# UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

# FACULDADE DE CIÊNCIAS ECONÔMICAS
# PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

## FERNANDO A. BOEIRA SABINO DA SILVA

## ENSAIOS EM CÓPULAS E FINANÇAS EMPÍRICAS

Porto Alegre

2017

# FERNANDO A. BOEIRA SABINO DA SILVA

# ENSAIOS EM CÓPULAS E FINANÇAS EMPÍRICAS

Tese submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Doutor em Economia, com ênfase em Economia Aplicada.

Orientador: Prof. Dr. Flávio A. Ziegelmann

Porto Alegre

2017

# FERNANDO A. BOEIRA SABINO DA SILVA

# ENSAIOS EM CÓPULAS E FINANÇAS EMPÍRICAS

Tese submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Doutor em Economia, com ênfase em Economia Aplicada.

Aprovada em: Porto Alegre, 25 de abril de 2017.

BANCA EXAMINADORA:

---

Prof. Dr. Flávio Augusto Ziegelmann – Orientador
UFRGS

---

Prof. Dr. André Alves Portela Santos
UFSC

---

Prof. Dr. Guilherme Valle Moura
UFSC

---

Prof. Dr. Tiago Pascoal Filomena
UFRGS

Para os meus pais (in memorian).

# AGRADECIMENTOS

Sou muito grato a todos que me ajudaram nesta caminhada, em especial aos meus pais, irmãos, Cristina e a sua família.

O amor é o alimento da vida. Cristina me ajudou a trilhar por caminhos com um número muito menor de espinhos, usando muito tempo para me deixar mais seguro de que estava indo na direção correta e sendo exemplo de disciplina.

Ao prof. Flávio pelos ensinamentos, incluindo os exemplos de justiça, correção e sensibilidade.

Aos membros da banca examinadora, pelo tempo dispensado na leitura deste trabalho e pelos comentários e críticas.

Aos meus colegas, em especial Hudson e Gabrielito.

A Tainan por toda a ajuda na editoração final da tese.

Aos professores e técnicos administrativos do PPGE pela dedicação e por todo o apoio.

A todos aqueles que não tenham sido citados aqui, mas que de algum modo colaboraram para a realização deste trabalho.

Por fim, que eu possa aplicar o que recebi para o propósito do bem, desviando de mim os maus pensamentos como o orgulho.

# RESUMO

Nesta tese discutimos abordagens que utilizam cópulas para descrever dependências entre instrumentos financeiros e avaliamos a performance destes métodos. Muitas crises financeiras aconteceram desde o final da década de 90, incluindo a crise asiática (1997), a crise da dívida da Rússia (1998), a crise da bolha da internet (2000), as crises após o 9/11 (2001) e a guerra do Iraque (2003), a crise do *subprime* or crise financeira global (2007-08), e a crise da dívida soberana europeia (2009). Todas estas crises levaram a uma perda maciça de riqueza financeira e a um aumento da volatilidade observada, e enfatizaram a importância de uma política macroprudencial mais robusta. Em outras palavras, perturbações financeiras tornam os processos econômicos altamente não-lineares, levando os principais bancos centrais a tomarem medidas contrárias para conter a angústia financeira. Devido aos complexos padrões de dependência dos mercados financeiros, uma abordagem multivariada em grandes dimensões para a análise da dependência caudal é seguramente mais perspicaz do que assumir retornos com distribuição normal multivariada. Dada a sua flexibilidade, as cópulas são capazes de modelar melhor as regularidades empiricamente verificadas que são normalmente atribuídas a retornos financeiros multivariados: (1) volatilidade condicional assimétrica com maior volatilidade para grandes retornos negativos e menor volatilidade para retornos positivos (HAFNER, 1998); (2) assimetria condicional (AIT-SAHALIA; BRANDT, 2001; CHEN; HONG; STEIN, 2001; PATTON, 2001); (3) excesso de curtose (TAUCHEN, 2001; ANDREOU; PITTIS; SPANOS, 2001); e (4) dependência temporal não linear (CONT, 2001; CAMPBELL; LO; MACKINLAY, 1997). A principal contribuição dos ensaios é avaliar se abordagens mais sofisticadas do que o método da distância e o tradicional modelo de Markowitz podem tirar proveito de quaisquer anomalias/fricções de mercado. Os ensaios são uma tentativa de fornecer uma análise adequada destas questões usando conjuntos de dados abrangentes e de longo prazo. Empiricamente, demonstramos que as abordagens baseadas em cópulas são úteis em todos os ensaios, mostrando-se benéficas para modelar dependências em diferentes cenários, avaliando as medidas de risco caudais mais adequadamente e gerando rentabilidade superior a dos *benchmarks* utilizados.

**Palavras chave**: Alocação de Ativos. Copula. Distância. Alta Frequência. Seleção de Portfólios. Medidas Realizadas. Gerenciamento de Riscos. S&P500. Bootstrap Estacionário. Arbitragem Estatística. WCVaR.

# ABSTRACT

In this thesis we discuss copula-based approaches to describe statistical dependencies within financial instruments and evaluate its performance. Many financial crises have occurred since the late 1990s, including the Asian crisis (1997), the Russian national debt crisis (1998), the dot-com bubble crisis (2000), the crises after 9-11 (2001) and Iraq war (2003), the subprime mortgage crisis or global financial crisis (2007-08), and the European sovereign debt crisis (2009). All of these crises lead to a massive loss of financial wealth and an upward in observed volatility and have emphasized the importance of a more robust macro-prudential policy. In other words, financial disruptions make the economic processes highly nonlinear making the major central banks to take counter-measures in order to contain financial distress. The methods for modeling uncertainty and evaluating the market risk on financial markets are now under more scrutiny after the global financial crisis. Due to the complex dependence patterns of financial markets, a high-dimensional multivariate approach to tail dependence analysis is surely more insightful than assuming multivariate normal returns. Given its flexibility, copulas are able to model better the empirically verified regularities normally attributed to multivariate financial returns: (1) asymmetric conditional volatility with higher volatility for large negative returns and smaller volatility for positive returns (HAFNER, 1998); (2) conditional skewness (AIT-SAHALIA; BRANDT, 2001; CHEN; HONG; STEIN, 2001; PATTON, 2001); (3) excess kurtosis (TAUCHEN, 2001; ANDREOU; PITTIS; SPANOS, 2001); and (4) nonlinear temporal dependence (CONT, 2001; CAMPBELL; LO; MACKINLAY, 1997). The principal contribution of the essays is to assess if more sophisticated approaches than the distance method and plain Markowitz model can take advantage of any market anomalies/fricctions. The essays are one attempt to provide a proper analysis in these issues using a long-term and comprehensive datasets. We empirically show that copula-based approaches are useful in all essays, proving beneficial to model dependencies in different scenarios, assessing the downside risk measures more adequately and yielding higher profitability than the benchmarks.

**Keywords**: Asset Allocation. Copula. Distance. High-Frequency. Portfolio Selection. Realized Measures. Risk Management. S&P500. Stationary Bootstrap. Statistical Arbitrage. WCVaR.

**Chapter 1**

**Introduction**

This thesis is composed by three essays. In the first and third essays, we investigate whether copulas can take advantage of any market anomalies/fricctions and thus improve the profitability of pairs trading. The second essay deals with an asset allocation problem in which we want to protect an investor against the worst possible scenario.

Pairs trading is an arbitrage strategy that involves identifying pairs of securities that historically move together and going short on the overvalued instrument and long on the undervalued as soon as a relative mispricing signal occurs and close the positions once the normalized prices cross. The distance method is the most known technique used to measure the spread or divergence between a pair of stocks. However, it is known that the distance method has a multivariate normal nature since it assumes a symmetric distribution of the spread between the normalized prices of the securities within a pair and it uses a single distance measure, which can be seen as an alternative measurement of the linear association, to describe the relationship between two stocks (XIE; LIEW, et al., 2016). We know that if the series have joint normal distribution, then the linear correlation fully describes the dependence between between stocks. However, it is well known that the dependence between two securities is rarely jointly normal (CAMPBELL; LO; MACKIN-LAY, 1997; CONT, 2001; ANE; KHAROUBI, 2003; MCNEIL; FREY; EMBRECHTS, 2015). The main feature of joint distributions characterized by tail dependence is the presence of heavy, and possibly asymmetric, tails. In this case, the traditional hypothesis of (multivariate) Gaussianity is completely inadequate. Therefore, a single distance measure may fail to catch the dynamics of the spread between a pair of securities, and thus initiate and close the trades at non-optimal positions.

In the first essay, we employ a mixed copula model, which consists in fitting, initially, the daily returns of the formation period using an ARMA(p,q)-GARCH(1,1) to model the marginals. For each pair we test the following elliptical and Archimedean copula functions: Gaussian, Student's t, Clayton, Frank, Gumbel, one Archimedean mixture copula consisting of the optimal linear combination of Clayton, Frank and Gumbel copulas and one mixture copula consisting of the optimal linear combination of Clayton, t and Gumbel copulas. These mixtures are chosen because these Archimedean copulas contain different tail dependence characteristics. It combines a copula with lower tail dependence and a copula with upper tail dependence to produce a more flexible copula capable of modeling the multivariate log returns. Hence, by using a mixture copula we cover a wider range of possible dependence structures within a single model. This allows us to capture better the dependence between the individual assets, which strongly influences the risk measures.

Using data from the S&P500 shares from 1990 to 2015, we find that the mixed copula strategy is able to generate a higher mean excess return than the traditional distance method under different weighting structures when the number of tradeable signals is equiparable. In addition, the mixed copula approach delivers economically larger alphas than the distance method for both weighting schemes, especially on fully invested capital, for Top 5 pairs. However, results suggest, since neither strategy is consistently superior in all subperiods, at least on committed capital, that the strategies may be combined to yield a viable long-short strategy.

In the second essay we compare, using data from the S&P500 shares from 1990 to 2015, the Worst Case Copula-CVaR with two benchmarks: (i) a multidimensional gaussian copula model, and (ii) an equally weighted portfolio.

The most well known portfolio optimization approach was introduced by Markowitz (1952). However, one of the assumptions of his work is that returns are normally distributed and that the risk is measured by the portfolio variance. In order to manage market risk we need to employ statistical models that are able to deal more adequately with empirically observed extreme events in financial markets than a Gaussian distribution does.

It is also known that mean-variance frontier is very sensitive to the inputs. Therefore, a robust optimization framework is required and we need to deal with stochastic uncertainty sets to protect investor against the worst case within this set. Copulas are also a very flexible tool for describing uncertainty, since they capture very well tail dependence in the presence of heavy and asymmetric tails, a common characteristic in the financial markets during financial distress co-movements. In this context, Rockafellar and Uryasev (2000) show how the Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) can be computed simultaneously, introducing auxiliary risk measures, and it can be used in conjunction with scenario based optimization algorithms to reduce the problem to a linear programming problem. This allows us to handle a portfolio with a large number of scenarios and instruments in addition to the adequate modeling of risks, constraints and utility functions. Zhu and Fukushima (2009) derive the Worst Case CVaR considering a mixture of distributions in a prescribed set, *i.e.*, the worst performing distribution combination in the set in the sense it produces the greatest CVaR. Finally, Kakouris and Rustem (2014) extend their framework considering the use of mixture copulas.

Our empirical analysis shows that the Mixed Copula-CVaR approach generates portfolios with better downside risk statistics for any rebalancing period and it is more profitable than the Gaussian Copula-CVaR and 1/N portfolios for daily and weekly rebalancing. To cope with the dimensionality problem we employ a similar approach to that used by Gatev, Goetzmann and Rouwenhorst (2006) to select a set of assets that are the most diversified, in some sense, to the S&P 500 index in the constituent set. We find that copula-based approaches offer better hedges against losses than the 1/N portfolio.

In the third essay, we use an intra-day realized measure to empirically evaluate a Copula-HEAVY pairs trading model against a Copula-GARCH and distance methods. We used 21 global stock market indexes from January 2000 to June 2016. Algorithmic trading systems can accommodate a wide range of different goals or risk-return preferences. Such environment take a big role in the financial markets today allowing one to operate in a much higher speed - since softwares perform real-time analysis and thus can identify price differentials very quickly - diversifying his/her portfolio more easily and, in particular, are very well suited to arbitrage. The necessary speed for making fast decision requires the use of more accurate methods, justifying the use of high frequency techniques and copulas in order to maximize profit. Moreover, these techniques can provide a better understanding of the market microstructure.

**Chapter 2**

**Performance of Pairs Trading on the S&P 500 using Distance and Mixed Copula models**

Fernando A. Boeira Sabino da Silva[1], Flavio A. Ziegelmann[2], and João F. Caldeira[3][4]

**Abstract.** We carry out a study to evaluate and compare the relative performance of the distance and mixed copula strategies. Using data from the S&P 500 shares from 1990 to 2015, we find that the mixed copula strategy is able to generate a higher mean excess return than the traditional distance method under different weighting structures when the number of tradeable signals is equiparable. Particularly, the mixed copula and distance methods show a mean annualized value-weighted excess returns after costs on committed and fully invested capital as high as 3.98% and 3.14% and 12.73% and 6.12%, respectively, with annual Sharpe ratios up to 0.88. The mixed copula strategy shows positive and significant alphas during the sample period after accounting for various risk-factors.

**Keywords:** Distance. Mixed Copula. Two-Dimensional Pairs Trading. S&P 500. Statistical Arbitrage.

**JEL Classifications.** C51, C58, G11.

[1]Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 91509-900, Brazil, e-mail: fsabino@ufrgs.br

[2]Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 91509-900, Brazil, e-mail: flavioz@ufrgs.br

[3]Department of Economics, Federal University of Rio Grande do Sul, Porto Alegre, RS 90040-000, Brazil, e-mail: joao.caldeira@ufrgs.br

## 2.1 Introduction

Pairs trading is a statistical arbitrage strategy that involves the simultaneous long/short of two relatively mispriced stocks which have strong historical co-movements. The performance of the strategies has been recently discussed in several studies with much interest in empirical finance, since the strategy has potential to generate sustained alpha through relatively low-risk positions. In addition, the strategy is claimed to be market neutral, which means that the investors are not exposed to market risk. The strategy was pioneered by Gerry Bamberger and later led by Nunzio Tartaglia's quantitative group at Morgan Stanley in the 1980s. However, it became popular through the study carried out by Gatev, Goetzmann and Rouwenhorst (2006), named distance method.

Currently, there are three main strategies for pairs trading: distance, cointegration and copula. In this paper, we will conduct an empirical investigation to offer some evidence of the behavior of the distance and mixed copula strategies under different investment scenarios.

The performance of the distance method has been measured thoroughly using different data sets and financial markets (GATEV; GOETZMANN; ROUWENHORST, 2006; PERLIN, 2009; DO; FAFF, 2010; 2012; BROUSSARD; VAIHEKOSKI, 2012; CALDEIRA; MOURA, 2013; RAD; LOW; FAFF, 2015). In an efficient market, strategies based on mean-reversion concepts should not generate consistent profits. However, Gatev, Goetzmann and Rouwenhorst (2006) find that pairs trading generates consistent statistical arbitrage profits in the U.S. equity market during 1962-2002 using CSRP data, although the profitability declines over the period. They obtain a mean excess return above 11% a year during the reported period. The authors attribute the abnormal returns to a non-identified systematic risk factor. They support their view showing that there is a high degree of correlation between the excess returns of no overlapping top pairs even after accounting for risk factors from an augmented version of Fama and French (1993)'s three factors. Do and Faff (2010) extend their work expanding the data sample and also find a declining trend - 33 basis points (bps) mean excess return per month for 2003-09 versus 124 basis points mean excess return per month for 1962-88. Do and Faff (2012) show that the distance method is unprofitable after 2002 when trading costs are considered. Broussard and Vaihekoski (2012) test the profitability of pairs trading under different weighting structures and trade initiation conditions using data from the Finnish stock market. They also find that their proposed strategy is profitable even after initiating the positions one day after the signal. Rad, Low and Faff (2015) evaluates distance, cointegration and copula methods using a long-term comprehensive data-set spanning over five decades. They find that the copula method has a weaker performance than the distance and cointegration methods in terms of excess returns and various risk-adjusted metrics.

The distance strategy (GATEV; GOETZMANN; ROUWENHORST, 2006) uses the distance between normalized security prices to capture the degree of mispricing between stocks. According to Xie, Liew, Wu and Zou (2016) the distance method has a multivariate normal nature since it assumes a symmetric distribution of the spread between the normalized prices of the stocks within a pair and it uses a single distance measure, which can be seen as an alternative measurement of the linear association, to describe the relationship between two stocks. We know that if the series have joint normal distribution, then the linear correlation fully describes the dependence between securities. However, a main feature of joint distributions characterized by tail dependence is the presence of heavy and possibly asymmetric tails. it is well known that the dependence between two securities are rarely jointly normal and thus the traditional hypothesis of (multivariate) gaussianity is completely inadequate (CAMPBELL; LO; MACKINLAY, 1997; CONT, 2001; ANÉ; KHAROUBI, 2003; MCNEIL; FREY; EMBRECHTS, 2015). Therefore, a single distance measure may fail to catch the dynamics of the spread between a pair of securities, and thus initiate and close the trades at non-optimal positions.

Due to the complex dependence patterns of financial markets, a high-dimensional multivariate approach to tail dependence analysis is surely more insightful than assuming multivariate normal returns. Since its flexibility, copulas are able to model better the empirically verified regularities normally attributed to multivariate financial returns:

(1) asymmetric conditional variance with higher volatility for large negative returns and smaller volatility for positive returns (HAFNER, 1998); (2) conditional skewness (AIT-SAHALIA; BRANDT, 2001; CHEN; HONG; STEIN, 2001; PATTON, 2001); (3) Leptokurdicity (TAUCHEN, 2001; ANDREOU; PITTIS; SPANOS, 2001); and (4) nonlinear temporal dependence (CONT, 2001; CAMPBELL; LO; MACKINLAY, 1997). Thus, to address these issues, Liew and Wu (2013) propose a pairs trading strategy based based on two-dimensional copulas to overcome the limitations of the distance method. However, they evaluate its performance using only three pre-selected pairs over a period of less than three years. Xie, Liew, Wu and Zou (2016) employ a similar methodology over a ten-year period with 89 stocks. Both studies find that the performance of copula strategy is superior to the distance strategy. Xie and Yuan (2013) set out the distance and cointegration approaches as special cases of copulas under certain regularity conditions. The authors also recommend further research on how to incorporate copulas in pairs selection. It is suggested there is a possibility of larger profits in terms of returns since copulas deals better with non-linear dependency structures. The approach may sound plausible but it may not lead to a viable standalone trading quantitative strategy due to issues as overfitting, hence do not justifying the marginal performance improvement given by a more complex model. As cited above Rad, Low and Faff (2015) use a more comprehensive data set consisting of all the shares in the US market from 1962 to 2014. Meanwhile, they find an opposite result. Particularly, the distance, cointegration and Copula-GARCH strategies show a mean monthly excess return of 36, 33, and 5 bps after transaction costs and 88, 83, and 43 bps before transaction costs.

The main novel contribution of this paper is to employ a mixed copula model in order to capture linear/nonlinear associations and at the same time cover a wider range of possible dependencies structures. We aim to assess if build a more sophisticated model can take advantage of any market frictions/anomalies uncovering relationships and patterns. We find that the mixture copulas are selected for more than 90% of the pairs in the scenarios investigated and that the strategy is able to generate a higher mean excess return than the distance method when the number of trading signals is equiparable. We also want to investigate the sensitivity of copula method to different opening thresholds and how trading costs and different measures for pairs selection affect the profitability of these strategies.

Our strategy consists in fitting, initially, the daily returns of the formation period using an ARMA(1,1)-GARCH(1,1) to model the marginals. For each pair, we test the following elliptical and Archimedean copula functions: Gaussian, t, Clayton, Frank, Gumbel, one Archimedean mixture copula consisting of the optimal linear combination of Clayton, Frank and Gumbel copulas and one mixture copula consisting of the optimal linear combination of Clayton, t and Gumbel copulas. Following Gatev, Goetzmann and Rouwenhorst (2006) we calculate returns using two weighting schemes: the return on committed capital and on fully invested capital. The former commits[1] equal amounts of capital to each one of the pairs even if the pair has not been traded[2], whereas the latter divides all capital between the pairs that are open.

We compare the performance out-of-sample of the strategies using a variety of criteria, all of which are computed using a rolling period procedure similar to that used by Gatev, Goetzmann and Rouwenhorst (2006) with the exception that the time horizon of formation and trading periods are rolled forward by six months as in Broussard and Vaihekoski (2012). The main criteria we focus are: (1) mean and cumulative excess return, (2) risk-adjusted metrics as Sharpe and Sortino ratios, (3) % of negative trades, (4) t-values for various risk factors, (5) maximum drawdown between two consecutive days and between two days within a period of maximum six month and (6) total number of pairs opened.

In order to find if pairs trading profitability is associated to exposure to different

---

[1] We assume zero return for non-open pairs, although in practice one could earn returns on idle capital.

[2] The commited capital is considered more realistic as it takes into account the opportunity cost of the capital that has been allocated for trading.

systematic risk factors[3], we regress daily excess returns on seven factors: daily Fama and French (2015)'s five research factors [4] plus momentum and long-term reversal. We find that the intercept is statistically greater than zero for all regressions at 1% level when considering the mixed copula strategy, showing that our results are robust to the augmented Fama and French (2015)'s risk adjustment factors. In addition, the share of observations with negative excess returns is smaller for the mixed copula than distance strategy.

To test the statistical significance of the returns, standard deviations and Sharpe ratio we use the stationary bootstrap of Politis and Romano (1994) using the automatic block-length selection of Politis and White (2004) and 10,000 bootstrap resamples. To compute the bootstrap p-values we use the methodology proposed by Ledoit and Wolf (2008). We aim to compare the results on a statistical basis to mitigate potential data snooping problems.

The remainder of the paper is organized as follows. The data, a general review of the distance and copula models and the trading strategies we have used are discussed in Section 2. Section 3 summarizes the empirical results of the analysis. Finally, Section 4 provides a brief conclusion. Additional results are reported in the Appendix.

## 2.2 Data and Methodology

Our data set consists of daily data of adjusted closing prices of all shares that belongs to the S&P500 market index from July 2nd, 1990 to December 31st, 2015. We obtain adjusted closing prices from Bloomberg and returns on the Fama and French factors from French's website[5]. The data set sample period is made up of 6,426 days and includes a total of 1100 stocks over all periods. Only stocks that are listed during the formation period are included in the analysis, i.e., around 500 stocks in each trading period.

Using data from the Center for Research in Security Prices (CRSP) from 1980 to 2006, French (2008) estimates that the cost of active investing, including total commissions, bid-ask spreads, and other cost investors pay for trading services, has dropped from 146 basis points in 1980 to 11 basis points in 2006. Considering the US stock live trades on the Nyse-Amex between August 1998 and September 2013 for a large institucional investor, Frazzini, Israel and Moskowitz (2015) estimate that the average trading costs for market impact (MI) and implementation shortfall methodology (IS) are 8.81 and 9.13 basis points, respectively, while the median trading costs are 6.24 and 7.63 basis points, respectively. For our data we assume trading costs on the order of 10 and 20 basis points.

### 2.2.1 Distance Framework

Our implementation of the distance strategy is similar to Gatev, Goetzmann and Rouwenhorst (2006) and Broussard and Vaihekoski (2012). We calculate the spread between the normalized daily closing prices (known as distance) of all combinations of stocks pairs during the next 12 months, named formation period, adjusting them by dividends, stock splits and other corporate actions. Specifically, the pairs are formed using data from January to December or from July to June. Prices are scaled to $1 at the beginning of

---

[3]The single-factor capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965), as well as its consumption based version (Breeden (1979)), among other extensions, has been empirically tested and rejected by numerous studies, which show that the cross-sectional variation in expected equity returns cannot be explained by the market beta alone, providing evidence that investors demand compensation for not being able to diversify firm-specific characteristics.

[4]Fama and French (2015) found evidences that the three factor model was not sufficient to explain a lot of the variation in average returns related to profitability and investment.

[5]http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_st_rev_factor_daily.html

each formation period and then evolve using the return series[6]. We then select the top 5, 10,...,35 of those combinations that have the smallest sum of squared spreads, allowing re-selection of a specific pair, during the formation period. These pairs are then traded in the next six months (trading period).

In Gatev, Goetzmann and Rouwenhorst (2006), when the spread diverges by two or more standard deviations (which is calculated in the formation period) from the mean, the stocks are assumed to be mispriced in terms of their relative value to each other and thus we open a short position in the outperforming stock and a long in the underperforming one.

The price divergence is expected to be temporary, i.e., the prices are expected to converge to its long-term mean value of 0 (mean-reverting behavior). Hence, the positions are closed once the normalized prices cross. The pair is then monitored for another divergence and thus a pair can complete multiple round-trip trades. Trades that do not converge can result in a loss if they are still open at the end of the trading period when they are automatically closed. This results in fat left tails. Therefore, since the conditional variance is empirically higher for large negative returns and smaller for positive returns, it may be inappropriate to use constant trigger points because the volatility differs at different price levels.

To calculate the daily percentage returns for a pair, we compute

$$r_{pt} = w_{1t}r_t^L - w_{2t}r_t^S, \tag{2.1}$$

where $L$ and $S$ stands for long and short, respectively. Returns $r_{pt}$ can be interpreted as excess returns since in (2.1) the riskless rate is canceled out when one calculates the long and short excess returns. The weights $w_{1t}$ and $w_{2t}$ are initially assumed to be one. After that, they change according to the changes in the value of the stocks, i.e., $w_{it} = w_{it-1}(1 + r_{it-1})$.

### 2.2.2 Copula Framework

The notion of copula was first introduced by Sklar (1959). Sklar's theorem states that any multivariate joint distribution can be written in terms of their univariate marginal distribution function and the dependence structure (represented in $C$) between the variables. We state the Sklar's Theorem below.

**Theorem 1** (Sklar's Theorem) 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,

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d)), \tag{2.2}$$

for all $\mathbf{x} = (x_1, ..., x_d) \in \mathbb{R}^d$. Conversely, if $C$ is an $d$-copula and $F_1, ..., F_d$ are distribution functions, then the function $F$ defined by (3.23) is a $d-$dimensional distribution function with margins $F_1, ..., F_d$. Furthermore, if the marginals $F_1, ..., F_d$ are all continuous, $C$ is unique. Otherwise $C$ is uniquely determined on $\text{Im} F_1 \times ... \times \text{Im} F_d$.

Using the Sklar's theorem we can derive an important relation between the marginal distributions and a copula. Let $f$ be a joint density function (of the $d-$dimensional distribution function $F$) and $f_1, ..., f_d$ univariate density functions of the margins $F_1, ..., F_d$. Assuming that $F(\cdot)$ and $C(\cdot)$ are differentiable, by (3.23) we have

---

[6]Missing values have been interpolated.

$$(2.3) \qquad \frac{\partial^d F\left(x_1,...,x_d\right)}{\partial x_1...\partial x_d} \;\equiv\; f\left(x_1,...,x_d\right) = \frac{\partial^d C\left(F_1\left(x_1\right),...,F_d\left(x_d\right)\right)}{\partial x_1...\partial x_d}$$

$$(2.4) \qquad\qquad = \; c\left(u_1,...,u_d\right)\prod_{i=1}^{d} f_i\left(x_i\right),$$

where $u_i = F_i\left(x_i\right)$, $i = 1,...,d$. Thus, copulas are functions that connect a multivariate distribution function and their marginal distributions. Thereafter, copulas accommodate various forms of dependence through suitable choice of the copula "correlation matrix" since they conveniently separate marginals from dependence component. These carry on all relevant information about the dependence structure between random variables and allow a greater flexibility in modeling multivariate distributions and their margins. The methodology allows one to derive joint distributions from marginals, even when these are not normally distributed. In fact, copulas allow the marginal distributions to be modeled independently from each other, and no assumption on the joint behavior of the marginals is required, which provides a great deal of flexibility in modeling joint distributions. From a modelling perspective, Sklar's Theorem allow us to estimate the multivariate distribution in two parts: (i) finding the marginal distributions; (ii) finding the dependency between the filtered data from (i).

The choice of the copula function is also not dependent on the marginal distributions. Thus, by using copulas, the linearity restriction that applies to the dependence structure of multivariate random variables in a traditional dependence setting is relaxed. Therefore, depending on the chosen copulas, different dependence structures can be modeled to allow for any asymmetries[7].

A further important property of copulas concerns the partial derivatives of a copula with respect to its variables. Let now $H$ be a bivariate function with marginal distribution functions $F$ and $G$. According to Sklar (1959) then there exists a copula $C : [0,1]^2 \rightarrow [0,1]$ such that $H(x_1, x_2) = C(F(x_1), G(x_2))$ for all for all $x_1, x_2 \in \mathbf{R^2}$. If $F$ and $G$ are continuous, then $C$ is unique; otherwise, $C$ is uniquely determined in $\operatorname{Im} F \times \operatorname{Im} G$. Conversely, if $C$ is a copula and $F$ and $G$ are distribution functions, then the function $H$ is a joint distribution function with marginals $F$ and $G$ and we can write

$$C(u_1, u_2) = H(F^{-1}(u_1), G^{-1}(u_2)), \qquad\qquad (2.5)$$

where $u_1 = F(x_1) \Rightarrow x_1 = F^{-1}(u_1)$, $u_2 = G(x_2)) \Rightarrow x_2 = G^{-1}(u_2)$ and $F^{-1}$ and $G^{-1}$ are the quasi-inverses of $F$ and $G$, respectively. For any copula $C$, $\frac{\partial C(u_1,u_2)}{\partial u_1}$ and $\frac{\partial C(u_1,u_2)}{\partial u_2}$ exist almost everywhere. The proposition below states that the partial derivatives of a copula function corresponds to the conditional probabilities of the random variables (see CHERUBINI; LUCIANO; VECCHIATO, 2004; NELSEN, 2006).

**Proposition 1** *Let $U_1$ and $U_2$ be two random variables with distribution $U(0,1)$. Then,*

---

[7]Copulas measures lower and upper tail dependencies and nonlinear and linear relationships in a rich set for describing dependencies between pairs. Copula is also invariant under strictly monotonic transformations (CHERUBINI; LUCIANO; VECCHIATO, 2004; NELSEN, 2006) and hence the same copula is obtained if we use price or return series, for example.

$$P\left(U_1 \le u_1 \,|U_2 = u_2\right) \quad = \quad \frac{\partial C\left(u_1, u_2\right)}{\partial u_2} = P\left(X_1 \le x_1 \,|X_2 = x_2\right),$$

$$P\left(U_2 \le u_2 \,|U_2 = u_1\right) \quad = \quad \frac{\partial C\left(u_1, u_2\right)}{\partial u_1} = P\left(X_2 \le x_2 \,|X_1 = x_1\right)$$

where

$$\frac{\partial C\left(u_1, u_2\right)}{\partial u_2} = P\left(U_1 \le u_1 \,|U_2 = u_2\right) = \lim_{h \to 0} P\left(U_1 \le u_1 \,|u_2 \le U_2 \le u_2 + h\right) \qquad (2.6)$$

and

$$\frac{\partial C\left(u_1, u_2\right)}{\partial u_1} = P\left(U_2 \le u_2 \,|U_1 = u_1\right) = \lim_{h \to 0} P\left(U_2 \le u_2 \,|u_1 \le U_1 \le u_1 + h\right). \qquad (2.7)$$

By using the fact that the partial derivative of the copula function gives the conditional distribution function, Xie, Liew, Wu and Zou (2016) define a measure to denote the degree of mispricing:

**Definition 1** *(Mispricing Index) 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*

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

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

Therefore, the conditional probabilities $MI_{X|Y}^t$ and $MI_{Y|X}^t$ indicate whether the return of $X$ is considered high or low at time $t$, given the information on the return of $Y$ on the time $t$ and the historical relation between the two stocks' returns, and vice-versa. For example, if the value of $MI_{X|Y}^t$ is equal to 0.5, $r_t^X$ is neither too high nor too low given $r_t^Y$ and their historical relation. In other words, the historical data indicate that, on average, there is an equal number of observations of the return of $X$ being larger or smaller than $r_t^X$ if the return of stock $Y$ is equal to $r_t^Y$ and therefore, a conditional value of 0.5 means that the two underlying stocks are considered fairly-valued. In this case, we can say that stock $X$ is fairly priced relative to stock $Y$ on day $t$.

Given current realizations $r_t^X$ and $r_t^Y$, if $F_X$ and $F_Y$ are the marginal distribution functions of $R_t^X$ and $R_t^Y$ and C is the copula connecting $F_X$ and $F_Y$, we define $u_1 = F_X\left(r_t^X\right)$ and $u_2 = F_Y\left(r_t^Y\right)$, and have

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

$$and \qquad\qquad\qquad\qquad (2.8)$$

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

Note that the conditional probabilities, $M_t^{X|Y}$ and $M_t^{Y|X}$, only measure the degrees of relative mispricing for a single day. To determine an overall degree of relative mispricing we follow Rad, Low and Faff (2015). Initially, let $m_{1,t}$ and $m_{2,t}$ be the overall mispricing indexes of stocks $X_1$ and $X_2$, defined by $\left(M_t^{X|Y} - 0.5\right)$ and $\left(M_t^{Y|X} - 0.5\right)$, respectively. At beggining of each trading period two cumulative mispriced indexed $M_1$ and $M_2$ are set to zero and then evolve for each day through

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

Positive (negative) $M_1$ and negative (positive) $M_2$ are interpreted as stock 1 (stock 2) being overvalued relative to stock 2 (stock 1).

We perform a sensitivity analysis to 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 for Top 5, 10,....,35 pairs. The positions are unwound when both cumulative mispriced indexes return to zero. The pairs are then monitored for other possible trades throughout the remainder of the trading period.

Rad, Low and Faff (2015) propose the following steps to obtain $M_{1,t}$ and $M_{2,t}$ using copulas: (1) First, we calculate daily returns for each stock during the formation period and estimate the marginal distributions of these returns separately by fitting an appropriate ARMA(p,q)-GARCH(1,1) model[8] to each univariate time series by obtaining the estimates $\widehat{\mu}_i$ and $\widehat{\sigma}_i$ of the conditional mean and variance of these processes, respectively. Moreover, using the estimated parametric models, we construct the standardized residuals vectors given, for each $i = 1, ..., t$, by

$$
\widehat{\varepsilon}_i = \frac{x_i - \widehat{\mu}_i}{\widehat{\sigma}_i}. \tag{2.9}
$$

The standardized residuals vectors are then converted to the pseudo-observations $z_i = \frac{n}{n+1} F_i(\widehat{\varepsilon}_i)$, where $F_i$ is estimated by using their empirical distribution function;

(2) After obtaining the estimated marginal distributions from the previous step , we estimate the two-dimensional copula model to data that has been transformed to [0,1] margins to connect the joint distributions with the marginals $F_X$ and $F_Y$, 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 that best fits the uniform marginals and estimate its parameter(s). Copulas that are tested in this step are Gaussian, t, Clayton, Frank, Gumbel, one Archimedean mixture copula consisting of the optimal linear combination of Clayton, Frank and Gumbel copulas and one mixture copula consisting of the optimal linear combination of Clayton, t and Gumbel copulas.[9].

Specifically, a mixture of Clayton, Frank and Gumbel copulas and a mixture of Clay-

---

[8] We look for the best ARMA(p,q) model up to order (1,1).

[9] Archimedean copulas contain different tail dependence characteristics. Clayton and Gumbel are nonsymmetric copulas that describe more accurately lower and upper tail dependence, respectively. The Frank copula is the only bivariate reflection symmetric Archimedean family but it has different properties when compared to the bivariate gaussian and bivariate t copulas. Hence, by using mixture copulas we cover a wider range of possible dependencies within a single model.

ton, t and Gumbel copulas can be written, respectively, as

$$\mathcal{C}_\theta^{CFG}(u_1, u_2) = \pi_1 \mathcal{C}_\alpha^C(u_1, u_2) + \pi_2 \mathcal{C}_\beta^F(u_1, u_2) + (1 - \pi_1 - \pi_2) \mathcal{C}_\delta^G(u_1, u_2), \qquad (2.10)$$

and

$$\mathcal{C}_\xi^{CFG}(u_1, u_2) = \pi_1 \mathcal{C}_\alpha^C(u_1, u_2) + \pi_2 \mathcal{C}_{\Sigma,\nu}^t(u_1, u_2) + (1 - \pi_1 - \pi_2) \mathcal{C}_\delta^G(u_1, u_2), \qquad (2.11)$$

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 $\pi_1, \pi_2 \in [0, 1]$. The estimates are obtained by the minimization of the negative log-likelihood consisting of the weighted densities of the copulas;

(3) Take the first derivative of the copula function to compute conditional probabilities and measure mispricing degrees $MI_{X|Y}$ and $MI_{Y|X}$ for each day in the trading period using the copula and estimated parameters;

(4) Build long and short positions of $Y$ and $X$ on the days that $M_{1,t} > \Delta_1$ and $M_{2,t} < \Delta_2$ if there is no positions in $X$ or $Y$. Conversely, build positions long/short of $X$ and $Y$ on the day that $M_{1,t} < \Delta_2$ and $M_{2,t} > \Delta_1$ if there is no positions in $X$ or $Y$;

(5) All positions are closed if $M_{1,t}$ reaches $\Delta_3$ or $M_{2,t}$ reaches $\Delta_4$, where $\Delta_1, \Delta_2, \Delta_3$ and $\Delta_4$ are predetermined thresholds or are automatically closed out on the last day of the trading period if they do not reach the thresholds. Here we use $\Delta_1 = 0.2, \Delta_2 = -0.2$ and $\Delta_3 = \Delta_4 = 0$.

## 2.3 Empirical Results

Next, we summarize the empirical results of the analysis.

### 2.3.1 Profitability of the Strategies

First, we analyze the sensitivity of the outcomes when the opening thresholds are changed to top 5-35 pairs we provide multiple boxplots in Figures 2.1 and 2.2. They report annualized excess returns (Figure 2.1) and annualized Sharpe ratios (Figure 2.2) for each of the strategies from 1991/2-2015 on commited capital and on fully invested capital after costs (10 bps)[10]. Pairs are formed based on the smallest sum of squared deviations. The last boxplot (from left to right) show the performance for the distance strategy, while the others report the outcomes using multiple opening trigger points for the cumulative mispriced indexes $M_1$ and $M_2$ (0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, and 0.55). Based on these outcomes we perform the subsequent analyses considering 0.2 as the opening threshold for the mixed copula strategy.

---

[10]The numerical experiments show that the performances out-of-sample stay very similar when we consider 20 bps. Since the results are very much alike they are not presented here and are available under request. Hereafter, we consider 10 basis points as transaction costs to report results for the remainder of this paper

**Figure 2.1: Annualized returns of pairs trading strategies after costs on committed and fully invested capital**
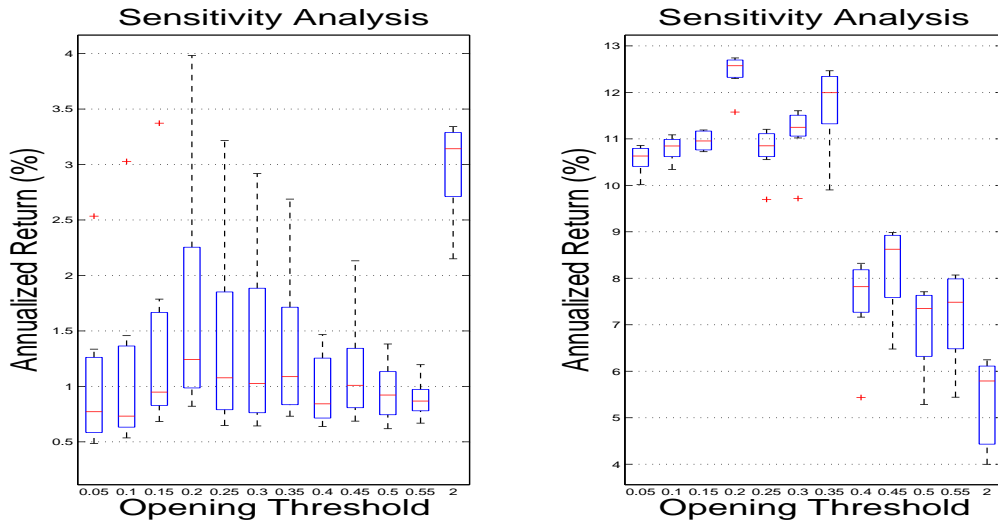


Source: Author's own elaboration (2017).

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. Pairs are formed based on the smallest sum of squared deviations.

**Figure 2.2: Sharpe ratio of pairs trading strategies after costs on committed and fully invested capital**
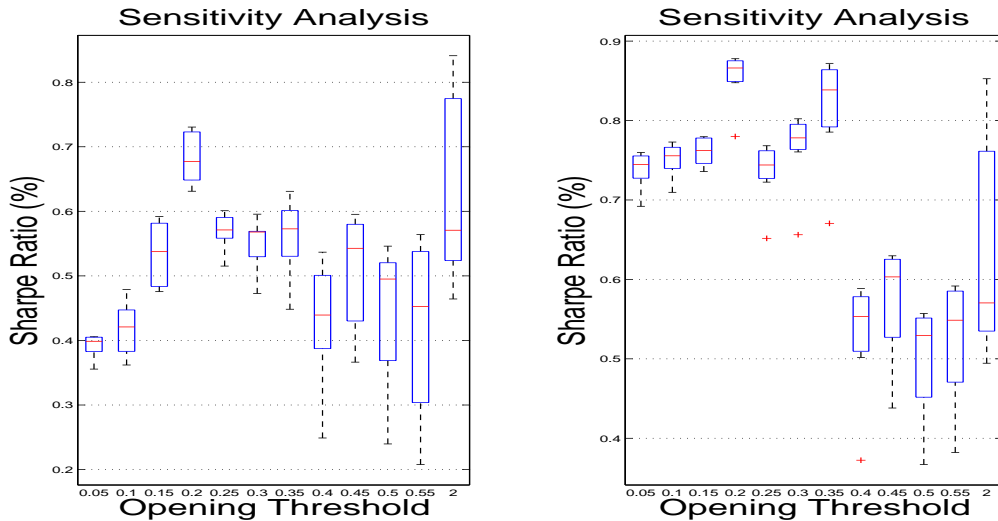


Source: Author's own elaboration (2017).

These boxplots show Sharpe ratios 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. Pairs are formed based on the smallest sum of squared deviations.

Table 2.1 reports annualized excess returns, annualized Sharpe and Sortino ratios, Newey and West (1987) adjusted t-statistics, share of negative observations, the maximum

drawdown in terms of maximum percentage drop between two consecutive days (MDD1) and between two days within a period of maximum six months (MDD2) for both strategies from 1991/2-2015, for Top 5 (Panel A), Top 20 (Panel B), and Top 35 (panel C) pairs after costs (10 bps) [11]. Furthermore, Section 1 shows the Return on Committed Capital and Section 2 on Fully Invested Capital.

By analyzing Table 2.1, it is possible to observe a series of important facts. First, note that the copula-based pairs strategy outperforms the distance method for all metrics for Top 5 pairs and committed capital. The mixed copula strategy yields the highest average excess returns and reach a Sharpe ratio of 0.63, slightly more than double what we get from investing in the tradicional distance method. The Sortino ratio confirms that the mixed copula model offers better risk-adjusted returns. The statistics also indicate that the mixed copula model delivers the highest t-statistics (statistically and economically significant at 1%) and a lower probability of a negative trade, where the share of days with negative returns (41.79%) is consistently less than the market performance (47.45% of negative returns over the period). Furthermore, the summary statistics also show that mixed copula method offers better hedges against losses than the distance strategy for Top 5 pairs on committed capital when considering the downside risk statistics $MDD1$ and $MDD2$. We find that the number of tradeable signals is only equiparable in this study for the different strategies for Top 5 pairs. We will explore this point further in the next subsection.

The listed results for Top 20 and Top 35 pairs on committed capital show that the distance strategy is more profitable than the mixed copula method, although the Sharpe ratios are similar, indicating that returns are alike when we take into account the risks taken. All profits are now economically and statistically significant at 1%. Overall, the copula method is again a less risky strategy regarding the drawdown measures.

Section 2 of Table 2.1 shows results on fully invested capital scheme. We can note that this approach yields a higher Sharpe and Sortino ratios for the copula-based strategy and the excess return of the portfolio averaged 11.6% a year (almost thrice the return of the committed capital approach), with large and significant Newey-West adjusted t-statistic of 4.26 for Top 5 pairs. Apart the tail risk (drawdown) measures, the mixed copula method consistently outperforms the distance strategy for all pairs considered.

---

[11]The numerical experiments show that the performances out-of-sample stay very similar when we consider 20 bps. The outcomes are also robust for the other number of pairs considered. Since the results are very much alike they are not presented here and are available under request.

**Table 2.1:** Excess returns 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 |
|---|---|---|---|---|---|---|---|
| | | | **Section 1: 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 | 3.06***$^{(<)}$ | 0.76 | 1.34 | 3.87*** | 48.02 | -2.70 | -7.52 |
| Mixed Copula | 0.82 | 0.73 | 1.19 | 3.95*** | 41.31 | -1.18 | -1.98 |
| | | | **Section 2: 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 | 6.12 | 0.67 | 1.20 | 3.58*** | 48.02 | -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 | 5.71 | 0.75 | 1.37 | 4.01*** | 48.02 | -4.23 | -15.07 |
| Mixed Copula | 12.73**$^{(>)}$ | 0.88 | 1.59 | 4.73*** | 41.28 | -9.00 | -25.68 |

%

*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). Pairs are formed based on the smallest sum of squared deviations. 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.

The symbol labels ($>$) and ($<$) indicate that the null hypothesis represented in 4.5 is rejected in favor of the alternative and that the average excess returns and/or the Sharpe Ratio of the mixed copula strategy are found greater ($>$) or lower ($<$) than the average excess returns and/or the Sharpe Ratio of the distance strategy, respectively.

Source: Author's own elaboration (2017).

Figure 2.3 show cumulative excess returns through the full dataset for both strategies for Top 5 (top), Top 20 (center) and Top 35 (bottom) pairs. The left panels displays cumulative returns on committed capital, whereas the right panels on fully invested capital. The patterns found in the figure strengthen the mean returns and t-statistics displayed in Table 2.1. It should be noted that the mixed copula strategy achieves a very favorable out-of-sample performance relative to the distance approach after the subprime mortgage crisis, especially after 2010 for Top 5 pairs (when the number of trades is comparable) on committed capital.

**Figure 2.3: Cumulative excess returns of pairs trading strategies after costs**
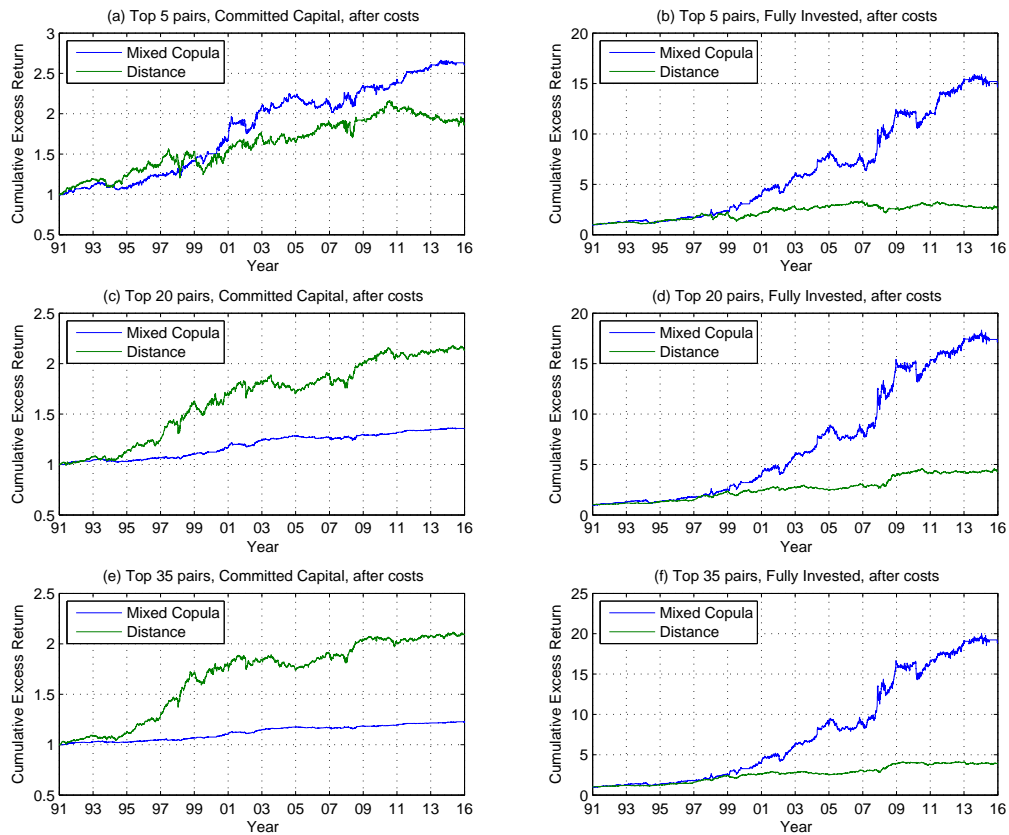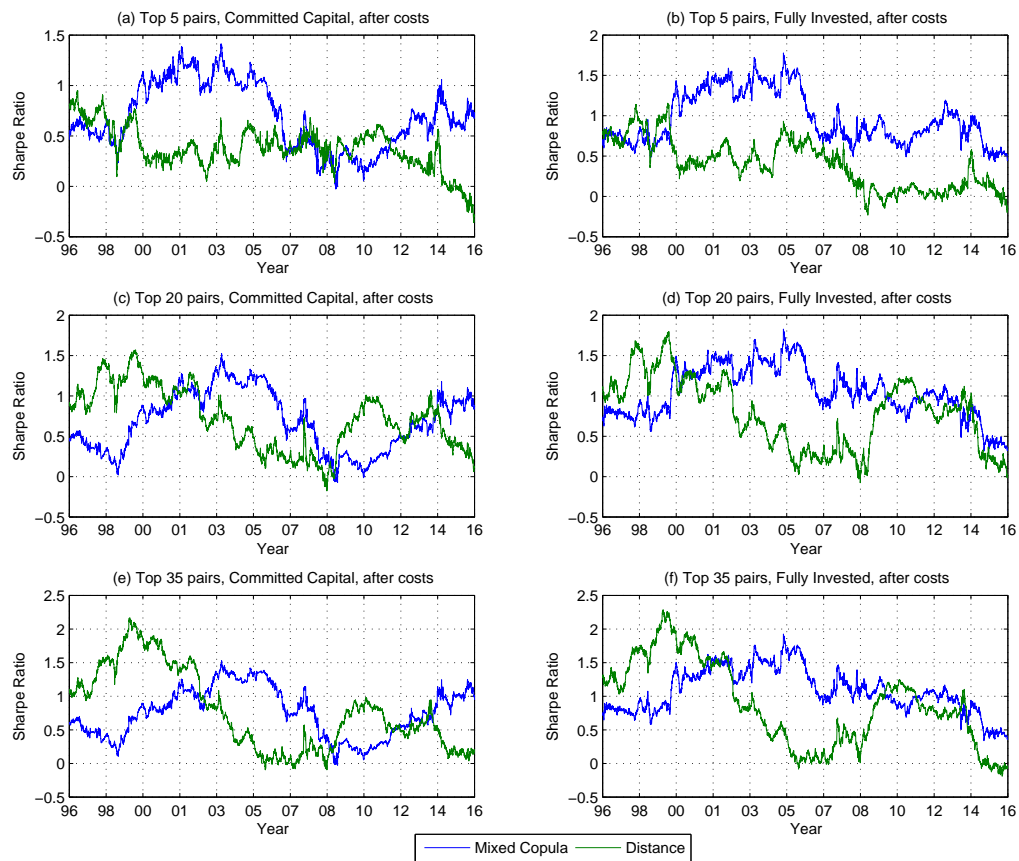


Source: Author's own elaboration (2017).

This figure shows how an investment of $1 evolves from July 1991 to December 2015 for each of the strategies.

Figure 2.4 shows five-year rolling window Sharpe ratio after costs. The figure reveals mixed results over the long-term period. However, when the number of tradeable signals is similar (Top 5 pairs), the copula-based approach yields the highest five-year Sharpe ratio (up to 1.41) on committed capital in 68.94% of the days. In 26.22% of the days over the period the copula method delivers a rolling Sharpe ratio above 1, whereas the distance strategy never attains 1. In fact, in only 25.4% of the full period the distance approach produces a five-year Sharpe ratio above 0.5 (and most often below zero after 2014), indicating that the strategy does not reward the risk taken.

For Top 20 and Top 35 pairs on committed capital the strategies show a more competitive pattern. The distance model presents a greater rolling window Sharpe ratio in 53.37% and 50.99% of the days, respectively. However, as we will explore further, the distance approach is a more volatile strategy, identifying a greater number of trading opportunities (more opportunities to make profit) than the copula approach, making the comparison less reliable for a larger number of pairs. The 5-year Sharpe ratios for distance and mixed copula methods surpass one in 24.97% and 24.04% for Top 20 pairs, and 31.92% and 30.33% for Top 35 pairs, respectively.

For fully invested weighting scheme the mixed copula approach achieves the highest five-year risk-adjusted statistic over the long-term period in 88.5%, 69.14% and 61.8% of the data sample for Top 5, Top 20 and Top 35 pairs, respectively.

**Figure 2.4: Cumulative excess returns of pairs trading strategies after costs**
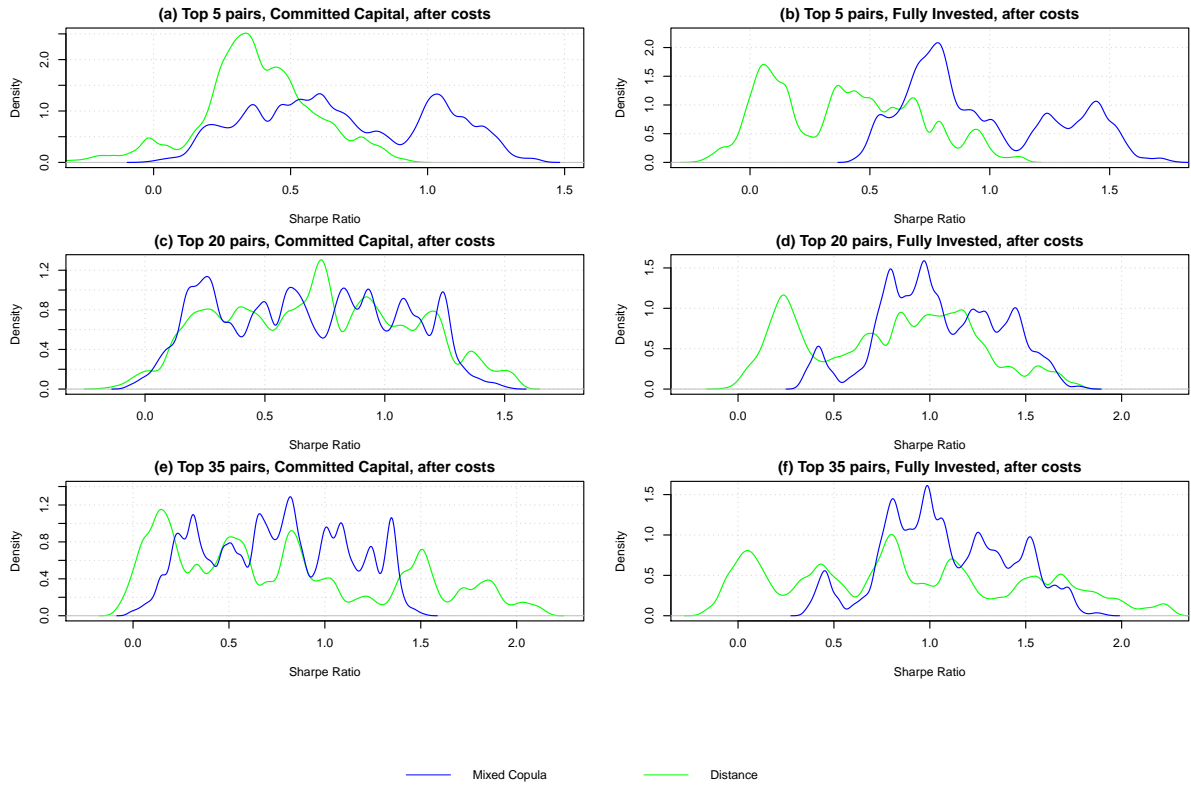


Source: Author's own elaboration (2017).

This figure shows how the 5-year rolling window Sharpe ratio evolves from July 1996 to December 2015 for each of the strategies.

Figure 2.5 shows the densities of the five-year Sharpe ratios after costs estimated by means of Sheather and Jones (1991) bandwidth. As can be observed, the densities reinforce our findings in Figure 2.4, showing that the right-hand tail of the distribution of the copula-based strategy remains long for Top 5 pairs.

## Figure 2.5: Densities of 5-year rolling window Sharpe ratio after costs



Source: Author's own elaboration (2017).

This figure shows how the 5-year rolling window Sharpe ratio densities evolve from July 1996 to December 2015 for each of the strategies using the method of Sheather and Jones (1991) to select the bandwidths.

One possible criticism might be that the conclusions are based on only one realization of the stochastic process of asset returns computed from the observed series of prices, since among thousands of different strategies is very likely that we find some that show superior performance in terms of excess return or Sharpe Ratio. In order to mitigate data-snooping criticisms, we use the stationary bootstrap[12] of Politis and Romano (1994) to compute the bootstrap p-values using the methodology proposed by Ledoit and Wolf (2008).

Our bootstrapped null distributions result from Theorem 2 of Politis and Romano (1994). We select the optimal block length for the stationary bootstrap following Politis and White (2004). As the optimal bootstrap block-length is different for each strategy, we average[13] the block-lengths found to proceed the comparisons between the mixed copula and the distance strategies.

To test the hypotheses that the average excess returns and Sharpe Ratios of the copula-based strategy are equal to that of distance method, that is,

$$H_0 : \mu_c - \mu_d = 0 \quad \text{and} \quad H_0 : \frac{\mu_c}{\sigma_c} - \frac{\mu_d}{\sigma_d} = 0, \tag{2.12}$$

---

[12]It allows for weakly dependent correlation over time.

[13]We also use the optimal block size for each strategy. We find that the results are robust to the optimal block size, and therefore, we do not report them here.

we compute, following Davison and Hinkley (1997), a two-sized $p$-value using $B = 10000$ (stationary) bootstrap re-samples as follows:

$$p_{sboot} = \begin{cases} 2\frac{\sum_{b=1}^{B} \mathbb{I}\{0 < t^{*(b)}\} + 1}{B+1}, & \text{if } median\left\{t^{*(1)}, ..., t^{*(B)}\right\} > 0, \\ 2\frac{\sum_{b=1}^{B} \mathbb{I}\{0 \geq t^{*(b)}\} + 1}{B+1}, & \text{otherwise,} \end{cases} \tag{2.13}$$

where $\mathbb{I}$ is the indicator function, $t^{*(b)}$ are the values in each block stationary bootstrap replication, and B denotes the number of bootstrap replications.

The symbol labels ($>$) and ($<$) in Table 2.1 indicate that the respective null hypothesis is rejected in favor of the alternative and that the average excess returns and/or the Sharpe Ratio of the mixed copula approach are found greater or lower than the average excess returns and/or the Sharpe Ratio of the distance method, respectively.

Overall, these results reinforce the ones previously obtained. As can be observed, the distance approach is more profitable than the copula method, at least at 10%, for Top 20 and Top 35 pairs on committed capital. On the other hand, the copula approach significantly outperforms the distance strategy in terms of mean excess returns and in risk-adjusted returns when the number of tradeable signals is comparable (Top 5 pairs) on fully invested weighting structure.

### 2.3.2 Trading statistics

Table 2.2 reports trading statistics. Panel A, B and C report results for Top 5, Top 20 and Top 35 pairs, respectively. The average price deviation trigger for opening pairs is listed in the first row of each panel. We can observe that, in average, we initiate the positions before when using the distance approach. The positions are initiated when prices have diverged by 5.94%, 6.81%, and 7.29% for Top 5, Top 20, and Top 35 pairs, respectively. Similar to Gatev, Goetzmann and Rouwenhorst (2006), the trigger spread increases with the number of pairs for all approaches[14].

The table reveals that the average number of pairs traded per six-month period is only equiparable among the strategies for Top 5 pairs. For Top 20 and Top 35 pairs the total number of pairs opened is about 75% and 138% when starting positions based on the distance approach. This suggests that a two standard deviation trigger as opening criterion is less conservative than the opening threshold suggested by Rad, Low and Faff (2015) using the cumulative mispriced indexed $M1$ and $M2$. Thus, the distance approach will be able to identify more trading opportunities to profit making the comparison less meaningful, although in practice the benefits are partly offset by the trading costs.

Finally, note that each pair is held open, in average, by 50.7 and 37.7 trading days (2.4 and 1.8 months) under the distance and copula approaches, respectively, for Top 5 pairs, which indicates that they are a medium-term investment under these strategies.

---

[14]citet*ggr06 explains that the standard deviation of the prices increases as the proximity of the securities in price space decreases, thus increasing the trigger spreads.

**Table 2.2:** 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 six-month period | 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: Top 20* | |
| Average price deviation trigger for opening pairs | 0.0681 | 0.0821 |
| Total number of pairs opened | 1312 | 749 |
| Average number of pairs traded per six-month period | 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 |
| | *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.
Source: Author's own elaboration (2017).

### 2.3.3 Regression on Fama-French asset pricing factors

In one attempt to understand the economic drivers behind our data and to evaluate whether pairs trading profitability is a compensation for risk, we regress daily excess returns onto various risk factors: daily Fama and French (2015)'s five research factors, the excess return on a broad market portfolio, $(R_m - R_f)$, the difference between the return on a portfolio of small stocks and the return on a portfolio of large stocks ($SMB$, small minus big), the difference between the return on a portfolio of high book-to-market stocks and the return on a portfolio of low book-to-market stocks ($HML$, high minus low), the difference between the return of the most profitable stocks and the return of the least profitable stocks ($RMW$, robust minus weak), the difference between the return of stocks that invest conservatively and the return of stocks that invest aggressively ($CMA$, conservative minus aggressive) plus momentum (Mom), short-term reversal (SRev), and long-term reversal (Rev) factors, *i.e.*,

$$
R_{i,d} - R_{f,d} = \alpha_i + \beta_i \left( R_{m,d} - R_{f,d} \right) + s_i SMB_d + h_i HML_d + r_i RMW_d
$$
$$
+ c_i CMA_d + m_i Mom_d + v_i SRev_d + l_i LRev_d + \varepsilon_{i,d}.
$$

(2.14)

All the data used to fit the above regressions are described in and obtained from Kenneth French's data library[15]. We select the model in terms of an approximation to the mean squared prediction error using Bayesian Information Criterion (BIC) (**Schwarz1978**). Based on this variable selection procedure we remove the short-term reversal factor from the model.

The main purpose of these regressions is to estimate the intercept alpha - the average excess return not explained by these factors. The standard errors have been adjusted for heteroscedasticity and autocorrelation by using Newey-West adjustment with six lags.

Tables 2.3, A.1 and A.2 report the coefficients and corresponding Newey-West t-statistics of regressing daily portfolio return series onto Fama and French (2015)'s five research factors plus momentum and long-term reversal factors for each of the strategies from 1991/2-2015, after transaction costs, for Top 5, Top 20 and Top 35 pairs, respectively. For each table, Section 1 lists the Return on Committed Capital and Section 2 on Fully Invested Capital. Panel A provides results after transaction costs and Panel B before transaction costs.

Table 2.3 shows the results for Top 5 pairs. It is clear that the mixed copula approach produces larger alphas than the distance method for both weighting schemes, especially on fully invested capital. It should also be noted that the alphas provided by copula and distance strategies are significant at 1% and 10%, respectively, after accouting for all the previously mentioned factors. Thus, the results show that the profits are not fully explained by the other factors.

From Table 2.3 one could also observe that the market premium slopes are larger for the distance method, especially on committed capital, and significant at 1% for both strategies. The SMB coefficients are only significant at 10% on committed capital for the mixed copula model, whereas the value effect premium (HML factor) slopes are significant at 5% for the distance approach. In addition, we find that returns are negatively correlated with momentum loadings for both methods at 1% on committed capital, although a larger negative effect on the performance is found when we use the 2.0 standard deviation threshold. We also find that excess returns are driven by the momentum factor on fully invested capital, although only significant at 5% for the copula approach. Furthermore, the exposures of pairs trading return strategies to the reversal factor are negative for both weighting schemes and significant at 5% for the distance approach on committed and fully invested capital, while it is only significant to the less conservative approach for the mixed copula strategy. The correlation of the excess returns with other traditional

---

[15]http://http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

equity risk premia factors (RMA and CMW) is nearly zero. Finally, It should be noted that the results show that the exposures to the various sources of systematic profile risk provide a low explanation of the average excess returns for any strategy as measured by the R-square and adjusted R-squared, particularly for the copula-based pairs strategy, indicating that the method is nearly factor-neutral over the whole sample period.

The regression on asset pricing factors for Top 5 pairs strengthen the patterns found in the Figure 2.3, indicating that the mixed copula strategy is able to produce relatively economically larger profits after costs than the distance approach when the number of tradeable signals is equiparable. Tables A.1 and A.2 provided in Appendix report the regressions for Top 20 and Top 35 pairs. As expected, one could observe that the results are in agreement the patterns found in the center and bottom of Figure 2.3.

31

**Table 2.3:** Systematic risk of Top 5 pairs: Fama and French(2016)'s five factors plus Momentum and Long-Term Reversal.

| Strategy | Intercept | t-stat | Rm-Rf | t-stat | SMB | t-stat | HML | t-stat | RMW | t-stat | CMA | t-stat | Mom | t-stat | LRev | t-stat | $R^2$ | $R^2_{adj}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Section 1: Return on Committed Capital** | | | | | | | | | | | | | | | | | | |
| Distance | 0.0116 | 1.8964* | 0.0422 | 4.2196*** | -0.0152 | -0.7099 | 0.0489 | 2.0508** | 0.0064 | 0.2537 | -0.0058 | -0.1822 | -0.0513 | -4.8063*** | -0.0433 | -1.9673** | 0.0282 | 0.0271 |
| Mixed Copula | 0.0165 | 3.5469*** | 0.0247 | 3.6774*** | -0.0207 | -1.8296* | 0.0186 | 1.2035 | -0.0166 | -0.9918 | 0.0131 | 0.6299 | -0.0258 | -2.9972*** | -0.0274 | -1.5702 | 0.0154 | 0.0143 |
| **Section 2: Return on Fully Invested Capital** | | | | | | | | | | | | | | | | | | |
| Distance | 0.0187 | 1.7536* | 0.0793 | 4.8800*** | -0.0138 | -0.4545 | 0.0774 | 2.2184** | 0.0319 | 0.7571 | 0.0030 | 0.0525 | -0.0780 | -4.3007*** | -0.0755 | -1.9656** | 0.0250 | 0.0239 |
| Mixed Copula | 0.0466 | 4.1685*** | 0.0702 | 3.5090*** | -0.0401 | -1.4507 | 0.0717 | 1.6355 | -0.0253 | -0.5956 | 0.0414 | 0.7471 | -0.0391 | -2.1901** | -0.1069 | -2.0807** | 0.0179 | 0.0168 |

*Note:* This table shows results of regressing daily portfolio return series onto Fama and French(2016)'s five factors factors plus momentum and long-term reversal over July 1991 and December 2015 (6173 observations). Section 1 shows the Return on Committed Capital and Section 2 on Fully Invested Capital after transaction costs. Pairs are formed based on the smallest sum of squared deviations. The t-statistics are computed using Newey-West standard errors with six lags.
***, **, * significant at 1%, 5% and 10% levels, respectively.
Source: Author's own elaboration (2017).

### 2.3.4 Sub-period analysis

We split the full sample period into five sub-periods: (1) July 1991 to December 1995, (2) January 1996 to December 2000, (3) January 2001 to December 2005, (4) January 2006 to December 2010, and (5) January 2010 to December 2015. The third sub-period corresponds to the bear market that comprises the dotcom crisis and the September 11th terrorist attack, whereas the fourth sub-period corresponds to the subprime mortgage financial crisis period.

Figures 2.6 and 2.7 show the profitability and risk-adjusted patterns of both strategies for Top 5 (top), Top 20 (center) and Top 35 (bottom) pairs after costs, respectively, for each sub-period on commited capital (left) and fully invested capital (right).

Overall, the mixed copula strategy achieves a favorable out-of-sample performance relative to the distance approach in the second and third subperiods (1996-2000 and 2001-2005) and after the subprime mortgage crisis, while the distance method delivers a significant better performance in the first (1991-1995) and fourth subperiods when the number of trades (Top 5 pairs) are similar on committed capital. It should be noted that the distance strategy shows a more constant pattern during the main volatile periods, and a very poor performance (negative mean excess returns) after the financial crisis. For fully invested weighting scheme the results are consistent with we have found in the full period analysis (see the right panels in Figure 2.3).

The results for Top 20 and Top 35 pairs are in agreement with what we expected from previous analyses. The main difference is the performance of the strategies in the second subperiod on committed capital.

The results suggest, since neither strategy is consistently superior in all subperiods, at least on committed capital, that the strategies may be combined to yield a viable long-short strategy.

**Figure 2.6: Average excess returns of pairs trading strategies after costs for each sub-period**
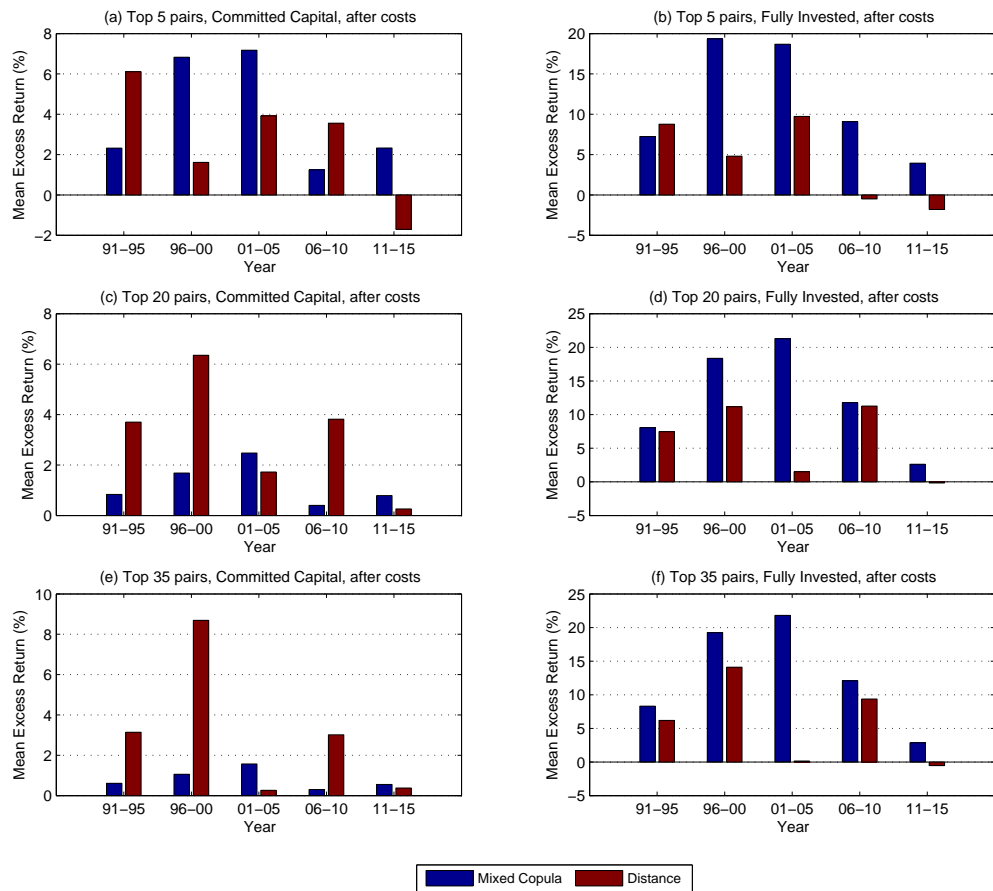


Source: Author's own elaboration (2017).

This figure shows how the 5-year rolling window Sharpe ratio densities evolve from July 1996 to December 2015 for each of the strategies using the method of Sheather and Jones (1991) to select the bandwidths.

**Figure 2.7: Sharpe Ratio of pairs trading strategies after costs for each sub-period**
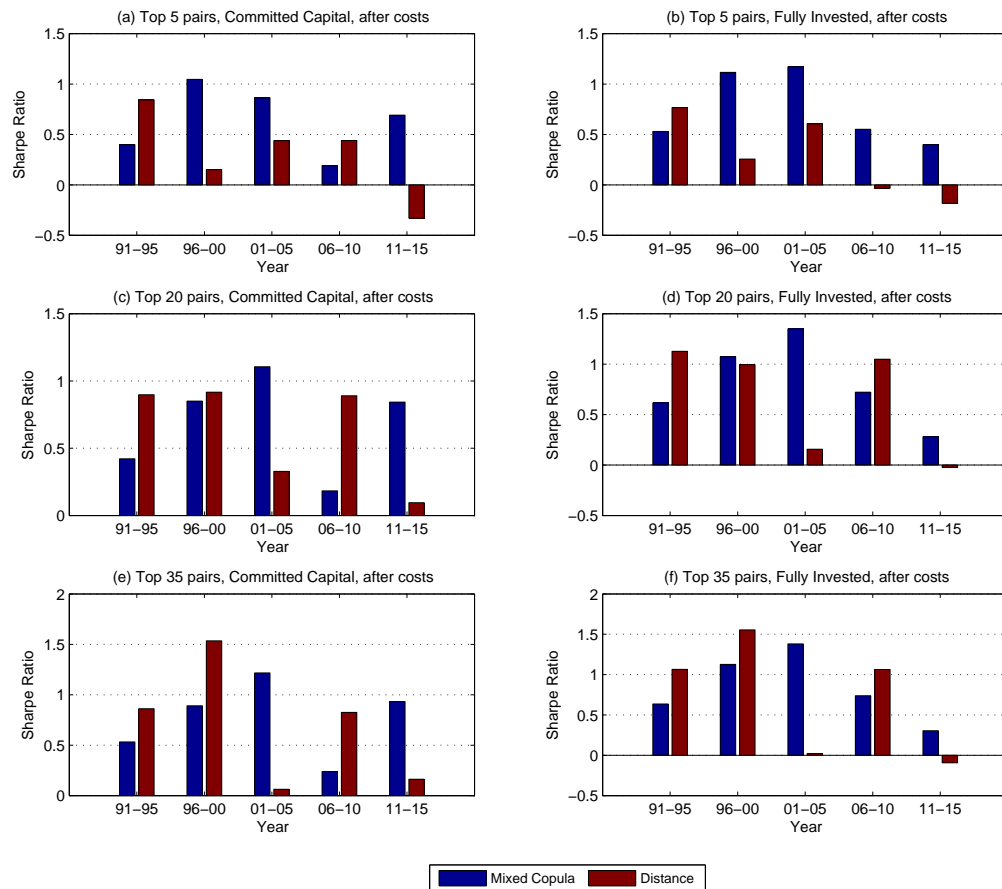


Source: Author's own elaboration (2017).

This figure shows how the 5-year rolling window Sharpe ratio densities evolve from July 1996 to December 2015 for each of the strategies using the method of Sheather and Jones (1991) to select the bandwidths.

## 2.4 Concluding Remarks

The main objective of the paper is to compare the copula and distance estimation procedures to understand better the factors that affect the profitability of the strategies and to find if the approaches can produce sustainable alpha.

The comparison made by means of an empirical investigation suggests that the copula strategy has a superior performance than the distance approach when the number of trades is comparable.

The main findings are summarized below.

1. By capturing linear/nonlinear associations and covering a wider range of possible dependencies structures, the mixed copula strategy is able to generate a higher mean excess return than the distance method when the number of trading signals is equiparable.

2. The mixed copula approach delivers economically larger alphas than the distance method for both weighting schemes, especially on fully invested capital, for Top 5 pairs. It should also be noted that the alphas provided by copula and distance strategies are significant at 1% and 10%, respectively, after accouting for several asset pricing factors such as momentum, liquidity, profitability and investment. Thus, the results show that the profits are not fully explained by the other factors.

3. The results suggest, since neither strategy is consistently superior in all subperiods, at least on committed capital, that the strategies may be combined to yield a viable long-short strategy.

Finally, we provide some suggestions for future research to improve our findings. First, we could create similar copula-based arbitrage for triplets to increase information dependency information and measure relative pricing more comprehensively. Second, since daily returns only capture close-to-close volatility, leaving much to be said regarding the actual volatility of the stocks intra-day, the inclusion of realized measures of volatility using higher frequency data may indeed prove beneficial. Third, we could borrow some ideas from statistical/machine learning and perform a model validation during the formation period rather than simply selecting the pairs. After choosing the pairs by whatever criterion has been adopted, we could check its predictive power by splitting data into two subsets. The idea is to select and fit the model based entirely on one part of the data, and then test it on the remaining part. Fourth, we could improve the method of pairs selection before using the copula approach. Since copulas capture better nonlinear dependence structures, we suspect that if we use a nonlinear association measurement such as the randomized dependency coefficient (LOPEZ-PAZ; HENNIG; SCHÖLKOPF, 2013), or a procedure based on data mining tools as random forest (GIOVANNI DE LUCA; ZUCCOLOTTO, 2010) to select the pairs, the performance of the copula method may be enhanced.

## REFERENCES

AIT-SAHALIA, Yacine; BRANDT, Michael W. Variable selection for portfolio choice. **The Journal of Finance**, v. 56.4, p. 1297–1351, 2001.

ANDREOU, Elena; PITTIS, Nikitas; SPANOS, Aris. On modelling speculative prices: the empirical literature. **Journal of Economic Surveys**, v. 15.2, p. 187–220, 2001.

ANE, Thierry; KHAROUBI, Cécile. Dependence Structure and Risk Measure. **The journal of business**, v. 76, n. 3, p. 411–438, 2003.

BREEDEN, Douglas T. An intertemporal asset pricing model with stochastic consumption and investment opportunities. **Journal of financial Economics**, v. 7.3, p. 265–296, 1979.

BROUSSARD, John Paul; VAIHEKOSKI, Mika. Profitability of pairs trading strategy in an illiquid market with multiple share classes. **Journal of International Financial Markets, Institutions and Money**, v. 22.5, p. 1188–1201, 2012.

CALDEIRA, João; MOURA, Guilherme V. Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy. **Available at SSRN**, v. 2196391, 2013.

CAMPBELL, John Y; LO, Andrew Wen-Chuan; MACKINLAY, Archie Craig. **The econometrics of financial markets**. [S.l.]: princeton University press, 1997.

CHEN, Joseph; HONG, Harrison; STEIN, Jeremy C. Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. **Journal of Financial Economics**, v. 61.3, p. 345–381, 2001.

CHERUBINI, Umberto; LUCIANO, Elisa; VECCHIATO, Walter. **Copula methods in finance**. Chichester, England: John Wiley & Sons, 2004.

CONT, Rama. Empirical properties of asset returns: stylized facts and statistical issues. Taylor & Francis, 2001.

DAVISON, Anthony Christopher; HINKLEY, David Victor. **Bootstrap methods and their application**. Cambridge, UK: Cambridge university press, 1997. v. 1.

DO, Binh; FAFF, Robert. Are pairs trading profits robust to trading costs? **Journal of Financial Research**, v. 35.2, p. 261–287, 2012.

_____. Does simple pairs trading still work? **Financial Analysts Journal**, v. 66.4, p. 83–95, 2010.

FAMA, Eugene F.; FRENCH, Kenneth R. A five-factor asset pricing model. **Journal of Financial Economics**, v. 116.1, p. 1–22, 2015.

_____. Common risk factors in the returns on stocks and bonds. **Journal of financial economics**, v. 33.1, p. 3–56, 1993.

FRAZZINI, Andrea; ISRAEL, Ronen; MOSKOWITZ, Tobias J. Trading Costs of Asset Pricing Anomalies. **Chicago Booth Research Paper, Fama-Miller Working Paper**, 2015.

FRENCH, Kenneth R. Presidential address: The cost of active investing. **The Journal of Finance**, Wiley Online Library, v. 63, n. 4, p. 1537–1573, 2008.

GATEV, Evan; GOETZMANN, William N.; ROUWENHORST, K. Geert. Pairs Trading: Performance of a Relative-Value Arbitrage Rule. **Review of Financial Studies**, v. 19, n. 3, p. 797–827, 2006.

GIOVANNI DE LUCA, Giorgia Rivieccio; ZUCCOLOTTO, Paola. Combining random forest and copula functions: a heuristic approach for selecting assets from a financial crisis perspective. **Intelligent Systems in Accounting, Finance and Management**, v. 17.2, p. 91–109, 2010.

HAFNER, Christian M. Estimating high-frequency foreign exchange rate volatility with nonparametric ARCH models. **Journal of Statistical Planning and Inference**, v. 68.2, p. 247–269, 1998.

KAKOURIS, Iakovos; RUSTEM, Berç. Robust portfolio optimization with copulas. **European Journal of Operational Research**, Elsevier, v. 235, n. 1, p. 28–37, 2014.

LEDOIT, Oliver; WOLF, Michael. Robust performance hypothesis testing with the Sharpe ratio. **Journal of Empirical Finance**, v. 15.5, p. 850–859, 2008.

LIEW, Rong Qi; WU, Yuan. Pairs trading: A copula approach. **Journal of Derivatives & Hedge Funds**, Springer, v. 19, n. 1, p. 12–30, 2013.

LINTNER, John. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. **The review of economics and statistics :** p. 13–37, 1965.

LOPEZ-PAZ, David; HENNIG, Philipp; SCHÖLKOPF, Bernhard. The Randomized Dependence Coefficient. In: ADVANCES in Neural Information Processing Systems. [S.l.: s.n.], 2013. p. 1–9.

MARKOWITZ, Harry. Portfolio selection. **The journal of finance**, Wiley Online Library, v. 7, n. 1, p. 77–91, 1952.

MCNEIL, Alexander J; FREY, Rüdiger; EMBRECHTS, Paul. **Quantitative risk management: Concepts, techniques and tools**. [S.l.]: Princeton university press, 2015.

NELSEN, R. B. **An introduction to copulas, 2nd**. New York: Springer Science Business Media, 2006.

NEWEY, Whitney K; WEST, Kenneth D. Hypothesis testing with efficient method of moments estimation. **International Economic Review**, JSTOR, p. 777–787, 1987.

PATTON, Andrew J. Modelling time-varying exchange rate dependence using the conditional copula. UCSD Discussion Paper, 2001.

PERLIN, Marcelo Scherer. Evaluation of pairs-trading strategy at the Brazilian financial market. **Journal of Derivatives & Hedge Funds**, v. 15.2, p. 122–136, 2009.

POLITIS, Dimitris N.; ROMANO, Joseph P. The stationary bootstrap. **Journal of the American Statistical association**, v. 89.428, p. 1303–1313, 1994.

POLITIS, Dimitris N.; WHITE, Halbert. Automatic block-length selection for the dependent bootstrap. **Econometric Reviews**, v. 23.1, p. 53–70, 2004.

RAD, Hossein; LOW, Rand Kwong Yew; FAFF, Robert W. The profitability of pairs trading strategies: Distance, cointegration, and copula methods. **Cointegration, and Copula Methods (June 3, 2015)**, 2015.

ROCKAFELLAR, R Tyrrell; URYASEV, Stanislav. Optimization of conditional value-at-risk. **Journal of risk**, v. 2, p. 21–42, 2000.

SHARPE, William F. Capital asset prices: A theory of market equilibrium under conditions of risk. **The journal of finance**, v. 19.3, p. 425–442, 1964.

SHEATHER, Simon J; JONES, Michael C. A reliable data-based bandwidth selection method for kernel density estimation. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 683–690, 1991.

SKLAR, M. **Fonctions de répartition à n dimensions et leurs marges**. Paris, France: Université Paris 8, 1959.

TAUCHEN, George. Notes on financial econometrics. **Journal of Econometrics**, v. 100.1, p. 57–64, 2001.

XIE, Wenjun; LIEW, Rong Qi, et al. Pairs Trading with Copulas. **The Journal of Trading**, v. 11, n. 3, p. 41–52, 2016.

XIE, Wenjun; YUAN, Wu. Copula-Based Pairs Trading Strategy. **Asian Finance Association (AsFA)**, v. 2013, 2013.

ZHU, Shushang; FUKUSHIMA, Masao. Worst-case conditional value-at-risk with application to robust portfolio management. **Operations research**, INFORMS, v. 57, n. 5, p. 1155–1168, 2009.

**Appendix A - Regressions on asset pricing factors**

This appendix contains the regressions on asset pricing factors for Top 20 and Top 35 pairs for both strategies.

**Table A.1:** Systematic risk of Top 20 pairs: Fama and French (2016)'s five factors plus Momentum and Long-Term Reversal.

| Strategy | Intercept | t-stat | Rm-Rf | t-stat | SMB | t-stat | HML | t-stat | RMW | t-stat | CMA | t-stat | Mom | t-stat | LRev | t-stat | $R^2$ | $R^2_{adj}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | **Section 1: Return on Committed Capital** | | | | | | | | | |
| Distance | 0.0131 | 3.4717*** | 0.0258 | 5.0243*** | -0.0078 | -0.6229 | 0.0045 | 0.3818 | -0.0136 | -0.8488 | 0.0325 | 1.8151* | -0.0342 | -4.9021*** | -0.0330 | -2.3743** | 0.0277 | 0.0266 |
| Mixed Copula | 0.0050 | 3.5317*** | 0.0071 | 3.5187*** | -0.0028 | -0.7514 | 0.0010 | 0.2195 | -0.0024 | -0.4938 | 0.0036 | 0.5762 | -0.0061 | -2.1883** | -0.0058 | -1.1018 | 0.0091 | 0.0079 |
| | | | | | | | | **Section 2: Return on Fully Invested Capital** | | | | | | | | | |
| Distance | 0.0255 | 3.6809*** | 0.0485 | 5.1809*** | -0.0128 | -0.5533 | 0.0210 | 0.8965 | -0.0099 | -0.3448 | 0.0604 | 1.7816* | -0.0710 | -4.9449*** | -0.0553 | -2.0788** | 0.0305 | 0.0294 |
| Mixed Copula | 0.0490 | 4.4947*** | 0.0674 | 3.3944*** | -0.0419 | -1.5363 | 0.0644 | 1.4682 | -0.0295 | -0.7013 | 0.0410 | 0.7465 | -0.0233 | -1.3647 | -0.1186 | -2.3591** | 0.0163 | 0.0152 |

*Note:* This table shows results of regressing daily portfolio return series onto Fama and French(2016)'s five factors factors plus momentum and long-term reversal over July 1991 and December 2015 (6173 observations). Section 1 shows the Return on Committed Capital and Section 2 on Fully Invested Capital after transaction costs. Pairs are formed based on the smallest sum of squared deviations. The t-statistics are computed using Newey-West standard errors with six lags.
***, **, * significant at 1%, 5% and 10% levels, respectively.
Source: Author's own elaboration (2017).

**Table A.2:** Systematic risk of Top 35 pairs: Fama and French (2016)'s five factors plus Momentum and Long-Term Reversal.

| Strategy | Intercept | t-stat | Rm-Rf | t-stat | SMB | t-stat | HML | t-stat | RMW | t-stat | CMA | t-stat | Mom | t-stat | LRev | t-stat | $R^2$ | $R^2_{adj}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Section 1: Return on Committed Capital** *Panel A: Minimum-Distance based on SSD* | | | | | | | | | | | | | | | | | | |
| Distance | 0.0122 | 3.9305*** | 0.0279 | 5.5504*** | -0.0046 | -0.5391 | -0.0048 | -0.4967 | 0.0050 | 0.4059 | 0.0283 | 1.9636** | -0.0322 | -5.3765*** | -0.0252 | -1.9921** | 0.0343 | 0.0332 |
| Mixed Copula | 0.0033 | 3.9574*** | 0.0044 | 3.8052*** | -0.0015 | -0.6902 | -0.0001 | -0.0400 | -0.0020 | -0.6919 | 0.0035 | 0.9766 | -0.0035 | -2.1733** | -0.0035 | -1.1366 | 0.0094 | 0.0082 |
| **Section 2: Return on Fully Invested Capital** *Panel A: Minimum-Distance based on SSD* | | | | | | | | | | | | | | | | | | |
| Distance | 0.0230 | 4.0660*** | 0.0525 | 5.3566*** | -0.0142 | -0.9048 | -0.0062 | -0.3052 | 0.0112 | 0.4935 | 0.0682 | 2.5580** | -0.0630 | -5.4671*** | -0.0515 | -2.0365** | 0.0372 | 0.0361 |
| Mixed Copula | 0.0504 | 4.6266*** | 0.0688 | 3.4792*** | -0.0408 | -1.5037 | 0.0615 | 1.3997 | -0.0335 | -0.7938 | 0.0476 | 0.8707 | -0.0233 | -1.3632 | -0.1209 | -2.4150** | 0.0167 | 0.0155 |

*Note:* This table shows results of regressing daily portfolio return series onto Fama and French(2016)'s five factors factors plus momentum and long-term reversal over July 1991 and December 2015 (6173 observations). Section 1 shows the Return on Committed Capital and Section 2 on Fully Invested Capital after transaction costs. Pairs are formed based on the smallest sum of squared deviations. The t-statistics are computed using Newey-West standard errors with six lags.

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

Source: Author's own elaboration (2017).

**Chapter 3**

**Worst Case CVaR Portfolio Optimization with Multidimensional Copulas**

Fernando A. Boeira Sabino da Silva,[1] and Flavio A. Ziegelmann[2]

**Abstract.** Using data from the S&P 500 shares from 1990 to 2015, we measure the downside market risk by Conditional Value-at-Risk (CVaR) subject to return constraints following the approach of Rockafellar and Uryasev (2000) and Rockafellar and Uryasev (2002) and the extended framework of Kakouris and Rustem (2014) through the use of multidimensional mixed archimedean copulas. We implement a dynamic investing viable strategy where the portfolios are optimized using three different length of rolling calibration windows. The out-of-sample performance is evaluated and compared against two benchmarks; a multidimensional gaussian copula model and a constant mix portfolio. Our empirical analysis shows that the Mixed Copula-CVaR approach generates portfolios with better downside risk statistics for any rebalancing period and it is more profitable than the Gaussian Copula-CVaR and 1/N portfolios for daily and weekly rebalancing. To cope with the dimensionality problem we employ a similar approach to that used by Gatev, Goetzmann and Rouwenhorst (2006) to select a set of assets that are the most diversified, in some sense, to the S&P 500 index in the constituent set. We find that copula-based approaches offer better hedges against losses than the 1/N portfolio. The accuracy of VaR forecasts is determined by how well they minimize a capital requirement loss function (CR). In order to mitigate data-snooping problems we apply a test for superior predictive ability (SPA) proposed by Hansen (2005) to determine which model significantly minimizes this expected loss function. The test is implemented via stationary bootstrap of Politis and Romano (1994) using the automatic block-length selection of Politis and White (2004) and 10,000 bootstrap resamples. We find that the minimum average loss of the mixed Copula-CVaR approach is smaller than the average performance of the Gaussian Copula-CVaR.

**Keywords:** Asset Allocation. Capital Requirement Loss. Copula. Portfolio Selection. Risk Management. S&P 500. Superior Predictive Ability. Stationary Bootstrap. WCC-VaR.

**JEL Classifications.** C15; C52; C53; C61; C63; G11.

[1]Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 91509-900, Brazil, e-mail: fsabino@ufrgs.br

[2]Department of Statistics, Federal University of Rio Grande do Sul, Porto Alegre, RS 91509-900, Brazil, e-mail: flavioz@ufrgs.br

## 3.1   Introduction

The seminal article Portfolio Selection published by Markowitz (1952) introduces the foundation for modern portfolio theory (MPT) or mean-variance analysis. Markowitz considered the problem of an agent who wishes to find the maximum (expected) return for a given level of risk or minimize risk for a given level of return. He identified that, by diversifying a portfolio among investments that have different return patterns, investors can build such an efficient portfolio.

Quantile functions are commonly used for measuring the market risk of models with parameter uncertainty and variability. Portfolio optimization involving a mean-value-at-risk (mean-VaR) portfolio and the CVaR have been analyzed by Alexander and Baptista (2002) and Rockafellar and Uryasev (2000), respectively. Akin to the classical Markowitz portfolio, in these approaches we want to determine the weights that maximize the portfolio return for a specified VaR or CVaR at a given confidence level or minimize these quantiles for a given confidence level subject to a fixed portfolio return.

Artzner, Delbaen, Eber and Heath (1999) show that VaR has undesirable properties such as lack of sub-additivity and thus it is not a coherent measure. Furthermore, Uryasev and Rockafellar (2001) show that VaR may be a non-convex function with respect to portfolio weights, which can yield to multiple portfolio local solutions, but CVaR is coherent both for continuous and discrete distributions and it is a convex function. In addition, they show that an outright optimization with respect to CVaR is numerically difficult due to the dependence of the CVaR on VaR. However, Rockafellar and Uryasev (2000) show that VaR and CVaR can be computed simultaneously, introducing auxiliary risk measures, and it can be used in conjunction with scenario based optimization algorithms reducing the problem to a linear programming problem which allows us to optimize a portfolio with very large dimensions and stable numerical implementations.

A well-known problem of the Markowitz model is its sensitivity to the input parameters. In practice, the implementation of strategies based on the risk-return trade-off remains a fundamental challenge in many areas of financial management, since estimation errors of the expected returns of the assets and the covariance matrix of these returns can significantly affect the asset allocations weights, no longer leading to an efficient portfolio. This issue can be overcome by employing robust optimization and worst case techniques Zhu and Fukushima (2009) in which assumptions about the distribution of the random variable are relaxed, and thus, we obtain the optimal portfolio solution by optimizing over a prescribed feasible set and possible densities. Kakouris and Rustem (2014) show how copula-based models can be introduced in the Worst Case CVaR (WCVaR) framework. This approach is motivated by an investor's desire to be protected against the worst possible scenario.

In this paper, we employ a similar methodology to that of Kakouris and Rustem (2014) and investigate the advantage of such dependence structure through an empirical study. We evaluate the out-of-sample performance of the Worst Case Copula-CVaR and compare the relative performance of the strategy to the Gaussian Copula-CVaR, a naive 1/N (equally-weighted) and the S&P 500 index in the long term in terms of wealth accumulation and downside risk.

The main novel contribution of this paper is to select a diversified set of assets that can be useful during crises and tranquil periods, i.e., that somehow involves hedging of decisions to protect the investors against any market conditions and evaluate the approaches using a 50-dimensional archimedean and gaussian copula models without constructing hierarchical copulas.

Our data set consists of daily data of adjusted closing prices of all shares that belong to S&P 500 market index from July 2st, 1990 to December 31st, 2015. We obtain the adjusted closing prices from Bloomberg. The data set sample encompasses 6426 days and includes a total of 1100 stocks over the whole period. We consider only stocks that are listed throughout a 48 plus 6-month period in the analysis, i.e., close to 500 stocks in each trading period. Moreover, we consider alternative portfolio rebalancing frequencies and return constraints. Returns are calculated in excess of a risk-free asset.

To be effective in dealing with uncertainty, we select, among all listed assets in each

formation period, a set of 50 stocks based on the ranked sum of absolute deviations (the twenty-five largest and smallest) between the normalized daily closing prices deviations of the S&P 500 index and all shares, adjusting them by dividends, stock splits and other corporate actions. By selecting a diversified cluster of assets that can be useful during crises and tranquil periods we address the issue of asset allocation taking into account the purpose of risk diversification. These stocks are then evaluated over the next six months.

Our backtest results indicate that the Worst Case Copula-CVaR approach outperforms consistently the benchmark strategies in the long term in terms of downside risk measured by VaR, CVaR and capital requirement loss.

The remainder of the paper is structured as follows. In Section 2 we present a general review, notations and definitions of the CVaR and mean-variance optimization methodologies and extends them to our Worst Case framework through the use of appropriate allowable and uncertainty sets. The data we use is briefly discussed in Section 3. Section 4 summarizes the empirical results of the analysis, while Section 5 concludes and provide further ideas for research.

## 3.2 Specifications of the Models under Analysis

### 3.2.1 Conditional Value-at-Risk

Let $Y$ be a stochastic vector standing for market uncertainties and $F_Y$ be its distribution function, i.e., $F_Y(u) = P(Y \leq u)$. Let also $F_Y^{-1}(v) = \inf\{u : F_Y(u) \geq v\}$ be its right continuous inverse and assume that it has a probability density function represented by $p(y)$[3]. Define the $VaR_\beta$ as the $\beta$-quantile by

$$
\begin{aligned}
(3.1) \qquad VaR_\beta(Y) &= \arg\min\{\alpha \in \mathbb{R} : P(Y \leq \alpha) \geq \beta\} \\
&= F_Y^{-1}(\beta),
\end{aligned}
$$

and the $CVaR_\beta$ as the solution to the following optimization problem (PFLUG, 2000):

$$
CVaR_\beta(Y) = \inf\left\{\alpha \in \mathbb{R} : \alpha + \frac{1}{1-\beta}E[Y - \alpha]^+\right\}, \qquad (3.2)
$$

where $[t]^+ = \max(t, 0)$.

Uryasev and Rockafellar (1999) have shown that the $CVaR$ is the conditional expectation of $Y$, given that $Y \geq VaR_\beta$, i.e.,

$$
CVaR_\beta(Y) = \mathbb{E}(Y \,|\, Y \geq VaR_\beta(Y)). \qquad (3.3)
$$

Let $f(x, y)$ be a loss function depending upon a decision vector $x$ that belongs to any arbitrarily chosen subset $X \in \mathbb{R}^n$ and a random vector $y \in \mathbb{R}^m$. In a portfolio optimization problem, the decision vector $x$ can be a vector of portfolios' weights, $X$ be a set of feasible portfolios, subject to linear constraints[4] and $y$ a vector that stands for market variables that can affect the loss.

---

[3]This assumption can be relaxed (URYASEV, 2013).

[4]For example, we can assume a portfolio $X$ that does not allow short-selling positions (all $x_i \geq 0$, for $i = 1, ..., n$), that be fully invested, i.e., the total portfolio weights sum up to unity and that the expected return be greater than an arbitrary value $R$.

For each $x$, the loss function $f(x, y)$ is a random variable defined on a probability space $(\Omega, \mathcal{F}, P_{\mathcal{F}})$ having a probability distribution on $(\mathbb{R}, \mathcal{B}, P_{\mathcal{B}})$ induced by that of $y$, where $(\mathbb{R}, \mathcal{B}, P_{\mathcal{B}})$ stands for a Borel probability space including, therefore, the open and closed intervals in $\mathbb{R}$. The underlying distribution of $y \in \mathbb{R}^m$ is assumed to have a density $p(y)$ and let the probability of $f(x, y)$ not exceeding some threshold $\alpha$ be denoted by

$$F(x, \alpha) = \int_{f(x,y) \leq \alpha} p(y) \, dy, \tag{3.4}$$

where $F(x, \alpha)$ is the cumulative distribution function for the loss function $f(x, y)$, non-decreasing and right-continuous with respect to $\alpha$.

Using the previously defined notation (3.2) we can write the $CVaR$ function, at confidence level $\beta$, by

$$CVaR_\beta(x) = \frac{1}{1 - \beta} \int_{f(x,y) \geq VaR_\beta(x)} f(x, y) p(y) dy, \tag{3.5}$$

The optimization of $CVaR$ is difficult because of the presence of the $VaR$ in its definition (it requires the use of the nonlinear function max) in this infinite dimensional problem. The main contribution of Rockafellar and Uryasev (2000) was to define a simpler auxiliary function

$$F_\beta(x, \alpha) = \alpha + \frac{1}{1 - \beta} \int_{f(x,y) \geq \alpha} (f(x, y) - \alpha) p(y) dy, \tag{3.6}$$

which can be used instead of $CVaR$ directly, without need to compute $VaR$ first due to the following proposition (Pflug (2000)):

**Proposition 2** *The function $F_\beta(x, \alpha)$ is convex with respect to $\alpha$. In addition, minimizing $F_\beta(x, \alpha)$ with respect to $\alpha$ gives $CVaR$ and $VaR$ is a minimum point of this function with respect to $\alpha$, i.e.*

$$\min_{\alpha \in \mathbb{R}} F_\beta(x, \alpha) = F_\beta(x, VaR_\beta(x)) = CVaR_\beta(x) \tag{3.7}$$

Moreover, we can use $F_\beta(x, \alpha)$ to find the optimal weights, $CVaR$ and $VaR$ simultaneously over an allowable feasible set, i.e.,

$$\min_{x \in X} CVaR_\beta(x) = \min_{(x,\alpha) \in X \times \mathbb{R}} F_\beta(x, \alpha). \tag{3.8}$$

Pflug (2000) show that under quite general conditions $F_\beta(x, \alpha)$ is a smooth function. In addition, if $f(x, y)$ is convex with respect to $x$, then $F_\beta(x, \alpha)$ is also convex with respect to $x$. Hence, if the allowable set $X$ is also convex, we then have to solve a smooth convex optimization problem.

### 3.2.2   Optimization Problem

Assume that the analytical representation for the density $p(y)$ is not available, but we can approximate $F_\beta(x, \alpha)$ using $J$ scenarios, $y_j$, $j = 1, ..., J$ which are sampled from the density function $p(y)$. Then, we approximate

$$
\begin{aligned}
F_\beta(x, \alpha) &= \alpha + \frac{1}{1-\beta} \int_{f(x,y) \geq \alpha} (f(x,y) - \alpha) \, p(y) dy \\
&= \alpha + \frac{1}{1-\beta} \int_{y \in \mathbb{R}^m} (f(x,y) - \alpha)^+ \, p(y) dy
\end{aligned}
$$
(3.9)

by its discretized version

$$
\widetilde{F}_\beta(x, \alpha) = \alpha + \frac{1}{(1-\beta) J} \sum_{j=1}^{J} (f(x, y_j) - \alpha)^+ .
$$

Assuming that the feasible set $X$ and the loss function $f(x, y_j)$ are convex and the same for the loss function $f(x, y_j)$, we solve the following convex optimization problem:

$$
\min_{x \in X, \alpha \in \mathbb{R}} \widetilde{F}_\beta(x, \alpha) .
$$
(3.10)

In addition, if the loss function $f(x, y)$ is linear with respect to $x$, then the optimization problem (3.10) reduces to the following linear programming (LP) problem:

$$
\underset{x \in \mathbb{R}^n, z \in \mathbb{R}^J, \alpha \in \mathbb{R}}{\text{minimize}} \ \alpha + \frac{1}{(1-\beta) J} \sum_{j=1}^{J} z_j
$$
(3.11)

$$
\text{subject to} \quad x \in X,
$$
(3.12)

$$
z_j \geq f(x, y_j) - \alpha,
$$
(3.13)

$$
z_j \geq 0, \ j = 1, \dots, J.
$$
(3.14)

where $z_j$ are indicator variables. By solving the LP problem above we find the optimal decision vector, $x^*$, the optimal $VaR$, $\alpha^*$, and consequently the optimal $CVaR$, $\widetilde{F}_\beta(x = x^*, \alpha = \alpha^*)$.

Therefore, the optimization problem can be solved using algorithms that are capable of solving efficiently very large-scale problems with any distribution within reasonable time and high reliability as, for example, simplex or interior point methods.

In the next subsections we assume that there are $n$ risky assets and denote by $r$ their random (log)returns vector, i.e., $\mathbf{r} = (r_1, ..., r_n)^\top$, with expected returns $\mu = (\mu_1, ..., \mu_n)^\top$ and covariance matrix $\Sigma_{nxn}$. Let also $r_f$ represent the risk-free asset returns and the decision (portfolio's weights) vector by $\mathbf{w} = (w_1, ..., w_n)^\top$. Also assume that $\mathbf{w} \in \mathcal{W}$, where $\mathcal{W}$ is a feasible set and that the portfolio return loss function $f_L(\mathbf{w}, \mathbf{r})$ is a convex (linear) function given by

$$
f_L(\mathbf{w}, \mathbf{r}) = \mathbf{w}^\top \mathbf{r}.
$$

By definition, the portfolio return is the negative of the loss, i.e., $-\mathbf{w}^\top\mathbf{r}$.

### 3.2.3   Worst Case CVaR

Assume now that we do not have precise information about the distribution of the random vector $\mathbf{r}$, but that its density belongs to a family of distributions $\mathcal{P}$ defined by

$$\mathcal{P} = \left\{ \mathbf{r} \,|\, \mathbb{E}(\mathbf{r}) = \mu, Cov(\mathbf{r}) = \Sigma \right\}, \tag{3.15}$$

where $\Sigma$ is a positive definite matrix.

Instead of assuming the precise distribution of the random vector $r$, Zhu and Fukushima (2009) consider the case where the probability distribution $\pi$ is only known to belong to a certain set, say $\mathcal{P}$, defined the worst-case CVaR (WCVaR) as the CVaR when the worst-case probability distribution in the set $\mathcal{P}$ occurs.

**Definition 2** *Given a confidence level $\beta$, $\beta \in (0,1)$, the worst-case CVaR for fixed $w \in \mathcal{W}$ with respect to the uncertainty set $\mathcal{P}$ is defined as*

$$
\begin{aligned}
WCVaR_\beta\left(\mathbf{w}\right) &\equiv \sup_{\pi\in\mathcal{P}} CVaR_\beta\left(\mathbf{w}\right) \\
&= \sup_{\pi\in\mathcal{P}} \min_{\alpha\in\mathbb{R}} F_\beta\left(\mathbf{w},\alpha\right)
\end{aligned}
$$
(3.16)

Zhu and Fukushima (2009) further investigated the WCVaR risk measure with several structures of uncertainty in the underlying distribution. In particular, Zhu and Fukushima (2009) consider the case where the distribution of $r$ belong to a set of distributions consisting of all mixtures of some possible distribution scenarios, i.e.,

$$p\left(\cdot\right) \in \mathcal{P}_M \equiv \left\{ \sum_{i=1}^{d} \pi_i p^i\left(\cdot\right) : \sum_{i=1}^{d} \pi_i = 1, \ \pi_i \geq 0, \ i = 1,...,d \right\}, \tag{3.17}$$

where $p^i\left(\cdot\right)$ denotes the $j$-th likelihood distribution and define

$$
\begin{aligned}
G_\beta\left(\mathbf{w},\alpha,\pi\right) &= \alpha + \frac{1}{1-\beta}\int_{\mathbf{r}\in\mathbb{R}^n}\left(f(\mathbf{w},\mathbf{r})-\alpha\right)^+\sum_{i=1}^{d}\pi_i p^i(\mathbf{r})d\mathbf{r}, \ j=1,...,J \\
&= \sum_{i=1}^{d}\pi_i G_\beta^i\left(\mathbf{w},\alpha\right), \ i=1,...,d,
\end{aligned}
$$
(3.18)

where

$$G_\beta^i\left(\mathbf{w},\alpha\right) = \alpha + \frac{1}{1-\beta}\int_{\mathbf{r}\in\mathbb{R}^n}\left(f(\mathbf{w},\mathbf{r})-\alpha\right)^+ p^i(\mathbf{r})d\mathbf{r}, \ i=1,...,d. \tag{3.19}$$

By theorem 1 of Zhu and Fukushima (2009) we can state that for each fixed $\mathbf{w} \in \mathcal{W}$,

$$WCVaR_\beta(\mathbf{w}) = \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} G_\beta(\mathbf{w}, \alpha, \pi)$$

(3.20)
$$= \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} \sum_{i=1}^{d} \pi_j G_\beta^i(\mathbf{w}, \alpha).$$

Thus, minimizing the worst-case CVaR over $w \in \mathcal{W}$ is equivalent to the following min-min-sup optimization problem:

$$\min_{\mathbf{w} \in \mathcal{W}} WCVaR_\beta(\mathbf{w}) = \min_{\mathbf{w} \in \mathcal{W}} \min_{\alpha \in \mathbb{R}} \sup_{\pi \in \Pi} G_\beta(\mathbf{w}, \alpha, \pi). \qquad (3.21)$$

Zhu and Fukushima (2009) also proved that WCVaR is a coherent risk measure and it clearly satisfies $WCVaR_\beta(\mathbf{w}) \geq CVaR_\beta(\mathbf{w}) \geq VaR_\beta(\mathbf{w})$. Thus, $WCVaR_\beta(\mathbf{w})$ can be effectively used as a risk measure.

In the next subsections we provide a brief review of copulas and consider the augmented framework of Kakouris and Rustem (2014) to compute WCVaR through the use of multidimensional mixture archimedean copulas.

### 3.2.4   Copulas

Copulas are often defined as multivariate distribution functions whose marginals are uniformly distributed on $[0, 1]$. In other words, a copula $C$ is a function such that

$$C(u_1, ..., u_n) = P(U_1 \leq u_1, ..., U_n \leq u_n), \qquad (3.22)$$

where $U_i \sim U[0, 1]$ and $u_i$ are realizations of $U_i$, $i = 1, ..., n$. The margins $u_i$ can be replaced by $F_i(x_i)$, where $x_i$, $i = 1, ..., n$ is a realization of a (continuous) random variable, since they both belong to the domain $[0, 1]$ and are uniformly distributed by its probability integral transform (note that $P(F(x) \leq u) = P(x \leq F^{-1}(u)) = F(F^{-1}(u)) = u$). Therefore, copulas can be used to model the dependence structure and margins separately, and therefore provide more flexibility.

Formally, we can define a copula function $C$ as follows.

**Definition 3** *An $n$-dimensional copula (or simply $n$-copula) is a function $C$ with domain* $[0, 1]^n$, *such that:*
    1. C is grounded an $n$-increasing;
    2. C has marginal distributions $C_k$, $k = 1, ..., n$, where $C_k(\mathbf{u}) = u$ for every $\mathbf{u} = (u_1, ..., u_n)$ in $[0, 1]^n$.

Equivalently, an $n$-copula is a function

$$C : [0, 1]^n \rightarrow [0, 1]$$

with the following properties:
    (i) (grounded) For all $\mathbf{u}$ in $[0, 1]^n$, $C(\mathbf{u}) = 0$, if at least one coordinate of $\mathbf{u}$ is 0 and $C(\mathbf{u}) = u_k$, if all the coordinates of $\mathbf{u}$ are 1 except $u_k$;

(ii) ($n$-increasing) For all **a** and **b** in $[0,1]^n$ such that $a_i \leq b_i$, for every $i$, $V_C\left([a,b]\right) \geq 0$, where $V_c$ is called $C-$volume.

One of the main results of the theory of copulas is Sklar's Theorem Sklar (1959).

**Theorem 2** *(Sklar's Theorem) Let $X_1, ..., X_n$ be random variables with distribution functions $F_1, ..., F_n$, respectively. Then, there exists an n-copula $C$ such that,*

$$F\left(x_1, ..., x_n\right) = C\left(F_1\left(x_1\right), ..., F_n\left(x_n\right)\right), \tag{3.23}$$

for all $\mathbf{x} = \left(x_1, ..., x_n\right) \in \mathbb{R}^n$. If $F_1, ..., F_n$ are all continuous, then the function $C$ is unique; otherwise $C$ is determined only on $\operatorname{Im} F_1 \times ... \times \operatorname{Im} F_n$. Conversely, if $C$ is an $n$-copula and $F_1, ..., F_n$ are distribution functions, then the function $F$ defined above is an $n-$dimensional distribution function with marginals $F_1, ..., F_n$.

**Corollary 2.1** *Let $F$ be an n-dimensional distribution function with continuous marginals $F_1, ..., F_n$, and copula $C$. Therefore, for any $\mathbf{u} = \left(u_1, ..., u_n\right)$ in $[0,1]^n$,*

$$C\left(u_1, ..., u_n\right) = F\left(F_1^{-1}\left(u_1\right), ..., F_n^{-1}\left(u_n\right)\right), \tag{3.24}$$

where $F_i^{-1}$, $i = 1, ..., n$ are the quasi-inverses of the marginals.

Using the Sklar's theorem and its corollary we can derive an important relation between the probability density functions and copulas. Let $f$ be a joint density function (of the $n-$dimensional distribution function $F$) and $f_1, ..., f_n$ univariate density functions of the margins $F_1, ..., F_n$. Assuming that $F\left(\cdot\right)$ and $C\left(\cdot\right)$ are differentiable, by (3.23) and (3.24)

$$(3.25) \qquad \frac{\partial^n F\left(x_1, ..., x_n\right)}{\partial x_1 ... \partial x_n} \equiv f\left(x_1, ..., x_n\right) = \frac{\partial^n C\left(F_1\left(x_1\right), ..., F_n\left(x_n\right)\right)}{\partial x_1 ... \partial x_n}$$

$$(3.26) \qquad = c\left(u_1, ..., u_n\right) \prod_{i=1}^{n} f_i\left(x_i\right).$$

From a modelling perspective, Sklar's Theorem allow us to separate the modeling of the marginals $F_i\left(x_i\right)$ from the dependence structure, represented in $C$. The copula probability density function

$$c\left(u_1, ..., u_n\right) = \frac{f\left(x_1, ..., x_n\right)}{\prod_{i=1}^{n} f_i\left(x_i\right)} \tag{3.27}$$

is the ratio of the joint probability function to what it would have been under independence. Thus, we can interpret the copula as the adjustment that we need to make to convert the independence probability density function into the multivariate density function. In other words, copulas decompose the joint probability density function from its margins. Now we can estimate the multivariate distribution in two parts: (i) finding the marginal distributions; (ii) finding the dependency between the filtered data from (i).

Thereafter, copulas accommodate various forms of dependence through suitable choice of the copula "correlation matrix" since they conveniently separate marginals from dependence component. The methodology allows one to derive joint distributions from marginals, even when these are not normally distributed. In fact, copulas allow the marginal distributions to be modeled independently from each other, and no assumption on the joint behavior of the marginals is required, which provides a great deal of flexibility in modeling joint distributions.

The reader interested on the history of copulas is referred to Schweizer and Sklar (2011) and Nelsen (2006). For an extensive list of parametric copula families see Joe

(1997) and Nelsen (2006) and references therein. After late 1990's the theory and application of copulas grew enormously. Cherubini, Luciano and Vecchiato (2004) accounts for developments in finance and stochastic processes.

Here we focus in a special class of copulas called Archimedean. An Archimedean copula has the form

$$C\left(u_1, ..., u_n\right) = \psi\left(\psi^{-1}\left(u_1\right) + \psi^{-1}\left(u_2\right) + ... + \psi^{-1}\left(u_n\right)\right), \tag{3.28}$$

for an appropriate generator $\psi\left(\cdot\right)$, where $\psi : [0, \infty] \to [0, 1]$ and satisfies (i) $\psi(0) = 1$ and $\psi(\infty) = 0$; (ii) $\psi$ is $d$-monotone, i.e., $(-1)^k \mathrm{d}^k \psi\left(s\right) \mathrm{d}s^k \geq 0$ for $k \in \{0, ..., d-2\}$ and $(-1)^{d-2} \mathrm{d}^{d-2} \psi\left(s\right) \mathrm{d}s^{d-2}$ is decreasing and convex. Most of the copulas in this class has a closed form. Moreover, each member has a single parameter that controls the degree of dependence, which allow modeling dependence in arbitrarily high dimensions with only one parameter.

The concept of tail dependence can be especially useful when dealing with co-movements of assets. Studies on financial markets show that financial assets have asymmetric distributions and heavy tails. Often elliptical copulas, like gaussian and t-student copulas are used in Finance. However, these copulas suffer from an absent or symmetric lower and upper tail dependence, respectively. Therefore, we need copulas that capture better the stylized facts and thus give a more complete description of the joint distribution.

In this paper, the worst case Copula-CVaR is achieved through the use of a convex linear combination of archimedean copulas consisting of the best mixture of Clayton, Frank and Gumbel copulas. This mixture is chosen because these archimedean copulas contain different tail dependence characteristics. It combines a copula with lower tail dependence, a copula with positive or negative dependence and a copula with upper tail dependence to produce a more flexible copula capable of modelling the multivariate log returns. Hence, by using a mixture copula we cover a wider range of possible dependencies structures within a single model capturing better the dependence between the individual assets which strongly influences the risk measures.

### 3.2.5   Worst Case Copula-CVaR

Similarly to Kakouris and Rustem (2014) and using the previously defined notations, let the decision vector be $\mathbf{w} = (w_1, ..., w_n)^\top$, $\mathbf{u} = (u_1, ..., u_n)$ in $[0, 1]^n$ be an stochastic vector, $h\left(\mathbf{w}, \mathbf{u}\right) = h\left(\mathbf{w}, F\left(\mathbf{x}\right)\right)$ the loss function and $F\left(\mathbf{x}\right) = \left(F_1\left(x_1\right), ..., F_n\left(x_n\right)\right)^\top$ a set of marginal distributions. Also, assume that $\mathbf{u}$ follows a continuous distribution with copula $C\left(\cdot\right)$.

Given a fixed $\mathbf{w} \in \mathcal{W}$, a random vector $\mathbf{x} \in \mathbb{R}^n$ and the equation (3.26), the probability that $h\left(\mathbf{w}, \mathbf{x}\right)$ does not exceed a threshold $\alpha$ is represented by

$$
\begin{aligned}
P\left(h\left(\mathbf{w}, \mathbf{x}\right) \leq \alpha\right) &= \int_{h(\mathbf{w}, \mathbf{x}) \leq \alpha} f\left(\mathbf{x}\right) d\mathbf{x} \\
&= \int_{h(\mathbf{w}, \mathbf{x}) \leq \alpha} c\left(F\left(\mathbf{x}\right)\right) \prod_{i=1}^n f_i\left(x_i\right) d\mathbf{x} \\
&= \int_{h\left(\mathbf{w}, \mathbf{F}^{-1}(\mathbf{u})\right) \leq \alpha} c\left(\mathbf{u}\right) d\mathbf{u} \\
&= C\left(\mathbf{u} \left| \widetilde{h}\left(\mathbf{w}, \mathbf{u}\right) \leq \alpha\right.\right),
\end{aligned}
$$

where $f_i\left(x_i\right) = \frac{\partial F_i(x_i)}{\partial x_i}$, $\mathbf{F}^{-1}\left(\mathbf{u}\right) = \left(F_1^{-1}\left(u_1\right), ..., F_n^{-1}\left(u_n\right)\right)^+$ and $\widetilde{h}\left(\mathbf{w}, \mathbf{u}\right) = h\left(\mathbf{w}, \mathbf{F}^{-1}\left(\mathbf{u}\right)\right)$. Thus, we can represent the $VaR_\beta$ by

$$VaR_\beta\left(\mathbf{w}\right) = \arg\min\left\{\alpha \in \mathbb{R} : C\left(\mathbf{u} \middle| \widetilde{h}\left(w, \mathbf{u}\right) \le \alpha\right) \ge \beta\right\} \tag{3.29}$$

and following (3.5) we can define $CVaR_\beta$ by

$$CVaR_\beta\left(\mathbf{w}\right) = \frac{1}{1-\beta} \int_{\widetilde{h}(\mathbf{w}, \mathbf{u}) \ge VaR_\beta(\mathbf{w})} \widetilde{h}\left(\mathbf{w}, \mathbf{u}\right) c\left(\mathbf{u}\right) d\mathbf{u}, \tag{3.30}$$

and by (3.9) we can write

$$H_\beta\left(\mathbf{w}, \alpha\right) = \alpha + \frac{1}{1-\beta} \int_{\mathbf{u} \in [0,1]^n} \left(\widetilde{h}\left(\mathbf{w}, \mathbf{u}\right) - \alpha\right)^+ c\left(\mathbf{u}\right) d\mathbf{u}. \tag{3.31}$$

Zhu and Fukushima (2009) derived the WCVaR considering a mixture of distributions in a prescribed set as we have seen in equations $(3.16) - (3.21)$. Kakouris and Rustem (2014) extended their framework considering a set of copulas $C\left(\cdot\right) \in \mathcal{C}$.

Let

$$C\left(\cdot\right) \in \mathcal{C}_M \equiv \left\{\sum_{i=1}^d \pi_i C_i\left(\cdot\right) : \sum_{i=1}^d \pi_i = 1, \ \pi_i \ge 0, \ i = 1, ..., d\right\}, \tag{3.32}$$

and similarly to (3.18)

$$\begin{aligned} (3.33) H_\beta\left(\mathbf{w}, \alpha, \pi\right) &= \alpha + \frac{1}{1-\beta} \int_{\mathbf{u} \in [0,1]^n} \left(\widetilde{h}\left(\mathbf{w}, \mathbf{u}\right) - \alpha\right)^+ \sum_{i=1}^d \pi_i c_i\left(\mathbf{u}\right) d\mathbf{u}, \ i = 1, ..., d \\ (3.34) \quad\quad &= \sum_{i=1}^d \pi_i H_\beta^i\left(\mathbf{w}, \alpha\right), \ i = 1, ..., d, \end{aligned}$$

where

$$H_\beta^i\left(\mathbf{w}, \alpha\right) = \alpha + \frac{1}{1-\beta} \int_{\mathbf{u} \in [0,1]^n} \left(\widetilde{h}\left(\mathbf{w}, \mathbf{u}\right) - \alpha\right)^+ c_i\left(\mathbf{u}\right) d\mathbf{u}, \ i = 1, ..., d. \tag{3.35}$$

Invoking theorem 1 of Zhu and Fukushima (2009) again we can state that for each fixed $\mathbf{w} \in \mathcal{W}$ the $WCVaR_\beta\left(\mathbf{w}\right)$ with respect to the set $\mathcal{C}$ is represented by

$$\begin{aligned} WCVaR_\beta\left(\mathbf{w}\right) &= \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} H_\beta\left(\mathbf{w}, \alpha, \pi\right) \\ (3.36) \quad\quad &= \min_{\alpha \in \mathbb{R}} \max_{\pi \in \Pi} \sum_{i=1}^d \pi_i H_\beta^i\left(\mathbf{w}, \alpha\right). \end{aligned}$$

Thus, the Worst Case Copula-CVaR with respect to $\mathcal{C}$ is the mixture copula that produces the greatest CVaR, i.e., the worst performing copula combination in the set $\mathcal{C}$. Moreover,

minimizing the worst-case Copula-CVaR over $\mathbf{w} \in \mathcal{W}$ can be defined as the following optimization problem:

$$\min_{\mathbf{w} \in \mathcal{W}} WCVaR_\beta (\mathbf{w}) = \min_{\mathbf{w} \in \mathcal{W}} \min_{\alpha \in \mathbb{R}} \sup_{\pi \in \Pi} H_\beta (\mathbf{w}, \alpha, \pi). \qquad (3.37)$$

Using the approach of Rockafellar and Uryasev (2000) the integral in (3.35) is approximated by sampling realizations from the copulas $C_i (\cdot) \in \mathcal{C}$ using as inputs the filtered uniform margins. If the sampling generates a collection of values $\left( \mathbf{u}_i^{[1]}, \mathbf{u}_i^{[2]}, ..., \mathbf{u}_i^{[J]} \right)$, where $\mathbf{u}_i^{[j]}$ and $S^i$ are the j-th sample drawn from copula $C_i (\cdot)$ of the mixture copula using as inputs the filtered uniform margins and its corresponding size, respectively, $i = 1, ..., d$, we can approximate $H_\beta^i (\mathbf{w}, \alpha)$ by

$$\widetilde{H}_\beta^i (\mathbf{w}, \alpha) = \alpha + \frac{1}{(1 - \beta) S^i} \sum_{j=1}^{S^i} \left( \widetilde{h}(\mathbf{w}, \mathbf{u}_i^{[j]}) - \alpha \right)^+, i = 1, .., d \qquad (3.38)$$

Assuming that the allowable set $\mathcal{W}$ is convex and the loss function $\widetilde{h} (\mathbf{w}, \mathbf{u})$ is linear with respect to $\mathbf{w}$ then optimization problem

$$\min_{\mathbf{w} \in \mathcal{W}, \alpha \in \mathbb{R}} \widetilde{H}_\beta (\mathbf{w}, \alpha). \qquad (3.39)$$

reduces to the following LP problem:

$$\operatorname*{minimize}_{\mathbf{w} \in \mathbb{R}^n, v \in \mathbb{R}^m, \alpha \in \mathbb{R}} \alpha + \frac{1}{(1 - \beta) S^i} \sum_{i=1}^{S^i} v_i \qquad (3.40)$$

$$\text{subject to} \quad \mathbf{w} \in \mathcal{W}, \qquad (3.41)$$

$$v_i^{[j]} \geq \widetilde{h}(\mathbf{w}, \mathbf{u}_i^{[j]}) - \alpha, \qquad (3.42)$$

$$v_i^{[j]} \geq 0, \; j = 1, \ldots, J \; i = 1, \ldots, d \qquad (3.43)$$

where $v_i$ are auxiliary indicator (dummy) variables and $m = \sum_{i=1}^d S^i$. By solving the LP problem we find the optimal decision vector, $\mathbf{w}^*$, and at "one shot" the optimal $VaR$, $\alpha^*$, and the optimal $CVaR$, $\widetilde{H}_\beta (\mathbf{w} = \mathbf{w}^*, \alpha = \alpha^*)$.

## 3.3  Methodology

Our data set consists of daily data of adjusted closing prices of all shares that belong to S&P 500 market index from July 2st, 1990 to December 31st, 2015. We obtain the adjusted closing prices from Bloomberg and log-returns are calculated in excess of a risk-

free asset[5]. The data set sample period is made up of 6426 days and includes a total of 1100 stocks over all periods. Only stocks that are listed throughout in-sample (48-month formation period) and out-of-sample (6 months) periods are included in the analysis.

We want a diversified set of stocks that can be useful during crises and tranquil periods. To attain a robust portfolio construction, we select a set of 50 stocks (only S&P 500 constituents) in each formation period, based on the ranked sum of absolute spreads[6] (the twenty-five largest and smallest) between the normalized daily closing prices deviations of the S&P 500 index and all shares. The distances are computed using data for January to December or from July to June. Prices are scaled to 1 at the beginning of each formation period and then evolve using the return series [7]. Specifically, the spread between the normalized closing prices at time $t$ is computed as

$$Spread_t = NP_{i,t} - NP_{SP500,t}, \tag{3.44}$$

where $NP_{i,t} = NP_{i,t-1}\left(1 + r_{i,t}\right),\ i = 1,...,N,\ t = 2,...,T$. We rebalance our portfolio every six months.

Our optimization strategy adopt a sliding window of calibration of T=4 years of daily data (1008 observations in average). Therefore, for example, we use day 1 to 1008 to estimate the parameters of all models and determine portfolio weights for day 1009 and then repeat the process including the latest observation and removing the oldest until reaching the end of the time series. We define $L$ as the number of days in the data set and thus, we compute $L - T - 1$ daily log-returns.

### 3.3.1   Strategies under Analysis

To apply Worst Case Copula-CVar Portfolio Optimization we go through the following steps : (1) First we fit an AR(1)-GARCH(1,1) model with skew t-distributed innovations to each univariate time series selected from distance method; (2) Using the estimated parametric model, we construct the standardized residuals vectors given, for each $i = 1,...,50$ and $t = 1,...,L - T - 1$, by

$$\frac{\widehat{\varepsilon}_{i,t}}{\widehat{\sigma}_{i,t}}.$$

The standardized residuals vectors are then converted to pseudo-uniform observations $z_{i,t} = \frac{n}{n+1}F_i\left(\widehat{\varepsilon}_{i,t}\right)$, where $F_i$ is their empirical distribution function; (3) Estimate the copula model, i.e., fits the multivariate Clayton-Frank-Gumbel (CFG) Mixture Copula to data that has been transformed to [0,1] margins by

$$C^{CFG}(\Theta, \mathbf{u}) = \pi_1 C^C(\theta_1, \mathbf{u}) + \pi_2 C^F(\theta_2, \mathbf{u}) + (1 - \pi_1 - \pi_2)C^G(\theta_3, \mathbf{u}) \tag{3.45}$$

where $\Theta = (\alpha, \beta, \delta)^\top$ are the Clayton, Frank and Gumbel copula parameters, respectively, and $\pi_1,\ \pi_2 \in [0, 1]$. The estimates are obtained by the minimization of the negative log-likelihood consisting of the weighted densities of the Clayton, Frank and Gumbel copulas. Probability density function for multivariate Archimedean copula is computed as described in Mcneil and Neshelova (2009); (4) Use the dependence structure determined by the estimated copula for generating $J$ scenarios. To simulate data from the three Archimedean copulas we use the sampling algorithms provided in Melchiori (2006); (5)

---

[5]We use 3-month Treasury Bill obtained at `https://fred.stlouisfed.org/series/TB3MS` as a proxy for the risk-free rate.

[6]A L1-norm minimization is used for increase robustness.

[7]Missing values have been interpolated.

compute t-quantiles for these Monte Carlo draws; (6) Compute the standard deviation $\widehat{\sigma}_{i,t}$ using the estimated GARCH model; (7) determine the simulated daily asset log-returns, i.e., determine the simulated daily log-returns as $r_{i,t}^{sim} = \widehat{\mu}_t + \widehat{\sigma}_{i,t} z_{i,t}$; (8) Finally, use the simulated data as inputs when optimizing portfolio weights by minimizing CVaR for a given confidence level and a given minimum expected return.

From Sklar's theorem the Gauss copula is given by

$$C_p\left(u_1, ..., u_d\right) = \mathbf{\Phi}_p\left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_d)\right), \tag{3.46}$$

where $\Phi$ denotes the standard normal distribution function, and $\mathbf{\Phi}_p$ denotes the multivariate standard normal distribution function with correlation matrix $P$. We simulate realizations from a Gaussian Copula repeating the following steps $J$ times. We first perform a Cholesky decomposition of $P$, and set $A$ as the resulting lower triangular matrix. Next we generate a vector $Z = (Z_1, ..., Z_d)'$ of independent standard normal variables and set $X = AZ$. Finally, we compute $U = (\Phi(X_1), ..., \Phi(X_d))'$.

For each of the four copulas, we run 1000 return scenarios from the estimated multivariate CFG Mixture Copula and multivariate Gaussian Copula models for each of the fifty assets. The weights are recalibrated at a daily, weekly and monthly basis. We assume that the feasible set $\mathcal{W}$ attends the budget and non-negativity constraints, i.e., the sum of the weights should be 1 ($\mathbf{w}^\top \mathbf{1} = 1$) and no short-selling is allowed ($\mathbf{w} \geq 0$).

The expected return should be bound below by an arbitrary value $\overline{r}$, i.e., $\mathbf{w}^\top \mathbb{E}\left(\mathbf{r}_{p,t+1}\right) \geq \overline{r}$, where $\overline{r}$ represents the target mean return.. The confidence level $\beta$ is set at $\beta = 0.95$.

We also consider three benchmarks: the Gaussian Copula portfolio, the equally weighted naive or 1/N portfolio, where $\mathbf{w} = \left(\frac{1}{N}, ..., \frac{1}{N}\right)^\top$ for any rebalancing date $t$ and the S&P 500 index as a proxy for market return.

### 3.3.2 Performance Measures

We assess out-of-sample portfolio allocation performance and its associated risks by means of the following statistics: mean excess returns, standard deviation, maximum drawdown between two consecutive days and between two days within a period of maximum six months, Sharpe ratio, Sortino ratio, turnover, breakeven costs, $\text{VaR}_{0.95}$, $\text{CVaR}_{0.95}$ and $\text{CR}_{0.95}$.

For each strategy, we compute the optimal weights $\mathbf{w}_t$ for each $t$ using the moving calibration window. We compute the portfolio excess return in time $t$, $t = T+1, ..., L-1$ by

$$\widehat{r}_{p,t} = \mathbf{w}_{t-1}^\top \mathbf{r}_{p,t} - r_{f,t}, \tag{3.47}$$

and the portfolio mean excess return as

$$\widehat{\mu}_p = \frac{1}{L-T}\sum\nolimits_{t=T+1}^{L-1} \widehat{r}_{p,t} \tag{3.48}$$

The portfolio standard deviation and Sharpe ratio are given, respectively, by

$$\widehat{\sigma}_p = \sqrt{\frac{1}{L-T-2}\sum\nolimits_{t=T}^{L-1}\left(w_t r_{p,t+1} - \widehat{\mu}_p\right)^2}, \tag{3.49}$$

and

$$SR = \frac{\widehat{\mu}_p}{\widehat{\sigma}_p}. \tag{3.50}$$

Denote now by $r_p(\mathbf{w},t)$ the cumulative portfolio return at time $t$, where $\mathbf{w}$ are asset weights in the portfolio. The drawdown function at time $t$ (Unger (2014)) is the defined as the difference between the maximum of the function $r_p(\mathbf{w},t)$ over the history preceding time $t$ and the value of this function at time $t$, i.e.,

$$D(\mathbf{w},t) = \max_{0 \leq \tau \leq t} \{r_p(\mathbf{w},t)\} - r_p(\mathbf{w},t). \qquad (3.51)$$

The maximum drawdown on the interval $[0,T]$ is defined as

$$MaxDD(\mathbf{w}) = \max_{0 \leq \tau \leq t} \{D(\mathbf{w},t)\}. \qquad (3.52)$$

In other words, the maximum drawdown over a period is the maximum loss from worst peak to a trough of a portfolio drop from the start to the end of the period.

The Sortino's ratio (Sortino and Price (1994)) is the ratio of the mean excess return to the standard deviation of negative asset returns, i.e.,

$$SoR = \frac{\widehat{\mu}_p}{\widehat{\sigma}_{p,n}}, \qquad (3.53)$$

where

$$\widehat{\sigma}_{p,n} = \sqrt{\frac{1}{L-T-1} \sum_{t=T}^{L-1} \left(\min\left(0, \mathbf{w}_t^\top \mathbf{r}_{t+1} - \mathbf{r}_{MAR}\right)\right)^2}, \qquad (3.54)$$

where $\mathbf{r}_{MAR}$ is the value of a minimal acceptable return (MAR), usually zero or the risk-free rate[8].

We define the portfolio turnover from time $t$ to time $t+1$ as the sum of the absolute changes in the $N$ risky asset weights, i.e., in the optimal values of the investment fractions:

$$Turnover = \frac{1}{L-T-1} \sum_{t=T}^{L-1} \sum_{j=1}^{N} \left(|w_{j,t+} - w_{j,t+1}|\right), \qquad (3.55)$$

where $w_{j,t+}$ is the actual weight in asset $j$ before rebalancing at time $t+1$, and $w_{j,t+1}$ is the optimal weight in asset $j$ at time $t+1$. Turnover measures the amount of trading required to implement a particular portfolio strategy and can be interpreted as the average fraction of wealth traded in each period.

We also report the break-even transaction cost proposed by Bessembinder and Chan (1995), which is the level of transaction costs leading to zero net profits, i.e., the maximum transaction cost that can be imposed before making the strategies less desirable than the buy-and-hold strategy (see Han (2006)). Following Santos and Tessari (2012) we consider the average net returns on transaction costs, $\widehat{\mu}_{TC}$, given by

$$\widehat{\mu}_{TC} = \frac{1}{L-T} \sum_{t=T}^{L-1} \left[\left(1 + \mathbf{w}_t^\top \mathbf{r}_{t+1}\right)\left(1 - c \sum_{j=1}^{N} \left(|w_{j,t+1} - w_{j,t+}|\right)\right) - 1\right], \qquad (3.56)$$

where $c$ is called breakeven cost when we solve $\widehat{\mu}_{TC} = 0$. Those portfolios that achieve

---

[8]Sortino Ratio is an improvement on the Sharpe Ratio since it is more sensitive to extreme risks or downside than measures Sharpe Ratio. Sortino contends that risk should be measured in terms of not meeting the investment goal. The Sharpe ratio penalizes financial instruments that have a lot of upward jumps, which investors usually view as a good thing.

a higher break-even cost are preferable, since the level required to make these portfolios non-profitable are higher.

The Capital requirement loss function (CR) is a regulatory loss function to evaluate VaR forecasts, given by:

$$CR_t = \max \left[ \left( \left( \frac{3+\delta}{60} \right) \sum_{i=0}^{59} VaR_{\beta,t-i} \right), VaR_{\beta,t} \right], \qquad (3.57)$$

where $\delta = \begin{cases} 0, \text{if } \zeta \leq 4 \\ 0.3 + 0.1\,(\zeta - 4), \text{if } 5 \leq \zeta \leq 6 \\ 0.65, \text{if } \zeta = 7 \\ 0.65 + 0.1\,(\zeta - 7), \text{if } 8 \leq \zeta \leq 9 \\ 1, \text{if } \zeta \geq 10 \end{cases}$

is a multiplicative factor that depends on the number of violations of the VaR in the previous 250 trading days $(\zeta)$. Note that the CR is based on the larger out of the current VaR estimate and a multiple of the average estimate over the past 60 days.

We compute the capital requirement losses based on the 95% daily VaR forecasts, and in order to mitigate data-snooping problems we apply a test for superior predictive ability (SPA) proposed by Hansen (2005) to test which model significantly minimizes the expected loss function (or has a superior predictive ability, in other words). The relative performance of the Mixed Copula-CVaR portfolio to the Gaussian Copula-CVaR portfolio may be defined as

$$d_{m,t} = CR_{g,t} - CR_{m,t},$$

where $m$ and $g$ stands for Mixed and Gaussian Copula CVaR portfolios, respectively. The null hypothesis is that the average performance of the Gaussian Copula CVaR portfolio is as small as the minimum average performance of the Worst Case Copula CVaR portfolio. Provided that $E(d_{m,t}) = \mu_{m,t}$ is well defined, we can formulate the null hypothesis as

$$\begin{cases} H_0 : \mu_{m,t} \leq 0 \\ H_1 : \mu_{m,t} > 0. \end{cases}$$

The test statistic is given by

$$T^{SPA} = \max \left[ \frac{\sqrt{T}\overline{d}_{k,t}}{\widehat{w}_k}, 0 \right],$$

where $\overline{d}_{k,t} = T^{-1} \sum_{t=1}^{T} d_{k,t}$ e $\widehat{w}_k^2$ is some consistent estimator of $w_k^2$, the variance of $\sqrt{T}\overline{d}_{k,t}$. To construct the distributions we bootstraped the original time series $B = 10000$ times and select the optimal block length for the stationary bootstrap following Politis and White (2004). Our bootstrapped null distributions result from Theorem 2 of Politis and Romano (1994). Since the optimal bootstrap block length is different for each strategy we average the block lengths found to proceed the comparisons between the strategies.

## 3.4   Empirical Results

Table 1 reports out-of-sample mean excess return, standard deviation, Sharpe ratio, Sortino ratio, turnover, break-even transaction costs, maximum drawdown between two consecutive days (MDD1) and between two days within a period of maximum six months (MDD2), $VaR_{0.95}$, $CVaR_{0.95}$ and $CR_{0.95}$ of the different portfolio strategies from $1991/2$ to $2015/2$ involving daily, weekly and monthly rebalacing frequencies. Returns, standard deviation, Sharpe ratio and Sortino ratio are summarized in an annualized basis.

By selecting a diversified set of assets over a long-term period we found that copula-based approaches offer better hedges against losses than the $1/N$ portfolio. The WCCVaR approach generates portfolios with better downside risk statistics for any rebalancing period and it is more profitable than the Gaussian Copula-CVaR for daily and weekly rebalancing.

By analyzing Table 1, it is clear that the copula-based allocations consistently outperform the other benchmarks in terms of risk-return adjusted metrics, volatility and maximum drawdown, although the results are somewhat sensitive to the rebalancing window. We can also note that the Worst Case Copula-CVaR has a superior performance than the Gaussian Copula-CVaR in terms of the downside (tail) risk metrics as VaR, CVaR and CR.

Although our main objective is to protect investors from the worst possible scenarios it is worthwhile noticing that the WCCVaR strategy consistently outperforms the other strategies since July 1994 in terms of cumulative excess returns for any daily and weekly rebalancing frequencies without a minimum expected return constraint [9]. However, the breakeven costs show that the profits of the copula strategies do not survive for transactions costs greater than 6.2 basis points.

---

[9]The numerical experiments show that the performances out-of-sample stay very similar when we consider a daily constraint $\bar{r} \geq VaR_{10\%}$. Since the resuls are very much alike they are not presented here and are available under request.

**Table 1:** Excess returns of Worst Case Copula-CVaR (WCCVaR), Gaussian Copula-CVaR (GCCVaR), Equal Weights (1/N) portfolios and S&P 500 market index without a minimum expected return constraint

| | WCCVaR | GCCVaR | 1/N | S&P 500 |
|---|---|---|---|---|
| *Daily Rebalancing* | | | | |
| Mean Return (%) | 11.29 | 10.59 | 10.52 | 6.61 |
| Standard Deviation (%) | 15.35 | 15.05 | 23.86 | 19.01 |
| Sharpe Ratio | 0.70 | 0.67 | 0.42 | 0.34 |
| Sortino Ratio | 1.14 | 1.09 | 0.68 | 0.54 |
| Turnover | 0.6868 | 0.6425 | 0.0001 | |
| Break-even (%) | 0.0618 | 0.0622 | 275.09 | |
| MDD1 (%) | -13.20 | -12.89 | -18.00 | -13.20 |
| MDD2 (%) | -38.94 | -41.66 | -71.84 | -52.61 |
| $VaR_{0.95}$ | -0.0091 | -0.0109 | | |
| $CVaR_{0.95}$ | -0.0109 | -0.0135 | | |
| $CR_{0.05}$(%) | 13.00*** | 15.63 | | |
| *Weekly Rebalancing* | | | | |
| Mean Return (%) | 10.57 | 10.04 | 10.52 | 6.61 |
| Standard Deviation (%) | 15.48 | 15.21 | 23.86 | 19.01 |
| Sharpe Ratio | 0.65 | 0.63 | 0.42 | 0.34 |
| Sortino Ratio | 1.06 | 1.02 | 0.68 | 0.54 |
| Turnover | 0.1535 | 0.1564 | 0.0001 | |
| Break-even (%) | 0.2596 | 0.2427 | 275.09 | |
| MDD1 (%) | -12.59 | -10.39 | -18.00 | -13.20 |
| MDD2 (%) | -39.78 | -41.32 | -71.84 | -52.61 |
| $VaR_{0.95}$ | -0.0091 | -0.0109 | | |
| $CVaR_{0.95}$ | -0.0109 | -0.0135 | | |
| $CR_{0.05}$ (%) | 12.86*** | 15.57 | | |
| *Monthly Rebalancing* | | | | |
| Mean Return (%) | 9.22 | 9.42 | 10.52 | 6.61 |
| Standard Deviation (%) | 15.69 | 15.55 | 23.86 | 19.01 |
| Sharpe Ratio | 0.56 | 0.58 | 0.42 | 0.34 |
| Sortino Ratio | 0.91 | 0.94 | 0.68 | 0.54 |
| Turnover | 0.0447 | 0.0499 | 0.0001 | |
| Break-even (%) | 0.7829 | 0.7155 | 275.09 | |
| MDD1 (%) | -13.21 | -13.22 | -18.00 | -13.20 |
| MDD2 (%) | -38.58 | -46.13 | -71.84 | -52.61 |
| $VaR_{0.95}$ | -0.0091 | -0.0109 | | |
| $CVaR_{0.95}$ | -0.0109 | -0.0134 | | |
| $CR_{0.05}$ (%) | 12.22*** | 14.67 | | |

*Note:* Out-of-sample performance statistics between July 1994 and December 2015 (5414 observations). The rows 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. Returns, standard deviation, Sharpe ratio and Sortino ratio are annualized.

\*\*\* , \*\*, \* indicate a rejection of the null hypothesis of equal predictive ability at 1%, 5% and 10% levels, respectively, according to the Hansen (2005)'s SPA test.

Source: Author's own elaboration (2017).

Up to now we did not consider transaction costs when we purchase and sell the assets (or "turns over" our portfolio). But, if we want to use our strategies for tactical asset allocation, transaction costs play a non-trivial role and must not be neglected. With this in mind, we compute the portfolio turnover of each strategy. The higher the turnover, the higher the transaction cost that the portfolio incurs at each rebalancing day. We

can note that the copula-based approaches present a similar turnover in the rebalancing frequencies analyzed. However, the WCCVaR has a relatively better performance when the rebalancing frequency decreases. As expected, the portfolio turnover decreases for both allocations when the holding portfolio period increases and the EWP (1/N) has the smallest turnover.

In addition, we report the break-even transaction cost in order to investigate if the profits are economically significant. The break-even values in Table 1 represent the level of transaction costs leading to zero excess return. Thus, those portfolios that achieve a higher break-even cost are preferable, since the level required to make these portfolios non-profitable are higher[10]. For the WCCVaR and GCCVaR allocations the break-even costs are between [0.062%-0.783%] and [0.062%-0.716%], respectively. Thus, for month rebalancing this means that if the trading costs are anything less than 78 basis points the excess profits for the WCCVaR allocations are still greater than zero. We can also note that the break-even costs for WCCVaR and GCCVaR are similar, but that the difference increases when the holding period lengthens, as expected after analyzing the portfolio turnover.

Finally, we reject for all rebalancing frequencies the null hypothesis of the Hansen (2005)'s SPA test (p-values=0.0000) that the average performance of the Gaussian Copula CVaR portfolio is as small as the minimum average performance of the Worst Case Copula CVaR portfolio. In other words, we find that WCCVaR allocations have a superior predictive ability when compared to the GCCVaR allocations.

Figure 3.1 depicts the wealth trajectories of the portfolios strategies for all rebalancing frequencies assuming an initial wealth of $1 monetary unit. Panels (a) to (c) shows the excess returns for daily, weekly and monthly rebalancing, respectively.

The copula-based approaches are more profitable for daily and weekly rebalancing over time, especially after 2009. We can note the copula methods present a hump-shaped pattern in 1999, while the other benchmarks show a sharp decline in the subperiod that corresponds to the bear market that comprises the dotcom crisis and the September 11th terrorist attack (2000-2002). All portfolios show a hump-shaped pattern during the subprime mortgage financial crisis in 2007-2008. Overall, we can observe that after 2002 the patterns are similar, but the figure indicates that the copula methods, even though the objective function is the minimization of CVaR under a constraint on expected return, preserve more wealth in the long-term period, particularly the WCCVaR portfolio, for daily and weekly rebalancing [11].

## 3.5 Conclusions

In this paper we combine robust portfolio optimization and copula-based models in a Worst Case CVaR framework. To cope with the large number of financial instruments we employ a procedure similar to that used by Gatev, Goetzmann and Rouwenhorst (2006) to select a set of diversified assets that can be useful during crises and tranquil periods. Using data from the S&P 500 shares from 1990 to 2015 we evaluate the performance of the WCCVaR (Worst Case Copula-CVaR) portfolio, considering different rebalancing strategies, and compare it against three benchmarks: a Gaussian Copula-CVaR (GCCVaR) portfolio, an equally weighted portfolio (1/N) and the S&P 500 index.

By selecting a diversified set of assets over a long-term period we found that copula-based approaches offer better hedges against losses than the 1/N portfolio. Moreover, the WCCVaR approach generates portfolios with better downside risk statistics for any rebalancing period and it is more profitable than the Gaussian Copula-CVaR for daily and weekly rebalancing.

Finally, we offer some suggestions for future research for improving our discoveries. First, we could improve the method of asset selection. We suspect that if we use a

---

[10]Jegadeesh and Titman (1993) consider a conservative transaction cost of 0.5% per trade, while Allen and Karjalainen (1999) consider 0.1%, 0.25% and 0.5%.

[11]Transaction costs should not be neglected as we could notice after the breakeven analysis

**Figure 3.1:** Cumulative excess returns of the portfolio strategies without daily target mean return
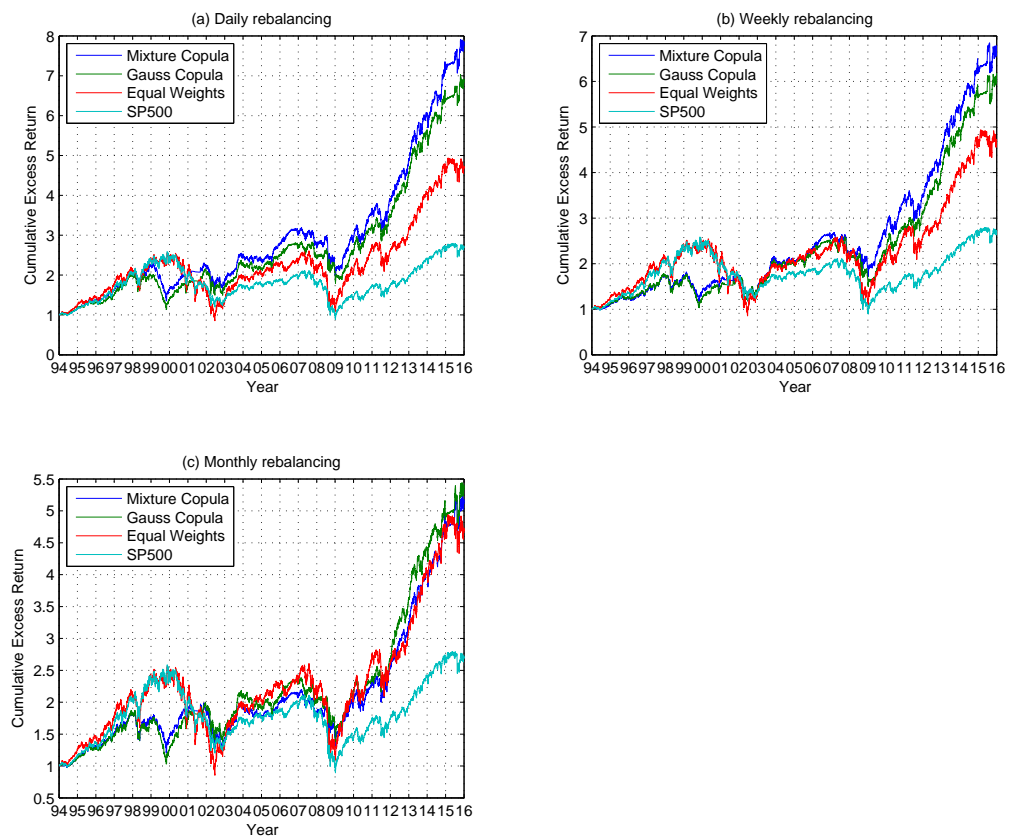


Source: Author's own elaboration (2017).

This figure shows how an investment of $1 evolves from July 1994 to December 2015 for each of the portfolios.

measure of nonlinear dependence between random variables of arbitrary dimension such as the randomized dependency coefficient (Lopez-Paz, Hennig and Schölkopf (2013)) or a procedure based on data mining tools as random forest Giovanni De Luca and Zuccolotto (2010) to select the stocks the portfolio performances would be even better. Additional suggestions include relaxing the assumption of no short selling and incorporate transaction cost as an additional constraint in the optimization problem as in Krokhmal, Palmquist and Uryasev (2002).

## REFERENCES

ALEXANDER, Gordon J; BAPTISTA, Alexandre M. Economic implications of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis. **Journal of Economic Dynamics and Control**, Elsevier, v. 26, n. 7, p. 1159–1193, 2002.

ALLEN, Franklin; KARJALAINEN, Risto. Using genetic algorithms to find technical trading rules. **Journal of financial Economics**, Elsevier, v. 51, n. 2, p. 245–271, 1999.

ARTZNER, Philippe et al. Coherent measures of risk. **Mathematical finance**, Wiley Online Library, v. 9, n. 3, p. 203–228, 1999.

BESSEMBINDER, Hendrik; CHAN, Kalok. The Profitability of Technical Trading Rules in the Asian Stock Markets. **Pacific-Basin Finance Journal**, Elsevier, v. 3, n. 2, p. 257–284, 1995.

CHERUBINI, Umberto; LUCIANO, Elisa; VECCHIATO, Walter. **Copula methods in finance**. Chichester, England: John Wiley & Sons, 2004.

GATEV, Evan; GOETZMANN, William N.; ROUWENHORST, K. Geert. Pairs Trading: Performance of a Relative-Value Arbitrage Rule. **Review of Financial Studies**, v. 19, n. 3, p. 797–827, 2006.

GIOVANNI DE LUCA, Giorgia Rivieccio; ZUCCOLOTTO, Paola. Combining random forest and copula functions: a heuristic approach for selecting assets from a financial crisis perspective. **Intelligent Systems in Accounting, Finance and Management**, v. 17.2, p. 91–109, 2010.

HAN, Yufeng. Asset Allocation with a High Dimensional Latent Factor Stochastic Volatility Model. **Review of Financial Studies**, Soc Financial Studies, v. 19, n. 1, p. 237–271, 2006.

HANSEN, Peter Reinhard. A test for superior predictive ability. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 23, n. 4, p. 365–380, 2005.

JEGADEESH, Narasimhan; TITMAN, Sheridan. Returns to buying winners and selling losers: Implications for stock market efficiency. **The Journal of finance**, Wiley Online Library, v. 48, n. 1, p. 65–91, 1993.

JOE, Harry. **Multivariate models and multivariate dependence concepts**. CRC Press, 1997.

KAKOURIS, Iakovos; RUSTEM, Berç. Robust portfolio optimization with copulas. **European Journal of Operational Research**, Elsevier, v. 235, n. 1, p. 28–37, 2014.

KROKHMAL, Pavlo; PALMQUIST, Jonas; URYASEV, Stanislav. Portfolio Optimization with Conditional Value-At-Risk Objective and Constraints. **Journal of Risk**, Citeseer, v. 4, p. 43–68, 2002.

LOPEZ-PAZ, David; HENNIG, Philipp; SCHÖLKOPF, Bernhard. The Randomized Dependence Coefficient. In: ADVANCES in Neural Information Processing Systems. [S.l.: s.n.], 2013. p. 1–9.

MARKOWITZ, Harry. Portfolio selection. **The journal of finance**, Wiley Online Library, v. 7, n. 1, p. 77–91, 1952.

MCNEIL, Alexander J; NESHELOVA, Johanna. Multivariate Archimedean Copulas, d-Monotone Functions and $\ell$1-norm Symmetric Distributions. **Annals of statistics**, Institute of Mathematical Statistics, v. 37, n. 5, p. 3059–3097, 2009.

MELCHIORI, Mario R. Tools for sampling multivariate archimedean copulas, 2006.

NELSEN, R. B. **An introduction to copulas, 2nd**. New York: Springer Science Business Media, 2006.

PFLUG, Georg Ch. Some remarks on the value-at-risk and the conditional value-at-risk. In: PROBABILISTIC constrained optimization. Boston, USA: Springer, 2000. p. 272–281.

POLITIS, Dimitris N.; ROMANO, Joseph P. The stationary bootstrap. **Journal of the American Statistical association**, v. 89.428, p. 1303–1313, 1994.

POLITIS, Dimitris N.; WHITE, Halbert. Automatic block-length selection for the dependent bootstrap. **Econometric Reviews**, v. 23.1, p. 53–70, 2004.

ROCKAFELLAR, R Tyrrell; URYASEV, Stanislav. Conditional value-at-risk for general loss distributions. **Journal of banking & finance**, Elsevier, v. 26, n. 7, p. 1443–1471, 2002.

_____. Optimization of conditional value-at-risk. **Journal of risk**, v. 2, p. 21–42, 2000.

SANTOS, André Alves Portela; TESSARI, Cristina. Técnicas Quantitativas de Otimização de Carteiras Aplicadas ao Mercado de Ações Brasileiro (Quantitative Portfolio Optimization Techniques Applied to the Brazilian Stock Market). **Revista Brasileira de Finanças**, Sociedade Brasileira de Finanças, v. 10, n. 3, p. 369, 2012.

SCHWEIZER, Berthold; SKLAR, Abe. **Probabilistic Metric Spaces**. Mineola, USA: Courier Corporation, 2011.

SKLAR, M. **Fonctions de répartition à n dimensions et leurs marges**. Paris, France: Université Paris 8, 1959.

SORTINO, Frank A.; PRICE, Lee N. Performance Measurement in a Downside Risk Framework. **The Journal of Investing**, Institutional Investor Journals, v. 3, n. 3, p. 59–64, 1994.

UNGER, Albina. **The Use of Risk Budgets in Portfolio Optimization**. Friedrichsdorf, Germany: Springer, 2014.

URYASEV, Stanislav. **Probabilistic Constrained Optimization: Methodology and Applications**. [S.l.]: Springer Science & Business Media, 2013. v. 49.

URYASEV, Stanislav; ROCKAFELLAR, R Tyrrell. Conditional value-at-risk: optimization approach. In: STOCHASTIC optimization: algorithms and applications. [S.l.]: Springer, 2001. p. 411–435.

_____. **Optimization of Conditional Value-At-Risk**. [S.l.]: Department of Industrial & Systems Engineering, University of Florida, 1999.

ZHU, Shushang; FUKUSHIMA, Masao. Worst-case conditional value-at-risk with application to robust portfolio management. **Operations research**, INFORMS, v. 57, n. 5, p. 1155–1168, 2009.

**Chapter 4**

**Performance of Copula-HEAVY, Copula-GARCH and Distance Pairs Trading**

**on global stock market equity indices**

**Abstract.** By using an intra-day realized measure we empirically evaluate a Copula-HEAVY pairs trading model against a Copula-GARCH and the distance methods using 21 global stock market indices from January 2000 to June 2016. We find that Copula-GARCH approach is highly profitable with annual excess returns up to 12.10% and 29.95% on committed and fully invested capital, respectively, and Sharpe ratios up to 3.16. On the other hand, the Copula-HEAVY and distance methods show an annualized excess returns on committed and fully invested capital as high as 7.02% and 17.19% and 3.10% and 3.62%, and Sharpe ratios up to 2.21 and 0.6, respectively. The copula-based pairs strategies also present more trading opportunities than the distance strategies. However, the average excess returns of copula strategies are more sensitive to transaction costs and speed of execution. To test the statistical significance of the Sharpe Ratio and excess returns we adopt the stationary bootstrap of Politis and Romano (1994) using the automatic block-length selection of Politis and White (2004). Results suggest that the Copula-GARCH is consistently the best rewarding strategy when a fast trade is made.

**Keywords:** Pairs Trading. Copula. Distance. Statistical Arbitrage. High-Frequency. Realized variance. Stationary Bootstrap.

**JEL Classifications.** C51, C58, G11, G12, G14.

## 4.1 Introduction

Pairs trading is a statistical arbitrage method that is often used by proprietary trading firms and hedge funds which exploits relative mispricing within two assets. The strategy attracts much interest in empirical finance, since it has potential to achieve abnormal returns through relatively low-risk positions. In addition, the strategy is claimed to be market neutral, which means that the investors are not exposed to market risk. The strategy was successfully pioneered by Gerry Bamberger and Nunzio Tartaglia's quantitative group at Morgan Stanley in the 1980s Gatev, Goetzmann and Rouwenhorst (2006). Since then many techniques to compute the spread between stock pairs were proposed; however, the strategy became popular through the study carried out by Gatev, Goetzmann and Rouwenhorst (2006), named distance method.

In an efficient market, strategies based on mean-reversion concepts should not generate consistent profits. However, Gatev, Goetzmann and Rouwenhorst (2006) found that pairs trading generates consistent statistical arbitrage profits in the U.S. equity market during 1962-2002 using all liquid U.S. stocks from the Center for Research in Security Prices (CRSP) data, although the profitability declined over the period. They obtain an annual mean excess return above 11% during the reported period. The authors attributed the abnormal returns to a non-identified systematic risk factor. They support their view by showing that there is a high degree of correlation between the excess returns of non-overlapping top pairs even after accounting for risk factors from an augmented version of Fama and French (1993)'s three-factor model. Do and Faff (2010) extend their work by expanding the data sample by more seven years and also found a declining trend (33 basis points (bps) mean excess return per month for 2003-09 versus 124 bps mean excess return per month for 1962-88). Do and Faff (2012) show that the distance method is unprofitable after 2002 when taking trading costs into account. Broussard and Vaihekoski (2012) test the profitability of pairs trading under different weighting structures and trade initiation conditions using data from the Finnish stock market. They also found that their proposed strategy is profitable even after initiating the positions one day after the signal.

Currently, there are three main methods for pairs trading: distance, cointegration, and copula. The first two rely on the assumption that the pair of stocks is linearly associated which holds as long as the data have a multivariate normal distribution. However, it is well known that the dependence between two securities is rarely jointly normal (CAMPBELL; LO; MACKINLAY, 1997; CONT, 2001; ANE; KHAROUBI, 2003; MCNEIL; FREY; EM-BRECHTS, 2015) and thus a high-dimensional multivariate approach to tail dependence analysis is surely more insightful than assuming multivariate normal returns. Therefore, the trading signals based on the distance and cointegration strategies may fail to catch the dynamics of the spread between a pair of securities, and thus initiate and close the trades at non-optimal positions. To overcome these issues, Liew and Wu (2013) propose a pairs trading strategy based on bivariate copulas. Because of its flexibility, copulas are able to model better the empirically verified non-linear relations normally attributed to multivariate financial returns such as asymmetric conditional variance with higher volatility for large negative returns and smaller volatility for positive returns (HAFNER, 1998) and leptokurdicity (TAUCHEN, 2001; ANDREOU; PITTIS; SPANOS, 2001). However, they evaluate its performance using only three pre-selected pairs over a period of less than three years. Xie, Liew, Wu and Zou (2016) employ a similar methodology on a daily frequency over a ten-year period selecting 89 stocks from the U.S. utility sector. Both studies found that the copula-based strategy generates higher average excess returns and more trading opportunities than the distance approach. However, Rad, Low and Faff (2015) found opposite results using daily U.S. equity data from 1962 to 2014. Particularly, the distance, cointegration, and Copula-GARCH strategies show an average monthly excess return of 36, 33, and 5 basis points (bps) after transaction costs, respectively, and 88, 83, and 43 bps before transaction costs, respectively. All the aforementioned papers test the distance and copula strategies on daily frequency data and the dependence parameters of the copula-based approaches are assumed to be time-invariant.

However, currently there is no extensive research on pairs trading strategies using realized measures computed by employing higher frequency data. Nath (2003) employs

a variant of the Gatev, Goetzmann and Rouwenhorst (2006)'s approach using intraday quotes from the U.S. treasury market from 1994 to 2000. He found that the strategy outperforms the benchmarks, especially regarding Sharpe Ratio. Bowen, Hutchinson and O'Sullivan (2010) test the method using a 60-minute frequency in a sample of the FTSE 100 constituent stocks from January to December 2007. They document that the average excess returns of the strategy is very sensitive to trading costs and speed of execution. In particular, the profitability of the strategy for Top 5 pairs using $2.0\sigma$ standard deviations is almost 20% more before costs and it is reduced by almost 50% when employing 10 basis points. In addition, when allowing a one-day waiting period restriction on execution, the profits are eliminated. They also describe that over a half of the excess returns is accomplished in the first hour of trading and 75% in the first and last hours of the trading day. Dunis, Giorgioni, Laws and Rudy (2010) test a cointegration pairs trading approach in intraday data of 5, 10, 20, 30 and 60 minutes in the constituent shares of the Euro Stoxx 50 index comprising the period from 3rd July 2009 to 17th November 2009. They found an annualized average Information Ratio above 3 for high frequency intervals and about 1.3 using daily data for the Top 5 pairs.

Standard univariate GARCH models specify the conditional volatility as a function of past low-frequency data (*i.e.*, daily returns). Andersen, Bollerslev, Diebold and Ebens (2001) show that underlying latent volatility can be computed accurately using intraday returns, whereas squared returns only provide a very noisy volatility indicator. It is shown that, when appropriately managed, incorporating realized volatility measures provides better volatility forecasts when compared to estimates constructed using daily data (SHEPHARD; SHEPPARD, 2010; HANSEN; HUANG; SHEK, 2012). Fleming, Kirby and Ostdiek (2003) estimate that a risk-averse investor would be willing to pay 50 to 200 basis points per year to capture the observed gains in performance.

A class of univariate forecasting model that are based on realized measures of high frequency volatility is called High Frequency Based Volatility (HEAVY). It can be shown that the HEAVY model exhibits short-run momentum and mean-reversion effects, adjusting quickly to structural breaks in the level of the volatility process than standard GARCH models (SHEPHARD; SHEPPARD, 2010). Thus, by using realized information about current levels of volatilities, the HEAVY model may be particularly useful during periods of fast changes in the underlying volatility process. The realized measure we use here is a simple realized variance (RV) estimator with a 5-minute sample frequency, namely 5-minute realized variance. Liu, Patton and Sheppard (2015) show that there is little evidence that a 5-minute RV is outperformed by a profusion of other realized measures.

In this chapter, we conduct an empirical investigation in order to offer some evidence of the behavior of the Copula-HEAVY, Copula-GARCH, and distance pairs trading strategies under different investment scenarios using 21 global stock market indexes from January 2000 to June 2016. All data we use are obtained from Oxford-Man Institute of Quantitative Finance Realized Library[1]. We are also interested in providing some evidence of how the delay to start the positions as well the trading costs affects the profitability of these strategies. Our aim is to provide investors with information that allows for a more accurate selection of which strategy should be used in a particular setting and understand any differences in performance.

Our results indicate that the Copula-GARCH strategy outperforms consistently the other strategies for Top 20 pairs in the long term in terms of wealth accumulation and risk-adjusted returns when a fast trade can be executed assuming four equally spaced levels of transaction costs from zero to 15 basis points and before costs for a wait one-day rule. The Copula-HEAVY approach also outperforms the distance methods for transaction costs up to 10 basis points when positions are initiated and closed in the same day the pair diverges and converges, respectively. However, the returns from copula strategies are very sensitive to the magnitude of transaction costs and speed of execution. This is probably due to the fact we measure the degrees of relative mispricing for a single day instead of determining an overall degree of mispricing, increasing the number of trading opportunities.

To test the statistical significance of the returns and Sharpe Ratios we adopt the sta-

---

[1] http://realized.oxford-man.ox.ac.uk/data/download

tionary bootstrap of Politis and Romano (1994) using the automatic block-length selection of Politis and White (2004) and 10,000 bootstrap resamples. To compute the bootstrap $p$-values we use the methodology proposed by Ledoit and Wolf (2008). Our objective is to compare the results on a statistical basis in order to mitigate potential data snooping problems. We find that the Copula-GARCH strategy significantly outperforms the other strategies (all p-values are less than 0.5%) when a rapid trade is made.

Additionally, we conduct the analysis using four sub-periods to avoid the concern that the results are good only for the whole period: (1) January 2001 to December 2002; (2) January 2003 to June 2007; (3) July 2007 to June 2009, (4) and July 2009 to June 2016. The first sub-period corresponds to the bear market that spans the dotcom crisis and the September 11th terrorist attack. The third sub-period corresponds to the subprime mortgage financial crisis. Our results indicate that the copula methods achieve its best performance during financial downturns when trades can be quickly executed after the signal.

The remainder of this chapter is organized as follows. In Section 2 we briefly discuss the data and introduce the HEAVY model. In Section 3, we present the results of our empirical analysis. Finally, in Section 4, we conclude. The Appendix contains additional results.

## 4.2   Data and the HEAVY model

Our data set consists of daily data of adjusted closing prices of 21 global stock market indexes obtained from the Oxford-Man Institute's realized library from January 2000 to June 2016. The data set sample period is made up of 4310 days and consists of the following indexes: S&P 500, NASDAQ-100, FTSE 100, STI, Nikkei 225, Russell 2000, DAX-30, AORD, DJI, CAC40, HSI, KS11, AEX, SSMI, IBEX 35, Nifty 50, MXX, Ibovespa, S&P/TSX, EURO STOXX 50, and FTSE MIB.

Our implementation of the pairs trading is similar to Gatev, Goetzmann and Rouwenhorst (2006) and is explained in more details in subsections 2.2.1 and 2.2.2. Potential security pairs are sorted based on the sum of squared differences in their normalized prices during the formation period. After the pairs are formed, their spread is monitored throughout the trading period and any deviations beyond a certain threshold in that spread would trigger the opening of two simultaneous long and short positions. However, as in Broussard and Vaihekoski (2012), we trade the pairs at the beginning of every six months, so we do not have a time series of overlapping six-month trading period as in Gatev, Goetzmann and Rouwenhorst (2006). We use the distance approaches as our benchmarks and evaluate Copula-HEAVY and Copula-GARCH strategies by comparing their performance. As realized measure, we use realized variance at 5-minute intervals without sub-sampling.

Pairs are formed based on the smallest sum of squared deviations between the two normalized price deviations among all shares listed throughout the 12 plus 6-month period (formation and trading periods) for both distance and copula approaches[2]. For the distance method, the positions are initiated when this spread is greater than 0.75 and 2.0 standard deviations (based on Vidyamurthy (2004) and Gatev, Goetzmann and Rouwenhorst (2006), respectively). If normally distributed, this would happen 45% and 5% approximately of the time, respectively. For the copula approach, we set 0.25 and 0.75 for the mispricing indexes. This would be similar to $0.75\sigma$ trigger point if data is normally distributed.

---

[2]Missing values have been interpolated.

### 4.2.1   The HEAVY model

Let us denote the time series of log daily return data and the realized 5-minute variance by $r_1, r_2, \ldots, r_T$ and $RV_1, RV_2, \ldots, RV_T$, respectively. With the realized measures computed for $T$ days, the HEAVY (High Frequency Based Volatility) model (SHEPHARD; SHEPPARD, 2010) is given by:

$$Var\left(r_t \left| \mathcal{F}_{t-1}^{HF} \right.\right) \equiv h_t = w_H + \alpha_H RV_{t-1} + \beta_H h_{t-1}, \tag{4.1}$$

$$E\left(RV_t \left| \mathcal{F}_{t-1}^{HF} \right.\right) \equiv \mu_t = w_R + \alpha_R RV_{t-1} + \beta_R \mu_{t-1}, \tag{4.2}$$

for $t = 2, \ldots, T$, where $\mathcal{F}_{t-1}^{HF}$ stands for high-frequency information (intraday returns) from the previous day, and $w_H, \alpha, w_R, \alpha_R, \beta_R \geq 0$, $\beta_H \in [0, 1]$ and $\alpha_R + \beta_R \in [0, 1]$ are stability conditions. The first equation models the known as close-to-close conditional variance and is alike to a standard GARCH model[3], whereas the second equation models the conditional expectation of the open-to-close variation[4]. The parameters estimates are obtained by using quasi-maximum likelihood.

The implementation of the Copula-HEAVY strategy is similar to the Copula-GARCH approach from subsection 2.2.2. In the first step, we fit an appropriate ARMA-HEAVY model[5] to each univariate time series (daily returns of the formation period) by obtaining the estimates $\widehat{\mu}_i$ and $\widehat{\sigma}_i$ of the conditional mean and variance of these processes, respectively. Moreover, using the estimated parametric models, we construct the standardized residuals vectors given, for each $i = 1, ..., t$, by

$$\widehat{\varepsilon}_i = \frac{x_i - \widehat{\mu}_i}{\widehat{\sigma}_i}. \tag{4.4}$$

The standardized residuals vectors are then converted to the pseudo-observations $z_i = \frac{n}{n+1} F_i\left(\widehat{\varepsilon}_i\right)$, where $F_i$ is estimated by using their empirical distribution function. In the second step, with the estimated parameters from the previous step, we nominate the copula that best fits the uniform marginals and estimate its parameter(s). The copulas that are tested are Gaussian, Student's t, Clayton, Frank, Gumbel and one Archimedean mixture copula consisting of the best linear combination of Clayton, Frank and Gumbel copulas. We consider as the optimal copula the one with the lowest negative log-likelihood.

Finally, we follow the approach of Xie, Liew, Wu and Zou (2016) to obtain $MI_{X|Y}$ and $MI_{Y|X}$ as already described in subsection 2.2.2.

### 4.3   Empirical Results

An empirical study is carried out to evaluate and compare the performance of the distance and copula strategies. Such study is necessarily restrictive because there are many possibilities regarding the selection of the degree of pair-divergence, the trading

---

[3]The standard GARCH(1,1) is given by

$$Var\left(r_t \left| \mathcal{F}_{t-1}^{LF} \right.\right) = \sigma_t^2 = w_G + \alpha_G r_{t-1}^2 + \sigma_{t-1}^2, \tag{4.3}$$

where $\mathcal{F}_{t-1}^{LF}$ denotes the low-frequency information (daily returns) from the previous day. The GARCH equation models the close-to-close conditional variance.

[4]A Matlab code is available from Kevin Sheppard at `https://www.kevinsheppard.com/MFE_Toolbox`.

[5]We look for the optimal ARMA(p,q) model up to order (1,1).

costs, the number of pairs, the type of volatility model, the probability density functions, the method of pair selection, among other factors.

### 4.3.1 Profitability of the Strategies

Tables 1 and 2 exhibit annualized excess returns, annualized Sharpe and Sortino ratios, Newey and West (1987) adjusted t-statistics, share of negative observations, the maximum drawdown in terms of maximum percentage drop between two consecutive days (MDD1) and between two days within a period of maximum six months (MDD2) for each of the four strategies for from 2001-2016/1 assuming four equally spaced levels of transaction costs for Top 20 pairs. Table 1 summarize results when positions are initiated and closed in the same day the pair diverges and Table 2 when we delay trades by one day. Section 1 describes the Return on Committed Capital and Section 2 on Fully Invested Capital.

A series of important observations can be made from Table 1. First, notice that there exist strong evidence of the superiority of the Copula-GARCH strategy concerning the profitability of the strategies. The method yields an annual excess return up to 12.10% and 29.95% before costs on committed and fully invested capital, respectively and shows the highest Sharpe ratios (up to 3.16). Also note that the approach is the only strategy which delivers Sharpe ratios above 1 for all levels of transaction costs, indicating that investors are compensated for taking on additional risk. The Sortino ratio also confirms that the Copula-GARCH strategy delivers better risk-adjusted performance across-the-board. Second, we can observe that the excess returns from the model are always statistically and economically significant at 1% (Newey and West (1987) adjusted t-statistics are between 5.94 and 13.57). Furthermore, the Copula-GARCH approach shows a lower probability of a negative trade and it also presents the best results on committed capital when we focus on capital preservation over a period of six months based on maximum drawdown (MDD2) measure.

The Copula-HEAVY strategy also shows a consistent better performance than the distance methods for transaction costs up to 10 basis points in terms of wealth accumulation, wealth preservation and risk-adjusted returns. However, the approach does not generated significant excess returns when we set the transaction costs to be 15 basis points. In this case the distance strategies show better results than the Copula-HEAVY approach, particularly for $0.75\sigma$ trigger point.

**Table 1:** Excess returns of pairs trading strategies on portfolios of Top 20 pairs without delay.

| Strategy | Mean Return | Sharpe ratio | Sortino ratio | t-stat | % of negative trades | MDD1 | MDD2 |
|---|---|---|---|---|---|---|---|
| | | | **Section 1: Return on Committed Capital** | | | | |
| | | | *Panel A - Before Transaction Costs* | | | | |
| Distance (2.0$\sigma$) | 1.49 | 0.32 | 0.58 | 1.76* | 49.80 | -2.55 | -6.25 |
| Distance (0.75$\sigma$) | 3.10 | 0.57 | 1.06 | 2.99*** | 49.88 | -3.05 | -7.43 |
| Copula-GARCH | 12.10 | 3.16 | 7.21 | 13.57*** | 42.99 | -2.37 | -2.70 |
| Copula-HEAVY | 7.02 | 2.21 | 4.38 | 9.81*** | 44.35 | -2.34 | -3.30 |
| | | | *Panel B - Transaction Costs: 5 basis points* | | | | |
| Distance (2.0$\sigma$) | 1.37 | 0.29 | 0.53 | 1.63 | 49.88 | -2.55 | -6.26 |
| Distance (0.75$\sigma$) | 2.86 | 0.53 | 0.98 | 2.78*** | 50.07 | -3.05 | -7.52 |
| Copula-GARCH | 10.05 | 2.68 | 5.99 | 11.52*** | 45.06 | -2.37 | -3.66 |
| Copula-HEAVY | 4.82 | 1.55 | 2.98 | 6.90*** | 46.81 | -2.35 | -3.87 |
| | | | *Panel C - Transaction Costs: 10 basis points* | | | | |
| Distance (2.0$\sigma$) | 1.25 | 0.27 | 0.49 | 1.50 | 49.95 | -2.55 | -6.29 |
| Distance (0.75$\sigma$) | 2.62 | 0.48 | 0.90 | 2.56** | 50.12 | -3.05 | -7.60 |
| Copula-GARCH | 8.05 | 2.19 | 4.79 | 9.42*** | 47.23 | -2.38 | -4.38 |
| Copula-HEAVY | 2.67 | 0.88 | 1.64 | 3.93*** | 49.83 | -2.37 | -4.58 |
| | | | *Panel D - Transaction Costs: 15 basis points* | | | | |
| Distance (2.0$\sigma$) | 1.13 | 0.24 | 0.45 | 1.36 | 50.02 | -2.55 | -6.34 |
| Distance (0.75$\sigma$) | 2.38 | 0.44 | 0.82 | 2.35** | 50.22 | -3.06 | -7.69 |
| Copula-GARCH | 6.08 | 1.69 | 3.61 | 7.27*** | 49.78 | -2.39 | -5.19 |
| Copula-HEAVY | 0.58 | 0.19 | 0.37 | 0.91 | 52.47 | -2.38 | -5.54 |
| | | | **Section 2: Return on Fully Invested Capital** | | | | |
| | | | *Panel A - Before Transaction Costs* | | | | |
| Distance (2.0$\sigma$) | 2.29 | 0.31 | 0.60 | 1.76* | 49.80 | -3.61 | -8.49 |
| Distance (0.75$\sigma$) | 3.62 | 0.58 | 1.10 | 3.00*** | 49.88 | -3.52 | -8.85 |
| Copula-GARCH | 29.95 | 3.07 | 6.55 | 13.04*** | 42.99 | -4.24 | -6.96 |
| Copula-HEAVY | 17.19 | 2.19 | 4.20 | 9.77*** | 44.35 | -3.86 | -9.68 |
| | | | *Panel B - Transaction Costs: 5 basis points* | | | | |
| Distance (2.0$\sigma$) | 2.13 | 0.29 | 0.56 | 1.65* | 49.88 | -3.61 | -8.50 |
| Distance (0.75$\sigma$) | 3.35 | 0.54 | 1.02 | 2.79*** | 50.07 | -3.53 | -8.94 |
| Copula-GARCH | 23.68 | 2.52 | 5.28 | 10.74*** | 45.06 | -4.25 | -8.40 |
| Copula-HEAVY | 11.29 | 1.49 | 2.81 | 6.70*** | 46.77 | -3.88 | -10.93 |
| | | | *Panel C - Transaction Costs: 10 basis points* | | | | |
| Distance (2.0$\sigma$) | 1.96 | 0.27 | 0.52 | 1.54 | 49.95 | -3.61 | -8.51 |
| Distance (0.75$\sigma$) | 3.08 | 0.50 | 0.94 | 2.59*** | 50.12 | -3.53 | -9.04 |
| Copula-GARCH | 17.72 | 1.96 | 4.04 | 8.37*** | 47.23 | -4.26 | -10.55 |
| Copula-HEAVY | 5.70 | 0.78 | 1.46 | 3.57*** | 49.78 | -3.91 | -12.25 |
| | | | *Panel D - Transaction Costs: 15 basis points* | | | | |
| Distance (2.0$\sigma$) | 1.80 | 0.24 | 0.48 | 1.42 | 50.02 | -3.62 | -8.53 |
| Distance (0.75$\sigma$) | 2.81 | 0.45 | 0.86 | 2.38** | 50.22 | -3.53 | -9.13 |
| Copula-GARCH | 12.04 | 1.38 | 2.81 | 5.94*** | 49.68 | -4.28 | -12.84 |
| Copula-HEAVY | 0.40 | 0.06 | 0.16 | 0.40 | 52.47 | -3.93 | -13.82 |

*Note:* Summary statistics of the annualized excess returns, annualized Sharpe and Sortino ratios on portfolios of top 20 pairs between January 2001 and June 2016 (4050 observations). The positions are initiated in the same day the pair diverges. The t-statistics are computed using Newey-West standard errors with six-lag correction. The positions are initiated in the same day the pair diverges. 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.
Source: Author's own elaboration (2016).

It should be noted that the excess returns of copula strategies are considerably more affected by increasing transactions costs. Overall, the average excess return of copula approaches are reduced by around 2% on committed capital for each extra 5 basis points attached to trade. On the other hand, the distance strategies profits are reduced by less than 0.25% for each 5 basis points added. Considering this trend, it would be required more than 25 basis points to make the Copula-GARCH model less profitable than the distance strategies.

**Speed of execution**

So far we have assumed that a trade can be started as soon as the trigger points are reached. However, one may face situations where the price of a stock bounces very quickly back and forth affecting how fast a trade can be executed. To investigate the robustness of our results to bid-ask bounce (JEGADEESH, 1990; JEGADEESH; TITMAN, 1995; CONRAD; KAUL, 1989) we follow Gatev, Goetzmann and Rouwenhorst (2006) and use a "one-day rule", *i.e.*, we evaluate the performance of the strategies when positions are initiated on the day after the price divergence and closed on the day after the convergence.

Table 2 displays that the copula strategies are, likewise after transaction costs, more sensitive to the speed of execution than the distance approaches. The Copula-GARCH model still shows a better performance than the other strategies and delivers profits before costs. However, the trading profits do not survive after costs. The $2.0\sigma$ trigger point shows very modest and not significant profits over the period.

Thus, our results imply that a significant portion of the returns for copula approaches occur in the first day of trading and therefore may be due to bid-ask bounce. However, as highlighted by Gatev, Goetzmann and Rouwenhorst (2006), it is difficult to quantify which portion of this reduction is due to bid-ask bounce and which portion is derived from the true mean reversion in prices because of fast market adjustment.

**Table 2:** Excess returns of pairs trading strategies on portfolios of Top 20 pairs with one day waiting period.

| Strategy | Mean Return | Sharpe ratio | Sortino ratio | t-stat | % of negative trades | MDD1 | MDD2 |
|---|---|---|---|---|---|---|---|
| **Section 1: Return on Committed Capital** | | | | | | | |
| *Panel A - Before Transaction Costs* | | | | | | | |
| Distance (2.0σ) | 0.38 | 0.09 | 0.18 | 0.56 | 49.90 | -2.21 | -6.93 |
| Distance (0.75σ) | 0.01 | 0.00 | 0.05 | 0.14 | 50.74 | -2.37 | -8.39 |
| Copula-GARCH | 1.31 | 0.39 | 0.67 | 1.86* | 49.21 | -2.04 | -5.75 |
| Copula-HEAVY | -0.34 | -0.12 | -0.16 | -0.50 | 49.31 | -2.34 | -4.98 |
| *Panel B - Transaction Costs: 5 basis points* | | | | | | | |
| Distance (2.0σ) | 0.26 | 0.06 | 0.13 | 0.42 | 49.98 | -2.21 | -7.01 |
| Distance (0.75σ) | -0.22 | -0.04 | -0.03 | -0.10 | 50.84 | -2.37 | -8.53 |
| Copula-GARCH | -0.54 | -0.16 | -0.24 | -0.67 | 52.54 | -2.07 | -6.61 |
| Copula-HEAVY | -2.40 | -0.82 | -1.23 | −3.95*** | 53.58 | -2.36 | -5.91 |
| *Panel C - Transaction Costs: 10 basis points* | | | | | | | |
| Distance (2.0σ) | 0.14 | 0.03 | 0.09 | 0.28 | 50.05 | -2.21 | -7.08 |
| Distance (0.75σ) | -0.44 | -0.09 | -0.11 | -0.34 | 50.94 | -2.37 | -8.66 |
| Copula-GARCH | -2.36 | -0.72 | -1.12 | −3.21*** | 55.16 | -2.10 | -7.51 |
| Copula-HEAVY | -4.40 | -1.53 | -2.25 | −7.36*** | 56.99 | -2.37 | -7.01 |
| *Panel D - Transaction Costs: 15 basis points* | | | | | | | |
| Distance (2.0σ) | 0.02 | 0.00 | 0.04 | 0.13 | 50.10 | -2.21 | -7.16 |
| Distance (0.75σ) | -0.67 | -0.14 | -0.19 | -0.59 | 51.14 | -2.37 | -8.79 |
| Copula-GARCH | -4.15 | -1.28 | -1.99 | −5.75*** | 57.48 | -2.13 | -8.46 |
| Copula-HEAVY | -6.36 | -2.23 | -3.24 | −10.71*** | 59.63 | -2.39 | -8.11 |
| **Section 2: Return on Fully Invested Capital** | | | | | | | |
| *Panel A - Before Transaction Costs* | | | | | | | |
| Distance (2.0σ) | 0.50 | 0.07 | 0.18 | 0.54 | 49.95 | -4.01 | -10.87 |
| Distance (0.75σ) | -0.21 | -0.04 | -0.02 | -0.05 | 50.35 | -2.94 | -9.78 |
| Copula-GARCH | 1.75 | 0.18 | 0.38 | 1.01 | 49.48 | -4.75 | -19.26 |
| Copula-HEAVY | -2.79 | -0.27 | -0.34 | -0.98 | 49.21 | -6.19 | -16.80 |
| *Panel B - Transaction Costs: 5 basis points* | | | | | | | |
| Distance (2.0σ) | 0.34 | 0.05 | 0.14 | 0.42 | 50.02 | -4.01 | -10.95 |
| Distance (0.75σ) | -0.45 | -0.08 | -0.09 | -0.27 | 50.40 | -2.95 | -9.89 |
| Copula-GARCH | -1.03 | -0.11 | -0.10 | -0.27 | 51.14 | -4.75 | -19.76 |
| Copula-HEAVY | -5.17 | -0.51 | -0.71 | −2.06** | 51.80 | -6.19 | -17.55 |
| *Panel C - Transaction Costs: 10 basis points* | | | | | | | |
| Distance (2.0σ) | 0.18 | 0.03 | 0.10 | 0.31 | 50.07 | -4.01 | -11.03 |
| Distance (0.75σ) | -0.69 | -0.12 | -0.16 | -0.49 | 50.47 | -2.95 | -10.01 |
| Copula-GARCH | -3.73 | -0.40 | -0.59 | -1.57 | 52.10 | -4.76 | -20.25 |
| Copula-HEAVY | -7.49 | -0.74 | -1.08 | −3.15*** | 52.86 | -6.19 | -18.52 |
| *Panel D - Transaction Costs: 15 basis points* | | | | | | | |
| Distance (2.0σ) | 0.02 | 0.00 | 0.06 | 0.19 | 50.12 | -4.01 | -11.11 |
| Distance (0.75σ) | -0.93 | -0.17 | -0.23 | -0.71 | 50.67 | -2.95 | -10.13 |
| Copula-GARCH | -6.36 | -0.70 | -1.07 | −2.88*** | 53.51 | -4.76 | -20.74 |
| Copula-HEAVY | -9.76 | -0.98 | -1.45 | −4.24*** | 54.10 | -6.19 | -19.48 |

*Note:* Summary statistics of the annualized excess returns, annualized Sharpe and Sortino ratios on portfolios of top 20 pairs between January 2001 and June 2016 (4050 observations). We assume a one day waiting period after the pair diverges. The t-statistics are computed using Newey-West standard errors with 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.
Author's own elaboration (2016).

Cumulative excess returns over the full data set for each of the strategies with no delay and with "one day waiting rule" are displayed in Figures 4.1 and 4.2, respectively. Panels (a) to (h) show the returns from zero to fifteen basis points as transaction costs on committed (left panels) and fully invested capital (right panels), respectively. The conclusions drawn from the graphical inspection of the results are consistent with our quantitative counterparts for the mean excess returns and t-statistics provided in Tables 1 and 2, thus reinforcing our empirical findings.

**Figure 4.1:** Cumulative excess returns of the pairs trading strategies on portfolios of Top 20 pairs with no delay
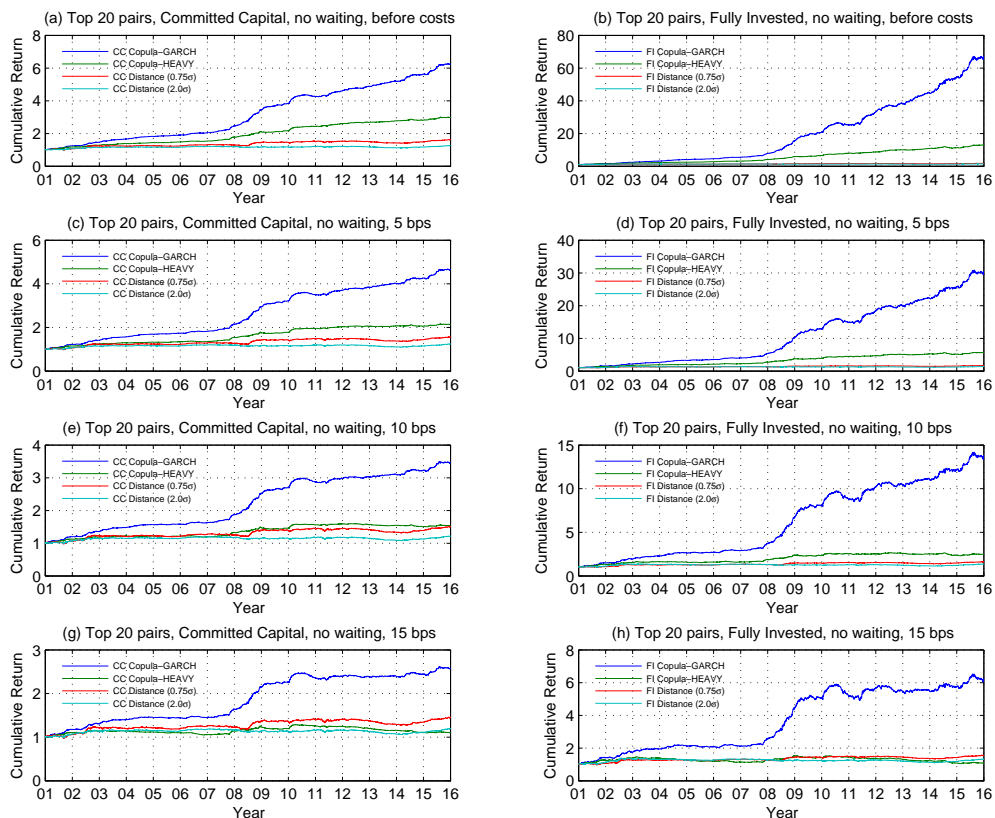


Source: Author's own elaboration (2016.)

This figure shows how an investment of $1 evolves from January 2001 to June 2016 for each of the strategies

**Figure 4.2:** Cumulative excess returns of pairs trading strategies on portfolios of Top 20 pairs with one day waiting period
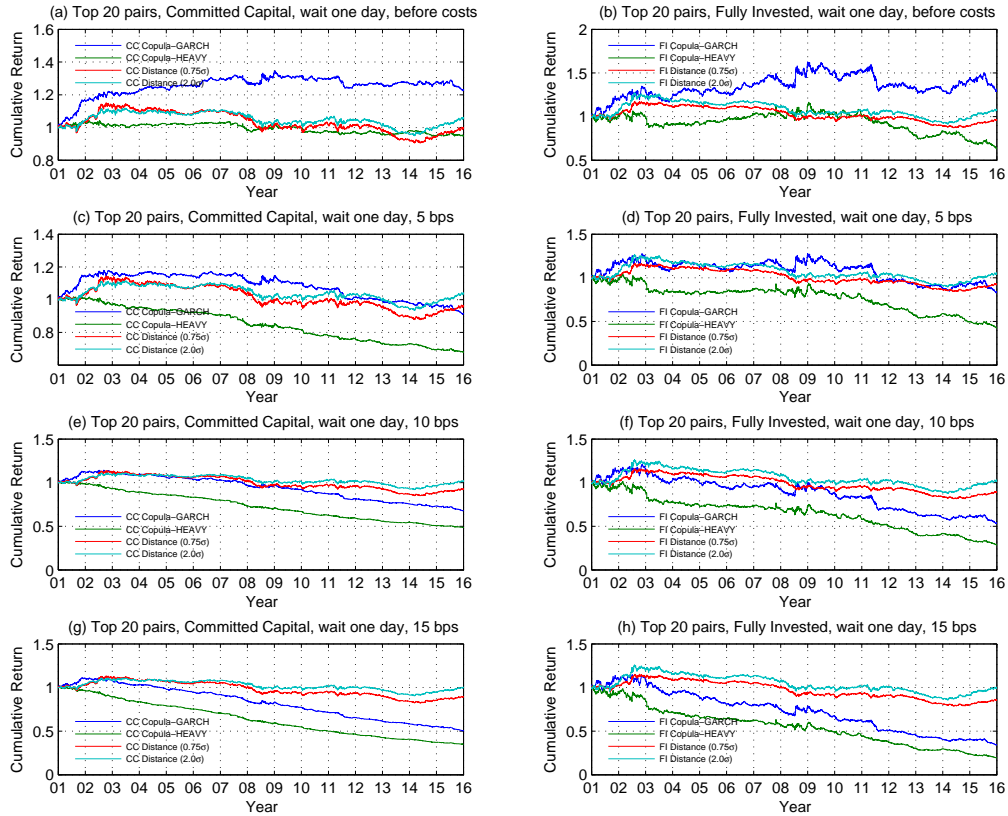


Source: Author's own elaboration (2016.)

This figure shows how an investment of $1 evolves from January 2001 to June 2016 for each of the strategies

### 4.3.2 Trading statistics

Table 3 provides trading statistics. The average price deviation trigger for opening pairs are listed in the first row of the panel for each strategy. The positions typically opens when prices have diverged by 5.64%, 3.15%, 5.32% and 5.23% or more for $2.0\sigma$, $0.75\sigma$, Copula-GARCH and Copula-HEAVY strategies, respectively. It is noticeable that the copula approaches generate more trading opportunities than the distance strategies, *i.e.*, the average number of pairs open per period is notably higher across the board for copula strategies. Furthermore, when using the copula framework, the pairs stay open for a shorter period. This is a very important finding, since every trading signal is an opportunity to profit. The average number of pairs traded per six-month period is at least 8.35 times more for our copula strategies than for the distance approaches considered here. In addition, under copula rules, each pair is held open, in average, 2.48 and 2.33 trading days, indicating that the copula approaches are short-term strategies. Meanwhile, the average holding period for the distance approaches is at least 2.37 trading months which indicate that the methods are medium-term strategies in this case.

**Table 3:** Trading statistics.

| Strategy | Distance (2.0$\sigma$) | Distance (0.75$\sigma$) | Copula-GARCH | Copula-HEAVY |
|---|---|---|---|---|
| Average price deviation trigger for opening pairs | 0.0564 | 0.0315 | 0.0532 | 0.0523 |
| Total number of pairs opened | 758 | 1433 | 11965 | 13456 |
| Average number of pairs traded per six-month period | 24.45 | 46.23 | 385.97 | 434.06 |
| Average number of round-trip trades per pair | 1.22 | 2.31 | 19.30 | 21.70 |
| Standard Deviation | 0.70 | 1.78 | 6.12 | 13.33 |
| Average time pairs are open in days | 69.92 | 49.83 | 2.48 | 2.33 |
| Standard Deviation | 44.49 | 47.37 | 2.32 | 3.13 |
| Median time pairs are open in days | 65.5 | 30 | 2 | 1 |

*Note:* Trading statistics for portfolio of top 20 pairs between January 2001 and June 2016 (31 periods). Pairs are formed over a 12-month period according to a minimum-distance 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.

Source: Author's own elaboration (2016).

### 4.3.3 Robustness Checks of the Performance of Excess Returns and Sharpe Ratios

In order to mitigate data-snooping criticisms, we use the stationary bootstrap of Politis and Romano (1994) to compute the bootstrap p-values using the methodology proposed by Ledoit and Wolf (2008).

We want to further investigate if the average excess return and the Sharpe Ratio of the Copula-GARCH method show a significative superior out-of-sample performance compared to the distance and Copula-HEAVY strategies.

In order to construct the distributions we bootstrapped the original time series $B = 10,000$ times. Our bootstrapped null distributions result from Theorem 2 of Politis and Romano (1994). We select the optimal block length for the stationary bootstrap following Politis and White (2004). As the optimal bootstrap block length is different for each strategy, we average the block lengths found to perform the comparisons.

To test the hypotheses that the average excess returns and Sharpe Ratios of the Copula-GARCH (cg) model are equal to that of distance (d) and Copula-HEAVY (ch) approaches, that is,

$$H_0 : \mu_{cg} - \mu_d = 0 \quad \text{and} \quad H_0 : \frac{\mu_{cg}}{\sigma_{cg}} - \frac{\mu_d}{\sigma_d} = 0, \tag{4.5}$$

and

$$H_0 : \mu_{cg} - \mu_{ch} = 0 \quad \text{and} \quad H_0 : \frac{\mu_{cg}}{\sigma_{cg}} - \frac{\mu_{ch}}{\sigma_{ch}} = 0, \tag{4.6}$$

we compute, following Davison and Hinkley (1997), a two-sized $p$-value using $B = 10000$ (stationary) bootstrap re-samples.

Table 4 provides the bootstrap p-values for testing the null hypotheses represented by (4.5) and (4.6). Sections 1 and 2 compare the Copula-GARCH strategy to the 2.0$\sigma$ and

$0.75\sigma$ trigger points, respectively, and Section 3 to the Copula-HEAVY approach from 2001-2016/1 for all investment scenarios, *i.e.*, without delay and waiting one day period and before and after costs.

As can be deduced from Table 4 the Copula-GARCH model significantly outperforms the distance strategies when positions are initiated and closed in the same day the pair diverges (p-values < 0.5%) for both weighting structures. On the other hand, the distance methods provide profits and Sharpe ratios significantly higher for wait one-day rule and 15 basis points at least at 10% for both weighting schemes. The Copula-GARCH method also show a significant superior performance compared to the Copula-HEAVY approach (p-values less than 0.0001) when a fast execution of the trade can be made and for one-day waiting trading strategy on committed capital, at least at 5%.

**Table 4:** Bootstrap p-values computed from B=10,000 replications for testing the null hypotheses of equality of the average excess returns and Sharpe Ratios over the period between January 2001 and June 2016.

| Scenario | Committed Capital | | Fully Invested Capital | |
|---|---|---|---|---|
| | Return | Sharpe Ratio | Return | Sharpe Ratio |
| Section 1: Copula-GARCH versus $2.0\sigma$ | | | | |
| d0c0 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c5 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c10 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c15 | 0.0006***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d1c0 | 0.3900 | 0.2792 | 0.5748 | 0.6934 |
| d1c5 | 0.4036 | 0.3918 | 0.6358 | 0.6130 |
| d1c10 | 0.0134**(<) | 0.0062***(<) | 0.1258 | 0.1522 |
| d1c15 | 0.0000***(<) | 0.0000***(<) | 0.0096***(<) | 0.0186**(<) |
| Section 2: Copula-GARCH versus $0.75\sigma$ | | | | |
| d0c0 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c5 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c10 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c15 | 0.0048***(>) | 0.0000***(>) | 0.0000***(>) | 0.0008***(>) |
| d1c0 | 0.2562 | 0.1510 | 0.3490 | 0.4118 |
| d1c5 | 0.7138 | 0.6324 | 0.9116 | 0.9874 |
| d1c10 | 0.0620*(<) | 0.0182**(<) | 0.2284 | 0.3592 |
| d1c15 | 0.0010***(<) | 0.0000***(<) | 0.0218**(<) | 0.0704*(<) |
| Section 3: Copula-GARCH versus Copula-HEAVY | | | | |
| d0c0 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c5 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c10 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c15 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d1c0 | 0.0134**(>) | 0.0138**(>) | 0.0826*(>) | 0.0820*(>) |
| d1c5 | 0.0028***(>) | 0.0010***(>) | 0.1026 | 0.1250 |
| d1c10 | 0.0010***(>) | 0.0000***(>) | 0.1246 | 0.1940 |
| d1c15 | 0.0000***(>) | 0.0000***(>) | 0.1650 | 0.2920 |

*Note:* This table reports the bootstrap p-values for testing the null hypothesis of equality of the average excess returns and the Sharpe Ratios of Copula-GARCH and distance and Copula-HEAVY strategies for Top 20 pairs over the period between January 2001 and June 2016 (4050 observations). The column labeled Scenario contains symbol labels for trading with no delay or one day waiting period (d0 and d1, respectively) and before or after costs (c0, c5, c10 and c15, respectively). The symbol labels (>) and (<) indicate that the respective null hypothesis is rejected in favor of the alternative and that the average excess returns and/or the Sharpe Ratio of the Copula-GARCH strategy are found greater or less than the the average excess returns and/or the Sharpe Ratio of the distance and Copula-HEAVY strategies, respectively.

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

Author's own elaboration (2016).

### 4.3.4 Sub-period analysis

Due to the sample horizon over 15 years, a sub-period analysis is inevitable to examine whether the profitability pattern has changed over time. We split the full sample period into four subperiods: (1) January 2001 to December 2002, (2) January 2003 to June 2007, (3) July 2007 to June 2009, and (4) July 2009 to June 2016. The first subperiod corresponds to the bear market that comprises the dotcom crisis and the September 11th terrorist attack. The third subperiod corresponds to the global financial crisis.

Figures 4.3 to 4.6 show annualized excess returns and Sharpe ratios of the four strategies for each subperiod with no delay and one-day waiting period, respectively. Panels (a) to (h) show the returns from zero to fifteen basis points as transaction costs on committed (left panels) and fully invested capital (right panels), respectively.

As can be discerned analyzing figures 4.3 and 4.4 the Copula-GARCH strategy is consistently the most profitable approach among the four strategies and which produces the highest risk-adjusted returns with no delay to execute the required trades for both weighting structures. Moreover, it is remarkable that the Copula-GARCH method achieves its best performance over the two periods of higher stock market volatility, *i.e.*, during financial downturns that comprise the first and the third subperiods. This is a desirable finding for a market neutral strategy. In addition, the Copula-GARCH method shows Sharpe ratios always above 2.5 over these two subperiods (they range between 2.56 and 3.92 for January 2001 to December 2002 and between 3.46 and 4.55 during the subprime mortgage crisis).

Further, as can be seen from figures 4.3 and 4.4, the Copula-HEAVY and the 0.75 standard deviation approaches also achieve their best results during the subperiods of higher volatility, although the magnitude of the gains are relatively smaller for the distance strategy during the subprime mortgage crisis.

**Figure 4.3:** Average excess returns of pairs trading strategies for each subperiod with no delay
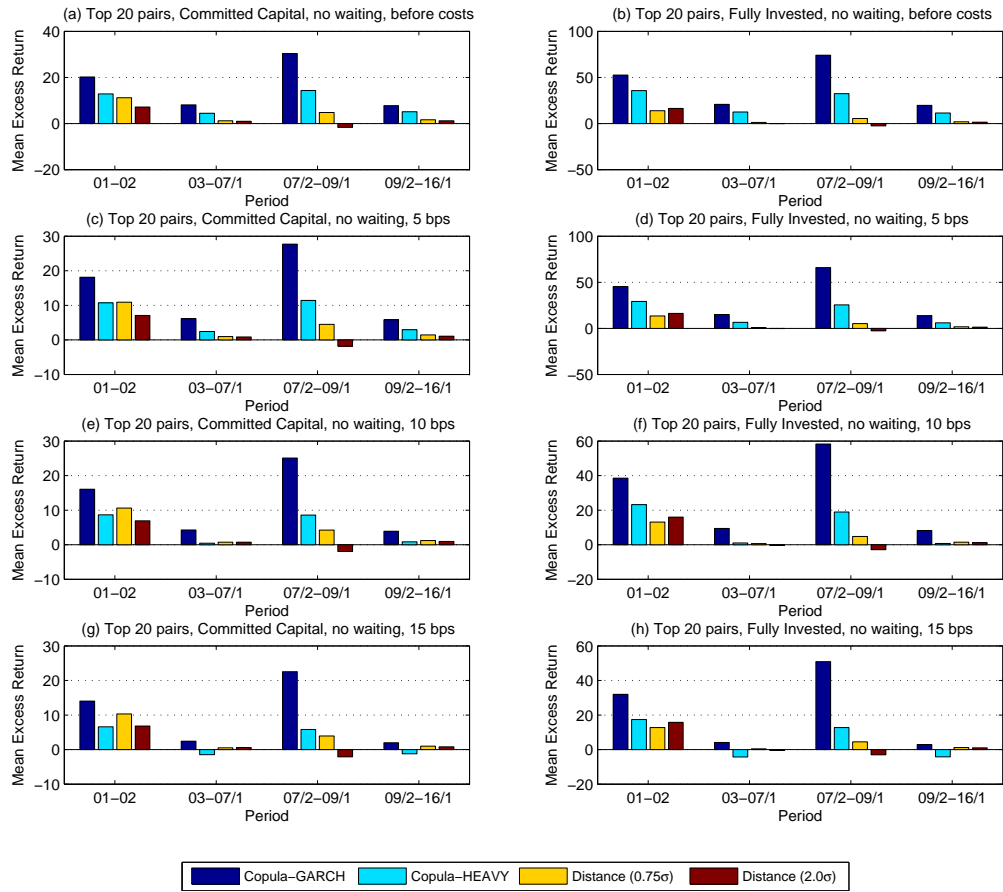


Source: Author's own elaboration (2016.)

**Figure 4.4:** Sharpe Ratio (annualized) of pairs trading strategies for each subperiod with no delay
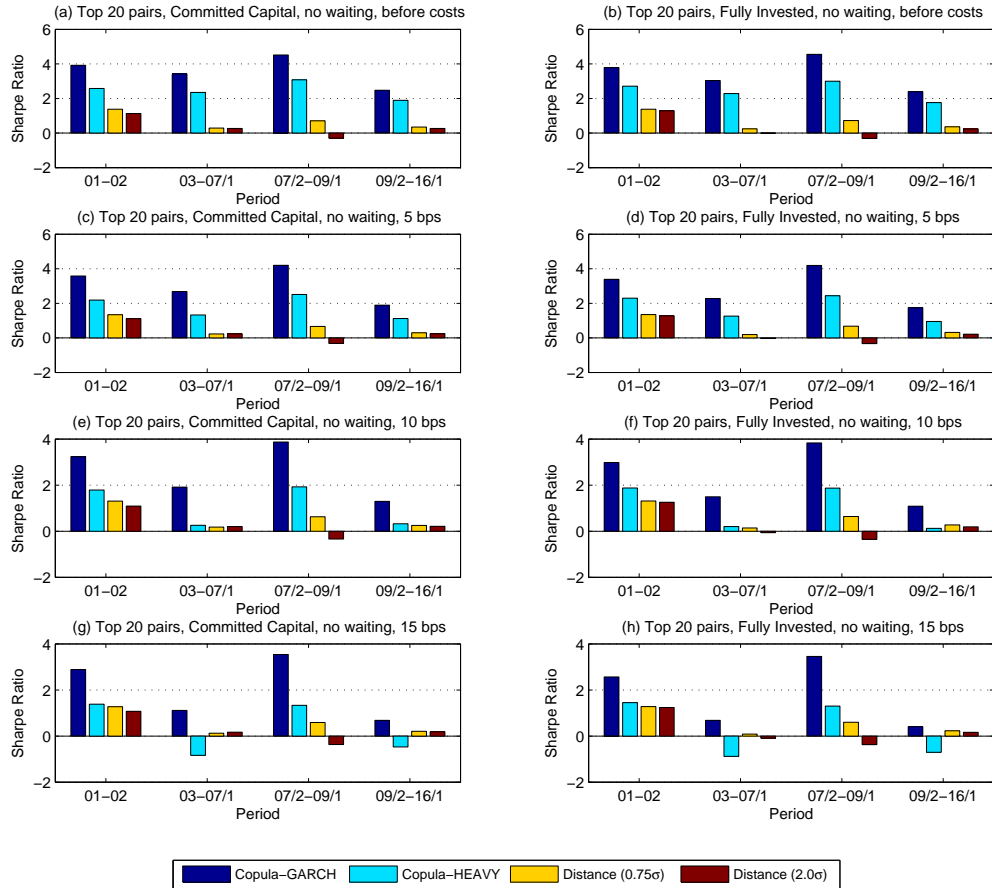

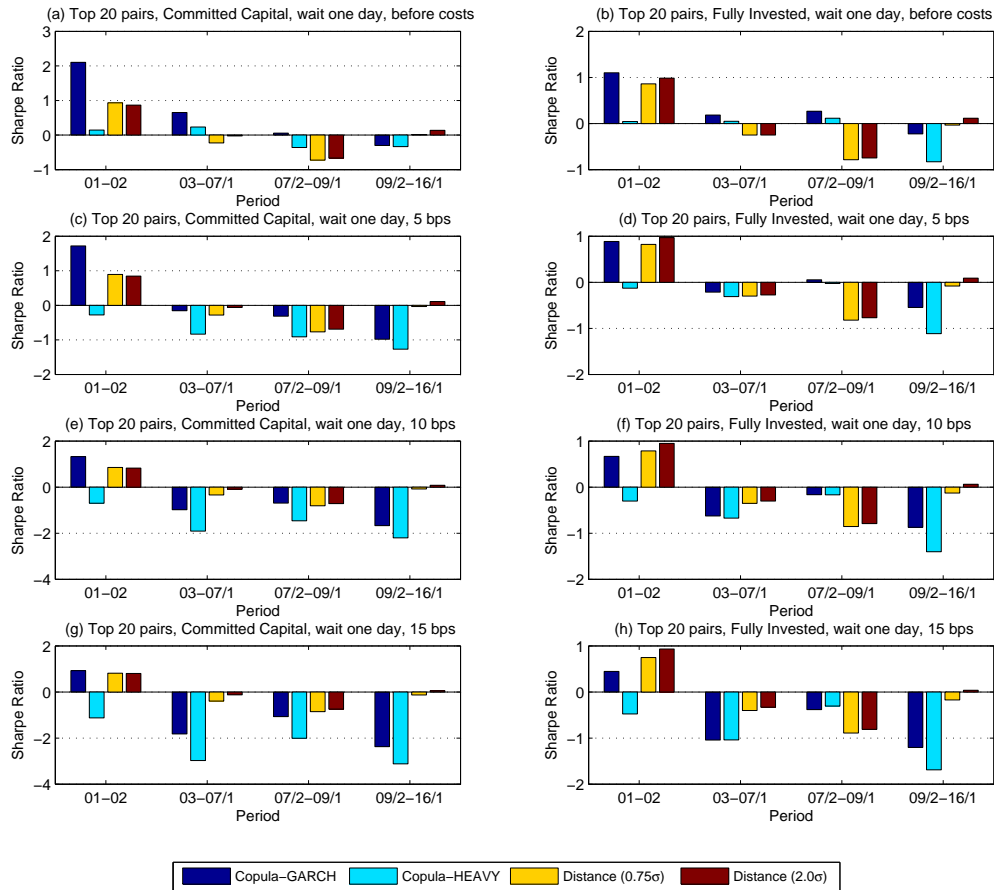
Source: Author's own elaboration (2016.)

Figures 4.5 and 4.6 show annualized excess returns and Sharpe ratios of the four strategies for each subperiod with one-day waiting period rule. As can be seen the Copula-HEAVY and the distance strategies are clearly more profitable in the first subperiod. However, there is a substantial decrease of profitability and risk-adjusted returns during the global financial crisis. Hence, it can be deduced that a significant part of the returns of the strategies occur in the first day of trading.

**Figure 4.5:** Average excess returns of pairs trading strategies for each subperiod with one day waiting period



(a) Top 20 pairs, Committed Capital, wait one day, before costs

(b) Top 20 pairs, Fully Invested, wait one day, before costs

(c) Top 20 pairs, Committed Capital, wait one day, 5 bps

(d) Top 20 pairs, Fully Invested, wait one day, 5 bps

(e) Top 20 pairs, Committed Capital, wait one day, 10 bps

(f) Top 20 pairs, Fully Invested, wait one day, 10 bps

(g) Top 20 pairs, Committed Capital, wait one day, 15 bps

(h) Top 20 pairs, Fully Invested, wait one day, 15 bps

Copula–GARCH   Copula–HEAVY   Distance (0.75σ)   Distance (2.0σ)
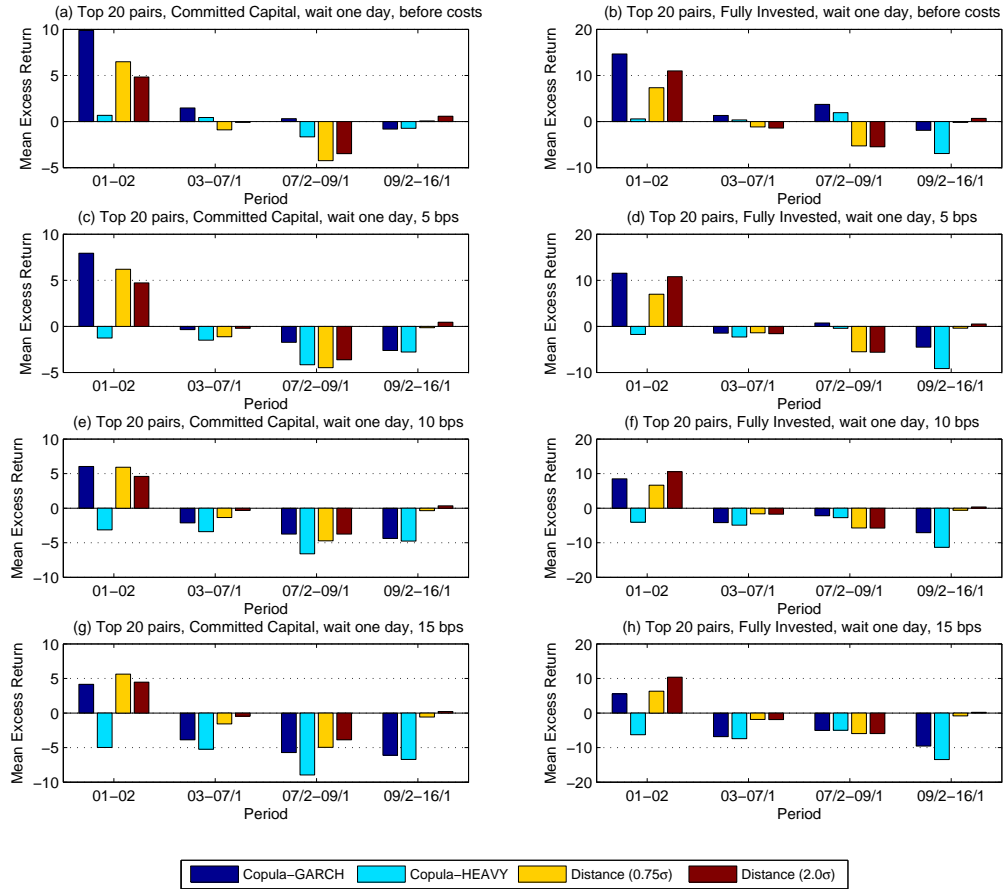
Source: Author's own elaboration (2016.)

**Figure 4.6:** Sharpe Ratio (annualized) pairs trading strategies for each subperiod with one day waiting period



Source: Author's own elaboration (2016.)

Robustness checks of the performance of annualiazed excess returns and Sharpe ratios over the four subperiods are presented in in Tables A.1-A.4 in the Appendix.

The results appear to remain consistent over the subperiods enhancing that the Copula-GARCH pairs strategy outperforms the other strategies when we initiate and close positions in the same day the pair diverges and converges, respectively.

## 4.4 Conclusions

The main objective of this chapter is to compare alternative copula pairs trading estimation procedures to understand better the factors that affect the profitability of the strategies. The main findings are summarized below.

1. Results suggest that the Copula-GARCH pairs trading strategy is consistently the best rewarding strategy when a fast trade can be made.

2. The Copula-HEAVY pairs trading strategy also shows a consistent better performance than the distance strategies for transaction costs up to 10 basis points in terms of wealth accumulation, wealth preservation and risk-adjusted returns when positions are initiated and closed in the same day the pair diverges and converges, respectively.

3. Results suggest that in the whole sample period the high-frequency based volatility model does not produce higher mean excess returns than the standard GARCH(1,1) model. It is possible that the estimation of more parameters leads to high estimation errors and therefore false transaction signals, thus decreasing the profitability of the Copula-HEAVY method.

4. The excess returns from the Copula-GARCH method are always statistically and economically significant at 1% when a fast execution of the trade can be made for all levels of transaction costs in this study. For the Copula-HEAVY approach the excess returns are not significant when the transaction costs are set to 15 basis points.

5. The Copula-GARCH strategy shows a lower probability of a trade with negative returns compared to the other approaches and it also displays the best results on committed capital when we focus on a steady performance and capital preservation over a period of six months.

6. Without delay to start and close the positions the Copula-GARCH approach is the only strategy which produces Sharpe ratios above 1 for all levels of transaction costs, indicating that investors are compensated for taking on additional risk. The subperiod analysis also endorses that the strategy yields better risk-adjusted performance.

7. The copula-based pairs strategies also present more trading opportunities than the distance strategies. However, the average excess returns of copula strategies are more sensitive to timing and transaction costs. When allowing a one-day waiting period restriction on execution, the profits of the copula-based pairs strategy are eliminated after costs. Hence, it can be deduced that a significant part of the returns of the strategies occur in the first day of trading.

8. It is remarkable that the copula approaches achieved its very best performance over periods of higher volatility, *i.e.*, during financial downturns. This is a desirable finding for a market neutral strategy.

9. The copula approaches are able to identify more trading opportunities than the distance strategies. Moreover, when using the copula methods the pairs stay open for a shorter period of time. This is a very important finding, since every trading signal is an opportunity to profit.

Finally, we provide some suggestions for future research to improve our findings. First, we must become the comparisons more reliable and meaningful making the number of tradeable signals more comparable. We should investigate an overall degree of relative mispricing as in Rad, Low and Faff (2015). Second, in one attempt to understand the economic drivers behind our data and to evaluate whether the profitability is a compensation for risk, we could regress daily excess return onto various asset pricing factors (see Fama and French (2012)). Third, we could employ a mixed copula model to cover a wider range of possible dependencies structures. Fourth, we could create similar copula-based arbitrage for triplets to increase information dependency information and measure relative pricing more comprehensively. Fifth, we could borrow some ideas from statistical learning and perform a model validation during the formation period. After choosing the pairs, we could check its predictive power by splitting data into two subsets. The idea is to select and fit the model based entirely on one part of the data, and then test it on the remaining part. Sixth, we could improve the method of pairs selection before using the copula approach. Since copulas capture better nonlinear dependence structures, we suspect that if we use a nonlinear association measurement such as the randomized dependency coefficient (Lopez-Paz et al., 2013), or a procedure based on data mining tools as random forest (Giovanni De Luca and Zuccolotto, 2010) to select the pairs, the performance of the copula method may be enhanced.

**REFERENCES**

ANDERSEN, Torben G et al. The distribution of realized stock return volatility. **Journal of financial economics**, Elsevier, v. 61, n. 1, p. 43–76, 2001.

ANDREOU, Elena; PITTIS, Nikitas; SPANOS, Aris. On modelling speculative prices: the empirical literature. **Journal of Economic Surveys**, v. 15.2, p. 187–220, 2001.

ANE, Thierry; KHAROUBI, Cécile. Dependence Structure and Risk Measure. **The journal of business**, v. 76, n. 3, p. 411–438, 2003.

BOWEN, David; HUTCHINSON, Mark C; O'SULLIVAN, Niall. High-frequency equity pairs trading: transaction costs, speed of execution, and patterns in returns. **The Journal of Trading**, v. 5, n. 3, p. 31–38, 2010.

BROUSSARD, John Paul; VAIHEKOSKI, Mika. Profitability of pairs trading strategy in an illiquid market with multiple share classes. **Journal of International Financial Markets, Institutions and Money**, v. 22.5, p. 1188–1201, 2012.

CAMPBELL, John Y; LO, Andrew Wen-Chuan; MACKINLAY, Archie Craig. **The econometrics of financial markets**. [S.l.]: princeton University press, 1997.

CONRAD, Jennifer; KAUL, Gautam. Mean reversion in short-horizon expected returns. **Review of Financial Studies**, v. 2.2, p. 225–240, 1989.

CONT, Rama. Empirical properties of asset returns: stylized facts and statistical issues. Taylor & Francis, 2001.

DAVISON, Anthony Christopher; HINKLEY, David Victor. **Bootstrap methods and their application**. Cambridge, UK: Cambridge university press, 1997. v. 1.

DO, Binh; FAFF, Robert. Are pairs trading profits robust to trading costs? **Journal of Financial Research**, v. 35.2, p. 261–287, 2012.

_____. Does simple pairs trading still work? **Financial Analysts Journal**, v. 66.4, p. 83–95, 2010.

DUNIS, Christian L et al. Statistical arbitrage and high-frequency data with an application to Eurostoxx 50 equities. **Liverpool Business School, Working paper**, 2010.

FAMA, Eugene F.; FRENCH, Kenneth R. Common risk factors in the returns on stocks and bonds. **Journal of financial economics**, v. 33.1, p. 3–56, 1993.

FAMA, Eugene F; FRENCH, Kenneth R. Size, value, and momentum in international stock returns. **Journal of financial economics**, Elsevier, v. 105, n. 3, p. 457–472, 2012.

FLEMING, Jeff; KIRBY, Chris; OSTDIEK, Barbara. The economic value of volatility timing using "realized" volatility. **Journal of Financial Economics**, Elsevier, v. 67, n. 3, p. 473–509, 2003.

GATEV, Evan; GOETZMANN, William N.; ROUWENHORST, K. Geert. Pairs Trading: Performance of a Relative-Value Arbitrage Rule. **Review of Financial Studies**, v. 19, n. 3, p. 797–827, 2006.

HAFNER, Christian M. Estimating high-frequency foreign exchange rate volatility with nonparametric ARCH models. **Journal of Statistical Planning and Inference**, v. 68.2, p. 247–269, 1998.

HANSEN, Peter Reinhard; HUANG, Zhuo; SHEK, Howard Howan. Realized garch: a joint model for returns and realized measures of volatility. **Journal of Applied Econometrics**, Wiley Online Library, v. 27, n. 6, p. 877–906, 2012.

JEGADEESH, Narasimhan. Evidence of predictable behavior of security returns. **The Journal of Finance**, v. 45.3, p. 881–898, 1990.

JEGADEESH, Narasimhan; TITMAN, Sheridan. Short-horizon return reversals and the bid-ask spread. **Journal of Financial Intermediation**, Elsevier, v. 4, n. 2, p. 116–132, 1995.

LEDOIT, Oliver; WOLF, Michael. Robust performance hypothesis testing with the Sharpe ratio. **Journal of Empirical Finance**, v. 15.5, p. 850–859, 2008.

LIEW, Rong Qi; WU, Yuan. Pairs trading: A copula approach. **Journal of Derivatives & Hedge Funds**, Springer, v. 19, n. 1, p. 12–30, 2013.

LIU, Lily Y; PATTON, Andrew J; SHEPPARD, Kevin. Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. **Journal of Econometrics**, Elsevier, v. 187, n. 1, p. 293–311, 2015.

MCNEIL, Alexander J; FREY, Rüdiger; EMBRECHTS, Paul. **Quantitative risk management: Concepts, techniques and tools**. [S.l.]: Princeton university press, 2015.

NATH, Purnendu. High frequency pairs trading with us treasury securities: Risks and rewards for hedge funds, 2003.

NEWEY, Whitney K; WEST, Kenneth D. Hypothesis testing with efficient method of moments estimation. **International Economic Review**, JSTOR, p. 777–787, 1987.

POLITIS, Dimitris N.; ROMANO, Joseph P. The stationary bootstrap. **Journal of the American Statistical association**, v. 89.428, p. 1303–1313, 1994.

POLITIS, Dimitris N.; WHITE, Halbert. Automatic block-length selection for the dependent bootstrap. **Econometric Reviews**, v. 23.1, p. 53–70, 2004.

RAD, Hossein; LOW, Rand Kwong Yew; FAFF, Robert W. The profitability of pairs trading strategies: Distance, cointegration, and copula methods. **Cointegration, and Copula Methods (June 3, 2015)**, 2015.

SHEPHARD, Neil; SHEPPARD, Kevin. Realising the future: forecasting with high-frequency-based volatility (HEAVY) models. **Journal of Applied Econometrics**, v. 25.2, p. 197–231, 2010.

TAUCHEN, George. Notes on financial econometrics. **Journal of Econometrics**, v. 100.1, p. 57–64, 2001.

VIDYAMURTHY, Ganapathy. **Pairs Trading: quantitative methods and analysis**. [S.l.]: John Wiley & Sons, 2004. v. 217.

XIE, Wenjun et al. Pairs Trading with Copulas. **The Journal of Trading**, v. 11, n. 3, p. 41–52, 2016.

## Appendix A - Robustness Checks

Tables contain the results of robustness checks over the four subperiods.

**Table A.1:** Bootstrap p-values computed from B=10,000 replications for testing the null hypotheses of equality of the average excess returns and Sharpe Ratios over the period between January 2001 and June 2002.

| Scenario | Committed Capital | | Fully Invested Capital | |
|---|---|---|---|---|
| | Return | Sharpe Ratio | Return | Sharpe Ratio |
| Section 1: Copula-GARCH versus $2.0\sigma$ | | | | |
| d0c0 | 0.0022***(>) | 0.0000***(>) | 0.0046***(>) | 0.0028***(>) |
| d0c5 | 0.0084***(>) | 0.0006***(>) | 0.0158**(>) | 0.0084***(>) |
| d0c10 | 0.0306**(>) | 0.0028***(>) | 0.0544*(>) | 0.0318**(>) |
| d0c15 | 0.1068 | 0.0152**(>) | 0.1438 | 0.0898*(>) |
| d1c0 | 0.2906 | 0.1600 | 0.7338 | 0.8582 |
| d1c5 | 0.5048 | 0.3252 | 0.9454 | 0.9402 |
| d1c10 | 0.7614 | 0.5726 | 0.8950 | 0.7920 |
| d1c15 | 0.9320 | 0.9036 | 0.6904 | 0.6168 |
| Section 2: Copula-GARCH versus $0.75\sigma$ | | | | |
| d0c0 | 0.0714*(>) | 0.0000***(>) | 0.0006***(>) | 0.0036***(>) |
| d0c5 | 0.1622 | 0.0022***(>) | 0.0040***(>) | 0.0114**(>) |
| d0c10 | 0.2768 | 0.0064***(>) | 0.0164**(>) | 0.0408**(>) |
| d0c15 | 0.4706 | 0.0226**(>) | 0.0562*(>) | 0.1128 |
| d1c0 | 0.4744 | 0.1304 | 0.4578 | 0.7448 |
| d1c5 | 0.7210 | 0.3012 | 0.6228 | 0.9024 |
| d1c10 | 0.9850 | 0.5594 | 0.8278 | 0.9150 |
| d1c15 | 0.6986 | 0.9022 | 0.9832 | 0.7670 |
| Section 3: Copula-GARCH versus Copula-HEAVY | | | | |
| d0c0 | 0.0068***(>) | 0.0136**(>) | 0.1020 | 0.1156 |
| d0c5 | 0.0040***(>) | 0.0086***(>) | 0.1006 | 0.1056 |
| d0c10 | 0.0040***(>) | 0.0066***(>) | 0.0862*(>) | 0.0868*(>) |
| d0c15 | 0.0044***(>) | 0.0052***(>) | 0.0954*(>) | 0.0914*(>) |
| d1c0 | 0.0004***(>) | 0.0004***(>) | 0.1846 | 0.1508 |
| d1c5 | 0.0004***(>) | 0.0004***(>) | 0.2116 | 0.1882 |
| d1c10 | 0.0008***(>) | 0.0010***(>) | 0.2226 | 0.2004 |
| d1c15 | 0.0004***(>) | 0.0004***(>) | 0.2292 | 0.2196 |

*Note:* This table reports the bootstrap p-values for testing the null hypothesis of equality of the average excess returns and the Sharpe Ratios of Copula-GARCH and distance and Copula-HEAVY strategies for Top 20 pairs over the period between January 2001 and December 2002 (520 observations). The column labeled Scenario contains symbol labels for trading with no delay or one day waiting period (d0 and d1, respectively) and before or after costs (c0, c5, c10 and c15, respectively). The symbol labels (>) and (<) indicate that the respective null hypothesis is rejected in favor of the alternative and that the average excess returns the Sharpe Ratio of the Copula-GARCH strategy are found greater or less than the the average excess returns and/or the Sharpe Ratio of the distance and Copula-HEAVY strategies, respectively.

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

Author's own elaboration (2016).

**Table A.2:** Bootstrap p-values computed from B=10000 replications for testing the null hypotheses of equality of the average excess returns and Sharpe Ratios over the period between January 2003 and June 2007.

| Scenario | Committed Capital | | Fully Invested Capital | |
|---|---|---|---|---|
| | Return | Sharpe Ratio | Return | Sharpe Ratio |
| Section 1: Copula-GARCH versus $2.0\sigma$ | | | | |
| d0c0 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c5 | 0.0010***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c10 | 0.0176**(>) | 0.0008***(>) | 0.0022***(>) | 0.0022***(>) |
| d0c15 | 0.2162 | 0.0622*(>) | 0.1254 | 0.1256 |
| d1c0 | 0.2756 | 0.1716 | 0.4100 | 0.3882 |
| d1c5 | 0.9218 | 0.8570 | 0.9550 | 0.8878 |
| d1c10 | 0.1788 | 0.0728*(<) | 0.4678 | 0.5444 |
| d1c15 | 0.0134**(<) | 0.0004***(<) | 0.1218 | 0.1774 |
| Section 2: Copula-GARCH versus $0.75\sigma$ | | | | |
| d0c0 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c5 | 0.0016***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c10 | 0.0270**(>) | 0.0010***(>) | 0.0008***(>) | 0.0036***(>) |
| d0c15 | 0.2358 | 0.0518*(>) | 0.1592 | 0.2068 |
| d1c0 | 0.1122 | 0.0664*(>) | 0.4092 | 0.3638 |
| d1c5 | 0.6222 | 0.8090 | 0.9752 | 0.8340 |
| d1c10 | 0.5658 | 0.1946 | 0.4272 | 0.6156 |
| d1c15 | 0.1040 | 0.0032***(<) | 0.1118 | 0.2272 |
| Section 3: Copula-GARCH versus Copula-HEAVY | | | | |
| d0c0 | 0.0004***(>) | 0.0170**(>) | 0.0116**(>) | 0.1266 |
| d0c5 | 0.0000***(>) | 0.0030***(>) | 0.0064***(>) | 0.0360**(>) |
| d0c10 | 0.0000***(>) | 0.0004***(>) | 0.0034***(>) | 0.0072***(>) |
| d0c15 | 0.0000***(>) | 0.0000***(>) | 0.0016***(>) | 0.0014***(>) |
| d1c0 | 0.2620 | 0.3456 | 0.8122 | 0.7996 |
| d1c5 | 0.1990 | 0.1312 | 0.8218 | 0.8568 |
| d1c10 | 0.1490 | 0.0416**(>) | 0.8520 | 0.9324 |
| d1c15 | 0.1192 | 0.0140**(>) | 0.8676 | 0.9830 |

*Note:* This table reports the bootstrap p-values for testing the null hypothesis of equality of the average excess returns and the Sharpe Ratios of of Copula-GARCH and distance and Copula-HEAVY strategies for Top 20 pairs over the period between January 2003 and June 2007 (1177 observations). The column labeled Scenario contains symbol labels for trading with no delay or one day waiting period (d0 and d1, respectively) and before or after costs (c0, c5, c10 and c15, respectively). The symbol labels (>) and (<) indicate that the respective null hypothesis is rejected in favor of the alternative and that the average excess returns and/or Sharpe Ratio of the Copula-GARCH strategy are found greater or less than the the average excess returns and/or the Sharpe Ratio of the distance and Copula-HEAVY strategies, respectively.

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

Author's own elaboration (2016).

**Table A.3:** Bootstrap p-values computed from B=10000 replications for testing the null hypotheses of equality of the average excess returns and Sharpe Ratios over the period between July 2007 and June 2009.

| Scenario | Committed Capital | | Fully Invested Capital | |
|---|---|---|---|---|
| | Return | Sharpe Ratio | Return | Sharpe Ratio |
| Section 1: Copula-GARCH versus $2.0\sigma$ | | | | |
| d0c0 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c5 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c10 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c15 | 0.0006***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d1c0 | 0.3064 | 0.3038 | 0.2782 | 0.1908 |
| d1c5 | 0.6198 | 0.6104 | 0.4452 | 0.2830 |
| d1c10 | 0.9790 | 0.9930 | 0.6520 | 0.4256 |
| d1c15 | 0.6022 | 0.6186 | 0.8888 | 0.5970 |
| Section 2: Copula-GARCH versus $0.75\sigma$ | | | | |
| d0c0 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c5 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c10 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c15 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0008***(>) |
| d1c0 | 0.2148 | 0.2366 | 0.2678 | 0.1528 |
| d1c5 | 0.4442 | 0.5028 | 0.4372 | 0.2468 |
| d1c10 | 0.8036 | 0.8930 | 0.6752 | 0.3896 |
| d1c15 | 0.8154 | 0.7090 | 0.9046 | 0.5432 |
| Section 3: Copula-GARCH versus Copula-HEAVY | | | | |
| d0c0 | 0.0000***(>) | 0.0292**(>) | 0.0000***(>) | 0.0134**(>) |
| d0c5 | 0.0004***(>) | 0.0098***(>) | 0.0000***(>) | 0.0030***(>) |
| d0c10 | 0.0000***(>) | 0.0038***(>) | 0.0000***(>) | 0.0018***(>) |
| d0c15 | 0.0000***(>) | 0.0010***(>) | 0.0000***(>) | 0.0004***(>) |
| d1c0 | 0.4244 | 0.4092 | 0.9108 | 0.8628 |
| d1c5 | 0.3420 | 0.2516 | 0.9610 | 0.9466 |
| d1c10 | 0.2494 | 0.1436 | 0.9918 | 0.9616 |
| d1c15 | 0.1858 | 0.0784*(>) | 0.9486 | 0.8682 |

*Note:* This table reports the bootstrap p-values for testing the null hypothesis of equality of the average excess returns and the Sharpe Ratios of Copula-GARCH and distance and Copula-HEAVY strategies for Top 20 pairs over the period between July 2007 and June 2009 (522 observations). The column labeled Scenario contains symbol labels for trading with no delay or one day waiting period (d0 and d1, respectively) and before or after costs (c0, c5, c10 and c15, respectively). The symbol labels (>) and (<) indicate that the respective null hypothesis is rejected in favor of the alternative and that the average excess returns and/or the Sharpe Ratio of the Copula-GARCH strategy are found greater or less than the the average excess returns and/or the Sharpe Ratio of the distance and Copula-HEAVY strategies, respectively.

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

Author's own elaboration (2016).

**Table A.4:** Bootstrap p-values computed from B=10000 replications for testing the null hypotheses of equality of the average excess returns and Sharpe Ratios over the period between July 2009 and June 2016.

| Scenario | Committed Capital | | Fully Invested Capital | |
|---|---|---|---|---|
| | Return | Sharpe Ratio | Return | Sharpe Ratio |
| Section 1: Copula-GARCH versus 2.0$\sigma$ | | | | |
| d0c0 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c5 | 0.0060***(>) | 0.0008***(>) | 0.0004***(>) | 0.0008***(>) |
| d0c10 | 0.0916*(>) | 0.0212**(>) | 0.0302**(>) | 0.0614*(>) |
| d0c15 | 0.5020 | 0.2958 | 0.5432 | 0.6136 |
| d1c0 | 0.3362 | 0.3022 | 0.4524 | 0.4410 |
| d1c5 | 0.0342**(<) | 0.0118**(<) | 0.1072 | 0.1356 |
| d1c10 | 0.0018**(<) | 0.0004***(<) | 0.0114**(<) | 0.0228**(<) |
| d1c15 | 0.0000***(<) | 0.0000***(<) | 0.0002***(<) | 0.0032**(<) |
| Section 2: Copula-GARCH versus 0.75$\sigma$ | | | | |
| d0c0 | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) | 0.0000***(>) |
| d0c5 | 0.0152**(>) | 0.0004***(>) | 0.0004***(>) | 0.0024***(>) |
| d0c10 | 0.1512 | 0.0270**(>) | 0.0250**(>) | 0.0756*(>) |
| d0c15 | 0.5942 | 0.3150 | 0.5288 | 0.6858 |
| d1c0 | 0.5398 | 0.4508 | 0.6374 | 0.6910 |
| d1c5 | 0.1042 | 0.0272**(<) | 0.1758 | 0.2832 |
| d1c10 | 0.0084***(<) | 0.0002***(<) | 0.0250**(<) | 0.0822*(<) |
| d1c15 | 0.0000***(<) | 0.0000***(<) | 0.0036***(<) | 0.0188**(<) |
| Section 3: Copula-GARCH versus Copula-HEAVY | | | | |
| d0c0 | 0.0036***(>) | 0.0636*(>) | 0.0052***(>) | 0.0994*(>) |
| d0c5 | 0.0026***(>) | 0.0144**(>) | 0.0032***(>) | 0.0324**(>) |
| d0c10 | 0.0010***(>) | 0.0024***(>) | 0.0048***(>) | 0.0116**(>) |
| d0c15 | 0.0004***(>) | 0.0004***(>) | 0.0050***(>) | 0.0038***(>) |
| d1c0 | 0.9486 | 0.9242 | 0.1444 | 0.1492 |
| d1c5 | 0.8242 | 0.4358 | 0.1508 | 0.161 |
| d1c10 | 0.6372 | 0.1756 | 0.1792 | 0.1992 |
| d1c15 | 0.476 | 0.0624*(>) | 0.2104 | 0.2396 |

*Note:* This table reports the bootstrap p-values for testing the null hypothesis of equality of the average excess returns and the of Copula-GARCH and distance and Copula-HEAVY strategies for Top 20 pairs over the period between July 2009 and June 2016 (1831 observations). The column labeled Scenario contains symbol labels for trading with no delay or one day waiting period (d0 and d1, respectively) and before or after costs (c0, c5, c10 and c15, respectively). The symbol labels (>) and (<) indicate that the respective null hypothesis is rejected in favor of the alternative and that the average excess returns and/or the Sharpe Ratio of the Copula-GARCH strategy are found greater or less than the the average excess returns and/or the Sharpe Ratio of the distance and Copula-HEAVY strategies, respectively.

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

Author's own elaboration (2016).

# Chapter 5

# Concluding Remarks

In this thesis we discuss copula-based approaches to describe statistical dependencies within pairs of instruments and in a robust optimization framework. We empirically evaluate their performance using comprehensive and long-term data sets and compare them to competitive models.

Chapter 2 and 4 cover the key concepts and definitions of a market neutral strategy, known as pairs trading, and an overview of two alternative estimation procedures, involving the distance and copula-based approaches. In Chapter 3 we apply a copula-based model within a robust convex portfolio optimization. Finally, in Chapter 4 we use realized measures computed employing higher frequency data.

Of particular interest is a family of copulas, known as Archimedean copulas, that has good mathematical properties that are useful in inference procedures and that cover a wide range of dependence structures. By using a convex linear combination of three copulas from this family with different tail dependence characteristics we produce a more flexible copula capable of capturing better the stylized facts and thus giving a more complete description of the joint distribution of multivariate log returns. The mixture copula also allows us to introduce copulas in a robust optimization framework in Chapter 3.

Our findings in Chapter 2 suggests that the copula strategy has a superior performance than the distance approach when the number of trades is comparable. However, the results also suggest, since neither strategy is consistently superior in all subperiods, at least on committed capital, that the strategies may be combined to yield a viable long-short strategy.

In Chapter 3 we combine robust portfolio optimization and copula-based models in a Worst Case CVaR framework. To cope with the large number of financial instruments we employ the distance method of Gatev, Goetzmann and Rouwenhorst (2006) to select a set of diversified assets that can be useful during crises and tranquil periods and compare the proposed model against three benchmarks: a Gaussian Copula-CVaR (GCCVaR) portfolio, an equally weighted portfolio (1/N) and the S&P 500 index . Our empirical analysis shows that the Copula-CVaR approach generates portfolios with better downside risk statistics for any rebalancing period and it is more profitable than the Gaussian Copula-CVaR for daily and weekly rebalancing.

Our findings in Chapter 4 reveal that copula methods are more profitable than the distance strategy when fewer restrictions are attached to trade. We also found that the copula based-approaches are more sensitive to timing and transactions costs than the distance model using a measure of relative mispricing for a single day. An overall degree of mispricing should be investigated to make the comparisons more meaningful and reliable.

Copula-based approaches empirically proved to be beneficial and quite useful for those interested in managing their risks in Chapters 2-4.