Aalto University
School of Science

Master's programme in Mathematics and Operations Research

# Comparison of Options for Including Nonlinear Dependencies in Portfolio Credit Risk Modeling

**Elias Ylä-Jarkko**

Master's Thesis
2025

Aalto University
School of Science

**Author** Elias Ylä-Jarkko

**Title** Comparison of Options for Including Nonlinear Dependencies in Portfolio Credit Risk Modeling

**Degree programme** Mathematics and Operations Research

**Major** Applied Mathematics

**Supervisor** Prof. Ahti Salo

**Advisor** D.Sc. (Econ.) Jaakko Sääskilahti

**Collaborative partner** OP Financial Group

**Date** 17 April 2025 **Number of pages** 56 **Language** English

**Abstract**

Banks prepare for crises by estimating the amount of capital they would need in extreme situations. That capital amount is called economic capital. Before the 2008 financial crisis, when estimating economic capital, it was common for banks and regulators to assume that the dependence of credit losses in a portfolio is linear according to the multivariate normal distribution. In other words, the dependence was estimated with the Gaussian copula. However, several studies have observed that the Gaussian copula underestimates risk in extreme scenarios. Therefore, many banks seek to replace it with another copula that can model nonlinear dependencies, but with many alternatives, making the choice can be difficult.

This thesis helps to choose the copula for an economic capital model by comparing the Student's t copula and the Gumbel copula as alternatives to the Gaussian copula. The copulas are compared with qualitative criteria based on factors that were found important when developing economic capital calculation in a large bank. In particular, copulas are compared by their theoretical justifiability, the complexity of their simulation algorithms, and their explainability to model users. Empirical performance is assessed by citing the literature.

The results of the comparisons suggest that no copula can be declared to be the best in terms of the criteria. The Gumbel copula is theoretically better in modeling stock market data, and therefore credit losses, than the other copulas. On the other hand, the Student's t copula also has desired dependence modeling ability, and it is easier to explain and has a simpler implementation than the Gumbel copula.

The results of this thesis imply that the Student's t copula is unconditionally better than the Gaussian copula for modeling economic capital. The Gumbel copula may not always be better than the Gaussian or Student's t copula, but it should be considered in some cases, for example if theoretical justifiability is highly valued and the implementation could be simplified.

**Keywords** Copula, economic capital, credit risk, nonlinear

Aalto-yliopisto
Perustieteiden
korkeakoulu

**Tekijä** Elias Ylä-Jarkko

**Työn nimi** Comparison of Options for Including Nonlinear Dependencies in Portfolio Credit Risk Modeling

**Koulutusohjelma** Mathematics and Operations Research

**Pääaine** Applied Mathematics

**Työn valvoja** Prof. Ahti Salo

**Työn ohjaaja** KTT Jaakko Sääskilahti

**Yhteistyötaho** OP

**Päivämäärä** 17. huhtikuuta 2025 **Sivumäärä** 56 **Kieli** englanti

**Tiivistelmä**

Pankit varautuvat kriiseihin laskemalla arvion taloudelliseta pääomasta, eli pääomasta joka vaaditaan kattamaan tappiot erittäin harvinaisessa tilanteessa. Ennen vuoden 2008 finanssikriisiä monet pankit ja valvovat viranomaiset olettivat taloudellista pääomaa arvioidessaan, että luottotappioiden todennäköisyydet lainasalkussa riippuvat toisistaan lineaarisesti kuin moniulotteisessa normaalijakaumassa, eli normaalikopulan mukaisesti. Useissa tutkimuksissa on kuitenkin havaittu, että normaalikopulan käyttäminen aliarvioi äärimmäisiä riskejä. Monet pankit haluavat siksi käyttää jotakin toista, epälineaarisia ilmiöitä huomioivaa kopulaa riippuvuuksien mallintamiseen. Vaihtoehtoja on kuitenkin lukuisia, eikä niiden vertailu ole helppoa.

Tässä opinnäyttetyössä helpotetaan kopulan valintaa luottosalkun riskimallinnukseen vertailemalla Studentin t -kopulaa ja Gumbel-kopulaa vaihtoehtona normaalikopulalle. Kopuloita vertaillaan kvalitatiivisista näkökulmista, jotka on havaittu tärkeiksi, kun kehitetään suuren pankin taloudellisen pääoman laskentaa. Erityisesti työssä vertaillaan kopuloita kolmesta näkökulmasta: kuinka hyvin kutakin kopulaa hyödyntävä malli kykenee arviomaan tappioiden suuruutta, kuinka monimutkainen mallin laskenta-algoritmista tulisi ja kuinka helposti mallin rakenne on selitettävissä mallin käyttäjille. Empiiristä suorituskykyä ei mallinneta tässä työssä, vaan sitä arvioidaan nojautumalla tutkimuskirjallisuuteen.

Vertailun tuloksista havaitaan, että yksiselitteisesti parasta kopulaa ei ole. Gumbel-kopulan ominaisuudet kuvaavat luottotappioiden käyttäytymistä teoriassa parhaiten. Toisaalta Studentin t -kopulalla kyetään huomioimaan riippuvuuksia halutusti, ja sen selitettävyys todettiin paremmaksi ja toteutus yksinkertaisemmaksi kuin Gumbel-kopulan.

Tämän työn tuloksista voidaan päätellä, että Studentin t -kopula on tilanteesta riippumatta parempi vaihtoehto taloudellisen pääoman mallintamiseen kuin normaalikopula. Gumbel-kopula ei välttämättä ole parempi kuin Studentin t -kopula tai normaalikopula, mutta se voi olla hyvä vaihtoehto, jos esimerkiksi teoreettista suorituskykyä pidetään tärkeänä ja toteutusta onnistutaan yksinkertaistamaan.

**Avainsanat** Kopula, taloudellinen pääoma, luottoriski, epälineaarinen

# Preface

When first introducing the topic to family members, there were curious and creative guesses about the interpretation of the thesis title. One guesser must have been hungry, judging by the connection of *copula* to the Finnish word *kapula*.

> Different types of copulas are used to cook different types of soups. Bigger copulas are needed for family soups, whereas smaller ones are suitable for one person meals. This thesis will compare different copulas and their usage.

Another guesser realized that they are already familiar with nonlinear tails - a key to understanding nonlinear tail dependence.

> For decades, it has been a mystery why animals have different tails. Some of them have linear tails, such as cows, but some of them have nonlinear tails, for example pigs. The aim of this thesis is to solve this huge mystery.

If I were to rewrite this thesis, I might consider choosing one of these interpretations instead of the one currently presented. However, I express my deepest gratitude to my family, friends, advisor, supervisor, and coworkers for their support throughout the process of writing this less cute but arguably more scientifically sound thesis.

Espoo, 17 April 2025

Elias A. Ylä-Jarkko

# Contents

# Abbreviations

| | |
|---|---|
| CCF | credit conversion factor |
| EAD | exposure at default |
| EC | economic capital |
| EL | expected loss |
| ICAAP | internal capital adequacy assessment process |
| i.i.d | independent and identically distributed |
| IRB | internal ratings based |
| LGD | loss given default |
| PCA | principal component analysis |
| PD | probability of default |
| PiT | point-in-time |
| RWA | risk-weighted assets |
| TTC | through-the-cycle |
| VaR | value at risk |
| WCDR | worst-case default rate |

# 1 Introduction

Banks want to have enough capital to cover credit losses, and thus avoid bankruptcy. A tool for assessing this capital amount is called economic capital [1]. Economic capital models aim to measure the effect of unexpected phenomena in the economy by estimating the amount of capital that is needed to cover losses up to a selected confidence level within the modeling horizon, for example, one year [2]. This corresponds to the rarest event that the bank is willing to reserve capital for [3, Ch. 2]. After modeling the loss distribution and calculating the size of the losses at the selected confidence level, banks aim to hold at least that amount of capital and manage it as part of their Internal Capital Adequacy Assessment Process (ICAAP). Preparing for every possible scenario is not feasible, as banks could then only lend out as much as they hold in equity. Moreover, accurate modeling of economic capital is important because capital reserves have to be large enough, but excessive reserves decrease a bank's profitability.

Modeling the loss distribution is difficult because observations of crises, especially large ones, are extremely rare [4]. Therefore, some assumptions are needed to model even the extreme losses. A common solution in corporate portfolio modeling is to build a structural model that makes two important assumptions. The first is that defaults, that is, failure to pay back loans, are triggered by a decrease in the firm's asset value and that value follows a geometric Brownian motion, also called the Merton model [5]. Although greatly simplifying, the Merton model is popular due to its influence in credit risk modeling [3, Ch. 8]. The other major simplification is to assume that the value of a firm's assets is driven by one or more macroeconomic factors and an independent factor [6]. Together, these make it possible to model the total loss in stages. Loss distributions can first be estimated for each obligor independently. Then, the individual distributions can be combined with a dependence structure that can be inferred from stock market data, for instance. Furthermore, the total loss can be attributed partly to the factors that were considered in modeling, for example industries, geographical regions, or economic factors. This helps explain the model and the drivers of risk for model users.

An alternative to structural models is a reduced form model. Instead of connecting defaults to asset values, reduced form models assume a hazard rate process that determines probability of default. They perform well when sufficient historical data from credit default swaps or defaults is available, but this often is unrealistic for anything but the largest corporations. On the other hand, defining a reduced form model can be easier than determining the dependence structure of a structural model. Choosing a structural model and a dependence structure may still be beneficial, especially for smaller or newer banks due to lower data requirements and better explainability [7, Ch. 7].

In this thesis, we explore alternatives for nonlinear dependence structures for structural models. We focus on those that can be easily represented by copulas and implemented in a factor model. Similar comparisons have been made previously [8], but this thesis contributes to the literature by presenting and evaluating the theoretical justification, practical implementation in a simulation model, and explainability for model users. Model accuracy has been comprehensively studied by, e.g. [8–10] and remains a major area of research.

Historically, the financial industry has used linear correlation with a Gaussian copula to estimate asset value dependence [11, 12]. However, several studies show evidence that asset correlation is nonlinear in some situations [13, 14]. Longin and Solnik showed that the correlation between asset prices increases in times of low asset values. However, they did not find a similar phenomenon for high asset values. This poses a challenge for modeling asset prices because the dependence is not only nonlinear but also asymmetric. Methodological deficiencies in accounting for these phenomena may not be dismissed, as underestimating risk has serious consequences. If a linear dependence structure is calibrated to the correlation levels of normal times, implying some level of diversification benefits, the model would apply similar diversification benefits to a crisis scenario, underestimating the amount of loss that could, in reality, occur at once. For example, the 2007-2009 subprime mortgage crisis demonstrated the consequences of not adequately estimating asset value correlations, since it was partly caused by underestimating the possibility for simultaneous losses [12]. Here lies the motivation for meticulous credit risk modeling and the reason for investigating nonlinear dependence structures in this thesis.

Different solutions for accurate modeling of asset dependence have been proposed to supersede the previously popular linear dependence model and the Gaussian copula, such as Student's t copula [15], Archimedean copulas [9], vine copulas [16, 17], Wishart sampling for the correlation matrix [18], Levý processes [19] and new algorithms for setting and selecting copulas [20]. We examine the methods that are most similar to the Gaussian copula, namely the Student's t copula and the Hierarchical Gumbel copula, because they would be the easiest replacements for a Gaussian copula in a portfolio credit risk model while avoiding a total model overhaul.

Specifically, we define what a copula is and how it can be used to model the dependence of asset prices in loss distributions. We then present two nonlinear options, the Gumbel copula and the Student's t copula, and the Gaussian copula for reference. The contributions of this thesis then follow, as the theoretical justification is evaluated, implementations of simulation algorithms are presented and examined, and finally, the explainability of the models is assessed. Finally, we conclude the pros and cons of each copula in economic capital modeling.

# 2 Background

When granting a loan, there is possibility that the counterparty does not pay back the loan. This is called *credit risk*. Even though lending is risky, banks have a large incentive to engage in it, to profit from interest payments by the counterparty. However, to keep the activity profitable, banks have to carefully assess the associated risk and how much loss they expect over certain time horizons. Furthermore, they also have to prepare for unexpected scenarios to remain operational and trustworthy in times of crisis. As a solution, banks carry out *credit risk modeling* to estimate the probability and magnitude of different credit loss scenarios.

This thesis studies the implementation of nonlinear dependence structures in economic capital models. Therefore, we cover the basics of economic capital and credit risk modeling, including common strategies, risk parameters, and regulation in credit risk modeling. Moreover, we introduce the reader to copulas to have a common ground for understanding the assessments in later sections.

## 2.1 Credit risk modeling

To model credit risk, one has to model the probability distribution of losses that could occur, given a loan portfolio and a time horizon. This probability distribution is called the loss distribution. Banks are interested in both the expected value of the loss distribution and the tail of the distribution because the loss has to be examined from the perspectives of day-to-day business, as well as risk estimation. The expected value is accounted for in, e.g. pricing and customer selection processes, as it resembles the amount of loss that can be expected in the long run, over a wide range of possible scenarios. On the other hand, the upper tail of the loss distribution, in other words, the extreme losses, are examined carefully because banks also have to prepare for the most unlikely events to avoid bankruptcy. This thesis focuses on estimating the capital needed to withstand these unexpected events, also called *economic capital*.

### 2.1.1 Economic capital

It is challenging to estimate how much capital a bank should hold because losses vary over time and there are various methods for taking it into account in calculations. Specifically, it is challenging to determine how much capital is needed in a crisis because the frequency and size of crises are difficult to estimate. Banking regulation aims to ensure that all banks are equally and adequately prepared for such crises by imposing minimum levels of capital held by each bank. This is called *regulatory capital*. However, as demonstrated in the following sections, the regulatory capital requirements are not always aligned with the views that banks have on their risk positions. Consequently, banks develop internal methodologies for estimating credit losses and thus the capital amount to absorb them. The internal estimates are called economic capital. There are alternative ways to measure economic capital [2], but we focus on the arguably most common, value-at-risk measure. The choice of measure and other methodological choices in this thesis aim to tailor the comparison of nonlinear dependence methods in the empirical section to risk analysts in large banks.

In a value-at-risk setting, banks set a certain confidence level, which corresponds to the most unlikely event they want to prepare for. To give an example, the Basel capital framework, which is presented later, requires banks to use a confidence level of 99,9% in regulatory capital calculation. This level implies that the bank would survive 99,9% of all possible scenarios, but would have a 100% − 99,9% = 0,1% probability of going bankrupt. It is not feasible to prepare for everything at a confidence level of 100% as the support of the loss distribution is infinite in most models, implying that the bank could only loan capital that it could lose immediately without going bankrupt. The confidence level that the bank chooses often coincides with the credit rating the bank desires from external agencies for financing because external ratings also imply a probability of default [21]. The terms default and probability of default are elaborated in Section 2.1.2. For example, the weighted long-term average global default rates have been 0,05% for Standard and Poor's rating A and 0,14% for BBB, the next best rating [22]. An institution using a confidence level of 99,9% would thus likely have a rating of A or BBB.

To calculate economic capital using a value-at-risk measure, we introduce the following notation. Let $L$ be a random variable that represents portfolio loss within the next prediction horizon, usually one year, $F_L(l)$ the cumulative distribution function of $L$, and $\alpha \in (0,1)$ be the selected confidence level. Then, we define similarly as McNeil, Frey and Embrechts [3]

$$\mathrm{VaR}_\alpha = \inf\{l \in \mathbb{R} : \mathbb{P}(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_L(l) \geq \alpha\} . \tag{1}$$

$\mathrm{VaR}_\alpha$ can be divided into two parts: expected and unexpected loss. Banks can write the expected credit loss as an expense in accounting according to the IFRS-9 standard, which is the reason why expected and unexpected losses are handled separately. Decomposing the value-at-risk to expected loss (EL) and unexpected loss (UL), we also define economic capital (EC) to mean the unexpected loss, and thus obtain

$$\begin{aligned} & \mathrm{VaR}_\alpha = \mathbb{E}(L) + \mathrm{UL} = \mathrm{EL} + \mathrm{UL} \\ \Leftrightarrow \quad & \mathrm{EC} \equiv \mathrm{UL} = \mathrm{VaR}_\alpha - \mathrm{EL} \end{aligned} \tag{2}$$

This definition is illustrated in Figure 1, with a loss distribution and its values for $\mathrm{VaR}_{99,9\%}$ and EL. Note that the illustration does not consider the profits that are expected from the loans, as they are often calculated separately from credit risk.

### 2.1.2 Parameters

In order to calculate the expected loss (EL) in the formula of the previous section, it is often parametrized into components that are easier to model separately. After the introduction of the Basel capital framework, it has been standard to parametrize expected credit loss to three components, probability of default (PD), loss given default (LGD) and exposure at default (EAD). Then, the maximum likelihood estimates of

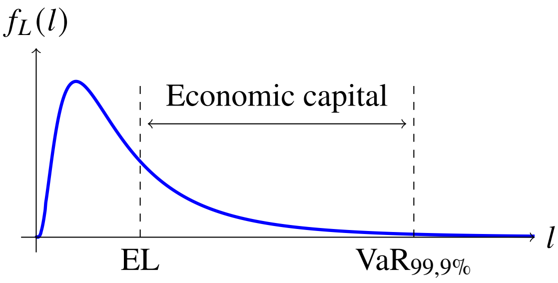



**Figure 1:** Economic capital (EC) is defined to be the difference of $\mathrm{VaR}_\alpha$ and expected loss (EL) of the loss distribution.

each parameter can be multiplied together to obtain a point estimate for the expected loss (EL) of an obligor [23]

$$\mathrm{EL} = \mathrm{PD} \cdot \mathrm{LGD} \cdot \mathrm{EAD} \tag{3}$$

When these parameters are modeled for all obligors, they can be combined to a portfolio model that predicts the loss distribution of the portfolio.

Probability of default (PD) is the likelihood that an obligor (the borrower) does not pay back their loan partly or at all. That event is called a default, and the definition of a default is set in the EU by the European Banking Authority: a default occurs when the obligor has missed payments for more than 90 days and the sum of missed payments is large enough or the obligor is credit-impaired in some other way, such as bankruptcy [21, 24]. When modeling PD, banks have to choose what the modeling horizon is and if they want the model to factor in the economic cycle. If the model estimates vary with the economic cycle, the model is said to be a point-in-time (PiT) estimate. Otherwise, the model is calibrated to be a through-the-cycle (TTC) estimate. A common modeling horizon is one year, which is a good compromise of ease of modeling and usefulness to business. In practice, a PD model could be, e.g. a logistic regression model where training data is the proportion of defaults to defaults for certain customer groups, along with information related to the obligors ability to pay.

Loss given default (LGD) is defined as the ratio between the loss of the bank and the total loan amount (EAD) at the time of default. Essentially, LGD is a percentage of total loan amount, a "loss rate given default". In addition to the outstanding amount, the loss includes direct and indirect expenses from handling the collection of the debt, plus any cash flows that are collected from the obligor after default [23].

LGD also depends on the process of handling defaults. It is heavily affected by the quality and amount of collaterals, costs of handling the default, and possible resolution or recovery plans. For example, the LGD for collateralized exposures, such as housing loans, depends significantly on housing prices and can thus be minimized with efficient collection processes. Therefore, LGD modeling is often segmented by product type, as well as customer type.

Exposure at default (EAD) is an estimate of the loan balance at the time of default. In the Basel framework, it is defined as the drawn loan balance plus the undrawn committed amount multiplied by a credit conversion factor (CCF). The CCF is included because it predicts the probability that the obligor draws more funds right before default, which is possible with some instruments, such as credit cards. The CCFs can be fixed values determined by the type of credit instrument or independently modeled values, depending on the regulatory approach the bank has selected.

These credit risk parameters are important not only because they are needed to calculate the expected loss but also because they are used in defining the capital requirement in regulation.

### 2.1.3 Regulation

Banks are bound by the Basel capital framework that sets rules on how banks have to manage their capital base, including credit risk. Relevant to this thesis and credit risk modeling is that the framework defines formulas for calculating the capital requirement that is supposed to shield the bank against unexpected losses. To comply with the requirement, banks have to calculate *risk-weighted assets* (RWA) and then hold parts of the RWA in different forms of capital. In Basel III, the requirement is, that Common Equity Tier 1 must cover at least 4,5%, Tier 1 capital 6% and total capital 8% of the RWA. The categories overlap but amount to a capital base of 10,5% including a 2,5% buffer on top of the 8% total capital requirement. Some banks are also subject to additional buffers, depending on their capital management process or their systematic importance.

Within the Basel III framework, banks can choose to follow the standardized approach or the internal ratings-based approach. In the standardized approach, the risk weight of each asset is determined by external rating agencies or solely by its category if external ratings are unavailable. For example, a corporate exposure could have a risk weight ranging from 20% to 150% if rated, otherwise 100% [25]. Several challenges have been identified in this approach by Y. Konno and Y. Itoh. For example, rating agencies have different criteria and thus produce different ratings, external ratings react slowly to changing economic circumstances, and few companies have been rated by external agencies [26].

The internal ratings-based (IRB) approach tries to patch the shortcomings of the standardized approach by allowing banks to rate the risk of each asset themselves. In the IRB approach, risk is parametrized by three parameters: probability of default (PD), loss given default (LGD) and exposure at default (EAD). These parameters are modeled by the bank and then risk-weighted assets are calculated by formulas given in the Basel framework. For example, risk-weighted assets for a corporate exposure are calculated as follows:

$$R = 0.12 \cdot \frac{1 - e^{-50 \cdot \mathrm{PD}}}{1 - e^{-50}} + 0.24 \cdot \left(1 - \frac{1 - e^{-50 \cdot \mathrm{PD}}}{1 - e^{-50}}\right) \tag{4}$$

$$b = [0.11852 - 0.05378 \cdot \ln(\mathrm{PD})]^2 \tag{5}$$

$$\mathrm{WCDR} = \Phi\left(\sqrt{\frac{1}{1-R}} \cdot \Phi^{-1}(\mathrm{PD}) + \sqrt{\frac{R}{1-R}} \cdot \Phi^{-1}(0.999)\right) \tag{6}$$

$$K = [\mathrm{WCDR} \cdot \mathrm{LGD} - \mathrm{PD} \cdot \mathrm{LGD}] \cdot \frac{1 + (M - 2.5) \cdot b}{1 - 1.5 \cdot b} \tag{7}$$

$$\mathrm{RWA} = K \cdot 12.5 \cdot \mathrm{EAD}\,, \tag{8}$$

where $R$ is correlation, $b$ is maturity adjustment, WCDR is the so-called *worst-case default rate*, $K$ is capital requirement, $M$ is loan maturity and RWA is the risk-weighted assets [27]. The framework does not explicitly mention the worst-case default rate, but it is a term sometimes used in industry. The intuition is that the worst-case default rate at 99,9% confidence corresponds to the frequency of defaults that would cause the 99,9% value at risk losses of the loss distribution.

The main deficiency of the IRB approach is that it is *portfolio invariant*, meaning that the structure and contents of the rest of the portfolio do not affect the economic capital of a single obligor. This can be thought of as perfect diversification: if the portfolio is diverse with no concentrations, then the added risk of a new loan can be quantified with only the characteristics of the new obligor [28]. However, assuming perfect diversification is often not justified for any, but for the largest global banks for several reasons. First, many smaller banks operate more exclusively in one region, making them susceptible to shocks in the economy of the region or country. Second, specialization drives profits for banks, as they are able to better screen and monitor obligors in the application phase and during the lifetime of the loan [29].

The deficiencies in both approaches of the Basel framework lead to regulatory capital estimates that do not cover the actual risk of the portfolio sufficiently. To combat this, banks develop custom models for more accurate measurement of economic capital by taking into account portfolio concentrations and idiosyncratic risk [10]. The development of these economic capital models is guided by the instructions given by the Basel committee [2]. However, the structure of the models is not strictly regulated as they are not directly used in calculating the capital requirement. They may still be used as part of the internal capital adequacy assessment process (ICAAP) [1]. Ultimately, banks hold at least the amount of capital that is the larger of two figures: the amount mandated by regulation and the one estimated by their economic capital model.

This thesis seeks to provide tools for decision making, especially to those banks which are primarily bound by their economic capital model. In those cases, efforts towards developing an economic capital model will not go to waste because an accurate model is in the interest of both business and risk management. Note though, that the amount of capital the bank actually holds may still be considerably larger than determined using an economic capital model or regulation. Banks usually hold additional buffers to allow some fluctuation in the capital base because adverse consequences follow from capital falling below the regulatory limit, and raising more capital suddenly is hard or at least expensive.

### 2.1.4 Portfolio credit risk

Portfolio credit risk models can be divided into two groups by modeling philosophy: structural models and reduced-form models [3, Ch. 8]. The assumptions of both types of model are important as they are most prominent in the tail of the distribution where observation and data are scarce. The assumptions also affect the possible implementations of the dependence methods covered in this thesis.

In structural models, motivated by Merton [5], the idea is to model the underlying cause of default, which is usually assumed to be the firm's asset value falling below its liabilities. Structural models are more popular due to their interpret-ability in terms of economic variables and use of more commonly available data, compared to reduced form models [7, Ch. 7]. They are often built by implementing one of the major, well documented, portfolio credit risk frameworks, such as CreditMetrics by J.P Morgan, CreditRisk+ by Credit Suisse or Moody's KMV.

The structural portfolio models are further divided into two categories: default-mode models, e.g., CreditRisk+ and KMV, that consider only failure and success, and mark-to-market models such as CreditMetrics that model changes in a loan's market value [30]. They can also be characterized by data source, as others rely on asset value and volatility data, but CreditRisk+ relies mostly on default risk level information.

Reduced form models, on the other hand, are built on the idea that defaults are unpredictable but follow a given process that can be described with some latent variables. The advantage of this approach is that the model has nicer mathematical properties and makes fewer assumptions that are difficult to justify in practice [7, Ch. 7]. On the other hand, the tail of the distribution is then heavily determined by the chosen process. One practical implementation of reduced-form models is Kamakura's Public Firm Model by Kamakura's Risk Information Services (KRIS) [31].

## 2.2 Copulas

Copulas are functions that combine one-dimensional distributions to form a joint multivariate distribution. The word copula comes from Latin, meaning a link, a tie, or a bond. It was first used in a mathematical context by mathematician Abe Sklar in 1956 when he published a theorem that would later become Sklar's theorem [32]. The theorem will be examined carefully in the following chapters, as it is a fundamental definition of copulas.

The motivation behind studying copulas for credit risk modeling is that asset value dependencies, and thus loss dependencies, are the most significant driver of total loss in portfolio credit risk models [3, Ch.1]. Therefore, they are also major contributors to economic capital in banks that use such portfolio credit risk models in determining their economic capital. Low dependence between assets yields diversification benefits, while high dependence amplifies large losses in times of crisis, as large parts of a portfolio might default simultaneously.

Since value-at-risk measures are calculated from the extreme end of a loss distribution, small errors in dependence estimates can lead to large deviations in the final economic capital figure. Accurate modeling and measurement of dependence are therefore key in portfolio credit risk modeling. We introduce nonlinear copulas to model the observed nonlinear phenomena [13, 14] more accurately and to separate the discussion of dependence from the distributions of the individual losses and asset value changes. However, to calibrate nonlinear dependence structures, measures for detecting nonlinear dependence are needed.

### 2.2.1 Measuring dependence

Commonly, dependence between two random variables is called *correlation*. However, the term usually refers to Pearson correlation, which is a measure of linear correlation, basically meaning how well the observations fit onto a line. This type of dependence measure is not always suitable for credit risk modeling because the assumption of linear dependence may not be justified or desired.

Some authors have also found it unreasonable to assume that the correlation of asset values, and therefore defaults, stays constant over time [13, 14]. Instead, they saw increasing correlation in times of crisis, implying that asset values are not linearly dependent. Furthermore, the studies show asymmetric levels of asset value correlation since positive returns show lower levels of correlation than losses in the negative extremes. Not only is linear correlation insufficient for estimating asset price movements, symmetry of the distribution should not be assumed either. Therefore, it is evident that a measure other than linear correlation is needed to identify dependence between assets in a portfolio.

There are measures that do not assume linearity of dependence, most known of which are Spearman's rho ($\rho$) and Kendall's tau ($\tau$). Both of them can measure nonlinear, monotonic, dependence between variables. In practice, this means that both measures indicate perfect positive dependence (correlation of 1) in cases where an increase in one coordinate implies an increase in the other. None of the measures can indicate dependence that is not linear, e.g., second-degree polynomial dependence.

The features of the three dependence measures are summarized in Figure 2. Figure 2a displays a perfect linear dependence, which is indicated by all measures being equal to +1. Figure 2b shows a set of points that have monotonic and nonlinear dependence $y = x^5$. The points do not fit on a line $y = x$, which is why $r < 1$ but others recognize the dependence. Finally, Figure 2c shows symmetric and nonmonotonic dependence that has no correlation. This example highlights the fact that Spearman's rho ($\rho$) is slightly more unstable than Kendall's tau ($\tau$) with small datasets. On the other hand, Spearman's rho ($\rho$) is faster to compute on large datasets.

The correlation of defaults has a significant effect on portfolio credit risk, because low correlation yields diversification benefits. If defaults in a portfolio are uncorrelated, even unlikely events result in low amount of total loss percentage-wise. Moreover, highly correlated defaults mean that many defaults occur simultaneously in some events, resulting in large losses.

### 2.2.2 Defining copulas

The main motivation for using *copulas* is that they make it possible to separate the analysis of marginal distributions and from that of dependence. As established earlier, determining the amount of dependence within a portfolio is already a challenging task, and that is why it is beneficial to perform independently from any features or restrictions of the marginal distributions.

We first examine this topic in two dimensions. Theoretical basis is provided by Sklar's theorem, which is presented, for instance, in Nelsen [32].

**Theorem 2.1** (Sklar's theorem)**.** *Assume random variables $X_1$ and $X_2$. Let $H$ be a joint distribution function with cumulative distribution functions $F_1$ and $F_2$ for the margins such that $F_{X_i}(x_i) = \mathbb{P}(X_i \leq x_i)$. Then there exists a copula $C$ such that for all $x_1, x_2 \in \mathbb{R}$,*

$$H(x_1, x_2) = C(F_{X_1}(x_1), F_{X_2}(x_2)) \tag{9}$$

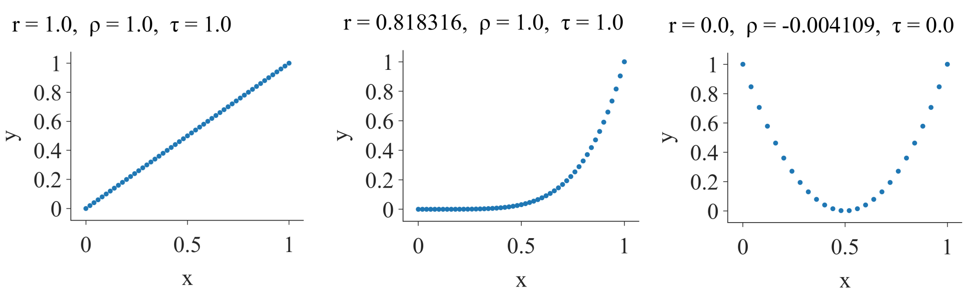
*If $F_{X_1}$ and $F_{X_2}$ are continuous, then $C$ is unique.*

To express Equation (9) in terms of $C$, we replace $X$ with a new random variable $U$, the *grade* of $X$: $U \equiv F_X(X)$.

**Corollary 2.1.1.** *Let $U_i \equiv F_{X_i}(X_i)$. Then, the grades $U_i \sim U_{[0,1]}$ are uniformly distributed in the unit interval and $C$ with marginals $U_i \sim U_{[0,1]}$ is equivalent to $H$ with marginals $X_i$ distributed according to the law $F_{X_i}$.*

*Proof.* Let us first prove that $U_i \sim U_{[0,1]}$.

$$\begin{aligned} F_U(u) &= \mathbb{P}(U \leq u) = \mathbb{P}(F_X(X) \leq u) \\ &= \mathbb{P}(X \leq F_X^{-1}(u)) = F_X(F_X^{-1}(u)) = u \end{aligned} \tag{10}$$




**(a)** Linear, $y = x$

**(b)** Monotonic, $y = x^5$

**(c)** Non-monotonic, $y = 4x^2 - 4x + 1$

**Figure 2:** Pearson correlation $r$ captures only linear dependence fully, while Spearman's $\rho$ and Kendall's $\tau$ also capture nonlinear but monotonic dependence. None of the measures capture non-monotonic dependence.

As $u \in [0,1]$, $F_U(u) = u$ is the uniform distribution in the unit interval [33]. For this and the following proofs, the existence of $F^{-1}$ must be guaranteed. For strictly increasing and continuous $F$ this is straightforward. For other continuous and non-decreasing $F$, a quasi-inverse is constructed, also known as a right-continuous inverse or generalized inverse. The method is described in [32, 34].The quasi-inverse is equal to the ordinary inverse in strictly increasing cases.

Then, using Equation (9), we obtain

$$
\begin{aligned}
C &: [0,1]^2 \mapsto [0,1] \\
H &: \mathbb{R}^2 \mapsto [0,1] \\
C(u_1,u_2) &= C(F_{X_1}(x_1), F_{X_2}(x_2)) && (11) \\
&= H(x_1,x_2) = H(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2)) && (12)
\end{aligned}
$$

□

Sklar's theorem and Corollary 2.1.1 deserve careful inspection as they contain the key to the usefulness of copulas. Firstly, Sklar's theorem shows that it is possible to isolate the dependence information of a joint distribution into one function: the copula. Secondly, Corollary 2.1.1 shows that it is always possible to define a copula from a joint distribution, assuming the joint distribution and the marginal distributions are known. These results imply that it is possible to model a joint distribution by choosing a suitable multivariate function $H$ and margins $F_i$.

Marginal distributions are often easy to select as their empirical distributions are easy to observe from stock market data, for instance. However, the joint distribution function $H$ is more difficult to choose as data of joint movements of corporate asset values might be limitedly available, especially data of extreme scenarios. Portfolio models for credit loss focus on examining the extreme tails of the loss distribution, which is why the joint distribution must be calibrated as well as possible. In that process, being able to swap the copula to a different one while preserving the well-defined marginal distributions is beneficial.

Dividing the modeling process into copula and marginals helps understand both the loss distribution and the overall model. Copula functions use uniformly [0,1]-distributed random variables as inputs, making it easier to visualize and understand the dependence structure of the model. For example, extreme movements in one marginal distribution might mask the dependence structure if one only looks at the final joint distribution given by the model.

It is often necessary to model distributions in more than two dimensions. Sklar's theorem can be extended to multiple dimensions with some effort, that is documented in the original work by Sklar and Schweizer [35]. We present the result and give an example of a multivariate copula.

**Theorem 2.2** (Sklar's theorem in multiple dimensions)**.** *Assume n random variables $X_i$, $1 \leq i \leq n$. Let $H$ be a joint distribution function with margins $F_{X_i}$ such that*

$F_{X_i}(x_i) = \mathbb{P}(X_i \leq x_i)$. *Then there exists a copula $C$ such that for all $x_i \in \mathbb{R}^n$,*

$$H(x_1, x_2, ..., x_n) = C(F_{X_1}(x_1), F_{X_2}(x_2), ..., F_{X_n}(x_n)) \tag{13}$$

$$C(u_1, u_2, ..., u_n) = H(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2), ..., F_{X_n}^{-1}(u_n)) \tag{14}$$

*If $F_{X_i}$ are continuous, then $C$ is unique.*

A basic example of a copula is the Gaussian copula, which is defined as follows. Let $u_i$ be standardized random variables that are correlated with a $n \times n$ correlation matrix $\mathbf{R}$. Then, the corresponding $n$-dimensional normal distribution with correlation matrix $\mathbf{R}$ is written as $\Phi_n(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n); \mathbf{R})$, where $\Phi$ is the univariate standard normal distribution. The copula can now be expressed as

$$C(u_1, ..., u_n) = \Phi_n(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n); \mathbf{R}). \tag{15}$$

Note that here $H$ and $F$ from Equation (14) are represented by the same symbol, but $\Phi_n$ is the multivariate density function that depends on the correlation structure.

### 2.2.3 Types of copulas

Copulas can be split into two categories based on fundamental mathematical properties and method of construction: elliptical and Archimedean copulas. Elliptical copulas are constructed from elliptical symmetric distributions [36]. The most famous example is the Gaussian copula, presented in Equation (15). To understand the definition of elliptical symmetry, we will have to define spherical symmetry because elliptical distributions are affine transformations of spherical distributions.

Spherically distributed random vector $\mathbf{X} \in \mathbb{R}^n$ can be constructed from random variables $\mathbf{U}$ that are uniformly distributed along the unit sphere in $\mathbb{R}^n$ by setting $\mathbf{X} = r\mathbf{U}$, where $r \geq 0$ is an independent random variable. In a way, spherically distributed variables are, therefore, variables that are first uniformly distributed on a sphere surface and spread to different radius values. To obtain an elliptically distributed vector $\mathbf{Y}$, an n-dimensional vector $\mu$ as the center of the ellipse and an $n \times n$ matrix $\Sigma$ are needed to complete the definition

$$\mathbf{Y} = \mu + A'\mathbf{X},$$

where $\mathbf{X}$ is spherically distributed and matrix $A$ is $k \times n$ and satisfies $\Sigma = A'A$ such that $\Sigma$ is symmetric with $\text{rank}(\Sigma) = k$. This step shifts and stretches the spherically distributed variable $\mathbf{X}$ into an ellipse. The symmetry implication here is that the probability density $f_Y$ is equal in all locations that sit on the same "ring" of the ellipse.

In addition to symmetry, another important property is that weighted sums of components of an elliptical distribution are also distributed elliptically [37]. For the Gaussian copula, this property corresponds to linear and summation invariance: linear transformations of normally distributed variables are also normal, as are sums of independent normally distributed variables. This summation property enables the use of factor models, which are presented in the following section.

Archimedean copulas are usually defined with a generator function instead of a certain distribution. Assuming a generator function $\phi(x)$, all Archimedean copulas are constructed as

$$C(u_1, ..., u_n)) = \phi^{-1}(\phi(u_1) + ... + \phi(u_n)). \tag{16}$$

The inverse of the generator $\phi^{-1}$ is required to belong to the family of Laplace transforms, thus most importantly, demanding $\phi^{-1}$ to be infinitely differentiable. The following list contains some well-known Archimedean copulas along with their generator functions [32, Ch.4].

$$\begin{aligned}
&\text{Clayton:} && \phi_\theta(t) = (t^{-\theta} - 1) && C(u_1, ..., u_n) = (u_1^{-\theta} + ... + u_n^{-\theta} - n + 1)^{-1/\theta} \\
&\text{Frank:} && \phi_\theta(t) = -\ln\left(\frac{e^{-\theta t} - 1}{e^{-\theta} - 1}\right) && C(u_1, ..., u_n) \\
& && && \quad = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u_1} - 1)...(e^{-\theta u_n} - 1)}{e^{-\theta} - 1}\right) \\
&\text{Gumbel:} && \phi_\theta(t) = (-\ln t)^{\alpha} && C(u_1, ..., u_n) = e^{-((-\ln u_1)^{\theta} + ... + (-\ln u_n)^{\theta})^{1/\theta}}
\end{aligned}$$

## 2.3 Copulas in portfolio models

Copulas are used in portfolio models to define the dependence between the loss distributions of individual obligors. Given individual losses and dependence information, it is possible to construct the total loss distribution. Ideally, the total loss distribution would be constructed by combining all $n$ loss distributions – one for each asset – of the portfolio using a $n$-dimensional copula. However, this is not possible for any but the smallest of portfolios, as the number of correlation parameters to be estimated grows polynomially with $O(n^2)$, where $n$ is the number of obligors. The rapid increase in the number of parameters makes estimation slower and more uncertain in large portfolios. Typical portfolios might consist of tens of thousands or more obligors, meaning that a full covariance matrix would then have more than 50 million parameters. Therefore, the drawbacks of this approach outweigh the benefits already in quite small banks.

An option to mitigate the number of parameters is to impose some structure on the correlation matrix of the portfolio model. It has been proposed to divide the correlation matrix into blocks of constant correlation by industry, where the diagonal blocks represent correlations within an industry, off-diagonal blocks are intersector correlations, and all correlations within a block are constant [38]. Another option is to use a filtering procedure, which keeps only the eigenvectors that are stronger than something obtained from random matrices [38]. Both options have their advantages: dividing the matrix into blocks greatly reduces the number of parameters, and the filtering procedure was shown to provide results that are more stable and reliable than a sample correlation matrix.

Despite the benefits of matrix adjustment methods, they may not be usable for a corporate portfolio of a large bank. Estimating pairwise correlations is nearly impossible for portfolios that involve significant amounts of privately owned companies.

This is because default events, credit spreads, or equity information is usually not publicly available, which means that dependence information is inferred from the most similar publicly traded companies. However, this level of inaccuracy may be too much to estimate pairwise correlations. Then again, the block matrix method, which fixes the first problem, suffers slightly from the inflexibility of having a constant correlation for all companies within an industry.

To address the challenges of complete correlation matrices, we introduce factor models that make it possible to model correlations between and within an industry (inter- and intra-correlations) while still estimating only one correlation parameter for each company. The correlation between industries also has to be estimated, but the increase in computation time can be assumed constant because the number of industries $K \ll N$ is significantly smaller than the number of companies $N$.

Factor models are commonly used in credit risk modeling [3, 10], and are also referred to as conditionally independent credit risk models. The benefit of factor models is that they "are among the few models that can replicate a realistic correlated default behaviour while still retaining a certain degree of analytical tractability." [39]. In other words, the factor model structure is easily understandable and does not require the same level of knowledge about matrix theory as the simplifications presented above.

### 2.3.1 Single-factor model

The simplest factor model is the single-factor model. It assumes that there is a determining factor, also known as a latent variable, behind the price movements of assets. This could be a valid assumption, as the single factor could represent the overall economic state, also known as the business cycle, which is observed to change periodically and affect companies widely. The business cycle is widely researched and easily observable, which makes the single-factor model an attractive choice, especially when not much data is available about the obligors or assets.

Methodologically, the single-factor model is built on the Merton model [5], which assumes that asset price movements follow a Brownian motion process. This implies that the change in asset prices between time 0 and some time $T$ follows a normal distribution and can thus be modeled.

In the general single-factor model, there is a portfolio of $N$ obligors such that each obligor $n$ has probability $p$ of defaulting before the time horizon $T$ of the model. Changes in asset values $\Delta V_n$ of obligors are assumed to be driven by a $G$-distributed latent variable, the single factor $Y$, and differences in asset prices are assumed to be solely explained by idiosyncratic factors $\epsilon_n$ that are independent and identically $H$-distributed. Then,

$$\Delta V_n = R_n Y + \sqrt{1 - R_n^2}\epsilon_n \ , \tag{17}$$

where $R_n$ is the linear correlation coefficient between assets of corporation $n$ and the latent variable $Y$ [39]. Coefficients $R_n$ are sometimes called *factor loadings*.

**Corollary 2.2.1.** *If $G$ and $H$ are standard normal distributions, then $\Delta V_n$ is standard normal distributed with correlation $R_n$ to $Y$.*

*Proof.* Assuming $G = \mathcal{N}(0,1)$ and $H = \mathcal{N}(0,1)$, we can write

$$\begin{aligned}\Delta V_n = R_n Y + \sqrt{1-R_n^2}\epsilon_n &\sim R_n\mathcal{N}(0,1) + \sqrt{1-R_n^2}\mathcal{N}(0,1)\\ &\sim \mathcal{N}(0,R_n^2) + \mathcal{N}(0,1-R_n^2)\\ &\sim \mathcal{N}(0,1)\ , \end{aligned}\tag{18}$$

which implies that $\Delta V_n$ is standard normal distributed. The correlation between $Y$ and $\Delta V_n$ is

$$\begin{aligned}\rho(\Delta V_n, Y) = \frac{\operatorname{Cov}(\Delta V_n, Y)}{\sigma_{\Delta V_n}\sigma_Y} &= \operatorname{Cov}(\Delta V_n, Y) = \mathbb{E}[(\Delta V_n - \mathbb{E}(\Delta V_n))(Y - \mathbb{E}(Y))]\\ &= \mathbb{E}\left[\left(R_n Y + \sqrt{1-R_n^2}\epsilon_n - 0\right)(Y-0)\right]\\ &= \mathbb{E}\left[R_n Y^2 + \sqrt{1-R_n^2}\epsilon_n Y\right] = R_n\mathbb{E}[Y^2] = R_n. \end{aligned}\tag{19}$$

Therefore, $\Delta V_n \sim \mathcal{N}(0,1)$ with correlation $\rho = R_n$ between $\Delta V_n$ and $Y$. □

Corollary 2.2.1 presents the single-factor model as a way to generate random variables that are correlated to the single factor $Y$ with individual correlation parameters $R_n$. Another way to generate such random variables is to sample each variable from a bivariate distribution that is also correlated with $R_n$. In that case, one dimension is the latent variable $Y$ and the other is $\Delta V_n$. The only restriction is that we need to sample the bivariate distribution given the $Y$-value. In other words, we need a method for sampling the conditional distribution $\mathbb{P}(\Delta V_n|Y = y)$ given the bivariate distribution $\mathbb{P}(\Delta V_n, Y)$ with parameter $R_n$.

Keeping in mind the requirements for the bivariate distribution, it is possible to insert various copulas into the model since it is possible to describe any bivariate distribution with a copula and some marginal distributions. The choice of copula and its dependence parameter value determine the dependence of each corporation's asset values on the single factor, giving the modeler control of model output while remaining computationally efficient.

Although the single-factor model does not explicitly model dependence between obligors, the obligors are correlated via $Y$. If the dependence of obligors on $Y$ can be expressed as a linear combination, the correlation $\rho(\Delta V_i, \Delta V_j)$ can be calculated. As an example, consider the case of Corollary 2.2.1 where $G$ and $H$ are standard normal distributions. Then,

$$\begin{aligned}\rho(\Delta V_i, \Delta V_j) &= \operatorname{Cov}(\Delta V_i, \Delta V_j) = \mathbb{E}[(\Delta V_i - \mathbb{E}(\Delta V_i))\left(\Delta V_j - \mathbb{E}(\Delta V_j)\right)]\\ &= \mathbb{E}\left[(R_i Y + \sqrt{1-R_i^2}\epsilon_i - 0)(R_j Y + \sqrt{1-R_j^2}\epsilon_j - 0)\right]\\ &= R_i R_j\mathbb{E}[Y^2] + R_i\sqrt{1-R_j^2}\mathbb{E}[Y\epsilon_j] + R_j\sqrt{1-R_i^2}\mathbb{E}[y\epsilon_i]\\ &\quad + \sqrt{1-R_i^2}\sqrt{1-R_j^2}\mathbb{E}[\epsilon_i\epsilon_j]\\ &= R_i R_j\mathbb{E}[Y^2] = R_i R_j \end{aligned}\tag{20}$$

However, the correlation between obligors, or the closed form of the implied loss distribution function, can rarely be expressed analytically because only a few choices of $G$ and $H$ give closed-form distributions for $\Delta V_n$ [40]. This is because it is a property of *stable* distributions that a sum of i.i.d random variables shares the distribution of the individual variables [41]. Common stable distributions are the Gaussian, Cauchy, and Levý distributions. For other distributions, correlations and likelihood could be calculated numerically. Either way, the induced correlation between corporations suggests that the single-factor model actually defines a multivariate distribution, where each dimension shares some variance with one of the dimensions.

### 2.3.2 Multi-factor model

The multi-factor model is able to explain asset values with several latent variables instead of one. Regardless of how the factors are chosen, the structure of the multifactor model remains the same. The structure resembles the single-factor model, but instead of loading onto one factor, the assets of each counterparty can load onto several factors in such weights that the total variance remains constant. Schönbucher presents the definition as an extension of the single-factor model [39]. Asset value changes are therefore defined as

$$\Delta V_n = \sum_{j}^{J} \beta_{n,j} Y_j + \epsilon_n \ , \tag{21}$$

where $J$ is the total number of latent variables, $Y_j$ is the $j$th latent variable and $\beta_{n,j}$ are the factor loadings (coefficients of determination) of counterparty $n$ towards latent variable $j$. The weights $\beta$ and variance of the noise components $\epsilon$ are scaled to conserve unit variance.

In the single-factor case, we suggested that a single-factor model can be defined using a bivariate copula that can be sampled conditionally. In the multivariate case, a necessary change is that the model requires a multivariate copula and we would sample the conditional distribution $\mathbb{P}(\Delta V_n | Y_1 = y_1, ..., Y_J = y_J)$.

An intuitive way to assign latent variables for a corporate credit portfolio would be to assess whether some attributes cause simultaneous defaults or, on the other hand, protect some companies from a crisis of other companies. It has been seen that crises are often shared between companies in a given industry and geographical area [4]. Thus, assigning latent variables that represent industries or countries could be a valid option. In such cases, most companies would still load onto only one latent variable, as most companies operate in a single industry or geographical region. For the largest corporations, loading onto multiple latent variables represents reality more closely, as those corporations are no longer as exposed to risk in a particular region or industry.

Latent variables can also be assigned statistically. Principal component analysis (PCA) on asset return data gives a set of principal components that can be used as latent variables. These principal components are orthogonal and usually explain variance in the data well, but the challenge in using them as latent variables is that they may not be easy to interpret as any underlying cause or trait of reality. This is because a latent variable formed by a weighted sum of certain asset returns could seem quite arbitrary or at least be difficult to interpret.

### 2.3.3 Linking to portfolio models

Factor models are used to model the defaults of corporate obligors. From the losses incurred, the portfolio loss distribution is then calculated. Specifically, we determine which companies default on their loans in the modeling horizon $T$ by adding the change in asset value $\Delta V_n$ to the starting asset value $V_n$, and then adding the resulting losses to form the loss distribution. Some portfolio modeling frameworks, for example, CreditMetrics, also include the losses incurred from rating downgrades [42].

Intuitively, corporate credit defaults would be determined in simulation by collecting the amount of liabilities $B_n$ each company has and determining whether $V_n(T) = V_n(0) + \Delta V_n \leq B_n$, indicating that the values of assets at time $T$ have decreased below the default threshold in the modeling horizon. However, in portfolio modeling, it is not necessary to go through the trouble of trying to obtain liability data for companies just to determine defaults. Instead, portfolio models use the default threshold implied by the PD of a company.

The implied default threshold of a factor model is calculated by setting the probability of crossing the default threshold equal to the PD of the obligor. This gives us

$$\mathrm{PD}_n = \mathbb{P}(V_n(0) + \Delta V_n \leq B_n) = \mathbb{P}(\Delta V_n \leq B_n - V_n(0)) = F_{\Delta V_n}(B_n - V_n(0)) \; , \quad (22)$$

where $F_{\Delta V_n}$ is the cumulative distribution of $\Delta V_n$ induced by the factor model. Solving for the threshold for the change in asset value, it is

$$\hat{B}_n = B_n - V_n(0) = F^{-1}_{\Delta V_n}(F_{\Delta V_n}(B_n - V_n(0))) = F^{-1}_{\Delta V_n}(\mathrm{PD}_n) \quad (23)$$

Therefore, a loss distribution can be simulated with a factor model by simulating random values for the asset value changes $\Delta V_n$ and calculating the hypothetical losses from defaults indicated by $\Delta V_n \leq F^{-1}_{\Delta V_n}(\mathrm{PD}_n)$. The benefit of this approach is that there is no need to know how asset values are distributed. In fact, it leaves some room for experimentation: changing the copula of the factor model, and thus the distribution of $\Delta V_n$ has no effect in determining defaults because changing the distribution of $\Delta V_n$ also changes $F^{-1}_{\Delta V_n}$, eliminating the need for additional compensation or customization.

Using Equation (23), it is also possible to calculate thresholds for rating migrations in addition to default. If probabilities for different rating migrations are given, they can be inserted in place of PD in Equation (23) to obtain the threshold corresponding to that migration. This method is used in the CreditMetrics framework when determining losses from rating migrations [42]. In Figure 3, we present an illustration of the method for calculating different migration thresholds, including migration to default. The illustration represents a corporate loan where the initial asset value corresponds to a CCC rating and the simulated asset value $V_n + \Delta V_n$ corresponds to a BBB rating, equaling an increase of three grades. Note that without loss of generality, it is possible to set the expected asset value to zero.Then $\Delta V_n$ can be drawn from a standardized distribution, which means a distribution with mean equal to zero.

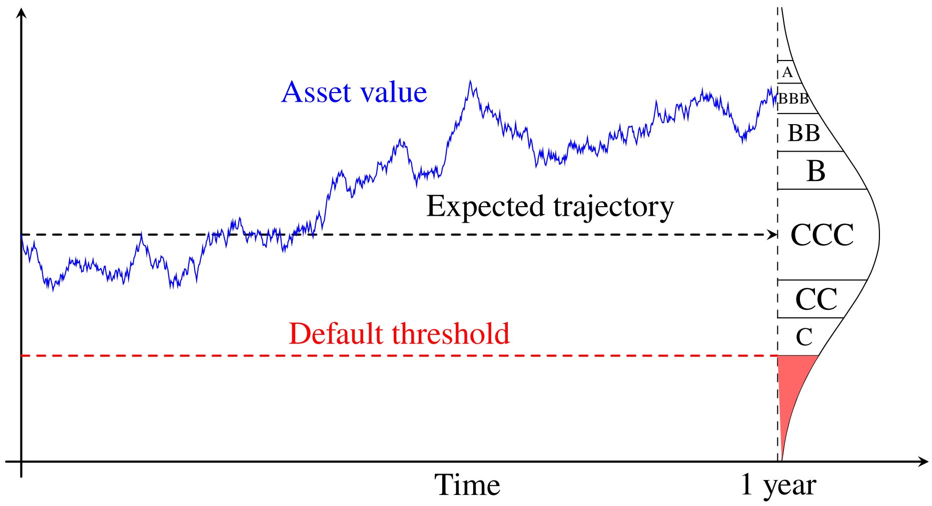



**Figure 3:** The asset value of an obligor increases, upgrading its credit rating from the initial CCC to BBB. Had the asset value decreased and crossed the default threshold, the company would have defaulted on their loan, meaning inability to pay back the loan. The Merton model assumes the rating distribution at the horizon $T = 1$ year is normal, which makes it possible to simulate the default of the company by drawing the change in asset value $\Delta V_n$ from a standard normal distribution.

# 3 Research material and methods

In this section, we present the nonlinear dependence structures that are compared in this thesis. Historically, the Gaussian copula has been a popular dependence structure to use in portfolio credit risk modeling [11] and it is still present in the Basel III framework, as seen in Equation (6), which is the reason for comparing the alternatives against it and each other. Since the Gaussian copula does not include nonlinear dependence, we do not, however, consider it one of the options. Instead, we compare the Student's t copula and the Gumbel copula against each other as a replacement for the Gaussian copula. The purpose of the comparisons in this thesis is not only to compare the copulas themselves, but also to compare *the effects of choosing* each copula for a portfolio model. The copulas may require modifications to the modeling frameworks, different assumptions, or other customizations, which contribute to the final choice of copula for a portfolio credit risk model.

Comparisons of the above copulas have previously been made [8], but the comparisons have focused on the accuracy of the estimates and the fit of the model. Instead of focusing on numerical accuracy, this thesis contributes to the literature by providing qualitative tools for other aspects of decision making, with the aim of helping risk analysts choose which dependence structure to implement in a portfolio credit risk model for corporate exposures. Towards this end, we consider criteria related to theoretical justification, the complexity of the implementations, and the explainability of the dependence structures. We also build on previous work on comparisons of the accuracy of the different alternatives.

The comparison criteria are based on factors that were found important when selecting the modeling methodologies for OP Financial Group. Theoretical justification is examined because while it is important to prepare for risks, it is harmful for profitability to be unreasonably careful. Therefore, all methodological choices must be justified by the risks they cover. Explainability is assessed because it increases a model's significance in a business environment: the users have limited time to spend understanding the model, but they need to trust the model and its results in order for the model to get used in processes such as capital adequacy assessment and evaluating riskiness of different business units. Finally, complexity of the implementation is analyzed because if an otherwise good model is difficult to build or maintain, it detracts attention from efforts to develop and analyze other credit risk models, reducing overall risk awareness of the bank. For instance, finding and fixing errors takes more effort and calculation consumes more resources for a complex model rather than a simple one.

We expand on Section 2 by presenting in detail the copulas to be compared and how they can be implemented into an economic capital model of a large bank with corporate credit portfolios. Finally, we compare the selected copulas using qualitative criteria that are motivated by the needs of such banks, such as OP Financial Group. Further contributions could apply similar qualitative criteria to different options to make comparison of dependence structures easier.

## 3.1 Scope

The comparisons in this thesis are limited by choices that tailor the results to a specific target audience. In detail, the results are targeted to risk analysts working with economic capital models for large corporate credit portfolios that use the value-at-risk measure in their economic capital estimates. The characteristics of this audience make some options out of question, and some more interesting than others. For example, the large size of the corporate loan portfolio makes it necessary to use a dimensionality reduction technique, which in this case is the factor model framework. The main limitations and choices of the thesis are covered in detail below. Overall, limitations in dependence structure options and comparison criteria are explained by the large portfolio size, the former popularity of the Gaussian copula, and the challenges in high-dimensional integration.

### 3.1.1 Factor model

In this thesis, the dependence structures to be compared are limited to those that can be implemented as copulas within a multifactor credit risk model. Firstly, they are required to be presented as copulas because the Gaussian copula is often the point of reference for other options due to previous popularity [11] and presence in the Basel III framework [27], see Equation (6). Secondly, they are required to be implementable within a multifactor credit risk model for performance reasons that are explained in more detail in Section 2.3. Otherwise, running the model would be computationally too expensive or it would not include the diversification that is present in reality.

The corporate credit portfolios of large banks contain thousands of obligors and are often diversified between some segments, for example, geographical regions or industries, which are heterogeneous in terms of default correlation [4]. Therefore, it is justified to require these types of diversification benefits in the model. Hence, it is required that the copulas can be implemented in a portfolio model with $K$ subgroups, represented by the latent variables $(Y_1, ..., Y_K)$ of the model. We choose to set the latent variables to represent industries, and for simplicity, assume that each obligor belongs only to one industry. Therefore, each obligor $n$ belongs to a group $k \in 1, ..., K$, which implies that changes in its asset prices are explained by a latent variable $Y_{k(n)}$ and an independent component $\epsilon_n$.

### 3.1.2 Implementation in a simulation model

The implementation complexity of the dependence structures is evaluated by implementing the structures in a Monte Carlo simulation model and comparing the complexity of the resulting simulation algorithms. It would be interesting to compare the implied multivariate distributions of the copulas along with an analytical solution for the VaR level, but performing the calculation would be computationally too demanding. Integrating the $N$-dimensional distribution, where $N$ is the number of obligors in the corporate portfolio, is already a challenging task, not to mention that it may not be possible to explicitly formulate the loss distribution in the first place. Limiting the comparison of implementations to simulation models is not seen as a drawback, as there are benefits to solving the VaR-level using simulation. For example, it is easy to add dynamic features that affect only certain rounds of the simulation.

## 3.2 Dependence structures

In this section, we present the mathematical definition of the copulas to be compared and how they are applied to a portfolio credit risk model that complies with the limitations presented in the previous section. To follow the limitations, we implement the model in the following structure, which is also illustrated in Figure 4:

1. A $K$-dimensional copula defines the relationship of the $K$ latent variables, where the variables represent some subgroups, for example industries. In other words, we define the distribution of the latent variables $(Y_1, ..., Y_K)$ in terms of a copula $C(u_1, ..., u_K)$.

2. $K$ single factor models are used to connect each obligor to its latent variable. This structure can also be thought of as a multifactor model where each obligor has only one non-zero loading coefficient $\beta$.

### 3.2.1 Gaussian copula - the baseline

The Gaussian copula has been the standard dependence structure in portfolio modeling and is therefore presented as a baseline for comparison. Let $u_i$ be uniform [0,1]-distributed random variables whose correlation is defined by a $n \times n$ correlation matrix $\mathbf{R}$. Individual correlations $R_{i,j}$ are Pearson correlation coefficients, but they can also be expressed using Kendall's tau coefficients as $R_{i,j} = \sin(\pi\tau_{i,j}/2)$. Then, $\Phi_n(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n); \mathbf{R})$ is the corresponding $n$-dimensional normal distribution with correlation matrix $\mathbf{R}$, and $\Phi$ is the univariate standard normal distribution. Then, the Gaussian copula has the following form [43, Ch. 6]

$$C(u_1, ..., u_n) = \Phi_n(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n); \mathbf{R}). \tag{24}$$

An illustration of the bivariate Gaussian copula is provided in Figure 5.

Using the definition, the desired portfolio model is built as follows:

1. Assume $K$ subgroups in the portfolio and let $Y_k$ be the latent variables for each group $k \in 1, ..., K$. Set the variables to be distributed according to the Gaussian copula by letting $Y_k \sim U(0, 1)$ and

$$\begin{aligned}\mathbb{P}(Y_1 \leq u_1, ..., Y_K \leq u_K) &= C_{\text{Gauss}}(u_1, u_2, ..., u_K) \\ &= \Phi_K(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_K); \mathbf{R}).\end{aligned} \tag{25}$$

Consider also an alternative definition, where we have $Y_k^{'} \sim \mathcal{N}(0, 1)$. Then,

$$\begin{aligned}\mathbb{P}(Y_1^{'} \leq n_1, ..., Y_K^{'} \leq n_K) &= \mathbb{P}(\Phi(Y_1^{'}) \leq \Phi(n_1), ..., \Phi(Y_K^{'}) \leq \Phi(n_K)) \\ &= \mathbb{P}(Y_1 \leq \Phi(n_1), ..., Y_K \leq \Phi(n_K)) \\ &= \Phi_K(n_1, ..., n_K); \mathbf{R})\ ,\end{aligned} \tag{26}$$

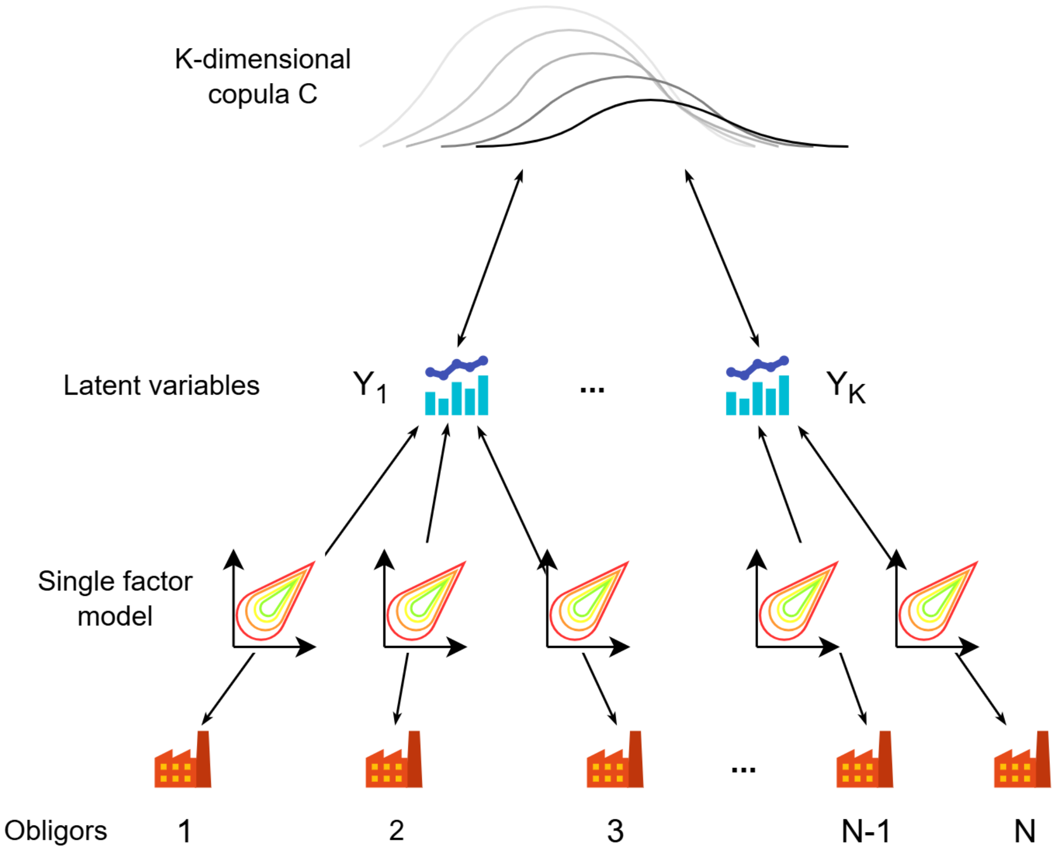
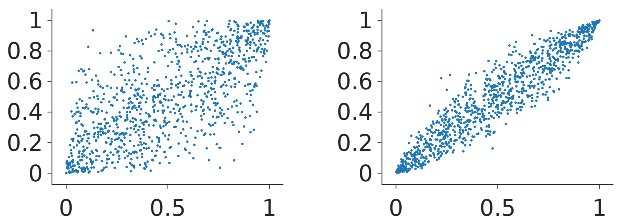



**Figure 4:** The dependence structures are implemented in two layers where the dependence of the latent variables is defined by copula $C$, and then the asset value changes of each obligor are calculated with a single factor model that connects each obligor with its industry, represented by a latent variable.




**(a)** $\tau = 0.5$ **(b)** $\tau = 0.8$

**Figure 5:** The bivariate Gaussian copula is illustrated by sampling 1000 samples from two Gaussian copulas with different dependence parameters. The Kendall's tau ($\tau$) values of 0.50 and 0.80 correspond to correlations of 0.70 and 0.95, respectively.

implying that $(Y_1^{'}, ..., Y_K^{'}) \sim \mathcal{N}_K(0, 1; \mathbf{R})$ and making it possible to sample $Y_k$ directly from the multivariate standard normal distribution that they follow. We use the definition in Equation (26) because the assumption $Y_k \sim \mathcal{N}(0, 1)$ helps determine which obligors default on their loan.

2. Set the asset value changes of each obligor $n$ to depend on the corresponding latent variable $Y_{k(n)}$ model with the Gaussian copula. This creates a bivariate relationship, as described in Section 2.3.1. To create a bivariate dependence between $Y_{k(n)}$ and $\Delta V_n$, Equation (26) can be applied in two dimensions.

   Assume $\Delta V_n \sim \mathcal{N}(0, 1)$ and $Y_k \sim \mathcal{N}(0, 1)$. Then, the dependence between the variables is

$$\mathbb{P}(Y_{k(n)} \leq n_1, \Delta V_n \leq n_2) = \Phi_2(n_1, n_2; R) \tag{27}$$
$$\Rightarrow (Y_{k(n)}, \Delta V_n) \sim \mathcal{N}_2(0, 1; R) \tag{28}$$

   As proven in Corollary 2.2.1, the relationship $\mathcal{N}_2(0, 1; R)$ between $Y_{k(n)}$ and $\Delta V_n$ is obtained with the single-factor definition

$$\Delta V_n = R_n Y_{k(n)} + \sqrt{1 - R_n^2}\epsilon_n = R_n Y_{k(n)} + \sqrt{1 - R_n^2}\epsilon_n \tag{29}$$

   given $\epsilon \sim \mathcal{N}(0, 1)$.

### 3.2.2 Student's t copula

Student's t copula has been found more accurate in modeling stock returns than the Gaussian copula due to heavier tails of the distribution [8, 15, 44]. The tail dependence of the t-copula is controlled with a single parameter $\nu$, called degrees of freedom. This simplicity makes the t-copula an attractive choice for adding nonlinear dependence to a portfolio model.

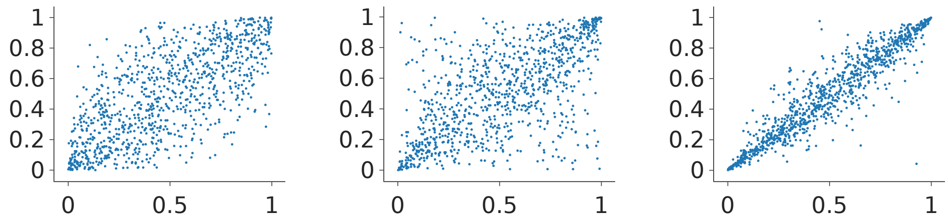
Conventionally, Student's t distribution has location $\mu = 0$ and scale $\sigma = 1$, thus generalizing the standard normal distribution. When $\mu$ and $\sigma$ are given other values, the distribution is called location-scale t distribution $\text{lst}(\mu, \sigma, \nu)$. At the limit $\nu \to \infty$, Student's t distribution and location-scale t distribution approach the standard normal and normal distributions, respectively. This is visible, for example, from the variance of the location-scale t distribution, which has variance $\text{var} = \sigma^2 \frac{\nu}{\nu-2}$.

To define the Student's t copula, let $T$ be the cumulative distribution function of the Student's t distribution with correlation matrix $\mathbf{R}$ and degrees of freedom $\nu$. Then, the Student's t copula has the following form [44, Ch. 4].

$$C_t(u_1, ..., u_n) = T_n(T^{-1}(u_1; \nu), ..., T^{-1}(u_n; \nu); \mathbf{R}, \nu). \tag{30}$$

An illustration of the bivariate Student's t copula is provided in Figure 6.

We construct the portfolio model using this definition:




**(a)** $\tau = 0.5, \nu = 2000$ **(b)** $\tau = 0.8, \nu = 3$ **(c)** $\tau = 0.8, \nu = 3$

**Figure 6:** The bivariate Student's t copula is illustrated by plotting 1000 samples from three Student's t copulas with different dependence parameters. The Kendall's tau ($\tau$) values of 0.50 and 0.80 correspond to correlations of 0.70 and 0.95, respectively. The degrees of freedom parameter ($\nu$) value of 2000 is very close to a normal distribution, while the value of 3 is a very high level of tail dependence.

1. Assume $K$ subgroups in the portfolio and let $Y_k$ be the latent variables for each group $k \in 1, ..., K$. Set the variables to be dependent according to the t-copula by letting $Y_k \sim U(0, 1)$ and

$$\begin{aligned}\mathbb{P}(Y_1 \leq u_1, ..., Y_K \leq u_K) &= C_{\mathrm{t}}(u_1, u_2, ..., u_K) \\ &= T_n(T^{-1}(u_1; \nu), ..., T^{-1}(u_n; \nu); \mathbf{R}, \nu) \end{aligned} \tag{31}$$

As in the Gaussian copula, we can also set $Y_k' \sim t(\nu)$ to obtain the latent variables directly from a multivariate t distribution.

$$\mathbb{P}(Y_1' \leq s_1, ..., Y_K' \leq s_K) = T_n(s_1, ..., s_n; \mathbf{R}, \nu) \tag{32}$$

$$\Rightarrow (Y_1', ..., Y_K') \sim t_K(\mathbf{R}, \nu) \tag{33}$$

2. Set the asset value changes of each obligor $n$ to depend on the corresponding latent variables $Y_{k(n)}$ using the t-copula. This defines a bivariate relationship that is a two-dimensional version of Equation (33).

Let $Y_k \sim t(\nu)$ and $\Delta V_n \sim t(\nu)$ for $k \in 1, ..., K$ and $n \in 1, ..., N$, with $\nu$ being the degree-of-freedom parameter and $R_{n,k}$ quantifying the correlation of $Y_{k(n)}$ and $\Delta V_n$. Then,

$$\mathbb{P}(Y_{k(n)} \leq s_1, \Delta V_n \leq s_2) = T_2(s_1, s_2; R_{n,k}, \nu) \tag{34}$$

$$\Rightarrow (Y_{k(n)}, \Delta V_n) \sim t_2(\mathbf{R}, \nu) \tag{35}$$

### 3.2.3 Hierarchical Archimedean (Gumbel copula)

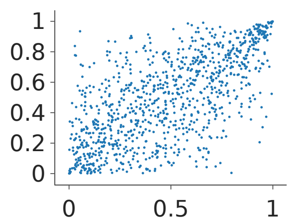
Archimedean copulas have been advocated over the Gaussian copula due to their ability to capture extremal dependence [9]. The benefit of Archimedean copulas is that they share many mathematical properties because they are defined with a generator function, as presented in Section 2.2.3. That is why Archimedean copulas are often presented together, instead of individually respective to the generator functions.

The drawback to the generator function property is that Archimedean copulas are parametrized by a single dependence parameter of the generator function, which is why there is a need to either use the same level of dependence for all dimensions of the copula or to build a more complex construction using several copulas. In this section, we present the Gumbel-Hougaard copula, which we will refer to as the Gumbel copula, and a hierarchical implementation [45] of several Gumbel copulas. Another alternative to increasing the flexibility of the dependence model is to build a pair-copula construction, also called a vine copula [44, 46].

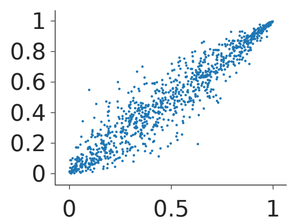
The multivariate Gumbel copula is defined using the generator $\phi(t;\theta) = (-\ln(t))^{\theta}$. This gives the $n$-dimensional Gumbel copula for $\theta \geq 1$ and any $n \geq 2$ [32, Ch. 4].

$$C_{\text{Gu}}(u_1, ..., u_n; \theta) = e^{-((-\ln u_1)^{\theta} + ... + (-\ln u_n)^{\theta})^{1/\theta}} \tag{36}$$

The dependence parameter $\theta$ in the Gumbel copula is defined using Kendall's tau as $\theta = 1/(1-\tau)$, where $\tau$ is the Kendall's tau coefficient estimated for a pair of variables $u_i, u_j$ [47]. Evidently, the same value of $\tau$ has to quantify the dependencies between all pairs of variables because the dependencies are expressed using only a single parameter in the Gumbel copula. Figure 7 illustrates the bivariate Gumbel copula.




**(a)** $\tau = 0.5$




**(b)** $\tau = 0.8$

**Figure 7:** The bivariate Gumbel copula is illustrated by sampling 1000 samples from two Gumbel copulas with different dependence parameters. The Kendall's tau ($\tau$) values of 0.50 and 0.80 correspond to $\theta$ parameter values of 2 and 5, respectively.

Combining many (Gumbel) copulas into a hierarchical one gives possibilities for assigning several $\theta$ parameters, in order to model more complex dependencies. In practice, copulas are inserted as arguments into other copulas to create a structure that can be modeled as a tree. A simple example of a hierarchical copula is the fully nested copula using bivariate copulas, that is illustrated in Figure 8 and used by Kole, Koedijk and Verbeek [8]. By definition, a $k$-variate fully nested copula has $k-1$ levels of hierarchy. It is formulated recursively as

$$C_{HA}(u_1, \ldots, u_k; \mathbf{\Theta}) = \begin{cases} C_{\text{Gu}}(u_1, u_2; \theta_1) & k = 2 \\ C_{\text{Gu}}[C_{HA}(u_1, ..., u_{k-1}; \mathbf{\Theta}_{[1,k-2]}), u_k; \theta_{k-1}] & k \geq 2 \end{cases} \tag{37}$$

$$\Rightarrow$$

$$C_{HA}(u_1, \ldots, u_k; \mathbf{\Theta}) = C_{\text{Gu}}(\ldots C_{\text{Gu}}(C_{\text{Gu}}(u_1, u_2; \theta_1), u_3; \theta_2) \ldots), u_k; \theta_{k-1}),$$

where $\mathbf{\Theta}$ refers to the full vector of dependence parameters $(\theta_1, ..., \theta_k)$ and $\mathbf{\Theta}_{[1,s]} = (\theta_1, ..., \theta_s)$ to elements in indices 1 to $s$, $s < k$ of the full vector. The variables are ordered such that $\theta_1 \geq \theta_2 \geq ... \geq \theta_{k-1}$.

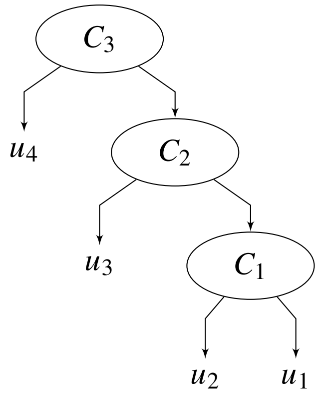



**Figure 8:** A fully nested $(k = 4)$ copula has four variables $u_i$ and therefore $(k - 1 = 3)$ levels of hierarchy, represented by copulas $C_i$. Dependence parameters $\theta_i$ of copulas $C_i$ must be ordered such that $\theta_1 \geq \theta_2 \geq \theta_3$, implying that the strongest dependence is between variables $u_1$ and $u_2$.

More complex hierarchical copulas can involve copulas with more than two dimensions, and therefore less than $k - 1$ levels in their hierarchy, given that the final copula has $k$ dimensions. An example by Joe is provided [44, Ch. 3.4]. Let

$$C_{HA}(u_1, u_2, u_3, u_4, u_5; \mathbf{\Theta}) = C_{\text{Gu}}(C_{\text{Gu}}(u_1, u_2; \theta_1), C_{\text{Gu}}(u_3, u_4, u_5; \theta_2); \theta_3) \; , \tag{38}$$

which is then the hierarchical copula in Figure 9 consisting of three-dimensional and two-dimensional Gumbel copulas combined with a third, two-dimensional, Gumbel copula. Hierarchical constructions like this could be attractive in cases where several variables share similar dependence characteristics, while a fully nested model could be more useful when the dependence strengths of all dimensions are different.

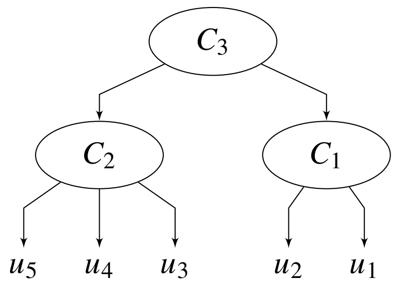



**Figure 9:** The two-dimensional copula $C_1$ and the three-dimensional copula $C_2$ are nested in a third copula $C_3$. The dependence parameters $\theta_i$ for $C_i$ must be such that $\theta_1 \geq \theta_2 \geq \theta_3$, meaning that strongest dependence is captured by the smallest inner copula and weakest dependence is captured by the root copula.

A portfolio model using hierarchical Gumbel copulas would therefore be constructed as follows.

1. Assume $K$ subgroups in the portfolio and let $Y_k$ be the latent variables for each group $k \in 1, ..., K$. Define a hierarchical Gumbel copula $C_{\text{Gu}}(u_1, ..., u_k; \mathbf{\Theta})$ with $\mathbf{\Theta} = (\theta_1, ...\theta_h)$ being the required dependence parameters, such that

$$\mathbb{P}(Y_1 \leq u_1, ..., Y_k \leq u_k) = C_{\text{Gu}}(u_1, ..., u_k; \mathbf{\Theta}). \tag{39}$$

2. Set the asset value changes of each obligor $n$ to depend on the corresponding latent variable $Y_{k(n)}$ using the bivariate Gumbel copula. Let $Y_k \sim U(0, 1)$ and $\Delta V_n \sim U(0, 1)$ for $k \in 1, ..., K$ and $n \in 1, ..., N$. Quantify the dependence between $\Delta V_n$ and its respective group $Y_k$ with parameter $\theta_n$. Then,

$$\mathbb{P}(\Delta V_n \leq u_1, Y_{k(n)} \leq u_2) = C_{\text{Gu}}(Y_{k(n)}, \Delta V_n; \theta_n). \tag{40}$$

   To calculate $\Delta V_n$ in practice and determine defaults in simulation, we use the conditional probability form

$$\mathbb{P}(\Delta V_n \leq u_1 | Y_{k(n)} = y) = C_{\text{Gu}}(\Delta V_n | (Y_{k(n)} = y; R_n). \tag{41}$$

   since the realizations of $Y_k$ from step 1 would be available in simulation. Writing the previous equation as a sum of random variables similarly to the Gaussian case is not available for Archimedean copulas because the equation relies on the summation property of stable distributions.

## 3.3 Comparison criteria

The copulas are compared using three qualitative criteria that aim to capture possible challenges in the development of a portfolio credit risk model. The criteria do not provide absolute scales or metrics, but instead are used to order the copula options so that a credit risk modeler can get a summary of the properties of the copulas at a glance. This goal is facilitated by summary tables that are presented in Section 4. The criteria are limited similarly to the copulas themselves, namely by the constraints listed in Section 3.1.

### 3.3.1 Theoretical justification

Copulas chosen in portfolio models must be theoretically justified to make them worth investigating in the first place. We examine how well the properties of each copula suit the modeling of observed asset value behavior. Specifically, we compare the features of the copulas with the observations made by Longin and Solnik [13], and Bae et. al.[14]:

1. Large losses occur more frequently together than smaller losses. This is also known as default clustering.

2. The dependence of large positive returns is different from the dependence of large negative returns.

In order to justify a copula other than the Gaussian copula, there should be some evidence or justification that the copula being evaluated would depict reality better than the Gaussian copula or the independence copula. Otherwise, the added complexity can be rejected by Occam's razor principle.

The ability to depict reality is assessed using the coefficient of tail dependence that is calculated for each copula. Specifically, we compare the upper and lower coefficients of tail dependence, which quantify the level of dependence in the positive and negative extremes of the copulas [3, Ch. 5]. For random variables $X_1$ and $X_2$ with continuous cumulative distribution functions $F_1$ and $F_2$, respectively, and dependence defined by copula $C$, the coefficient of lower tail dependence $\lambda_u$ is defined as

$$\begin{aligned}\lambda_l &= \lim_{q\to 0^+} \mathbb{P}(X_2 \leq F^{-1}(q)|X_1 \leq F_1^{-1}(q)) \\ &= \lim_{q\to 0^+} \frac{\mathbb{P}(X_2 \leq F^{-1}(q), X_1 \leq F_1^{-1}(q))}{\mathbb{P}(X_1 \leq F_1^{-1}(q))} \\ &= \lim_{q\to 0^+} \frac{C(q,q)}{q} \end{aligned} \tag{42}$$

assuming the limit exists [3, pp. 209]. Intuitively, the coefficient indicates the correlation of the values of $X_1$ and $X_2$ that are below the quantile $q$, while $q$ approaches the lower tail. The coefficient of upper tail dependence $\lambda_u$ given by

$$\lambda_u = \lim_{q\to 0^+} \frac{\hat{C}(q,q)}{q}$$

is defined similarly [3, pp. 209], but using the survival copula $\hat{C}(1-u, 1-v) = 1-u-v+C(u,v)$, which corresponds to the copula $C$, where the x and y axes have been flipped around.

The observations by Longin and Solnik [13] and Bae et. al.[14] may now be written more precisely in terms of the coefficient of tail dependence.

1. The coefficient of upper tail dependence does not equal zero, i.e. the losses are not asymptotically independent.

2. The coefficient of upper tail dependence is positive, while the coefficient of lower dependence is zero, i.e. the losses show extremal dependence, but "profits" are asymptotically independent.

### 3.3.2 Implementation in a portfolio model

We compare the implementations of the models by analyzing the length and complexity of the simulation algorithm that is required to implement the desired copula following the limitations established in earlier sections.

The models are assumed to be calibrated using equity price data from stock markets, as other data types are difficult to obtain [39]. As we focus on the perspective of large banks and their corporate portfolios, it is valid to assume the existence of such equity data, since the corporate portfolios would likely consist, in large parts, of companies that are listed in the stock market.

When assessing the implementation in a portfolio model, we write a pseudo-code that specializes the following steps to the copula in question. The bold parts of the code are specific to each copula.

For a large number of repetitions:

1. **Sample random values for each latent variable $Y_k$ in the portfolio.** We set these latent variables to be industries of the obligors, but they could also be something else, as described earlier.
2. **Calculate the asset value change $\Delta V_n$ for each company in the portfolio using independent random variables.**
3. Determine which companies default on their loan by comparing their asset value changes $\Delta V_n$ to the corresponding **default threshold** $F_{V_n}^{-1}$ of the same company. If a company defaults, add it's EAD*LGD to a variable that counts the total loss of that simulation round.
4. Save the total loss and information of which companies defaulted in order to distribute the economic capital among the obligors.

The simulation is run for many rounds, in the order of hundreds of thousands to millions, to keep the Monte Carlo error low. The required number of simulation rounds depends on the size of the portfolio.

### 3.3.3 Explainability

Finally, we compare how well the selected models and their implementations can be explained to model users and other stakeholders. Since explainability may be interpreted in different ways, we evaluate the explainability using three different categories defined by P. Mishra [48], presented in Table 1.

**Table 1:** Methods for explaining machine learning models can be grouped into three categories [48].

| Category | Method example |
|---|---|
| Textual Explainability | Natural Language Generation |
| | Summary Generation |
| Visual Explainability | Tree Based Flow Chart |
| | Rule Extraction |
| Example Based | Using Common Examples |
| | Business Scenarios |

Textual explainability is evaluated by creating textual descriptions of the mathematics involved in each copula implementation. Visual explainability is determined by drawing diagrams of each copula implementation and seeing which copula has the simplest one. Finally, example based explainability is evaluated by writing a hypothetical scenario that is implied by each copula. The scenarios are then compared by how easy they are to understand.

# 4 Results

In this section, the selected copulas are compared using the criteria described in Section 3. For each criterion, the copulas are ordered from best to worst with justification. Since the assessment is qualitative, the performance of two copulas may be declared equal if their performance is similar. A summary of the comparison results is in Section 4.4.

## 4.1 Theoretical justification

The justifiability of each copula is determined by tabulating their upper and lower tail dependence coefficients and comparing them with the two desired characteristics established in Section 3.3.1.

1. The coefficient of upper tail dependence does not equal zero, i.e. the losses are not asymptotically independent.

2. The coefficient of upper tail dependence is positive, while the coefficient of lower dependence is zero, i.e. the losses show extremal dependence, but "profits" are asymptotically independent.

The upper and lower tail dependence coefficients of each copula are in Table 2. The table shows that the Gaussian copula is asymptotically independent in both tails, while the Gumbel copula displays extremal dependence in the upper tail and the Student's t copula in both tails symmetrically. Therefore, the Student's t copula fulfills the first criterion, the Gumbel copula fulfills both, and the Gaussian copula fulfills neither. The more important characteristic is the first, which demands extremal dependence in the upper tail. It is significant because a lack of dependence in the upper tail can cause a serious underestimation of risk, as happened during the financial crisis of 2007-2008 [11, 12]. As long as the model can capture the upper tail dependence, it is a major improvement to the previously common but insufficient methodology.

The second characteristic is more related to the calibration of the models. Crises are scarce in stock market data, which means that the true extent of the upper tail

**Table 2:** Values for the coefficients of upper and lower tail dependence for each copula under comparison. The values are calculated for bivariate copulas, where the Gaussian copula has dependence parameter $\rho$, the Student's t copula has dependence parameter $\rho$ and degrees-of-freedom parameter $\nu$, and the Gumbel copula has dependence parameter $\theta$. Here $t_{\nu+1}$ is the univariate cdf of the Student's t distribution with $\nu + 1$ degrees of freedom, $\nu$ being the degrees-of-freedom of the copula.

| Copula | $\lambda_l$ | $\lambda_u$ |
|---|---|---|
| Gaussian | 0 | 0 |
| Student's t | $2t_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{(1+\rho)}}\right)$ | $2t_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{(1+\rho)}}\right)$ |
| Gumbel | 0 | $2 - 2^{1/\theta}$ |

of losses may not be present in the data. Therefore, symmetrical models, such as the Student's t copula, have their tail dependence parameters calibrated mostly to the lower tail (low losses, bull market), which has a lower level of dependence than the upper tail, leading to the underestimation of extreme risks that would be present in the upper tail of the loss distribution.

From the formulas in Table 2, we can also note that the Student's t copula has two parameters controlling the tail dependence coefficient, while the Gumbel copula only has one. Having two parameters is advantageous because the tail dependence strength of the copula can be controlled while keeping the estimated correlation parameter constant. A study that finds the Student's t copula more accurate than the Gumbel copula concludes that the additional parameter is a driver of the better performance [8].

Given the properties listed in Table 2, the Gumbel copula is the best in modeling observed phenomena. However, the additional parameter of the Student's t copula compared to the Gumbel compula makes it a valid contender. We conclude that the Gumbel copula is theoretically more justified, but empirical accuracy needs to be assessed when calibrating the selected model.

## 4.2 Implementation

Pseudo-code implementations of each copula in an economic capital simulation model are presented below. Each implementation follows the format presented in Section 3.3.2. The implementations require sampling random numbers, which is assumed to be available in the simulation environment. Specifically, we assume the ability to sample the Normal, Student's t, inverse Gamma, Exponential, and Stable distributions.

### 4.2.1 Gaussian copula

1. Sample random values for each latent variable $Y_k$ in the model.

   As established in Section 3.2.1, it is possible to obtain standard normal distributed variables $(Y_1, ..., Y_K)$ from a multivariate standard normal distribution $\mathcal{N}_K(0, 1; \mathbf{R})$. Sampling can be performed using Cholesky decomposition [44, chapter 6].

   (a) Find the $K \times K$ lower triangular Cholesky matrix $A$ such that $AA^\intercal = \mathbf{R}$.

   (b) Sample a vector $\mathbf{Z}^\perp$ of $K$ i.i.d. $\mathcal{N}(0, 1)$ variables.

   (c) $(Y_1, ..., Y_K) = A\mathbf{Z}^\perp$.

2. For each obligor $n$, calculate $\Delta V_n$ given the sample $(Y_1, ..., Y_K) = (y_1, ..., y_K)$. Given the bivariate relationship $(Y_{k(n)}, \Delta V_n) \sim \mathcal{N}_2(0, 1; R)$ in Section 3.2.1, it is possible to sample $\Delta V_n$ similarly to in the first step, using the Cholesky decomposition.

(a) Calculate the Cholesky lower triangular matrix

$$A = \begin{bmatrix} 1 & 0 \\ R_{n,k} & \sqrt{1 - R_{n,k}^2} \end{bmatrix},$$

where $R_{n,k}$ is the correlation between $\Delta V_n$ and the corresponding latent variable $Y_k$.

(b) For each obligor $n$, sample i.i.d. $\epsilon_n \sim \mathcal{N}(0, 1)$. Then, we have two independent $\mathcal{N}(0, 1)$ variables for each obligor, $\epsilon_n$ and $y_{k(n)}$, and we set $\mathbf{Z}^{\perp} = [y_{k_n}, \epsilon_n]^{\intercal}$

(c) For each obligor, calculate the realization

$$(y_{k(n)}, \Delta V_n) = A\mathbf{Z}^{\perp} = \left(y_{k(n)}, R_{n,k} \cdot y_{k(n)} + \epsilon_n \sqrt{1 - R_{n,k}^2}\right)$$
$$\Rightarrow \Delta V_n = R_{n,k} \cdot y_{k(n)} + \epsilon_n \sqrt{1 - R_{n,k}^2},$$

where $R_{n,k}$ is the correlation between the obligor $n$ and its industry $Y_{k(n)}$. This is equivalent to the single-factor formula proven in 2.2.1.

3. Determine which obligors defaulted on their loan by checking when $\Delta V_n \leq \Phi^{-1}(\mathrm{PD}_n)$. The tails of the distribution are symmetrical, which is why the upper tail comparison $-\Delta V_n \geq \Phi^{-1}(1 - \mathrm{PD}_n)$ is equivalent. Use the indicator function to denote the comparison

$$\mathbb{1}_{default}(n) = [n \in \{n \mid \Delta V_n \leq \Phi^{-1}(\mathrm{PD}_n)\}].$$

4. Calculate the total loss with

$$L = \sum_{n=1}^{N} \mathbb{1}_{default}(n) \cdot \mathrm{EAD}_n \cdot \mathrm{LGD}_n$$

### 4.2.2 Student's t copula

The implementation of the Student's t copula is similar to the Gaussian copula, and it can even be seen to be built on top of it. Samples from the Gaussian copula can be converted to samples from the Student's t copula by multiplying the samples with an independent inverse-gamma distributed variable.

1. Sample random values for each latent variable $Y_k$ in the model.

   It is possible to obtain Student's t distributed variables $(Y_1, \ldots, Y_K)$ from a multivariate t distribution $t(\mathbf{R}, \nu)$. Sampling can be performed using Cholesky decomposition similar to the Gaussian copula [44, chapter 6]. The only difference is the added variable $W$.

   (a) Find the $K \times K$ lower triangular Cholesky matrix $A$ such that $AA^{\intercal} = \mathbf{R}$.

(b) Sample a vector $\mathbf{Z}^{\perp}$ of $K$ i.i.d. $\mathcal{N}(0,1)$ variables.

(c) Calculate a sample from the normal distribution as in the case of the Gaussian copula $\mathbf{Z} = A\mathbf{Z}^{\perp} \sim \mathcal{N}_K(0, 1, \mathbf{R})$.

If we wanted to observe the sample of the latent variables at this stage, we would sample $W \sim \text{Inv-Gamma}(\nu/2, \nu/2)$ and multiply the previous multivariate normal sample with it to obtain the multivariate t distributed sample of $(Y_1, ..., Y_K) = \mathbf{Z}\sqrt{W} \sim t_K(\mathbf{R}, \nu)$. However, multiplication is performed only in the next stage because values of the normally distributed $\mathbf{Z}$ are needed to generate the conditional sample in the next stage. It is possible to observe the latent variables after the changes in asset values $\Delta V_n$ are computed.

2. For each obligor $n$, calculate $\Delta V_n$ given a sample of the latent variables $(Y_1, ..., Y_K) = (y_1, ..., y_K)$.

Note that in this case, the sample $\mathbf{Z} = (z_1, ..., z_K)$ is technically used in place of a sample of $\mathbf{Y}$ to avoid unnecessary computation, but the result is equivalent and the sample of $\mathbf{Y}$ is retrieved in the end.

Changes in asset values are coupled to the respective latent variables with the bivariate t copula $(Y_{k(n)}, \Delta V_n) \sim t_2(\mathbf{R}, \nu)$. Similarly to the Gaussian copula, the two-dimensional Cholesky matrix is created and used to form the samples.

(a) For each obligor $n$, calculate the lower triangular Cholesky matrix, which is

$$A = \begin{bmatrix} 1 & 0 \\ R & \sqrt{1-R^2} \end{bmatrix},$$

where $R_{n,k}$ is the correlation between $\Delta V_n$ and the corresponding latent variable $Y_k$.

(b) For each obligor $n$, sample i.i.d. $\epsilon_n \sim \mathcal{N}(0,1)$. Then we have two independent $\mathcal{N}(0,1)$ variables for each obligor, $\epsilon_n$ and $y_{k(n)}$, and we set $\mathbf{Z}^{\perp} = [z_{k_n}, \epsilon_n]^{\intercal}$, where $z_k$ are the normally distributed samples of the latent variables.

(c) Sample $W \sim \text{Inv-Gamma}(\nu/2, \nu/2)$. The same value is used for all obligors in one simulation round.

(d) For each obligor $n$, calculate the realization

$$\begin{aligned}(y_{k(n)}, \Delta V_n) = A\mathbf{Z}^{\perp}\sqrt{W} &= \left(z_{k(n)}\sqrt{W},\ R_{n,k}\cdot z_{k(n)}\sqrt{W} + \epsilon_n\sqrt{1-R_{n,k}^2}\sqrt{W}\right) \\ &= \left(y_{k(n)},\ R_{n,k}\cdot y_{k(n)} + \tilde{\epsilon_n}\sqrt{1-R_{n,k}^2}\right) \\ \Rightarrow \Delta V_n &= R_{n,k}\cdot y_{k(n)} + \tilde{\epsilon_n}\sqrt{1-R_{n,k}^2},\end{aligned}$$

where $R_{n,k}$ is the correlation between the obligor $n$ and its industry $Y_{k(n)}$ and $\tilde{\epsilon_n} \sim t(\nu)$ is the t distributed equivalent of $\epsilon_n$. This is equivalent to the single-factor formula proven in Section 2.2.1 but converted to the Student's t distribution. At this stage, the sample of the t distributed latent variables $Y_k = y_k$ can be observed.

3. Determine which obligors defaulted by seeing when $\Delta V_n \leq T^{-1}(\mathrm{PD}_n)$, where $T^{-1}$ is the inverse cdf of Student's t distribution. The tails of the distribution are symmetrical, which is why the upper tail comparison $-\Delta V_n \geq \Phi^{-1}(1 - \mathrm{PD}_n)$ is equivalent. Use the indicator function to denote the comparison

$$\mathbb{1}_{\mathrm{default}}(n) = [n \in \{n \mid \Delta V_n \leq T^{-1}(\mathrm{PD}_n)\}].$$

4. Calculate the total loss with

$$L = \sum_{n=1}^{N} \mathbb{1}_{\mathrm{default}}(n) \cdot \mathrm{EAD}_n \cdot \mathrm{LGD}_n$$

In comparison to the Gaussian copula, the Student's t copula causes two additions to the simulation implementation. In the first stage, where latent variables are calculated, and in the second stage, where asset values are simulated, one more random variable is added. The additional random variable $W$ is needed to convert the sampled distribution from normal to Student's t. However, the addition is quite minor, since sampling Gamma-distributed random variables, and therefore also inverse Gamma-distributed variables, is implemented to many popular programs, such as R, Python and Matlab.

### 4.2.3 Hierarchical Gumbel copula

1. Sample random values for each latent variable $Y_k$ in the model.

   Regardless of the structure of the hierarchy tree, latent variables are sampled recursively, following the algorithm of Hofert and Mächler [49]. The algorithm is based on algorithm 5 of McNeil [50] and is implemented in the R package `nacopula`. Let $C$ be the hierarchical Gumbel copula, where the root copula is $C_0$ with generator $\psi_0$.

   (a) Sample $V_0 \sim F_0 = \mathcal{LS}^{-1}[\psi_0^{-1}]$
   (b) For all child copulas $u$ of $C_0$:
      i. Set $C_1$ to be the child copula $u$ with generator $\psi_1$.
      ii. Sample $V_{01} \sim F_{01} = \mathcal{LS}^{-1}[\psi_{01}^{-1}(\cdot; V_0)]$
      iii. Set $C_0 := C_1$, $\psi_0 := \psi_1$ and $V_0 := V_{01}$ and continue from (1b).
   (c) For all other components $u$ of $C_0$, which are leaves of $C$ corresponding to some latent variable $Y_k$:
      i. Sample $R \sim \mathrm{Exp}(1)$
      ii. Set $y_k := \psi_0(R/V_0)$
   (d) Return the sample $(y_1, \ldots, y_k)$

The sampling distributions are defined as

$$F_0 = \mathcal{LS}^{-1}[\psi_0^{-1}] = S\left(\frac{1}{\theta}, 1, \cos^{\theta_0}\left(\frac{\pi}{2\theta_0}\right), \mathbb{1}_{\{\theta_0=1\}}; 1\right)$$

for the root copula, and

$$F_{01} = \mathcal{LS}^{-1}[\psi_{01}^{-1}(\cdot; V_0)] = S\left(\frac{\theta_0}{\theta_1}, 1, \cos\left(\frac{\pi\theta_0}{2\theta_1}\right)^{\theta_1/\theta_0}, V_0\mathbb{1}_{\{\theta_1/\theta_0=1\}}; 1\right)$$

for the child of the root copula, where $\psi_0^{-1}(t;\theta_0) = \exp(-t^{1/\theta_0})$ is the inverse generator of the root copula in the hierarchy tree and

$$\begin{aligned}\psi_{01}^{-1}(t;x) &= \exp(-x\psi_0(\psi_1^{-1}(t))) \\ &= \exp(-x(-\ln(\exp(-t^{1/\theta_1})))^{\theta_0}) \\ &= \exp(-xt^{\theta_0/\theta_1}),\ t \in [0,\infty],\ x \in (0,\infty)\end{aligned}$$

is the generator of the child copula, the nested one. The sampling distributions $S$ are stable distributions, derived using the inverse Laplace-Stieltjes transform of the inverse generator function, notated with $\mathcal{LS}^{-1}[\psi^{-1}]$. The sampling algorithm for stable distributions has been implemented, for example, in R, Python, and Matlab.

2. For each obligor $n$, calculate $\Delta V_n$ given the sample $(Y_1, ..., Y_K) = (y_1, ..., y_K)$ and determine if the obligor defaulted.

   It is not possible to calculate $\Delta V_n$ by directly sampling the Gumbel copula as in the case of the Gaussian and Student's t copula because the underlying random variables of the Gumbel sampling process do not correspond to the dimensions of the resulting copula. Instead, it is necessary to use the conditional distribution function of the copula.

   Sampling any random variable $X$ can be done by sampling a uniform random variable $U = u$ and feeding it into the inverse of the cumulative distribution function of $X$ to obtain the sample $F_X^{-1}(u) = x$.

   Following the idea, calculate the conditional distribution function $C_{\text{Gu}}(\Delta V_n | Y_{k(n)} = y; \theta_n)$, which is the cumulative distribution function of $\Delta V_n$, given the realization $Y_{k(n)} = y_{k(n)}$, by differentiating the copula by the conditioning variable [46].

$$\begin{aligned}C_{\text{Gu}}(\Delta V | Y = y; \theta) &= \frac{\partial C_{\text{Gu}}(\Delta V, Y = y; \theta)}{\partial Y} \\ &= C_{\text{Gu}}(\Delta V, y; \theta) \cdot \frac{(-\ln y)^{\theta-1}}{y} \cdot ((-\ln \Delta V)^{\theta} + (-\ln y^{\theta})^{1/\theta - 1}\end{aligned}$$

   However, the inverse of the conditional distribution function $C_{\text{Gu}}^{-1}(\Delta V | Y = y; \theta)$ can be calculated only numerically, for example, with the Newton-Raphson method [46].

(a) Sample $U = u$ from the uniform distribution $U \sim [0, 1]$.

(b) Solve $g(x) = C_{\text{Gu}}(\Delta V = x|Y = y; \theta) - u = 0$ using Newton-Raphson. Set $x_0 := 0.5$ and $i = 0$. While $g(x_i) > 0.01$:

i. $x_{i+1} := x_i - g(x_i)/g'(x_i)$

ii. $i := i + 1$

(c) Conclude $\Delta V_n = x_i$, when the iteration has converged within the tolerance $g(x_i) \leq 0.01$.

3. Determine which obligors defaulted. The upper tail of the distribution needs to represent high losses, so the change in asset values is actually $1 - \Delta V$ and the necessary comparison is therefore $1 - \Delta V_n \leq \text{PD}_n$, since $\Delta V_n$ is uniformly distributed in $[0, 1]$. Equivalently, the comparison can be $\Delta V_n \geq 1 - \text{PD}_n$ Use the indicator function to denote the comparison

$$\mathbb{1}_{\text{default}}(n) = [\{n \mid 1 - \Delta V_n \leq \text{PD}_n\}].$$

4. Calculate the total loss with

$$L = \sum_{n=1}^{N} \mathbb{1}_{\text{default}}(n) \cdot \text{EAD}_n \cdot \text{LGD}_n$$

In comparison to the Gaussian copula and the Student's t copula, the Gumbel copula has two major disadvantages. Firstly, the algorithm to calculate the latent variables in stage 1 is significantly more complex than the corresponding stage in the Gaussian or Student's t copula. It is easy to make mistakes when implementing the recursion. Secondly, the computation of $\Delta V_n$ is more complex and potentially slower because for each obligor, several iterations of Newton's method are needed. The derivative of the optimization function $g'(x)$ is also rather long.

If using R, the disadvantages of the hierarchical Gumbel approach are not that significant, as sampling the hierarchical copula is implemented in the R package `nacopula`. Overall, complexity of the implementation and potential challenges in computation speed make the hierarchical Gumbel copula less attractive than the Gaussian and Student's t copulas.

## 4.3 Explainability

The explainability of the copulas is presented by creating visual, verbal, and example-based explanations of economic capital models that utilize a given copula. Visual explanations aim to communicate the structure of each model with charts and symbols, verbal explanations describe the structure and capabilities of each model in general terms, and example-based explanations show how the structure of each model can be interpreted with business intuition.

### 4.3.1 Visual explainability

Visual illustrations of economic capital models involving each copula are presented in Figures 10, 11 and 12. The illustrations of the Gaussian copula in Figure 10 and the Student's t copula in Figure 11 are very similar because both of them can be presented as a $N$-variate version of a familiar distribution. The hierarchical Gumbel copula in Figure 12 has a considerably more complex illustration because the model cannot be expressed as a single distribution. The tree structure of the hierarchical copula becomes complex with increasing latent variables or hierarchy levels.

### 4.3.2 Textual explainability

The following verbal explanations describe the structure and capabilities of the models that incorporate each copula. The descriptions repeat some parts on purpose to clearly indicate which sections of the models work similarly.

**Description of the model with Gaussian copula:**

> The model assumes that the relationship between the latent variables is a multivariate normal distribution with correlation matrix $\mathbf{R}$. Each latent variable, referring to a dimension of the distribution, represents the performance of a single sector of the stock market. This multivariate distribution is sampled to generate values for latent variables, which can be thought of as scenarios of the economy. The change in the asset value of each company is conditionally sampled from a company-specific bivariate normal distribution, given the performance of the company's sector in a given scenario and the correlation between the company and its sector. If the asset value of a company falls too low, the company defaults on their loan and contributes to the total credit losses of the portfolio, which is used to calculate economic capital. In the high quantiles of the total loss distribution, which determine the economic capital estimate, correlation of losses tends to zero the closer the quantile is to 1. This implies that the extreme scenarios simulated by the model are primarily the result of many unlikely and independent events, instead of a single event that affects one or many sectors comprehensively.

**Description of the model with Student's t copula:**

> The model assumes that the relationship between the latent variables is a multivariate Student's t distribution with correlation matrix $\mathbf{R}$ and degrees-of-freedom parameter $\nu$. Each latent variable, referring to a dimension of the distribution, represents the performance of a single sector of the stock market. This multivariate distribution is sampled to generate values for latent variables, which can be thought of as scenarios of the economy. The change in the asset value of each company is conditionally sampled from a company-specific bivariate Student's t distribution, given the performance of the company's sector in a given scenario, the correlation between the company and its sector, and a calibrated value for $\nu$. If the asset value of a company falls too low, the company defaults on their loan and contributes to the total credit losses of the portfolio, which is used to calculate economic capital. In the high quantiles of the loss

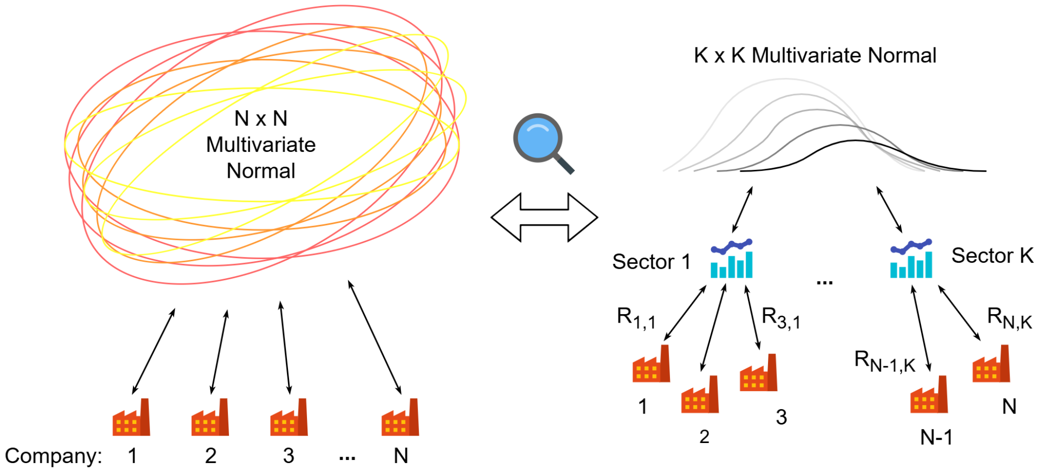



**Figure 10:** The model with the Gaussian copula can be expressed as a large multivariate normal distribution, as presented on the left. However, it is more informative to present the model with the full correlation structure present, like shown on the right. This is because pairwise correlations between companies are not actually estimated, even though they can be calculated and presented as one large distribution.

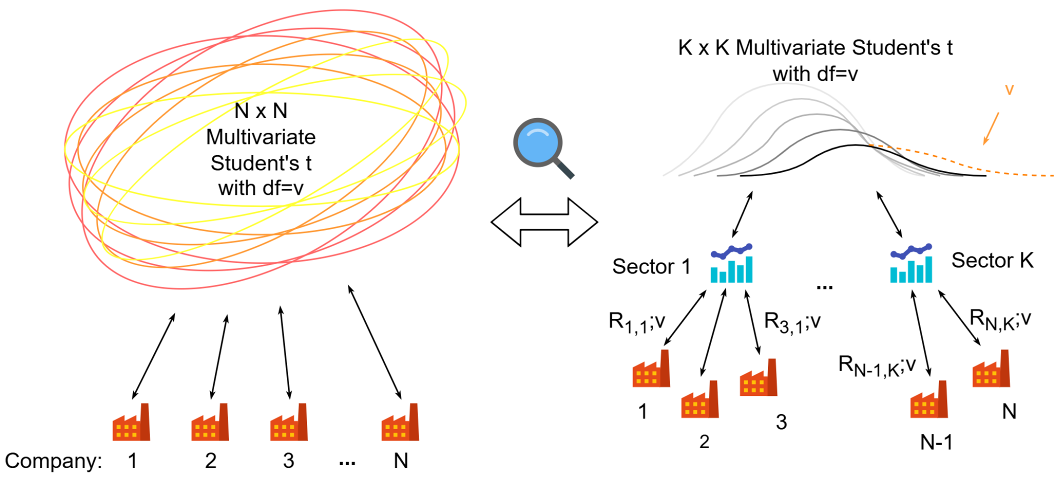



**Figure 11:** The model with the Student's t copula has a structure very similar to the one of the Gaussian copula. The only difference is that the Student's t copula enables controlling tail dependence with the degrees-of-freedom parameter $\nu$, like shown on the right side.

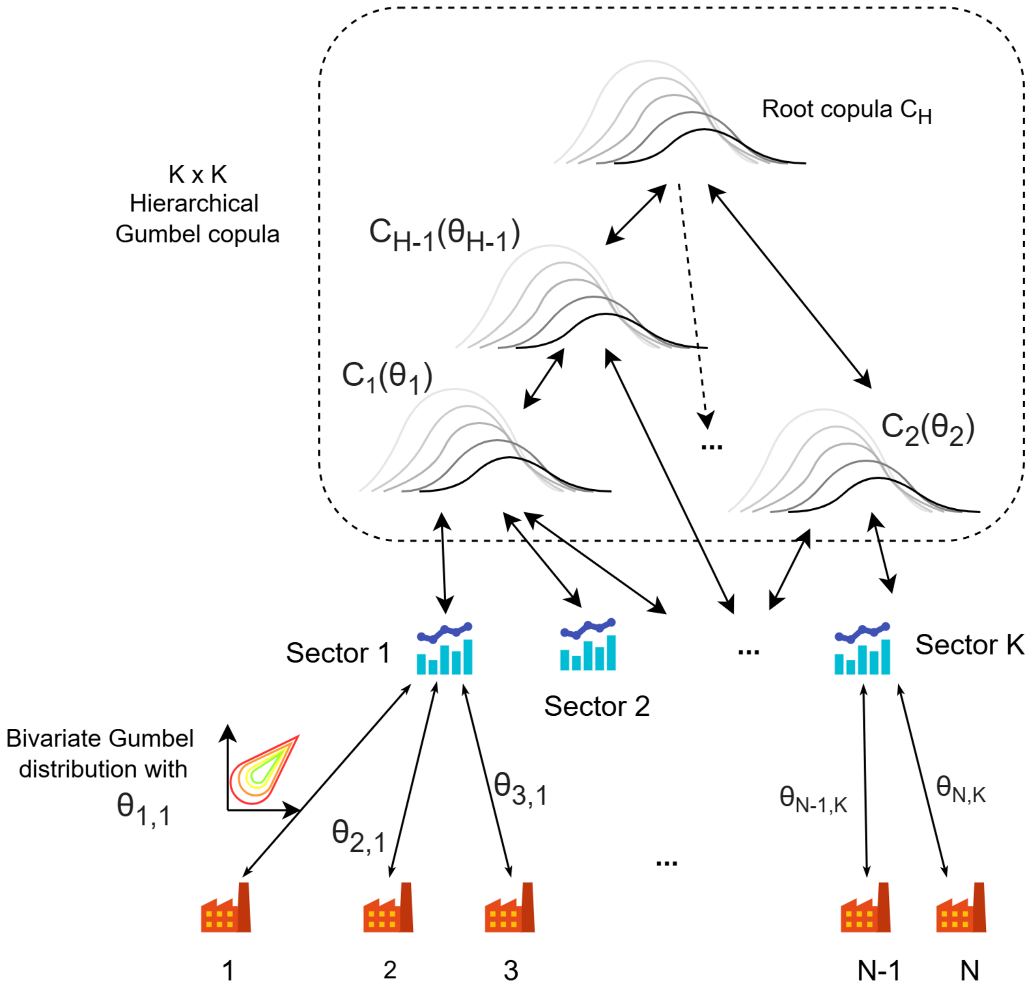



**Figure 12:** The model with the Gumbel copula has a more complex structure than the ones of the Gaussian and Student's t copulas. The main drivers of complexity are the tree structure of the hierarchy and the factor model of the obligors that does not simplify to a single distribution with the Gumbel copulas.

distribution, which determine the economic capital estimate, the correlation is lower than overall, but is controlled by the degrees-of-freedom parameter. This implies that the degrees-of-freedom parameter controls how probable it is to face an extreme situation that affects an entire sector or the economy.

**Description of the model with hierarchical Gumbel copula:**

The model describes the mutual dependence of latent variables with a tree structure, where each variable depends directly on only a few variables or groups of variables. However, these dependencies, defined with Gumbel copulas that are characterized by dependence parameters $\theta_i$, connect all latent variables together via the tree, and thus define a $K$-dimensional distribution, where $K$ is the number of latent variables. Each latent variable, referring to a dimension of the distribution, represents the performance of a single sector of the stock market.

> This distribution is sampled to generate values for latent variables, which can be thought of as scenarios of the economy. The change in the asset value of each company is conditionally sampled from a company-specific bivariate Gumbel copula, given the performance of the company's sector in a scenario, and the dependence parameter $\theta_{n,k}$ between company $n$ and its sector $k$. If the asset value of a company falls too low, the company defaults on their loan and contributes to the total credit losses of the portfolio, which is used to calculate economic capital. The $\theta$-parameter describes dependence strength asymmetrically, which means that the model assumes events that affect the economy widely in crises, but does not assume similar economy-wide effects that drive profits.

Overall, all descriptions feel quite similar because the copulas only change how parts of the model are interpreted, and not the entire model structure. The abstract tree structure of the Gumbel copula makes it the most difficult to understand, while the Gaussian and Student's t copulas rank equal.

### 4.3.3 Example based explainability

To explain the model to its users, examples with business terminology are needed instead of high-level descriptions of mathematics. Therefore, this section presents and compares examples of how the model structure can be explained to business experts. Since the process of calculating the asset value changes from the latent variables is similar for all copulas, the following business examples will focus on the interpretation of the copula for the latent variables. Examples can be crafted once the dependence parameters of the copula are calibrated, which is why arbitrary dependence parameters are assumed for the examples in this case. The dependencies are fictional, but they resulting examples look similar to what could be created with dependence parameters calibrated with actual stock market data.

To create the examples, assume that the models contain five latent variables $Y_1, ..., Y_5$ corresponding to five sectors in the Global Industry Classification Standard (GICS): Construction, Real Estate, Materials, Industrials, and Consumer Discretionary, respectively [51]. The Gaussian and Student's t copulas will then have a $5\times5$ correlation matrix $\mathbf{R}$ that describes the dependencies of the sectors. An arbitrary example of such matrix is presented in Table 3. The Student's t copula will also have a degrees-of-freedom parameter $\nu$. The hierarchical Gumbel copula will have a structure similar to Figure 9 with arbitrary parameter values, for example, $\theta_1 = 2.5$ being the dependence of Construction and Real Estate, $\theta_2 = 2.1$ being the dependence of Materials, Industrials and Consumer Discretionary, and $\theta_3 = 1.7$ being the dependence of the previously mentioned groups. For the structure to be valid, we must have $\theta_1 \geq \theta_2 \geq \theta_3$.

In models with the Gaussian or Student's t copula, pairwise correlation for each industry can be determined, and therefore used in explanations. In fact, the dependence structure is interpreted similarly for both copulas. The only difference is that correlations do not diminish in the tails of the Student's t copula.

**Hypothetical business scenario with the Gaussian copula or the Student's t copula:**

> In the example model, Real Estate is correlated with Construction because they both depend on the housing market. However, Real Estate is not very correlated with other sectors because it does not suffer from high material prices or issues in other countries like the other sectors. Construction, on the other hand, is rather dependent on the cost of construction materials, which is indicated by its correlation to Materials sector. Materials, Industrials and Consumer Discretionary are correlated with each other because they all depend on consumption and international supply chains.

Pairwise dependencies cannot be directly described in the hierarchical Gumbel copula because they are not explicitly defined for most pairs. Instead, scenario explanations must focus on finding common characteristics for sectors that are in the same copula. In our example, Construction and Real Estate are highly dependent on each other with parameter $\theta_1$ in the copula $C_1$. Materials, Industrials and Consumer Discretionary also depend on each other in the copula $C_2$.

**Hypothetical business scenario with the hierarchical Gumbel copula:**

> The example model estimates that large losses from Construction and Real Estate would occur simultaneously because they are both connected to the housing market. Materials, Industrials and Consumer Discretionary are also estimated to depend on each other, possibly because they all depend on the availability of certain materials, and interruptions in supply chains would affect all three industries. Dependencies between other pairs of industries, for example, Construction and Materials cannot be explicitly expressed because they are only implicitly set by the dependence of the two groups that represent the housing market and manufacturing sectors. Their dependence is explained by the remaining economic effects, such as consumer confidence or global stability.

The Gumbel copula example shows that in some cases its group-wise structure can even yield simpler explanations than the other copulas as there are fewer dependencies to examine. However, depending on the sectors in each group, it may be challenging to find common denominators within the group, which will lead to explanations that

**Table 3:** Arbitrary correlation matrix of latent variables for the example based explanations. The latent variables are assigned by the stock market sector of each obligor.

| | Construction | Real Estate | Materials | Industrials | Consumer Discr. |
|---|---|---|---|---|---|
| Construction | 1 | 0.9 | 0.7 | 0.4 | 0.4 |
| Real Estate | 0.9 | 1 | 0.3 | 0.3 | 0.5 |
| Materials | 0.7 | 0.3 | 1 | 0.7 | 0.7 |
| Industrials | 0.4 | 0.3 | 0.7 | 1 | 0.7 |
| Cons. Discr. | 0.4 | 0.5 | 0.7 | 0.7 | 1 |

are harder to understand. Furthermore, model users may find it insufficient that some dependencies, which may have significance in business sense, are not visible in the model.

## 4.4 Overall assessment

Conclusions of the previous subsections are gathered in Table 4. Each copula is given a grade: Best, Good, or Weak, in the three categories of comparisons that were made. The grades are relative, meaning that a copula that is rated good may still be objectively bad, when compared to something outside the scope of this thesis. Therefore, the table should be seen as a summary of the analyses performed in this thesis and should always be read, when accompanied by the details.

**Table 4:** The results of the comparisons are summarized with a grading to categories Best, Good, and Weak.

| Category | Gaussian | Student's t | Hierarchical Gumbel |
|---|---|---|---|
| Theoretical justification | Weak | Good | Best |
| Implementation | Best | Best | Weak |
| Explainability | Best | Best | Weak |

Theoretical justification of the Gumbel copula is declared best, because it was able to capture both theoretical properties found in the stock market, while the Student's t copula was able to capture one and the Gaussian copula neither. The implementation of the Gaussian and Student's t copulas was almost equal, while the Gumbel copula had severe drawbacks caused by its recursive sampling algorithm and the need to solve inverse distribution values numerically. The explainability of the Gaussian and the Student's t copulas is also very similar. The hierarchy structure of the Gumbel copula is hard to imagine and explain, which is why it was assessed to be the weakest in explainability.

Table 4 shows that the Gaussian copula is dominated by the Student's t copula. The t copula has equal or better grade than the Gaussian copula in each category, meaning that regardless of preferences between the comparison categories, a rational decision maker would choose the Student's t copula over the Gaussian. The table also shows that the Hierarchical Gumbel copula would only get selected if the decision maker would heavily prefer theoretical justifiability over the other criteria.

Based on the results, the Student's t copula should be chosen most often out of the covered options. The Gaussian copula could be chosen in situations where asset correlation is weaker and more linear, and the Gumbel copula would likely be chosen only if computational efficiency is not a concern and theoretical justification is weighted.

# 5 Discussion

There are numerous ways to find better dependence structures than the ones covered by this thesis. One option would be to compare new ones and another one is to further develop the options that were covered. For example, the Student's t copula could be combined with another copula in the lower tail of the loss distribution to make it asymmetrical and therefore improve its theoretical justifiability. The Gumbel copula could also be replaced with another Archimedean copula that supports conditional sampling, such as the Clayton copula. However, the fact that these possibilities are identified based on the comparisons in this thesis, is evidence of the usefulness of the comparison methods presented by this thesis. Ultimately, the results presented in this work highlight that there is no copula that would be perfect for all modeling needs. Instead, the choice of copula is a balance between different characteristics and constraints. This work provides an example of characteristics that could be considered and how they can be evaluated.

Since there is a large number of copulas to test, future work could focus on identifying groups of copulas that dominate others in terms of preference. The copulas within the dominating group could then be compared, or samples from each group could be compared to help risk analysts determine which group to search further. Alternatively, some characteristics could also be identified that would predict better ranking in comparisons. For example, the number of dependence parameters or the value of the tail dependence coefficients could quickly filter out some copulas that do not suit the model.

## References


[1] European Central Bank. *ECB Guide to the Internal Capital Adequacy Assessment Process (ICAAP)*. 2018.

[2] BIS Economic Capital. “Range of practices and issues in economic capital frameworks”. In: *Basel Committee on Banking Supervision, Bank for International Settlements, Basel* (2009).

[3] A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools - Revised Edition*. Princeton university press, 2015.

[4] M. Baron, E. Verner, and W. Xiong. *Banking crises without panics*. Working Paper 26908. National Bureau of Economic Research, 2020.

[5] R. C. Merton. “On the pricing of corporate debt: The risk structure of interest rates”. In: *The Journal of Finance* 29.2 (1974), pp. 449–470.

[6] O. Vasicek. “The distribution of loan portfolio value”. In: *Risk* 15.12 (2002), pp. 160–162.

[7] H. Fong and H. G. Fong. *The Credit Market Handbook Advanced Modeling Issues*. 1st edition. Wiley, 2006.

[8] E. Kole, K. Koedijk, and M. Verbeek. “Selecting copulas for risk management”. In: *Journal of Banking & Finance* 31.8 (2007), pp. 2405–2423.

[9] H. Cui, K. S. Tan, and F. Yang. “Portfolio credit risk with Archimedean copulas: Asymptotic analysis and efficient simulation”. In: *Annals of Operations Research* 332.1-3 (2024), pp. 55–84.

[10] A. Di Clemente. “Modeling portfolio credit risk taking into account the default correlations using a copula approach: Implementation to an Italian loan portfolio”. In: *Journal of Risk and Financial Management* 13.6 (2020), pp. 1–23.

[11] D. MacKenzie and T. Spears. “’The formula that killed Wall Street’: The Gaussian copula and modelling practices in investment banking”. In: *Social Studies of Science* 44.3 (2014), pp. 393–417.

[12] C. Donnelly and P. Embrechts. “The devil is in the tails: Actuarial mathematics and the subprime mortgage crisis”. In: *ASTIN Bulletin : The Journal of the IAA* 40.1 (2010), pp. 1–33.

[13] F. Longin and B. Solnik. “Extreme correlation of international equity markets”. In: *The Journal of Finance (New York)* 56.2 (2001), pp. 649–676.

[14] K.-H. Bae, G. A. Karolyi, and R. M. Stulz. “A new approach to measuring financial contagion”. In: *The Review of Financial Studies* 16.3 (2003), pp. 717–763.

[15] D. Cossin, H. Schellhorn, N. Song, and S. Tungsong. "A theoretical argument why the t-copula explains credit risk contagion better than the Gaussian copula". In: *Advances in Decision Sciences* 2010 (2010), pp. 1–29.

[16] M. Geidosch and M. Fischer. "Application of vine copulas to credit portfolio risk modeling". In: *Journal of Risk and Financial Management* 9.2 (2016), pp. 1–15.

[17] R. K. Y. Low, J. Alcock, R. Faff, and T. Brailsford. "Canonical vine copulas in the context of modern portfolio management: Are they worth it?" In: *Journal of Banking & Finance* 37.8 (2013), pp. 3085–3099.

[18] A. Mühlbacher and T. Guhr. "Credit risk meets random matrices: Coping with non-stationary asset correlations". In: *Risks* 6.2 (2018).

[19] A. Lipton and A. Rennie. *Credit Correlation: Life After Copulas*. World Scientific Publishing Co. Pte. Ltd, 2008.

[20] A. Meucci. "A new breed of copulas for risk and portfolio management". In: *Risk (London. 1987)* 24.9 (2011), pp. 110–.

[21] B. Baesens, D. Rösch, and H. Scheule. *Credit Risk Analytics: Measurement Techniques, Applications, and Examples in SAS*. 1st ed. Wiley, 2016.

[22] Standard & Poor's Financial Services LLC. *Default, Transition, and Recovery: 2023 Annual Global Corporate Default and Rating Transition Study*. 2023.

[23] BCBS. *An Explanatory Note on the Basel II IRB Risk Weight Functions*. 2005.

[24] EBA. *Final Report on Guidelines on Default Definition*. 2016.

[25] BCBS. *Calculation of RWA for Credit Risk, Standardized Approach: Individual Exposures*. 2022.

[26] Y. Konno and Y. Itoh. "An alternative to the standardized approach for assessing credit risk under the Basel accords". In: *Cogent Economics & Finance* 4.1 (2016), pp. 1–13.

[27] BCBS. *Calculation of RWA for Credit Risk, IRB Approach: Risk Weight Functions*. 2020.

[28] M. B. Gordy. "A risk-factor model foundation for ratings-based bank capital rules". In: *Journal of Financial Intermediation* 12.3 (2003), pp. 199–232.

[29] L. D. Adzobu, E. K. Agbloyor, and A. Aboagye. "The effect of loan portfolio diversification on banks' risks and return: Evidence from an emerging market". In: *Managerial Finance* 43.11 (2017), pp. 1274–1291.

[30] B. Kollár and B. Gondžárová. "Comparison of current credit risk models". In: *Procedia Economics and Finance* 23 (2015), pp. 341–347.

[31] A. Saunders and L. Allen. *Credit Risk Management In and Out of the Financial Crisis: New Approaches to Value at Risk and Other Paradigms*. 3rd Edition. Vol. 528. Wiley, 2010.

[32] R. B. Nelsen. *An Introduction to Copulas*. Springer, 2006.

[33] A. Meucci. “A short, comprehensive, practical guide to copulas”. In: *GARP Risk Professional* (2011), pp. 22–27.

[34] T. Kämpke and F. J. Radermacher. “The generalized inverse of distribution functions”. In: *Income Modeling and Balancing: A Rigorous Treatment of Distribution Patterns*. Springer International Publishing, 2015, pp. 9–28.

[35] B. Schweizer and A. Sklar. *Probabilistic Metric Spaces*. Courier Corporation, 2011.

[36] U. Cherubini. *Dynamic Copula Methods in Finance*. 2nd ed. Wiley, 2011.

[37] J. Owen and R. Rabinovitch. “On the class of elliptical distributions and their applications to the theory of portfolio choice”. In: *The Journal of Finance (New York)* 38.3 (1983), pp. 745–752.

[38] S. Pafka and I. Kondor. “Estimated correlation matrices and portfolio optimization”. In: *Physica A* 343 (2004), pp. 623–634.

[39] P. J. Schoenbucher. “Factor models: Portfolio credit risks when defaults are correlated”. In: *The Journal of Risk Rinance* 3.1 (2001), pp. 45–56.

[40] D. H. Oh and A. J. Patton. “Time-varying systemic risk: Evidence from a dynamic copula model of CDS spreads”. In: *Journal of Business & Economic Statistics* 36.2 (2018), pp. 181–195.

[41] B. Mandelbrot. “The Pareto-Lévy law and the distribution of income”. In: *International Economic Review (Philadelphia)* 1.2 (1960), pp. 79–106.

[42] J. Morgan. “CreditMetrics technical document”. In: *JP Morgan, New York* 1 (1997), pp. 102–127.

[43] S. M. Ross. *Simulation*. Sixth Edition. Academic Press, 2023.

[44] H. Joe. *Dependence Modeling with Copulas*. CRC press, 2014.

[45] W. Zhu, C.-W. Wang, and K. S. Tan. “Structure and estimation of Lévy subordinated hierarchical Archimedean copulas (LSHAC): Theory and empirical tests”. In: *Journal of Banking & Finance* 69 (2016), pp. 20–36.

[46] K. Aas, C. Czado, A. Frigessi, and H. Bakken. “Pair-copula constructions of multiple dependence”. In: *Insurance: Mathematics and Economics* 44.2 (2009), pp. 182–198.

[47] K. Goda. “Statistical modeling of joint probability distribution using copula: Application to peak and permanent displacement seismic demands”. In: *Structural Safety* 32.2 (2010), pp. 112–123.

[48] P. Mishra. *Explainable AI Recipes : Implement Solutions to Model Explainability and Interpretability with Python*. 1st ed. 2023. Apress Media LLC, 2023.

[49] M. Hofert and M. Mächler. “Nested Archimedean copulas meet R: the nacopula package”. In: *Journal of Statistical Software* 39 (2011).

[50] A. J. McNeil. “Sampling nested Archimedean copulas”. In: *Journal of Statistical Computation and Simulation* 78.6 (2008), pp. 567–581.

[51] MSCI and S. D. J. Indices. *The Global Industry Classification Standard (GICS)*. 2023.