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

Fernando A. B. Sabino da Silva[1], Flavio A. Ziegelmann[1,2], Joao F. Caldeira[2]

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

Pairs Trading is one type of Statistical “Arbitrage”

- Identify a pair of stocks whose prices tend to move together
- When they diverge
  - short the “winner”
  - buy the “loser”
- Reverse your positions when the two prices converge $\Rightarrow$ Profit from the reversal in trend

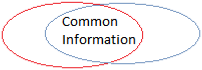
Pairs Trading is one type of Statistical “Arbitrage”

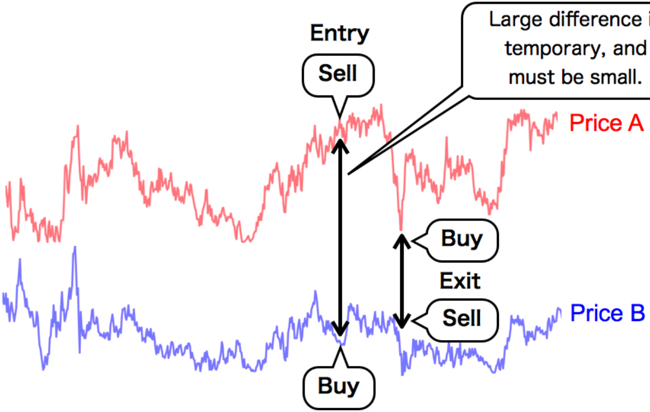
- Identify a pair of stocks whose prices tend to move together
- When they diverge
  - short the “winner"
  - buy the “loser"
- Reverse your positions when the two prices converge $\Rightarrow$ Profit from the reversal in trend

Pairs Trading is one type of Statistical “Arbitrage”

- Identify a pair of stocks whose prices tend to move together
- When they diverge
  - short the “winner”
  - buy the “loser”
- Reverse your positions when the two prices converge $\Rightarrow$ Profit from the reversal in trend

Univariate Pairs Trading





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“(Hansell *et al.*, 1989)
- Imperfect Markets?
  - Well-planned assault on the Efficient Market Hypothesis?
- Overaction?
  - Contrarian profits are in part due to over-reaction to firm-specific factors (Jegadeesh and Titman's, 1995)

Relative Pricing

- Pairs Trading does not seek to determine the absolute price of any stock
- Approximate APT Models
  - Long-short "arbitrage in expectations"
  - Eliminate relative mispricing
  - 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 ⇒ 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 ⇒ 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 ⇒ 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 ⇒ Linear correlation fully describes the dependence
- Tail dependence
  - Heavy tails
  - Possibly Asymmetric
- A single distance measure ⇒ fail to catch the dynamics of the spread between a pair of securities?
  - Volatility differs at different price levels ⇒ 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 ⇒ Linear correlation fully describes the dependence
- Tail dependence
  - Heavy tails
  - Possibly Asymmetric
- A single distance measure ⇒ fail to catch the dynamics of the spread between a pair of securities?
  - Volatility differs at different price levels ⇒ 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 ⇒ Linear correlation fully describes the dependence
- Tail dependence
  - Heavy tails
  - Possibly Asymmetric
- A single distance measure ⇒ fail to catch the dynamics of the spread between a pair of securities?
  - Volatility differs at different price levels ⇒ 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 $\Rightarrow$ 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 $X_1, ..., X_d$ *be random variables with distribution functions* $F_1, ..., F_d$*, respectively. Then, there exists an d-copula* C *such that,*

$$F\left(x_1, ..., x_d\right) = C\left(F_1\left(x_1\right), ..., F_d\left(x_d\right)\right), \tag{1}$$

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

Why should we care about copulas?

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

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

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

where $u_i{=}F_i(x_i)$, $i = 1, ..., d$.

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

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

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

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

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

Copula

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

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

Copula Method

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

$$
\begin{aligned}
MI_{X|Y}^t &= P(R_t^X < r_t^X \mid R_t^Y = r_t^Y) \\
&\text{and} \\
MI_{Y|X}^t &= P(R_t^Y < r_t^Y \mid R_t^X = r_t^X),
\end{aligned}
$$

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

$$
\begin{aligned}
MI^t_{X|Y} = \quad & \frac{\partial C(u_1, u_2)}{\partial u_2} = \quad P(R^X_t < r^X_t \mid R^Y_t = r^Y_t) \\
\text{and} & \\
MI^t_{Y|X} = \quad & \frac{\partial C(u_1, u_2)}{\partial u_1} = \quad P(R^X_t < r^X_t \mid R^Y_t = r^Y_t).
\end{aligned}
\tag{4}
$$

Copula Method

$$\begin{aligned} MI^t_{X|Y} = \quad & \frac{\partial C(u_1,u_2)}{\partial u_2} = \quad P(R^X_t < r^X_t \mid R^Y_t = r^Y_t) \\ \text{and} \quad & \\ MI^t_{Y|X} = \quad & \frac{\partial C(u_1,u_2)}{\partial u_1} = \quad P(R^X_t < r^X_t \mid R^Y_t = r^Y_t). \end{aligned} \tag{4}$$

- 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
  - Two underlying stocks are considered fairly-valued

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

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

  - Cumulative mispricing indexes

$$\begin{aligned} M_{1,t} &= M_{1,t-1} + m_{1,t} \\ M_{2,t} &= M_{2,t-1} + m_{2,t} \end{aligned}$$

    - Positive M1 and negative M2 $\Rightarrow$ Stock 1 is overvalued relative to stock 2
    - Note: M_1 and M_2 are set to zero at the beggining 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 M1 and negative M2 $\Rightarrow$ Stock 1 is overvalued relative to stock 2
    - Note: M_1 and M_2 are set to zero at the beggining of the trading period

Copula

- 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.

Copula

- 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.,

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

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.,

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

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.,

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

   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.,

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

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

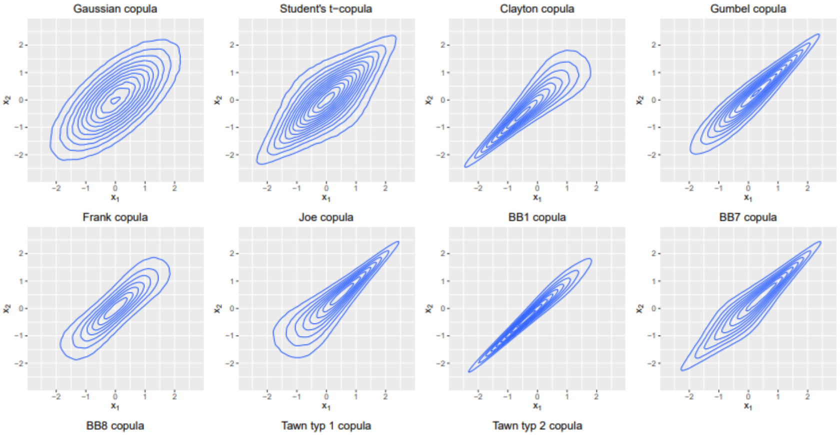
$$\mathcal{C}_{\theta}^{CFG}\left(u_1, u_2\right) = \pi_1 \mathcal{C}_{\alpha}^{C}\left(u_1, u_2\right) + \pi_2 \mathcal{C}_{\beta}^{F}\left(u_1, u_2\right) + \left(1 - \pi_1 - \pi_2\right) \mathcal{C}_{\delta}^{G}\left(u_1, u_2\right),$$

and

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

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 π_1, $\pi_2 \in [0, 1]$.

Tail Dependence



Data

- **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

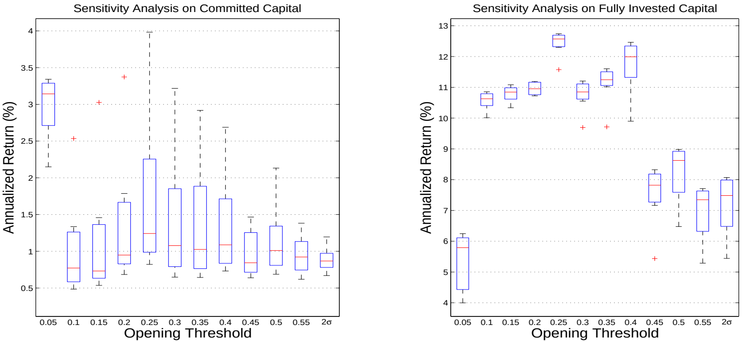


Figure 1: **Annualized returns of pairs trading strategies after costs on committed and fully invested capital**

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

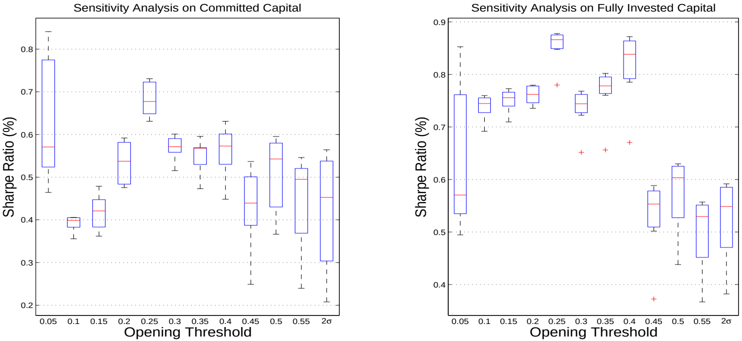


Figure 2: **Sharpe ratio of pairs trading strategies after costs on committed and fully invested capital**

Trading Statistics

Table 1: Trading statistics.

Strategy	Distance	Mixed Copula
	Panel A: Top 5	
Average price deviation trigger for opening pairs	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
	Panel B: Top20	
Average price deviation trigger for opening pairs	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

Trading Statistics

Table 2: Trading statistics.

Strategy	Distance	Mixed Copula
		Panel C: Top 35
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				
			Panel A - Top 5 pairs				
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
			Panel B - Top 20 pairs				
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
			Panel C - Top 35 pairs				
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.

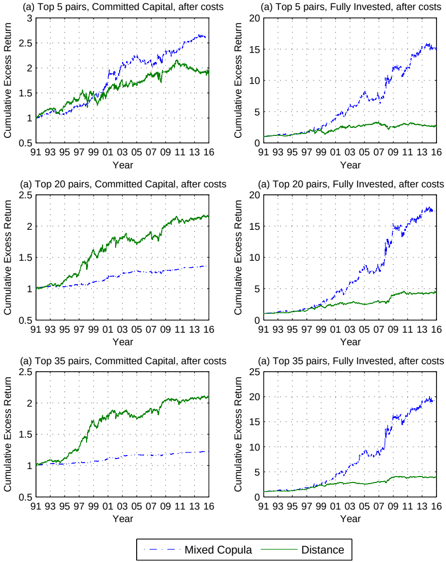
Trading Performance

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

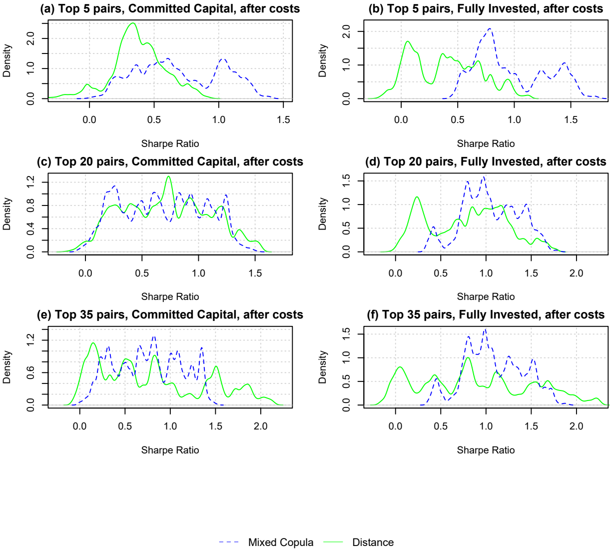
Strategy	Mean Return (%)	Sharpe ratio	Sortino ratio	t-stat	% of negative trades	MDD1	MDD2
			Return on Fully Invested Capital				
			Panel A - Top 5 pairs				
Distance	4.01	0.28	0.57	1.81*	46.98	8.70	38.36
Mixed Copula	11.58	0.78	1.43	4.26***	41.79	9.00	25.68
			Panel B - Top 20 pairs				
Distance	6.07	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
			Panel C - Top 35 pairs				
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



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	0.0091	-0.0032	0.0113	0.0003	-0.0029	-0.0107	-0.0084	0.028	0.027
	(1.89)*	(4.22)***	(-0.71)	(2.05)**	(0.25)	(-0.18)	(−4.80)***	(−1.96)**		
Mixed Copula	0.0035	0.0052	-0.0043	0.0039	-0.0035	0.0027	-0.0054	-0.0057	0.015	0.014
	(3.55)***	(3.68)***	(−1.83)*	(1.20)	(-0.99)	(0.63)	(−2.99)***	(-1.57)		
Section 2: Return on Fully Invested Capital										
Distance	0.0040	0.0170	-0.0031	0.0185	0.0049	-0.0018	-0.0161	-0.0150	0.025	0.024
	(1.75)*	(4.88)***	(-0.45)	(2.22)**	(0.76)	(0.05)	(−4.30)***	(−1.97)**		
Mixed Copula	0.0098	0.0148	-0.0084	0.0152	-0.0053	0.0087	-0.0082	-0.0222	0.018	0.017
	(4.17)***	(3.51)***	-1.45	1.6355	-0.60	0.75	(−2.19)**	(−2.08)**		

***, **, * 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

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	0.0091	-0.0032	0.0113	0.0003	-0.0029	-0.0107	-0.0084	0.028	0.027
	(1.89)*	(4.22)***	(-0.71)	(2.05)**	(0.25)	(-0.18)	(−4.80)***	(−1.96)**		
Mixed Copula	0.0035	0.0052	-0.0043	0.0039	-0.0035	0.0027	-0.0054	-0.0057	0.015	0.014
	(3.55)***	(3.68)***	(−1.83)*	(1.20)	(-0.99)	(0.63)	(−2.99)***	(-1.57)		
Section 2: Return on Fully Invested Capital										
Distance	0.0040	0.0170	-0.0031	0.0185	0.0049	-0.0018	-0.0161	-0.0150	0.025	0.024
	(1.75)*	(4.88)***	(-0.45)	(2.22)**	(0.76)	(0.05)	(−4.30)***	(−1.97)**		
Mixed Copula	0.0098	0.0148	-0.0084	0.0152	-0.0053	0.0087	-0.0082	-0.0222	0.018	0.017
	(4.17)***	(3.51)***	-1.45	1.6355	-0.60	0.75	(−2.19)**	(−2.08)**		

***, **, * 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

Summary of findings

1. 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.

2. 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.

Summary of findings

1. 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.

2. 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.

Summary of findings

1. 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.

2. 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.

Extensions

- Copula-based arbitrage for baskets to increase information dependency and measure relative pricing more comprehensively.

Extensions

- Machine Learning and AI-based solutions
  - Deep Reinforcement Learning

- News Sentiment
  - Enhances a pairs-trading strategy using an abnormal news volume and sentiment overlay
  - Effect of negative news is bigger than positive news

Thank you! Questions?

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		
	Panel A: 1991-1995		
S&P 500	7.17	0.72	1.30
Mixed Copula	2.66	0.45	0.74
	Panel B: 1996-2000		
S&P 500	10.03	0.51	1.01
Mixed Copula	6.90	1.05	1.77
	Panel C: 2001-2005		
S&P 500	-2.28	-0.13	-0.06
Mixed Copula	6.84	0.83	1.44
	Panel D: 2006:2010		
S&P 500	-1.71	-0.07	0.09
Mixed Copula	1.56	0.24	0.46
	Panel E: 2011:2015		
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			
Panel A: 1991-1995			
S&P 500	7.17	0.72	1.30
Mixed Copula	7.69	0.56	1.02
Panel B: 1996-2000			
S&P 500	10.03	0.51	1.01
Mixed Copula	19.61	1.13	1.96
Panel C: 2001-2005			
S&P 500	-2.28	-0.13	-0.06
Mixed Copula	18.07	1.14	2.07
Panel D: 2006:2010			
S&P 500	-1.71	-0.07	0.09
Mixed Copula	9.42	0.57	1.16
Panel E: 2011:2015			
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		
	Panel A: 1991-1995		
S&P 500	7.17	0.72	1.30
Mixed Copula	0.93	0.46	0.70
	Panel B: 1996-2000		
S&P 500	10.03	0.51	1.01
Mixed Copula	1.67	0.84	1.37
	Panel C: 2001-2005		
S&P 500	-2.28	-0.13	-0.06
Mixed Copula	2.43	1.09	1.86
	Panel D: 2006:2010		
S&P 500	-1.71	-0.07	0.09
Mixed Copula	0.49	0.22	0.38
	Panel E: 2011:2015		
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			
Panel A: 1991-1995			
S&P 500	7.17	0.72	1.30
Mixed Copula	8.18	0.63	1.10
Panel B: 1996-2000			
S&P 500	10.03	0.51	1.01
Mixed Copula	18.48	1.08	1.85
Panel C: 2001-2005			
S&P 500	-2.28	-0.13	-0.06
Mixed Copula	21.07	1.34	2.42
Panel D: 2006:2010			
S&P 500	-1.71	-0.07	0.09
Mixed Copula	12.09	0.74	1.48
Panel E: 2011:2015			
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			
Panel A: 1991-1995			
S&P 500	7.17	0.72	1.30
Mixed Copula	0.70	0.60	0.93
Panel B: 1996-2000			
S&P 500	10.03	0.51	1.01
Mixed Copula	0.99	0.84	1.37
Panel C: 2001-2005			
S&P 500	-2.28	-0.13	-0.06
Mixed Copula	1.59	1.23	2.11
Panel D: 2006:2010			
S&P 500	-1.71	-0.07	0.09
Mixed Copula	0.35	0.28	0.46
Panel E: 2011:2015			
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.

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		
	Panel A: 1991-1995		
S&P 500	7.17	0.72	1.30
Mixed Copula	8.50	0.65	1.14
	Panel B: 1996-2000		
S&P 500	10.03	0.51	1.01
Mixed Copula	19.10	1.12	1.93
	Panel C: 2001-2005		
S&P 500	-2.28	-0.13	-0.06
Mixed Copula	21.81	1.38	2.50
	Panel D: 2006:2010		
S&P 500	-1.71	-0.07	0.09
Mixed Copula	12.39	0.76	1.51
	Panel E: 2011:2015		
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.