Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: https://www.tandfonline.com/loi/rquf20

Pairs trading: optimal thresholds and profitability

Zhengqin Zeng & Chi-Guhn Lee

To cite this article: Zhengqin Zeng & Chi-Guhn Lee (2014) Pairs trading: optimal thresholds and profitability, Quantitative Finance, 14:11, 1881-1893, DOI: 10.1080/14697688.2014.917806

To link to this article: https://doi.org/10.1080/14697688.2014.917806

Published online: 08 Oct 2014.

Quantitative Finance, 2014
Vol. 14, No. 11, 1881–1893, http://dx.doi.org/10.1080/14697688.2014.917806

Pairs trading: optimal thresholds and profitability

ZHENGQIN ZENG and CHI-GUHN LEE*

Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, ON M5S 3G8, Canada

(*Received 14 January 2014; accepted 17 April 2014*)

1. Introduction

Since its birth in the 1980s, pairs trading have been popular as a statistical arbitrage strategy among major investment banks and hedge funds. Despite the high average annualized excess return, which has been as high as 11%, the idea behind the strategy is simple. If the two prices of a pair of stocks move together in the past, they are likely to continue in the future. So when the prices diverge, a trader can simply take a short position with the over-priced stock and a long position with the under-priced one, and wait for the prices to converge in the future. When they do, the trader clears the positions and makes a profit.

For practitioners, the common practice of pairs trading can be summarized in three steps: (1) select a pair of stocks and calculate the mean and standard deviation of the price ratio for the pair; (2) when the ratio deviates from the mean by two standard deviations, short the over-priced stock and long the under-priced one; and (3) when the ratio reverts to the mean, clear the positions to make a profit. Note that there are alternative quantities to price for triggering transactions, such as the squared Euclidean norm and the log difference of the prices. Also, the investor may not have to follow the rule of two standard deviations in the second step, but can strategically change the trigger point according to the market conditions and the actual movement of the pair. Similarly, it is not a fixed rule to clear positions exactly when the spread reverts back to the mean. The investor might want to wait longer to gain larger profit, but then also bears more risk.

Various quantitative methods have been developed and applied to pairs trading in the literature. Three commonly used techniques are: distance method, co-integration and stochastic spread. The distance method is mostly used by practitioners. Nath (2003) used the 15th percentile of the distribution of distance as a trigger for trading and the 5th percentile as the stop-loss barrier. Gatev *et al.* (2006) selected pairs and generated trading signals based on this method. Despite its model-free feature that prevents misestimation of parameters, the distance method provides little help in forecasting according to Do *et al.* (2006). Instead, Vidyamurthy (2004) developed a framework for forecasting using the co-integration method, and analysed the mean reversion of the residuals. The method of co-integration was further applied by Lin *et al.* (2006) to develop a loss protection for pairs trading,

*Corresponding author. Email: cglee@mie.utoronto.ca

and Puspaningrum *et al.* (2010) to develop algorithms to estimate trade duration and find optimal preset boundaries. The stochastic spread method, on the other hand, models the mean reverting process of pairs trading as an Ornstein–Uhlenbeck (OU) process. Elliott *et al.* (2005) provided an analytic framework of pairs trading, which laid the ground for prediction and decision-making based on the hidden OU process. Ekström *et al.* (2011) explored optimal liquidation of pairs trading in the framework of the OU process and analysed the sensitivity of the model parameters. Vladislav (2004) and Boguslavsky and Boguslavskaya (2004) also based their research on the OU process. Bertram (2010) argued strongly for the role of time and derived analytic formula for the thresholds of a synthetic asset whose price is assumed to follow an OU process. He showed that the optimal thresholds were symmetric around the mean both for maximizing the return per unit time and the Sharpe ratio. However, he did not allow short selling of the synthetic asset.

To maximize the expected profit per unit time in the long run, the investor should choose the right entry and exit thresholds. If the thresholds are narrow, then the time it needs to complete a trade is small, but so is the profit in each trade. On the other hand, if thresholds are too wide, the profit in each trade is larger, but so is the total time needed to complete a trade. The interplay between the profit per trade and the length of a trade gives rise to an interesting optimization problem, which, to the best of our knowledge, has not been previously studied. Therefore, the main objective of this paper is to find the optimal thresholds as functions of the transaction cost and parameters of the OU process for the objective of maximizing the long run expected average profit.

This paper contributes both to theory and practice. From a theoretical point of review, we derive a polynomial expression for the expectation of the first-passage time of an OU process with two-sided boundary. Though derivation of the optimal thresholds using the Laplace transform is still possible theoretically, our expression can greatly simplify the proof and calculation. This polynomial expression can be easily applied in other research problems. From a practical point of view, we obtain the analytic formula of optimal thresholds for pairs trading, and the results are counter-intuitive. To compare with common practice, we also show a step-by-step procedure on the daily data of Coca-Cola and Pepsi. Results show that the new optimal strategy developed in this paper performs better than the common practice. Also useful to practitioners is a profitability indicator measuring theoretical return of a pair, by which investors can identify more profitable pairs.

The structure of this paper is as follows. Section 2 discusses the model of pairs trading and derives the objective functions for long-run profit per unit time. In section 3, we give a brief overview of the first-passage time over a one-sided boundary and derive the expectation of the two-sided boundary first-passage time. In section 4, we present the optimal thresholds and compare the return with the optimized conventional pairs trading strategy. Section 5 uses the optimal trading strategy on real data and discusses the results. Section 6 concludes the paper and propose topics for future research.

2. Model description

In Avellaneda and Lee (2010), the co-integration is modelled as:

$$\ln(P_t)-\ln(P_0) = \alpha(t-t_0)+\beta[\ln(Q_t)-\ln(Q_0)]+\epsilon_t, \quad t \geq 0, \tag{1}$$

where P_t and Q_t are the stock prices of a pair of assets at time t. Notice that the drift rate α is usually ignorable compared to fluctuation of the residual ϵ_t. The above model suggests that if we take a long position of 1 dollar in stock P at time t, we should short Q for β dollars, and vice versa. In this paper, we continue to use the relationship above and assume that the mean reverting process ϵ_t follows an OU process. For simplicity, define $X_t = \epsilon_t + \ln(P_0) - \beta\ln(Q_0)$ in equation (1). Note that X_t is still an OU process since $\ln(P_0) - \beta\ln(Q_0)$ is only a constant. A trading signal is generated when X_t reaches a preset threshold. We have the following two equations for the correlation of the pair and the dynamics of the residual X_t:

$$\ln(P_t) - \beta\ln(Q_t) = X_t, \tag{2}$$

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t, \tag{3}$$

where θ is the mean reversion rate, μ is the mean of X_t, W_t is the standard Wiener process, and σ is the standard deviation for the Wiener process in equation (3).

Similar to Bertram (2010), we can transform equation (3) into the dimensionless system by $\tau = \theta t$ and $Y_\tau = \frac{\sqrt{2\theta}}{\sigma}(X_t - \mu)$. Hence, we have:

$$dY_\tau = -Y_\tau d\tau + \sqrt{2}dW_\tau \tag{4}$$

We call equation (4) dimensionless system because Y_τ is not dependent on the model parameters. Notice that the above transformation is linear, so that each value of X_t corresponds to a unique value of Y_τ.

We generate trading signals when Y_τ reaches a preset threshold. For example, when $Y_{\tau_1} = a$ $(a > 0)$, we short 1 dollar of stock P and long β dollars of stock Q, and when $Y_{\tau_2} = b$ $(b < a)$, we clear positions and make profit. The profit on P is $r_1 = \frac{P_{\tau_2}-P_{\tau_1}}{P_{\tau_1}}$, or $r_1 = \ln(P_{\tau_1}) - \ln(P_{\tau_2})$ in terms of the continuous compound rate of return. Similarly, $r_2 = \beta[\ln(Q_{\tau_2}) - \ln(Q_{\tau_1})]$. From equation (2), we can express the return as $r = r_1+r_2 = X_1 - X_2 = \tilde{a}-\tilde{b}$, where $\tilde{a} = a\frac{\sigma}{\sqrt{2\theta}}+\mu$ and $\tilde{b} = b\frac{\sigma}{\sqrt{2\theta}} + \mu$. Assume the transaction cost is $\tilde{c}$ and let $c = \tilde{c}\frac{\sqrt{2\theta}}{\sigma}$ be the transaction cost in the dimensionless system, so the net profit for each transaction is $\tilde{a}-\tilde{b}-\tilde{c}$, or $a-b-c$ in the dimensionless system. Similarly, if we trade in at $Y_{\tau_1} = -a$, at which we go long 1 dollar of P and short β dollars of stock Q, then we trade out at $Y_{\tau_2} = -b$. As one can compute, the net profit in each trade in the dimensionless system is again $a-b-c$. Without loss of generality, we will assume positions are first taken at $Y_{\tau_1} = a$. It is intuitive that $b \in [-a, a]$. If $b > a$, then the trader will always lose since $a-b-c < 0$ for any $c \geq 0$. To rule out the case $b < -a$, a rigorous analysis will be given in section 4.

Each trading cycle is composed of two parts: the first part is from taking positions to clearing positions, and the second part is simply waiting until the next trading opportunity. Notice that the profit is made only in the first part. Let t_1 and t_2 be the durations of the two parts, and τ_1 and τ_2 be the corresponding time in the dimensionless system. Similar to Bertram (2010), τ_1 is the first-passage time from a to b, whereas, τ_2 is the time it takes from b to escape the range $[-a, a]$. Mathematically, τ_1 and τ_2 are defined as follows:

$$\tau_1 = \inf\left\{t;\, Y_t = b \;\middle|\; Y_0 = a\right\} \tag{5}$$

$$\tau_2 = \inf\left\{t;\, |Y_t| = a \;\middle|\; Y_0 = b\right\} \tag{6}$$

The total time for each trading cycle is $T = \tau_1 + \tau_2$. Suppose there are N_τ transactions completed in $[0, \tau]$, so the net profit is $NP_\tau = (a-b-c)N_\tau$. By the elementary renewal theorem, the expected profit per unit time is given by

$$\mu = \lim_{\tau\to\infty} \frac{E[NP_\tau]}{\tau} = (a-b-c) \lim_{\tau\to\infty} \frac{E[N_\tau]}{\tau} = \frac{a-b-c}{E[T]}, \tag{7}$$

where $E[T] = E[\tau_1] + E[\tau_2]$. Also, we know that the expected time of one cycle in the real system is $E[\tilde{T}] = \frac{E[T]}{\theta}$. In this paper, our objective is to find optimal thresholds to maximize the expected return per unit time μ.

Notice that the expected return per unit time in real system is $\tilde{\mu} = \frac{\tilde{a}-\tilde{b}-\tilde{c}}{E[\tilde{T}]} = \frac{\sigma\sqrt{\theta}}{\sqrt{2}} \frac{a-b-c}{E[T]} = \sqrt{\frac{\theta}{2}}\sigma\mu$. The coefficient $\sigma\sqrt{\theta/2}$ is only determined by the prices of the pairs, and is a constant once the model parameters are known. Therefore, maximizing the real return is the same as maximizing the return in the dimensionless system. The constant $\sigma\sqrt{\theta/2}$ contains intuitive and important information: a larger mean reversion rate θ means a higher trading frequency, and a larger σ means a bigger fluctuation of X_t, both leading to a higher profit in each trade.

Since both the time and scale are linearly transformed into the dimensionless system, we can first obtain the optimal thresholds in the dimensionless system and then transform back to the real system. For simplicity, we will only write in the notation of the dimensionless system afterwards.

3. First-passage times

It is crucial to find the expectation of the first-passage time over one-sided and two-sided boundaries in order to find the optimal thresholds. In this section, we will give a brief review on the first-passage time over one-sided boundary, and derive the expectation of the first-passage time over two-sided boundaries. A major contribution of this model lies in finding a polynomial form of the expectation over two-sided boundary.

3.1. First-passage time over a one-sided boundary

For one-sided boundary, Thomas (1975), Sato (1977) and Ricciardi and Sato (1988) expressed the expectation as an infinite sum of polynomials. To summarize, for $x > 0$ and $y > 0$, the expectation of $T_{x,0}$, the first-passage time from x to 0 is:

$$E[T_{x,0}] = \frac{1}{2}\sum_{k=1}^{\infty} (-1)^{k+1} \frac{(\sqrt{2}x)^k}{k!} \Gamma\left(\frac{k}{2}\right), \tag{8}$$

and the expectation of $T_{0,y}$, the first-passage time from 0 to y is:

$$E[T_{0,y}] = \frac{1}{2}\sum_{k=1}^{\infty} \frac{(\sqrt{2}y)^k}{k!} \Gamma\left(\frac{k}{2}\right). \tag{9}$$

Hence, the expectation $E[T_{a,b}]$ for the case $a > 0$ can be written as:

$$E[T_{a,b}] = \begin{cases} E[T_{a,0}] - E[T_{b,0}], & \text{for } b > 0 \\ E[T_{a,0}] + E[T_{0,-b}], & \text{for } b \le 0 \end{cases} \tag{10}$$

By symmetry of an OU process, we can also get the expectation for the case $a < 0$ by $E[T_{a,b}] = E[T_{-a,0}] + E[T_{0,b}]$ for $b > 0$ and $E[T_{a,b}] = E[T_{-a,0}] - E[T_{-b,0}]$ for $b < 0$. Similarly, there are explicit results for the variance of the first-passage time over a one-sided boundary (shown in section 1). One can find the variance of this type of first-passage time between any two points by the symmetric property of an OU process.

3.2. First-passage time over a two-sided boundary

For the first-passage time over a two-sided symmetric boundary, Darling and Siegert (1953) derived the Laplace transform of $T_{-a,a,b}$, the first-passage time from b to cross the boundary $(-a, a)$ as given by

$$E\left[e^{-\lambda T_{-a,a,b}}\right] = \frac{D_{-\lambda}(b) + D_{-\lambda}(-b)}{D_{-\lambda}(a) + D_{-\lambda}(-a)} \exp\left(\frac{b^2 - a^2}{4}\right), \tag{11}$$

where $D_{-\lambda}(b)$ is the Weber function, which can be shown as:

$$D_{-\lambda}(x) = \sqrt{\frac{2}{\pi}} \exp\left(\frac{x^2}{4}\right) \int_0^\infty t^{-\lambda} \exp\left(-\frac{t^2}{2}\right) \times \cos\left(xt + \frac{\lambda\pi}{2}\right) dt, \quad \text{for } \lambda < 1. \tag{12}$$

Define $m(\lambda, x) = D_{-\lambda}(x) + D_{-\lambda}(-x)$. We have the following from the Weber function (12)

$$\begin{aligned} m(\lambda, x)\big|_{\lambda=0} &= 2\sqrt{\frac{2}{\pi}} \exp\left(\frac{x^2}{4}\right) \int_0^\infty \exp\left(\frac{-t^2}{2}\right) \cos(xt) dt \\ &= 2\exp\left(-\frac{x^2}{4}\right), \end{aligned} \tag{13}$$

$$\frac{\partial m(\lambda, x)}{\partial \lambda}\bigg|_{\lambda=0} = -2\sqrt{\frac{2}{\pi}} \exp\left(\frac{x^2}{4}\right) \int_0^\infty \ln(t) \exp\left(-\frac{t^2}{2}\right) \times \cos(xt) dt. \tag{14}$$

To get equation (13), we need to use the fact that $\int_0^\infty y^{2n} \exp(-\frac{y^2}{a^2}) dy = \sqrt{\pi}\frac{(2n)!}{n!}(\frac{a}{2})^{2n+1}$ for $n = 1, 2, 3, \ldots$. Therefore, if we let $y = xt$ in equation (13) and use Taylor expansion on $\cos(xt)$, we can get:

$$\begin{aligned} &\int_0^\infty \exp\left(\frac{-t^2}{2}\right) \cos(xt) dt \\ &= \frac{1}{x} \int_0^\infty \exp\left(-\frac{y^2}{2x^2}\right) \cos(y) dy \\ &= \frac{1}{x} \int_0^\infty \exp\left(-\frac{y^2}{2x^2}\right) \sum_{n=0}^\infty (-1)^n \frac{y^{2n}}{(2n)!} dy \end{aligned}$$

$$= \frac{1}{x}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}\int_0^{\infty}\exp\left(-\frac{y^2}{2x^2}\right)y^{2n}dy$$

$$= \sqrt{\frac{\pi}{2}}\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\left(\frac{\sqrt{2}x}{2}\right)^2$$

$$= \sqrt{\frac{\pi}{2}}\exp\left(-\frac{x^2}{2}\right)$$

Hence we have equation (13). Taking the first derivative on both sides of equation (11) and setting $\lambda = 0$, we have:

$$E[-T_{-a,a,b}] = \frac{\frac{\partial m(\lambda,b)}{\partial\lambda}\Big|_{\lambda=0} m(\lambda,a)\Big|_{\lambda=0} - \frac{\partial m(\lambda,a)}{\partial\lambda}\Big|_{\lambda=0} m(\lambda,b)\Big|_{\lambda=0}}{(m(\lambda,a)|_{\lambda=0})^2} \times \exp\left(\frac{b^2-a^2}{4}\right)$$

By using equations (13) and (14), and multiplying both sides by -1, we can get the expectation of $T_{-a,a,b}$ as follows:

$$E[T_{-a,a,b}] = \sqrt{\frac{2}{\pi}}[h(b) - h(a)], \tag{15}$$

where

$$h(x) = \exp\left(\frac{x^2}{2}\right)\int_0^{\infty}\ln(t)\exp\left(-\frac{t^2}{2}\right)\cos(xt)dt. \tag{16}$$

The following proposition will further simplify the expression and make it handy to find the optimal thresholds in the next section.

Proposition 3.1 *The integral form of $h(x)$ shown above can be expressed as an infinite sum of polynomials and a constant:*

$$h(x) = -\frac{1}{2}\sqrt{\frac{\pi}{2}}\sum_{n=1}^{\infty}\frac{(\sqrt{2}x)^{2n}}{(2n)!}\Gamma(n) + C, \tag{17}$$

where $C = \int_0^{\infty}\ln(t)\exp\left(-\frac{t^2}{2}\right)dt$.

Proof see Appendix A.

From equations (15) and (17), we can simplify the expectation as

$$E[T_{-a,a,b}] = \frac{1}{2}\sum_{n=1}^{\infty}\frac{(\sqrt{2}a)^{2n} - (\sqrt{2}b)^{2n}}{(2n)!}\Gamma(n) \tag{18}$$

Similarly, we can find the second moment of $T_{-a,a,b}$, which is used in computing the variance per unit time in section 4. We have:

$$\frac{\partial^2 m(\lambda,x)}{\partial\lambda^2}\Big|_{\lambda=0} = 2\sqrt{\frac{2}{\pi}}\exp\left(\frac{x^2}{4}\right)\int_0^{\infty}\ln(t)^2\exp\left(-\frac{t^2}{2}\right) \times \cos(xt)dt - \frac{\pi^2}{2}\exp\left(-\frac{x^2}{4}\right),$$

and the expectation of the second moment becomes:

$$E[T^2_{-a,a,b}] = \exp\left(\frac{b^2-a^2}{4}\right)[g_1(a,b) - g_2(a,b)]$$

where

$$g_1(a,b) = \frac{\frac{\partial^2 m(\lambda,b)}{\partial\lambda^2}\Big|_{\lambda=0} m(\lambda,a)\Big|_{\lambda=0} - \frac{\partial m(\lambda,a)}{\partial\lambda}\Big|_{\lambda=0}\frac{\partial m(\lambda,b)}{\partial\lambda}\Big|_{\lambda=0}}{(m(\lambda,a)|_{\lambda=0})^2}$$

and

$$g_2(a,b) = \frac{\frac{\partial^2 m(\lambda,a)}{\partial\lambda^2}\Big|_{\lambda=0} m(\lambda,b)\Big|_{\lambda=0} + \frac{\partial m(\lambda,a)}{\partial\lambda}\Big|_{\lambda=0}\frac{\partial m(\lambda,b)}{\partial\lambda}\Big|_{\lambda=0}}{(m(\lambda,a)|_{\lambda=0})^2} - 2\frac{\left(\frac{\partial m(\lambda,a)}{\partial\lambda}\Big|_{\lambda=0}\right)^2 m(\lambda,b)\Big|_{\lambda=0}}{\left(m(\lambda,a)|_{\lambda=0}\right)^3}$$

Unlike the first moment, the integral form cannot be simplified to the polynomial form, leaving the second moment difficult to use. The variance can be found by the first two moments, but only in a very complicated integral form.

4. Optimal thresholds

With the polynomial form of the expectation in section 3, we are now ready to find the optimal thresholds for pairs trading. The main goal is to maximize the expected return per unit time. As explained in section 2, we take positions when Y_τ reaches the opening threshold a (or $-a$), clear positions when it reaches the closing threshold b (or $-b$) and wait for the next opportunity until Y_τ reaches an opening threshold again. In this section, we will discuss three cases with the values of a and b. Throughout this section, we assume $a \geq 0$. Since the OU process is symmetric, the case for $a < 0$ will be exactly the same.

Case 1 $0 \leq b \leq a$

From equation (7), our objective function is given by $f(a,b) = \frac{a-b-c}{E[\tau_1]+E[\tau_2]}$, where $E[\tau_1]$ and $E[\tau_2]$ are explicitly shown by equations (10) and (18). In order for $f(a,b)$ to be non-negative, we have to restrict $a - b - c \geq 0$. The optimization problem is:

$$\begin{aligned} \operatorname*{Max}_{a,b} \quad & f(a,b) \\ &= \frac{a-b-c}{E[\tau_1]+E[\tau_2]} \\ &= \frac{a-b-c}{\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}-(\sqrt{2}b)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)} \\ \text{subject to} \quad & 0 \leq b \leq a - c \end{aligned} \tag{19}$$

To find the optimal solution in the domain $0 \leq b \leq a - c$, we need to use the fact that $\frac{\partial f(a,b)}{\partial b} < 0$ for any a in the domain. To prove this, we have

$$\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1} - (\sqrt{2}b)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)$$

$$= \sum_{n=0}^{\infty}\frac{(\sqrt{2}a - \sqrt{2}b)\sum_{k=0}^{2n}(\sqrt{2}a)^{2n-k}(\sqrt{2}b)^k}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)$$

$$\geq \sqrt{2}(a-b-c)\sum_{n=0}^{\infty}\frac{\sum_{k=0}^{2n}(\sqrt{2}a)^{2n-k}(\sqrt{2}b)^k}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)$$

$$\geq \sqrt{2}(a-b-c)\sum_{n=0}^{\infty}\frac{(2n+1)(\sqrt{2}b)^{2n}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)$$

$$= \sqrt{2}(a-b-c)\sum_{n=0}^{\infty}\frac{(\sqrt{2}b)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)$$

The first inequality is due to the fact that $c \geq 0$ and the second inequality is due to the fact that $a \geq b \geq 0$. With the above inequality, we can get

$$\frac{\partial f(a,b)}{\partial b} = \frac{-\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}-(\sqrt{2}b)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right) + (a-b-c)\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}b)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)}{\left(\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}-(\sqrt{2}b)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)\right)^2} \leq 0$$

Equality only holds when $a = b$ and $c = 0$. So for any given a and c, the optimal value of b is $b^* = 0$. Therefore, the original maximization problem is now:

$$f(a) = f(a,0) = \frac{a-c}{\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)}$$

Setting $\frac{df(a)}{da} = 0$, we can find the optimal value a^* by solving the equation:

$$\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right) = (a-c)\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right) \tag{20}$$

The existence and uniqueness of the solution to equation (20) can be easily shown. When $c = 0$, $a = 0$ is a solution. When $c > 0$, if we let $a \to c$, we will have

$$\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right) > (a-c)\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)$$

If we let $a \to \infty$, we will have

$$\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right) < (a-c)\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right),$$

which proves the existence of the solution. To prove the uniqueness, we take derivative of both sides of equation (20) with respect to a. We have

$$c'(a) = \frac{(a-c)\sum_{n=1}^{\infty}\frac{(\sqrt{2}a)^{2n-1}}{(2n-1)!}\Gamma\left(\frac{2n+1}{2}\right)}{\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)} > 0$$

Since $c(a)$ is an increasing function of a, there is a unique value of a that satisfies equation (20) for any given $c > 0$.

To see that a^* is the maximizer rather than the minimizer, we let $a \to c$ and $a \to \infty$. For any $c > 0$, when $a \to c$, it is easy to see that $f(a) \to 0$. Similarly, when $a \to \infty$, we will have:

$$f(a) = \frac{a-c}{\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)} = \frac{1-\frac{c}{a}}{\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)} \to 0$$

because $0 \leq 1-\frac{c}{a} < 1$ and $\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right) \to \infty$.

Also we know that for any $c \leq a < \infty$, $f(a) \geq 0$. Therefore we conclude that for $c > 0$, a^* maximizes $f(a)$.

When $c = 0$, we have

$$f(a) = \frac{a}{\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)} = \frac{\sqrt{2}}{\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)}$$

is a decreasing function of a. When $a \to 0$, we have $f(a) \to \sqrt{\frac{2}{\pi}}$. In this case, by solving equation (20), we can still get $a^* = 0$.

Remark The optimal solutions in this case are exactly the optimal thresholds for the conventional way of the pairs trading: take positions when the spread widens ($Y_t = a^*$) and clear positions when the spread reverts to the mean ($Y_t = b^* = 0$). Note that when there is no transaction cost ($c = 0$), the gap between a and b should be infinitely close to 0, which means that the trader should constantly adjust his positions to make as many trades as possible to in a given time. In this case, the trader values the trading frequency more than the profit per trade. This is also consistent with Bertram (2010).

Case 2 $-a \leq b \leq 0$

Here, we do not exclude $b = 0$ for the feasibility of our optimal solution. For $b \leq 0$, the optimization problem is written as:

$$\begin{aligned} &\max_{a,b} \quad f(a,b) \\ &= \frac{a-b-c}{E[\tau_1]+E[\tau_2]} = \frac{a-b-c}{\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}-(\sqrt{2}b)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)} \\ &\text{subject to} \quad -a \leq b \leq \min\{0, a-c\} \end{aligned} \tag{21}$$

Here, we require $a \geq \frac{c}{2}$ for feasibility. Notice that $f(a,b)$ is bounded inside the domain. First of all, for any a, $f(a,b)$ is boundary since b is bounded and $f(a,b)$ is continuous in b. To prove $f(a,b)$ is bounded in a, we discuss two cases: $c > 0$ and $c = 0$. When $c > 0$, if we let $a \to \frac{c}{2}$, then $b \to -\frac{c}{2}$ in the domain. So we will have $f(a,b) \to 0$. If we let $a \to \infty$, we will get $f(a,b) \to 0$ for any b in the domain. When $c = 0$, if we let $a \to 0$, we will have $b \to 0$, and $f(a,b) \to \sqrt{\frac{2}{\pi}}$. Again, if we let $a \to \infty$ when $c = 0$, we will get $f(a,b) \to 0$ for any b in the domain. So for both cases, $f(a,b)$ is bounded in a and the minimal value $f(a,b) \to 0$ appears when a approaches its boundary. Since $f(a,b)$ is continuous in both a and b, the maximal value exists on the closed set of the domain.

Setting the gradient of $f(a,b)$ to 0, we have:

$$E[\tau_1]+E[\tau_2]=(a-b-c)\left(\frac{\partial E[\tau_1]}{\partial a}+\frac{\partial E[\tau_2]}{\partial a}\right)$$

$$E[\tau_1]+E[\tau_2]=-(a-b-c)\left(\frac{\partial E[\tau_1]}{\partial b}+\frac{\partial E[\tau_2]}{\partial b}\right)$$

Therefore, we have:

$$\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)=\sum_{n=0}^{\infty}\frac{(\sqrt{2}b)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right) \tag{22}$$

Since $g(x)=\sum_{n=0}^{\infty}\frac{(\sqrt{2}x)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)$ is an increasing function, for equation (22) to hold, we must have $a^2=b^2$. Since $-a\le b\le\min\{0,a-c\}$, the optimal solution can only be $b^*=-a^*$, where a^* can be found by solving the equation:

$$\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)=\left(a-\frac{c}{2}\right)\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right) \tag{23}$$

With the same argument in case 1, we can show the existence, uniqueness of the solution a^* in equation (23).

However, we still have to check that $b^*=-a^*$ is the global maximal by showing that $f(a^*,b^*)\ge f(a,b)$ for any a,b on the boundary. For any $b=-a$, we can prove that $f(a,b)\le f(a^*,b^*)$ by the same argument as in case 1. For any $b\to a-c$ where $c>0$, we have $f(a,b)\to 0<f(a^*,b^*)$. When $c=0$, we have $a^*=b^*=0$ and it is easy to check that $f(a^*,b^*)\ge f(a,b)$ for any $b\to a-c$. When $b\to 0$, we can show that $\max_{a\ge\frac{c}{2}}f(a,b)\le f(a^*,b^*)$ by Proposition 4.1.

Remark The only difference between equations (23) and (20) is the term of c. Equation (23) will be the same as equation (20) if the transaction cost in equation (20) is reduced to a half. Therefore, we can expect case 2 to have a higher return than case 1 for a given value of c. A formal statement and rigorous proof is given by Proposition 4.1 at the end of this section.

Case 3 $b<-a$

In this case, one may expect more profit in each trading cycle, but the expected time in each cycle is longer. Different from the two earlier cases, only the first-passage time over the one-sided boundary is used. The optimization problem is:

$$\begin{aligned}\max_{a,b}\quad & f(a,b)\\ &=\frac{a-b-c}{E[\tau_1]+E[\tau_2]}=\frac{a-b-c}{\sum_{n=0}^{\infty}\frac{(-\sqrt{2}b)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)}\\ \text{subject to}\quad & b<-a\end{aligned} \tag{24}$$

where τ_1 and τ_2 are both first-passage times over a one-sided boundary.

This time, the expected time for one trading cycle $E[\tau_1]+E[\tau_2]=\sum_{n=0}^{\infty}\frac{(-\sqrt{2}b)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)$ does not depend on a. For a given value of b, the objective function $f(a,b)=\frac{a-b-c}{\sum_{n=0}^{\infty}\frac{(-\sqrt{2}b)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)}$ is a linearly increasing function of a. To maximize $f(a,b)$, a should be as large as possible. In this case, since $b<-a$, the largest a tends to the boundary $a^*=-b$ for any fixed value of b. If $b>0$, the optimal solution will be infeasible since we restrict $a\ge 0$. When $b\le 0$, the problem goes back to case 2 and we only need to solve equation (23) to get the value of a^* and thus get b^* by $b^*=-a^*$.

Out of the three cases, we have seen two different optimal rules which gives two different values of a^* and b^*. We call the optimal rule in case 1 the 'Conventional Optimal Rule' since it clears position exactly when the spread reverts to the mean at $b^*=0$, which is consistent with the common practice. In contrast, we call the rule in case 2 the 'New Optimal Rule', which basically allows no waiting time between the two trades. Since the 'New Optimal Rule' cuts the transaction costs in a half compared to the 'Conventional Optimal Rule', it is intuitive that the 'New Optimal Rule' performs better than the 'Conventional Optimal Rule'. Formally, we state the proposition below:

Proposition 4.1 *When there is no transaction cost ($c=0$), the maximal return in case 1 is the same as the maximal return in case 2. When transaction cost exists ($c>0$), the maximal return in case 1 is strictly smaller than the maximal return in case 2.*

Proof see Appendix B.

Graphically, the comparison between the two rules in the theoretical level are shown in figure 1. The 'New Optimal Rule' (red curve) is always better than the 'Conventional Optimal Rule' (blue curve). The advantage is more apparent when the transaction cost increases, despite the fact that the expected returns for both rules decrease as the transaction cost increases.

The comparison is only made in terms of the profit per unit time since it is our objective in this paper. So 'better' only means more profit per unit time. For traders who are more concerned about the risk, we show the variance per unit time for these two methods in figure 2. Naturally, the risk of 'New Optimal Rule' is always higher than the 'Conventional Optimal Rule' since the expected return is higher. To take risk into account, Sharpe ratio or the mean-variance optimization can be considered. In figure 2, we can see that when the transaction cost increases to a very large value, the change of variances of both rules is small, but the change of the expected return is relatively large. Even when we consider the risk, 'New Optimal Rule' can be more preferable when transaction cost is large enough. However, in this paper, we will only focus on the expected return per unit time.

5. Numerical examples

In this section, we will apply the two optimal rules derived in section 4 and compare them with the common practice using actual daily data. Comparison is made in two aspects. Firstly, for the same pair of stocks, we compare the profitability for different trading rules. Secondly, we compare the profitability among different pairs under the same rule.

5.1. *Comparison of different trading rules*

One of the most commonly used pairs is Coca-Cola (KO) and Pepsi (PEP). We collected 756 daily prices of the pair KO-PEP from Yahoo-Finance from 30 November 2009 to 29 November 2012. As shown in figure 3, their prices moved together.

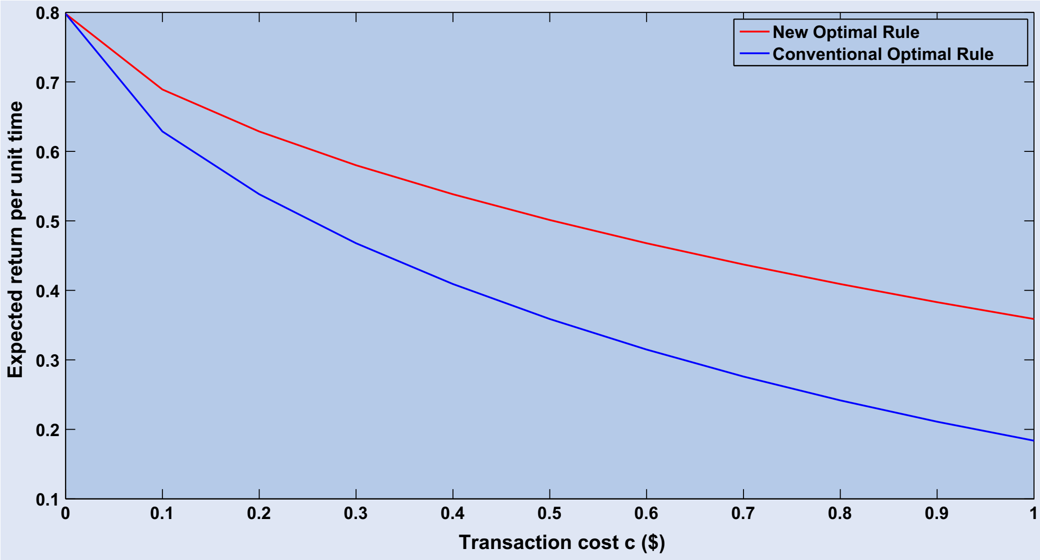

Figure 1. Comparison between the ‘Conventional Optimal Rule’ (case 1), and ‘New Optimal Rule’ (case 2). Optimal thresholds of the two rules developed in this paper are dependent on the transaction cost, thus they are shown to be curves instead of straight lines.

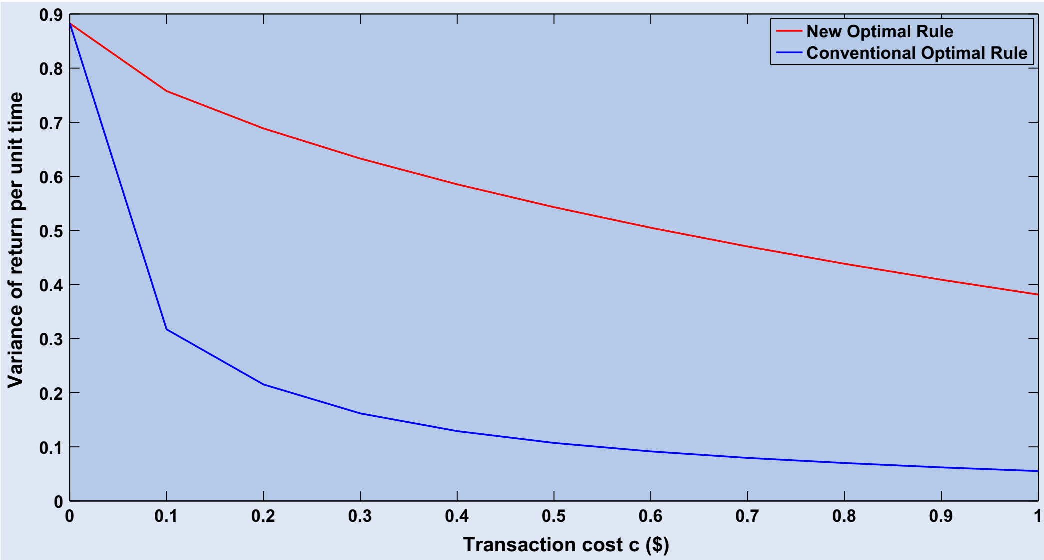

Figure 2. Comparison of the variance between the ‘Conventional Optimal Rule’ (case 1), and ‘New Optimal Rule’ (case 2).

Let the prices of PEP and KO be P_t and Q_t, respectively. Applying linear regression, we get $\ln(P_t) - \beta \ln(Q_t) = X_t$, where $\beta = 0.2187$. The residual X_t is assumed to follow an OU process $dX_t = \theta(\mu - X_t)dt + \sigma dW_t$. In this paper, we use the Maximum-Likelihood (ML) method to estimate the parameters based on Hu and Long (2007). The log likelihood for the process X_t is given by:

$$L(X|\mu,\theta,\sigma) = -\frac{n}{2} - \frac{1}{2}\sum_{i=1}^{n} \ln(1 - e^{-2\theta(t_i - t_{i-1})}) - \frac{\theta}{\sigma^2}\sum_{i=1}^{n} \frac{X_{t_i} - \mu - (X_{t_{i-1}} - \mu)e^{-\theta(t_i - t_{i-1})}}{1 - e^{-2\theta(t_i - t_{i-1})}}$$

Maximizing $L(X|\mu,\theta,\sigma)$, we get the estimation for the parameters: $\mu = 3.4241$, $\theta = 0.0237$ and $\sigma = 0.0081$. Assuming that the parameters are constant during the data collection period, we can apply our optimal pairs trading rules. We compare in figure 4 our ‘New Optimal Rule’ and ‘Conventional Optimal Rule’ with two common practices, which take positions at one standard deviation (we call it ‘1-σ Rule’) or two standard deviations (‘2-σ Rule’) and clear positions when the spread reverts back to the mean. Since thresholds of common practices do not change with the transaction cost, the total return should be straight lines. Similarly, since thresholds of the two optimal rules vary with the transaction cost, the total return should be curves.

As predicted in section 4, the ‘New Optimal Rule’ performs best. There is a trend of decreasing profit for all of the rules as the transaction cost c increases, but the ‘New Optimal Rule’ performs increasingly better as c increases. The ‘Conventional Optimal Rule’ does not distinguish itself from the ‘1-σ Rule’ when the transaction cost is small, but tends to perform better as c increases. However, the result is not exactly as we expected. For example, there is a sudden drop in return at $c = 0.006$ dollars with both the ‘New Optimal Rule’ and ‘Conventional Optimal Rule’, and their profits are even less than the ‘1-σ

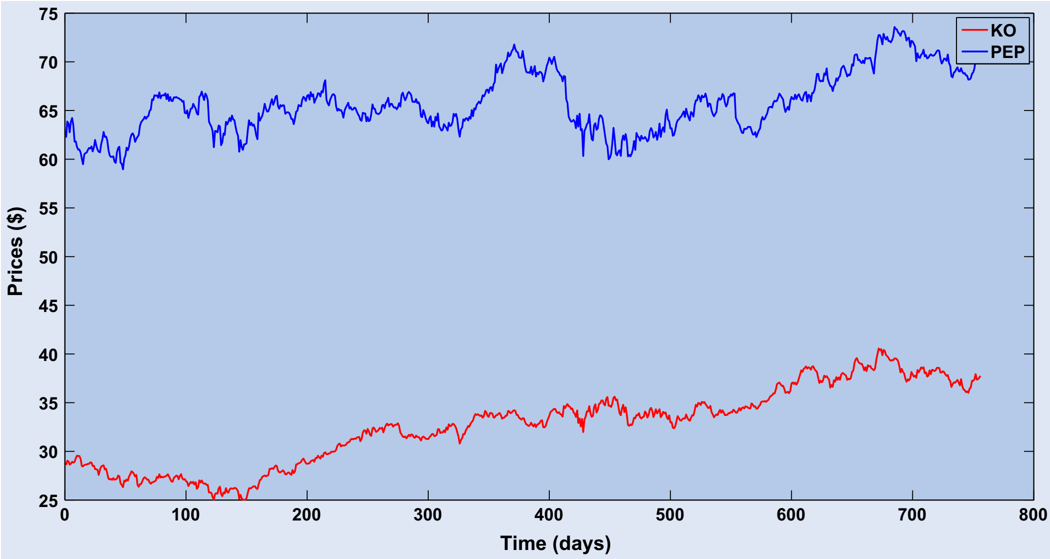

Figure 3. Actual adjusted daily prices of KO and PEP are shown in this graph. Time = 0 is the starting date on 30 November 2009.

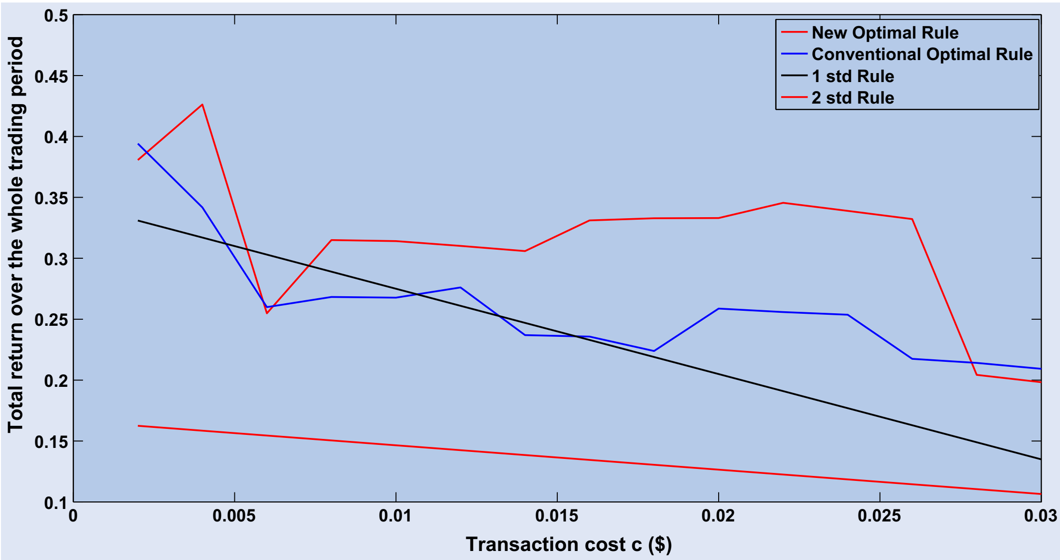

Figure 4. Comparison between the four rules in this paper using daily prices of KO and PEP. Generally, 'New Optimal Rule' performs best. 'Conventional Optimal Rule' performs slightly better than the '1-σ Rule' when c is small, and significantly better when c is larger. '2-σ Rule' performs worst.

Rule'. A possible cause might be that model parameters might have been poorly estimated, and/or even that X_t might not have been an OU process. To see the impact of model parameters, we conduct sensitivity analysis with each of the three parameters μ, θ, and σ, which is presented in Appendix D. We find the return to be very sensitive to the mean of the spread but not to the reversion rate θ nor to the standard deviation σ.

We also show the actual trading process for the 'New Optimal Rule' in table 1 and figure 5. In each trade, we assume that the transaction cost $c = 0.02$ dollars for each dollars invested. We transform the transaction cost into the dimensionless system and obtain the optimal thresholds as $a^* = 0.991$ and $b^* = -0.991$. Then, we transform back to get the real thresholds as $\tilde{a^*} = a^* \frac{\sigma}{\sqrt{2\theta}} + \mu = 3.4611$ and $\tilde{b^*} = b^* \frac{\sigma}{\sqrt{2\theta}} + \mu = 3.3871$. A trading is triggered whenever X_t reaches $\tilde{a^*}$ or $\tilde{b^*}$. The last trade is not counted since positions cannot be cleared within our trading period.

There have been a total of five trades over the three years with an average earning per trade at 6% and the total earning over the whole period at 33.33%.

5.2. *Comparison between different pairs*

We considered five pairs: Coca-Cola and Pepsi (KO_PEP), Target and Wal-mart (TGT_WMT), Dell and Hewlet-Packard (DELL_HPQ), RWE AG and E.On Se† (RWE_EOAN), and Chevron and Exxon Mobile (CVX_XOM). We computed the net returns of the five pairs under four trading rules given the actual daily prices between 30 November 2009 and 29 November 2012. The profitability indicator $\sigma\sqrt{\theta/2}$ and the

†RWE and E. On Se are German utility companies.

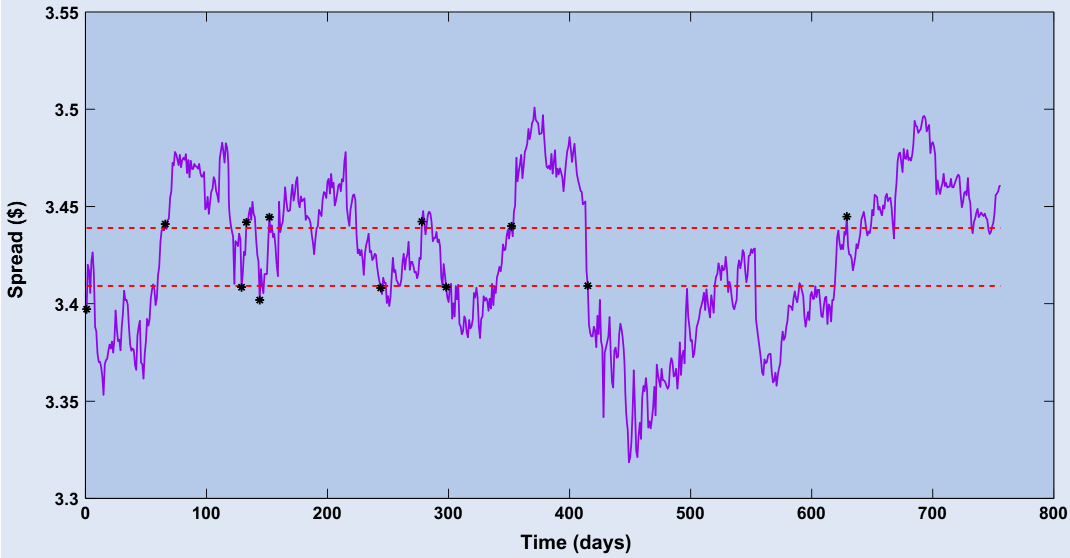

Figure 5. Dynamics of the spread is shown in the blue curve. Dashed lines are the optimal trading thresholds a and b. Block points are the day of opening and closing positions. They are not exactly on the dashed lines because trading is discrete. The graph shows that there are more trading opportunities in the first year.

Table 1. Details of transaction for each trade.

Trades	Status	Date	KO		PEP		Returns (%)	
			Prices($)	Action	Prices($)	Action	Total	Net
Trade 1	Open	10 December 2009	29.290	Sell $0.22	61.84	Buy $1	8.67	6.67
	Close	15 March 2010	26.825	Clear positions	66.15	Clear positions		
Trade 2	Open	15 March 2010	26.825	Buy $0.22	66.15	Sell $1	8.85	6.85
	Close	23 February 2011	31.955	Clear positions	62.93	Clear positions		
Trade 3	Open	23 February 2011	31.955	Sell $0.22	62.93	Buy $1	9.08	7.08
	Close	28 April 2011	33.705	Clear positions	69.72	Clear positions		
Trade 4	Open	28 April 2011	33.705	Buy $0.22	69.72	Sell $1	9.02	7.02
	Close	26 July 2011	34.595	Clear positions	64.07	Clear positions		
Trade 5	Open	26 July 2011	34.595	Sell $0.22	64.07	Buy $1	7.72	5.72
	Close	26 July 2012	39.425	Clear positions	71.22	Clear positions		

Table 2. Profitability indicator and returns for different pairs.

	$\sigma\sqrt{\theta/2}$	New (%)	Conventional (%)	1-σ (%)	2-σ (%)	Average (%)
RWE_EOAN	0.00142	39	38	39	23	35
TGT_WMT	0.00126	66	28	29	28	38
KO_PEP	0.00088	33	26	20	13	23
DELL_HPQ	0.00083	54	40	40	0	34
CVX_XOM	0.00071	32	26	26	26	28
Average		45	32	31	18	**32**

net returns for the five pairs under the four trading rules are summarized in table 2.

The last row of table 2 is the average return of the trading rules. It is clear that the 'New Optimal Rule' out-performs the rest. Moreover, if we invest only on the two most profitable pairs – RWE_EOAN and TGT_WMT – according to $\sigma\sqrt{\theta/2}$, the average return will increase significantly. That is, if the investor invests only on RWE_EOAN and TGT_WMT, the return under the "New Optimal Rule" would have been 52.5% $(= (39\% + 66\%)/2)$, an increase of 7.5% over the average 45% shown in the last row of table 2. The increases under other trading rules are 1.0% for the 'Conventional Optimal Rule', 3.0% for the '1-σ Rule' and 7.5% for the '2-σ Rule'. This demonstrates that the profitability indicator $\sigma\sqrt{\theta/2}$ can be used as a guideline to select a profitable pair for trading.

6. Conclusions

In this paper we derived a polynomial form of the expectation of the first-passage time over two-sided boundaries,which allows relatively straightforward optimization of pairs trading strategies. The main focus of the paper is the maximization of the expected return per unit time in the long run. The log returns of a pair are co-integrated, and we assume the co-integration residuals follow an OU process. Contrary to common practice, we show that it is optimal to set the trading thresholds symmetric around the mean of the residuals X_t. In other words, it is optimal for the trader to take the opposite positions exactly when she used to clear the positions under the conventional trading rule.

We discussed four trading rules in this paper, and compared their maximum expected returns. In theory, we showed that the 'New Optimal Rule' outperforms the 'Conventional Optimal Rule'. Both rules proposed in this paper perform better than the common practice, especially when the transaction cost is high. We also proposed the quantity $\sigma\sqrt{\theta/2}$ as a measure for potential profit.

There are limitations to these results, which we leave for future studies. While the strategies we developed are directly applicable, the application is limited since we have assumed the model

parameters are constant. Though model parameters may remain unchanged in a short period of time, it is unrealistic to apply the strategies in the long run. The parameters are subject to various factors such as the overall market states and internal management issues within individual firms. It remains to be investigated how the changing model parameters impact on the optimal thresholds.

References

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

Bertram, W., Analytic solutions for optimal statistical arbitrage trading. *Physica A*, 2010, **389**(11), 2234–2243.

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

Darling, D. and Siegert, A., The first passage problem for a continuous Markov process. *Ann. Math. Stat.*, 1953, **24**(4), 624–639.

Do, B., Faff, R. and Hamza, K., A new approach to modeling and estimation for pairs trading, In *Proceedings of 2006 Financial Management Association European Conference*, Stockholm, June, 2006.

Ekström, E., Lindberg, C. and Tysk, J., Optimal liquidation of a pairs trade. In *Advanced Mathematical Methods for Finance*, edited by J. Di Nunno and B. Oksendal, pp. 247–255, 2011 (Springer-Verlag: Berlin).

Elliott, R., Van Der Hoek, J. and Malcolm, W., Pairs trading. *Quant. Finance*, 2005, **5**(3), 271–276.

Gatev, E., Goetzmann, W. and Rouwenhorst, K., Pairs trading: Performance of a relative-value arbitrage rule. *Rev. Financ. Stud.*, 2006, **19**(3), 797–827.

Hu, Y. and Long, H., Parameter estimation for Ornstein–Uhlenbeck processes driven by α-stable Lévy motions. *Commun. Stoch. Anal.*, 2007, **1**(2), 175–192.

Lin, Y., McCrae, M. and Gulati, C., Loss protection in pairs trading through minimum profit bounds: A cointegration approach. *J. Appl. Math. Decis. Sci.*, 2006, **2006**, 1–14.

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

Puspaningrum, H., Lin, Y. and Gulati, C., Finding the optimal preset boundaries for pairs trading strategy based on cointegration technique. *J. Stat. Theory Pract.*, 2010, **4**(3), 391–419.

Ricciardi, L. and Sato, S., First-passage-time density and moments of the Ornstein–Uhlenbeck process. *J. Appl. Probab.*, 1988, **25**, 43–57.

Sato, S., Evaluation of the first-passage time probability to a square root boundary for the Wiener process. *J. Appl. Probab.*, 1977, **14**, 850–856.

Thomas, M., Some mean first-passage time approximations for the Ornstein–Uhlenbeck process. *J. Appl. Probab.*, 1975, **12**, 600–604.

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

Vladislav, K., Optimal convergence trading. Technical report, EconWPA, 2004.

Appendix A. Proof of proposition 4.1

Proof We use the symbolic calculation from Mathematica to get the following results:

$$\exp\left(\frac{x^2}{2}\right)\int_0^\infty \ln(t)\left(-\frac{t^2}{2}\right)\cos(xt)dt = -\frac{1}{2}\sqrt{\frac{\pi}{2}}\left[\gamma+\ln(2)+Hypergeometric_1F_1^{\{1,0,0\}}\left(0,\frac{1}{2},\frac{x^2}{2}\right)\right]$$

where $\gamma=\int_0^\infty \exp(-t)\ln(t)dt$ is the Euler constant. Since the definite integral $\int_0^\infty \exp(-x^2)\ln(x)dx=-\frac{\sqrt{\pi}}{4}[\gamma+2\ln(2)]$, it is easy to verify that $C=\int_0^\infty \ln(t)\left(-\frac{t^2}{2}\right)dt=-\frac{1}{2}\frac{\sqrt{\pi}}{2}[\gamma+2\ln(2)]$. The polynomial form of the Kummer confluent hypergeometric function is $Hypergeometric_1F_1(a,b,z)=\sum_{k=0}^\infty \frac{(a)_k}{(b)_k}\frac{z^k}{k!}$, where $(x)_n=x(x+1)(x+2)\ldots(x+n-1)$. By the chain rule we have $\frac{d(x)_n}{dx}=(n-1)!$ at $x=0$. Therefore it is easy to verify that:

$$Hypergeometric_1F_1^{\{1,0,0\}}\left(0,\frac{1}{2},\frac{x^2}{2}\right)=\sum_{n=1}^\infty \frac{2^n(n-1)!}{(2n)!}x^{2n}$$

Therefore, we get:

$$\exp\left(\frac{x^2}{2}\right)\int_0^\infty \ln(t)\left(-\frac{t^2}{2}\right)\cos(xt)dt = -\frac{1}{2}\sqrt{\frac{\pi}{2}}\sum_{n=1}^\infty \frac{(\sqrt{2}x)^{2n}}{(2n)!}\Gamma(n)+\int_0^\infty \ln(t)\left(-\frac{t^2}{2}\right)dt$$

□

Appendix B. Proof of proposition 4.2

Let $f^i(a_i,b_i)$ be the maximal expected return per unit time in case i with the optimal thresholds a_i and b_i, for $i=1,2$. We only have to show:

$$f^1(a_1,b_1)=f^2(a_2,b_2),\quad \text{for } c=0 \tag{B1}$$
$$f^1(a_1,b_1)<f^2(a_2,b_2),\quad \text{for } c>0 \tag{B2}$$

Proof In both case 1 and case 2, the expected times of one trading cycle are both $E[T]=\frac{1}{2}\sum_{n=0}^\infty \frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)-\frac{1}{2}\sum_{n=0}^\infty \frac{(\sqrt{2}b)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)$. In case 1, since $b_1=0$, the objective function is given by:

$$F^1(a,c)=\frac{a-c}{\frac{1}{2}\sum_{n=0}^\infty \frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)} \tag{B3}$$

and the optimal function for case 2 is:

$$F^2(a,c)=\frac{a-\frac{c}{2}}{\frac{1}{2}\sum_{n=0}^\infty \frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)} \tag{B4}$$

Obviously, when $c=0$, we have $F^1(a,c)=F^2(a,c)$ for any value of a. Therefore B1 is proved.

To prove B2, we make use of the fact that the optimal values of a_1 and a_2 satisfy the equations (20) and (23), respectively. Therefore, in the optimal solution we have:

$$F^1(a_1,c) = \frac{a_1 - c}{\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a_1)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right)} = \frac{1}{\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a_1)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)} \tag{B5}$$

and similarly

$$F^2(a_2,c) = \frac{1}{\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a_2)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)} \tag{B6}$$

Since $\frac{1}{\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}x)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)}$ is a decreasing function of x, to prove $F^2(a_2,c) > F^1(a_1,c)$, we only need to prove $a_2 < a_1$. Let $a(x)$ satisfy the following equation:

$$\frac{1}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n+1}}{(2n+1)!}\Gamma\left(\frac{2n+1}{2}\right) = (a-x)\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)$$

It is easy to see that $a(x) > x$ for any $x > 0$. Obviously, $a_2 = a(\frac{c}{2})$ and $a_1 = a(c)$. Since $c > 0$, all we need is to prove $a(x)$ is a strictly increasing function. Taking the derivative of x on both sides of the equation above, we get:

$$a'(x) = \frac{\frac{\sqrt{2}}{2}\sum_{n=0}^{\infty}\frac{(\sqrt{2}a)^{2n}}{(2n)!}\Gamma\left(\frac{2n+1}{2}\right)}{(a-x)\sum_{n=1}^{\infty}\frac{(\sqrt{2}a)^{2n-1}}{(2n-1)!}\Gamma\left(\frac{2n+1}{2}\right)} > 0$$

Appendix C. Comparison of simulation and analytic solution

We have obtained the polynomial form for the expectation of first-passage time over two-sided boundaries in section 3. Now we want to verify these results by comparing it with simulations. For given a and b, we simulate 10^6 paths and for each path we get a τ_2 defined in section 2. Then we calculate $E[\tau_2]$ by taking the average of all the simulated values of τ_2. On the other hand we calculate the expectation directly from the polynomial form. Comparison is shown in figures C1 and C2.

The blue lines are the simulation results and thus vibration is expected. The red lines are the results from the polynomial form and therefore it is very smooth. The simulation and the polynomial form are very close, which verify our polynomial form.

Appendix D. Sensitivity analysis

Fix θ and σ and we impose a small shock on μ, we find the sensitivity analysis for μ in table D1 and figure D1.

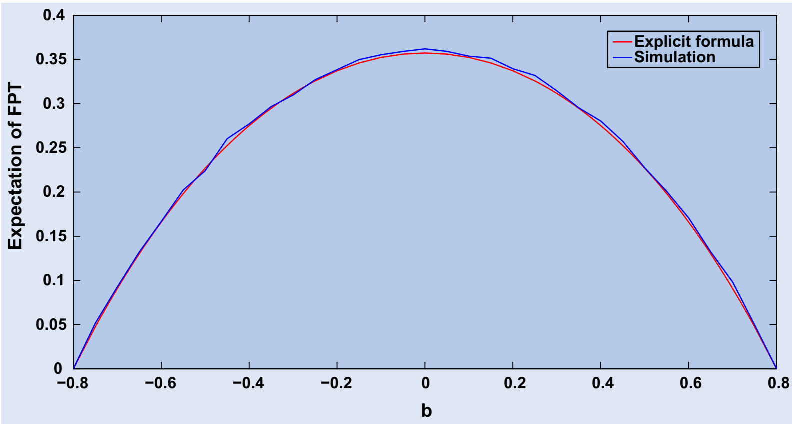

Figure C1. The red curve is from our polynomial form of the expectation in (18), and the blue curve is from the simulation. We fix the upper at $a = 0.8$ and lower boundaries at $-a = -0.8$. For any starting point $b \in (-0.8, 0.8)$ inside the boundaries, we can get the expected time for the OU process to reach the boundaries either by simulation or using our polynomial form.

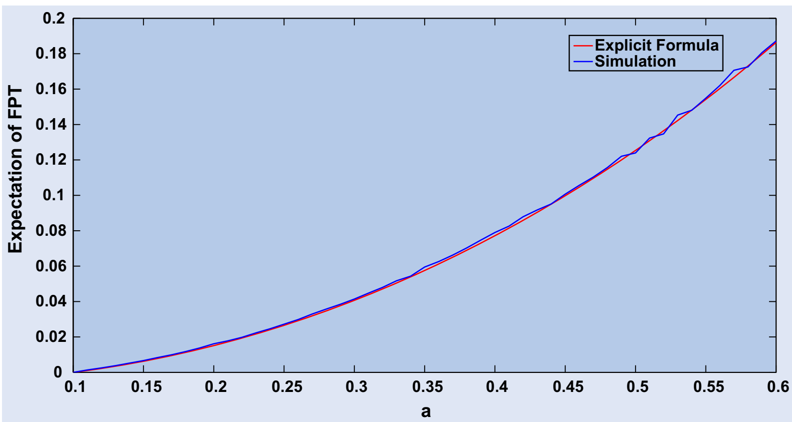

Figure C2. The red curve is from our polynomial form of the expectation in (18), and the blue curve is from the simulation. We fix our starting point at $b = 0.1$ and let the upper boundary $a \in (0.1, 0.6)$.

Table D1. Sensitivity analysis for μ.

Mean μ	3.36	3.37	3.39	3.41	**3.42**	3.44	3.45	3.475477	3.49
Percentage of change	−2%	−1.50%	−1%	−0.50%	**0%**	0.50%	1%	1.50%	2%
Total Return	0.12	0.12	0.14	0.24	**0.33**	0.32	0.21	0	0
Percentage of change	−65%	−64%	−59%	−27%	**0%**	−5%	−38%	−100%	−100%

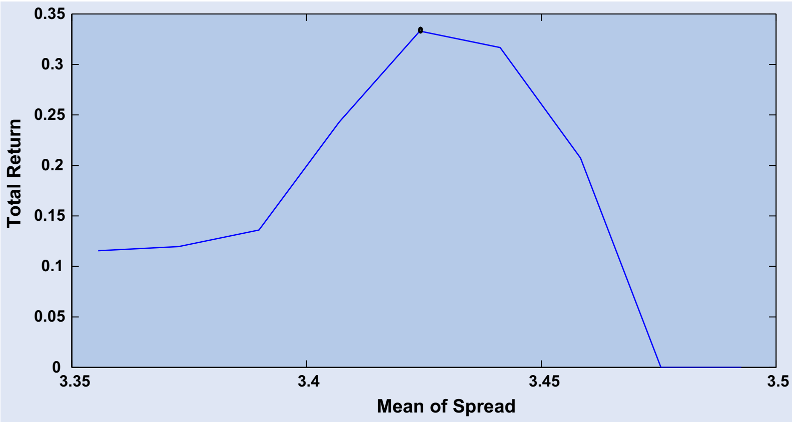

Figure D1. Sensitivity analysis for μ. The black spot is the estimated value of μ.

Table D2. Sensitivity analysis for θ.

Reversion rate θ	0.014	0.017	0.019	0.021	**0.024**	0.026	0.028403	0.031	0.033
Percentage of change	−40%	−30%	−20%	−10%	**0%**	10%	20%	30%	40%
Total Return	0.23	0.37	0.36	0.34	**0.33**	0.33	0.32	0.31	0.28
Percentage of change	−31%	12%	8%	3%	**0%**	−1%	−3%	−7%	−15%

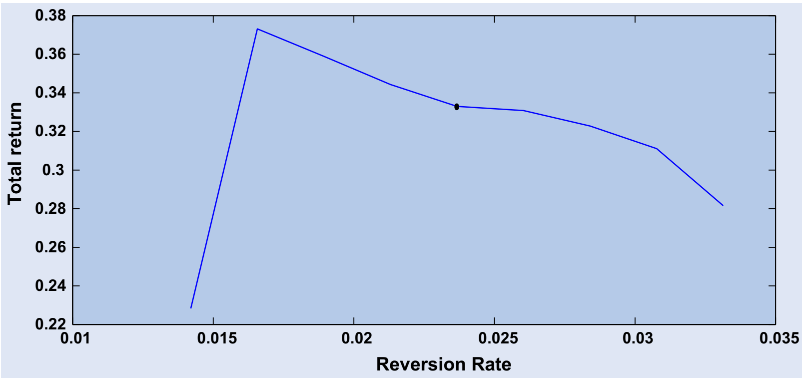

Figure D2. Sensitivity analysis for θ. The black spot is the estimated value of θ.

Table D3. Sensitivity analysis for σ.

Standard deviation σ	0.0049	0.0057	0.0065	0.0073	**0.0081**	0.0089	0.0098	0.011	0.011
Percentage of change	−40%	−30%	−20%	−10%	**0%**	10%	20%	30%	40%
Total Return	0.26	0.26	0.28	0.31	**0.33**	0.36	0.37	0.23	0.23
Percentage of change	−21%	−21%	−17%	−6%	**0%**	8%	12%	−31%	−31%

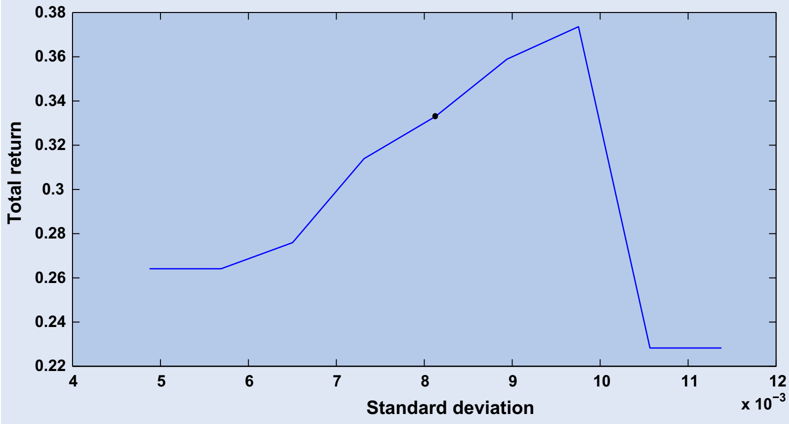

Figure D3. Sensitivity analysis for σ. The black spot is the estimated value of σ.

For a small change of μ, the change for the return is very large. Therefore, it is crucial to have a very accurate estimation of μ. In our example of KO and PEP, we have a relatively accurate estimation of μ since the change in each direction results in a huge loss.

Then we fix μ and σ and conduct sensitivity analysis for θ shown in table D2 and figure D2.

The change of return is relatively inactive to the change of the reversion rate θ. In fact, small changes of θ does not affect the return at all. It only affects the return when the change of θ is large enough. From figure D2, we can see that a smaller value of θ actually gives a higher return.

Lastly, we show the sensitivity analysis for σ in table D3 and figure D3.

Similar as θ, the change of σ is not influential to the change of the total return. A larger value of σ may result in a higher value in our example.