UNIVERSIDADE ABERTA

Modelling and Estimation of Systemic Risk: new perspectives

Jorge Manuel Lopes Basilio

Doctorate in Applied Mathematics and Modelling

2021

UNIVERSIDADE ABERTA

Modelling and Estimation of Systemic Risk: new perspectives

Jorge Manuel Lopes Basilio

Doctorate in Applied Mathematics and Modelling

Thesis supervised by Doctor Amílcar Manuel do Rosário Oliveira and Doctor Rahim Mahmoudvand

2021

Resumo

A Análise de Risco tem-se assumido como um tema de investigação recorrente, atraindo a atenção dos investigadores de forma consistente. Mais recentemente e motivado pelo último colapso do sistema financeiro, o risco sistémico tem sido alvo de especial atenção por parte da comunidade académica, tendo-se tornado numa ferramenta amplamente aplicada para identificar a contribuição para o risco sistémico de instituições financeiras.

Esta tese é baseada no trabalho de investigação de Adrian and Brunnermeier (2011), onde foi introduzirdo o conceito de $CoVaR$ e $\Delta CoVaR$ de uma instituição financeira, bem como apresentada uma metodologia para estimar o $\Delta CoVaR$ usando dados públicos do mercado financeiro .

Nesta tese serão discutidos os pressupostos da metodologia original, analisadas as características de cada medida de risco utilizada, e discutidas alternativas para medir a contribuição individual de uma entidade para o risco sistémico do sistema financeiro Benoit et al. (2017). Será proposta uma nova metodologia baseada em funções de cópula de modo a evidenciar o papel da dependência e da dependência nos extremos entre o retorno da instituição financeira e o retorno do sistema financeiro.

Iremos destacar as diferenças entre a qualidade do ajustamento nas abas da distribuição dos retornos financeiros assim como o ajustamento para toda a distribuição.

Plavras chave: Risco Sistémico, Copulas, Modelos de Mistura, Valores Extremos, $CoVaR$

Abstract

Risk Analysis is becoming a recurrent subject in research, attracting the attention of researchers in a consistent way. More recently and motivated by the last collapse of the financial system, systemic risk is getting special attention from academic community as well as becoming a tool widely applied for detecting financial institutions systemic risk contributions.

We started with Adrian and Brunnermeier (2011) work, where they introduced first time the concept of $CoVaR$, and $\Delta CoVaR$ of a financial institution, as well as a methodology to estimate $\Delta CoVaR$ using financial market public data.

This thesis will discuss the assumptions taken along the original methodology, analyse the characteristics of each risk measured used, and discussed alternatives to measure the individual contribution of a single entity to the systemic risk of a financial system Benoit et al. (2017). It will be proposed a new methodology based on copula functions to evidence the role of dependence and tail dependence between financial institution returns and financial system returns.

We will highlight the differences between the quality of fitting in the tails of the financial returns distribution and the fitting for all the distribution.

Keywords: Systemic Risk, Copulas, Mixture Models, Extreme Values, $CoVaR$

Acknowledgements

I would like to thanks to my supervisors, Doctor Amílcar Oliveira for all the dedicated support and guidance throughout this project as well as all thoughtful comments and recommendations, and Doctor Rahim Mahmoudvand for providing guidance and feedback.

Thanks also to my wife, for providing guidance and unconditional support when required.

Contents

List of Tables

List of Figures

Nomenclature

ARCH Autoregressive Conditional Heteroskedasticity

BCBS Basel Committee on Banking Supervision

BM Block Maxima

cdf cumulative distribution function

CDS Credit Default Swaps

CoVaR Conditional Value at Risk

CVaR Conditional Value at Risk

EM Expectation-Maximization

ES Expected Shortfall

ETL Expected Tail Loss

EVA Extreme Value Analysis

EVD Extreme Value Distribution

EVT Extreme Value Theory

FSB Financial Stability Board

GARCH Generalized Autoregressive Conditional Heteroskedasticity

GEV Generalized Extreme Value

GLS Generalised Least Squares

GPD Generalized Pareto Distribution

IMF International Monetary Fund

MES Marginal Expected Shortfall

MLE Maximum Likelihood Estimator

MVA Market Value of Assets

OLS Ordinary Least Squares

pdf probability density function

POT Pick Over Threshold

SES Systemic Expected Shortfall

SIB Systemically Important Banks

TDC Tail-dependence Coefficient

VaR Value at Risk

Chapter 1

Introduction

1.1 Preface

Risk Analysis is becoming a recurrent subject in research, attracting the attention of researchers in a consistent way.

Risk is often defined as the real outcome of an event can differ from the expected or desired outcome. This definition, when applied to finance domain is associate with the probability that actual results of an investment will differ from expected results, and can also be associated with the volatility of those returns. Two main categories can therefore be identified on risk, the systemic and unsystemic risk. While unsystemic risk relates to the risk of the activity of each agent in a system, the systemic risk is related to the risk of the entire system, and the risk of the system eventually collapse. On other turn sytemic risk is a type of risk that relates sytemic and unsystemic risk, as it is the risk associated with the possibility that an event at the agent level could cause an impact in the system Nils et al. (2018).

More recently and motivated by the last collapse of the financial system, systemic risk is getting special attention as well as becoming a tool widely applied for detecting financial institutions systemic risk contributions.

We start with Adrian and Brunnermeier (2016) work, where they introduced first time the concept of Conditional Value at Risk ($CoVaR$), and ΔConditional Value at Risk ($\Delta CoVaR$) of a financial institution, as well as a methodology to estimate $\Delta CoVaR$ using financial market public data.

In this thesis we will then discuss assumptions taken along the methodology proposed, analyse the characteristics of each risk measured used, and discuss alternatives to measure the individual contribution of a single entity to the systemic risk of a financial system Benoit et al. (2017).

At the moment, there is not yet a consensus in accepting existing measures and methodologies as good enough to correctly identify the biggest contributors to systemic risk Delbaen (2002). As conclusion, a modified methodology to estimate individual contributions for systemic risk using market data is presented.

1.2 Background and Motivation

In the wake of the last financial crisis, research on systemic risk in financial markets has been intensified. The adequate monitoring of systemic risk related to the financial system becomes a major requirement for the regulators across the world. This needs lead to design new systems and indicators more suitable for such monitoring. The process to identify and collect the necessary data in order to permit to compute those indicators become also a huge challenge for the regulators Data analysis of data on previous financial crises did not provide always the valuable insights for the task. From this become also clear that, in order to obtain and design the relevant indicators and reporting tools, it will be required more than just collect the data, and it will be required a process to identify the propagation dynamics across the financial system to be monitored. This way, the process of designing a system to monitoring the financial system requires an effective interaction between theory and also empirical research.

Those crisis exposed the vulnerability of the financial system, with the major reference in the bankruptcy of Lehman Brothers in 2008, followed by a chain of events that generated panic and undermining the confidence required by financial systems to work in a proper way. It was recognized by the International Monetary Fund (IMF) the inadequacy of current mechanisms to address properly all the complexity inherent to a global financial system and also that the lack of effective mechanisms for such purpose lead to situations of significant risk (Risk, 2009).

Looking at the case of 2008 financial crisis that was initiated in the United States subprime mortgage market, at the time a relatively negligible, in terms of the size when compared with the financial sector, this localized problem quickly spread and contaminated the global financial system.

This impact in the financial systems, expressed then in excessive imbalances in the world economies soon undermined the investors confidence forced the world economy to a prolonged period of financial stress closely followed by economic recession.

This catastrophic sequence of events showed also that financial professional were unprepared in terms of empirical models and tools, that could allow for a suitable monitoring of all the complexity involved in the financial environment at the time.

Also the risk management framework offered by Basel II regulations, focused on Value at Risk (*VaR*) as a base for risk management policies reveals inadequate, underestimate risk in scenarios of extreme rare events. The role of the supervisor in identifying systemic financial institutions was in this way compromised and the eventual preventive actions in order to mitigate undesired outcomes were not taken.

1.3 Systemic Risk and Related Metrics

One of the strategies to prevent future systemic financial crisis is implementing policies to limit systemic risk in the financial system, as it was shown by empirical evidence from the past that systemic risk has a huge potential to spread thought economy. Controlling systemic risk allows to contain the domino effect that puts at risk of bankruptcy in financial institutions one after the other.

Basel Committee on Banking Supervision Young (2011) establishes that each institution should have a surplus of capital in line with the negative impact it generates, this means in relation to its individual contribution to the risk of the global financial system.

Currently the Financial Stability Board (FSB) publishes a list of banks denominated as Systemically Important Banks (SIB), based on the potential systemic risk they add to the entire financial system as an individual contribution for systemic risk. The methodology in place used to rank the financial institutions by FSB is based on balance-sheet figures periodically communicated by each financial institution.

However the data used on the process are not publicly available, and the reasons that support the regulator decisions are not always clear and transparent to financial markets and operators. These was noticed as a weakness of such practise.

These limitation has been a source of motivation for scientific community and several researchers (Acharya et al., 2010), (Benoit et al., 2017), Adrian and Brunnermeier (2016) worked in alternative methodologies that could be based instead on public market data and this way add the required transparency and clarity to the process of identify the SIB in the financial system.

With this in mind new systemic risk measures were proposed such as Expected Shortfall (ES), Conditional Expected Shortfall (CES) and $\Delta CoVaR$. The same way our first purpose is also to contribute for this discussion and contribute to make available more accurate methods to identify systemic important financial institutions.

1.4 Thesis Overview

This thesis is composed by of six chapters. The main research efforts and the ensuing empirical findings and policy prescriptions are described in chapter two to six. Here, a short description of the aforementioned is provided.

Chapter 2. A description of the most significant research work in the subject covered in this thesis, in special the work developed on systemic risk is provided. The walk through over the state of the art in terms of systemic risk research including a reference to the main advances in the area.

Chapter 3 covers the main theoretical principles used to estimate systemic important financial institutions. It starts with an introduction to the concept of risk measures and an introductory description of the most important risk measures in the context of systemic risk. Some of methodologies to estimate those risk measures are also covered with examples for:

- Historical (*VaR*)
- Parametric *VaR*
- Extreme Value Theory (EVT) *VaR*

as different options to calculate *VaR*.

The role and the impact of normal distribution assumption to estimate risk measures is also approached. Extreme value theory as a theory to model rare events is covered as well, as financial crisis and introduces the problematic of modelling heavy tailed data. Mixture models are also included as a flexible technique to model data with heavy tailed.

Copula theory is also introduced as a base theoretical foundation used to support the main results and findings described on this thesis. Starting with the copula definition introduced by Sklar (1973), the measures of dependence of a random variable are covered. Detail on the most significant copula families are included as well.

Also at Chapter 4 is described the details and theoretical foundations of the methodology proposed to identify systemic risky financial institutions.

Chapter 5 presents in detail the steps to implement the methodology proposed in the previous Chapter, followed by a discussion of each result obtained.

Chapter 6 concludes this thesis. First, a short overview of the research aim and scope is provided. Finally, a few recommendations for future research efforts are made.

Chapter 2

Literature Review

2.1 Introduction

Researchers have been paid a considerable amount of attention to study risk measures and risk management methodologies and bibliography is abundant on the subject.

Systemic risk research focuses mainly on the point of view of crisis and a series of financial crises erupted from the bank system point out the contagion issues during the crises due to the spillover effect and turn the systemic risk widely recognized.

A considerable effort has being made to develop and propose new tools and methodologies to measure systemic risk, as the β coefficient and VaR reveals limitations to assess contagion risk and systemic risk and the need for a more effective tools is urging as it was demonstrated during and on the aftermath of the most recent financial crisis.

One of the most important drawback by using VaR, the most important classical risk measure, is it's inability to capture the systemic nature of the risk involved, as it is focused on the individual agent in the system as a single institution, not incorporate the complexity of the interactions between that particular institution and the others institutions in the system. It is believed that those complex interaction and relations are a source of new risks (Danielsson et al., 2012)

The most recent research work is now presenting and looking for new quantification measures that could cope with the need to include those interactions on systemic risk measurement (Acharya et al., 2012a) as it is a source of debate on academia and as well within the regulators which tools are adequate and which tools are effective choices to accurately estimate systemic risk across financial systems (Schwaab, 2010).

The systemic risk we focus in this work is the risk associated with a failure of one individual financial institution and the risk of a contagion and spillover on other institutions through the connections established between these institutions.

2.2 Risk Measures

Value at risk has an important role in risk management research and practise, due to its popularity, playing also an important role in research and its concept is often present in the base of more complex and elaborated risk measures.

Although value at risk popularity as risk measure, there is also criticism as it has been pointed out some undesirable mathematical characteristics, such as for example the lack of subadditivity and convexity.

Hendricks (1996) describes several different methods to estimate the VaR and compared the results by using a simulation technique to obtain random foreign exchange portfolios. They obtained measures of price risk for the portfolios at both 95 percent and 99 percent confidence levels over one-day holding periods. Considering only the context of market risk, the methodology does not consider portfolios with nonlinear price behavior.

Duffie and Pan (1997) provided an overview of value at risk and discussed some of the econometric modelling required to estimate it, this work contributed with a description of some of the basic issues involved in measuring the market risk of a financial institution. Although this article provides an accessible overview of VaR, it does not make any claims to include new research results. It includes a discussion of the econometric modelling requirements to estimate VaR.

Artzner et al. (1999) proposed a list of desired properties to be found in a coherent risk measure and showed the value at risk has missing on subadditivity and convexity. At the same time this paper defined the consistency conditions desired for a coherent risk measure also it raised that VaR does not meat such criterion and in this context new risk measures that could satisfy these consistency conditions and also being easy to compute were needed.

To respond these need new risk metrics where proposed such as $CoVaR$ by Rockafellar and Uryasev (2002) and ES, developed by Acerbi and Tasche (2002), in order to mitigate the limitations of *VaR*, specially those concerns with risk metrics coherence. These measures will use the α percent worst cases of losses to estimate the expectation for those worst losses.

Föllmer and Schied (2008) contributed with the concept of Convex Risk Measures. These type of risk measures remove the axiom of positive homogeneity of coherent risk measures and allowed for the introduction of conditional convex risk measures as well. Starting by discussing the limitations of industry standard for risk measures, they search for alternatives and started by specifying a set of desirable axioms for risk measures. Axioms for monetary, convex, and coherent risk measures were provided. A dual representation for convex and coherent risk measures were included followed with examples.

2.3 Systemic Risk Measures in Financial Systems

Research on systemic risk has been intensified specially after the last financial crisis, named also as the subprime 2008 crises. All the events related, and the implosion of the Lehman Brothers exposed vulnerabilities within the financial system combined with other events such as the European sovereign debt crisis that affected Eurozone countries. Such events demonstrated also that the tools available at the time showed ineffective to measure and manage systemic risk, as well as a lack of effective mechanisms for dealing with these events, that exposed those economies to important risks (Risks and Soundness, 2008).

This topic was approached by Oort (1990) where is identified three possible sources of vulnerability to the financial system:

- a larger institution failure causes an extended financial crisis via an complex network of interconnections with the other financial institutions;

- the systemic risks alleged to be inherent to certain "new" bank products;
- impact of external events, such as debt crises, strong changes in interest or exchange rates, deregulation, and recession.

Even though Oort (1990) considers the possibility of a major financial crisis could be avoided and tuned to be small, due to improvements in regulation, including adequate policies by the monetary authorities and effective international coordination, the recent events were not avoided.

The first source risk category identified, which refers to the risk of a bankruptcy of one financial institution that causes a contagion effect on other institutions in the system. Several researchers have focused their attention on this type of source of risk and risk propagation.

In this regard, Segoviano Basurto and Goodhart (2009) proposed a banking stability index that assesses interbank dependence based on tail events, where the financial system is set as a portfolio of individual financial institutions. They analyse how individual firms contribute to financial system distress by using the Credit Default Swaps (CDSs) of each financial institution.

Acharya et al. (2010) measures systemic risk using the joint propensity to failure that arises from correlated returns, by applying systemic and marginal expected shortfall measures to quantify the contribution of individual financial institutions to systemic risk. The authors demonstrated empirically the ability of systemic expected shortfall risk measure to predict emerging risks during the financial crisis of 2007-2009. The predictions included the outcome of stress tests performed by regulators, the decline in equity valuations of systemic important financial institutions during the financial crises and also the widening of their credit default swap spreads. Additionally is also proposed a systemic risk measure, named Systemic Expected Shortfall, to measures the conditional capital shortfall of a financial institution.

Adrian and Brunnermeier (2011) introduced *CoVaR* as systemic risk measure that links the systemic risk contribution of a financial institution with the raise of the *Var* of the financial system when that financial institution is under distress to capture the effect of risk spillovers between the financial institutions in the system.

Allen et al. (2012) proposes an aggregated measure of systemic risk and systemic risk index designated CATFIN, associating systemic risk to financial system *Var*. The authors demonstrate that combining several risk measures results in a unique measure it could acquire significant predictive power. They use the generalized Pareto distribution to model extreme losses. It was also considered the skewed generalized error distribution, in order to investigate the shape of the entire distribution of excess returns on financial institutions in a given time period. CATFIN can explain liquidity from either a demand perspective or from a supply perspective. The authors advocate that CATFIN risk measure is able to forecast financial market volatility by considering index options and credit default swap spreads and that with CATFIN it is possible to predict bank lending activity. An increase in CATFIN risk measure is interpreted as a reduction in bank profits.

Brownlees and Engle (2012) presented SRISK as an empirical methodology to measure the systemic risk contribution of financial institutions. SRISK is a systemic risk measure that estimate the amount of capital needed to restore the defined minimum capital requirements of a financial institution. SRISK measure is a function of the size of the firm, its degree of leverage, and its expected equity loss conditional on the market decline. SRISK allows to build rankings of systemically risky institutions, where institutions with the highest SRISK are the largest contributors to the the risk and undercapitalization of the financial system. The sum of SRISK of all financial institutions is systemic risk of the entire financial system.

Billio et al. (2010) focused on identify and measure contagion and exposure effects that arises from in the relationship verified between financial institutions based on principal components analysis and Granger-causality networks.

Girardi and Ergün (2013) proposed a new approach to quantifying *CoVaR*, initially proposed by Adrian and Brunnermeier (2011), by using a multivariate GARCH model to estimate the joint density of the financial system and financial institution returns.

Acharya et al. (2012b) proposed a new systemic measure based on the premise that the systemic risk should not be described only in terms of a financial firm's failure, but also in the context of a firm's overall contribution to system failure. The authors introduced a method to estimate the capital that a financial firm would need to raise to face another financial crisis. This measure of capital shortfall is based on publicly available data and is in is form similar to the stress tests conducted by regulators. The authors advocate that this measure summarizes the most important characteristics of systemic risk.

2.4 Copulas Theory

Copulas are functions that allows to work separately the marginal distributions from the dependency structure associated to a multivariate distribution.

The copula concept was introduced by Sklar (1959), in an article written in French. The same concept was after discussed in an article in English (Sklar, 1973). Both articles described copulas as functions that describe multivariate distributions to the one-dimensional margins.

Joe and Hu (1996) provided details and information about copulas and related properties and applications to multivariate distributions.

Nelsen (2003) provides an extended survey where copula properties are described in detail. It is also included copula applications and the study of dependence and measures of association as well as methods of construction of copula families for bivariate distributions.

Trivedi and Zimmer (2007) explores the potential of copulas in econometric modelling by using joint parametric distributions, discussing the possible advantages of copulas and scenarios where applying copulas could be advantageous, as it allows for construct the joint distribution based on the marginal distributions.

The interest in copulas as an approach to modelling joint distributions aroused in a diversity of areas. Therefore copulas gain reputation as a valuable tool to analyse and model problems involving complex dependence structures as risk (Cherubini et al., 2004). This work developed by Cherubini et al. (2004) is focused on financial applications and extends the analysis also to copula theory foundations.

Embrechts et al. (2001) used copula function to model dependence in financial data covering topics as extreme values and the behaviour of correlation subject to extreme market movements, while discusses the hypothesis of linear correlation as measure of dependence.

Schweizer (1991) provided a historic background in copulas detailing the most significant achievements in copula theory. Important results such as the developments in the theory of probabilistic metric spaces, its connection with the study of families of binary operations over probability distribution functions, the results that allowed for the use of bi-dimensional copulas to define measures of dependence of pairs of random variables.

The *copula* R package (Hofert et al., 2014), also described in detail by Yan et al. (2007) was used as the base software tool during the implementation steps.

2.5 Extreme Value Theory

Both *VaR* and also ES are making use only of the extreme quantiles of the distribution in use, being it losses distribution or returns distribution. This way, also in both cases the center of the distribution is disregarded in favor of the extremes and Extreme Value

Theory (EVT) proved to be a useful tool in such scenarios. EVT is a sound and proven probability theory with the ability to study the behavior of the extremes, maximums or minimums of other probability distributions. This theory was first developed by Fisher and Tippett (1928) as also referenced in Beirlant et al. (2006) and de Oliveira (2013). Those statisticians established the foundations of EVT and were followed by others like Gumbel (1958) and Gumbel (2012). Gumbel formalized the theory and statistic methods involved in EVT and contributed significantly to the popularity of EVT.

The Generalized Extreme Value distribution (GEV), that models the behavior of block maxima was presented by Jenkinson (1955) in the popular functional form that combines the 3 distributions, Gumbel, Fréchet and Weibull.

Pareto distribution is a well known statistical distribution that also exhibits heavy tails and is widely used across different areas with many applications in economy, physics, engineering, biology, finance and others. In 1975 Pickands III (1975) introduced a generalization of the Pareto model with special interest in application in finance to model the excess above a defined level. This generalisation has been applied to model extreme values modelling problems in areas such as insurance and finance, among others (de Zea Bermudez and Kotz, 2010).

Embrechts et al. (1998) discuss the importance of EVT as analytic tool in risk management and compare the EVT approach with other option to illustrate the potential of EVT to model risk.

The increasing interest in EVT motivated also researchers to develop several R packages to facilitate the application of EVT (Gilleland et al., 2016). The extRemes R package has a set of functions for extreme values analysis by using the block maxima or excesses over threshold.

Additional examples of R packages for extreme values are also *evir* R package (Pfaff et al., 2018) and *evd* R package (Stephenson, 2018) that provides a range of functions for fitting EVT using maximum likelihood estimates for univariate and bivariate models with both maxima threshold models.

2.6 Mixture Distributions

Mixture models are probabilistic models used in machine learning and statistics to represent the presence of sub-populations over observed data. Those sub-populations or groups of data are then represented by distinct model distributions.

The process of implementing this type of statistical models often involves multi-steps processes and iterative processes.

As Scott and Symons (1971) described in their paper, this concept is also related to cluster analysis, where they discussed a cluster analysis process involving several normal distributions.

Also Day (1969) described the process of estimate parameters for a mixture of two normal distributions and discussed the challenges of doing it by applying the maximum likelihood method. The case involving more than two normal distributions is also approached in this work.

An earlier specification of mixture models was provided by Teicher et al. (1960), covering the most important theoretical concepts involved.

However, as mixtures models are in a summation, the process required for parameter estimation reveals a complex task. This complexity motivated the research that led to the EM algorithm, described by Dempster et al. (1977). The EM algorithm is now accepted as the standard tool to estimate the parameter in mixture distribution.

Several mixture models were proposed to model financial risk and financial returns, combining several Gaussian distributions or combining a Gaussian distribution with other distributions more able to deal with the heavy-tails. In order to manage financial risk, Li (2012) proposed a Gaussian-Cauchy mixture model exploring the capabilities of Cauchy distribution to deal with heavy-tails.

Some authors also proposed mixture-models to deal with time-varying type of problems including an autoregressive component as Wong and Li (2000) and Zhang et al. (2006) who also used a GARCH model with mixture models to model real data with a heavy-tailed distribution in a time-varying context.

MacDonald et al. (2011) motivated by the challenge involved in modelling the lower tails of a distribution, developed and proposed a new extreme value model able to deal with rare events as a flexible mixture model, by combining a non-parametric kernel density estimator, in order to estimate the bulk component of the mixed distribution, combined with an appropriate tail model.

Several researchers contributed also with a large number of R packages to facilitate the use and estimation of mixture distribution. Some of the examples are Young et al. (2020) and Hu and Scarrott (2018) more focused on extreme value mixtures, but many others could be referenced as well.

Chapter 3

Theoretical Foundation

3.1 Risk Measures

3.1.1 Introduction

Uncertainty plays a role in the most part of human activities and also has a role in decision making.

Risk relates to the future uncertainty and implies a deviation from an expected outcome. In finance, risk is often defined as the probability that actual results will differ in future from the expected results. Capital Asset Pricing Model (CAPM) defines risk as the volatility of the returns of an investment.

Risk measures the uncertainty that an investor is willing to take to realize a gain from an investment.

Risk metrics are becoming an essential tool to express, communicate and use the results obtained from risk analysis in a proper decision-making process.

A risk measure is used to calculate and determine the amount of an asset or several assets (usually currency) to be maintained as a reserve in order to cover for unexpected losses. Risk measure tuns then in statistical measures used to predict financial risk of volatility. It can be seen as the minimum extra cash that has to be added to the risky position and invested in a risk free asset to make the risky position acceptable (Duffie and Pan, 1997).

Several authors have defined risk in terms of changes in the value of a position between two dates, it can also be established that the risk is related to the variability the future value of a position due to uncertain events, so it will be preferable, consider only future values (Chengli and Yan, 2012). In fact, what really matters, from the point of view of applicability, are the future values of a position compared to the current value.

Thus, the study of risk should focus on the random variables in the set of states in the future, which in turn can be interpreted as possible values positions currently held by the agent. A first, but crucial position's risk measurement will be whether its future value belongs or not to the subset risks deemed acceptable, according to the decision maker, who has the duty to assess the overall cost of these positions.

There is a trade-off between the severity of risk measurement, and the level of activity in the supervised domain. The axioms and characterizations presented alone do not indicate a risk measure to follow, so there are also strategic decisions to be taken into account when choosing the risk measure to be adopted.

The positions of agents may change due to the actions of the agent or the counterparties. In general, one can consider the risk of following a strategy (which specifies the position held on each date according to the different events and counterparty actions) over an arbitrary period of time.

3.1.2 Value at Risk - *VaR*

Value at Risk, or *VaR* is a central tool in risk, asset and portfolio risk. And it also plays a key role in systemic risk. *VaR* is defined as the maximum loss an asset/portfolio/institution can incur at a defined quantile α (Basilio et al., 2020).

VaR represents the value in potential risk of loss, regarding an asset portfolio, in a specific time period and specific probability α. This probability represents a quantile for risk. With a random variable X and a cumulative distribution function F that models the losses verified for a specific asset in a time period. Let F be a continuous function and X a continuous random variable, with $0 \leq \alpha \leq 1$, then VaR_α is defined as:

$$VaR_\alpha = F^{-1}(1 - \alpha) \tag{3.1}$$

A *VaR* of d days at α% confidence level means that on $(1 - \alpha)$% of d days, we will not see a loss higher than the *VaR*, but for the $(1 - \alpha)$% of times the loss will be higher.

For example, a 1 day *VaR* at the 99% confidence level of 5% means that on only 1 of every 100 days we will see a loss higher that 5% of the initial capital.

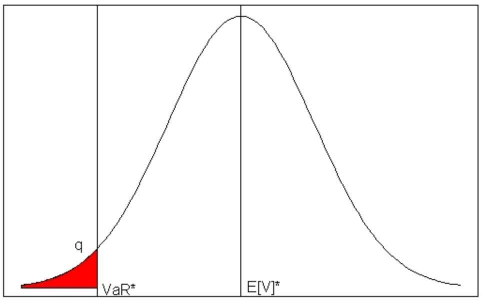


Figure 3.1: Graphic representation of *VaR*

The concept of *VaR* becomes central to the study of systemic risk, constituting a fundamental concept for the definition of several of the most significant systemic risk measures mentioned in the literature. Yet, this method faces challenges dealing with risk associated with events involving volatility component with dependencies between extremes values in distinct data sets and modelling extreme values with volatility.

3.1.3 Expected Shortfall

Shortfall Risk or Expected Shortfall (Brownlees and Engle, 2016) measure is an extension to the MES in that it includes in the calculation of this systemic risk the liabilities incurred and the size of the agent. Expected Shortfall corresponds to the expected value of the capital deficit of a given agent in the presence of an event (crisis) that affects the entire system. The agent with the highest Expected Shortfall value is the agent considered to be the agent that contributes most to the systemic risk, that is, the agents with the highest systemic risk. Expected Shortfall is formally defined as:

$$ES_{it\alpha} = E(X|X \geq VaR_{it\alpha}(X)) \tag{3.2}$$

where X represents the losses for an agent i at a confidence level α for time period of t and $0 \leq \alpha \leq 1$.

This definition also allows to express ES as:

$$ES_{it\alpha} = \frac{1}{1-\alpha}\int_{\alpha}^{1} VaR_{itq}(X))dq \tag{3.3}$$

The Expected Shortfall is an average of the worst $(1-\alpha)\%$ of losses. As $E[X|A]$ is the expectation of loss X occurs, given event A. The Expected Shortfall is therefore the probability-weighted average of all losses beyond *VaR*, and therefore the expected loss beyond *VaR*.

The Expected Shortfall is often considered a preferable risk measure than the *VaR* because, while *VaR* only tells us the maximum threshold for losses at a given confidence level, it does not tell us how much that loss could be. The Expected Shortfall tells us how much we expect to lose once a catastrophic loss happens. Also Expected Shortfall always satisfies subadditivity.

3.1.4 Coherent Risk Measures

According to the literature, for the measurement of risk, four characteristics are desirable for any proposed risk measure: monotonicity, translation invariance, positive homogeneity and subadditivity, which can be replaced by convexity.

These characteristics, known as the basic axioms of risk measures, are also the required properties to qualify a risk measure as coherent risk measure. Those axioms provide the conditions to define what are known as coherent risk measures.

The list of properties, proposed by Artzner et al. (1999) for a risk measure establishes the concept of a coherent risk measure, in case the risk measure complies with the requirements of that list, as opposed to the incoherent risk measure.

Considering X a random variable of returns and $\mathbb{M}$ the set of all risks defined for a real valued measurable functions on a probability space Ω, a coherent risk measure ρ is an application $\rho : \mathbb{M} \rightarrow \mathbb{R} \cup \{+\infty\}$ satisfying the following properties:

- Monotonicity - For all L_1, $L_2 \in \mathbb{M}$ and $L_1 \leq L_2$, then $\rho(L_1) \leq \rho(L_2)$
- Sub-additivity - If L_1, $L_2 \in \mathbb{M}$ then $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$. This property allows some important conclusions, such as: merge doesn't cause additional risk; diversification can reduce the risk; allows decentralized risk management.
- Translation Invariance - For all $L \in \mathbb{M}$ and all constant $m \in \mathbb{R}$ then also verify $\rho(L+m) = \rho(L)-m$. It means that removing the initial quantity m at this initial position and applying that same quantity to the reference instrument (without risk) decreases the risk measure by m.
- Positive homogeneity - For all $L \in \mathbb{M}$ and for $\lambda > 0$ then also verify $\rho(\lambda L) = \lambda\rho(L)$.
- Relevance - For $L < 0$, then $\rho(L) > 0$.

The concept of convex risk measure is an extension to the definition introduced by Artzner, as that is they required positive homogeneity as well, to which Frittelli and Gianin (2002) added a new axiom of convexity:

For all $L_1, L_2 \in \mathbb{M}$ and $0 \leq \lambda \leq 1$, one verifies that $\rho(\lambda L_1 + (1-\lambda)L_2) \leq \lambda\rho(L_1) + (1-\lambda)\rho(L_2)$.

The Convexity axiom states that risk can not be increased through diversification.

Axioms of Homogeneity and Sub-additivity are relaxed and replaced by Convexity axiom as presented. Monotonicity and Translation Invariance axioms remain valid.

A Standardization axiom is also added: $\rho(0) = 0$.

Similar to what has been described for coherent risk measures, also for convex risk measures is it possible to define a set of acceptance, such as $A_\rho = \{L \in \mathbb{M} : \rho(L) \leq 0\}$.

Even though *VaR* is one of the most popular risk measures, is not a coherent one, because it does not have sub-additivity property (Dowd, 2003). As risk measure, *VaR*, as shown by Dowd (2003), is not an adequate risk measure for tail losses, which are the kind of risks more important in terms of risk management. One possible way to measure tail risk is consider the expected loss above *VaR*. This will lead to the measure of Expected Tail Loss (ETL). This measure is formalized as:

$$ETL = E\Big[L|L > VaR\Big]$$

where L is a random variable of the loss.

Risk Measure	Coherence	Convexity	Monotonicity	Sub-additivity	Translation Invariance	Positive homogeneity	Relevance
VaR	×	✓	✓	×	✓	✓	✓
ES	✓	✓	✓	✓	✓	✓	✓
MSE	×	×	✓	✓	✓	✓	✓
CoVaR	✓	✓	✓	✓	✓	✓	✓
Δ *CoVaR*	×	✓	✓	✓	✓	✓	✓
ETL	×	×	✓	✓	✓	✓	✓

Table 3.1: Risk measures and related properties.

3.1.5 Systemic Risk Measures

Systemic risk can be defined as the underlying risk on a complex system compound by a set of agents which interact with each other. This interaction increases risk dissemination (Smaga, 2014). From this point of view, it is of particular interest to evaluate and model the impact of a failure of a particular agent over the remaining agents as a whole. Systemic failure can begin with the failure of some agents, being amplified both by the interaction mechanisms and by the feedback system itself, which can lead to successive failures that affect and cause a significant impact in part or even throughout the system.

To quantify systemic risk in a system at a point in time, or the occurrence of a certain event with systemic impact, one should consider the modelling, the measure to be used, data accessibility and collection conditions (Eijffinger, 2011). Modelling the interactions that occurs between agents in a given system, will answer three fundamental questions in a systemic risk model (Martínez-Jaramillo et al., 2010):

- Identify the contagion channels;
- Quantify the importance of systemic impact;
- Compute a failure probability.

Defined a probability space $(\Omega, \mathbb{F}, \mathcal{P})$ and let $\mathbb{M}$ be a set of risks which random variables are defined under that space over a fixed time interval Δ. $\mathbb{M}$ is assumed to be linear with constants such as $L_1, L_2 \in \mathbb{M}, m \in \mathbb{R}$ and $k > 0$, then $L_1 + L_2, L_1 + m$ and $kL_1 \in \mathbb{M}$. A risk measure is defined as a function $\rho \to \mathbb{R}$ where $\rho(L)$ is the amount needed to recover from a loss L for a position with an acceptable risk. This means a portfolio with $\rho(L) < 0$ is in an acceptable position, and does not need additional funds.

3.1.5.1 Marginal Expected Shortfall - MES

MES risk measure measures the marginal contribution of an institution i, in period t to the systemic risk. This systemic risk corresponds to the Expected Shortfall of the system.

Expected Shortfall represents the probability that the return falls below a given value, C. Thus, the ES is a conditional value of the loss, since the loss is greater than a certain value, in this case the MES. In this way, let r_{it} be the return for given agent i in a time period t, the conditional value of the ES is defined as:

$$ES_{it}(C) = E_t(r_{it} | r_{it} < C) \tag{3.4}$$

and for the entire system, ES comes therefore as:

$$ES_{system\ t}(C) = \sum_{i=1}^{N} w_i E_t(r_{it}|r_{it} < C) \tag{3.5}$$

where w_i is the relative weight of the agent i in the system. MES is the partial derivative of the ES taking into account the weight of the agent i in the system.

$$MES_{it}(C) = \frac{\partial ES_{mt}(C)}{\partial w_i} = E_t(r_{it}|r_{it} < C) \tag{3.6}$$

Marginal Expected Shortfall can be interpreted as an extension to the concept of marginal *VaR* from Expected Shortfall. In this risk measure the increase in the system risk due to the increase of the relative weight, determined by w_i, is measured. Thus, the higher the agent's MES, the greater the risk that the agent represents to the system.

3.1.5.2 Δ *CoVaR*

Δ *CoVaR* is a systemic risk measure defined by Adrian and Brunnermeier (2016), based on the Value-at-Risk concept, *VaR*. *VaR* measures the risk based on maximum expected loss for a given confidence interval *VaR*(α). *CoVaR* is the *VaR* risk *subject* to the occurrence of a specific event (typically a crisis), affecting the agent i at a point in time t and the returns of the agent i is r^{it}, as defined by $C(r^{it})$. There are several options to define a critical event $C(r^{it})$. One is to consider the *VaR*. This way, *CoVaR* is defined for a single agent i, as $CoVaR^i_\alpha$ by:

$$CoVaR^i_\alpha = Pr(r^{it} \leq VaR^{it}_\alpha) = \alpha \tag{3.7}$$

where r^{it} represents the returns of the agent j at a point in time t.

The $CoVaR^{j|i}_\alpha$ is the *VaR* of agent j conditional on agent i is in distress, this means at it's *VaR* level.

$$CoVaR^{j|i}_\alpha = Pr(r^{jt} \leq CoVaR^{j|i}_\alpha | r^{it} = VaR^i_\alpha) = \alpha. \tag{3.8}$$

The concept of Δ *CoVaR* is a representation of the difference between *VaR* of an entire system subject to that a critical event affects agent i and the *VaR* of the entire system when that event does not occur.

There are several options to define a critical event $C(r^{it})$. One is to consider a loss equivalent to *VaR*:

$$Pr(r^{jt} \leq CoVaR_{\alpha}^{j|C(r^{it})}|r^{it} = VaR_{\alpha}^{it}) = \alpha \tag{3.9}$$

and then the Δ *CoVaR* comes as:

$$\Delta CoVaR^{j|it}(\alpha) = CoVaR_t^{j|r^{it}=VaR_{\alpha}^{it}} - CoVaR_t^{j|r^{it}=Median(r^{it})}. \tag{3.10}$$

Other option is to consider a critical event when losses exceed *VaR* (Girardi and Ergün, 2013)

$$P(r^{jt} \leq CoVaR_{\alpha}^{j|C(r^{it})}|r^{it} \leq VaR_{\alpha}^{it}) = \alpha \tag{3.11}$$

$$\Delta CoVaR^{it}(\alpha) \leq CoVaR_t^{j|r_{it} \leq VaR_{it}(\alpha)} - CoVaR_t^{j|r_{it}=Median(r_{it})}. \tag{3.12}$$

This definition of *CoVaR*, by applying this slightly change, can now represent, instead the conditional *VaR* when the financial institution is at most at its *VaR* level. In this case the condition in relation to the *VaR* of the financial institution is now *less or equal than*, to represent the case when the loss exceeds the *VaR* of the financial institution.

By including the change in the way the financial distress of the agent s now defined, as it was shown by Mainik and Schaanning (2014), turns *CoVaR* risk measure in a increasing function of the dependence parameter, established between i and j. The initial definition proposed by Adrian and Brunnermeier (2011) does not comply with this property.

If we consider the special case where j is representing the system, in this way that $j \equiv s$, the $CoVaR^{s|i}_{\alpha,t}$ is equivalent to the *VaR* of the financial system , under the condition of institution i being in a situation of financial distress at point in time t.

The Δ*CoVaR* is then the individual contribution of agent i for the systemic risk at time t.

3.1.6 Methods to Estimate Risk Measures

The core of estimating a risk measure involves to manage how to describe the Profit and Loss (P&L) distribution of the portfolio, institution or in more generic terms an asset. There are many ways to estimate the Profit ans Loss but we can in general encompass them in three main categories as:

- Non Parametric;
- Parametric;
- Monte Carlo.

The Non Parametric Methods are based on the historical simulation approach, in it is most basic forms this is the simplest and most straightforward *VaR* method. To make this calculation, observations are simply ordered from largest to smallest. The observation that follows the threshold level denotes the *VaR* limit.

This method has several know disadvantages:

- Implicit assumption that you expect future performance to follow the same pattern as in the past.
- Is unable to adjust for changing economic conditions.
- Observations that happened a long time ago have the same weight as a recent one.

Parametric Approach consists on adjusting a theoretical probability distribution to the return distribution, most commonly a normal or a t-student distribution.

The advantages of using this method are:

- Not limited to only past scenarios;
- Fast and easy to calculate;

and the disadvantages normally associated with the parametric approach are:

- Normality assumption for modelling risk, as it assumes that the returns are normally distributed;
- The returns are assumed to be serially independent, meaning that no prior return should influence the current return;

The Monte Carlo method consists in describing the asset prices diffusion in order to get the simulated returns, instead of trying to describe those returns distribution. The advantages of using this method are:

- Accurate for non-linear instruments;
- Full distribution of potential distributions;

and the disadvantages normally associated with the Monte Carlo method are:

- Complexity of the risk model involved;
- Computational power required.

3.1.7 Conclusion

As discussed, a coherent risk measure has several desirable properties. Perhaps the most interesting of all is the property of sub-additivity which encourages diversification, an intuitive way to reduce risk. Another important conclusion is that a coherent risk measure provides an upper limit of the total risk, when analysing the sum of the risk of the individual actions. A consistent set of well defined axioms are vital to allow to compare between different risk measures.

3.2 Normal Distribution Assumption

There is an open ongoing discussion about the application of normal distribution to model financial related data and the fairness of its inclusion in risk models in finance.

In fact, the use of the normal distribution in order to model financial returns is considered a traditional assumption in finance, since Markowitz developed his portfolio theory in 1952 (Haugen and Haugen, 2001), as it is the backbone of traditional (mean-variance) premise (Fabozzi et al., 2002).

Despite that, some recent research work has been rejecting this assumption based on the study of skewness, kurtosis and in particular in heavy tail, in the distribution of financial returns (Jondeau et al., 2007).

Even though it is noticeable that financial returns distributions are at least close to a bell shaped curve, even if this does not translate directly for a normal distribution.

By using the graphs below, we can notice a bell shape pattern on the returns of financial institutions. We also can notice an heavy tail behavior, best described in this case, in terms of extreme tails by the t-student distribution curve than by the normal distribution curve.

The following figures, are comparing the histogram of the returns sample corresponding to the period in analysis, 1998 to 2018, for two financial institutions HSBA, and BBVA, that will be compared with three theoretical distributions, the normal distribution curve, a t-student distribution curve, and a Cauchy distribution curve as well. These two graphs helps to interpret how the financial institution returns distributions compares with the theoretical distribution.

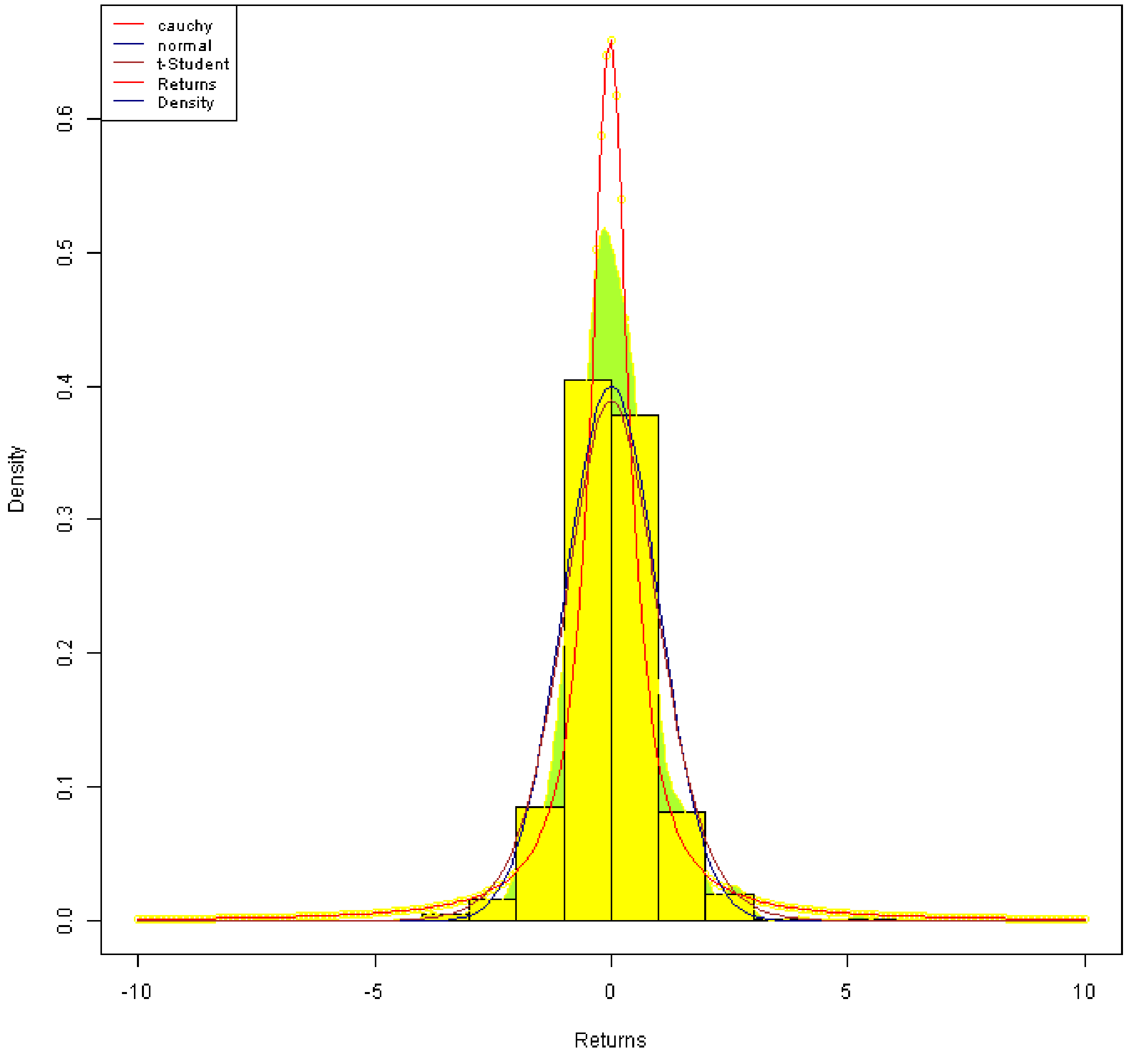


Figure 3.2: HSBA financial institution returns histogram vs Normal, Cauchy and *t*-student theoretical distributions

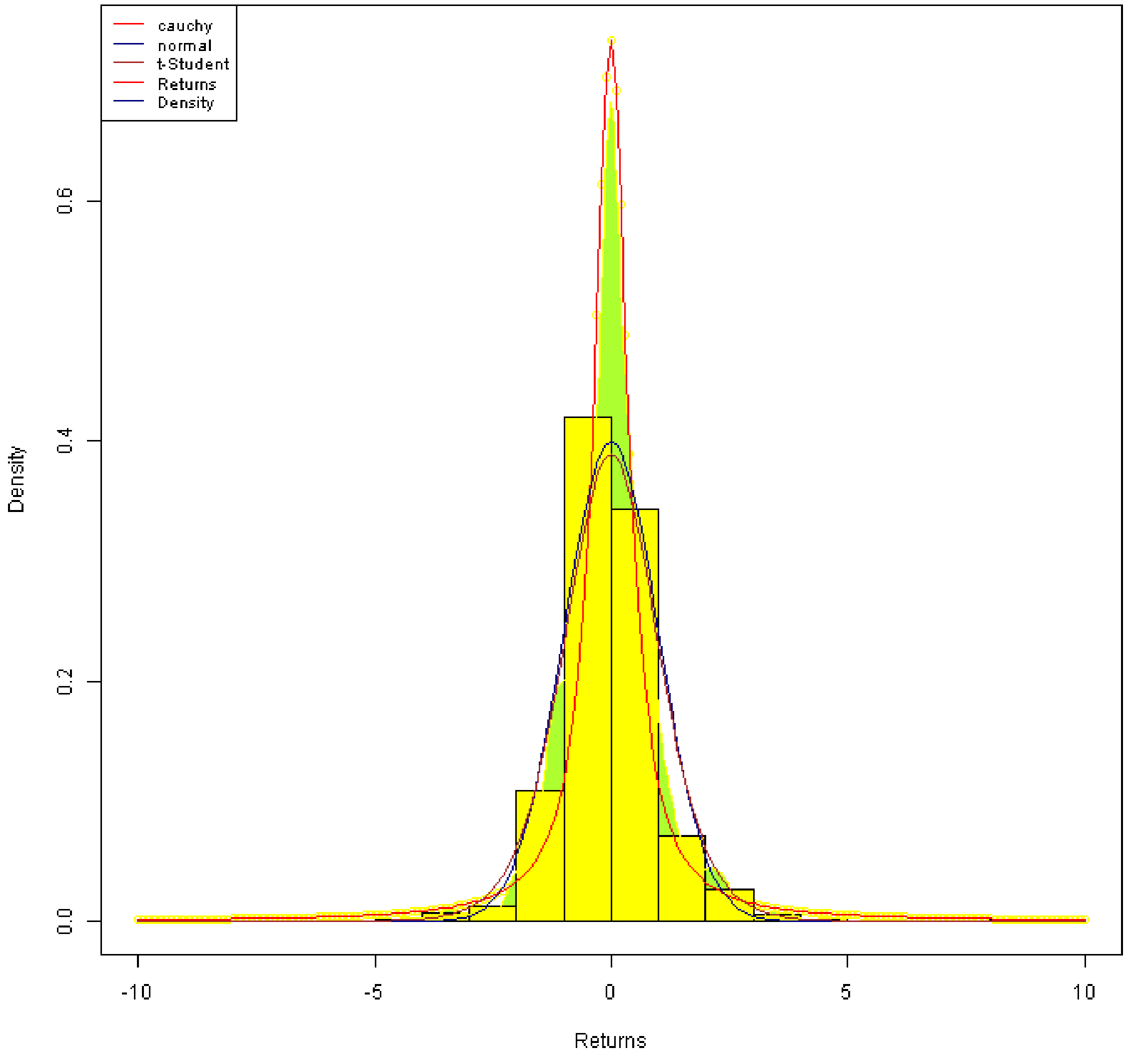
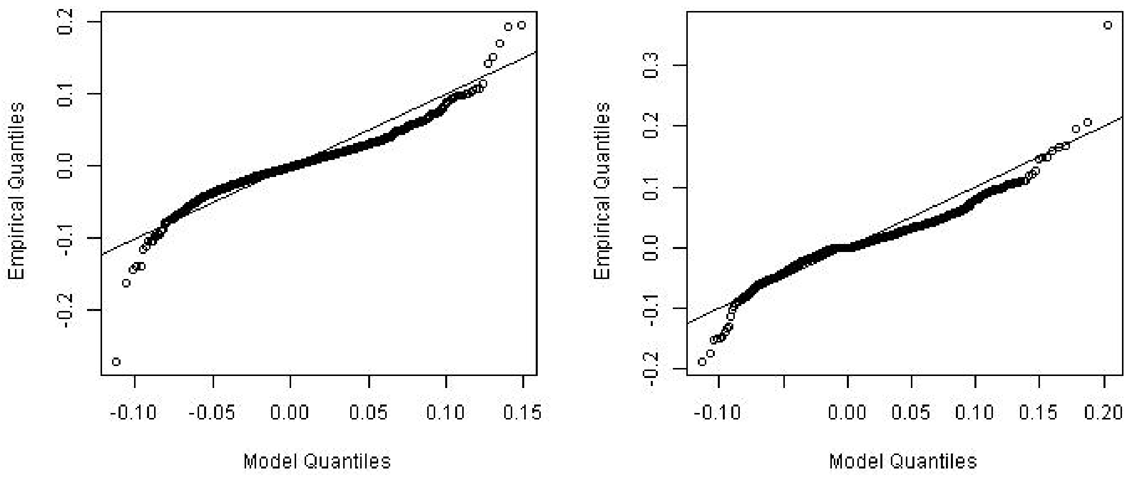


Figure 3.3: BBVA financial institution returns histogram vs Normal, Cauchy and *t*-student theoretical distributions

An additional complementary graphical analysis in order to compare the behavior of the returns across the distinct quantiles and compare that behavior with the expect behavior for a Normal population is to use a Q-Q plot.

The Q-Q plot, or quantile-quantile plot, is a type of graph that allows us to assess if it is plausible to assume that a data set came from some theoretical distribution such as a Normal. Even if it is just a visual check, and somewhat subjective, it allows us to see at-a-glance if our assumption is plausible, and how the assumption is eventually violated and which data points cause that violation.



a HSBA returns vs normal distribution b BBVA returns vs normal distribution

Figure 3.4: Financial institution returns Q-Q plot vs normal distribution

As shown on the above results, we should consider the financial institutions returns (and financial system returns too) as not normally distributed.

It becomes also clear we have a different behavior on extreme tails of returns distribution, and in the tail it is not following a normal behavior. This pattern must be included in the modelling.

3.3 Extreme Values Theory

Extreme value theory is used to model unusually low or high value data, when compared with the expected values, that is observed in the tail of the distribution.

These unusual events represented by extreme data points are often complex to model and requires advanced techniques to fit a distribution that includes the heavy tail with satisfactory results.

In statistics the concept of mixture distribution concerns the combination of two or more distributions.

Mixture distributions are of importance in order to model complex processes allowing for a more flexible approach than using a single distribution.

Normal distributions mixtures are a possible choice to model this type of complex processes and are formed by linear combinations of two or more Normal distributions, as a weighted sum of that Normal distributions, in order to form a new distribution.

One promising application for mixture models, is to model heavy tailed distributions. Even though financial returns are usually modeled as normally distributed this assumption proved to be inconsistent with empirical evidence. Asset returns are considered heavy tailed which means that extreme values are more likely to happen in practise than suggested by a Normal distribution. These discrepancies could cause estimation errors of major impact if for example we are using those assumptions to build risk measures such as *VaR*. Then normality assumption can lead to inappropriate risk management measures (Kimball et al., 2000).

Some research work involving mixture models assumes that several distributions are drawn from the same probability density functions. On the other hand heterogeneous mixture models can be a valid option for modelling more complex phenomena and improving modelling capabilities.

An option is to model asset returns distribution with a mixture of distributions, a mixture of Normal distributions, a mixture of Normal and Gumbel distributions and also Normal distribution and GEV as the extreme value distributions.

The modelling approach presented explores the capabilities of a mixture model in order to fit the financial returns (Razzaghi and Kodell, 2000).

3.3.1 Rare Events

One of the *VaR* method limitations is its dependence on estimates of the extreme values (in the tails) of loss distribution. Traditional methods used to calculate the *VaR* are based on the entire distribution of the data, which shows difficulties estimating the distribution at the tails. One possibility is to use extreme values theory.

In modeling the maximum of a random variable, extreme value theory plays the same fundamental role that the central limit theorem plays in modeling the sums of random variables. In either case, the theory tells us what the limits of distribution are. One of the fundamental assumptions underlying the application of the extreme value theory is that the data that make up the sample are all independent and identically distributed.

One method is identify the extreme values (maximum or minimum) verified in each period (block), called *block maxima* method. Another method is to define a level that splits the sample into extreme values and *"standard"* values.

The *block maxima* method is usually associated with data series seasonally while the limit value method has a broader applicability. The mathematical construction that supports each approach is however different as follows.

3.3.1.1 *Block maxima* method

The distribution associated to *block maxima*, designated as M_n, where n represents the block dimension is given by Fisher and Tippett, (Fisher and Tippett, 1928) and Embrechts et al. (1998).

Theorem 3.3.1 (Fisher)**.** *Let* X_n *be a sequence of random variables i.i.d. If there exists constants* $c_n > 0, d_n \in \mathbb{R}$ *and some non-degenerated distribution function F,* *such as:*

$$\lim_{n\to\infty} P\Big(\frac{\max\{X_1, X_2, \ldots, X_n\} - d_n}{c_n} \leq x\Big) = F(x) \tag{3.1}$$

then the limit function F belongs to one of the three extreme value distributions:

$$\textit{Fréchet:} \qquad \Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ e^{-x^{-\alpha}}, & x > 0 \end{cases} \tag{3.2}$$

$$\textit{Weibull:} \qquad \Psi_\alpha(x) = \begin{cases} e^{-(-x^{\alpha})}, & x \leq 0 \\ 1, & x > 0 \end{cases} \tag{3.3}$$

$$\textit{Gumbel:} \qquad \Lambda(x) = e^{-e^{-x}}, x \in R \tag{3.4}$$

Based on these three distributions it is possible to obtain a generalization, which is called the Generalized Extreme Value Distribution (GEVD), developed by Von Mises and Jenkinson (Bücher and Segers, 2017). The generalization is obtained by a transformation of the parameters, considering ε, the shape parameter of GEV, in which:

- in case of Fréchet distribution $\varepsilon = \alpha^{-1}$
- in case of Weibull distribution $\varepsilon = -\alpha^{-1}$
- finally, in the case of Gumbel distribution, it is interpreted as the case where $\varepsilon = 0$

The distribution of the extreme values is represented by a parameter only:

$$\text{GEV:} \qquad H_\varepsilon(x) = \begin{cases} e^{-(1+\varepsilon x)^{-\frac{1}{\varepsilon}}}, & \varepsilon \neq 0 \\ e^{-e^{-x}}, & \varepsilon = 0 \end{cases} \tag{3.5}$$

Usually the limit distribution is not known a priori, which makes the generalized representation of the extreme distribution particularly useful. The previously defined GEV function is the limit distribution of normalized extremes. Since in practice we do not know the true distribution of the returns and therefore there is no information on the normalization constants c_n and d_n, it is possible to use a three parameter specification, which in turn corresponds to the non-standard *maxima* distribution limit.
GEV distribution (Coles, 2013):

$$H_{\varepsilon,\lambda,\nu}(x) = H_{\varepsilon}\Big(\frac{x-\nu}{\lambda}\Big) \qquad x \in D, \qquad D = \begin{cases}]-\infty, \nu - \frac{\lambda}{\varepsilon}[, & \varepsilon < 0 \\ ]-\infty, \infty[, & \varepsilon = 0 \\ ]\nu - \frac{\lambda}{\varepsilon}, \infty[& \varepsilon > 0 \end{cases} \tag{3.6}$$

The two additional parameters ν and λ correspond to the location and scale, representing here the normalization constants. However, in analyzing one of the extreme value distributions, the values that are usually relevant are the percentiles, also referred to as GEV return values. Applying a variable substitution again such that:

$$R^k = H^{-1}_{\varepsilon,\lambda,\nu}(1 - \frac{1}{k}) \tag{3.7}$$

Replacing the parameters $\varepsilon, \lambda, \nu$ and their expected values $\hat{\varepsilon}, \hat{\lambda}$ e $\hat{\nu}$ we then obtain:

$$\hat{R^k} = \begin{cases} \hat{\nu} - \frac{\hat{\lambda}}{\hat{\varepsilon}}\Bigg(1 - \Big(1 - \log\Big(1 - \frac{1}{k}\Big)\Big) - \varepsilon\Bigg), & \varepsilon \neq 0 \\ \hat{\mu} - \hat{\lambda}\log\Bigg(-\log\Big(1 - \frac{1}{k}\Big)\Bigg), & \varepsilon = 0 \end{cases} \tag{3.8}$$

The interpretation of R^k has a very specific meaning. Taking as an example $k = 5$, which represents the number of periods under analysis, for example years, we have $\hat{R}^5$ of 10 means that the annual loss observed over a 5-year period will be greater than 10% for at least one time on average.

3.3.1.2 Excess Distribution

An alternative to model extreme values is the *Peak Over Threshold* (POT) distribution, an excess distribution function, where the distribution of observed values above a certain value, is considered.

Considering a F distribution function of a random variable X, the function F_u represents the distribution function of x values above the u limit. F_u, is the conditional excess distribution function, formalized by:

$$F_u(y) = P(X - u \leq y | X > u), \qquad 0 \leq y \leq x_F - u \tag{3.9}$$

X is a random variable, u is a previously defined limit value, y represents the excess, and $x_F \leq \infty$ is the limit of F on the right tail. The extremes occur "near" the upper end of distribution support, hence intuitively asymptotic behavior of the excesses must be related to the distribution function F in its right tail near the right endpoint. Thus,

$$x_F = \sup\{x \in R : F(x) < 1\} \tag{3.10}$$

represents the right end point of F.

Establishing the maximum domain of attraction of F as the function extreme value function H if exists $c_n > 0$ and $d_n \in R$ such as

$$\lim_{n\to\infty} P\Big(\frac{\max\{X_1, X_2, \ldots, X_n\} - d_n}{c_n} \leq x\Big) = H(x) \tag{3.11}$$

Then we say $F \in MDA(H)$

Extreme value theory permit to apply a theorem to estimation F_u (Pickands III, 1975):

Theorem 3.3.2 (Pickands)**.** *Let F be a excess distribution. The conditional excess distribution function $F_u(y)$, for u large, is well approximated by $G_{\varepsilon,\sigma}(y)$, also called Generalized Pareto Distribution (GPD):*

$$F \in MDA(H) \iff \lim_{u \to x_F} \sup_{0<x<x_F-u} |F_u(y) - G_{\varepsilon,\sigma}(x)| = 0 \tag{3.12}$$

where:

$$G_{\varepsilon,\sigma}(y) = \begin{cases} 1 - \left(1 + \frac{\varepsilon}{\sigma} y\right)^{-\frac{1}{\varepsilon}}, & \varepsilon \neq 0 \\ 1 - e^{y/\sigma}, & \varepsilon = 0 \end{cases} \tag{3.13}$$

for $y \in [0, x_F - u]$ *if* $\varepsilon \geq 0$ *and* $y \in [0, -\frac{\sigma}{\varepsilon}]$, *if* $\varepsilon < 0$.

This theorem establishes that for some σ function, to be estimated from the data, the excess distribution function F_u converges to GEV for a large u.

3.4 Heavy Tails Distributions

In statistics the term heavy tail is associated with distributions with a relatively high probability of extreme outcomes.

Even though there is not a definitive and formal definition of heavy tail distribution usually it is assumed that a distribution has a heavy tail when the probability in the tail is thicker when compared with a normal distribution. Taken as example Cauchy distribution, we can notice a ticker tail in Cauchy distribution when compared with a Normal distribution.

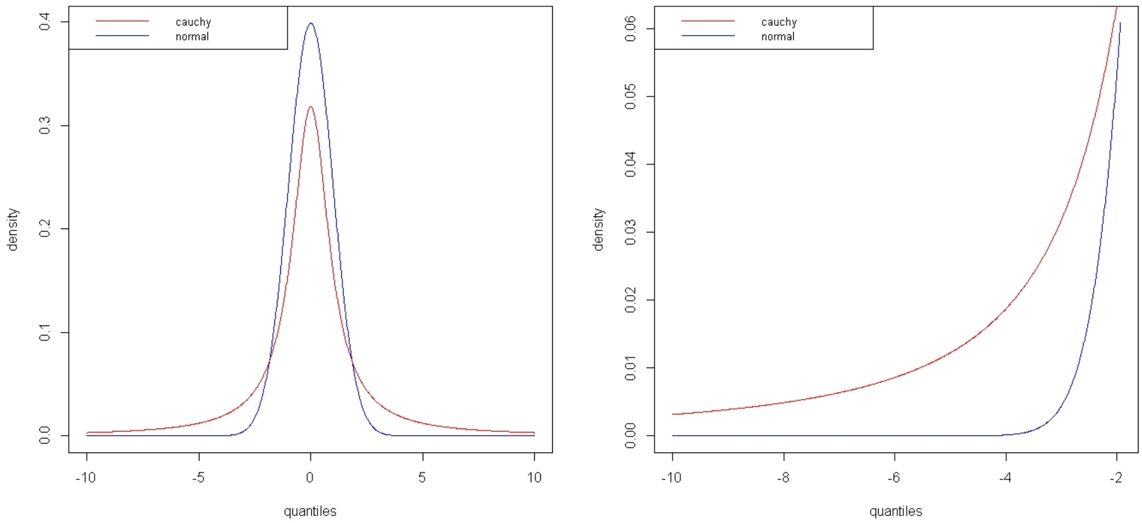


Figure 3.5: Heavy tail distribution versus normal

Cauchy among others, like t-Student distribution for example, are known as heavy tailed distributions.

In order to close the gap in terms of modelling the extreme tail of financial institution returns as a first approach we will study an approach of univariate fitting where we will include the modelling of the tails of the distribution.

To evaluate how heavy tails impact our financial data set, we will compare results from fitting three theoretical distributions: Normal, Cauchy and t-Student.

3.4.1 Cauchy Distribution

In probability theory, Cauchy distribution is the probability distribution whose probability density function is defined by:

$$f(x; x_0, \gamma) = \frac{1}{\pi\gamma \left[1 + \left(\frac{x - x_0}{\gamma}\right)^2\right]} = \frac{1}{\pi\gamma}\left[\frac{\gamma^2}{(x - x_0)^2 + \gamma^2}\right] \tag{3.14}$$

where x_0 is the location parameter and γ is the scale parameter and x $\in \mathbb{R}$, In its standard form, Cauchy distribution represents the special case where $x_0 = 0$ and $\gamma = 1$ and is represented by:

$$f(x) = \frac{1}{\pi(1 + x^2)}$$

In its standard form, the median is 0, and first and third quantiles are -1 and 1 respectively snd we can write Cauchy(0,1).

If X is a random variable with standard Cauchy distribution then, let $x_0 \in \mathbb{R}$ be an arbitrary value and $\sigma > 0$. The random variable Y, defined as:

$$Y = x_0 + \sigma X$$

also follows a Cauchy distribution with median x_0 and whose first quantile is $\mu - \sigma$ and the third quantile is $\mu + \sigma$. The probability density function is therefore defined as:

$$f(y) = \frac{1}{\pi\sigma\Big(1 + \frac{(x-x_0)^2}{\sigma^2}\Big)}.$$

Standard Cauchy distribution can also be defined as a ratio of two normal distributions. Let X and Y be two independent random variables. If $X \sim N(0,1)$ an $Y \sim N(0,1)$ then:

$$\frac{X}{Y} \sim \text{Cauchy}(0,1).$$

In addition to its application in physics (Alzaatreh et al., 2016), Cauchy distribution is commonly used in models in finance to represent deviations in returns from the predictive model (Harris, 2017). The reason for this is that practitioners in finance are wary of using models that have light-tailed distributions as Normal, on their returns, and they generally prefer to go the other way and use a distribution with very heavy tails as Cauchy. The history of finance has a vast record of catastrophic predictions based on models that did not have heavy enough tails in their distributions. The Cauchy distribution has sufficiently heavy tails as its moments does not exist, and so it is an ideal candidate to give an error term with extremely heavy tails (Guerrero-Cusumano, 1996).

3.4.2 *t*-Student Distribution

Also *t*-Student distribution is an option to deal with heavy tails and have been an option for researchers too (Terzić et al., 2014).

The *t*-Student distribution with v degrees of freedom can be defined as the random variable T as:

$$T = Z.\sqrt{\frac{v}{W}} \sim t$$

where Z is a random variable that follows a standard normal distribution, $Z \sim N(0,1)$ and W is a random variable that follows a chi-squared distribution with v degrees of freedom, $W \sim X_v^2$, then the standardized quotient of the two follows a *t*-Student distribution with v degrees of freedom.

The *t*-Student distribution could be very useful for financial analysis as we can adapt to the tail behavior of the data. In its conventional form, *t*-Student could not be a very flexible model because of the absence of a location and a scale parameter. An alternative definition can then be described as:

$$T \sim \nu \Longrightarrow S = \nu + \lambda T \sim t_v(\nu, \lambda^2)$$

with ν as location parameter and $\lambda^2(\frac{v}{v-2})$ as scale parameter. The tail decay is polynomial, that is, the density function goes to zero proportional to $x^{-(v+1)}$ for $x \to \infty$. For low values of v this is a much slower rate than for the Normal distribution (Theodossiou, 1998).

As the typical assumption involving the Normal distribution had failed is now hardly accepted due to the probabilities at the extremes are much larger than those supported by Normal distributions. This invalid assumption is specially dangerous for risk management related applications.

3.5 Modelling Financial Returns with a Mixture Distributions

As empirical evidence and results suggest, the normality assumption of financial institution returns is not verified as it is heavy tailed, and modelling this behavior using only one distribution has shown limitation, one option is to consider an approach that uses more than one distribution.

Heavy tailed distributions can be modeled instead by a mixture of distribution. In this case, we will first approach the problem by applying a mixture of normal distributions.

Assuming the returns are following a stochastic process for a financial institution i as:

$$R_{it} = \lambda_{it} R_{it}^{\alpha} + (1 - \lambda_{it}) R_{it}^{\beta} \tag{3.15}$$

where $R_{it}^{\alpha} \sim N(\mu_{\alpha}, \sigma_{\alpha})$, $R_{it}^{\beta} \sim N(0, \sigma_{\beta})$, and λ_{it} is 1 with probability p and 0 otherwise. These three random variables R_{it}^{α}, R_{it}^{β} and λ_{it} are independent of each other.

Depending on λ and with probability p the distribution to apply will be $N(\mu_{\alpha}, \sigma_{\alpha})$, for example for the most normal situations. With probability (1 - p), λ will be equal to 0 and the distribution to apply is $N(0, \sigma_{\beta})$ and it could be interpreted as an exceptional case.

The challenge is now centered in the estimation of the parameters involved; $p, \mu_{\alpha}, \sigma_{\alpha}, \sigma_{\beta}$.

Despite several alternative methods that are possible to use to estimate the parameters of a mixture of normal distribution, if we consider the traditional maximum likelihood method, it could then be formulated as:

$$l\Big((p,\mu_\alpha,\sigma_\alpha,\sigma_\beta)|R_{it}\Big) = \sum_t \log\left[\frac{p}{\sigma_\alpha}\exp\left(-\frac{\Big(R_t-\mu_\alpha\Big)^2}{2\mu_\alpha^2}\right)+\frac{1-p}{\sigma_\beta}\exp\left(-\frac{(R_t^2)}{-\mu_\beta}\right)\right]$$

Due to the existence of both poles and saddle points, the maximization of the mixture of normals likelihood could be challenging and the global maximum for that function could not exist (Hamilton, 1991).

This problem however could be described as an incomplete data problem since the data we observe in our sample can be viewed as a subset of the "complete" data.

3.5.1 Expectation-Maximization Algorithm

The Expectation-Maximization (EM) Algorithm is an appropriate tool for that type of problems. EM Algorithm is an approach for maximum likelihood estimation in the presence of latent variables and can be used to predict the latent variables values with the condition that the general form of the probability distribution governing those latent variables is known.

The algorithm is implemented as an iterative procedure given a set of incomplete data and considering a set of starting parameters will iterate on two steps

- Expectation step (E – step). Using the available observed data and the current model parameters the missing or latent variables are estimated by.
- Maximization step (M – step). After estimating missing values, this step will be used to update the parameters by computing the parameters that maximize the expected log-likelihood of the model based on the values estimated on E-step.

EM Algorithm includes statistical considerations to compute the maximum-likelihood (ML), source distribution that would have created the observed data, including the effects of counting statistics. Specifically, it assigns greater weight to high-count elements of a profile and less weight to low-count regions (Dempster et al., 1977).

3.5.2 EM Algorithm and Mixture of Normal Distributions

Considering the case of a mixture of Normal distributions let $\mathbf{x} = (\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n)$ be a sample of n i.i.d. observations of a mixture of two Normal and $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ the latent variables that determine the component where the observation originates (Reynolds, 2009).

$$X_i|(Z_i = 1) \sim \mathcal{N}_d(\boldsymbol{\mu}_1, \Sigma_1) \quad \text{and} \quad X_i|(Z_i = 2) \sim \mathcal{N}_d(\boldsymbol{\mu}_2, \Sigma_2),$$

where

$$\mathrm{P}(Z_i = 1) = \tau_1 \quad \text{and} \quad \mathrm{P}(Z_i = 2) = \tau_2 = 1 - \tau_1$$

The goal of this process is to estimate the parameters for the mixture of Normal distributions:

$$\theta = \Big(\boldsymbol{\tau_1}, \boldsymbol{\tau_2}, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \Sigma_1, \Sigma_2\Big)$$

The likelihood function therefore is:

$$L(\theta; \mathbf{x}, \mathbf{z}) =$$
$$\exp\left\{\sum_{i=1}^{n}\sum_{j=1}^{2}\mathbb{I}(z_i = j)\Big[\log\tau_j - \tfrac{1}{2}\log|\Sigma_j| - \tfrac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_j)^\top\Sigma_j^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_j) - \tfrac{d}{2}\log(2\pi)\Big]\right\}.$$

However, while the maximum likelihood estimators will provide good results for short tailed distributions, this in not true for heavy tailed distribution or even with the presence of outliers as demonstrated by (Schuster and Gregory, 1981), leading to inconsistent estimates.

In order to mitigate this inconvenience in the maximum likelihood method, was proposed as alternative the application of Markov chain Monte Carlo methods (Levine and Casella, 2001).

3.5.3 Extreme Value Mixture Models

The idea behind the use of Mixture of Extreme Value distribution is to combine the flexibility of using a distribution to capture the main component, also identified as the bulk of the distribution (Fúquene Patiño, 2015), that could be for example a Normal, and also the tails, as extreme values. With this mixture model one will get an entire distribution function by splitting the distribution in a bulk component and the tail components.

There are several approaches that consider only one tail and also approaches that consider both, lower and upper tail, the tail on the left and the tail on the right respectively. In this case the mixture function will be compounded potentially by a mixture of distribution from distinct families.

In our case we are specially interested in exploring a mixture of a Normal distribution as a bulk distribution, with two Gamma tail distributions in both, upper and lower tail.

MacDonald et al. (2011) proposed a two tailed mixture model where the standard kernel density estimator is spliced with two extreme value tail models.

This model uses a kernel density estimator to estimate the non-extreme value distribution and GPD to estimate the tail distribution. A boundary-corrected kernel density estimator is also used in the case of a population with bounded support. This kernel density estimator assumes a particular kernel, in this case the normal density, which is centered at each data point, and uses only one parameter to define bandwidth. The model uses also the standard cross-validation likelihood to define bandwidth, combined with the likelihood for the peaks over threshold tail model, to give a full likelihood for all of the observations.

The term tail fraction refers to the proportion of the distribution above the threshold. This parameter will be identified by Φ_u and u represents the threshold.

The distribution function comes as:

$$F(x|\Theta) = \begin{cases} \phi_{u_l} 1 - G(-x| - u_l, \sigma_{u_l}, \epsilon_l), & x < u_l \\ \\ H(x|\mu, \sigma), & u_l \leq x \leq u_r \\ \\ (1 - \phi_{u_r}) + \phi_{u_l} G(x|u_r, \sigma_{u_r}, \epsilon_r), & x > u_r \end{cases}$$

where $\phi_{u_l} = H(u_l|\mu, \sigma)$ and $\phi_{u_r} = 1 - H(u_r|\mu, \sigma)$ and $H(\cdot|\mu, \sigma)$ is the normal distribution with mean μ and standard deviation σ. $G(\cdot| - u_l, \sigma_{u_l}, \epsilon_l)$ and $G(\cdot| - u_r, \sigma_{u_r}, \epsilon_r)$ are GPD distributions for lower and upper tails respectively.

3.5.4 Conclusion

The results obtained by comparing the goodness of fit obtained by applying distinct statistics modelling techniques, highlighted a concern regarding the quality of the global adjustment versus the quality of the adjustment on the tails of the distribution. In certain applications the analyses of the tail of the distributions is of major importance, as for example in risk analyses. The results obtained for *VaR* estimates for each model implemented also showed that more complex models could be advantageous as they are more flexible in adapting to the tail of the distribution providing better adjustments.

Complex phenomena also requires more complex models and the complexity of certain phenomena, like behavior of financial returns, requires more complex and versatile models.

3.6 Copulas

3.6.1 Introduction

Copulas can be described as functions that formalize the dependence structure between random vectors and the joint distributions generated from the marginals of the given random vector (Nelsen, 2003).

Copula properties are similar to properties of joint distributions but they allow us to split the marginal behavior from the dependence structure between the variables in the respective joint distribution function. This property gives copulas unique modeling flexibility, which explains the wide interest in copulas for modeling the dependence structure between variables.

By using copulas it will be possible to isolate the dependency structure in a multivariate distribution.

3.6.2 Copula Definition

As a multivariate function Nelsen (2007), a copula is a multivariate distribution function from the unit d-cube $[0,1]^d$ to the unit interval [0,1], as C : $[0, 1]^d : \to [0, 1]$ is a cumulative distribution function (CDF) with uniform marginals and satisfying a set of properties as follows:

- C(1,...,1, u_i, 1, ..., 1) $= u_i$. The i^{th} marginal distribution is obtained by setting $u_j = 1$ for $j \neq i$ and assuming u_j is uniformly distributed. This property allows to conclude that, if the value of the *d-1* variables are known with marginal probability one, then the d outcomes of the joint probability is equal to the one with uncertain outcome (u_i).

- C(u_1, ..., u_d) $= 0$ if $\exists i \in \{1, \ldots, d\}$ such that $u_i = 0$. By this property we have that if one variable has the marginal probability zero, then the joint probability of all outcomes is zero. This property is also known as the grounded property.
- C(u_1, ..., u_d) is non-decreasing in each component, u_i. This property assures that the joint probability will never be negative, as C of any d-dimensional interval is also non-negative.
- C is bounded by the so called Fréchet bounds as follows:

$$max\left\{\sum_{i=1}^{d} u_i + 1 - d, 0\right\} \leq C(u) \leq min\{u_1, ..., u_d\} \tag{3.16}$$

Theorem 3.6.1 (Sklar)**.** *Sklar (1959) Let F be a joint distribution function with marginals* $F_1, ..., F_n$. *There exist a copula C such that for all* $x_1, ..., x_d$ *in* $[-\infty, \infty]$ *and* $i = 1, ..., d$ *that*

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_1(x_d)) \tag{3.17}$$

If F_i is continuous for all $i = 1, ..., d$, then C is unique, otherwise C is uniquely determined only on $Ran(F_1) \times ... \times Ran(F_d)$ where $Ran(F_i)$ denotes the range of the CDF, F_i. In other way, considering a copula, C, and univariate CDF's, $F_1, ..., F_d$, then F is a multivariate CDF with marginals $F_1, ..., F_d$.

If the marginal distributions, $F_1, ..., F_n$, are continuous, it is possible to shown that

$$F_i(F_i^{-1}(y)) = y \tag{3.18}$$

doing now $x_i = F_i(u_i)$ and using the last result we obtain

$$C(u) = F(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)) \tag{3.19}$$

By using this theorem, it is possible to conclude that copulas are joint distribution functions and also that these joint distribution functions can be written in terms of a copula and of the marginal distributions. This way the exercise of modeling joint distributions can be reduced to an exercise of modeling copulas (Schweizer, 1991).

Another consequence of Sklar's theorem is that copulas also represent the dependence between the variables that come as result of the splitting the joint distribution into a copula and the marginals. This is the reason why copulas are also called dependence functions.(Hürlimann, 2003).

Since copulas are also dependence measures, they allow us to distinguish the perfect dependence and the independence as well.

The following useful copula properties can this way be established:

- Independence: The random variables X and Y are said to be independent if they have the product copula, i.e., $C(F_1(x), F_2(y)) = F_1(x) \times F_2(y)$.
- Invariant property: The invariant property can be considered as the most important property of copulas. This property establishes that the dependence structure that the copula describes is invariant under monotone transformations of the marginal distributions. As an example, if a logarithmic transformation is applied to the marginal distribution, this will not affect the copula. Because copulas have invariant property under monotonic transformations, this makes them a powerful tool in several applications.

3.6.3 Measures of Dependence

Since the copula of a multivariate distribution describes its dependence structure, it might be appropriate to use measures of dependence which are copula-based. The bivariate concordance measures Kendall's tau and Spearman's rho, as well as the coefficient of tail dependence, can, as opposed to the linear correlation coefficient, be expressed in terms of the underlying copula alone.

Two random variables X and Y are said to be dependent or associated if they do not satisfy the independence property:

$$(X, Y) = F_1(x) \times F_2(y), \tag{3.20}$$

where $F_1(x)$ and $F_2(y)$ are the marginal distributions functions of the random variables X and Y.

It is desirable that a dependence measures complies with a set of four properties (Straumann, 2001). Let δ express a simple scalar measure of dependence. Then :

- $\delta(X, Y) = \delta(Y, X)$, known as the condition of symmetry.
- $-1 \leq \delta(X, Y) \leq 1$, known as the condition of normalization.
- $\delta(X, Y)$= 1 ,then *(X, Y)* are co-monotonic and if $\delta(X, Y)$= 1, then *(X, Y)* are counter-monotonic.
- $\delta(T(X), Y)$= $\delta(Y, X)T$ and $-\delta(Y, X)$ where T is a strictly monotonic transformation of T: $\mathbb{R} \rightarrow \mathbb{R}$ of X

Understanding the dependence structure of copulas is vital to understanding their properties. There are three principal measures of dependence:

- The usual Pearson, i.e. linear, correlation coefficient is only defined if second moments exist. It is invariant under positive linear transformations, but not under general strictly increasing transformations. Moreover, there are many fallacies associated with the Pearson correlation.
- Rank correlations only depend on the unique copula of the joint distribution and are therefore invariant to strictly increasing transformations. Rank correlations can also be very useful for calibrating copulas to data.
- Coefficients of tail dependence are a measure of dependence in the extremes of the distributions.

3.6.3.1 Pearson Linear Correlation

Pearson linear correlation is the most widely used type of dependence measures. The Pearson linear correlation measures the direction and the degree to which one variable is linearly related to the other variable. For non-degenerating random variables X and Y, the linear correlation coefficient is defined by:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{3.21}$$

where $cov(X,Y)$is the covariance between X and Y, σ_X and σ_Y are the standard deviations of X and Y respectively. The Pearson linear correlation takes values between -1 and 1. When $\rho_{(X,Y)} = -1$ then the variables X and Y are perfectly dependent by an increasing relationship. When $\rho_{(X,Y)} = 1$ then the variables X and Y are perfectly dependent by a decreasing relationship. If the X and Y are independent, the correlation between them is equal to zero (Sedgwick, 2012).

Even though Pearson linear correlation is a popular correlation and dependence measure it doesn't complies entirely with those properties of dependence measures enunciated above. Independence implies that the correlation is equal to zero, but zero correlation does not imply that the random variables are independent.

Another Pearson linear correlation shortcoming is that it is not defined when the variance of X or Y is not finite. This way linear correlation is not a suitable dependence measure to deal with distributions with fatter tails, which is the case in the most financial time series data (Cherubini et al., 2004).

Pearson linear correlation also does not satisfy the invariance property. This way Pearson linear correlation is not invariant under non-linear monotone transformations since linear correlation does not only depend on the joint distribution of random variables, but also on their marginals.

In order to solve Pearson linear correlation limitations we need to use others dependence measures as Spearman's rank correlation and the Kendall's rank correlation. Before going through them it is also needed to introduce the concept of concordance.

3.6.3.2 Concordance

The observations (X_1, Y_1) and (X_2, Y_2) are said to be concordant if $(X_1 > Y_1)$ and $(X_2 > Y_2)$ or if $(X_1 < Y_1)$ and $(X_2 < Y_2)$. This means that large (small) values of the random variable X are associated with large (small) values of the random variable Y. If the opposite is true, discordance arises.

So,let $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$ be a random sample of n observations from a vector *(X, Y)* of continuous random variables. There are $\binom{n}{2}$ distinct pairs $(X_i, X_i), (X_j, Y_j)$ of observations each pair is either concordant or discordant. Kendall's tau is given by (Nelsen, 2003) :

$$\frac{\text{(number of concordant pairs)–(number of discordant pairs)}}{\text{total number of pairs}} \tag{3.22}$$

3.6.3.3 Spearman Correlation Coefficient

Spearman's correlation coefficient is defined as a statistical measure of the strength of a monotonic relationship that occurs between paired data. In a sample it is identified by and is by design defined as follows (Akoglu, 2018):

$$\rho_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \tag{3.23}$$

Where n is the number of the paired ranks, and d is the difference between the paired ranks. Based on this definition, the Spearman's rank correlation coefficient can be interpreted as the Pearson linear correlation coefficient between the two ranked variables. The variables are ranked by assigning the highest rank to the highest value.

The most attractive property of Spearman's rank correlation is that it does not make any assumption in relation to the frequency distribution of the two variables involved. Another important feature associated to Spearman's rank correlation is the ability to describe the non-linear dependence between the two variables.

3.6.3.4 Kendall's tau

Tau correlation coefficient is defined as the probability of concordance minus the probability of discordance for given a pair, $(X_i, Y_i), (X_j, Y_j)$ of observations randomly selected from the sample. This definition can therefore be extended to the entire population and we obtain the population version of this measure. Analogous to the sample version, we let $(X_1, Y_1), (X_2, Y_2)$ be independent random vectors with a common joint distribution. The population version of Kendall's tau is:

$$\tau = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \tag{3.24}$$

Kendalls's rank correlation is this way a non-parametric correlation measure that measures the difference between the probability of concordance and the one of discordance between the r.v.s X and Y. Except that Kendall's rank correlation represents a probability, it is considered equivalent to Spearman's rank correlation. Kendall's tau is given by:

$$\tau_K = \frac{2(C - D)}{n(n-1)} \tag{3.25}$$

Where C is the number of concordant pairs and D is the number of discordant pairs. Like the Spearman's rho, Kendall's rank correlation is invariant under monotonic non-linear transformations of the underlying variables. Tail dependence is when the correlation between two variables increases as you get "further" in the tail (either or both) of the distribution as tail dependence measures the dependence between X and Y in the upper-right and lower-left quadrant of the joint distribution function. This dependence measure is the most appropriate when interested in the probability that one variable exceeds or falls below some given threshold.

The parameter to represent the asymptotic lower tail dependence, noted by, is the conditional probability in the limit that one variable takes a very low value, given that the other also takes a very low value.

3.6.3.5 Tail-dependence Coefficient

Tail dependence exists when the correlation between two variables increases as we move "further" in the tail (either or both) of the distribution and Tail dependence describes the limiting proportion that one margin exceeds a certain threshold given that the other margin has already exceeded that threshold. It is also a dependence measure that takes in account the concordance between extreme values (tail) of the joint distribution.

- The parameter of asymptotic lower tail dependence, λ_L, is the conditional probability in the limit that one variable takes a very low value, given that the other also takes a very low value.
- The parameter of the asymptotic upper tail dependence λ_U, is the conditional probability in the limit that one variable takes a very high value, given that the other also takes a very high value.

The asymptotic tail dependence parameters for copula function are given by (Joe, 1997):

$$\lambda_L = \lim_{\alpha \to 0} P(Y > F_Y^{-1}(\alpha) | X > F_X^{-1}(\alpha)) \tag{3.26}$$

$$\lambda_U = \lim_{\alpha \to 1} P(Y > F_Y^{-1}(\alpha) | X > F_X^{-1}(\alpha)). \tag{3.27}$$

Tail dependence measures are independent of the marginal distributions of the variables and that they are invariant under strictly monotone transformations of X and Y. There is a saying in finance that in times of stress, correlations will increase. With the bivariate tail dependence it is possible to measure the amount of dependence in the upper and lower quadrant tail of a bivariate distribution.

3.6.4 Methods for Generating Copulas

As sanctioned in the literature (Nelsen, 2007), exists several approaches to constructing copulas. Some of these methods generate copulas related to specific applications but other ones are also suitable in order to obtain more generic results.

3.6.4.1 Inversion of Marginals

The simplest method to generate a copula function is perhaps the Inversion Method, which results directly from the Skal's theorem. In this method the copula function is generated from a given join distribution.

Let H be a joint continuous distribution with marginals F and G, and let F^{-1} and G^{-1} be the existing respective inverse functions.

Since the join distribution function can be expressed as a function of the marginals as:

$$H(x, y) = C(F(x), G(y)). \tag{3.28}$$

The corresponding copula function is then generated by applying the inverse transformation $x = F^{-1}(u)$ and $y = F^{-1}(v)$, then the copula function can also be expressed as a function of the respective inverse function of the marginals as:

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)). \tag{3.29}$$

As example of Inverse Marginals method we can use the Gaussian Copula. Let $\phi_\rho(x, y)$ be a standard bivariate normal distribution function and ρ the correlation coefficient.

The Gaussian copula is then represented as:

$$C(u, v; \rho) = \phi_\rho(\phi^{-1}(u), \phi^{-1}(v)). \tag{3.30}$$

where $\phi(x)$ and $\phi(y)$ represents the univariate standard normal margin respectively.

3.6.4.2 Algebraic Method

The derivation of a copula can be obtained from relationship between marginals based on independence. This independence will after be modified in a way to include a dependence parameter. The process can be illustrated using as example the process to derive the Gumbel's bivariate logistic distribution. For this case the join distribution $H(x, y)$ is:

$$H(x, y) = (1 + e^{-x} + e^{-y})^{-1}. \text{ let } \frac{1 - H(x, y)}{H(x, y)} \tag{3.31}$$

Let $\frac{1 - H(x, y)}{H(x, y)}$ be the bivariate survival odds ratio. Then,

$$\frac{1 - H(x, y)}{H(x, y)} = e^{-x} + e^{-y} \tag{3.32}$$

$$= \frac{1 - F(x)}{F(x)} + \frac{1 - G(y)}{G(y)} \tag{3.33}$$

where $F(x)$ and $G(y)$ are the univariate marginals. If we consider first the case of independence, with $H(x, y) = F(x)G(y)$ we can rewrite the equation above as:

$$\frac{1 - H(x, y)}{H(x, y)} = \frac{1 - F(x)G(y)}{F(x)G(y)} \tag{3.34}$$

$$= \frac{1 - F(x)}{F(x)} + \frac{1 - G(y)}{G(y)} + \frac{1 - F(x)}{F(x)} \frac{1 - G(y)}{G(y)}. \tag{3.35}$$

Considering a dependence parameter θ, the dependence bivariate odds ratio can be expressed as:

$$\frac{1-H(x,y)}{H(x,y)} = \frac{1-F(x)}{F(x)} + \frac{1-G(y)}{G(y)} + (1-\theta)\frac{1-F(x)}{F(x)}\frac{1-G(y)}{G(y)} \tag{3.36}$$

Applying the transformation $u = F(x)$ and $v = G(y)$ we can obtain in a similar way:

$$\frac{1-C(u,v;\theta)}{C(u,v;\theta)} = \frac{1-u}{u} + \frac{1-v}{v} + (1-\theta)\frac{1-u}{u}\frac{1-v}{v}. \tag{3.37}$$

and consequently:

$$C(u,v;\theta) = \frac{uv}{1-\theta(1-u)(1-v)}. \tag{3.38}$$

3.6.4.3 The Mixtures Method

Considering a copula C, the correspondent lower and upper bounds can be defined as C_L and C_U respectively, and $C^{\perp}$ a product copula. It is possible obtain a new copula by applying a convex sum. As Fréchet upper bound is also a copula (Trivedi and Zimmer, 2007), the constant λ_1, defined as $0 \leq \lambda_1 \leq 1$, then the convex sum , in the upper bound can be expressed as:

$$C^M = \lambda_1 C^{\perp} + (1+\lambda_1)C_U \tag{3.39}$$

and C^M is also a copula, denoted as a mixture copula. This copula is a also a special case in terms of Fréchet copulas, C^F that are also defined as:

$$C^F = \lambda_1 C_L + (1-\lambda_1-\lambda_2)C^{\perp} + \lambda_2 C_U \tag{3.40}$$

where the constants λ_1 and λ_2 are defined in such way that $0 \leq \lambda_1$, $0 \leq \lambda_2$ and $\lambda_1 + \lambda_2 \leq 1$.

3.6.4.4 The Generator Function

A function $\phi : \mathbb{R}_+ \rightarrow I$ is said to be a generator if it is continuous, decreasing and $\phi(0) = 1$, $\underset{t\rightarrow+\infty}{lim}\phi(t) = 0$ and is strict decreasing on $[0,\ t_0]$, where $t_0 > 0$. If the function ϕ is invertible or strictly decreasing on $\mathbb{R}_+$ then the generator is defined as *strict* (Najjari, 2018). If ϕ is strict then $\phi(t) > 0$ for every $t > 0$, and $\underset{t\rightarrow+\infty}{lim}\phi(t) = 0$:

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t) & 0 \leq t \leq \phi(0) \\ 0 & \phi(0) \leq t \leq \infty \end{cases} \tag{3.41}$$

Let ϕ be a convex function defined on I and valued in $[0, \infty]$ such that ϕ is strictly decreasing and fulfills $\phi(1) = 0$. Let $\phi^{[-1]}$ be the pseudo-inverse of ϕ, that is, $\phi^{[-1]} =$ [0,∞], and $\phi^{[-1]} =$ I.

Depending on the form assumed by ϕ(t), different copulas exhibiting different properties are defined. More details and a complete review of these copulas can be found at Nelsen (2003) and Joe (1997).

3.6.5 Copula Families

There are two parametric families of copulas namely implicit copulas and explicit copulas. Implicit copulas do not have a simple closed form. Copulas from this family are implied by well-known multivariate distribution functions. The most known copulas from this class of copulas are the Gaussian copula and the Student's t-copula. Explicit copulas also called Archimedian copulas represent a class of copulas that are broadly used to model the dependence structure in the data. This class of copulas became very popular due to the easiness of the construction and the implementation of their copulas (simple closed form) next to the wide range of dependence that they allow for. We will focus on bivariate Archimedean copulas, the most important copulas within this class and some examples will be discussed with details as the Clayton copula, the Gumbel copula and the Frank copula. The main properties of each copula will be discussed and how these copulas are related to dependence measures.

3.6.5.1 Archimedian Copulas

An important class of copulas is provided by Archimedean copulas. A (bivariate) copula C is said Archimedean if a generator functions ϕ exists, then a bi-variate Archimedean copula exists and C is given by:

$$C(u_1, u_2) = \phi^{[-1]}(\phi(u_1) + \phi(u_2)). \tag{3.42}$$

One advantage demonstrated by Archimedean copulas is the way they relate with dependence measures.

The relation between the copula generator function and Kendall's in the bivariate case can be given by Joe (1997):

$$\tau = 1 + 4\int_0^1 \frac{\phi(v)}{\phi'(v)} dv. \tag{3.43}$$

Archimedean copulas is of major interest. Indeed, it makes possible to estimate Kendall's tau from a data sampled and subsequently compute the dependence parameter of various generators through equation 3.43 and then to construct the corresponding copulas C.

It was also demonstrated a relation between Archimedian copulas and Tail dependency (Joe, 1997).

Let ϕ a strict generator. If $\phi'(0)$ is finite and $\neq 0$ then:

$$C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)). \tag{3.44}$$

If the copula has upper tail dependence, then $\frac{1}{\phi'(0)} = -\infty$

The coefficients of upper tail dependence and lower tail dependence are given by:

$$\lambda_U = 2 \lim_{s \to \infty} \frac{\phi'(s)}{\phi'(2s)} \tag{3.45}$$

$$\lambda_L = 2 - 2 \lim_{s \to 0} \frac{\phi'(s)}{\phi'(2s)} \tag{3.46}$$

Frank copula. Lets use the Frank copula as an example of application for the generator method. Assuming in this case the generator defined as (Cherubini et al., 2004):

$$\phi(t) = -log\left(\frac{1-e^{\delta t}}{1-e^{\delta}}\right) \text{ and the inverse:}$$
$$\phi(t)^{-1} = -\frac{1}{\delta} log\left[1-\left(1-e^{\delta}\right)e^{-t}\right]$$

Based on the generator function defined above, the Frank's copula family results as:

$$C(u,v) = -\frac{1}{\delta} log\left(1 - \frac{\left(1-e^{\delta u}\right)\left(1-e^{\delta v}\right)}{\left(1-e^{\delta}\right)}\right) \tag{3.47}$$

Frank copula is used for modeling data characterized by weak tail dependence and is given by:

$$C(u_1,u_2;\theta) = -\theta^{-1} log\left(1 + \frac{(e^{-\theta u_1}-1)(e^{-\theta u_2}-1)}{(e^{-\theta}-1)}\right). \tag{3.48}$$

Frank copula does not have lower neither upper tail dependence, and in this case we have $\lambda_L = \lambda_U = 0$.

Clayton Copula. One of the main application of Clayton copula is in the research and analysis of correlated risks because of a recognised ability to capture lower tail dependence. Clayton copula is not derived from a multivariate distribution function, but does have simple closed forms and is designed also as an explicit copula. The Bivariate Clayton copula expression is then given by:

$$C(u_1,u_2;\theta) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{\frac{1}{\theta}} \tag{3.49}$$

where θ is the copula parameter restricted on the interval $(0,\infty\ )$. If $\theta = 0$ then the marginal distributions become independent.

A relation between Kendal tau and the Clayton Copula parameter can be established as (Nelsen, 2003):

$$\rho_\tau(X, Y) = \frac{\theta}{\theta + 2}. \tag{3.50}$$

The parameter of lower tail dependence for this copula can be established as (Cherubini et al., 2004):

$$\lambda_L(X, Y) = 2^{\frac{1}{\theta}} \tag{3.51}$$

$$\lambda_U(X, Y) = 0, \tag{3.52}$$

so Clayton copula is lower tail dependent copula.

Gumbel copula. Gumbel copula is used to model asymmetric dependence. This copula is able to capture strong upper tail dependence and weak lower tail dependence. These copula also is not derived from a multivariate distribution function, but do have simple closed forms and is designed also as an explicit copula. The bivariate Gumbel copula is given by:

$$C(u_1, u_2; \theta) = e^{-[(-\log u_1)^{-\theta} + (-\log u_2)^{-\theta}]^{\frac{1}{\theta}}} \tag{3.53}$$

where θ is the copula parameter restricted on the interval $[1, \infty[$. When θ goes to 1, marginals become independent. As in the case of Clayton copula, the Gumbel copula represents only the case of independence and positive dependence. The relation with the Kendall's tau is given by:

$$\rho_\tau(X, Y) = \frac{\theta - 1}{\theta}. \tag{3.54}$$

The parameter for upper tail dependence is:

$$\lambda_U = 2 - 2^{\frac{1}{\theta}}. \tag{3.55}$$

The parameter for lower tail dependence is given by $\lambda_L = 0$.

3.6.5.2 Implicit Copulas

Some copulas are implied by well-known multivariate distribution functions, and are said to be implicit copulas. Two examples are the Gaussian and the Student's t-Copula.

Gaussian copula. The bivariate Gaussian copula is given by:

$$C(u_1, u_2; \theta) = \Phi(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \theta) \tag{3.56}$$

$$= \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi(1-\theta)^(1/2)} exp(-\frac{x^2 - 2\theta xy + y^2}{2(1-\theta^2)}) dxdy \tag{3.57}$$

where Φ^{-1} is the inverse of the standard univariate Gaussian cumulative distribution function and θ is the parameter of the copula that represents the linear correlation between $\Phi^{-1}(u_1)$ and $\Phi^{-1}(u_2)$.

t-Copula. The bivariate t-Copula relates with the bivariate t-distribution where two dependent variables following a t-student distribution with v degrees of freedom and a correlation coefficient of θ. The bivariate t-Copula is given by:

$$C(u_1, u_2; \theta) = \Phi(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \theta) \tag{3.58}$$

$$= \int_{-\infty}^{t_v^{-1}(u_1)} \int_{-\infty}^{t_v^{-1}(u_2)} \frac{1}{2\pi(1-\theta)^(1/2)} \left(1 + \frac{x^2 - 2\theta xy + y^2}{v(1-\theta^2)}\right)^{-\frac{v+2}{2}} dxdy \tag{3.59}$$

where $t_v^{-1}(\dot{)}$ represents the inverse of the cumulative distribution function t_v, θ is the linear correlation coefficient between $t_v^{-1}(u_1)$ and $t_v^{-1}(u_2)$.

3.6.5.3 Joe Copula

Joe copula or asymmetric negative logistic copula was introduced by Joe and Hu (1996). The generator function for Joe Copula is given by:

$$\phi(t) = -\log[1 - (1-t)^\theta]$$

The copula is defined as:

$$C_\theta(u_1, u_2) = 1 - [(1-u_1)^\theta + (1-u_2)^\theta - (1-u_1)^\theta(1-u_2)^\theta]^{1/\theta} \ ,$$

where $1 \leq \theta < \infty$

The Joe copula is not compatible with negative dependence and τ in this case can be expressed based on θ parameter as:

$$\tau = 1 + \frac{4}{\theta} D_j(\theta)$$
$$D_j(\theta) = \int_t^1 \frac{[\log(1-t^\tau)](1-t^\tau)}{t^{\theta-1}} dt \ ,$$

where $0 \leq \tau \leq 1$.

Joe copula is then a right tailed copula.

3.6.5.4 EVT Copulas

Copulas are also an option to model extreme events, and joint extreme events. Extreme-value copulas provides appropriate models to deal with dependence structure between rare events (Capéraà et al., 2000).

Let $X_i = (X_{i1}, ..., X_{id}), i \in 1, ..., n$, be a *i.i.d.* random vector with a joint distribution F with d margins defined as $F_1, ..., F_d$, and a related copula C_F. Considering now the vector of component maxima, as:

$$M_n = (M_{n,1}, ..., M_{n,d}), \text{ where } M_{n,j} = Max_{i=1}^n X_{ij} \tag{3.60}$$

The joint distribution of a maxima is then M_n a and the marginals of that joint distribution becomes as $F_1^n, ..., F_d^n$, and the copula C_n of M_n comes as:

$$C_n(u_1, ..., u_d) = C_F(u_1^{\frac{1}{n}}, ..., u_d^{\frac{1}{n}})^n, \text{ where } (u_1, ..., u_d) \in [0,1]^d \tag{3.61}$$

This way, as n grows to infinity, the extreme copula become as the limit of C_n.

A copula C is an extreme-copula if exists a copula C_F that

$$C_F(u_1^{\frac{1}{n}}, ..., u_d^{\frac{1}{n}})^n \to C_n(u_1, ..., u_d),\ (n \to \infty)$$

for all $(u_1, ..., u_d) \in [0, 1]^d$. C_F is then a *domain of attraction* of C.

The extreme copula definition presented above also implies and matches with the definition of copula max stable (Gudendorf and Segers, 2010).

Performing a linear expansion of the logarithm and the exponential function, the domain-of-attraction equation can be written as:

$$\lim_{t \to 0} t^{-1}(1 - C_F(1 - t_{x_1}, ..., 1 - t_{x_d})) = -\log\ C(e^{-x_1}, ..., e^{-x_d}) = l(x_1, ..., x_d)$$

for all $(x_1, ..., xd) \in [0, \infty[^d$ which is a tail dependence function convex and homogeneous of order one (Drees and Huang, 1998). If we characterize this homogeneity using the *Pickands dependence function*, $A : \delta_{d-1} \to [\frac{1}{d}, 1]$. We can interpret A as a function defined on the $(d-1)$-dimensional unit, and $\delta_(d-1)$ as:

$$\delta_{d-1} := \Big\{(w_1, ..., w_d) \in [0, 1]^d : \sum_{i=1}^{d} w_i = 1\Big\}$$

and l become:

$$l(x_1, ..., x_d) = (x_1 + ... + x_d)A(w_1, ..., w_d) \text{ where } w_j = \frac{x_j}{x_1 + ... + x_d}$$

for $(x_1, ..., x_d) \in]0, \infty[^d$. With these results, extreme copula can be expressed, using A as:

$$C(u_1, ..., u_d) = exp\Big[(\sum_{j=1}^{d} \log u_j)A(\frac{\log u_1}{\sum_{j=1}^{d} \log u_j}, ..., \frac{\log u_d}{\sum_{j=1}^{d} \log u_j})\Big]$$

As function A is convex also satisfies the condition:

$$max(w_1, ..., w_d) \leq A(w_1, ..., w_d) \leq 1 \text{ for all } (w_1, ..., w_d) \in \delta_{d-1}$$

These results will be very useful to study and define the bivariate extreme copula.

For the bivariate case we have then $\delta_1 = (1-t,t) : t \in [0,1]$ form $R^2 \rightarrow [0,1]$ as definition, we can conclude then a bivariate copula is an extreme-value copula only if (Mai and Scherer, 2011)

$$C(u,v) = (uv)^{A\frac{\log(v)}{\log(uv)}}, \text{ where } (u,v) \in]0,1]^2 \tag{3.62}$$

where $A : [0,1] \rightarrow [\frac{1}{2},1]$ is convex and satisfies the fowling condition

$$t \vee (1-t) \leq A(t) \leq 1 \text{ for all } t \in [0,1]$$

and in this case:

- Upper bound with $A(t) = 1$ means independence, where $C(u,v) = uv$
- Lower bound represent $A(t) = t \vee (1-t)$ means perfect dependence, where $C(u,v) = u \wedge v$

where $\vee$ means maximum and $\wedge$ mean minimum. Also $A(t) \leq 1$ implies that $C(u,v) \geq uv$, and this means extreme-value copulas are necessarily positive quadrant dependent.

3.6.6 Conditional Probability Based on Copulas

As we are interested in dependence, conditional probabilities play an important role in modelling systemic risk too (Ahooyi, 2015).

Also, the conditional probability can be defined by using a copula (Käärik et al., 2011).

Let C denote the copula function defined for a bivariate function H, in such as that $H(x,y) = C(F_X(x), F_Y(y))$.

Given the bivariate random variables (X,Y) one can define the bivariate joint distribution defined as:

$$F_{XY}(x,y) = P(X \leq x, Y \leq y)$$

By Sklar theorem (Nelsen, 2003), we have a joint distribution dependent in the terms of the marginal distributions F_X and F_Y. This way there is a cumulative distribution such as:

$$F_{XY}(x, y) = C(F_X(x), F_Y(y))$$

where C represents a copula function on $[0, 1]^2$ space, and defined for all *(x, y)*.

By applying a variable transformation such as

$$u = F_X(x)$$
$$v = F_Y(y)$$

with $F_X(x)$ and $F_Y(y)$ uniformly distributed in [0,1], we have that

$$C(u, v) = F_{XY}\left(F_X^{-1}(u), F_Y^{-1}(v)\right)$$

where $F_X^{-1}(u)$ and $F_Y^{-1}(v)$ are the marginal generalized inverse distributions.

By definition the conditional probabilities is defined as

$$P(X \leq x, Y \leq y | X = x)$$

and by Bayes theorem it comes that:

$$P(X \leq x, Y \leq y | X = x) = \frac{P(X \leq x, Y \leq y)}{P(X = x)}$$

and

$$P(X \leq x, Y \leq y | X \leq x) = \frac{P(X \leq x, Y \leq y)}{P(X \leq x)}$$

As we have a bivariate probability distribution, this expressions can also be established by using copulas as:

$$Pr(X \leq x, Y \leq y | X \leq x) = \frac{C\Big(F_X(x), F_Y(y)\Big)}{F_X(x)}$$

As prove, we have that

$$\begin{aligned} P(X \leq x, Y \leq y | X = x) &= \lim_{\epsilon \to 0} P\Big(X \leq x, Y \leq y | x \leq X \leq x + \epsilon) \\ &= \lim_{\epsilon \to 0} \frac{F(x+\epsilon, y) - F(x, y)}{F_X(x+\epsilon) - F_X(x)} \\ &= \lim_{\epsilon \to 0} \frac{C\Big(F_X(x+\epsilon), F_Y(y)\Big) - C\Big(F_X(x), F_Y(y)\Big)}{F_X(x+\epsilon) - F_X(x)} \\ &= \frac{\partial C\Big(F_X(x), F_Y(y)\Big)}{\partial F_X(x)} \end{aligned}$$

if we apply the variable transformation discussed previously:

$$u = F_X(x) \text{ and } v = F_Y(y)$$

then it comes that

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

and, the same way

$$\begin{aligned} P(X \leq x, Y \leq y | X \leq x) &= \frac{P(X \leq x, Y \leq y)}{P(X \leq x)} \\ &= \frac{C(F - Y(y), F_X(x)}{F_X(x)} \end{aligned}$$

applying the variable transformation:

$$\begin{aligned} u &= F_X(x) \text{ and } v = F_Y(y) \\ P(X \leq x, Y \leq y | X \leq x) &= \frac{C(u, v)}{v} \end{aligned}$$

The conditional probability results presented above are extremely important when we want to establish and analyse the dependence, and in special the dependence in the tail, of the bivariate distribution.

The existence of tail dependence relates to clustering of extreme events, that is of major interest for risk management. Generically it relates with the propensity of huge losses occurring together, which also relates to dependence and conditional concepts.

We can consider two distinct tails, that we will designate as Upper tail and Lower tail.

Upper tail is defined as :

$$\lambda_u = \lim_{\alpha \to 1} Pr\Big(Y > F_Y^{-1}(\alpha) | X > F_X^{-1}(\alpha)\Big)$$

If $\lambda_u \in]0, 1]$ then X and Y are asymptotically dependent otherwise, if $\lambda_u = 0$ then X and Y are asymptotically independent.

Assuming that F_X and F_Y are continuous functions and $\overline{C}$ represents the survival copula, the previous equation can be rewritten as:

$$\begin{aligned}\lim_{\alpha \to 1} P\Big(Y > F_Y^{-1}(\alpha) | X > F_X^{-1}(\alpha)\Big) &= \lim_{\alpha \to 1} \frac{P\Big(X > F_X^{-1}, Y > F_Y^{-1}\Big)}{P(X > F_X^{-1})} \\ &= \lim_{\alpha \to 1} \frac{\overline{C}(\alpha, \alpha)}{1 - \alpha} \\ &= \lim_{\alpha \to 1} \frac{2\alpha - 1 + C(1 - \alpha, 1 - \alpha)}{1 - \alpha}\end{aligned}$$

where a function $\overline{C}(u, v)$ represents the survival copula function which is defined as

$$\overline{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$$

As we have the parameter for a specific copula function, we can then proceed to the substitution of those parameters and derive the upper tail.

On other hand we have Lower tail defined in similar way:

$$\lambda_l = \lim_{\alpha \to 0} P\Big(Y > F_Y^{-1}(\alpha) | X > F_X^{-1}(\alpha)\Big)$$

Like before X and Y are asymptotically dependent in lower tail. If $\lambda_l \in]0,1]$ and asymptotically independent in lower tail if $\lambda_l = 0$.

Applying the same arguments as in upper tail dependence, the lower tail dependence can also be determined by using a copula approach such as:

$$\lambda_l = \lim_{\alpha \to 0} \frac{C(\alpha, \alpha)}{\alpha}$$

It can be proved that this limit indeed exists for every known parametric model (Joe and Kurowicka, 2011) and this can be considered one of the most popular ways to estimate lower tail dependence.

3.6.7 Conclusion

One advantage of using copulas is the possibility to separate univariate marginal distributions from the multivariate dependence structure that describes how they are coupled (Joe, 1997). While learning these marginals is easy, learning the copula is more difficult and requires models that represent a broad range of dependence patterns. Two-dimensional copulas have a set of parametric copula models available (Nelsen, 2003). As some examples of the two-dimensional copulas, one can mention the Gaussian copula, Student copula, Clayton, Gumbel or Frank copula. In each of those families we can find a different dependence structure between two random variables.

If there exist many options for parametric models for two dimensional copulas, for more than two dimensions the number and expressiveness of families of parametric copulas is more limited. A solution to this problem is given by pair copula constructions, vine copulas (Bedford et al., 2002). Vines can be used to specify multivariate distributions by specifying various marginal distributions and the ways in which these marginals are to be coupled.

Chapter 4

Proposed Methodology

4.1 Introduction

The application of copulas in the area of risk and finance already has considerable amount of literature and research work. In the case of systemic risk and when it becomes necessary for the copula model to include asymmetries and dependence at the extremes the choice has fallen also on vine copulas (Joe and Kurowicka, 2011) in some of previous work.

The systemic risk measure *CoVaR* can also be calculated using a conditioned probability, defined based on two random variables, one representing the return of a financial institution i and the other representing the global return of the financial system:

- Y representing the financial system returns
- X_i representing the financial institution i returns

Considering the random variables X and Y, *CoVaR* is defined as the *VaR* of Y conditioned by X, more specifically because X is faced with a relevant event that impacted its returns and its *VaR* has been reached or violated. In a formal way this statement can be represented as:

$$CoVaR_{\alpha\beta t}(Y|X_i) = VaR_{\beta}(Y|X_i \in E) \tag{4.1}$$

where E represents a set of critical events affecting financial institution i, considered for instance as a crisis. Those critical events are usually defined based on a threshold over which we will consider that event as a critical impact over the financial institution i returns.

It follows from the current definition of *CoVaR* that there is a dependency relationship between the random variables X and Y.

As discussed previously, this dependency situation can be modeled using a copula function approach.

4.2 Dependence

As discussed previously, tail dependence is a measure of the dependence established between two random variables in the tail of their joint distribution. As if tail dependence can be included as a representation of systemic risk, however, there are copula structures that do not include this type of dependence and so the probability of simultaneous extreme events.

The proposed methodology takes into account the different dependency structures that occur in different quantiles of the joint distribution that fits the series of returns under analysis. Even the traditional copula approach is not flexible enough to modulate different dependency structures, which change according to the quantile in question, so it is necessary to include a second adjustment and include dependency structures in the model that represent dependency extremes. in times of stress from the financial institution or crisis. Assuming that the structure of dependence and correlation in extreme situations will change, it is necessary to reflect this change in the modeling process:

- without crisis, in the median
- in the presence of a crisis, on the tail of return distribution

The dependence on the tail (extreme) of the joint distribution of X and Y is of special interest since *VaR* refers to the value at risk or the maximum expected loss. In this way *VaR* will normally (but not necessarily always) be located on the left tail of the return distribution.

The tail dependence is defined as the limiting proportion that one of the margins exceeds a specific threshold assuming the other margin has already been in a position that exceeded a given threshold.

There are plenty of possible definitions for tail dependence. The following approach was provided, by Joe (1997).

Let $(Y, X)^\top$ be a two-dimensional random vector. We say that $(Y, X)^\top$ is (bi-variate) upper tail-dependent if:

$$\lambda_U \stackrel{\text{def}}{=} \lim_{v \uparrow 1} \mathrm{P}\left\{Y > F^{-1}(v) \mid Y > G^{-1}(v)\right\} > 0 \tag{4.2}$$

in the case that the limit exists.

F^{-1} and G^{-1} represents the generalized inverse distribution functions of Y and X, respectively. This way we say $(Y, X)^\top$ is upper tail-independent if λ_U equals 0. Further, we call λ_U the upper tail-dependence coefficient (upper TDC). The same way, we define the lower tail-dependence coefficient, if it exists, by:

$$\lambda_L \stackrel{\text{def}}{=} lim_{v \downarrow 0} \mathrm{P}\left\{Y \leq F^{-1}(v) \mid X \leq G^{-1}(v)\right\}. \tag{4.3}$$

$$\lambda_L = \lim_{u \searrow 0} \frac{C(u, u)}{u} \tag{4.4}$$

For the purpose of inference and also in order to facilitate interpretation it is useful to reduce the information of this relationship, between bi-variate variables that the copula evidences, to a one-dimensional function or to a single parameter.

Below we will use two dependence measures χ and $\bar{\chi}$, that allow us to measure distinct aspects of dependence and extremal dependence (dependence on the tails of the distribution).

Usually both measures are useful in order to have a fair summary on dependence, asymptotically dependence and asymptotically independence between the two variables.

Dependence measure χ summarised in χ coefficient is a measure of dependence as

$$\chi = \lim_{u \to 1} P\Big(V \geq u | U \geq u\Big)$$

where the pair of random variables (U, V) are assumed to be obtained by uniform transformation of the margin of non-identically random variables (X, Y).

The χ coefficient can be obtained as the limit of an asymptotically equivalent function as:

$$\begin{aligned} P\Big(V \geq u | U \geq u\Big) &= \frac{P\Big(V \geq u, U \geq u\Big)}{P\Big(V \geq u\Big)} \\ &= \frac{1 - 2u + C(u,u)}{1-u} \\ &= 2 - \frac{1 - C(u,u)}{1-u} \\ &\simeq 2 - \frac{\log \ C(u,u)}{\log \ u} \end{aligned}$$

Since $u \to 1$, then

$$\chi(u) = 2 - \frac{\log \ P(U \leq u, V \leq u)}{\log \ P(U \leq u)}$$

for $0 \leq u \leq 1$ and then we have that:

$$\chi = \lim_{u \to 1} \chi(u)$$

The $\chi(u)$ function can be used to obtain and analyse dependence over the quantiles and can be interpreted as a quantile dependent measure of dependence.

$\chi(u)$ as dependence measure can be interpreted as follows:

- $\chi(u) = 0$ for independent variables
- $\chi(u) = 1$ for perfect positive dependent variables
- $\chi(u) = -1$ for perfect negative dependent variables

The dependence measure $\bar{\chi}$ will allows us to analyse a situation of asymptotically independence, in the case of multivariate extreme value distributions (Bernard and Czado, 2015). While $\chi(u)$ is unable to provide details on the dependence of such type of models, we need to use a different dependence measure based on the joint survivor function defined as $P(X \geq x, Y \geq y)$ and represented as $\bar{F}(x, y)$ it could be expressed as:

$$\begin{aligned} \bar{F}(x, y) &= 1 - F_X(x) - F_Y(y) + F(x, y) \\ &= \bar{C}\Big(F_X(x), F_Y(y)\Big) \\ &\text{where} \\ \bar{C}(u, v) &= 1 - u - v + C(u, v) \end{aligned}$$

The dependence measure coefficient $\bar{\chi}$ can then be defined as:

$$\begin{aligned} \bar{\chi} &= \frac{2 \log P(U \geq u)}{\log P(U \geq u, V \geq v)} - 1 \\ &= \frac{2 \log(1 - u)}{\log \bar{C}(u, u)} - 1 \\ &\text{for } 0 \leq u \leq 1 \end{aligned}$$

where $-1 \leq \bar{\chi} \leq 1$ for all $0 \leq u \leq 1$.

In similar way we used to define above the χ dependence measure we also define now dependence measure $\bar{\chi}$ as:

$$\bar{\chi} = \lim_{u \to 1} \bar{\chi}(u)$$

The values of $\bar{\chi}$ coefficient can be interpreted as for values of $\bar{\chi} = 1$ we have asymptotically dependent variables. Also for independence case we have $\bar{\chi} = 0$.

In fact the case of asymptotically independent distributions is the one of greater interest as $\bar{\chi}$ as provides a dependence measure that is increasing on side of dependence strength is also increasing.

As shown in Bernard and Czado (2015), the copula tail dependence is of major importance in conditional quantile estimation, tacking in account that conditioning involves the tail of one of the marginal distribution.

As estimating *VaR* at a extreme (high) confidence level, meaning a quantile in the tail of the distribution, is a challenge and so it is estimate the conditional *VaR*, is thus even more challenging such estimation is required in order to calculate systemic risk measure *CoVaR* on the condition of an extreme event.

As *CoVaR* relates the returns of a financial institution with the returns of system, it is of interest to analyse how this relation is established across the distinct quantiles of the returns distribution.

We will take as an example three financial institution and compare the dependence between the returns of each of those financial institutions and the returns of the system. The details about the data set used and how it was collected can also be consulted in detail in sections A.3 and 5.3.

To do so was applied a Q–Q (quantile-quantile) plot in order to build a visual representation of the measures of dependence regarding the correlation. By using this Chi-plot technique it is possible to look at local dependence signs and moreover analyse and interpret the measure of dependence locally.

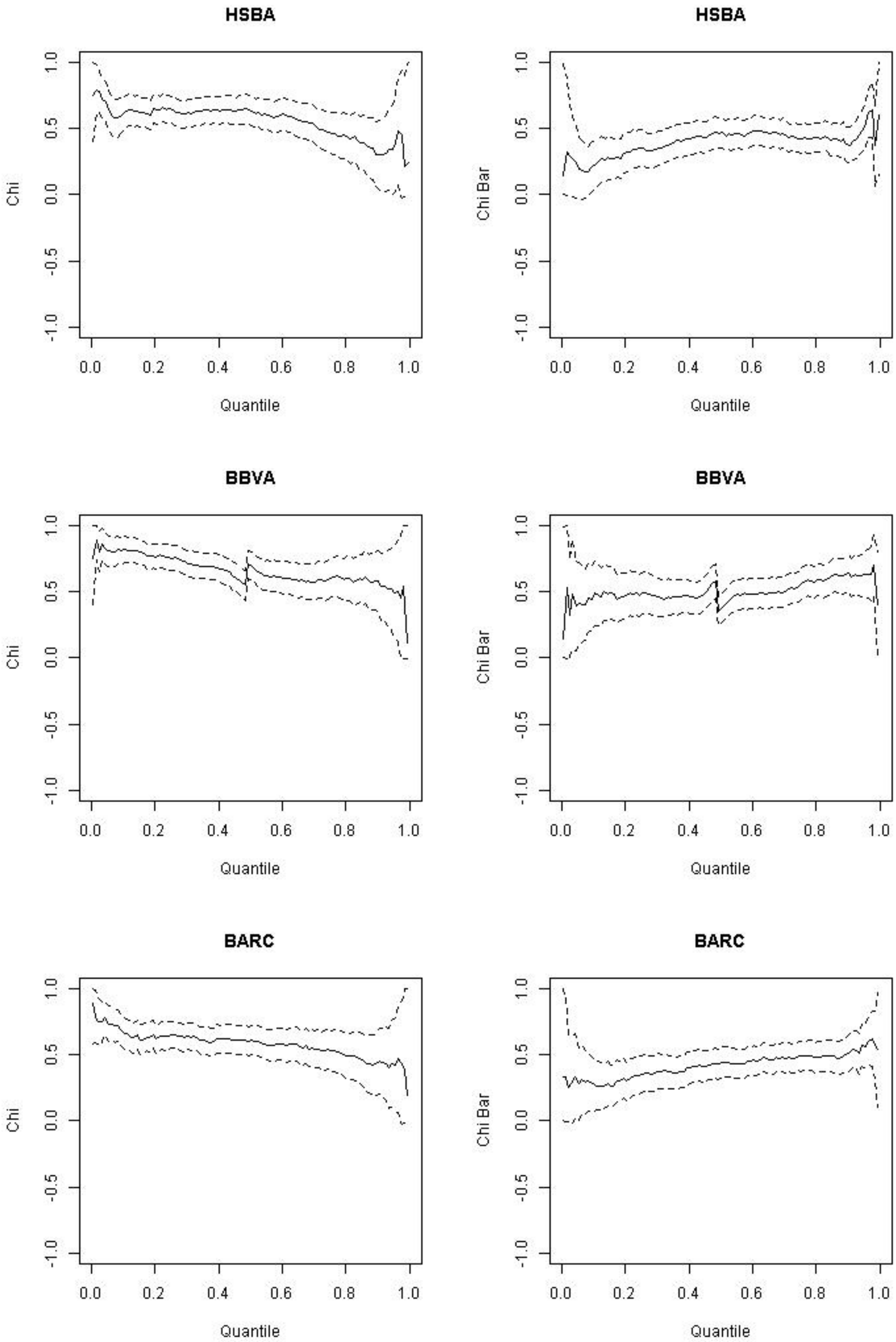


Figure 4.1: Dependence between a financial institution and the system by quantile

By analysing the results above we can notice a difference on the pattern of dependence, depending on the quantile.

The structure and behavior of dependence show considerable difference if we are looking at tail quantiles when compared with central quantiles.

The pattern exhibited on these sample graphs shows that in the extremes the values are more spread out, this means the tail is longer or heavier, which suggests distinct dependence regions and more dependence on the extremes. Another characteristic from χ and $\bar{\chi}$ function analysis is the dependence between financial institution and system is positive. For example, HSBC, BARC and BBVA financial institutions exhibit a stronger dependence in the left tail as shown in the the figure above.

This distinct structure in dependence and correlation related to the quantile is an important characteristic inherent to systemic risk that needs to be included in the model.

From the point of view of the dependency analysis between the two random variables X and Y, it is therefore of particular interest to look with special detail for the dependency that can be seen in the left tail of the distribution of returns (profit and loss), and include this feature in the model.

The behavior on the tails of a distribution has been described as behaving differently from the rest of the distribution (Phillips, 1985).

4.2.1 Normal Distribution Assumption

As mentioned in section **??** some recent research work has been advocate that the normal assumption does not fit for financial returns and even though it is noticeable that financial returns distributions are at least close to a bell shaped curve, even if this doesn't translate directly for a normal distribution.

By applying a simple normality test to financial institution returns series, it becomes very clear that we should strongly consider other options to model the financial institution returns.

Let us start by applying some conventional statistical tests for normality, such as Shapiro-Wilk's test and Kolmogorov-Smirnov (K-S) normality test.

Financial Institution	W statistic	K-S statistic	W p-value	K-S p-value
HSBA	0.94344	0.06652	2.2e-16	0.00017
BBVA	0.93228	0.07789	2.2e-16	5.249e-06
BARC	0.77149	0.11839	2.2e-16	2.565e-13
Financial System	0.69337	0.11109	2.2e-16	8.891e-12

Table 4.1: Normality testing

Based on the above results, we should consider the financial institutions returns (and financial system returns too) as not normally distributed.

However it is also known that normality tests are in fact very sensitive to what happens on the extreme tails. This fact can then restrain all the conclusions based on those type of tests.

As we have a relatively large sample of data on our data set, it will also worthwhile try a visual approach to investigate normality. Lets then compare the histogram of returns for some financial institutions and the system.

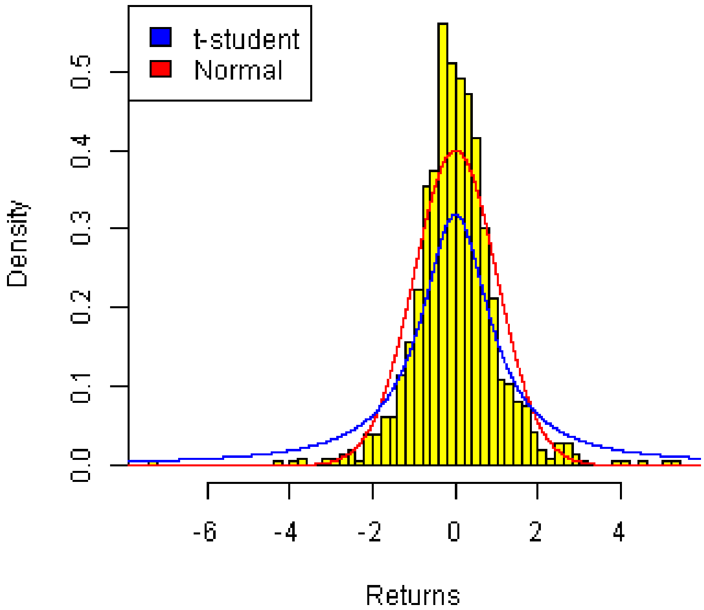


Figure 4.2: HSBA return frequency compared with normal and t-student distributions

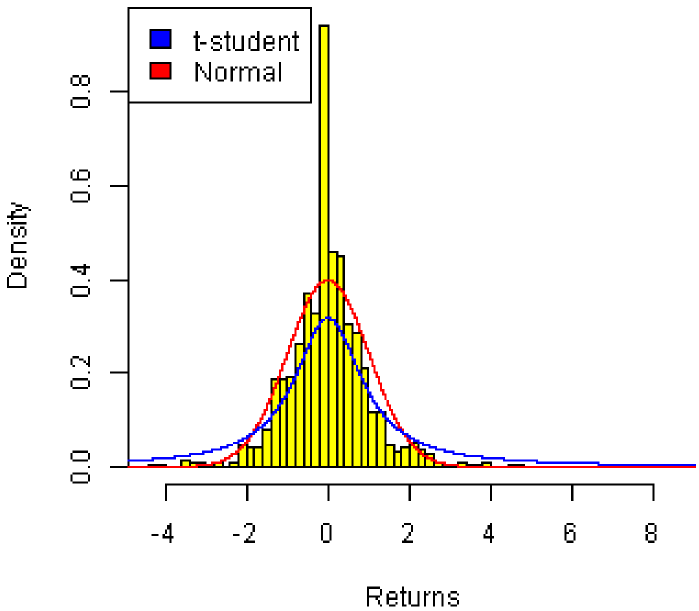


Figure 4.3: BBVA return frequency compared with normal and t-student distributions

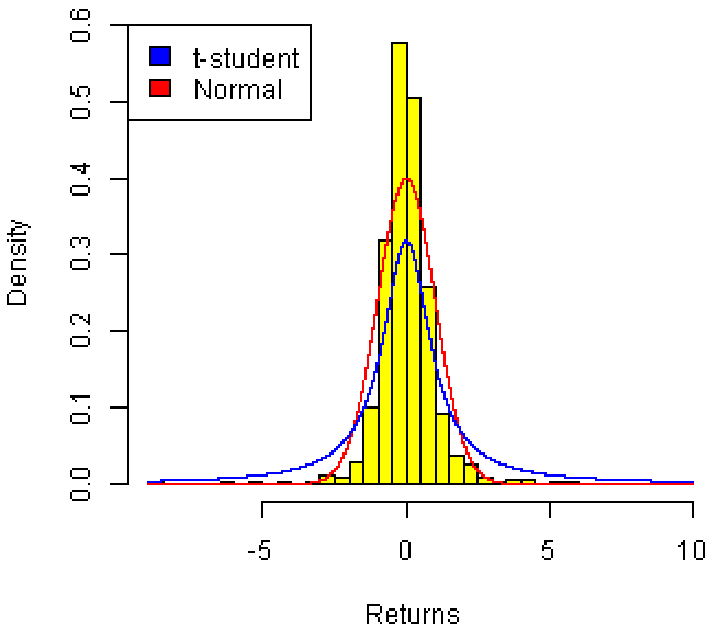


Figure 4.4: BARC return frequency compared with normal and t-student

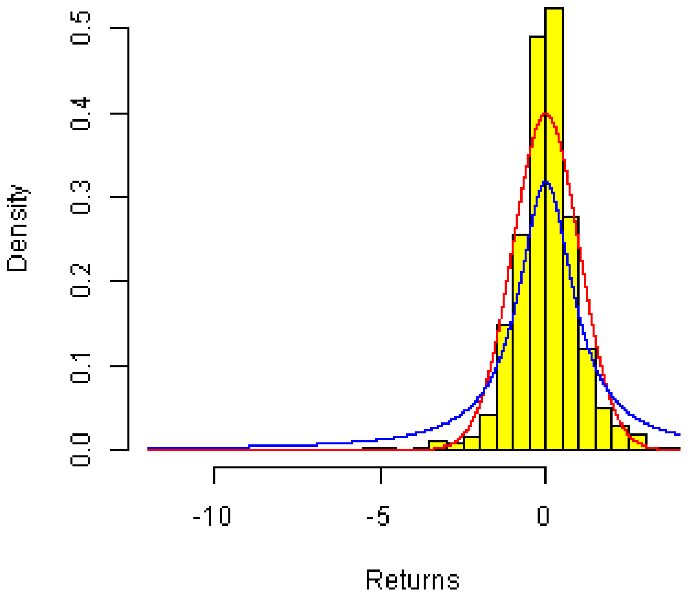


Figure 4.5: Financial System return frequency compared with normal and t-student distributions

By using the graphs above, we can notice a bell shape pattern on the returns of financial institutions as of the returns of the financial system. Specially in the case of the returns of the financial institutions we also can notice a fat tail behavior, best described in this case, in terms of extreme tails by the t-student curve than by the normal curve.

An additional complementary graphical analysis in order to compare the behavior of the returns across the distinct quantiles and compare that behavior with the expected behavior for a Normal population is to use a Q-Q plot.

The Q-Q plot, or quantile-quantile plot, is a type of graph that allows us to assess if it is plausible to assume that a data set came from some theoretical distribution such as a Normal. Even if it is just a visual check, and somewhat subjective, it allows us to see at-a-glance if our assumption is plausible, and how the assumption is eventually violated and which data points cause that violation.

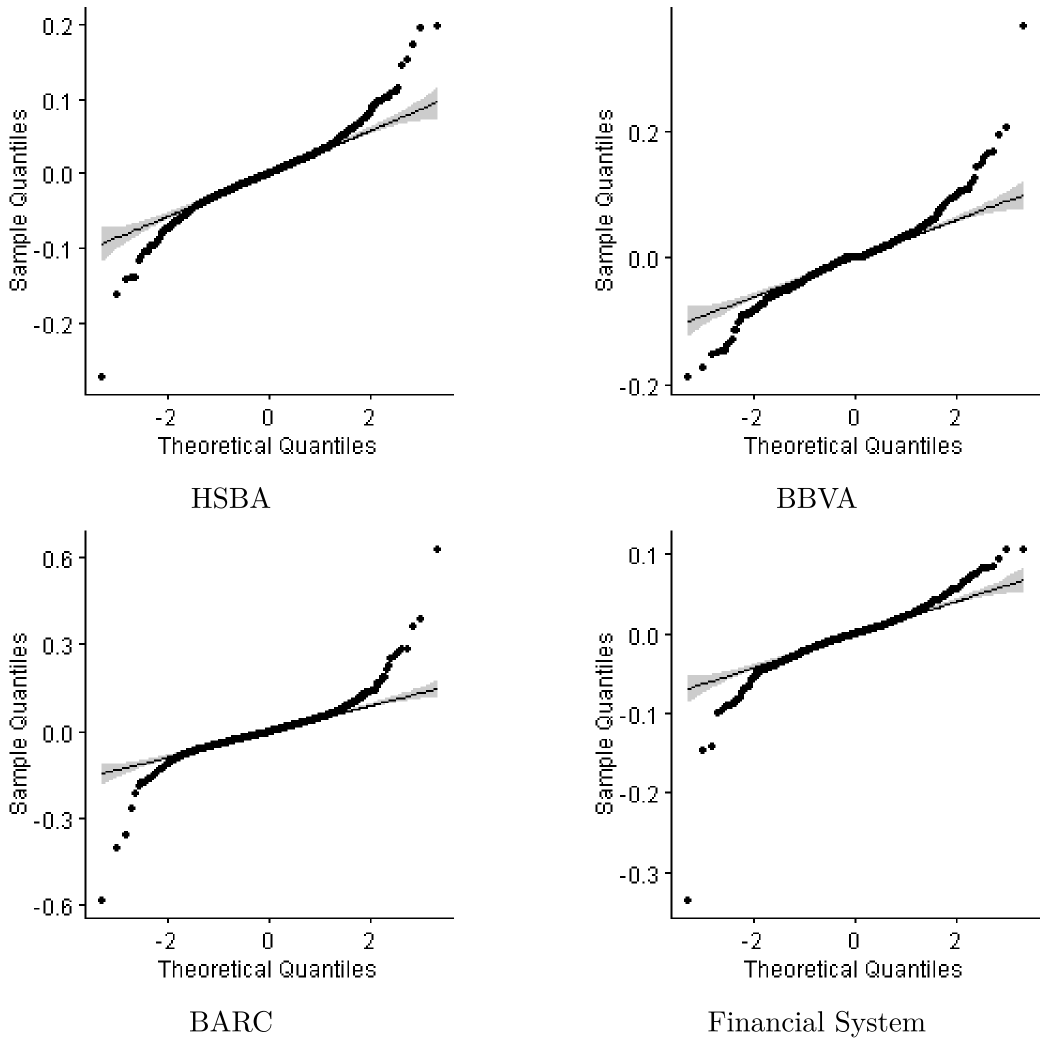


Figure 4.6: Financial institution and System returns Q-Q plot vs normal distribution with confidence intervals

Now, it becomes clear that we have a different behavior on extreme tails of returns distribution, and in the tail it is not following a normal behavior. This pattern must be included in the modelling.

4.2.2 Modeling the Extremes of Returns Distribution

The theory of extreme values provides options for modeling the behavior of a random variable in these conditions (at the extremes) and in combination with the rectory of the copulas it is also possible to model the dependence at the extremes with greater rigor, taking into account the particularity of the behavior at the extremes .

The original formulation of *CoVaR* elaborated by Adrian and Brunnermeier (2011) uses only one point to define the set of critical events E. This way we have:

$$CoVaR_{\alpha\beta t}(Y|X_i) = VaR_\beta(Y|X_i = VaR_\alpha(X_i)) \tag{4.5}$$

As a risk measure, *CoVaR* requires that a real number be uniquely associated with a pair of values (Y, X), in order to guarantee the properties of a coherent risk measure (Delbaen, 2000).

Assuming that the distribution of X and Y are continuous, through copula theories it is possible to obtain a process to select a conditional probability version in order to redefine *CoVaR* as a univocal risk measure.

As a measure of systemic risk, the *CoVaR* must also enjoy the property of concordance order which means that for higher levels of dependency there is also a higher level of systemic risk (Mai and Scherer, 2014).

Taken as assumption that the both return variables, X and Y have continuous distribution functions, but without include any additional assumption on the linking copula function, we can make use of the principles of Dini derivatives (Fernandez-Sanchez and Ubeda-Flores, 2016), as our assumption is only on the continuity of the functions F_X and F_Y.

Let define $D_U C$ as the partial left-sided derivative Copula $C(u, v)$. Based on Dini derivatives (Durante and Jaworski, 2010) $D_U C$ comes as:

$$D_U C(u, v) = \lim_{h \to 0^+} \sup \frac{C(u, v) - C(u - h, v)}{h} \tag{4.6}$$

As $u \in [0; 1[$ we can obtain the following results:

$$
\begin{aligned}
D_U C(u,0) &= \lim_{h\to 0^+} \sup \frac{C(u,0)-C(u-h,0)}{h} \\
&= \lim_{h\to 0^+} \sup \frac{0-0}{h} = 0
\end{aligned} \tag{4.7}
$$

and

$$
\begin{aligned}
D_U C(u,1) &= \lim_{h\to 0^+} sup \frac{C(u,1)-C(u-h,1)}{h} \\
&= \lim_{h\to 0^+} \sup \frac{u-(u-h)}{h} = 1
\end{aligned} \tag{4.8}
$$

For all other possible cases $D_U C(u,v)$ is decreasing in v (Durante and Jaworski, 2010).

Therefore, if we rewrite $D_U C(u,v)$ as

$$
D_U C(u,v) = \lim_{v\to 0^+} \sup_{0<h\leq v} \frac{C(u,v)-C(u-h,v)}{h} \tag{4.9}
$$

we then obtain that $D_U C(F_X(x),v)$ is a mensurable random variable and $D_U C(F_X(x),F_Y(y))$ is a version of the conditional expected value of the characteristic function (Ke and Yin, 2019), identified as:

$$
D_U C(u,v) = E(1_{Y\leq y}|X) \tag{4.10}
$$

from where we can define and fix the conditional probability (Bernard and Czado, 2015) as:

$$
F_{Y|X=x} = P(Y\leq y|X=x) = \lim_{h\to y^+} D_u C(F_X(x),F_Y(h)) \tag{4.11}
$$

CoVaR is this way defined as:

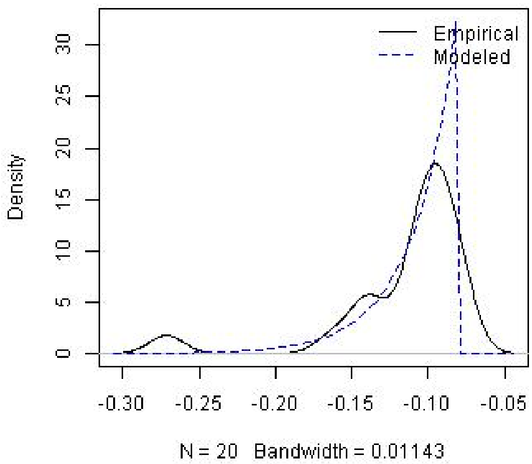
$$
CoVaR_{\alpha\beta} = \sup\left\{ y : F_Y(y) = \inf\{v : D_U C(\alpha,v) > \beta\}\right\} \tag{4.12}
$$

As C is continuous and differentiable we obtain from previous result that

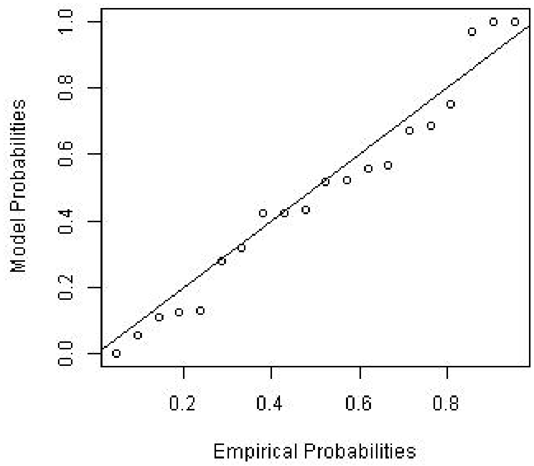
$$
\frac{\partial C}{\partial u} = F_Y((CoVaR_{\alpha,\beta}(Y|X))) = \beta \tag{4.13}
$$

In order to verify and analyse the possibility to use the extreme value theory to model the extreme tail of financial institutions returns we will also use some graphical tools to verify if that hypothesis is plausible as well.

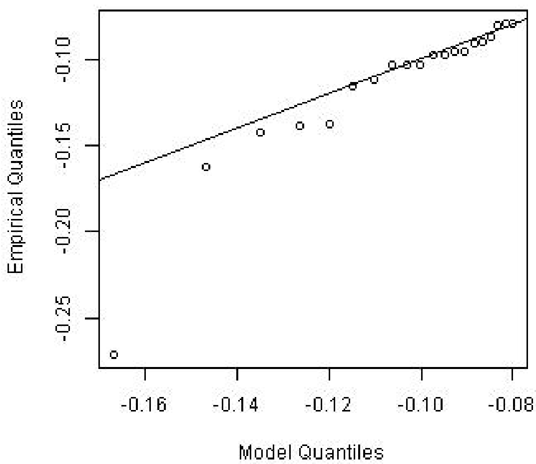
Based on R package extRemes (Gilleland et al., 2016) it was produced several graphical analyses to compare extreme value theoretical distributions such as generalized extreme value (GEV) and how they adjust to financial returns (Basilio and Oliveira, 2020).



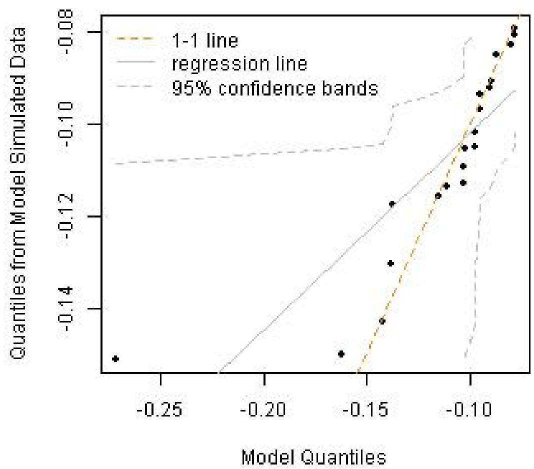
a



b



c



d

Figure 4.7: GEV fit diagnosis for HSBA extremes data series

Figure 4.7 shows a set of graphical tools to analyse the behavior on extremes of financial institution HSBA series of returns:

a) Density plots of empirical data and fitted GEV for HSBA

b) Theoretical distribution vs observed extremes for HSBA

c) Quantile-quantile plots for the GEV fit to the HSBA data

d) Quantile-quantile plot, quantiles from a sample drawn from the fitted GEV df against the empirical data quantiles with 95% confidence bands

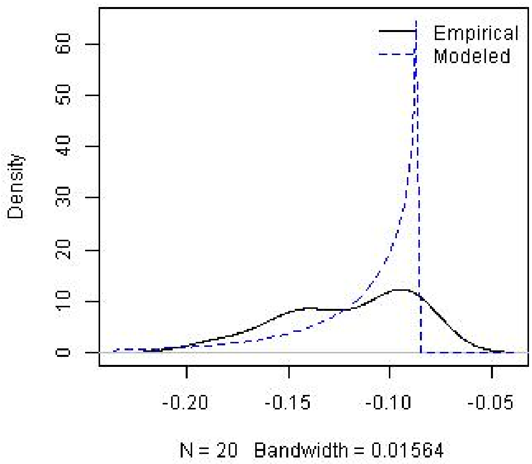
Figure 4.8 shows a set of graphical tools to analyse the behavior on extremes of financial institution BBVA series of returns:

a) Density plots of empirical data and fitted for BBVA

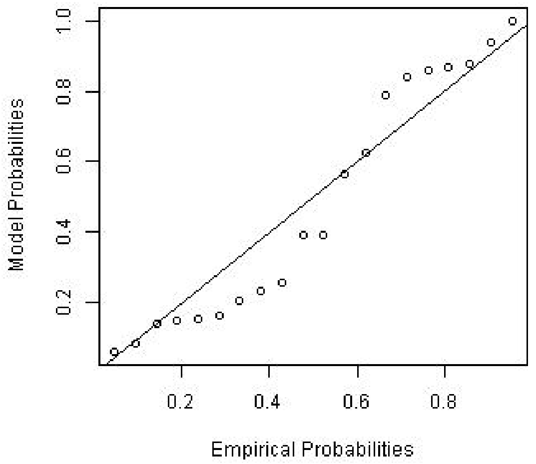
b) Theoretical distribution vs observed extremes for BBVA

c) Quantile-quantile plots for the GEV fit to the BBVA data

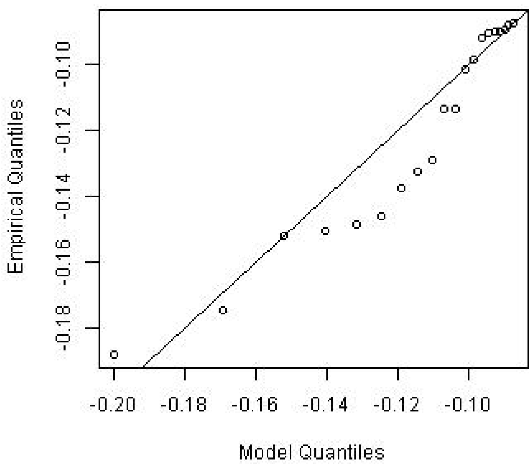
d) Quantile-quantile plot, quantiles from a sample drawn from the fitted GEV df against the empirical data quantiles with 95% confidence bands for BBVA.



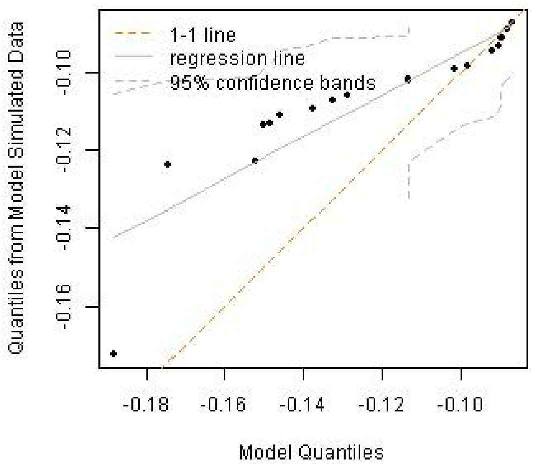
a



b

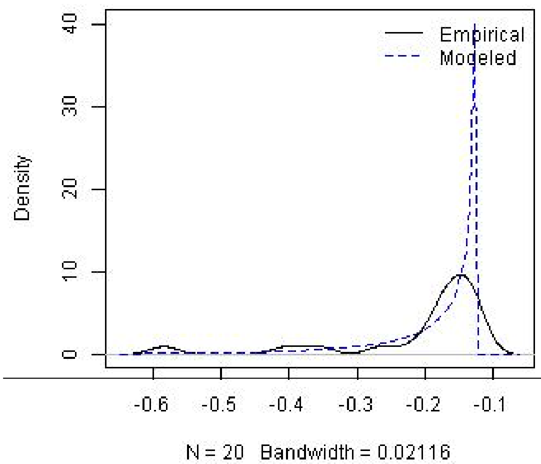


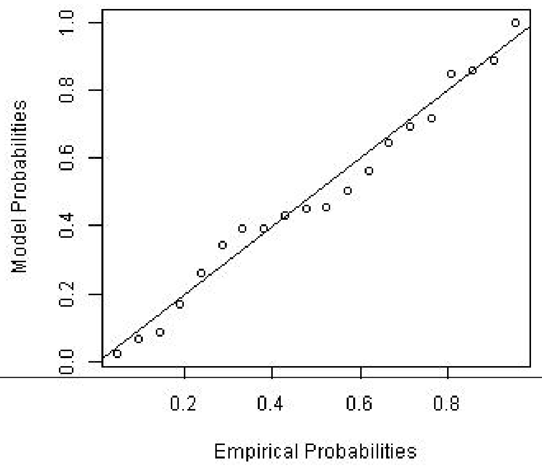
c



d

Figure 4.8: GEV fit diagnosis for BBVA extremes data series





a

b

Figure 4.9 shows a set of graphical tools to analyse the behavior on extremes of financial institution BARC series of returns:

a) Density plots of empirical data and fitted for BARC

b) Theoretical distribution vs observed extremes for BARC

c) Quantile-quantile plots for the GEV fit to the BARC data

d) Quantile-quantile plot, quantiles from a sample drawn from the fitted GEV df against the empirical data quantiles with 95% confidence bands for BARC.

Based on the graphs above it is possible to see how the extreme tail series of returns for these three financial institutions are fitting to a GEV distribution.

By analysing the Model Quantile plot, also a quantile-quantile plot, we can see now that GEV shows a more plausible fit then the normal distribution in the tail.

Even though we are using here the MLE method, we can observe that all observations are included inside the confidence interval provided.

Additionally we can also use Kolmogorov-Smirnov and Chi-square Goodness-of-Fit Test:

Financial Institution	χ statistic	K-S statistic	χ p-value	K-S p-value
HSBA	0.89676	0.18265	0.10124	0.63427
BBVA	0.91594	0.21124	0.09523	0.29111
BARC	0.94473	0.18544	0.32031	0.44402

Table 4.2: Results of Goodness-of-Fit Test

By applying extreme value theory we could now obtain a much better fitting on the left tail of return distributions, highlighted here by the results from Goodness-of-Fit Test applied on the reduced sample of financial institutions. Both tests applied, Chi-square and Kolmogorov-Smirnov, allow us to conclude that it is acceptable that all the return distributions follow a GEV distribution.

The same conclusion is also possible to obtain by a graphical analysis of the histogram of returns in the tail and by comparing the histogram of the returns in the tail with the modeled distribution that was a GEV (with distinct parameters for each financial institution).

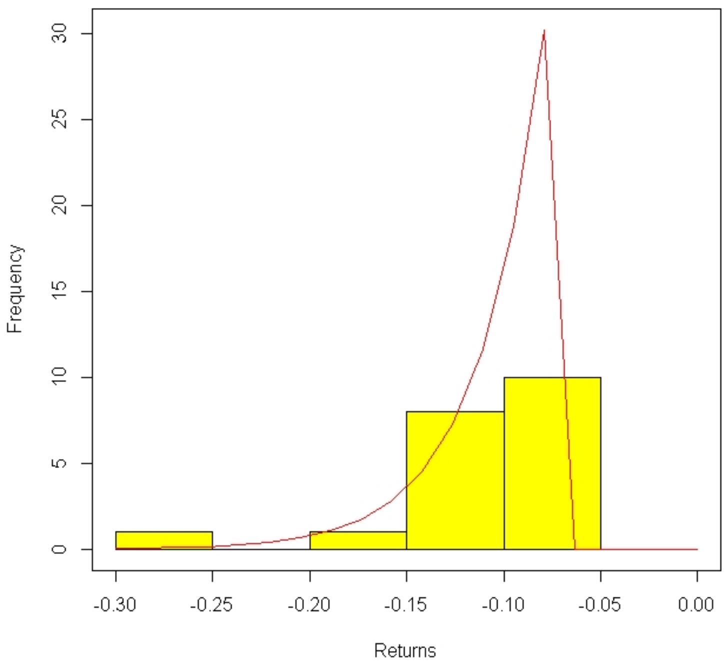


Figure 4.10: Left Tail Histogram for HSBA

In figure 4.10 the return values for financial institution HSBA are distributed along the line representing GEV. Even though there is a lack of concentration on the right. This behavior is confirmed by the results of the goodness-of-fit test where we have a p-value of 0.10124 for Chi square test.

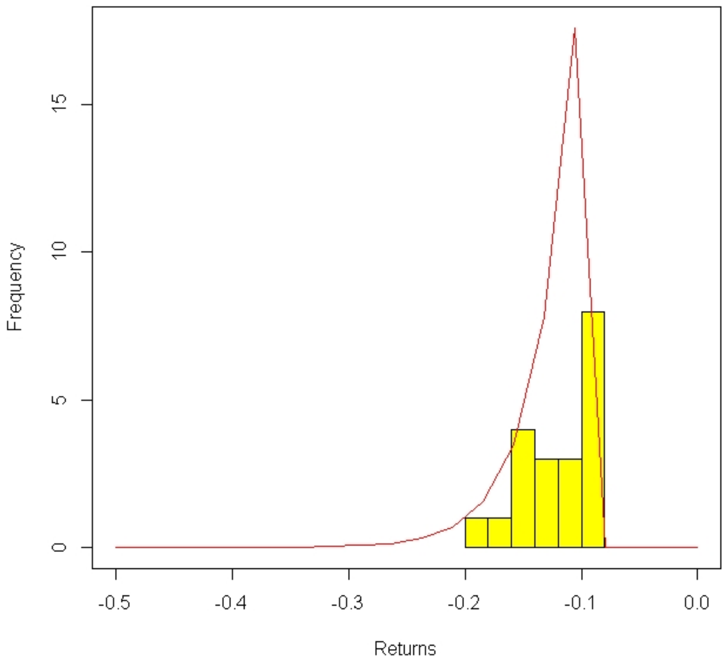


Figure 4.11: Left Tail Histogram for BBVA

In figure 4.11 the return values for financial institution BBVA are more concentrated on the right tail, and a significant part of GEV line is not followed by the returns series. This behavior is confirmed by the results of the goodness-of-fit returned a p-value of 0.09523 for Chi square test, slightly lower than we had obtained in the previous case.

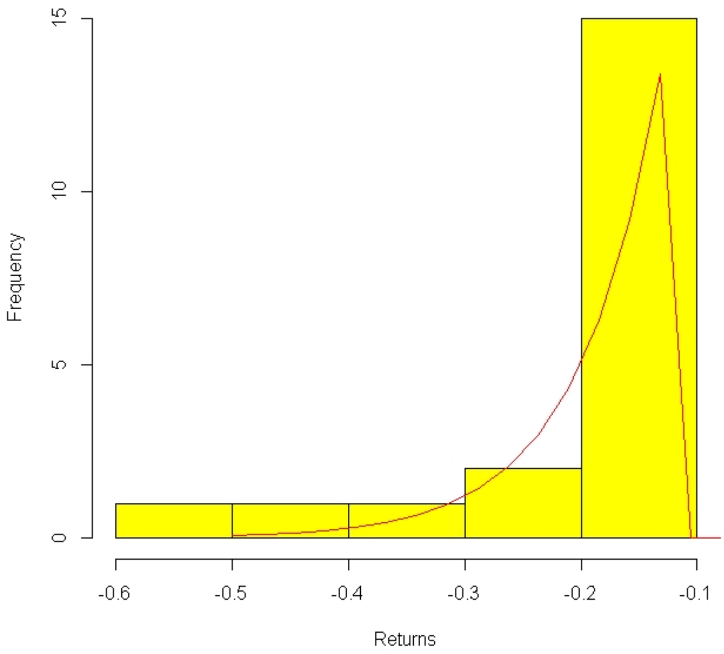


Figure 4.12: Left Tail Histogram for BARC

In figure 4.12 the return values for financial institution BARC are now following completely the GEV line. This behavior is confirmed by the results of the goodness-of-fit returned a p-value of 0.32031 for Chi square test, higher than the previous case.

Graphically we can notice that the best fit for the extreme returns is shown by financial institution BARC, as the histogram follows closely the line representing the GEV distribution, also confirmed by goodness of fit test. On other hand the worst case is shown by financial institution BBVA, as all the extremes values of returns are concentrated on the right tail of the histogram, and on the far left there are no occurrences following the GEV distribution.

4.3 *CoVaR* Copula

Conditional Value-at-Risk (*CoVaR*) can also be estimated using copula functions. Taking advantage of the intrinsic property of copula functions that permit the isolation of dependence from the copula marginal distribution functions, copula approach to *CoVaR* calculation provides flexibility in the specification of the marginals and dependence structure.

Let (X,Y) be as a pair of random vectors, the joint distribution is given by $F_{XY}(x,y) = P(X \leq x, Y \leq y)$, where F_{XY} represents the bivariate cumulative distribution function and F_X, F_Y represents the marginal distribution, then by Skar's theorem (Sklar, 1973) exists a two dimensional copula cumulative distribution function $C \in [0,1]^2$ that $F_{XY}(x,y) = C(F_X(x), F_Y(y))$. If F_X and F_Y are continuous then $C(u,v) = F_{XY}(F_X^{-1}(u), F_Y(v)^{-1})$

The conditional probability distribution as described in section 3.6.6, can be expressed by using a copula function (Salmon and Bouyé, 2008) as:

$$P(Y \leq y | X = x) = \frac{\partial C(u,v)}{\partial v} \tag{4.14}$$

and

$$P(Y \leq y | X \leq x) = \frac{C(u,v)}{v} \tag{4.15}$$

As discussed previously, a Archimedian copula is defined as :

$$C(u,v) = \varphi^{-1}[\varphi(u), \varphi(v)] \tag{4.16}$$

where the function φ is the generator function of the copula C.

By representing $CoVaR_{\alpha,\beta,t}$ as an Archimedian copula, we have, from previous results:

$$P(Y \leq y|X \leq x) = \frac{C(u,v)}{v} = \frac{\varphi'(v)}{\varphi'(C(u,v))} = \frac{\varphi'(v)}{\varphi'(\varphi^{-1}(\varphi(u)+\varphi(v)))} \tag{4.17}$$

If random variables Y represents the system returns, $R_{s,t}$ and X represents institutions i returns, $R_{i,t}$ with distributions $F_{s,t}$ and $F_{i,t}$, the conditional distribution of *CoVaR* can be write as

$$P(R_{s,t} \leq CoVaR_{\alpha,\beta,t}|R_{i,t} = VaR^{j}_{\alpha,t}) = \frac{\varphi'(v)}{\varphi'(\varphi^{-1}(\varphi(u)+\varphi(v)))} = \beta \tag{4.18}$$

By assumption $\frac{\delta C(u,v)}{\delta v}$ is partial invertible in order to u, the copula conditional quantile is given by :

$$u = \varphi^{-1}\left(\varphi\left((\varphi')^{-1}\left(\frac{1}{\beta}\varphi'(v)\right)\right) - \varphi(v)\right) \tag{4.19}$$

By using the probability integral transformation the $CoVaR_{\alpha,\beta,t}$ expression become :

$$CoVaR_{\alpha,\beta,t} = F^{-1}_{s,t}\left((\varphi')^{-1}\left(\varphi\left((\varphi')^{-1}\left(\frac{1}{\beta}\varphi'\left(F_{i,t}(VaR^{i}_{\alpha,t})\right)\right)\right) - \varphi\left(F_{i,t}(VaR^{i}_{\alpha,t})\right)\right)\right) \tag{4.20}$$

From *VaR* definition $F_{i,t}(VaR^{i}_{\alpha,t}) = \alpha$, the last expression can be simplified as:

$$CoVaR_{\alpha,\beta,t} = F^{-1}_{s,t}\left((\varphi')^{-1}\left(\varphi\left((\varphi')^{-1}\left(\frac{1}{\beta}\varphi'(\alpha)\right)\right) - \varphi(\alpha)\right)\right) \tag{4.21}$$

It is also possible to derive an analytical expression for $CoVaR_{\alpha,\beta,t}$, and it will take the form as:

$$P(R_{s,t} \leq CoVaR_{\alpha,\beta,t}|R_{i,t} \leq VaR^{j}_{\alpha,t}) = \frac{P(R_{s,t} \leq CoVaR_{\alpha,\beta,t})}{P(R_{i,t} \leq VaR^{j}_{\alpha,t})} \tag{4.22}$$

4.4 Contribution to Systemic Risk

CoVaR as a conditional risk measure can be used to estimate the individual contribution of each financial institution to the systemic risk.

Lets define that marginal contribution to risk as a difference between two estimates for *CoVaR*, taken with distinct assumptions:

- when financial institution i is under stress
- when financial institution i is in a *"normal"* situation and then not under stress

The marginal contribution of financial institution i is then interpreted as the difference of the system *CoVaR* in the two mentioned situations.

For a time period t this marginal contribution is represented as :

$$\Delta CoVaR_{\alpha\beta t} = CoVaR_{\alpha\beta t} - CoVaR_{0.5\beta t} \tag{4.23}$$

Based on the procedure described by Adrian and Brunnermeier (2016) one will need to estimate the *CoVaR* for the two above distinct situations. When we are facing a stress situation impacting on financial institution i returns for time period t and where the *CoVaR* corresponds to:

$$CoVaR_{\alpha\beta t} \tag{4.24}$$

and for a situation that we define as normal, without any stress situation impacting on financial institution i returns for time period t

$$CoVaR_{0.5\beta t} \tag{4.25}$$

which is the estimated value of the *VaR* of the financial system, conditioned by the *VaR* of the financial institution at the median, what means that the financial institution is not under stress. This way $\Delta CoVaR_{\alpha\beta t}$ represents the difference between the *VaR* estimate for financial system when:

- $Return_{it} < VaR_{\alpha t}$
- $Return_{it} < VaR_{0.5t}$

In order to calculate the estimate for *CoVaR* in both of the described situations, we need to take in account the dependence structure and correlation that exists in each one of those situations.

The literature has been proficuous in presenting cases and demonstrating that, for financial data the structure of dependence and correlation at the tail of the distribution is different from the dependence structure and correlation that occurs in the rest of the distribution. In a simple way, for a moment of stress or crisis, the dependence structure between the distribution of returns of financial assets and/or agents of the financial system changes, and consequently with more aggressive movements usually being observed.

The proposed approach for calculating the estimate of the financial institution individual contribution to systemic risk implies the calculation of *CoVaR* in two very different areas of the distribution of returns, from the point of view of the quantiles used. One component involves calculating the probability at the extreme, the left tail in the specific case of the risk, of the distribution, while the other component involves calculating the probability at the median, the middle point of the returns distribution. Modeling the dependency structure at the extremes is best achieved by applying the theory of extreme values (Fougères et al., 2009).

In this context, we are faced with the joint modulation of two distributions of returns, typically of a financial institution and the financial system, and the Gumbel copula with heavy tails proves to be an adequate choice .

On other hand, we have the dependency and correlation structure in the median, which should also take into account the heavy tails, but in this case should be better approximated by a t copula.

Consequently, the $\Delta CoVaR$ results from the difference between two copulas with distinct correlations structure and eventually different copula linking functions.

Let us look at first instance for $CoVaR$ on the tail. In the extremes or tails of the returns distribution, is expected to obtain an extreme value copula where we will fit the minimums distribution, obtained by selecting the k lowest values verified for the financial institution's returns.

The most well known methods to analyse extremes are the Block Maxima (BM) and the Peak Over Threshold (POT) (Bücher et al., 2019).

Instead those methods we will use a distinct approach based on the k smallest order statistics. One of the difficulties in analysing extremes is exactly the limited amount of available data labeled as extremes to make possible to estimate the parameters. Extreme values due to its own nature rare, have high volatility and this way providing less information about the probability associated to the occurrence of a phenomena. An alternative to mitigate this scarce availability of data is to use the k smallest (or largest) order statistics.

As we select k minimum values of financial institution i returns, we need to ensure that we will respect the concomitant for each of the k_{th} order statistics in regards of financial system returns data series.

After selecting the k minimum values, this series of values will be adjusted with the series, together with the returns to the system verified in the same time period, and which do not necessarily correspond to the minimums in the series of system returns, in a copula function.

4.5 Methodology Description

The proposed methodology to identify and rank the financial institutions in order to the highest systemic risk, implies to define a given time window to use as a basis for calculating VaR and, consequently, $CoVaR$.

This method will make it possible to estimate VaR and $CoVaR$ based on a given time interval for a specific point in time.

Moving this time window to start in a subsequent time period will make it possible to obtain a time series for the estimated values of $CoVaR$ and $\Delta CoVaR$ and, consequently allowing the temporal analysis of the results and comparison of $CoVaR$ and $\Delta CoVaR$ between different periods.

4.5.1 Time Series Analysis

Most of the studies involving financial time series make use of returns time series than asset price time series and, this is usually due to two reasons. The first one relates to the investor awareness as it is easier for the average investor to understand the returns as a complete characterisation of the performance of an investment rather than the prices. The second is more related with a more technical point of view as from a mathematical point of view, return series usually exhibits more desirable statistical properties than prices series.

4.5.2 Returns

Assets returns can be defined by using two distinct assumptions resulting in also two distinct definitions. Assuming simple returns and let $V(t)$ the value of a financial asset at a time period t, then the one period return is defined as:

$$R(t) = \frac{V(t)}{V(t-1)} - 1 \tag{4.26}$$

Based on the above definition one can write the expression for a k - periods returns as :

$$R^{(k)}(t) = \frac{V(t)}{V(t-k)} - 1 \tag{4.27}$$

Assuming continuous compounded returns also known as log returns. In this case the one period return can be expressed as:

$$r(t) = log(1 + R(t)) = log\left(\frac{V(t)}{V(t-1)}\right) \tag{4.28}$$

The k-period expression can be written as:

$$r^{(k)}(t) = log\left(\prod_{i=1}^{k-1}(1 + R(t-i))\right) = \sum_{i=0}^{k-1} log(1 + R(t-i)) = \sum_{i=0}^{k-1} r(t-i) \tag{4.29}$$

Comparing both results one can notice that log-returns are additive compared to simple returns, which are multiplicative. This will results in that log-returns become more convenient to use.

As the additive property is a desirable property and therefore the will consider log-returns.

4.5.3 Rolling-Window Analysis of Time-Series Models

There are several advantages that we can recognise in rolling-windows techniques to analyse and model time-series (Zivot and Wang, 2003).

For instance when working with financial time series data one important assumption is that the parameters of the statistical model applied are constant over time. But in the real world the economic environment and context changes considerably over time, and so assuming constant parameters over time could not be a reasonable assumption. By using rolling windows we are incorporate those differences in our model.

By defining a time window with a constant size we are also allowing to compare the results across time.

4.5.4 Methodology Steps

Several methodologies have been proposed to identify systemic institutions and to rank financial institutions in a financial system, in order to the systemic risk and systemic impact related to those financial institution.

1. Estimate the assets value. One of the objectives of this research is also to use public available data to identify systemic risky institutions. The first step on the methodology consists in collecting stock prices and calculate for each financial institution, the market capitalization. As assumption, we have that the market capitalization each institution reflects the book value of the assets of that financial institution, assumed as the market value of assets (MVA). The system valuation is obtained by a simple aggregation of the market capitalization of all financial institutions included in the system. Based on the the market value in each point in time, we will compute the returns:

$$X_i^t = \frac{MVA_i^t - MVA_i^{t-1}}{MVA_i^{t-1}} \tag{4.30}$$

   where X_i^t represents the return of financial institution i in the time period t.

2. Define the time horizon. The time horizon must be long enough to allow for the occurrence of different relevant critical events or crises.

3. Calculate the $CoVaR$ for the extreme distribution of returns. Define a minimum number of k of the financial institution's return (maximum loss verified), verified

within the time window taken into consideration. A copula function will be fitted to the two series:

a) Fit the copula to the two series on the extremes by selecting the copula that evidence more tail dependence.

b) The series of returns to the system in the same periods in which the minimum of the financial institution returns were verified, in order to preserve the distribution of concomitants of the extremes as induced from the order statistic, according to the of financial institution returns.

c) After adjusting the copula to both series of returns, obtained in the above step, the partial derivative of the copula function will be estimated for the marginal related to the series of returns of the financial system.

$$\frac{\partial C(u,v)}{\partial v} = CoVaR \tag{4.31}$$

Let us assume this partial derivative is invertible in relation to v, then we have the following result:

$$CoVaR^i_{\alpha\beta t} = F^{-1}_{s,t}(g^{-1}(\beta, F^{-1}_{i,t}(VaR^i_{\alpha,t}))) = F^{-1}_{s,t}(g^{-1}(\beta,\alpha)) \tag{4.32}$$

Assuming in this case that X^t_i follows a EVT distribution, and focusing in left tail of returns distribution we obtain the following result:

$$CoVaR^i_{\alpha\beta t} = \begin{cases} u_s + \frac{\beta_s}{\xi_s}\left(1 - \left(\left(-\frac{1}{\theta}ln\left(\frac{1-\left(1-e^{-\theta}\right)}{1+e^{-\theta\alpha}\left(\beta^{-1}-1\right)}\right)\right)\frac{N}{N-N_u}\right)^{-\xi_s}\right), \\ \text{if} \\ -\frac{1}{\theta}ln\left(1-\left(1-e^{-\theta}\right)\left[1+e^{-\theta\alpha}\left(\beta^{-1}-1\right)\right]^{-1}\right) < 1-\frac{N_u}{N} \\ u_s + \frac{\beta_s}{\xi_s}\left(1 - \left(\left(-\frac{1}{\theta}ln\left[1-\frac{\left(1-e^{-\theta}\right)\left(1-e^{-\theta\alpha}\right)}{e^{-\theta\alpha}}\right]\right)\frac{N}{N-N_u}\right)^{-\xi_s}\right), \\ \text{if} \\ -\frac{1}{\theta}ln\left[1-\frac{\left(1-e^{-\theta}\right)\left(1-e^{-\theta\alpha}\right)}{e^{-\theta\alpha}}\right] < 1-\frac{N_u}{N} \end{cases} \tag{4.33}$$

d) To obtain the *CoVaR* in the median of the financial institution returns we will take into account all the values of the returns in the time window. Both series will be adjusted to a copula. After adjusting the copula to both series of returns, the partial derivative of the copula function will be estimated for the marginal corresponding to the series of returns of the financial institution.

$$\frac{\partial C(u,v)}{\partial v} = CoVaR \tag{4.34}$$

In the median we will take a t copula. In this case *CoVaR* is obtained by:

$$CoVaR^{\beta}_{\alpha} = F^{-1}_s\left(t_v\left(\rho t^{-1}_v(\beta) + \sqrt{\frac{(1-\rho^2)(v+t^{-1}_v(\beta)^2)}{v+1}}t^{-1}_{v+1}(\alpha)\right)\right) \tag{4.35}$$

e) Determine Δ*CoVaR* by the simple difference between the extreme *CoVaR* and median *CoVaR*.

4. After obtaining the $\Delta CoVaR$ for each financial institution one can build a ranking based on $\Delta CoVaR$ where a higher value of $\Delta CoVaR$ corresponds to a higher contribution to systemic risk.

4.6 Conclusion

This new proposed methodology to calculate $\Delta CoVaR$ will use two different approaches for the purpose of select the copula function in the interest of align with also two distinct situation. First when the financial institution is not in stress, this mean a β parameter of 0.5, in the median of the returns distribution, and a different approach for select a copula to better describe the behavior on the tails, and in particularly, the tail dependence between the financial institution returns and financial system returns.

Chapter 5

Empirical Results

5.1 Introduction

In the next section we will presented the results obtained, by applying the methodology proposed.

We had used a modified version of Adrian and Brunnermeier's *CoVaR* methodology , presented in a 2011 paper, (Adrian and Brunnermeier, 2011), and also reviewed in 2016, (Adrian and Brunnermeier, 2016), which is defined as the *VaR* of the whole financial system, given that one of the financial institutions is in distress. Quantile regression is employed to estimate the daily *VaR* and then *CoVaR*. In these processes, we use the equity market return and market volatility.

Using data collected for the period 1998 to 2018, we will identify the financial institutes that are the largest contributors to the banking sector's systemic risk in Europe.

5.2 The Data Set

The proposed approach only relies on publicly available market data, such as stock returns as they are believed to reflect all information about publicly traded firms.

Based on the list of banks that are part of the *STOXX Europe 600* Banks index, corresponding to the biggest and most important banks in Europe, information on daily quotations for each title, public available for consultation at *https://finance.yahoo.com/*.

The original data set used was compounded by the daily quotations of the 50 biggest banks in Europe over the last 20 years, from January 1998 to June 2018. Subsequently, information on the number of shares issued was also collected for each of the institutions. This data, together with other information such as daily exchange rates, available at *https://www.bankofengland.co.uk/statistics/exchange-rates* were also collected in order to build a capitalization daily series of each institution (in euro).

This data set was later reduced to include only weekly closing prices for each financial institution. From this series, a series corresponding to the weekly returns (in percentage) was obtained from each of the institutions and for the entire system as well.

The overlapping rolling window was applied over periods of three years each. For each one of these periods was estimated *VaR*, *CoVaR* and $\Delta CoVaR$, forming a new time series related to each institution, representing the risk position for each institution in each point in time (week).

5.2.1 Calculating *VaR*

There are plenty of different ways to obtain *VaR* and distinct methods to calculate them. All these methods have a common base but then will diverge in how they, in fact, calculate *VaR*. Usually, those methods also have a common problem in assuming that the future will follow the past.

What are the options to obtain *VaR*? We will discuss here three different approaches:

- Historical *VaR*.
- Analytic *VaR* (Variance CoVaRiance, Parametric)
- Extreme Value *VaR*

While the first ones are standard in literature and across the industry, the last one mentioned is not yet so broadly used (Fallon et al., 1996).

With the Historical *VaR* method, we will be looking at the data, that is the returns during a period of time and check off for the value at specific quantile q previously defined. The advantage in this method is we do not need to have also any special assumption or knowledge about returns distribution. So it is straight forward methodology to apply and implement, and namely:

- Normality Assumption is not required.
- Works on historical returns.

Historical *VaR* Calculation:

- Step 1: Collect data on historical returns for an institution. These returns over a time interval = desired *VaR* time period.
- Step 2: From this info, make a histogram of historical return data.
- Step 3: *VaR* is the return associated with the cumulative probability from the left tail of the histogram that equals q quantile.

Therefore, one can associate the following advantages with this methodology (Jadhav and Ramanathan, 2009):

- Because it is non-parametric, the historical method does not require normality assumption.
- Easy to understand and implement.
- Based only on historical information.
- Is consistent with the risk factor changes being from any distribution.

Parametric *VaR* is also a popular way to calculate *VaR*. Actually, this method will be using returns information in order to estimate the parameters, as average and variance for a theoretical distribution that will then be fitted to returns data series.

The most common distribution associated with the returns under this method is perhaps the normal distribution, and in this case, the method is also called the variance-covariance method where the returns are assumed to be normally distributed.

- The most common measure of risk is standard deviation of the distribution of returns.
- Higher volatility = higher risk = potential for higher losses.
- Using standard deviation and some assumptions about returns, we can derive a probability distribution for returns.

With this methodology we are taken the following assumptions:

- Variance-Covariance *VaR* assumes that asset returns are normally distributed with known mean and standard deviation over a specified time-period.
- Covariances (correlations) among assets are known for the same time interval.

Inputs into the *VaR* calculation:

- Market values of all securities in the portfolio.
- Their volatilities.

The assumption is that the movement of the components of the portfolio are random, and drawn from a normal distribution.

Combining EVT and *VaR*. Extreme value theory can be used to investigate the properties of the Left tail of the empirical distribution of a variable X^i.

By applying extreme value theory we do not have to make assumptions on returns distribution as well. In fact, we don't need to know the distribution either. As, in terms of *VaR*, we are looking for the behaviour on the extremes, and what we really need here is to modulate the behaviour on the tail of the distribution, in this case in the left tail Levine (2009).

By using EVT to model extremes behavior it also means:

- Follows mathematical theory of the behaviour of extremes.
- The body and the tail of data do not necessarily belong to the same underlying distribution.
- Does not require particular assumptions on the nature of the original underlying distribution of all the observations.

Additionally, with GPD we can consider the following properties:

- GPD is an appropriate distribution for independent observations of excesses over defined thresholds.
- GPD can be used to predict extreme portfolio losses.

The two methods to model extremes, GPD and GEV could be proved equivalent, and both methods require to set an arbitrary value, the time interval in the GEV and the threshold in the GPD. Analyzing differences between the two methods, we have that, while the GPD method requires only two parameters, GEV method requires three parameters. The most relevant difference between the methods relies upon in the way it identifies the extremes. In the case of GEV, it relies on *T-maxima* (peaks in time intervals of duration T), which can include observations of lower magnitude than the threshold defined for GPD, and this way obtains more data. On another hand, if in the same interval we have several observations over the threshold, all will be considered with GPD, but some could be discarded with GEV.

One drawback of GPD application, in this case, is related to the definition of a threshold, which means also to establish an arbitrary definition of a *"crisis"*. With GEV we don't need to make this assumption.

So, by applying Fisher-Trippett theorem we can get an expression for the extreme value distribution. Then we can select only a few data samples, estimate parameters, tail, scale, and location parameters, and fit this extreme value distribution. The idea here is to select only a small number of the most extreme values we have on the data set, the worst 5 returns, for instance in each time window has been used.

5.2.2 Comparing *VaR*, *CoVaR* and Δ*CoVaR*

In order to compare the performance of each one of the models, we will use a simple measure to count the *breaks* verified on *VaR* estimates by using each method. Let us consider a break on *VaR* whenever $loss_{t+1} > VaR_t^q$.

$$break_{t-1}(i) = \begin{cases} 1 \text{ if } loss_t > VaR_{t-1}^q \\ 0 \text{ otherwise} \end{cases} \tag{5.1}$$

where i represents institution i, and t a point in time.

The measure used will be simple the count of break observed over the period in analyse, or:

$$\sum break_{t-1}(i) \tag{5.2}$$

As of when we are analyzing risk and in special systemic risk, we are interested in extreme cases localized on the extreme of left tail of the distribution.

5.3 Financial Data Set Analyse

We will analyse financial institutions returns time series and look for patterns in the behavior of those time series.

Harvesting data to support any research is usually a challenging process. Institution financial details frequently are not available on the public domain and are informed only on a periodic basis. This methodology consist therefore in an option to obtain *Var* and *CoVaR*, based on market public data, which brings additional clarity to the process as well as allows us to access and calculate those risk measures at any point in time, independently when the data for each financial institution is made available or published. As an assumption, to use publicly available data, the market value, the market capitalization of each institution reflects the book value of the assets. Also, the market value of the system is assumed as the aggregation of the market value of all the institutions belonging to that system. So, as step 1, for this method, we have to collect data and calculate the respective market capitalization for each day and for each institution, and estimate the capitalization for all the system as well. The process could be summarised as follows:

- Obtain the data:
  - Collect stock prices.
  - Balance sheet equity (BVE) and total assets (BVA).
  - In this work we will use stock prices, in weekly base (Fridays price).
- Market value of equity (MVE):
  - Stock price $\times$ shares outstanding.
- Assume market value of assets (MVA):
  - Book value of assets (BVA) * (MVE / BVE).
  - Means market-to-book ratios for equity and assets are equal.
- Define system asset value as the sum of institutions $MVA_t^{sys} = \sum MVA_t^i$, for a pointy in time t.
- Getting returns as: $X_t^i = \frac{(MVA_t^i - MVA_{t-}^i)}{MVA_{t-}^i}$; for financial institutions and financial system.

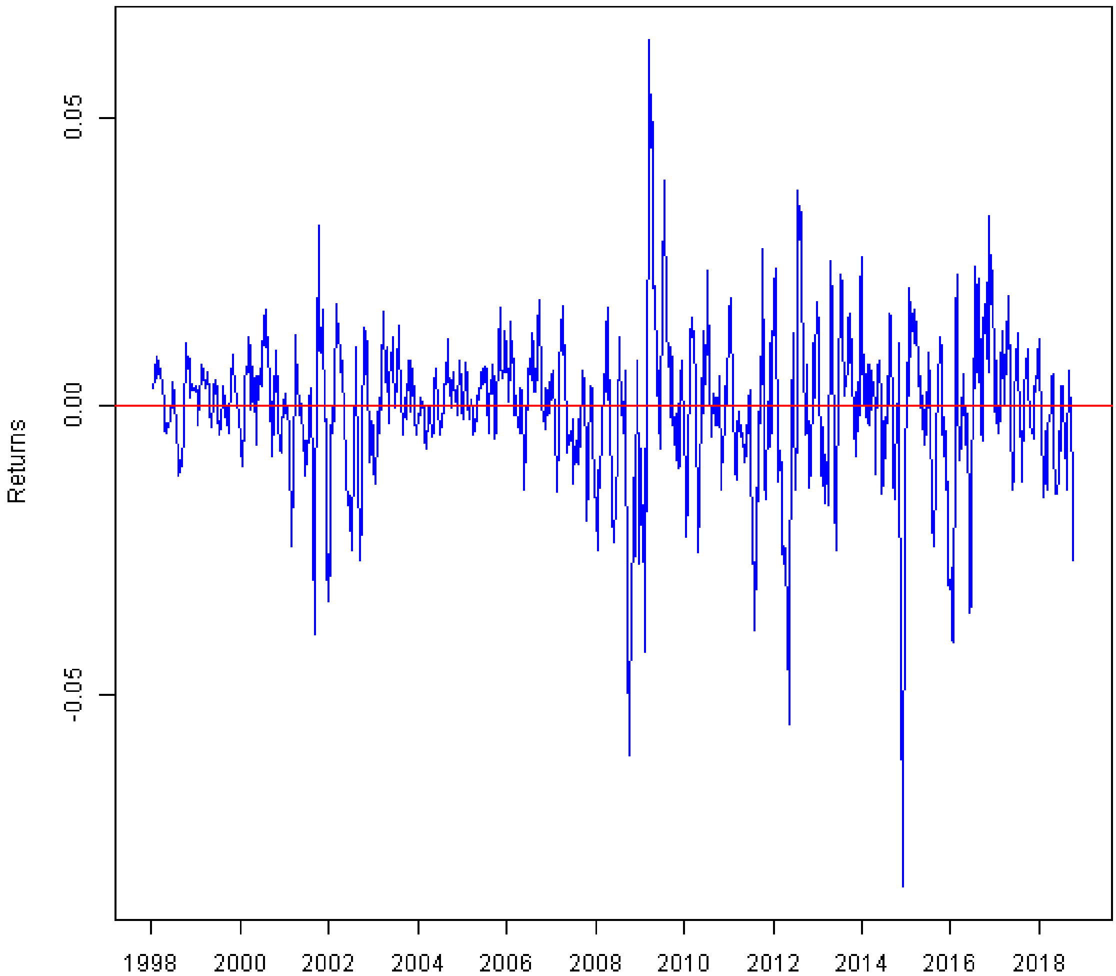


Figure 5.1: Financial System returns from 1998 - 2018

By applying the procedure described above, we will get series of the results for each individual financial institution ans for the financial system as well. In figure 5.1 it is shown the series os system results obtained as described for the period 1998 to 2018, according to the data set collect.

5.3.1 Data Analyses

The data exploratory analysis and data preparations steps were implement in R (Ihaka and Gentleman, 1996) as well as all the graphics and data visualization works (Tippmann, 2015). The data integration was implemented in Talend ® (Barton, 2013) and the data set was stored in a PostGreSql ® database (Drake and Worsley, 2002).

From the returns time series obtained with the methodology described previously we identify some patterns of relationship between the several series.

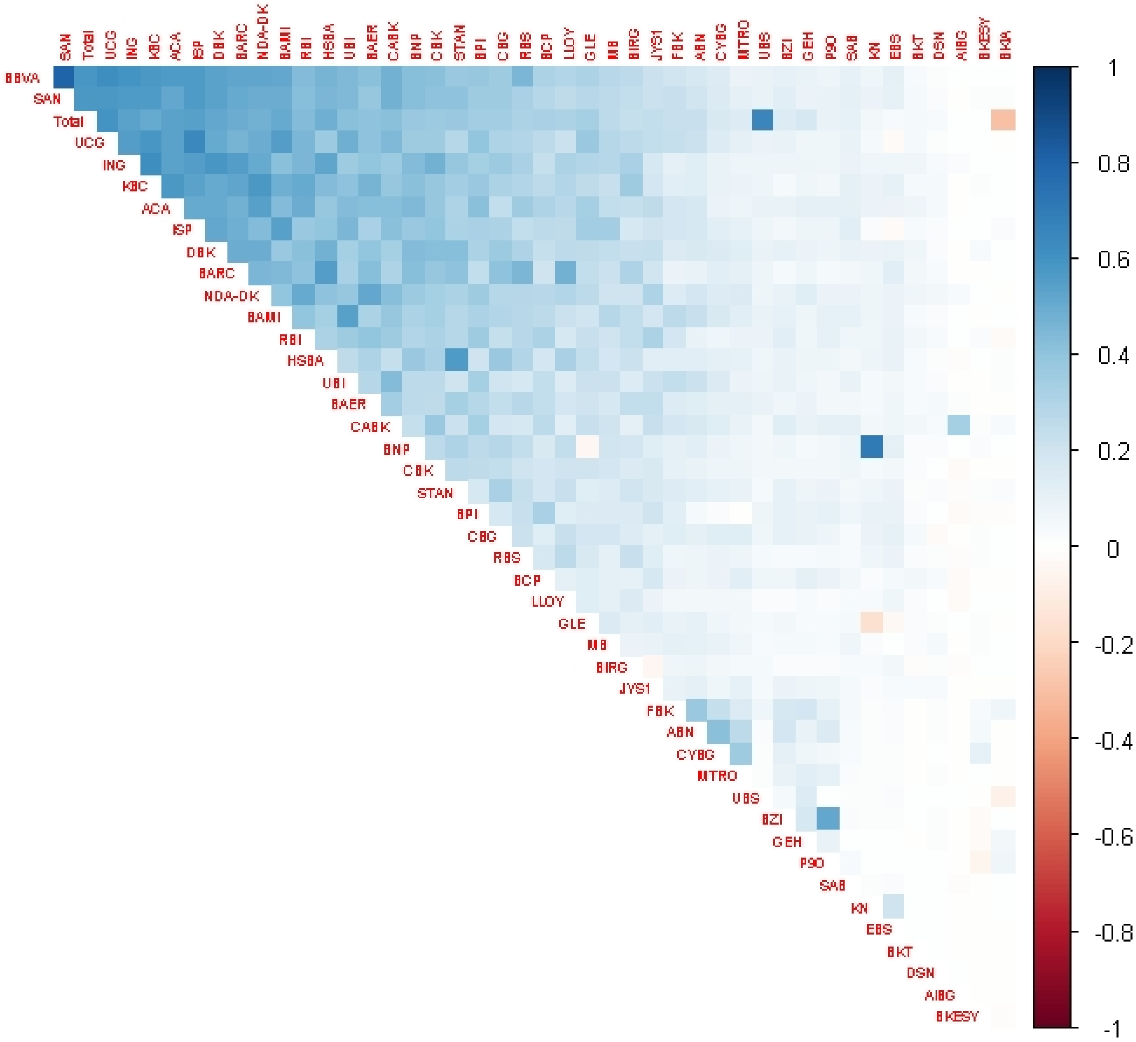


Figure 5.2: Correlation matrix of all financial institution returns

By analyzing the graph above one can notice that the correlations are predominately positive.

Also the correlation between the financial system and the financial institutions are between the most significant correlations detected, as shown bellow:

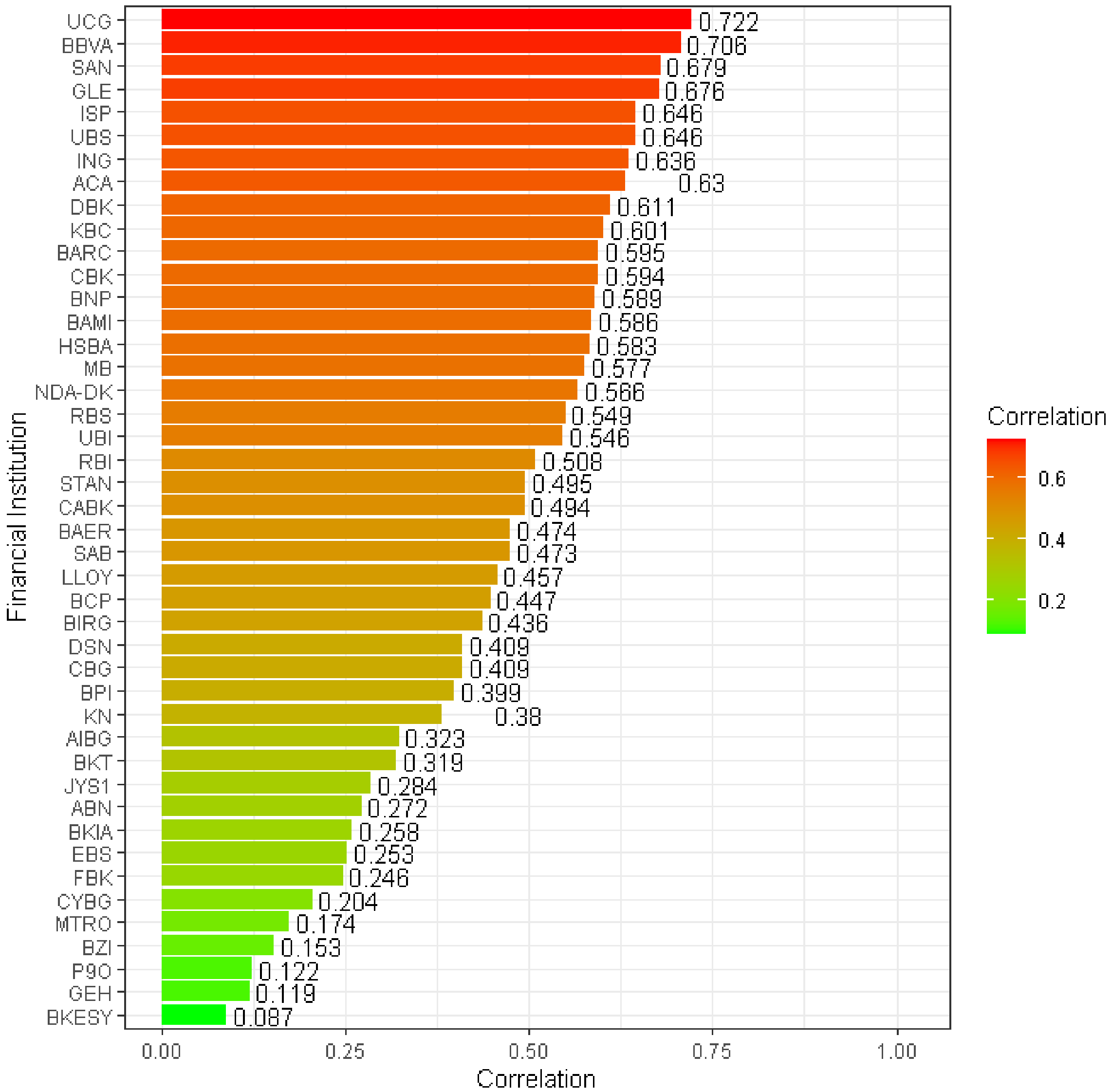


Figure 5.3: Observed correlation between the financial institution returns and financial system returns

The table following summarizes a set of statistics in order to characterize better each one of the financial institution returns:

Financial Institution	Min	Max	Mean	Variance	Correlation With System
ABN	-0.1305	0.0832	0.0019	0.0012	0.2718
ACA	-0.4276	0.2969	0.0011	0.0029	0.6302
AIBG	-0.5804	0.5702	-0.0036	0.0090	0.3225
BAER	-0.3000	0.2141	0.0021	0.0022	0.4743
BAMI	-0.2530	0.5443	-0.0021	0.0041	0.5858
BARC	-0.5836	0.6274	0.0015	0.0044	0.5948
BBVA	-0.1881	0.3673	0.0012	0.0021	0.7064
BCP	-0.2485	0.6412	-0.0006	0.0049	0.4472
BIRG	-0.4943	0.6666	-0.0001	0.0085	0.4357
BKESY	-0.6994	0.6797	-0.0565	0.1245	0.0874
BKIA	-0.5200	0.3666	-0.0114	0.0071	0.2583
BKT	-0.3395	0.6838	0.0040	0.0041	0.3187
BNP	-0.5370	0.3544	0.0010	0.0033	0.5891
BPI	-0.1887	0.3717	0.0003	0.0028	0.3985
BZI	-0.1318	0.1254	0.0001	0.0017	0.1527
CABK	-0.1723	0.3352	0.0013	0.0022	0.4937
CBG	-0.2089	0.2491	0.0020	0.0022	0.4088
CBK	-0.2870	0.5317	-0.0006	0.0041	0.5937
CYBG	-0.1192	0.1778	0.0019	0.0016	0.2044
DBK	-0.3665	0.4980	0.0014	0.0036	0.6106
DSN	-0.3033	0.5520	0.0017	0.0028	0.4089
EBS	-0.4001	0.3722	0.0026	0.0029	0.2528
FBK	-0.1241	0.1021	0.0044	0.0012	0.2462
GEH	-0.0836	0.1461	0.0023	0.0007	0.1190
GLE	-0.2357	0.2885	0.0011	0.0031	0.6762

Financial Institution	Min	Max	Mean	Variance	Correlation With System
HSBA	-0.2712	0.1969	0.0007	0.0013	0.5829
ING	-0.3487	0.6346	0.0021	0.0040	0.6362
ISP	-0.2507	0.2490	0.0012	0.0027	0.6457
JYS1	-0.2570	0.6563	0.0023	0.0027	0.2838
KBC	-0.4058	0.5457	0.0027	0.0044	0.6009
KN	-0.2480	0.5711	0.0030	0.0040	0.3802
LLOY	-0.5571	0.6220	-0.0010	0.0060	0.4567
MB	-0.2250	0.3876	0.0013	0.0023	0.5766
MTRO	-0.1343	0.1519	0.0025	0.0025	0.1741
NDA-DK	-0.2291	0.2500	0.0008	0.0016	0.5659
P9O	-0.0927	0.1307	0.0013	0.0011	0.1223
RBI	-0.4517	0.4217	0.0027	0.0045	0.5080
RBS	-0.5404	0.5056	0.0004	0.0036	0.5492
SAB	-0.1535	0.4461	0.0036	0.0031	0.4733
SAN	-0.2109	0.3116	0.0017	0.0021	0.6789
STAN	-0.4286	0.2784	0.0017	0.0029	0.4953
UBI	-0.1900	0.5846	0.0009	0.0030	0.5457
UBS	-0.2299	0.3019	0.0008	0.0022	0.6457
UCG	-0.3157	0.3019	-0.0007	0.0035	0.7216

Table 5.1: Descriptive statistics

5.4 Implementation Details

As the main objective in this thesis is to detail a methodology to identify systemic risky financial institutions we will illustrate the process to calculate the $\Delta CoVaR$ in detail first taken one financial institution as example and then by computing the measure $\Delta CoVaR$ for all financial institutions.

5.4.1 Joint Distribution and Copula

Considering a joint distribution to model the dependence between the financial institution i and the entire system we will have from copula theory as exposed in detail in section 3.6 :

$$H(x, y) = C(F_i(x), F_s(y))$$

Therefore, the first step is to fit a copula to both series of returns, the financial institution returns, and the overall returns of the system.

Also from copula theory, we can recall that if X is a random variable with distribution F, then F(X) is uniformly distributed in the interval [0, 1]. Here, we can start by adjust a distribution function, F(X), to the financial institution i returns and a F(Y) to the system returns.

Assuming then a pair of random variables (X,Y) then the transformed variable U = F_i(X) and V = F_s(Y) will have a standard uniform distribution and the copula C(U,V) is a joint distribution of (U,V), also C(u,v) = $C(P(U \leq u), P(V \leq v))$ where (u,v) $\in$ $[0, 1]^2$.

Assuming H as a bivariate distribution with margins F_i and F_s exists a copula such as $C(u, v) = C(F_i^{-1}(u), F_s^{-1}(v))$.

An aspect of major importance in this modeling is the relationship between the two variables. The nature of that relationship will determinate the the structure of dependence in cause. By using some graphical tools it is possible to visualize that relationship and representing graphically the copula.

Those, the joint distribution of the random variables (X,Y) is represented by the marginal distributions F_i(X) and F_s(Y) in conjunction with the copula C. This copula function C will then establish a link between these marginal distributions in order to construct the join distribution.

The copula can be also interpreted as the adjustment that we need to apply in order to convert from a situation of independence to joint distribution of the random variables (X,Y).

5.4.2 Select the Copula

Copulas can be selected according to the Akaike and Bayesian Information Criteria (AIC and BIC, respectively) and we can also perform a statistical test to assess how this function performs in a goodness-of-fit test based on White's information matrix equality.

The copula selection method consists in computing the criteria for all possible copulas choices and then the copula family that evidences the minimum value is chosen.

We are in this case interested in exploring a positive relation between a financial institution returns and financial system returns.

The process to select the copula function and the copula family that better reflects the dependence between both return series needs to make sure not allow us to adequately capture the full range of behaviour in the distribution of the dependent variable. The dependence on the tails should be taken in account as well, specially for a purpose of systemic risk modelling.

When we pursuit the best fit based on AIC/BIC criteria we are obtaining the following results for each pair :

Financial Institution	Copula Family	AIC	τ	Lower Tail Dependence	p-Value
ABN	Gumbel	-106.59	0.52	0	0.94
ACA	t	-731.56	0.55	0.42	0.17
AIBG	t	-228.76	0.36	0.31	0.86
BAER	t	-297.12	0.41	0.29	0.02
BAMI	t	-538.16	0.46	0.23	0.87
BARC	t	-720.93	0.49	0.37	0.71
BBVA	t	-994.6	0.59	0.51	0.13
BCP	t	-313.39	0.36	0.22	0.11
BIRG	t	-329.7	0.38	0.27	0.24
BKESY	Frank	-13.32	0.13	0	0.5
BKIA	t	-191.5	0.49	0.39	0.67
BKT	t	-295.66	0.47	0.31	0.05
BNP	t	-967.0	0.57	0.47	0.02
BPI	t	-226.86	0.3	0.18	0.52
BZI	Gumbel	-25.3	0.21	0	0.29
CABK	t	-292.13	0.49	0.3	0.1
CBG	t	-251.23	0.31	0.03	0.78
CBK	t	-690.04	0.49	0.37	0.17
CYBG	BB7	-45.72	0.37	0.37	0.95
DBK	t	-854.57	0.53	0.44	0.57
DSN	t	-254.59	0.32	0.26	0.83
EBS	t	-399.84	0.4	0.32	0.22
FBK	Survival BB8	-104.9	0.45	0	0.28

Financial Institution	Copula Family	AIC	τ	Lower Tail Dependence	p-Value
GEH	t	-22.12	0.2	0.12	0.72
GLE	t	-1024.55	0.61	0.5	0.08
HSBA	t	-629.12	0.47	0.22	0.01
ING	t	-888.43	0.55	0.42	0.06
ISP	t	-752.5	0.51	0.32	0.02
JYS1	Survival BB1	-168.54	0.32	0.35	0.23
KBC	t	-700.01	0.51	0.43	0.44
KN	t	-578.17	0.47	0.41	0.18
LLOY	t	-544.65	0.44	0.19	0.03
MB	t	-628.62	0.48	0.3	0.03
MTRO	Gaussian	-33.41	0.34	0	0.8
NDA-DK	t	-428.18	0.49	0.33	0.67
P9O	Frank	-23.56	0.21	0	0.61
RBI	Survival BB1	-347	0.44	0.46	0.62
RBS	t	-618.03	0.46	0.28	0.19
SAB	t	-445.27	0.45	0.32	0.01
SAN	t	-941.38	0.58	0.52	0.07
STAN	t	-460.65	0.41	0.23	0.5
UBI	t	-511	0.51	0.37	0
UBS	t	-689.48	0.51	0.42	0.43
UCG	t	-1034.08	0.61	0.51	0

Table 5.2: Copula Goodness-of-fit

The results obtained shows a predominance of t-copula to model the correlation structure between each financial institution returns and the system returns, but there is no guarantee that the copula family with the best fit for each financial institution will be the same.

The goodness-of-fit obtained in the most part of the cases point us for a copula function with tail dependence as it is the case for a t-Copula, Gumbel, BB7, Survival BB1 and Survival BB8.

Before we decide for a copula family it is important analyse some the result obtained from fitting the bi-variate data to a copula and analyse how well this fit works on the tails.

Taken as example $LLOY$ financial institution, we will select the best fit copula based on AIC criterion wit the following results:

Copula Family	t	
Parameter	rho:	0.64
	df:	8.47
Dependence measures	Kendall's tau:	0.44
	Upper TD:	0.18
	Lower TD:	0.18
Fit statistics	logLik:	273.23
	AIC:	-542.47
	BIC:	-532.54

Table 5.3: Copula fitted for LLOY

This copula fit represents a dependence structure with symmetric tails as we can also visualize in the figure below where we can identify the symmetric tails.

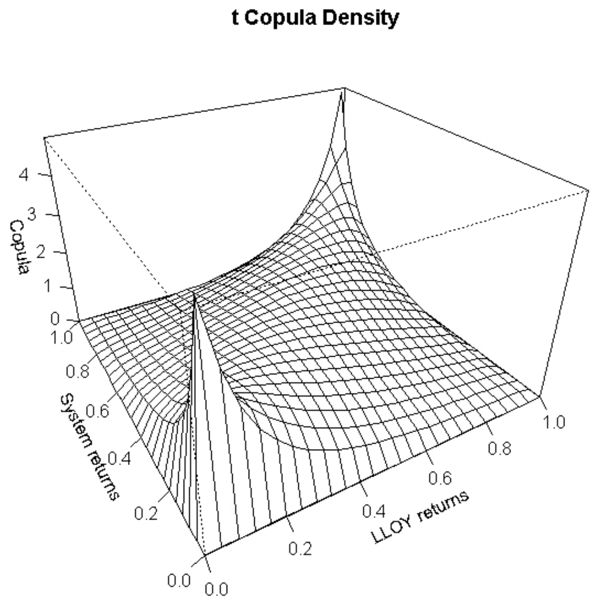


Figure 5.4: t-Copula density for LLOY

In order to visualize how the fit is close to the original data we can plot both bivariate data together in a 2-dimensional graph.

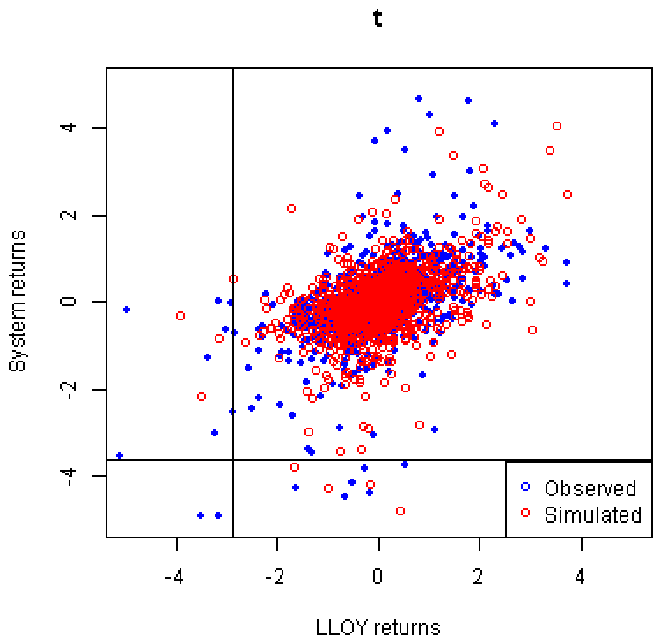


Figure 5.5: t-Copula vs empirical data for LLOY

We can denote here a good fit on the most dense section of the distribution what turns in an optimized value for AIC criterion, but if we put more attention on the tails we can also observe that the most extreme values are not followed with this copula function, in this case a t-Copula.

As the dependence concept is described as measure of association between, in this case two random variables, considering all their range, this dependence measures could not reflect properly the behavior on the tails and at the same time the dependence in the central component of those distributions.

Tail dependence, on the other hand, is related to the association of the variables over the extremes measuring that association on the tails of the joint distribution. This way, dependence and tail dependence are not the same, and two random variables could exhibit dependence without dependence in the tail.

The graphs below are split in quadrants where the vertical and horizontal line represents a quantile in the left tail (1%). This way the lower left quadrant is where we expect to found the extreme movements involving systemic risk represented.

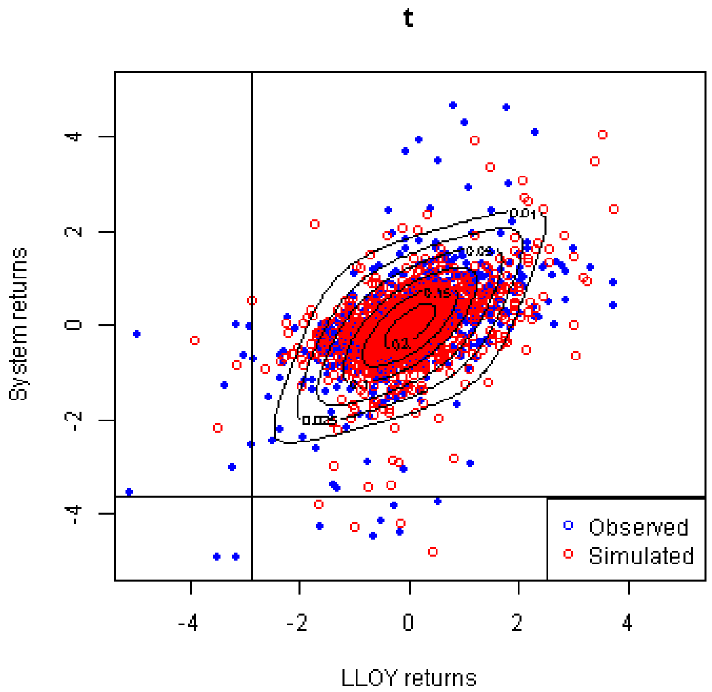


Figure 5.6: Generated t-Copula vs empirical data for LLOY

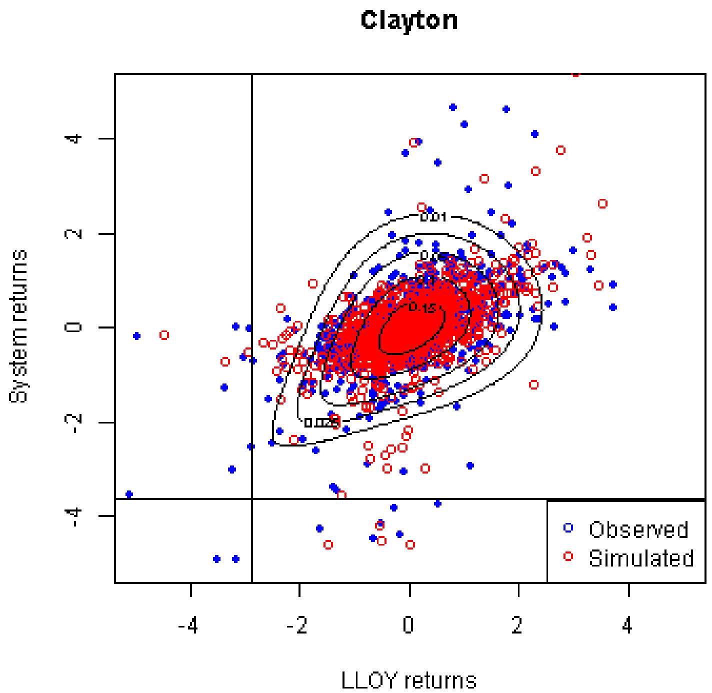


Figure 5.7: Generated Clayton Copula vs empirical data for LLOY

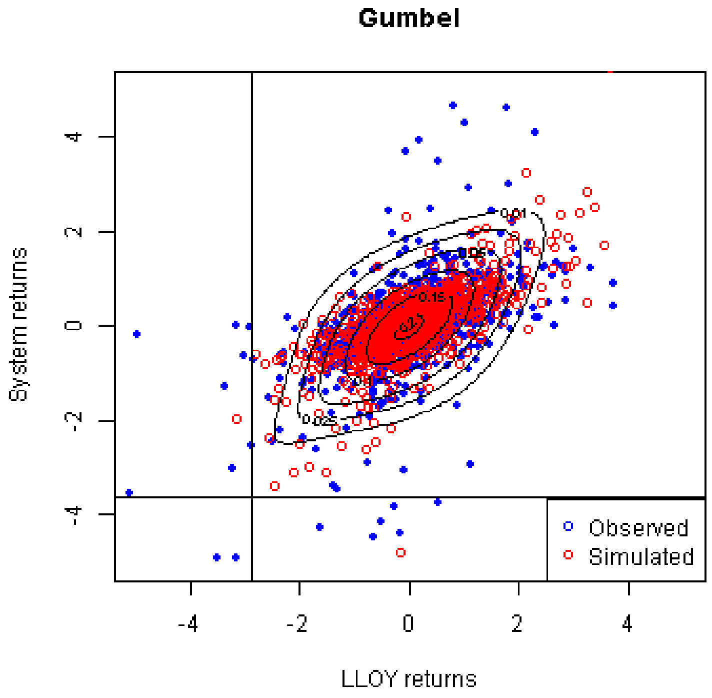


Figure 5.8: Generated Gumbel Copula vs empirical data for LLOY

Even though the all the different copulas were fit to the same data, the results on the extreme left tail are different. The t-Copula only generates one observation in this quadrant, while Clayton copula generates eight. Gumbel copula on other hand generates no data on that quadrant.

One option is to select a copula from a different copula family, that could adjust better to extreme values and fit better the tails of the distribution.

		Tail Dependence	
Copula Family	τ	Lower	Upper
t	0.442	0.1814	0.1814
Clayton	0.3381	0.5073	0
Gumbel	0.4176	0	0.5027
tev	0.39	0	0.5066
Husler Reiss	0.3914	0	0.468
Galambos	0.4135	0	0.4948
Frank	0.4571	0	0
Joe	0.6776	0	0.8515

Table 5.4: Dependence and Tail dependence for different Copula fitted for LLOY

The options to model lower tail dependence will resume to the case of t-Copula, which has symmetric tail dependence, and Clayton copula which has all tail dependence on lower tail and stronger than in t-Copula.

If we choose a copula with higher tail dependence, in this case lower tail dependence, we are associating a higher probability of an extreme event in both, in the financial institution and in the Financial System simultaneously.

Even though usually, there is a tendency to focus on the central moments ignoring the behavior in the tails, in a context of risk management the tail behavior and tail dependence is critical.

Another aspect in consideration is the copula symmetry. Some copulas are symmetric as t-Copula for example and others will not be symmetric as Clayton copula or other extreme value copulas. But in a context that involves risk management could not be the most appropriate to model extreme loss and extreme profits in the same way to avoid any undervaluation of the risk involved.

Event though the process to select a copula will rely on the data to determinate the shape of the copula, with copulas this shape is also a product of theoretical considerations and not only a result of data parameterisation. In this case, before to start the fitting process we should also decide on which copula family is more adequate to a natural interpretation and makes more sense given any prior knowledge of the risks.

5.4.3 Copula Selection and Tail Dependence

An unavoidable component of systemic risk analyses is the presence of tail dependence.

The inclusion of tail dependence on the model could be achieved by selecting the appropriate copula family. Here the strategy is not locking only looking for the usual selection criterion as AIC or BIC, but include here also the theoretical tail dependence coefficient of the copula as a results of the following expression obtained for each copula family:

$$\lambda_L = \lim_{u \searrow 0} \frac{C(u,u)}{u}$$
$$\lambda_U = \lim_{u \nearrow 1} \frac{1 - 2u + C(u,u)}{1-u}$$

As we are mainly interested in lower tail dependence, the selection criterion applied will compare lower tail coefficient.

This selection criterion will be used also with a goodness-of fit test as specified and implemented in R package *VineCopula* (Schepsmeier et al., 2015).

The hypothesis test is defined as :

$$H_0 = \boldsymbol{H}(\theta) + \boldsymbol{C}(\theta) = 0$$
$$H_1 = \boldsymbol{H}(\theta) + \boldsymbol{C}(\theta) \neq 0$$

where $\boldsymbol{H}(\theta)$ is the expected Hessian matrix and $\boldsymbol{C}(\theta)$ is the expected product of the score function.

As an example we will go through the results obtained for BBVA financial institution.

Copula Family	AIC	BIC	Lower Tail Dependence	p-Value
Gaussian	-926.4388	-921.5898	0.0000	0.42
t	-1003.3092	-993.6111	0.4997	0.15
Clayton	-830.1377	-825.2886	0.7190	0.00
Gumbel	-972.1559	-967.3069	0.6686	0.00
Frank	-898.8010	-893.9519	0.0000	0.00
Joe	-821.4519	-816.6028	0.7323	1.00

Copula Family	AIC	BIC	Lower Tail Dependence	p-Value
BB1	-993.3541	-983.6560	0.6149	0.61
BB6	-970.0841	-960.3860	0.6687	0.00
BB7	-966.5711	-956.8729	0.6754	0.00
BB8	-902.6389	-892.9408	0.0000	0.00

Table 5.5: Copula Goodness-of-fit for BBVA

From these results only the copula families that exhibit a significant p-Value will be considered. From those we will take the one with higher lower tail coefficient. This criterion leads us to Joe copula, that looks like the best approach to model the tails of the copula. As the AIC criterion represents the best fit for all the distribution we will not necessarily have a match between AIC criterion and tail dependence coefficient.

In similar way, we can proceed with the choice for copula to be used to model the sections outside of the tail, the bulk section of the bivariate distribution, but now we should use the AIC (or BIC) criterion, as we are looking now for a better fit over all the distribution (not only in the tail). We have as better fit for the purpose the t copula. We can also notice that Gaussian copula also have similar results with stronger significance (0.42 against 0.15 on t-copula).

In conclusion, by running this sample bootstrap on an wide number of copula families, we conclude that:

- By use a Joe copula for the tail, when the financial institution is in situation of stress.

- By use a t copula to model the case where the financial institution is not in stress ($\alpha = 0.5$).

5.4.4 Copula Conditional Distribution Function

In order to obtain the conditional probabilities that involved in the calculation of Financial System *CoVaR*, we need to determine also the conditional distribution function of the bivariate copula, we fit to the data:

$$h_1(u|v;\theta) := P(U \leq u|V = v) = \frac{\partial C(v,u;\theta)}{\partial v},$$

where $(U, V) \sim C$ is a bivariate copula distribution function with parameters θ.

In terms of systemic risk measure this results could be translated as:

$$P(R_{s,t} \leq CoVaR_{\alpha,\beta,t}|R_{i,t} = VaR^{j}_{\alpha,t}) = \frac{P(R_{s,t} \leq CoVaR_{\alpha,\beta,t})}{P(R_{i,t} \leq VaR^{j}_{\alpha,t})}$$

as described in detail in A.4.2 and A.4.3 sections.

This process will be used then twice, one will estimate the *CoVaR* when the financial institution faces a stress situation and other when the financial institution is average returns. The*CoVaR* corresponds to:

$$CoVaR_{\alpha\beta t} \tag{5.3}$$

in the first situation and in the second situation by:

$$CoVaR_{0.5\beta t} \tag{5.4}$$

To finalize the calculation process for *CoVaR* we need yet to calculate the inverse of the cumulative distribution function of the copula marginal for the financial system returns.

5.4.5 Fitting the Margins

As the last step to obtain the CoVaR estimate, we have to fit the financial system returns series in order to, with that probability function, use the correspondent inverse of the cumulative distribution function to finally calculate the CoVaR value. Any assumption on this regard and the decision on which probability distribution function to use will have a huge impact on the result.

In order to mitigate the inconveniences of using a normal assumption, the option to model the financial system returns will fall on a mixture model.

The idea behind the use of Mixture of Extreme Value distribution is to combine the flexibility of using a distribution to capture the main component, corresponding to the central quantiles, also referred as the bulk distribution, that could be for example a Normal, and also the tails, as extreme values. With this mixture model one will get an entire distribution function by splitting the distribution in a bulk component and a tail component. The mixture function allows for a mixture of distribution from distinct families as well.

For the purpose of modelling systemic risk we are interested in exploring a mixture of a Normal distribution as bulk distribution with two Gamma tail distributions in both upper and lower tail (MacDonald et al., 2011).

This model uses kernel density estimators to estimate the non-extreme value distribution and GPD to estimate the tail distribution. This kernel density estimator assumes a normal density, which is centered at each data point, and uses only one parameter to define bandwidth.

The tail fractions refers to the proportion of the distribution above the threshold. This parameter will be identified by Φ_u and u represents the threshold.

The distribution function comes as:

$$F(x|\Theta) = \begin{cases} \phi_{u_l}\Big(1 - G(-x| - u_l, \sigma_{u_l}, \epsilon_l)\Big), & x < u_l \\ \\ H(x|\mu, \sigma), & u_l \leq x \leq u_r \\ \\ (1 - \phi_{u_r}) + \phi_{u_l} G(x|u_r, \sigma_{u_r}, \epsilon_r) & x > u_r \end{cases}$$

where $\phi_{u_l} = H(u_l|\mu, \sigma)$ and $\phi_{u_r} = 1 - H(u_r|\mu, \sigma)$ and $H(.|\mu, \sigma)$ is the normal distribution with mean μ and standard deviation σ. $G(.| - u_l, \sigma_{u_l}, \epsilon_l)$ and $G(.| - u_r, \sigma_{u_r}, \epsilon_r)$ are GPD distributions for lower and upper tails respectively.

By applying the extreme value mixture model to Financial System returns series we obtained the following estimates for the parameters:

- The graph shows the results for a mixture of a normal N(-0.017,0.787) bounded at left by parameter u_l = -0.940 and on right by parameter u_r = 0.909.
- The Gamma parameters obtained are respectively :

left tail	**right tail**
ϕ_{u_l} = 0.0669	ϕ_{u_r} = 0.129
μ_l =0.3858	μ_r = 0.2218
σ_l =0.0139	σ_r = 0.0174

The goodness of fit for this model is also slightly better than previous ones with an BIC criterion value estimated as -2356.741. The advantage and flexibility of this mixture model is essentially in the tails of the distribution as it is able to take advantage of the capabilities of Gamma distribution to adapt to the tail. The graphs obtained for the extreme value mixture are as follows:

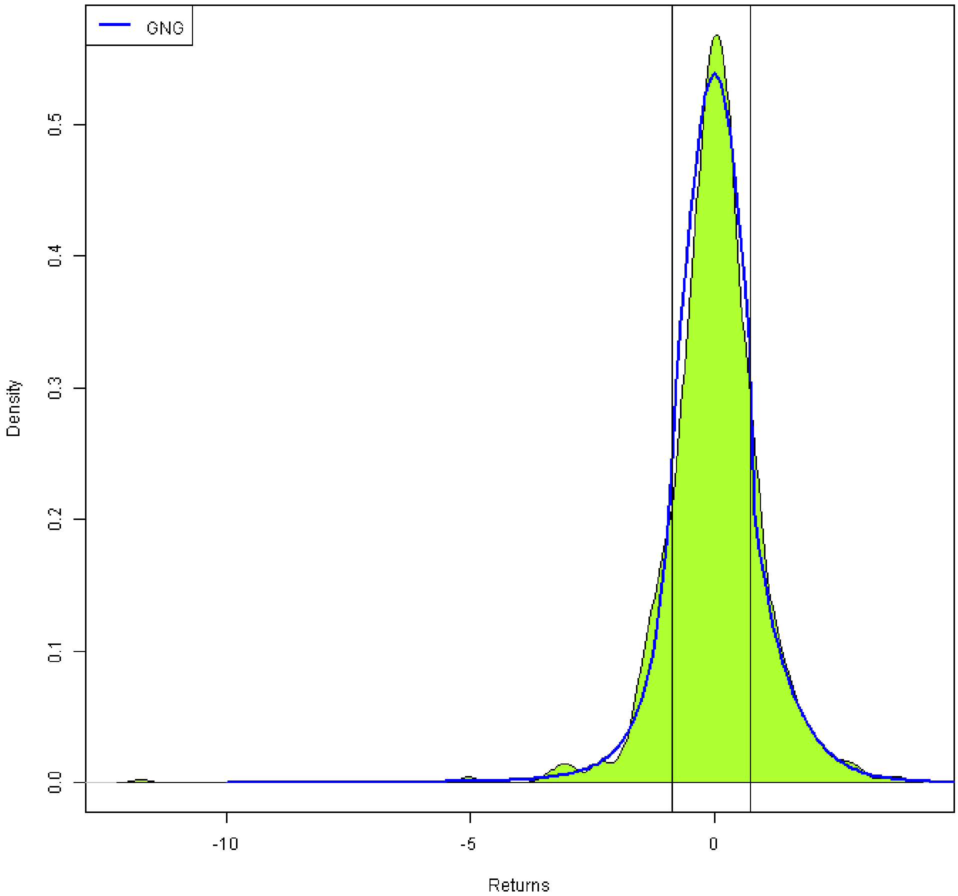


Figure 5.9: Fit of Financial System weekly returns with a mixture of Gamma-Normal-Gamma

By visual analysis of the graph obtained with the Gamma-Normal-Gamma (GNG) fitting it is possible to identify a very close adjustment.

Also in the tail of the distribution we can notice a good approximation including a decay of the distribution function when it goes further on the left.

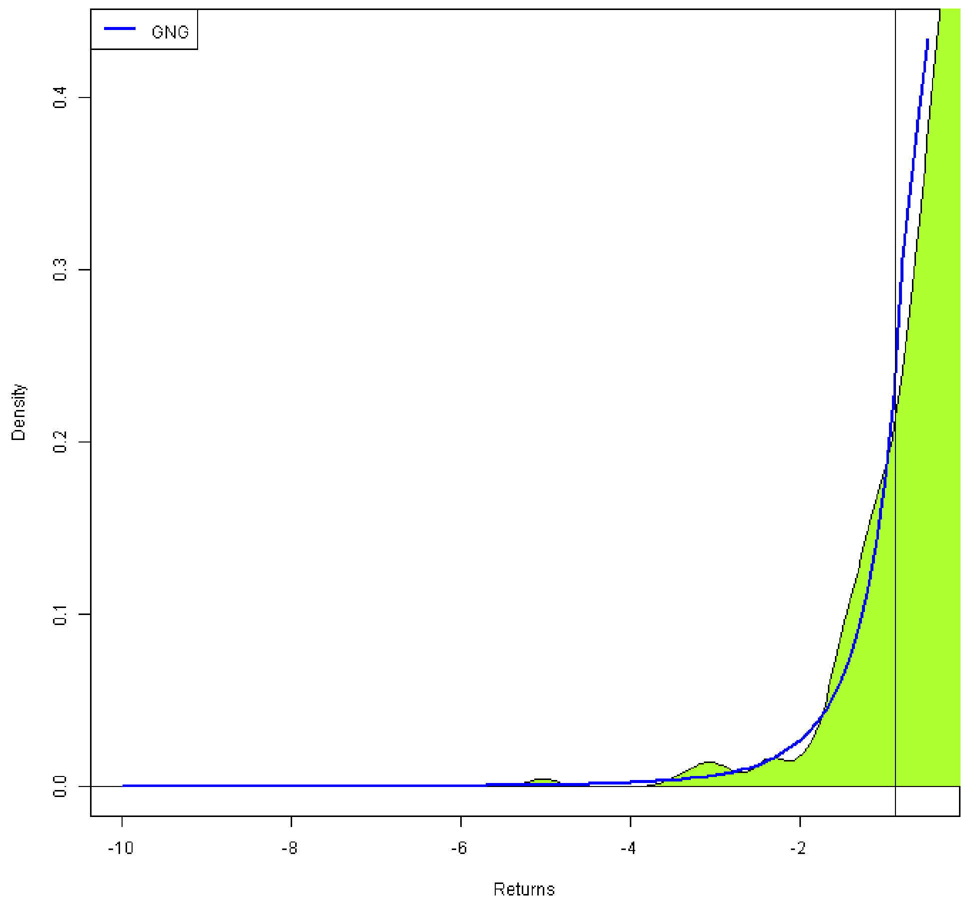


Figure 5.10: Fit of Financial System weekly returns with a mixture of Gamma-Normal-Gamma on the left tail

Instead of using a single distribution to adjust the financial returns series over all the quantiles, we can than use a mixture model to adjust distinct distributions accordingly to the quantiles.

5.4.6 $\Delta CoVaR$ Results

On the following table, we can compare the impact that each modeling option, by selecting distinct copula families has in terms of the resulted calculated for $CoVaR$ and $\Delta CoVaR$.

Copula Family	$CoVaR_{0.5}$	$CoVaR_{\beta=0.01}$	$CoVaR_{\beta=0.001}$	$\Delta CoVaR_{\beta=0.001}$
t	-0.0445	-0.2111	-0.4990	-0.4545
Clayton	-0.0742	-0.0982	-0.0986	-0.0244
Gumbel	-0.0570	-0.1827	-0.2760	-0.2190
t & Clayton	-0.0445	-0.0982	-0.0986	-0.0541
t & Gumbel	-0.0445	-0.1827	-0.2760	-0.2315

Table 5.6: $CoVaR$ and $\Delta CoVaR$ results for LLOY ($\alpha = 0.01$)

Considering three copula families, t as a symmetric example with symmetric weak tail dependence, and on other side the Clayton copula that is left tailed and Gumbel copula that is right tailed.

The t-copula, as it is symmetric and will allow for better results if used to estimate only the $CoVaR_{0.5}$, on the median, representing a situation where the financial institution is no subject to distress, also interpreted as the the expected situation in normal circumstances.

On the tails of the joint distribution of financial institution returns and financial system returns, t-copula tends to overestimate the risk due to the difficulties in dealing with tail dependence.

5.4.7 Identifying Financial Institutions with Systemic Risk

The risk codependence relations between the financial institutions considered were estimated by using the methodology described, by applying two distinct ways of selecting the copula. This approach reveals it adequate to estimate the systemic relations to incorporate tail dependence versus dependence over the median (Embrechts et al., 2001).

$\Delta CoVaR$ represents an estimate of the magnitude of each financial entity contribution to the systemic market risk.

The table above shows the estimated values of $\Delta CoVaR$ for June 2018, by applying a 3 years rolling window.

Rank	Financial Institution	$\Delta CoVaR_{\beta=0.01}$	τ	Copula (median)	Copula (tail)
1	BNP	6.94%	0.49	t	Clayton-Gumbel (BB1)
2	ISP	6.20%	0.51	t	Clayton-Gumbel (BB1)
3	BIRG	6.00%	0.44	t	Clayton-Gumbel (BB1)
4	GLE	5.97%	0.48	t	Clayton-Gumbel (BB1)
5	CBK	5.93%	0.46	t	Clayton-Gumbel (BB1)
6	MB	5.90%	0.47	t	Clayton-Gumbel (BB1)
7	ABN	5.86%	0.51	t	Clayton-Gumbel (BB1)
8	SAN	5.79%	0.41	t	Clayton-Gumbel (BB1)
9	RBI	5.66%	0.44	t	Clayton-Gumbel (BB1)
10	NDA-DK	5.61%	0.47	t	Clayton-Gumbel (BB1)

Table 5.7: Top 10 Financial institutions by $\Delta CoVaR$ - June 2018

This systemic importance rank for financial institutions, as systemic risk measure is not necessarily linked to, or even shows a possible situation of the distress of a particular financial institution. Instead, the systemic risk measure rank reflects the expected additional impact as a cost to the financial system, given the fact that such an event occurs in the specific financial institution.

The above table lists financial institutions ordered by its systemic risk estimated impact. In brackets is the % of $\Delta CoVaR$ in system VaR, representing the impact of a significant event in that financial institution reflected in system VaR.

Similarly and for comparison purposes, the table below also has a lists of financial institutions ordered by its systemic risk estimated now by three methodologies based on quantile linear regression.

Table 5.8: Top 10 Financial institutions by $\Delta CoVaR$ - June 2018

Rank	Historic	Parametric	EVT
1	HSBA (7,4%)	HSBA (9,9 %)	HSBA (8,9%)
2	SAN (5,6%)	SAN (6,6 %)	SAN (7,7%)
3	ING (5,2%)	ING (5,9 %)	BANK45 (6,3%)
4	UCG (5,1%)	UCG (5,2 %)	ING (5,8%)
5	ISP (4%)	DBK (5 %)	ISP (5,1%)
6	DBK (3,6%)	ISP (4,4 %)	DBK (4,5%)
7	STAN (3,3%)	STAN (3,6 %)	STAN (3,8%)
8	RBS (2,5%)	RBS (3,5 %)	RBS (3,3%)
9	GLE (2,5%)	GLE (3,1 %)	GLE (3,2%)
10	ACA (2,2%)	ACA (2,7 %)	UBS (2,8%)

[a]

[a]*in () $\Delta CoVaR$ % on system VaR

Because the Copula will establish a non linear relationship between the two series in opposition to the quantile linear regression, there is a significant difference between results from Copula methods and the other $\Delta CoVaR$ calculation methods.

5.5 Adding Time Variation

In the original paper Adrian and Brunnermeier (2011) included an additional layer of assumptions making the institution returns, X dependent on a set of state variables and assuming an underlying factor model for asset returns, where the return on each asset depends linearly on these factors:

- A set of lagged state variables $M - t - 1$ (to be defined shortly)
- The system-wide growth in assets X^{sys}

This way, the asset growth of each financial institution will depend on selected lagged state variables, while the growth rate of system assets depends on individual bank asset growth and lagged state variables.

As we aim is to remove additional assumptions from the model, an alternative to avoid this extra layer of assumptions is to apply a rolling windows technique as a way to include time variance to have an analysis through time (Chong and Hurn, 2016). The caveat here is we will lose a part of the initial data set. Defining a window length, for example, 3 years of data, we will be moving this window day by day, or week by week and applying all the previous steps to each one of those time windows. In the end we will get a time series for VaR, $CoVaR$ and $\Delta CoVaR$.

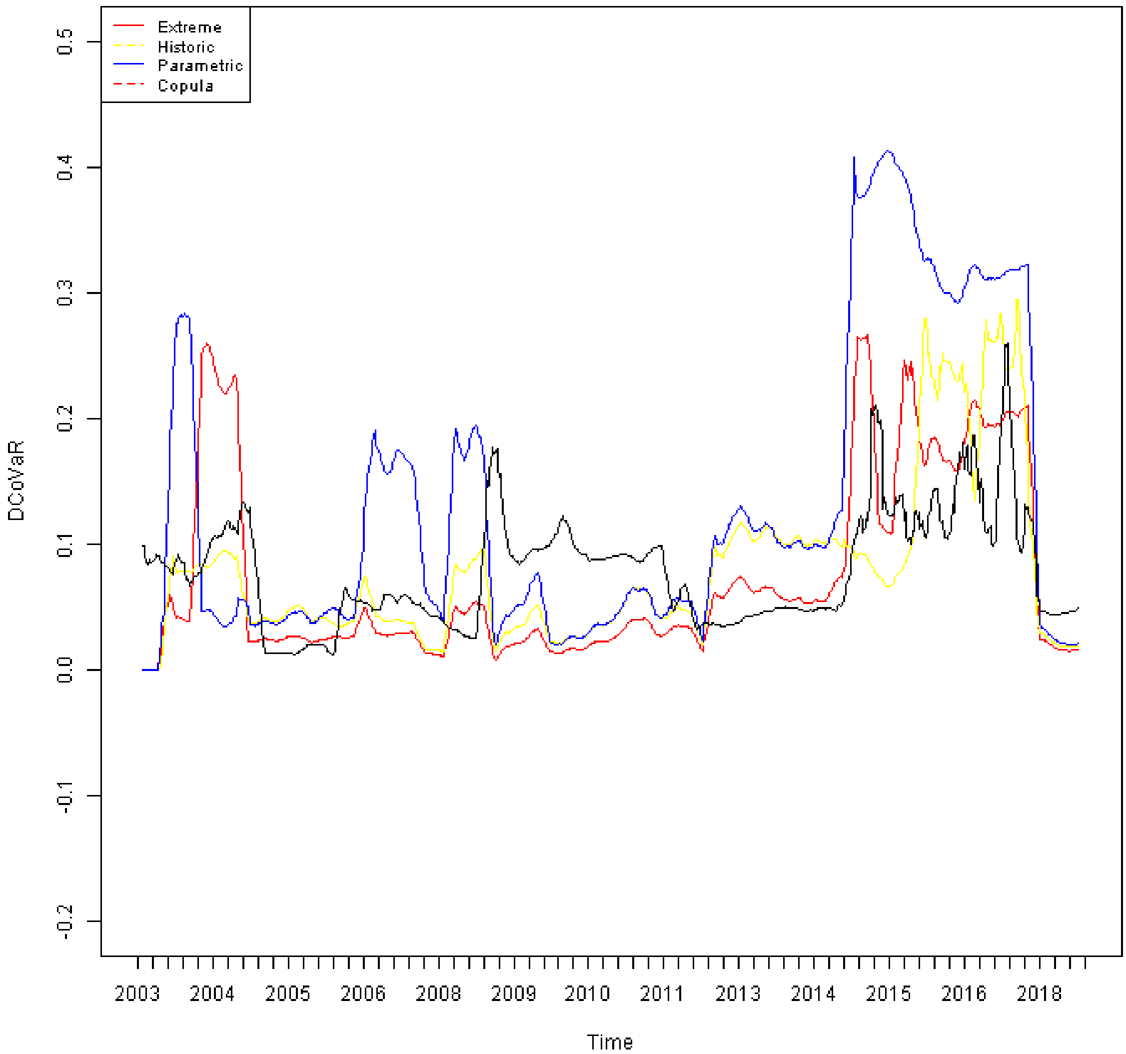


Figure 5.11: $\Delta CoVaR$ for UBS using four methodologies: Extreme Value, Parametric, Historic and Copula

Another caveat we can refer to regarding using this method is also the high computational cost of it. Despite some of its shortcomings, the rolling-window procedure is accessible to implement and easy to interpret also without including additional assumptions.

By applying the rolling window to estimate the $\Delta CoVaR$, for each point in time is possible then to have a perception of the evolution of $\Delta CoVaR$ thought time.

In this case $\Delta CoVaR$ is being estimated for each point in time (week) and the copula parameters are being adjusted to each time window defined as a way to update the structure of dependence between both returns series.

5.6 Discussion of the Results

The systemic risk $\Delta CoVaR$, as calculated in the previous sections, can be used to provide a comprehensive and unified statistical profile of Banks according to their implication level (contribution and exposure) in systemic risk. Thus, we set a detailed map to show the relative positioning of all banks according to their implication into systemic risk.

We had used Adrian and Brunnermeier (2016) $CoVaR$ methodology which is defined as the VaR of the whole financial system given that one of the financial institutions is in distress.

Using data collected for the period 1998 to 2018, our results indicate that UBS is the largest contributor to the banking sectors systemic risk in Europe.

By comparing the three different methods to obtain Var, Historic and Parametric methods are given similar results. When comparing with those methodologies, VaR calculated by using EVT offers a better fit (less violations of VaR value by loss).

By applying the methodology suggested in this thesis, we are exploring modelling solutions that are more flexible, as it is not dependent of linear regression assumptions between the returns of the financial institution and the returns of the financial system. As we also showed, this relation is not linear neither constant across a distribution function, so we include also a component of mixture of distribution in order to improve the fitting of the returns in the tail of the distribution.

By applying an approach of fitting two copula functions, one in the tails another in the central qauntiles also proved to be advantageous in order to capture better the dependence between the two return series, as the nature of this dependence tends to change depending the quantiles of each distribution.

The results obtained in terms of the adjustment also showed that while in the middle of the distribution (the central quantiles) a t-Copula proved to be the best fit for the generality of the financial institutions in the data set used, in the tails the Gumbel copula proved to be more suitable.

Chapter 6

Conclusion

One of the aims of this thesis was to explore additional options to identify systemic important financial institutions, the so called SIB, by using market public available data.

There are several advantages that one can associate, by using market public available data as a source of information to the process of identify the systemic important financial institutions, starting by the clarity and transparency in the process and also to make possible to have this type of insight over the financial system, independently of the programmed periodically reports that each institution have to communicate to the regulator.

Even though the multiple possibilities and techniques to calculate the $\Delta CoVaR$ systemic risk measure, most recent research work have been arguing on the advantages of using copula related techniques, in order to model the dependence between two random variables, and in special as an option to be considered when the relation between those variables is not well defined as a linear regression. Often this is the case with financial returns.

We have used as a start point Adrian and Brunnermeier (2011) $CoVaR$ methodology which is defined as the VaR of the whole financial system given that one of the financial institutions is in distress. This distress event is also usually associated with the financial institution ave reached its VaR as well.

In the original $CoVaR$ methodology the authors have applied a quantile regression technique to estimate the daily $CoVaR$ and then $\Delta CoVaR$. The quantile regression is easy to implement and also provide a a easy way to interpret the results, but it could be limited in order to capture correctly the dependence structure between the financial institution returns and the financial system returns.

Using financial data, collected for the period 1998 to 2018, the results obtained first by applying a quantile regression approach on three different methods to obtain the VaR, based on historical information, with a parametric approach and also with a methodology based on extreme values, focused on the behavior on the tails, we have identified HSBA as the largest contributor to the banking sectors systemic risk in Europe and it was also observed that:

- Historic, parametric and EVT methods identified a very similar top ten of risky institutions.
- Historic and parametric approaches more close.
- Some changes on positions but with a similar set of institutions.

Based our findings and results it also indicates that the contribution of institutions to systemic risk is linked to the size of the institution, with the larger institutions contributing more than the smaller ones.

By comparing the three different methods to obtain *VaR*, the Historic and Parametric methods are given similar results. When comparing with those mythologies, EVT *VaR* offers a better fit (less violations of *VaR* value by loss).

However, EVT requires very few data as input (only used 5 records in each window) on our example. Depending on the scenario and the use case, this feature could be an advantageous on. This characteristic makes it suitable to be used, for example, in scenario analyses, for instance since it will be easier to get reliable results using a limited data set as input.

The results obtained when we replaced the quantile regression based approach by a copula function approach evidenced some difference in the results. This difference in the approach followed means we are replacing the assumption based on a linear relationship between the two series, financial institution returns and financial system returns, by a more elaborated way to model the dependence, now based on the join distribution defined by the two marginal distributions, defined also in terms of the observations supplied on both return series, and the copula function we construct based on these marginal distributions.

CoVaR was calculated also by using a methodology based on bivariate copulas theory, due to the copula capabilities to model dependence and in this case tail dependence. We pointed out the importance of tail dependence in a scenario of risk analysis, and in particular in systemic risk analysis as the more rare cases are localized in the tail of the distributions, and even with small probability of occurrence, these rare events are the motivation for risk analysis due to the potential impact.

The copula approach allows for results focused in the marginal impact on the risk estimated for each financial institution.

The results obtained highlighted a concern regarding the quality of the global adjustment versus the quality of the adjustment on the tails of the distribution. The quality of the adjustment can also vary depending on the segment or quantile of a distribution one is analysing. this leads to the conclusion that to obtain a good fit, an approach that applies with several distinct distribution functions can be preferable in opposition to a single distribution function approach.

In some applications the analyses of the tail of the distributions is of major importance, as for example in risk analyses as mentioned above. When we bring a copula approach to the solution model, it is desirable that that model is also able to deal with tail dependence, as we should also account for a special dependence relation in the tail. In terms of financial returns, it means that for example in extreme conditions, or in the extremes of the distribution of returns, the dependence is usually stronger in the context of a financial crises than in a normal scenario. A suitable model needs then to be able to accommodate also these kind of different situations.

In this use case we applied two criteria to select the copula function, one based on the best overall adjustment, to estimate the *CoVaR* in the scenario without financial stress in the financial institution, and another based on the best fit to model tail dependence.

Additionally, in order to obtain a correct copula function fitting is also of major importance to assure a suitable fit on the marginal distributions, in special the marginal that represents the financial system returns, in such a way that that fitting provides good results not only on the overall distribution but also in the tails of the distribution.

In order to accomplish with those requirements, we have implemented an extreme value mixture model to fit the financial results, which gave us the flexibility to adapt to the tail of the distribution providing better adjustments.

Complex phenomenons require also more complex models and the complexity of certain phenomenons like behavior of financial returns requires more complex and versatile models. In this case, to identify systemic riskiest financial institutions based on $CoVaR$ is fundamental to dispose of methodologies that allow for a correct tail dependence model, and the fitting methods that could work well for an overall distribution fitting purpose could not be the most appropriated for fitting the tails of the distribution.

In future developments and improvements of this model, the results obtained here with the copula approach could also be generalized with a copula regression function as a solutions to model the dependence between the two series, financial system returns and financial institution returns. From a copula regression function, the methodology could also be generalized by applied copula quantile regression in order to obtain a more generic and flexible models.

Bibliography

Acerbi, C. and Tasche, D. (2002). Expected shortfall: a natural coherent alternative to value at risk. *Economic notes*, 31(2):379–388.

Acharya, V., Brownless, C., Engle, R., Pedersen, L., Philippon, T., and Richardson, M. (2012a). How to measure and regulate systemic risk. *NYU Stern School of Business*.

Acharya, V., Engle, R., and Richardson, M. (2012b). Capital shortfall: A new approach to ranking and regulating systemic risks. *American Economic Review*, 102(3):59–64.

Acharya, V. V., Pedersen, L. H., Philippon, T., and Richardson, M. (2010). Measuring systemic risk. *Working Paper*.

Adrian, T. and Brunnermeier, M. K. (2011). Covar. Technical report, National Bureau of Economic Research.

Adrian, T. and Brunnermeier, M. K. (2016). Covar. *American Economic Review*, 106(7):1705–41.

Ahooyi, T. M. (2015). *Probabilistic Modeling of Process Systems with Application to Risk Assessment and Fault Detection*. Drexel University.

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish journal of emergency medicine*.

Allen, L., Bali, T. G., and Tang, Y. (2012). Does systemic risk in the financial sector predict future economic downturns? *The Review of Financial Studies*, 25(10):3000–3036.

Alzaatreh, A., Lee, C., Famoye, F., and Ghosh, I. (2016). The generalized cauchy family of distributions with applications. *Journal of Statistical Distributions and Applications*, 3(1):1–16.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical finance*, 9(3):203–228.

Barton, R. (2013). *Talend Open Studio Cookbook*. Packt Publishing Ltd.

Basilio, J. and Oliveira, A. (2020). Fitting heavy tail distributions with mixture models. In *Proceedings 6th Stochastic Modeling Techniques and Data Analysis International Conference with Demographics Workshop*, pages 53–68. ISAST: International Society for the Advancement of Science and Technology.

Basilio, J., Oliveira, A., and Mahmoudvand, R. (2020). An overview of the systemic risk measures. In *AIP Conference Proceedings*, volume 2293. AIP Publishing LLC.

Bedford, T., Cooke, R. M., et al. (2002). Vines–a new graphical model for dependent random variables. *The Annals of Statistics*, 30(4):1031–1068.

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. L. (2006). *Statistics of extremes: theory and applications*. John Wiley & Sons.

Benoit, S., Colliard, J.-E., Hurlin, C., and Pérignon, C. (2017). Where the risks lie: A survey on systemic risk. *Review of Finance*, 21(1):109–152.

Bernard, C. and Czado, C. (2015). Conditional quantiles and tail dependence. *Journal of Multivariate Analysis*, 138:104–126.

Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2010). Econometric measures of systemic risk in the finance and insurance sectors. *NBER WORKING PAPER SERIES*. "http://www.nber.org/papers/w16223".

Brownlees, C. and Engle, R. F. (2016). Srisk: A conditional capital shortfall measure of systemic risk. *The Review of Financial Studies*, 30(1):48–79.

Brownlees, C. T. and Engle, R. (2012). Volatility, correlation and tails for systemic risk measurement. *Available at SSRN*, 1611229. "http://vlab.stern.nyu.edu/".

Bücher, A. and Segers, J. (2017). On the maximum likelihood estimator for the generalized extreme-value distribution. *Extremes*, 20(4):839–872.

Bücher, A., Volgushev, S., and Zou, N. (2019). On second order conditions in the multivariate block maxima and peak over threshold method. *Journal of Multivariate Analysis*, 173:604–619.

Capéraà, P., Fougères, A.-L., and Genest, C. (2000). Bivariate distributions with given extreme value attractor. *Journal of Multivariate Analysis*, 72(1):30–49.

Chengli, Z. and Yan, C. (2012). Coherent risk measure based on relative entropy. *Appl. Math*, 6(2):233–238.

Cherubini, U., Luciano, E., and Vecchiato, W. (2004). *Copula methods in finance*. John Wiley & Sons.

Chong, J. and Hurn, A. (2016). Testing for speculative bubbles: Revisiting the rolling window. *School of Economics Discussion Paper Series*.

Coles, S. (2013). *An Introduction to Statistical Modeling of Extreme Values*. Springer Science & Business Media.

Danielsson, J., Shin, H. S., and Zigrand, J.-P. (2012). Endogenous and systemic risk. In *Quantifying systemic risk*, pages 73–94. University of Chicago Press.

Day, N. E. (1969). Estimating the components of a mixture of normal distributions. *Biometrika*, 56(3):463–474.

de Oliveira, J. T. (2013). *Statistical extremes and applications*, volume 131. Springer Science & Business Media.

de Zea Bermudez, P. and Kotz, S. (2010). Parameter estimation of the generalized pareto distribution—part i. *Journal of Statistical Planning and Inference*, 140(6):1353–1373.

Delbaen, F. (2000). Coherent risk measures. *Blätter der DGVFM*, 24(4):733–739.

Delbaen, F. (2002). Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, pages 1–37. Springer.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22.

Dowd, K. (2003). *An introduction to market risk measurement.* John Wiley & Sons.

Drake, J. D. and Worsley, J. C. (2002). *Practical PostgreSQL.* " O'Reilly Media, Inc.".

Drees, H. and Huang, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *Journal of Multivariate Analysis*, 64(1):25–46.

Duffie, D. and Pan, J. (1997). An overview of value at risk. *Journal of derivatives*, 4(3):7–49.

Durante, F. and Jaworski, P. (2010). A new characterization of bivariate copulas. *Communications in Statistics—Theory and Methods*, 39(16):2901–2912.

Eijffinger, S. (2011). Defining and measuring systemic risk. In *Handbook of Central Banking, Financial Regulation and Supervision.* Edward Elgar Publishing.

Embrechts, P., Lindskog, F., and McNeil, A. (2001). Modelling dependence with copulas. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich*, 14.

Embrechts, P., Resnick, S., and Samorodnitsky, G. (1998). Living on the edge. *Risk*, 11(1):96–100.

Fabozzi, F. J., Gupta, F., and Markowitz, H. M. (2002). The legacy of modern portfolio theory. *The Journal of Investing*, 11(3):7–22.

Fallon, W. et al. (1996). *Calculating value-at-risk.* Wharton School, University of Pennsylvania.

Fernandez-Sanchez, J. and Ubeda-Flores, M. (2016). Proving the characterization of archimedean copulas via dini derivatives. *Kybernetika*, 52(5):785–790.

Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge University Press.

Föllmer, H. and Schied, A. (2008). Convex and coherent risk measures. *Encyclopedia of Quantitative Finance*, pages 355–363.

Fougères, A.-l., Nolan, J. P., and Rootzén, H. (2009). Models for dependent extremes using stable mixtures. *Scandinavian Journal of Statistics*, 36(1):42–59.

Frittelli, M. and Gianin, E. R. (2002). Putting order in risk measures. *Journal of Banking & Finance*, 26(7):1473–1486.

Fúquene Patiño, J. A. (2015). A semi-parametric bayesian extreme value model using a dirichlet process mixture of gamma densities. *Journal of Applied Statistics*, 42(2):267–280.

Gilleland, E., Katz, R. W., et al. (2016). extremes 2.0: An extreme value analysis package in r. *Journal of Statistical Software*, 72(8):1–39.

Girardi, G. and Ergün, A. T. (2013). Systemic risk measurement: Multivariate garch estimation of covar. *Journal of Banking & Finance*, 37(8):3169–3180.

Gudendorf, G. and Segers, J. (2010). Extreme-value copulas. In *Copula theory and its applications*, pages 127–145. Springer.

Guerrero-Cusumano, J.-L. (1996). An asymptotic test of independence for multivariate t and cauchy random variables with applications. *Information sciences*, 92(1-4):33–45.

Gumbel, E. J. (1958). Statistics of extremes, columbia univ. *Press, New York*, 201.

Gumbel, E. J. (2012). *Statistics of extremes*. Courier Corporation.

Hamilton, J. D. (1991). A quasi-bayesian approach to estimating parameters for mixtures of normal distributions. *Journal of Business & Economic Statistics*, 9(1):27–39.

Harris, D. E. (2017). The distribution of returns. *Available at SSRN 2828744*.

Haugen, R. A. and Haugen, R. A. (2001). *Modern investment theory*, volume 5. Prentice Hall Upper Saddle River, NJ.

Hendricks, D. (1996). Evaluation of value-at-risk models using historical data. *Economic policy review*, 2(1).

Hofert, M., Kojadinovic, I., Maechler, M., Yan, J., Maechler, M. M., and Suggests, M. (2014). Package 'copula'. *URL http://ie.archive.ubuntu. com/disk1/disk1/cran.r-project. org/web/packages/copula/copula.pdf*.

Hu, Y. and Scarrott, C. (2018). evmix: An r package for extreme value mixture modeling, threshold estimation and boundary corrected kernel density estimation. *R News*.

Hürlimann, W. (2003). Hutchinson-lai's conjecture for bivariate extreme value copulas. *Statistics & probability letters*, 61(2):191–198.

Ihaka, R. and Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of computational and graphical statistics*, 5(3):299–314.

Jadhav, D. and Ramanathan, T. (2009). Parametric and non-parametric estimation of value-at-risk. *The Journal of Risk Model Validation*, 3(1):51.

Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348):158–171.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts*. Chapman and Hall/CRC.

Joe, H. and Hu, T. (1996). Multivariate distributions from mixtures of max-infinitely divisible distributions. *Journal of multivariate analysis*, 57(2):240–265.

Joe, H. and Kurowicka, D. (2011). *Dependence modeling: vine copula handbook*. World Scientific.

Jondeau, E., Poon, S.-H., and Rockinger, M. (2007). *Financial modeling under non-Gaussian distributions*. Springer Science & Business Media.

Käärik, M., Selart, A., Käärik, E., and Liivi, J. (2011). The use of copulas to model conditional expectation for multivariate data. In *ISI Proc. 58th World Statistical Congress*, pages 5533–5538.

Ke, C. and Yin, X. (2019). Expected conditional characteristic function-based measures for testing independence. *Journal of the American Statistical Association*.

Kimball, R. C. et al. (2000). Failures in risk management. *New England Economic Review*, pages 3–12.

Levine, D. (2009). Modeling tail behavior with extreme value theory. *Risk Management*, 17:14–18.

Levine, R. A. and Casella, G. (2001). Implementations of the monte carlo em algorithm. *Journal of Computational and Graphical Statistics*, 10(3):422–439.

Li, J. (2012). *A Cauchy-Gaussian Mixture Model For Basel-Compliant Value-At-Risk Estimation In Financial Risk Management*. PhD thesis, Lehigh University. "http://preserve.lehigh.edu/etd".

MacDonald, A., Scarrott, C. J., Lee, D., Darlow, B., Reale, M., and Russell, G. (2011). A flexible extreme value mixture model. *Computational Statistics & Data Analysis*, 55(6):2137–2157.

Mai, J. and Scherer, M. (2014). *Financial engineering with copulas explained*. Springer.

Mai, J.-F. and Scherer, M. (2011). Bivariate extreme-value copulas with discrete pickands dependence measure. *Extremes*, 14(3):311–324.

Mainik, G. and Schaanning, E. (2014). On dependence consistency of covarand some other systemic risk measures. *Statistics & Risk Modeling*, 31(1):49–77.

Martínez-Jaramillo, S., Pérez, O. P., Embriz, F. A., and Dey, F. L. G. (2010). Systemic risk, financial contagion and financial fragility. *Journal of Economic Dynamics and Control*, 34(11):2358–2374.

Najjari, V. (2018). On generators in archimedean copulas. *Sahand Communications in Mathematical Analysis*, 12(1):155–166.

Nelsen, R. B. (2003). Properties and applications of copulas: A brief survey. In *Proceedings of the first brazilian conference on statistical modeling in insurance and finance*, pages 10–28. Citeseer.

Nelsen, R. B. (2007). *An introduction to copulas*. Springer Science & Business Media.

Nils, M. et al. (2018). The contribution of large banking institutions to systemic risk: What do we know? a literature review. *Review of Economics*, 69(3):231–257.

Oort, C. J. (1990). Banks and the stability of the international financial system. *De Economist*, 138(4):451–463.

Pfaff, B., McNeil, A., and Pfaff, M. B. (2018). Package 'evir'. *R News*. https://CRAN.R-project.org/package=evir.

Phillips, P. (1985). A theorem on the tail behaviour of probability distributions with an application to the stable family. *Canadian Journal of Economics*, pages 58–65.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131.

Razzaghi, M. and Kodell, R. L. (2000). Risk assessment for quantitative responses using a mixture model. *Biometrics*, 56(2):519–527.

Reynolds, D. A. (2009). Gaussian mixture models. *Encyclopedia of biometrics*, 741.

Risk, M. S. (2009). Global financial stability report. *International Monetary Fund, Washington*.

Risks, C. S. and Soundness, R. F. (2008). Global financial stability report. *International Monetary Fund, Washington*.

Rockafellar, R. T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443–1471.

Salmon, M. and Bouyé, E. (2008). Dynamic copula quantile regressions and tail area dynamic dependence in forex markets. *Available at SSRN 1129855*.

Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., Erhardt, T., Almeida, C., Min, A., Czado, C., Hofmann, M., et al. (2015). Package 'vinecopula'. *R News*, 2(5). "https://CRAN.R-project.org".

Schuster, E. F. and Gregory, G. G. (1981). On the nonconsistency of maximum likelihood nonparametric density estimators. In *Computer Science and Statistics: Proceedings of the 13th Symposium on the interface*, pages 295–298. Springer.

Schwaab, B. (2010). New quantitative measures of systemic risk. *Financial Stability Review Special Feature E, European Central Bank*.

Schweizer, B. (1991). Thirty years of copulas. In *Advances in probability distributions with given marginals*, pages 13–50. Springer.

Scott, A. J. and Symons, M. J. (1971). Clustering methods based on likelihood ratio criteria. *Biometrics*, pages 387–397.

Sedgwick, P. (2012). Pearson's correlation coefficient. *Bmj*, 345:e4483.

Segoviano Basurto, M. and Goodhart, C. (2009). Banking stability measures. *IMF working papers*, pages 1–54.

Sklar, A. (1973). Random variables, joint distribution functions, and copulas. *Kybernetika*, 9(6):449–460.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 8:229–231.

Smaga, P. (2014). The concept of systemic risk. *Systemic Risk Centre Special Paper*, 5.

Stephenson, A. G. (2018). Package evd: Extreme value distributions. *R News*.

Straumann, D. Y. (2001). Correlation and dependency in risk management: properties and pitfalls. In *Risk Management: Value at Risk and Beyond*. Cambridge University Press.

Teicher, H. et al. (1960). On the mixture of distributions. *The Annals of Mathematical Statistics*, 31(1):55–73.

Terzić, I. et al. (2014). Empirical estimation and comparison of normal and student t linear var on the belgrade stock exchange. *Sinteza 2014-Impact of the Internet on Business Activities in Serbia and Worldwide*, pages 298–302.

Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science*, 44(12-part-1):1650–1661.

Tippmann, S. (2015). Programming tools: Adventures with r. *Nature News*, 517(7532):109.

Trivedi, P. K. and Zimmer, D. M. (2007). *Copula modeling: an introduction for practitioners*. Now Publishers Inc.

Wong, C. S. and Li, W. K. (2000). On a mixture autoregressive model. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(1):95–115.

Yan, J. et al. (2007). Enjoy the joy of copulas: with a package copula. *Journal of Statistical Software*, 21(4):1–21.

Young, D., Benaglia, T., Chauveau, D., Hunter, D., Elmore, R., Hettmansperger, T., Thomas, H., Xuan, F., and Young, M. D. (2020). Package 'mixtools'.

Young, K. (2011). Basel committee on banking supervision. *Handbook of Transnational Governance: Institutions and Innovations (Cambridge: Polity, 2011)*, pages 39–44.

Zhang, Z., Li, W. K., and Yuen, K. C. (2006). On a mixture garch time-series model. *Journal of Time Series Analysis*, 27(4):577–597.

Zivot, E. and Wang, J. (2003). Rolling analysis of time series. In *Modeling Financial Time Series with S-Plus®*, pages 299–346. Springer.

Appendix A

Resumo em Português

A.1 Introdução

O risco sistémico pode ser definido como o risco subjacente a um sistema complexo composto por um conjunto de agentes que interagem entre si. Esta interacção aumenta a disseminação do risco (Smaga, 2014). Deste ponto de vista é de particular interesse avaliar e modelar o impacto de uma falha ocorrida num determinado agente sobre os restantes agentes como um todo. Uma falha sistémica pode começar como uma falha em alguns agentes, ser posteriormente amplificada, através dos mecanismos de interacção estabelecidos no sistema, o que pode conduzir a falhas sucessivas que podem vir a causar um impacto significativo em todo o sistema. Para quantificar o risco sistémico num determinado momento, ou a ocorrência de um determinado evento com impacto sistémico, deve-se considerar a modelo a utilizar, os dados disponíveis e as condições para a recolha dos mesmos (Eijffinger, 2011). Modelar as interacções ocorridas entre os agentes em determinado sistema permitem responder a três questões fundamentais num modelo de risco sistémico: (Martínez-Jaramillo et al., 2010):

- Identificar os canais de contagio;
- Quantificar a importância do impacto sistémico;
- Calcular a a probabilidade de falha.

Conditional Value at Risk (*CoVaR*) e Δ *CoVaR* são medidas de risco sistémico definidas por Adrian and Brunnermeier (2016), baseadas na medida de risco *Value at Risk* (*VaR*) . O *VaR* mede o risco com base no máximo da expectativa do risco para um dado intervalo de confiança $VaR(\alpha)$. *CoVaR* é o *VaR* condicionado pela ocorrência de um determinado evento (tipicamente uma crise), que afecta o agente i em determinado momento t, e em que o retorno do agente i é r_{it}, definido por $C(r_{it})$.

A aplicação de copulas na área do riso e na área financeira conta com extensa literatura e trabalho de investigação. Também no caso do risco sistémico a escolha pode recair na teoria de copulas de modo a incluir assimetrias e dependência nos extremos (Joe and Kurowicka, 2011).

A medida de risco sistémico *CoVaR* pode também ser determinada através de probabilidade condicionada, com base em duas variáveis aleatórias, uma que representa o retorno da instituição financeira i e outra que presenta o retorno de todo o sistema financeiro:

- Y representa o retorno do sistema financeiro.
- X_i representa o retorno da instituição financeira i

Considerando as variáveis aleatórias aleatórias X e Y, *CoVaR* é definida como o *VaR* de Y condicionado por X, ou seja devido X estar sujeito a um evento que causou impacto nos seus retornos e o seu *VaR* ter sido atingido. Mais formalmente, a definição pode ser representada da seguinte forma:

$$CoVaR_{\alpha\beta t}(Y|X_i) = VaR_\beta(Y|X_i \in E) \tag{A.1}$$

onde E representa o conjunto de eventos críticos que afectam a instituição financeira i, considerados para este efeito como uma crise. Decorre da definição de *CoVaR* que existe uma relação de dependência entre as variáveis X e Y.

A.2 $\Delta CoVaR$ Bases Teóricas

A.2.1 Teoria das Copulas

Uma função de copula é uma função distribuição multivariada, definida no cubo unitario de dimensão d $[0,1]^d$ para o intervalo unitário [0,1], as C : [0, 1]d : $\rightarrow$ [0, 1] é uma função de distribuição comulativa (FDC) com as marginais que seguem uma distribuição uniforme e que satisfaz as seguintes propriedades:

- C(1,...,1, u_i, 1, ..., 1) = u_i . A i^{th} distribuição marginal é obtida fazendo u_j = 1 para $j \neq i$ e desde que seja uniformemente distribuída. Esta propriedade estabelece que, se o valor de *d-1* variáveis é conhecido, com probabilidade 1, então as d possibilidades da probabilidade conjunta é igual á da variável desconhecida (u_i).

- C(u_1, ..., u_d) = 0 se u_i = 0 $\forall i \leq$ d. Esta propriedade estabelece que se uma variável tem probabilidade marginal 0, então todos os resultados possíveis para a probabilidade conjunta são também 0.

- C(u_1, ..., u_d) é não-decrescente em cada componente u_i. Esta propriedade assegura que a probabilidade conjunta nunca será negativa, uma vez que o volume (C) para qualquer intervalo d-dimensional é também não negativo.

- C is limitada pelos limites de Fréchet:

$$max\left\{\sum_{i=1}^{d} u_i + 1 - d, 0\right\} \leq C(u) \leq min\{u_1, ..., u_d\} \tag{A.2}$$

  onde o limite superior é o limite superior de Frechet – Hoeffding, e o limite inferior é limite inferior de Frechet – Hoeffding.

Theorem A.2.1 (Sklar)**.** *Seja F a função distribuição conjunta com margens $F_1, ..., F_n$. Existe uma copula C tal que para todo o $x_1, ..., x_d$ in $[-\infty, \infty]$ e $i = 1, ..., d$ verifica*

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_1(x_d)) \tag{A.3}$$

Se F_i é continua para todo o $i = 1, ..., d$, então C é única, caso contrario C é unicamente definida em $Ran(F_1) \times ... \times Ran(F_d)$ onde $Ran(F_i)$ representa o intervalo da FDC, F_i. Caso contrário, considerando a copula, C, e uma função FDC univariada, $F_1, ..., F_d$. Então F é uma função FDC multivariada com marginais $F_1, ..., F_d$.

Se as distribuições marginais $F_1, ..., F_n$, são continuas é possível demonstrar que:

$$F_i(F_i^{-1}(y)) = y \tag{A.4}$$

fazendo $x_i = F_i(u_i)$ e usando o resultado anterior, obtemos:

$$C(u) = C(F_1^{-1}(u_1), ..., F_d^{-1}(u_d)) \tag{A.5}$$

Este teorema permite concluir que copulas são distribuições conjuntas e também que estas distribuições podem ser representadas através de copulas e das suas função de distribuição marginais. Deste modo, o exercício de modulação de distribuições conjuntas pode ser reduzido a um exercício de modulação de copulas (Schweizer, 1991).

Outro resultado do teorema de Sklar é que as cópulas também representam a dependência entre as variáveis que vem como resultado da divisão da distribuição conjunta de uma cópula e nas respectivas marginais. Devido a este resultado as cópulas são também designadas de funções de dependência (Hürlimann, 2003).

Como as cópulas também são medidas de dependência, permitem distinguir a dependência perfeita e também a independência.

A.2.2 Distribuição de Abas Largas

Em estatística, o termo cauda pesada está associado a distribuições com uma probabilidade relativamente elevada de resultados extremos.

Mesmo que não haja uma definição definitiva e formal de distribuição de abas largas, geralmente é assumido que uma distribuição tem abas largas quando a probabilidade na cauda é mais densa quando comparada com uma distribuição normal.

A.2.3 Modelação de Retornos Financeiros com Distribuições de Mistura

Como a evidência empírica e os resultados sugerem, a suposição de normalidade dos retornos das instituições financeiras não é verificada por possuirem abas largas. Modelar esse comportamento utilizando apenas uma distribuição demonstrou limitações, e uma opção será considerar uma abordagem que usa mais de uma distribuição.

Distribuições de abas largas podem ser modeladas por uma mistura de distribuições. Nesse caso, primeiro abordaremos o problema aplicando uma mistura de distribuições normais.

Assumindo que os retornos seguem um processo estocástico para a instituição financeira i tal que:

$$R_{it} = \lambda_{it} R_{it}^{\alpha} + (1 - \lambda_{it}) R_{it}^{\beta} \tag{A.6}$$

onde $R_{it}^{\alpha} \sim N(\mu_\alpha, \sigma_\alpha)$, $R_{it}^{\beta} \sim N(0, \sigma_\beta)$, e λ_{it} é 1 com probabilidade p e 0 caso contrário.

Estas três variáveis aleatórias R_{it}^{α}, R_{it}^{β} e λ_{it} são independentes entre si.

Dependendo de λ e com probabilidade p a distribuição a aplicar será $N(\mu_\alpha, \sigma_\alpha)$, para as situações mais expectáveis. Com probabilidade (1 - p), λ será igual a 0 e a distribuição a aplicar será $N(0, \sigma_\beta)$, o que pode ser interpretado como os casos de excepção.

O desafio está agora em estimar os parâmetros envolvidos; $p, \mu_\alpha, \sigma_\alpha, \sigma_\beta$.

Apesar de vários métodos alternativos que são possíveis de usar para estimar os parâmetros de uma mistura de distribuições normal, se considerarmos o método de máxima verosimilhança tradicional, ele poderia então ser formulado do seguinte modo:

$$l\Big((p, \mu_\alpha, \sigma_\alpha, \sigma_\beta)|R_{it}\Big) = \sum_t \log\left[\frac{p}{\sigma_\alpha}\exp\left(-\frac{\big(R_t - \mu_\alpha\big)^2}{2\mu_\alpha^2}\right) + \frac{1-p}{\sigma_\beta}\exp\left(-\frac{(R_t^2)}{-\mu_\beta}\right)\right]$$

Devido à existência de pólos e pontos de sela, a maximização da mistura de probabilidade de normais pode não constituir um desafio trivial e o máximo global para essa função poderia mesmo não existir (Hamilton, 1991).

Este problema, entretanto, pode ser descrito como um problema de dados incompletos, uma vez que os dados que observamos na nossa amostra podem ser vistos como um subconjunto dos dados "completos".

A.2.4 Algoritmo de Maximização de Expectativa

O algoritmo *Expectation-Maximization* (EM) é uma abordagem para estimativa de máxima verosimilhança na presença de variáveis latentes e pode ser usado para estimar os valores das variáveis latentes, com a condição de que a forma geral da distribuição de probabilidade que governa essas variáveis latentes seja conhecida.

O algoritmo é implementado como um procedimento iterativo. Dado um conjunto de dados incompletos e considerando um conjunto de parâmetros iniciais o processo irá iterar em duas etapas:

- Passo *Expectation* (E – step). As variáveis ausentes ou latentes são estimadas com base nos dados observados disponíveis e nos parâmetros do modelo actual,
- Passo *Maximization* (M – step). Depois de estimar os valores ausentes, esta etapa será usada para actualizar os parâmetros, calculando os parâmetros que maximizam a probabilidade logarítmica esperada do modelo com base nos valores estimados no passo E.

O algoritmo EM inclui considerações estatísticas para calcular a distribuição da fonte de máxima probabilidade (ML) que teria criado os dados observados, incluindo os efeitos das estatísticas de densidade. Especificamente, o algoritmo atribui peso maior aos elementos de alta densidade de um perfil e menos peso às regiões de baixa densidade (Dempster et al., 1977).

A.2.5 Algoritmo EM e Mistura de Distribuições Normais

Considerando o caso de uma mistura de distribuições normais, seja $\mathbf{x} = (\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_n)$ ser uma amostra de n i.i.d. observações de uma mistura de duas normais e $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ as variáveis latentes que determinam o componente de onde se origina a observação (Reynolds, 2009).

$$X_i|(Z_i = 1) \sim \mathcal{N}_d(\boldsymbol{\mu}_1, \Sigma_1) \quad \text{and} \quad X_i|(Z_i = 2) \sim \mathcal{N}_d(\boldsymbol{\mu}_2, \Sigma_2),$$

where

$$\mathrm{P}(Z_i = 1) = \tau_1 \quad \text{and} \quad \mathrm{P}(Z_i = 2) = \tau_2 = 1 - \tau_1$$

O objectivo deste processo é estimar os parâmetros para a mistura de distribuições normais:

$$\theta = \Big(\boldsymbol{\tau_1}, \boldsymbol{\tau_2}, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \Sigma_1, \Sigma_2\Big)$$

A função de verosimilhança, portanto, é:

$$L(\theta; \mathbf{x}, \mathbf{z}) =$$
$$\exp\left\{\sum_{i=1}^{n}\sum_{j=1}^{2} \mathbb{I}(z_i = j)\Big[\log \tau_j - \tfrac{1}{2}\log|\Sigma_j| - \tfrac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_j)^\top \Sigma_j^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_j) - \tfrac{d}{2}\log(2\pi)\Big]\right\}.$$

No entanto, embora os estimadores de máxima verosimilhança forneçam bons resultados para distribuições de cauda curta, isso não é verdade para distribuição de cauda pesada ou mesmo com a presença de outliers conforme demonstrado por (Schuster and Gregory, 1981), levando a estimativas inconsistentes.

A fim de mitigar este inconveniente no método de máxima verosimilhança, foi proposta como alternativa a aplicação dos métodos de Monte Carlo através de cadeias de Markov (Levine and Casella, 2001).

A.2.6 Modelos de Mistura para Valores Extremos

A ideia por trás do uso da distribuição de Mistura para Valores Extremos é combinar a flexibilidade de usar uma distribuição para capturar o componente principal, também identificado como o bloco central da distribuição (Fúquene Patiño, 2015), o que poderia ser, por exemplo, uma distribuição Normal, e também as abas, como valores extremos. Com este modelo de mistura, obter-se-á uma função de distribuição completa, dividindo a distribuição num componente a central e nos componentes das abas.

Existem várias abordagens que consideram apenas uma aba e também abordagens que consideram ambas, a aba superior e inferior. Nesse caso, a função de mistura será potencialmente composta por uma mistura de distribuição de famílias distintas.

Em nosso caso, estamos especialmente interessados em explorar uma mistura de uma distribuição Normal como uma distribuição do componente principal, com duas distribuições Gama em ambas as abas.

MacDonald et al. (2011) propôs um modelo de mistura de duas abas onde o estimador do parâmetro de limite entre as diferentes distribuições é dividido pelos dois modelos, uma para cada uma das abas de valor extremo.

Este modelo usa um estimador de densidade do núcleo para estimar a distribuição de valores não extremos e a Distribuição de Pareto Generalizada (DPG) para estimar a distribuição nos extremos. Um estimador de densidade do núcleo com correcção de limite também é usado. Este estimador de densidade do núcleo assume um núcleo particular, neste caso a densidade normal, que é centralizada em cada ponto dos dados e usa apenas um parâmetro para definir a amplitude da banda.

O componente da aba refere-se à proporção da distribuição acima do limite. Este parâmetro será identificado por Φ_u onde u representa esse limite.

The distribution function comes as:

$$F(x|\Theta) = \begin{cases} \phi_{u_l} 1 - G(-x| - u_l, \sigma_{u_l}, \epsilon_l), & x < u_l \\ \\ H(x|\mu, \sigma), & u_l \leq x \leq u_r \\ \\ (1 - \phi_{u_r}) + \phi_{u_l} G(x|u_r, \sigma_{u_r}, \epsilon_r), & x > u_r \end{cases}$$

onde $\phi_{u_l} = H(u_l|\mu, \sigma)$ e $\phi_{u_r} = 1 - H(u_r|\mu, \sigma)$ e $H(.|\mu, \sigma)$ é uma distribuição normal com média μ e desvio padrão σ. $G(.| - u_l, \sigma_{u_l}, \epsilon_l)$ e $G(.| - u_r, \sigma_{u_r}, \epsilon_r)$ são distribuições GPD para caudas inferior e superior, respectivamente.

A.3 Dados Utilizados

A abordagem proposta depende apenas de dados de mercado disponíveis publicamente, como a cotação das acções, pois assume-se que estas refletem correctamente todas as informações sobre empresa cotada.

Com base na lista dos bancos que integram o índice *STOXX Europe 600*, correspondente aos maiores e mais importantes bancos da Europa foram recolhidas as cotações diárias de cada título, tal como disponíveis em *https://finance.yahoo.com/*.

O conjunto de dados original utilizado foi composto pelas cotações diárias dos 50 maiores bancos da Europa nos últimos 20 anos, de Janeiro de 1998 a Junho de 2018. Posteriormente, também foram recolhidas informações sobre a quantidade de acções emitidas de cada uma das instituições. Esses dados, juntamente com outras informações, como taxas de câmbio diárias, disponíveis em *https://www.bankofengland.co.uk/statistics/exchange-rates* foram também recolhidos com o objectivo de construir uma série diária de capitalização de cada instituição (em euros).

Posteriormente, esse conjunto de dados foi reduzido para incluir apenas os preços de fecho semanais de cada instituição financeira. A partir dessa série, obteve-se uma série correspondente aos retornos semanais (em percentagem) de cada uma das instituições e para todo o sistema.

A janela móvel sobreposta foi aplicada ao longo de períodos de três anos cada. Para cada um desses períodos foi estimado o VaR, $CoVaR$ e $\Delta CoVaR$, formando uma nova série temporal relacionada a cada instituição, representando a posição de risco de cada instituição em cada momento (semana).

A recolha de dados para suportar um trabalho de investigação é normalmente um processo árduo. Os detalhes financeiros das instituição financeiras muito frequentemente não estão disponíveis para o público em geral e são informados apenas numa base periódica. Esta metodologia propõe obter VaR e $CoVaR$, com base em dados públicos de mercado, o que traz clareza adicional ao processo, bem permite aceder e calcular essas medidas de risco em qualquer momento, independentemente de quando os dados de cada uma das instituições financeiras são disponibilizados ou publicados. Como pressuposto, para utilizar dados publicamente disponíveis, assume-se que o valor de mercado, a capitalização bolsista de cada instituição reflecte o valor contabilístico dos activos. Além disso, o valor de mercado do sistema é assumido como a agregação do valor de mercado de todas as instituições pertencentes a esse sistema.

A.4 Metodologia para o Cálculo do *CoVaR*

A.4.1 Descrição da Metodologia e Medidas de Risco

$CoVaR$ é definido como o VaR do sistema financeiro, dado que uma das Instituições está em stress (Adrian and Brunnermeier, 2016).

Value at Risk, ou VaR é uma ferramenta central em risco, risco de activos e carteiras e também desempenha um papel fundamental no risco sistémico. *VaR* é definido como a perda máxima que um activo/carteira/instituição pode incorrer num quantil definido por α (Basilio et al., 2020), e representa o valor em risco potencial de perda em um período de tempo específico e probabilidade específica α.

Considerando uma variável aleatória X, o retorno financeiro em um determinado momento t e uma função de distribuição cumulativa F que modelam os retornos financeiros. Seja x F uma função continua e X uma variável aleatória contínua, com $0 \leq \alpha \leq 1$, então VaR_α é definido para um momento no tempo t e para um intervalo de confiança $1 - \alpha$ como:

$$P(X_t \leq VaR_\alpha) = \alpha \tag{A.7}$$

$$VaR_\alpha = F^{-1}(1 - \alpha) \tag{A.8}$$

Δ *CoVaR* é uma medida de risco sistémico definida por Adrian and Brunnermeier (2016), baseado no *VaR*. *VaR* mede o risco com base na perda máxima esperada para um determinado intervalo de confiança $VaR(\alpha)$. *CoVaR* corresponde ao *VaR condicionado* à ocorrência de um evento específico (normalmente uma crise), que afecta o agente i num momento t e o retorno do agente i é r_{it}, definido por $C(r_{it})$. Existem várias opções para definir um evento crítico $C(r_{it})$. Uma é considerar o *VaR*. Deste modo, *CoVaR* e definido para um agente i, como $CoVaR^i_\alpha$ by:

$$CoVaR^i_\alpha = P(r^{it} \leq VaR^{it}_\alpha) = \alpha \tag{A.9}$$

onde r^{it} representa o retorno do agente j num momento t.

O $CoVaR^{j|i}_\alpha$ e o *VaR* do agente j condicionado pelo facto do agente i estar em stress, ou seja ter atingido o seu *VaR* .

$$CoVaR^{j|i}_\alpha = P(r_{jt} \leq CoVaR^{j|i}_\alpha | r_{it} = VaR^i_\alpha) = \alpha. \tag{A.10}$$

O conceito de Δ $CoVaR$ representa a diferença entre o VaR de todo o sistema no caso em que um evento crítico afecta o agente i e o VaR de todo o sistema quando esse evento não ocorre.

A.4.2 *CoVaR* com Base em Cópulas

O risco condicionado, ($CoVaR$) também pode ser estimado através de funções de cópula. Aplicando a propriedade intrínseca das funções de cópula que permitem o isolamento da dependência das funções de distribuição marginal da cópula, a abordagem da cópula para o cálculo do $CoVaR$ fornece flexibilidade na especificação das estruturas marginais e de dependência.

Seja (X, Y) um par de vectores aleatórios, a distribuição conjunta é dada por $F_{XY}(x, y) = P(X \leq x, X \leq y)$, onde F_{XY} representa a função de distribuição cumulativa bi-variada e F_X, F_Y representa a distribuição marginal, então pelo teorema de Skar's(Sklar, 1973) existe uma função de distribuição cumulativa de cópula bidimensional $C \in [0, 1]^2$ que verifica $F_{XY}(x, y) = C(F_X(x), F_Y(y))$. Se F_X e F_Y são continuas, então $C(u, v) = F_{XY}(F_X^{-1}(u), F_Y(v)^{-1})$

A distribuição de probabilidade condicional pode ser expressa usando uma função de cópula (Salmon and Bouyé, 2008):

$$P(Y \leq y | X = x) = \frac{\partial C(u, v)}{\partial v} \tag{A.11}$$

e

$$P(Y \leq y | X \leq x) = \frac{C(u, v)}{v} \tag{A.12}$$

A copula Archimediana é definida por:

$$C(u, v) = \varphi^1[\varphi(u), \varphi(v)] \tag{A.13}$$

onde a função φ é a função geradora da cópula C.

Representando $CoVaR_{\alpha,\beta,t}$ como uma copula Archimediana, temos:

$$P(Y \leq y|X \leq x) = \frac{\partial C(u,v)}{\partial v} = \frac{\varphi'(v)}{\varphi'(C(u,v))} = \frac{\varphi'(v)}{\varphi'((\varphi^{-1}[\varphi(u)+\varphi(v)])} \tag{A.14}$$

Se as variáveis aleatórias Y representam os retornos do sistema, $R_{s,t}$ e X representa os retornos da instituição i, $R_{i,t}$ com distribuições $F_{s,t}$ e $F_{i,t}$, a distribuição condicional do *CoVaR* pode ser escrito como:

$$P(R_{s,t} \leq CoVaR_{\alpha,\beta,t}|R_{i,t} = VaR^{j}_{\alpha,t}) = \frac{\varphi'(v)}{\varphi'(\varphi^{-1}[\varphi(u)+\varphi(v)])} = \beta \tag{A.15}$$

Assumindo que $\frac{\partial C(u,v)}{\partial v}$ é parcialmente derivável em ordem a u, o quantil condicional da cópula é dado por:

$$u = \varphi^{-1}\left(\varphi\left(\varphi'^{-1}\left(\frac{1}{\beta}\varphi'(v)\right)\right) - \varphi(v)\right) \tag{A.16}$$

Usando a transformação integral de probabilidade, a expressão de $CoVaR_{\alpha,\beta,t}$ vem:

$$CoVaR_{\alpha,\beta,t} = F^{-1}_{s,t}\left(\varphi'^{-1}\left(\varphi\left(\varphi'^{-1}\left(\frac{1}{\beta}\varphi'\left(F_{i,t}(VaR^{i}_{\alpha,t})\right)\right)\right) - \varphi\left(F_{i,t}(VaR^{i}_{\alpha,t})\right)\right)\right) \tag{A.17}$$

Com base na defnição de *VaR*, $F_{i,t}(VaR^{i}_{\alpha,t}) = \alpha$, A.17 pode ser simplificada:

$$CoVaR_{\alpha,\beta,t} = F^{-1}_{s,t}\left(\varphi'^{-1}\left(\varphi\left(\varphi'^{-1}\left(\frac{1}{\beta}\varphi'(\alpha)\right)\right) - \varphi(\alpha)\right)\right) \tag{A.18}$$

A expressão geral de *CoVaR* representada em A.18 requer que $F^{-1}_{s,t}$, a função de distribuição cumulativa inversa dos retornos do sistema esteja definida. Além do ajuste da cópula a ambas as séries dos retorno financeiros, é necessário ajustar a série dos retornos a uma função que represente a distribuição cumulativa.

A.4.3 Contribuição para o Risco Sistémico

CoVaR como medida de risco condicional pode ser usada para estimar a contribuição individual de cada instituição financeira para o risco sistémico.

Definamos essa contribuição marginal para o risco como uma diferença entre duas estimativas para *CoVaR*, tomadas com suposições distintas:

- quando instituição financeira i está sob stress financeiro,
- quando instituição financeira i está numa situação *"normal"*.

A contribuição marginal da instituição financeira i é então interpretada como a diferença do sistema *CoVaR* nas duas situações mencionadas.

Para determinado momento t esta contribuição marginal e representada por:

$$\Delta CoVaR_{\alpha\beta t} = CoVaR_{\alpha\beta t} - CoVaR_{0.5\beta t} \tag{A.19}$$

Com base no procedimento descrito por Adrian and Brunnermeier (2016) será necessário estimar o *CoVaR* para as duas situações mencionadas. Quando enfrentamos uma situação de stress financeiro que afecta os retornos da instituição financeira i no momento t e onde o *CoVaR* corresponde a:

$$CoVaR_{\alpha\beta t} \tag{A.20}$$

e para uma situação que definimos como normal, sem qualquer situação de stress que afecta os retornos da instituição financeira i para o período t

$$CoVaR_{0.5\beta t} \tag{A.21}$$

que corresponde ao valor estimado de *VaR* para o sistema financeiro condicionado por o *VaR* da instituição financeira na mediana, o que significa que a instituição financeira não está sob stress. Deste modo$\Delta CoVaR_{\alpha\beta t}$ representa a diferença entre 0 *VaR* estimado para o sistema financeiro quando:

- $Retorno_{it} < VaR_{\alpha t}$
- $Retorno_{it} < VaR_{0.5t}$

Para calcular a estimativa para $CoVaR$ em ambas as situações descritas, devemos ter em consideração a estrutura de dependência e correlação que existe em cada uma dessas situações.

A literatura tem sido profícua em apresentar casos e demonstrar que, para dados financeiros, a estrutura de dependência e correlação na cauda da distribuição é diferente da estrutura de dependência e correlação que ocorre no resto da distribuição. De forma simples, por um momento de stress ou crise, a estrutura de dependência entre a distribuição dos retornos das instituições financeiras do sistema financeiro muda, e consequentemente com movimentos mais agressivos usualmente sendo observados.

A abordagem proposta para o cálculo da estimativa da contribuição individual de uma instituição financeira para o risco sistémico implica o cálculo do $CoVaR$ em duas áreas muito distintas da distribuição dos retornos, do ponto de vista dos quantis utilizados. Um componente envolve o cálculo da probabilidade no extremo, a aba esquerda no caso específico do risco, da distribuição, enquanto o outro componente envolve o cálculo da probabilidade na mediana da distribuição dos retornos. A modelação da estrutura de dependência nos extremos pode ser obtida aplicando a teoria dos valores extremos (Fougères et al., 2009).

Neste contexto, deparamo-nos com a modulação conjunta de duas distribuições de retornos, tipicamente de uma instituição financeira e do sistema financeiro, e a cópula de Gumbel com cauda pesada mostra-se uma escolha adequada (Basilio and Oliveira, 2020).

Por outro lado, temos a estrutura de dependência e correlação na mediana, que também deve levar em consideração as abas largas, mas neste caso poderá ser melhor aproximada por uma cópula t. Consequentemente, o $\Delta CoVaR$ resulta da diferença entre duas cópulas com estruturas de correlações distintas e, eventualmente, funções de ligação de cópulas diferentes.

A.4.4 Metodologia Proposta

A nova metodologia proposta compreende as seguintes passos:

1. Estimar o valor dos activos. Um dos objectivos deste trabalho de investigação é também usar dados públicos disponíveis para identificar instituições de risco sistémico. A primeira etapa da metodologia consiste em recolher os preços das acções e calcular, para cada instituição financeira, a capitalização de mercado. A avaliação do sistema é obtida por uma simples agregação da capitalização bolsista de todas as instituições financeiras incluídas no sistema. Com base no valor de mercado em cada momento, calcularemos a série de retornos para a instituição financeira

2. Definir um horizonte temporal. O horizonte de tempo deve ser longo o suficiente para permitir a ocorrência de diferentes eventos críticos ou crises relevantes.

3. Calcular o $CoVaR$ no extremo da distribuição dos retornos. Definir um número mínimo k do retorno da instituição financeira (perda máxima verificada), verificada dentro da janela de tempo considerada. Uma função de cópula será ajustada às duas séries:

    a) Ajustar a cópula às duas séries nos extremos e selecionar a cópula que evidencia uma maior dependência na aba.

    b) A série de retornos do sistema nos mesmos períodos em que foram verificados os retornos mínimos da instituição financeira, de forma a preservar a distribuição dos concomitantes dos extremos induzidos a partir da estatística de ordem, de acordo com os retornos da instituição financeira.

c) Após ajustar a cópula às duas séries de retornos, obtidas na passo anterior, a derivada parcial da função de cópula será estimada para a marginal relativa à distribuição dos retornos do sistema financeiro, modelada a partir da série de retornos usando um modelo de mistura de distribuições.

$$\frac{\partial C(u,v)}{\partial v} = CoVaR \tag{A.22}$$

Assumindo que a derivada parcial é invertível em relação a v, temos os seguinte resultado:

$$CoVaR^{i}_{\alpha\beta t} = F^{-1}_{s,t}(g^{-1}(\beta, F^{-1}_{i,t}(VaR^{i}_{\alpha,t}))) = F^{-1}_{s,t}(g^{-1}(\beta,\alpha)) \tag{A.23}$$

d) Para obter o *CoVaR* na mediana dos retornos da instituição financeira, levaremos em consideração todos os valores dos retornos na janela de tempo. Ambas as séries serão ajustadas a uma cópula. Após ajustar a cópula às duas séries de retornos, a derivada parcial da função de cópula será estimada para a marginal correspondente à distribuição dos retornos da instituição financeira, modelada a partir das séries de retornos usando um modelo de mistura de distribuições.

$$\frac{\partial C(u,v)}{\partial v} = CoVaR \tag{A.24}$$

Na mediana, tomaremos uma cópula t. Neste caso *CoVaR* é obtido por:

$$CoVaR^{\beta}_{\alpha} = F^{-1}_{s}\left(t_v\left(\rho t^{-1}_v(\beta) + \sqrt{\frac{(1-\rho^2)(v + t^{-1}_v(\beta)^2)}{v+1}}\, t^{-1}_{v+1}(\alpha)\right)\right) \tag{A.25}$$

e) Determinar Δ*CoVaR* através da diferença simples entre o *CoVaR* o extremo e o *CoVaR* na mediana.

4. Após obter $\Delta CoVaR$ para instituição financeira é possivel é possível construir um ranking baseado no $\Delta CoVaR$ onde os valores mais elevados do $\Delta CoVaR$ correspondem a uma contribuição mais elevada para o risco sistémico.

A nova metodologia proposta para calcular $\Delta CoVaR$ usará duas abordagens diferentes com o propósito de seleccionar a função de cópula no interesse de alinhar também com duas situações distintas. Quando a instituição financeira não está em stress corresponde a um parâmetro β de 0,5, na mediana da distribuição dos retornos, e uma abordagem diferente para seleccionar uma cópula para melhor descrever o comportamento na aba e, em particular, a dependência na aba entre os retornos da instituição financeira e os retornos do sistema financeiro.

A.5 Resultados Obtidos

A.5.1 Detalhes da Implementação

Como o objectivo principal deste trabalho de pesquisa é detalhar uma metodologia para identificar instituições financeiras de risco sistémico, ilustraremos o processo para calcular o $\Delta CoVaR$ em detalhe, primeiro tomando uma instituição financeira como exemplo e depois calculando a medida $\Delta CoVaR$ para todas as instituições financeiras.

A.5.1.1 Seleccionar a Cópula

As funcção de Cópulas podem ser seleccionadas de acordo com os critérios de informação de Akaike e de Bayes (AIC e BIC, respectivamente) e também podemos realizar um teste estatístico para avaliar o desempenho dessa função num teste de adequação com base na igualdade da matriz de informação de White.

O método de selecção de cópulas consiste em calcular os critérios para todas as escolhas de cópulas possíveis e de seguida é escolhida a família de cópulas que evidencia o valor mínimo para o crirério de informação aplicado.

Neste caso, estamos interessados em explorar uma relação positiva entre o retorno de uma instituição financeira e o retorno do sistema financeiro.

O processo para seleccionar a função de cópula e a família de cópula que melhor reflicta a dependência entre as duas séries de retorno não se pode limitar a capturar adequadamente comportamento em toda a distribuição da variável dependente. A dependência das caudas também deve ser levada em consideração, em especial no caso de modelaão do risco sistémico.

Quando buscamos o melhor ajuste com base nos critérios AIC/BIC, obtemos os seguintes resultados para cada par:

Instituição Financeira	Copula	AIC	τ	Dependência na Aba esquerda	p-Value
ABN	Gumbel	-106.59	0.52	0	0.94
ACA	t	-731.56	0.55	0.42	0.17
AIBG	t	-228.76	0.36	0.31	0.86
BAER	t	-297.12	0.41	0.29	0.02
BAMI	t	-538.16	0.46	0.23	0.87
BARC	t	-720.93	0.49	0.37	0.71
BBVA	t	-994.6	0.59	0.51	0.13
BCP	t	-313.39	0.36	0.22	0.11
BIRG	t	-329.7	0.38	0.27	0.24
BKESY	Frank	-13.32	0.13	0	0.5
BKIA	t	-191.5	0.49	0.39	0.67
BKT	t	-295.66	0.47	0.31	0.05
BNP	t	-967.0	0.57	0.47	0.02
BPI	t	-226.86	0.3	0.18	0.52
BZI	Gumbel	-25.3	0.21	0	0.29
CABK	t	-292.13	0.49	0.3	0.1
CBG	t	-251.23	0.31	0.03	0.78

Financial Institution	Copula Family	AIC	τ	Lower Tail Dependence	p-Value
CBK	t	-690.04	0.49	0.37	0.17
CYBG	BB7	-45.72	0.37	0.37	0.95
DBK	t	-854.57	0.53	0.44	0.57
DSN	t	-254.59	0.32	0.26	0.83
EBS	t	-399.84	0.4	0.32	0.22
FBK	Survival BB8	-104.9	0.45	0	0.28
GEH	t	-22.12	0.2	0.12	0.72
GLE	t	-1024.55	0.61	0.5	0.08
HSBA	t	-629.12	0.47	0.22	0.01
ING	t	-888.43	0.55	0.42	0.06
ISP	t	-752.5	0.51	0.32	0.02
JYS1	Survival BB1	-168.54	0.32	0.35	0.23
KBC	t	-700.01	0.51	0.43	0.44
KN	t	-578.17	0.47	0.41	0.18
LLOY	t	-544.65	0.44	0.19	0.03
MB	t	-628.62	0.48	0.3	0.03
MTRO	Gaussian	-33.41	0.34	0	0.8
NDA-DK	t	-428.18	0.49	0.33	0.67
P9O	Frank	-23.56	0.21	0	0.61
RBI	Survival BB1	-347	0.44	0.46	0.62
RBS	t	-618.03	0.46	0.28	0.19
SAB	t	-445.27	0.45	0.32	0.01
SAN	t	-941.38	0.58	0.52	0.07
STAN	t	-460.65	0.41	0.23	0.5
UBI	t	-511	0.51	0.37	0

Financial Institution	Copula Family	AIC	τ	Lower Tail Dependence	p-Value
UBS	t	-689.48	0.51	0.42	0.43
UCG	t	-1034.08	0.61	0.51	0

Table A.1: Qualidade o ajustamento da Cópula

Os resultados obtidos mostram uma predominância de t-Copula para modelar a estrutura de correlação entre os retornos de cada instituição financeira e os retornos do sistema, mas não há garantia de que a família de cópula com melhor ajuste para cada instituição financeira seja a mesma .

O ajuste obtido na maior parte dos casos aponta-nos para uma função de cópula com dependência na aba, como é o caso de a t-Copula, Gumbel, BB7, Survival BB1 e Survival BB8.

Antes de decidirmos por uma família de cópula, é importante analisar alguns dos resultados obtidos do ajuste dos dados bivariados a uma cópula e analisar como esse ajuste funciona nas abas.

Tomada como exemplo a instituição financeira $LLOY$, seleccionamos a cópula de melhor ajuste com base no critério AIC com os seguintes resultados:

Copula Family	t	
Parameter	rho:	0.64
	df:	8.47
Dependence measures	Kendall's tau:	0.44
	Upper TD:	0.18
	Lower TD:	0.18
Fit statistics	logLik:	273.23
	AIC:	-542.47
	BIC:	-532.54

Table A.2: Copula ajustada para LLOY

Este ajuste de cópula representa uma estrutura de dependência com abas simétricas como também podemos visualizar na figura abaixo onde podemos identificar as abas simétricas.

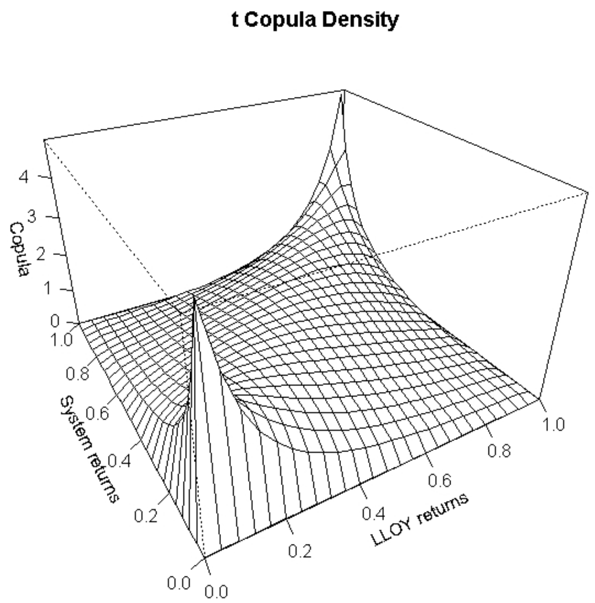


Figure A.1: t- Densidade da Copula para LLOY

Para visualizar a aproximação do ajuste aos dados originais, podemos mostrar os dados bivariados conjuntos num gráfico bidimensional.

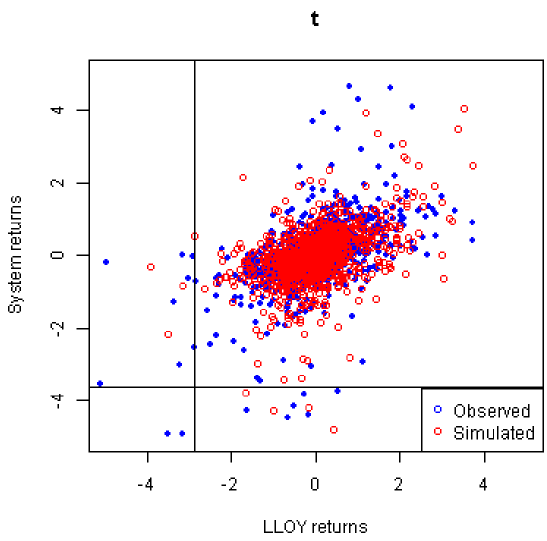


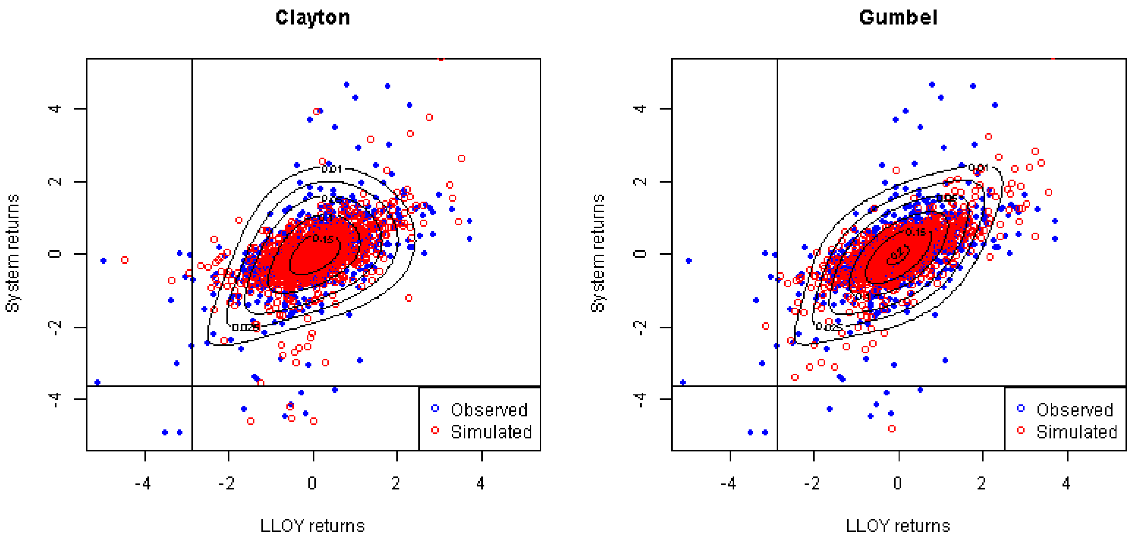
Figure A.2: t-Copula vs dados reais para LLOY

Podemos denotar aqui um bom ajuste na secção mais densa da distribuição o que se transforma num valor optimizado para o critério AIC, mas se colocarmos mais atenção nas caudas também podemos observar que os valores mais extremos não são seguidos com esta função de cópula, neste caso a t-Copula.

Como o conceito de dependência é descrito como medida de associação entre, neste caso duas variáveis aleatórias, considerando todo o seu alcance, essas medidas de dependência não reflectem adequadamente o comportamento nas abas e ao mesmo tempo a dependência no componente central dessas mesmas distribuições.

A dependência na aba, por outro lado, está relacionada à associação das variáveis nos extremos que medem essa associação nas abas da distribuição conjunta. Dessa forma, dependência e dependência na aba não coincidem, e duas variáveis aleatórias poderiam exibir dependência sem dependência na aba.

Os gráficos abaixo são divididos em quadrantes onde as linhas vertical e horizontal representam um quantil na cauda esquerda (1 %). Dessa forma, o quadrante inferior esquerdo é onde esperamos encontrar os movimentos extremos do risco sistémico representados.



a b

Figure A.3:
(a) Valores gerados para a Copula Clayton vs dados observados para LLOY
(b) Valores gerados para a Copula Gumbel vs dados observados para LLOY

Mesmo que todas as diferentes cópulas fossem ajustadas aos mesmos dados, os resultados na aba esquerda são diferentes. Uma opção é seleccionar uma cópula de uma família de cópula diferente, que poderia se ajustar melhor a valores extremos nas abas da distribuição.

Copula	τ	Dependência na Aba Esquerda	Dependência na Aba Direita
t	0.442	0.1814	0.1814
Clayton	0.3381	0.5073	0
Gumbel	0.4176	0	0.5027
tev	0.39	0	0.5066
Husler Reiss	0.3914	0	0.468
Galambos	0.4135	0	0.4948
Frank	0.4571	0	0
Joe	0.6776	0	0.8515

Table A.3: Dependência e dependência na aba para diferentes cópulas ajustadas para LLOY

As opções para modelar a dependência da aba inferior serão resumidas no caso da t-Copula, que tem dependência simétrica na aba, e cópula de Clayton que tem toda a dependência da aba na aba esquerda e mais forte do que na t-Copula.

Se escolhermos uma cópula com maior dependência na aba estaremos a associar uma maior probabilidade de um evento extremo em ambos, na instituição financeira e no Sistema Financeiro simultaneamente. Num contexto de gestão de risco, o comportamento da aba e a dependência da aba são críticos.

Outro aspecto a ser considerado é a simetria da cópula. Algumas cópulas são simétricas como t-Cópula por exemplo e outras não serão simétricas como cópulas de Clayton ou outras cópulas de valores extremos. Porém, num contexto que envolve gestão de risco, não será o mais adequado modelar perdas e lucros extremos da mesma forma, para que se evite a subvalorização do risco envolvido.

O processo de selecção de uma cópula contará com os dados para determinar a forma da cópula, mas resulta também de considerações teóricas e não apenas da parametrização desses dados. Nesse caso, antes de iniciar o processo de adaptação também devemos decidir qual família de cópula que é mais adequada, dado o conhecimento prévio dos riscos envolvidos.

A.5.1.2 Selecção da Cópula e Dependência na Aba

Incluir a dependência na aba no modelo pode ser obtida através da escolha da família de cópula apropriada. Aqui a estratégia não é identificar a cópula apenas com base num critério de selecção como o AIC ou BIC, mas incluir aqui também o coeficiente de dependência teórica da aba da cópula como resultado da seguinte expressão obtida para cada família de cópula:

$$\lambda_L = \lim_{u \searrow 0} \frac{C(u,u)}{u}$$
$$\lambda_U = \lim_{u \nearrow 1} \frac{1 - 2u + C(u,u)}{1 - u}$$

Como estamos principalmente interessados na dependência da cauda inferior, o critério de selecção aplicado irá comparar o coeficiente da cauda inferior.

Este critério de selecção será usado também com um teste de adequação conforme especificado e implementado no *package* R *VineCopula* (Schepsmeier et al., 2015).

O teste de hipótese é definido como:

$$H_0 = \boldsymbol{H}(\theta) + \boldsymbol{C}(\theta) = 0$$
$$H_1 = \boldsymbol{H}(\theta) + \boldsymbol{C}(\theta) \neq 0$$

onde $\boldsymbol{H}(\theta)$ é a matriz Hessiana esperada e $\boldsymbol{C}(\theta)$ é o produto esperado da função de pontuação.

A título de exemplo, passaremos pelos resultados obtidos para a instituição financeira BBVA.

Copula			Dependência	
	AIC	BIC	Aba esquerda	p-Value
Gaussian	-926.4388	-921.5898	0.0000	0.42
t	-1003.3092	-993.6111	0.4997	0.15
Clayton	-830.1377	-825.2886	0.7190	0.00
Gumbel	-972.1559	-967.3069	0.6686	0.00
Frank	-898.8010	-893.9519	0.0000	0.00
Joe	-821.4519	-816.6028	0.7323	1.00
BB1	-993.3541	-983.6560	0.6149	0.61
BB6	-970.0841	-960.3860	0.6687	0.00
BB7	-966.5711	-956.8729	0.6754	0.00
BB8	-902.6389	-892.9408	0.0000	0.00

Table A.4: Qualidade do ajustamento da Copula para BBVA

A partir desses resultados, apenas as famílias de cópula que exibem um valor p significativo serão consideradas. Destes, tomaremos aquele com coeficiente de cauda inferior mais alto. Este critério leva-nos à cópula de Joe, que parece ser a melhor abordagem para modelar as abas da cópula. Como o critério AIC representa o melhor ajuste para toda a distribuição, não teremos necessariamente uma correspondência entre o critério AIC e o coeficiente de dependência da cauda.

De maneira semelhante, podemos prosseguir com a escolha da cópula a ser utilizada para modelar as secções fora das abas, a secção mais densa da distribuição bivariada, mas agora devemos usar o critério AIC (ou BIC), Como estamos a procurar um melhor ajuste em toda a distribuição (não apenas nas abas). Temos como melhor ajuste para o propósito a t-Cópula. Podemos notar que a cópula gaussiana também tem resultados semelhantes com significância mais forte (0,42 contra 0,15 em t - copula).

Em conclusão, executando este processo num grande número de famílias de cópula, concluímos que:

- Aplicando uma cópula de Joe para a cauda, quando a instituição financeira está em situação de stress.
- Aplicando a t-Cópula para modelar o caso em que a instituição financeira não está em stress ($\alpha = 0.5$).

A.5.1.3 Função de Distribuição Condicional com Cópulas

Para obter as probabilidades condicionais envolvidas no cálculo do Sistema Financeiro $CoVaR$, é também necessário determinar a função de distribuição condicional da cópula bivariada, ajustando os dados:

$$h_1(u|v;\boldsymbol{\theta}) := P(U \leq u|V = v) = \frac{\partial C(v,u;\boldsymbol{\theta})}{\partial v},$$

onde $(U,V) \sim C$ é uma função de distribuição de cópula bivariada com parâmetro θ.

Em termos de medida de risco sistémico, este resultado pode ser traduzido como:

$$P(R_{s,t} \leq CoVaR_{\alpha,\beta,t}|R_{i,t} = VaR^j_{\alpha,t}) = \frac{P(R_{s,t} \leq CoVaR_{\alpha,\beta,t})}{P(R_{i,t} \leq VaR^j_{\alpha,t})}$$

como descrito em A.4.2 e A.4.3.

Este processo será utilizado duas vezes, um estimará o *CoVaR* quando a instituição financeira enfrentar uma situação de stress e outro quando a instituição financeira estiver com retornos médios. O *CoVaR* corresponde a:

$$CoVaR_{\alpha\beta t} \tag{A.26}$$

na primeira situação e na segunda situação por:

$$CoVaR_{0.5\beta t} \tag{A.27}$$

Para finalizar o processo de cálculo do *CoVaR*, precisamos ainda calcular o inverso da função de distribuição cumulativa da cópula marginal para os retornos do sistema financeiro.

A.5.2 Ajuste das Margens

Como último passo para obter a estimativa do *CoVaR*, temos que ajustar a série de retornos do sistema financeiro para, com aquela função de probabilidade, usar a inversa da função de distribuição cumulativa para finalmente calcular o valor do *CoVaR*.

Para mitigar os inconvenientes do uso de uma premissa normal, a opção de modelar os retornos do sistema financeiro recairá em um modelo misto.

A ideia por trás do uso da distribuição de Mistura de Valores Extremos é combinar a flexibilidade de usar uma distribuição para capturar o componente principal, correspondendo aos quantis centrais, também chamados de distribuição do núcleo, que podem ser, por exemplo, uma Normal, e também as abas, como valores extremos. Com este modelo de mistura, obter-se-á uma função de distribuição completa, dividindo a distribuição em um componente em nuclear e em componentes da aba. A função de mistura também permite uma mistura de distribuição de famílias distintas.

Com o propósito de modelar o risco sistémico, estamos interessados em explorar uma mistura de uma distribuição normal como distribuição no núcleo (central) com duas distribuições Gama na aba direita e esquerda (MacDonald et al., 2011).

Este modelo usa estimadores de densidade no núcleo para estimar a distribuição de valores não extremos e DPG para estimar a distribuição na aba. Este estimador de densidade no núcleo assume uma densidade normal, que é centralizada em cada ponto de dados e usa apenas um parâmetro para definir a largura da componente central. As componentes nas abas referem-se à proporção da distribuição acima do limite determinado. Este parâmetro será identificado por Φ_u e u representa o limite. A função de distribuição vem como:

$$F(x|\Theta) = \begin{cases} \phi_{u_l}\Big(1 - G(-x| - u_l, \sigma_{u_l}, \epsilon_l)\Big), & x < u_l \\ \\ H(x|\mu, \sigma), & u_l \leq x \leq u_r \\ \\ (1 - \phi_{u_r}) + \phi_{u_l} G(x|u_r, \sigma_{u_r}, \epsilon_r) & x > u_r \end{cases}$$

onde $\phi_{u_l} = H(u_l|\mu, \sigma)$ e $\phi_{u_r} = 1 - H(u_r|\mu, \sigma)$ e $H(.|\mu, \sigma)$ é uma distribuição Normal com média μ e desvio padrão σ. $G(.| - u_l, \sigma_{u_l}, \epsilon_l)$ e $G(.| - u_r, \sigma_{u_r}, \epsilon_r)$ são distribuições DPG para abas esquerda e direita, respectivamente.

Aplicando o modelo de mistura de valores extremos às séries de retornos do Sistema Financeiro, obtivemos as seguintes estimativas para os parâmetros:

- O gráfico mostra os resultados para uma mistura de um normal N(-0.017,0.787) limitado à esquerda pelo parâmetro u_l = -0.940 e à direita por parâmetro u_r = 0.909.
- Os parâmetros Gama obtidos são respectivamente:

left tail	**right tail**
$\phi_{u_l} = 0.0669$	$\phi_{u_r} = 0.129$
μ_l =0.3858	μ_r = 0.2218
σ_l =0.0139	σ_r = 0.0174

A qualidade do ajuste para este modelo também é ligeiramente melhor do que os anteriores com um valor de critério BIC estimado em -2356,741. A vantagem e flexibilidade deste modelo de mistura está essencialmente nas caudas da distribuição, visto que é capaz de tirar partido das capacidades da distribuição Gama para se adaptar à cauda. Os gráficos obtidos para a mistura de valores extremos são os seguintes:

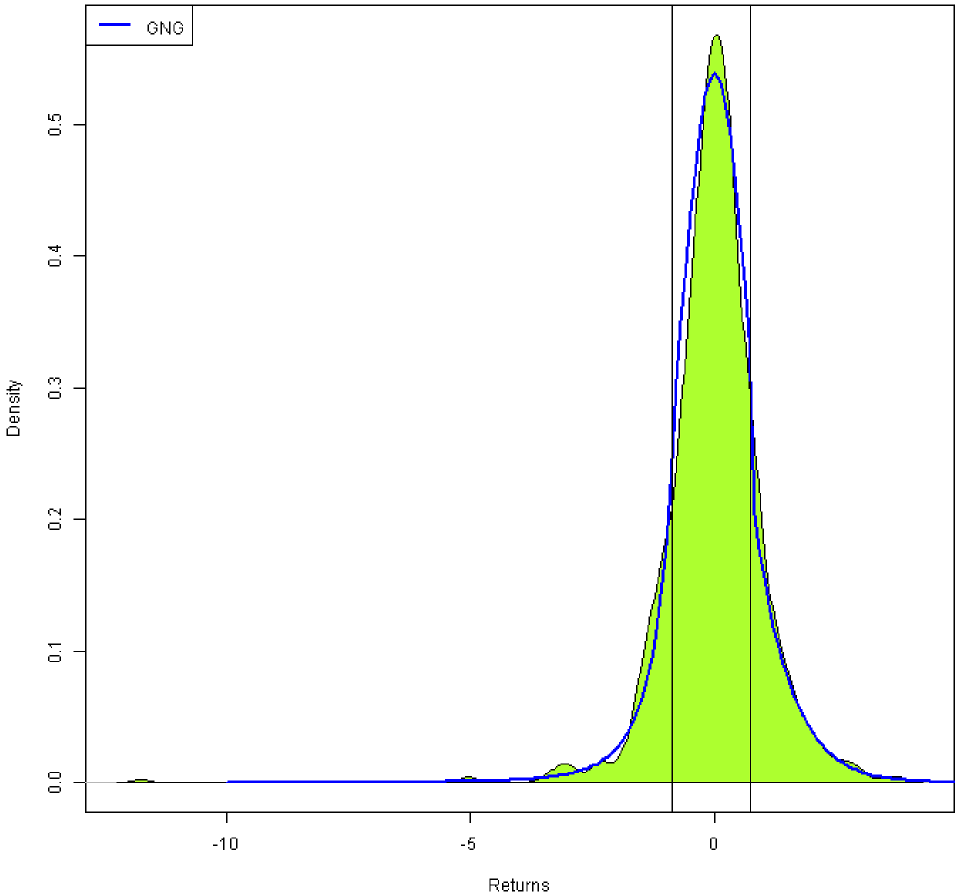


Figure A.4: O ajuste para os retornos semanais do sistema financeiro com uma mistura de distribuições Gamma-Normal-Gamma

Pela análise visual do gráfico obtido com o ajuste a uma Gamma-Normal-Gamma (GNG) é possível identificar um ajuste muito próximo.

Também na aba da distribuição, podemos notar uma boa aproximação incluindo um decaimento da função de distribuição mais á esquerda. Em vez de usar uma única distribuição para ajustar a série de retornos financeiros sobre todos os quantis, podemos usar um modelo de mistura para ajustar distribuições distintas de acordo com os quantis (Basilio and Oliveira, 2020).

A.5.3 Resultados $\Delta CoVaR$

Na tabela a seguir, podemos comparar o impacto que cada opção de modelação, ao seleccionar famílias de cópula distintas, tem em termos do resultado calculado para for $CoVaR$ and $\Delta CoVaR$.

Copula Family	$CoVaR_{0.5}$	$CoVaR_{\beta=0.01}$	$CoVaR_{\beta=0.001}$	$\Delta CoVaR_{\beta=0.001}$
t	-0.0445	-0.2111	-0.4990	-0.4545
Clayton	-0.0742	-0.0982	-0.0986	-0.0244
Gumbel	-0.0570	-0.1827	-0.2760	-0.2190
t & Clayton	-0.0445	-0.0982	-0.0986	-0.0541
t & Gumbel	-0.0445	-0.1827	-0.2760	-0.2315

Table A.5: Resultados obtidos para $CoVaR$ e $\Delta CoVaR$ para LLOY ($\alpha = 0.01$)

Considerando três famílias de cópula, t como um exemplo simétrico com dependência na aba fraca simétrica, e por outro lado a cópula de Clayton com aba à esquerda e a cópula de Gumbel com aba à direita.

A cópula t, por ser simétrica e permitirá melhores resultados se utilizada para estimar apenas o $CoVaR_{0,5}$, na mediana, representando uma situação onde a instituição financeira não está sujeita a stress, também interpretada como a situação esperada em circunstâncias normais.

Nas caudas da distribuição conjunta dos retornos das instituições financeiras e dos retornos do sistema financeiro, t-Cópula tende a sobrestimar o risco devido às dificuldades em lidar com a dependência na aba.

A.5.4 Identificação de Instituições Financeiras com Risco Sistémico

As relações de co-dependência de risco entre as instituições financeiras consideradas foram estimadas utilizando a metodologia descrita, aplicando-se duas formas distintas de selecção da cópula. Esta abordagem mostra-se adequada para estimar as relações sistémicas e incorporar a dependência na aba versus dependência sobre a mediana (Embrechts et al., 2001).

Δ *CoVaR* representa uma estimativa da magnitude da contribuição de cada entidade financeira para o risco sistémico de mercado.

A tabela acima mostra os valores estimados de Δ *CoVaR* para Junho de 2018, aplicando uma janela móvel de 3 anos.

Ordem	Instituição Financeira	$\Delta CoVaR_{\beta=0.01}$	τ	Copula (mediana)	Copula (aba)
1	BNP	6.94%	0.49	t	Clayton-Gumbel (BB1)
2	ISP	6.20%	0.51	t	Clayton-Gumbel (BB1)
3	BIRG	6.00%	0.44	t	Clayton-Gumbel (BB1)
4	GLE	5.97%	0.48	t	Clayton-Gumbel (BB1)
5	CBK	5.93%	0.46	t	Clayton-Gumbel (BB1)
6	MB	5.90%	0.47	t	Clayton-Gumbel (BB1)
7	ABN	5.86%	0.51	t	Clayton-Gumbel (BB1)
8	SAN	5.79%	0.41	t	Clayton-Gumbel (BB1)
9	RBI	5.66%	0.44	t	Clayton-Gumbel (BB1)
10	NDA-DK	5.61%	0.47	t	Clayton-Gumbel (BB1)

Table A.6: Top 10 das Instituições Financeiras por $\Delta CoVaR$ - June 2018

A classificação da importância sistémica para as instituições financeiras, como medida de risco sistémico, não está necessariamente vinculada, ou mesmo evidencia uma possível situação de stress de uma determinada instituição financeira. Em vez disso, a classificação da medida de risco sistémico reflecte o impacto adicional esperado como um custo para o sistema financeiro, no caso que tal evento ocorrer na instituição financeira específica.

A tabela acima lista as instituições financeiras ordenadas pelo impacto estimado de risco sistémico como % de *Delta CoVaR* no sistema *VaR*, representando o impacto de um evento significativo naquela instituição financeira reflectido no sistema *VaR*.

A.5.5 Adicionar Variação Temporal

O trabalho original de Adrian and Brunnermeier (2011) incluiu uma camada adicional de pressupostos que fazem o retorno da instituição, X dependente de um conjunto de variáveis de estado e assumindo um modelo de factorização subjacente para retornos de activos, onde o retorno de cada activo depende linearmente destes factores:

- Um conjunto de variáveis de estado M_{t-1}
- O crescimento de activos em todo o sistema X^{sys}

Dessa forma, o crescimento dos activos de cada instituição financeira dependerá de variáveis de estado desfasadas seleccionadas, enquanto a taxa de crescimento dos activos do sistema depende do crescimento dos activos bancários individuais e das variáveis de estado desfasadas.

Como nosso objetivo é remover suposições adicionais do modelo, uma alternativa para evitar essa camada extra de pressupostos é aplicar uma técnica de janelas de tempo deslizantes como uma forma de incluir a variação de tempo para ter uma análise ao longo do tempo (Chong and Hurn, 2016). Definindo uma amplitude temporal, por exemplo, 3 anos de dados, iremos mover esta janela de tempo dia a dia, ou semana a semana e aplicar todos os passos anteriores a cada uma dessas janelas de tempo. No final, obtemos uma série temporal para *VaR*, *CoVaR* and $\Delta CoVaR$.

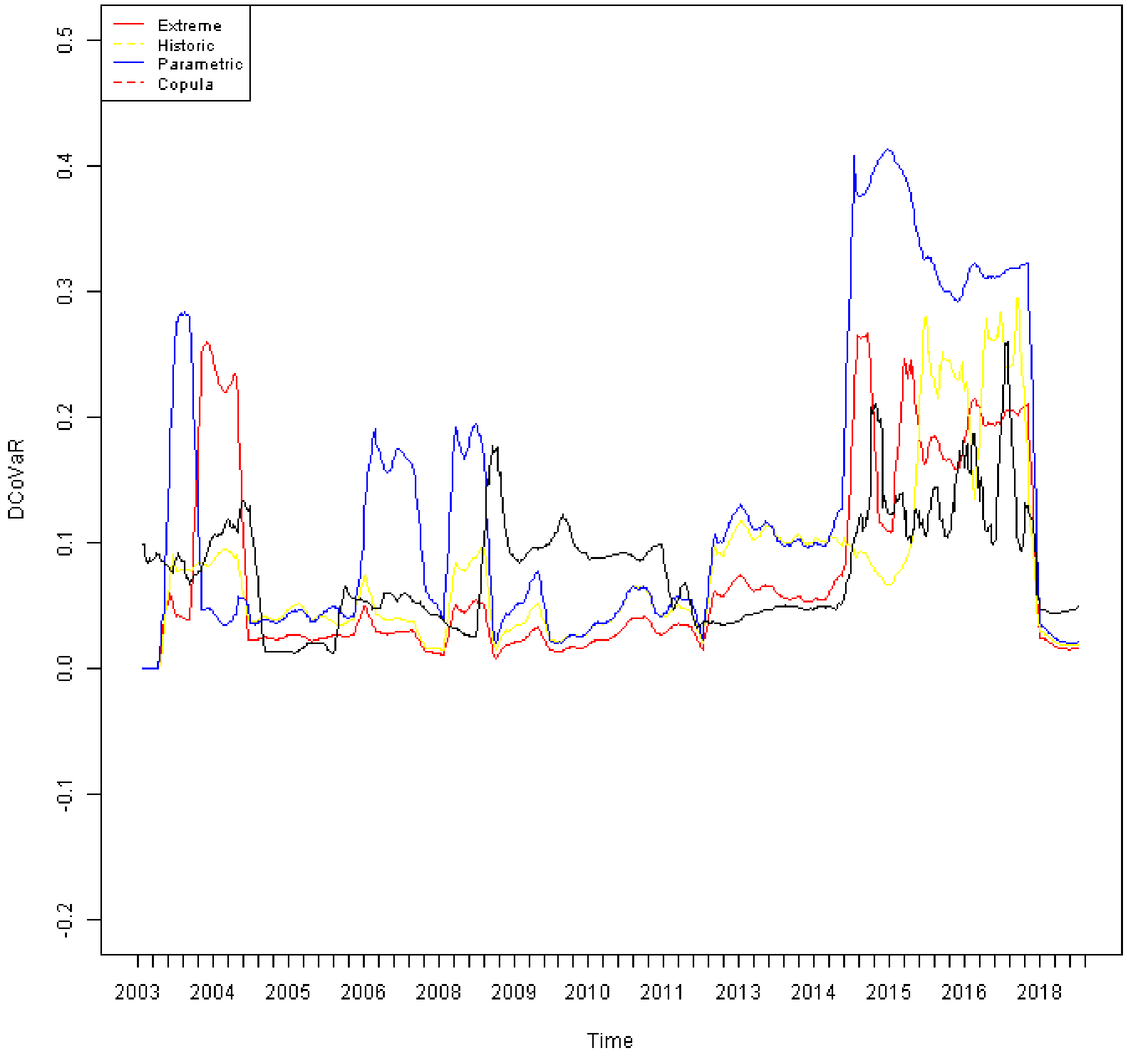


Figure A.5: $\Delta CoVaR$ para UBS aplicado quatro metodologias: Valores Extremos, Paramétrico, Histórico e Copula

Existem várias vantagens que podemos reconhecer nas técnicas de janelas de tempo contínuas (*rolling-windows*) para analisar e modelar séries temporais (Zivot and Wang, 2003), uma vez que é acessível de implementar, fácil de interpretar e também sem incluir suposições adicionais. Ao aplicar a janela de tempo deslizante para estimar o $\Delta CoVaR$, para cada ponto no tempo é possível ter uma percepção da evolução de $\Delta CoVaR$ através do tempo.

Neste caso $\Delta CoVaR$ está a ser estimado para cada ponto no tempo (semana) e os parâmetros da cópula são ajustados para janela de tempo definida como forma de actualizar a estrutura de dependência entre as duas séries de retornos.

A.6 Conclusão

O risco sistémico Δ $CoVaR$, conforme calculado nos capítulos anteriores, pode ser usado para fornecer um perfil estatístico abrangente e unificado dos bancos de acordo com seu nível de implicação (contribuição e exposição) no risco sistémico. Deste modo, montamos um mapa detalhado para mostrar o posicionamento relativo de todos os bancos de acordo com sua implicação no risco sistémico.

Ao aplicar a metodologia sugerida neste trabalho de investigação, estamos a explorar soluções de modelação mais flexíveis, uma vez que não dependem de premissas de regressão linear entre os retornos da instituição financeira e os retornos do sistema financeiro. Como esta relação não é linear nem constante ao longo de uma função de distribuição, incluímos no modelo uma mistura de distribuições a fim de melhorar o ajuste dos retornos nas abas.

Ao aplicar uma abordagem de ajuste de duas funções de cópula, uma nas abas e outra nos quantis centrais permitiu captar melhor a dependência entre as duas séries de retornos, uma vez que a natureza dessa dependência tende a se alterar em função dos quantis da aba distribuição.

Os resultados obtidos em termos de ajuste também mostraram que enquanto no centro da distribuição (os quantis centrais) uma t-Cópula mostrou ser o mais adequado para a generalidade das instituições financeiras. Nas caudas, a cópula de Gumbel mostrou-se mais adequada. É evidenciada ainda uma preocupação quanto à qualidade do ajuste global versus a qualidade do ajuste nas caudas da distribuição. A qualidade do ajuste também pode variar dependendo do quantil que se está a analisar.

Em algumas aplicações, como no caso da análise de risco, a análise das abas das distribuições é de suma importância. É desejável que esses modelos também sejam capazes de lidar com a dependência nos extremos, e devemos também levar em consideração uma relação especial de dependência nos extremos da distribuição. Em termos de retorno financeiro, significa que, por exemplo, em condições extremas, ou nos extremos da distribuição de retornos, a dependência costuma ser mais forte no contexto de uma crise financeira do que em um cenário normal. Um modelo adequado deve ser capaz de acomodar esta situação.

Além disso, para se obter um correto ajuste da função de cópula também é de grande importância garantir um ajuste adequado nas distribuições marginais, em especial a marginal que representa os retornos do sistema financeiro, de forma que o ajuste proporcione bons resultados não apenas no distribuição geral, mas também nas caudas da distribuição. Para atender a esse objectivo aplicamos um modelo de mistura de valores extremos para adequação aos resultados financeiros, o que nos deu flexibilidade para nos adaptarmos à cauda da distribuição proporcionando melhores ajustes.

Fenómenos complexos, como o comportamento dos retornos financeiros, requerem modelos mais complexos e versáteis. Neste caso, para identificar instituições financeiras de maior risco sistémico com base em $CoVaR$ é fundamental dispor de metodologias que permitem um modelo de dependência nas abas da distribuição correto, e os métodos de ajuste que poderiam ser adequados para o ajuste da distribuição geral pode não ser o mais apropriado para o ajuste nas abas da distribuição.

Appendix B

List of Financial Institution

Table B.1: List of Financial Institutions used

ID	COMPANY NAME	CONTRY CODE	CURR.	QUOTE CODE
1	HSBC	GB	GBP	HSBA.L
2	BCO SANTANDER	ES	EUR	SAN.MC
3	BNP PARIBAS	FR	EUR	BNP.PA
4	LLOYDS BANKING GRP	GB	GBP	LLOY.L
5	UBS GROUP	CH	USD	UBS
6	ING GRP	NL	USD	ING
7	BCO BILBAO VIZCAYA ARGENTARIA	ES	EUR	BBVA.MC
8	INTESA SANPAOLO	IT	EUR	ISP.MI
9	BARCLAYS	GB	GBP	BARC.L
10	CREDIT SUISSE GRP	CH	USD	CS
11	GRP SOCIETE GENERALE	FR	EUR	GLE.PA
12	NORDEA BANK	FI	DKK	NDA-DK.CO
13	UNICREDIT	IT	EUR	UCG.MI
14	SWEDBANK	SE	SEK	SWED-A.ST
15	STANDARD CHARTERED	GB	GBP	STAN.L
16	DEUTSCHE BANK	DE	EUR	DBK.DE
17	KBC GRP	BE	EUR	KBC.BR
18	DNB	NO	NOK	DNB.OL
19	SKANDINAVISKA ENSKILDA BK A	SE	SEK	SEB-A.ST

ID	COMPANY NAME	CONTRY CODE	CURR.	QUOTE CODE
20	SVENSKA HANDELSBANKEN A	SE	SEK	SHB-A.ST
21	CREDIT AGRICOLE	FR	EUR	ACA.PA
22	DANSKE BANK	DK	EUR	DSN.F
23	CAIXABANK	ES	EUR	CABK.MC
24	ROYAL BANK OF SCOTLAND GRP	GB	GBP	RBS.L
25	ERSTE GROUP BANK	AT	EUR	EBS.VI
26	ABN AMRO GROUP	NL	EUR	ABN.AS
27	JULIUS BAER GRP	CH	CHF	BAER.VX
28	COMMERZBANK	DE	EUR	CBK.DE
29	PKO BANK	PL	EUR	P9O.MU
30	BCO SABADELL	ES	EUR	SAB.MC
31	BANK OF IRELAND GROUP	IE	EUR	BIRG.IR
32	PEKAO	PL	PLN	PEO
33	MEDIOBANCA	IT	EUR	MB.MI
34	BANKINTER	ES	EUR	BKT.MC
35	NATIXIS	FR	EUR	KN.PA
36	FINECOBANK	IT	EUR	FBK.MI
37	CYBG PLC	GB	GBP	CYBG.L
38	BANKIA	ES	EUR	BKIA.MC
39	RAIFFEISEN BANK INTERNATIONAL	AT	EUR	RBI.VI
40	AIB GROUP	IE	EUR	AIBG.L
41	BANCO BPM	IT	EUR	BAMI.MI
42	JYSKE BANK	DK	EUR	JYS1.F

ID	COMPANY NAME	CONTRY CODE	CURR.	QUOTE CODE
43	CLOSE BROTHERS GRP	GB	GBP	CBG.L
44	SANTANDER BANK POLSKA	PL	EUR	BZI.F
45	UBI BCA	IT	EUR	UBI.MI
46	METRO BANK	GB	GBP	MTRO.L
47	CEMBRA MONEY BANK	CH	EUR	GEH.BE
48	BCO COMERCIAL PORTUGUES	PT	EUR	BCP.LS
49	Banco Espírito Santo	PT	USD	BKESY
50	Banco BPI	PT	EUR	BPI.LS