever the accumulated spread over a given time frame exceeds certain threshold values. However, these thresholds and the expected holding time are not further specified and have to be evaluated for potential implementation. The same applies for the time frame over which they suggest to accumulate the residual spread to detect deviations from equilibrium. Here, it is insightful to draw parallels to Galenko *et al.* (2012), suggesting an infinite time frame to accumulate returns. If the pairs of Do *et al.* (2006) were cointegrated in the sense of Galenko *et al.* (2012), this notion is also applicable here. The authors demonstrate their strategy in a simulation study and in a small empirical application to selected securities. However, as Puspaningrum (2012) points out, they use a down-sized version of their model, relying on one-factor Capital Asset Pricing Model (CAPM) instead of multiple factor APT. Applying the fully fledged version on a large data set may thus yield further insights.

Triantafyllopoulos and Montana (2011) also build on the model of Elliott *et al.* (2005) and enhance it in two key respects: First, they introduce time dependency in the parameters, thus improving flexibility. Second, the authors replace the EM algorithm with a Bayesian estimation framework. The latter is useful for high-frequency applications due to faster convergence.

4.2 *Applications of the Ornstein–Uhlenbeck Process*

Bertram (2010) develops a statistical arbitrage model for a spread X_t between two log price time series. The latter is assumed to follow a zero-mean, symmetric OU-process:

$$dX_t = -\rho X_t dt + \sigma dW_t \tag{14}$$

A trade cycle is defined as follows: Let a be the threshold to enter and m the threshold to exit a trade. Thereby, we assume that $a < m$, so we go long the spread at level a and reverse the positions at level m. The cycle time $\mathcal{T}$ can be split up in two subperiods:

$$\mathcal{T} = \mathcal{T}_1 + \mathcal{T}_2 \tag{15}$$

where $\mathcal{T}_1$ is the time the process takes to transition from entry to exit and $\mathcal{T}_2$ the time from exit to a subsequent entry. The times $\mathcal{T}_1$ and $\mathcal{T}_2$ are independent, due to the Markovian property of the OU-process. Considering relative transaction costs c, the total log return per trade cycle can be defined as a function of the trigger thresholds and the transaction costs: $r(a, m, c) = m - a - c$. Since the OU-process is stationary, this return is deterministic. In contrast, the associated cycle time is stochastic. With trade frequency following a renewal process, Bertram (2010) uses renewal theory to derive the following expressions for expected return and variance per unit time:

$$\mu(a, m, c) = \frac{r(a, m, c)}{\mathbb{E}[\mathcal{T}]} \tag{16}$$

$$\sigma^2(a, m, c) = \frac{r^2(a, m, c)\mathbb{V}[\mathcal{T}]}{\mathbb{E}^3[\mathcal{T}]} \tag{17}$$

Next, Itô's lemma is used to transform the OU-process to a dimensionless system in order to simplify the analysis. Leveraging first passage time theory about the OU-process finally allows Bertram (2010) to derive analytic expressions for the expected trade length and its variance. With the help of these formulas, closed-form solutions for the expected return and the Sharpe ratio of the strategy are developed, both per unit time. Applying straightforward optimization routines to these equations results in optimal entry a^* and exit thresholds m^*, corresponding to maximum return or Sharpe ratio. As Bertram (2010) himself points out, the downside of this approach lies in a Gaussian OU-process being applied to non-Gaussian financial data. On the other hand, the upside lies in the availability of closed-form solutions. The latter allows for analytic investigations of the spread dynamics and for implementations in high-frequency settings, requiring computationally efficient solutions. Bertram (2010) empirically applies his strategy to a pair of dual-listed securities. Cummins and Bucca (2012) perform a large-scale implementation on 861 energy futures spreads from 2003 until 2010. After rigorously controlling for data snooping bias following procedures of Romano and Wolf (2007) as well as Romano *et al.* (2010), they find that the daily returns of their top strategies range between 0.07% and 0.55%. The corresponding Sharpe ratios are often larger than two. This study is striking evidence for the profit potential of this approach and it should be applied to other asset classes.[10]

4.3 *Further Concepts from Time-Series Analysis*

The field of time-series analysis is vast and equally numerous are the methods that could potentially be applied to develop trading systems for mean-reverting spreads. Bock and Mestel (2009) use a Markov switching model to develop a pairs trading framework. Kanamura *et al.* (2010) derive a profit model for spread trading based on a mean-reverting process. They empirically apply their strategy to the energy futures market. Bogomolov (2013) develops an innovative nonparametric approach for pairs trading originating from Japanese charting indicators. Chen *et al.* (2014) construct a pairs trading strategy with three-regime threshold autoregressive models with GARCH effects.

5. Stochastic Control Approach

5.1 *Modeling Asset Pricing Dynamics with the Ornstein-Uhlenbeck Process*

Jurek and Yang (2007) provide the paper with the highest impact in this domain.[11] In their setup, they allow nonmyopic arbitrageurs to allocate capital to a mean-reverting spread or a risk-free asset. The former follows an OU-process and the latter is compounded continuously with the risk-free rate. Two scenarios for investor preferences are considered over a finite time horizon: constant relative risk aversion and the recursive Epstein–Zin utility function. Utilizing the asset price dynamics, Jurek and Yang (2007) develop budget constraints and wealth dynamics for the arbitrageurs' assets. Applying stochastic control theory, the authors are able to derive the Hamilton–Jacobi–Bellmann (HJB) equation and subsequently find closed-form solutions for the value and policy functions, allowing for three contributions: First, the OU-process captures the uncertainties of an arbitrage opportunity in the form of horizon and divergence risk. This is a novelty compared to existing models in this domain, such as the Brownian Bridge. Second, the incorporation of different utility functions as opposed to log utility makes it possible to split the optimal policy function in two demand components: Myopic demand is a short-term component and exclusively oriented on the current magnitude of the mispricing. Conversely, intertemporal hedging demand addresses the investor's need for a hedge against the risk stemming from the state variables. Jurek and Yang (2007) show that intertemporal hedging demand can explain a significant part of the allocation to the arbitrage opportunity. Third, arbitrageurs do not always perform arbitrage. In line with the numerical findings of Xiong (2001), Jurek and Yang (2007) analytically identify the boundaries of a "stabilization region," in which the arbitrageur trades against divergences of the spread. On the outside, he decreases his positions to avoid negative wealth effects. The authors apply their strategy in a simulation study and to a pair of stocks. In comparison to the simple threshold rule, the optimal strategy shows significant outperformance in case of highly mean-reverting spreads. The effect is less pronounced in case of slow mean-reversion and when estimation errors are assumed. Also, no market frictions are considered. In reality, daily rebalancing would result in substantial transaction costs compared to GGR's relatively passive trading rule.

5.2 *Modeling Asset Pricing Dynamics with Error Correction Models*

Liu and Timmermann (2013) build on the results of Jurek and Yang (2007). They also derive optimal portfolio holdings for convergence trades under recurring and nonrecurring arbitrage opportunities for an investor with power utility over terminal wealth. The authors use a cointegration framework for the asset price dynamics and allow for nondelta-neutral positions in a pairs trade. In their model, the market index $P_{m,t}$ evolves according to a geometric random walk:

$$\frac{dP_{m,t}}{P_{m,t}} = (r + \mu_m)\,dt + \sigma_m dB_t \tag{18}$$

with market risk premium μ_m, market volatility σ_m, the risk-free rate r, and a standard Brownian motion B_t. Moreover, the prices $P_{1,t}$ and $P_{2,t}$ of two risky assets exhibit the following dynamics:

$$\frac{dP_{1,t}}{P_{1,t}} = (r + \beta\mu_m)\,dt + \beta\sigma_m dB_t + \sigma dZ_t + b dZ_{1,t} - \lambda_1 X_t dt \tag{19}$$

$$\frac{dP_{2,t}}{P_{2,t}} = (r + \beta\mu_m)\,dt + \beta\sigma_m dB_t + \sigma dZ_t + b dZ_{2,t} + \lambda_2 X_t dt \tag{20}$$

$$X_t = \ln(P_{1,t}) - \ln(P_{2,t}) \tag{21}$$

Thereby, λ_i, β, b, and σ are constants, and Z_t, $Z_{i,t}$ are mutually independent standard Brownian motions for $i = 1, 2$, and X_t is the error term. Further, the sum of λ_1 and λ_2 is assumed to be greater than zero, so X_t is stationary and the log prices are cointegrated. The investor has the choice of allocating his funds to the market portfolio and to the risky assets with shares ϕ_m, ϕ_1, and ϕ_2, respectively. The authors follow Jurek and Yang (2007) and derive the HJB equation for an investor under power utility over terminal wealth. They find the value and optimal policy functions for this stochastic control problem, providing the optimal portfolio weights. By examining the arbitrage opportunity in this portfolio maximization context, the strategy not only incorporates the arbitrage opportunity, but also diversification benefits—resulting in two new relevant insights: First, it can be optimal to hold both risky assets long (or short) at the same time, even if prices eventually converge. Second, it can also be optimal to only hold one of the two assets. This optimal investment policy is in stark contrast to standard delta-neutral long-short pairs trades. Also, the work of Jurek and Yang (2007) did not allow for nondelta-neutral positionings. Liu and Timmermann (2013) empirically compare the optimal unconstrained with the delta-neutral strategy on a set of Chinese banking stocks. Their findings are in line with their theories—the unconstrained strategy can result in economically significant gains over the standard arbitrage strategy. Lei and Xu (2015) expand on the methodology of Liu and Timmermann (2013) and include transaction costs, thereby significantly affecting the optimal policy of the arbitrageur.[12] It would be interesting to see how such strategies fare in large-scale empirical applications.

6. Other Approaches

6.1 *Machine Learning and Combined Forecasts Approach*

The main studies in this domain are Huck (2009) and Huck (2010), respectively. The methodology is based on three steps: forecasting, outranking, and trading. In the forecasting step, Huck (2009) uses Elman neural networks to generate one-week ahead return forecasts $\hat{x}_{i,T+1}|\mathcal{F}_{ij,T}$ for each security i, conditional to the past return information $\mathcal{F}_{ij,T}$ of securities i and j, with $i, j \in \{1, ..., n\}$. Thus, in total, $n - 1$ return forecasts are generated per period for each security i. In the outranking step, Huck (2009) uses a multicriteria decision method (MCDM) called ELECTRE III. This method ranks a set of alternatives according to a set of criteria. Here, the n stocks represent the alternatives as well as the criteria. The performance $\hat{x}_{ij,T+1}$ of each stock i to criterion j is calculated as

$$\hat{x}_{ij,T+1} = \hat{x}_{i,T+1}|\mathcal{F}_{ij,T} - \hat{x}_{j,T+1}|\mathcal{F}_{ij,T} \tag{22}$$

Thus, the performance is the anticipated spread, i.e., the difference in return forecasts of securities i and j, conditional to their past information. These values are collected in an antisymmetric $n \times n$ matrix. Its rows correspond to the n alternatives and its columns to the n criteria. In each cell, we find the anticipated spread of stock i versus criterion j. With ELECTRE III, an outranking is created, so that undervalued stocks are at the top and overvalued stocks at the bottom. In the trading step, the top stocks of the ranking are bought and the bottom stocks sold short. After a trading period of one week, the positions are closed, a new ranking is created, and the process repeated. It is important to notice that these pairs do not share any kind of equilibrium model. Instead, trading actions are simply triggered based on the position in the final ranking. An empirical application on the S&P 100 constituents from 1992 to 2006 produces impressive results: Buying the top five stocks and selling short the bottom five stocks of the ranking lead to a 54% forecasting accuracy and more than 0.8% weekly excess returns. However, these findings should be handled with care. First, Huck (2009) did not eliminate the survivor bias. Second, the value-add of the relatively complex MCDM needs further investigation. Normally, the key advantages of ELECTRE III are its fuzzy logic to account for uncertainties and its ability to outrank alternatives across criteria denoted in different units. In the present application, all performance values are return differences of the same dimension. A small increase in anticipated spread relative to one criteria is clearly better from an economic perspective. Thus, a rational investor would buy (sell short) the stocks with the highest (lowest) anticipated spreads. Hence, the ELECTRE III trading results should be compared to a simpler ranking to prove its superiority. Nevertheless, this equilibrium-free approach constitutes a promising direction of further research, since it defines a completely new direction for pairs trading.[13]

6.2 *Copula Approach*

Two substreams can be found in the literature, return-based versus level-based copula methods.

Return-based copula method: Representatives of this category are Ferreira (2008), Liew and Wu (2013), Stander *et al.* (2013), and Krauss and Stübinger (2015). In a formation period, pairs are built based on previously discussed correlation or cointegration criteria. Next, log returns $R_i = (R_{i,t})_{t \in T}$ and $R_j = (R_{j,t})_{t \in T}$ for the two legs i and j of each pair are considered as well as the marginal distributions F_{R_i} and F_{R_j}. Stander *et al.* (2013) discuss parametric and nonparametric approaches to obtain the marginal distributions, and Ferreira (2008) and Liew and Wu (2013) opt for fitting parametric distribution functions. Applying probability integral transform by plugging the returns into their own distribution functions creates two uniform variables $U_i = F_{R_i}(R_i)$ and $U_j = F_{R_j}(R_j)$. Now, an adequate copula function can be identified. Ferreira (2008) just uses one particular copula, and Stander *et al.* (2013) rely on a set of 22 different Archimedean copulas and determine the best-fitting one with the

Kolmogorov–Smirnov goodness-of-fit test. Liew and Wu (2013) start with five copulas most commonly used in financial applications and determine the best-fitting one with information criteria. The trading strategy is described following Stander *et al.* (2013) and Liew and Wu (2013). They use the best-fitting copula to calculate the conditional marginal distribution functions as first partial derivatives of the copula function $C(u_i, u_j)$:

$$h_i\left(u_i|u_j\right) = P\left(U_i \leq u_i|U_j = u_j\right) = \frac{\partial C\left(u_i, u_j\right)}{\partial u_j} \tag{23}$$

$$h_j\left(u_j|u_i\right) = P\left(U_j \leq u_j|U_i = u_i\right) = \frac{\partial C\left(u_i, u_j\right)}{\partial u_i}$$

If the conditional probability is greater (less) than 0.5, a stock can be considered relatively overvalued (undervalued). The authors suggest to trade when the conditional probabilities are well in the tail regions of their conditional distribution functions, i.e., below their 5% and above their 95% confidence level. To be specific, stock i is bought and stock j is sold short when their transformed returns fall outside both confidence bands derived by $P(U_i \leq u_i|U_j = u_j) = 0.05$ and $P(U_j \leq u_j|U_i = u_i) = 0.95$. Informally speaking, this corresponds to the extreme regions in the northwest quadrant of a scatter plot of u_i and u_j. Conversely, stock j is bought and stock i is sold short, when inverse conditions apply (extreme regions in the southeast quadrant). Stander *et al.* (2013) suggest to exit a trade as soon as it is profitable or after one week. Liew and Wu (2013) reverse their positions once the conditional probabilities cross the boundary of 0.5 again. The application of copulas to pairs trading research is promising. Stylized facts such as negative skewness and excess kurtosis are regularly observed. Copulas are an adequate way of modeling such complex dependence structures and thus have the potential to identify better trading opportunities. The key issues inherent to the approaches of Ferreira (2008), Liew and Wu (2013), and Stander *et al.* (2013) are the loss of time structure and the lack of copula-based pairs selection.

Initial solutions are proposed by Krauss and Stübinger (2015), who develop an integrated copula-based pairs trading framework and apply it to the S&P 100 constituents from 1990 until 2014. The authors use copulas for selection and trading. First, in a 12-month formation period, t-copulas are fitted to all pairs. Next, in a 48-month in-sample pseudotrading period, the profitability of copula-based mispricing signals derived from equation (23) is assessed. The most profitable pairs fulfilling minimum dependence criteria are considered in a 12-month trading period. In particular, a differentiation is made between pairs exhibiting mean-reversion and momentum effects and idiosyncratic take-profit and stop-loss rules are applied to each pair. The study shows promising results, with out-of-sample returns of almost 8% per year and Sharpe ratios above 1.5.

Level-based copula method: Xie and Wu (2013), Xie *et al.* (2014), and Rad *et al.* (2015) develop the level-based copula method. We briefly discuss the approach of Rad *et al.* (2015), reflecting all recent developments. First, during a formation period, the authors identify the 20 most similar pairs based on minimum SSD. Second, they fit marginal distribution functions to the return series of each nominated pair, choosing the best-fitting one among extreme value, generalized extreme value, logistic, or normal distribution. Third, uniformly distributed variables are created by feeding the returns into their own marginals and a copula is fitted for each pair. The best-fitting one among Clayton, rotated Clayton, Gumbel, and rotated Gumbel copulas based on AIC and BIC is selected. Fourth, a mispricing index is constructed, justifying the name "level-based copula method." Essentially, the authors calculate the conditional probabilities $h_i(u_{i,t}|u_{j,t})$, $h_j(u_{j,t}|u_{i,t})$ as time series with $t \in T$, which are transformed into mispricing incides $m_{i,t}$ and $m_{j,t}$ as follows:

$$m_{i,t} = h_i\left(u_{i,t}|u_{j,t}\right) - 0.5, \quad m_{j,t} = h_j\left(u_{j,t}|u_{i,t}\right) - 0.5 \tag{24}$$

According to Rad *et al.* (2015), positive values of $m_{i,t}$ and negative values of $m_{j,t}$ are interpreted as stock i being overvalued with respect to stock j at time t, and vice versa. Next, cumulative mispricing indices $cm_{i,t}$ and $cm_{j,t}$ are calculated as

$$cm_{i,t} = cm_{i,t-1} + m_{i,t}, \quad cm_{j,t} = cm_{j,t-1} + m_{j,t}, \quad t \in T \tag{25}$$

with $cm_{i,0} = cm_{j,0} = 0$. Essentially, the cumulative mispricing index $cm_{i,t}$ indicates the level of mispricing of stock i, given the information of stock j, aggregated over several periods. The interpretation of $cm_{j,t}$ follows in analogy. Positive values of $cm_{i,t}$ and negative values of $cm_{j,t}$ are interpreted as stock i being overvalued with respect to stock j, taking into account all past realizations of the respective mispricing indices. The interpretation for stock j follows in analogy. Xie *et al.* (2014) discuss the time-series properties of these mispricing indices and the conditions under which it may be mean-reverting. Rad *et al.* (2015) open a short position in stock i and a long position in stock j at time t if $cm_{i,t} > 0.4$ and simultaneously $cm_{j,t} < -0.4$ and vice versa. The position is unwound when both cumulative mispricing indices revert to zero. This approach is appealing, as it reflects multiperiod mispricings and thus retains the time structure. However, there are a few downsides associated with it: First, pairs selection is not copula-based, but the top 20 pairs with minimum sum of squared distances are chosen, introducing a severe selection bias. Along these lines, the copula method in Rad *et al.* (2015) is merely an alternative way of determining entry and exit signals. Second, the authors rely on the theoretical framework of Xie *et al.* (2014), who derive the conditions under which the cumulative mispricing indices are mean-reverting. However, Rad *et al.* (2015) do not test if the cumulative mispricing indices really are mean-reverting—a quintessential prerequisite for profitable pairs trading. Third, no differentiation is made between mean-reversion and momentum pairs, as in Krauss and Stübinger (2015).

6.3 *Principal Components Analysis Approach*

Avellaneda and Lee (2010) develop a statistical arbitrage strategy for the U.S. equity market and apply it to stocks exceeding USD 1 billion in market capitalization at the time of trading. In their formation period, they use two approaches to decompose stock returns into systematic and idiosyncratic components. In the first approach, Avellaneda and Lee (2010) regress the returns R_i of each stock i on the return F of its corresponding sector ETF:

$$R_i = \beta_i F + \varepsilon_i \tag{26}$$

so $\beta_i F$ denotes the systematic component of the portfolio (β_i is the factor loading and F the factor return). Conversely, ε_i represents the idiosyncratic component. In the second approach, a multifactor model with m factors is considered:

$$R_i = \sum_{j=1}^{m} \beta_{ij} F_j + \varepsilon_i \tag{27}$$

Avellaneda and Lee (2010) use PCA to create m eigenportfolios in line with this statistical factor model. Next, they develop a relative value model for equity valuation. Based on the multifactor model above, it is assumed that stock returns satisfy the differential equation

$$\frac{dP_{i,t}}{P_{i,t}} = \mu_i dt + \sum_{j=1}^{m} \beta_{ij} \frac{dI_{j,t}}{I_{j,t}} + dX_{i,t} \tag{28}$$

whereby μ_i represents stock price drift and the residual $X_{i,t}$ is assumed to follow an OU-process. These two components correspond to the idiosyncratic returns of stock i. The remaining summand represents the systematic returns, stemming from the corresponding sector ETF ($m = 1$) or from the statistical factor model ($m > 1$). In the trading period, Avellaneda and Lee (2010) further consider the idiosyncratic returns of equation (28) and apply a trading model similar to Elliott *et al.* (2005). The results are impressive with annualized Sharpe ratios of 1.44 from 1997 to 2007 for the PCA-based strategies and 1.1 for the ETF-based strategies. Going forward, the following improvements could be considered: First, the results are not robust to data mining, especially since Avellaneda and Lee (2010) experiment with different entry and exit thresholds for trading. An approach as in Cummins and Bucca (2012) is suggested. Second, as the authors point out, "there are considerably more entries in the correlation matrix than data points" when performing PCA (Avellaneda and Lee, 2010, p. 764). Considering asymptotic PCA instead as suggested in Tsay (2010) could be an adequate solution. Third, it may be beneficial to use a cointegration framework instead of PCA. PCA delivers a limited set of series that can be used to approximate a much larger one, i.e., a small number of eigenportfolios are used to represent the systematic risk of the entire stock universe. Conversely, "cointegration gives all possible stationary linear combinations of a set of random walks" (Alexander, 2001, p. 353). Since Avellaneda and Lee (2010) aim for stationary residuals, a cointegration framework is considered more appropriate.

Montana *et al.* (2009) develop an adjacent approach for the S&P 500, relying on dimensionality reduction via PCA and flexible least squares.

7. Pairs Trading in the Light of Market Frictions

In this section, we focus on selected studies featuring research on pairs trading in the light of market frictions, i.e., trading costs, delays in implementation, short-selling restrictions, and liquidity.

Gatev *et al.* (2006) already perform a set of robustness checks. First, they address the issue of bid-ask bounce (Jegadeesh, 1990). Pairs trading sells short winners and buys losers. After divergence, the price of the short leg is more likely to be an ask quote and the price of the long leg a bid quote. Implicitly, upon divergence, GGR are thus selling at ask quotes and buying at bid quotes. For convergence, the opposite holds, so results may systematically be biased upwards. To alleviate this effect, GGR introduce a one-day-waiting rule after each trading signal, reducing average excess returns by 0.54% per month for the top 20 pairs. Second, GGR provide a rough estimation for transaction costs. The 0.54% reduction in monthly excess returns constitutes an estimate for the average bid-ask spread. Provided a six-month trading period and two round-trip trades per pair, GGR estimate transaction costs of 3.24% for two round-trips (effective spread of 0.81%) compared to profits of 5.49%, leading to statistically and economically significant excess returns of 2.25% after transaction costs. Third, the authors consider explicit short-selling costs in the spirit of D'Avolio (2002). These explicit costs in the form of specials have a minimal effect on large stocks, so GGR apply pairs trading to the largest three size deciles with basically unchanged profitability. Also, they simulate implicit short-selling costs in the form of recalls, leading to a slight reduction in pairs trading profitability. Overall, the authors regard pairs trading profits as robust in the light of market frictions.

Do and Faff (2012) analyze the impact of trading costs on pairs trading profitability. In particular, they account for commissions, market impact, and short-selling fees. First, institutional commissions are approximated following Jones (2002). The authors consider time-varying commissions, with average one-way costs of 0.34% over the full sample period. Second, market impact or slippage can be defined as difference between expected price and price actually paid (Zhang and Zhang, 2008). Do and Faff (2012) perform an estimation by monitoring the spread at divergence and for the two subsequent days. On average, it narrows by 0.26% on the first day and by an additional 0.12% on the second day after divergence. The authors assume that an institutional trader can achieve a volume-weighted average price equal to the closing price on day 1, leading to an average one-way market impact of 0.26%. If an investor faces slower execution, slippage may be even higher. In summary, the authors consider average one-way costs of 0.60%. Since pairs trading consists of two round-trips, the impact is significant. Third, following D'Avolio (2002), a short-selling fee of 1% p.a., payable over the duration of each trade, is applied. Considering all three cost dimensions, pairs trading in the sense of GGR becomes unprofitable.

Only refined portfolios still achieve positive excess returns, albeit at clearly diminished levels of 0.30% per month. It is interesting to note that these portfolios are constructed with refined selection metrics, i.e., combinations of low SSD with a high number of zero-crossings and intraindustry pair formation. Particularly, the second SSD vigintile performs well compared to the first SSD vigintile, indicating that trading-excessive pairs at low standard deviations of the spread become unprofitable in the light of market frictions. Finally, Do and Faff (2012) show that pairs trading profits are negatively related to liquidity. Accordingly, the authors find profitability to be highest during bear markets, i.e., the dot-com bust and the global financial crisis. According to Khandani and Lo (2007, 2011), the latter constitutes a low-liquidity environment, essentially caused by the massive unwinding of factor-based long–short portfolios. Pairs trading as mean-reverting, liquidity providing strategy usually benefits from such environments. However, as Khandani and Lo (2011) show, this is only the case if levered positions can be held until potential convergence.

Bowen and Hutchinson (2014) analyze pairs trading profitability in the U.K. from 1979 to 2012. They first consider transaction costs following the one-day waiting strategy and find effective spreads of 0.35%. The authors cross-validate their results following Roll (1984). Further elements of trading costs, as in Do and Faff (2012), are not considered. Equally, the authors provide evidence for significant outperformance of pairs trading during the global financial crisis and show that the strategy positively relates to illiquidity measures. After transaction costs and when controlling for time-varying risk exposures, Bowen and Hutchinson (2014) show that risk-adjusted returns are no longer significantly different from zero.

Jacobs and Weber (2015) provide a comprehensive study on international markets and the U.S. stock universe. They show that pairs trading profits are strongly related to limits to arbitrage, proxied by volatility, bid-ask spreads, firm size, and liquidity. Generally speaking, increasing impediments to arbitrage have a positive impact on pairs trading profitability. In a cross-country analysis of pairs trading profitability, the authors provide evidence that returns are higher in emerging markets and markets with large stock universes. Further results confirm that limits to arbitrage are yet again related to profitability in the cross section. Surprisingly, short-selling restrictions lead to higher pairs trading profitability, albeit not statistically significant. This effect may be caused by collinearities, as markets prohibiting short-selling generally exhibit higher limits to arbitrage. Finally, an analysis of more than 100,000 round-trip trades in the U.S. market powerfully depicts that the lion's share of pairs trading profits are typically captured on the first few days after divergence. Clearly, fast execution is key for practitioners.

8. Conclusion

We have comprehensively reviewed literature closely related to the umbrella term of pairs trading, covering both univariate and multivariate strategies. Clustered by pairs trading approach, we can summarize our findings and suggestions for further research as follows.

8.1 *Distance Approach*

The distance approach relies on a simple algorithm that is easy to implement and robust to data snooping in its original implementation. The SSD has several deficiencies, leading to low variance spreads with limited profit potential and substantial divergence risk. Pearson correlation on return levels performs slightly better. Quasi-multivariate pairs trading leverages information from a whole portfolio of matching partners in one synthetic asset and generally outperforms the univariate version. Pairs trading profitability has low exposure to systematic sources of risk, declines over time, and can partially be explained by information diffusion, pairs visibility, and market frictions, such as liquidity factors. There are applications to other asset classes (bonds, commodities) or time frames (daily data, high-frequency data) and GGR's initial findings are usually confirmed.

Further research could aim at improving the selection metric to avoid the creation of minimum variance spreads without benefiting from increasing mean-reversion strength. Especially, combinations with the cointegration approach are promising. The chase for common factors explaining pairs trading profitability could be implemented in a truly global setting, for example, following Asness *et al.* (2013) across multiple markets and asset classes. The presented research leads to the conjecture that pairs trading return premia may be consistent across diverse markets with a strong common factor structure—as it is the case for value-momentum trading (Asness *et al.*, 2013).

8.2 *Cointegration Approach*

The frameworks presented in the cointegration approach are more diverse than in the distance approach. We can summarize as follows: Cointegration constitutes a more rigorous framework for pairs trading compared to the distance approach due to the econometrically sound identification of equilibrium relationships. Vidyamurthy (2004) is the most cited author in this domain. He has proposed a set of heuristics instead of cointegration testing that have not yet been empirically applied. Rad *et al.* (2015) provide the first large-scale empirical application, albeit with a strong bias induced by preselecting pairs via SSD. Lin *et al.* (2006) and Puspaningrum *et al.* (2010) develop improved trading rules specifically designed for cointegrated securities. Their concepts have the potential to improve trading results and yet need to be tested on a large data set. Further empirical applications of the univariate frameworks are frequently limited to smaller groups of securities. In most cases, the specifics of multiple comparison problems are not considered, i.e., the cumulation of type I errors through repeated testing on the same data set and the resulting high number of false positives. A combination of the preselection heuristics of Vidyamurthy (2004) with adequate statistical procedures to control the familywise error rate as suggested in Cummins and Bucca (2012) would allow for large-scale empirical applications. The multivariate-enhanced indexation strategies are highly susceptible to identifying spurious relationships. More appealing are the multivariate statistical arbitrage strategies. The approach of Galenko *et al.* (2012) leads to positive profits in expectation—it should be applied to a larger stock universe.

8.3 *Times-Series Approach*

Elliott *et al.* (2005) introduce state space models and appropriate estimation algorithms to parametrically deal with mean-reverting spreads. Several authors have capitalized on their findings. Avellaneda and Lee (2010) successfully apply a variant of this approach to mean-reverting portfolios developed with PCA. This application clearly indicates that dynamic trading rules based on time-series analysis can improve trading returns. Yet, this paper remains the only larger empirical application so far. Bertram (2010) presents an optimal statistical arbitrage trading rule for mean-reverting portfolios. Initial empirical applications by Cummins and Bucca (2012) to the energy futures markets are highly promising. Also, in this case, the strategy has not yet been deployed to other samples. Future research should especially bridge the gap between the distance/cointegration approach and the time-series approach. The former focus on the identification of pairs and only apply simple trading rules, mostly on a standard deviation logic in the sense of GGR. The latter do not address the issue of actually finding matching pairs, but instead develop complex trading systems aiming for improved profitability. Hence, combining strong pairs selection algorithms with suitable trading strategies from the time-series approach may yield powerful empirical results.

8.4 *Stochastic Control Approach*

The stochastic control approach is primarily focused on finding the optimal investment in the two legs of a pair when other assets are available. The model of Jurek and Yang (2007) clearly outperforms a standard threshold rule as in GGR. However, using a cointegration framework and allowing for flexible investment positions, Liu and Timmermann (2013) show that standard delta-neutral strategies may also be suboptimal relative to their unconstrained investment policy. Lei and Xu (2015) enhance this framework and incorporate transaction costs. A large-scale comparison of these strategies compared to GGR's algorithm is key. It would show how these concepts fare in light of actual market frictions and model misspecifications.

8.5 *Other Approaches*

Huck (2009, 2010) introduces a combined approach, building on artificial neural networks and ELECTRE III. The results are impressive, but it is unclear if this is due to the combined forecasts or the MCDM. Cutting the complexity and replacing the MCDM with a more transparent outranking system may even show improvements. Alternatively, a rigorous approach of ensemble learning as discussed in Mendes-Moreira *et al.* (2012) could be applied to better manage the bias-variance trade-off. Concretely, Huck does not perform ensemble pruning (i.e., no models are discarded) and it is unclear if ELECTRE III is the optimal mechanism for ensemble integration. In any case, despite the computational complexity, such studies should be conducted on a larger database. Several authors apply copulas to pairs trading. A key direction for further research is the development of an integrated copula-based selection and trading algorithm in the spirit of Krauss and Stübinger (2015), but for the level-based copula method of Rad *et al.* (2015). Avellaneda and Lee (2010) use PCA to create seemingly mean-reverting time series, which they model as OU-processes similar to Elliott *et al.* (2005). Their empirical application is state-of-the-art, except for a rigorous control mechanism with respect to data mining. Methodological improvements should be focused on replacing standard PCA with more advanced methods, such as asymptotic PCA or a multivariate cointegration model.

8.6 *Pairs Trading in the Light of Market Frictions*

Several studies discuss the effect of market frictions on pairs trading profitability. Generally, returns are driven by limits to arbitrage. Incorporating all dimension of trading costs renders classical approaches in the spirit of GGR virtually unprofitable—only refined portfolios continue to deliver positive excess returns at much diminished levels.

Future research should address the question how short-selling constraints impact pairs trading profitability, given that several large capitalization equity markets have implemented such restrictions in the course of the financial crisis (Gregoriou, 2012). Optimal trading models in the spirit of Liu and Timmermann (2013) and Lei and Xu (2015) could be extended by all three dimensions of trading costs, i.e., commissions, market impact, and short-selling fees. The resulting trading thresholds would truly incorporate the trade-off between trading frequency and net profit per trade, thus optimizing total profit over the trading period. Finally, market frictions are mainly discussed for the distance approach. Similar analyses could still be performed for the other approaches.

Acknowledgements

I am grateful to Matthew Clegg, Nicolas Huck, Ingo Klein, Benedikt Mangold, Walter Scherb, Johannes Stübinger, and two anonymous referees for many helpful discussions and suggestions on this topic.

Notes

1. GGR have published their research in two stages, i.e., Gatev *et al.* (1999); Gatev *et al.* (2006).

2. As of 1st of March 2016, there are 1.855 citations on Google Scholar for the key contrarian paper by Jegadeesh (1990) and 7.780 citations for the key momentum paper by Jegadeesh and Titman (1993) as opposed to 428 citations for Gatev *et al.* (2006).
3. In some cases, returns are annualized. When several variants of the strategy are tested, we select a representative return or provide a range. Please note that the calculation logic for the returns differs between papers, so they are not necessarily directly comparable. This applies to all return metrics provided in this paper. Furthermore, if not indicated otherwise, the respective samples refer to stock markets. The latter applies to all subsequent tables in this paper.
4. For the rest of this paper, price denotes the cumulative return index, with reinvested dividends.
5. Pairs exhibiting very low SSD are also highly correlated, see Section 2.3.
6. See Chen *et al.* (2012): Raw returns of panel A of table 1 on p. 32 amount to 1.40% for the long-short portfolio formed with the comover portfolio logic. Raw returns of panel B of table 5 on p. 36 amount to 0.95% for the long-short portfolio formed with classical stock pairs.
7. Consider average SSD of the realizations $p_{i,t}$ and $p_{j,t}$ with $t \in T$ of two standardized price processes to see that high sample correlation $\hat{\rho}_{P_i,P_j}$ leads to low SSD:

$$\overline{ssd}_{P_i,P_j} = \frac{1}{T}\sum_{t=1}^{T}\left(p_{i,t} - p_{j,t}\right)^2 = \frac{1}{T}\sum_{t=1}^{T}\left(p_{i,t}^2 + p_{j,t}^2\right) - \frac{1}{T}\sum_{t=1}^{T} 2p_{i,t}p_{j,t} = \frac{1}{T}\sum_{t=1}^{T}\left(p_{i,t}^2 + p_{j,t}^2\right) - 2\hat{\rho}_{P_i,P_j}$$

8. As Girma and Paulson (1999) point out, there is no uniform approach to calculate returns on future investments. Hence, the authors have provided profits in terms of USD and an estimation for annualized returns based on the required initial investment to run such a strategy.
9. For further details, see Do *et al.* (2006, p. 10 ff.)
10. Further discussions of the OU-process in the context of pairs trading can be found in Rampertshammer (2007). Kim (2011) provides an empirical implementation in a high-frequency setting in the Korean stock market. Zeng and Lee (2014) develop an extension of Bertram (2010).
11. Boguslavsky and Boguslavskaya (2004) develop the optimal investment strategy for a single risky asset following an OU-process for an arbitrageur under power utility. Mudchanatongsuk *et al.* (2008) also solve the stochastic control problem for pairs trading under power utility for terminal wealth. Their approach mostly differs in the assumed asset pricing dynamics, but the spread also relies on an OU-process. Kim *et al.* (2008) extend the stochastic control problem to multiple spreads and provide tractable solutions. Zhang and Zhang (2008) develop optimal buy and sell thresholds. The works of Ekström *et al.* (2011), Larsson *et al.* (2013), Song and Zhang (2013), Lindberg (2014), and Kuo *et al.* (2015) focus on how to optimally liquidate a pairs trade when incorporating stop-loss thresholds.
12. Further developments of optimal strategies for cointegrated risky assets can be found in Tourin and Yan (2013) and Chiu and Wong (2015).
13. There are further authors who apply machine learning to pairs trading. Dunis *et al.* (2006a) model the gasoline crack spread with artificial neural networks. Dunis *et al.* (2006b) apply recurrent and higher order networks to the soybean-oil crush spread, and Dunis *et al.* (2008) to a portfolio of oil futures spreads. Thomaidis *et al.* (2006) propose an experimental statistical arbitrage system based on neural network GARCH models. Lin and Cao (2008) and Huang *et al.* (2015) use genetic algorithms for pairs mining and Dunis *et al.* (2015) develop pairs trading strategies for the corn-ethanol crush spread with different neural network types and genetic algorithms. Finally, Montana and Parrella (2009) use an ensemble of support vector regressions to develop a pairs trading strategy for the iShares S&P 500 ETF.

References

Alexander, C. (1999) Optimal hedging using cointegration. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* 357(1758): 2039–2058.

Alexander, C. (2001) *Market Models: A Guide to Financial Data Analysis*. Chichester, UK and New York, NY: Wiley.

Alexander, C. and Dimitriu, A. (2005) Indexing and statistical arbitrage. *The Journal of Portfolio Management* 31(2): 50–63.

Andrade, S., Di Pietro, V. and Seasholes, M. (2005) Understanding the profitability of pairs trading. Working paper, UC Berkeley, Northwestern University.

Asness, C.S., Moskowitz, T.J. and Pedersen, L.H. (2013) Value and momentum everywhere. *Journal of Finance* 68(3): 929–985.

Avellaneda, M. and Lee, J.-H. (2010) Statistical arbitrage in the US equities market. *Quantitative Finance* 10(7): 761–782.

Baronyan, S.R., Boduroğlu, İ.İ. and Şener, E. (2010) Investigation of stochastic pairs trading strategies under different volatility regimes. *The Manchester School* 78(1): 114–134.

Bertram, W.K. (2010) Analytic solutions for optimal statistical arbitrage trading. *Physica A: Statistical Mechanics and its Applications* 389(11): 2234–2243.

Bianchi, R., Drew, M. and Zhu, R. (2009) Pairs trading profits in commodity futures markets. In *Proceedings of Asian Finance Association 2009 International Conference*, pp. 1–26.

Bock, M. and Mestel, R. (2009) A regime-switching relative value arbitrage rule. In *Operations Research Proceedings 2008*, pp. 9–14. Springer.

Bogomolov, T. (2011) Pairs trading in the land down under. In *Finance and Corporate Governance Conference*.

Bogomolov, T. (2013) Pairs trading based on statistical variability of the spread process. *Quantitative Finance* 13(9): 1411–1430.

Boguslavsky, M. and Boguslavskaya, E. (2004) Arbitrage under power. *Risk* 17(6): 69–73.

Bossaerts, P. (1988) Common nonstationary components of asset prices. *Journal of Economic Dynamics and Control* 12(2–3): 347–364.

Bowen, D.A. and Hutchinson, M.C. (2014) Pairs trading in the UK equity market: risk and return. *The European Journal of Finance*, doi: 10.1080/1351847X.2014.953698.

Bowen, D.A., Hutchinson, M.C. and O'Sullivan, N. (2010) High frequency equity pairs trading: transaction costs, speed of execution and patterns in returns. *The Journal of Trading* 5(3): 31–38.

Broumandi, S. and Reuber, T. (2012) Statistical arbitrage and FX exposure with South American ADRs listed on the NYSE. *Financial Assets and Investing* 3(2): 5–18.

Broussard, J.P. and Vaihekoski, M. (2012) Profitability of pairs trading strategy in an illiquid market with multiple share classes. *Journal of International Financial Markets, Institutions and Money* 22(5): 1188–1201.

Burgess, A.N. (1999) A computational methodology for modelling the dynamics of statistical arbitrage. PhD thesis, University of London, London Business School.

Burgess, A.N. (2003) Using cointegration to hedge and trade international equities. In C.L. Dunis, J. Laws and P. Naïm (eds.), *Applied Quantitative Methods for Trading and Investment* (pp. 41–69). Chichester, UK: John Wiley & Sons, Ltd.

Caldeira, J.F. and Moura, G.V. (2013) Selection of a portfolio of pairs based on cointegration: a statistical arbitrage strategy. *Brazilian Review of Finance* 11(1): 49–80.

Chen, C.W.S., Chen, M. and Chen, S.-Y. (2014) Pairs trading via three-regime threshold autoregressive GARCH models. In: V.-N. Huynh, V. Kreinovich and S. Sriboonchitta (eds.), *Modeling Dependence in Econometrics: Selected Papers of the Seventh International Conference of the Thailand Econometric Society , Faculty of Economics, Chiang Mai University, Thailand* (pp. 127–140). Springer International Publishing, Cham.

Chen, H., Chen, S.J. and Li, F. (2012) Empirical investigation of an equity pairs trading strategy. Working paper, University of British Columbia, University of Michigan.

Cheng, X., Yu, P.L.H. and Li, W.K. (2011) Basket trading under co-integration with the logistic mixture autoregressive model. *Quantitative Finance* 11(9): 1407–1419.

Chiu, M.C. and Wong, H.Y. (2015) Dynamic cointegrated pairs trading: mean–variance time-consistent strategies. *Journal of Computational and Applied Mathematics* 290: 516–534.

Cohen, L. and Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance* 63(4): 1977–2011.

Cummins, M. and Bucca, A. (2012) Quantitative spread trading on crude oil and refined products markets. *Quantitative Finance* 12(12): 1857–1875.
D'Aspremont, A. (2011) Identifying small mean-reverting portfolios. *Quantitative Finance* 11(3): 351–364.
D'Avolio, G. (2002) The market for borrowing stock. *Journal of Financial Economics* 66(2–3): 271–306.
Do, B. and Faff, R. (2010) Does simple pairs trading still work? *Financial Analysts Journal* 66(4): 83–95.
Do, B. and Faff, R. (2012) Are pairs trading profits robust to trading costs? *Journal of Financial Research* 35(2): 261–287.
Do, B., Faff, R. and Hamza, K. (2006) A new approach to modeling and estimation for pairs trading. In *Proceedings of 2006 Financial Management Association European Conference*.
Dunis, C. and Lequeux, P. (2010) Intraday data and hedging efficiency in interest spread trading. *The European Journal of Finance* 6(4): 332–352.
Dunis, C.L., Giorgioni, G., Laws, J. and Rudy, J. (2010) Statistical arbitrage and high-frequency data with an application to Eurostoxx 50 equities. Working paper, Liverpool Business School.
Dunis, C.L. and Ho, R. (2005) Cointegration portfolios of European equities for index tracking and market neutral strategies. *Journal of Asset Management* 6(1): 33–52.
Dunis, C.L., Laws, J. and Evans, B. (2006a) Modelling and trading the gasoline crack spread: a non-linear story. *Derivatives Use, Trading & Regulation* 12(1/2): 126–145.
Dunis, C.L., Laws, J. and Evans, B. (2006b) Modelling and trading the soybean-oil crush spread with recurrent and higher order networks: a comparative analysis. *Neural Network World* 16(3): 193–213.
Dunis, C.L., Laws, J. and Evans, B. (2006c) Trading futures spreads: an application of correlation and threshold filters. *Applied Financial Economics* 16(12): 903–914.
Dunis, C.L., Laws, J. and Evans, B. (2008) Trading futures spread portfolios: applications of higher order and recurrent networks. *The European Journal of Finance* 14(6): 503–521.
Dunis, C.L., Laws, J., Middleton, P.W. and Karathanasopoulos, A. (2015) Trading and hedging the corn/ethanol crush spread using time-varying leverage and nonlinear models. *The European Journal of Finance* 21(4): 352–375.
Ekström, E., Lindberg, C. and Tysk, J. (2011) Optimal liquidation of a pairs trade. In: G. Di Nunno and B. Øksendal (eds.) *Advanced Mathematical Methods for Finance* (pp. 247–255). Springer, Berlin, Heidelberg.
Elliott, R.J., Van Der Hoek, J. and Malcolm, W.P. (2005) Pairs trading. *Quantitative Finance* 5(3): 271–276.
Emery, G.W. and Liu, Q.W. (2002) An analysis of the relationship between electricity and natural–gas futures prices. *Journal of Futures Markets* 22(2): 95–122.
Engelberg, J., Gao, P. and Jagannathan, R. (2009) An anatomy of pairs trading: the role of idiosyncratic news, common information and liquidity. In *Proceedings of the 3rd Singapore International Conference on Finance*.
Ferreira, L. (2008) New tools for spread trading. *Futures: News, Analysis & Strategies for Futures, Options & Derivatives Traders* 37(12): 38–41.
Galenko, A., Popova, E. and Popova, I. (2012) Trading in the presence of cointegration. *The Journal of Alternative Investments* 15(1): 85–97.
Gatarek, L.T., Hoogerheide, L.F., van Dijk, H.K. and Verbeek, M. (2011) A simulation-based Bayes' procedure for robust prediction of pairs trading strategies. Working paper, Econometric Institute and Tinbergen Institute, Erasmus University Rotterdam.
Gatarek, L.T., Hoogerheide, L.F. and van Dijk, H.K. (2014) Return and risk of pairs trading using a simulation-based Bayesian procedure for predicting stable ratios of stock prices. Tinbergen Institute Discussion Papers, 14-039/III.
Gatev, E., Goetzmann, W.N. and Rouwenhorst, K.G. (1999) Pairs trading: performance of a relative value arbitrage rule. Working paper, Yale School of Management's International Center for Finance.
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.
Girma, P.B. and Paulson, A.S. (1999) Risk arbitrage opportunities in petroleum futures spreads. *Journal of Futures Markets* 19(8): 931–955.
Gregoriou, G.N. (2012) *Handbook of Short Selling*. Amsterdam and Boston, MA: Academic Press.
Gutierrez, J.A. and Tse, Y. (2011) Illuminating the profitability of pairs trading: a test of the relative pricing efficiency of markets for water utility stocks. *The Journal of Trading* 6(2): 50–64.

Hansen, P.R. (2005) A test for superior predictive ability. *Journal of Business & Economic Statistics* 23(4): 365–380.
Hong, G. and Susmel, R. (2003) Pairs-trading in the Asian ADR market. Working paper, University of Houston.
Huang, C.-F., Hsu, C.-J., Chen, C.-C., Chang, B.R. and Li, C.-A. (2015) An intelligent model for pairs trading using genetic algorithms. *Computational Intelligence and Neuroscience* 501. DOI: 10.1155/2015/939606.
Huck, N. (2009) Pairs selection and outranking: an application to the S&P 100 index. *European Journal of Operational Research* 196(2): 819–825.
Huck, N. (2010) Pairs trading and outranking: the multi-step-ahead forecasting case. *European Journal of Operational Research* 207(3): 1702–1716.
Huck, N. (2013) The high sensitivity of pairs trading returns. *Applied Economics Letters* 20(14): 1301–1304.
Huck, N. (2015) Pairs trading: does volatility timing matter? *Applied Economics* 47(57): 6239–6256.
Huck, N. and Afawubo, K. (2015) Pairs trading and selection methods: is cointegration superior? *Applied Economics* 47(6): 599–613.
Jacobs, H. (2015) What explains the dynamics of 100 anomalies? *Journal of Banking & Finance* 57: 65–85.
Jacobs, H. and Weber, M. (2013) Losing sight of the trees for the forest? Attention shifts and pairs trading. Working paper, University of Mannheim.
Jacobs, H. and Weber, M. (2015) On the determinants of pairs trading profitability. *Journal of Financial Markets* 23: 75–97.
Jegadeesh, N. (1990) Evidence of predictable behavior of security returns. *The Journal of Finance* 45(3): 881–889.
Jegadeesh, N. and Titman, S. (1993) Returns to buying winners and selling losers: implications for stock market efficiency. *The Journal of Finance* 48(1): 65–91.
Jones, C.M. (2002) A century of stock market liquidity and trading costs. Working paper, Columbia University.
Jurek, J.W. and Yang, H. (2007) Dynamic portfolio selection in arbitrage. Working paper, Harvard University.
Kanamura, T., Rachev, S.T. and Fabozzi, F.J. (2010) A profit model for spread trading with an application to energy futures. *The Journal of Trading* 5(1): 48–62.
Karakas, O. (2009) Mean reversion between different classes of shares in dual-class firms: evidence and implications. Working paper, London Business School.
Khandani, A.E. and Lo, A.W. (2007) What happened to the quants in August 2007? *Journal of Investment Management* 5(4): 5–54.
Khandani, A.E. and Lo, A.W. (2011) What happened to the quants in August 2007? Evidence from factors and transactions data. *Journal of Financial Markets* 14(1): 1–46.
Kim, K. (2011) Performance analysis of pairs trading strategy utilizing high frequency data with an application to KOSPI 100 equities. Working paper, Harvard University.
Kim, S.-J., Primbs, J. and Boyd, S. (2008) Dynamic spread trading. Working paper, Stanford University.
Krauss, C. and Stübinger, J. (2015) Nonlinear dependence modeling with bivariate copulas: statistical arbitrage pairs trading on the S&P 100. IWQW Discussion Paper Series, University of Erlangen-Nürnberg, Nürnberg.
Kuo, K., Luu, P., Nguyen, D., Perkerson, E., Thompson, K. and Zhang, Q. (2015) Pairs trading: an optimal selling rule. *Mathematical Control and Related Fields* 5(3): 489–499.
Larsson, S., Lindberg, C. and Warfheimer, M. (2013) Optimal closing of a pair trade with a model containing jumps. *Applications of Mathematics* 58(3): 249–268.
Lehmann, B.N. (1990) Fads, martingales, and market efficiency. *The Quarterly Journal of Economics* 105(1): 1–28.
Lei, Y. and Xu, J. (2015) Costly arbitrage through pairs trading. *Journal of Economic Dynamics and Control* 56: 1–19.
Li, M.L., Chui, C.M. and Li, C.Q. (2014) Is pairs trading profitable on China AH-share markets? *Applied Economics Letters* 21(16): 1116–1121.
Liew, R.Q. and Wu, Y. (2013) Pairs trading: a copula approach. *Journal of Derivatives & Hedge Funds* 19(1): 12–30.
Lin, L. and Cao, L. (2008) Mining in-depth patterns in stock market. *International Journal of Intelligent Systems Technologies and Applications* 4(3–4): 225–238.
Lin, Y.-X., McCrae, M. and Gulati, C. (2006) Loss protection in pairs trading through minimum profit bounds: a cointegration approach. *Journal of Applied Mathematics and Decision Sciences* 2006: 1–14.
Lindberg, C. (2014) Pairs trading with opportunity cost. *Journal of Applied Probability* 51(1): 282–286.

Liu, J. and Timmermann, A. (2013) Optimal convergence trade strategies. *Review of Financial Studies* 26(4): 1048–1086.

Liu, S.-M. and Chou, C.-H. (2003) Parities and spread trading in gold and silver markets: a fractional cointegration analysis. *Applied Financial Economics* 13(12): 899–911.

Mendes-Moreira, J., Soares, C., Jorge, A.M. and De Sousa, J.F. (2012) Ensemble approaches for regression. *ACM Computing Surveys* 45(1): 1–40.

Miao, G.J. (2014) High frequency and dynamic pairs trading based on statistical arbitrage using a two-stage correlation and cointegration approach. *International Journal of Economics and Finance* 6(3): 96–110.

Montana, G. and Parrella, F. (2009) Data mining for algorithmic asset management. In: L. Cao, P.S. Yu, C. Zhang and H. Zhang (eds.) *Data Mining for Business Applications* (pp. 283–295). Springer, US, Boston, MA.

Montana, G., Triantafyllopoulos, K. and Tsagaris, T. (2009) Flexible least squares for temporal data mining and statistical arbitrage. *Expert Systems with Applications* 36(2): 2819–2830.

Mori, M. and Ziobrowski, A.J. (2011) Performance of pairs trading strategy in the U.S. REIT market. *Real Estate Economics* 39(3): 409–428.

Mudchanatongsuk, S., Primbs, J.A. and Wong, W. (2008) Optimal pairs trading: a stochastic control approach. In *American Control Conference, 2008*, pp. 1035–1039.

Nath, P. (2003) High frequency pairs trading with U.S. treasury securities: risks and rewards for hedge funds. Working paper, London Business School.

Papadakis, G. and Wysocki, P. (2007) Pairs trading and accounting information. Working paper, Boston University and MIT.

Perlin, M.S. (2007) M of a kind: a multivariate approach at pairs trading. Working paper, ICMA/Reading University.

Perlin, M.S. (2009) Evaluation of pairs-trading strategy at the Brazilian financial market. *Journal of Derivatives & Hedge Funds* 15(2): 122–136.

Peters, G.W., Kannan, B., Lasscock, B., Mellen, C. and Godsill, S. (2011) Bayesian cointegrated vector autoregression models incorporating alpha-stable noise for inter-day price movements via approximate Bayesian computation. *Bayesian Analysis* 6(4): 755–792.

Pole, A. (2008) *Statistical Arbitrage: Algorithmic Trading Insights and Techniques*. Hoboken, NJ: John Wiley & Sons.

Puspaningrum, H. (2012) Pairs trading using cointegration approach. PhD thesis, University of Wollongong.

Puspaningrum, H., Lin, Y.-X. and Gulati, C.M. (2010) Finding the optimal pre-set boundaries for pairs trading strategy based on cointegration technique. *Journal of Statistical Theory and Practice* 4(3): 391–419.

Rad, H., Low, R.K.Y. and Faff, R.W. (2015) The profitability of pairs trading strategies: distance, cointegration, and copula methods. Working paper, University of Queensland.

Rampertshammer, S. (2007) An Ornstein-Uhlenbeck framework for pairs trading. Working paper, University of Melbourne.

Roll, R. (1984) A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance* 39(4): 1127–1139.

Romano, J.P., Shaikh, A.M. and Wolf, M. (2010) Hypothesis testing in econometrics. *Annual Review of Economics* 2(1): 75–104.

Romano, J.P. and Wolf, M. (2007) Control of generalized error rates in multiple testing. *The Annals of Statistics* 35(4): 1378–1408.

Simon, D.P. (1999) The soybean crush spread: empirical evidence and trading strategies. *Journal of Futures Markets* 19(3): 271–289.

Song, Q. and Zhang, Q. (2013) An optimal pairs-trading rule. *Automatica* 49(10): 3007–3014.

Stander, Y., Marais, D. and Botha, I. (2013) Trading strategies with copulas. *Journal of Economic and Financial Sciences* 6(1): 83–107.

Thomaidis, N.S., Kondakis, N. and Dounias, G.D. (2006) An intelligent statistical arbitrage trading system. In: G. Antoniou, G. Potamias, C. Spyropoulos and D. Plexousakis (eds.) *Advances in Artificial Intelligence: 4th Helenic Conference on AI, SETN 2006, Heraklion, Crete, Greece* (pp. 596–599). Springer, Berlin, Heidelberg.

Tourin, A. and Yan, R. (2013) Dynamic pairs trading using the stochastic control approach. *Journal of Economic Dynamics and Control* 37(10): 1972–1981.

Triantafyllopoulos, K. and Montana, G. (2011) Dynamic modeling of mean-reverting spreads for statistical arbitrage. *Computational Management Science* 8(1–2): 23–49.
Tsay, R.S. (2010) *Analysis of Financial Time Series* (3rd ed.). Wiley Series in Probability and Statistics. Hoboken, NJ: John Wiley & Sons.
Vidyamurthy, G. (2004) *Pairs Trading: Quantitative Methods and Analysis*. Hoboken, NJ: John Wiley & Sons.
Wahab, M. and Cohn, R. (1994) The gold-silver spread: integration, cointegration, predictability, and ex-ante arbitrage. *Journal of Futures Markets* 14(6): 709–756.
White, H. (2000) A reality check for data snooping. *Econometrica* 68(5): 1097–1126.
Xie, W., Liew, Q.R., Wu, Y. and Zou, X. (2014) Pairs trading with copulas. Working paper, Nanyang Technological University.
Xie, W. and Wu, Y. (2013) Copula-based pairs trading strategy. In *Asian Finance Association (AsFA) 2013 Conference*.
Xiong, W. (2001) Convergence trading with wealth effects: an amplification mechanism in financial markets. *Journal of Financial Economics* 62(2): 247–292.
Zeng, Z. and Lee, C.-G. (2014) Pairs trading: optimal thresholds and profitability. *Quantitative Finance* 14(11): 1881–1893.
Zhang, H. and Zhang, Q. (2008) Trading a mean-reverting asset: buy low and sell high. *Automatica* 44(6): 1511–1518.

Appendix

Table A.1. Distance Approach

Study	Sample	Objective
Gatev *et al.* (1999) Gatev *et al.* (2006)	U.S. CRSP 1962–1997 U.S. CRSP 1962–2002	Baseline approach in U.S. equity markets: Pairs trading is profitable; returns are robust
Do and Faff (2010) Do and Faff (2012) Rad *et al.* (2015)	U.S. CRSP 1962–2009 U.S. CRSP 1963–2009 U.S. CRSP 1962–2014	Expanding on GGR: Profitability is declining and not robust to transaction costs; improved formation based on industry, number of zero crossings
Perlin (2007) Perlin (2009) Chen *et al.* (2012)	Brazil 2000–2006 Brazil 2000–2006 U.S. CRSP 1962–2002	Improvements: Correlation-based formation outperforms SSD rule; quasi-multivariate pairs trading variants outperform univariate pairs trading
Andrade *et al.* (2005) Papadakis and Wysocki (2007) Engelberg *et al.* (2009) Jacobs and Weber (2013) Jacobs (2015) Jacobs and Weber (2015)	Taiwan 1994–2002 U.S. subset 1981–2006 U.S. CRSP 1993–2006 Intl'; U.S. CRSP 1960–2008 U.S. CRSP 1962–2008 Intl'; U.S. CRSP 1962–2008	Sources of pairs trading profitability: Uninformed demand shocks, accounting events, common versus idiosyncratic information, market frictions, etc.
Huck (2013) Huck (2015)	U.S. S&P 500 2002–2009 U.S.; Japan 2003–2013	Sensitivity of pairs trading profitability to duration of formation period and to volatility timing
Nath (2003) Bowen *et al.* (2010)	U.S. GovPX 1994–2000 U.K. FTSE 100 2007–2007	High frequency: Pairs trading profitability in the U.S. bond market and in the U.K. equity market
Bianchi *et al.* (2009) Mori and Ziobrowski (2011) Broussard and Vaihekoski (2012) Bowen and Hutchinson (2014)	Commodities 1990-2008 U.S. REITS 1987–2008 Finland 1987–2008 U.K. 1979–2012	Further out-of-sample tests: Pairs trading profitability in the commodity markets, the REIT sector, the Finnish market, the U.K. equity market

Table A.2. Cointegration Approach

Study	Sample	Objective
Vidyamurthy (2004)	–	Most widely cited cointegration-based concept
Rad *et al.* (2015)	U.S. CRSP 1962–2014	Largest empirical study of the cointegration approach
Lin *et al.* (2006) Puspaningrum *et al.* (2010) Puspaningrum (2012)	Selected stocks 2001–2002 Selected stocks 2004–2005 Selected applications	Entry/exit signals: Development of minimum profit bounds and of optimal preset boundaries for cointegration-based pairs trading
Wahab and Cohn (1994) Girma and Paulson (1999) Simon (1999) Emery and Liu (2002) Dunis *et al.* (2006c)	Gold-silver spread 1988–1992 Crack spread 1983–1994 Soy crush spread 1985–1995 Spark spread 1996–2000 Energy futures 1995–2004	Futures spread trading: Cointegration-based trading of the gold-silver spread, the crack spread, the soy crush spread, the spark spread, and the WTI-brent spread
Hong and Susmel (2003) Broumandi and Reuber (2012)	64 Asian shares 1991–2000 Selected stocks 2003–2009	ADRs: Cointegration-based pairs trading strategies for ADRs and the local stocks
Dunis and Lequeux (2010) Gutierrez and Tse (2011) Caldeira and Moura (2013) Li *et al.* (2014) Miao (2014)	EuroStoxx 50 2003/2009–2009 Selected stocks 1997–2008 Brazil 2005–2012 38 Chinese stocks 2009–2013 2100 U.S. stocks 2012–2013	Common stocks: Cointegration-based pairs trading frameworks with applications in European and U.S. high-frequency settings and applications to further international markets; improved selection via Granger-causality

(*Continued*)

Table A.2. *Continued*

Study	Sample	Objective
Baronyan *et al.* (2010)	U.S. DJIA 1999–2008	Comparison studies: Comparison of univariate
Bogomolov (2011)	Australia ASX 1996–2010	pairs trading strategies—most notably distance
Huck and Afawubo (2015)	U.S. S&P 500 2000–2011	versus different variants of cointegration approach
Alexander (1999)	Intl' indexes 1990–1998	Passive index tracking/enhanced indexation:
Alexander (2001)	U.S. S&P 100 1995–2000	Development of multivariate cointegration-based
Dunis and Ho (2005)	EuroStoxx 50 1999–2003	strategies for tracking indices or artificial benchmarks
Alexander and Dimitriu (2005)	U.S. DJIA 1990–2003	
Burgess (2003)	EuroStoxx 50 1998–2002	Cointegration-based multivariate statistical arbitrage approaches
Galenko *et al.* (2012)	Intl' indexes 2003–2009	
Burgess (1999)	U.K. FTSE 100 1997–1999	Statistical arbitrage based on cointegration and machine learning
D'Aspremont (2011)	U.S. Swap rates 1998–2005	Identification of sparse mean-reverting portfolios
Liu and Chou (2003)	Gold-silver futures 1983–1995	Fractional cointegration: Pairs trading approaches
Karakas (2009)	U.S. dual class firms 1980–2006	based on fractional cointegration
Peters *et al.* (2011)	Selected securities 1999–2005	Bayesian approach: Development of Bayesian approaches
Gatarek *et al.* (2011, 2014)	Selected securities 2009–2009	for cointegration-based pairs trading
Cheng *et al.* (2011)	Selected stocks 2005–2008	Cointegration approach with logistic mixture AR equilibrium errors

Table A.3. Time-Series Approach

Study	Sample	Objective
Elliott *et al.* (2005)	–	Modeling the spread in state space at the price level
Do *et al.* (2006)	–	Modeling the spread in state space at the return level
Triantafyllopoulos and Montana (2011)	Selected stocks 1980–2008	Modeling the spread in state space with a Bayesian approach
Rampertshammer (2007)	–	Modeling the spread with OU-processes; development of optimal entry and exit thresholds; selected empirical applications, also in high-frequency settings
Bertram (2010)	–	
Kim (2011)	Korea KOSPI 2008–2010	
Cummins and Bucca (2012)	Energy futures 2003–2010	
Zeng and Lee (2014)	Selected stocks	
Bock and Mestel (2009)	DJ STOXX 600 2006–2007	Further time-series approaches for spread modeling: Markov regime-switching; profit model based on OU-process; nonparametric charting approach; three regime TAR-GARCH
Kanamura *et al.* (2010)	Energy futures 2000–2008	
Bogomolov (2013)	U.S.; Australia; 1996–2011	
Chen *et al.* (2014)	U.S. DJIA 2006–2013	

Table A.4. Stochastic Control Approach

Study	**Sample**	**Objective**
Boguslavsky and Boguslavskaya (2004)	–	OU-process: Derivation of the optimal strategy for
Jurek and Yang (2007)	Selected stocks 1962–2006	a risky asset following an OU-process under various utilities
Mudchanatongsuk *et al.* (2008)	–	
Kim *et al.* (2008)	Selected stocks 2002–2008	
Zhang and Zhang (2008)	–	Optimal stopping theory: Derivation of the optimal opening
Ekström *et al.* (2011)	–	and closing of a pairs trade under different conditions
Larsson *et al.* (2013)	–	(OU-process; Lévy processes with jumps; opportunity costs, etc.)
Song and Zhang (2013)	Selected stocks 1992–2012	
Lindberg (2014)	–	
Kuo *et al.* (2015)	Selected stocks 2001–2012	
Liu and Timmermann (2013)	Selected stocks 2006–2012	Cointegration: Derivation of the optimal strategy for two
Tourin and Yan (2013)	Selected stocks 2011–2011	cointegrated risky assets under various utilities
Chiu and Wong (2015)	–	
Lei and Xu (2015)	Selected stocks	

Table A.5. Other Approaches

Study	Sample	Objective
Huck (2009)	U.S. S&P 100 1992–2006	Multivariate pairs trading based on Elman artificial neural networks and ELECTRE III
Huck (2010)	U.S. S&P 100 1993–2006	
Dunis *et al.* (2006a)	Crack spread 1995–2005	Pairs trading frameworks based on different machine learning techniques (artificial neural networks, higher order neural networks, recurrent neural networks, genetic algorithms, support vector regression, etc.)
Thomaidis *et al.* (2006)	Selected stocks 2005–2005	
Dunis *et al.* (2008)	Energy future spreads 1995–2005	
Lin and Cao (2008)	Australia ASX selection 2000–2002	
Montana and Parrella (2009)	U.S. S&P 500 EFT 2000–2007	
Huang *et al.* (2015)	Selected stocks 2003–2012	
Dunis *et al.* (2015)	Corn/eth. crush spread 2005–2010	
Ferreira (2008)	Selected stocks 2007–2008	Return-based copula methods for pairs trading
Liew and Wu (2013)	Selected stocks 2009–2012	
Stander *et al.* (2013)	Selected stocks, SSFs 2007–2009	
Krauss and Stübinger (2015)	U.S. S&P 100 1990–2014	
Xie and Wu (2013)	Selected stocks 2009–2012	Level-based copula methods for pairs trading
Xie *et al.* (2014)	CRSP utility stocks 2003–2012	
Rad *et al.* (2015)	U.S. CRSP 1962–2014	
Montana *et al.* (2009)	U.S. S&P 500 1997–2005	Multivariate pairs trading frameworks based on PCA
Avellaneda and Lee (2010)	U.S. subset 1997–2007	