Pricing multiasset equity options with copulas: an empirical test

Francesco Saita
(corresponding/presenting author)
SDA Bocconi and Newfin Research Center
Università Bocconi
E-mail: francesco.saita@unibocconi.it
Tel. +39 02 5836 6102 – 5908
Fax + 39 02 5836 5921 – 5920
Address: Viale Isonzo 25 – 20135 Milano

Maria Egle Romano
Capital Allocation and Risk Management-Market Risk Control Unit
Unicredito Italiano
E-mail: mariaegle.romano@unicredit.it

Francesca Campolongo
Joint Research Center of the European Commission
E-mail: francesca.campolongo@jrc.it

EFM classification codes: 410, 450, 530

Pricing multiasset equity options with copulas: an empirical test

Abstract

The market for equity exotic options on a basket of underlying assets (whether single stocks or stock indexes) has grown significantly through time thanks to the development of structured equity-linked bonds. When evaluating those options, the way in which dependence among underlying assets' returns is critical, and so is the definition of input parameters. Nevertheless, there is little empirical literature that tries to quantify the dispersion of theoretical fair price a trader may face when pricing an option the first time or when revaluing the option in his or her portfolio. This paper tries to analyze the effects of uncertain correlation inputs and of the choice between traditional standard methods assuming joint normality of asset returns and copula-based methods, through a Monte Carlo simulation applied to many different exotic contracts on a basket of five US stocks. Implications for traders and risk managers and auditors, and the potential for further use of copulas in exotic options pricing are then discussed.

1. Introduction

The market for structured equity-linked bonds, that started with a guaranteed equity linked note in the United States in 1987, has since then developed both in terms of volumes and of sophistication. Derivatives desks inside major investment banks, in a quest for financial innovation and for the higher margins that innovators may sometimes attain, have invented each year new exotic options that could produce various types of payoffs for the final investors. Internal market reports from investment banks' research departments now consider a huge number of different alternative exotic equity-linked products: some are able to gain worldwide success, while others may become common and successful in certain countries and almost unknown in others.

In the great variety of the exotic options that an observer may find in the market one feature that is very common is the presence of more than one underlying asset. The kinds of multiasset exotic structures may vary from simple basket options whose payoff is linked to the overall performance of a basket of stock indices or single stocks, to cases such as the so-called conditional coupon structure where the investor receives a fixed coupon each year provided that none of a basket of stocks trespassed a certain barrier (e.g. none of the stocks went below 70% of the initial price). Whether or not closed-end pricing formulas are available, a potentially crucial issue in pricing these options is correlation among the different underlying assets. For instance, it is intuitive that the value of a basket call option would increase if correlation among underlying stocks or indices increase, since this would increase the volatility of the basket: therefore, the basket call would react as a simple call option whose value grows if implied volatility increases. The value of a conditional coupon structure would grow too for a very different reason. In fact, if underlying assets have low positive (or, in theory, even negative) correlation it would be more likely that at least one could touch the barrier and make the coupons disappear; if instead their correlation is higher it is more likely that all of them may grow together remaining distant from the barrier. The critical role of correlation in pricing these options is why they are often labeled as "correlation products", and raise two main problems.

The first problem is that the choice of correlation inputs becomes important for many different players inside the bank. Traders, risk managers and internal auditors are for different reasons interested in using the right correlation estimates, so to guarantee that the option is priced correctly. The fact that risk managers should control traders' work, and internal auditors should control risk managers is relevant since it means that traders' choice of correlation inputs must not only be right, but also must be clearly explainable to other parties inside the bank. Risk managers and auditors know in fact that a wrong set of correlation inputs would alter the value of the portfolio: at least potentially, the trader might hide losses by modifying correlation inputs so to increase the value of its position while using perfectly fair and certified implied volatility inputs.

Of course, if it were possible to extract implied correlations from traded multiasset options as implied volatilities are extracted from traded plain vanilla options, then risks would be much smaller. Unfortunately, while implied volatility can at least up to a certain extent be extracted from traded options' prices[1], it is almost impossible in practice to extract implied correlations. In fact, while when extracting implied volatility from a plain vanilla option's price the trader has one equation to be satisfied (the pricing formula) and one unknown term, in the case of a basket option with five underlying stocks the trader has only one equation (the pricing equation of the basket option), five unknown implied volatilities and ten unknown correlation coefficients. Moreover, the pricing algorithm may not be a closed-end formula, and the trader may also be uncertain about whether all market participants are using the same pricing technique. Despite the fact that the trader cannot infer them from the market, correlation values are critical for him since they influence the price of the option. Therefore, wrong correlation inputs would produce (a) a wrong price when the option is issued and offered inside a structured bond to the institutional client of the investment bank, (b) a wrong mark-to-market (or more precisely, mark-to-model) evaluation of the exotic at the end of each day when it has already been issued and (c) a wrong assessment of its risk profile and its Greeks (Delta, Gamma, Vega, Theta and Rho), since their values derive from the pricing formula and are equally sensible to input data. All the problems are relevant for the trader; the risk manager is concerned especially with the second and the third one. For the internal auditor the second is always crucial and the third is important as well if,

[1] *The problem of implied volatility remains an issue for long term exotic options where it is impossible to extract data from prices of traded options, whose maturity is typically much shorter. In this case, however, some information may be obtained from the OTC market. Some information providers have also tried through time to produce "average" implied volatilities by receiving data from individual investment banks and giving back an aggregate average*

as it should be, he is actively controlling the efficiency of the risk management systems in place inside the bank.

The second problem is that if the dependence among underlying assets' returns plays a key role, one should question whether the classic assumption of multivariate normal distribution is suitable to price this products. The recent stream of contributions concerning copulas and their application to risk management can clearly be applied also to the pricing of multiasset exotic options. Yet, despite the growing literature on copulas on one hand and the relevance of the equity-linked exotic options market on the other hand, there have been very few contributions aimed at testing empirically the role of correlation and dependence when pricing these products. In our view, it is important both to test whether and how much using copulas may significantly change the fair price estimates for exotic products, and to discuss the implementation problems that their use may raise. It is in fact surprising the fact that while they are widely recognized as a theoretically superior means to model dependence among returns, they do not appear to be really used in practice to price multiasset equity derivatives. One possible explanation could be that they might have only a modest impact on fair prices estimates (so that a simpler even if approximate method could be preferred). If this were not the case, other reasons should be found in order to explain why despite the growing interest also in practice as far as both risk measurement issues and credit derivatives evaluation are concerned, the diffusion of copulas in multiasset equity derivatives pricing decisions is still very limited. The aims of the paper are therefore the following:

(1) to evaluate the impact that using copulas may have on fair price estimates of different exotic options on a given basket of stocks;

(2) to discuss the problem of modeling dependence and defining proper input parameters in the context of equity exotic pricing;

(3) to analyze which are the risks deriving from uncertain dependence structure among assets and how the different players inside an investment bank may try to handle the problem.

The structure of the paper is the following. Section 2 defines the key elements about copulas and the main kinds of copulas that may be applied in a generic multivariate (and not only bivariate) setting. Section 3 describes the empirical test analyzing its aims, the

volatility. This solution may be used for instance by the risk manager if he or she wants to control whether the choice of options' payoff that have been tested, the underlying asset and the data and procedures on which parameters' calibration has been based. Section 4 presents the results of the test, while Section 5 discusses its implications for the different players inside the bank. Section 6 concludes.

2. Copula functions

2.1. Definition of copula functions

An n-dimensional copula[2] is a multivariate distribution function (d.f.), C, with uniform distributed margins in [0,1] (U(0,1)) and the following properties:

1. C: $[0,1]^n \rightarrow [0,1]$
2. $\forall u_i \in [0,1] \quad C(u_1,\ldots,u_n)=0$ *if at least one of the* u_i *equals zero*
3. C is n-increasing
4. C has margins C_i which satisfy $C_i(u)=C(1,\ldots,1,u,1,\ldots1)=u$ for all $u \in$ [0,1].

It is clear from the definition above that if $F_1,\ldots,F_n$ are univariate distribution functions, $C(F_1(x_1),\ldots,F_n(x_n))$ is a multivariate d.f. with margins $F_1,\ldots,F_n$ because $u_i = F_i(x_i)$ is a uniform random variable, so the copulas are a useful tool to construct and simulate multivariate distributions.

The following theorem is known as Sklar's Theorem and it is the most important one about copulas because many practical applications are based on it.

Let F be an n-dimensional d.f. with continuous margins $F_1,\ldots,F_n$*, then it has the follow unique copula representation:*

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

The following corollary can be obtained from the expression above.

Let F be an n-dimensional d.f. with continuous margins $F_1,\ldots,F_n$*, and copula , then, for any* $(u_1,\ldots,u_n)$ *in* $[0,1]^n$:

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

---

implied volatility estimate used by the trader is correct.

[2] *See Nelsen (1998)*

where F_i^{-1} is the generalized inverse of F_i.

Therefore the use of copula function allows to overcome the issue of multivariate d.f. estimate, dividing it into two steps:

- determine the margins $F_1, \ldots, F_n$ which represent the distribution of each marginal distribution (in our case, of each risk factor) and estimate their parameters;
- determine the copula function which completely describes the dependence structure of random variables.

2.2. Elliptical copulas: the Gaussian copula

Elliptical distributions class provides a great range of multivariate distribution functions that share many of the tractable properties of the multivariate normal distribution and allow to design different dependence structures. Elliptical copulas are the copulas of elliptical distributions. Simulation from elliptical distributions is easy to perform, therefore, as a consequence of the Sklar's theorem, the simulation of elliptical copulas is also easy.

The most frequently used elliptical copulas are the Gaussian copula and the t-Student copula.

The Gaussian or normal copula is simply the copula derived from the multivariate normal distribution. Let f the standard univariate Gaussian d.f. and $f_{r,n}$ the standard multivariate normal d.f. with linear correlation matrix r, then the n-dimensional copula with correlation matrix r is the following:

$$C_r(u_1, \ldots, u_n) = f_{r,n}\left(f^{-1}(u_1), \ldots, f^{-1}(u_n)\right)$$

The Gaussian copula does not have upper tail dependence and, since elliptical copulas are symmetric, does not even have lower tail dependence.

The Gaussian copula is completely determined by the knowledge of the correlation matrix r and the parameters involved are simple to estimate.

To simulate random variables from Gaussian copula it is enough to simulate a vector from the standard multivariate normal distribution with correlation matrix Σ and then to trasform this vector through a univariate d.f so that you can obtain a vector from the chosen copula.

The matrix Σ, positive definite, can be easily determined with the Cholesky decomposition in order to calculate a matrix A such as $AA^{T} = \Sigma$. Let $Z_1, \dots, Z_n$ be indipendent standard normal variable and the vector $m \in \Re^n$, then the vector $m + AZ$ is multinormally distributed with mean vector m and matrix Σ.

It is then possible to generate random variates from the n-dimensional Gaussian copula running the following algorithm:

- calculate the Cholesky decomposition A of the matrix Σ;
- simulate n indipendent standard normal random variates $z_1, \dots, z_n$;
- set $x = Az$;
- determine the components $u_i = f(x_i), \qquad i = 1, \dots, n$;
- the vector $(u_1, \dots, u_n)^{T}$ is a random variate from the n-dimensional Gaussian copula

2.3. Elliptical copulas: the t-Student Copula

The copula of the multivariate t-Student distribution is the t-Student copula. Defining by $T_{r,n}$ a multivariate t-Student distribution with n degrees of freedom and correlation matrix r, the corresponding copula is the following:

$$C_{r,n}(u_1, \dots, u_n) = T_{r,n}\left(t_n^{-1}(u_1), \dots, t_n^{-1}(u_n)\right)$$

where t_n is the univariate t-Student distribution with n degrees of freedom.

Because the t-Student distribution tends to the normal distribution when n goes to infinity, so the t-Student copula tends to the normal copula when $n \to +\infty$.

In contrast to the Gaussian copula, the t-Student copula has upper tail dependence increasing in r and decreasing in n. Therefore, the t-Student copula is more suitable to simulate events like stock market crashes or the joint default. Besides, for quite large values for n, the tail dependence is significantly different from 0 only when the correlation coefficient is close to 1. This suggests that, for moderate values of the correlation coefficient, a Student copula with a large number of degrees of freedom may be difficult to separate from the Gaussian copula.

The description of a Student copula is defined by two parameters: the correlation matrix r and the number of degrees of freedom n. The estimation of the parameter n is rather difficult and has an important role in the estimation of the correlation matrix, as we will see in the following section. Therefore the t-Student copula is more difficult to calibrate and use than the Gaussian copula.

Random variates from the n-dimensional t-Student copula can be generated through the following steps:

- calculate the Cholesky decomposition A of the matrix Σ;
- simulate n indipendent standard normal random variates $z_1, \dots, z_n$;
- simulate a random variate, s, from c_n^2 distribution, indipendent of z;
- set $y = Az$
- set $x = \frac{\sqrt{n}}{\sqrt{s}} y$;
- determine the components $u_i = t_n(x_i), \quad i = 1, \dots, n$;
- the vector $(u_1, \dots, u_n)^{\mathrm{T}}$ is a random variate from the n-dimensional t-Student copula with n degrees of freedom

2.4. Archimedean copulas

Elliptical copulas are not the only possible type of copulas. Another important family is represented by Archimedean copulas, that include for instance the Gumbel and Clayton copula. This class of copulas, in contrast to elliptical copulas, have closed form expressions, because these copulas are not derived from a multivariate distribution function using the Sklar's theorem. As a consequence the Archimedean copulas are originally defined on two dimensions and their multivariate extension need some technical conditions to assert the n-copulas are proper. Therefore, even if they allow to model the dependence structure between variables in different and even more flexible ways than elliptical copulas, their application is currently confined in practice to bivariate problems (an example being the valuation of a credit derivative whose price depends also on the risk of joint default of the underlying bond and of the protection seller, that can be modelled through copulas). Unfortunately, multiasset equity options typically imply much more than two underlying assets, and therefore our test will be restricted to elliptical copulas.

3. The design of the empirical test

3.1. General aims of the test

The purpose of the empirical test is to check the differences among fair prices for a set of multiasset exotic options that can be obtained according to copulas as opposed to the standard assumption of joint normality of asset returns. Therefore, we will check how close or distant copula-based prices will be relative to the prices obtained through the simpler method, and how close or distant they are among themselves, if different copulas are used. In particular, we have tested t-Student copulas with 4, 12 and 20 degrees of freedom. The choice to avoid Archimedean copulas derives from the fact that fitting those copulas on joint distributions of more than two underlying assets is extremely complex, and at least at present it is more than unlikely that they might be applied in practice to price multiasset options where the number of underlying assets may range from three to even twenty or more assets. At the same time, as far as the traditional pricing method is concerned, we will check the impact of different estimates of the linear correlation matrix, depending on the size of the sample and on the frequency of return data that are used. As a whole, the test will give a picture of how stable or unstable fair prices may be depending on the assumptions the trader (or risk manager) is making, and of the degree of uncertainty that the different players inside the bank who are concerned with dependence and correlation on equity exotic products may typically face.

3.2. Underlying assets, data sample, and the set of exotic options

The underlying assets we have considered for our test are five US stocks, and precisely Microsoft, General Electric, Coca Cola, IBM and JPMorgan Chase. In order to estimate parameters we used a five years historical time series of daily closing prices and returns from October 1st, 1999 to September 30th, 2004. All option valuations were conducted with market data on October 1st 2004; zero coupon risk-free rates were derived from the US swap curve on that date, while dividend yields were estimated based on historical average dividend yields for the stocks in the sample.

As far as the sample of exotic options is concerned, we tried to build a sample of different payoff structures, a large part of which is actually commonly used in practice. All options had a remaining maturity of 5 years (i.e., all options were assumed to expire on October 1st, 2009). Yet, for most options we distinguished between a brand new option evaluated at the date of issue, and an already existing option with 8 years of initial maturity, issued on October 1st 2001 and evaluated after 3 years. Those options will be labelled as having a 3+5 maturity. The reason why we differentiated between new and existing options is that we wanted to test whether they showed a different sensitivity to the correlation among underlying assets. For instance, an option whose payoff is based on the worst performing asset within a basket could be expected to have a different sensitivity to changes in correlation among assets when it is issued (and the worst performing asset is still unknown) or after three years (when there will be a *pro tempore* ranking of assets' performance so that the option's payoff is likely to be linked especially to the correlation of the two or three worst performing assets, while being relatively less sensitive to correlation between the best performing ones). The underlying research question is therefore whether – at least for some kinds of payoffs – sensitivity to correlation or to dependence structure among asset returns is only a temporary effect or whether instead it is a permanent one.

The specific types of options we considered is detailed in Table 1. In some cases the name of the option is consistent with a very well-defined industry standard, while in others a single widely accepted name for the particular payoff did not exist. In any case, Table 1 should provide a clear enough description of the payoff to avoid misunderstandings.

Table 1. A description of the exotic options analyzed in the empirical test

Option	Option description
Asian basket option (5 yrs, strikes equal to initial prices)	New Asian basket options with 5 years of maturity. The option's payoff is the maximum between the mean percentage return of each stock (if the mean is positive) and zero (basket feature). The performance of each stock is calculated as the difference between the average price of the stock at the end of each month (Asian feature) and the stock's initial price.
Asian basket option (3+5 yrs, strikes equal to initial prices)	Already issued Asian basket options; the option has been issued 3 years before valuation date and has 5 years of remaining maturity. The option's payoff is calculated as in the previous case, with the only difference that relevant initial prices are not prices on October 1st, 2004 but prices on October 1st, 2001, since the option has been issued on that day.
Asian basket option (3+5 yrs, strikes equal to average prices during the first 3 years)	Already issued Asian basket option. The option is identical to the preceding one, apart from the fact that the strike price is set equal for each stock to the average price at the end of the month during the first free years. Therefore, the option is at the money at the beginning of the simulation as the first 5-year Asian option.

Table 1 (continued). A description of the exotic options analyzed in the empirical test

Option	Option description
Asian best option (5 yrs, 40% participation rate)	New option with 5 years of maturity where the final payoff is 40% of the return of the highest performing stock (if positive). The performance of each stock is calculated as the difference between the average price of the stock at the end of each month (Asian feature) and the stock's initial price.
Asian best option (3+5 yrs, 40% participation rate)	Already issued Asian best option identical to the preceding one but with 3 years of past history and 5 years of remaining maturity
Napoleon option (based on monthly basket returns, annual coupon 12%)	5-year option where the investor receives each year a coupon equal to 12% minus the worst monthly performance of the stock basket during the year. The coupon is floored at zero if the basket has a minimum monthly performance lower than - 12%.
Conditional coupon structure (5 yrs, 8% annual coupon, barrier=60%)	New 5-year option that pays each year a fixed 8% coupon if none of the underlying assets at the end of each month has ever touched a barrier equal to 60% of the initial price on October 1st, 2004.
Conditional coupon structure (3+5 yrs, 8% annual coupon, barrier=70%)	Already issued option with 5 years of remaining maturity that pays each year a fixed 8% coupon if none of the underlying assets at the end of each month has ever touched a barrier equal to 70% of the initial price on October 1st, 2001.
Fixed 80% coupon minus worst performance (5 yrs)	New 5-year option that pays at maturity a coupon equal to 80% plus the negative performance of the worst performing stock within the basket. The coupon is capped at 80% (if all stocks had positive returns) and floored at 0 (if one stock decreased by more than 80%)[3]
Fixed 80% coupon minus worst performance (3+5 yrs)	Already issued option with 3 years of past history and 5 years of remaining maturity that pays at maturity a coupon equal to 80% plus the negative performance of the worst performing stock within the basket. The coupon is capped at 80% (if all stocks had positive returns) and floored at 0 (if one stock decreased by more than 80%)
Fixed 25% coupon minus put on basket performance (5 yrs)	New 5-year option that pays at maturity a coupon equal to 25% plus the performance (only if it is negative) of the basket of stocks. The coupon is floored at 0 (if the basket value decreased by more than 25%)[4]
Fixed 25% coupon minus put on basket performance (3+5 yrs)	Already issued option with 3 years of past history and 5 years of remaining maturity that pays at maturity a coupon equal to 25% plus the performance (only if it is negative) of the basket of stocks. The coupon is floored at 0 (if the basket value decreased by more than 25%)

3.3. Pricing methodology

All the options described in the previous paragraph have been priced through a Monte Carlo simulation, calculating therefore the fair value of the option as the discounted

[3] *The zero floor can be interpreted as the result of a spread position between two different worst put options: the investor would be selling a put on the worst of all assets and then buying another put on the worst of all assets whose strike price is fixed 80% below current prices.*

[4] *Again, the position can be conceived as the result between a short ATM basket put and a long OTM basket put whose strike price is fixed at a 25% decrease from current basket prices.*

value of the expected payoff in a risk-neutral world. In each of the 10,000 simulation runs we have simulated daily correlated returns for each of the stocks and reproduced the daily price of each asset so to precisely calculate the payoff of each contract. In each simulation run we have compared seven different pricing alternatives: four of them were based on a simple multivariate Monte Carlo simulation under a joint normality assumption, and three were based on t-Student copulas with different degrees of freedom. The risk-neutral expected return for each asset was obviously the same under each pricing method. In all the cases, since the simulation requires to extract first a vector of uncorrelated standard normal random variables and then to transform them into correlated returns, we have used the same uncorrelated vectors so to guarantee that differences in prices may not derive from uncorrelated random variables sampling errors. In practice, for each extraction of a daily random uncorrelated vector we have transformed the same vector into different correlated vectors according to each of the seven method tested, and reproduced jointly seven alternative return and price paths according to each methodology. Uncorrelated random variables were extracted using a Latin Hypercube algorithm, that enabled us to reproduce in the best possible way the whole joint multivariate distribution in our simulation.

The four alternative payoffs under the standard assumption of multivariate normality (MVN) were obtained by using four different input historical correlation matrixes: two 5-year linear correlation matrixes based respectively on monthly and weekly stock returns, and two equivalent linear correlation matrixes based on 3 years of data only. The purpose was to check the effects of the uncertainty that even if the simpler model is adopted a trader or a risk manager may face in feeding the simple model with the "right" inputs. Volatility was set equal to historical volatility in all four cases. In fact, even if one could argue that consistent volatilities should have been used (e.g. weekly 3-year sample volatilities should be combined with weekly 3-year sample correlation coefficients), our assumption about the practical trader's behaviour is that the trader would know an implied volatility value for all the underlying assets, and would then have to decide which correlation inputs he should use. Since we wanted to investigate the correlation problem only, we decided to test it by maintaining the same level of implied volatility throughout the four correlation scenarios. We simply used 5-year, monthly data historical volatility as a proxy for implied volatility, but this should not alter results in any way. Annualized historical 5-year volatilities based on monthly returns for Microsoft, General Electric, Coca Cola, IBM and JPMorgan Chase were equal respectively to 44.15%, 27.25%, 25.83%, 35.67% and 39.65%.

For any of the four MVN cases, historical correlation matrixes have been reproduced by simulating random uncorrelated standard normal variables and then transforming them into correlated random variables through the classic Cholesky decomposition. Therefore we had four different Cholesky matrixes (one for each historical correlation matrix) that were applied at the same time to produce different multivariate return paths. Correlation matrixes in the four different cases are reported in Tables 2 through 5.

Table 2. Correlation matrix based on monthly returns (5-year sample)

	Microsoft	GE	Coca Cola	IBM	JPM
Microsoft	1,000	0,446	-0,039	0,587	0,392
GE	0,446	1,000	-0,068	0,455	0,342
Coca Cola	-0,039	-0,068	1,000	-0,124	0,154
IBM	0,587	0,455	-0,124	1,000	0,509
JPM	0,392	0,342	0,154	0,509	1,000

Table 3. Correlation matrix based on monthly returns (3-year sample)

	Microsoft	GE	Coca Cola	IBM	JPM
Microsoft	1,000	0,458	0,062	0,745	0,532
GE	0,458	1,000	-0,078	0,605	0,327
Coca Cola	0,062	-0,078	1,000	0,049	0,402
IBM	0,745	0,605	0,049	1,000	0,586
JPM	0,532	0,327	0,402	0,586	1,000

Table 4. Correlation matrix based on weekly returns (5-year sample)

	Microsoft	GE	Coca Cola	IBM	JPM
Microsoft	1,000	0,321	0,014	0,393	0,382
GE	0,321	1,000	0,119	0,418	0,537
Coca Cola	0,014	0,119	1,000	0,010	0,103
IBM	0,393	0,418	0,010	1,000	0,398
JPM	0,382	0,537	0,103	0,398	1,000

Table 5. Correlation matrix based on weekly returns (3-year sample)

	Microsoft	GE	Coca Cola	IBM	JPM
Microsoft	1,000	0,457	0,112	0,619	0,598
GE	0,457	1,000	0,056	0,503	0,520
Coca Cola	0,112	0,056	1,000	-0,041	0,142
IBM	0,619	0,503	-0,041	1,000	0,542
JPM	0,598	0,520	0,142	0,542	1,000

As far as copulas are concerned, we tested three different t-Student copulas with 4, 12 and 20 degrees of freedom. The choice of the degrees of freedom has been arbitrary, due to the fact that it is very difficult in practice to optimize jointly the number of degrees of freedom and the correlation coefficients' matrix. The t-20 copula is of course closer to the normal case, since as the number of degrees of freedom grow the t-Student copula gets closer and closer to the normal copula. The cases of 4 and 12 degrees of freedom represent instead two markedly different cases[5].

3.4. Estimating parameters for t-Student copulas

To run our test we have chosen to use the t-Student copula with t-Student margins. The t-Student margins have been chosen in order to consider the fat tail and the t-Student copula in order to have a copula with tail dependence (lower and upper) and simply to use in a more than 2 assets dimension.

The first step has therefore been to estimate the marginal distribution parameters. The density function of a t-Student distribution with n degrees of freedom is given by

$$h_n(x) = \frac{\Gamma(n+1/2)}{\Gamma(n/2)\Gamma(1/2)\sqrt{n}}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

and in the general case $x \to \frac{x-m}{s}$ the distribution becomes

$$f_n(x) = \frac{\Gamma(n+1/2)}{\Gamma(n/2)\Gamma(1/2)}\left(\sqrt{n}\,s\right)^n\left(ns^2+(x-m)^2\right)^{-\frac{n+1}{2}}$$

In $f_n(x)$, m is the mean of the distribution and can be simply estimate from the historical series through the estimator $\hat{m}$

$$\hat{m}_i = \frac{1}{T}\sum_{j=1}^{T} S_i(t_j) \qquad i = 1,\ldots,5 .$$

In $f_n(x)$, s is not the standard deviation of t-Student distribution, because the variance of this distribution is infinite. s is an additional parameter that affects the shape of the distribution and can be valued from the data using the maximum likelihood method.

---

5 *In the different context of risk integration among different risks, Rosenberg and Schuermann (2004) have tested for instance the impact of a t-Student copula with 5 or 10 degrees of freedom against the normal copula. Even in that case, the choice of the number of degrees of freedom was subjective.*

For every series of asset returns it is then necessary to:

1. determine the number of degrees of freedom n ;
2. estimate the mean setting $\hat{m} = \frac{1}{T}\sum_{j=1}^{T} S(t_j)$;
3. estimate the parameter s using the likelihood function give by

$$L(s) = \prod_{j=1}^{T} f_n\left(S(t_j)\right)$$

   from which it is possible to obtain the estimator

$$\hat{s} = \{s_{\max} / \ln L(s_{\max}) \geq \ln L(s) \quad \forall s > 0\}.$$

In order to determine n it is possible to follow a procedure which, among various t-Student distributions with different degrees of freedom n, selects the one that best fits the empirical data.

The procedure in this case is represented by the following steps:

1. estimate the mean of the empirical data;
2. assume that the data are described by $f_1(x)$, $f_2(x)$, $f_3(x)$, …, and estimate s for every supposed number of degrees of freedom;
3. standardize the data;
4. divide the data into some intervals and determine the frequency of each class (N_j) through *m* indipendent extractions;
5. determine the frequency of every class for every t-Student distribution (n_j);
6. for every n calculate the quadratic deviation between N_j and n_j through the formula $Q = \sum_{j=1}^{t} \frac{(N_j - n_j)^2}{n_j}$ where *t* is the number of classes;
7. calculate c^2 for every Q: you will choose the t-Student distribution with the smallest c^2.

The limit in this method is that results may partly depend on arbitrary choices, such as for instance the definition of the classes into which data are divided. We decided to test parameters under the assumption of three different degrees of freedom, respectively equal to 4, 12 and 20. The final results that were obtained as far as marginal distribution parameters are concerned are reported in Table 6.

Table 6. Parameters of marginal distribution functions for the t-Student copula

	Microsoft	GE	Coca Cola	IBM	JPM
Mean	-0,00043	-0,00016	-0,00021	-0,00023	-0,00017
Shape ($u = 4$)	0,00033	0,000263	0,000169	0,000543	0,000372
Shape ($u = 12$)	0,00047	0,000362	0,000243	0,000379	0,000523
Shape ($u = 20$)	0,000522	0,000396	0,00027	0,000422	0,000575

After estimating parameters of the marginal distribution functions, the second logical step is represented by estimating copula parameters. The classical estimation method is again the maximum likelihood method. The density of the joint distribution F is given by

$$f_q(x_1,\ldots,x_n) = c(F_1(x_1),\ldots,F_n(x_n))\prod_{i=1}^{n} f_i(x_i)$$

where f_i is the density of the marginal distribution F_i and c is the copula density.

Set $\mathrm{X} = \{(x_1^t,\ldots,x_n^t)\}_{t=1}^{T}$, the likelihood function will be

$$L(q) = \prod_{t=1}^{T} f_q(x_1^t,\ldots,x_n^t)$$

from which the estimator

$$\hat{q} = \{q_{\max} / \ln L(q_{\max}) \geq \ln L(q) \quad \forall q \in \Theta\}.$$

The function that should be maximized is represented by the logarithm of the likelihood function $l(q)$. So we obtain

$$l(q) = \sum_{t=1}^{T} \ln c(F_1(x_1^t),\ldots,F_n(x_n^t)) + \sum_{t=1}^{T}\sum_{i=1}^{n} \ln f_i(x_i^t)$$

where q is the vector including the parameters of the n marginal distributions and the parameters of the copula.

For a t-Student copula the function is

$$l(q) = -\frac{T}{2}\ln|r| - \left(\frac{n+n}{2}\right)\sum_{t=1}^{T}\ln\left(1+\frac{1}{n}V_t^{\mathrm{T}} r^{-1} V_t\right) + \left(\frac{n+1}{2}\right)\sum_{t=1}^{T}\sum_{i=1}^{n}\ln\left(1+\frac{V_n^2}{n}\right)$$

where $V_t = (t_n^{-1}(u_1^t),\ldots,t_n^{-1}(u_n^t))$.

This method can be computational intensive in the case of high dimensional distributions, because it requires to estimate jointly the parameters of the marginal distributions and the parameters of the dependence structure. Since copulas allow to split the parameters in *specific* parameters for the margins and in *common* parameters for the joint structure the log of the maximum likelihood function can be written in the following way:

$$l(q) = \sum_{t=1}^{T} \ln c\left(F_1\left(x_1^t; q_1\right), \ldots, F_n\left(x_n^t; q_n\right); a\right) + \sum_{t=1}^{T} \sum_{i=1}^{n} \ln f_i\left(x_i^t; q_i\right)$$

where $q = (q_1, \ldots, q_n, a)$. q_i and a are respectively the parameters of the margins and the copula.

So we can perform the estimation of the univariate marginal distributions at a first steps and then determine a given the previous estimates through

$$\hat{a} := \arg\max \sum_{t=1}^{T} \ln c\left(F_1\left(x_1^t; \hat{q}_1\right), \ldots, F_n\left(x_n^t; \hat{q}_n\right); a\right)$$

This two-steps method is called the method of *inference functions for margins* or IMF method.

However in the case of a t-Student copula the estimation of the parameters could require numerical optimisation of the likelihood function because it does not exist a closed form expression as in the case of a Gaussian copula.

The best ideal choice would have been to estimate jointly the number of degrees of freedom and the correlation matrix using a simulation. The procedure would have implied to simulate some matrixes through extractions of random numbers between 0 and 1 and a number between 4 and 21 (for n), to calculate the likelihood function for each matrix and n, and to select the combination with the maximum output. Unfortunately the positive constraint on the logarithm and the need to obtain a positive definite matrix would have forced to an excessively high number of simulations in order to peg the estimate. For this reason, and considering that we wanted to replicate a method that could be applied in practice, we decided to resort to a simpler method, even if it does not allow to estimate the degrees of freedom and correlation parameters jointly. More precisely, we decided to apply the IFM method and, after the estimation of the margins parameters, to calculate the correlation matrix using an iterative algorithm, that does not require optimisation:

1. let $\hat{r}_0$ the IFM estimate of the correlation matrix for the Gaussian copula[6]
2. $\hat{r}_{m+1}$ is obtained using the following equation

$$\hat{r}_{m+1} = \frac{1}{T}\left(\frac{n+n}{n}\right)\sum_{t=1}^{T}\frac{V_t^{\mathrm{T}}V_t}{1+\frac{1}{n}V_t^{\mathrm{T}}\hat{r}_m^{-1}V_t}$$

3. repeat the second step until convergence $\hat{r}_{m+1} = \hat{r}_m (:= \hat{r}_\infty)$
4. the IFM estimate of the correlation matrix for the Student copula is $\hat{r}_{IFM} = \hat{r}_\infty$

This procedure does not allow to estimate at the same time the number of degrees of freedom for copula, which are considered as given. So we arbitrary chose three different number of degrees of freedom (4, 12, 20) and we estimated three different correlation matrixes, used to compare the simulated option prices. The same numbers of degrees of freedom have been used for the t-Student marginal distributions in order to simplify the simulation procedure. The final correlation matrix outcomes under the three different degrees of freedom are reported in Tables 7 through 9.

Table 7. Correlation matrix for the t-Student copula with 4 degrees of freedom

	Microsoft	GE	Coca Cola	IBM	JPM
Microsoft	1,000	0,477	0,191	0,533	0,442
GE	0,477	1,000	0,281	0,446	0,555
Coca Cola	0,191	0,281	1,000	0,180	0,225
IBM	0,533	0,446	0,180	1,000	0,419
JPM	0,442	0,555	0,225	0,419	1,000

Table 8. Correlation matrix for the t-Student copula with 12 degrees of freedom

	Microsoft	GE	Coca Cola	IBM	JPM
Microsoft	1,000	0,434	0,157	0,469	0,404
GE	0,434	1,000	0,248	0,412	0,529
Coca Cola	0,157	0,248	1,000	0,147	0,194
IBM	0,469	0,412	0,147	1,000	0,386
JPM	0,404	0,529	0,194	0,386	1,000

[6] $\hat{r}_{IFM} = \frac{1}{T}\sum_{t=1}^{T} V_t^{\mathrm{T}} V_t$ where $V_t = \left(f^{-1}(u_1^t),\ldots,f^{-1}(u_n^t)\right)$ and $u_1^t = F_1(x_1^t),\ldots,u_n^t = F_n(x_n^t)$.

Table 9. Correlation matrix for the t-Student copula with 20 degrees of freedom

	Microsoft	GE	Coca Cola	IBM	JPM
Microsoft	1,000	0,430	0,150	0,459	0,404
GE	0,430	1,000	0,243	0,413	0,537
Coca Cola	0,150	0,243	1,000	0,137	0,189
IBM	0,459	0,413	0,137	1,000	0,387
JPM	0,404	0,537	0,189	0,387	1,000

3.5. Generation of random correlated returns

After defining the payoff structures, the different pricing methodologies or input sets and the proper parameters, we started running the simulation. While the generation of correlated returns in the MVN case was performed through a simple Cholesky decomposition method, a few words may be useful about the generation of random correlated returns with copulas.

Let us consider first the bidimensional case. The objective is to generate random number pairs (u,v) from the variables U and V, being uniformly distributed on $[0,1]$ and having C as joint distribution function. So we have

$$C(u,v) = C(F_u(u), F_u(v)) = \mathrm{P}(U \leq u, V \leq v)$$

where F_u is the uniform d.f on $[0,1]$.

The conditional distribution function of variable V, for a setted value u of U is given by

$$\mathrm{P}(U = u, V \leq v) = \frac{\partial \mathrm{P}(U \leq u, V \leq v)}{\partial u} = \frac{\partial C(u,v)}{\partial u} = C_u(v)$$

where $C_u(v)$ is the partial derivative of the copula as to u. To simplify, we assume that $C_u(v)$ exists for any $v \in [0,1]$ and that it is strictly monotonic.

Using now the method of transformation of variables it is possible to generate the desired pair of random numbers through the following steps:

- generate two random numbers u and w, indipendent and uniformly distributed on $[0,1]$; u is already the first number we are looking for
- put $v = C_u^{-1}(w)$

Then (u,v) is a pair of random numbers that are uniformly distributed on $[0,1]^2$ and have C as a joint distribution function. It is worthwhile to note that in some cases the inverse function of the partial derivative of the copula cannot be obtained. It is therefore necessary to use a numerical algorithm to calculate v.

To simulate a n-dimensional vector $(u_1,\ldots,u_n)$[7] the only difference is that the algorithm above should be run recursively:

- to simulate u_1 from the uniform distribution on $[0,1]$;
- to simulate u_2 from $C_2(u_2/u_1)$ as described before;
- to simulate u_3 from $C_3(u_3/u_1,u_2)$;
- $\vdots$
- to simulate u_n from $C_n(u_n/u_1,\ldots,u_{n-1})$.

4. Empirical results

4.1. A comparison between copula-based and multivariate normal simulations

The empirical test described in section 3 has therefore produced the prices of a set of different exotic options, that have been obtained as the present value of the expected payoff in a risk-neutral world over 10,000 simulations. The overall results can be summarized in Table 10, which reports the fair price for all the options depending on whether the dependence structure between assets' returns had been modelled either by assuming multivariate normally distributed returns or through t-Student copulas. In the case of the multivariate normality assumption (MVN) there is then a distinction based on how the historical linear correlation matrix had been estimated, while copulas differ depending on the number of the degrees of freedom.

[7] *Cfr. Embrechts-McNeil-Straumann(1999).*

Table 10. Fair prices for all option types depending on simulation method.

Simulation method / Option type	Multivariate normal				t-Student copula		
	5 yrs, monthly returns	3 yrs, monthly returns	5 years, weekly returns	3 yrs, weekly returns	4 D.F.	12 D.F.	20 D.F.
Asian basket option (5 yrs, strikes equal to initial prices)	14,907	15,948	14,742	15,705	34,801	18,950	17,303
Asian basket option (3+5 yrs, strikes equal to initial prices)	5,445	6,067	5,314	5,942	16,198	7,509	6,628
Asian basket option (3+5 yrs, strikes equal to average prices during the first 3 years)	10,690	11,309	10,588	11,195	23,754	13,322	12,231
Asian best option (5 yrs, 40% participation rate)	24,351	22,531	24,845	23,139	48,495	28,721	26,631
Asian best option (8 yrs, 40% participation rate)	12,088	11,306	12,252	11,305	26,300	14,700	13,460
Napoleon option (based on monthly basket returns, annual coupon 12%)	10,971	8,701	11,318	9,086	3,774	7,929	8,740
Conditional coupon structure (5 yrs, 8% annual coupon, barrier=60%)	10,866	12,327	10,536	11,996	7,895	10,069	10,419
Conditional coupon structure (3+5 yrs, 8% annual coupon, barrier=70%)	7,404	8,945	7,123	8,341	5,660	7,203	7,384
Fixed 80% coupon minus worst performance (5 yrs)	24,935	27,734	24,386	27,144	21,893	24,323	24,762
Fixed 80% coupon minus worst performance (3+5 yrs)	21,951	24,831	21,387	23,833	19,543	21,654	21,968
Fixed 25% coupon minus put on basket performance (5 yrs)	14,055	13,476	14,197	13,581	14,801	14,243	14,057
Fixed 25% coupon minus put on basket performance (3+5 yrs)	12,307	11,877	12,438	11,918	13,645	12,749	12,478

The first clear result is that the prices produced by the t4-copula are strikingly different from the others, while the remaining methods produce values which apparently remain within a much closer range. Anyway, to clearly appreciate results it is useful to consider the differences from a real-world perspective in which the situations that may actually occur could be either (a) the case of a risk manager (or trader) who has decided to use a MVN assumption, but has to choose which estimate of historical correlation should be used, or (b) the case of a more sophisticated risk manager (or trader) who wants to decide whether to use copulas instead than a MVN assumption and tries to evaluate the potential effects of different copula specifications (in our case, the number of degrees of freedom for the t-Student copula).

The first situation can be analyzed through Table 11.

Table 11. The effect of different estimates of linear correlations under multivariate normality assumption (effect of different sample size and return frequency).

Option type	Fair price (5-yr monthly sample)	Weekly vs monthly returns		3-yr vs 5-yr sample		$(P_{max}-P_{min}) / P_{avg}$
		5-yr sample	3-yr sample	Monthly returns	Weekly returns	
Asian basket option (5 yrs, strikes equal to initial prices)	14,907	-1,11%	-1,53%	6,98%	6,53%	7,87%
Asian basket option (3+5 yrs, strikes equal to initial prices)	5,445	-2,42%	-2,06%	11,42%	11,83%	13,24%
Asian basket option (3+5 yrs, strikes equal to average prices during the first 3 years)	10,690	-0,96%	-1,01%	5,79%	5,73%	6,59%
Asian best option (5 yrs, 40% participation rate)	24,351	2,03%	2,70%	-7,47%	-6,87%	9,75%
Asian best option (8 yrs, 40% participation rate)	12,088	1,36%	0,00%	-6,47%	-7,72%	8,06%
Napoleon option (based on monthly basket returns, annual coupon 12%)	10,971	3,16%	4,42%	-20,69%	-19,72%	26,12%
Conditional coupon structure (5 yrs, 8% annual coupon, barrier=60%)	10,866	-3,04%	-2,69%	13,45%	13,85%	15,67%
Conditional coupon structure (3+5 yrs, 8% annual coupon, barrier=70%)	7,404	-3,80%	-6,75%	20,81%	17,10%	22,90%
Fixed 80% coupon minus worst performance (5 yrs)	24,935	-2,20%	-2,13%	11,23%	11,31%	12,85%
Fixed 80% coupon minus worst performance (3+5 yrs)	21,951	-2,57%	-4,02%	13,12%	11,43%	14,97%
Fixed 25% coupon minus put on basket performance (5 yrs)	14,055	1,01%	0,78%	-4,12%	-4,34%	5,21%
Fixed 25% coupon minus put on basket performance (3+5 yrs)	12,307	1,06%	0,35%	-3,49%	-4,18%	4,62%
Average		*-0,62%*	*-0,99%*	*3,38%*	*2,91%*	*12,32%*
Average of differences in absolute value		*2,06%*	*2,37%*	*10,42%*	*10,05%*	-
Maximum		*3,16%*	*4,42%*	*20,81%*	*17,10%*	*26,12%*
Minimum		*-3,80%*	*-6,75%*	*-20,69%*	*-19,72%*	*4,62%*

First of all, it is clear that the choice of the frequency of return data is not as critical as the choice of sample length, since changes between the 3-year and the 5-year sample are substantial. Anyway, the choice between monthly versus weekly returns to estimate correlation may maintain a significant impact on prices. The fact that average differences may be close to zero should not lead to wrong conclusions, since this is due to the fact that using weekly data there is a mild reduction in correlation. As a consequence, while the value of options which are (intuitively) positively linked with assets' correlation such as Asian baskets or conditional coupon structures diminishes, other options such as Asian best options which can profit from lower assets' correlation register an increase in their value. In practice, therefore, the average difference in absolute value gives a clearer picture of the uncertainty the trader or the risk manager may face in pricing the option.

Second, at least in the specific case and analyzing the impact in terms of percentage change, there is no systematic difference between the impact of correlation inputs changes over new (5 years) as opposed to already issued ("3+5") options. This result should in any case be taken with remarkable caution considering that (i) pre-issued options may be particularly sensitive to one or two coefficients alone, and (ii) correlation coefficients for each pair of assets do not change in the same direction when moving from an historical sample to another.

It must be noted that it might be wise to evaluate differences deriving from data frequency and sample size separately. In fact, while the choice between weekly against monthly data might represent a pure discretionary one, most traders would likely prefer to use a 5-year rather than a 3-year sample if they were to price a long-dated option. Yet, the difference between the prices that can be obtained with different data samples points out the further problems that might be faced when pricing an option where some of the underlying assets have only a short return series (e.g. because they are stocks that were listed through an IPO only a couple of years before the exotic option is issued).

In any case, even if we consider the most favorable case when any sample size is available, the differences between the "fair" prices that can be obtained remains remarkable, especially if we consider that (a) we are assuming no uncertainty about implied volatility inputs (i.e. volatility is the same across all simulations) and dividend yields and (b) we are adopting the same pricing method with the same, easy assumption on the shape of the joint distribution of assets' returns. Both conditions may not hold true in practice, where traders may use different implied volatilities and might be tempted to use alternative models to define the dependence structure among assets' returns.

This introduces us to the second issue, that is the difference among options' prices obtained through different t-Student copulas against the simpler standard multivariate normal method (that will be represented here through the 5-year sample, monthly data case).

Table 12. Fair values obtained through copulas versus standard multivariate normal method.

	Differences versus multivariate normal method (MVN)			$(P_{max}-P_{min}) / P_{avg}$		
	t-Student copula, 4 DF	t-Student copula, 12 D.F.	t-Student copula, 20 D.F.	All copulas	All copulas + MVN case	t-12 and t-20 copulas +MVN case
Asian basket option (5 yrs, strikes equal to initial prices)	133,5%	27,1%	16,1%	73,88%	92,57%	23,71%
Asian basket option (3+5 yrs, strikes equal to initial prices)	197,5%	37,9%	21,7%	94,64%	120,20%	31,61%
Asian basket option (3+5 yrs, strikes equal to average prices during the first 3 years)	122,2%	24,6%	14,4%	70,11%	87,09%	21,78%
Asian best option (5 yrs, 40% participation rate)	99,1%	17,9%	9,4%	63,16%	75,33%	16,45%
Asian best option (8 yrs, 40% participation rate)	117,6%	21,6%	11,4%	70,73%	85,42%	19,47%
Napoleon option (based on monthly basket returns, annual coupon 12%)	-65,6%	-27,7%	-20,3%	72,88%	91,64%	33,02%
Conditional coupon structure (5 yrs, 8% annual coupon, barrier=60%)	-27,3%	-7,3%	-4,1%	26,67%	30,28%	7,63%
Conditional coupon structure (3+5 yrs, 8% annual coupon, barrier=70%)	-23,6%	-2,7%	-0,3%	25,54%	25,23%	2,74%
Fixed 80% coupon minus worst performance (5 yrs)	-12,2%	-2,5%	-0,7%	12,12%	12,69%	2,48%
Fixed 80% coupon minus worst performance (3+5 yrs)	-11,0%	-1,4%	0,1%	11,51%	11,39%	1,43%
Fixed 25% coupon minus put on basket performance (5 yrs)	5,3%	1,3%	0,0%	5,18%	5,22%	1,33%
Fixed 25% coupon minus put on basket performance (3+5 yrs)	10,9%	3,6%	1,4%	9,01%	10,46%	3,54%

Results from Table 12 can be analyzed from two perspectives. On one hand, they should be considered as a whole to evaluate whether and how much the choice of copulas to model dependence can change estimated fair values. On the other hand, it is possible to study and discuss the impact on the price of each kind of option.

The first of the two perspectives is reasonably the most important, and clearly shows that using t-Student copulas may significantly change fair prices' estimates. We can note first that, not surprisingly, the choice of the degrees of freedom is extremely critical; differences from the MVN method systematically grow as the number of degrees of freedom reduces, as this implies a stronger departure from the Gaussian copula. Since the t-20 copula is close to the Gaussian case, it can be easily noticed that the dispersion between the t-4, t-12 and t-20 case is much wider than the dispersion between the t-12, t-20 and MVN case. Even if we excluded the t-4 copula as an extreme choice, anyway, the difference in fair prices depending on the choice of the degrees of freedom is so remarkable that it would be a clear problem for a trader or a risk manager to adopt the method without being able to clearly justify on an empirical ground which is the optimal number of degrees of freedom he or she should really consider.

As far as individual options are considered, an interesting point is that differences between MVN method and copulas are relatively smaller for options whose payoff is somewhat constrained (as it happens for all the options from the Napoleon option to the bottom end of Table 12, since the maximum payoff is given by a fixed coupon or series of coupons and the minimum payoff is zero). This can be explained intuitively by noting that the ability of t-Student distributions and copulas to model tails and tail dependence better than in the MVN may be less relevant if the payoff is never "extreme" due to the existence of a clear predefined (albeit wide, as in the case of the 80% coupon) fluctuation range. This also suggests that if the first five options in Table 12 had been capped, by determining a maximum payoff, differences would have been lower. This also suggests that a trader willing to reduce mispricing risk if he is using a MVN method for simplicity, or because he simply lacks the time to recalibrate a more complex copula-based simulation, may first of all try to offer exotics with a capped payoff, since the cap, however out-of-the money it may be, might reduce the risk of mis-modelling the tail dependence effect.

In general, it can be noted that the impact of using copulas is clearly different from a simple change in correlation parameters, since it implies a completely different "mechanics" of price co-movements, where again the different way in which tail dependence is modeled is critical. This is evident, for instance, by noting that while an increase in (linear) correlation had an opposite effect on Asian basket and Asian best options (as observed earlier in Table 11), using copulas increases both types of options. In fact, their payoff is unconstrained and in the copula-based simulation they can clearly fully profit from higher tail dependence in the case of strong upward moves of underlying assets' prices.

5. Implications for traders, risk managers and auditors

The empirical test whose results have been just displayed has many implications for at least three different players who may be concerned about correlation and correlation risk in a multiasset exotic options' desk.

Let us consider the trader first. The trader's typical activity is to price and then sell the exotic option (e.g. when the investment bank builds a structured bond for a defined counterparty) and then either hedge the position directly or hedging just part of the risk and transferring directional risk linked to individual equities to the desks that are in charge of it through an internal deal (so to make a sort of internal Delta hedge). Therefore, he may be concerned with (a) being wrong in pricing the exotic option when he sells it and (b) being wrong in hedging the position. Correlation is then important for two reasons. First, a poor correlation estimate may lead the trader to misprice the option and to hedge it poorly, since the hedging coefficients would be wrong as well. Second, correlations among underlying assets may change through time, and therefore a change in the level of correlation may produce a change in the option's price (and in hedging coefficients) that is very hard to hedge for the trader[8]. The huge bid-ask spread that is typical for the long-term exotic equity options which are embedded into structured equity-linked bonds can be at least partially explained with the need to compensate for those risks.

As far as the trader's viewpoint is concerned, our test suggests that if the trader believes in a multivariate normally distributed world but is concerned with correlation changes, a natural way to reduce the risk would be to try to balance inside the portfolio different kind of options with opposite exposure to correlation changes. For instance, while the value of an Asian best option would increase if correlation among underlying assets increases, a conditional coupon structure that pays a fixed coupon provided that the price of all the stocks included in the basket does not fall below a certain prespecified barrier has the opposite exposure. Anyway, this solution may not work in practice since it may be difficult to persuade many final customers to buy options on almost identical

[8] *A similar point has been made by Rebonato (1999, p. xiv) who stated (with reference to the different issue of OTC options' volatility smiles) that "a trader can hope to make money from a non-plain-vanilla options strategy if her*

baskets so to hedge correlation risk almost entirely. Moreover, even if the trader sold an Asian best option and a conditional coupon structure on the same set of underlying assets, the price of the former option would be driven after a while by the group of best performing assets, while the value of the latter would mostly depend on the price of the worst performing ones. The hedging problem could therefore be only partially tempered by proposing to final customers those options which either typically have a smoother behaviour in terms of Greeks (e.g. Asian-type options) or are likely to depend after a certain period mostly on a limited number of underlying assets (such as options with a best or worst feature, where even if the stock basket is large at the beginning after a while the performance is linked to a small subset only).

If instead the trader is not so convinced about the real joint return distribution, we have seen that the impact on the fair price of different alternative assumptions reduces when the option's payoff is bounded in some manner, as it happens for instance for the last four options that can be conceived as spreads between a long position in an exotic option and a short position in a similar option with a different strike. Yet, if the trader uses more sophisticated methods – at least in order to control the prices he is quoting for new options – than the market, he could also plan to exploit the asymmetric information he has in his favour, so to suggest those options that may be overvalued under a simpler MVN pricing algorithm. The effect could be similar to what happens sometimes as far as volatility is concerned, so that some investment banks are said to suggest structured bonds where the investor is buying an option (e.g. an Asian basket call) when historical volatilities are higher than long-term implied volatilities, and to suggest structured bonds where the investor is selling an option (e.g. a put on the worst stock, or on the worst monthly return, as in the conditional coupon or in the Napoleon structure) when historical volatilities are lower than long-term implied volatilities. If this were the case, the uninformed customer who cannot observe implied volatilities but only historical ones would be lead to overestimate the bond's price as opposed to the "real" price the trader knows.

Obviously, a key issue is whether and to such extent the trader may be able to exploit this private superior information even internally, when providing the inputs to price the option. This is one of the key concerns for the risk manager. The results of our test

---

view about the future evolution of some un-tradable key quantities, on which her hedge is based, turns out to be unfortunately tell to the risk manager that the choice of both the set of inputs, even in the MVN, and especially of the assumptions on the joint distribution of asset returns may be critical. Therefore, the risk manager has in practice three alternative solutions. The first one is to strive to detect both the best method and the best set of parameters in order to price any single option in his or her portfolio, continuously revising his or her optimal choice, and force the trader to use the "best" model. Unfortunately, this task may be too computationally intensive for a bank that has to revaluate a huge portfolio of options on different sets of underlying assets.

The second possible alternative solution is to define the best model once and then adopt it with only small "maintenance" costs based on a limited input revision. This would imply for instance to choose once the distribution function for marginal distributions and the copula to be adopted (e.g., a t-Student copula with 10 degrees of freedom) to simulate dependence among assets' returns. After the initial choice, only a relatively minor estimation effort aimed at redefining marginal distribution and t-10 copula parameters would be needed. The same result could be achieved even if through time an industry standard emerged; in this case, many risk managers would probably assume the industry standard as a benchmark and refine only parameters within the "champion" model. In this case, of course, it remains questionable how frequently parameters could be revised. While in a very simplified setting a vector of implied volatilities and a correlation matrix would be sufficient to resume the trader's estimates, in a more complicated setting the communication of all the underlying pricing assumption may become more complex, and inputs revision may inadvertently become slower.

The third possible solution would be to maintain the simple model for day close portfolio repricing and use the most sophisticated models as a pricing control tool when the option is issued on one side and as a measure to capture model risk on the other. Maximum differences or dispersion measures among theoretical prices (at single option level) or theoretical portfolio values (at desk level) may be considered as a rough proxy of the amount of model risk the trader is assuming, and is reflected on the chance of making both pricing mistakes and hedging mistakes. Since the trader responds to the existence of the greater pricing and hedging risk that a multiasset exotic option may generate by overpricing the option he sells relative to its fair value, a part of the

---

similar what she assumed when pricing the option".

overpricing should be considered as a provisioning against mispricing and mishedging. As a consequence, the markup against the theoretical fair price should be attributed to the derivatives desk only *pro rata* throughout the option's life span. To a certain extent, using the complex methodology just at the date of issue, using then a simpler (and easier to check) pricing methodology and forcing the trader to split the initial markup over time may be for the risk manager a more viable and less risky solution than attributing to the trader a result which is derived by the comparison of theoretical prices obtained through complex model whose inputs would be very hard to verify in an efficient manner for the risk manager.

The constraints deriving from the computational costs of (i) adapting to changing market conditions more sophisticated pricing models and (ii) having an independent check of the same inputs by an independent risk manager are even greater if we consider that the internal auditor too could be assigned the responsibility to check the consistency of pricing algorithms and risk management procedures. Again, the more complex are the methodologies and the tougher is the task for the auditor. Since complex methodologies typically require very skilled people whose cost may be higher than average, it is quite unlikely to assume that a certain set of competencies can be easily double or tripled inside the bank. Yet, a certain understanding of the most sophisticated method might be useful for a few auditors too, so to be able to check from time to time the methodology and the process that the front-office is adopting. Moreover, auditors too may sponsor the adoption of a provisioning system that may cover the bank through time from the risks deriving from inherently uncertain pricing and the consequent potentially poor hedging that is typical for most correlation products.

6. Concluding remarks

What are the implications of our study for the diffusion of copulas for pricing multiasset equity options? In a sense, the substantial differences that copulas may produce in fair prices that have been documented in Section 4 are at the same time good news and bad news. It is good news since if differences had been small then most banks still using a standard multivariate normal assumptions would have been justified in maintaining a simpler and cheaper method as a fast approximation to the theoretically superior copula approach. Instead, the existence of wide differences suggests that using the right method can provide a substantial advantage. It is at the same time bad news since differences are so huge that on one hand the trader may face the risk to quote options so distant from other competitors' prices to be virtually excluded from the market and on the other hand a risk manager and an auditor may hardly accept to confirm day-close exotic options' mark-to-model prices if there is uncertainty or even subjectivity in defining critical parameters such as, in particular, the number of the degrees of freedom. After all, plain vanilla options are typically priced through a Black-Scholes model and a proper implied volatility estimate despite the fact that no trader or risk manager really believes in the distributional assumptions of the model. Yet, when the trader's evaluation can be summarized in an implied volatility surface, then it may become relatively easier at least for the risk manager to check the consistency of the price with the market and through time.

As a consequence, further diffusion of copulas for pricing exotic options may become possible especially if clear best practices about parameters' estimation and "leading" copula functions will emerge. There is probably at present a lack in empirical literature on efficient methods for calibrating copulas in a high dimensional setting, as it is typically required in order to use them consistently and continuously to price multiasset derivatives.

At the same time, and even experimentally, the use of copulas may be important to point out to risk managers and auditors where the greatest mispricing risks may lie within the bank's derivatives portfolio. While a lot of improvements have been made in checking implied volatility inputs, the way in which dependence among returns is modeled, relevant parameters are calibrated, correlation risk is quantified and – if not reduced through a proper balanced portfolio with "correlation-bullish" and "correlation-bearish" options – at least covered by appropriate provisioning, represent a set of topics that need great attention for those banks that run huge portfolios of complex equity derivatives.

References

BOUYÉ E. DURRLEMAN V., NIKEGHBALI A., RIBOULET G. and RONCALLI T. (2001), "Copulas : An Open Field for Risk Management", Groupe de Recherche Opérationnelle, Crédit Lyonnais, working paper.

BOUYÉ E., DURRLEMAN V., NIKEGHBALI A., RIBOULET G. and RONCALLI T. (2000), "Copulas for Finance: A Reading Guide and Some Applications", Groupe de Recherche Opérationnelle, Crédit Lyonnais, working paper.

BOUYÉ E., GAUSSEL N. and SALMON S. (2002), "Investigating Dynamic Dependence Using Copulae", Financial Econometrics Research Centre, Cass Business School, London, working paper.

COSTINOT A., RONCALLI T. and TEILETCHE J. (2000), "Revisiting the Dependence between Financial Markets with Copulas", Groupe de Recherche Opérationnelle, Crédit Lyonnais, working paper.

EMBRECHTS P., LINDSKOG F. and McNEIL A.J. (1999), "Modelling Dependence with Copulas and Applications to Risk Management", Department of Mathematics, ETH, Zurich, working paper.

EMBRECHTS P., McNEIL A.J. and STRAUMANN D. (1999), "Correlation and Dependence in Risk Management: Properties and Pitfalls", Department of Mathematics, ETH, Zurich, *Working paper*.

FREES E. e VALDEZ E. (1998), Understanding relationship using copula, *North American Actuarial Journal* **2**, 1-25.

LINDSKOG F. (2000), "Linear Correlation Estimation", *RiskLab Research Paper*.

MALEVERGNE Y. and SORNETTE D. (2001), "Testing the Gaussian Copula Hypothesis for Financial Asset Dependences", University of Nice, working paper.

NELSEN R.B. (1998), *An Introduction to Copulas*, Springer Verlag, New York.

REBONATO R. (1999), *Volatility and Correlation,* J. Wiley and Sons, Chichester, UK.

ROMANO C. (2002), "Calibrating and Simulating Copula Functions: An Application to the Italian Stock Market", *working paper*, November.

ROSENBERG J., T. SCHUERMANN, *A General Approach to Integrated Risk Management with Skewed, Fat-Tailed Risk*, Federal Reserve Bank of New York Staff Reports, n. 185, May 2004.