

## A good pair: alternative pairs-trading strategies

R. Todd Smith<sup>1</sup> · Xun Xu<sup>2</sup>

Published online: 9 January 2017  
© Swiss Society for Financial Market Research 2017

**Abstract** This paper studies alternative techniques for identifying stock pairs in a pairs-trading strategy over 1980–2014. We consider two main techniques: the distance approach and the cointegration approach. We also consider a range of parameterizations of the trading system design. Parameterization of the trading system matters for the profitability of pairs trading. We find that the cointegration approach, despite using an optimal in-sample parameterization, yields significant returns only in the 1980s. The distance approach performs better, producing significantly positive risk-adjusted returns in all sub-periods. However, when transaction costs are properly taken into account, the returns largely disappear in recent years.

**Keywords** Stocks · Pairs trading · Investing · Arbitrage

**JEL Classification** G1

### 1 Introduction

Pairs trading is the practice of taking simultaneous short and long positions in two similar securities (pairs) when their price spread exceeds a specified threshold. It is a common investment strategy: a Google News search of “pairs trade” on August 3, 2016 yields 154,000 articles, many recommending specific pair trades. While there has surely been analysis of alternative pairs-trading strategies and methods by investors

---

✉ R. Todd Smith  
smithrt@ualberta.ca

<sup>1</sup> Department of Economics, University of Alberta, 8-14 Tory Building, Edmonton, AB T6G 2H4, Canada

<sup>2</sup> Vestcor Investment Management Corporation, Fredericton, NB, Canada

that use them, such work is largely proprietary. The academic literature on this issue is small. The best-known work is [Gatev et al. \(2006\)](#). They find that, for one technique (the “distance approach”) with a specific parameterization, trading pairs of US equities produces excess returns of over 10% annually. Study of a second technique (the “cointegration approach”) has reached similar broad conclusions, although studies that use this technique are fewer and more specialized. The purpose of this paper is twofold. First, we study both approaches and compare their relative strengths and weaknesses for pairs trading using data from 1980 through 2014. Second, for both techniques, we do not limit ourselves to a specific parameterization of the pairs matching and trading algorithm, in contrast to much of the existing literature, but instead explore a wide range of possible parameterizations. This provides insight into the design of an optimal pairs-trading system.

There are three broad dimensions in the design of a pairs-trading system that we explore. The first is the pairs matching algorithm. The literature focuses on two main approaches. The most studied from the perspective of profitability of pairs trading is the distance approach, which is a simple method that selects pairs for trading based on their historical closeness in (normalized) price paths. This approach is reportedly (e.g., [Gatev et al. 2006](#)) commonly used by industry practitioners. The second approach—the cointegration approach—has greater theoretical foundation (i.e., a good pair will have cointegrated prices and their spread will be a mean-reverting process), but is much more complicated to implement.<sup>1</sup> There is no comprehensive analysis of pairs trading using both the distance and cointegration approaches. This paper fills that gap.

The second aspect of designing pairs-trading systems is specifying values for a series of parameters that affect both the pairs selection stage and the precise conditions under which pairs trades are executed. The literature focuses mainly on one such parameterization due to [Gatev et al. \(2006\)](#): pairs are chosen based on 12 months of historical data (the “formation period” length); pair trades are executed when price spreads diverge by two standard deviations (computed from the 12-months formation period); and a trade is closed when prices converge or after 6 months if prices have not converged (the “trading period” length). What is interesting is that this specific parameterization has become standard in analyses of pairs trading using both approaches, but appears to lack rigorous justification. As [Gatev et al. \(2006, 803\)](#) state: “We form pairs over a 12-months period (formation period) and trade them in the next 6-months period (trading period). Both 12 and 6 months are chosen arbitrarily and have remained our horizons since the beginning of the study” (emphasis added). The two-standard-deviation value for the pair trade trigger is similarly chosen arbitrarily.

In this study, we measure the sensitivity of return distributions from pairs trading to these three parameters for each of the two main approaches to pairs trading. This analysis requires the examination of billions of possible pairs for each of the two approaches. Our broad conclusion is that the parameterization can matter greatly for the return distribution for pairs trading, and the best parameterization is different for the two approaches, in different sub-periods, and across different sized portfolios.

<sup>1</sup> Specifically, the cointegration approach is often motivated in terms of common stochastic trends that arise from common fundamentals for the two securities. [Vidyamurthy \(2004\)](#) and others relate the cointegration model to the arbitrage pricing theory.

Taken together, return distributions are generally substantially more favorable under the optimal parameterization than with the standard parameterization. Of course, one should be mindful in such analyses of data-snooping bias, but our results, nevertheless, suggest that more consideration should be given to parameterization of pairs-trading methods.

One further aspect of pairs trading we consider is the sensitivity of pairs-trading profitability to the number of pairs that are tracked at any given point in time, which, therefore, sets an upper bound on the number of pairs trades that can be open at any time. While there is little direct evidence on how many pairs practitioners monitor, the academic literature considers this issue to some extent. We more thoroughly explore this aspect of designing a pairs-trading system, namely, we study systems that monitor the best 5, 10, 20, 50, 100, and 200 pairs, in all our analyses described above.

In general, we find that, ignoring transaction costs, the returns from both approaches are very large in the 1980s, 1990s, and for 2000–2014. However, once bid-ask spreads are taken into account, the cointegration approach does not produce positive excess returns in the post 1980s period, whereas the distance approach produces statistically positive risk-adjusted returns in all sub-periods. Once other transactions costs (commissions and market impact costs) are properly taken into account, the returns from the distance approach largely disappear in recent years, save perhaps periods of extreme market turmoil, such as the 2007 financial crisis, although even then they do not appear large.

The paper proceeds as follows. Section 2 discusses the literature. Section 3 describes pairs-trading methods, and Sect. 4 discusses parameterization of pairs-trading systems. Section 5 presents the results, and Sect. 6 concludes and suggests several extensions to explore in future research.

## 2 Literature

Reportedly, pairs trading originated in the mid-1980s with the quantitative group at Morgan Stanley ([Bookstaber 2006](#)). Although pairs trading is reportedly a common strategy at hedge funds and proprietary-trading operations, the academic literature is limited.

The most well-known study of pairs trading is by [Gatev et al. \(2006\)](#). The authors define and use the distance approach. The authors use daily US stock prices from 1962–2002, a 12-months formation period, a 6-months trading period, and a fixed trading threshold (a two-standard-deviation divergence in normalized prices). Their study finds that pairs trading generates a substantial excess return and a monthly Sharpe ratio six times larger than that of the overall market. They also show that their pairs-trading strategy has a high risk-adjusted Jensen alpha, low exposure-to-several systematic risk factors, covers reasonable transaction costs, and is not simply capturing short-term return reversals (e.g., [Lehmann 1990](#)). They also find that the profitability of pairs trading has declined over time, which they attribute to the strategy becoming more prevalent. Restricting pairs to industries does not raise profitability relative to the unrestricted strategy.

[Do and Faff \(2010, 2012\)](#) extend the sample period of [Gatev et al. \(2006\)](#) from 2002 through mid-2009. They find that the profitability of pairs trading has shown a downward trend, as [Gatev et al. \(2006\)](#) also note, although profitability is markedly higher in the bear market of the early 2000s and in the portion of the global financial crisis included in their sample. They attribute the higher profitability during bear markets to less market efficiency during these periods. [Do and Faff \(2012\)](#) argue that transaction costs significantly reduce the profitability of pairs trading.

[Papadakis and Wysocki \(2007\)](#) examine the impact of accounting information events on the profitability of pairs-trading strategies in US equity markets. They find that earning announcements and analyst forecasts can cause drift in relative prices, which often triggers the opening of pair trades. However, pairs trading based on such events is less profitable than non-event-triggered pairs trading.

[Engelberg et al. \(2009\)](#) investigate how information and liquidity influence the profitability of pairs trading. These researchers find that profit is lower when the news is specific to only one stock in the pairs as idiosyncratic news increases divergence risk and trading-period horizon risk. They also find that trading on large and liquid pairs tends to underperform trading on smaller and less liquid pairs.

The distance approach has also been studied for non-US equity markets, for example, [Broussard and Vaihekoski \(2012\)](#) for Finland, [Perlin \(2009\)](#) for the Brazilian market, [Bolgun et al. \(2010\)](#) for the Istanbul stock market, [Lucey and Walshe \(2013\)](#) for Germany and France, and [Deaves et al. \(2013\)](#) for Canada. These studies typically reach the same broad conclusions as [Gatev et al. \(2006\)](#). [Jacobs and Weber \(2013\)](#) use Gatev et al.'s [\(2006\)](#) approach to examine whether pairs trading returns are explained by the investor-inattention hypothesis ([Peng and Xiong 2006](#)). They find that this strategy is more profitable when their measure of investor inattention is high. This is confirmed using larger firms from US markets with daily data through 2008, and also generally for eight other developed-country stock markets.

The distance approach has been used to study pairs trading in markets other than equities. These studies include [Nath \(2003\)](#), who studies the secondary market for US Treasury securities, and [Kanemura et al.'s \(2008\)](#) study of energy futures markets.

A second branch of the literature uses the cointegration approach. This technique—but little or no empirical analysis—is discussed by, for example, [Vidyamurthy \(2004\)](#), [Gregory et al. \(2011\)](#), and [Herlemont \(2004\)](#). [Lin et al. \(2006\)](#) add a minimum profit constraint and study this modification using two Australian bank stocks. [Caldeira and Moura \(2013\)](#) use 50 stocks from Brazil and choose (cointegrated) pairs using the top-20 Sharpe ratios for in-sample returns on pairs trades using data for 2005–2012.

[Baronyan et al. \(2010\)](#) examine pairs trading by comparing the distance approach and cointegration approaches using weekly data for the 30 stocks in the DJIA over the period 1999–2008. This study refines the pair selection method by selecting (cointegrated) pairs using criteria similar to that explained in the present study and explained in Sect. 3. The authors find considerably lower returns on average than do most other studies.

Some studies consider parameterizations alternative to the standard one of [Gatev et al. \(2006\)](#). [Baronyan et al. \(2010\)](#) consider different formation period lengths, but their data are weekly for only 30 stocks, which could be a major drawback for analysis of what is typically considered as a fairly high-frequency trading strategy. Broussard

and Vaihekoski's (2012) study for Finland considers alternative values of the trigger parameter, Perlin (2009) considers the trigger parameter value in his study of 50 Brazilian stocks, and Lucey and Walshe (2013) do so in their analysis of pairs trading for a subset of stocks on the German and French exchanges.

The literature also contains some alternative techniques to the main ones discussed above. Huck (2009, 2010) develops a methodology that combines forecasting techniques with multi-criteria decision-making methods. The author ranks stocks in the S&P 100 with weekly data according to expected returns and pairs assets based on the highest over-valuations and under-valuations. Elliott et al. (2005), Tourin and Yan (2013), and Mudchanatongsuk et al. (2008) consider various technical models of the spread between two cointegrated prices, but in large part do not study these models empirically. Chen et al. (2012) use monthly data on the CRSP stocks from 1931–2010. Their pairs selection strategy is to create portfolios of 50 pairs that have the highest cross correlation of returns in the preceding 5 years.

### 3 Pairs-trading methodology

Pairs trading consists of two stages. The first stage is the formation period, where pairs of stocks are selected. The second stage is the trading period, where trades are opened and closed. We first discuss the formation period and, in particular, the two main techniques for selecting pairs.

#### 3.1 Pair formation: the distance approach

The first step in the distance approach is to normalize the price of all stocks in the sample to unity at the beginning of the formation period. Let  $T_{fp}$  denote the number of trading days in the formation period. The normalized price of each stock at the end of day  $t = 1, 2, \dots, T_{fp}$ , is, therefore:

$$P_t^i = \prod_{\tau=1}^t \left(1 + r_{\tau}^i\right), \quad (1)$$

where  $r_{\tau}^i$  is the stock's daily return inclusive of dividends.

The second step is to compute the distance,  $D_{i,j}$ , between two stocks  $i$  and  $j$  over the formation period:

$$D_{i,j} = \frac{\sum_{t=1}^{T_{fp}} \left(P_t^i - P_t^j\right)^2}{T_{fp}}. \quad (2)$$

If there are  $N$  stocks, then there are  $(N \times (N - 1))/2$  distances. The final step in selecting pairs with this method is to rank the pairs on the basis of the distances and create portfolios of pairs with the smallest distances (e.g., the “top 5” and “top 10” pairs).

### 3.2 Pair formation: the cointegration approach

A preliminary first step in the cointegration approach is determining the order of integration of each stock price using the Augmented Dickey–Fuller (ADF) test. Stocks can be potential pairs only if their order of integration is the same. The second step is to calculate the spread between two stock prices,  $\text{PR}_t^{ij} = \log P_t^i - \log P_t^j$ , and use the ADF to test for mean reversion of the spread. That is, estimate the regression equation:

$$\Delta \text{PR}_t^{ij} = \gamma \text{PR}_{t-1}^{ij} + \varepsilon_t, \quad (3)$$

and test the null hypothesis that  $\gamma = 0$ . If the null hypothesis can be rejected (we use a 99% confidence level), it indicates that the spread is mean reverting. The third step is to test for cointegration using the Johansen test.<sup>2</sup>

As might be expected, this procedure produces an extraordinarily large number of possible pairs, as many stock prices are cointegrated. To improve the pairs matching algorithm, further refinements are introduced. The fourth step we use is based on Granger causality. As Baronyan et al. (2010, 118) explain: “Specifically, the pairs must pass through the filter of satisfying two-way Granger causality tests. ... Obviously, two-way Granger causality is stronger than one-way Granger causality. A pair selected as such makes it less likely for the aforementioned structural breakdown [spurious cointegration] to take place before the trade is timed out at the end of the year”.

The cointegration approach does not provide any obvious metric for how to rank pairs in terms of their possible profitability for pairs trading. In contrast, with the distance approach, the measure (2) readily translates into forming pairs-trading portfolios for the “top 5” or “top 100” pairs, for example. To allow for a similar construction with the cointegration approach, we follow the literature and use a fifth step that involves specifying a metric for ranking pairs. Specifically, the metric used is the “market factor spread” (MFS) (see Herlemont 2004; Baronyan et al. 2010), where  $\text{MFS}(i,j) = |\beta_i - \beta_j|$  and  $\beta_i$  and  $\beta_j$  denote the market factors (the CAPM betas, computed over the formation period) for the pair. Pairs are ordered by their MFSs and pairs-trading portfolios are created using pairs having the lowest MFSs. The intuition behind the MFS criterion is that a pairs-trading strategy should be nearly market neutral. We acknowledge that there are clearly other possible criteria to use in place of the MFS. This is a topic for future research.

### 3.3 Opening a pairs trade

After choosing pairs in the formation period, all prices are normalized to unity at the beginning of a subsequent period, the “trading period”, and the spread of normalized prices for each pair is tracked. A pair trade is opened only if the spread of normalized

<sup>2</sup> Specifically, we use the algorithm “jcitest” in Matlab, the description of which is: “Johansen tests assess the null hypothesis  $H(r)$  of cointegration rank less than or equal to  $r$  among the numDims-dimensional time series in Y against alternatives  $H(\text{numDims})$  (trace test) or  $H(r+1)$  (maxeig test). The tests also produce maximum likelihood estimates of the parameters in a vector error-correction (VEC) model of the cointegrated series”.

prices exceeds a threshold, called the “trigger”. Opening a trade involves longing \$1 of the lower priced stock and shorting \$1 of the higher priced one. Formally, a trade is opened when  $|P_t^i - P_t^j| \geq \text{trigger}(i, j)$ , where:

$$\text{Trigger}(i, j) = n \times \text{stdev}(i, j). \quad (4)$$

Here,  $n$  is a scalar and  $\text{stdev}(i, j)$  is the standard deviation of the spread of normalized prices computed from the formation period. Because prices are normalized in both approaches, we can write:

$$\text{stdev}(i, j) = \sqrt{\frac{1}{T_{\text{fp}} - 1} \sum_{t=1}^{T_{\text{fp}}} \left[ (P_t^i - P_t^j)^2 - D_{i,j} \right]^2}. \quad (5)$$

The standard value in the literature for the scalar  $n$  that pins down the trigger is  $n = 2$ .

### 3.4 Closing a pairs trade

A pairs trade is closed when the normalized price spread returns to a non-positive value. If that does not happen during the trading period, and thus, a pairs trade is still open at the end of the trading period, the position is automatically closed at that time.

### 3.5 Calculation of return on pairs trading

We use the same method to calculate portfolio return as in [Gatev et al. \(2006\)](#) (and most of the subsequent literature), called the “fully-invested return” or “return on capital employed”. Specifically, for any pair  $k$ , let  $p^k = \{l^k, s^k\}$  represent a simultaneous long and short position of \$1 in the two stocks of pair  $k$ , where  $l^k$  and  $s^k$  are indicator variables that take the value 1 when a pair trade has occurred and 0 otherwise. Let  $d^k$  indicate the day during a specific trading period that a trade was most recently opened on pair  $k$ . Because we are using end-of-day prices (see below), then  $d^k \in \{1, 2, \dots, T_{\text{fp}} - 1\}$ , where  $T_{\text{fp}}$  is the number of days in a trading period. If  $R_t(l^k)$  and  $R_t(s^k)$ , respectively, represent 1-day returns on the long and short positions of the pair, then the daily return for the pair trade is

$$R_t(p^k) = R_t(l^k) - R_t(s^k). \quad (6)$$

The daily return on a portfolio of  $N_t^*$  pairs is

$$R_t^{\text{port}} = \sum_{k=1}^{N_t^*} W_t^k R_t(p^k), \quad (7)$$

where  $W_t^k = w_t^k / \sum_{j=1}^{N_t^*} w_t^j$  and

$$w_t^k = \left[ 1 + R_{t-1}(p^k) \right] \times \left[ 1 + R_{t-2}(p^k) \right] \times \cdots \times \left[ 1 + R_{d^k+1}(p^k) \right], \quad (8)$$

for  $t \geq d^k + 2$  and  $\omega_t^k = 1$  for  $t = d^k + 1$ . That is, we use the  $N_t^*$  open pairs that are held in the portfolio on day  $t$  to calculate the daily return of the portfolio, which is equal to the weighted average return of the pairs. The weight given to a pair is determined by its cumulative return relative to the sum of cumulative returns of all pairs in the portfolio. Note that because the strategy is based on a long–short position of \$1, the return of the portfolio has the interpretation of excess return (i.e., the risk-free rate cancels from the two sides of the pair trade).

## 4 Parameterization of the trading system

These approaches can be implemented once some parameter values are specified. The parameters in question determine the formation period length ( $T_{fp}$ ), the trading-period length ( $T_{tp}$ ), the number of pairs to include in the pairs-trading portfolio, and the trigger. Note that the trigger is pinned down by the parameter  $n$  and we will henceforth simply refer to  $n$  as the “trigger value”.

There is no intuitively obvious parameterization. The literature, however, focuses mainly on the single parameterization used by Gatev et al. (2006) in their analysis using the distance approach. This “standard parameterization” is a 12-month formation period, a 6-month trading period, and a trigger value  $n = 2$ . The optimality of this particular parameterization for the profitability of pairs-trading techniques has received little attention in the academic literature.<sup>3</sup> In addition, there is no obvious reason to think this parameterization should work equally well for the cointegration and distance approaches, or that it should be time invariant.

We, therefore, consider a range of parameterizations. We consider the standard 12-months formation period as well as a 9-month formation period. For the trigger value, we consider  $n \in \{0.3, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0\}$ . The trading period is set at 6 months throughout as we find that a shorter trading-period results in many pairs trades remaining open at the end of the trading period, and a longer trading period has minimal effects as pairs that converge do so by 6 months.<sup>4</sup> For portfolio sizes, we consider those with the top 5, 10, 20, 50, 100, and 200 pairs.

We use CRSP daily data from January 1980–December 2014. We consider only common stocks. To avoid survivorship bias, we drop stocks from a given formation

<sup>3</sup> Gatev et al. (2006) suggest that data-snooping bias is a reason to avoid exploring alternative parameterizations. Data-snooping bias is a valid argument, of course, but surely institutions that trade pairs, such as hedge funds, have studied alternative parameterizations. However, those studies would be largely proprietary. The present study is not.

<sup>4</sup> This is consistent with anecdotal evidence from the industry. For example, Schmerken (2006) reports that “[p]airs trades usually are meant for investors looking for a short-term profit in the market”. “The average time for a trade to be open is 15 days”, says LCM’s Gaubert. “A position could be held as long as a month”.

and trading period only when they have missing data during that particular formation or trading period. The number of stocks under consideration is 4258 in the 1980s, 9522 in the 1990s, and 9644 for 2000–2014. The number of possible pairs to examine, therefore, ranges from about 9 million in the 1980s to more than 45 million in each of the two later sub-periods for a given parameterization. The dimensionality of the problem, however, is much larger: with two alternative formation period lengths and seven alternative trigger values, in each month, we must examine 126, 884, and 142 possible pairs for the 1980s, 634, 612, and 734 possible pairs in the 1990s, and 650, 979, and 644 possible pairs in the 2000s—for each of the distance and cointegration approaches. Furthermore, this pairs matching and trading problem is repeated monthly using a rolling window of formation and trading periods: the first formation period begins on the first trading day of January 1980, the second period begins on the first trading day of February 1980, and both formation and trading periods roll forward by 1 month throughout the sample. Thus, there are multiple overlapping portfolios in any given month and returns are averaged across these portfolios.

## 5 Results

### 5.1 Profitability of pairs trading and the role of parameterization of the trading system

We begin by presenting returns from pairs trading without adjustments for transaction costs of any sort, and study to what extent the parameterization of the trading system matters for the profitability of pairs trading. A main focus later in the paper is the effect of transaction costs, but the qualitative implications we discuss here in regard to the role of parameterization are valid in all cases.

The obvious question to address in regard to parameterization is whether the standard parameterization used in the literature is more or less the best one, or whether there are gains in profitability to be had with different ones. We first present results for the distance approach for the three sub-periods. For each portfolio of a given number of pairs, the results presented in Table 1 correspond to the portfolio, where the trigger value and formation period are chosen optimally on an in-sample basis. The displayed results are for the case in which pairs trades are opened and closed on the same day as the trigger is breached (“no waiting”). Below, we consider an alternative, where trading is delayed 1 day after the trigger is breached (“one-day-later rule”), because the no-waiting case is arguably contaminated with some degree of bid-ask bounce. For the purposes of assessing the consequences of alternative parameterizations, the conclusions are similar in both cases.

Note three main observations from the results with the distance approach. First, both excess returns and market-risk-adjusted excess returns (Jensen alphas) are very large and statistically significant for virtually all portfolio sizes and sub-periods, reaching as high as nearly 40% annually for the portfolios with fewer pairs; moreover, most of the

**Table 1** Pairs trading with distance approach—no waiting

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
1980s						
Trigger	2	1.5	1.5	3	1.5	3
Formation period	12	9	12	12	12	9
Mean excess return	0.0248	0.0235	0.0214	0.0171	0.0159	0.0130
Standard deviation	0.0278	0.0206	0.0258	0.0294	0.0327	0.0360
t-statistic	9.0351	11.6725	8.3817	5.8806	4.9164	3.6878
Jensen's alpha	0.0250	0.0234	0.0212	0.0167	0.0165	0.0142
t-statistic	8.1144	10.6974	8.1104	5.1645	4.9639	4.0363
Market beta	-0.0178	0.0100	0.0217	0.0438	-0.0631	-0.1228
t-statistic	-0.3154	0.2534	0.5748	0.7058	-0.7536	-1.6587
Skewness	-1.0538	-1.0094	-2.2857	-1.7164	-1.9219	-0.3089
Kurtosis	4.5755	2.0397	8.3936	4.6336	4.5584	2.5510
Minimum	-0.1036	-0.0434	-0.1002	-0.1011	-0.1025	-0.0991
Maximum	0.0994	0.0659	0.0688	0.0690	0.0696	0.1152
Positive returns (%)	0.8922	0.9333	0.9216	0.8529	0.8333	0.7714
1990s						
Trigger	2	2	2	2	3	2.5
Formation period	12	12	12	12	12	12
Mean excess return	0.0275	0.0245	0.0191	0.0148	0.0129	0.0120
Standard deviation	0.0202	0.0174	0.0106	0.0066	0.0103	0.0154
t-statistic	13.7685	14.2308	18.1840	22.6255	12.6431	7.8599
Jensen's alpha	0.0267	0.0244	0.0185	0.0145	0.0130	0.0112
t-statistic	10.9855	9.3072	11.3071	18.1068	9.7704	8.4139
Market beta	0.0204	-0.0164	0.0230	0.0152	0.0074	0.0190
t-statistic	0.2401	-0.1642	0.4691	0.5691	0.2319	0.5473
Skewness	1.0145	1.5134	0.6647	0.3712	1.4984	-2.2542
Kurtosis	0.6178	3.7143	0.3388	-0.1158	5.3065	24.6889
Minimum	-0.0036	-0.0009	0.0007	0.0005	-0.0085	-0.0942
Maximum	0.0861	0.1025	0.0546	0.0319	0.0641	0.0685
Positive returns (%)	0.9608	0.9608	1.0000	1.0000	0.9510	0.9608
2000s						
Trigger	2	2	3	2	2.5	3
Formation period	12	12	12	9	12	9
Mean excess return	0.0199	0.0267	0.0207	0.0138	0.0113	0.0071
Standard deviation	0.0273	0.0199	0.0291	0.0216	0.0299	0.0432
t-statistic	9.2821	17.0733	9.0474	8.1931	4.8294	2.1059
Jensen's alpha	0.0199	0.0269	0.0206	0.0136	0.0115	0.0072
t-statistic	7.3287	14.1915	7.5481	6.3703	5.2003	2.1746
Market beta	0.0265	-0.0175	0.0119	0.0492	-0.0227	-0.0607
t-statistic	0.2401	-0.1642	0.4691	0.5691	0.2319	0.5473

**Table 1** continued

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
Skewness	-1.6236	-0.4600	-1.1257	-0.9677	-0.8970	-2.4571
Kurtosis	10.4878	7.3558	5.1604	7.8353	3.7100	12.2453
Minimum	-0.1361	-0.0863	-0.1025	-0.0985	-0.0998	-0.2622
Maximum	0.1176	0.0877	0.1094	0.0900	0.1033	0.1132
Positive returns (%)	0.9259	0.9506	0.9012	0.9212	0.8333	0.7697

Statistics are for the 1-month excess return distributions for pairs trading from the distance approach with the optimal strategy. Pairs are opened/closed on the same day the trigger is reached (“no waiting”)

excess return is market neutral.<sup>5</sup> Second, the number of pairs matters for profitability, with portfolios with fewer pairs having the highest returns. This is intuitive as the more pairs monitored for trading, the lower will be the average distance for a pair in the portfolio measured during the formation period, and thus, the pairs will generally be of lower quality in terms of matched stock pairs. Third, there is considerable variation in the optimal parameterization, with the standard parameterization being optimal in just seven of the 18 portfolios studied. The optimal trigger is generally larger than the standard value. The optimal parameterization, however, is reasonably close to the standard parameterization for the most profitable portfolio sizes (five and ten pairs).

The results for the cointegration approach (Table 2) are broadly similar to those for the distance approach, with a few exceptions: (1) they are somewhat lower than with the distance approach, but still large; (2) the standard parameterization is less often optimal than with the distance approach (in just two of 18 cases is it optimal); and (3) the optimal trigger value is generally lower than the standard value. Note that the standard value of the trigger is attributed to Gatev et al. (2006), who study only the distance approach, but this value has been adopted in the cointegration literature also. Our results suggest that, for both approaches, more thought may need to be given to simply assuming a trigger value of  $n = 2$  as best practice for actually trading pairs or for researching the profitability of this strategy.

To explore the extent to which parameterization of the trading system is important, we study how profitability varies with trigger values and two alternative formation periods (Figs. 1, 2, 3, 4). To do this, we vary one of these parameters, holding the other at its standard value. For presentation purposes, we average results across sub-periods. Clearly, the trigger value can matter a great deal for profitability of the trading system, and this is without any regard to transaction costs. This is true also for risk-adjusted returns as they are very similar to raw returns regardless of the parameterization (not shown).

<sup>5</sup> Jensen's alpha is estimated for each sub-period by regressing the monthly pairs-portfolio returns on a constant (Jensen's alpha) and the market return net of the risk-free rate. The same procedure is used below to estimate the Fama-French four-factor model. Data for the market excess return, as well as the other variables used in the Fama-French four-factor model considered below, are obtained from Kenneth French's website ([http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)). Reported t-statistics in the market model and Fama-French model are robust to heteroscedasticity and autocorrelation.

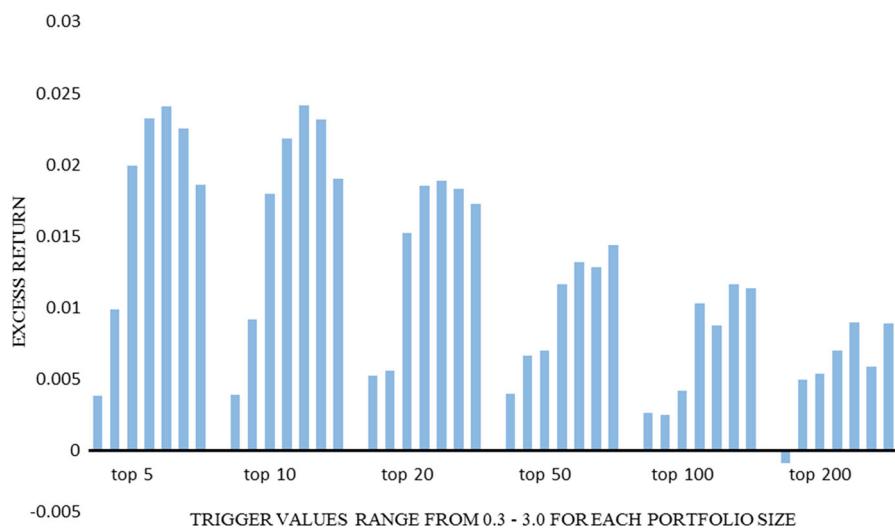
**Table 2** Pairs trading with cointegration approach—no waiting

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
1980s						
Trigger	2	2.5	2	2	1	1.5
Formation period	12	9	9	12	9	9
Mean excess return	0.0159	0.0169	0.0149	0.0161	0.0131	0.0129
Standard deviation	0.0290	0.0227	0.0300	0.0293	0.0397	0.0527
t-statistic	5.5532	7.6255	5.0990	5.5596	3.3760	2.5115
Jensen's alpha	0.0166	0.0173	0.0148	0.0165	0.0145	0.0131
t-statistic	4.8457	7.8414	4.3517	5.4572	3.8224	2.4689
Market beta	-0.0339	-0.0351	0.0323	-0.0280	-0.1303	-0.0142
t-statistic	-0.4518	-0.6846	0.4778	-0.6203	-1.7163	-0.1308
Skewness	0.1681	-0.1150	-1.9105	-1.7562	-0.6796	-0.5485
Kurtosis	-0.0607	-0.0395	5.4951	4.9522	2.3109	0.2470
Minimum	-0.0660	-0.0437	-0.1091	-0.1049	-0.1349	-0.1273
Maximum	0.0864	0.0691	0.0623	0.0683	0.1187	0.1208
Positive returns (%)	0.7157	0.7905	0.8095	0.8529	0.7810	0.7048
1990s						
Trigger	1	1	1.5	1	1.5	2.5
Formation period	12	12	12	12	12	12
Mean excess return	0.0161	0.0168	0.0167	0.0189	0.0191	0.0162
Standard deviation	0.0278	0.0222	0.0216	0.0123	0.0228	0.0323
t-statistic	5.8451	7.6428	7.7865	15.4916	8.4728	5.0764
Jensen's alpha	0.0152	0.0174	0.0166	0.0186	0.0196	0.0165
t-statistic	5.2979	6.8308	6.2995	15.6960	9.8770	4.6982
Market beta	0.0572	-0.0493	0.0097	0.0238	-0.0377	-0.0184
t-statistic	0.8681	-0.6954	0.1250	0.8832	-0.7411	-0.2096
Skewness	0.1652	-0.1396	-0.1452	0.9535	-0.9927	-0.5505
Kurtosis	0.2933	0.5795	-0.2420	1.7936	11.3208	2.3330
Minimum	-0.0506	-0.0444	-0.0339	-0.0079	-0.1007	-0.1028
Maximum	0.0952	0.0848	0.0740	0.0667	0.1168	0.1074
Positive returns (%)	0.7157	0.7941	0.8137	0.9706	0.9216	0.7843
2000s						
Trigger	1.5	1.5	1.5	1.5	1.5	2
Formation period	9	9	9	9	9	9
Mean excess return	0.0108	0.0136	0.0131	0.0162	0.0146	0.0197
Standard deviation	0.0355	0.0320	0.0284	0.0238	0.0238	0.0307
t-statistic	3.9123	5.4619	5.9142	8.7287	7.8956	8.2614
Jensen's alpha	0.0076	0.0126	0.0138	-0.0036	0.0142	0.0105
t-statistic	1.6129	4.0249	5.4750	-1.4372	6.3279	3.0285
Market beta	-0.0766	-0.0699	-0.0359	-0.0752	-0.0160	-0.0555
t-statistic	-0.6491	-0.7113	-0.4336	-1.0640	-0.2436	-0.6164

**Table 2** continued

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
Skewness	0.0087	0.3466	0.0232	0.5045	0.2755	-0.3274
Kurtosis	0.1993	0.7244	1.7468	1.3684	2.6497	4.0014
Minimum	-0.0812	-0.0703	-0.0875	-0.0453	-0.0631	-0.1136
Maximum	0.1068	0.1086	0.1110	0.1133	0.1186	0.1213
Positive returns (%)	0.6485	0.6848	0.7030	0.7697	0.7697	0.8242

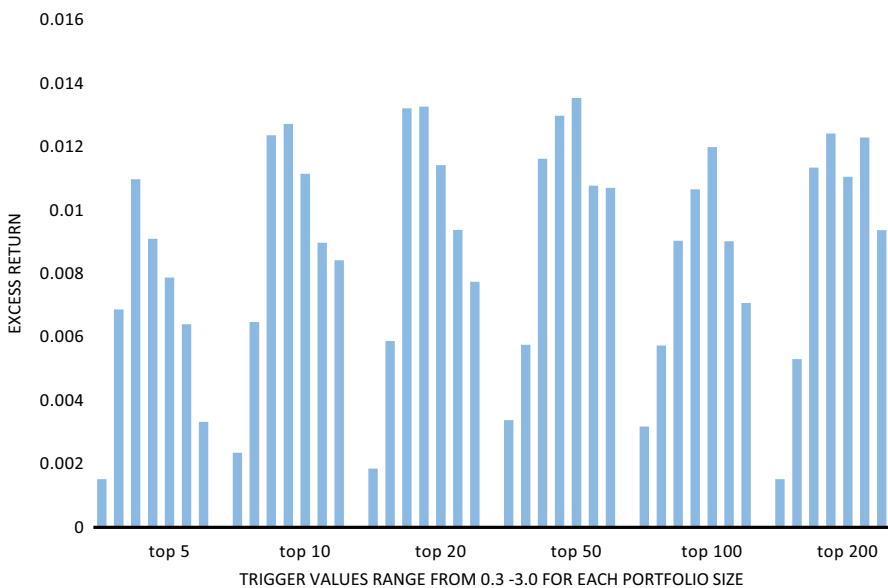
Statistics are for the 1-month excess return distributions for pairs trading from the cointegration approach with the optimal strategy. Pairs are opened/closed on the same day the trigger is reached (“no waiting”)



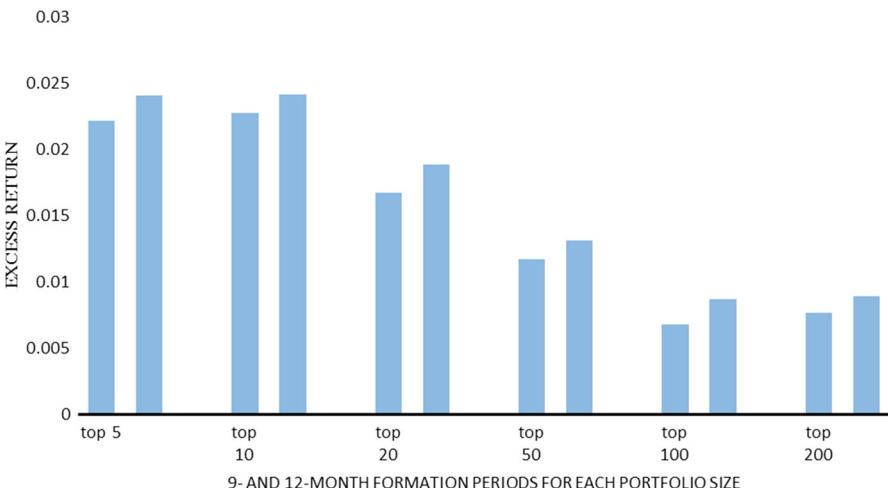
**Fig. 1** Trigger values and excess returns—distance approach. The figure shows monthly excess returns averaged across sub-periods, 1980–2014

Turning to specific parameter values, profitability is generally a hump-shaped function of the trigger value. Further investigation reveals an intuitive reason for this finding. First, increasing the trigger captures only pairs that have diverged by a greater amount, which raises profitability from a given trade, but reduces the number of pair trades. A very high trigger, such as three standard deviations ( $n = 3$ ), produces high profitability but from relatively few trades open at any given time. On the other hand, a low trigger results in more trading opportunities, but the returns on a given trade are smaller as the price divergence is smaller when the trade is initiated.

To give an example of the importance of the trigger value from these results, consider the cointegration approach and the top-5 pairs portfolio. The optimal trigger value is one-half the standard value of  $n = 2$ . Most importantly, with the (in-sample) optimal trigger value  $n = 1$ , profits are 40% larger than they would be if one were instead to use the standard trigger value.

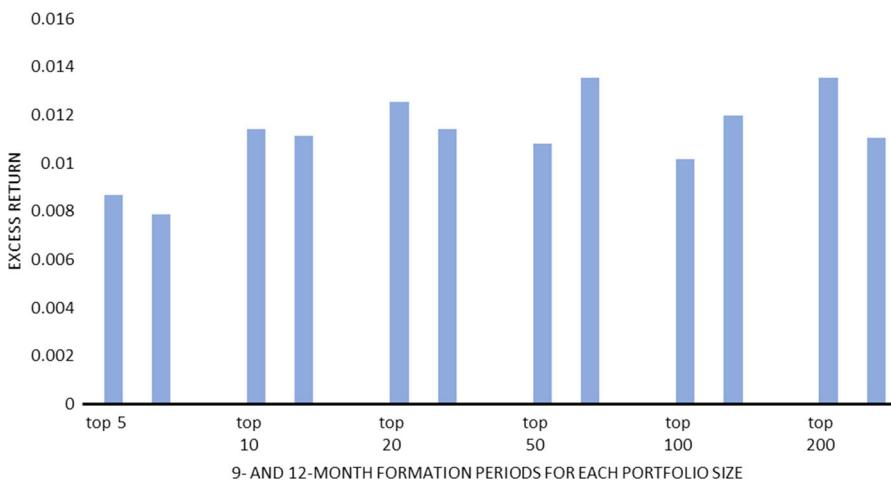


**Fig. 2** Trigger values and excess returns—cointegration approach. The figure shows monthly excess returns averaged across sub-periods, 1980–2014



**Fig. 3** Formation period and excess returns—distance approach. The figure shows monthly excess returns averaged across sub-periods, 1980–2014

Finally, we note that formation period length also matters, but to a lesser extent quantitatively than the trigger value, at least for our limited comparison of 9 and 12 months. It would be of interest in future research to examine longer formation periods. A longer formation period may well be desirable; perhaps especially with the cointegration approach as there the first objective is to discover long-run relationships with a statistical technique designed for this purpose.



**Fig. 4** Formation period and excess returns—cointegration approach. The figure shows monthly excess returns averaged across sub-periods, 1980–2014

## 5.2 Trading costs: one-day-later rule

Many studies of pairs trading make an adjustment for transaction costs, the most common of which is due to [Gatev et al. \(2006\)](#) and termed the “one-day-later rule”. This amounts to opening and closing trades 1 day later than when the trigger is first breached. The reasoning behind this assumption is to account for bid-ask spreads in computing returns from pairs trading. This can be a major component of transaction costs, particularly if some stocks have low liquidity.

More specifically, the one-day-later rule is employed to reduce the effects on calculated returns of bid-ask bounce associated with using daily closing stock prices from the Center for Research in Security Prices (CRSP) database. CRSP uses the daily closing price as the index of the daily stock price, which will be the price for the last transaction of the day, which in turn can either be at the bid or the ask depending on the order. The excess return calculated from these prices could be biased upward because pairs trading is a contrarian investment strategy, and thus, when we open the position, we may be buying underpriced stocks at the bid and selling overpriced ones at the ask (see [Gatev et al. 2006](#)). Delaying trades until 1 day after the trigger is breached for a stock pair will correct for this bid-ask bias, assuming that closing prices 1 day later are equally likely to be at bid or ask, which is, arguably, a reasonable assumption. Of course, this assumption will not pick up returns on pairs that converge extremely rapidly (i.e., in 1 day). Note also that this adjustment does not capture other types of trading costs—commissions, market impact costs of trading, and short-selling costs. Such costs are rarely accounted for in studies of pairs trading; we will consider those additional costs later in this section.

Tables 3 and 4 present the results for the one-day-later rule. Note first that with this adjustment, there is considerably more variation in the optimal in-sample parameterization, with many fewer instances, where the standard parameterization is optimal.

**Table 3** Pairs trading with distance approach—one-day-later rule

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
1980s						
Trigger	1.5	3	2	1.5	1.5	1
Formation period	9	12	12	12	12	9
Mean excess return	0.0118	0.0623	0.0130	0.0110	0.0106	0.0114
Standard deviation	0.0192	0.0947	0.0164	0.0216	0.0298	0.0397
t-statistic	5.0000	6.6420	8.0070	5.1534	3.6118	2.9356
Jensen's alpha	0.0130	0.0096	0.0129	0.0114	0.0106	0.0115
t-statistic	7.5345	6.3340	7.4501	5.4521	3.4334	2.8988
Market beta	-0.1122	-0.0062	0.0042	-0.0390	0.0034	-0.0133
t-statistic	-1.9862	-0.1770	0.1309	-1.0218	0.0707	-0.1911
Skewness	-1.8131	0.5815	-1.9456	-1.1490	-1.7974	-0.6568
Kurtosis	8.5121	1.1723	9.9275	4.0632	4.9247	3.1078
Minimum	-0.0879	-0.1498	-0.0753	-0.0736	-0.1040	-0.1430
Maximum	0.0642	0.4006	0.0586	0.0660	0.0686	0.1190
Positive returns (%)	0.8286	0.7745	0.8725	0.8431	0.8627	0.8095
1990s						
Trigger	2	1.5	1.5	2.5	1	1.5
Formation period	12	12	9	9	12	9
Mean excess return	0.0099	0.0095	0.0074	0.0063	0.0050	0.0056
Standard deviation	0.0104	0.0098	0.0075	0.0100	0.0050	0.0129
t-statistic	9.6523	9.8807	10.0483	6.4907	10.0346	4.4413
Jensen's alpha	0.0093	0.0094	0.0074	0.0059	0.0049	0.0056
t-statistic	6.6360	7.6196	9.9101	6.2875	8.3891	3.5389
Market beta	0.0405	0.0067	0.0085	0.0329	0.0049	-0.0007
t-statistic	1.0896	0.1992	0.4909	1.0547	0.3417	-0.0274
Skewness	0.3943	0.3139	-0.2090	1.6110	0.0091	6.0797
Kurtosis	0.6338	0.2739	0.2898	7.3076	-0.2962	48.8413
Minimum	-0.0150	-0.0140	-0.0166	-0.0197	-0.0082	-0.0153
Maximum	0.0402	0.0382	0.0241	0.0563	0.0164	0.1144
Positive returns (%)	0.8627	0.8627	0.8381	0.8000	0.8333	0.7333
2000s						
Trigger	1.5	1	0.5	0.5	0.5	2
Formation period	12	12	9	9	9	9
Mean excess return	0.0037	0.0066	0.0033	0.0036	0.0023	0.0012
Standard deviation	0.0202	0.0133	0.0093	0.0137	0.0201	0.0226
t-statistic	2.3051	6.3070	4.6287	3.3869	1.4555	0.6713
Jensen's alpha	0.0037	0.0066	0.0034	0.0036	0.0023	0.0011
t-statistic	2.3940	5.6460	4.5939	3.0483	1.5400	0.6637
Market beta	-0.0178	0.0164	-0.0055	0.0074	0.0131	0.0272
t-statistic	-0.4543	0.5502	-0.3660	0.4087	0.4245	0.6354

**Table 3** continued

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
Skewness	-6.9494	0.2663	0.3407	1.4616	1.1464	-1.4349
Kurtosis	70.5919	9.2239	8.0846	8.0939	11.5110	7.1697
Minimum	-0.2032	-0.0583	-0.0455	-0.0549	-0.0892	-0.1045
Maximum	0.0634	0.0669	0.0398	0.0678	0.1206	0.0623
Positive returns (%)	0.8827	0.8704	0.7515	0.6182	0.5879	0.5697

Statistics are for the 1-month excess return distributions for pairs trading from the distance approach with the optimal strategy and the one-day-later rule

Most striking, however, is that profitability is dramatically lower in both approaches, and especially in the more recent sub-samples. Specifically, for the distance approach, returns in the 1980s and 1990s drop by roughly 50% relative to the no-waiting case, and in the 2000s period, they drop by an even larger amount: from upward of 2% with no waiting to less than 0.5% with the one-day-later rule. Nevertheless, excess returns and Jensen alphas remain statistically significant for the distance approach.

The drop in returns after the 1980s, however, is most dramatic for the cointegration approach. Specifically, pairs trading returns in the 1990s and 2000s are economically and statistically insignificantly different from zero for every portfolio size in both sub-periods. These facts are evident in Figs. 5, 6, and 7, which show the cumulative returns from the two approaches for the top-20 portfolios versus that of the equity premium (using the S&P 500 returns net of the risk-free rate). In this example, the most striking observation is that the cointegration approach would have been a money-losing strategy in both the 1990s and 2000s, and this is using an optimal in-sample parameterization of the trading system—the performance of the pairs-trading portfolio would have been even worse than shown had one used the standard parameterization as it was not optimal in any sub-period for the top-20 portfolio.

We next turn to discussion of risk-adjusted returns. We begin with Sharpe ratios. Standard Sharpe ratios may be problematic for our purposes, because the return distributions (see Tables 1, 2, 3, and 4) are highly non-normal.<sup>6</sup> For this reason, we present in Table 5 both standard and modified Sharpe ratios. The latter differ from the standard ones in that the standard deviation in the denominator of the standard Sharpe ratio is replaced by the modified value at risk for a non-normal return distribution. What is important about this correction is that both skew and kurtosis enter directly into the risk-adjustment term via a Cornish-Fisher asymptotic expansion for the quantiles of a non-normal distribution. This modified Sharpe ratio is known as the Cornish-Fisher Sharpe ratio (e.g., Bramante 2013).

The main implications from both standard and modified Sharpe ratios are consistent with the above inferences from simple excess returns and Jensen alphas: (1) in the 1980s, Sharpe ratios were much larger than for the equity premium with either pairs-trading approach; (2) for the distance approach, they remain higher than the equity

<sup>6</sup> A normal distribution has a skew of 0 and a kurtosis of 3.

**Table 4** Pairs trading with cointegration approach—one-day-later rule

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
1980s						
Trigger	1.5	2	1.5	3	2	0.3
Formation period	9	9	9	12	12	12
Mean excess return	0.0126	0.0131	0.0138	0.0113	0.0119	0.0112
Standard deviation	0.0307	0.0277	0.0243	0.0277	0.0256	0.0548
t-statistic	4.2070	4.8492	5.7996	4.1256	4.6955	2.0621
Jensen's alpha	0.0127	0.0137	0.0153	0.0108	0.0116	0.0137
t-statistic	4.6365	5.1287	6.6287	3.7058	3.9880	2.0629
Market beta	0.0243	-0.0403	-0.1302	0.0934	0.0464	-0.0674
t-statistic	0.4661	-0.8008	-2.1447	1.9173	0.8914	-0.8352
Skewness	-0.0226	0.1623	-1.2251	-0.8983	-1.4723	-3.4804
Kurtosis	1.1035	1.6690	4.3166	2.1876	4.6581	23.3804
Minimum	-0.0861	-0.0635	-0.0959	-0.0847	-0.0965	-0.4704
Maximum	0.1078	0.1169	0.0682	0.0719	0.0678	0.1211
Positive returns (%)	0.6952	0.7143	0.7619	0.7549	0.8431	0.7451
1990s						
Trigger	3	1	1.5	1.5	2.5	2.5
Formation period	12	12	12	12	12	12
Mean excess return	0.0030	-0.0018	-0.0027	-0.0016	0.0011	0.0021
Standard deviation	0.0399	0.0296	0.0274	0.0205	0.0228	0.0260
t-statistic	0.7492	-0.6208	-1.0003	-0.8119	0.4842	0.8028
Jensen's alpha	0.0031	-0.0019	-0.0039	-0.0031	0.0002	0.0015
t-statistic	0.6813	-0.5901	-1.5329	-1.6130	0.0724	0.5793
Market beta	-0.0270	0.0037	0.0825	0.0982	0.0453	0.0360
t-statistic	-0.2796	0.0418	1.0016	2.0637	0.6838	0.7224
Skewness	-0.6415	-0.1033	-0.0748	-0.3258	0.4909	-0.9279
kurtosis	2.4533	0.5586	1.2281	0.5409	1.8462	3.4085
Minimum	-0.1259	-0.0797	-0.0886	-0.0601	-0.0565	-0.0970
Maximum	0.1130	0.0815	0.0772	0.0566	0.0932	0.0678
Positive returns (%)	0.5784	0.4706	0.5098	0.4902	0.4902	0.5686
2000s						
Trigger	2.5	1.5	1	1.5	1.5	3
Formation period	9	9	9	9	9	9
Mean excess return	0.0002	0.0000	-0.0012	-0.0018	-0.0001	0.0021
Standard deviation	0.0484	0.0389	0.0268	0.0257	0.0241	0.0291
t-statistic	0.0453	-0.0114	-0.5905	-0.8866	-0.0532	0.9415
Jensen's alpha	0.0006	-0.0010	-0.0052	-0.0036	-0.0024	-0.0063
t-statistic	0.1850	-0.2699	-1.7874	-1.4372	-0.9356	-1.6545
Market beta	-0.1627	-0.1193	0.0152	-0.0752	-0.0221	0.0097
t-statistic	-1.2747	-1.0973	0.1548	-1.0640	-0.3325	0.1090

**Table 4** continued

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
Skewness	-0.9344	-0.7949	-0.0163	-0.7707	0.0148	0.0587
Kurtosis	1.8955	3.7900	1.3856	2.5024	3.3880	0.4961
Minimum	-0.1979	-0.1992	-0.0814	-0.1134	-0.0852	-0.0828
Maximum	0.1117	0.1062	0.0931	0.0639	0.1021	0.0795
Positive returns (%)	0.5818	0.5394	0.4545	0.5030	0.5091	0.5212

Statistics are for the 1-month excess return distributions for pairs trading from the cointegration approach with the optimal strategy and the one-day-later rule

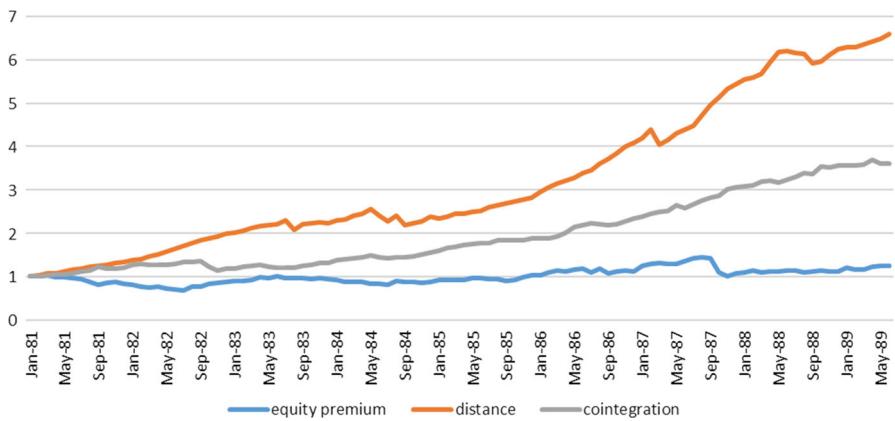
premium in the 1990s and 2000s; and (3) the cointegration approach produces lower Sharpe ratios than the market throughout the 1990s and 2000s periods.

Figures 5, 6, and 7 suggest that an important reason for the superior risk-adjusted performance of pairs trading with the distance approach relative to the equity premium is not only that raw returns are higher but also that pairs trading clearly produces much less volatile returns relative to the overall equity market. However, returns from pairs trading are not risk free—the standard deviation is not zero. A remaining question is what underlies the risk in pairs trading. To investigate this, we next decompose excess returns using the Fama-French factor approach (Tables 6, 7). It is strikingly apparent that none of the four Fama-French factors are driving returns from pairs trading with either approach, in any sub-period: Fama-French alphas are very similar quantitatively to Jensen alphas and, more importantly, to mean excess returns. We conclude that the variation in pairs trading excess returns is not explained by standard factor models, and is arguably simply “arbitrage risk” (Gatev et al. 2006).

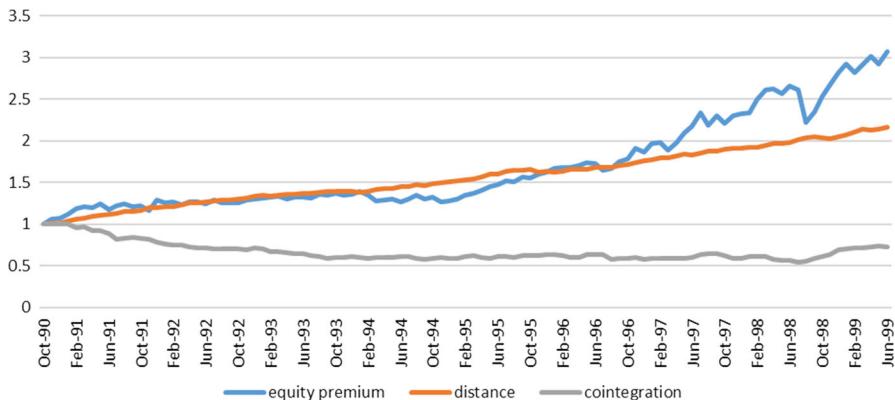
In sum, the cointegration approach has not been a profitable strategy throughout the past 25 years even when the parameterization of the trading system is chosen optimally on an in-sample basis. The distance approach performs better, but excess returns have clearly fallen considerably over time: the top-20 portfolio excess returns halved in the 1990s from the 1980s, and halved again from the 1990s to the 2000s. In the full 2000–2014 period, the excess returns on this portfolio are just 33 bp. The main unresolved question is whether such returns with the distance approach are still economically significant if, for example, we break this time period down into finer sub-periods and, more importantly, when we account for other costs associated with pairs trading.

### 5.3 Is pairs trading still profitable?

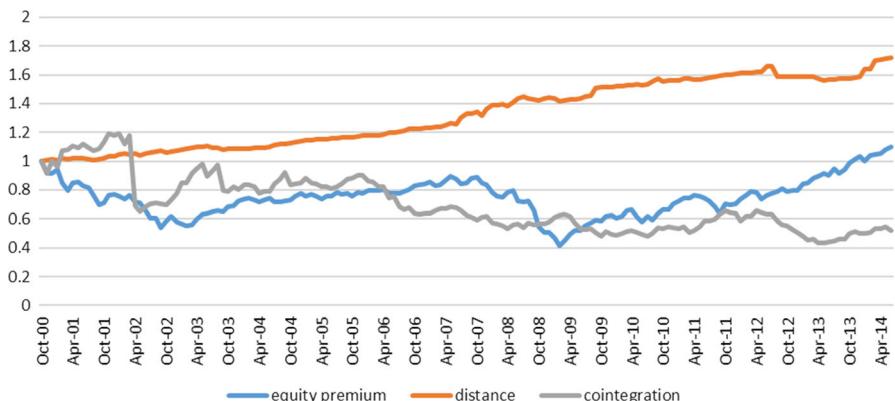
The evidence presented so far clearly shows the decline in pairs-trading returns across three sub-periods. In the full 2000–2014 period, excess returns with the distance approach range from a low of 12 bp to a high of 66 bp on the top-200 and top-10 portfolios, respectively, and risk-adjusted returns (Fama-French alphas) are nearly identical. As the objective here is to assess whether pairs trading with the distance approach is profitable with the most recent data and with account taken of additional transaction costs, it is instructive first to examine the temporal pattern of returns in



**Fig. 5** Cumulative returns for the 1980s. Cumulative returns for the pairs portfolios refer to the top-20 portfolio



**Fig. 6** Cumulative returns for the 1990s. Cumulative returns for the pairs portfolios refer to the top-20 portfolio



**Fig. 7** Cumulative returns for the 2000s. Cumulative returns for the pairs portfolios refer to the top-20 portfolio

**Table 5** Sharpe ratios

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
Basic sharpe ratio						
1980s						
Distance	0.615	0.658	0.793	0.510	0.358	0.286
Cointegration	0.411	0.473	0.566	0.408	0.465	0.189
S&P 500	0.084	0.084	0.084	0.084	0.084	0.084
1990s						
Distance	0.956	0.978	0.981	0.633	0.994	0.433
Cointegration	0.074	-0.061	-0.099	-0.080	0.048	0.079
S&P 500	0.302	0.302	0.302	0.302	0.302	0.302
2000s						
Distance	0.181	0.496	0.360	0.264	0.113	0.052
Cointegration	0.004	-0.001	-0.046	-0.069	-0.004	0.073
S&P 500	0.064	0.064	0.064	0.064	0.064	0.064
Modified sharpe ratio						
1980s						
Distance	0.162	0.474	0.205	0.164	0.111	0.093
Cointegration	0.187	0.224	0.182	0.147	0.145	0.034
S&P 500	0.023	0.023	0.023	0.023	0.023	0.023
1990s						
Distance	0.817	0.855	0.632	0.511	0.791	-0.085
Cointegration	0.024	-0.024	-0.036	-0.029	0.021	0.023
S&P 500	0.102	0.102	0.102	0.102	0.102	0.102
2000s						
Distance	0.032	0.132	0.101	0.127	0.032	0.012
Cointegration	0.001	0.000	-0.017	-0.021	-0.001	0.032
S&P 500	0.022	0.022	0.022	0.022	0.022	0.022

the 2000s period. For the top-20 portfolio, for example, average excess returns were 33 bp for the full period 2000–2014, 26 bp for 2000–2004, 48 bp for 2005–2009, and 23 bp for 2010–2014. The global financial crisis in 2007–2008 clearly presented more favorable pairs-trading opportunities—raising excess returns by roughly half relative to the full 2000s sample. This fact has been pointed out by others—market turmoil presents better prospects for pairs-trading strategies (see [Do and Faff 2010, 2012; Deaves et al. 2013](#)). Before discussing transaction costs, we also note that it is important to recognize that these return calculations are generous measures as they use an optimal in-sample parameterization. For instance, if one were to simply use the standard parameterization, the returns drop significantly: on the top-20 portfolio, for example, they drop from 36 bp to just 13 basis points for the full 2000–2014 period. The key question is whether, even apart from the relatively small residual risk that is involved in pairs trading—arbitrage risk—these returns are economically meaningful.

**Table 6** Factor analysis—pairs trading with distance approach

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
1980s						
FF alpha	0.0117	0.0106	0.0127	0.0100	0.0115	0.0119
t-statistic	5.5307	5.4290	6.9162	3.9834	3.7931	2.7208
FF beta	-0.1076	-0.0464	-0.0045	0.0081	0.0073	-0.0040
t-statistic	-1.8500	-1.1124	-0.1184	0.1759	0.1134	-0.0492
FF size	0.0510	-0.0303	-0.0201	-0.0577	0.1125	0.0182
t-statistic	0.6148	-0.5519	-0.3236	-0.5578	1.1959	0.1281
FF value	0.0471	-0.1096	0.0001	0.1332	-0.0391	0.0314
t-statistic	0.6165	-1.4221	0.0011	1.3007	-0.3140	0.1875
FF momentum	0.0607	0.0213	0.0291	-0.0488	-0.1493	-0.1346
t-statistic	1.2873	0.3791	0.6313	-0.9048	-1.8197	-1.2307
1990s						
FF alpha	0.0097	0.0098	0.0070	0.0063	0.0046	0.0057
t-statistic	6.9268	7.8414	8.3597	6.1724	6.9171	4.0863
FF beta	0.0206	-0.0131	0.0042	0.0397	0.0070	-0.0200
t-statistic	0.4406	-0.3339	0.2171	1.0975	0.4617	-0.5078
FF size	-0.0330	-0.0196	0.0283	0.0577	0.0185	-0.0175
t-statistic	-0.6310	-0.4166	0.9457	1.4975	0.9306	-0.2676
FF value	-0.0553	-0.0515	0.0058	0.0125	0.0168	-0.0567
t-statistic	-0.8909	-1.0416	0.1599	0.3662	0.7857	-1.1326
FF momentum	0.0098	0.0018	0.0411	-0.0606	0.0338	0.0300
t-statistic	0.2639	0.0556	1.3725	-1.6759	1.9277	0.4601
2000s						
FF alpha	0.0038	0.0063	0.0036	0.0038	0.0024	0.0013
t-statistic	2.3523	5.0938	4.6496	3.1686	1.5562	0.6933
FF beta	-0.0373	0.0094	-0.0057	0.0032	0.0190	0.0513
t-statistic	-0.6946	0.2630	-0.3033	0.1262	0.5899	0.9090
FF size	0.0183	0.0256	-0.0526	-0.0080	-0.0411	-0.0780
t-statistic	0.2595	0.6380	-1.9334	-0.1804	-1.0941	-1.3767
FF value	-0.0193	0.0455	-0.0111	-0.0499	-0.0111	-0.0009
t-statistic	-0.3798	0.9992	-0.3944	-1.5942	-0.3117	-0.0189
FF momentum	-0.0263	-0.0006	-0.0141	-0.0020	0.0085	0.0214
t-statistic	-1.3976	-0.0377	-1.0118	-0.0906	0.2691	0.3862

It is impossible to directly measure the additional transaction costs from each pair trade. The best that can be done is to provide some evidence on the magnitudes of such costs on average for trades executed by professional investors. [Do and Faff \(2012\)](#) estimate three transaction costs for pairs trading by professional investors: commission expenses, market impact costs associated with relative value arbitrage, and costs associated with short selling. They report that the costs associated with

**Table 7** Factor analysis—pairs trading with cointegration approach

	Top 5	Top 10	Top 20	Top 50	Top 100	Top 200
1980s						
FF alpha	0.0121	0.0135	0.0140	0.0112	0.0107	0.0102
t-statistic	3.9484	4.5209	4.8041	4.0363	3.3809	1.4227
FF beta	-0.0466	-0.0680	-0.1219	0.0938	0.0842	-0.0415
t-statistic	-0.8304	-1.0789	-1.9117	1.3983	1.2578	-0.4506
FF size	0.0585	-0.0311	-0.0427	-0.1056	-0.2052	0.2524
t-statistic	0.4639	-0.2475	-0.3071	-0.9278	-2.2638	1.2227
FF value	-0.0954	-0.0483	0.0445	-0.0298	0.0991	0.1852
t-statistic	-0.7801	-0.4407	0.4015	-0.2317	0.6833	0.6709
FF momentum	0.2515	0.0726	0.0283	-0.0250	-0.0046	0.3263
t-statistic	3.5226	1.1001	0.4446	-0.3102	-0.0663	2.2676
1990s						
FF alpha	0.0028	-0.0015	-0.0014	-0.0034	0.0020	0.0008
t-statistic	0.5248	-0.4274	-0.5107	-1.6868	0.8677	0.2369
FF beta	-0.0236	-0.0845	-0.0304	0.1076	-0.0234	0.0680
t-statistic	-0.2112	-1.0356	-0.4400	2.0453	-0.4075	1.3128
FF size	-0.0347	-0.1176	-0.0871	-0.0490	-0.0125	0.1205
t-statistic	-0.2811	-1.1861	-0.9093	-0.6021	-0.1871	1.0057
FF value	-0.0008	-0.2215	-0.3251	0.0310	-0.2227	0.1198
t-statistic	-0.0048	-1.8921	-2.9916	0.3314	-2.4574	0.9587
FF momentum	0.0605	0.1436	-0.0365	0.0298	-0.0147	0.0321
t-statistic	0.3798	1.4314	-0.3258	0.3659	-0.1404	0.3436
2000s						
FF alpha	0.0002	0.0020	-0.0032	-0.0029	-0.0020	-0.0039
t-statistic	0.0616	0.5213	-1.1134	-1.0661	-0.8033	-0.9514
FF beta	-0.2272	-0.2402	0.0576	-0.1189	-0.0303	-0.0349
t-statistic	-1.6573	-2.0170	0.5154	-1.4349	-0.4339	-0.3591
FF size	-0.2756	-0.1614	-0.1618	-0.0669	-0.2073	-0.1688
t-statistic	-2.0015	-1.2987	-1.6253	-0.9369	-2.4012	-1.7786
FF value	-0.1573	-0.3673	0.0673	-0.1276	-0.0550	-0.1900
t-statistic	-0.9357	-2.2743	0.5128	-1.1639	-0.7293	-1.5374
FF momentum	-0.0068	-0.1558	-0.3746	-0.0491	-0.0996	-0.2279
t-statistic	-0.0375	-0.9802	-3.8742	-0.6962	-1.4830	-2.0181

short sales are very small (less than 1% loan fee per annum), and [Gatev et al. \(2006\)](#) reach similar conclusions. Thus, short-selling costs will likely not greatly alter the profitability of pairs trading.

For each one-way trade, [Do and Faff \(2012\)](#) report average commission fees of 9 bp for the 2000s (their sample ends in 2009) and market impact trading costs of around 20 bp. Each pair trade involves two roundtrip trades (four one-way trades).

Thus, using these estimates for the 2000s period, total transaction costs per pair trade are on the order of 116 bp (i.e.,  $4 \times (9 + 20)$ ). On average, we trade each pair twice in a 6-month trading period, and thus, transaction costs are on the order of 232 bp per 6 months, which translates to 39 bp per month. Based on these estimates of transaction costs for the 2000s period, and returns ranging from 12–66 bp on a monthly basis, returns from pairs trading with the distance approach and using an optimal in-sample parameterization are at best 17 bp monthly (which corresponds to the top-10 portfolio).

In sum, it is difficult to conclude that the returns to pairs trading in the 2000s with any of the two main approaches, and accounting for transaction costs, are on average economically significant. The most favorable case is that the distance approach can yield economically significant returns mainly during market disruptions. Although some market disruptions are obvious (such as the 2007–2008 financial crisis), which could provide a guide for pairs traders, other market disruptions may not be. For instance, we experimented with using the VIX implied volatility index as an indicator of market stress and conditioned pairs trading on it. In the 2000s period, we found no correlation between the VIX and the returns to pairs trading. Further research into what types of market conditions are most favorable to pairs trading would be worthwhile.

## 6 Conclusion

This paper studies alternative techniques for identifying stock pairs in a pairs-trading strategy and explores a range of parameterizations of the trading system. Parameterization of the trading system matters for the profitability of pairs trading, and the standard one used in the literature does not appear to be the optimal one in the majority of cases studied.

The profitability of pairs trading has declined over time, and this has been noted in the literature. However, we find that the cointegration approach to pairs trading would not have been a profitable investment strategy throughout the 1990s and 2000s, and quite likely would have returned a negative amount when all transaction costs are accounted for properly. The distance approach also shows declining returns over time, but the crucial question is whether the very low, but statistically significant, returns in recent years are high enough to cover all relevant transaction costs. Our analysis suggests they may not be.

Our analysis leads to a number of suggestions for future research. One is to experiment with alternative implementations of the cointegration approach. This approach readily identifies cointegrated stock prices, but the number of such “pairs” identified just with cointegration is very large. As a result, some additional structure must be employed to reduce the set of pairs to a set of the best ones. In this paper, we used a strategy that narrows this set by requiring mean reversion of the price spreads, two-way Granger causality in the prices of a given pair, and quantitatively similar market betas. The point is that there are other refinements to the cointegration approach that should be explored with the aim of seeing whether they produce better pairs-trading returns.

A second avenue worth considering is the role of the formation period length. We have studied a shorter formation period than the standard 12-months period, but it could be that a longer formation period might be beneficial for pairs-trading profitability. In

the cointegration approach, for instance, it could very well be the case that more than 1 year of historical price data would help in identifying cointegrated prices and their possible mean reversion. These are important topics for future work.

Finally, it has come to be an accepted belief that difficult market conditions produce the best opportunities for pairs trading. More research into exactly what types of market conditions, and how they are identified in real time, are most favorable to pairs trading could produce important findings.

**Acknowledgements** We are grateful to two referees for suggestions. This paper represents the views of the authors and does not represent the views, practices, or policies of Vestcor Investment Management Corporation.

## References

- Baronyan, S., Boduroglu, I.I., Sener, E.: Investigation of Stochastic Pairs Trading Strategies Under Different Volatility Regimes. *Manch. School* **78**(s1), 114–134 (2010)
- Bolgun, K.E., Kurun, E., Guven, S.: Dynamic Pairs Trading Strategy for the Companies Listed in the Istanbul Stock Exchange. *Int. Rev. Appl. Financ. Issues Econ.* **1**, 37–57 (2010)
- Bookstaber, R.: *A Demon Of Our Own Design*. Wiley, New York (2006)
- Bramante, R.: An Approach to Ranking the Hedge Fund Industry. In: Giudici, P., Ingrassia, S., Vichi, M. (eds.) *Statistical Models for Data Analysis*, pp. 63–71. Springer, New York (2013)
- Broussard, J.P., Vaihekoski, M.: Profitability of Pairs Trading Strategy in an Illiquid Market with Multiple Share Classes. *J. Int. Financ. Mark. Inst. Money* **22**(5), 1188–1201 (2012)
- Caldeira, J.F., Moura, G.V.: Selection of a Portfolio of Pairs Based on Cointegration: A Statistical Arbitrage Strategy. *Rev. Braz. Financ.* **11**(1), 49–80 (2013)
- Chen, H., Chen, S., Li, F.: Empirical Investigation of an Equity Pairs Trading Strategy. Mimeo, New York (2012)
- Deaves, R., Liu, J., Miu, P.: Pairs Trading in Canadian Markets: Pay Attention to Inattention. *Canadian Investment Review, Analysis and Research* (2013)
- Do, Binh: Faff, Robert: Does Simple Pairs Trading Still Work? *Financ. Anal. J.* **66**(4), 83–95 (2010)
- Do, B., Faff, R.: Are Pairs Trading Profits Robust to Trading Costs? *J. Financ. Res.* **35**(2), 261–287 (2012)
- Elliott, R., Van Der Hoek, J., Malcolm, W.: Pairs Trading. *Quant. Financ.* **5**(3), 271–276 (2005)
- Engelberg, J., Gao, P., Jagannathan, R.: An Anatomy of Pairs Trading: The Role of Idiosyncratic News. Common Information and Liquidity. Mimeo, New York (2009)
- Gatev, E., Goetzmann, W., Rouwenhorst, G.: Pairs Trading: Performance of a Relative-Value Arbitrage Rule. *Rev. Financ. Stud.* **19**(3), 797–827 (2006)
- Gregory, I., Knox, P., Oliver-Wald, C.: Analytical Pairs Trading Under Different Assumptions on the Spread and Ratio Dynamics. Mimeo, New York (2011)
- Herlemont, D.: Pairs Trading, Convergence Trading, Cointegration. Working paper, YATS Finances & Technologies, USA (2004)
- Huck, N.: Pairs Selection and Outranking: An Application to the S&P 100 Index. *Eur. J. Oper. Res.* **196**(2), 819–825 (2009)
- Huck, N.: Pairs Trading and Outranking: The Multi-Step-Ahead Forecasting Case. *Eur. J. Oper. Res.* **207**(3), 1702–1716 (2010)
- Jacobs, H., Weber, M.: Losing Sight of the Trees for the Forest? Attention Allocation and Anomalies. Mimeo, New York (2013)
- Kanemura, T., Rachev, S.T., Fabozzi, F.: The Application of Pairs Trading to Energy Futures Markets. mimeo, New York (2008)
- Lehmann, B.: Fads, Martingales and Market Efficiency. *Q. J. Econ.* **105**(1), 1–28 (1990)
- Lin, Y., McCrae, M., Gulati, C.: Loss Protection in Pairs Trading Through Minimum Profit Bounds: A Cointegration Approach. *J. Appl. Math. Decis. Sci.* **2006**, 1–14 (2006)
- Lucey, M., Walshe, D.: European Equity Pairs Trading: The Effect of Data Frequency on Risk and Return. *J. Bus. Theory Pract.*, 329–341 (2013)

- Mudchanatongsuk, S., Primbs, J., Wong, W.: Optimal Pairs Trading: A Stochastic Control Approach. In: American Control Conference, pp. 1035–1039 (2008)
- Nath, P.: High Frequency Pairs Trading with US Treasury Securities: Risks and Rewards for Hedge Funds. Mimeo, New York (2003)
- Papadakis, G., Wysocki, P.: Pairs Trading and Accounting Information. Mimeo, New York (2007)
- Peng, L., Xiong, W.: Investor Attention, Overconfidence and Category Learning. *J. Financ. Econ.* **80**, 563–602 (2006)
- Perlin, M.S.: Evaluation of Pairs Trading Strategy at the Brazilian Financial Market. *J. Deriv. Hedge Funds* **15**(2), 122–136 (2009)
- Schmerken, I.: Finding the Perfect Pair. *Wall Street Technol.* **24**(7), 1 (2006)
- Tourin, A., Yan, R.: Dynamic Pairs Trading using the Stochastic Control Approach. *J. Econ. Dyn. Control* **37**, 1972–1981 (2013)
- Vidyamurthy, G.: Pairs Trading-Quantitative Methods and Analysis. Wiley, New York (2004)

**R. Todd Smith** is a Professor in the Department of Economics. He received a BA Honors (Economics) from the University of Saskatchewan and a Ph.D. (Economics and Finance) from Queen's University. He served as an Economist in the Capital Markets and Financial Studies Division of the Research Department of the International Monetary Fund during 1994–1997 and as Senior Economist in the Financial Stability Division and the Emerging Markets Division of the International Capital Markets Department of the IMF during 2000–2002. Smith's research and teaching interests are in financial economics (investments), macroeconomics, and international finance.

**Xun Xu** is currently working at Vestcor Investment Management Corporation (previous New Brunswick Investment Management Corporation) as a Quantitative Investment Analyst. She received a BA (Accounting) from the University of Tongji in Shanghai, a Masters (Economics) from the Lakehead University in Canada and a Ph.D. (Economics) from the University of Alberta. She is interested in financial economics and quantitative modeling of financial markets.