Value-at-Risk Estimation of Aggregated Risks Using Marginal Laws and Some Dependence Information

Andrés Cuberos

SCOR

joint work with Véronique Maume-Deschamps and Esterina Masiello (Univ Lyon 1)

The Second International Congress on
Actuarial Science and Quantitative Finance

June 17, 2016 Cartagena

Agenda

- Introduction
- Dependence in a risk vector
- Empirical, checkerboard and checkermin copulas
- Numerical Examples: VaR estimation
- Checkerboard copula with information on the tail
- Conclusions

Introduction (1/4)

Consider:

- A random vector $\mathbf{X} = (X_1, \ldots, X_d)$ on $\mathbb{R}^d$, we call it the risk vector.
- A function $\Psi : \mathbb{R}^d \to \mathbb{R}^+$, measurable and non-decreasing on each variable, called the aggregation function
  - Some examples of aggregation functions are: the sum, max, weighted sums or a slightly more complex function that may include stop-loss reinsurance type function on each of the marginals.
- The aggregated risk $\Psi(\mathbf{X})$, that represents the total future position of the company.

In the case of an insurance or reinsurance company, each component of the risk vector may be the losses of a particular line of business in the year and $\Psi(\mathbf{X})$ the total losses.

Introduction (2/4)

- To estimate the total future position $\Psi(\mathbf{X})$ we must interest not only on the individual risks X_i of the vector $\mathbf{X}$ but also on the dependence between them.
- For example, in some cases the losses of different lines of business of a company could have a tendency to be produced simultaneously.
- A good understanding of the dependence between the different risk is important to estimate the aggregated risk, moreover that in some cases the dependence is stronger when extreme events arrive.

Introduction (3/4)

Agrégation des risques

- We are interested here in the estimation of the aggregated risk.
- If we now the joint distribution function $\mathbf{F}$ of the risk vector $\mathbf{X}$, then the c.d.f. of $\Psi(\mathbf{X})$, denoted by F_Ψ, is given by

$$F_\Psi(s) = P(\Psi(\mathbf{X}) \leq s) = \int_{\mathcal{S}} d\mathbf{F}(x_1, \ldots, x_d).$$

  where $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^d \; : \; \Psi(\mathbf{x}) \leq s\}$.

- The Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR) of the aggregated risk is given by:

$$\mathrm{VaR}_p(\Psi(\mathbf{X})) = \inf\{x \in \mathbb{R}, F_\Psi(x) \geq p\}$$

$$\mathrm{TVaR}_p(\Psi(\mathbf{X})) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_u(\Psi(\mathbf{X}))\,\mathrm{d}u$$

Introduction (4/4)

- In practice, neither the marginals nor the dependence of the risk vector $\mathbf{X} = (X_1, \ldots, X_d)$ will be known.
- However, in many cases, the information available on the marginal distributions is much more important than the one on the dependence structure.
- For example, when some observations of the vector $\mathbf{X}$ are available, inferences one can do on the marginal distributions give better results than inferences one can do on the multivariate distribution.
- Samples available for marginal laws may be much larger than those available for the joint distribution.
- Moreover, on each marginal risk, some extra information may be available: for example expert opinion.

Dependence in a risk vector (1/3)

Copulas

- A *copula* $C : [0,1]^d \to [0,1]$ is the c.d.f. of a random vector on d-dimension with uniform marginals in $[0,1]$.

Sklar's Theorem:

- If $\mathbf{F}$ is a distribution function over $\mathbb{R}^d$ with marginals $F_1, \ldots, F_d$. Then, there exists a copula C such that

$$C(F_1(x_1), \ldots, F_d(x_d)) = \mathbf{F}(x_1, \ldots, x_d),$$

  for all $\mathbf{x} \in \mathbb{R}^d$. If moreover the marginals are continuous C is unique.

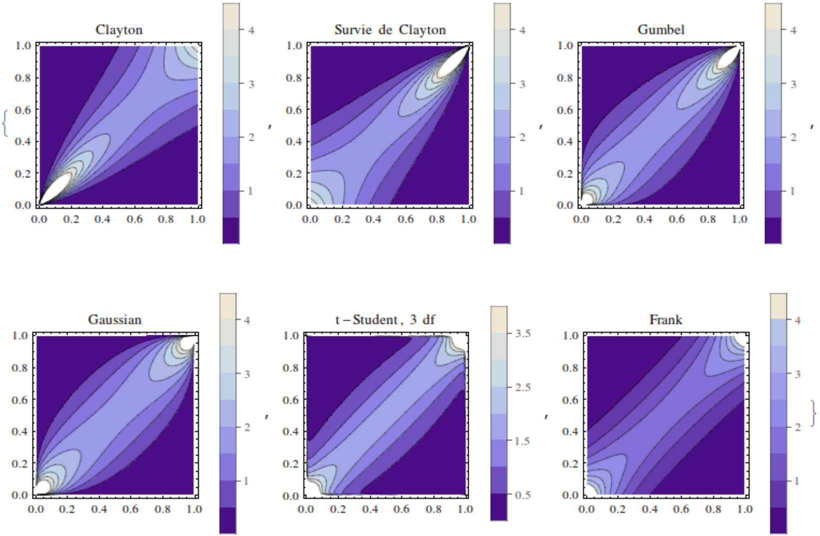
- Conversely, given a copula C and a set of marginals c.d.f. $F_1, \ldots, F_d$ then the function $\mathbf{F} : \mathbb{R}^d \to [0,1]$ defined by

$$\mathbf{F}(x_1, \ldots, x_d) := C(F_1(x_1), \ldots, F_d(x_d)),$$

  is a c.d.f on dimension d with marginals $F_1, \ldots, F_d$.

Dependence in a risk vector (2/3)

Example: Densities of some copula families



Dependence in a risk vector (3/3)

- By Sklar's Theorem, all random vector $\mathbf{X}$ is completely characterized by the marginal variables and its copula.

Sklar's Theorem allows us to separate the modeling of the marginal distributions F_i from the dependence structure.

- The marginal risks and the copula of the vector $\mathbf{X}$ determines the distribution of the aggregated risk $\Psi(\mathbf{X})$.

$$F_\Psi(s) = \mathrm{P}\left(\Psi(\mathbf{X}) \leq s\right) = \int_{\mathcal{S}} dC(u_1, \ldots, u_d).$$

where $\mathcal{S} = \{\mathbf{u} \in [0,1]^d \ : \ \Psi(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)) \leq s\}$.

Why the copula estimation is not needed to estimate aggregation

Let C be a copula and $\mathbf{X}^C = (X_1, \ldots, X_d)^C$ be a risk vector with same marginal risks as $\mathbf{X}$ but with dependence given by C.

- If Ψ is an aggregation function as the sum, the maximum or the minimum function, and $\mathbf{X}$ admits a density f then there exist a family of infinite copulas such that $\Psi(\mathbf{X}) \stackrel{\mathcal{L}}{=} \Psi(\mathbf{X}^C)$
- If $\mathbf{X}$ admits a density with identical marginals and Ψ is a symmetrical aggregation function then there exists a symmetrical copula C such that $\Psi(\mathbf{X}) \stackrel{\mathcal{L}}{=} \Psi(\mathbf{X}^C)$.

Even if the examples above may seems trivial it shows that the full knowledge of the copula distribution is unnecessary when studying an aggregation: <u>there is redundant information.</u>

Summarizing...

- The exact copula estimation can be considered as a redundant exercise when estimating the distribution of an aggregation of $\mathbf{X}$.
- In some practical cases, the information available on the marginal risks may be much more important than the one on the dependence structure.

We will provide a non-parametric method to estimate the distribution of the aggregated risk $\Psi(\mathbf{X})$, and in particular $\mathrm{VaR}_p(\Psi(\mathbf{X}))$,

- Without estimating precisely the copula or the dependence structure of the risk vector
- Assuming that robust and precise estimations of the marginal risks are provided

Principal source of information for the dependence: Observations of the joint distribution. Additionally we will explain how expert opinion on the tail could be introduced.

Empirical copula

Definition (Deheuvels 1979)

The empirical copula may be used to estimate non parametrically a copula.

- Let $\mathbf{X}^1, \dots, \mathbf{X}^n$ be n independent copies of $\mathbf{X}$. Each of them writes $\mathbf{X}^j = (X_1^j, \dots, X_d^j)$.
- Let $R_i^1, \dots, R_i^n$, $i = 1, \dots, d$ be their marginals ranks, i.e.,

$$R_i^j = \sum_{k=1}^{n} 1\{X_i^{(j)} \geq X_i^{(k)}\},\ i = 1, \dots, d,\ j = 1, \dots, n$$

  where $X_i^{(1)} < \cdots < X_i^{(n)}$ are the order statistics associated to the ith coordinate sample $(X_i^1, \dots X_i^n)$.
- The empirical copula $\widehat{C}_n$ of $\mathbf{X}^1, \dots \mathbf{X}^n$ is defined as

$$\widehat{C}_n(u_1, \dots, u_d) = \frac{1}{n} \sum_{k=1}^{n} 1 \left\{ \frac{1}{n} R_1^k \leq u_1, \dots, \frac{1}{n} R_d^k \leq u_d \right\}.$$

Checkerboard copulas (1/2)

Let μ_C be the probability measure associated to a copula C, i.e such that:

$$\mu_C\left(\prod_{i=1}^{d}[0,u_i]\right) = C(u_1,\ldots,u_d)$$

for any $\mathbf{u} = (u_1,\ldots,u_d) \in [0,1]^d$.

By a μ-decomposition of a set $A \subset \mathbb{R}^d$ we mean a finite family of measurable sets $\{A_i \subset A\}$ such that

- $\mu(A_i \cap A_j) = 0$ whenever $i \neq j$
- $\sum_i \mu(A_i) = \mu(A)$.

Checkerboard copulas (2/2)

Defintion (Li et al 1998)

A measure μ^* is a checkerboard approximation for a copula C if there exists a λ-decomposition $\mathcal{A} = \{(a_i, b_i)\}$ of $[0,1]^d$, the d-dimensional unit cube, made out of d-intervals such that for all i,

- μ^* is uniform on (a_i, b_i);
- $\mu^*(A) = \mu_C(A)$ for any $A \in \mathcal{A}$,

where λ is the d-dimensional Lebesgue measure on $[0,1]^d$.
For $m \in \mathbb{N}$, let us consider the regular λ-decomposition of the unite cube $[0,1]^d$ denoted as $\mathcal{I}_m$ and consisting of m^d d-cubes with side length $1/m$:

$$I_{i,m} = \prod_{j=1}^{d} \left[\frac{i_j - 1}{m}, \frac{i_j}{m} \right], \; i = (i_1, \ldots, i_d), \; i_j \in \{1, \ldots, m\}.$$

μ^*_m is the checkerboard approximation associated to the regular decomposition $\mathcal{I}_m$.

Checkermin copulas

Defintion (Mikunsinski et al 2010)

A measure μ^* is a checkermin approximation for a copula C if there exists a regular λ-decomposition $\mathcal{A} = \{(a_i, b_i)\}$ of I^d, the d-dimensional unit cube, made out of d-intervals such that for all i:

- μ^* is uniform on the principal diagonal of (a_i, b_i);
- $\mu^*(A) = \mu_C(A)$ for any $A \in \mathcal{A}$.

Empirical copula + Checker copula

The empirical checkerboard and checkermin copula

Let $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}$ be n independent copies of $\mathbf{X}$. Let $\widehat{\mu}$ be the probability measure associated to the empirical copula $\widehat{C_n}$ of $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}$. We define the following copulas:

- The empirical checkerboard copula (ECBC) $\widehat{C}^*_m$ is defined by

$$\widehat{C}^*_m(\mathbf{x}) = \sum_i m^d \widehat{\mu}(I_{i,m}) \lambda([\mathbf{0}, \mathbf{x}] \cap I_{i,m}).$$

- The empirical checkermin copula (ECMC) $\widehat{C}'_m$ is defined by

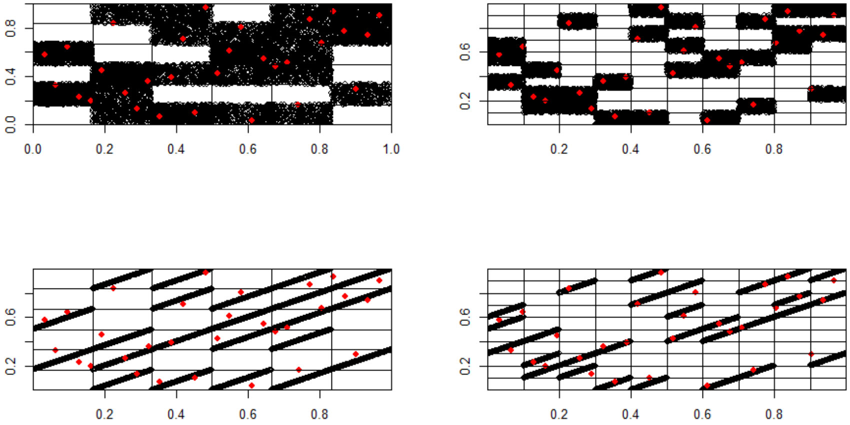
$$\widehat{C}'_m(\mathbf{x}) = \sum_i m \widehat{\mu}(I_{i,m}) 1_{(\frac{i}{m},1]}(\mathbf{x}) \min\left(x_1 - \frac{i_1}{m}, \ldots, x_d - \frac{i_d}{m}, \frac{1}{m}\right).$$

- In Li (1998) it has been shown that when C is a copula its checkerboard approximation on a regular partition is always a copula.
- This is not true in general for the empirical copulas which are not copulas.

The empirical checkerboard and checkermin copulas defined on the regular partition $\mathcal{I}_m$ and based on an i.i.d sample of size n is a copula if and only if the size of the regular partition m divides n.

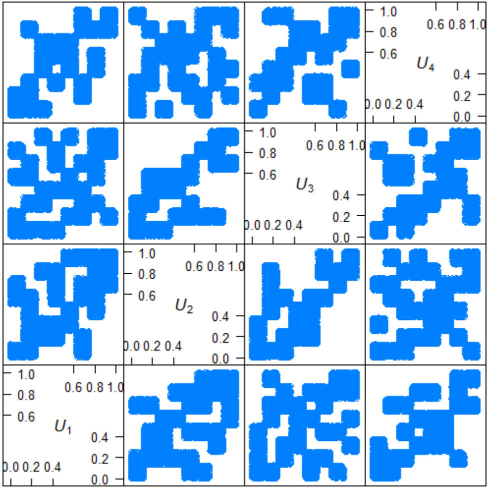
Simulation of ECBC and ECMC copulas (1/3)

At the top 2 checkerboard copulas, at the bottom 2 checkermin copulas. At the left side $m = 6$, at the right side $m = 10$. All copulas are based in a same sample of size $n = 30$ in dimension 2.



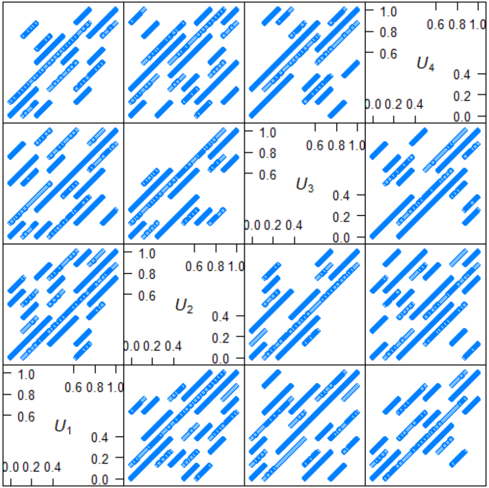
Simulation of ECBC and ECMC copulas (2/3)

Simulation of a checkerboard copulas with $m = 8$ based in a same sample of size $n = 40$ in dimension 4.



Simulation of ECBC and ECMC copulas (3/3)

Simulation of a checkermin copulas with $m = 8$ based in a same sample of size $n = 40$ in dimension 4.



Some asymptotic results

- Let m divide n, then:

$$\sup_{t\in[0,1]} |\widehat{C}_m^*(t) - C(t)| \leq O_{\mathrm{P}}\left(\frac{1}{\sqrt{n}}\right) + \frac{d}{2\,m}.$$

- Let $F_\Psi(t) = \mathrm{P}\left(\Psi(\mathbf{X}) \leq t\right)$ and $F_{\Psi,m}^* = \mathrm{P}\left(\Psi(\mathbf{X}^{\widehat{C}_m^*}) \leq t\right)$. Then, under some regularity conditions on the copula C

$$\sup_{t\in\mathbb{R}} |F_\Psi(t) - F_{\Psi,m}^*| = O_{\mathrm{P}}\left(\frac{1}{\sqrt{n}}\right) + O\left(\frac{1}{m}\right).$$

Numerical examples (1/3)

The Pareto-Clayton Model

- Let Λ be a positive random variable and let $\mathbf{X} = (X_1, \ldots, X_d)$ be a random vector such that

$$\mathrm{P}\left(X_1 > x_1, \ldots, X_d > x_d \,|\, \Lambda = \lambda\right) = \prod_{i=1}^{d} e^{-\lambda x_i},$$

  for each $x_1, \ldots, x_d \geq 0$.
- When Λ is Gamma(α, β) distributed, then the marginals of $\mathbf{X}$ are Pareto(α, β) type with dependence given by a survival Clayton copula $1/\alpha$.
- The VaR of $S = X_1 + \cdots + X_d$ can be explicitly calculated by:

$$\mathrm{VaR}_p(S) = \frac{F^{-1}(p)}{1 - F^{-1}(p)},$$

  where F is the c.d.f. of a Beta distribution with parameters $(d\beta, \alpha)$.

Numerical examples (2/3)

RMSE in % of the VaR based on 1000 estimations of the VaR of Pareto-Clayton sum in dimension 25, with a sample of size 200.

		VaR 50%	VaR 70%	VaR 80%	VaR 90%	VaR 95%	VaR 99%	VaR 99.5%	VaR 99.9%
Empirical Checkerboard copula	m=8	11%	21%	24%	18%	5%	35%	46%	62%
	m=10	8%	17%	22%	20%	9%	30%	43%	61%
	m=20	6%	8%	12%	18%	18%	17%	31%	55%
	m=40	5%	6%	8%	13%	18%	12%	20%	48%
	m=100	5%	5%	7%	9%	14%	20%	18%	38%
	m=200	5%	6%	7%	9%	12%	20%	24%	33%
Empirical Checkermin copula	m=8	6%	9%	8%	2%	4%	11%	14%	21%
	m=10	5%	8%	8%	4%	4%	11%	13%	21%
	m=20	5%	6%	8%	8%	5%	11%	14%	23%
	m=40	5%	6%	7%	9%	9%	10%	15%	23%
	m=100	5%	5%	7%	9%	11%	12%	14%	26%
	m=200	5%	6%	7%	9%	12%	14%	15%	27%
Parametric Copula	Gauss	19%	17%	13%	5%	5%	26%	33%	48%
	Survival Clayton	1%	1%	1%	2%	3%	5%	7%	17%
	Clayton	44%	30%	17%	4%	21%	49%	56%	66%
	Empirical	5%	6%	7%	9%	11%	28%	35%	59%

Numerical examples (2/3)

RMSE in % of the VaR based on 1000 estimations of the VaR of a sum of lognormal distributions with a Gaussian copula with correlations between 0.25 and 0.75 in dimension 25, with a sample of size 60.

		VaR 50%	VaR 70%	VaR 80%	VaR 90%	VaR 95%	VaR 99%	VaR 99.5%	VaR 99.9%
Empirical Checkerboard copula	m=8	8%	10%	9%	4%	10%	30%	37%	47%
	m=10	7%	9%	8%	5%	8%	28%	35%	46%
	m=20	5%	7%	7%	7%	6%	22%	29%	42%
	m=40	5%	6%	6%	7%	7%	16%	23%	37%
	m=100	5%	6%	6%	7%	8%	15%	21%	35%
	m=200	5%	6%	6%	7%	10%	18%	21%	33%
Empirical Checkermin copula	m=8	3%	2%	2%	5%	5%	5%	9%	19%
	m=10	4%	3%	2%	4%	5%	5%	8%	17%
	m=20	4%	4%	3%	3%	5%	8%	8%	12%
	m=40	4%	5%	5%	4%	5%	10%	12%	15%
	m=100	5%	5%	5%	5%	6%	11%	13%	18%
	m=200	5%	6%	6%	6%	8%	14%	18%	25%
Parametric Copula	Gauss	1%	1%	1%	2%	2%	4%	5%	9%
	Survival Clayton	3%	7%	7%	6%	4%	10%	15%	28%
	Clayton	18%	7%	2%	13%	23%	39%	44%	52%
	Empirical	6%	7%	6%	9%	11%	21%	29%	49%

Checkerboard copula with information on the tail (1/3)

- We say that a set of the form $E_p = [0,1]^d \setminus [0,p]^d$ corresponds to the tail of the copula when p is close to 1.

- Consider that information is given on $\mathcal{E}_p$, the λ-decomposition of E_p consisting of the hyper rectangles $[a_1, b_1] \times \cdots \times [a_d, b_d]$ where $[a_i, b_i] = [0, p]$ or $[a_i, b_i] = [p, 1]$ for all $i = 1, \ldots d$ with at least one of $[a_i, b_i] = [p, 1]$.

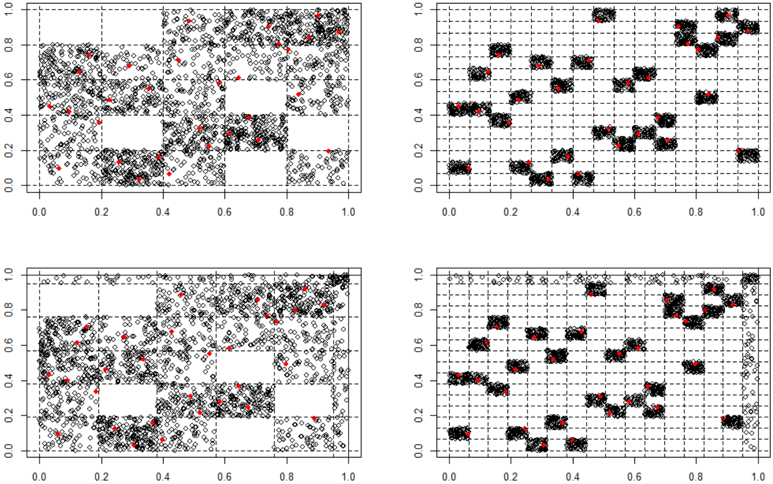
- The ECBC with information on $\mathcal{E}_p$ is given by

$$\widehat{C}_m^{\mathcal{E}_p}(\mathbf{x}) = (1 - \mu_C(E_p))\widehat{C}_m^{*}(\mathbf{x}_p) + \sum_{E \in \mathcal{E}_p} \frac{\mu_C(E)}{\lambda(E)} \lambda([\mathbf{0}, \mathbf{x}] \cap E),$$

where $\mathbf{x}_p = (\min\{x_1/p, 1\}, \ldots, \min\{x_d/p, 1\})$.

Checkerboard copula with information on the tail (2/3)

At the top 2 ECBC copulas without information on the tail, at the bottom the same copulas with information on $\mathcal{E}_{0.95}$. At the left side $m = 5$, at the right side $m = 15$. All copulas are based in a same sample of size $n = 30$ in dimension 2.



Checkerboard copula with information on the tail (3/3)

Mean and RMSE in % of the exact VaR based on 1000 estimations of the VaR of Pareto-Clayton sum in dimension 2, with a sample of size 30.

VaR	80%	90%	95%	99%	99.5%	99.9%
Exact value	2.5	4.1	6.4	16.0	23.2	53.4
Empiric	2.5	4.0	6.1	12.2	13.2	14.0
	(26%)	(31%)	(39%)	(72%)	(70%)	(78%)
ECBC (m=6)						
No information	2.6	4.4	6.6	14.8	20.8	45.7
	(9%)	(8%)	(6%)	(8%)	(11%)	(15%)
With information on $\mathcal{E}_p$	2.6	4.4	6.4	14.2	22.7	49.5
p=0.99	(9%)	(8%)	(5%)	(11%)	(3%)	(8%)
With information on $\mathcal{E}_p$	2.7	4.1	6.1	15.6	21.8	46.8
p=0.95	(10%)	(5%)	(4%)	(3%)	(6%)	(13%)
ECBC (m=15)						
No information	2.5	4.2	6.8	15.5	21.5	46.4
	(12%)	(13%)	(11%)	(9%)	(10%)	(14%)
With information on $\mathcal{E}_p$	2.5	4.3	6.8	14.3	22.7	49.5
p=0.99	(12%)	(12%)	(12%)	(11%)	(3%)	(8%)
With information on $\mathcal{E}_p$	2.6	4.3	6.2	15.6	21.8	46.8
p=0.95	(11%)	(10%)	(4%)	(3%)	(6%)	(13%)
ECBC (m=30)						
No information	2.5	4.2	6.6	15.8	22.0	47.0
	(13%)	(15%)	(17%)	(13%)	(12%)	(14%)
With information on $\mathcal{E}_p$	2.5	4.2	6.7	14.3	22.7	49.5
p=0.99	(13%)	(16%)	(16%)	(11%)	(3%)	(8%)
With information on $\mathcal{E}_p$	2.6	4.4	6.2	15.6	21.8	46.8
p=0.95	(13%)	(11%)	(4%)	(3%)	(6%)	(13%)

Conclusions

- If the principal objective is to estimate the aggregated risk, all the information of the dependence structure given by a copula is not necessary
- We have introduced a set of copulas that combines the concepts of the empirical copulas with the checkerboard and checkermin copulas
- In comparison to alternative methods, as the use of the empirical copula or the estimation of a parametric copula from data, the empirical checkermin give good results when estimating the VaR of the aggregated risk
- Moreover, if additional information can be given in the tail of the joint distribution, the method could be adapted in order to integrate this information. The estimation of the VaR at confidence levels close to 1 improves.