UNIVERSITÀ
DEGLI STUDI
DI BRESCIA

DOTTORATO DI RICERCA IN
Modelli e Metodi per l'Economia e il Management
(Analytics for Economics and Management - AEM)
Dipartimento di Economia e Management
Università degli Studi di Brescia
Settore scientifico – disciplinare
SECS-S/06 METODI MATEMATICI DELL'ECONOMIA E
DELLE SCIENZE ATTUARIALI E FINANZIARIE

CICLO XXXIII - A.A. 2019/2020

Topics in Dependence Modelling in Finance

DOTTORANDO

Dott. Paolo Neri

Paolo Neri

RELATORE
Prof. Umberto Cherubini

CO- RELATORE
Prof. Cristian Pelizzari

Pelizzari Cristian

Abstract

L'obiettivo della tesi è quello di analizzare due principali temi di finanza quantitativa, quali il risk management e lo studio dei rendimenti composti. Tali tematiche sono di grande interesse da parte di soggetti intermediari finanziari, quali banche e assicurazioni perchè riguardano principalmente il capitale da sottoporre a riserva e le previsione dei rendimenti futuri sulle attività finanziarie.

All'interno della tesi il filo che lega questi due temi da un punto di vista statistico è l'utilizzo delle *funzioni di Copula* come modello per la dipendenza tra le attività finanziarie.

Più nello specifico, nel capitolo 2 applichiamo la convoluzione di copula, nota come *C-convoluzione*, al problema dell'aggregazione del Value-at-Risk (VaR) per i rischi con code pesanti. Come esempio rappresentativo di distribuzioni a code spesse, è stata considerata la classe delle α-stabili. Con la C-convoluzione, confermiamo i risultati sulla super additività del VaR. Si mostra che il VaR è non sub-additivo per le α-stabili con livelli del parametro $\alpha < 1$. In questi casi l'aggregazione di due rischi i.d. in portafoglio produrrà sempre una super-additività: il capitale necessario per assicurare insieme i due rischi sarà sempre superiore alla somma del capitale per i rischi marginali. Questo è vero per tutte le strutture di dipendenza, cioè le funzioni di copula applicate nell'analisi (ellittiche e archimedee), e confermato sopratutto per livelli di dipendenza positivi.

Nel Capitolo 3, si considera l'analisi su una proprietà delle funzioni di copula, la dipendenza di coda. Quando il VaR è super-additivo, l'aumento di capitale dovuto all'aggregazione risulta inferiore per le funzioni di copula con maggiore dipendenza dalla coda. Questo risultato sembra controintuitivo, contradditorio, poiché ci si aspetta che la dipendenza di coda sia un fattore in grado far accrescere la quantità di rischio associato, in altri termini la quantità di capitale a riserva. Si mostra che questo "puzzle" sulla dipendenza di coda si presenta anche nei casi in cui il VaR è sub-additivo. Quindi, tale puzzle persiste anche nei casi con rischi marginali α-stabili con $\alpha > 1$, ovvero t-Student con comportamento di coda simile, per i quali il rischio misurato in termini di VaR aggregato è sub-additivo. In particolare, la funzione di copula t-Student con una maggiore dipendenza di coda mostra una riduzione del requisito patrimoniale calcolato utilizzando il framework del VaR, rispetto alle funzioni di copula, come la Gaussiana, per le quale, come noto, i rischi nelle code sono asintoticamente indipendenti. Questo è chiaramente un paradosso perché sarebbe naturale aspettarsi che maggiore probabilità condizionata di eventi estremi, cioè una maggiore probabilità che due rischi estremi si materializzino insieme, debba sempre tradursi in requisiti di capitale più elevati. Documentiamo che non è così. Osserviamo che il puzzle scompare quando i rischi marginali sono Gaussiani. Quindi, il puzzle della dipendenza dalla coda è determinato in modo cruciale dalle code sul rischio marginale. Mentre l'aggregazione del rischio con marginali a code "sottili" assume un costo del capitale maggiore, se i rischi dipendono dalla coda, una dipendenza di coda accentuata consente di risparmiare capitale se i rischi marginali hanno code sufficientemente "pesanti".

Nel Capitolo 4 affrontiamo un argomento più classico nella teoria delle funzioni di copula, ovvero la generalizzazione e la distorsione della funzione prodotto. È noto che se i fattori del prodotto sono variabili casuali uniformi nell'intervallo unitario, questa procedura genera la famiglia delle copule di Archimedee. Sfruttiamo la stessa idea per un'applicazione diversa. Il prodotto è il nostro riferimento per la composizione geometrica utilizzata nella valutazione delle attività. Si propone una distorsione di questo principio di composizione geometrica. La logica economica della distorsione è fornita dal concetto di *compounding generalizzato* e basato su modelli di cambio di tempo stocastico. L'idea è che se i rendimenti sono composti o scontati a tempo discreto secondo un orologio stocastico, ciò determina una distorsione che può essere rappresentata in termini della trasformata di Laplace. Quindi, con lo stesso strumento proposto da Marshall and Olkin (1988) per le funzioni copula, viene esaminata una famiglia di copule Archimedee per la attualizzazione e lo sconto dei rendimenti.

Contents

1 Introduction **5**

1.1 Thesis purpose 5

1.2 Copula Functions 6

1.2.1 The Mathematics of Copula Functions 7

1.2.2 Copula Functions and Joint Probability Distributions 9

1.2.3 Conditional Probabilities 11

1.2.4 Non Parametric Dependence Measures 12

1.2.5 Main Copula Functions 14

1.2.6 Tail dependence measure 18

1.3 Value-at-Risk and Risk Measures (VaR) 20

1.4 Stochastic Time Models 21

1.5 Outline of the main results 22

2 Value-at-Risk on α-stable Risks: The Limits of Diversification **25**

2.1 Introduction 26

2.2 Preliminaries: Convolution of α-stable Risks 27

2.2.1 Stable Distributions 27

2.2.2 Convolution of independent variables 30

2.2.3 Copula Functions and Dependence 33

2.2.4 Convolution of dependent variables 35

2.3 Application 37

2.3.1 Convolution through copulas 37

2.3.2 Convolution and general results 40

2.3.3 Diversification failure and tail dependence 43

2.4 Conclusions 46

3 Value-at-Risk and the Tail Dependence Puzzle 49
3.1 Introduction . . . 50
3.2 Preliminaries: Convolution . . . 52
3.2.1 Copula functions and Dependence . . . 52
3.2.2 Convolution of dependent variables . . . 54
3.2.3 Student's t representation . . . 57
3.3 Tail-dependence analysis through C-convolution . . . 59
3.3.1 t-copula convolution distribution . . . 59
3.3.2 t-copula convolution derivatives . . . 60
3.3.3 Conditions for the inversion of the risk ranking . . . 63
3.4 A Real World Example . . . 69
3.5 Conclusions . . . 70
3.6 Appendix A . . . 71
3.6.1 The derivative of the t-copula convolution w.r.t. ν . . . 71
3.6.2 Partial Derivative $\frac{\partial x_1}{\partial z}$ $x_1 = t_\nu^{-1}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right)$ . . . 72
3.6.3 Partial Derivative $\frac{\partial x_1}{\partial \alpha}$ $x_1 = t_\nu^{-1}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right)$ . . . 73
3.6.4 Derivative of $\frac{\partial x_1}{\partial \nu}$ w.r.t. α . . . 73
3.7 Appendix B . . . 74
3.7.1 Mixed derivative of the t-distribution w.r.t. (ν, α) . . . 74
3.7.2 Derivative of the t-convolution density w.r.t. ν . . . 75
3.7.3 Derivative of the t-convolution w.r.t. ν at z=0 is negative . . . 77
3.7.4 Derivative of the t-density function w.r.t. ν . . . 80
3.7.5 First derivative of the t-density function . . . 81
3.7.6 Mixed derivative of the t-density function w.r.t. (z, ν) . . . 81
3.7.7 The sign of $G(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega)$ for $\alpha \to 0$, not correlated case $\rho = 0$ . . . 83
3.7.8 The sign of $G(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega)$ for $\alpha \to 0$, correlated case $\rho > 0$ . . . 84

4 Time change, Generalized Compounding and the Term Structure 87
4.1 Introduction . . . 88
4.2 A SDF Model with Time Change . . . 90
4.3 A Lognormal SDF Model Quadratic in Risk Factors . . . 92
4.4 Generalisations . . . 94
4.5 Term Structure . . . 97
4.6 Empirical Analysis . . . 100
4.7 Conclusions . . . 101

Chapter 1

Introduction

1.1 Thesis purpose

Copula functions have mainly been known as probability tools for the analysis of multivariate distributions. On one side in this thesis we show that many problems in the application to probability remain open, and are problems that at first sight may look quite basic. On the other side, we show that copula functions as mathematical objects can be applied to problems that are more general than probability theory.

We start with the key issue that links all the chapters of this thesis. Consider the following problem: assume X and Y are random variables with assigned probability distributions F and G. We ask what is the distribution of

$$Z \equiv X + Y$$

At first glance, the problem looks basic and easy to solve because we are used to think in terms of elliptical or stable random variables. Indeed, recovering the distribution of Z in full generality is a formidable task.

If X and Y are independent, the solution is known as the *convolution*, that is written as

$$F_Z(z) = \int F(z - G^{-1}(\omega))d\omega$$

This tool was generalised to the case in which the variables are dependent, with dependence represented by copula function C.

Definition 1.1.1 (Cherubini et al. (2011))**.** Lef F, G be two continuous c.d.f's and C a copula function. The *C-convolution* of G and F is defined as the c.d.f.

$$F \overset{C}{*} G(z) = \int_0^1 D_1C\Big(\omega, F(z - G^{-1}(\omega))\Big)\ d\omega$$

Chapter 2 and 3 apply the concept of C-convolution to the aggregation of risks.

If the distribution of a sum is an open problem of copula functions theory, the product is not solved as well. In fact, the two problems are clearly linked:

$$\log XY = \log X + \log Y$$

Of course, if one had a solution for one problem, he would have solved the other as well. Nevertheless, there is another way to address the product. In fact, the product can also be written as

$$XY = \exp\left(-[(-\log X) + (-\log Y)]\right)$$

Notice that if we set $\varphi(x) = -\log x$, this can be written

$$XY = \varphi^{-1}\left(\varphi(X) + \varphi(Y)\right)$$

This representation, for general functions with behaviour similar to $\varphi(x) = -\log x$ reminds of aggregation operators called *Archimedean*. In the case in which $X, Y \in [0, 1]$, the aggregation operator is a t-norm and the function $\varphi(x)$ is called the *generator*. Two pioneering papers by Schweizer and Sklar (1961) and Ling (1965) proved that generators exist for all Archimedean t-norms. The case in which the generator is convex defines the family of Archimedean copulas, that will be introduced later on. The case in in which X, Y is extended to the domain $[0, \infty]$ was addressed by Sugeno and Murofushi (1987).

In the third paper of this thesis the Archimedean structure will be applied to the problem of compounding and discounting. If one denotes P_k the one period discounting function, the corresponding n-period discounting is defined as

$$P^{(n)} = P_1 P_2 \dots P_{n-1} P_n$$

This product function, that does not have almost anything to do with multivariate probability, can be addressed with the same copula function tools that are applied to probability theory.

The rest of this Chapter will be devoted to provide a reference of the main concepts and tools used throughout the thesis: copula functions, VaR and risk measures, and the basic principles behind time change models to which we will refer in the final chapter.

1.2 Copula Functions

Copula functions are mostly known as statistical tools that allow to break joint distribution functions in marginal distributions and a function representing dependence. More to the point, let $F(x) = P[X \le x]$ and $G(y) = P[Y \le y]$ be two distribution functions related to the joint distribution $H(x, y) = P[X \le x, Y \le y]$. Since by the probability integral transformation the variables can be transformed into uniform random variables, in $[0, 1]$, the joint distribution $H(x, y) = P[X \leq x, Y \leq y]$ can be written in terms of uniforms. In statistics, this joint distribution taking uniform marginals is called *Copula function.*

However, copula functions are mathematical objects before being statistical tools, and the representation of joint distributions is only the main application. In this work, the same object will also be applied to a different concept, that is compounding and discounting. For this reason we first introduce the mathematics of copula functions.

1.2.1 The Mathematics of Copula Functions

We start describing the main concepts and properties necessary for the definition of a copula function as a mathematical object.

Definition 1.2.1 (H-Volume)**.** Let S_1, S_2 be two non-empty sets on $\overline{\mathbb{R}}$ and assume a function H of two real variables such that $DomH = S_1 \times S_2$. Let $B = [x_1, x_2] \times [y_1, y_2]$ the rectangle in $DomH$. Then, the *H-Volume* of B is given by

$$V_H(B) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1)$$

Definition 1.2.2 (2-increasing)**.** Let H be a two-place real function. H is *2-increasing* if $V_H(B) \geq 0$ for each rectangle B within $DomH$

Lemma 1.2.3. *Let S_1, S_2 two non-empty subsets of $\overline{\mathbb{R}}$ and let H be a 2-increasing function with domain $S_1 \times S_2$. Given the points $x_1, x_2 \in S_1$ and $y_1, y_2 \in S_2$, then the function $t \to H(t, y_2) - H(t, y_1)$ is not decreasing on S_1 and $t \to H(x_2, t) - H(x_1, t)$ is not decreasing on S_2*

From this lemma we can determine, adding a further hypothesis, that each function H, defined as above, is non-decreasing on its arguments. Such hypothesis is *groundedness.*

Definition 1.2.4 (grounded)**.** Let H be a two-place real function on $DomH = S_1 \times S_2$. Let a_1 the minimum of S_1 and a_2 the minimum of S_2. THen H is *grounded* if $H(x, a_2) = H(a_1, y) = 0 \; \forall \; (x, y) \; \in (S_1 \times S_2)$

Lemma 1.2.5. *Let S_1, S_2 two non-empty subsets of $\overline{\mathbb{R}}$ and H a 2-increasing, grounded function on $S_1 \times S_2$. Then H is non increasing on each argument.*

A useful property of copula that can be used in what follow is:

Definition 1.2.6. Let S_1, S_2 be two non-empty subsets of $\overline{\mathbb{R}}$ with greatest elements b_1 and b_2 respectively, and H a function from $S_1 \times S_2$ into $\overline{\mathbb{R}}$. We say that a function H has margins F, G if

$$DomF = S_1 \quad F(x) = H(x, b_2) \; \forall x \in S_1, \; b_2 = \max(S_1)$$

$$DomG = S_2 \quad G(y) = H(b_1, y) \; \forall y \in S_2, \; b_1 = \max(S_2)$$

Lemma 1.2.7. *Let S_1, S_2 two non-empty subsets of $\overline{\mathbb{R}}$ and H a grounded 2-increasing on $S_1 \times S_2$. Let (x_2, y_2) and (x_1, y_1) points on $S_1 \times S_2$. It follows*

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|$$

Proof in (Nelsen, 2006, Lemma 2.1.5) . □

We approach now the definition of copula starting from a restriction on the marginals domain that is *subcopula*.

Definition 1.2.8 (subcopula)**.** A *subcopula* $C^{'}$ is a function with the following properties

1. $DomC^{'} = S_1 \times S_2$ where $S_1, S_2 \subseteq \mathbf{I}$
2. $C^{'}$ is *2-increasing, grounded*
3. $\forall u \in S_1 \; e \; v \in S_2,\; C^{'}(u, 1) = u \; e \; C^{'}(1, v) = v$

Definition 1.2.9 (copula)**.** A two-dimensional *copula* (2-copula) is a subcopula C defined on the domain $DomC = S_1 \times S_2 = \mathbf{I^2}$, such that

1. $\forall\, u, v \in \mathbf{I}$

$$C(u, 0) = C(0, v) = 0 \tag{1.1}$$

$$C(u, 1) = u \qquad C(1, v) = v \tag{1.2}$$

2. $\forall\, u_1, u_2, v_1, v_2 \in \mathbf{I}$ tale che $u_1 \leq u_2 \; e \; v_1 \leq v_2$

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0 \tag{1.3}$$

Next theorem characterizes copulas functions, providing an upper and lower bound to the set of copulas.

Theorem 1.2.10. *Let $C^{'}$ a subcopula. Then $\forall (u, v) \in DomC^{'}$*

$$\max(u + v - 1, 0) \leq C^{'}(u, v) \leq \min(u, v)$$

See (Nelsen, 2006, Theorem 2.2.3) . □

Since the theorem is true for subcopulas, it remains true for copulas because each copula is a subcopula. Furthermore, the upper and lower bounds showed in the previous result are themselves copulas. Given $M(u, v) =$

$\min(u,v)$ the *maximum* copula, $W(u,v) = \max(u+v-1,0)$ the *minimal* copula, that refers to the *upper and lower bound or Fréchet-Hoeffding bounds.*

Then the previous theorem (1.2.10) can be extended to the copulas.

Theorem 1.2.11. *Let C a copula. Then $\forall(u,v) \in DomC$, holds*

(1.4)

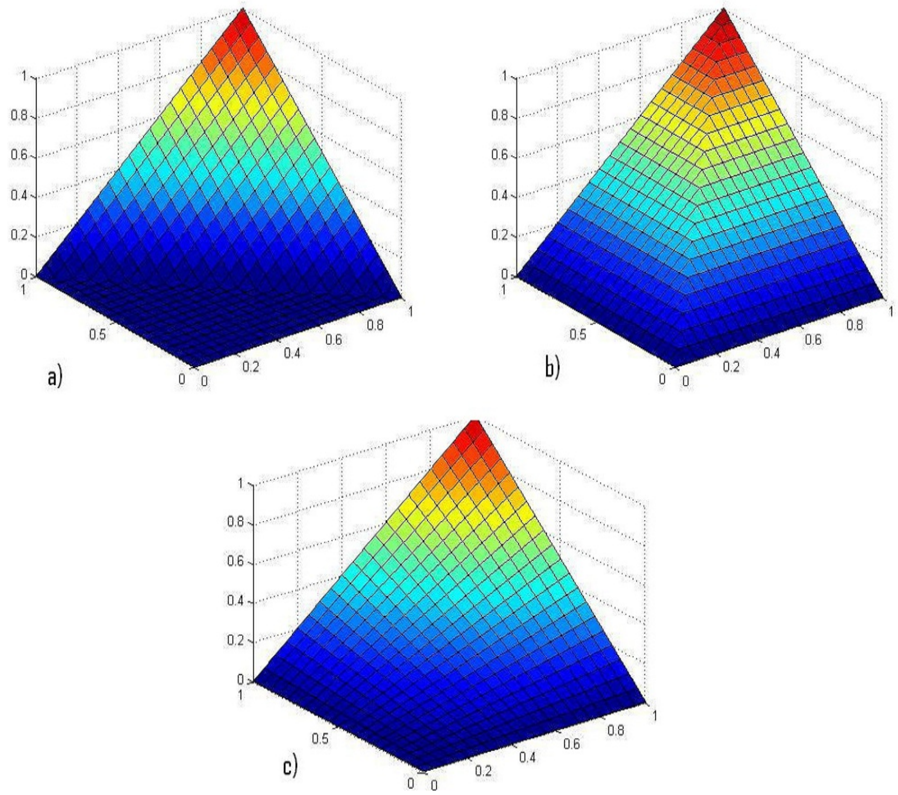


Figure 1.1: Copulas cdf
a) W(u,v) = max(u + v - 1, 0) b) M(u,v) = min(u,v) c) $\prod$(u,v) = uv

1.2.2 Copula Functions and Joint Probability Distributions

We now apply the results above to the analysis of joint distributions.

Definition 1.2.12. A *distribution function* is a function F with domain $\overline{\mathbb{R}}$ such that

1. F is non-decreasing
2. $F(-\infty) = 0$ and $F(\infty) = 1$.

Definition 1.2.13. A *joint distribution function* is a function F with domain $\overline{\mathbb{R}}^2$ such that

1. H is 2-increasing
2. $H(-\infty, y) = H(x, -\infty) = 0$ and $H(\infty, \infty) = 1$.

We can then now state the main result of copula function application to probability theory.

Theorem 1.2.14 (Sklar, 1959)**.** *Let H the join distribution with margins F, G . Then exists a copula C such that* $\forall\, x, y \in \overline{\mathbb{R}}$

$$H(x, y) = C\left(F(x), G(y)\right) \tag{1.5}$$

If F and G are continuous, then C is unique: otherwise, it is uniquely determined on RanF X RanG*. Conversely, if C is a copula and F and G are distribution functions, then the function H defined in (1.5) is a joint distribution function with margins F and G.*

There is then a one to one relationship between copula functions with given margins and joint distributions. A straigtforward way to extract a copula function from an assigned joint distribution is to use the inverse of the margins. However, first we must introduce a general definition of inverse to allow for cases when marginals are not strictly increasing.

Definition 1.2.15 (quasi-inverse)**.** Let F be a distribution function. Then the *quasi-inverse* of F is any function $F^{(-1)}$ on **I** such that

1. if $t \in DomF$ then $F^{(-1)}$ is any number $\in \overline{\mathbb{R}}$ where

$$F(x) = t \qquad \forall t \in ImF, \qquad F(F^{(-1)}(t)) = t$$

2. if $t \notin DomF$ then

$$F^{(-1)}(t) = \inf\{x|F(x) \geq t\} = \sup\{x|F(x) \leq t\}$$

Intuitively, F strictly increasing implies the quasi-inverse function is equal to the ordinary inverse F^{-1}, while when the distribution is flat it picks up the first point of the flat part. Now, one can express the copula by means a quasi-inverse transformation on the marginals.

Lemma 1.2.16. *Let H, F, G distributions as in Sklar's theorem (1.5). The sub-copula C' and marginal quasi-inverse $F^{(-1)}$, $G^{(-1)}$ are given. Then for all $(u, v) \in DomC'$,*

$$C'(u, v) = H(F^{(-1)}(u), G^{(-1)}(v))$$

Notice that by the *probability integral transformation* theorem, u and v are uniform random variables in $[0, 1]$. So, the above results allows to get a copula function from any joint distribution functions once that the marginal distributions F and G are known.

1.2.3 Conditional Probabilities

Copula functions allow to recover in full generality important results for conditional probability.

The first result will be useful later on to define the *tail dependence index*.

Theorem 1.2.17. *Assume $H(x, y) = C(F(x), G(y)) \equiv C(u, v)$, then*

$$P(X \leq x | Y \leq y) = \frac{C(u, v)}{v} \tag{1.6}$$

Proof.

$$\begin{aligned} P(X \leq x | Y \leq y) &= \frac{H(x, y)}{F(y)} \\ &= \frac{C(F(x), G(y))}{F(y)} \\ &= \frac{C(u, v)}{v} \end{aligned}$$

□

The second result has plenty of applications, including Markov processes and the concept of C-convolution, used in this work.

Theorem 1.2.18. *Assume $H(x, y) = C(F(x), G(y)) \equiv C(u, v)$ where F, G are strictly continuous marginal distributions. Then*

$$P(X \leq x | Y = y) = \frac{\partial}{\partial v} C(u, v) \tag{1.7}$$

Proof.

$$\begin{aligned} P(X \leq x | Y = y) &= \lim_{h \to 0} P(X \leq x | y \leq Y \leq y + h) \\ &= \lim_{h \to 0} \frac{H(x, y + h) - H(x, y)}{F(y + h) - F(y)} \\ &= \lim_{h \to 0} \frac{C(F(x), G(y + h)) - C(F(x), G(y))}{F(y + h) - F(y)} \\ &= \lim_{h \to 0} \frac{C(F(x), G(y) + \Delta(h)) - C(F(x), G(y))}{\Delta(h)} \\ &= \lim_{h \to 0} \frac{C(u, v + \Delta(h)) - C(u, v)}{\Delta(h)} \\ &= \frac{\partial}{\partial v} C(u, v) \end{aligned}$$

□

1.2.4 Non Parametric Dependence Measures

Copula functions are linked to non parametric association measures among random variables. These measures fix the well known problems of the standard dependence measure, that is Pearson correlation. This measure is limited by the following issues

1. It is a linear measure of dependence

2. It is affected by the shape of marginals

Linear correlation can model very well the case of elliptical models such as multivariate normal and Student's t distributions. In fact, in these cases the distribution is fully described by a vector of means and a covariance matrix. Instead, in the context of copulas in general we deal with non-linear transformations of random variables. Except for the case of elliptical models, cited above, then, Pearson correlation does not preserve invariance under non-linear transformations, and it is not appropriate to represent dependence among these random variables. Measures built on non parametric transformations of the marginal distributions are instead naturally related to copula functions.

Rank correlation measure We start recalling the definition of Pearson linear correlation

Definition 1.2.19. Let (X, Y) a random vector with finite positive variance. Then linear correlation is defined as

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

Rank correlation is the most straight way to overcome the shortcomings in linear correlation. It consists of applying the Pearson correlation definition to the ranks of the variables rather than the variables themselves. Ranks are defined as the empirical estimator of the marginal distributions of the random variables.
Non parametric measures based on copulas have the following properties

- they take values in [-1,1]

- they give value zero to independence,

- extreme values $1, -1$ imply comonotonic and countermonotonic r.v.'s respectively.

To understand these measures, we introduce the concepts of *concordance* and *discordance* among r.v.'s. Intuitively, a couple of random vector $(\mathbf{X}, \mathbf{Y})$ are *concordant* whether large values of $\mathbf{X}$ are linked by big values $\mathbf{Y}$ and the same can occurs for low values.

Definition 1.2.20 (Concordance)**.** Two couple of observations (x_1, y_1) and (x_2, y_2) of a random vector(X, Y) are concordant if $x_1 < x_2$ and $y_1 < y_2$ or $x_1 > x_2$ and $y_1 > y_2$. Similarly (x_1, y_1) and (x_2, y_2) are discordant if

$x_1 < x_2$ and $y_1 > y_2$ or $x_1 > x_2$ e $y_1 < y_2$. Equivalently, (x_1, y_1) and (x_2, y_2) are concordant if $(x_2 - x_1)(y_2 - y_1) > 0$ are discordant if $(x_2 - x_1)(y_2 - y_1) < 0$

Kendall's τ. Consider a random sample of n observation $(x_1, y_1), \dots, (x_n, y_n)$ from a vector $(\mathbf{X}, \mathbf{Y})$ of r.v.'s. There exist $\binom{n}{2}$ distinct pairs from which to compute concordance: in general 'c' couple are concordant and 'd' discordant, where d= n-c. Kendall's τ is defined as

$$\tau = \frac{c - d}{c + u} = \frac{(c - d)}{\binom{n}{2}}$$

Kendall's τ can be defined as the probability of concordance minus the probability of discordance.

$$\tau = P[(X_2 - X_1)(Y_2 - Y_1) > 0] - P[(X_2 - X_1)(Y_2 - Y_1) < 0]$$

The link to copula functions can be recovered in the theorem below.

Theorem 1.2.21 (Nelsen (2006, 5.1.1))**.** *Let (X_1, Y_1) and (X_2, Y_2) continuous random vectors with joint distributions H_1, H_2 respectively and same marginals F on (X_1, X_2) and G on (Y_1, Y_2). Let C_1, C_2 be the copula functions linking (X_1, Y_1) and (X_2, Y_2) respectively, $H_1(x, y) = C_1(F(x), G(y))$ e $H_2(x, y) = C_2(F(x), G(y))$. Let Q defined as above*

$$Q = P[(X_2 - X_1)(Y_2 - Y_1) > 0] - P[(X_2 - X_1)(Y_2 - Y_1) < 0]$$

Then

$$Q = Q(C_1, C_2) = 4 \iint_{\mathbf{I}^2} C_2(u, v)\, dC_1(u, v) - 1$$

If the copula function is unique, $C_1 = C_2 = C$, then

$$Q = Q(C) = 4 \iint_{\mathbf{I}^2} C(u, v)\, dC(u, v) - 1 \tag{1.8}$$

The result can be rewritten as the expectation of the copula function $C(U, V)$ with uniform r.v.'s:

$$\tau(X, Y) = 4\mathbb{E}[C(U, V)] - 1$$

Spearman's ρ. We introduce the other main rank correlation measure of association.

Definition 1.2.22. Let $(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)$ three independent random vectors, where (X_i, Y_i) has common join distribution H and marginals F and G. and copula C. The rank correlation *Spearman's* ρ is a quantity proportional to the difference between concordance and discordance probabilities on (X_1, Y_1) e (X_2, Y_3).

$$\rho(X,Y) = P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]$$

Theorem 1.2.23. *Let X, Y continuous r.v.'s with joint distribution H and copula C. Then the correlation Spearman's ρ is given by*

$$\rho(X,Y) = 3Q(C,\Pi) = 12 \iint_{\mathbf{I}^2} C(u,v)\, dC(u,v) - 3$$

1.2.5 Main Copula Functions

Here we describe the main copula functions used in applications. For each copula the main properties are described.

Gaussian copula The *Gaussian* copula is derived from the multivariate normal distribution. Let Φ^n and Φ^{-1} the multivariate standard normal distribution with Pearson correlation matrix Σ and the inverse of univariate standard respectively. From Sklar's theorem, the multivariate Gaussian copula is given as

$$C(u_1, u_2, ..., u_n) = \Phi^n(\Phi^{-1}(u_1), \Phi^{-1}(u_2), ..., \Phi^{-1}(u_n); \Sigma)$$

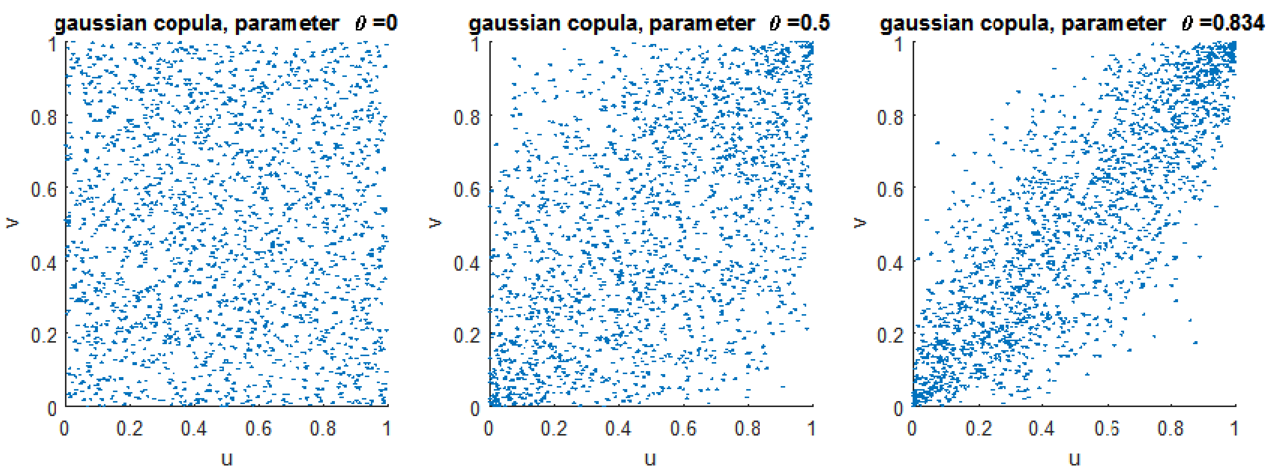


Figure 1.2: Simulation of the bivariate Gaussian copula for different parameter θ con N= 2000, uniform marginals

It is easy to show that this copula is *comprehensive*, meaning that it includes the upper and lower Frechet-Hoeffding bounds as well as the independence copula

$$\lim_{\theta \to -1^+} C_\theta^{Ga}(u,v) = W(u,v) \qquad C_0^{Ga}(u,v) = \Pi(u,v) \qquad \lim_{\theta \to 1} C_\theta^{Ga}(u,v) = M(u,v)$$

t copula It is derived from the Student's-t distribution and, as in the univariate case, the degree of freedom parameter ν represents departure from the Gaussian copula. In analogy with the Gaussian case, the t copula

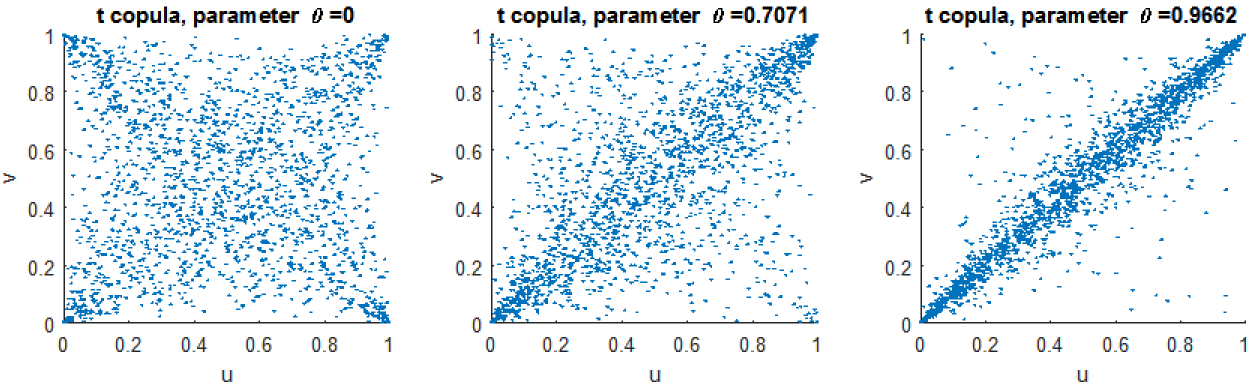


Figure 1.3: Simulation of bivariate t copula ($\nu = 1$) for different level of θ N= 2000, uniform marginals

is given by

$$C(u_1, u_2, ..., u_n) = \mathbf{t}_\nu^n(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2), ..., t_\nu^{-1}(u_n); \Sigma)$$

where ν is the degrees of freedom parameter. When ν is large the Student-t copula approaches the Gaussian copula.

Archimedean copulas Copulas in the family of Archimedean copulas are completely described in analytical form by the so called *generator*. This univariate continuous strictly decreasing function provides a way to define the relationship between the r.v.'s.
Archimedean structure implies that there exists $\lambda(t)$ such that

$$\lambda(H(x, y)) = \lambda(F(x))\lambda((G(x)), \qquad \lambda(t) > 0, \ \forall t \in [0, 1]$$

Setting $\varphi(t) = -\ln \lambda(t)$, the bivariate joint distribution can be expressed as the sum of the marginals $\varphi(H(x, y)) = \varphi(F(x)) + \varphi(G(y))$. By the probability integral transformation one gets

$$\varphi\left(C(u, v)\right) = \varphi(u) + \varphi(v)$$

The definition of generalised inverse of the generator is finally needed to recover the formal definition of Archimedean copulas.

Definition 1.2.24. Let φ an continuous strictly decreasing function, $\varphi : \mathbf{I} \to [0, \infty]$ such that $\varphi(1) = 0$. The pseudo-inverse of φ, expressed as $\varphi^{[-1]} : [0, \infty] \to \mathbf{I}$ is given by

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \le t \le \varphi(0) \\ 0, & \varphi(0) \le t \le \infty \end{cases} \tag{1.9}$$

Based on this definition, one assumes $\varphi^{[-1]}$ is continuous and non-increasing on $[0, \infty]$ and strictly decreasing on $[0, \varphi(0)]$

Noting that $\varphi^{[-1]}(\varphi(u)) = u \ \forall u \in \mathbf{I}$, therefore

$$\varphi\left(\varphi^{[-1]}(t)\right) = \begin{cases} t, & 0 \leq t \leq \varphi(0) \\ \varphi(0), & \varphi(0) \leq t \leq \infty \end{cases}$$
$$= \min(t, \varphi(0))$$

Lemma 1.2.25. *Let φ a continuous strictly decreasing function, $\varphi : \mathbf{I} \to [0, \infty]$ such that $\varphi(1) = 0$. the pseudo-inverse $\varphi^{[-1]}$ is given in (1.9). Let's define $C : \mathbf{I^2} \to \mathbf{I}$ such that*

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v))$$

Then C satisfies (1.1) and (1.2)

The *2-increasing* property (1.3), condition is guaranteed by the following theorem.

Theorem 1.2.26. *Let φ a continuous strictly decreasing function, $\varphi : \mathbf{I} \to [0, \infty]$ such that $\varphi(1) = 0$ and let $\varphi^{[-1]}$ its pseudo-inverse, as above. Let's define $C : \mathbf{I^2} \to \mathbf{I}$. C is a copula function if and only if φ is a convex function.*

Archimedean copulas have the following properties

- given a generator $\varphi \Rightarrow c\varphi$ is a generator $\forall c \in \mathbb{R}$. Furthermore, each convex combination of generators becomes a generator, that is $c\varphi + (1-c)\varphi \ \forall c \in \mathbf{I}$
- C is exchangeable, that is $C(u, v) = C(v, u) \ \forall u, v \in [0, 1]$
- C is associative, namely $\forall u, v, w \in [0, 1] \Rightarrow C(C(u, v), w) = C(u, C(v, w))$

A feature of Archimedean copulas is the analytic expression of the rank correlation measure. Kendall's τ can be computed as a function of the generator φ from

$$\rho_\tau = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} \, \mathrm{d}t$$

Remark 1.2.5.1. It is worth noting that in modern research Archimedean copula functions are often denoted with the notation

$$C(u, v) = \psi\left(\psi^{-1}(u) + \psi^{-1}(v)\right) \tag{1.10}$$

with $\psi(x) \equiv \varphi^{-1}(x)$. Care must be taken in recognizing the difference, when needed.

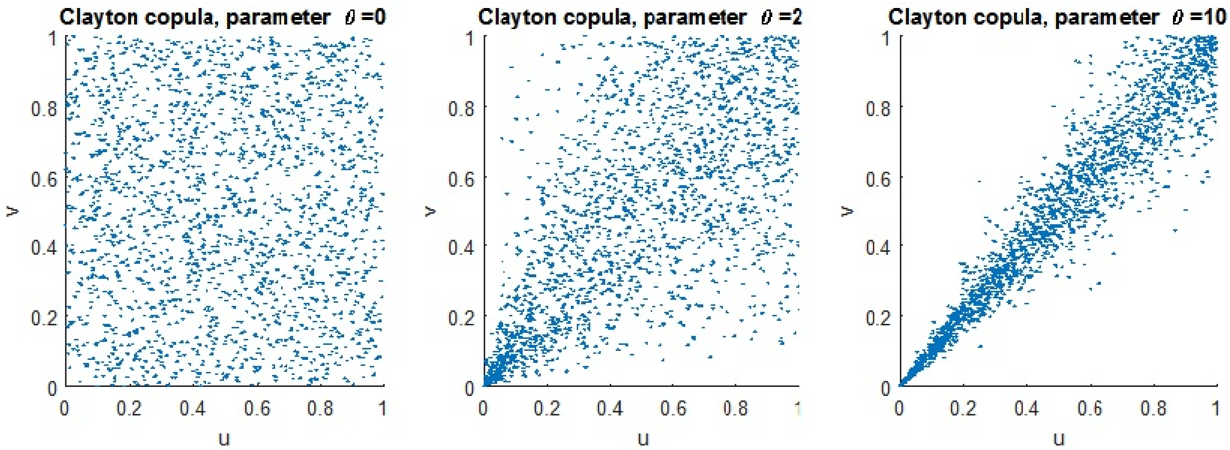


Figure 1.4: Simulation on bivariate Clayton copula for different parameters θ and N= 2000, marginals uniform

Clayton copula

The *Clayton* generator is $\varphi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$, where θ is the copula parameter on the domain $[-1, \infty]$. Its explicit representation is

$$C_\theta(u, v) = \max([u^{-\theta} + v^{-\theta} - 1]^{-\frac{1}{\theta}}, 0)$$

This copula is also *comprehensive*, so that it can reach the whole dependence structure.

$$C_{-1}(u, v) = W(u, v) \qquad \lim_{\theta \to 0} C_\theta(u, v) = \Pi(u, v) \qquad \lim_{\theta \to \infty} C_\theta(u, v) = M(u, v)$$

The Kendall's τ is $\rho_\tau = \theta/(\theta + 2)$.

Frank copula

Frank copula is defined by the generator $\varphi(t) = -\ln \frac{\exp^{-\theta t} - 1}{\exp^{-\theta} - 1}$, con $\theta \in \mathbb{R} \backslash \{0\}$. The corresponding bivariate copula is

$$C_\theta(u, v) = -\frac{1}{\theta} \ln \left(1 + \frac{(\exp^{-\theta u} - 1)(\exp^{-\theta v} - 1)}{\exp^{-\theta} - 1} \right)$$

Gumbel copula

Gumbel copula is defined by the generator $\varphi(t) = (-\ln t)^\theta$, con $\theta \geq 1$. The corresponding bivariate copula holds

$$C_\theta(u, v) = \exp^{-[(-\ln u)^{-\theta} + (-\ln v)^{-\theta}]^{\frac{1}{\theta}}}$$

Kendall's τ is $(\rho_\tau = 1 - 1/\theta)$.

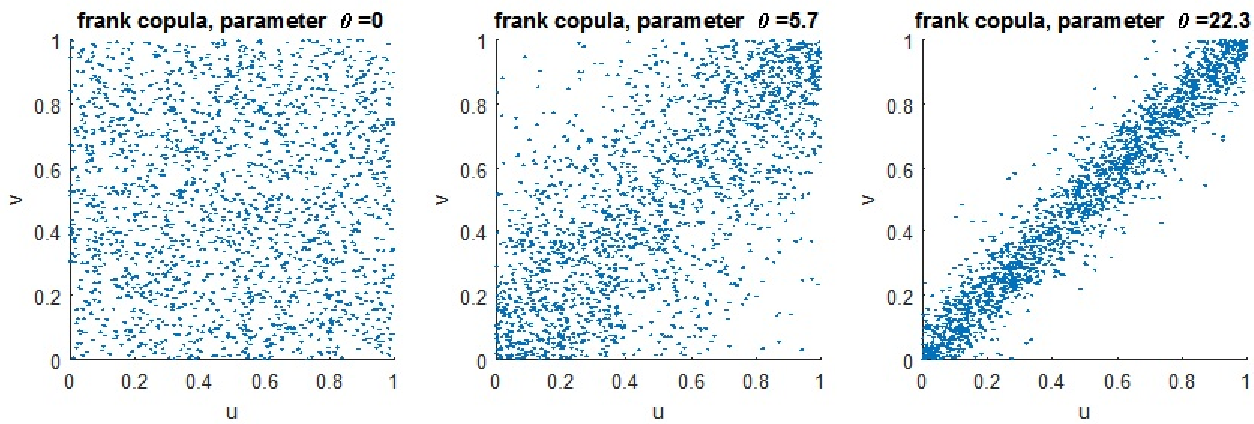


Figure 1.5: Simulation on bivariate Frank copula for different parameters θ and N= 2000, marginals uniform

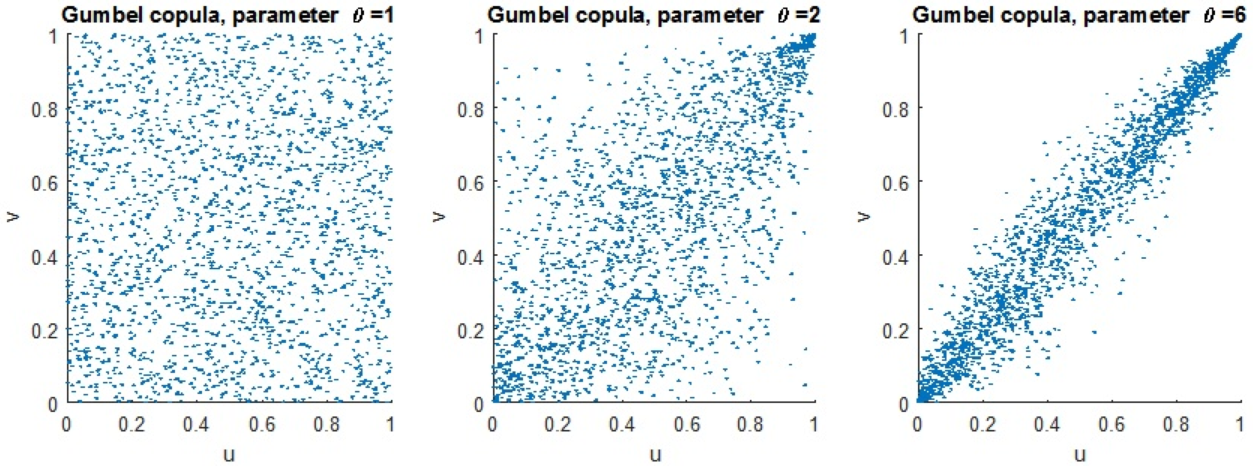


Figure 1.6: Simulation on bivariate Gumbel copula for different parameters θ and N= 2000, marginals uniform

Focusing on the range of dependence, one obtains

$$C_1(u,v) = \Pi(u,v) \qquad \lim_{\theta\to\infty} C_\theta(u,v) = M(u,v)$$

Therefore Gumbel is not comprehensive because does not attain the countermonotonic case.

1.2.6 Tail dependence measure

Tail dependence measures the strength of pairwise dependence in the extreme tails of bivariate distribution (X_1, X_2).

Definition 1.2.27. Let X e Y be two continuous r.v's with marginals F and G. Define λ_U the upper-tail dependence parameter as the limit, if it exists, as t approaches 1, of the conditional probability that Y assumes values greater than 100-t-th percentile of G, given that X is greater than 100-t-th percentile of F.

$$\lambda_U = \lim_{t\to 1^-} P\Big[Y > G^{(-1)}(t) \Big| X > F^{(-1)}(t)\Big] \tag{1.11}$$

For $0 < \lambda_U < 1$ one can state that there exists upper-tail dependence in the pair (X, Y). On the contrary $\lambda_U = 0$ means that r.v's are asymptotically independent on the upper-tail. In analogy, the lower-tail dependence coefficient λ_L can be represented as

$$\lambda_L = \lim_{t \to 0^+} P\Big[Y \leq G^{(-1)}(t) \Big| X \leq F^{(-1)}(t)\Big] \tag{1.12}$$

If F,G are the continuous distribution functions of the bivariate r.v's, then λ_U, λ_L can be expressed in terms of the unique copula function. The link comes from the definition of conditional probability in Theorem (1.2.17). From that, the result below shows how to express the upper and lower tail dependence coefficients in terms of the corresponding bivariate copula.

Theorem 1.2.28. *Let X, Y, F, G, λ_U, λ_L defined as in (1.2.27) and let C the bivariate copula of (X, Y). If the limits (1.11) and (1.12) exist then the coefficients of tail dependence can be formulated as*

$$\lambda_U = 2 - \lim_{t \to 1^-} \frac{1 - C(t,t)}{1-t}$$
$$\lambda_L = \lim_{t \to 0^+} \frac{C(t,t)}{t}$$

Theorem (1.2.28) implies the following result for Archimedean copulas.

Corollary 1.2.29. *Let C an Archimedean copula with generator $\varphi(t)$. Then*

$$\lambda_U = 2 - \lim_{t \to 1^-} \frac{1 - \varphi^{[-1]}(2\varphi(t))}{1-t} = \lim_{x \to 0^+} \frac{1 - \varphi^{[-1]}(2x)}{1 - \varphi^{[-1]}(x)}$$
$$\lambda_L = 2 - \lim_{t \to 0^+} \frac{\varphi^{[-1]}(2\varphi(t))}{t} = \lim_{x \to \infty} \frac{\varphi^{[-1]}(2x)}{\varphi^{[-1]}(x)}$$

Example 1.2.30 (Copulas and tail dependence).

- Gaussian copula $C_\rho(u_1, u_2) = \Phi_n(\Phi^{-1}(u_1), \Phi^{-1}(u_2), ..., \Phi^{-1}(u_n), \rho)$, $\lambda_U = \lambda_L = 0$, where Φ_n, Φ^{-1} are the n-variate and inverse standard Gaussian cdf respectively.

- t-Student's copula $C^\nu_\rho(u_1, u_2, ..., u_n) = t^n_\nu(t^{-1}_\nu(u_1), t^{-1}_\nu(u_2), ..., t^{-1}_\nu(u_n), \rho)$

  $$\lambda_U = \lambda_L = 2\, t_{\nu+1}\left(\sqrt{\nu+1}\,\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right)$$

  where t_ν, t^n_ν, t^{-1}_ν are the univariate, n-variate and inverse student's t-cdf with ν DoF respectively.

- Clayton copula $C_\theta(u, v) = [u^{-\theta} + v^{-\theta} - 1]^{-\frac{1}{\theta}}$
  $\lambda_L = \lim_{t \to \infty} \frac{\varphi^{-1}(2t)}{\varphi^{-1}(t)} = \lim_{t \to \infty} \frac{(1+2t\theta)^{-\frac{1}{\theta}}}{(1+t\theta)^{-\frac{1}{\theta}}} = 2^{-\frac{1}{\theta}}, \quad \lambda_U = 0$

- Frank copula $C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1+\frac{(\mathrm{e}^{-\theta u}-1)(\mathrm{e}^{-\theta v}-1)}{\mathrm{e}^{-\theta}-1}\right)$ like the Gaussian
  $\lambda_U = \lambda_L = 0$
- $C_\theta(u,v) = \mathrm{e}^{-[(-\ln u)^{-\theta}+(-\ln v)^{-\theta}]^{\frac{1}{\theta}}}$
  $\lambda_U = \lim_{t\to 0^+} \frac{1-\varphi^{-1}(2t)}{1-\varphi^{-1}(t)} = \lim_{t\to 0^+} \frac{1-\mathrm{e}^{-2t^{\frac{1}{\theta}}}}{1-\mathrm{e}^{-t^{\frac{1}{\theta}}}} = 2-2^{\frac{1}{\theta}}, \quad \lambda_L = 0$

1.3 Value-at-Risk and Risk Measures (VaR)

We introduce here Value-at-Risk, the main risk measure that has been used in the industry for the evaluation of market and credit risk and for the definition of capital requirements. We will also introduce the main object of discussion that has taken place among the industry, the academia and the regulators, on the failure of VaR to meet the requirements of so called *coherent measures*.

Definition 1.3.1. Given a probability space on a σ-algebra and a time horizon Δ, define $L^0(\Omega,\mathcal{F},P)$ the set of random variables on $(\Omega,\mathcal{F})$. Financial risks are defined as a subset $\mathcal{M}\subset L^0(\Omega,\mathcal{F},P)$ of r.v's. Moreover, it is assumed that $\mathcal{M}$ be a convex cone, such that $L_1\in\mathcal{M}$ and $L_2\in\mathcal{M}$ imply $(L_1+L_2)\in\mathcal{M}$ e $\lambda L_1\in\mathcal{M}$, $\forall\lambda>0$.

Definition 1.3.2. A risk measure ϱ is a real valued function $\varrho:\mathcal{M}\to\mathbb{R}$ on the r.v. cone $\mathcal{M}$.

Coherent Measures Artzner et al. (1999) proposed the following requirements to define a proper risk measure. Risk measures satisfying these axioms are defined *coherent*.

Axiom 1.3.3 (Monotonicity).
For any $L_1, L_2\in\mathcal{M}$ t.c. $L_1\le L_2$ we have, a.e.

$$\varrho(L_1)\le\varrho(L_2)$$

Axiom 1.3.4 (Positive homogeneity).
For any $L\in\mathcal{M}$ and any $\lambda>0$ we have

$$\varrho(\lambda L)=\lambda\varrho(L)$$

Axiom 1.3.5 (Translation invariance).
For any $L\in\mathcal{M}$ and any $l\in\mathbb{R}$ we have $\varrho(L+l)=\varrho(L)-l$.

Axiom 1.3.6 (Subadditivity).
For all $L_1, L_2\in\mathcal{M}$ si ha

$$\varrho(L_1+L_2)\le\varrho(L_1)+\varrho(L_2)$$

Based on the properties above we can state the following:

Definition 1.3.7 (Coherent Risk Measures).

A risk measure ϱ is defined coherent on the convex cone domain$\mathcal{M}$ if and only if the axioms (1.3.5 - 1.3.3) are satisfied.

Value at Risk *Value at Risk (VaR)* has been the measure massively used in financial risk management.

Definition 1.3.8 (Value at Risk). Let consider a confidence level $\alpha \in (0,1)$, The VaR of a portfolio L at level α is the smallest number l such that the portfolio loss probability exceeds the value l is not greater than $(1-\alpha)$. Formally, if the portfolio distribution is denoted as $F_L(l)$, one gets

$$VaR_\alpha = \inf\{l \in R : P(L > l) \leq \ 1-\alpha\} = \inf\{l \in R : F_L(l) \geq \ \alpha\} \tag{1.13}$$

In terms of probability theory VaR is nothing but a quantile of the probability distribution of losses L that is $\inf\{l \in R : \ l \geq F_L^{-1}(\alpha) \ \}$.

It is well known that the main criticism about VaR has to do with violation of the subadditivity axiom. The following theorem shows an interesting characterisation of VaR with respect to subadditivity.

Theorem 1.3.9 (VaR subadditivity for elliptical risk factors).

Let $\mathbf{X}$ *be a d-dimension vector of r.v.'s,* $\mathbf{X} \backsim \mathcal{N}_d(\mu, \Sigma, \rho)$*, and define a set* $\mathcal{M}$ *of losses as follows*

$$\mathcal{M} = \left\{L : L = \lambda_0 + \sum_{i=1}^{d} \lambda_i X_i \ , \ \ \lambda_i \in \mathbb{R}\right\}$$

Then, for any pair of distributions $L_1, L_2 \in \mathcal{M}$ *and for* $0.5 \leq \alpha < 1$

$$VaR_\alpha(L_1 + L_2) \leq VaR_\alpha(L_1) + VaR_\alpha(L_2)$$

Another important result of VaR is *comonotonic additivity.*

Proposition 1.3.10. *Let* $0 < \alpha < 1$ *and let* L_1, L_2 *be comonotonic r.v.'s. with distributions* F_1, F_2 *absolutely continuous. Then,*

$$VaR_\alpha(L_1 + L_2) = VaR_\alpha(L_1) + VaR_\alpha(L_2)$$

1.4 Stochastic Time Models

In the 1960s and 70s a line of research developed, according to which various kinds of stochastic processes can be generated by an arithmetic Brownian motion and a sequence of stopping times. The seminal result for this literature is known as *Skorokhod embedding theorem.*

Theorem 1.4.1. *Skorokhod and Slobodenyuk (1965). Suppose that* $X_1, X_2, \ldots, X_n$ *are independently and identically distributed random variables with mean zero and finite variance, and put* $S_n = X_1 + \ldots + X_n$. *There is a non decreasing sequence of stopping times* $\tau_1, \tau_2 \ldots$ *such that the* W_{τ_n} *of a Wiener process have the same joint distributions as the* S_n *and* $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \ldots$ *are independent and identically distributed random variables with* $E(\tau_n - \tau_{n-1}) = E(X_1^2)$ *and* $E[(\tau_n - \tau_{n-1})^2] \leq 4E(X_1^4)$.

About at the same time when this result was published, Dubins and Schwarz (1965) proved that the same result holds for martingale processes.

Theorem 1.4.2. *Dubins and Schwarz (1965). Let* X *be a continuous stochastic processes, almost all of whose paths are nowhere constant. There exists a mapping* Π *such that the composition* $\Pi(X_t)$ *is a Brownian motion.*

Monroe (1978) proved the extension to local semimartingales. We collect here below his two theorems.

Theorem 1.4.3. *Monroe (1978)*

1. *Let* $(X_t, \mathcal{G}_t)$ *be a local semimartingale and let* X_{T_s} *a time change of* X_t. *Let* $\mathcal{F}_s$ *be the family of* σ*-fields generated by* $X_{T_v}, v \leq s$. *Then,* $(X_t, \mathcal{F}_s)$ *is a local semimartingale.*
2. *The local semimartingale* X_s *is equivalent to a time change of a Brownian motion*

1.5 Outline of the main results

We summarize here the problems addressed and the main results obtained in the thesis.

Value-at-Risk on α-stable risk: limits to diversification In Chapter 2 we apply the tool of C-convolution to the problem of Value-at-Risk aggregation for risks with heavy fat tails. As a representative example of heavy tail distributions we address α-stable risks. By C-convolution, we confirm the results on VaR super-additivity that have been obtained by Monte Carlo and we move forward to far percentiles of the distribution. We show that VaR fails to be sub-additive for α-stable distributions with levels of the parameter α that are around or lower than 1. These are the cases for which the tails are so heavy that not even the first moment, the mean, is finite. In these cases aggregating two i.i.d risks in an equally weighted portfolio will always produce super-additivity: the capital required to provide insurance to the two risks together is always higher than the sum of the capital for the marginal risks. This is true for all the dependence structures, that is copula functions, applied in the analysis (elliptical and Archimedean), and across most of the dependence levels.
We also document a new unexpected result. When VaR is super-additive, the increase in capital due to aggregation is lower for copula functions with higher tail dependence. This result appears counter-intuitive since one would expect tail dependence to be a factor that should increase the amount of risk, and the amount of capital to absorb it. Moreover, if we push the analysis further in the tail, that is for very low percentiles, it

turns out that at some point in the tail the VaR becomes additive above some level of dependence, but this critical dependence level is lower for copula function with lower tail dependence.

Value-at-Risk and the tail dependence puzzle In Chapter 3 we extend and focus the analysis on the tail dependence puzzle that was retrieved in Chapter 2. Differently from what had been documented in that chapter, we show that the tail dependence puzzle also shows up in cases in which the Value-at-Risk is sub-additive. So, also in cases with α-stable marginal risks with $\alpha > 1$ and Student-t with similar tail behaviour, for which the aggregated VaR measure is subadditive, the tail dependence puzzle persists. Student-t copula functions with higher tail dependence showcase a reduction of the capital requirement computed using VaR with respect to copula functions, like the Gaussian, for which risks are asymptotically independent in the tails. This is clearly a paradox and a puzzle because it would be natural to expect that higher conditional probability of extreme events, that is higher probability that two extreme risks materialize together, should always result into higher capital requirements. We document that this is not the case. We do observe that the puzzle disappears when marginal risks are gaussian. So, the tail dependence puzzle is crucially determined by the tails of the marginal risk. While the aggregation of thin tailed risk costs more capital if risks are tail dependent, tail dependence allows to save capital if marginal risks are sufficiently heavy tailed.

Two main questions are worth being addressed: the first is which degree of marginal fat tails brings about the puzzle; the second is under which conditions the puzzle showing up at percentile z persists for all tail levels lower than z. The latter question is particularly relevant because it provides a sufficient condition for the paradox to be extended to the Expected Shorfall risk measure as well, that is actually the integral of VaR measures on the tail segment below z. These questions are not generally solvable entirely analytically, and in Chapter 3 we collect first analytical and simulation results, trying to give first tentative answers to these issues.

Time change, generalised compounding and the term structure In Chapter 4 we address a more classical topic in copula function theory, that is the generalisation and distortion of the product function. It is well known that if the factors of the product are uniform random variables in the unit interval, this procedure generates the family of Archimedean copulas. Here we exploit the same idea for a different application. The product is our reference for geometric compounding or discounting used in the valuation of assets. Here we propose an Archimedean distortion of this geometric compounding principle. The economic rationale for the distortion is provided by the concept of *generalised compounding* proposed by Carr and Cherubini (2020) and based on time change models. The idea is that if the returns are compounded or discounted in discrete time according to a stochastic clock, this brings about a distortion that can be represented in terms of Laplace transforms. Then, the same technical tool proposed by Marshall and Olkin (1988) for copula function, we design a family of Archimedean compounding/discounting functions.

From a technical point of view, compounding functions are isomorphic to copula functions. They are more general in the sense that the marginal variables are not necessarily uniform, and in general they are not. In compounding functions the marginals represent one period return accrual or discount. In particular, we give conditions for a class of compounding/discounting functions to have lognormal margins. We show that this specification can be usefully applied to the analysis of the risk free term structure in a model based on the Stochastic Discount Factor.

Chapter 2

Value-at-Risk on α-stable Risks: The Limits of Diversification

Abstract

We apply the concept of convolution copulas to the problem of risk aggregation of α-stable heavy tailed dependent risks. While we confirm results of *Value-at-Risk* (VaR) super-additivity for cases with α below 1 we focus on the role of tail dependence. We provide evidence of a new puzzle: α-stable risks, the VaR at the common confidence level (i.e. 1%) is lower for copulas with higher tail dependence, except for the Gaussian case. The result is confirmed as we move towards extreme points in the tail. We also show that at some point in the tail the aggregated VaR becomes additive above some level of dependence, but this critical dependence level is lower for copulas with lower tail dependence.

2.1 Introduction

Starting with the seminal paper by Artzner et al. (1999), a long debate has developed concerning risk measures and diversification, with particular focus on the fact that *Value-at-Risk* (VaR) may fail to represent risk diversification. Following this long debate, the trend of regulation has drifted in favour of alternative risk measures ensuring sub-additivity. So, in market risk regulation a measure resembling *Expected Shortfall* (ES) is bound to replace VaR in the new FRTB (*Fundamental Review of the Trading Book*). The trend has been much less clear-cut in the insurance regulation, where the new Solvency II framework still relies on the VaR measure.

If one considers the nature of the *ES* measure and the tail behaviour of some actuarial risks, it is quite straightforward to explain the different trend in insurance. In fact, since *ES* is the expected loss in the tail, this measure can only exist if the first moment of the loss distribution is well defined in the first place. Unfortunately, existence of first moment is very often challenged in the world of actuarial risks. This is particularly true for catastrophic risks, for which the fat tail phenomenon is so extreme to destroy the integrability requirement for existence of the first moment. For these kinds of risk, a plausible shape that has been suggested for the distribution of losses is the α-stable family. In particular, Ibragimov et al. (2008) [IJW] show that seismic risks and other risk events typical of the catastrophe insurance market can be described by stable when $\alpha < 1$, so that the first moment is not defined. It is well known that stable distribution is a Pareto distribution–like (power law) and vary regularly at infinity. The same result is documented by Chavez-Demoulin et al. (2006) for a number of operational risks.

Concerning diversification, the IJW paper also shows that diversification dramatically fails, even in the case of independent risks, that is where diversification should be highest. Beyond the discussion on the property of the risk measures, this finding points to a real relevant problem of the catastrophe insurance market, so that putting catastrophic risks together dramatically increases the ruin probability of the reinsurance policy of such risks, even if such risks are independent. This adds an important issue in the debate about the possibility of a private reinsurance market for catastrophic risks versus insurance schemes based on public funds. This also raises the question of how the results extend to the case in which catastrophic risks are dependent. This extension was provided by Ibragimov and Prokhorov (2016) using Monte Carlo analysis applied to copula functions. Here we investigate the problem with C-convolution copulas, proposed by Cherubini et al. (2011) with the purpose to extend the analysis further in the tail and to explore the relevance of different tail dependence indexes.

The contribution of this paper to this stream of research is the twofold. First, we want to explore the shape of the relationship between a dependence measure of the risks and a sub-additivity measure of the VaR. Of course, since it is well known that perfectly dependent risks are additive, the IJW result implies that this relationship must be non monotone for heavy tailed α-stable risks. Second, we want to explore whether this diversification failure is made more severe by the tail-dependence feature of the risks.

The plan of the paper is as follows. In Section 2.2 we define the convolution α-stable risks under general dependence structures represented by copulas. In Section 2.3 we present our results analysis addressing the two issues described above. Section 2.4 concludes.

2.2 Preliminaries: Convolution of α-stable Risks

In this section, we review the main concepts that will be used in our numerical analysis. First, we describe a pair of α-stable i.i.d. risks and we review the results on VaR aggregation (corresponding to the standard concept of convolution). Second, we assume a general dependence structure represented by a copula function with different degrees of tail-dependence. Third, we review the concept of convolution of dependent risks, called C-convolution.

2.2.1 Stable Distributions

Stable distributions are part of a wide class of probability distributions often used to represent heavy tailed models, particularly when heavy tails are so extreme that even first and second moments may fail to be well-defined. From a mathematical point of view, stable distribution is uniquely characterized by the property of closure under sum of independent variables.

In insurance applications, this class of distributions is sometimes applied to catastrophic risks, including earthquakes, floods, wind damage, etc... In these cases stable distributions are used to create loss distribution functions to model rare events with huge losses. Here we recall definitions and basic useful properties of univariate stable distributions, referring the reader to Samorodnitsky and Taqqu (1997) and Nolan (2012) for a detailed treatment.

Definition 2.2.1 (Stable)**.** A random variable X is stable if for $X_1, X_2, .., X_n$ independent copies of X and $S = \sum_{i=1}^{n} X_i$, there exist a, b $\in \mathbb{R}$ such that $S = aX + b$ holds. The random variable is called *strictly* stable if $b = 0$.

The stable family is denoted by $\mathcal{S}(\alpha, \beta, \gamma, \delta)$. The four parameters α, β, γ, and δ determine the density function. These parameters can be interpreted as follows:

(α) is the basic stability parameter, determining the weight in the tails: the lower the α value, the greater the frequency and size of extreme events. Its range is $(0 < \alpha \leq 2)$

(β) It is the skewness distribution parameter and $(-1 \leq \beta \leq 1)$, with a zero value indicating that the distribution is symmetric. A negative / positive β implies that the distribution is skewed to the left or right respectively

(γ) The parameter γ is positive and represents a measure of dispersion. It determines the width of the density. The higher the γ the higher the dispersion around the δ parameter.

(δ) The parameter δ determines the location of the distribution. It generalizes the concept of mean of the distribution

Closed forms for α-stable distributions generally do not exist, except for specific cases of the parameters pair (α, β). As showed below, when $\alpha = 2$ the Stable distribution is a normal distribution (standard normal has variance $\sigma^2 = 2\gamma^2$). When $\alpha = 1$ and $\beta = 0$ the distribution becomes a Cauchy. Another case concerning Lévy ($\alpha = 0.5$ and $\beta = 1$), which is a totally asymmetric distribution. For these analytic distributions, (γ, δ) parameters are set to $(1, 0)$ respectively to achieve standard case.

Lévy: $\mathcal{S}(0.5, 1, \gamma, \delta)$ $\qquad f(x) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(x-\delta)^{3/2}} \exp\left(-\frac{\gamma}{2(x-\delta)}\right), \qquad \delta \le x \le \infty$

Cauchy: $\mathcal{S}(1, 0, \gamma, \delta)$ $\qquad f(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x-\delta)^2}, \qquad -\infty \le x \le \infty$

Gaussian: $\mathcal{S}(2, 0, \gamma = \frac{\sigma}{\sqrt{2}}, \delta)$ $\qquad f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\delta)^2}{2\sigma^2}\right), \qquad -\infty \le x \le \infty$

The most direct way to describe all possible stable distributions is through the characteristic function.

Definition 2.2.2 (Stable random variable)**.** A r.v. X is stable if and only if it is expressed by the following characteristic function $\varphi_X(t)$

$$\varphi_X(t) = \mathbb{E}\left[e^{\imath tX}\right] = \begin{cases} \exp\left\{\imath\delta t - \gamma^\alpha |t|^\alpha (1 - \imath\beta sign(t) \tan(\pi\alpha/2))\right\}, \ \alpha \neq 1 \\ \exp\left\{\imath\delta t - \gamma|t|(1 + \imath\beta sign(t)\frac{2}{\pi} \ln|t|)\right\}, \ \alpha = 1 \end{cases} \tag{2.1}$$

Proposition 2.2.3. *The Family of stable distribution $\mathcal{S}(\alpha, \beta, \gamma, \delta)$ is described by the following properties*

a) Let $X \in \mathcal{S}(\alpha, \beta, \gamma, \delta)$, $\forall a \neq 0, b \in \mathbb{R}$

$$aX + b \backsim \begin{cases} \mathcal{S}(\alpha, sign(a)\beta, |a|\gamma, a\delta + b), & \alpha \neq 1 \\ \mathcal{S}(\alpha, sign(a)\beta, |a|\gamma, a\delta + b - \frac{2}{\pi}\beta\gamma a \log|a|), & \alpha = 1 \end{cases} \tag{2.2}$$

b) Let $X \in \mathcal{S}(\alpha, \beta_1, \gamma_1, \delta_1)$ and $Y \in \mathcal{S}(\alpha, \beta_2, \gamma_2, \delta_2)$ independent r.v.'s. Then, the sum $X + Y \backsim \mathcal{S}(\alpha, \beta, \gamma, \delta)$ where

$$\beta = \frac{\beta_1\gamma_1^\alpha + \beta_2\gamma_2^\alpha}{\gamma_1^\alpha + \gamma_2^\alpha}, \quad \gamma = (\gamma_1^\alpha + \gamma_2^\alpha)^{1/\alpha}, \quad \delta = \delta_1 + \delta_2 \tag{2.3}$$

c) For all values in $0 < \alpha < 2$, it follows the reflection property of X

$$X \backsim \mathcal{S}(\alpha, \beta, \gamma, 0) \Leftrightarrow -X \backsim \mathcal{S}(\alpha, -\beta, \gamma, 0)$$

d) An important result involves p-order moments of a Stable r.v. Let $X \backsim \mathcal{S}(\alpha, \beta, \gamma, 0), 0 < \alpha < 2$. *It holds*

$$\mathbb{E}|X|^p < \infty, \qquad \forall\, 0 < p < \alpha$$
$$\mathbb{E}|X|^p = \infty, \qquad \forall\, p \geq \alpha$$

It follows that the stable family does not admit finite second moment for $\alpha < 2$ *and so variance does not exist, and if* $\alpha \leq 1$ *the first moment does not exist either.*

Linear combination

When $X_i \sim S(\alpha, \beta_i, \gamma_i, \delta_i)$, are independent stable r.v.'s then for all arbitrary $c_i \in \mathbb{R}, i = 1, ..., n$

$$\sum_{i=1}^{n} c_i X_i \backsim S(\alpha, \beta, \gamma, \delta) \qquad \text{where}$$
$$\beta = \frac{\sum_{i=1}^{n} \beta_i \left(sign(c_i)\right) |c_i \gamma_i|^\alpha}{\gamma^\alpha}$$
$$\gamma^\alpha = \sum_{i=1}^{n} |c_i \gamma_i|^\alpha$$
$$\delta = \begin{cases} \sum_{i=1}^{n} c_i \delta_i, & \alpha \neq 1 \\ \sum_{i=1}^{n} c_i \delta_i - \dfrac{2}{\pi} \sum_{i=1}^{n} \beta_i\, c_i\, \gamma_i\, ln(c_i), & \alpha = 1 \end{cases}$$

This results can be obtained by induction on the property (2.3) applying the γ computed in the property (2.2.3).

A special case is represented if $\beta_i = 0, \gamma_i = 1, \delta_i = 0\ \forall\ i = 1, ..., n$. In that case for all $c_i \geq 0$ such that $\sum_{i=1}^{n} c_i \neq 0$, one obtains

$$\sum_{i=1}^{n} c_i X_i \Bigg/ \left(\sum_{i=1}^{n} c_i^\alpha\right)^{1/\alpha} \stackrel{d}{=} X_1 \backsim S(\alpha, 0, 1, 0) \tag{2.4}$$

This means that a linear combination of independent standard stable r.v.'s behaves like the same standard stable r.v. X_1 times a function of either the linear combination of coefficient c_i or the parameter α. Samorodnitsky and Taqqu (1997) and Nolan (2012) further describe the Stable properties.

Tail behaviour. If $\alpha < 2$ and $-1 < \beta \leq 1$, then the density $f(x)$ and cumulative distribution function $F(x)$ have an asymptotic power law: as $x \to \infty$

$$1 - F(x) = P(X > x) \backsim \gamma^\alpha c_\alpha (1 + \beta) x^{-\alpha}$$
$$f(x|\alpha, \beta, \gamma, \delta) \backsim \alpha \gamma^\alpha c_\alpha (1 + \beta) x^{-(\alpha+1)}$$

where $c_\alpha = \frac{1}{\pi} \sin \frac{\pi\alpha}{2} \Gamma(\alpha)$. Using the reflection property, the lower tail properties are similar. When $\alpha < 2$, the variance is infinite and the tails are asymptotically equivalent to a Pareto law, i.e. they exhibit a power-law behaviour.

2.2.2 Convolution of independent variables

In the finance literature there are many examples of stable variables applications, particularly at the crossroad of insurance. As for speculative price dynamics, heavy-tailed distributions were first modelled with α-stable processes by Mandelbroot and Fama in the 1960s (Fama (1965)). More recently, Chavez-Demoulin et al. (2006) discuss that tail indices less than one are observed for empirical loss distributions of a number of operational risks.

Since the parameter γ generalizes the concept of variance, it is natural to start from this as a risk measure. This gives the first evidence that diversification may increase the risk of the portfolio when the distribution of the convolution shows heavy-tails. To see this, let us take into account the cases of stable distributions listed above, looking at the parameter γ for a portfolio. For the Lévy distribution, the p.d.f. with location parameter δ and dispersion γ is

$$f(x) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(x-\delta)^{3/2}} \exp\Big(-\frac{\gamma}{2(x-\delta)}\Big), \qquad \delta \leq x \leq \infty$$

When the standard case is considered, i.e. $\gamma = 1, \delta = 0$, then the homogeneous portfolio of k standard Levy r.v.'s $X_1, ..., X_k$ each of weight $\dfrac{1}{k}$ holds $S = \sum_{i=1}^{k} \frac{1}{k} X_i$. Then, applying the linear portfolio property in (2.4), one obtains

$$S = \sum_{i=1}^{k} \frac{1}{k} X_i =^d X_1 \left(\sum_{i=1}^{k} \frac{1}{\sqrt{k}} \right)^2 = k \ X_1 \backsim \mathcal{S}(\alpha, 0, k, 0)$$

$$\text{For general scale } \gamma's: \quad \gamma_S^\alpha = \sum_{i=1}^{k} \Big(\frac{\gamma}{k}\Big)^\alpha = \sum_{i=1}^{k} \Big(\frac{\gamma}{k}\Big)^{\frac{1}{2}} = k \Big(\frac{\gamma}{k}\Big)^{\frac{1}{2}} \implies \gamma_S = k\gamma$$

We conclude that the scaling factor is growing k times, proportional to the number of portfolio asset. To put it in different words, the riskiness of the equally-weighted portfolio is k times of the risk of a single asset.

In the Cauchy case, with $\alpha = 1$, the p.d.f. is

$$f(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x-\delta)^2}, \qquad -\infty \leq x \leq \infty$$

Now the property (2.4) allows to show that

$$S = \sum_{i=1}^{k} \frac{1}{k} X_i =^d X_1 \left(\sum_{i=1}^{k} \frac{1}{k} \right) = X_1 \backsim \mathcal{S}(\alpha, 0, 1, 0)$$

$$\text{For general scale } \gamma' s: \quad \gamma_S^\alpha = \sum_{i=1}^{k} \Big(\frac{\gamma}{k}\Big)^\alpha = \sum_{i=1}^{k} \Big(\frac{\gamma}{k}\Big) = k\Big(\frac{\gamma}{k}\Big) \implies \gamma_S = \gamma$$

In this case portfolio riskiness neither increases nor reduces its magnitude since the γ parameter remains the same of the marginal Cauchy variables.

Diversification reduces the risk with gaussian marginal distributions, that is $\alpha = 2$. The density is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\Big(-\frac{(x-\delta)^2}{2\sigma^2}\Big), \quad -\infty \leq x \leq \infty \quad \text{where} \quad \sigma = \frac{\gamma}{2}$$

and it follows that

$$S = \sum_{i=1}^{k} \frac{1}{k} X_i =^d X_1 \left(\sum_{i=1}^{k} \frac{1}{k^2} \right)^{\frac{1}{2}} = \frac{1}{\sqrt{k}} X_1 \sim \mathcal{S}(\alpha, 0, \tfrac{1}{\sqrt{k}}, 0)$$

$$\text{For general scale } \gamma' s: \quad \gamma_S^\alpha = \sum_{i=1}^{k} \Big(\frac{\gamma}{k}\Big)^\alpha = \sum_{i=1}^{k} \Big(\frac{\gamma}{k}\Big)^2 = k\Big(\frac{\gamma}{k}\Big)^2 \implies \gamma_S = \frac{\gamma^2}{k}$$

The latter is the case of diversification reduction because the γ parameter is scaled by an order factor k.

Ibragimov (2009) provides a rigorous analysis of the study of portfolio diversification with heavy tails in the *Value-at-Risk* (VaR) framework. He shows that closure under convolution of independent stable r.v.'s and the positive homogeneity property of VaR allow to describe the relationship between portfolio weights with Schur-convex/concave functions.

A vector a is said to be majorized by $a \in \mathbb{R}^\ltimes$ vector $b \in \mathbb{R}^\ltimes$; written $a \prec b$ if

$$\sum_{i=1}^{k} a[i] \leq \sum_{i=1}^{k} b[i], \quad k = 1, ..., n-1, \quad \sum_{i=1}^{n} a[i] = \sum_{i=1}^{n} b[i]$$

where $a[1], ..., a[n]$ and $b[1], ..., b[n]$ denote the components of a and b in decreasing order.

The relation $a \prec b$ implies that the components of the vector **a** are less diverse than those of **b**. See Marshall et al. (1979) for a full reference on the majorization theory.

A function $\phi : \mathbf{R}_+^n \to \mathbf{R}$ is called *Schur-convex* (resp. *Schur-concave*) if $a \prec b \Rightarrow (\phi(a) \leq \phi(b))$ (resp. $a \prec b \Rightarrow \phi(a) \geq \phi(b)$). If a is not a permutation of b the above definition holds *strictly*, that is $(a_1, a_2, ...a_n) \notin (b_{\pi(1)}, b_{\pi(2)}, ..., b_{\pi(n)})$, for all permutations π set $\{1, 2, ..., n\}$.

A typical example of Schur-convex function is given by the class $\phi_\alpha(w_1, ..., w_n) = \sum_{i=1}^n w_i^\alpha, \alpha \geq 1$. Conversely if $\alpha \leq 1$ the class is Schur-concave.

Given weights $w = (w_1, w_2, ..., w_n)$, denote the return portfolio Z_w of risks $X_1, X_2, ..., X_n$ with weights belonging to the set $\mathrm{I}_n : \{w = (w_1, w_2, ..., w_n) \in \mathbf{R}_+^n,\ \sum_{i=1}^n w_i = 1,\ w_i \geq 0,\ \forall i = 1, ..., n\}$.

Denoting $\underline{w} = \big(1/n, ..., 1/n, 1/n\big)$ and $\overline{w} = (1, ..., 0, 0) \Longrightarrow \underline{w},\ \overline{w} \in \mathrm{I}_n$. it is natural to think of $\underline{w}$ as being more diversified that $\overline{w}$ because the latter consists of one risk only and it is the least diversified portfolio in the class of all portfolios $w \in \mathrm{I}_n$. Then, let's consider $VaR_q(\underline{w})$ and $VaR_q(\overline{w})$ as the value-at-risk of the equal weights and the undiversified portfolios. Then, majorization theory can be used to assess the effects of diversification policies in the framework of i.i.d. stable risks.

In this framework, Ibragimov (2009) defines the class of convolution-stable log-concave distribution $\mathcal{CSLC}$ of risks and proves

Theorem (4.1 in Ibragimov (2009))**.** Let $q \in (0, 1/2)$ and let $X_i, i = 1, ..., n$ be i.i.d. risk such that $X_i \sim \overline{\mathcal{CSLC}}, i = 1, ..., n$ Then

- $VaR_q(Z_v) < VaR_q(Z_w)$ if $v \prec w$ and v is not a permutation of w, i.e. $\psi(w, q) = VaR_q(Z_w)$ is strictly Schur-Convex in $w \in \mathbf{R}_+^n$
- Further, $VaR_q(Z_{\underline{w}}) < VaR_q(Z_w) < VaR_q(Z_{\overline{w}})$ for all $q \in (0, 1/2)$ and all weights $w \in \mathcal{I}_n$

While it may be shown that the class of risks for which the Theorem holds are characterized by $\alpha > 1$, it is easy to show that VaR diversification fails in the Lévy distribution. Here the stable parameters are $\alpha = 1/2, \beta = 1$. Let us $X_1, X_2, ..., X_n \backsim \mathcal{S}_{1/2}(1, \sigma, 0)$ with density $f(x) = (\sigma/(2\pi))^{1/2} \exp(-\sigma/(2\pi))x^{-3/2}$. Considering equal weights $w_i = 1/n$, from the linear convolution property (2.4) one obtains $Z_{\underline{w}} = (1/n) \sum_{i=1}^n X_i =^d nX_1$. By positively homogeneity: $VaR_q(Z_{\underline{w}}) = nVaR_q(X_1) = nVaR_q(Z_{\overline{w}}) > VaR_q(Z_{\overline{w}})$. It follows that the VaR for the most diversified portfolio $\underline{w}$ is higher than that $\overline{w}$ the least diversified one.

Theorem (4.2 in Ibragimov (2009))**.** Let $q \in (0, 1/2)$ and let $X_i, i = 1, ..., n$ be i.i.d. convolution stable risks. Then,

- $VaR_q(Z_v) > VaR_q(Z_w)$ if $v \prec w$ and v is not a permutation of w, i.e. $\psi(w, q) = VaR_q(Z_w)$ is strictly Schur-Concave in $w \in \mathbf{R}_+^n$
- Further, $VaR_q(Z_{\underline{w}}) > VaR_q(Z_w) > VaR_q(Z_{\overline{w}})$ for all $q \in (0, 1/2)$ and all weights $w \in \mathcal{I}_n$

Theorems 4.1 and 4.2 give an accurate bound to the given portfolio value at risk with the borderline case $\alpha = 1$ which corresponds to i.i.d. risks $X_1, ..., X_n$ with a symmetric Cauchy distribution.

Starting from the results above in the independence framework, we now deal with the more realistic case of dependence among risks, represented in full generality by a copula function.

2.2.3 Copula Functions and Dependence

We here recall the basic theory and useful results on copula functions and the non parametric dependence structure of random variables. A complete reference on copula functions is Nelsen (2006), while Frees and Valdez (1998) describes copula simulation methods. A reference on copula methods applied in finance is Cherubini et al. (2004).

Definition 2.2.4 (Copula). A n-dimensional Copula $\mathcal{C}$ is a function $\mathcal{C} : U_1 \times U_2 \cdots \times U_n \to \mathbb{I}, U_i \in \mathbb{I}$ the unit space, $\mathbb{I} \in (0,1)$, that satisfies these properties

1. $\forall i = 1...n,\ \mathcal{C}(1,1,...,U_i,...,1,1) = U_i$

2. $\forall i = 1...n,\ \mathcal{C}(U_1,U_2,...,U_i = 0,...,U_{n-1},U_n) = 0$

3. $\forall\ (u_1,...,u_n),(v_1,...,v_n)\ \in [0,1]^n$ where $u_i \leq v_i$, holds

$$\sum_{i_1=1}^{2} \cdots \sum_{i_n=1}^{2} (-1)^{i_1,...,i_n} \mathcal{C}(s_{1,i_1},...,s_{n,i_n}) \geq 0 \qquad \text{where } s_{j,1} = u_j \text{ and } s_{j,2} = v_j$$

Copula functions allow to separate the marginal behaviour of individual random variables from their dependence structure. Due to the integral probability transform, copula functions single out the dependence structure by reducing the representation of dependence to a joint probability function taking uniform random variables as arguments.

Theorem 2.2.5 (Sklar). *Let $F_1, F_2, ..., F_n$ be a set of n continuous univariate distribution functions of the random variables $(X_1, X_2, \ldots, X_n)$. Then, $H(X_1, X_2, \ldots, X_n)$ is a joint distribution function if and only if* $\forall\ (x_1,x_2,...,x_n) \in \mathbb{R}^n$

$$H(x_1,x_2,...,x_n) = \mathcal{C}\left(F_1(x_1),F_2(x_2),...,F_n(x_n)\right) \tag{2.5}$$

where $\mathcal{C}$ is a copula function.

It may be proved that copula functions are naturally linked to the concepts and measures of non parametric association. The maximal copula is defined as $\mathcal{M}(u_1,u_2,...,u_n) = \min\{u_1,u_2,...,u_n\}$, and is known as upper Fréchet bound. This is related to the concept of perfectly positive dependence, or *co-monotonicity*, that is r.v.'s $X_1,...,X_n$ such that exists strictly increasing functions linking each one to the other. As for perfect negative dependence, the Fréchet bound is meaningful only in the bivariate dimension, and reads $\mathcal{W}(u_1,u_2) = \max\left\{u_1+u_2-1,0\right\}$. Finally, it is well known that the independence case is embedded with the *Product Copula* or independence copula: $\prod(u_1,u_2,..u_n) = \prod_{i=1}^{n} u_i$

Copula functions are naturally linked to non parametric association measures such as Spearman's ρ and Kendall's τ.

Proposition 2.2.6. *Suppose (X, Y) have continuous marginal distributions and unique copula $\mathcal{C}$. Then the Kendall's τ are given by*

$$\rho_\tau(X,Y) = 4\int_0^1\int_0^1 \mathcal{C}(u,v)\,d\mathcal{C}(u,v) - 1$$

and rank correlation measure Spearman's ρ is defined as

$$\rho_S(X,Y) = 12\int_0^1\int_0^1 \mathcal{C}(u,v)\,d\mathcal{C}(u,v) - 3$$

Copula functions are also useful to represent the dependence structure of extreme events, that is measured by the so called *tail dependence* coefficients. These are linked to the concept of conditional distribution, and how this can be written in terms of copula functions. We have the following definition.

Definition 2.2.7. Let X,Y be two r.v.'s with corresponding continuous d.f.'s F, G. We define the upper tail dependence parameter λ_U as the limit,

$$\lambda_U = \lim_{t\to 1^-} \mathbb{P}\left[Y > G^{(-1)}(t) \middle| X > F^{(-1)}(t)\right] \tag{2.6}$$

Accordingly, the lower tail dependence coefficient λ_L is defined as

$$\lambda_L = \lim_{t\to 0^+} \mathbb{P}\left[Y \le G^{(-1)}(t) \middle| X \le F^{(-1)}(t)\right] \tag{2.7}$$

Next theorem shows how tail dependence can be expressed in terms of copulas.

Theorem 2.2.8. *Given X, Y, F, G, λ_U, λ_L defined as above in (2.2.7), let $\mathcal{C}$ be the bivariate copula of (X, Y). If the limits (2.6) and (2.7) exist, then tail dependence coefficients can be expressed as follows*

$$\lambda_U = 2 - \lim_{t\to 1^-} \frac{1-\mathcal{C}(t,t)}{1-t}$$
$$\lambda_L = \lim_{t\to 0^+} \frac{\mathcal{C}(t,t)}{t}$$

Here below we show some tail dependence formulas for the most commonly used copulas: elliptical (Gaussian and t-Student's) and Archimedean (Clayton, Frank and Gumbel). We note that t-Student's, Gumbel and Clayton copula show tail dependence.

Example 2.2.9 (Copulas and corresponding tail dependencies).

- Gaussian copula $C_\rho(u_1, u_2, ..., u_n) = \phi_n(\phi^{-1}(u_1), \phi^{-1}(u_2), ..., \phi^{-1}(u_n), \rho) \qquad \lambda_U = \lambda_L = 0$
  where $\phi_n,\ \phi^{-1}$ are the n-variate and inverse standard Gaussian cdf respectively.
- t-Student's copula $C_\rho^\nu(u_1, u_2, ..., u_n) = t_\nu^n(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2), ..., t_\nu^{-1}(u_n), \rho, \nu) \qquad \lambda_U = \lambda_L = 2\, t_{\nu+1}\left(\sqrt{\nu+1}\,\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right)$
  where $t_\nu,\ t_\nu^n,\ t_\nu^{-1}$ are the univariate, n-variate and inverse t-student's cdf with ν DoF respectively.

- Clayton copula $C_\theta(u,v) = [u^{-\theta} + v^{-\theta} - 1]^{-\frac{1}{\theta}}$ $\lambda_L = \lim_{t\to\infty} \frac{\varphi^{-1}(2t)}{\varphi^{-1}(t)} = \lim_{t\to\infty} \frac{(1+2t\theta)^{-\frac{1}{\theta}}}{(1+t\theta)^{-\frac{1}{\theta}}} = 2^{-\frac{1}{\theta}}$, $\lambda_U = 0$

- Frank copula $C_\theta(u,v) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right)$ like the Gaussian $\lambda_U = \lambda_L = 0$

- Gumbel copula $C_\theta(u,v) = e^{-[(-\ln u)^{-\theta} + (-\ln v)^{-\theta}]^{\frac{1}{\theta}}}$ $\lambda_U = \lim_{t\to 0^+} \frac{1-\varphi^{-1}(2t)}{1-\varphi^{-1}(t)} = \lim_{t\to 0^+} \frac{1-e^{-2t^{\frac{1}{\theta}}}}{1-e^{-t^{\frac{1}{\theta}}}} = 2 - 2^{\frac{1}{\theta}}$, $\lambda_L = 0$

2.2.4 Convolution of dependent variables

We now formally introduce the representation of the VaR aggregation problem for dependent variables. Under the assumption of independence, the problem of computing the VaR of a sum of risks would merely leads to the standard concept of convolution of random variables. Then, the natural extension of this concept to the case of dependent variables is the definition of *C-convolution* proposed by Cherubini et al. (2011). The term C in the definition reminds of the copula function representing a general relationship of a pair of variables X and Y. The finding was obtained as a by-product of a more general result for the characterization of the dependence structure between the variables X and $X+Y$ from the dependence between X and Y. We report here the main proposition.

Proposition 2.2.10. *Let X, Y be two real-valued random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with corresponding copula $C_{X,Y}$ and continuous marginals F_X, F_Y. Then,*

$$C_{X,X+Y}(u,v) = \int_0^u D_1 C_{X,Y}\Big(\omega, F_Y(F^{-1}_{X+Y}(v) - F_X^{-1}(\omega))\Big) d\omega \tag{2.8}$$

$$F_{X+Y}(t) = \int_0^1 D_1 C_{X,Y}\Big(\omega, F_Y(t - F_X^{-1}(\omega))\Big) d\omega \tag{2.9}$$

where $C_{i,j}(u,v)$ denotes the copula functions between the variables reported in the underscore, $D_1C(u,v)$ represents the derivative with respect to u, F_i denotes the distribution function of the variable reported in the underscore(X, Y and $X+Y$). Notice that in this result we have implicitly defined te concept of *C-convolution*.

Definition 2.2.11. Let F, H be two continuous c.d.f's and C a copula function. The *C-convolution* of H and F is defined as the c.d.f.

$$H \overset{C}{*} F(t) = \int_0^1 D_1 C\Big(\omega, F(t - H^{-1}(\omega))\Big)\, d\omega$$

These results are related to the conditional copula and pretty derived

$$\begin{aligned}
F_{X,X+Y}(s,t) &= P\left(X \le s, X+Y \le t\right) \\
&= \int_{-\infty}^{s} P\left(X+Y \le t | X = x\right)\ dF_X(x) \\
&= \int_{-\infty}^{s} P\left(Y \le t-x | X = x\right)\ dF_X(x) \\
&= \int_{-\infty}^{s} D_1 C_{X,Y}\left(F_X(x), F_Y(t-x)\right)\ dF_X(x) \\
&= \int_{0}^{F_X(s)} D_1 C_{X,Y}\left(\omega, F_Y(t - F_X^{-1}(\omega))\right)\ d\omega \quad \text{where } \omega = F_X(x) \in (0,1)
\end{aligned}$$

When $F_X(s)$ approach to infinity, one gets a definition of convolution generalised to arbitrary dependence between X and Y

$$F_X \overset{C}{*} F_Y(t) = F_{X+Y}(t) = \lim_{s\to\infty} F_{X,X+Y}(s,t) = \int_0^1 D_1 C_{X,Y}\left(\omega, F_Y(t - F_X^{-1}(\omega))\right)\ d\omega$$

Therefore, the copula linking $(X, X+Y)$ is

$$C_{X,X+Y}(u,v) = F_{X,X+Y}(F_X^{-1}(u), F_{X+Y}^{-1}(v))$$

This result provides a formal definition and an alternative computation approach to the convolution of dependent variables, with respect to the standard Monte Carlo. In some cases an important property of the *C-convolution* operator can be used to make the computation easier. In fact, it can be easily proved that the *C-convolution* operator is closed with respect to mixture of copula functions. In other words, it can be shown that if for some bivariate copula functions A and B we have

$$C(u,v) = \lambda A(u,v) + (1-\lambda)B(u,v),\ \forall\ \lambda\ \in\ [0,1]$$

then, for all c.d.f's H, F it holds

$$H \overset{C}{*} F(t) = H \overset{\lambda A+(1-\lambda)B}{*} F = \lambda H \overset{A}{*} F + (1-\lambda) H \overset{B}{*} F$$

Remark 2.2.4.1. Notice that the *C-convolution* can be formulated in terms of either the first or the second derivative. The related choice depends on the conditioning random variable that we take into account. For this reason, it is wise to select copula functions that are exchangeable. For these copulas, we have $C(u,v) = C(v,u)$, and are almost all the copula functions that are used in practice. This choice makes sure that, just like in the standard convolution of independent risk, the order in which the risks are aggregated does not matter. For exchangeable copulas we have then

$$F_X \overset{C}{*} F_Y(t) = \int_0^1 D_1 C_{X,Y}\Big(\omega, F_Y(t - F_X^{-1}(\omega))\Big) d\omega = \int_0^1 D_2 C_{X,Y}\Big(F_X(t - F_Y^{-1}(\omega)), \omega\Big) d\omega$$

2.3 Application

In this section, we consider the problem of risk diversification in a VaR framework with heavy-tailed α-stable risks and arbitrary dependence captured by a copula function. As motivated in the introduction, VaR may be the only available measure of risk in α-stable cases when the first moment is not defined, that is when $\alpha \leq 1$. We also remind that we are interested in exploring two issues:

- the shape of the non-monotonic relationship between dependence and the degree of super-additivity
- the role played by tail-dependence in the degree of super-additivity.

We consider two identically distributed α-stable r.v.'s X, Y and their sum $S = (X + Y)$ with dependence structure represented by copulas with different tail-dependence. We use the Kendall ρ_τ concordance measure to span a wide range of dependence levels. For each level, we compute a super-additivity index, as in Ibragimov and Prokhorov (2016)

$$SR = \frac{VaR_q(X+Y)}{VaR_q(X) + VaR_q(Y)}$$

Super-additivity is detected whether $SR \geq 1$, so that $VaR_q(X+Y)$, the risk corresponding to the portfolio, is greater than the sum of separated risks $VaR_q(X) + VaR_q(Y)$. When instead $SR < 1$, diversification is at work and the aggregation of risks reduces the probability of losing the capital allocated to support them.

2.3.1 Convolution through copulas

C-convolution computation provides a direct way the obtain the VaR of a bivariate portfolio $S = X + Y$ with dependence between the r.v.'s represented by their copula function. In what follows, we provide the analytical formulas of the C-convolution on the copulas reported in (1.2.30) through the definition in (2.9), that is

$$F_X \overset{C}{*} F_Y(t) = \int_0^1 D_2 C_{X,Y}\left(F_Y(t - F_X^{-1}(\omega)), \omega\right) d\omega$$

- **Gaussian Convolution**

  Through the Leibniz's rule, the partial derivative of $C_{Ga}(u_1, u_2; \rho)$ wrt u_2 is computed to be

$$D_2 C_{Ga}(u_1, u_2; \rho) = \Phi\left(\frac{\Phi^{-1}(u_1) - \rho\Phi^{-1}(u_2)}{\sqrt{1-\rho^2}}\right)$$

and the Gaussian C-convolution holds

$$F_X \overset{C_{Ga}}{*} F_Y(t) = \int_{-\infty}^{\infty} \Phi\left(\frac{\Phi^{-1}\left(F_X(t-y)\right) - \rho\Phi^{-1}\left(F_Y(y)\right)}{\sqrt{1-\rho^2}}\right) dF_Y(y) = \int_0^1 \Phi\left(\frac{\Phi^{-1}\left(F_X(t-F_Y^{-1}(\omega))\right) - \rho\Phi^{-1}(\omega)}{\sqrt{1-\rho^2}}\right) d\omega$$

- **t-Student's Convolution**

The bivariate Student's t copula is defined as

$$C_{St}^{\rho,\nu}(u_1,u_2) = \mathbf{t}_\nu^\rho(t_\nu^{-1}(u_1),\ t_\nu^{-1}(u_2)) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{\Gamma(\frac{\nu+2}{2})}{\Gamma(\frac{\nu}{2})\pi\nu\sqrt{1-\rho^2}} \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)}\right)^{-\frac{(\nu+2)}{2}} dx_1 dx_2$$

where $\Gamma(\cdot)$ is the Gamma function and $\mathbf{t}_\nu$ is the bivariate Student's t distribution function with ν degree.

The partial derivative of $C_{St}^{\rho,\nu}(u_1,u_2)$ wrt u_2 is

$$D_2C\left(u_1,u_2;\rho,\nu\right) = t_{\nu+1}\left(\frac{t_\nu^{-1}\left(u_1\right) - \rho\ t_\nu^{-1}\left(u_2\right)}{\sqrt{\frac{\left(\nu + t_\nu^{-1}(u_2)^2\right)(1-\rho^2)}{\nu+1}}}\right) \tag{2.10}$$

and the Student's t C-convolution with ν degree is

$$F_X \overset{C_{St}^{\nu}}{*} F_Y(t) = \int_{-\infty}^{\infty} t_{\nu+1}\left(\frac{t_\nu^{-1}\left(F_X(t-y)\right) - \rho\ t_\nu^{-1}\left(F_Y(y)\right)}{\sqrt{\frac{(\nu+y^2)(1-\rho^2)}{\nu+1}}}\right) dF_Y(y) = \int_0^1 t_{\nu+1}\left(\frac{t_\nu^{-1}\left(F_X(t-F_Y^{-1}(\omega))\right) - \rho\ t_\nu^{-1}\left(\omega\right)}{\sqrt{\frac{\left(\nu+t_\nu^{-1}(\omega)^2\right)(1-\rho^2)}{\nu+1}}}\right) d\omega$$

- **Clayton Convolution**

The Clayton copula is expressed by the formula below:

$$C_{Cl}(u_1,u_2;\theta) = max\left(\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-\frac{1}{\theta}}, 0\right) \qquad \text{with copula parameter } \theta \in (0,+\infty)$$

The partial derivative w.r.t. u_2 is

$$D_2C_{Cl}\left(u_1,u_2;\theta\right) = u_2^{-\frac{1+\theta}{\theta}}\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)$$

Therefore, provided that $u_1 = F_X(t-y)$ and $u_2 = F_Y(y)$ the C-convolution has the form

$$F_X \overset{C_{Cl}}{*} F_Y(t) = \int_{-\infty}^{\infty} \left(F_Y(y)^{-\frac{1+\theta}{\theta}}\left(F_X(t-y)^{-\theta} + F_Y(y)^{-\theta} - 1\right)\right) dF_Y(y) = \int_0^1 \omega^{-\frac{1+\theta}{\theta}}\left(F_X(t-F_Y^{-1}(\omega))^{-\theta} + \omega^{-\theta} - 1\right) d\omega$$

- **Frank Convolution**

The Frank copula has the following expression

$$C_{fr}(u_1, u_2; \theta) = -\frac{1}{\theta} \ln\left(1 + \frac{(\mathrm{e}^{-\theta u_1} - 1)(\mathrm{e}^{-\theta u_2} - 1)}{\mathrm{e}^{-\theta} - 1}\right) \qquad \text{with copula parameter } \theta \in (-\infty, +\infty)$$

The partial derivative w.r.t. u_2 is

$$D_2 C_{fr}\,(u_1, u_2; \theta) = \frac{\mathrm{e}^{-\theta u_2}(\mathrm{e}^{-\theta u_1} - 1)}{\mathrm{e}^{-\theta} - 1 + (\mathrm{e}^{-\theta u_1} - 1)(\mathrm{e}^{-\theta u_2} - 1)}$$

Therefore, provided that $u_1 = F_X(t - y)$ and $u_2 = F_Y(y)$ the C-convolution has the form

$$F_X \overset{C_{fr}}{*} F_Y(t) = \int_{-\infty}^{\infty} \left(\frac{\mathrm{e}^{-\theta F_Y(y)}(\mathrm{e}^{-\theta F_X(t-y)} - 1)}{\mathrm{e}^{-\theta} - 1 + (\mathrm{e}^{-\theta F_X(t-y)} - 1)(\mathrm{e}^{-\theta F_Y(y)} - 1)}\right) dF_Y(y) = \int_0^1 \frac{\mathrm{e}^{-\theta\omega}(\mathrm{e}^{-\theta F_X(t - F_Y^{-1}(\omega))} - 1)}{\mathrm{e}^{-\theta} - 1 + (\mathrm{e}^{-\theta F_X(t - F_Y^{-1}(\omega))} - 1)(\mathrm{e}^{-\theta\omega} - 1)}\, d\omega$$

- **Gumbel Convolution**

  The Gumbel copula is

$$C_{Gum}(u_1,u_2;\theta) = \mathrm{e}^{-\left((-\ln u_1)^{-\theta}+(-\ln u_2)^{-\theta}\right)^{\frac{1}{\theta}}} = \mathrm{e}^{-(z_1+z_2)^{\frac{1}{\theta}}} \text{ where } z_i = (-\ln u_i)^{-\theta}\ i=1,2 \text{ and } \theta \in (1,+\infty)$$

  The partial derivative w.r.t. u_2 is

$$\begin{aligned} D_2 C_{Gum}(u_1,u_2;\theta) &= \frac{\partial}{\partial u_2}\mathrm{e}^{-(z_1+z_2)^{\frac{1}{\theta}}} = -\mathrm{e}^{-(z_1+z_2)^{\frac{1}{\theta}}}\frac{1}{\theta}(z_1+z_2)^{\frac{1-\theta}{\theta}}\frac{\partial z_2}{\partial u_2} \\ \frac{\partial z_2}{\partial u_2} &= \frac{\partial}{\partial u_2}(-\ln u_2)^{-\theta} = \frac{\partial}{\partial u_2}\mathrm{e}^{\ln\left((-\ln u_2)^{-\theta}\right)} = \frac{\partial}{\partial u_2}\mathrm{e}^{-\theta\ln(-\ln u_2)} = z_2\frac{\partial}{\partial u_2}\ln(-\ln u_2) \\ \frac{\partial}{\partial u_2}\ln(-\ln u_2) &= \frac{-1/u_2}{-\ln u_2} = \frac{1}{u_2\ln u_2} \end{aligned}$$

  The result for partial derivatives becomes

$$\begin{aligned} D_2 C_{Gum}(u_1,u_2;\theta) &= -\mathrm{e}^{-(z_1+z_2)^{\frac{1}{\theta}}}\frac{1}{\theta}(z_1+z_2)^{\frac{1-\theta}{\theta}}\frac{z_2}{u_2\ln u_2} \\ &= -\mathrm{e}^{-((-\ln u_1)^{-\theta}+(-\ln u_2)^{-\theta})^{\frac{1}{\theta}}}\frac{1}{\theta}((-\ln u_1)^{-\theta}+(-\ln u_2)^{-\theta})^{\frac{1-\theta}{\theta}}\frac{(-\ln u_2)^{-\theta}}{u_2\ln u_2} \end{aligned}$$

  The C-convolution for Gumbel copula gives

$$\begin{aligned} F_X \overset{C_{Gum}}{*} F_Y(t) &= \int_{-\infty}^{\infty} -\mathrm{e}^{-\left((-\ln F_X(t-y))^{-\theta}+(-\ln F_Y(y))^{-\theta}\right)^{\frac{1}{\theta}}}\frac{1}{\theta}((-\ln F_X(t-y))^{-\theta}+(-\ln F_Y(y))^{-\theta})^{\frac{1-\theta}{\theta}}\frac{(-\ln F_Y(y))^{-\theta}}{F_Y(y)\ln F_Y(y)}\,dF_Y(y) \\ &= \int_0^1 -\mathrm{e}^{-\left((-\ln F_Y^{-1}(\omega))\right)^{-\theta}+(-\ln\omega)^{-\theta})^{\frac{1}{\theta}}}\frac{1}{\theta}\left((-\ln F_X(t-F_Y^{-1}(\omega)))^{-\theta}+(-\ln\omega)^{-\theta}\right)^{\frac{1-\theta}{\theta}}\frac{(-\ln\omega)^{-\theta}}{\omega\ln\omega}\,d\omega \end{aligned}$$

2.3.2 Convolution and general results

The C-convolution representations provides an alternative to the standard tool that has been applied to the problem of risk aggregation, that is Monte Carlo simulation. It is in fact well known that Monte Carlo requires intensive computational effort, and the problem is compounded in applications, like VaR analysis, in which simulation should be focussed on the tails. While importance sampling techniques can be applied to ease this flaw, the C-convolution approach provides a direct way to address the problem. So, C-convolution analysis offers a better tool to investigate the aggregation analysis quite further in the tail.
Therefore, here we consider a portfolio of two assets X, Y drawn from heavy-tailed marginal distribution with several tail indexes α ranging between $0 < \alpha \leq 2$ and linked by a bivariate copula determining their dependence structure. The general scheme for computing the super-additivity ratio for any given couple of identically

distributed stable r.v.'s (X, Y) with common c.d.f. F, follows the following three steps

1) We select a given the Copula function $C_{\rho_\tau} = C_{X,Y}^{\rho_\tau}$ and we assume a Kendall dependence ρ_τ such that

$$F_{X+Y}(t) = F \overset{C_{\rho_\tau}}{*} F(t) = \int_0^1 D_2 C_{\rho_\tau}\left(F(t - F^{-1}(\omega)), \omega\right) d\omega$$

2) For a given quantile level q, we look for the point $t \in \mathbb{R} : F_{X+Y}(t) = q$. Because of the increasing property of the distribution of $(X + Y)$, the inverse convolution function $t = F_{X+Y}^{-1}(q)$ exists and we run a root-finding method, searching for the point $t^* = \underset{t\in\mathbb{R}}{inf}\left\{F_{X+Y}(t) - q = 0\right\}$. To this aim, we chose the Brent-Dekker's method, that realizes a mixed-search through Bisection, Secant and inverse quadratic interpolation with a zero-precision level $\epsilon = 10^{-8}$. For a complete description of the method, see Press et al. (2007, p. 454) .

3) We compute the S-Ratio. The VaR of the portfolio $X + Y$ with Kendall ρ_τ at level q is

$$VaR_q(X + Y) = \underset{t\in\mathbb{R}}{inf}\left\{F_{X+Y}(t) = q\right\} \quad \text{that is} \quad t = F_{X+Y}^{-1}(q)$$

Since it is well known that VaR is comonotonic additive, the perfect dependence VaR of the portfolio is

$$VaR_q(X) + VaR_q(Y) = 2VaR_q(X) \qquad \text{since X,Y are identically distributed r.v.'s}$$

The ratio $SR_q(X + Y) = VaR_q(X + Y)\big/2\ VaR_q(X)$ provides a measure of super-additivity for $\rho_S < 1$.

The tables 1 to 5 report the results for different α values of the marginal stable distributions and the main families copula functions typically used, that is the elliptical (gaussian and Student-t) and the Archimedean ones (Clayton, Gumbel and Frank). As it is well known, only the Gaussian and Frank copulas do not have tail dependence, the Gumbel has only upper tail dependence, the Clayton has only lower tail dependence and t-Student's has both upper and lower tail dependence.

All these results for the different copulas seem to design a consistent picture. Some diversification benefits typically appear in the upper and right parts of the tables. At first sight, sub-additivity generally holds when the value of the tail index α is higher than 1. In some cases, diversification may also hold for α values lower than 1 if there is negative association. This is particularly true for elliptical copulas.

As for super-additivity, the failure of diversification is a general result for cases in which the first moment does not exist. The results are instead mixed in the region in which the first moment is defined, where diversification may still fail for higher level of dependence. However, even in the positive dependence region in which diversification benefits materialize, their amount is of a limited order.

Finally, a casual look at the comparison of different dependence models suggests that the red super-additivity region seems to be wider in the gaussian dependence than in the Student-t copula case. The same evidence seems to appear from a comparison of the Frank copula with the Clayton one. The common feature of these two comparisons is that super-additivity seems to be less extended in the cases in which the copula functions has positive tail dependence. On the role of tail dependence we focus in the analysis that follows.

Table 2.1: C-convolution integration: Super-additivity Ratio, Gaussian copula

q=5%	α	**0.2**	**0.3**	**0.4**	**0.6**	**0.9**	**1**	**1.1**	**1.3**	**1.6**	**1.8**	**2**
Param. θ	Kendall-τ	**Gaussian copula**										
-1	-1	5.0e-06	3.5e-06	2.7e-06	1.9e-06	1.3e-06	1.2e-06	1.1e-06	9.8e-07	8.8e-07	8.4e-07	7.9e-07
-0.99	-0.9	0.5443	0.3593	0.2754	0.1929	0.1368	0.1241	0.1142	0.0980	0.0878	0.0835	0.0785
-0.89	-0.7	**2.4852**	**1.2751**	0.8931	0.5832	0.4034	0.3676	0.3378	0.2920	0.2608	0.2483	0.2334
-0.59	-0.4	**8.7629**	**3.2337**	**1.9524**	**1.1392**	0.7621	0.6946	0.6416	0.5661	0.5063	0.4794	0.4543
-0.45	-0.3	**11.3880**	**3.9143**	**2.2594**	**1.2959**	0.8598	0.7878	0.7324	0.6497	0.5813	0.5505	0.5225
-0.16	-0.1	**15.3299**	**4.8540**	**2.7295**	**1.5192**	**1.0226**	0.9452	0.8810	0.7938	0.7136	0.6786	0.6493
-0.02	-0.01	**16.0090**	**5.0211**	**2.8224**	**1.5814**	**1.0740**	0.9935	0.9344	0.8469	0.7656	0.7291	0.7015
0	0	**15.8676**	**5.0450**	**2.8237**	**1.5858**	**1.0792**	**1.0009**	0.9386	0.8522	0.7698	0.7350	0.7072
0.02	0.01	**16.1089**	**5.0292**	**2.8408**	**1.5927**	**1.0840**	**1.0043**	0.9430	0.8570	0.7765	0.7406	0.7127
0.16	0.1	**15.6895**	**4.9868**	**2.8299**	**1.6148**	**1.1193**	**1.0417**	0.9832	0.9015	0.8215	0.7865	0.7604
0.31	0.2	**14.0857**	**4.7032**	**2.7219**	**1.5978**	**1.1398**	**1.0676**	**1.0132**	0.9409	0.8661	0.8323	0.8088
0.45	0.3	**11.9989**	**4.2562**	**2.5444**	**1.5526**	**1.1439**	**1.0826**	**1.0343**	0.9672	0.9032	0.8724	0.8528
0.59	0.4	**9.5069**	**3.6756**	**2.3091**	**1.4737**	**1.1361**	**1.0804**	**1.0418**	0.9869	0.9332	0.9070	0.8909
0.89	0.7	**3.2120**	**1.8769**	**1.4633**	**1.1810**	**1.0556**	**1.0358**	**1.0230**	**1.0053**	0.9870	0.9773	0.9725
0.99	0.9	**1.2798**	**1.1202**	**1.0652**	**1.0234**	**1.0072**	**1.0047**	**1.0038**	**1.0011**	**0.9990**	0.9973	0.9969
1	1	**1**	**1.0000**	**1.0000**	**1**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1.0000**	**1**

Table 2.2: C-convolution integration: Super-additivity Ratio, Student-t($\nu = 5$) copula

q=5%	α	**0.2**	**0.3**	**0.4**	**0.6**	**0.9**	**1**	**1.1**	**1.3**	**1.6**	**1.8**	**2**
Param. θ	Kendall-τ	**Student-t copula** ($\nu = 5$)										
-1	-1	4.2e-06	2.9e-06	2.3e-06	1.7e-06	1.2e-06	1.1e-06	1e-06	9.3e-07	8.6e-07	8.2e-07	7.7e-07
-0.99	-0.9	0.4496	0.3004	0.2311	0.1658	0.1198	0.1104	0.1027	0.0927	0.0858	0.0816	0.0769
-0.89	-0.7	**1.8897**	**1.0421**	0.7376	0.4997	0.3557	0.3281	0.3061	0.2763	0.2550	0.2429	0.2294
-0.59	-0.4	**6.1498**	**2.5362**	**1.6017**	0.9842	0.6844	0.6331	0.5929	0.5358	0.4920	0.4700	0.4484
-0.45	-0.3	**7.9169**	**3.0426**	**1.8669**	**1.1213**	0.7806	0.7217	0.6767	0.6141	0.5638	0.5401	0.5172
-0.16	-0.1	**10.5219**	**3.7700**	**2.2448**	**1.3357**	0.9367	0.8715	0.8222	0.7521	0.6935	0.6670	0.6446
-0.02	-0.01	**10.9465**	**3.8993**	**2.3270**	**1.3924**	0.9894	0.9217	0.8736	0.8050	0.7456	0.7187	0.6970
0.00	0	**11.0378**	**3.9264**	**2.3414**	**1.3982**	0.9942	0.9271	0.8787	0.8102	0.7511	0.7241	0.7028
0.02	0.01	**11.0536**	**3.9473**	**2.3387**	**1.4050**	0.9961	0.9327	0.8839	0.8157	0.7561	0.7296	0.7088
0.16	0.1	**10.6819**	**3.9086**	**2.3511**	**1.4276**	**1.0338**	0.9721	0.9250	0.8599	0.8015	0.7760	0.7572
0.31	0.2	**9.9593**	**3.7179**	**2.2905**	**1.4310**	**1.0600**	**1.0041**	0.9607	0.9005	0.8466	0.8225	0.8059
0.45	0.3	**8.5283**	**3.3891**	**2.1585**	**1.3972**	**1.0739**	**1.0241**	0.9848	0.9332	0.8860	0.8641	0.8503
0.59	0.4	**6.8569**	**2.9793**	**1.9817**	**1.3481**	**1.0743**	**1.0328**	**1.0005**	0.9585	0.9195	0.9002	0.8894
0.89	0.7	**2.6423**	**1.6706**	**1.3521**	**1.1307**	**1.0340**	**1.0180**	**1.0087**	0.9957	0.9820	0.9752	0.9717
0.99	0.9	**1.2236**	**1.0999**	**1.0502**	**1.0177**	**1.0045**	**1.0028**	**1.0018**	**0.9997**	0.9982	0.9972	0.9968
1	1	**1.0000**	**1**	**1.0000**	**1**	**1**	**1.0000**	**1**	**1.0000**	**1.0000**	**1.0000**	**1.0000**

Table 2.3: C-convolution integration: Super-additivity Ratio, Frank copula

q=5%	α	**0.2**	**0.3**	**0.4**	**0.6**	**0.9**	**1**	**1.1**	**1.3**	**1.6**	**1.8**	**2**
Param. θ	Kendall-τ	**Frank copula**										
-709	-1	0.0698	0.0483	0.0376	0.0261	0.0181	0.0163	0.0147	0.0120	0.0088	0.0074	0.0065
-38.28	-0.9	**2.3807**	**1.1773**	0.7939	0.4945	0.3189	0.2846	0.2555	0.2065	0.1507	0.1257	0.1088
-11.41	-0.7	**8.4532**	**3.1117**	**1.8384**	**1.0425**	0.6610	0.5911	0.5312	0.4402	0.3435	0.3029	0.2727
-4.16	-0.4	**13.5038**	**4.4572**	**2.4888**	**1.3722**	0.8904	0.8077	0.7394	0.6404	0.5492	0.5088	0.4756
-2.92	-0.3	**14.7779**	**4.6550**	**2.6111**	**1.4407**	0.9428	0.8596	0.7968	0.6990	0.6102	0.5701	0.5383
-0.91	-0.1	**15.8484**	**4.9234**	**2.7937**	**1.5473**	**1.0357**	0.9593	0.8947	0.8046	0.7215	0.6845	0.6553
-0.09	-0.01	**15.9420**	**5.0413**	**2.8390**	**1.5899**	**1.0759**	0.9937	0.9357	0.8467	0.7656	0.7316	0.7014
0	0	**16.0998**	**5.0095**	**2.8282**	**1.5900**	**1.0811**	**0.9993**	0.9373	0.8514	0.7709	0.7347	0.7071
0.09	0.01	**15.9468**	**5.0584**	**2.8299**	**1.5919**	**1.0835**	**1.0070**	0.9424	0.8582	0.7769	0.7392	0.7113
0.91	0.1	**15.8797**	**5.0614**	**2.8471**	**1.6093**	**1.1162**	**1.0421**	0.9804	0.8948	0.8162	0.7812	0.7546
1.86	0.2	**15.7188**	**5.0186**	**2.8649**	**1.6352**	**1.1483**	**1.0732**	**1.0178**	0.9373	0.8604	0.8240	0.7987
2.92	0.3	**15.0588**	**4.9029**	**2.8181**	**1.6495**	**1.1761**	**1.1040**	**1.0505**	0.9729	0.8979	0.8626	0.8370
4.16	0.4	**14.2756**	**4.7217**	**2.7806**	**1.6555**	**1.1998**	**1.1306**	**1.0806**	**1.0088**	0.9337	0.8970	0.8730
11.41	0.7	**9.9142**	**3.8242**	**2.4189**	**1.5835**	**1.2305**	**1.1763**	**1.1376**	**1.0810**	**1.0153**	0.9836	0.9625
38.28	0.9	**3.4557**	**2.0352**	**1.5906**	**1.2843**	**1.1380**	**1.1143**	**1.0941**	**1.0657**	**1.0322**	**1.0162**	**1.0067**
709	1	**1.0122**	**1.0054**	**1.0036**	**1.0006**	**1.0005**	**1.0007**	**1.0003**	**1.0004**	**1.0001**	**1.0002**	**1.0001**

Table 2.4: C-convolution integration: Super-additivity Ratio, Clayton copula

q=5%	α	**0.2**	**0.3**	**0.4**	**0.6**	**0.9**	**1**	**1.1**	**1.3**	**1.6**	**1.8**	**2**
Param. θ	Kendall-τ	**Clayton copula**										
-1	-1	5.8e-06	4.2e-06	3.3e-06	2.4e-06	1.8e-06	1.7e-06	1.6e-06	1.4e-06	1.3e-06	1.2e-06	1.1e-06
-0.95	-0.9	0.9751	0.5325	0.3778	0.2560	0.1832	0.1681	0.1562	0.1387	0.1223	0.1149	0.1083
-0.82	-0.7	**3.0610**	**1.4229**	0.9407	0.5954	0.4177	0.3842	0.3590	0.3208	0.2916	0.2801	0.2703
-0.57	-0.4	**6.7280**	**2.6453**	**1.6368**	**0.9999**	0.6877	0.6338	0.5909	0.5353	0.4923	0.4763	0.4625
-0.46	-0.3	**8.5693**	**3.2123**	**1.9400**	**1.1556**	0.7896	0.7265	0.6812	0.6152	0.5638	0.5416	0.5240
-0.18	-0.1	**13.8704**	**4.5987**	**2.6015**	**1.4699**	0.9976	0.9247	0.8652	0.7817	0.7071	0.6741	0.6479
-0.02	-0.01	**15.7198**	**4.9752**	**2.8251**	**1.5697**	**1.0741**	0.9948	0.9318	0.8455	0.7651	0.7291	0.7016
0.00	0	**16.0693**	**4.9913**	**2.8240**	**1.5876**	**1.0786**	0.9981	0.9422	0.8518	0.7711	0.7348	0.7072
0.02	0.01	**15.6811**	**4.9562**	**2.7989**	**1.5823**	**1.0837**	**1.0037**	0.9454	0.8576	0.7779	0.7424	0.7150
0.22	0.1	**12.3384**	**4.2727**	**2.5377**	**1.5152**	**1.0929**	**1.0222**	0.9694	0.9004	0.8354	0.8054	0.7848
0.50	0.2	**7.9749**	**3.2764**	**2.1102**	**1.3780**	**1.0620**	**1.0143**	0.9801	0.9312	0.8860	0.8654	0.8544
0.86	0.3	**4.9692**	**2.4106**	**1.7135**	**1.2439**	**1.0379**	**1.0030**	0.9806	0.9532	0.9263	0.9154	0.9102
1.33	0.4	**3.2074**	**1.8440**	**1.4295**	**1.1446**	**1.0178**	**0.9989**	0.9865	0.9716	0.9582	0.9512	0.9490
4.67	0.7	**1.3079**	**1.1236**	**1.0621**	**1.0201**	**1.0018**	**1.0000**	0.9984	0.9966	0.9952	0.9945	0.9941
18	0.9	**1.0255**	**1.0099**	**1.0057**	**1.0016**	**0.9998**	**1.0001**	**0.9999**	**0.9997**	**0.9998**	**0.9995**	**0.9997**
2e06	1	**1**	**1**	**1**	**1**	**1.0000**	**1**	**1**	**1**	**1.0000**	**1**	**1**

Table 2.5: C-convolution integration: Super-additivity Ratio, Gumbel copula

q=5%	α	**0.2**	**0.3**	**0.4**	**0.6**	**0.9**	**1**	**1.1**	**1.3**	**1.6**	**1.8**	**2**
Param. θ	Kendall-τ	**Clayton copula**										
1.00	0	**16.0952**	**5.0635**	**2.8016**	**1.5854**	**1.0807**	0.9984	0.9403	0.8519	0.7711	0.7349	0.7077
1.01	0.01	**16.0678**	**5.0694**	**2.8456**	**1.5901**	**1.0820**	**1.0034**	0.9438	0.8545	0.7738	0.7386	0.7112
1.11	0.1	**15.9600**	**5.0510**	**2.8431**	**1.6124**	**1.1137**	**1.0306**	0.9742	0.8887	0.8084	0.7738	0.7459
1.25	0.2	**15.2526**	**4.9554**	**2.8090**	**1.6174**	**1.1346**	**1.0620**	**1.0070**	0.9236	0.8465	0.8094	0.7845
1.43	0.3	**14.0699**	**4.6658**	**2.7374**	**1.6078**	**1.1502**	**1.0810**	**1.0275**	0.9535	0.8804	0.8460	0.8232
1.67	0.4	**12.4740**	**4.2964**	**2.5933**	**1.5756**	**1.1609**	**1.0940**	**1.0461**	0.9807	0.9125	0.8800	0.8596
2	0.5	**10.1809**	**3.8287**	**2.3758**	**1.5114**	**1.1523**	**1.0994**	**1.0567**	**0.9995**	0.9416	0.9122	0.8946
3.33	0.7	**5.1980**	**2.5551**	**1.8028**	**1.3201**	**1.1080**	**1.0779**	**1.0533**	**1.0180**	0.9833	0.9662	0.9556
5	0.8	**3.1194**	**1.8558**	**1.4688**	**1.1913**	**1.0671**	**1.0495**	**1.0356**	**1.0163**	0.9950	0.9842	0.9787
10	0.9	**1.6575**	**1.2853**	**1.1554**	**1.0637**	**1.0233**	**1.0195**	**1.0127**	**1.0060**	**0.9996**	0.9964	0.9941
100	0.99	**1.0086**	**1.0043**	**1.0017**	**1.0009**	**1.0004**	**1.0002**	**1.0002**	**1.0003**	**1.0001**	**1.0000**	**0.9999**
1e06	1	**1.0000**	**1**	**1**	**1.0000**	**1.0000**	**1**	**1.0000**	**1.0000**	**1.0000**	**1**	**1**

2.3.3 Diversification failure and tail dependence

In this section we explore the impact of tail dependence on the limits to diversification. The intuitive idea seems to be that higher tail dependence should worsen the scope for diversification. Surprisingly, our results point out exactly the opposite. In Figure 2.1 we report the graph of the relationship between the dependence and the super-additivity indexes for two different levels of α, both in the region in which the first moment does not exist. The relationship is drawn for several copulas of the elliptical class with different tail-dependence indexes, that is the Gaussian copula (t-Student's with $\nu = \infty$), for which the tail dependence is zero, and t-Student's, for which the tail dependence increases for lower degrees of freedom. The highest tail dependence is then represented by the Student-t copula with 1 degree of freedom. We notice that for the case $\alpha = 0.2$ the relationship is non monotone for all copulas, and the maximum of super-additivity is reached around the independence case. The interesting finding, that is new to the best of our knowledge, is that super-additivity is uniformly lower for copulas with higher tail dependence. The evidence is confirmed in the case $\alpha = 0.8$, in which for the copula with highest tail dependence VaR is even sub-additive across the whole range of dependence.

So, there seems to be a puzzle in the result that higher tail dependence appears to ease the limits to diversification, rather than making them more binding. A very high degree of tail dependence may even destroy the non-monotonic nature of relationship between dependence and super-additivity, and so create a space for diversification.

The first conjecture that comes to mind to explain this result is that there could be a point in the tail where the expected ranking is established. The conjecture is suggested by what happens for gaussian risks, as documented in Figure 2.2. In this case, the relationship between dependence and diversification shows the same ranking as that in the previous cases if we consider the 10% percentile, while the order is completely reversed at the 1% percentile.

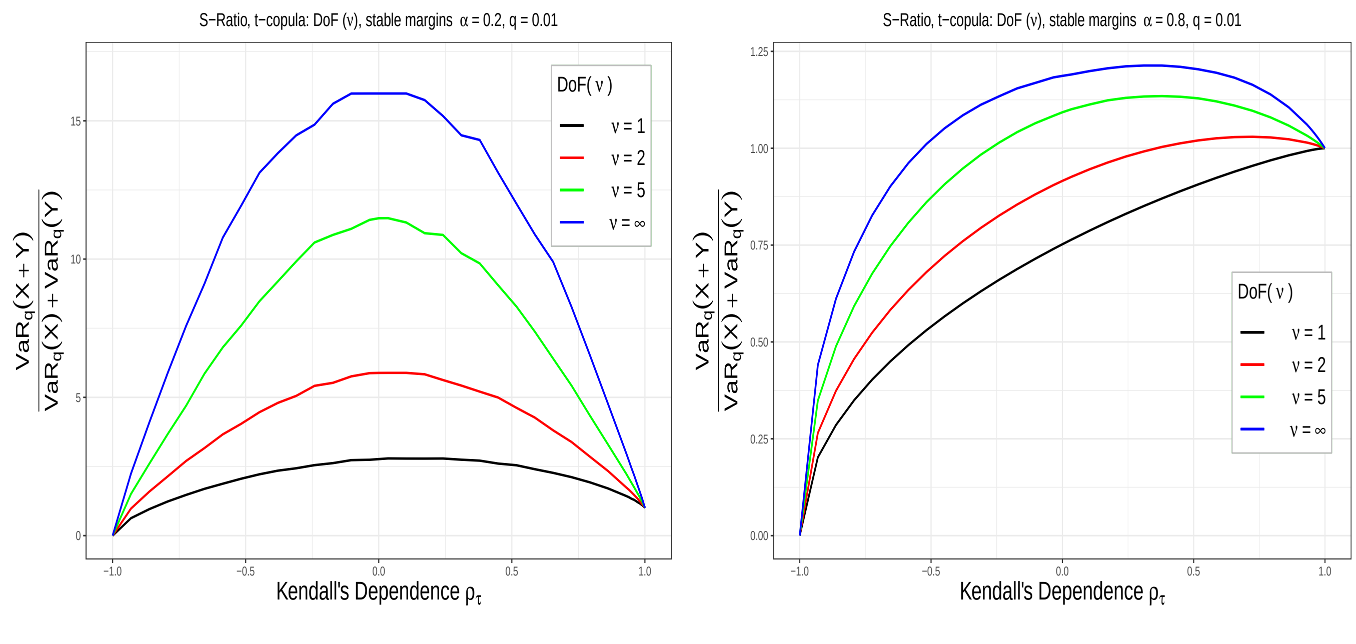


Figure 2.1: Risk structure on different Stable marginals ($\alpha = 0.2, \alpha = 0.8$), elliptical copulas

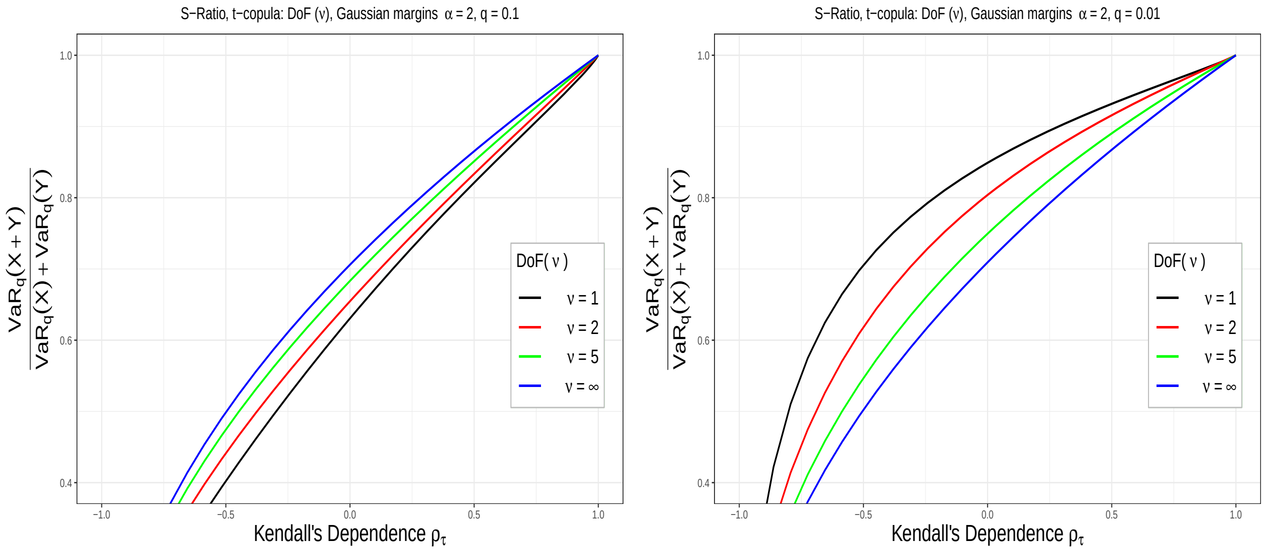


Figure 2.2: Risk structure on different quantile, Gaussian marginals, elliptical copulas

Our conjecture is not confirmed if we perform the same analysis for Cauchy marginals ($\alpha = 1$). Figure 2.3 depicts the ranking of the elliptical copulas with different tail dependence for quantiles further in the tails. We verify that the unexpected ranking persists. Dependence structures with higher tail dependence allow to save VaR capital. Moreover, a comparison of the different tails reveals another facet of this tail dependence paradox. Not only the schedule of the relationship between dependence and diversification is uniformly lower for copulas with increasing tail dependence. It also emerges that beyond some point in the tail the relationship is monotone. In other words, at some point in the tail a high degree of tail dependence creates diversification benefits. Even though the figure shows that, far in the tail, numerical integration error blurs the schedule of the relationship, the general path emerges quite clear. Deep in the tails, all the relationships become weakly monotonic, with an upper ceiling given by the comonotonic bound, $SR = 1$, beyond a critical dependence level. Moreover, this additivity ceiling is reached for lower levels of dependence, the lower the tail dependence value of the copula function.

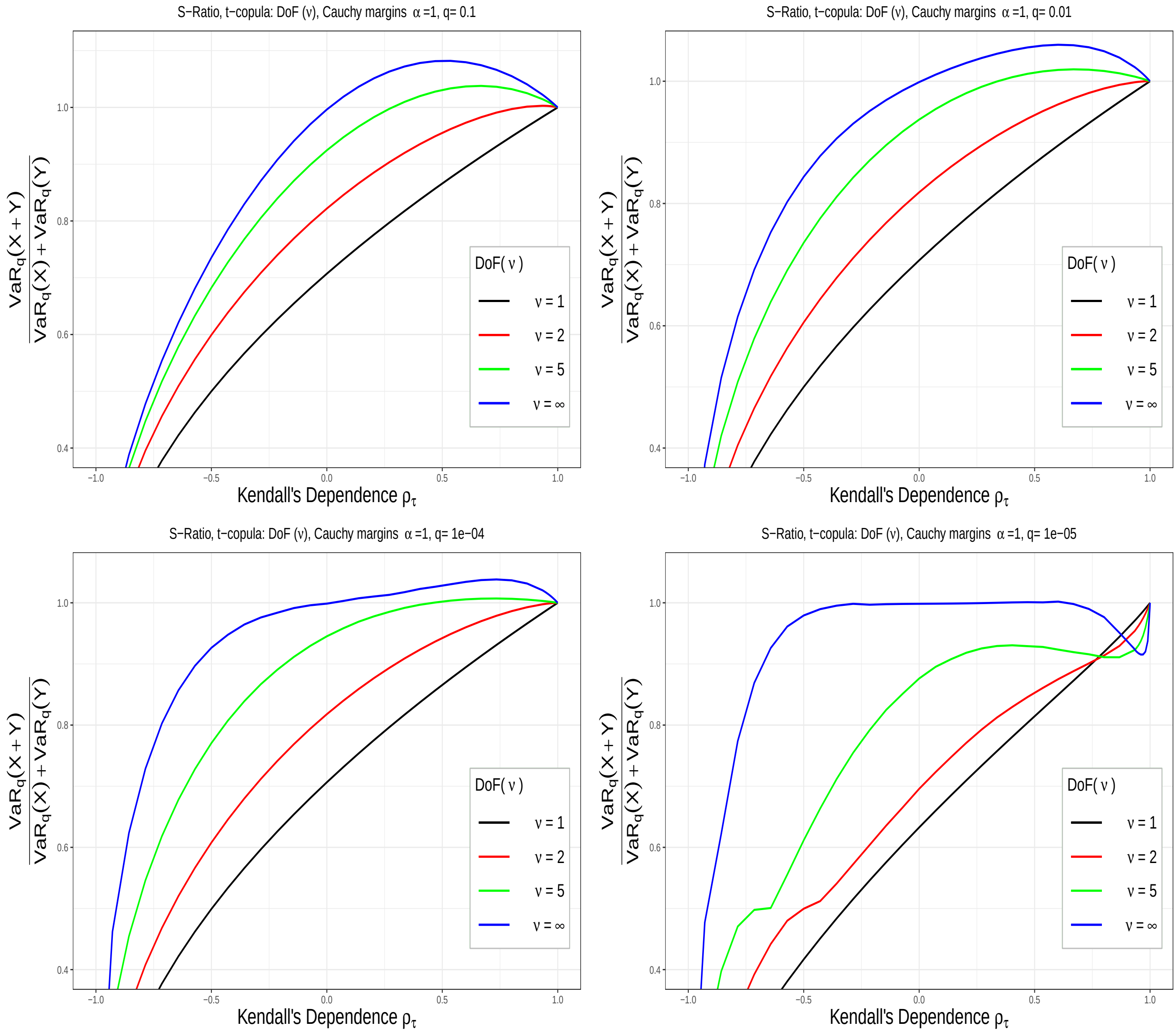


Figure 2.3: Risk structure on different quantile, Cauchy marginals, elliptical copulas

2.4 Conclusions

Heavy tailed α-stable risks have been largely used in the representation of actuarial risks, particularly in the field of catastrophe insurance, and operational risk, which is the kind of financial risk which is closest to the kinds of risk that are typical of insurance. Since the 60s, the α-stable distribution has also been proposed as a possible model to represent the dynamics of speculative prices. In this paper, we provide numerical analysis to study the aggregate behavior of losses for α-stable risks, with reference to a *super-additivity* index.

We contribute to the literature on this kind of risk in two directions. First, we confirm the results of super-additivity for stable risks with α lower than 1, that can only be mitigated with negative dependence. For what concerns the insurance market, where these kinds of risk are often found, our results support evidence concerning the impossibility of re-insurance policies. In fact, any reinsurance policy of α-stable risks would call for more capital than that needed to shield the losses on individual risks. Moreover, the super-additivity curse mainly hits cases in which the α parameter is lower than 1 and the mean is not defined. This implies that in this case coherent risk measures like the *Expected Shortfall* are not available and the Value-at-Risk is the obliged choice.

As for the analysis of diversification in the tail of the distribution, our simulations provide evidence that in our view represents a new puzzle. If we rank the dependence structures according to their tail dependence indexes, we find that structures with higher tail dependence provide more space for diversification. In particular, we find that: i) the super-additivity ratio for every level of dependence is lower for dependence structures with higher tail dependence; ii) for extreme points in the tail the aggregate VaR of risks becomes additive beyond a critical dependence level, and this critical level is lower for copula functions with lower tail dependence.

Concerning the latter point, our semi-analytic approach based on numerical integration opens a new research path that applies analysis to explore the very extreme tail behaviour of aggregated risks, parallel to a stream of literature that instead has studied the asymptotic tail behaviour of aggregated VaR from an analytical point of view (see for example Embrechts et al. (2009)). A more extensive empirical investigation of the actual convergence of the aggregated VaR to this limit is left as a topic for future research. The analytical study of the behaviour of VaR aggregation in model with tail dependence remains the other main avenue for future research.

Chapter 3

Value-at-Risk and the Tail Dependence Puzzle

Abstract

We document a new paradox in VaR aggregation. When marginal distribution have sufficiently heavy tails an increase in the tail dependence index may reduce the aggregated Value-at-Risk. We address by simulation conditions under which this result may persist for all finite percentiles in the tail. When this is case, the tail dependence puzzle is also extended to other law invariant measures, such as Expected Shortfall.

Keywords

risks aggregation, risk measures, copula functions, tail dependence structure.

3.1 Introduction

Consider a portfolio of two risks identically distributed and dependent. Assume the dependence structure is a Gaussian copula and you compute a 5536 $ risk capital requirement using Value-at-Risk. Say you are worried how this capital charge could change if the true dependence structure would actually exhibit tail dependence, such as with a Student-t copula with same correlation but 1 degree of freedom. You would think that tail dependence is bad and would cost you more capital, wouldn't you? Well, it may happen that the new capital charge would be now 4826 $ and heavy tail dependence would save you money. We call this the *tail dependence puzzle*. This paper is devoted to investigate this puzzle using both analytical and simulation means.

One could argue that we are putting together different concepts. On one side we computed 5536 $ risk capital without tail dependence against 4826 $ with heavy tail dependence on a *finite* quantile. On the other side tail dependence is an asymptotic concept, defined as a limit of a conditional distribution:

$$\lambda_L = \lim_{z \to 0} \frac{C(z,z)}{z}$$

where λ_L denotes the tail index of the lower tail and $C(u,v)$ is a copula function of the two risks in the portfolio. This contrast between finite and asymptotic does not provide an answer the paradox, but it even makes it even more evident. How could an *asymptotic* measure of risk provide a saving of risk capital at the standard finite percentiles? Is is then good and safer to seek tail dependent risks? Or are finite measures simply misleading and we should design capital requirement on asymptotic measures?

As soon as capital requirements will be computed on finite points of the distribution tail, the puzzle that we point out would survive. In Table 3.1 we give a more detailed account of the instance of the puzzle. The table reports the shape of the two identically distributed risks, X_1, X_2. We assume two cases of marginal distributions: α-stable symmetric distribution with $\alpha = 1.5$ and Student-t distribution with 1.86 degrees of freedom. The two tail indexes of the marginal distributions were chosen to ensure the same marginal VaR measure, equal to 3094 $. Let us notice that the two marginal distributions used in the experiment share the common feature of being moderately heavy-tailed. However, the tails are not heavy enough to violate subadditivity. We see this comparing the marginal VaR and the VaR of the portfolio of the two risks, $Z = X_1 + X_2$. In all cases, we assume elliptical copulas with the same correlation parameter, equal to 0.25. *En passant*, we notice that that the aggregation of α-stable risks costs more capital than Student-t risks. However, as we claimed above, we see that the capital charge is decreasing with lower degrees of freedom of the copulas, from a level of 25 that makes the copula indistinguishable from the Gaussian one, down to 1. This raises the paradox. One presumes that large joint negative outcomes are much more probable when tail-dependence is higher. In the Student-t copula framework, the lower the degrees of freedom parameter ν, the higher this probability, but in spite of this the risk capital computed is lower.

Table 3.1: Tail dependence and VaR: heavy tails marginals

marginals	degrees	$\mathbf{VaR_{1\%}(X_1)}$	Copula DOF	Corr.ρ	$\mathbf{VaR_{1\%}(Z)}$
Stable	$\alpha = 1.5$	3094 \$	1	0.25	4826 \$
Stable	$\alpha = 1.5$	3094 \$	3	0.25	5151 \$
Stable	$\alpha = 1.5$	3094 \$	25	0.25	5336 \$
t-Student	$\beta = 1.86$	3094 \$	1	0.25	4780 \$
t-Student	$\beta = 1.86$	3094 \$	3	0.25	4971 \$
t-Student	$\beta = 1.86$	3094 \$	25	0.25	5069 \$

One may wonder whether this is the exception or the rule. Table 3.2 shows that there is a standard case in which one obtains what he expects, that is higher capital requirements for higher tail dependence. This happens when marginal risks are Gaussian, as it is the case in the Table. Now the aggregate VaR figure increases from 2121 \$ to 2370 \$ as the degrees of freedom of the Student-t copula decrease from 25 down to 1.

Table 3.2: Tail dependence and VaR: Gaussian marginals

marginals	degrees	$\mathbf{VaR_{1\%}(X_1)}$	Copula DoF	Corr.ρ	$\mathbf{VaR_{1\%}(Z)}$
Stable	$\alpha = 2$	1328 \$	1	0.25	2370 \$
Stable	$\alpha = 2$	1328 \$	3	0.25	2243 \$
Stable	$\alpha = 2$	1328 \$	25	0.25	2121 \$

From the two examples above it emerges that a critical condition for the tail dependence paradox to show up is the tail behaviour of the marginals. On one hand, if marginal tails are thick enough, the aggregated risk is higher for weaker tail-dependence (with maximum obtained when the case of zero tail dependence is reached, that is the Gaussian copula). On the other hand, thin marginals allow for "regular" behaviour of the capital requirement, meaning that stronger dependence on the tails induces higher Value-at-Risk.

The relevance of marginal tails for aggregate VaR emerges clearly from the definition of C-convolution, that is the concept of convolution extended to dependent variable. This was first proposed in (Cherubini et al., 2011) and reads

$$F_{X_1+X_2}(z) = \int_0^1 D_1 C\Big(\omega, F(z - F^{-1}(\omega))\Big)\, \mathrm{d}\omega$$

where $C(.,.)$ is the copula function linking X_1 and X_2, D_1 denotes partial derivative with respect to the first argument and $F(x)$ is the probability distribution that is common to the (identically distributed)risks X_1 and X_2.

The z-level aggregated VaR of X_1 and X_2 crucially depends on both copula function $C(.,.)$ and the marginal distribution F, and we have seen in the examples before that tail dependence may play opposite roles for different tail shapes of F. For thin marginal tails, tail dependence increases the risk and the capital requirement, For heavier tails, at some point the paradoxical behaviour arise, and tail dependence sort of moderates risk and the capital required to absorb it.

The puzzle raises a further question. Say that the tail index paradox shows up at the z-percentile. Will it persist for lower percentiles as well? The question, that may seem only theoretical, has an important practical fallout, particularly in the current period in which VaR is being substituted in many applications and regulatory provisions by the alternative *Expected Shortfall*, ES measure (see for example the forthcoming market risk regulation for financial intermediaries called Fundamental Review of the Trading Book, FRTB). It is in fact well known that ES can be written as the integral of VaR measures

$$ES(X_1 + X_2) = -\frac{1}{z}\int_0^z F^{-1}_{X_1+X_2}(u)\,\mathrm{d}u$$

It is then clear that if the tail dependence reduces the risk for all VaR measures in the $(0, z]$ tail, the ES measure is also decreased. The condition is only sufficient because it may happen that the tail dependence paradox disappears at some percentile in the $(0, z]$ region, but this is not enough to compensate for the effect of the paradox in the rest of the region.

The plan of the paper is as follows. Section 3.2 introduces the $C-$convolution and then focus to the case of t-copula. In Section 3.3 we give the analysis on the tail dependence and provide the main results on the Value-at-Risk ranking condition. To confirm the result, Section 3.4 provides a real application on stock financial market. Section 3.5 concludes.

3.2 Preliminaries: Convolution

3.2.1 Copula functions and Dependence

We here recall the basic theory and useful results on copula functions and the non parametric dependence structure of random variables. In what follows we introduce the bivariate case, that we use in this analysis. A complete reference on copula functions is Nelsen (2006), while Frees and Valdez (1998) describes copula simulation methods. A reference on copula methods applied in finance is Cherubini et al. (2004).

Definition 3.2.1 (Copula)**.** A 2-dimensional Copula $\mathcal{C}$ is a function $\mathcal{C} : \mathbb{I} \times \mathbb{I} \to \mathbb{I}$, with $\mathbb{I} = [0, 1]$ the unit space, that satisfies these properties

1. $\mathcal{C}(u_1, 1) = u_2,\ \mathcal{C}(1, u_2) = u_1$
2. $\mathcal{C}(u_1, 0) = \mathcal{C}(0, u_2) = 0$
3. for $u_1 \geq u_2$ and $v_1 \geq v_2$,

$$\mathcal{C}(u_1, v_1) - \mathcal{C}(u_1, v_2) - \mathcal{C}(u_2, v_1) + \mathcal{C}(u_2, v_2) \geq 0$$

Copula functions allow to separate the marginal behaviour of the individual random variables from their dependence structure. Due to the integral probability transform, copula functions single out the dependence structure by reducing the representation of dependence to a joint probability function taking uniform random variables as arguments.

Theorem 3.2.2 (Sklar). *Let F_1, F_2 be the continuous univariate distribution functions of the random variables (X_1, X_2). Then, $H(x_1, x_2)$ is a joint distribution function if and only if* $\forall\,(x_1, x_2) \in \mathbb{R}^2$

$$H(x_1, x_2) = \mathcal{C}\left(F_1(x_1), F_2(x_2)\right) \tag{3.1}$$

where $\mathcal{C}$ is a copula function.

It may be proved that copula functions are naturally linked to the concepts and measures of non parametric association. The maximal copula is defined as $\mathcal{M}(u_1, u_2) = \min\{u_1, u_2\}$, and is known as upper Fréchet bound. This is related to the concept of perfectly positive dependence, or *co-monotonicity*, that is r.v.'s X_1, X_2 such that there exists a strictly increasing function linking one variable to the other. As for perfect negative dependence, the Fréchet bound reads $\mathcal{W}(u_1, u_2) = \max\left\{u_1 + u_2 - 1, 0\right\}$ and it is meaningful only in the bivariate dimension. Finally, it is well known that the independence case is represented by the *Product Copula*: $u_1 u_2$

Copula functions are naturally linked to non parametric association measures such as Spearman's ρ and Kendall's τ.

Proposition 3.2.3. *Suppose (X, Y) have continuous marginal distributions and unique copula $\mathcal{C}$. Then the Kendall's τ are given by*

$$\rho_\tau(X, Y) = 4\int_0^1\int_0^1 \mathcal{C}(u, v)\ \mathrm{d}\mathcal{C}(u, v) - 1$$

and rank correlation measure Spearman's ρ is defined as

$$\rho_S(X, Y) = 12\int_0^1\int_0^1 \mathcal{C}(u, v)\ \mathrm{d}\mathcal{C}(u, v) - 3$$

Copula functions are particularly well suited to represent the dependence structure of extreme events, that is measured by the so called *tail dependence* coefficients, and are the main object of our analysis. The tail index coefficient is linked to the concept of conditional distribution, and can be easily written in terms of copula functions. We have the following definition.

Definition 3.2.4. Let X,Y be two r.v.'s with corresponding continuous d.f.'s F, G. We define the upper tail dependence parameter λ_U as the limit,

$$\lambda_U = \lim_{t \to 1^-} \mathbb{P}\left[Y > G^{(-1)}(t) \middle| X > F^{(-1)}(t)\right] \tag{3.2}$$

Accordingly, the lower tail dependence coefficient λ_L is defined as

$$\lambda_L = \lim_{t \to 0^+} \mathbb{P}\Big[Y \leq G^{(-1)}(t) \Big| X \leq F^{(-1)}(t)\Big] \tag{3.3}$$

Next theorem shows how tail dependence can be expressed in terms of copulas.

Theorem 3.2.5. *Given X, Y, F, G, λ_U, λ_L defined as above in (3.2.4), let $\mathcal{C}$ be the bivariate copula of (X, Y). If the limits (3.2) and (3.3) exist, then tail dependence coefficients can be expressed as follows*

$$\lambda_U = 2 - \lim_{t \to 1^-} \frac{1 - \mathcal{C}(t,t)}{1 - t}$$
$$\lambda_L = \lim_{t \to 0^+} \frac{\mathcal{C}(t,t)}{t}$$

Here below we show the tail dependence formulas for the copulas used in this paper: elliptical (Gaussian and Student's).

Example 3.2.6 (Copulas and corresponding tail dependencies).

- Gaussian copula $C_\rho(u_1, u_2) = \Phi_2(\phi^{-1}(u_1), \Phi^{-1}(u_2), \rho) \qquad \lambda_U = \lambda_L = 0$
  where Φ_2 and Φ are the bivariate and univariate standard Gaussian cdf respectively.

- Student's t-copula $C^\nu_\rho(u_1, u_2) = \mathbf{t}_\nu(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2), \rho, \nu) \qquad \lambda_U = \lambda_L = 2\, t_{\nu+1}\left(\sqrt{\nu+1}\,\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right)$
  where $\mathbf{t}_\nu$ and t_ν are the bivariate and univariate Student's t-cdf respectively.

3.2.2 Convolution of dependent variables

We now formally introduce the representation of the VaR aggregation problem for dependent variables. Under the assumption of independence, the problem of computing the VaR of a sum of risks would merely leads to the standard concept of convolution of random variables. Then, the natural extension of this concept to the case of dependent variables is the definition of *C-convolution* proposed by Cherubini et al. (2011). The term C in the definition reminds of the copula function representing a general relationship of a pair of variables X and Y. The finding was obtained as a by-product of a more general result for the characterization of the dependence structure between the variables X and $X + Y$ from the dependence between X and Y. We report here the main proposition.

Proposition 3.2.7. *Let X, Y be two real-valued random variables on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with* *corresponding copula $C_{X,Y}$ and continuous marginals F_X, F_Y. Then,*

$$C_{X,X+Y}(u,v) = \int_0^u D_1 C_{X,Y}\Big(\omega, F_Y(F_{X+Y}^{-1}(v) - F_X^{-1}(\omega))\Big)\, \mathrm{d}\omega \tag{3.4}$$

$$F_{X+Y}(t) = \int_0^1 D_1 C_{X,Y}\Big(\omega, F_Y(t - F_X^{-1}(\omega))\Big)\, \mathrm{d}\omega \tag{3.5}$$

where $C_{i,j}(u,v)$ denotes the copula functions between the variables reported in the underscore, $D_1C(u,v)$ represents the derivative with respect to u, F_i denotes the distribution function of the variable reported in the underscore(X, Y and $X+Y$).

Notice that in this result we have implicit defined te concept of C-convolution

Definition 3.2.8 (C-convolution)**.** Lef F, H be two continuous c.d.f's and C a copula function. The C-convolution of H and F is defined as the c.d.f.

$$H \overset{C}{*} F(t) = \int_0^1 D_1 C\Big(\omega, F(t - H^{-1}(\omega))\Big)\ \mathrm{d}\omega$$

This result provides a formal definition and an alternative computation approach to the convolution of dependent variables, with respect to the standard Monte Carlo. In some cases, moreover, an important property of the *C-convolution* operator can be used to make the computation easier. In fact, it can be easily proved that the *C-convolution* operator is closed with respect to mixture of copula functions. In other words, it can be shown that if for some bivariate copula functions A and B we have

$$C(u,v) = \lambda A(u,v) + (1-\lambda)B(u,v),\ \forall\ \lambda\ \in\ [0,1]$$

then, for all c.d.f's H, F it holds

$$H \overset{C}{*} F(t) = H \overset{\lambda A + (1-\lambda)B}{*} F = \lambda H \overset{A}{*} F + (1-\lambda) H \overset{B}{*} F$$

Remark 3.2.2.1. Notice that the C-convolution can be formulated in terms of either the first or the second derivative.

$$F_{X+Y}(t) = \int_0^1 D_1 C_{X,Y}\Big(\omega, F_Y(t - F_X^{-1}(\omega))\Big)\, \mathrm{d}\omega = \int_0^1 D_2 C_{X,Y}\Big(F_X(t - F_Y^{-1}(\omega)), \omega\Big)\, \mathrm{d}\omega \tag{3.6}$$

The related choice depends of the conditioning random variable that we take into account. The C-convolution gives the same result if the joint distribution is exchangeable. Notice that in our case the definition of exchangeability involves both the copula function and the marginal distributions, that is $C(u,v) = C(v,u), \forall u, v$ and $F_X = F_Y = F$. This is the case of our analysis.

Example 3.2.9. **Gaussian *C*-convolution**

Recall the Gaussian Copula as

$$C_{Ga}(u_1,u_2;\rho) = \Phi_2(\Phi^{-1}(u_1),\Phi^{-1}(u_2);\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\Phi^{-1}(u_1)}\int_{-\infty}^{\Phi^{-1}(u_2)} \exp\left(-\frac{x_1^2-2\rho x_1x_2+x_2^2}{2(1-\rho^2)}\right)\mathrm{d}x_1\,\mathrm{d}x_2$$

where $x_1 \sim F_{X_1},\ x_2 \sim F_{X_2}$ are r.v.'s such that $u_i = \Phi(x_i), i = 1,2$ and $\Phi_2(\cdot,\cdot,\rho)$ is the cdf of two standard normally distributed random variables with correlation $\rho \in (-1,1), \Phi$ is the cdf of $\mathcal{N}(0,1)$ (standard normal) and Φ^{-1} is the quantile function.

The partial derivative of $C_{Ga}(u_1,u_2;\rho)$ wrt u_2 is

$$D_2C\left(u_1,u_2;\rho\right) = \Phi\left(\frac{\Phi^{-1}(u_1)-\rho\Phi^{-1}(u_2)}{\sqrt{1-\rho^2}}\right)$$

under the result in (3.6) and a change of variable $\omega = F_{X_2}(x_2)$, the Gaussian C-convolution gives

$$\begin{aligned} F_{X_1} \overset{C_{Ga}}{*} F_{X_2}(t) &= \int_{-\infty}^{\infty} \Phi\left(\frac{\Phi^{-1}\left(F_{X_1}(t-x_2)\right)-\rho\Phi^{-1}\left(F_{X_2}(x_2)\right)}{\sqrt{1-\rho^2}}\right)\mathrm{d}F_{X_2}(x_2) \\ &= \int_0^1 \Phi\left(\frac{\Phi^{-1}\left(F_{X_1}(t-F_{X_2}^{-1}(\omega))\right)-\rho\Phi^{-1}(\omega)}{\sqrt{1-\rho^2}}\right)\mathrm{d}\omega \end{aligned} \tag{3.7}$$

Example 3.2.10. **Student's t *C*-convolution**

The bivariate Student's t copula is defined as

$$\begin{aligned} C_{St}(u_1,u_2;\rho,\nu) =& \mathbf{t}_\nu(t_\nu^{-1}(u_1),\ t_\nu^{-1}(u_2),\rho) \\ =& \int_{-\infty}^{t_\nu^{-1}(u_1)}\int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{\Gamma(\frac{\nu+2}{2})}{\Gamma(\frac{\nu}{2})\pi\nu\sqrt{1-\rho^2}}\left(1+\frac{x_1^2+x_2^2-2\rho x_1x_2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}}\mathrm{d}x_1\,\mathrm{d}x_2 \end{aligned} \tag{3.8}$$

where $\Gamma(\cdot)$ is the Gamma function, $\Gamma\left(\frac{\nu+2}{2}\right)/\,\Gamma\left(\frac{\nu}{2}\right) = \frac{\nu}{2}$ and $\mathbf{t}_\nu$ is the bivariate Student's t distribution with ν degrees.

The partial derivative of $C_{St}(u_1,u_2;\rho,\nu)$ wrt u_2 is obtained applying the Leibniz's rule, as

$$D_2C_{St}\left(u_1,u_2;\rho,\nu\right) = t_{\nu+1}\left(\frac{t_\nu^{-1}\left(u_1\right)-\rho\ t_\nu^{-1}\left(u_2\right)}{\sqrt{\frac{\left(\nu+t_\nu^{-1}(u_2)^2\right)(1-\rho^2)}{\nu+1}}}\right) \tag{3.9}$$

as above, applying (3.6) the Student's t C-convolution with ν degree holds

$$F_{X_1} \overset{C^{\nu}_{St}}{*} F_{X_2}(t) = \int_{-\infty}^{\infty} t_{\nu+1}\left(\frac{t_{\nu}^{-1}\left(F_{X_1}(t-x_2)\right) - \rho\, t_{\nu}^{-1}\left(F_{X_2}(x_2)\right)}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}} \right) \mathrm{d}F_{X_2}(x_2)$$

$$= \int_{0}^{1} t_{\nu+1}\left(\frac{t_{\nu}^{-1}\left(F_{X_1}(t-F_{X_2}^{-1}(\omega))\right) - \rho\, t_{\nu}^{-1}(\omega)}{\sqrt{\frac{(\nu+t_{\nu}^{-1}(\omega)^2)(1-\rho^2)}{\nu+1}}} \right) \mathrm{d}\omega \tag{3.10}$$

3.2.3 Student's t representation

We start studying the sign of the derivative of the Student's t-distribution with respect to the degrees of freedom. According to Abramowitz and Stegun (1992) we write the univariate student's t-cumulative distribution in terms of the regularized beta function I

$$t_{\nu}(x) = \begin{cases} 1 - \frac{1}{2} I_{\frac{\nu}{\nu+x^2}}\left(\frac{\nu}{2}, \frac{1}{2}\right) & x > 0 \\ \frac{1}{2} I_{\frac{\nu}{\nu+x^2}}\left(\frac{\nu}{2}, \frac{1}{2}\right) & x \leq 0 \end{cases} \tag{3.11}$$

By definition the regularized beta function holds

$$I_{\frac{\nu}{\nu+x^2}}\left(\frac{\nu}{2}, \frac{1}{2}\right) = \frac{\beta_{\frac{\nu}{\nu+x^2}}\left(\frac{\nu}{2}, \frac{1}{2}\right)}{\beta\left(\frac{\nu}{2}, \frac{1}{2}\right)} = \frac{\int_0^{\frac{\nu}{\nu+x^2}} t^{\frac{\nu}{2}-1}(1-t)^{-\frac{1}{2}}\,\mathrm{d}t}{\beta\left(\frac{\nu}{2}, \frac{1}{2}\right)}$$

where
$$\beta(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}\,\mathrm{d}t$$
$$\beta_y(a,b) = \int_0^y t^{a-1}(1-t)^{b-1}\,\mathrm{d}t$$

are integral representations of the beta and incomplete beta functions respectively.

Following the approach of Dakovic and Czado (2009), the partial derivative of the Student's t-distribution in (3.11) wrt the degree-of-freedom ν can be expressed as follows

$$\frac{\partial}{\partial \nu} t_{\nu}(x) = -\frac{1}{2}\frac{\partial}{\partial \nu} I_{\frac{\nu}{\nu+x^2}}\left(\frac{\nu}{2}, \frac{1}{2}\right) = \mathbf{sgn}(x)\left(\frac{\beta_{\frac{\nu}{\nu+x^2}}\left(\frac{\nu}{2}, \frac{1}{2}\right) \frac{\partial}{\partial \nu}\beta\left(\frac{\nu}{2}, \frac{1}{2}\right)}{2\beta\left(\frac{\nu}{2}, \frac{1}{2}\right)^2} - \frac{\frac{\partial}{\partial \nu}\beta_{\frac{\nu}{\nu+x^2}}\left(\frac{\nu}{2}, \frac{1}{2}\right)}{2\beta\left(\frac{\nu}{2}, \frac{1}{2}\right)} \right), \quad \forall\, x \in \mathbb{R}$$

It follows

$$\frac{\partial}{\partial \nu} t_{\nu}(x) = \mathbf{sgn}(x)\left(\frac{1}{2} I_{\frac{\nu}{\nu+x^2}}\left(\frac{\nu}{2}, \frac{1}{2}\right) \frac{\frac{\partial}{\partial \nu}\beta\left(\frac{\nu}{2}, \frac{1}{2}\right)}{\beta\left(\frac{\nu}{2}, \frac{1}{2}\right)} - \frac{\frac{\partial}{\partial \nu}\beta_{\frac{\nu}{\nu+x^2}}\left(\frac{\nu}{2}, \frac{1}{2}\right)}{2\,\beta\left(\frac{\nu}{2}, \frac{1}{2}\right)} \right), \quad \forall\, x \in \mathbb{R} \quad \text{where} \quad \mathbf{sgn}(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x = 0 \\ -1, & \text{if } x < 0 \end{cases}$$

After few manipulations and the β's property such that $\frac{\partial}{\partial\nu}\beta\left(\frac{\nu}{2},\frac{1}{2}\right)=\frac{1}{2}\beta\left(\frac{\nu}{2},\frac{1}{2}\right)\left(\Psi\left(\frac{\nu}{2}\right)-\Psi\left(\frac{\nu+1}{2}\right)\right)$, it results

$$\frac{\partial}{\partial\nu}t_\nu(x)=\mathbf{sgn}(x)\left(\frac{1}{4}I_{\frac{\nu}{\nu+x^2}}\left(\frac{\nu}{2},\frac{1}{2}\right)\left(\Psi\left(\frac{\nu}{2}\right)-\Psi\left(\frac{\nu+1}{2}\right)\right)-\frac{1}{2\,\beta(\frac{\nu}{2},\frac{1}{2})}\left[|x|\left(\frac{1}{x^2+\nu}\right)^{\frac{\nu+1}{2}}\nu^{\frac{\nu}{2}-1}+\frac{1}{2}\int_0^{\frac{\nu}{\nu+x^2}}t^{\frac{\nu}{2}-1}(1-t)^{-\frac{1}{2}}\ln(t)\,\mathrm{d}t\right]\right),\ \forall\, x\in\mathbb{R} \tag{3.12}$$

Note that, by properties of equation (3.12) there exists a unique point $(x=0)$ where the derivative function on ν changes its sign. Therefore the function is negative (positive) when $x \overset{(>)}{<} 0$.

The Student's t-inverse derivative function on ν is involved in a relationship between the derivative of the t-distribution $\frac{\partial}{\partial\nu}t_\nu(x)$ and the density function $\varphi_\nu(x)$. Given $x=t_\nu^{-1}(u)$, by absolute continuity of the Student's t-distribution, one can consider the equality $t_\nu(t_\nu^{-1}(u))=u,\ \forall\, u\in(0,1)$. This equation allows to compute the total derivative wrt ν.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\nu}\left(t_\nu(t_\nu^{-1}(u))\right)&=\frac{\partial}{\partial\nu}t_\nu(t_\nu^{-1}(u))+\varphi_\nu(t_\nu^{-1}(u))\frac{\partial}{\partial\nu}t_\nu^{-1}(u)=0\Rightarrow\frac{\partial}{\partial\nu}t_\nu(x)+\varphi_\nu(x)\frac{\partial}{\partial\nu}t_\nu^{-1}(u)=0\\&\Rightarrow\frac{\partial}{\partial\nu}t_\nu^{-1}(u)=-\frac{\frac{\partial}{\partial\nu}t_\nu(x)}{\varphi_\nu(x)}\quad\forall\, u\in(0,1),\quad\text{given } x=t_\nu^{-1}(u)\end{aligned} \tag{3.13}$$

It is easy to check that the derivatives w.r.t. ν of the student's t-inverse cdf is positive for all $u<\frac{1}{2}$ and negative for the opposite sign. On the other side, as confirmed in the above formula, the t-cdf derivatives is positive for $z>0$ and negative for $z<0$. As consequence when $z<0$, the higher the degrees of freedom the lower the t-cdf, implying t-cdf never crosses others curves. This is illustrated in Figure 3.1.

Figure 3.1: Comparison between t-cdf and inverse t-cdf for different degrees of freedom $\nu=0.5,1,2,25$. The left side shows the idea of the partial derivative w.r.t. ν. Fixing a point $z<0$, the smaller the DoF the higher the value assumed by t-cdf. The reverse happens when $z>0$ is considered. On the right side, the t-quantile function is drawn. The derivative of the inverse t-cdf is positive(negative) for $q \overset{(>)}{<} \frac{1}{2}$ and increasing (decreasing) as DOF grows up (fall down).

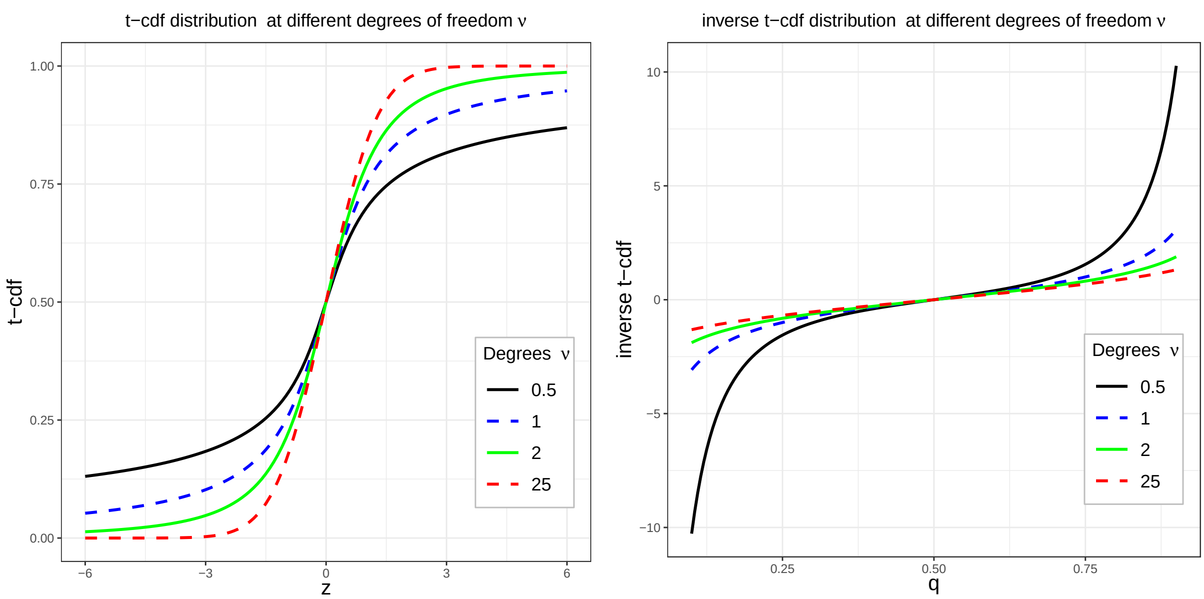
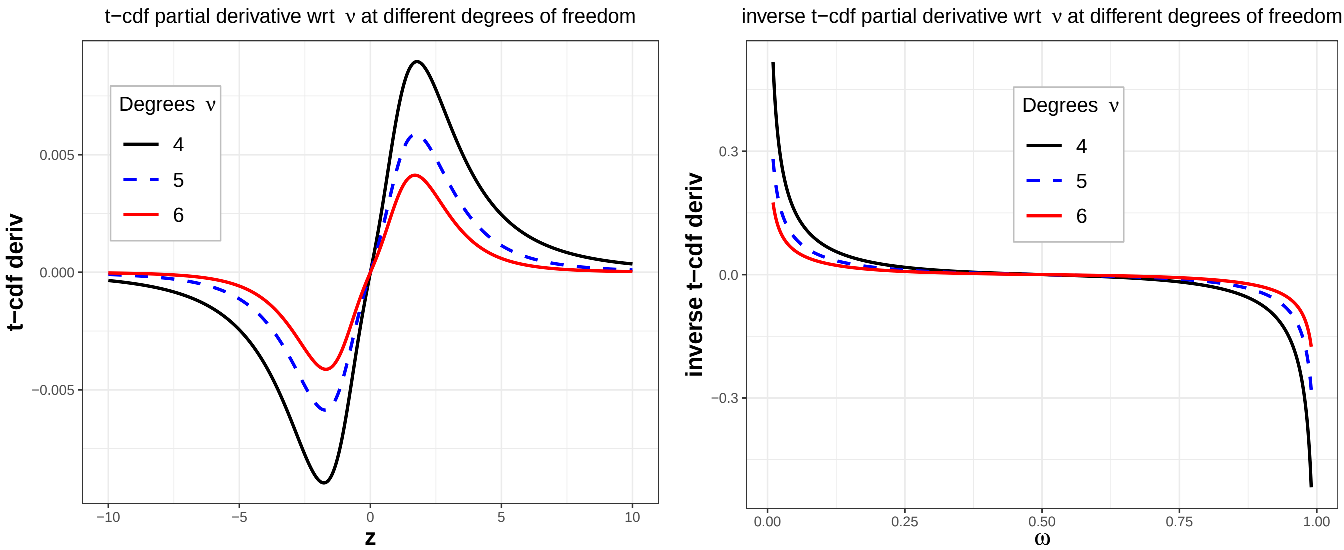


Figure 3.2: Graphics of the t-cdf and inverse t-cdf partial derivatives w.r.t. ν for different degrees of freedom $\nu = 4, 5, 6$. The left side shows the idea of the partial derivative w.r.t. ν. The derivative is negative (positive) for negative (positive) arguments. Furthermore, for $z < 0$ the smaller the DoF the higher the value assumed. The reverse happens when $z > 0$ is considered. The right side takes into account the derivative of the quantile function. The framework show higher value for lower degrees of freedom provided $\omega < \frac{1}{2}$



3.3 Tail-dependence analysis through C-convolution

3.3.1 t-copula convolution distribution

The Gaussian and Student's t convolutions with identically distributed marginals $F_{X_1}, F_{X_2} \sim F$ allow to write

$$F \overset{C_{Ga}}{*} F(z) = \int_0^1 \Phi\left(\frac{\Phi^{-1}\left(F_\alpha(z - F_\nu^{-1}(\omega))\right) - \rho\Phi^{-1}(\omega)}{\sqrt{1-\rho^2}}\right) \mathrm{d}\omega \tag{3.14}$$

$$F \overset{C_{St}^{\nu}}{*} F(z) = \int_0^1 t_{\nu+1}\left(\frac{t_\nu^{-1}\left(F_\alpha(z - F_\nu^{-1}(\omega))\right) - \rho t_\nu^{-1}(\omega)}{\sqrt{\frac{\left(\nu + t_\nu^{-1}(\omega)^2\right)(1-\rho^2)}{\nu+1}}}\right) \mathrm{d}\omega \tag{3.15}$$

The special case of no-correlation, i.e. $\rho = 0$ can be considered. After a change of variable one gets

$$F \overset{C_{Ga}}{*} F(z) = \int_0^1 F(z - F^{-1}(\omega))\,\mathrm{d}\omega = \int_{-\infty}^{+\infty} F(z-x)\,\mathrm{d}F(x) \tag{3.16}$$

$$F \overset{C_{St}^{\nu}}{*} F(z) = \int_0^1 t_{\nu+1}\left(\frac{t_\nu^{-1}\left(F\left(z - F^{-1}(\omega)\right)\right)}{\left(\frac{\nu + t_\nu^{-1}(\omega)^2}{\nu+1}\right)^{\frac{1}{2}}}\right) \mathrm{d}\omega = \int_{-\infty}^{+\infty} t_{\nu+1}\left(\frac{t_\nu^{-1}\left(F(z-x)\right)}{\left(\frac{\nu + t_\nu^{-1}(F(x))^2}{\nu+1}\right)^{\frac{1}{2}}}\right) \mathrm{d}F(x) \tag{3.17}$$

The above formulas confirm that Gaussian C-convolution behaves like standard convolution, that is the case of convolution of independent variables, differently from the Student's t-copula for which the DoF ν preserves asymptotic dependence in the tails. Hence, even in the case of zero-correlation, the t-copula can not be expressed by *product copula*. For the rest, it is straightforward to show that t-copula convolution converges to the Gaussian case as $\nu \to \infty$.

From now on, we assume student's t-marginals, with α, governs the depth of the tail. Then, we compare t-copula convolutions of two identically distributed student's r.v.'s, that with the same tail-index, that we denote α (the same tail index notation that is used for stable distributions), in order to distinguish marginal tails from the degrees of freedom of copulas, that we denote ν. Moreover, since we focus on C-convolutions of i.d. risks linked by Student-t copula, we will often use the term t-convolution.

Figure 3.3 compares two kind of t-marginals, namely $\alpha = 1, \infty$ corresponding to the Cauchy and Gaussian respectively. In the upper panel, notice the structure of the t-convolution: given unimodal, symmetric marginals around zero, the convolution gets perfect symmetry around zero but in opposite DoF order if compared with the univariate Student's t-case shown in Figure 3.1. In this case higher degrees of freedom are associated to lower quantile values.

Let's examine the case $\alpha = \infty$ (bottom-right), where marginals are Gaussian and tails are thin. Here, exists a point where the t-convolution with higher ν parameter crosses any curve with lower ν. Therefore, as the quantile z is sufficiently negative, the picture yields the "expected" quantile ranking induced by tail dependence: the higher the copula degrees of freedom, the lower the quantile values. More to the point, when t-convolution $F^{\nu_1,\rho}_{X+Y}(z)$ crosses the curve of $F^{\nu_2,\rho}_{X+Y}(z)$ the corresponding quantile q_{ν_1} level becomes lower than q_{ν_2}, for all $\nu_1 < \nu_2$. As a consequence, we may conclude that t-convolution of thin-tailed marginals are linked into higher negative quantile values by higher tail-dependence.

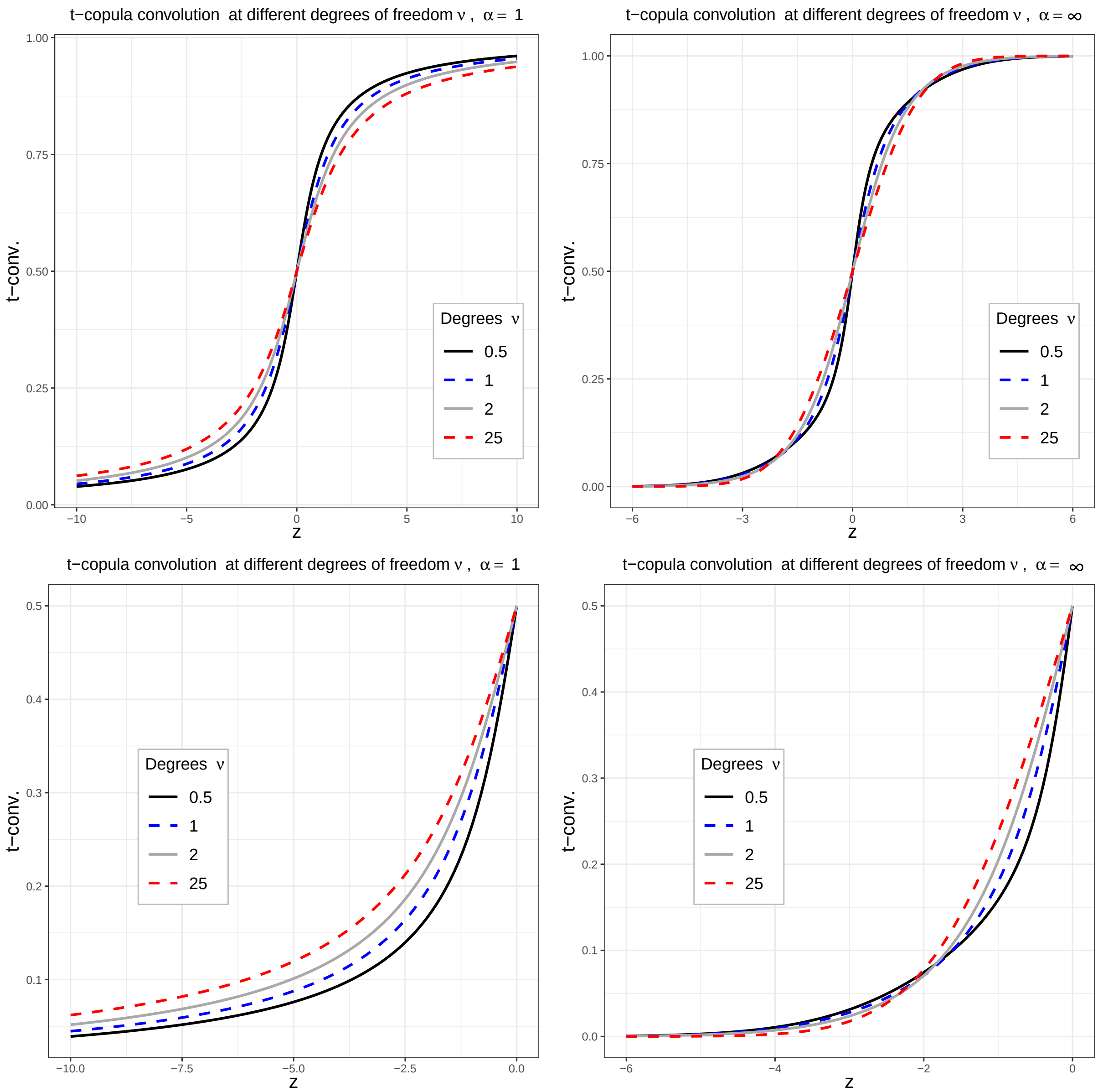
Now, look at t-convolution with degree $\alpha = 1$ (bottom-left) corresponding to the thick-tailed marginals case. Here, it is easy to see that the above order inversion point does not exist any longer. Then, the counter-intuitive case in which lower tail dependence dominates higher tail dependence persists down to the far end of the tail. In this case, t-convolution of thick-tailed marginals gives higher lower tail quantile values for lower tail-dependence.

3.3.2 t-copula convolution derivatives

In what follows, we try an analytical inspection of the tail dependence puzzle investigating the t-copula convolution derivative with respect to ν of two identically distributed r.v.'s with rank correlation ρ. We seek to explain when the DoF t-copula parameter delivers an anomalous quantile order when the marginal distributions have heavy tails. Using univariate Student's t-derivatives (3.13) we can compute the corresponding t-copula convolution derivative w.r.t. ν.

Lemma 3.3.1. *Given the Student's t-copula with ν DoF and univariate t-margins with α DoF, its partial*

Figure 3.3: Comparison between t-convolution distribution with t-marginal at different degrees of freedom $\nu = 0.5, 1, 2, 25$. On left part, it is showed the structure of the t-convolution with heavy tailed marginals ($\alpha = 1$). Fixing a point $z < 0$, the smaller the DoF the higher the cdf value. The inverted framework applies when positive z is considered. On the right side, when gaussian marginals are given, ($\alpha = \infty$), for a sufficient negative z, the structure is reversed and the higher the DoF, the larger the cdf. Bottom figures highlights this structure for negative values.



derivative with respect to ν is given by

$$\frac{\partial}{\partial\nu}\int_0^1 D_2 C_{St}^{\rho,\nu}(u_1,u_2)\,\mathrm{d}\omega = \int_0^1 \frac{\partial}{\partial\nu} t_{\nu+1}\left(\frac{x_1-\rho x_2}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}}\right) + \varphi_{\nu+1}\left(\frac{x_1-\rho x_2}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}}\right)$$
$$\times\left(\frac{\frac{\partial x_1}{\partial\nu}-\rho\frac{\partial x_2}{\partial\nu}}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}} - \frac{1}{2}\frac{x_1-\rho x_2}{\left(\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}\right)^{\frac{3}{2}}}(1-\rho^2)\left[\frac{1+2x_2\frac{\partial x_2}{\partial\nu}}{\nu+1} - \frac{\nu+x_2^2}{(\nu+1)^2}\right]\right)\mathrm{d}\omega$$

where $x_i = t_\nu^{-1}(u_i)$, $i = 1,2$

Proof. See Appendix 3.6.1. □

Figure 3.3 describes the relationship between the speed at which the t-convolution decreases and the degrees of freedom. More to the point, given copula DoF $\nu_1 < \nu_2$ if t-convolution $F_{X+Y}^{\nu_1,\rho}(z)$ decreases faster than $F_{X+Y}^{\nu_2,\rho}(z)$, then the latter dominates the former in distribution and provides higher quantile values. The speed at which a distribution function decreases is given by first derivative, and corresponds to the probability density function. Thus, considering $z < 0$, the higher the density, the faster the decreasing rate.

The C-convolution density function is given by

$$f_{X+Y}(z) = \frac{\partial}{\partial z} F \overset{C}{*} F(z) = \int_0^1 c_{X,Y}\left(F_X(z - F_Y^{-1}(\omega),\omega)\right) f_X\left(z - F_Y^{-1}(\omega)\right)\mathrm{d}\omega \quad \text{where } c_{X,Y}^{\nu}(u_1,u_2) = \frac{\partial^2 C_{X,Y}(u_1,u_2)}{\partial u_1 \partial u_2}$$

In our setting, we compute the t-copula convolution density as

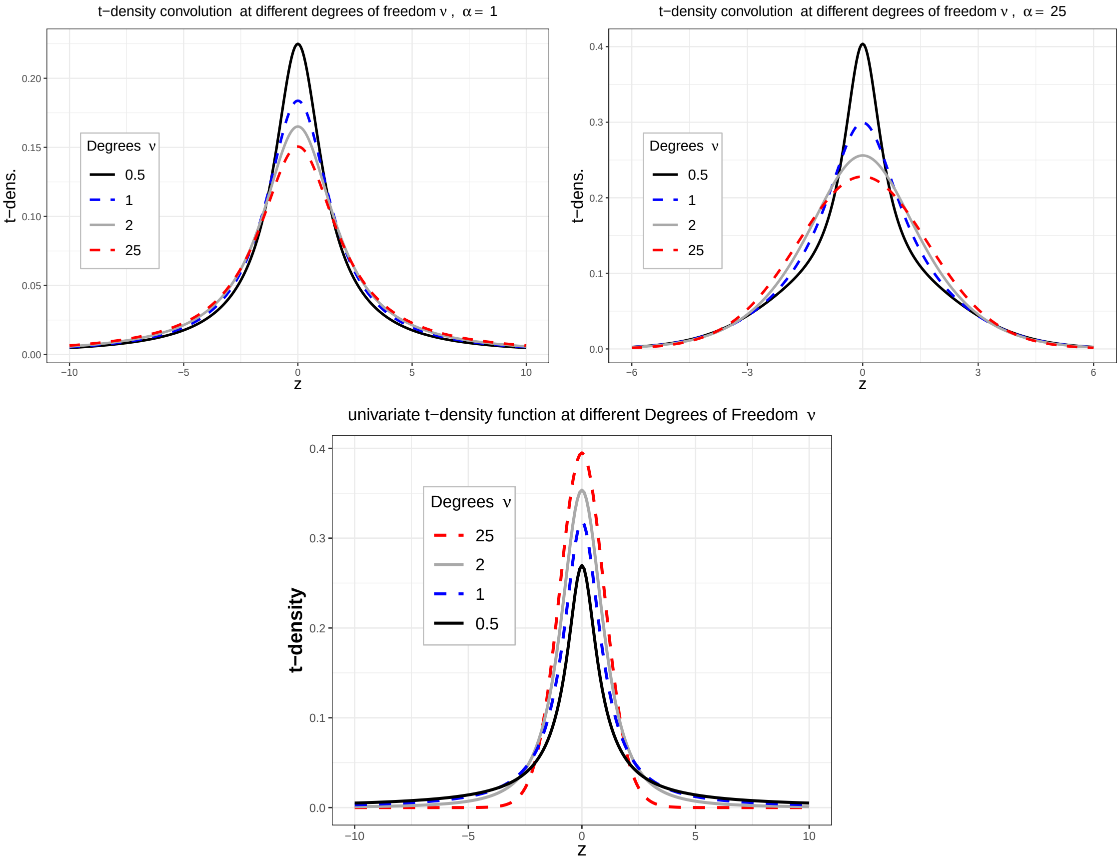
$$f_{X+Y}^{c_{St}^{\nu}}(z) = \int_0^1 c_{X,Y}^{\nu}\left(t_\alpha(z - t_\alpha^{-1}(\omega)),\omega\right) f_\alpha(z - t_\alpha^{-1}(\omega))\,\mathrm{d}\omega \tag{3.18}$$

with the t-copula density defined as

$$c_{X,Y}^{\nu}(u_1,u_2) = \frac{1}{2\pi\sqrt{1-\rho^2}}\frac{A^{-\frac{\nu+2}{2}}}{\varphi_\nu(x_1)\varphi_\nu(x_2)} \quad x_i = t_\nu^{-1}(u_i),\ i=1,2,\quad A = \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)}\right) \tag{3.19}$$

Figure 3.4 compares the densities of univariate Student-t with different DoF with the densities of the t-convolution. We notice that the ranking of the t-convolution density with respect to the degrees of freedom is opposite with respect to that one of the univariate.

Figure 3.4: Comparison between t-convolution density with student's t-marginal and univariate t-density. In the upper panel, we consider either different marginals degrees $\alpha = 1, 25$ and t-copula degrees of freedom $\nu = 0.5, 1, 2, 25$. In left part, the structure of the density assuming heavy-tailed marginals ($\alpha = 1$). Noting that, above a negative cut-off point $z < 0$, the structure shows higher density for higher t-convolution DoF. On the other side, with thin marginals ($\alpha = 25$), the density changes the tendency looking for a second cut-off point, where the structure is reversed. According the standard t-density function (lower panel) the structure on degrees ν is reversed.



3.3.3 Conditions for the inversion of the risk ranking

In this session we investigate the conditions under which the risk ranking that we call tail dependence puzzle shows up and remains verified down to the end of the tail. In other words, we want to know under which conditions if the tail dependence puzzle is observed at the point z in the tail, it is also verified for any $z < u < 0$. We start from a formal definition of the tail dependence puzzle.

Definition 3.3.2 (Odd Ranking Condition)**.**
Given a Student-t copula, we define the *Odd Ranking Condition* (ORC) as the state under which the portfolio riskiness grows-up as t-copula DoF increases, or equivalently, when tail-dependence reduces.

The ORC denotes the inverted risk structure, meaning that higher conditional probability in the tails does not correspond to higher portfolio risk but the converse. In terms of parameter of the t-copula convolution model, it can be stated that ORC exists if $F^{\nu_1,\rho}_{X+Y}(z) < F^{\nu_2,\rho}_{X+Y}(z), \quad \forall\, \nu_1 < \nu_2$.
The next statement guarantees the existence of ORC at sufficiently high percentiles.

Lemma 3.3.3 (Odd Ranking Condition existence)**.** *Assume a t-copula convolution. Then, when z approach zero from below, ORC holds.*

Proof. In Appendix 3.7.3 it is showed that the t-density convolution derivative w.r.t. ν is negative when evaluating at $z = 0$. This suggests near to zero, the speed of decreasing is higher for lower t-convolution DoF ν. Therefore

$$F_{X+Y}^{\nu_l,\rho}(z) < F_{X+Y}^{\nu,\rho}(z), \qquad \forall\ \nu_l < \nu, \qquad \text{when} \quad z \to 0^-$$

□

We now analyse whether the ORC persists further in the tail or is reversed. It clearly all depends on the derivative of the t-convolution with respect to the degrees of freedom. We start observing that this can behave in two possible ways:

1. $\frac{\partial}{\partial \nu} F_{X+Y}^{\nu,\rho}(z)$ becomes negative and converges to the end tail from 0^-
2. $\frac{\partial}{\partial \nu} F_{X+Y}^{\nu,\rho}(z)$ remains positive and converges to end tail from 0^+

It is obvious that this makes the difference for the inversion of the risk ranking: in the first case the derivative remains negative and for any low value ν_1 such that with $\nu_1 < \nu_2$ it is possible to find a value $z^*_{\nu_1}$ such that the corresponding convolution value is higher than the one with parameter ν, that is

$$F_{X+Y}^{\nu_1,\rho}(z^*_{\nu_1}) < F_{X+Y}^{\nu_2,\rho}(z^*_{\nu_2}), \qquad \forall\ \nu_1 < \nu_2$$

In the second case, instead, the risk ranking will remain inverted associating higher quantile values to lower tail dependence.

We then have to prove that the derivative of the t-copula with respect to ν may become negative.

We first recall the partial derivative w.r.t. ν of the t-convolution and show that it can be decomposed into the following sum of integrals

$$\begin{aligned} \frac{\partial}{\partial \nu} F_{X+Y}^{\nu,\rho}(z) = & \int_0^1 \frac{\partial t_{\nu+1}}{\partial \nu} \left(\frac{x_1 - \rho x_2}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}} \right) \mathrm{d}\omega \\ & + \int_0^1 \varphi_{\nu+1} \left(\frac{x_1 - \rho x_2}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}} \right) \times \left[\frac{\frac{\partial x_1}{\partial \nu} - \rho \frac{\partial x_2}{\partial \nu}}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}} - \frac{1-\rho^2}{2} \frac{x_1 - \rho x_2}{\left(\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}\right)^{\frac{3}{2}}} \left(\frac{1 + 2x_2 \frac{\partial x_2}{\partial \nu}}{\nu+1} - \frac{\nu + x_2^2}{(\nu+1)^2} \right) \right] \mathrm{d}\omega \end{aligned} \tag{3.20}$$

where $x_1 = t_\nu^{-1}(t_\alpha(z - t_\alpha^{-1}(\omega)))$, $\quad x_2 = t_\nu^{-1}(\omega)$.

Now, let's consider the case where the derivative may become negative, corresponding to the risk inversion. To see this, we study the mixed partial derivatives. Consider first a framework in which equation (3.20) starts at zero, then takes positive value and, going down on $z < 0$, becomes and remains negative approaching zero from below. In this case the mixed derivative $\frac{\partial^2}{\partial z \partial \nu} F^{\nu,\rho}_{X+Y}(z)$ initially is positive as long as the minimum is reached, then it assumes negative values as far as $z \to \infty$. On the other side, when (3.20) assumes only positive values, it approaches zero from above and the derivative of $\frac{\partial}{\partial \nu} F^{\nu,\rho}_{X+Y}(z)$ is positive for all $z < 0$, corresponding to $\frac{\partial^2}{\partial z \partial \nu} F^{\nu,\rho}_{X+Y}(z)$ always positive.

In can be easily noted that the derivative of (3.20) w.r.t. z corresponds to the t-density derivative on ν.

$$\frac{\partial}{\partial z}\left(\frac{\partial}{\partial \nu} F^{\nu,\rho}_{X+Y}(z)\right) = \frac{\partial}{\partial \nu}\left(\frac{\partial}{\partial z} F^{\nu,\rho}_{X+Y}(z)\right) = \frac{\partial}{\partial \nu}\left(\frac{\partial}{\partial z}\int_0^1 D_2 C^{\nu,\rho}_{St}\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right) \mathrm{d}\omega\right) = \frac{\partial}{\partial \nu} f^{c^{\nu}_{St}}_{X+Y}(z)$$

Furthermore, based on the result in Appendix 3.7.2, the t-convolution density partial derivative w.r.t. ν is stated as follows

$$\begin{aligned}\frac{\partial}{\partial \nu} f^{c^{\nu}_{St}}_{X+Y}(z) &= -\int_0^1 c^{\nu}_{X,Y}\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right)\left[\frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)} + \frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)} + \frac{1}{2}\left(\log A + (\nu+2)\,\frac{\frac{\partial}{\partial \nu} A}{A}\right)\right] f_\alpha(z - t_\alpha^{-1}(\omega))\ \mathrm{d}\omega \\ &= -\int_0^1 c^{\nu}_{X,Y}\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right) G\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right) f_\alpha(z - t_\alpha^{-1}(\omega))\ \mathrm{d}\omega \qquad (3.21)\end{aligned}$$

$$\text{where} \quad x_1 = t_\nu^{-1}(t_\alpha(z - t_\alpha^{-1}(\omega))), \quad x_2 = t_\nu^{-1}(\omega), \quad A = \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)}\right)$$

Thus, the function G is given by

$$G(u_1, u_2) = \frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)} + \frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)} + \frac{1}{2}\left(\log A + (\nu+2)\,\frac{\frac{\partial}{\partial \nu} A}{A}\right) \qquad (3.22)$$

As consequence, negative values on the partial derivatives on ν of the t-convolution density contradicts ORC and so the expected risk structure is obtained, that is, risk is positivity related to the tail dependence. We then now look for conditions that ensure ORC, that is a parametric set that allow positive values only.

Lemma 3.3.4. *Let's assume the t-copula framework. As $G(u_1, u_2) < 0 \Rightarrow$ ORC is satisfied*

Proof. The t-convolution density partial derivative w.r.t. ν (3.21) is made up by 3 parts: the first and the third are the bivariate t-copula density $c^{\nu}_{X,Y}(u_1, u_2)$ with ν DoF and the univariate t-density $\varphi_\alpha(z)$ with α DoF, respectively. Then, they are both positive and the sign of the equation (3.21) is only related to the sign of the $G(u_1, u_2)$ function. □

According to this result, then, it all boils down to study when the function $G(u_1, u_2)$. If we give conditions for negative values, we ensure positiveness of (3.21) and consequently we have that ORC holds.

Since we look for condition on the marginals DoF one can argue that for sufficiently small values on α, i.e. thick-tailed marginals, the risk framework is reversed, so the higher the tail dependence, the lower the Value-at-Risk. Therefore one investigate the sign of $G(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega)$ checking whether small enough marginal α-DoFs allow the function G to take negative values.

Remark 3.3.3.1.

Because we evaluate the function $G(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega)$ outside the integral, one have to deal with the average value of ω that is $\bar{\omega} = \omega(z, \alpha, \nu, \rho)$. This implies that $\bar{\omega}$ is no longer an independent variable but a function of all the variables of the t-copula framework. Since the integral in (3.20) cannot be solved analytically, $\bar{\omega}$ has not an analytical closed form. It follows one can only compute it by means a numerical procedure inverting the integral equation. In any case, since we are in the lower tail we may safely assume $\bar{\omega} < \frac{1}{2}$. Thus, we look at the behaviour of α to examine the change in $t_\alpha^{-1}(\bar{\omega})$. By result, as α decreases, ω increases.

Following the remark, we compute a couple of results of interest on $\bar{\omega}$, that is

$$\frac{\partial \bar{\omega}}{\partial z} > 0 \quad \text{and} \quad \frac{\partial \bar{\omega}}{\partial \alpha} < 0 \tag{3.23}$$

Based on this, we can first prove the existence of ORC in the simplest case $\rho = 0$ and then look at positive correlation ρ that leads to positive G.

Lemma 3.3.5. *Given the t-convolution partial derivative defined in* (3.21) *and assume* $\rho = 0$. *Then for sufficiently negative z,* $G\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right)$ *becomes positive as* $\alpha \to 0$.

Proof. See Appendix 3.7.7 □

The above Lemma implies that, for sufficiently negative z, the decrease of α causes the G function to decreases as well, suggesting that t-convolution density partial derivative rises-up. Because that derivative is not bounded, it exists an α-level such that the derivative assumes only positive values.

Lemma 3.3.6. *Given G defined in* (3.22) *and assume* ρ *positive. Then for sufficiently negative z and for* α *small enough* $G\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right)$ *rises up as* ρ *increases.*

Proof. See Appendix 3.7.8 □

We consider the relationship among the r.v's up to the case of comonotonicity. In fact, we know that if risks are comonotonic, the additivity property of the quantiles reduces the problem to a univariate distribution, and the ORC cannot be reached. The meaning of the Lemma 3.3.6 here is that the G-function increases its value and becomes positive for ρ sufficient large. Furthermore, the effect of the correlation ρ increases the value of α at which G becomes positive, therefore it is sufficient deal with thinner marginal to reject ORC.

At this purpose, in Figure 3.6 we report the behaviour of function G for different levels of correlation and in Figure 3.5 it is showed the derivative on ν of the t-density convolution that contains G and move on on the opposite side: it becomes negative when ORC is denied.

Figure 3.5: t-convolutions derivative w.r.t. ν as function of marginal α-level for different ρ curves. Left side shows the case of $\nu = 1$. Noting that all the curves grows-up constantly up to a maximum. On the right, we deal with $\nu = 25$. The tendency of that curves is slightly decreasing to negative values as $\alpha \to 0$.

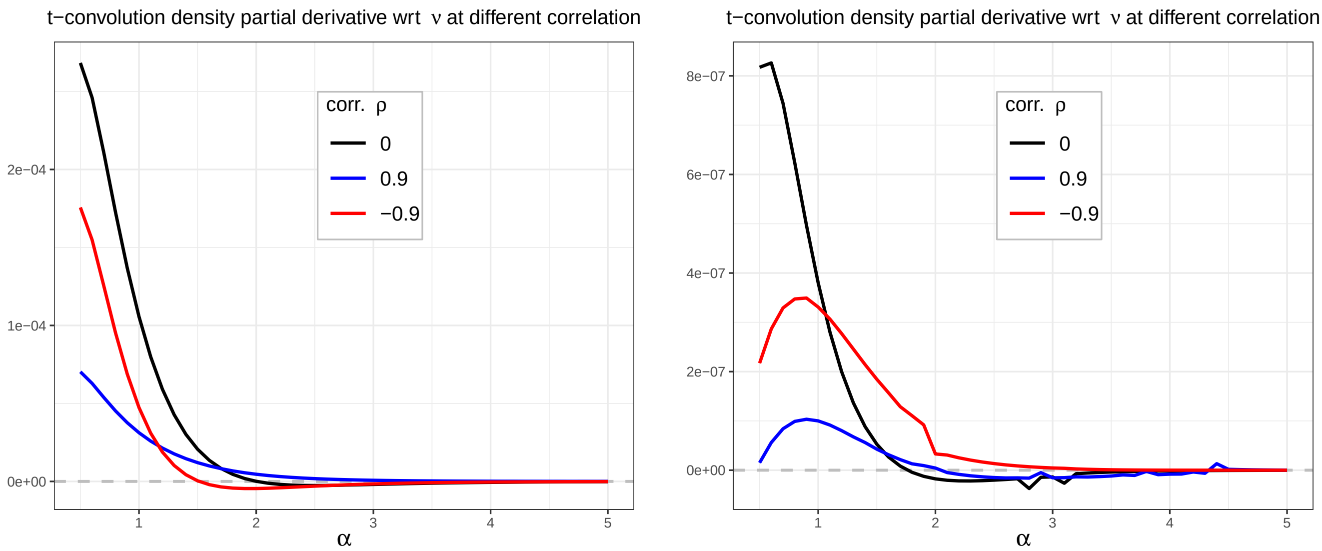
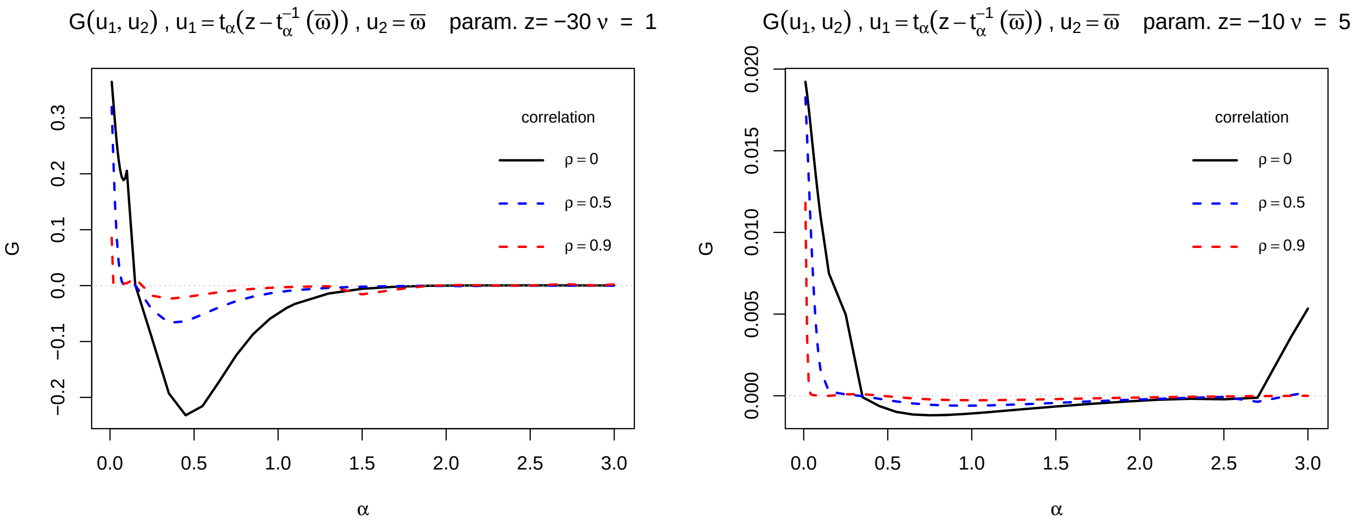


Figure 3.6: $G(u_1, u_2)$ as function of marginal α-level for different ρ curves. As can be seen the function becomes positive as $\alpha \to 0$



In the below tables, we examine t-convolution formula in (3.15). Given two positive real number ν_1, ν_2 corresponding to the DoFs of two t-convolutions on the r.v.'s. $X_1, X_2 \backsim t_\alpha$ with common tail-index α, we look at the zero of the function, that is the point z^* such that

$$F_{X+Y}^{\nu_1,\rho}(z) - F_{X+Y}^{\nu_2,\rho}(z) =$$

$$= \int_0^1 t_{\nu_1+1}\left(\frac{t_{\nu_1}^{-1}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right) - \rho t_{\nu_1}(\omega)}{\sqrt{\frac{\nu_1 + t_{\nu_1}^{-1}(\omega)^2(1-\rho^2)}{\nu_1+1}}}\right) - t_{\nu_2+1}\left(\frac{t_{\nu_2}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right) - \rho t_{\nu_2}(\omega)}{\sqrt{\frac{\nu_2 + t_{\nu_2}^{-1}(\omega)^2(1-\rho^2)}{\nu_2+1}}}\right) \mathrm{d}\omega = 0 \qquad (3.24)$$

Let's consider a negative value close to zero. In that case the ORC is verified as stated above, so it results $F_{X+Y}^{\nu_1,\rho}(z) < F_{X+Y}^{\nu_2,\rho}(z)$. To reject the condition, we decrease the variable z down to a sufficient negative value obtaining, if exists, a value z^* such that either convolutions with higher ν_2 and lower ν_1 DoFs evaluated at that point get the same value. As consequence at z^* we have the same VaR for both the convolution and $F_{X+Y}^{\nu_2,\rho}(z) < F_{X+Y}^{\nu_1,\rho}(z)$ for all $z < z^*$. Therefore one obtains the higher VaR for lower t-convolution DoF and so the ORC is denied because lower DOF are related to higher tail dependence.

In what follows, we show the application of the (3.24) setting various parameters. Results below confirm what we stated. First we observe a clear trend on optimal points. The lower the α, the more negative the z^* and this relationship is valid on the whole space parameters. We focus on Table 3.3 where we look at the inversion of the ORC, considering the couple of t-convolutions with DoFs $\nu_1 = 1$ and $\nu_2 = 50$. So in this setting, the parameter space that denies ORC allows to invert the risk structure on the tail-index interval $(1, 50)$ in which the ORC will not hold. More to the point, we observe when α decreases z decreases as well and this issue is true for all the dependence ρ.

Now, we give a counterexample. Let consider $\alpha = 1$. In that case the optimal is not accomplished so we did not find a z^* that keeps equal the cdf of the two above convolutions. The side effect is that VaR remains higher for lower tail dependence. Since that issue needs a investigation with marginal DoF lower than 1, we fail to reach to the result related to the Lemma 3.3.5 because numerical instability.

Now, we focus on the positive dependence ρ. First one note the role of ρ extends the value of z^*, so the optimal point is farther in the tail. As a matter of fact, one can study the case of $\alpha = 1.5, \rho = 0$, where we haven't found the z^*. But regards higher correlation level we are able to get that optimal point: at $\rho = 0.5 \Rightarrow z = -326.37$ and for $\rho = 0.9 \Rightarrow z = -120.76$. Therefore, as highlighted in the Lemma 3.3.6, the positive dependence allows to find an optimal point with reduced value as ρ increases. As evidence of this issue Figure 3.5 shows the gradual reduction of the t-convolution density derivative on ν when $\rho \to 1$.

Table 3.3: Optimal z^*-search, for various parameters, $\nu_1 = 1, \nu_2 = 50$

ν_1	ν_2	ρ	α	$\mathbf{z}^*$
1	50	0	25	-2.35
1	50	0	15	-2.46
1	50	0	5	-3.22
1	50	0	3	-4.89
1	50	0	2	-784.00
1	50	0	1.5	N.A.
1	50	0	1	N.A.

ν_1	ν_2	ρ	α	$\mathbf{z}^*$
1	50	0.5	25	-2.9
1	50	0.5	15	-3.04
1	50	0.5	5	-4.05
1	50	0.5	3	-6.32
1	50	0.5	2	-17.83
1	50	0.5	1.5	-326.37
1	50	0.5	1	N.A.

ν_1	ν_2	ρ	α	$\mathbf{z}^*$
1	50	0.9	25	-3.64
1	50	0.9	25	-3.83
1	50	0.9	25	-5.23
1	50	0.9	25	-8.38
1	50	0.9	25	-21.44
1	50	0.9	1.5	-120.76
1	50	0.9	1	N.A.

3.4 A Real World Example

Here we give a concrete example of the VaR of a portfolio of two of assets that turned out to be among the riskiest during the 2008 financial crisis. The two stocks are AIG and Freddie Mac. We decompose the change in VaR at the time of the crisis in the marginal risk component and the dependence structure component, splitting the effect of dependence and tail dependence.

Table 3.4 reports the estimates of the marginal distributions of the two stocks and their dependence structure before the crisis (2003-2006) and during the crisis (2007-2009). Both the marginal distributions and the copula functions are Student-t. Concerning dependence, we observe a sizeable increase of of the correlation parameter of the copula function that almost doubles from 0.369 to 0.622. As for tails, we notice a sharp decrease of the degrees of freedom both of the marginals and the copula function. Even though the marginal distributions were showing fat tails already before the crisis, we observe a massive increase of that feature in the crisis period and the degrees of freedom of both the stocks tumble down to values between 1 and 1.5. The same happens for the tail dependence that rises from 0.075 to 0,468 as the joint effect of the increase in correlation and the decrease in the tail index (form above 7 to slightly below 2.

Table 3.4: Marginal distribution and copula estimates

Period	**Marginals**	**Mean**	**St. Dev.**	**DoF**	**Corr.ρ**
2003-2006	AIG	2.32e-05	9.22e-03	3.219	
	FMCC	3.29e-04	0.01	5.134	
	Copula			7.0429	0.369
2007-2009	AIG	1.71e-02	1.99e-02	1.138	
	FMCC	-8.31e-04	0.03103	1.449	
	Copula			1.9608	0.622

Table 3.5 reports the change of VaR in the two periods of an equally weighted portfolio of the two stocks. As it was expected, the VaR rises by more than 10 times, from 600 \$ to over 7 500 \$. In both periods the VaR remains sub-additive, with a ratio of diversified to undiversified VaR that however increases from 0.82 to 0.92. The decomposition of the change in VaR in the changes in the marginal components and in the dependence

structure shows a huge relevance of the former. The VaR would have risen to 7 481 $ instead of 7 507 $ if the dependence structure had remained the same. However, a further decomposition of the dependence structure in correlation and degree of freedom shows that the impact of the change in the copula function is the net effect of the increase in correlation, that would raise the VaR by 308 $, and the decrease in the degrees of freedom, that instead decreased VaR by 282 $. So, the tail dependence effect goes in opposite direction with respect to the increase in correlation and is almost as big as that. The curse of dependence increase in bad times is then eased by the strange blessing of the tail dependence increase.

Table 3.5: VaR Decomposition

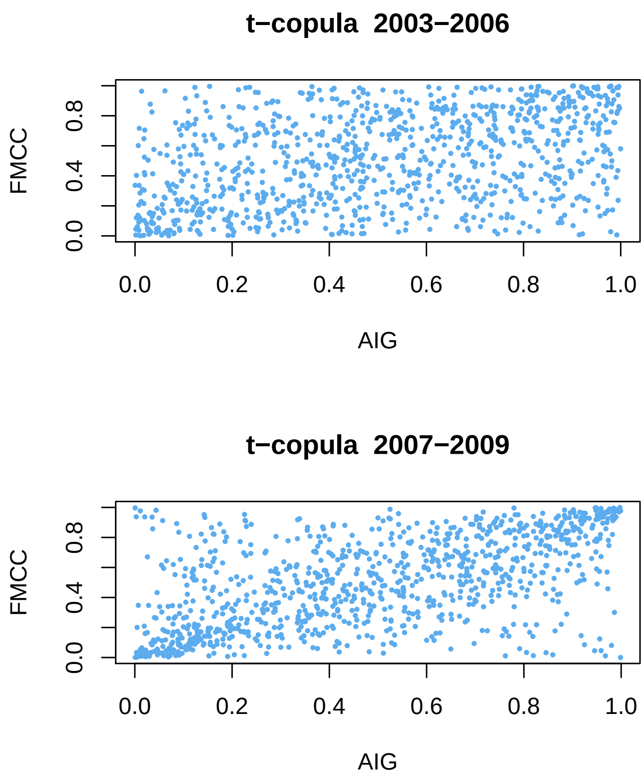
Period	**AIG**	**FMCC**	ρ	ν	Δ-VaR	VaR
2003-2006	396.80	3463	0.369	7.0429		604.16
	4307.12	3811.01	0.369	7.0429	6877.10	7481.26
	4307.12	3811.01	0.622	7.0429	307.72	7788.98
2007-2009	4307.12	3811.01	0.622	1.9608	-282.06	7506.92

Figure 3.7 draws the dependence structure among the two stocks on the early period (upper) and on the crisis (lower). One can easily recognize the transition of the t-copula relationship from lower to higher positive dependence. The graphical representation confirms also the increasing of tail-dependence on lower-left corner. One can highlighted on the drawback of the t-copula model concerning the symmetric tail-dependence. Indeed as can be seen, the model provides also higher dependence in the right-upper corner and tiny association in the countermonotonic cases (upper-left and lower-right corners).

3.5 Conclusions

We documented and explored a new paradox in VaR aggregation. When marginal distribution have sufficiently heavy tails an increase in the tail dependence index may reduce the aggregated Value-at-Risk. This is counterintuitive since one would naturally expect that the increase of the joint probability of extreme events should require more capital. While the paradox is well documented, it is difficult to prove whether it may persist for all finite percentiles in the tail. Here we give the first conditions, obtained by numerical integration to ensure that this may happen.

Figure 3.7: t-copula dependence structure in the period (2003-2006) and (2007-2009)



3.6 Appendix A

3.6.1 The derivative of the t-copula convolution w.r.t. ν

Lemma 3.3.1. *Given the Student's t-copula with ν DoF and univariate t-margins with α DoF, its partial derivative with respect to ν is given by*

$$\frac{\partial}{\partial \nu}\int_0^1 D_2 C_{St}^{\rho,\nu}(u_1,u_2)\,\mathrm{d}\omega = \int_0^1 \frac{\partial}{\partial \nu} t_{\nu+1}\left(\frac{x_1-\rho x_2}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}}\right) + \varphi_{\nu+1}\left(\frac{x_1-\rho x_2}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}}\right)$$
$$\times\left(\frac{\frac{\partial x_1}{\partial \nu}-\rho\frac{\partial x_2}{\partial \nu}}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}} - \frac{1}{2}\frac{x_1-\rho x_2}{\left(\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}\right)^{\frac{3}{2}}}(1-\rho^2)\left[\frac{1+2x_2\frac{\partial x_2}{\partial \nu}}{\nu+1} - \frac{\nu+x_2^2}{(\nu+1)^2}\right]\right)\mathrm{d}\omega$$

where $x_i = t_\nu^{-1}(u_i)$, $i=1,2$

Proof

Let's denote $k_\nu(u_2) = \frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}$, $\quad y = y(u_1,u_2,\nu) = \frac{t_\nu^{-1}(u_1)-\rho t_\nu^{-1}(u_2)}{k_\nu(u_2)^{\frac{1}{2}}}$

Recall the r.v. x_i depends either by the degree-of-freedom and u_i, that is, $x_i = t_\nu^{-1}(u_i), i=1,2$

The total derivative of the h-function of the t-copula depends on the r.v.'s (ν, y). Therefore

$$\frac{\partial}{\partial \nu} D_2 C_{St}^{\rho,\nu}(u_1, u_2) = \frac{\partial}{\partial \nu}\left(t_{\nu+1}(y)\right) = \frac{\partial t_{\nu+1}(y)}{\partial \nu} + \frac{\partial t_{\nu+1}(y)}{\partial y}\frac{\partial y}{\partial \nu}$$

Let's derive $y = y(u_1, u_2, \nu)$ and apply the transformation on $x_i = t_\nu^{-1}(u_i), i = 1, 2$.

$$\frac{\partial y}{\partial \nu} = \frac{\partial}{\partial \nu}\left(\frac{t_\nu^{-1}(u_1) - \rho t_\nu^{-1}(u_2)}{k_\nu(u_2)^{\frac{1}{2}}}\right) = \frac{\frac{\partial}{\partial \nu}x_1 - \rho \frac{\partial}{\partial \nu}x_2}{k_\nu(u_2)^{\frac{1}{2}}} - \frac{(x_1 - \rho x_2)\ \frac{\partial}{\partial \nu}k_\nu(u_2)^{\frac{1}{2}}}{k_\nu(u_2)}$$

$$\tfrac{\partial}{\partial \nu}k_\nu(u_2)^{\frac{1}{2}} = \tfrac{1}{2}k_\nu(u_2)^{-\frac{1}{2}}(1-\rho^2)\left(\frac{1 + 2x_2\frac{\partial x_2}{\partial \nu}}{\nu+1} - \frac{\nu + x_2^2}{(\nu+1)^2}\right)$$

Substituting the latter expression into the former

$$\frac{\partial y}{\partial \nu} = \frac{\frac{\partial}{\partial \nu}x_1 - \rho \frac{\partial}{\partial \nu}x_2}{k_\nu(u_2)^{\frac{1}{2}}} - \frac{1}{2}\frac{x_1 - \rho x_2}{k_\nu(u_2)^{\frac{3}{2}}}(1-\rho^2)\left(\frac{1 + 2x_2\frac{\partial x_2}{\partial \nu}}{\nu+1} - \frac{\nu + x_2^2}{(\nu+1)^2}\right)$$

Finally, summing-up all the parties

$$\frac{\mathrm{d}}{\mathrm{d}\nu}\, t_{\nu+1}\left(y\right) = \frac{\partial t_{\nu+1}(y)}{\partial \nu} + \varphi_{\nu+1}(y)\frac{\partial y}{\partial \nu} = \frac{\partial t_{\nu+1}}{\partial \nu}\left(\frac{x_1 - \rho x_2}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}}\right) + \varphi_{\nu+1}\left(\frac{x_1 - \rho x_2}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}}\right)$$
$$\times\left(\frac{\frac{\partial x_1}{\partial \nu} - \rho\frac{\partial x_2}{\partial \nu}}{\sqrt{\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}}} - \frac{1}{2}\frac{x_1 - \rho x_2}{\left(\frac{(\nu+x_2^2)(1-\rho^2)}{\nu+1}\right)^{\frac{3}{2}}}(1-\rho^2)\left(\frac{1 + 2x_2\frac{\partial x_2}{\partial \nu}}{\nu+1} - \frac{\nu + x_2^2}{(\nu+1)^2}\right)\right)$$

3.6.2 Partial Derivative $\frac{\partial x_1}{\partial z}$ $\quad x_1 = t_\nu^{-1}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right)$

Given $x_1 = t_\nu^{-1}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right)$, with $\omega = \omega(z)$, its partial with respect z is

$$\frac{\partial x_1}{\partial z} = \frac{\partial}{\partial z}\, t_\nu^{-1}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right) = \frac{\frac{\partial}{\partial z}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right)}{\varphi_\nu\left(t_\nu^{-1}(t_\alpha(z - t_\alpha^{-1}(\omega)))\right)}$$
$$= -\frac{\varphi_\alpha\left(z - t_\alpha^{-1}(\omega)\right)\frac{\partial}{\partial z}\left(z - t_\alpha^{-1}(\omega)\right)}{\varphi_\nu\left(t_\nu^{-1}(t_\alpha(z - t_\alpha^{-1}(\omega)))\right)}$$

It follows $\frac{\partial}{\partial z}\left(z - t_\alpha^{-1}(\omega)\right) = 1 - \frac{\frac{\partial \omega}{\partial z}}{\varphi_\alpha(t_\alpha^{-1}(\omega))} = \begin{cases} > 0, & \frac{\partial \omega}{\partial z} < \varphi_\alpha(t_\alpha^{-1}(\omega)) \\ < 0, & \frac{\partial \omega}{\partial z} > \varphi_\alpha(t_\alpha^{-1}(\omega)) \end{cases}$

Since densities are always positive, one obtains $\dfrac{\partial x_1}{\partial z} = \begin{cases} > 0, & \frac{\partial \omega}{\partial z} > \varphi_\alpha(t_\alpha^{-1}(\omega)) \\ < 0, & \frac{\partial \omega}{\partial z} < \varphi_\alpha(t_\alpha^{-1}(\omega)) \end{cases}$

3.6.3 Partial Derivative $\frac{\partial x_1}{\partial \alpha}$ $\quad x_1 = t_\nu^{-1}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right)$

Given $x_1 = t_\nu^{-1}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right)$, its partial with respect α is

$$\frac{\partial x_1}{\partial \alpha} = \frac{\partial}{\partial \alpha} t_\nu^{-1}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right) = \frac{\frac{\partial}{\partial \alpha}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right)}{\varphi_\alpha\left(t_\nu^{-1}(t_\alpha(z - t_\alpha^{-1}(\omega)))\right)}$$

The numerator is $\frac{\partial}{\partial \alpha}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right) = \frac{\mathrm{d}}{\mathrm{d}\alpha}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right) = \frac{\partial}{\partial \alpha} t_\alpha(z - t_\alpha^{-1}(\omega)) + \varphi_\alpha\left(z - t_\alpha^{-1}(\omega)\right)\frac{\partial}{\partial \alpha}(z - t_\alpha^{-1}(\omega))$

By Result in (3.13) one obtains

$$\frac{\partial}{\partial \alpha}\left(z - t_\alpha^{-1}(\omega)\right) = -\frac{\partial}{\partial \alpha} t_\alpha^{-1}(\omega) = \frac{\frac{\partial}{\partial \alpha} t_\alpha\left(t_\alpha^{-1}(\omega)\right)}{\varphi_\alpha\left(t_\alpha^{-1}(\omega)\right)}$$

Rearranging, the expression becomes

$$\frac{\partial x_1}{\partial \alpha} = \frac{\frac{\partial}{\partial \alpha} t_\alpha(z - t_\alpha^{-1}(\omega)) + \varphi_\alpha\left(z - t_\alpha^{-1}(\omega)\right)\frac{\partial}{\partial \alpha}(z - t_\alpha^{-1}(\omega))}{\varphi_\nu\left(t_\nu^{-1}(t_\alpha(z - t_\alpha^{-1}(\omega)))\right)} = \frac{\overbrace{\frac{\partial}{\partial \alpha} t_\alpha(z - t_\alpha^{-1}(\omega))}^{<0 \ \Leftrightarrow \ z < t_\alpha^{-1}(\omega)} + \varphi_\alpha\left(z - t_\alpha^{-1}(\omega)\right)\overbrace{\frac{\frac{\partial}{\partial \alpha} t_\alpha\left(t_\alpha^{-1}(\omega)\right)}{\varphi_\alpha\left(t_\alpha^{-1}(\omega)\right)}}^{<0 \ \Leftrightarrow \omega < 1/2}}{\varphi_\nu\left(t_\nu^{-1}(t_\alpha(z - t_\alpha^{-1}(\omega)))\right)}$$

Under the hypothesis $\omega \leq \frac{1}{2} \Longrightarrow t_\alpha^{-1}(\omega) < 0 \Longrightarrow \frac{\partial}{\partial \alpha} t_\alpha\left(t_\alpha^{-1}(\omega)\right) < 0$ and $\frac{\partial}{\partial \alpha} t_\alpha(z - t_\alpha^{-1}(\omega)) < 0 \Longleftrightarrow z < t_\alpha^{-1}(\omega)$. Finally, we can assert that if a sufficient conditions $z < t_\alpha^{-1}(\omega)$ holds, it follows $\frac{\partial x_1}{\partial \alpha} < 0$.

3.6.4 Derivative of $\frac{\partial x_1}{\partial \nu}$ w.r.t. α

Let $x_1 = t_\nu^{-1}\left(t_\alpha(z - t_\alpha^{-1}(\omega))\right)$. Recall the previous result for the partial derivative w.r.t. α, that is

$$\frac{\partial x_1}{\partial \alpha} = \frac{\frac{\partial}{\partial \alpha} t_\alpha(z - t_\alpha^{-1}(\omega)) + \varphi_\alpha\left(z - t_\alpha^{-1}(\omega)\right)\frac{\frac{\partial}{\partial \alpha} t_\alpha\left(t_\alpha^{-1}(\omega)\right)}{\varphi_\alpha\left(t_\alpha^{-1}(\omega)\right)}}{\varphi_\nu(x_1)}$$

Therefore, one could compute $\frac{\partial}{\partial \alpha}\left(\frac{\partial x_1}{\partial \nu}\right)$ inverting the order of the derivatives as $\frac{\partial}{\partial \nu}\left(\frac{\partial x_1}{\partial \alpha}\right)$. It can be easily seen that the numerator of the above equation does not depend on ν so it can be consider as a constant. Furthermore, when $z < t_\alpha^{-1}(\omega)$ the numerator becomes negative. So under this assumption, and denoting the numerator as K, one gets

$$\frac{\partial}{\partial \nu}\left(\frac{\partial x_1}{\partial \alpha}\right) = \frac{\partial}{\partial \nu}\left(\frac{K}{\varphi_\nu(x_1)}\right) = -\frac{K\,\frac{\partial}{\partial \nu}\left(\varphi_\nu(x_1)\right)}{\varphi_\nu(x_1)^2} = -\frac{\partial x_1}{\partial \alpha}\,\frac{\frac{\partial}{\partial \nu}\left(\varphi_\nu(x_1)\right)}{\varphi_\nu(x_1)}$$

where $\frac{\partial}{\partial \nu}\left(\varphi_\nu(x_1)\right) = \frac{\mathrm{d}}{\mathrm{d}\nu}\varphi_\nu(x_1) = \frac{\partial}{\partial \nu}\varphi_\nu(x_1) + \underbrace{\varphi_\nu'(x_1)\frac{\partial x_1}{\partial \nu}}_{>0}$. Recall that the derivative of the density $\frac{\partial}{\partial \nu}\varphi_\nu(x_1)$ is positive up to a given point $z_\nu^* < 0$ and then becomes negative with second derivative $\frac{\partial^2}{\partial \nu^2}\varphi_\nu(x_1) < 0$, meaning convergence to zero. Finally, we can assert that sufficient condition to obtain $\frac{\partial}{\partial \alpha}\left(\frac{\partial x_1}{\partial \nu}\right) < 0$ is $\frac{\partial}{\partial \nu}\left(\varphi_\nu(x_1)\right) < 0$ and therefore $\frac{\partial}{\partial \nu}\varphi_\nu(x_1) < -\varphi'_\nu(x_1)\frac{\partial x_1}{\partial \nu}$

3.7 Appendix B

3.7.1 Mixed derivative of the t-distribution w.r.t. (ν, α)

$$\frac{\partial}{\partial \nu}\frac{\partial}{\partial \alpha}F_{X+Y}^{\nu,\rho} = \frac{\partial}{\partial \nu}\int_0^1 c\left(\omega, t_\alpha(z - t_\alpha^{-1}(\omega))\right)\frac{\varphi_\alpha(z - t_\alpha^{-1}(\omega))}{\varphi_\alpha(t_\alpha^{-1}(\omega))}\,\mathrm{d}\omega$$

where t_α, φ_α are t-distribution and t-density respectively. The expression of the bivariate Student's t-copula density $c_{X,Y}^\nu(u,v)$ evaluated at point $u = \omega, v = t_\alpha(z - t_\alpha^{-1}(\omega))$ is given by

$$c_{X,Y}^\nu(u,v) = \frac{1}{2\pi\sqrt{1-\rho^2}}\frac{1}{\varphi_\nu(x_1)}\frac{1}{\varphi_\nu(x_2)}A^{-\frac{\nu+2}{2}}$$

where we have defined $x_1 = t_\nu^{-1}(u),\ x_2 = t_\nu^{-1}(v)$ and $A = \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)}\right)$. Recall the univariate t-density

$$\varphi_\nu(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}}\left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \quad \text{where} \quad \Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}\,dt$$

Since the term $\dfrac{\varphi_\alpha(z - t_\alpha^{-1}(\omega))}{\varphi_\alpha(t_\alpha^{-1}(\omega))}$ does not depend on ν, we focus on the derivative of the bivariate t-copula density on ν

$$\frac{\partial}{\partial \nu}c_{X,Y}^\nu(u,v) = \frac{1}{2\pi\sqrt{1-\rho^2}}\frac{\partial}{\partial \nu}\left[\frac{1}{\varphi_\nu(x_1)}\frac{1}{\varphi_\nu(x_2)}A^{-\frac{\nu+2}{2}}\right]$$

To perform the above derivative, we follow two steps

1. The derivative of the reciprocal t-density is $\dfrac{\partial}{\partial \nu}\dfrac{1}{\varphi_\nu(x_i)} = -\dfrac{\frac{\partial}{\partial \nu}\varphi_\nu(x_i)}{\varphi_\nu(x_i)^2},\ i = 1,2$
   The term $\frac{\partial}{\partial \nu}\varphi_\nu(x_i)\,,\ i = 1,2$, it is computed as the total derivative w.r.t. (ν, x_i)

$$\frac{\partial}{\partial \nu}\left(\varphi_\nu(x_i)\right) = \frac{\mathrm{d}\varphi(\nu, x_i)}{\mathrm{d}\nu} = \frac{\partial}{\partial \nu}\varphi_\nu(x_i) + \varphi'_\nu(x_i)\frac{\partial x_i}{\partial \nu} \qquad i = 1,2$$

2. The derivative of $A^{-\frac{\nu+2}{2}}$ is given by $\frac{\partial}{\partial \nu}A^{-\frac{\nu+2}{2}} = -\frac{1}{2}A^{-\frac{\nu+2}{2}}\frac{\partial}{\partial \nu}\left((\nu+2)\log A\right) = -\frac{1}{2}A^{-\frac{\nu+2}{2}}\left(\log A + (\nu+2)\frac{\frac{\partial}{\partial \nu}A}{A}\right)$
   where

$$\frac{\partial}{\partial \nu}A = \frac{2\left(x_1\frac{\partial x_1}{\partial \nu} + x_2\frac{\partial x_2}{\partial \nu} - \rho(x_2\frac{\partial x_1}{\partial \nu} + x_1\frac{\partial x_2}{\partial \nu})\right)}{\nu(1-\rho^2)} - \underbrace{\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu^2(1-\rho^2)}}_{> 0\ \forall\ x_1, x_2} \tag{3.25}$$

Remark 1. Noting that $x_i \partial_\nu x_i < 0, i = 1, 2$. Further, the expression

$$\left[x_1\partial_\nu x_1 + x_2\partial_\nu x_2 - \rho(x_2\partial_\nu x_1 + x_1\partial_\nu x_2)\right] \text{ can be restated as } \left[(x_1 - \rho x_2)\partial_\nu x_1 + (x_2 - \rho x_1)\partial_\nu x_2\right].$$

Under the hypothesis $x_1 = t_\nu^{-1}(u_1) < 0$ and $\forall \rho \geq 0$, one can consider two cases for $x_2 = t_\nu^{-1}(u_2)$

1. $x_2 > 0 \Rightarrow \left[(x_1 - \rho x_2)\partial_\nu x_1 + (x_2 - \rho x_1)\partial_\nu x_2\right] < 0$
2. $x_2 < 0 \Rightarrow \left[(x_1 - \rho x_2)\underbrace{\partial_\nu x_1}_{>0} + (x_2 - \rho x_1)\underbrace{\partial_\nu x_2}_{>0}\right] < 0 \Leftrightarrow 0 \leq \rho < \min\left\{\frac{x_2}{x_1}, \frac{x_1}{x_2}\right\}$

By the derivative product rule, the whole derivative w.r.t. ν holds

$$\frac{\partial}{\partial \nu} c^\nu_{X,Y}(u,v) = \frac{1}{2\pi\sqrt{1-\rho^2}}\left(-\underbrace{\frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)^2}}_{\frac{\partial}{\partial \nu}\frac{1}{\varphi_\nu(x_1)}} \frac{A^{-\frac{\nu+2}{2}}}{\varphi_\nu(x_2)} - \underbrace{\frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)^2}}_{\frac{\partial}{\partial \nu}\frac{1}{\varphi_\nu(x_2)}} \frac{A^{-\frac{\nu+2}{2}}}{\varphi_\nu(x_1)} - \frac{\overbrace{\boxed{\frac{1}{2}A^{-\frac{\nu+2}{2}}\left(\log A + (\nu+2)\frac{\partial_\nu A}{A}\right)}}^{\frac{\partial}{\partial \nu}A^{-\frac{\nu+2}{2}}}}{\varphi_\nu(x_1)\varphi_\nu(x_2)} \right)$$

Rearranging, one obtains

$$\begin{aligned}\tfrac{\partial}{\partial \nu} c^\nu_{X,Y}(u,v) &= -\frac{1}{2\pi\sqrt{1-\rho^2}}\frac{A^{-\frac{\nu+2}{2}}}{\varphi_\nu(x_1)\varphi_\nu(x_2)}\left[\frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)} + \frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)} + \frac{1}{2}\left(\log A + (\nu+2)\frac{\partial_\nu A}{A}\right)\right] \\ &= -c^\nu_{X,Y}(u,v)\left[\frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)} + \frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)} + \frac{1}{2}\left(\log A + (\nu+2)\frac{\partial_\nu A}{A}\right)\right]\end{aligned}$$

Finally the partial derivative of the mixed derivative on (α, ν) holds

$$\begin{aligned}\frac{\partial}{\partial\alpha\partial\nu}F^{\nu,\rho}_{X+Y} = &\int_0^1 \tfrac{\partial}{\partial \nu} c^\nu_{X,Y}\left(\omega, t_\alpha(z - t_\alpha^{-1}(\omega))\right)\frac{\varphi_\alpha(z - t_\alpha^{-1}(\omega))}{\varphi_\alpha(t_\alpha^{-1}(\omega))}\,\mathrm{d}\omega \\ = -\int_0^1 c^\nu_{X,Y}\left(\omega, t_\alpha(z - t_\alpha^{-1}(\omega))\right)&\left[\frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)} + \frac{\frac{\partial}{\partial \nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)} + \frac{1}{2}\left(\log A + (\nu+2)\frac{\partial_\nu A}{A}\right)\right]\frac{\varphi_\alpha(z - t_\alpha^{-1}(\omega))}{\varphi_\alpha(t_\alpha^{-1}(\omega))}\,\mathrm{d}\omega\end{aligned}$$

where $x_1 = t_\nu^{-1}(t_\alpha(z - t_\alpha^{-1}(\omega)))$, $x_2 = t_\nu^{-1}(\omega)$, $A = \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)}\right)$

3.7.2 Derivative of the t-convolution density w.r.t. ν

$$\frac{\partial}{\partial \nu} f^{c^\nu_{St}}_{X+Y}(z) = \frac{\partial}{\partial \nu}\int_0^1 c\left(\omega, t_\alpha(z - t_\alpha^{-1}(\omega))\right) f_\alpha(z - t_\alpha^{-1}(\omega))\,\mathrm{d}\omega$$

where F, f are distribution and density functions respectively. The expression of the bivariate Student's t-copula density $c^{\nu}_{X,Y}(u,v)$ evaluated at point $u=\omega, v=t_\alpha(z-t_\alpha^{-1}(\omega))$ is given by

$$c^{\nu}_{X,Y}(u,v)=\frac{1}{2\pi\sqrt{1-\rho^2}}\,\frac{1}{\varphi_\nu\left(x_1\right)}\,\frac{1}{\varphi_\nu\left(x_2\right)}\,A^{-\frac{\nu+2}{2}}$$

where we have defined $x_1=t_\nu^{-1}(u),\ x_2=t_\nu^{-1}(v)$ and $A=\left(1+\frac{x_1^2+x_2^2-2\rho x_1x_2}{\nu(1-\rho^2)}\right)$. Recall the univariate t-density

$$\varphi_\nu\left(x\right)=\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}\quad\text{where}\quad\Gamma(x)=\int_0^{+\infty}t^{x-1}e^{-t}\,dt$$

Since the term $f_\alpha(z-t_\alpha^{-1}(\omega))$ does not depend on ν, we focus on the derivative of the bivariate t-copula density w.r.t. ν

$$\frac{\partial}{\partial\nu}c^{\nu}_{X,Y}(u,v)=\frac{1}{2\pi\sqrt{1-\rho^2}}\frac{\partial}{\partial\nu}\left(\frac{1}{\varphi_\nu\left(x_1\right)}\,\frac{1}{\varphi_\nu\left(x_2\right)}\,A^{-\frac{\nu+2}{2}}\right)$$

To perform the above derivative, we follow two steps

1. The derivative of the reciprocal t-density is $\dfrac{\partial}{\partial\nu}\dfrac{1}{\varphi_\nu\left(x_i\right)}=-\dfrac{\frac{\partial}{\partial\nu}\varphi_\nu\left(x_i\right)}{\varphi_\nu\left(x_i\right)^2},\ i=1,2$
   The term $\frac{\partial}{\partial\nu}\varphi_\nu\left(x_i\right),\ i=1,2$, it is computed as the total derivative w.r.t. (ν,x_i)

$$\frac{\partial}{\partial\nu}\left(\varphi_\nu\left(x_i\right)\right)=\frac{\mathrm{d}\varphi(\nu,x_i)}{\mathrm{d}\nu}=\frac{\partial}{\partial\nu}\varphi_\nu\left(x_i\right)+\varphi_\nu'(x_i)\frac{\partial x_i}{\partial\nu}\qquad i=1,2 \tag{3.26}$$

2. The derivative of $A^{-\frac{\nu+2}{2}}$ is given by $\frac{\partial}{\partial\nu}A^{-\frac{\nu+2}{2}}=-\frac{1}{2}A^{-\frac{\nu+2}{2}}\frac{\partial}{\partial\nu}\left((\nu+2)\log A\right)=-\frac{1}{2}A^{-\frac{\nu+2}{2}}\left(\log A+(\nu+2)\,\frac{\partial_\nu A}{A}\right)$
   where

$$\frac{\partial}{\partial\nu}A=\frac{2\left(x_1\frac{\partial}{\partial\nu}x_1+x_2\frac{\partial}{\partial\nu}x_2-\rho(x_2\frac{\partial}{\partial\nu}x_1+x_1\frac{\partial}{\partial\nu}x_2)\right)}{\nu(1-\rho^2)}-\underbrace{\frac{x_1^2+x_2^2-2\rho x_1x_2}{\nu^2(1-\rho^2)}}_{>0\ \forall\ x_1,x_2} \tag{3.27}$$

Remark 2. Noting that $x_i\partial_\nu x_i<0, i=1,2$. Further, the expression

$$\left[x_1\partial_\nu x_1+x_2\partial_\nu x_2-\rho(x_2\partial_\nu x_1+x_1\partial_\nu x_2)\right]\text{ can be restated as }\left[(x_1-\rho x_2)\partial_\nu x_1+(x_2-\rho x_1)\partial_\nu x_2\right].$$

Under the hypothesis $x_1=t_\nu^{-1}(u_1)<0$ and $\forall\rho\geq0$, one can consider two cases for $x_2=t_\nu^{-1}(u_2)$

1. $x_2>0\Rightarrow\left[(x_1-\rho x_2)\partial_\nu x_1+(x_2-\rho x_1)\partial_\nu x_2\right]<0$

2. $x_2<0\Rightarrow\left[(x_1-\rho x_2)\underbrace{\partial_\nu x_1}_{>0}+(x_2-\rho x_1)\underbrace{\partial_\nu x_2}_{>0}\right]<0\Leftrightarrow0\leq\rho<\min\left\{\dfrac{x_2}{x_1},\dfrac{x_1}{x_2}\right\}$

By the derivative product rule, the whole derivative w.r.t. ν holds

$$\frac{\partial}{\partial\nu}c^{\nu}_{X,Y}(u,v)=\frac{1}{2\pi\sqrt{1-\rho^2}}\left(-\underbrace{\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)^2}}_{\frac{\partial}{\partial\nu}\frac{1}{\varphi_\nu(x_1)}}\frac{A^{-\frac{\nu+2}{2}}}{\varphi_\nu(x_2)}-\underbrace{\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)^2}}_{\frac{\partial}{\partial\nu}\frac{1}{\varphi_\nu(x_2)}}\frac{A^{-\frac{\nu+2}{2}}}{\varphi_\nu(x_1)}-\frac{\overbrace{\boxed{\frac{1}{2}A^{-\frac{\nu+2}{2}}\left(\log A+(\nu+2)\frac{\partial_\nu A}{A}\right)}}^{\frac{\partial}{\partial\nu}A^{-\frac{\nu+2}{2}}}}{\varphi_\nu(x_1)\varphi_\nu(x_2)}\right)$$

Rearranging, one obtains

$$\begin{aligned}\tfrac{\partial}{\partial\nu}c^{\nu}_{X,Y}(u,v)&=-\frac{1}{2\pi\sqrt{1-\rho^2}}\frac{A^{-\frac{\nu+2}{2}}}{\varphi_\nu(x_1)\varphi_\nu(x_2)}\left[\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)}+\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)}+\frac{1}{2}\left(\log A+(\nu+2)\frac{\partial_\nu A}{A}\right)\right]\\&=-c^{\nu}_{X,Y}(u,v)\left[\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)}+\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)}+\frac{1}{2}\left(\log A+(\nu+2)\frac{\partial_\nu A}{A}\right)\right]\end{aligned}$$

Finally the partial derivative of the t-convolution density holds

$$\begin{aligned}\frac{\partial}{\partial\nu}f^{c^{\nu}_{St}}_{X+Y}(z)=&\int_0^1\tfrac{\partial}{\partial\nu}c^{\nu}_{X,Y}\left(\omega,t_\alpha(z-t_\alpha^{-1}(\omega))\right)f_\alpha(z-t_\alpha^{-1}(\omega))\,\mathrm{d}\omega\\=-&\int_0^1c^{\nu}_{X,Y}\left(\omega,t_\alpha(z-t_\alpha^{-1}(\omega))\right)\left[\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)}+\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)}+\frac{1}{2}\left(\log A+(\nu+2)\frac{\partial_\nu A}{A}\right)\right]f_\alpha(z-t_\alpha^{-1}(\omega))\,\mathrm{d}\omega\\&\text{where}\quad x_1=t_\nu^{-1}(t_\alpha(z-t_\alpha^{-1}(\omega))),\quad x_2=t_\nu^{-1}(\omega),\quad A=\left(1+\frac{x_1^2+x_2^2-2\rho x_1x_2}{\nu(1-\rho^2)}\right)\end{aligned}$$

3.7.3 Derivative of the t-convolution w.r.t. ν at z=0 is negative

Recall the derivative of the t-convolution w.r.t. ν as

$$\frac{\partial}{\partial\nu}f^{c^{\nu}_{St}}_{X+Y}(z)=\int_0^1\left(\tfrac{\partial}{\partial\nu}c_{X,Y}\right)\left(t_\alpha(z-t_\alpha^{-1}(\omega)),\omega\right)\varphi_\alpha(z-t_\alpha^{-1}(\omega))\,\mathrm{d}\omega$$

As already showed the bivariate t-copula derivative is given by

$$\frac{\partial}{\partial\nu}c_{X,Y}(u,v)=-\frac{1}{2\pi\sqrt{1-\rho^2}}\frac{A^{-\frac{\nu+2}{2}}}{\varphi_\nu(x_1)\varphi_\nu(x_2)}\left[\frac{\frac{\partial}{\partial\nu}\varphi_\nu(x_1)}{\varphi_\nu(x_1)}+\frac{\frac{\partial}{\partial\nu}\varphi_\nu(x_2)}{\varphi_\nu(x_2)}+\frac{1}{2}\left(\log A+(\nu+2)\frac{\frac{\partial}{\partial\nu}A}{A}\right)\right]$$

where $x_1=t_\nu^{-1}(t_\alpha(z-t_\alpha^{-1}(\omega))),\ x_2=t_\nu^{-1}(\omega)$ and $A=\left(1+\frac{x_1^2+x_2^2-2\rho x_1x_2}{\nu(1-\rho^2)}\right)$.

We consider the expression of the bivariate t-copula density at $z=0$ and evaluated at $u=\omega,\ v=t_\alpha(-t_\alpha^{-1}(\omega))$.

When the case of zero-symmetric and unimodal distribution functions are considered, it holds $v = t_\alpha(-t_\alpha^{-1}(\omega)) = (1-\omega)\ \forall\ \omega \in (0,1)$. This implies $x_2 = -x_1$ and follows $A = 1 + \frac{2x_2^2}{\nu(1-\rho)}$ and $\frac{\partial}{\partial\nu}A = \frac{4x_2\frac{\partial x_2}{\partial\nu}}{\nu(1-\rho)} - \frac{2x_2^2}{\nu^2(1-\rho)} = \frac{4\nu x_2\frac{\partial}{\partial\nu}x_2 - 2x_2^2}{\nu^2(1-\rho)}$.

Noting that the expression $(\log A + \frac{\nu+2}{A}\frac{\partial}{\partial\nu}A)$ can be negative. More to the point, $A > 1 \Rightarrow \log A > 0$, therefore the term $\frac{\partial}{\partial\nu}A = 2x_2\left(\frac{2\nu\frac{\partial x_2}{\partial\nu} - x_2}{\nu^2(1-\rho)}\right)$ is the only one assuming a negative value. To get evidence of the issue, let's consider $\omega \in (0,1)$ and determine the $\mathbf{sgn}(x_2)$ where $x_2 = t_\nu^{-1}(\omega)$. It is easy to check that $\forall\ \omega \in (0, \frac{1}{2}) \Rightarrow t_\nu^{-1}(\omega) < 0 \Rightarrow \frac{\partial}{\partial\nu}t_\nu^{-1}(\omega) > 0$ and the reverse signs hold $\forall\ \omega \in (\frac{1}{2}, 1)$. Therefore this implies

$$\mathbf{sgn}\left(\frac{\partial}{\partial\nu}A\right) = \begin{cases} \omega > \frac{1}{2} \Rightarrow x_2 > 0 \Rightarrow \frac{\partial}{\partial\nu}x_2 < 0 \Rightarrow (2\nu\frac{\partial}{\partial\nu}x_2 - x_2) < 0 \Rightarrow 2x_2\left(\frac{2\nu\frac{\partial}{\partial\nu}x_2 - x_2}{\nu^2(1-\rho)}\right) < 0 \Rightarrow \mathbf{-1} \\ \omega < \frac{1}{2} \Rightarrow x_2 < 0 \Rightarrow \frac{\partial}{\partial\nu}x_2 > 0 \Rightarrow (2\nu\frac{\partial}{\partial\nu}x_2 - x_2) > 0 \Rightarrow 2x_2\left(\frac{2\nu\frac{\partial}{\partial\nu}x_2 - x_2}{\nu^2(1-\rho)}\right) < 0 \Rightarrow \mathbf{-1} \end{cases}$$

By unimodal and symmetric density properties, when $z = 0$, one gets $\frac{\partial}{\partial\nu}\varphi_\nu(x_2) = \frac{\partial}{\partial\nu}\varphi_\nu(x_2)$. Thus, the whole expression of the derivative w.r.t. ν can be restated as

$$\frac{\partial}{\partial\nu}c_{X,Y}(\omega, t_\alpha(-t_\alpha^{-1}(\omega))) = -\frac{1}{2\pi\sqrt{1-\rho^2}}\frac{A^{-\frac{\nu+2}{2}}}{\varphi_\nu(x_2)^2}\left[\frac{2\frac{\partial}{\partial\nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)} + \frac{1}{2}\left(\log A + (\nu+2)\frac{\frac{\partial}{\partial\nu}A}{A}\right)\right]$$

and if the integrand is positive the partial derivative will be negative. Then, we have to show that the following inequality holds

$$\underbrace{\frac{2\frac{\partial}{\partial\nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)}}_{\text{I PART}} > \underbrace{-\frac{1}{2}\left(\log A + (\nu+2)\frac{\frac{\partial}{\partial\nu}A}{A}\right)}_{\text{II PART}} \tag{3.28}$$

Then, the expression $\frac{\partial}{\partial\nu}(\varphi_\nu(x_2)) = \frac{\partial}{\partial\nu}\varphi_\nu(x_2) + \varphi_\nu'(x_2)\frac{\partial x_2}{\partial\nu}$ in (3.28) is a total derivative computed as

By (3.33) $$\frac{\partial}{\partial\nu}\varphi_\nu(x_2) = \frac{\varphi_\nu(x_2)}{2}\left[\Psi\left(\frac{\nu+1}{2}\right) - \Psi\left(\frac{\nu}{2}\right) - \frac{1}{\nu} + \frac{(\nu+1)x_2^2}{\nu(\nu+x_2^2)} - \log\left(1 + \frac{x_2^2}{\nu}\right)\right]$$

By (3.34) $$\varphi_\nu'(x_2) = -x_2\frac{\nu+1}{\nu+x_2^2}\varphi_\nu(x_2)$$

Rearranging we have

$$
\begin{aligned}
2\frac{\frac{\partial}{\partial\nu}\left(\varphi_\nu\left(x_2\right)\right)}{\varphi_\nu\left(x_2\right)} &= \left[\underbrace{\Psi\left(\frac{\nu+1}{2}\right)-\Psi\left(\frac{\nu}{2}\right)-\frac{1}{\nu}}_{>0}+\frac{(\nu+1)x_2^2}{\nu(\nu+x_2^2)}-\log\left(1+\frac{x_2^2}{\nu}\right)\right]-\frac{x_2(\nu+1)}{\nu+x_2^2}\frac{\partial x_2}{\partial\nu} \\
&= \underbrace{\Psi\left(\frac{\nu+1}{2}\right)-\Psi\left(\frac{\nu}{2}\right)-\frac{1}{\nu}}_{>0}-\log\left(1+\frac{x_2^2}{\nu}\right)+\underbrace{\frac{x_2(\nu+1)}{(\nu+x_2^2)\nu}\left(x_2-2\nu\frac{\partial x_2}{\partial\nu}\right)}_{\text{since } x_2<0 \implies >0}
\end{aligned}
\tag{3.29}
$$

For what concerns the second part in (3.28) and since $x_2<0$, $A=1+\frac{2x_2^2}{\nu(1-\rho)}$, $\frac{\partial}{\partial\nu}A=2x_2\left(\frac{2\nu\frac{\partial x_2}{\partial\nu}-x_2}{\nu^2(1-\rho)}\right)$, we compute the ratio

$$
\frac{\partial A}{\partial\nu}\Big/A=2\frac{x_2}{\nu}\left(\frac{2\nu\frac{\partial x_2}{\partial\nu}-x_2}{\nu(1-\rho)+2x_2^2}\right)<0 \tag{3.30}
$$

Therefore one obtains

$$
\frac{1}{2}\left(\log A+(\nu+2)\frac{\frac{\partial}{\partial\nu}A}{A}\right)=\frac{1}{2}\log\left(1+\frac{2x_2^2}{\nu(1-\rho)}\right)+\frac{\nu+2}{\nu}\left(\frac{2\nu x_2\frac{\partial x_2}{\partial\nu}-x_2^2}{\nu(1-\rho)+2x_2^2}\right) \tag{3.31}
$$

The whole expression is the sum of (3.29) and (3.31). We stated the sum is always positive regardless the values of ν,ρ.

Now, summing the two log terms in (3.29) and (3.31), we have $\log\left(\left(\frac{2x_2^2}{\nu(1-\rho)}\right)^{\frac{1}{2}}\frac{\nu}{\nu+x_2^2}\right)=\log\left(\left(\frac{2x_2^2}{1-\rho}\right)^{\frac{1}{2}}\frac{\sqrt{\nu}}{\nu+x_2^2}\right)$

and summing the following

$$
\frac{\nu+1}{\nu}\left(\frac{x_2^2-2\nu x_2\frac{\partial x_2}{\partial\nu}}{\nu+x_2^2}\right)+\frac{\nu+2}{\nu}\left(\frac{2\nu x_2\frac{\partial x_2}{\partial\nu}-x_2^2}{\nu(1-\rho)+2x_2^2}\right)=\frac{\nu+1}{\nu}\left(\frac{x_2^2-2\nu x_2\frac{\partial x_2}{\partial\nu}}{\nu+x_2^2}\right)-\left(\frac{\nu+1}{\nu}+\frac{1}{\nu}\right)\left(\frac{x_2^2-2\nu x_2\frac{\partial x_2}{\partial\nu}}{\nu(1-\rho)+2x_2^2}\right) \tag{3.32}
$$

One can easily note that (3.32) is positive for all the cases where $\nu(1-\rho)+2x_2^2>\nu+x_2^2\Rightarrow\rho\nu<x_2^2$ holds. All other cases happen when either $|\rho|$ and ν are sufficient large to denied the previous inequality. On the other side, relating to the expression $\log\left(\left(\frac{2x_2^2}{1-\rho}\right)^{\frac{1}{2}}\frac{\sqrt{\nu}}{\nu+x_2^2}\right)$, we obtain an opposite effect, since $x_2=t_\nu^{-1}(\omega)$ for ν increasing, x_2^2 decreases and for $|\rho|$ large enough the log expression becomes positive.

Further recall the the term $\Psi\left(\frac{\nu+1}{2}\right)-\Psi\left(\frac{\nu}{2}\right)-\frac{1}{\nu}>0$ where $\Psi(x)=\frac{\Gamma'(x)}{\Gamma(x)}$ increases faster when x is smaller.

In conclusion, the negative outcome provided by (3.32) is compensated with the opposite sign by either the log expression of the Digamma function Ψ that produce a positive value for the whole derivative (3.28).

3.7.4 Derivative of the t-density function w.r.t. ν

Recall the univariate t-density function as

$$\varphi_\nu(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \quad \text{where} \quad \Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}\, dt$$

The partial derivative w.r.t. ν is given by the sum of the following three expressions

$$\frac{\partial}{\partial\nu}\varphi_\nu(x) = \frac{\partial}{\partial\nu}\left(\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\right)\frac{1}{\sqrt{\pi\nu}}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} + \frac{\partial}{\partial\nu}\left(\frac{1}{\sqrt{\pi\nu}}\right)\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} + \frac{\partial}{\partial\nu}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}\frac{1}{\sqrt{\pi\nu}}\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}$$

1.

$$\begin{aligned}
\frac{\partial}{\partial\nu}\left(\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\right) &= \frac{1}{2}\left(\frac{\Gamma'(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} - \frac{\Gamma(\frac{\nu+1}{2})\Gamma'(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2})^2}\right) = \frac{1}{2}\left(\frac{\Gamma'(\frac{\nu+1}{2})\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu+1}{2})\Gamma(\frac{\nu}{2})} - \frac{\Gamma(\frac{\nu+1}{2})\Gamma'(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2})\Gamma(\frac{\nu}{2})}\right) \\
&= \frac{1}{2}\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\left[\Psi\left(\frac{\nu+1}{2}\right) - \Psi\left(\frac{\nu}{2}\right)\right] \text{ where } \Psi(x) = \frac{\Gamma'(x)}{\Gamma(x)} \text{ is the Digamma function}
\end{aligned}$$

2. $\displaystyle \frac{\partial}{\partial\nu}\frac{1}{\sqrt{\pi\nu}} = \frac{\partial}{\partial\nu}(\pi\nu)^{-\frac{1}{2}} = -\tfrac{1}{2}(\pi\nu)^{-\frac{3}{2}}\pi = -\tfrac{1}{2}(\pi\nu)^{-\frac{1}{2}}(\pi\nu)^{-1}\pi = -\frac{1}{2\nu}\frac{1}{\sqrt{\pi\nu}}$

3. $\displaystyle \frac{\partial}{\partial\nu}\left(1+\tfrac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} = -\left(1+\tfrac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}\frac{1}{2}\left(\log(1+\tfrac{x^2}{\nu}) + (\nu+1)\frac{-\frac{x^2}{\nu^2}}{1+\frac{x^2}{\nu}}\right) = -\frac{1}{2}\left(\log(1+\tfrac{x^2}{\nu}) - \tfrac{(\nu+1)x^2}{\nu(\nu+x^2)}\right)\left(1+\tfrac{x^2}{\nu}\right)^{-}$

Summing all the three derivative expressions

$$\begin{aligned}
\frac{\partial}{\partial\nu}\varphi_\nu(x) &= \frac{1}{2}\Bigg[\underbrace{\Psi\left(\frac{\nu+1}{2}\right) - \Psi\left(\frac{\nu}{2}\right)}_{>0}\Bigg]\underbrace{\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}\frac{1}{\sqrt{\pi\nu}}\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}}_{\varphi_\nu(x)} - \frac{1}{2\nu}\underbrace{\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}\frac{1}{\sqrt{\pi\nu}}\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}}_{\varphi_\nu(x)} \\
&\quad - \frac{1}{2}\left(\log\left(1+\frac{x^2}{\nu}\right) - \frac{(\nu+1)x^2}{\nu(\nu+x^2)}\right)\underbrace{\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}\frac{1}{\sqrt{\pi\nu}}\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}}_{\varphi_\nu(x)} \\
&= \frac{\varphi_\nu(x)}{2}\left[\Psi\left(\frac{\nu+1}{2}\right) - \Psi\left(\frac{\nu}{2}\right) - \frac{1}{\nu} + \frac{(\nu+1)x^2}{\nu(\nu+x^2)} - \log\left(1+\frac{x^2}{\nu}\right)\right] \qquad (3.33)
\end{aligned}$$

Noting that $\Psi\left(\frac{\nu+1}{2}\right) - \Psi\left(\frac{\nu}{2}\right) - \frac{1}{\nu} > 0$

3.7.5 First derivative of the t-density function

Recall the univariate t-density function as

$$\varphi_\nu\left(z\right)=\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}}\left(1+\frac{z^2}{\nu}\right)^{-\frac{\nu+1}{2}} \quad \text{where} \quad \Gamma(x)=\int_0^{+\infty}t^{x-1}e^{-t}\,dt$$

The partial derivative w.r.t. z $\frac{\partial}{\partial z}\varphi_\nu\left(z\right)$ is given by the following expression

$$\varphi'_\nu(z)=-\frac{\nu+1}{2}\left(1+\frac{z^2}{\nu}\right)^{-\frac{\nu+3}{2}}\frac{2z}{\nu}\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\frac{1}{\sqrt{\pi\nu}}=-z\underbrace{\frac{\nu+1}{\nu\sqrt{\pi\nu}}\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}\left(1+\frac{z^2}{\nu}\right)^{-\frac{\nu+3}{2}}}_{>0}=-z\frac{\nu+1}{(\nu+z^2)}\varphi_\nu(z) \tag{3.34}$$

As expected, it results that the derivative is positive when $z<0$, equal to zero at $z=0$ and negative otherwise.

3.7.6 Mixed derivative of the t-density function w.r.t. (z,ν)

The partial derivative of the t-density function w.r.t. ν is given by

$$\frac{\partial}{\partial\nu}\varphi_\nu\left(z\right)=\frac{\varphi_\nu\left(z\right)}{2}\left[\Psi\left(\frac{\nu+1}{2}\right)-\Psi\left(\frac{\nu}{2}\right)-\frac{1}{\nu}+\frac{(\nu+1)z^2}{\nu(\nu+z^2)}-\log\left(1+\frac{z^2}{\nu}\right)\right]$$

We compute the derivative of the above expression w.r.t. z, namely the partial derivative of the student's t first derivative

$$\frac{\partial}{\partial\nu}\varphi'_\nu(z) = \frac{\partial}{\partial z}\frac{\partial}{\partial\nu}\varphi_\nu(z) = \frac{\partial}{\partial z}\left(\frac{\varphi_\nu(z)}{2}\left[\Psi\left(\frac{\nu+1}{2}\right) - \Psi\left(\frac{\nu}{2}\right) - \frac{1}{\nu} + \frac{(\nu+1)z^2}{\nu(\nu+z^2)} - \log\left(1+\frac{z^2}{\nu}\right)\right]\right)$$

$$= \frac{\varphi'_\nu(z)}{2}\left[\Psi\left(\frac{\nu+1}{2}\right) - \Psi\left(\frac{\nu}{2}\right) - \frac{1}{\nu} + \frac{(\nu+1)z^2}{\nu(\nu+z^2)} - \log\left(1+\frac{z^2}{\nu}\right)\right]$$

$$+ \frac{\varphi_\nu(z)}{2}\left[\frac{\partial}{\partial z}\left(\frac{(\nu+1)z^2}{\nu(\nu+z^2)}\right) - \frac{\partial}{\partial z}\log\left(1+\frac{z^2}{\nu}\right)\right]$$

$$\frac{\partial}{\partial z}\left(\frac{(\nu+1)z^2}{\nu(\nu+z^2)}\right) = \frac{(\nu+1)2z}{\nu(\nu+z^2)} - \frac{\nu(\nu+1)2z^3}{\nu^2(\nu+z^2)^2}$$

$$= \frac{2z(\nu+1)(\nu+z^2) - 2z^3(\nu+1)}{\nu(\nu+z^2)^2} = \frac{2z\nu(\nu+1)}{\nu(\nu+z^2)^2} = \frac{2z(\nu+1)}{(\nu+z^2)^2}$$

$$\frac{\partial}{\partial z}\log\left(1+\frac{z^2}{\nu}\right) = \frac{2z}{\nu+z^2}$$

$$\frac{\partial}{\partial z}\left(\frac{(\nu+1)z^2}{\nu(\nu+z^2)} - \log\left(1+\frac{z^2}{\nu}\right)\right) = \frac{2z(\nu+1)}{(\nu+z^2)^2} - \frac{2z}{\nu+z^2} = \frac{2z(\nu+1) - 2z(\nu+z^2)}{(\nu+z^2)^2} = \frac{2z(1-z^2)}{(\nu+z^2)^2}$$

substituting the expression for the first derivative and the above difference, it holds

$$\frac{\partial}{\partial\nu}\varphi'_\nu(z) = -\frac{\varphi_\nu(z)}{2}\frac{z(\nu+1)}{(\nu+z^2)}\left[\Psi\left(\frac{\nu+1}{2}\right) - \Psi\left(\frac{\nu}{2}\right) - \frac{1}{\nu} + \frac{(\nu+1)z^2}{\nu(\nu+z^2)} - \log\left(1+\frac{z^2}{\nu}\right)\right] + \varphi_\nu(z)\frac{z(1-z^2)}{(\nu+z^2)^2} \tag{3.35}$$

$$= z\frac{\varphi_\nu(z)}{\nu+z^2}\left[\underbrace{\frac{1-z^2}{\nu+z^2}}_{\text{I TERM}} - \frac{\nu+1}{2}\left(\underbrace{\Psi\left(\frac{\nu+1}{2}\right) - \Psi\left(\frac{\nu}{2}\right) - \frac{1}{\nu}}_{\text{II TERM} > 0} + \underbrace{\frac{(\nu+1)z^2}{\nu(\nu+z^2)} - \log\left(1+\frac{z^2}{\nu}\right)}_{\text{III TERM}}\right)\right] \tag{3.36}$$

The first term in square bracket has derivative $\frac{\partial}{\partial z}(\frac{1-z^2}{\nu+z^2}) = -z\frac{2(\nu+1)}{(1+z^2)^2} > 0, \ \forall z < 0$. We compare this result with the corresponding derivative of the term provided by (3.35) adjusted by a negative factor $-\frac{\nu+1}{2} \times z\frac{2(z^2-1)}{(\nu+z^2)^2}$ that reverse the sign. It easy to check that the first term has a decreasing order of $o(z^3)$ higher than the third one $o(z)$, therefore, for sufficiently negative z, it decreases faster that the other one, so the combined effect brings to a positive value for the whole bracket expression. Finally, taking into account that the second term is a positive constant w.r.t. z, one obtains that for a sufficient negative z, the mixed derivative $\frac{\partial}{\partial\nu}\varphi'_\nu(z)$ becomes and remains negative.

3.7.7 The sign of $G(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega)$ for $\alpha \to 0$, not correlated case $\rho = 0$

In what follows, we compute the sign of the function G through the evaluation of the t-convolution density derivative wrt ν in (3.21) given by

$$\frac{\partial}{\partial \nu} f_{X+Y}^{c_{St}^{\nu}}(z) = -\int_0^1 c_{X,Y}^{\nu}\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right)\left[\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)} + \frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)} + \frac{1}{2}\left(\log A + (\nu+2)\,\frac{\frac{\partial}{\partial\nu}A}{A}\right)\right] f_\alpha(z - t_\alpha^{-1}(\omega))\;\mathrm{d}\omega$$

$$= -\int_0^1 c_{X,Y}^{\nu}\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right) G\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right) f_\alpha(z - t_\alpha^{-1}(\omega))\;\mathrm{d}\omega \tag{3.37}$$

$$\text{where}\quad x_1 = t_\nu^{-1}(t_\alpha(z - t_\alpha^{-1}(\omega))), \quad x_2 = t_\nu^{-1}(\omega), \quad A = \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)}\right)$$

where the expression of $G(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega)$ is given in (3.22) as

$$G(u_1, u_2) = \underbrace{\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)} + \frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)}}_{PART\ I} + \underbrace{\frac{1}{2}\left(\log A + (\nu+2)\,\frac{\frac{\partial}{\partial\nu}A}{A}\right)}_{PART\ II} \tag{3.38}$$

We look at the sensitivity of (3.37) with respect to α, especially when α is small, close to zero.

Let's consider the upper bound of α by the limit

$$\lim_{\alpha\to 0}\frac{\partial}{\partial \nu} f_{X+Y}^{c_{St}^{\nu}}(z) = -\int_0^1 \lim_{\alpha\to 0}\; c_{X,Y}^{\nu}\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right) G\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right) f_\alpha(z - t_\alpha^{-1}(\omega))\;\mathrm{d}\omega \tag{3.39}$$

To solve the limit, we focus on the third term. It is useful to know that $f_\alpha(z - t_\alpha^{-1}(\omega))$ approaches to the *Dirac Delta* function as the marginal tail index goes to zero.

Therefore approaching $\alpha \to 0$ (3.37) becomes

$$\lim_{\alpha\to 0}\frac{\partial}{\partial \nu} f_{X+Y}^{c_{St}^{\nu}}(z) = -\int_0^1 c_{X,Y}^{\nu}\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right) G\left(t_\alpha(z - t_\alpha^{-1}(\omega)), \omega\right) \delta(z - t_\alpha^{-1}(\omega))\;\mathrm{d}\omega \tag{3.40}$$

where $\delta(x)$ is the Dirac Delta function. First we set a change of variable $y = t_\alpha^{-1}(\omega)$ on equation (3.40), obtaining

$$\lim_{\alpha\to 0}\frac{\partial}{\partial \nu} f_{X+Y}^{c_{St}^{\nu}}(z) = -\int_{-\infty}^{\infty} c_{X,Y}^{\nu}\left(t_\alpha(z-y), t_\alpha(y)\right) G\left(t_\alpha(z-y), t_\alpha(y)\right) \delta(z-y)\;\mathrm{d}y \tag{3.41}$$

Then, we apply the *Sifting* property of the Delta Dirac, meaning that, for any function $f(x)$ continuous at x_0, we have $\int_{-\infty}^{\infty} f(x)\delta(x - x_0)\,\mathrm{d}x = f(x_0)$, see Mack (2008). Therefore, by *Symmetry* and *Sifting* properties of the

Delta Dirac function, one obtains

$$\lim_{\alpha\to 0}\frac{\partial}{\partial\nu}f^{c^{\nu}_{St}}_{X+Y}(z) = -\int_{-\infty}^{\infty} c^{\nu}_{X,Y}\left(t_\alpha(z-y),t_\alpha(y)\right)G\left(t_\alpha(z-y),t_\alpha(y)\right)\delta(y-z)\ \mathrm{d}y$$
$$= -c^{\nu}_{X,Y}\left(t_\alpha(0),t_\alpha(z)\right)G\left(t_\alpha(0),t_\alpha(z)\right)$$
$$= -c^{\nu}_{X,Y}\left(\tfrac{1}{2},\tfrac{1}{2}\right)G\left(\tfrac{1}{2},\tfrac{1}{2}\right) \tag{3.42}$$

where the last identity is true because the univariate Student's t-cdf $t_\alpha(x)$ approaches $\frac{1}{2}$ as $\alpha\to 0$ for any finite value x.

Recall the t-copula density is positive for all values, so $c^{\nu}_{X,Y}\left(\frac{1}{2},\frac{1}{2}\right)>0$. Concerning the G-function, we have $u_1=u_2=\frac{1}{2}$ and therefore (3.38) changes into

$$G\left(\tfrac{1}{2},\tfrac{1}{2}\right) = \sum_{i=1}^{2}\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_i))}{\varphi_\nu(x_i)}+\frac{1}{2}\left(\log A+(\nu+2)\,\frac{\frac{\partial}{\partial\nu}A}{A}\right),\qquad x_i=t_\alpha^{-1}(\tfrac{1}{2}),\quad i=1,2$$
$$= 2\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(0))}{\varphi_\nu(0)} \tag{3.43}$$

Because $x_1=x_2=t_\alpha^{-1}(\frac{1}{2})=0$, it follows

$$A=\left(1+\frac{x_1^2+x_2^2-2\rho x_1x_2}{\nu(1-\rho^2)}\right)=1 \qquad \text{and}$$
$$\tfrac{\partial}{\partial\nu}A=\frac{2\left(x_1\frac{\partial x_1}{\partial\nu}+x_2\frac{\partial x_2}{\partial\nu}-\rho(x_2\frac{\partial x_1}{\partial\nu}+x_1\frac{\partial x_2}{\partial\nu})\right)}{\nu(1-\rho^2)}-\frac{x_1^2+x_2^2-2\rho x_1x_2}{\nu^2(1-\rho^2)}=0.$$

Since $\frac{\partial}{\partial\nu}(\varphi_\nu(0))=\frac{\partial}{\partial\nu}\varphi_\nu(0)>0$, it results that G-function assumes positive values for all $\nu>0$.

3.7.8 The sign of $G(t_\alpha(z-t_\alpha^{-1}(\omega)),\omega)$ for $\alpha\to 0$, correlated case $\rho>0$

Recall the expression of $G(t_\alpha(z-t_\alpha^{-1}(\omega)),\omega)$ in (3.22) is given by

$$G(u_1,u_2)=\underbrace{\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_1))}{\varphi_\nu(x_1)}+\frac{\frac{\partial}{\partial\nu}(\varphi_\nu(x_2))}{\varphi_\nu(x_2)}}_{PART\ I}+\underbrace{\frac{1}{2}\left(\log A+(\nu+2)\,\frac{\frac{\partial}{\partial\nu}A}{A}\right)}_{PART\ II} \tag{3.44}$$
$$\text{where}\quad x_1=t_\nu^{-1}(t_\alpha(z-t_\alpha^{-1}(\omega))),\quad x_2=t_\nu^{-1}(\omega),\quad A=\left(1+\frac{x_1^2+x_2^2-2\rho x_1x_2}{\nu(1-\rho^2)}\right)$$

We look at the sensitivity of (3.44) with respect to α close to zero, when the correlation $\rho\to 1$. In that case, the value of the correlation involves the second part of the expression (3.44) related to the term A.

Therefore we consider $\frac{1}{2}\Big(\log A + (\nu+2)\ \frac{\frac{\partial}{\partial\nu}A}{A}\Big)$ and the expression of A and its derivative on ν

$$
\begin{aligned}
A &= \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)}\right) = \frac{\nu(1-\rho^2) + x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)} \\
\frac{\partial}{\partial\nu}A &= \frac{2\left(x_1\frac{\partial x_1}{\partial\nu} + x_2\frac{\partial x_2}{\partial\nu} - \rho(x_2\frac{\partial x_1}{\partial\nu} + x_1\frac{\partial x_2}{\partial\nu})\right)}{\nu(1-\rho^2)} - \underbrace{\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu^2(1-\rho^2)}}_{>0\ \forall\ x_1, x_2}
\end{aligned} \tag{3.45}
$$

We have already clarified as α reduces the values of $|x_1|$ and $|x_2|$ decrease as well. As consequence of that diminishing the value of ρ becomes meaningful. The way to describe that issue takes into account the numerator and the denominator of A. Basically, both decreases as ρ increases but $\nu(1-\rho^2)$ provides an increments of A. For a given fixed ν, A becomes a ρ's parametrized function of x_1 and x_2. One can prove that higher values of $|x_1|$ and $|x_2|$ the effect of ρ on A is limited: larger values of $|x_1|$ and $|x_2|$ enlarge the spread between the sum $x_1^2 + x_2^2$ and the value of $2\rho x_1 x_2$, thus ρ does not lead to a large changing (in absolute terms) with respect to A. The opposite is true when both $|x_1|$ and $|x_2|$ are smaller and the numerator of A decreases as well. Now evaluate $\frac{\partial}{\partial\nu}A$ in (3.45). Noting that is decreasing on ρ with the same rules of A, that is, decreases in absolute value as $\rho > 0$ grows up. The noteworthy difference relies in terms $\frac{\partial x_i}{\partial\nu}$. It can be easily evaluated that for not-extreme negative $x_i, i = 1, 2$ the value of $\frac{\partial x_i}{\partial\nu}$ decreases faster that $|x_i|$ as $|x_i| \to 0, i = 1, 2$. Under this condition, one can compute the ratio

$$
\frac{\frac{\partial}{\partial\nu}A}{A} = \frac{2(x_1\frac{\partial x_1}{\partial\nu} + x_2\frac{\partial x_2}{\partial\nu}) - 2\rho(x_1\frac{\partial x_2}{\partial\nu} + x_2\frac{\partial x_1}{\partial\nu})}{x_1^2 + x_2^2 - 2\rho x_1 x_2 + \nu^2(1-\rho^2)} - \left(\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{x_1^2 + x_2^2 - 2\rho x_1 x_2 + \nu^2(1-\rho^2)}\right)\frac{1}{\nu} \tag{3.46}
$$

it is worth noting that in the first part of (3.46), the numerator decreases faster than denominator because

$$
2|x_i|\ \frac{\partial x_i}{\partial\nu} < x_i^2 \quad i = 1, 2 \qquad \text{and} \qquad 2\rho x_1 x_2 < 2\rho\left(|x_1|\ \frac{\partial x_2}{\partial\nu} + |x_2|\ \frac{\partial x_1}{\partial\nu}\right)
$$

The second part of (3.46) is positive and bounded by $\frac{1}{\nu}$ and clearly reduces the expression.

This entails $\dfrac{\frac{\partial}{\partial\nu}A}{A}$ reduces its value.

In conclusion as α sufficiently small the increasing in ρ increases the value of G.

At the limit case $\rho \to 1$, we obtain

$$
\begin{aligned}
\lim_{\rho\to1}\frac{\frac{\partial}{\partial\nu}A}{A} &= \frac{2(x_1\frac{\partial x_1}{\partial\nu} + x_2\frac{\partial x_2}{\partial\nu}) - 2(x_1\frac{\partial x_2}{\partial\nu} + x_2\frac{\partial x_1}{\partial\nu})}{x_1^2 + x_2^2 - 2x_1x_2} - \frac{1}{\nu} \\
&= 2\frac{x_1(\frac{\partial x_1}{\partial\nu} - \frac{\partial x_2}{\partial\nu}) + x_2(\frac{\partial x_2}{\partial\nu} - \frac{\partial x_1}{\partial\nu})}{x_1^2 + x_2^2 - 2x_1x_2} - \frac{1}{\nu} = 2\frac{x_1(\frac{\partial x_1}{\partial\nu} - \frac{\partial x_2}{\partial\nu}) - x_2(\frac{\partial x_1}{\partial\nu} - \frac{\partial x_2}{\partial\nu})}{(x_1 - x_2)^2} - \frac{1}{\nu} \\
&= \frac{2(x_1 - x_2)(\frac{\partial x_1}{\partial\nu} - \frac{\partial x_2}{\partial\nu})}{(x_1 - x_2)^2} - \frac{1}{\nu} = \frac{2(\frac{\partial x_1}{\partial\nu} - \frac{\partial x_2}{\partial\nu})}{x_1 - x_2} - \frac{1}{\nu}
\end{aligned} \tag{3.47}
$$

that for what we said, as $|x_1|$,$|x_2| \to 0$, if follows (3.47) reaches to the maximum of the expression (3.46) and the G function rises up.

As consequence of ρ increasing and by result in (3.7.7), the value of G becomes positive early.

Chapter 4

Time change, Generalized Compounding and the Term Structure

Abstract

We apply the concept of "generalised compounding" proposed in Carr and Cherubini (2020) to the Stochastic Discount Factor dynamics and term structure models. We propose a class of time change models that preserve the lognormal distribution of one period SDFs while changing the compounding structure. These models decompose the SDF into lognormal marginal SDFs and a compounding function, that takes the shape of an Archimedean copula function. We show that stochastic clocks generally leads to a decrease of the yields, particularly pronounced on long term maturities. Preliminary evidence on US swap rates shows mild evidence in favour of the time changed model.

Keywords

Time change models, Generalized Compounding, Stochastic Discount Factor, Term structure, Copula functions

4.1 Introduction

Affine econometric models of the term structure are generally based on two main ingredients: lognormal stochastic discount factors (SDF) and geometric compounding. The main reference for the econometric parametrization of the affine model based on SDF is Ang and Piazzesi (2003) (see also Cochrane and Piazzesi (2005, 2008)) . A similar approach have been applied to the estimation of the term structures of risky assets, and particularly the equity market (see Giglio et al. (2020), and the literature review therein). In a recent paper, Carr and Cherubini (2020) show that models with time change may destroy the geometric compounding rule in discrete time. Aim of this paper is to investigate the impact of non-geometric compounding rules arising in time change models on the dynamics of the SDF.

For more clarity, define S_1 the one-period SDF from time $t = 0$ to time 1 and call M_N the compounded SDF for time N. Under the standard geometric compounding rule we have

$$M_N = S_0 S_1 S_2 \ldots S_{N-1} S_N$$

with $S_0 = 1$. Clearly, if the one-period SDF S_i is modelled as

$$S_i = \exp(-X_i)$$

with X_i normal random variables for all i, then both the S_i and the M_i SDF's are lognormal. The straightforward argument from which we draw our research question is that M_i are lognormal due to both the conditions above. In a time change model, the length of the time period, that in the standard model is set equal to 1, is assumed to be a stochastic variable, positive and possibly with mean equal to 1. So, the unit period SDF is modelled as

$$S_i = \exp(-X_i Z_i) \tag{4.1}$$

where Z_i is the positive random variable that we call *stochastic clock*.

The idea of time change in mathematics and probability theory is rooted in early work in the 60s and 70s of last century, when Dubins and Schwarz (1965), Dambis (1965), Follmer et al. (1973),Monroe (1978) proved that general stochastic processes, in general all semimartingales, can be generally obtained by applying a stopping time, that is a stochastic clock process, to a Brownian motion. Time change models have a long history in finance, starting from Clark (1973), who suggested a subordinated process for the dynamics of asset prices where the subordinator, i.e. the stochastic clock, was represented by the trading volume, an argument updated by Ané and Geman (2000) that instead identified the number of transactions per period as the relevant clock. Finally, many contributions can be found in mathematical finance and option pricing (Madan et al. (1998), Carr et al. (2003), Carr and Wu (2004), Madan and Yor (2008)). It is much less frequent to find applications of time change models in econometrics and macro-finance, with the exception of Stock (1988) (with application to GDP) and Shaliastovich and Tauchen (2005) (with application to the business cycle index and asset pricing). It must be remarked that most, if not all, the applications of time change models have been in continuous time. Carr and Cherubini (2020) showed that in a discrete time setting the compounding process with time changed can be written as in equation (4.1), that is as a Laplace transform (or a moment generating function, depending on the sign of the conditioning factor X_i). Here we apply this same idea to the SDF model.

Obviously the subordinated model in equation (4.1) implies that in general S_i is no longer lognormal and compounding is not geometric. Since we want to focus on compounding, it would be helpful to focus on a model that preserves lognormality. We will show that in a wide set of models (i.e. for many distributions of Z_i) we may find a function $f(y)$ such that S_i in

$$S_{t+i} = \exp(f(X_i)Z_i) \tag{4.2}$$

is lognormal. In particular, we will show two very simple cases in which this result obtains. In the first $f(y)$ is quadratic and Z_i is Inverse Gaussian (IG) distributed (one of the two models proposed in Carr and Cherubini (2020)). So, if the stochastic clock is IG-distributed and S_i is required to be lognormal, $f(y)$ must be quadratic (with a specific parametric form that will depend on the variance of the stochastic clock). In the second model the clock is assumed to be α-stable, with $\alpha \leq 1$. In this case one must have $f(y) = \exp(-y^{\alpha})$.

We can then turn our attention to the compounding/discounting rule, that is our main goal. Here assume that that the Z_i's are identically distributed and comonotonic. To put it in plain words, the stochastic clock is actually a single variable and we may set: $Z_i = Z$. Intuitively, this means that even if X_{t+i} are independent for all i, the dependence of all of them on the common variable Z will ensure that all of them are dependent. The SDF process will then be non-ergodic.

The intuition of dependence can be proved, and actually provides a link to a stream of literature quite well known in statistics, that addresses dependence by means of copula functions theory. Loosely speaking, copula functions are tools used to represent join distributions in terms of uniform marginal distributions (see Nelsen (2006) for an introduction and Cherubini et al. (2004) for financial applications). Here a specific family of these functions is found to represent the compounding structure linking one period SDFs. The family is called *Archimedean* copula and may be obtained using one of the most well known representation theorems, due to Marshall and Olkin (1988). In fact, that theorem uses multivariate Laplace transforms with respect to a common random variable (that in our case is Z).

The fact that we derive copula functions to represent compounding instead of dependence brings a notable innovation to the way we use copula functions here. The marginal are not uniformly distributed in the unit hypercube, but they are simply i.i.d. (assuming that the $f(X_i)$ elements are independent). In the cases that will be addressed, the one period SDFs will be lognormal. This will also be an original contribution to the copula functions literature, since this specific kind of Archimedean copula is novel, to the best of our knowledge.

To summarize our research program, we will model the SDF in a time change model, breaking down the specification of one periods SDFs and their compounded value. Both the distribution of one periods SDFs and their compounding structure will be affected by the distribution of the stochastic clock. The compounding structure is represented by an Archimedean copula function. By Archimedean structure, the SDF process is non ergodic, and can be also shown to be non Markovian. It may be said that this persistence obtained with the common stochastic clock substitutes for the martingale component used in the modelling of the SDF dynamics to represent long term returns (Hansen and Scheinkman (2009), Qin and Linetsky (2017).

The plan of the paper is as follows. In section 4.2 we lay out a Stochastic Discount Factor Model (SDF) with time change. In section 4.3 and 4.4 we give conditions for the time changed model to preserve the lognormal distribution of marginal one period SDFs. In section 4.5 we extract term structure from the model. Section 4.6 described the empirical implications and preliminary estimates on US data.

4.2 A SDF Model with Time Change

SDF is a stochastic process that is used to price financial products and contingent claims. Pricing is carried out under the physical measure so that SDF takes into account both expectations of future cash flows and the risk premium. Loosely speaking, the SDF is defined as a sequence of positive random variables M_t initialised at $M_0 = 1$. The price of each payoff G_n is obtained by

$$G_t = E_t\left(\frac{M_n}{M_t}G_n\right)$$

so that M_tG_t is a martingale under the physical measure $\mathbb{P}$. In this paper we focus on the risk free term structure, that is the set of zero coupon bonds paying $G_n = 1$. Denoting $P_t(n)$ the price of the zero coupon bond expiring at time $t+n$ we have

$$P_t(n) = E_t(M_n) \tag{4.3}$$

and the interest rate is defined as

$$r_t(t+n) = -\frac{\log P_t(n)}{n-t} = -\frac{1}{n-t}\log E_t(M_n) \tag{4.4}$$

One period forward rates can be defined from the logarithm of the expectation of the one period SDF, that we denote S_n

$$f_t(n) = -\log E_t(S_n) \tag{4.5}$$

In affine term structure models the SDFs are modelled as lognormal random variables, that is

$$S_t = \exp(-X_t) \tag{4.6}$$

where X_t is a normal random variable, possibly including a factor structure.

Example 4.2.1. A famous and typical example of lognormal specification is the Ang and Piazzesi (2003). In their parametrization

$$X_{t+1} = \delta_0 + \delta_1' F_t + \frac{1}{2}\lambda' \mathbb{R} \lambda + \lambda' \nu_{t+1}$$

where δ_0 is the intercept and δ_1 a vector of slope parameters of factors F_t, and ν_{t+1} is a vector of standard normal disturbances with correlation matrix $\mathbb{R}$. The vector λ denotes volatilities (that in their model are also considered conditionally heteroskedastic). We then have that the short term interest rate is affine in the factors

$$r_t(1) = \delta_0 + \delta_1' F_t$$

and so are the rates for all maturities.

We now introduce time deformation into the picture. Let us observe that the model on which we are working is specified at equally spaced intervals of time. In fact, we have in mind a specific frequency of observation Δt that we implicitly set equal to 1. Let us assume now that such time frequency is substituted by a positive random variable Z, with mean equal to 1, that is not observed and is meant to represent the degree of activity of the economy. The effect of this assumption is that the SDF is now modelled as

$$S_t = \exp(X_t Z) \tag{4.7}$$

and, if we maintain the assumption that X_t is normally distributed, the first effect is that S_t is no longer lognormal.

Before we go on, it is necessary to restrict the choice of our variable Z, with distribution F_Z. In fact, since it must be such that

$$E_Z(e^{-sZ}) = \int_0^\infty e^{-sZ} dF_Z \equiv \psi(s)$$

is well defined for s positive and negative. We can now define the building block of our model

$$S_t = E_Z(e^{-X_t Z}) = \psi(X_t)$$

We can now introduce our version of Marshall and Olkin (1988) theorem.

Theorem 4.2.2. *Assume $S_i, i = 1, 2, \ldots, n$ are i.i.d. positive random variables representing one-period SDFs,* *and a stochastic clock Z with Laplace transform $\psi(s)$. Then,*

$$E_Z(M_n) = E_Z(S_1 S_2 \dots S_{n-1} S_n) = \psi(\psi^{-1}(S_1) + \psi^{-1}(S_2) \dots + \psi^{-1}(S_{n-1}) + \psi^{-1}(S_n)) \tag{4.8}$$

Proof. We first define the Laplace transform

$$S_k = E_Z(-X_k Z) = \psi(X_k)$$

Since the S_k are independent we have

$$\begin{aligned} E_Z(M_n) &= E_Z(S_1 S_2 \dots S_{n-1} S_n) \\ &= E_Z\left(e^{-(X_1+X_2+\dots+X_{n-1}+X_n)Z}\right) \\ &= \psi(X_1 + X_2 + \dots + X_{n-1} + X_n) \\ &= \psi(\psi^{-1}(S_1) + \psi^{-1}(S_2) \dots + \psi^{-1}(S_{n-1}) + \psi^{-1}(S_n)) \end{aligned}$$

where the last line follows because the Laplace transform is invertible, so that $X_k = \psi^{-1}(S_k)$. □

Notice that the geometric compounding, which would obtain conditional on any value of the stochastic clock, is actually destroyed once that we conditioned out the clock. The compounding structure that is obtained is fully determined by the generator $\psi(s)$.
It is easy to verify that the compounding formula is actually the one obtained in copula functions theory, for the family of copula functions called Archimedean. In fact in the actual Marshall-Olkin theorem the variables X_k are assumed to be exponentially distributed with intensity parameter equal to 1 the variables S_k turn out to be uniformly distributed in the unit interval and we obtain standard Archimedean copulas. Here X_k have other distributions and one must be careful to handle the copula function. Moreover, in the compounding application one would be interested in preserving the lognormal distribution for S_k, that is typically assumed in affine models. We are going to show that this could be done, even though not always, of course at the cost of dropping the normality assumption on X_k. This will be shown in the next paragraph.

4.3 A Lognormal SDF Model Quadratic in Risk Factors

Since we saw that stochastic clocks destroy the geometric compounding structure, one would like to choose a model in which the lognormal distribution is preserved, at least, for the one period SDFs S_k. This way, one could compare two different term structures with same one period SDF focussing on the impact of the new compounding rule.
Here we show that this can be done keeping the model very simple, and as close as possible to the standard

literature. We first introduce the Inverse Gaussian distribution, and its Laplace transform

Definition 4.3.1. The Inverse Gaussian distribution $IG(\mu,\gamma)$ has density

$$g(z;\mu,\gamma)=\sqrt{\frac{\gamma}{2\pi z^3}}\exp\left(-\frac{\gamma(z-\mu)^2}{2\mu^2 z}\right)$$

with $\mu,\gamma>0$. The first two moments of the distribution are $E(Z)=\mu$ and $VAR(Z)=\mu^3/\gamma$. The Laplace transform is

$$E_Z\left(e^{-sZ}\right)=\exp\left[\frac{\gamma}{\mu}\left(1-\sqrt{1+\frac{2\mu^2 s}{\gamma}}\right)\right] \tag{4.9}$$

We now assume that the stochastic clock Z is Inverse Gaussian distributed and prove the following theorem.

Theorem 4.3.2. *Assume*

$$S_k=E_Z\left(\exp\left(-X_k+\frac{1}{2}VX_k^2\right)Z\right) \tag{4.10}$$

with $X_k \backsim N(m,\sigma^2)$ *and* Z *a positive random variable. Then,* S_k *is LogNormal*(m,σ^2) *if and only if* $Z \backsim IG(1,1/V)$.

Proof. Assume $Z \backsim IG(1,1/V)$. Using equation (4.9),

$$\begin{aligned}
S_k &= E_Z\left(\exp\left(-X_k+\frac{1}{2}VX_k^2\right)Z\right)\\
&= \exp\left[\frac{1}{V}\left(1-\sqrt{1+2V\left(-X_k+\frac{1}{2}VX_k^2\right)}\right)\right]\\
&= \exp\left[\frac{1}{V}\left(1-\sqrt{1-2VX_k+V^2X_k^2}\right)\right]\\
&= \exp(X_k)
\end{aligned}$$

Assume S_k is lognormal

$$\begin{aligned}
S_k &= E_Z\left(\exp\left(-X_k+\frac{1}{2}VX_k^2\right)Z\right)\\
&= \psi\left(-X_k+\frac{1}{2}VX_k^2\right)
\end{aligned}$$

Since $\psi(y)$ is invertible

$$\psi^{-1}(S_k)=-X_k+\frac{1}{2}VX_k^2$$

Since S_k is *Lognormal*(m, σ^2) and $X_k \backsim N(m, \sigma^2)$ the only possible shape of $\psi^{-1}(x)$ is

$$\psi^{-1}(S_k) = -\log(S_k) + \frac{1}{2}V(\log S_k)^2 = -X_k + \frac{1}{2}VX_k^2$$

and it may be verified that $\psi(s)$ is the Laplace transform of a variable IG$(1, 1/V)$. □

To make clear what we have did, we introduced a non linear function $f(X_k)$ to adjust or absorb the departure from lognormality induced by the stochastic clock Z. In this case, the shape of non linearity is the simplest one, since we only added a squared term multiplied by a parameter equal to half of the variance of the clock. So, in this model lognormality is preserved even though the compounding structure is changed to

$$E_Z(M_n) = \psi(\psi^{-1}(S_1) + \psi^{-1}(S_2) \ldots + \psi^{-1}(S_{n-1}) + \psi^{-1}(S_n))$$

that is more general than geometric compounding. When the variance of the clock V is set to 0 the function $f(X_k)$ returns linear and the compounding rule is geometric. One interesting question is the generality of this approach, that we address here below.

4.4 Generalisations

A question that is worth addressing is whether the IG clock model described above is the only model that accomplish the task of separating a multiperiod SDF into lognormal marginal one period SDFs and in case the answer is positive if it can always be done. An first sight the answer seems to be yes to the former question and no to the latter.

In fact, if we try to generalise the result above, it turns out that the model should work whenever the Laplace transform with respect to Z could be written as

$$E_Z\left(e^{-sZ}\right) = \psi(s) = \exp(\phi(s)) \tag{4.11}$$

In this case one can define $f(X_k) = \phi^{-1}(X_k)$ to obtain

$$E_Z\left(e^{-f(X_k)Z}\right) = \exp(\phi(f(X_k))) = e^{X_k}$$

and setting X_k normal accomplishes the task.

Now it is quite easy to find examples of Laplace transforms for which condition (4.11) holds. The first case that comes to mind is if one assumes that S_k be restricted in the unit interval. This assumption is not strange since in the term structure applications one would use $E(S_k)$ which is the forward price of a zero coupon bond, and then is bounded between 0 and 1. However, we will provide below the conditions for the model to be well behaved even for cases in which $S_k > 1$. The model is based on the assumption that Z be endowed with a positive stable distribution, for which the Laplace transform is

$$E_Z\left(e^{-sZ}\right) = \psi(s) = \exp\left(-s^{1/\theta}\right) \tag{4.12}$$

with $\theta \geq 1$. Notice that we stick to the parametrisation used by Marshall and Olkin (1988) and popular in the copula function literature. A more rigorous parametrisation would use $\alpha \equiv 1/\theta$, denoting the tail index of the stochastic clock. Obviously, the requirement $\theta > 1$ turns into $\alpha \in (0, 1]$, so that the distribution of the clock is such that neither first nor any higher order moments exist.

Now, if we define $f(X_k) = -X_k^\theta$ with X_k normal we clearly obtain that S_k is lognormal for all k and all values θ. As for compounding, it is clear that the case $\theta = 1$ would generate geometric compounding. The general compounding rule instead generates a compounding function that is very well known to any scholar in copula function theory

$$E_Z(M_n) = \exp\left(-\left[(-\log S_1)^\theta + (-\log S_2)^\theta + \ldots + (-\log S_{n-1})^\theta + (-\log S_n)^\theta\right]^{1/\theta}\right) \tag{4.13}$$

This is the well known Gumbel-Hougaard family, in the case in which S_k are in the unit interval. To extend the case to $S_k > 1$ we need a regularization. This is needed not only to give generality to the compounding function, but also because even in term structure applications as of today it is frequent to find $E_t(S_k) > 1$ due to negative interest rates.

Regularisation has to do with the terms $-\log(S_k)$ which are exponentiated to power θ. Allowing for $S_k > 1$ and $-\log(S_k)$ negative would do no harm if $\theta = 1$ but it would affect the symmetry property of the function $-\log(y)$ with $\theta > 1$. To see this, consider the very different cases that one would encounter with $\theta = 2$ and $\theta = 3$. Then we intervene defining the power function of $-\log(S_k) \geq 0$ and flipping the function for negative values. Formally, we define a power function ($y^{[\theta]}$ as

$$y^{[\theta]} = \begin{cases} y^\theta, & y \geq 0 \\ -\left(|y|^\theta\right), & y < 0 \end{cases} \tag{4.14}$$

Notice that this definition also takes care of another problem that we may run into when the argument of the sum of exponentiated $-\log(S_k)$ terms is negative and has to be exponentiated to $1/\theta < 1$.

We now introduce a definition that would helps us to simplify notation:

Definition 4.4.1. The operator $\oplus_\theta$, denoted pseudo-sum is defined as

$$y_1 \oplus_\theta y_2 \oplus_\theta \ldots \oplus_\theta y_{n-1} \oplus_\theta y_n = \left(y_1^{[\theta]} + y_2^{[\theta]} + \ldots + y_{n-1}^{[\theta]} + y_n^{[\theta]}\right)^{[1/\theta]}$$

We can then summarize these results in the following

Theorem 4.4.2. *Assume*

$$S_k = E_Z\left(-X_k^{[\theta]} Z\right) \tag{4.15}$$

with $X_k \backsim N(m,\sigma^2)$ *and* Z *a positive random variable. Then,* S_k *is Lognormal*(m,σ^2) *if and only if positive stable with parameter* $\alpha = 1/\theta \in (0,1]$. *Moreover, the compounding function is*

$$\log E_Z(M_n) = -((-\log S_1) \oplus_\theta (-\log S_2) \oplus_\theta \ldots \oplus_\theta (-\log S_{n-1}) \oplus_\theta (-\log S_n)) \tag{4.16}$$

We have then shown another example of SDF model with lognormal marginal (one period) SDF. We are tempted to check the other famous Archimedean copula function: the Clayton copula that would arise using a Gamma distribution for Z. Indeed, the procedure works also in this case, by a straightforward manipulation of the Laplace transform, as we show below.

Example 4.4.3. Consider a Gamma distributed clock. The Laplace transform is

$$S_k = E_Z\left(e^{-X_k Z}\right) = \psi(X_k) = (1+\theta X_k)^{-1/\theta}$$

We can rewrite the Laplace transform as

$$S_k = \psi(X_k) = \exp(\phi(X_k)) = \exp\left(\log(1+\theta X_k)^{-1/\theta}\right)$$

and we can compute

$$f(X_k) = \phi^{-1}(X_k) = \frac{e^{-\theta X_k} - 1}{\theta}$$

Now it is easy to verify that if $\theta > 0$

$$\begin{aligned} S_k =& \psi\left(\frac{e^{\theta X_k} - 1}{\theta}\right) \\ =& \left(1 + \theta \frac{e^{-\theta X_k} - 1}{\theta}\right)^{-1/\theta} \\ =& e^{X_k} \end{aligned}$$

If instead $\theta \to 0$ we have

$$\lim_{\theta \to 0} f(X_k) = \lim_{\theta \to 0} \frac{e^{-\theta X_k} - 1}{\theta} = X_k$$

and

$$\lim_{\theta \to 0} (1 + \theta f(X_k))^{-1/\theta} = e^{X_k}$$

4.5 Term Structure

We now apply the models described above to a generalisation of term structure models beyond the affine class. This will also show the utility of preserving lognormal marginal SDFs in order to maintain the model tractable, or at least only slightly more complex than affine models.

We start again from the basic case of geometric compounding remind that under the physical measure $\mathbb{P}$ we have $P_0(n) = E_0(M_n)$ and $P_0^f(k) = E_0(S_k)$, where $P_0^f(k)$ is the forward price established at time 0 to be paid at time $k-1$ for 1 dollar to be received at time k.

Since we aim to partition the zero coupon bond term structure in the product of forward prices, we switch measure from the physical measure $\mathbb{P}$ to the *forward martingale measure* (FMM) $\mathbb{Q}(k)$. We remind that under such measure we have

$$P_0(k) = E_0(M_k) = E_0(M_{k-1}S_k) = P_0(k-1)E_{Q(k-1}(1) = P_0(k-1)P_0^f(k)$$

Going back recursively, we have

$$P_0(k) = E_0(M_k) = P_0(1)P_0^f(2)\dots P_0^f(k-1)P_0^f(k)$$

We now introduce the stochastic clock Z. In order to maintain marginal lognormal we assume we may write the Laplace transform $\psi(s) = \exp(-\phi(s))$. The minus sign is immaterial. Notice that $\psi^{-1}(y) = \phi^{-1}(-\log y)$. Equipped with this we can write the same Marshall-Olkin style result.

$$\begin{aligned}
E_Z\left(E(M_n)\right) =& E_Z\left(E(M_n)|Z=z\right)\\
=& E_Z\left(P_0(1)P_0^f(2)\dots P_0^f(k-1)P_0^f(k))|Z=z\right)\\
=& \psi\left(\psi^{-1}\left(P_0(1)\right) + \psi^{-1}\left(P_0^f(2)\right) + \dots + \psi^{-1}\left(P_0^f(n-1)\right) + \psi^{-1}\left(P_0^f(n)\right)\right)\\
=& \exp\left[-\phi\left(\phi^{-1}\left(-\log P_0(1)\right) + \phi^{-1}\left(-\log P_0^f(2)\right) + \dots + \phi^{-1}\left(-\log P_0^f(n)\right) + \phi^{-1}\left(-\log P_0^f(n)\right)\right)\right]
\end{aligned}$$

If we now remind that

$$r_0(n) \equiv -\frac{1}{n}\log P_0^n \qquad f_0(k) = -\log P_0^f(k)$$

we recover the term structure

$$r_0(n) = \frac{1}{n}\phi\left(\phi^{-1}(r_0(1)) + \phi^{-1}(f_0(2)) + \dots + \phi^{-1}(f_0(n-1)) + \phi^{-1}(f_0(n))\right)$$

The short term rate $r_0(1)$ is observed on the market. As for the forward rates, they can be estimated by exploiting their lognormality:

$$-\log E_0(S_k) = f_0(k) = m_k + \frac{1}{2}\sigma_k^2$$

where m_k is the mean of the Gaussian factor X_k and σ_k^2 is its variance.

We now give the specific formulas for the models discussed in the previous sections.

Quadratic model. Remember that in the quadratic model the marginal SDFs are lognormal IFF the stochastic clock Z is $IG(1, 1/V)$. In this case the function $\phi(s)$ is

$$\phi(s) = \frac{1}{V}\left(1 - \sqrt{1 + 2Vs}\right)$$

and is

$$\phi^{-1}(\log(y)) = -\log(y) + \frac{1}{2V}(\log(y))^2$$

We then have

$$\begin{aligned} r_0(n) &= \frac{1}{nV}\left(\sqrt{1 + 2V\sum_{k=1}^{n}\left(f_0(k) + \frac{1}{2V}f_0(k)^2\right)} - 1\right) \\ &= \frac{1}{nV}\left(\sqrt{1 + \sum_{k=1}^{n}\left(2Vf_0(k) + f_0(k)^2\right)} - 1\right) \end{aligned} \tag{4.17}$$

Pseudo-sum model. We remind that in this model the stochastic clock Z has positive stable distribution with $\alpha = 1/\theta \leq 1$. We have $\phi(s) = s^{-1/\theta}$ and $\psi^{-1}(\log(y)) = (-\log(y))^{\theta}$. It is then easy to obtain:

$$r_0(n) = \frac{1}{n}((f_0(1)) \oplus_\theta (f_0(2)) \oplus_\theta \ldots \oplus_\theta (f_0(n-1)) \oplus_\theta (f_0(n))) \tag{4.18}$$

where we remind the definition of pseudo-sum

$$a \oplus_\theta b = \left(a^\theta + b^\theta\right)^{1/\theta}$$

for $a, b \geq 0$ (ad exponentiation is defined symmetrically for negative values).

Figure 4.1, Figure 4.2 reports the impact of stochastic clocks with considered on a flat forward term structure at 2% with increasing parameters. Figure 4.3 compares the two clocks assuming that they give the same rate for the 30 year maturity. We see that in all cases the stochastic clock reduces the interest rates, particularly on long term maturities. The main difference between the two models is that in the Pseudo-sum model the term structure seems to stabilize in the long run, in the quadratic model it is downward sloping along all the very long maturity horizon that was considered, that stretches up to 200 years.

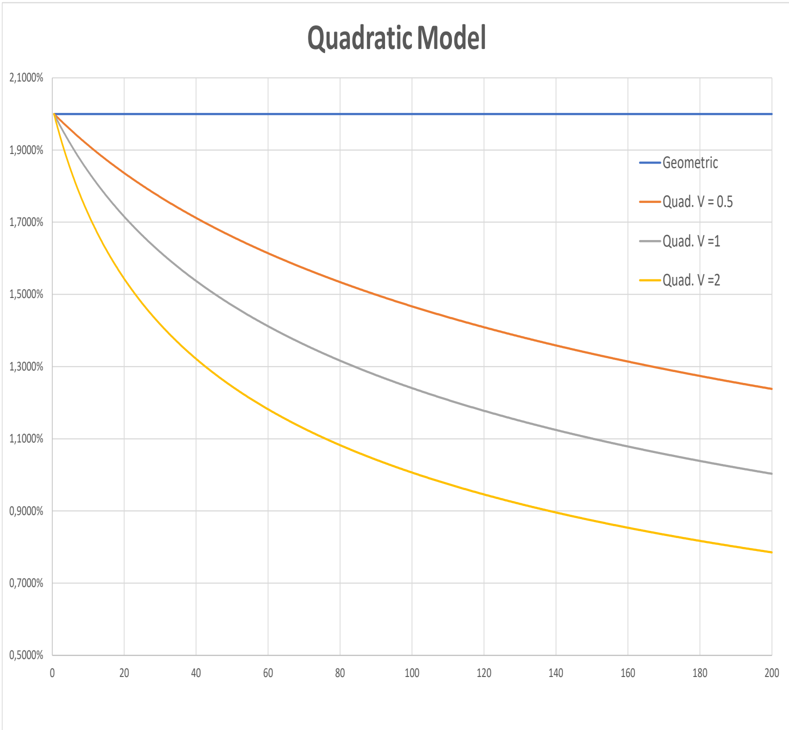


Figure 4.1: Quadratic model. One period forward rate flat at 2%.

Figure 4.2: Quadratic model. One period forward rate flat at 2%.

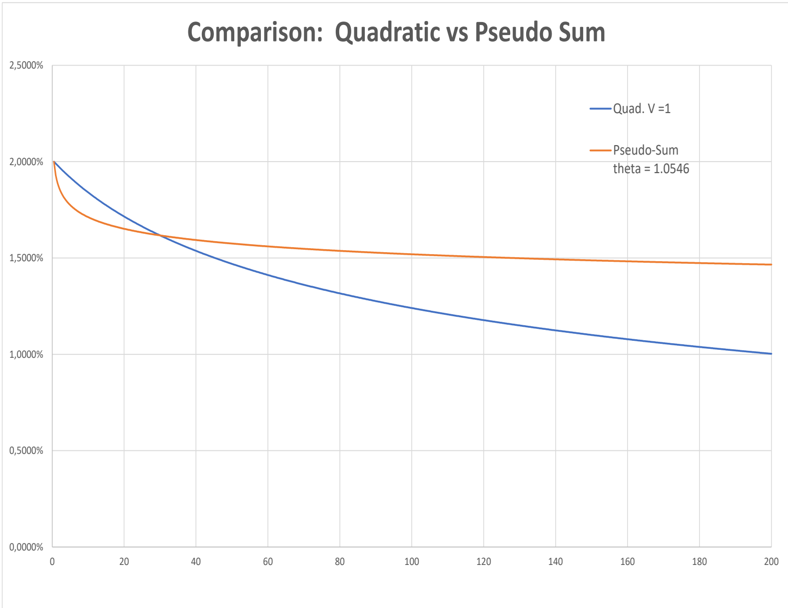


Figure 4.3: Term structure models: quadratic vs pseudo-sum. One period forward rate flat at 2%.

4.6 Empirical Analysis

We present here some preliminary results of an analysis of the evidence of the presence of a stochastic clock recovered from the term structure. The dataset used refers to the swap rate term structure in the US market. The data spans a period running from March 2005 to April 2021. The market and the period was selected because it includes quotes on the 50 year interest rate.

The model used is particularly simple. We assume a mean reverting risk factor. The model is calibrated using

$$m_k = (1 - b^{k-1})\bar{X} + b^{k-1}X_0$$

where $\bar{X}$ is estimated and X_0 is set equal to the 3-month rate. The variance is estimated as

$$\sigma_k^2 = \frac{1 - b^{2(k-1)}}{1 - b^2}$$

We selected the Pseudo-Sum model as the workhorse for the analysis. Table 4.1 reports the estimates of the model with geometric ($\theta = 1$) and Pseudo-Sum compounding. The results are mildly in favour of the time changed model, particularly in the periods before the financial crisis and during the crisis. The evidence does not seem to be clear also on the estimate on the whole sample. Here the stochastic clock parameter is very small and the MSE is slightly worse, but the θ parameter is statistically significant. Of course, further research and evidence will be needed to corroborate the results.

Table 4.1: Geometric ($\theta = 1$) and Pseudo-sum models estimation over different periods

Period	$\bar{\mathbf{X}}$	**b**	σ_ϵ	θ	**MSE**
03/01/2005 - 12/31/2007	0.0143	0.9454	0.01	1	6.0609
03/01/2005 - 12/31/2007	0.0169	0.9449	0.0029	1.387	6.0300
01/01/2008 - 12/31/2010	0.0262	0.8297	0.0949	1	55.750
01/01/2008 - 12/31/2010	0.0717	0.9713	1e-07	1.810	55.430
03/01/2005 - 04/10/2021	0.0114	0.9700	1e-07	1	1047.95
03/01/2005 - 04/10/2021	0.0115	0.9697	0.0267	1.002	1049.00

4.7 Conclusions

In this paper we have proposed an Archimedean distortion of the geometric compounding and discounting principle. The economic rationale for the distortion is provided by the concept of *generalised compounding* proposed by Carr and Cherubini (2020) and based on time change models. The idea is that if the returns are compounded or discounted in discrete time according to a stochastic clock, this brings about a distortion that can be modelled in terms of Laplace transforms. Then, the same technical tool proposed by Marshall and Olkin (1988) for copula functions allow to design a family of Archimedean compounding/discounting functions.

From a technical point of view, compounding functions are isomorphic to copula functions. They are more general in the sense that the marginal variables are not necessarily uniform, and in general they are not. In compounding functions the marginals represent one period return accrual or discount. In particular, we give conditions for a class of compounding/discounting functions to have lognormal margins. We propose two examples: an Inverse Gaussian clock, for which a quadratic factor allows to preserve lognormal margins; a positive stable clock that generates a Gumbel-like compounding/discounting function, and a term structure model that is specified in terms of a pseudo-sum.

We show that stochastic clocks generally affect the term structures by causing a decrease of the yields, particularly pronounced on long term maturities. Preliminary evidence on the US swap rates market shows mild evidence in favour of time changed models.

We show that this specification can be usefully applied to the analysis of the risk free term structure in a model based on the Stochastic Discount Factor.

Bibliography

Abramowitz, M. and Stegun, A. (1992). Handbook of mathematical functions (re-print of the 1972 edition).

Ané, T. and Geman, H. (2000). Order flow, transaction clock, and normality of asset returns. *The Journal of Finance*, 55(5):2259–2284.

Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary economics*, 50(4):745–787.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical finance*, 9(3):203–228.

Carr, P. and Cherubini, U. (2020). Generalized compounding and growth optimal portfolios: Reconciling kelly and samuelson. *Available at SSRN 3529729*.

Carr, P., Geman, H., Madan, D. B., and Yor, M. (2003). Stochastic volatility for lévy processes. *Mathematical finance*, 13(3):345–382.

Carr, P. and Wu, L. (2004). Time-changed lévy processes and option pricing. *Journal of Financial economics*, 71(1):113–141.

Chavez-Demoulin, V., Embrechts, P., and Nešlehová, J. (2006). Quantitative models for operational risk: extremes, dependence and aggregation. *Journal of Banking & Finance*, 30(10):2635–2658.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). *Copula methods in finance*. John Wiley & Sons.

Cherubini, U., Mulinacci, S., Gobbi, F., and Romagnoli, S. (2011). *Dynamic Copula methods in finance*, volume 625. John Wiley & Sons.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica: journal of the Econometric Society*, pages 135–155.

Cochrane, J. H. and Piazzesi, M. (2005). Bond risk premia. *American economic review*, 95(1):138–160.

Cochrane, J. H. and Piazzesi, M. (2008). Decomposing the yield curve.

Dakovic, R. and Czado, C. (2009). Comparing point and interval estimates in the bivariate t-copula model with application to financial data. *Statistical Papers*, 52(3):709–731.

Dambis, K. E. (1965). On the decomposition of continuous submartingales. *Theory of Probability & Its Applications*, 10(3):401–410.

Dubins, L. E. and Schwarz, G. (1965). On continuous martingales. *Proceedings of the National Academy of Sciences of the United States of America*, 53(5):913.

Embrechts, P., Nešlehová, J., and Wüthrich, M. V. (2009). Additivity properties for value-at-risk under archimedean dependence and heavy-tailedness. *Insurance: Mathematics and Economics*, 44(2):164–169.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, 38(1):34–105.

Follmer, H. et al. (1973). On the representation of semimartingales. *the Annals of Probability*, 1(4):580–589.

Frees, E. W. and Valdez, E. A. (1998). Understanding relationships using copulas. *North American actuarial journal*, 2(1):1–25.

Giglio, S., Maggiori, M., Stroebel, J., and Utkus, S. (2020). Inside the mind of a stock market crash. Technical report, National Bureau of Economic Research.

Hansen, L. P. and Scheinkman, J. A. (2009). Long-term risk: An operator approach. *Econometrica*, 77(1):177–234.

Ibragimov, R. (2009). Portfolio diversification and value at risk under thick-tailedness. *Quantitative Finance*, 9(5):565–580.

Ibragimov, R., Jaffee, D. M., and Walden, J. (2008). Insurance equilibrium with monoline and multiline insurers. *Fisher Center for Real Estate & Urban Economics.*

Ibragimov, R. and Prokhorov, A. (2016). Heavy tails and copulas: Limits of diversification revisited. *Economics Letters*, 149:102–107.

Ling, C.-H. (1965). Representation of associative functions. *Publ. Math. Debrecent*, 12:189–212.

Mack, C. (2008). Appendix c: The dirac delta function. https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470723876.app3.

Madan, D. B., Carr, P. P., and Chang, E. C. (1998). The variance gamma process and option pricing. *Review of Finance*, 2(1):79–105.

Madan, D. B. and Yor, M. (2008). Representing the cgmy and meixner lévy processes as time changed brownian motions. *Journal of Computational Finance*, 12(1):27.

Marshall, A. W. and Olkin, I. (1988). Families of multivariate distributions. *Journal of the American statistical association*, 83(403):834–841.

Marshall, A. W., Olkin, I., and Arnold, B. C. (1979). *Inequalities: theory of majorization and its applications*, volume 143. Springer.

Monroe, I. (1978). Processes that can be embedded in brownian motion. *The Annals of Probability*, pages 42–56.

Nelsen, R. B. (2006). *An Introduction to Copulas.* Springer Science & Business Media.

Nolan, J. P. (2012). *Stable distributions*, volume 1177108605. ISBN.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical recipes 3rd edition: The art of scientific computing.* Cambridge university press.

Qin, L. and Linetsky, V. (2017). Long-term risk: A martingale approach. *Econometrica*, 85(1):299–312.

Samorodnitsky, G. and Taqqu, M. S. (1997). Stable non-gaussian random processes. *Econometric Theory*, 13:133–142.

Schweizer, B. and Sklar, A. (1961). The algebra of functions. ii. *mathematical annals*, 143(5):440–447.

Shaliastovich, I. and Tauchen, G. (2005). Pricing implications of stochastic volatility, business cycle time change and non-gaussianity. *Duke University.*

Skorokhod, A. V. and Slobodenyuk, N. (1965). Limit theorems for random walks, i. *Theory of Probability & Its Applications*, 10(4):596–606.

Stock, J. H. (1988). Estimating continuous-time processes subject to time deformation: an application to postwar us gnp. *Journal of the American Statistical Association*, 83(401):77–85.

Sugeno, M. and Murofushi, T. (1987). Pseudo-additive measures and integrals. *Journal of Mathematical Analysis and Applications*, 122(1):197–222.