INTERNATIONAL
REVIEW OF FINANCE

International Review of Finance, 16:2, 2016: pp. 307–319
DOI: 10.1111/irfi.12074

Statistical Arbitrage with Pairs Trading

AHMET GÖNCÜ[†,‡] AND ERDİNÇ AKYILDIRIM[§,¶]

[†]Institute of Quantitative Finance, Xian Jiaotong Liverpool University, Suzhou, China,
[‡]Bogazici University, Center for Economics and Econometrics, Istanbul, Turkey,
[§]Department of Banking and Finance, Akdeniz University, Antalya, Turkey, and
[¶]Bogazici University, Center for Applied Research in Finance, Istanbul, Turkey

ABSTRACT

We analyze statistical arbitrage with pairs trading assuming that the spread of two assets follows a mean-reverting Ornstein–Uhlenbeck process around a long-term equilibrium level. Within this framework, we prove the existence of statistical arbitrage and derive optimality conditions for trading the spread portfolio. In the existence of uncertainty in the long-term mean and the volatility of the spread, statistical arbitrage is no longer guaranteed. However, the asymptotic probability of loss can be bounded as a function of the standard error of the model parameters. The proposed framework provides a new filtering technique for identifying best pairs in the market. Backtesting results are given for some of the pairs of stocks that are studied in the literature.

I. INTRODUCTION

In today's financial markets, there is an increasing tendency to utilize algorithmic trading in the presence of large number of assets and rapid flow of information. The existing trend of algorithmic trading forces the investors to use advanced quantitative techniques in order to generate trading signals. In the general sense, statistical arbitrage refers to trading strategies that generate almost sure profits asymptotically via trading signals generated from quantitative models, whereas pure arbitrage is a special case of statistical arbitrage that generates almost sure profits in finite time. In this study, we consider one of the most commonly used statistical arbitrage techniques known as pairs trading.

Pairs trading can be considered as the first generation of statistical arbitrage strategies that are used to exploit financial markets that are out of equilibrium. Pairs trading assumes that while markets may not be in equilibrium, over time they move to a rational equilibrium, and the trader has an interest to take maximum advantage of the deviations from the equilibrium. For interested readers, more detailed description of pairs trading can be found in Vidyamurthy (2004).

Whistler (2004), Ehrman (2006), Gatev et al. (2006), and Litterman (2004). Amongst others, empirical studies such as Avellaneda and Lee (2010), Gatev et al. (2006), Do and Faff (2010, 2012), and Huck and Afawubo (2015) show that pairs trading is still a profitable statistical arbitrage strategy.

In this study, we utilize a mean-reverting stochastic process to model the spread of two stocks, where the examples of this approach can be found in Zeng and Lee (2014), Elliot et al. (2005), Cummins and Bucca (2012), and Bertram (2010). In Bertram (2010), optimal buy/sell thresholds are derived by maximizing the expected return and Sharpe ratio of a synthetic asset that follows a mean-reverting Ornstein–Uhlenbeck (OU) process, whereas in Zeng and Lee (2014), optimal thresholds for pairs trading are derived by maximizing the expected profits per unit of time. In this study, optimal thresholds are derived from the maximization of the probability of successful termination of the pairs trading strategy.

To the best of our knowledge, this is the first study that proves the existence of statistical arbitrage for the pairs trading strategy considering the mean-reverting OU process for a spread portfolio. Different from the existing literature, we propose a new criteria to derive the optimal buy/sell trigger levels.

The article is organized as follows. In the next section, we introduce the mean-reverting spread model. In Section III, we show that statistical arbitrage opportunities exist via pairs trading, whereas in Section IV, we show that whenever there is uncertainty in the long-term mean level, statistical arbitrage condition is not satisfied. In Section V, optimal thresholds for executing the pairs trading strategies are derived, and the empirical backtesting results are presented in Section VI. We conclude in Section VII.

II. SPREAD MODEL

In Avellaneda and Lee (2010), the co-integration between two stocks is defined as

$$\ln\left(S_t^A/S_0^A\right) = \alpha(t - t_0) + \gamma \ln\left(S_t^B/S_0^B\right) + \varsigma_t, \quad t \geq 0, \tag{1}$$

where the drift is often insignificant compared with the residual term ς_t. We define the excess return of the spread position as

$$X_t = \ln\left(S_t^A/S_0^A\right) - \gamma \ln\left(S_t^B/S_0^B\right) - r_f t, \tag{2}$$

where the risk-free rate is denoted by $r_f > 0$; however, in practice, owing to the short holding periods of the spread positions and the effect of the risk-free rate is negligible. One can ignore the term $r_f t$ and the constant term $-ln\left(S_0^A\right) + \gamma ln\left(S_0^B\right)$. Therefore, in empirical applications, we simply use $X_t = ln\left(S_t^A\right) - \gamma ln\left(S_t^B\right)$ for generating the buy and sell signals (e.g., Zeng and Lee (2014)).

Remark 1. Note that if one of the stocks is expected to outperform the other stock significantly, which implies an increasing spread, then this situation can be considered in the framework of statistical arbitrage discussed in Göncü (2015).

There are two widely used choices for the pre-determined ratio γ. In the first case, the trader opens the spread position at time t^* with $\gamma = S_A(t^*)/S_B(t^*)$, so that the long and short positions in assets A and B offset each other, whereas in the second case, the trader chooses a market-neutral spread. These two cases are given as

$$\gamma = \begin{cases} S_A(t^*)/S_B(t^*), t^* > 0 \text{ (no borrowing/lending case)} \\ \beta_A/\beta_B \text{ (market neutral case)}, \end{cases} \tag{3}$$

In pairs trading, when the spread between two assets or portfolios becomes larger, the trader longs the relatively cheaper stock and shorts the expensive one. Market neutrality can be achieved by taking long and short positions proportional to the ratio of the betas of the two stocks.

Given the stochastic process for the discounted cumulative trading profits, denoted as $\{v(t) : t \geq 0\}$, which is defined on a probability space (Ω, F, P), the statistical arbitrage is defined as follows (Hogan et al. 2004).

Definition 2. A statistical arbitrage is a zero initial cost, self-financing trading strategy $\{v(t) : t \geq 0\}$ with cumulative discounted value $v(t)$ such that

1. $v(0) = 0$
2. $\lim_{t \to \infty} E[v(t)] > 0$,
3. $\lim_{t \to \infty} P(v(t) < 0) = 0$, and
4. $\lim_{t \to \infty} \frac{var(v(t))}{t} = 0$ if $P(v(t) < 0) > 0$, $\forall\, t < \infty$.

In pairs trading, the dynamics of the spread is often assumed to follow a mean-reverting OU process given by (Elliot et al. 2005; Avellaneda and Lee 2010)

$$dX_t = -\rho(X_t - \mu)dt + \sigma dW_t, \tag{4}$$

where ρ is the speed of mean reversion, W_t is the standard Brownian motion, and μ is the long-term equilibrium level of the spread. The solution of Equation 4 is given by

$$X_t = X_0 e^{-\rho t} + \mu\left(1 - e^{-\rho t}\right) + \sigma \int_0^t e^{-\rho(t-s)} dW_s, \tag{5}$$

where X_t is normally distributed with $E[X_t] = X_0 e^{-\rho t} + \mu(1 - e^{-\rho t})$ and $\text{var}(X_t) = \frac{\sigma^2}{2\rho}\left(1 - e^{-2\rho t}\right)$. The stationary mean and variance are given as μ and $\sigma^2/2\rho$ as $t \to \infty$, respectively.

We use the first passage time density and the Laplace transform given in Finch and Yt (2004) for the standardized OU process, that is, $Z\bar{t} = (X_t - \mu)/\left(\sigma/\sqrt{2\rho}\right)$ and $\bar{t} = \rho t$; however, with a slight abuse of notation, we still use t instead of $\bar{t}$, although it is now scaled by ρ. The standardized OU process is given as

$$dZ_t = -Z_t dt + \sqrt{2} dW_t. \tag{6}$$

Let the first passage time to a starting from c be denoted as

$$\tau_{a,c} = min\{t \geq 0 : Z_t = a | Z_0 = c\}, \tag{7}$$

where in the standardized spread process, long-term mean becomes zero. Mean reversion of the real process in Equation 4 to the long-term mean level μ is equivalent to the mean reversion of the dimensionless process to zero (i.e., $a=0$). The probability density of the first passage time $\tau_{0,c}$ starting from the dimensionless deviation level c is given by (see Finch and Yt (2004) for details)

$$f_{0,c} = \sqrt{\frac{2}{\pi}} \frac{|c|e^{-t}}{(1-e^{-2t})^{3/2}} \exp\left(-\frac{c^2 e^{-2t}}{2(1-e^{-2t})}\right), \quad \text{for } t > 0. \tag{8}$$

The first passage time can also be represented in terms of the maximum or minimum of the OU process. Let $\tau = \min\{t \geq 0 : Z_t = 0\}$ and $M_T = max_{0 \leq t \leq T} Z_t$ and as given in Finch and Yt (2004) discarding $Z_0 = c$ the maximum of the standard OU process is given by

$$P(\tau > T) \equiv P(M_T < 0) = \frac{1}{\pi} arcsin\left(e^{-T}\right), \tag{9}$$

where for $T \to \infty$, we have $P(\tau > T) \to 0$. This means that for the mean-reverting spread process, the probability of hitting to the long-term mean level converges to one as time increases. Therefore, by defining a pairs trading strategy that exploits the deviations from the long-term equilibrium level, we can obtain statistical arbitrage profits.

Next, we describe the statistical arbitrage strategies that lead to statistical arbitrage profits in the sense of Definition 2.[1]

III. STATISTICAL ARBITRAGE VIA PAIRS TRADING

In the next theorem, we show that there are statistical arbitrage opportunities in pairs trading via long/short positions in the spread portfolio during deviations from the long-term equilibrium level.

Theorem 3. Assume that the spread of two assets follows the model given in Equation 4, then there exists statistical arbitrage opportunities via pairs trading in the sense of Definition 2.

Proof Without loss of generality, we consider the case $X_{t^*} > \mu$, whereas for $X_{t^*} < \mu$, similar arguments follow.

First, note that the initial position satisfies $v(0) = 0$ (Condition 1 in Definition 2), because the trader can set $\gamma = S_A(t^*)/S_B(t^*)$ to remove the need for initial

[1]In addition to Definition 2, pairs trading strategies discussed in this study also satisfy the description of statistical arbitrage given in Avellaneda and Lee (2010).

borrowing or lending. To show the existence of statistical arbitrage opportunities in the sense of Definition 2, we consider the following strategies.

i If $X_{t^*} > \mu$ for $t^* > 0$, then this means that in the long-term equilibrium level, the discounted return of asset B is expected to exceed the return on asset A as the spread reverts back to its long-term mean level. Therefore, at time $t = t^*$, short one unit of asset A and simultaneously long $S_A(t^*)/S_B(t^*)$ units of asset B. Giving a zero initial cost spread portfolio, that is, $v(0) = 0$. We close the spread portfolio at the first passage time to the long-term mean level μ, where $au = min\{t \geq 0 : X_t = \mu\}$. At time τ, we sell $S_A(t^*)/S_B(t^*)$ units of asset B and buy one unit of asset A. The discounted rate of return on each \$1 invested in the spread portfolio is given by $X_\tau - X_{t^*} = X_{t^*} - \mu$.

ii If $X_{t^*} < \mu$ for $t^* > 0$, then this means that in the long-term equilibrium, the discounted return of asset A is expected to exceed the return on asset B. Therefore, at time $t = t^*$, we short one unit of asset B and simultaneously long $S_A(t^*)/S_B(t^*)$ units of asset B. Giving a zero initial cost spread portfolio, that is, $v(0) = 0$. We close the spread portfolio at the first passage time to the long-term mean level μ, where $\tau = min\{t \geq 0 : X_t = \mu\}$. At time τ, we sell $S_A(t^*)/S_B(t^*)$ units of asset A and buy one unit of asset B. The discounted rate of return on each \$1 invested in the spread portfolio is given by $X_\tau - X_{t^*} = \mu - X_{t^*}$.

Without loss of generality, consider the case $\mu > X_{t^*}$, and we can write the return generated from this strategy as

$$v(t) = \begin{cases} X_t - X_{t^*} \text{ for } t < \tau \\ \mu - X_{t^*} \text{ for } t \geq \tau. \end{cases} \tag{10}$$

From the definition of X_t, the termination of the spread position implies $\mu - X_{t^*} > r_f(\tau - t^*)$, which means we generate positive excess return.
The expected value of the trading return is given by

$$(\mu - X_{t^*})P(\tau \leq t) + (1 - P(\tau \leq t))E[X_t - X_{t^*} | \tau > t], \tag{11}$$

where as $t \to \infty$, we have $lim_{t\to\infty} E[v(t)] > 0$ because $lim_{t\to\infty} P(v(t) < 0) = 0$. Thus, Condition 2 in Definition 2 is satisfied.

The variance of the cumulative discounted trading profits is bounded by

$$\text{var}(v(t)) \leq \text{var}(X_t) = \left(\frac{\sigma^2}{2\rho}\left(1 - e^{-2\rho t}\right)\right), \tag{12}$$

where as $t \to \infty$, we have $\text{var}(X(t)) = \frac{\sigma^2}{2\rho}$ and thus $lim_{t\to\infty} \text{var}(v(t))/t = 0$ as required by Condition 4 in Definition 2.

We calculate the probability of loss as

$$P(v(t) < 0) = P(v(t) < 0|\tau \leq t)P(\tau \leq t) + P(v(t) < 0|\tau > t)P(\tau > t), \quad (13)$$

$$= P(X_t - X_{t^*} < 0|\tau > t)P(\tau > t) \quad (14)$$

where $lim_{t\to\infty}P(v(t) < 0) = 0$ as $lim_{t\to\infty}P(\tau > t) = 0$ as a result of the arcsine law of the maximum in Equation 9.

Therefore, all the conditions required in the definition of statistical arbitrage are satisfied.

IV. UNCERTAINTY IN THE SPREAD MODEL PARAMETERS

Suppose the trader has an educated guess on the long-term equilibrium level of the spread portfolio. However, his guess might not be accurate, and we introduce a noise term to capture this error in the long-term mean level with a random variable η, that is, $\mu = \widehat{\mu} + \eta$, where $\eta \sim N(0, \sigma_\eta)$. We repeat our analysis to check the existence of statistical arbitrage in the presence of uncertainty in model parameters.

First, we can easily verify that the $E[v(t)]$ with the noise term still satisfies $lim_{t\to\infty}E[v(t)] > 0$ because $E[\widehat{\mu}] = \mu$. Second, the condition for the variance follows, because $\text{var}(X_t) = \sigma_\eta^2(1 - e^{-\rho t})^2 + \frac{\sigma^2}{2\rho}(1 - e^{-2\rho t})$, with $lim_{t\to\infty}\text{var}(v(t))/t = 0$.

Owing to the uncertainty in the long-term mean level, there is a positive probability of loss in the limit, which is because of the fact that the size of the error in guessing the long-term mean level can be greater than $\mu - X_{t^*}$. Thus, the limiting probability of loss is calculated as

$$\lim_{t\to\infty} P(v(t) < 0) = P(v(t) < 0|\tau \leq t) \equiv P(\epsilon > \mu - X_{t^*}) = 1 - \Phi\left(\frac{\mu - X_{t^*}}{\sigma_\eta}\right) > 0, \quad (15)$$

where $\Phi(.)$ is the standard normal cumulative distribution function. This implies that we fail to satisfy the condition of statistical arbitrage. However, the probability of loss becomes smaller as we have better estimates for the long-term mean level, and the trader might decide to implement the pairs trading by calculating the probability of loss with respect to the standard deviation of his estimate for the long-term equilibrium level μ.

Similarly, we can introduce uncertainty to the volatility parameter. Denoting the estimate of the standard deviation with $\widehat{\sigma}$, we assume a normally distributed error term $\xi \sim N(0, \sigma_\xi)$, that is, $\sigma = \widehat{\sigma} + \xi$, which does not change the fact that the time-averaged variance decays to zero, that is, $lim_{t\to\infty}\text{var}(v(t))/t = 0$.

Introducing uncertainty in the volatility parameter with the noise term ξ does not effect the fact that the OU process reverts back to its long-term mean level μ, but uncertainty in the volatility estimation has an important effect on the expected mean reversion time. Therefore, even with an error in the estimation of

the volatility, we still have $lim_{t\to\infty}P(\tau<t)=1$. However, if the estimate for the long-term mean is not accurate enough, the probability of loss does not decay to zero. Therefore, it is clear that the estimation of the long-term mean is more crucial for generating statistical arbitrage profits in pairs trading.

V. OPTIMAL PAIRS TRADING

In this section, we consider optimal statistical arbitrage rules under the spread model for pairs trading. Traders are often interested in maximizing the probability of successful termination of the spread portfolio and realizing the profit. Usually, the trader has thousands of possible pairs and rapid flow of information regarding the current spreads between possible pairs. Therefore, it is a natural choice to find the optimal thresholds for starting to implement the optimal pairs trading for a given investment horizon.

For a given investment horizon T, the objective function is given by

$$\max_c P(\tau < T) = \max_c \int_0^T f_{0,c}\,dt, \tag{16}$$

where the $f_{0,c}$ is given in Equation 8.

From the first-order condition $\partial P(\tau<T)/\partial c=0$, the optimal c as a function of the investment horizon T is obtained as (see Appendix for details)

$$c^*(T) = \sqrt{\frac{1-e^{-2T}}{e^{-2T}}}, \quad T>0 \tag{17}$$

which is an increasing function with respect to T.

In Figure 1, we plot the evolution of the probability of first passage time with respect to the investment horizon and the initial spread distance represented via the standardized quantity c. This plot gives important insight to the trader to identify the pairs that have highest likelihood of statistical arbitrage for a given investment horizon. For example, suppose the trader has an investment horizon of 3 months and identified $j=1,2,\ldots,N$ number of pairs with currents spreads denoted by c_j with distance to the closest threshold is given by $d_j = min\left(\left(c_j-\overline{c}^*\right)^2,,\left(c_j-\underline{c}^*\right)^2\right)$ for $j=1,2,\ldots,N$. Then, the trader can prioritize pairs trading starting with the j-th pair that is closest to the optimal buy or sell level.

This means our optimality condition can be used as a filter for screening the best pairs available in the market for a given investment horizon. The trader simply needs a large class of pairs with the standardized c values, and then he can find the pairs that are optimal to trade by maximizing the probability of successful termination.

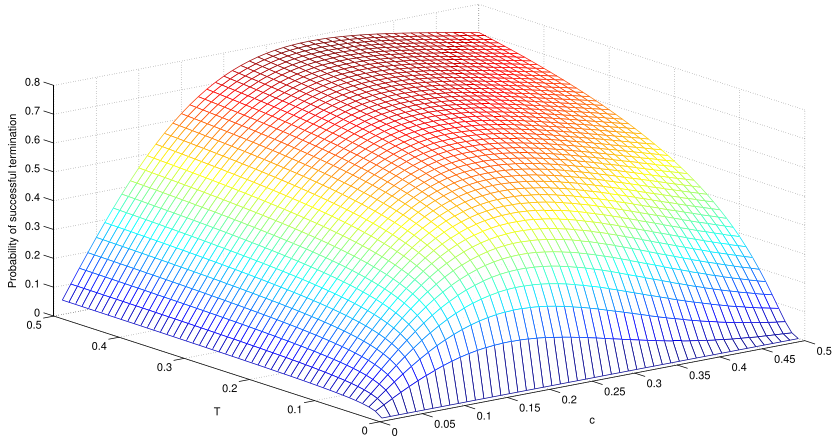

Figure 1 Probability of successful termination of the spread position with respect to the initial deviation and investment horizon.

Uncertainty in the estimation of the long-term equilibrium level causes uncertainty in the current distance to the optimal thresholds. However, using the standard error of the parameter estimate, the investor can still filter the best pairs of stocks for statistical arbitrage.

n ~ $N(0, \sigma_\varepsilon)$. In this case, the trader can construct the $(1-\alpha)$ % confidence interval around $\hat{c}$ denoted as $\left[\widehat{c} - z_{\alpha/2}\sigma_\varepsilon\sqrt{2\rho}/\sigma, \widehat{c} + z_{\alpha/2}\sigma_\varepsilon\sqrt{2\rho}/\sigma\right]$ and find N pairs that satisfy $C_j = \left|\widehat{c}_j - z_{\alpha/2}\sigma_{\varepsilon,j}\sqrt{2\rho_j}/\sigma_j\right| > c^*_{1-\text{month}}$ for $j = 1, 2, \ldots, N$. Similarly, the trader can prioritize implementing pairs trading by ranking the pairs in terms of the distance between the left end of the confidence interval and the optimal threshold.

VI. EMPIRICAL IMPLEMENTATION

In our empirical examples, we consider the adjusted closing prices of the same pairs of stocks as given in Zeng and Lee (2014), and these are, namely, Coca-Cola (KO)/Pepsi (PEP), Target (TGT)/Walmart (WMT), Dell (DELL)/Hewlet-Packard (HPQ), RWE AG (RWE.DE)/E.OnSe (EOAN.DE). Compared with Zeng and Lee (2014), we utilize a larger dataset that consists of 2000 trading days for the period of January 2007 to March 2014. For the pairs RWE and EOAN, which are German utility companies, the number of trading days is more than the US equities considered; however, for backtesting purposes, we still consider the last 2000 trading days counting from the common ending date of the dataset. For the Dell/HPQ pair, the common ending date is 29 October 2013, and thus, we include that last 2000 common trading days of the Dell/HPQ pair. Following Zeng and Lee (2014), we set $X_t = ln(S_t^B) - \gamma ln(S_t^A)$, where the coefficient γ is obtained from a least squares regression of $ln(S_t^B)$ on a constant term and $ln(S_t^A)$.

For the estimation of model parameters, we consider the maximum likelihood and least squares estimators. The maximum likelihood estimator (MLE) for the stochastic model given in Equation 4 can be easily derived from the solution of X_t given the previous time step X_{t-1} as

$$X_t = X_{t-1}e^{-\rho\Delta t} + \mu\left(1 - e^{-\rho\Delta t}\right) + \sigma\int_{t-1}^{t} e^{-\rho(t-s)}dW_s, \tag{18}$$

where $X_t|X_{t-1} \sim N\left(X_{t-1}e^{-\rho\Delta t} + \mu(1 - e^{-\rho\Delta t}), \frac{\sigma^2}{2\rho}\left(1 - e^{-2\rho\Delta t}\right)\right)$. Thus, the log-likelihood function is given as

$$\mathcal{L}(\mu, \sigma, \rho|X) = -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\left(\frac{\sigma^2}{2\rho}\left(1 - e^{-2\rho\Delta t}\right)\right) - \frac{\rho}{\sigma^2}\sum_{i=1}^{N}\left(X_t - \mu - (X_{t-1} - \mu)e^{-\rho\Delta t}\right)^2/\left(1 - e^{-2\rho\Delta t}\right). \tag{19}$$

Alternatively, using Equation 5, we can write the least squares (LS) regression as

$$X_t = a + bX_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d.}(0, \sigma_\varepsilon) \tag{20}$$

where $a = \mu(1 - e^{-\rho\Delta t})$, $b = e^{-\rho\Delta t}$, and $\sigma_\varepsilon^2 = \frac{\sigma^2}{2\rho}\left(1 - e^{-2\rho\Delta t}\right)$. Therefore, the model parameters are given by $\widehat{\rho} = -\ln\left(\widehat{b}\right)/\Delta t$, $\widehat{\mu} = \widehat{a}/\left(1 - \widehat{b}\right)$, and $\widehat{\sigma} = \frac{\widehat{\sigma}_\varepsilon\sqrt{2\widehat{\rho}}}{\sqrt{1 - e^{-2\widehat{\rho}\Delta t}}}$, respectively.

In pairs trading, it is crucial to be able to compute the optimal thresholds in real time with estimated parameters; otherwise, until the algorithm gives a trade signal, the price will change, and the trader will not able to buy or sell the asset from the optimal threshold level. As an example, we compare the performance of the MLE and LS methods for the case of the KO-PEP spread; however, similar results are valid for other spreads as well. The estimation results obtained from the MLE and LS methods are given in Table 1. We observe that the LS and the MLE methods give very close results with very different computational speeds. Note that the MLE method uses an optimization algorithm that requires the gradient vector for the parameter space search, and thus, optimization algorithm is much slower compared with the matrix operations needed in obtaining the LS estimates. Therefore, for real trading applications, we suggest the least squares estimation for model calibration.

In order to verify the performance of our method in comparison with optimal thresholds given by Zeng and Lee (2014) (denoted as ZL method), an in-sample backtesting is conducted. Our optimal threshold for the 6 months investment horizon is calculated using the formula in Equation 17 with respect to the

Table 1 Parameter estimation for the Coca-Cola-Pepsi spread: maximum likelihood and least squares estimation results presented together with the computational time required for each method

Method	$\widehat{\mu}$	$\widehat{\sigma}$	$\widehat{\rho}$	Comp. time (s)
MLE	1.7795	0.1380	1.3054	0.2113
LS	1.7795	0.1380	1.3053	0.0032

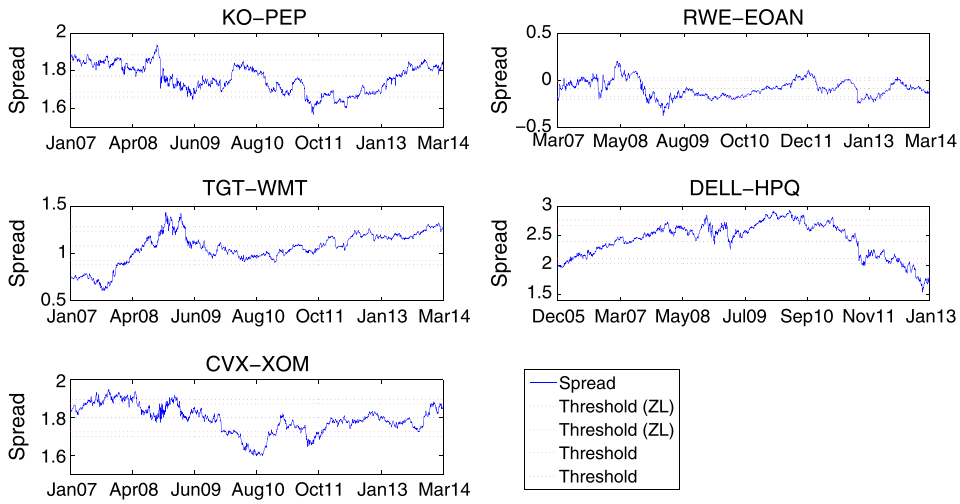

Figure 2 Optimal thresholds and spreads of five pairs with 2000 observations in each pair.

dimensionless system in Equation 6. Real thresholds for the Stochastic Differential Equation (SDE) in Equation 4 are obtained by $\mu \pm c^* \frac{\sigma}{\sqrt{2\rho}}$ for the upper and lower thresholds, respectively.

In Figure 2, the spreads of the pairs and optimal thresholds are plotted. We observe that for the given 6 months investment horizon, our optimal thresholds are slightly wider around the mean compared with the ZL (Zeng and Lee 2014) threshold levels. Our pairs trading strategy is in line with the conventional form of pairs trading where a spread position is opened whenever buy or sell threshold is reached. Upon successful reversion to the long-term mean, the spread position is closed. In the strategy assumed in Zeng and Lee (2014), a short (long) position is opened at the upper (lower) threshold level, and the position is closed at the lower (upper) threshold. This way transaction costs are reduced per trade. However, it is clear that ZL method requires a much longer time, compared with the conventional pairs trading strategy to reach the lower (upper) threshold starting from the upper (lower) threshold, which can be seen in Table 2. Furthermore, the

Table 2 In-sample backtesting results for the pairs considered in Zeng and Lee (2014): average profits and average profit per trading day is computed using our pairs trading model and the ZL method

Stock	Mean return		Average hitting time		Mean ret./exp hit time	
	ZL (%)	New rule (%)	ZL (days)	New rule (days)	ZL (%)	New rule (%)
KO-PEP	16.63	11.57	875.3	206.8	0.019	0.056
RWE-EOAN	18.50	11.13	233	111.2	0.079	0.10
TGT-WMT	35.06	26.56	458.3	237	0.076	0.112
DELL-HPQ	62.28	41.69	724	257.3	0.0860	0.162
CVX-XOM	13.46	8.97	840	167.7	0.016	0.054

results given in Zeng and Lee (2014) focus on the total return generated; however, when we take into account the long holding periods, the expected return per holding day is higher in our approach. Termination of the spread position within the given investment horizon is a practical necessity since the short positions are difficult and risky to maintain for long periods of time.

We should also note that there are many pairs of stocks that have historically high correlations in many industries; therefore, it is more important to terminate each pairs position successfully in a limited time because a successful trade implies that the resources can be invested into other potential pairs as well.

VII. CONCLUSION

In this study, assuming an OU spread process, we prove that pairs trading in its most common form satisfies the definition of statistical arbitrage given by Hogan et al. (2004). However, we also show that a time-independent error in trader's guess or forecast of the long-term mean level implies that the statistical arbitrage is no longer guaranteed. In other words, a perfect statistical arbitrage with the probability of loss decaying to zero is not available whenever there is uncertainty in the model parameters. The good news is that the probability of loss can be bounded as a function of the estimation error and for sufficiently good estimates, and the trader can still implement pairs trading knowing the potential probability of loss involved.

We derive optimal thresholds that maximize the probability of successful pairs trading for a given investment horizon. Derived optimal thresholds can also be used by traders as a new pairs screening tool that filters the pairs that are closest to the optimal buy/sell thresholds. Owing to the large number of potential pairs of stocks or portfolios in stock markets, it is crucial for the traders to maximize the probability of successful termination of the spread portfolio within the given time horizon.

In our empirical examples, model estimation is carried out using the maximum likelihood and least squares estimation techniques, which yield very similar parameter estimates. However, we show that the least squares estimation of our model parameters is much faster relative to the maximum likelihood estimation. Backtesting experiments are conducted for the same five pairs of stocks considered in Zeng and Lee (2014), which indicate the profitability of our optimal thresholds in pairs trading in terms of higher returns per holding days.

Erdinç Akyıldırım
Department of Banking and Finance
Akdeniz University
Antalya
Turkey
erdinc.akyildirim@gmail.com

REFERENCES

Avellaneda, M., and J.-H. Lee (2010), 'Statistical arbitrage in the US equities market', *Quantitative Finance*, 10, 761–82.

Bertram, W. K. (2010), 'Analytical solutions for optimal statistical arbitrage trading', *Physica A: Statistical Mechanics and its Applications*, 389, 2234–43.

Cummins, M., and A. Bucca (2012), 'Quantitative spread trading on crude oil and refined products markets', *Quantitative Finance*, 12, 1857–75.

Do, B., and R. Faff (2010), 'Does simple pairs trading still work?', *Financial Analysts Journal*, 66.

Do, B., and R. Faff (2012), 'Are pairs trading profits robust to trading costs?', *Journal of Financial Research*, 35, 261–87.

Ehrman, D. S. (2006), *The Handbook of Pairs Trading: Strategies Using Equities, Options, and Futures*, John Wiley, Hoboken, NJ.

Elliot, R. J., J. V. D. Hoek, and W. P. Malcolm (2005), 'Pairs trading', *Quantitative Finance*, 5, 271–6.

Finch, S., and Y. Yt (2004), *Ornstein–Uhlenbeck Process*. Available online at: http://pauillac.inria.fr/algo/bsolve/constant/constant.html.

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

Göncü, A. (2015), 'Statistical arbitrage in the Black–Scholes framework', *Quantitative Finance*, 15, 1489–99.

Hogan, S., R. Jarrow, M. Teo, and M. Warachka (2004), 'Testing market efficiency using statistical arbitrage with applications to momentum and value strategies', *Journal of Financial Economics*, 73, 525–65.

Huck, N., and K. Afawubo (2015), 'Pairs trading and selection methods: is cointegration superior?', *Applied Economics*, 47, 599–613.

Litterman, B. (2004), *Modern investment management: an equilibrium approach*, Vol. 246. New York: John Wiley & Sons.

Vidyamurthy, G. (2004), *Pairs Trading–Quantitative Methods and Analysis*. Wiley: New York.

Whistler, M. (2004), *Trading Pairs–Capturing Profits and Hedging Risk with Statistical Arbitrage Strategies*. New York: Wiley.

Zeng, Z., and C.-G. Lee (2014), 'Pairs trading: optimal thresholds and profitability', *Quantitative Finance*, 14, 1881–93.

APPENDIX: Derivation of the optimal threshold level c^*

The probability of successful pairs trading for a given investment horizon T, that is, successful mean reversion and closing the spread position, is given by

$$P(\tau < T) = \int_0^T f_{0,c}\,dt \tag{A1}$$

$$= \int_0^T \sqrt{\frac{2}{\pi}} \frac{|c|e^{-t}}{\left(1-e^{-2t}\right)^{3/2}} \exp\left(-\frac{c^2 e^{-2t}}{2(1-e^{-2t})}\right) dt, \quad \text{for } t > 0. \tag{A2}$$

We can write the first-order condition that maximizes the probability of successful pairs trading as follows

$$\frac{\partial}{\partial c} P(\tau < T) = \frac{\partial}{\partial c} \int_0^T f_{0,c}\,dt = \int_0^T \frac{\partial f_{0,c}}{\partial c}\,dt = 0 \quad \text{since } \frac{\partial f_{0,c}}{\partial c} \text{ is continuous.} \tag{A3}$$

The sufficient condition for the First Order Condition (FOC) to be satisfied is given by

$$\frac{\partial f_{0,c}}{\partial c} = \sqrt{\frac{2}{\pi}} \frac{e^{-t}}{(1-e^{-2t})^{3/2}} \exp\left(-\frac{c^2 e^{-2t}}{2(1-e^{-2t})}\right)\left(1-\frac{c^2 e^{-2t}}{1-e^{-2t}}\right) = 0 \tag{A4}$$

which yields

$$c^*(t) = \sqrt{\frac{1-e^{-2t}}{e^{-2t}}}, \quad t > 0. \tag{A5}$$

We check the second-order condition in order to verify that we obtain the deviation level c^* that maximizes the probability of successful pairs trade. The second-order partial derivative is given by

$$\frac{\partial^2 f_{0,c}}{\partial c^2} = \sqrt{\frac{2}{\pi}} \frac{e^{-t}}{(1-e^{-2t})^{3/2}} \left[\left(\frac{-ce^{-2t}}{1-e^{-2t}}\right)\exp\left(-\frac{c^2 e^{-2t}}{2(1-e^{-2t})}\right)\left(1-\frac{c^2 e^{-2t}}{1-e^{-2t}}\right) + \exp\leq ft\left(-\frac{c^2 e^{-2t}}{2(1-e^{-2t})}\right)\left(\frac{-2ce^{-2t}}{1-e^{-2t}}\right)\right. \tag{A6}$$

$$= \sqrt{\frac{2}{\pi}} \frac{e^{-t}}{(1-e^{-2t})^{3/2}} \exp\left(-\frac{c^2 e^{-2t}}{2(1-e^{-2t})}\right)\left[\frac{\left(1-e^{-2t}-c^2 e^{-2t}\right)\left(-ce^{-2t}\right)}{\left(1-e^{-2t}\right)^2} + \frac{-2ce^{-2t}}{1-e^{-2t}}\right], \tag{A7}$$

and if we plug in the aforementioned equation $c^*(t)$, then we obtain

$$\frac{\partial^2 f_{0,c}}{\partial c^2} = \sqrt{\frac{2}{\pi}} \frac{e^{-t}}{(1-e^{-2t})^{3/2}} \exp\left(-\frac{c^{*2} e^{-2t}}{2(1-e^{-2t})}\right)\left[0 + \frac{-2c^* e^{-2t}}{1-e^{-2t}}\right] < 0 \text{ since } c^* > 0. \tag{A8}$$