Copula-Based Pairs Trading Strategy

Wenjun Xie and Yuan Wu
Division of Banking and Finance, Nanyang Business School,
Nanyang Technological University, Singapore

ABSTRACT

Pairs trading is a technique that is widely used in the financial industry and its profitability has been constantly documented for various markets under different time periods. The two most commonly used methods in pairs trading are distance method and co-integration method. In this research work, we propose an alternative approach for pairs trading using copula technique. The proposed method can capture the dependency structure of co-movement between the stocks and is more robust and accurate. Distance method and co-integration method can be generalized as special cases of the proposed copula method under certain dependency structures.

Keywords: pairs trading; copula; dependency structure; trading strategy

1. Introduction

"Pairs trading" is a well-known Wall Street investment strategy. It was first established by the Wall Street quant Nunzio Tartaglia's team in Morgan Stanley in 1980s. It was first documented in literature by Gatev et al. (2006) where they tested the pairs trading strategy using the daily data from 1962 to 2002 in the U.S equity market. It was documented that this simple trading rule yields average annualized excess returns of up to 11% for self-financing portfolios of pairs. Later, other people extended the analysis to various international markets under different time periods and confirmed the profitability of this strategy. (Perlin, 2009; Simone et al., 2010; Binh Do et al., 2010).

The main idea of pairs trading strategy is to find a pair of stocks which have strong co-movement in prices in their history, and identify the relative over-valued and under-valued positions between them in the following period. Such relative mispricing occurs if the spread between the two stocks deviates from its equilibrium, and excess returns will be generated if the pair is mean-reverting. The trading strategy is to long a stock that is under-valued and short a stock that is over-valued at the same time, and close the positions when they return to fair values.

There are two most commonly used methods in pairs trading. The first is called distance method. The main idea of this method is to use the distance between the standardized prices of two stocks as the criteria to select pairs and form trading opportunities (Gatev et al., 2006; Perlin, 2009; Simone et al., 2010; Binh Do et al., 2010). Another method is called co-integration method. It uses the idea of co-integrated pairs proposed by Nobel laureate Clive Granger (1987). If two stocks are co-integrated with each other, the difference in their returns should perform mean-reverting property. Then long/short positions are constructed when the difference is relatively large and closed when the difference reverts to its mean level (Lin et al., 2006; Galenko et al., 2012).

Despite the success of this trading strategy documented in the literature, recent research conducted by Binh Do et al. (2010) confirmed the downward trend in profitability of pairs trading strategy. They proposed alternative methods such as choosing pairs within confined industry to increase the profitability.

Both distance method and co-integration method involves the idea of spread. They use spread either between standardized prices or returns to capture the idea of mispricing. Equal spreads are assumed to have same levels of mispricing regardless of the stock prices and returns. If the data is jointly normally distributed, then linear correlation completely describes dependency, and the distribution of spread is stationary under different prices and returns. On the other hand, the dependence between two variables has both linear and non-linear components, and it has many other forms of expressions rather than simple linear correlation coefficient, e.g. tail dependence. It is widely accepted that the financial data are rarely normally distributed in reality, and therefore correlation cannot completely describe the dependency structure (Kat, 2003; Crook and Moreira, 2011).

In this research work, a new algorithm for pairs trading using copula technique is proposed, and it will capture the dependency structure of the stocks. Liew and Wu (2013) documented several empirical evidences of extra profits using copula technique in pairs trading. The purpose of this research is to set up the theoretical foundation of copula-based pairs trading strategy and make comparison with traditional methods. We are not claiming that it is superior to distance method or co-integration method, but it provides an alternative way to looking at pairs trading.

The idea of copula arose as early as 19th century in the context of discussions regarding non-normality in multivariate cases. Sklar (1959) set the foundation of copula method which links the joint distribution of two variables to their separate margins. The advantage of copula method is that it describes the structure of dependence between several variables, and thus is more robust and accurate.

As mentioned earlier, linear correlation can fully explain the dependency structure if the two stocks follow joint normal distribution. In this research, it is shown that distance method can be generalized as a special case of copula method if the data do follow joint normal distribution. It is also shown that the co-integration method will generate the same trading opportunities as copula method under certain restrictions. In reality, the data structure is often complicated and no one knows the true marginal distributions and joint relation. Copula method has the advantage of estimating optimal marginal distributions and optimal copula separately, and thus has the optimal estimation of the joint distribution between two stocks, resulting in more trading opportunities. Several examples will be given to illustrate these points. Note therefore that, we only focus on the trading strategy at the current stage and assume that the pairs are already selected out.

The rest of the paper is organized as follows: Section 2 describes the traditional methods and introduces the new trading algorithm using copula technique. Section 3 generalizes current methods as special cases of copula methods under certain conditions and real examples are given. Section 4 concludes the paper.

2. Trading algorithm using copula method

2.1 Two commonly used methods in pairs trading

Assuming the pair, stock X and stock Y, is already picked out as the candidate for pairs trading, either by distance method or co-integration method. Under the general framework of pairs trading, the prices of the stocks from the selected pair are assumed to move together over time. In other words, under the arbitrage pricing model (APT), the two stocks should have exactly same risk factors exposures, and the expected returns in a given time duration are the same.

Now, we define P_t^X and P_t^Y to be the prices of stock X and stock Y at time t respectively (P_0^X and P_0^Y at time 0), and R_t^X and R_t^Y to be the cumulative returns of the stock X and stock Y at time t .Then

$$P_t^X = P_0^X e^{R_t^X}$$
$$P_t^Y = P_0^Y e^{R_t^Y}.$$

By taking logarithm at both sides, we obtain that

$$R_t^X = \log(P_t^X) - \log(P_0^X)$$
$$R_t^Y = \log(P_t^Y) - \log(P_0^Y).$$

According to the APT model, the expected return of stock X should equal to the expected return of stock Y at time t as long as the stock X and stock Y share the same risk factors exposures, i.e. $E(R_t^X) = E(R_t^Y) = R_t$. Moreover, R_t^X and R_t^Y can be expressed in the following form:

$$R_t^X = E(R_t^X) + \varepsilon_t^X = R_t + \varepsilon_t^X$$
$$R_t^Y = E(R_t^Y) + \varepsilon_t^Y = R_t + \varepsilon_t^Y.$$

ε_t^X and ε_t^Y represent the error terms which make the returns of stock X and stock Y deviate from the expected returns, and $E(\varepsilon_t^X) = E(\varepsilon_t^Y) = 0$.

Two commonly used methods in pairs trading are distance method and co-integration method. For the distance method, it uses the spread between the standardized prices of the two stocks as the measurement of degree of mispricing. For the co-integration method, it uses the spread between the returns of the two stocks as the measurement. The idea of spread captures how close the two stocks are. If the two stocks follow joint normal distribution, it can be easily proved that the distribution of the spread in both distance method and co-integration method are robust, and the spread itself fully captures the dependency structure of the two stocks. However, in reality, the true distribution of the error terms ε_t^X and ε_t^Y are unknown, and the dependency structure of the error terms and the expected return R_t is also unobservable, which makes the assumption of joint normal distribution invalid.

The disadvantages of the current measurements of degree of mispricing under distance method or co-integration method are similar and can be illustrated from two aspects, although our illustration is centered on co-integration method here. For example, if the trading strategy is to construct long/short positions when the spread between the returns of two stocks,

$R_t^X - R_t^Y$, is larger than Δ or smaller than $-\Delta$ and close the positions when the spread return to 0. The following two situations may happen:

Situation 1
Although the spread between the returns of two stocks has reached Δ, however, conditioning on the level R_t^X and R_t^Y, the variance of the spread tends to be larger than the mean sample variance. Thus, a potential scenario is that the spread will continue to grow larger. Consequently, the long/short position entering at spread equal to Δ may reach the pre-determined stop-loss position and it results in a loss of capital. Another scenario is that it may require longer time to revert to the 0 spread position. Neither of the scenarios is preferred by investors.

Situation 2
Next, we consider a time point which the spread is still less than Δ, however, conditioning on the return level R_t^X and R_t^Y, the spread tends to have smaller variance than the mean sample variance, so it may already be a good opportunity to trade. Considering the criteria that we only trade when the spread is larger than Δ or smaller than $-\Delta$, this trading opportunity is missed.

Two simulated series of stock prices are used to illustrate the above two situations. Figure 1(a) shows the movement of the two simulated series of stock prices. The vertical line divides the whole sample into formation period and trading period. Dickey-Fuller test has been used to confirm that they are co-integrated pairs. Figure 1(b) shows the spread of the return between them. It is simulated with the property that the variance of the spread is higher when the stock prices are at high level, which is σ_H, and the variance of the spread is lower when the stock prices are at low level, which is σ_L. The bounds refer to the $\pm 1.65\sigma$ level lines, where σ is the variance of the spread during formation period. It is easily known that $\sigma_H > \sigma > \sigma_L$. We use this value $\pm 1.65\sigma$ as the threshold value Δ. Consequently, at high price levels, after reaching Δ, the spread is going to grow larger and results in unfavorable situation. This is because the true σ_H at high price level is larger than σ, so it is optimal to trade when the spread reach $\pm 1.65\sigma_H$ instead of $\pm 1.65\sigma$. Again, at low price levels, some opportunities are missed by co-integration method since it is optimal to trade when the spread reach $\pm 1.65\sigma_L$ instead of $\pm 1.65\sigma$.

Through this example, it can be said that traditional pairs trading methods are not optimal if there are different dependency pattern conditioning on different price levels. It is common that the real data does not have such apparent patterns of variances at different price levels, but different patterns of spreads do exist for different price levels. This example is used as a magnifying version of the potential real situation. The distance method may produce the same situations as above.

2.2 A new measurement of degree of mispricing

Considering the situations discussed above, measurements of mispricing used in traditional methods (distance method and co-integration method) are robust only under the validity of joint normal distribution assumption. A new measurement of mispricing is desirable if it is able to remove the restrictions required by traditional methods, and should also have the following properties:

Comparability – For a given time point, the numerical value of this measurement can reflect the degree of mispricing. By comparing the values of this measurement in two realizations, it can be decided which realization is more mispriced.

Consistency - This measurement has to be consistent and comparable over different time periods and return levels. This means that two situations are of same degree of mispricing if and only if their measurements are of the same numerical values no matter what the time period and expected return levels are.

To satisfy the above properties, a new measurement of the degree of mispricing is proposed.

Definition Given time t, the returns of the two stocks are $R_t^X = r_t^X$ and $R_t^Y = r_t^Y$. Define $P(R_t^X < r_t^X | R_t^Y = r_t^Y)$, the conditional probability of stock X's return smaller than the current realization r_t^X given the stock Y's return equal to current realization r_t^Y as the **Mispricing Index of X given Y**, denoted as $\boldsymbol{MI_{X|Y}}$. $\boldsymbol{MI_{Y|X}}$ is defined in the same way.

Theorem 1 Assuming E and E' are two possible scenarios of the stock prices at time t. $R_t^X = r_t^X$ and $R_t^Y = r_t^Y$ in scenario E while $R_t^X = r_t^{X'}$ and $R_t^Y = r_t^{Y'}$ in scenario E'. $E(R_t^X) = E(R_t^Y) = R_t$. $r_t^Y = r_t^{Y'}$, $\varepsilon_t^X = \delta_t^X$, $\varepsilon_t^Y = \delta_t^Y$, $\varepsilon_t^{X'} = \delta_t^{X'}$ and $\varepsilon_t^{Y'} = \delta_t^{Y'}$. $MI_{X|Y} > MI_{X|Y}'$ if and only if $\delta_t^X - \delta_t^Y > \delta_t^{X'} - \delta_t^{Y'}$.

<u>Proof.</u>

$$\begin{aligned} MI_{X|Y} &= P(R_t^X < r_t^X | R_t^Y = r_t^Y) \\ &= P(R_t^X - R_t^Y < r_t^X - R_t^Y | R_t^Y = r_t^Y) \\ &= P(R_t^X - R_t^Y < r_t^X - r_t^Y | R_t^Y = r_t^Y) \\ &= P((R_t^X - R_t) - (R_t^Y - R_t)) < r_t^X - r_t^Y | R_t^Y = r_t^Y) \\ &= P(\varepsilon_t^X - \varepsilon_t^Y < r_t^X - r_t^Y | R_t^Y = r_t^Y) \\ &= P(\varepsilon_t^X - \varepsilon_t^Y < (r_t^X - r_t) - (r_t^Y - r_t) | R_t^Y = r_t^Y) \\ &= P(\varepsilon_t^X - \varepsilon_t^Y < \delta_t^X - \delta_t^Y | R_t^Y = r_t^Y) \end{aligned}$$

Although the detailed numerical values of ε_t^X and ε_t^Y are not observed, the difference, however, is captured by the difference of returns. It can be seen that $MI_{X|Y}$ is a monotonic increasing function of $\delta_t^X - \delta_t^Y$ given that $R_t^Y = r_t^Y$. This suggests that a higher value of $MI_{X|Y}$ indicates a higher degree of mispricing, and thus satisfies the comparability property. #

The consistency of this new measurement can be understood intuitively. Although the idea of spread is not used directly in $MI_{X|Y}$, it still explores the similar idea. The conditional probability of return of stock X smaller than the current realization given the return of stock Y fixed can be seen as the conditional probability that the spread smaller than the current realization given the return of stock Y fixed. As discussed above, the traditional methods ignore the non-linear correlation between the spread and the return levels. For the new measurement $MI_{X|Y}$, it is able to obtain the conditional distribution of the spread given each level of returns. $MI_{X|Y}$ indicates the likelihood that the spread is smaller than the current realization when given the information. Since the likelihood is comparable across different time periods and return levels, it can be concluded that $MI_{X|Y}$ is consistent.

In the traditional methods, the mean-reverting property of the spread is assumed. For the new measurement, though we condition the spread on the return levels, it is assumed that it only affects the shape of the distribution, and $MI_{X|Y}$ will still perform mean-reverting property.

Theorem 2 If the current realizations of returns of stock X and stock Y equal to the expected return $R_t = r_t$, then $MI_{X|Y} = 0.5$.

Proof.

$$\begin{aligned}
MI_{X|Y} &= P(R_t^X < r_t^X | R_t^Y = r_t^Y) \\
&= P(R_t^X < r_t | R_t^Y = r_t) \\
&= P(R_t^X - R_t^Y < r_t - r_t^Y | R_t^Y = r_t) \\
&= P(\varepsilon_t^X - \varepsilon_t^Y < r_t - r_t | R_t^Y = r_t) \\
&= P(\varepsilon_t^X - (R_t^Y - R_t) < r_t - r_t | R_t^Y = r_t) \\
&= P(\varepsilon_t^X - (r_t^Y - r_t) < r_t - r_t | R_t^Y = r_t) \\
&= P(\varepsilon_t^X - 0 < r_t - r_t | R_t^Y = r_t) \\
&= P(\varepsilon_t^X < r_t - r_t | R_t^Y = r_t) \\
&= P(\varepsilon_t^X < 0 | R_t^Y = r_t) \\
&= P(\varepsilon_t^X < 0) \\
&= 0.5
\end{aligned}$$
#

This suggests that when the error terms of the returns of the two stocks equal 0, $MI_{X|Y}$ should be equal to 0.5, which means that under fair prices, $MI_{X|Y}$ is 0.5. Similarly, it can be proved that under fair prices, $MI_{Y|X}$ is also 0.5.

Corollary If $MI_{X|Y} > 0.5$, then stock X is overpriced relative to stock Y.
If $MI_{X|Y} = 0.5$, then stock X is fairly priced relative to stock Y.
If $MI_{X|Y} < 0.5$, then stock X is underpriced relative to stock Y.

It has been discussed that the new measurement $MI_{X|Y}$ has the property of comparability and consistency. Under fair prices, the $MI_{X|Y}$ is 0.5, and this means that stock X is fairly priced relative to stock Y when $MI_{X|Y} = 0.5$. However, this does not only apply to the situation that stock Y is really fairly priced. In the real market, we never observe what the true expected return is. However, due to consistency, we can conclude that stock X is fairly priced relative to stock Y if $MI_{X|Y} = 0.5$. When $MI_{X|Y} > 0.5$, stock X is overpriced relative to stock Y, and when $MI_{X|Y} < 0.5$, the stock X is underpriced than stock Y.

2.3 Mispricing index under copula framework

The procedure to obtain $MI_{X|Y}$ or $MI_{Y|X}$ using copula technique is provided below. The marginal distributions of the returns of stock X and Y are assumed to be F_1 and F_2, while joint distribution is H.

Theorem 3 (Sklar's Theorem) If $F(\cdot)$ is a n-dimensional cumulative distribution function for the random variables $X_1, X_2, \dots, X_n$ with continuous margins $F_1, \dots, F_n$, then there exists a copula function C, such that

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)).$$

If the margins are continuous, then C is unique. If any distribution is discrete, then C is uniquely defined on the support of the joint distribution.

Let $U = F_1(R^X)$ and $V = F_2(R^Y)$. According to Sklar's theorem, we can always identify a unique copula C which satisfies the following equation:

$$H(r^X, r^Y) = C(u, v) = C[F_1(r^X), F_2(r^Y)].$$

Theorem 4 Given current realization of returns of stock X and stock Y (r^X, r^Y), we have:

$$MI_{X|Y} = \frac{\partial C(u,v)}{\partial v} \; and \; MI_{Y|X} = \frac{\partial C(u,v)}{\partial u}.$$

<u>Proof.</u> Given (r^X, r^Y), the mispricing index of X given Y can be calculated as:

$$MI_{X|Y} = P(R^X < r^X | R^Y = r^Y) = P[F_1(R^X) < F_1(r^X) | F_2(R^Y) = F_2(r^Y)].$$

Let $u = F_1(r^X)$ and $v = F_2(r^Y)$, then:

$$MI_{X|Y} = P(U < u | V = v) = \frac{\partial C(u,v)}{\partial v}.$$

Similarly, we obtain that

$$MI_{Y|X} = P(V < v | U = u) = \frac{\partial C(u,v)}{\partial u}. \qquad \#$$

2.4 Copula-based pairs trading algorithm

Similar to the distance method and co-integration method, the trading strategy using copula technique still consists of two periods, which are formation period and trading period. Below is the trading algorithm assuming pair X and Y is already selected.

Step 1. Calculate daily returns for each stock during formation period. Estimate the marginal distributions of returns of stock X and stock Y, which are F and G separately.

Step 2. After obtaining the marginal distributions, estimate the copula described by Sklar's theorem to link the joint distribution H with margins F and G, denoted as C.

Step 3. During trading period, daily closing prices p_t^X and p_t^Y are used to calculate the daily returns (r_t^X, r_t^Y). Therefore, $MI_{X|Y}$ and $MI_{Y|X}$ can be calculated for every day in the trading period by using estimated copula C.

Step 4. Construct short position in stock X and long position in stock Y on the day that $MI_{X|Y} > \Delta_1$ and $MI_{Y|X} < \Delta_2$ if no positions in X or Y is held. Construct long position in stock X and short position in stock Y on the day that $MI_{X|Y} < \Delta_2$ and $MI_{Y|X} > \Delta_1$ if no positions in X and Y is held. The market value of short positions should always equal to the market value of long positions at the start. All positions are closed if $MI_{X|Y}$ reaches Δ_3 or $MI_{Y|X}$ reaches Δ_4, where $\Delta_1, \Delta_2, \Delta_3, \Delta_4$ all pre-determined threshold values are.

3. Generalization of traditional methods

It has been explained that the traditional methods use the idea of spread, and are more robust when linear correlation captures the whole dependency structure of two stocks. In this section, we show that the traditional methods can be generalized into special cases of copula method under certain assumptions about the marginal distributions and the joint relation.

3.1 A special Case: Distance Method

Under assumptions of normal marginal distributions of the stock prices and normal copula of the joint relation between the stock returns, it can be shown that the distance method will generate the same result as the copula-based trading strategy.

The distance method uses standardized prices in the calculation. Denote SP_t^X and SP_t^Y as the standardized prices of stock X and stock Y. Then they follow standard normal distributions with mean 0 and variance 1 if the marginal distributions of stock prices are assumed to be normal. Denote Δ_t^{XY} as the spread between them, and thus $\Delta_t^{XY} = SP_t^X - SP_t^Y$. Denote ρ as the linear correlation between the two stocks, then:

$$\begin{aligned} E(\Delta_t^{XY}) &= E(SP_t^X) - E(SP_t^Y) \\ &= 0 - 0 \\ &= 0 \\ Var(\Delta_t^{XY}) &= Var(SP_t^X - SP_t^Y) \\ &= Var(SP_t^X) + Var(SP_t^Y) - 2Cov(SP_t^X, SP_t^Y) \\ &= 2 - 2\rho. \end{aligned}$$

This shows that Δ_t^{XY} follows normal distribution with mean 0 and variance $2 - 2\rho$ under the distance method framework. It is stationary, and neither the mean nor the variance changes with the time t or the price levels SP_t^X and SP_t^Y.

Now, let us look at the copula method under assumptions of normal marginal distributions of the stock prices and normal copula linking them.

Theorem 5 Let X and Y be continuous random variables with copula C_{XY}. If functions α and β are strictly increasing on RanX and RanY, respectively, then $C_{\alpha(X),\beta(Y)} = C_{XY}$. Thus C_{XY} is invariant under strictly increasing transformations of X and Y.

Proof of this theorem can be found in Nelson (2006).

Since $R_t^X = \log(P_t^X) - \log(P_0^X)$ and $R_t^Y = \log(P_t^Y) - \log(P_0^Y)$, R_t^X and R_t^Y are all strictly increasing transformations of P_t^X and P_t^Y. Therefore, $C_{R_t^X, R_t^Y} = C_{P_t^X, P_t^Y}$.

It is assumed that the copula linking R_t^X and R_t^Y is normal, so that the copula linking P_t^X and P_t^Y is also normal. Combing with the information that P_t^X and P_t^Y have normal marginal distributions, it can be proved that P_t^X and P_t^Y are joint normally distributed with ρ being the linear correlation coefficient between X and Y.

Consequently, the mispricing index $MI_{X|Y}$ can be calculated in the following manner, which is related to the distribution of $SP_t^X - SP_t^Y$.

$$\begin{aligned} MI_{X|Y} &= P(R_t^X < r_t^X | R_t^Y = r_t^Y) \\ &= P(P_t^X < p_t^X | P_t^Y = p_t^Y) \\ &= P(SP_t^X < sp_t^X | SP_t^Y = sp_t^Y) \\ &= P(SP_t^X - SP_t^Y < sp_t^X - sp_t^Y | SP_t^Y = sp_t^Y) \end{aligned}$$

Theorem 6. If random variables X and Y follow standard normal distributions, and are jointly distributed with linear correlation coefficient ρ, then X|Y=y follows normal distribution with mean ρy and variance $1 - \rho^2$.

Proof. Assuming $f(x|y)$ is the density function of the random variable X|Y=y. A~B means A is linearly proportional to B.

$$f(x|y) = \frac{f(x,y)}{f(y)}$$
$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}}\exp(-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)})$$
$$f(x|y) \sim f(x,y) \sim \exp(-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}),$$
$$f(x|y) \sim \exp(-\frac{(x-\rho y)^2}{2(1-\rho^2)})$$

Thus, X|Y=y follows normal distribution with mean ρy and variance $1 - \rho^2$. #

Since P_t^X and P_t^Y are jointly normally distributed, SP_t^X and SP_t^Y should also be jointly normally distributed. Then the conditional distribution of SP_t^X given $SP_t^Y = sp_t^Y$ should follow normal distribution with mean ρsp_t^Y and variance $1 - \rho^2$. Then

$$E(SP_t^X - SP_t^Y | SP_t^Y = sp_t^Y) = \rho sp_t^Y - sp_t^Y = (\rho - 1) sp_t^Y$$
$$Var(SP_t^X - SP_t^Y | SP_t^Y = sp_t^Y) = 1 - \rho^2.$$

This shows that $SP_t^X - SP_t^Y | SP_t^Y = sp_t^Y$ follows normal distribution with mean $(\rho-1)sp_t^Y$ and variance $1-\rho^2$. In other words, the spread between two standardized prices Δ_t^{XY} conditional on $SP_t^Y = sp_t^Y$ is assumed to follow normal distribution with mean $(\rho-1)sp_t^Y$ and variance $1-\rho^2$.

For pairs trading, the linear correlation coefficient ρ should be close to 1, which means $E(\Delta_t^{XY}|SP_t^Y = sp_t^Y)$ approaches 0. Moreover, since the variance of $\Delta_t^{XY}|SP_t^Y = sp_t^Y$ is $1-\rho^2$, which is independent of the price levels, it can be concluded that the distribution of the spread is robust across different price levels. Therefore, distance method can capture the optimal trading opportunities the same as copula method under assumptions of normal marginal distributions and normal copula. However, if these assumptions do not hold for the real data, distance method may not be optimal.

Two real data examples are used to illustrate the above analysis. The first one shown in Figure 2 uses the pair HES (Hess Corporation) and MUR (Murphy Oil Corporation). It is one of the stock pairs studied in Chiu et al. (2011), and is listed on PairsLog.com as a basic materials/oil pair. Daily stock prices from January 2010 to December 2011 are downloaded from Yahoo Finance. The first year data is used as the formation period and the second year data is used as the trading period. The movement of the stock prices is shown in Figure 2(a), and we can see that they have a strong co-movement. The standardized prices and the spread between them are shown in Figure 2(b), the two bounds represent the $\pm 1.65\sigma_{HES-MUR}$ levels corresponding to 95%/5% probability level in the copula method shown in Figure 2(c) while $\sigma_{HES-MUR}$ is the variance of the spread during formation period. In Figure 2(c), normal distribution is chosen to estimate the marginal distributions of stock prices and normal copula is chosen to best fit the joint relation between stocks. By doing so, the result calculated from copula method matches exactly the result from distance method. The line in Figure 2(b) has the same shape as the line in Figure 2(c).

Another example uses the pair BKD (Brookdale Senior Living) and ESC (Emeritus), which is also listed in PairsLog.com. The time period is from December 2009 to November 2012. The first 2 years data are used to form the formation period while the rest are used to form the trading period. Price co-movement is shown in Figure 3(a). The results calculated from distance method and copula method are shown in Figure 3(b) and (c) respectively. Here, the marginal distributions are best fitted by generalized extreme value distributions, and the estimated copula is gumbel copula. The bounds here have the same meaning as the previous sample.

Different from the previous sample, it is found that the two methods give different results. It is found that the copula method captures more trading opportunities that are overlooked by the distance method. This is consistent with the analysis above. Distance method only produces same result as the copula method when the data structure is indeed estimated by normal margins and normal copula. When the data structure is best fitted by other marginal distributions and copulas, the copula method gives more trading opportunities.

From the above analysis, it can be seen that the distance method is a special case of the proposed copula-based trading strategy under certain assumptions. If the real data structure is indeed captured by normal marginal distributions and normal copula, the result generated by distance method is the same as the copula method. However, in the real data structure, the normal assumptions are not always valid. Copula method can capture the structure of dependence rather a single measurement of spread, and is more robust and consistent.

3.2 **A special case: co-integration Method**

As mentioned earlier, the co-integration method assumes that same spread across different return levels has the same degree of mispricing. It is argued that the detailed dependency structure of error terms and expected return levels are complex and cannot be generalized to one fixed distribution across all price levels. Thus, the co-integration method may not be optimal under some real data structure.

From mathematical point of view, if the data structure has the intrinsic property

$$P(R^X < r^X | R^Y = r^Y) = P\left(R^X < r^{X'} \middle| R^Y = r^{Y'}\right)$$

$$\text{if } r^X - r^Y = r^{X'} - r^{Y'} \text{ for any } r^X, r^Y, r^{X'}, r^{Y'},$$

in other words,

$$MI_{X|Y} = MI_{X|Y}{}' \text{ if } r^X - r^Y = r^{X'} - r^{Y'}, \forall\, r^X, r^Y, r^{X'}, r^{Y'},$$

then the co-integration method will generate the same trading opportunity as the copula method if the threshold values are chosen carefully. It is hard to tell the assumptions of the marginal distributions and joint relation exactly for co-integration method, however, it can be seen that the co-integration method do put on more restriction in the data structure compared to the copula method. This kind of restriction sometimes overlooks the non-linear correlation between stock prices, and thus misses some trading opportunities.

It is difficult to give examples to show that co-integration method produces the same result as the copula method under some assumptions like the first example in the previous section. The reason is that the restriction described by co-integration method is hard to be satisfied using several simple assumptions about the margins and copula. However, an example on the difference of copula method and co-integration method can still be given. BKD-ESC pair is used to illustrate the idea here again. The same time period as the previous section is used. The results calculated from the co-integration method and the copula method are shown in Figure 4(b) and (c). It is obvious that copula method captures more trading opportunities. The reason for this is that the data structure has strong tail dependence, and it shows different patterns of correlation at different price levels. BKD tends to be largely higher than ESC at high price levels while it is not that obvious at low price levels. However, co-integration method does not capture this characteristic, and thus overlooks potential trading opportunities.

4. Conclusions

In this research work, an algorithm for pairs trading using copula technique is proposed. On the one hand, traditional methods such as distance and co-integration methods use the idea of spread in their trading strategy. When the two stocks follow joint normal distributions, the linear correlation can capture the whole dependency information and the distribution of spread is robust under different price levels. On the other hand, copula method captures the dependency structure of the two stocks which is more robust and accurate. Under certain assumptions, traditional methods can be generalized as special cases of copula method.

Overall, we provide an alternative way to look at pairs trading. The future research could focus on testing the profitability of the copula method by implementing the real data and incorporate this method in pairs selection process.

References

Chiu, J., Lukman, D. W., Modarresi, K., and Velayutham, A. S. (2011) High-frequency Trading. Standford, California, U.S.: Stanford University.

Crook, J., and Moreia, F. (2011) Checking for asymmetric default dependence in a credit card portfolio: A copula approach. *Journal of Empirical Finance* 18(4): 728-742.

Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error-correction: Representation, estimation and testing. *Econometrica* 55 (2): 251–276.

Do, B., and Faff, R. (2010) Does Simple Pairs Trading Still Work? *Financial Analysts Journal* 66(4): 83-95.

Galenko, A., Popova, E., and Popova, I. (2012) Trading in the Presence of Co-integration. *The Journal of Alternative Investments* 15(1): 85-97.

Gatev, E., Goetzmann, W. N., and Rouwenhorst, K. G. (2006) Pairs Trading: Performance of a Relative-Value Arbitrage Rule. *The Review of Financial Studies* 19(3):797-827.

Kat, H. M. (2003) The Dangers of Using Correlation to Measure Dependence. *The Journal of Alternative Investments* 6(2): 54-58.

Liew, R. Q. and Wu, Y. (2013). Pairs trading: A copula approach. *Journal of Derivatives and Hedge Funds*, Vol 19(1), pp12-30.

Lin, Y.-X., MaCrae, M., and Gulati, C. (2006) Loss Protection in Pairs Trading Through Minimum Profit Bounds: A Co-integration Approach. Journal of Applied Mathematics and Decision Sciences 2006: 1-14.

Nelson, R. B. (2006) *An Introduction to Copulas* (2nd Ed.). New York: Springer.

Perlin, M. S. (2009) Evaluation of Pairs-Trading Strategy at the Brazilian Financial Market. *Journal of Derivatives& Hedge Funds* 15(2): 122-136.

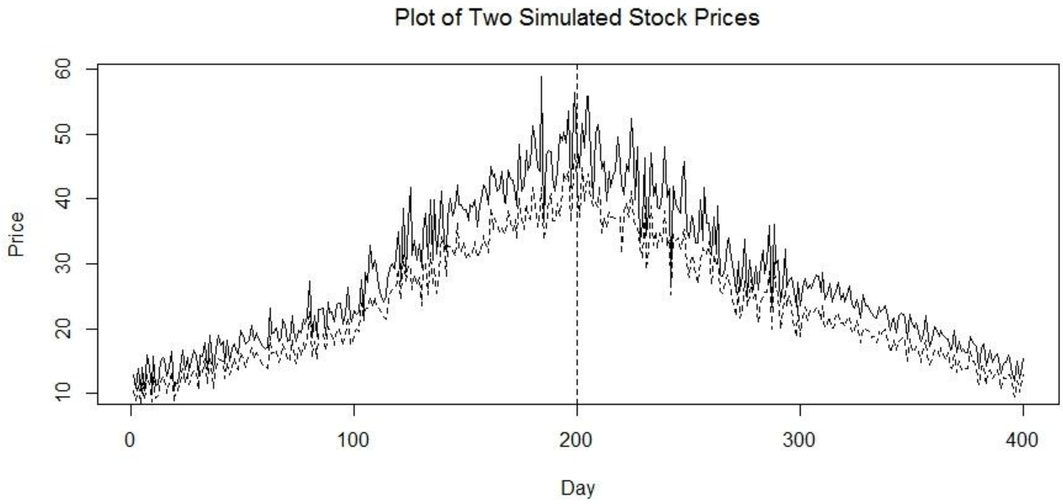
Simone, B. and Alessandro, G. (2010) A Pairs Trading Strategy Applied to the European Banking Sector. *Lecture "Asset Allocation and Performance Measurement"*, University of Zurich, spring 2010.

Sklar, A. (1959) Functions de répartition a n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Universite de Paris* 8:229-231.

Figure 1

Figure 1(a) shows the prices movement of two simulated stocks. (b) shows the spread between their returns. The solid line refers to stock X and the dotted line refers to stock Y. The simulation is designed to satisfy that the variance of the spread is larger when the stock prices are higher and lower when the stock prices are lower. The vertical line divides the formation period and the trading period.

a)



b)

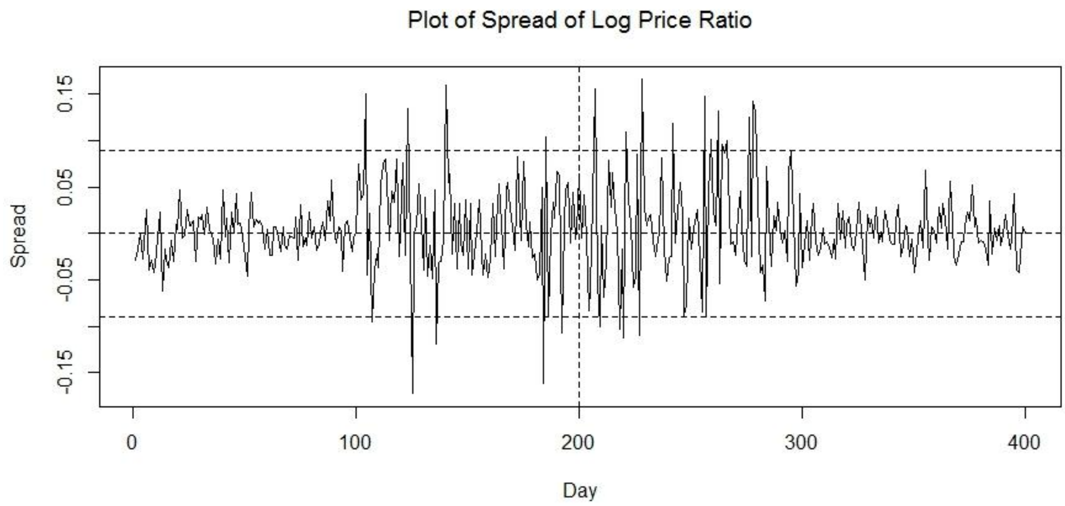
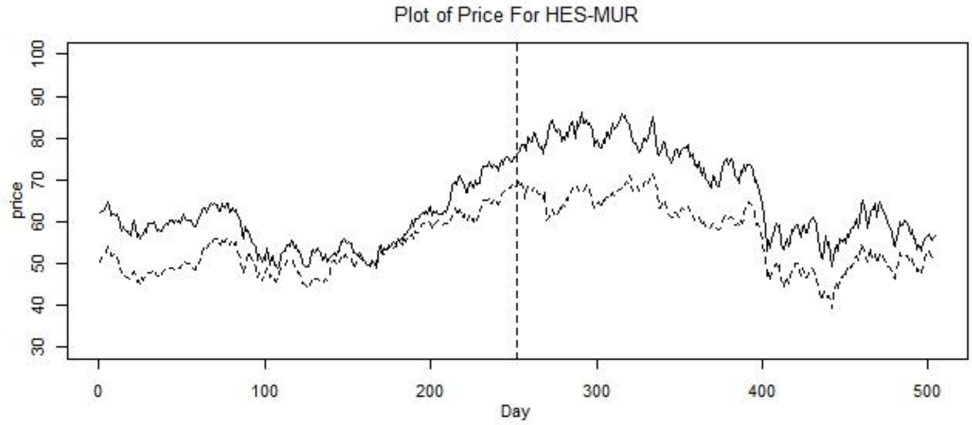


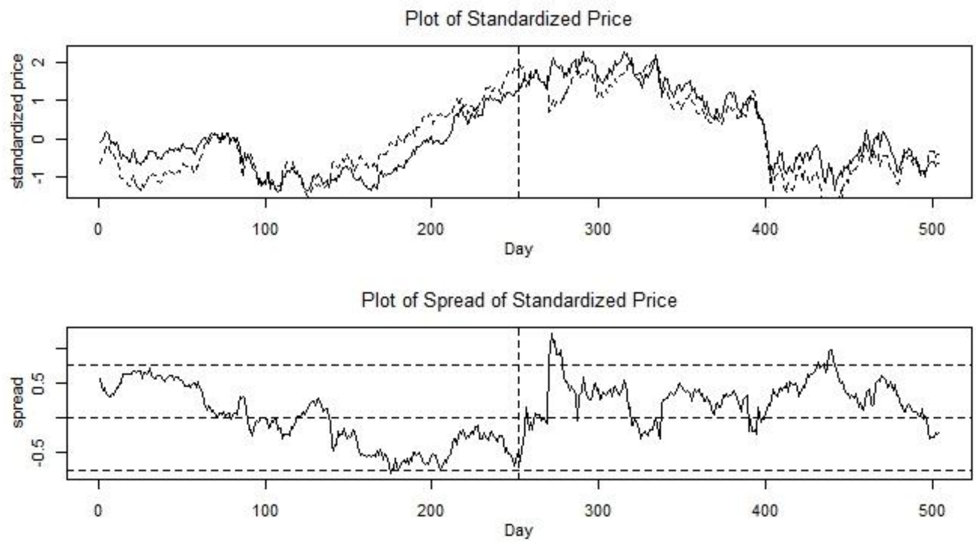
Figure 2

Figure 2(a) shows the prices movement of the pair HES-MUR. The solid line refers to HES and the dotted line refers to MUR. (b) shows the result estimated by distance method, the upper one shows the standardized prices and the lower one shows the spread between them. The bounds here refer to $\pm 1.65\sigma_{HES-MUR}$ levels where $\sigma_{HES-MUR}$ is the variance of the spread during formation period. (c) shows the result estimated by copula method, the solid line shows the $MI_{HES|MUR}$ and the dotted line shows the $MI_{MUR|HES}$.The vertical line divides the formation period and the trading period.

a)



b)



c)

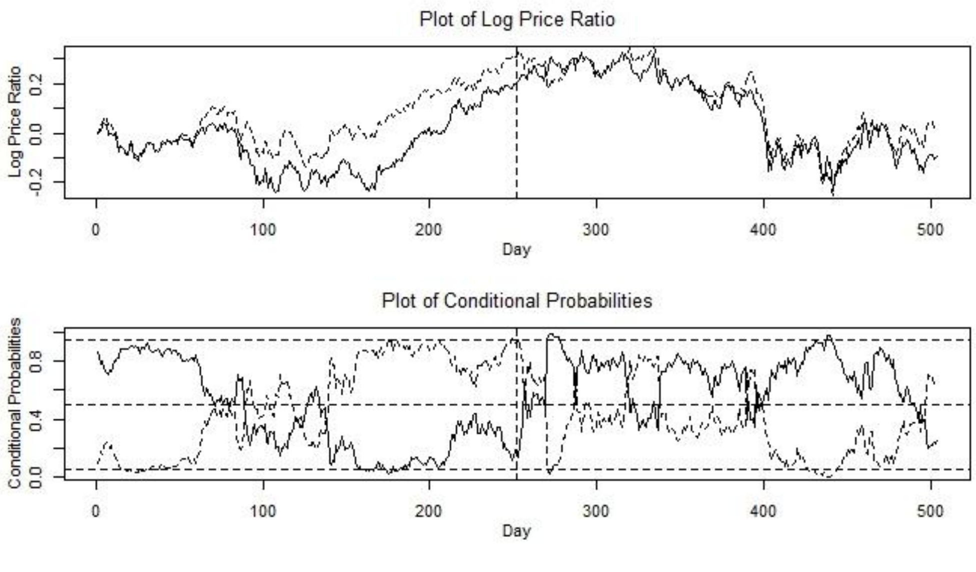
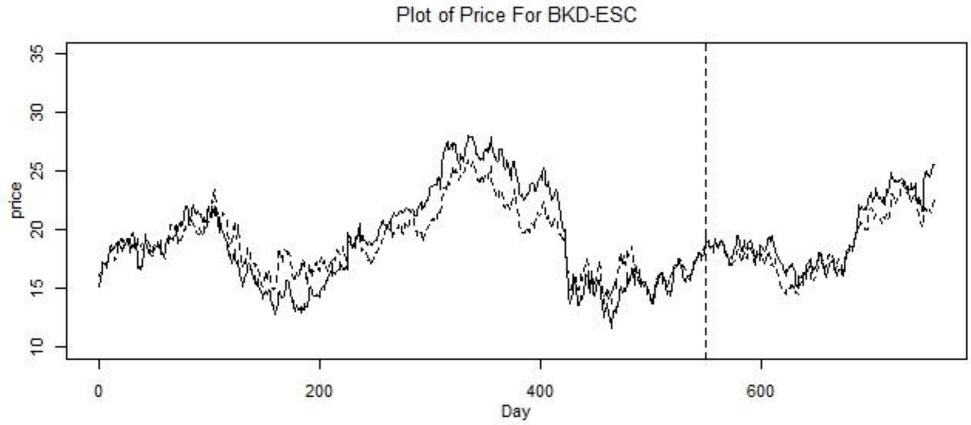


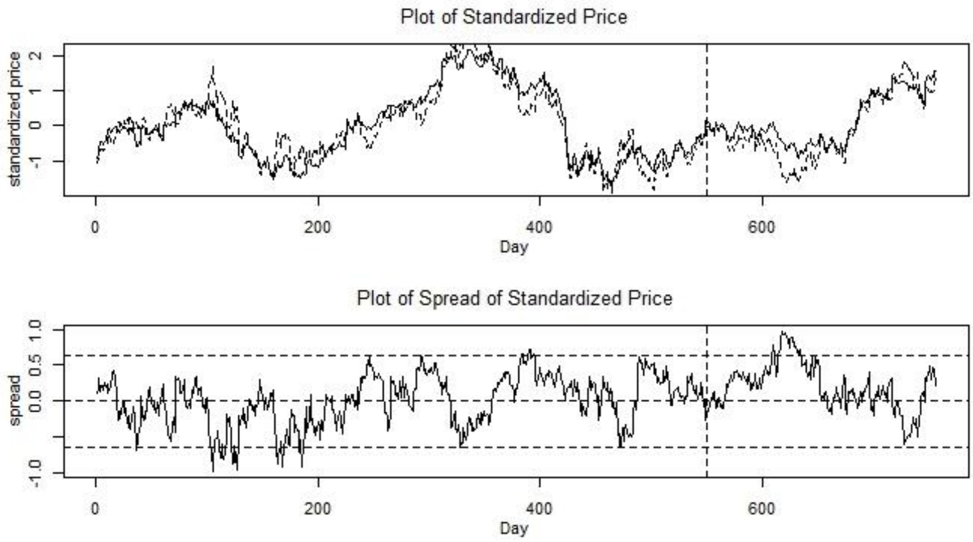
Figure 3

Figure 3(a) shows the prices movement of the pair BKD-ESC. The solid line refers to BKD and the dotted line refers to ESC. (b) shows the result estimated by distance method, the upper one shows the standardized prices and the lower one shows the spread between them. The bounds here refer to $\pm 1.65\sigma_{BKD-ESC}$ levels where $\sigma_{BKD-ESC}$ is the variance of the spread during formation period. (c) shows the result estimated by copula method, the solid line shows the $MI_{BKD|ESC}$ and the dotted line shows the $MI_{BKD|ESC}$.The vertical line divides the formation period and the trading period.

a)



b)



c)

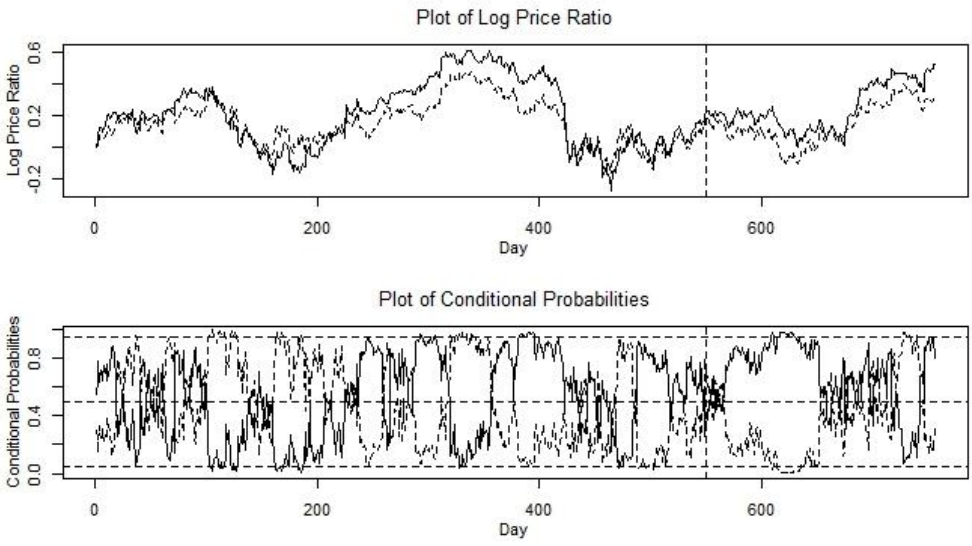
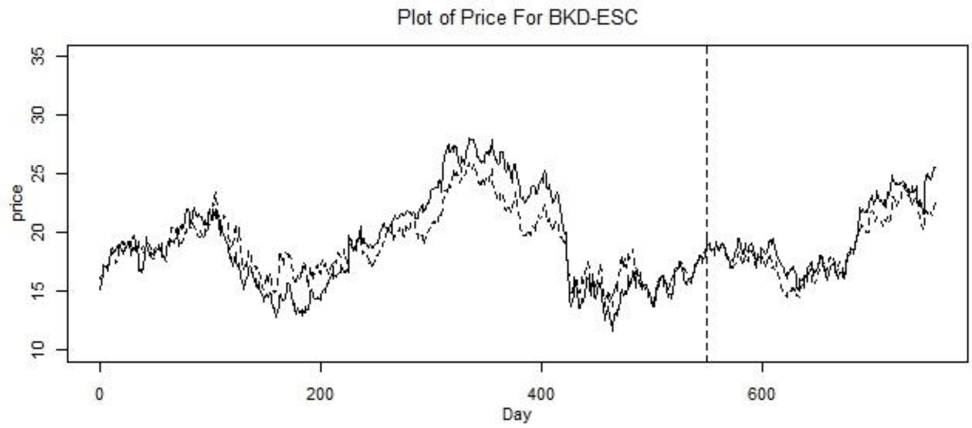


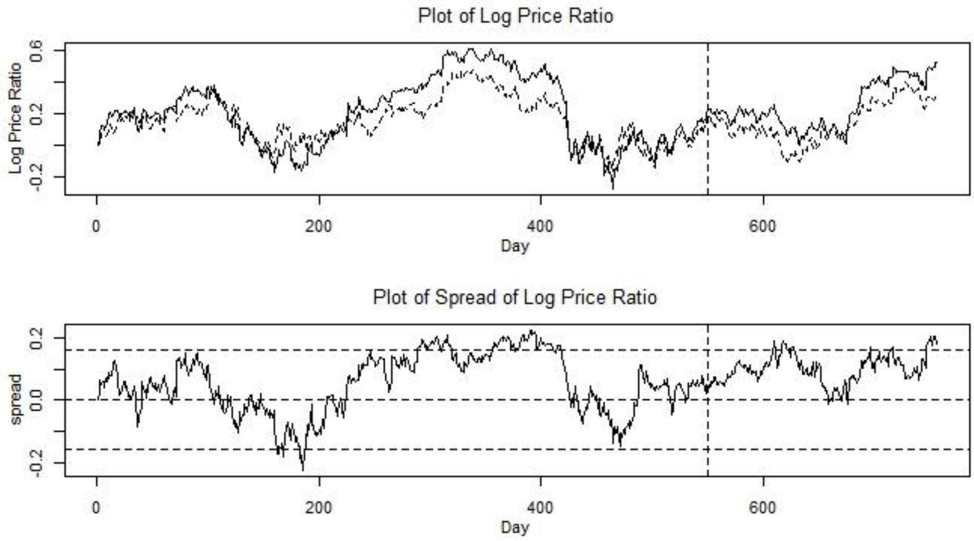
Figure 4

Figure 4(a) shows the prices movement of the pair BKD-ESC. The solid line refers to BKD and the dotted line refers to ESC. (b) shows the result estimated by co-integration method, the upper one shows the log price returns and the lower one shows the spread between them. The bounds here refer to $\pm 1.65\sigma_{BKD-ESC}$ levels where $\sigma_{BKD-ESC}$ is the variance of the spread during formation period. (c) shows the result estimated by copula method, the solid line shows the $MI_{BKD|ESC}$ and the dotted line shows the $MI_{BKD|ESC}$.The vertical line divides the formation period and the trading period.

a)



b)



c)