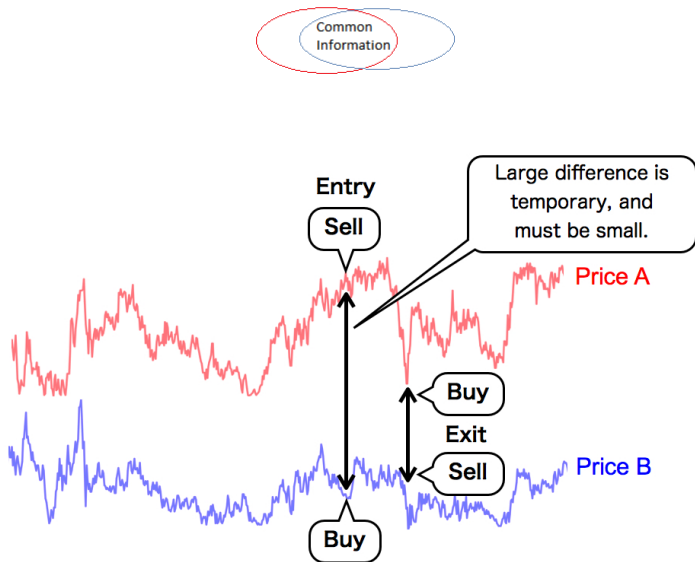


Pairs Trading: Optimizing via Mixed Copula versus Distance Method for S&P 500 Assets

Fernando Sabino da Silva¹, Flavio A. Ziegelmann^{1,2}, Joao F. Caldeira²

¹Department of Statistics, ²Graduate Program in Economics, Federal University of Rio Grande do Sul

Pairs Trading



Main Idea

- Pairs Trading is one type of Statistical “Arbitrage”.
- Identify a pair of securities whose prices tend to move together.
 - When they diverge
 - Buy the “cheap”
 - Sell the “expensive”
 - Reverse your positions when the two prices converge \Rightarrow Reversal profits.

Main Idea

- Pairs Trading is one type of Statistical “Arbitrage”.
- Identify a pair of securities whose prices tend to move together.
- When they diverge
 - ▶ short the “winner”
 - ▶ buy the “loser”.
- Reverse your positions when the two prices converge \Rightarrow Reversal profits.

Main Idea

- Pairs Trading is one type of Statistical “Arbitrage”.
- Identify a pair of securities whose prices tend to move together.
- When they diverge
 - ▶ short the “winner”
 - ▶ buy the “loser”.
- Reverse your positions when the two prices converge \Rightarrow Reversal profits.

Background

- Developed in the mid 1980's by Nunzio Tartaglia and his group at Morgan Stanley.
 - ▶ They report that the black box strategy made over 50 million profit for the firm in 1987.
- Who does it? Hedge funds, Proprietary trading desks.

Economic Rationale

- Tartaglia:
 - ▶ "Human beings don't like to trade against human nature, which wants to buy stocks after they go up, not down".
- Over-reaction?
 - ▶ Sentiment-based explanation: Evidence that market prices may reflect investor overreaction to information, or fads, or simply cognitive errors (Da et al., 2013).
 - ▶ Contrarian profits are in part due to overreaction to firm-specific factors (Jegadeesh and Titman's, 1995).
- Liquidity-based explanation
 - ▶ Campbell et al., 1993 ⇒ Uninformed traders lead to a temporary price concession that, when absorbed by liquidity providers ⇒ reversal in price that serves as compensation for those who provide liquidity.
 - ▶ Avramov et al., 2006 ⇒ Reversal profits mainly derive from positions in small, high turnover, and illiquid stocks.

Relative Pricing

- Pairs Trading does not seek to determine the absolute price of any stock.
- Long-short "arbitrage in expectations".
- Market-neutral.
- Self-financing.

Relative Pricing

- Pairs Trading does not seek to determine the absolute price of any stock.
- Long-short "arbitrage in expectations".
- Market-neutral.
- Self-financing.

Distance Method

- Distance method ([Gatev et al., 2006](#)).
 - ▶ Evidence that a simple strategy produced statistically significant excess returns for the period 1962-2002 in the US market.
 - ▶ Matching partner (12-month): Minimize the sum of squared deviations (distance) between normalized daily prices \Rightarrow Capture the degree of mispricing stocks.
 - ★ Trading period (6-month): A trade is initiated when the distance exceeds 2σ and exits when the distance is 0, or at the end of six-month.
 - ▶ Equivalent to matching on state-prices.
 - ★ Each day is a different state.
 - ★ Assumes stationarity.
 - ★ Assumes a year capture all states.
- [Xie et al. \(2014\)](#): "distance method has a multivariate normal nature..."

Distance Method

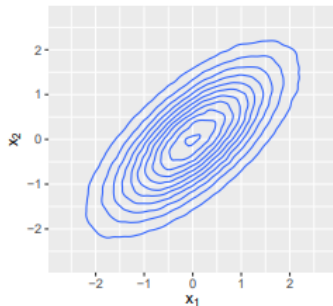
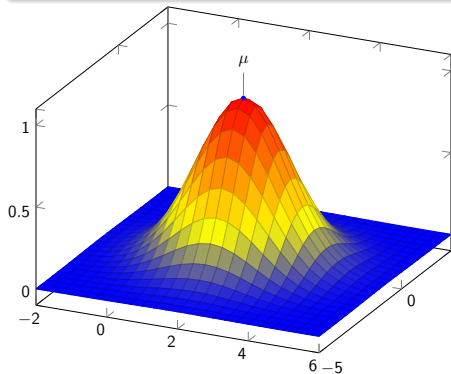
- Distance method ([Gatev et al., 2006](#)).
 - ▶ Evidence that a simple strategy produced statistically significant excess returns for the period 1962-2002 in the US market.
 - ▶ Matching partner (12-month): Minimize the sum of squared deviations (distance) between normalized daily prices \Rightarrow Capture the degree of mispricing stocks.
 - ★ Trading period (6-month): A trade is initiated when the distance exceeds 2σ and exits when the distance is 0, or at the end of six-month.
 - ▶ Equivalent to matching on state-prices.
 - ★ Each day is a different state.
 - ★ Assumes stationarity.
 - ★ Assumes a year capture all states.
- [Xie et al. \(2014\)](#): "distance method has a multivariate normal nature..."

Distance Method

- Distance method ([Gatev et al., 2006](#)).
 - ▶ Evidence that a simple strategy produced statistically significant excess returns for the period 1962-2002 in the US market.
 - ▶ Matching partner (12-month): Minimize the sum of squared deviations (distance) between normalized daily prices \Rightarrow Capture the degree of mispricing stocks.
 - ★ Trading period (6-month): A trade is initiated when the distance exceeds 2σ and exits when the distance is 0, or at the end of six-month.
 - ▶ Equivalent to matching on state-prices.
 - ★ Each day is a different state.
 - ★ Assumes stationarity.
 - ★ Assumes a year capture all states.
- [Xie et al. \(2014\)](#): "distance method has a multivariate normal nature..."

Bivariate Normal Distribution

$$f(x, y) = \frac{\exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_x}{\sigma_x} \right)^2 - 2\rho \left(\frac{x-\mu_x}{\sigma_x} \right) \left(\frac{y-\mu_y}{\sigma_y} \right) + \left(\frac{y-\mu_y}{\sigma_y} \right)^2 \right] \right\}}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}$$



Motivation

- Joint normal distribution \Rightarrow Linear correlation fully describes the dependence.
- Tail dependence
 - ▶ Heavy tails
 - ▶ Possibly Asymmetric
- A single distance measure may fail to catch the dynamics of the spread between a pair of securities.
 - ▶ Volatility differs at different price levels \Rightarrow inappropriate to use constant trigger points.
 - ▶ We may initiate and close the trades at non-optimal positions.
- Lie and Wu (2013): pairs trading strategy based on 2-dimensional copulas .

Motivation

- Joint normal distribution \Rightarrow Linear correlation fully describes the dependence.
- Tail dependence
 - ▶ Heavy tails
 - ▶ Possibly Asymmetric
- A single distance measure may fail to catch the dynamics of the spread between a pair of securities.
 - ▶ Volatility differs at different price levels \Rightarrow inappropriate to use constant trigger points.
 - ▶ We may initiate and close the trades at non-optimal positions.
- Lie and Wu (2013): pairs trading strategy based on 2-dimensional copulas .

Motivation

- Joint normal distribution \Rightarrow Linear correlation fully describes the dependence.
- Tail dependence
 - ▶ Heavy tails
 - ▶ Possibly Asymmetric
- A single distance measure may fail to catch the dynamics of the spread between a pair of securities.
 - ▶ Volatility differs at different price levels \Rightarrow inappropriate to use constant trigger points.
 - ▶ We may initiate and close the trades at non-optimal positions.
- Lie and Wu (2013): pairs trading strategy based on 2-dimensional copulas .

Motivation

- Joint normal distribution \Rightarrow Linear correlation fully describes the dependence.
- Tail dependence
 - ▶ Heavy tails
 - ▶ Possibly Asymmetric
- A single distance measure may fail to catch the dynamics of the spread between a pair of securities.
 - ▶ Volatility differs at different price levels \Rightarrow inappropriate to use constant trigger points.
 - ▶ We may initiate and close the trades at non-optimal positions.
- Lie and Wu (2013): pairs trading strategy based on 2-dimensional copulas .

Sklar's Theorem (1959)

Theorem 1

Let X_1, \dots, X_d be random variables with distribution functions F_1, \dots, F_d , respectively. Then, there exists an d -copula C such that,

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)), \quad (1)$$

for all $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$. If F_1, \dots, F_d are all continuous, then the function C is unique; otherwise C is determined only on $\text{Im } F_1 \times \dots \times \text{Im } F_d$.

Why should we care about copulas?

- Assuming that $F(\cdot)$ and $C(\cdot)$ are differentiable, by (1) we have

$$\frac{\partial^d F(x_1, \dots, x_d)}{\partial x_1 \dots \partial x_d} \equiv f(x_1, \dots, x_d) = \frac{\partial^d C(F_1(x_1), \dots, F_d(x_d))}{\partial x_1 \dots \partial x_d} \quad (2)$$

$$= c(u_1, \dots, u_d) \prod_{i=1}^d f_i(x_i), \quad (3)$$

where $u_i = F_i(x_i)$, $i = 1, \dots, d$.

multivariate = "1-dim" (marginals) + "joint" (copula).

Why should we care about copulas?

- Assuming that $F(\cdot)$ and $C(\cdot)$ are differentiable, by (1) we have

$$\frac{\partial^d F(x_1, \dots, x_d)}{\partial x_1 \dots \partial x_d} \equiv f(x_1, \dots, x_d) = \frac{\partial^d C(F_1(x_1), \dots, F_d(x_d))}{\partial x_1 \dots \partial x_d} \quad (2)$$

$$= c(u_1, \dots, u_d) \prod_{i=1}^d f_i(x_i), \quad (3)$$

where $u_i = F_i(x_i)$, $i = 1, \dots, d$.

multivariate = "1-dim" (marginals) + "joint" (copula).

Copula

| Strategy | Associations Captured | Required Marginal Distributions |
|-----------------|--------------------------------|---------------------------------|
| Distance Copula | Linear Linear and Nonlinear | Gaussian No assumption |

In practical terms, the copula provides an effective tool to monitor and hedge the risks in the markets.

Xie *et al.* (2014) define a measure to denote the degree of mispricing.

Definition 2

- Let R_t^X and R_t^Y represent the random variables of the daily returns of stocks X and Y on time t , and the realizations of those returns on time t are r_t^X and r_t^Y , we have

$$MI_{X|Y}^t = P(R_t^X < r_t^X \mid R_t^Y = r_t^Y)$$

and

$$MI_{Y|X}^t = P(R_t^Y < r_t^Y \mid R_t^X = r_t^X),$$

where $MI_{X|Y}$ and $MI_{Y|X}$ are named the mispricing indexes.

Copula Method

- Partial derivative of the copula function gives the conditional distribution function

$$MI_{X|Y}^t = \frac{\partial C(u_1, u_2)}{\partial u_2} = P(R_t^X < r_t^X \mid R_t^Y = r_t^Y)$$

and

$$MI_{Y|X}^t = \frac{\partial C(u_1, u_2)}{\partial u_1} = P(R_t^X < r_t^X \mid R_t^Y = r_t^Y). \quad (4)$$

$$MI_{X|Y}^t = \frac{\partial C(u_1, u_2)}{\partial u_2} = P(R_t^X < r_t^X \mid R_t^Y = r_t^Y)$$

and

$$MI_{Y|X}^t = \frac{\partial C(u_1, u_2)}{\partial u_1} = P(R_t^X < r_t^X \mid R_t^Y = r_t^Y). \quad (4)$$

- A value of 0.5 \Rightarrow 50% chance for the price of stock 1 to be below its current realization given the current price of stock 2.

Copula Method

- $M_t^{X|Y}$ and $M_t^{Y|X} \Rightarrow$ measure the degrees of relative mispricing for a single day.

- Overall degree of relative mispricing (Rad *et al.* (2016)).

- Mispricing indexes of stocks

$$m_{1,t} = \left(M_t^{X|Y} - 0.5 \right)$$

$$m_{2,t} = \left(M_t^{Y|X} - 0.5 \right)$$

- Cumulative mispricing indexes

$$M_{1,t} = M_{1,t-1} + m_{1,t}$$

$$M_{2,t} = M_{2,t-1} + m_{2,t}$$

- * Positive M_1 and negative $M_2 \Rightarrow$ Stock 1 is overvalued relative to stock 2

- * Note: M_1 and M_2 are set to zero at the beginning of the trading period

Copula Method

- $M_t^{X|Y}$ and $M_t^{Y|X} \Rightarrow$ measure the degrees of relative mispricing for a single day.
- Overall degree of relative mispricing ([Rad et al. \(2016\)](#)).
 - ▶ Mispricing indexes of stocks

$$m_{1,t} = \left(M_t^{X|Y} - 0.5 \right)$$
$$m_{2,t} = \left(M_t^{Y|X} - 0.5 \right)$$

- ▶ Cumulative mispricing indexes

$$M_{1,t} = M_{1,t-1} + m_{1,t}$$
$$M_{2,t} = M_{2,t-1} + m_{2,t}$$

- ★ Positive M_1 and negative $M_2 \Rightarrow$ Stock 1 is overvalued relative to stock 2
- ★ Note: M_1 and M_2 are set to zero at the beginning of the trading period

- Sensitivity analysis: open a long-short position once one of the cumulative indexes is above 0.05, 0.10, ..., 0.55 and the other one is below -0.05, -0.10, ..., -0.55 at the same time.
- How many pairs do we use?
 - ▶ 5, 10, 15, 20, 25, 30 and 35.
- The positions are unwound when both cumulative mispriced indexes return to zero.

- Sensitivity analysis: open a long-short position once one of the cumulative indexes is above 0.05, 0.10, ..., 0.55 and the other one is below -0.05, -0.10, ..., -0.55 at the same time.
- How many pairs do we use?
 - ▶ 5, 10, 15, 20, 25, 30 and 35.
- The positions are unwound when both cumulative mispriced indexes return to zero.

Pairs Implementation: Copula

① Estimate the marginal distributions of returns.

- ▶ ARMA(p,q)-GARCH(1,1).

② Estimate the two-dimensional copula model to data that has been transformed to $[0,1]$ margins, i.e.,

$$H(r_t^X, r_t^Y) = C(F_X(r_t^X), F_Y(r_t^Y)),$$

where H is the joint distribution, r_t^X e r_t^Y are stock returns and C is the copula.

- ▶ Gaussian, t, Clayton, Frank, Gumbel.

Mixed copula models to cover a wider range of dependence structures are proposed.

- ★ Archimedean mixture copula consisting of the optimal linear combination of Clayton, Frank and Gumbel copulas.
- ★ Mixture copula consisting of the optimal linear combination of Clayton, t and Gumbel copulas.

Pairs Implementation: Copula

① Estimate the marginal distributions of returns.

- ▶ ARMA(p,q)-GARCH(1,1).

② Estimate the two-dimensional copula model to data that has been transformed to $[0,1]$ margins, i.e.,

$$H(r_t^X, r_t^Y) = C(F_X(r_t^X), F_Y(r_t^Y)),$$

where H is the joint distribution, r_t^X e r_t^Y are stock returns and C is the copula.

- ▶ Gaussian, t, Clayton, Frank, Gumbel.

Mixed copula models to cover a wider range of dependence structures are proposed.

- ★ Archimedean mixture copula consisting of the optimal linear combination of Clayton, Frank and Gumbel copulas.
- ★ Mixture copula consisting of the optimal linear combination of Clayton, t and Gumbel copulas.

Pairs Implementation: Copula

- ① Estimate the marginal distributions of returns.
 - ▶ ARMA(p,q)-GARCH(1,1).
- ② Estimate the two-dimensional copula model to data that has been transformed to $[0,1]$ margins, i.e.,

$$H(r_t^X, r_t^Y) = C(F_X(r_t^X), F_Y(r_t^Y)),$$

where H is the joint distribution, r_t^X e r_t^Y are stock returns and C is the copula.

- ▶ Gaussian, t, Clayton, Frank, Gumbel.

Mixed copula models to cover a wider range of dependence structures are proposed.

- ★ Archimedean mixture copula consisting of the optimal linear combination of Clayton, Frank and Gumbel copulas.
- ★ Mixture copula consisting of the optimal linear combination of Clayton, t and Gumbel copulas.

Pairs Implementation: Copula

① Estimate the marginal distributions of returns.

- ▶ ARMA(p,q)-GARCH(1,1).

② Estimate the two-dimensional copula model to data that has been transformed to $[0,1]$ margins, i.e.,

$$H(r_t^X, r_t^Y) = C(F_X(r_t^X), F_Y(r_t^Y)),$$

where H is the joint distribution, r_t^X e r_t^Y are stock returns and C is the copula.

- ▶ Gaussian, t, Clayton, Frank, Gumbel.

Mixed copula models to cover a wider range of dependence structures are proposed.

- ★ Archimedean mixture copula consisting of the optimal linear combination of Clayton, Frank and Gumbel copulas.
- ★ Mixture copula consisting of the optimal linear combination of Clayton, t and Gumbel copulas.

Mixed Copula

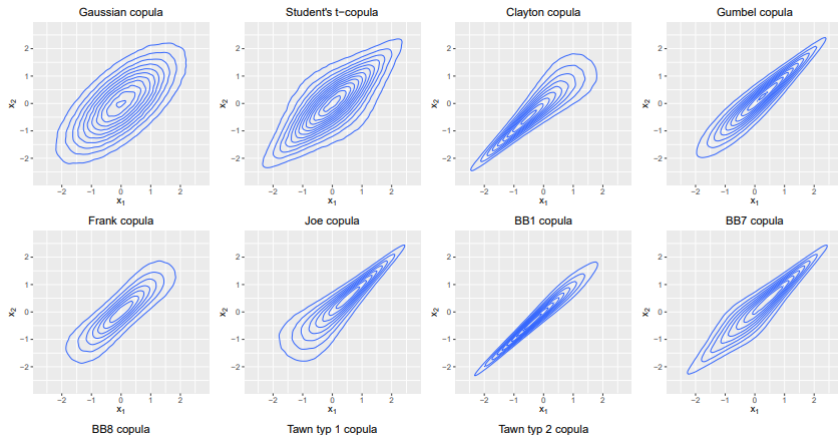
$$\mathcal{C}_{\theta}^{CFG}(u_1, u_2) = \pi_1 \mathcal{C}_{\alpha}^C(u_1, u_2) + \pi_2 \mathcal{C}_{\beta}^F(u_1, u_2) + (1 - \pi_1 - \pi_2) \mathcal{C}_{\delta}^G(u_1, u_2),$$

and

$$\mathcal{C}_{\xi}^{CtG}(u_1, u_2) = \pi_1 \mathcal{C}_{\alpha}^C(u_1, u_2) + \pi_2 \mathcal{C}_{\Sigma, \nu}^t(u_1, u_2) + (1 - \pi_1 - \pi_2) \mathcal{C}_{\delta}^G(u_1, u_2),$$

where $\theta = (\alpha, \beta, \delta)'$ are the Clayton, Frank and Gumbel copula (dependence) parameters and $\xi = (\alpha, (\Sigma, \nu), \delta)'$ are the Clayton, t and Gumbel copula parameters, respectively, and $\pi_1, \pi_2 \in [0, 1]$.

Tail Dependence



- **Sources:** Adjusted closing prices, Fama-French factors
 - ▶ Cumulative total return index for each stock.
- **Universe:** All shares that belongs to the S&P 500 market index.
- **Dates:** July 2nd, 1990 to December 31st, 2015.
- **Totals:** 1100 stocks during 6426 days.

Risk-Return characteristics

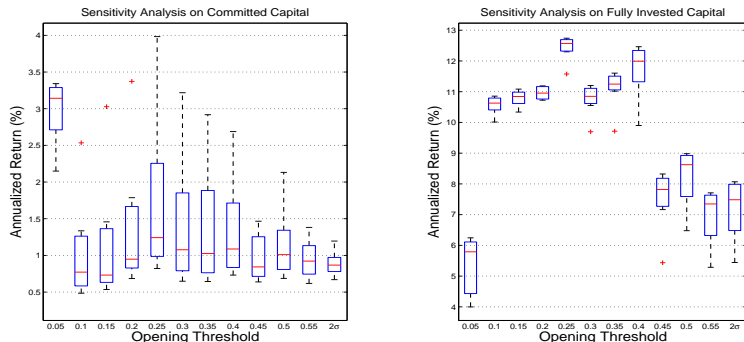


Figure 1: Annualized returns of pairs trading strategies after costs on committed and fully invested capital

These boxplots show annualized returns on committed (left) and fully invested (right) capital after transaction cost to different opening thresholds from July 1991 to December 2015 for Top 5 to Top 35 pairs.

Risk-Return characteristics

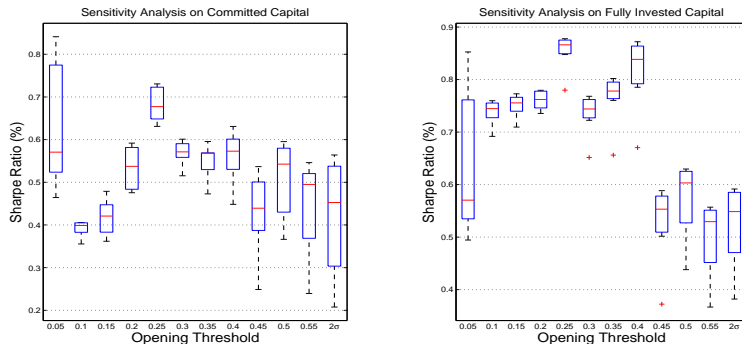


Figure 2: Sharpe ratio of pairs trading strategies after costs on committed and fully invested capital

Trading Statistics

Table 1: Trading statistics.

| Strategy | Distance | Mixed Copula |
|--|----------|--------------|
| <i>Panel A: Top 5</i> | | |
| Average price deviation trigger for opening | 0.0594 | 0.0665 |
| Total number of pairs opened | 352 | 348 |
| Average number of pairs traded per 6-month | 7.18 | 7.10 |
| Average number of round-trip trades per pair | 1.44 | 1.42 |
| Standard Deviation | 1.0128 | 1.33 |
| Average time pairs are open in days | 50.70 | 37.70 |
| Standard Deviation | 39.24 | 38.93 |
| Median time pairs are open in days | 38.5 | 19 |
| <i>Panel B: Top20</i> | | |
| Average price deviation trigger for opening | 0.0681 | 0.0821 |
| Total number of pairs opened | 1312 | 749 |
| Average number of pairs traded per 6-month | 26.78 | 15.29 |
| Average number of round-trip trades per pair | 1.34 | 0.76 |
| Standard Deviation | 0.99 | 0.99 |
| Average time pairs are open in days | 51.65 | 23.60 |
| Standard Deviation | 39.62 | 32.90 |
| Median time pairs are open in days | 41 | 9 |

Table 2: Trading statistics.

| Strategy | Distance | Mixed Copula |
|---|----------|--------------|
| <i>Panel C: Top 35</i> | | |
| Average price deviation trigger for opening pairs | 0.0729 | 0.0893 |
| Total number of pairs opened | 2238 | 941 |
| Average number of pairs traded per six-month period | 45.68 | 19.20 |
| Average number of round-trip trades per pair | 1.30 | 0.55 |
| Standard Deviation | 1.02 | 0.84 |
| Average time pairs are open in days | 52.72 | 19.35 |
| Standard Deviation | 40.48 | 30.56 |
| Median time pairs are open in days | 42 | 6 |

Note: Trading statistics for portfolio of top 5, 20 and 35 pairs between July 1991 and December 2015 (49 periods). Pairs are formed over a 12-month period according to a minimum-distance (sum of squared deviations) criterion and then traded over the subsequent 6-month period. Average price deviation trigger for opening a pair is calculated as the price difference divided by the average of the prices.

Trading Performance

Table 3: Excess returns on committed capital of pairs trading strategies on portfolios of Top 5, 20 and 35 pairs after costs.

| Strategy | Mean Return (%) | Sharpe ratio | Sortino ratio | t-stat | % of negative trades | MDD1 | MDD2 |
|------------------------------------|-----------------|--------------|---------------|---------|----------------------|-------|-------|
| Return on Committed Capital | | | | | | | |
| <i>Panel A - Top 5 pairs</i> | | | | | | | |
| Distance | 2.60 | 0.31 | 0.58 | 1.86* | 46.98 | 6.73 | 19.62 |
| Mixed Copula | 3.98 | 0.63 | 1.08 | 3.49*** | 41.79 | 4.36 | 9.29 |
| <i>Panel B - Top 20 pairs</i> | | | | | | | |
| Distance | 3.14 | 0.65 | 1.13 | 3.32*** | 48.02 | 3.88 | 9.69 |
| Mixed Copula | 1.24 | 0.64 | 1.04 | 3.52*** | 41.33 | 2.07 | 3.43 |
| <i>Panel C - Top 35 pairs</i> | | | | | | | |
| Distance | 3.12 | 0.77 | 1.36 | 3.92*** | 47.97 | 2.70 | 7.52 |
| Mixed Copula | 0.82 | 0.73 | 1.19 | 3.95*** | 41.31 | 1.18 | 1.98 |
| S&P 500 | 4.36 | 0.23 | 0.52 | 1.79* | 47.45 | 12.42 | 46.74 |

Note: Summary statistics of the annualized excess returns, annualized Sharpe and Sortino ratios on portfolios of top 5, 20 and 35 pairs between July 1991 and December 2015 (6,173 observations). The t-statistics are computed using Newey-West standard errors with a six-lag correction. The columns labeled MDD1 and MDD2 compute the largest drawdown in terms of maximum percentage drop between two consecutive days and between two days within a period of maximum six months, respectively.

***, **, * significant at 1%, 5% and 10% levels, respectively.

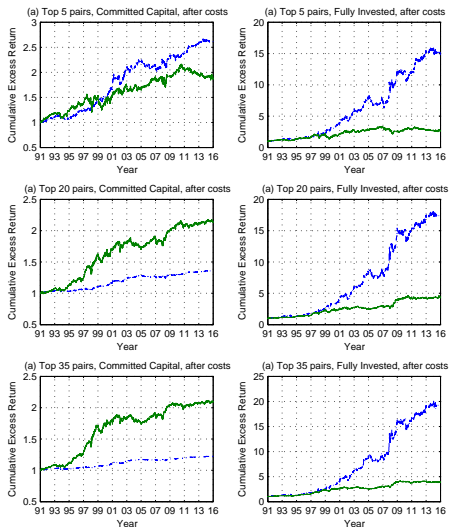
Trading Performance

Table 4: Excess returns on fully invested capital of pairs trading strategies on portfolios of Top 5, 20 and 35 pairs after costs.

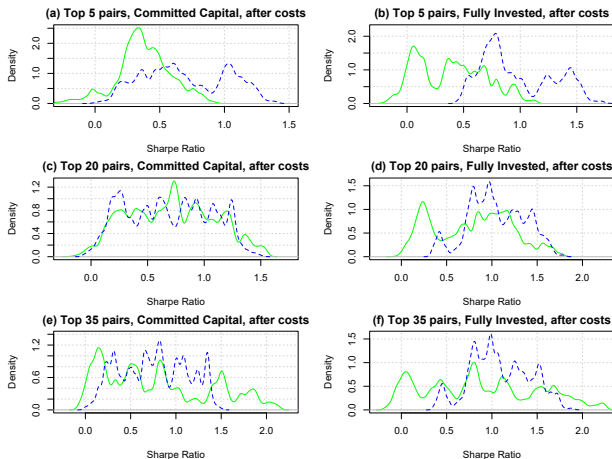
| Strategy | Mean Return (%) | Sharpe ratio | Sortino ratio | t-stat | % of negative trades | MDD1 | MDD2 |
|---|-----------------|--------------|---------------|---------|----------------------|------|-------|
| Return on Fully Invested Capital | | | | | | | |
| <i>Panel A - Top 5 pairs</i> | | | | | | | |
| Distance | 4.01 | 0.28 | 0.57 | 1.81* | 46.98 | 8.70 | 38.36 |
| Mixed Copula | 11.58 | 0.78 | 1.43 | 4.26*** | 41.79 | 9.00 | 25.68 |
| <i>Panel B - Top 20 pairs</i> | | | | | | | |
| Distance | 6.07 | 0.66 | 1.19 | 3.55*** | 48.06 | 5.43 | 20.03 |
| Mixed Copula | 12.30 | 0.85 | 1.54 | 4.60*** | 41.31 | 9.00 | 25.68 |
| <i>Panel C - Top 35 pairs</i> | | | | | | | |
| Distance | 5.76 | 0.76 | 1.38 | 4.05*** | 47.97 | 4.24 | 15.07 |
| Mixed Copula | 12.73 | 0.88 | 1.59 | 4.73*** | 41.28 | 9.00 | 25.68 |

***, **, * significant at 1%, 5% and 10% levels, respectively.

Cumulative excess returns of pairs trading strategies after costs



Kernel density estimation of 5-year rolling window Sharpe ratio after costs



--- Mixed Copula — Distance

Systematic Risk Exposure

Table 5: Monthly risk profile of Top 5 pairs: Fama and French (2016)'s five factors plus Momentum and Long-Term Reversal.

| Strategy | Intercept | Rm-Rf | SMB | HML | RMW | CMA | Mom | LRev | R ² | R ² _{adj} |
|---|---------------------|---------------------|---------------------|--------------------|--------------------|--------------------|-----------------------|----------------------|----------------|-------------------------------|
| Section 1: Return on Committed Capital | | | | | | | | | | |
| Distance | 0.0025 (1.89)* | 0.0091 (4.22)*** | -0.0032 (-0.71) | 0.0113 (2.05)** | 0.0003 (0.25) | -0.0029 (-0.18) | -0.0107 (-4.80)*** | -0.0084 (-1.96)** | 0.028 | 0.027 |
| Mixed Copula | 0.0035 (3.55)*** | 0.0052 (3.68)*** | -0.0043 (-1.83)* | 0.0039 (1.20) | -0.0035 (-0.99) | 0.0027 (0.63) | -0.0054 (-2.99)*** | -0.0057 (-1.57) | 0.015 | 0.014 |
| Section 2: Return on Fully Invested Capital | | | | | | | | | | |
| Distance | 0.0040 (1.75)* | 0.0170 (4.38)*** | -0.0031 (-0.45) | 0.0185 (2.22)** | 0.0049 (0.76) | -0.0018 (0.05) | -0.0161 (-4.30)*** | -0.0150 (-1.97)** | 0.025 | 0.024 |
| Mixed Copula | 0.0098 (4.17)*** | 0.0148 (3.51)*** | -0.0084 (-1.45) | 0.0152 (1.6355) | -0.0053 (-0.60) | 0.0087 (0.75) | -0.0082 (-2.19)** | -0.0222 (-2.08)** | 0.018 | 0.017 |

***, **, * significant at 1%, 5% and 10% levels, respectively.

- Alphas are significantly positive and higher than the raw excess returns by about 2-7 bps per month.
 - Only a small part of the excess returns can be attributed to their exposures to the seven risk determinants.

Conclusions

- 1 We examine two pairs trading strategies in a reasonably efficient market.
- 2 By capturing linear/nonlinear associations and covering a wider range of possible dependencies structures, the mixed copula strategy outperforms the distance method when the number of trading signals is equiparable, especially after the subprime mortgage crisis.
- 3 The mixed copula pairs trading strategy generates large and significant (at 1%) abnormal returns.
- 4 Only a small part of the pairs trading profits can be explained by market risk, portfolio (beta), size (SMB), value (HML), investment (CMA), profitability (RMW), momentum (MOM) and several D.R.-based factors.

Conclusions

- 1 We examine two pairs trading strategies in a reasonably efficient market.
- 2 By capturing linear/nonlinear associations and covering a wider range of possible dependencies structures, the mixed copula strategy outperforms the distance method when the number of trading signals is equiparable, especially after the subprime mortgage crisis.
- 3 The mixed copula pairs trading strategy generates large and significant (at 1%) abnormal returns.

4 Our research part of the pairs trading strategy can be explained by market portfolio (market), size (SIZE), value (VALUE), investment (INVT), profitability (PROF), momentum (MOM) and spread (D.P.A.) based factors.

Conclusions

- ① We examine two pairs trading strategies in a reasonably efficient market.
- ② By capturing linear/nonlinear associations and covering a wider range of possible dependencies structures, the mixed copula strategy outperforms the distance method when the number of trading signals is equiparable, especially after the subprime mortgage crisis.
- ③ The mixed copula pairs trading strategy generates large and significant (at 1%) abnormal returns.
 - ▶ Only a small part of the pairs trading profits can be explained by market portfolio (beta), size (SMB), value (HML), investment (CMA), profitability (RMW), momentum (Mom) and reversal (LRev) based factors.

- Machine Learning and AI-based solutions
 - ▶ Reinforcement Learning
- News Sentiment
 - ▶ Effect of negative news is bigger than positive news.

Thank you! Questions?

Subperiod Analysis

Table 6: Excess returns on committed capital on portfolios of Top 5 pairs after costs.

| Strategy | Mean Return (%) | Sharpe ratio | Sortino ratio |
|------------------------------------|--------------------|-----------------|------------------|
| Return on Committed Capital | | | |
| <i>Panel A: 1991-1995</i> | | | |
| S&P 500 | 7.17 | 0.72 | 1.30 |
| Mixed Copula | 2.66 | 0.45 | 0.74 |
| <i>Panel B: 1996-2000</i> | | | |
| S&P 500 | 10.03 | 0.51 | 1.01 |
| Mixed Copula | 6.90 | 1.05 | 1.77 |
| <i>Panel C: 2001-2005</i> | | | |
| S&P 500 | -2.28 | -0.13 | -0.06 |
| Mixed Copula | 6.84 | 0.83 | 1.44 |
| <i>Panel D: 2006-2010</i> | | | |
| S&P 500 | -1.71 | -0.07 | 0.09 |
| Mixed Copula | 1.56 | 0.24 | 0.46 |
| <i>Panel E: 2011-2015</i> | | | |
| S&P 500 | 9.91 | 0.61 | 1.09 |
| Mixed Copula | 2.01 | 0.61 | 1.08 |

***, **, * significant at 1%, 5% and 10% levels, respectively.

Subperiod Analysis

Table 7: Excess returns on fully invested capital on portfolios of Top 5 pairs after costs.

| Strategy | Mean Return (%) | Sharpe ratio | Sortino ratio |
|---|--------------------|-----------------|------------------|
| Return on Fully Invested Capital | | | |
| <i>Panel A: 1991-1995</i> | | | |
| S&P 500 | 7.17 | 0.72 | 1.30 |
| Mixed Copula | 7.69 | 0.56 | 1.02 |
| <i>Panel B: 1996-2000</i> | | | |
| S&P 500 | 10.03 | 0.51 | 1.01 |
| Mixed Copula | 19.61 | 1.13 | 1.96 |
| <i>Panel C: 2001-2005</i> | | | |
| S&P 500 | -2.28 | -0.13 | -0.06 |
| Mixed Copula | 18.07 | 1.14 | 2.07 |
| <i>Panel D: 2006:2010</i> | | | |
| S&P 500 | -1.71 | -0.07 | 0.09 |
| Mixed Copula | 9.42 | 0.57 | 1.16 |
| <i>Panel E: 2011:2015</i> | | | |
| S&P 500 | 9.91 | 0.61 | 1.09 |
| Mixed Copula | 3.62 | 0.37 | 0.69 |

***, **, * significant at 1%, 5% and 10% levels, respectively.

Subperiod Analysis

Table 8: Excess returns on committed capital on portfolios of Top 20 pairs after costs.

| Strategy | Mean Return (%) | Sharpe ratio | Sortino ratio |
|------------------------------------|--------------------|-----------------|------------------|
| Return on Committed Capital | | | |
| <i>Panel A: 1991-1995</i> | | | |
| S&P 500 | 7.17 | 0.72 | 1.30 |
| Mixed Copula | 0.93 | 0.46 | 0.70 |
| <i>Panel B: 1996-2000</i> | | | |
| S&P 500 | 10.03 | 0.51 | 1.01 |
| Mixed Copula | 1.67 | 0.84 | 1.37 |
| <i>Panel C: 2001-2005</i> | | | |
| S&P 500 | -2.28 | -0.13 | -0.06 |
| Mixed Copula | 2.43 | 1.09 | 1.86 |
| <i>Panel D: 2006:2010</i> | | | |
| S&P 500 | -1.71 | -0.07 | 0.09 |
| Mixed Copula | 0.49 | 0.22 | 0.38 |
| <i>Panel E: 2011:2015</i> | | | |
| S&P 500 | 9.91 | 0.61 | 1.09 |
| Mixed Copula | 0.70 | 0.77 | 1.30 |

***, **, * significant at 1%, 5% and 10% levels, respectively.

Subperiod Analysis

Table 9: Excess returns on fully invested capital on portfolios of Top 20 pairs after costs.

| Strategy | Mean Return (%) | Sharpe ratio | Sortino ratio |
|---|--------------------|-----------------|------------------|
| Return on Fully Invested Capital | | | |
| <i>Panel A: 1991-1995</i> | | | |
| S&P 500 | 7.17 | 0.72 | 1.30 |
| Mixed Copula | 8.18 | 0.63 | 1.10 |
| <i>Panel B: 1996-2000</i> | | | |
| S&P 500 | 10.03 | 0.51 | 1.01 |
| Mixed Copula | 18.48 | 1.08 | 1.85 |
| <i>Panel C: 2001-2005</i> | | | |
| S&P 500 | -2.28 | -0.13 | -0.06 |
| Mixed Copula | 21.07 | 1.34 | 2.42 |
| <i>Panel D: 2006:2010</i> | | | |
| S&P 500 | -1.71 | -0.07 | 0.09 |
| Mixed Copula | 12.09 | 0.74 | 1.48 |
| <i>Panel E: 2011:2015</i> | | | |
| S&P 500 | 9.91 | 0.61 | 1.09 |
| Mixed Copula | 2.33 | 0.25 | 0.49 |

***, **, * significant at 1%, 5% and 10% levels, respectively.

Subperiod Analysis

Table 10: Excess returns on committed capital on portfolios of Top 35 pairs after costs.

| Strategy | Mean Return (%) | Sharpe ratio | Sortino ratio |
|------------------------------------|--------------------|-----------------|------------------|
| Return on Committed Capital | | | |
| <i>Panel A: 1991-1995</i> | | | |
| S&P 500 | 7.17 | 0.72 | 1.30 |
| Mixed Copula | 0.70 | 0.60 | 0.93 |
| <i>Panel B: 1996-2000</i> | | | |
| S&P 500 | 10.03 | 0.51 | 1.01 |
| Mixed Copula | 0.99 | 0.84 | 1.37 |
| <i>Panel C: 2001-2005</i> | | | |
| S&P 500 | -2.28 | -0.13 | -0.06 |
| Mixed Copula | 1.59 | 1.23 | 2.11 |
| <i>Panel D: 2006:2010</i> | | | |
| S&P 500 | -1.71 | -0.07 | 0.09 |
| Mixed Copula | 0.35 | 0.28 | 0.46 |
| <i>Panel E: 2011:2015</i> | | | |
| S&P 500 | 9.91 | 0.61 | 1.09 |
| Mixed Copula | 0.50 | 0.86 | 1.56 |

***, **, * significant at 1%, 5% and 10% levels, respectively.

Subperiod Analysis

Table 11: Excess returns on fully invested capital on portfolios of Top 35 pairs after costs.

| Strategy | Mean Return (%) | Sharpe ratio | Sortino ratio |
|---|--------------------|-----------------|------------------|
| Return on Fully Invested Capital | | | |
| <i>Panel A: 1991-1995</i> | | | |
| S&P 500 | 7.17 | 0.72 | 1.30 |
| Mixed Copula | 8.50 | 0.65 | 1.14 |
| <i>Panel B: 1996-2000</i> | | | |
| S&P 500 | 10.03 | 0.51 | 1.01 |
| Mixed Copula | 19.10 | 1.12 | 1.93 |
| <i>Panel C: 2001-2005</i> | | | |
| S&P 500 | -2.28 | -0.13 | -0.06 |
| Mixed Copula | 21.81 | 1.38 | 2.50 |
| <i>Panel D: 2006:2010</i> | | | |
| S&P 500 | -1.71 | -0.07 | 0.09 |
| Mixed Copula | 12.39 | 0.76 | 1.51 |
| <i>Panel E: 2011:2015</i> | | | |
| S&P 500 | 9.91 | 0.61 | 1.09 |
| Mixed Copula | 2.56 | 0.27 | 0.53 |

***, **, * significant at 1%, 5% and 10% levels, respectively.