Master Thesis

Modelling Dependent Defaults in Static Credit Portfolios

Sonny Johansson and Tommy von Brömsen

Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
GÖTEBORG UNIVERSITY
Göteborg, Sweden 2011
THESIS FOR THE DEGREE OF MASTER OF SCIENCE (30 ECTS)

Modelling Dependent Defaults in Static Credit Portfolios

Sonny Johansson and Tommy von Brömsen

Tutor: Alexander Herbertsson
Examiner: Holger Rootzen

Department of Mathematical Sciences
Chalmers University of Technology and Göteborg University
41296 Göteborg, Sweden
Göteborg 2011
Abstract

This thesis studies the modelling of credit risk in static credit portfolios, where the emphasis lies on the modelling of the dependence between defaults of the obligors. Finding a proper way to model dependence between defaults is a central, however non-trivial, task in order to incorporate the possibilities of extreme events. Further, this default dependence greatly influences the portfolio loss distribution, and thus is vital to realistically capture the true credit risk in the portfolio. We will present several different approaches in order to introduce a default dependence, and the core ideas of these approaches are present in the most common industry models. Furthermore, the asymptotic behavior in static credit portfolios will be explored and we will derive large portfolio approximations in several model settings. Finally, it will be concluded that the different approaches of modelling dependence yield similar results for non-extreme scenarios. This implies that the choice of model is of less concern than the process of estimating the individual default probability and default correlation.
Acknowledgments

We would like to thank our supervisor Alexander Herbertsson for taking his time and making this thesis possible. Throughout the thesis he has provided us with excellent discussions and guidance.
# Contents

1 Introduction .................................................. 1

2 Credit Risk Management ........................................... 3
   2.1 Financial Risk ............................................ 4
   2.2 Measuring and Managing Credit Risk ....................... 4
   2.3 Financial Crisis ........................................... 7
   2.4 The Basel Accords ......................................... 9
   2.5 The Importance of Dependence Modelling .................. 10

3 Static Credit Risk Models ......................................... 12
   3.1 Fundamentals of Static Credit Risk Modelling ............. 13
   3.2 Binomial Model ............................................ 15
   3.3 Bernoulli Mixture Model .................................. 17
      3.3.1 One-Factor Exchangeable Bernoulli Mixture Model .... 18
      3.3.2 Asymptotic Behavior in Large Portfolios ............. 21
   3.4 Threshold Models Using The Mixture Approach .......... 23
      3.4.1 The Merton Framework ................................. 24
      3.4.2 Mixture Models in the Merton Framework ............. 27
   3.5 Threshold Models Using Copulas ................................
      3.5.1 Copulas .............................................. 36
      3.5.2 Simulations from Copulas in Static Credit Risk Portfolios .... 45
      3.5.3 Bernoulli Mixture Models based on the Archimedean Copula .. 50

4 Risk Measuring in Static Credit Risk Models .................. 54
   4.1 Value-at-Risk and Expected Shortfall ..................... 55
   4.2 Bernoulli Mixture Model .................................. 60
Chapter 1

Introduction

Credit risk management is a central concern for most financial institutions, since insufficient credit risk management can lead to severe unexpected losses. Financial institutions and other companies within the finance industry are not only forced by law to engage in risk management, but there are also beneficial consequences in terms of higher profitability. Further, risk management induces a stability to the organization and contributes to an overall economic stability of the society. However, to manage credit risk one must first be able to model and quantify this risk. When modelling risk in credit portfolios it is of uttermost importance to find a proper way to model the dependence structure between the defaults of companies.

In reality, extreme events such as those during the financial crisis of 2008, occur more often than most models suggest. These extreme events indicate that there is a strong default dependence between companies. In the last decade we have seen an increased globalization, stronger contagion effects mainly caused by the introduction of new advanced derivatives, and more regulations such as higher capital requirements. These are all factors that have been contributing to an increased need for enhanced credit risk management.

Hence, the main problem of realistically capturing the credit risk in a portfolio is the modelling of default dependence. However, this task is far from trivial and even the Basel II framework has been criticized for proposing a too simplified approach of describing the dependence structure. In this thesis we will explore different approaches
for modelling credit risk in static portfolios, where each approach has its own way of modelling the dependence structure. Further, we will investigate the implications that the different approaches entail, and thus expose any potential model risk. Ultimately, one of the main reasons for engaging in credit risk modelling is to find accurate estimates of the credit risk in a portfolio. Thus, we will explore how the choice of model affects the final risk measure result.

The rest of this thesis is organized as follows. First in Chapter 2 we introduce the concepts of risk and risk management, with an emphasis on credit risk. Further, we will discuss the role of credit derivatives in the recent financial crisis, and we will briefly mention the Basel regulations. Chapter 2 is concluded with a discussion of the importance of implementing dependence between obligors in a credit portfolio. Continuing, Chapter 3 outlines different approaches to modelling defaults and dependence in static credit portfolios. The two main types of models are the Bernoulli mixture model and the threshold model. Next, Chapter 4 treats the measuring of risk in static credit risk models. The two risk measures Value-at-Risk and Expected Shortfall will be introduced and applied to the credit models discussed in Chapter 3.
Chapter 2

Credit Risk Management

When mentioning the word "risk" people often refer to something negative, e.g. the risk of losing something valuable. However, the concept of risk does not only involve possible downside loss, but also the chance of an upside gain. In financial terms for instance, the volatility of a stock price is often referred to as the risk associated with the stock. Thus, one can speak of the risk of the stock price going both up and down. Depending on the position (long or short), either the appreciation or depreciation of the stock price can yield a profit for the investor. In general, the higher the volatility/risk is, the higher expected return on the investment. In the end, any investment strategy, whether it is in the stock market, housing market, or any other market, is concerned with the relationship between risk and expected return. Thus, in order to gain one must be willing to take on risk.

There are many types of risks, though the focus of this thesis will be on financial risk in general, and credit risk in particular. Being able to manage and control risk is of crucial importance to any practitioner. Subsequently, in this introductive, non-mathematical chapter the concepts of risk and risk management will be further explored. In Section 2.1 different types of financial risk will be introduced, e.g. credit risk, equity risk, interest rate risk and exchange rate risk. Further, Section 2.2 will discuss some common ways to measure risk and present the benefits of risk management. In an effort to connect these concepts to reality, Section 2.3 deals with the recent financial crisis, where the primary focus will be on credit derivatives and structured products. Furthermore, Section 2.4 discusses the regulatory capital requirements and the Basel accords, presenting
even more reasons to explore the importance of risk management. Finally, based on the
discussion in sections 2.1-2.4, the chapter will be concluded with a section emphasizing
the importance of modelling dependence between obligors.

2.1 Financial Risk

Financial risk is the common name for different types of risks present in the financial
industry. Usually one divides financial risk into three major components; market risk, operational risk and credit risk. Market risk is a wide definition of risks involving the
change in the value of a financial asset. Therefore, this type of risk is often divided
further into more specific parts, such as equity risk, interest rate risk and exchange rate
risk. Further, operational risk is the second main category, and is defined as "the risk
that existing technology or support systems may malfunction, that fraud may affect
the financial activities, and/or external shocks such as hurricanes and floods occur”
(Saunders and Cornett, 2007, p. 535). One important part of operational risk is model
risk, which is the risk that the applied model does not capture, or reflect, the reality
well enough. The third and final type of financial risk is credit risk, which is the risk
that a debtor will not fulfill his promised payments to the lender. If a debtor, or an
obligor, fails to uphold his part of the financial agreement and pay his debt we say that
the obligor defaults. In addition to credit risk, this thesis will also take into account
some of the other risks mentioned, especially model risk.

2.2 Measuring and Managing Credit Risk

This section is concerned with the reasons and benefits for managing risk. In order to
manage risk one must find an appropriate way to measure the risk, and there are several
approaches in doing so. Regardless of the chosen approach, it will involve the notion
of randomness, as the crucial part of the definition of risk is the uncertainty of the
possible gain or loss. This thesis is primarily interested in the credit risk of portfolios
of financial assets, i.e. the concern that one or more obligors might default, and thus
impose a credit loss in the portfolio. Therefore, in the context here, credit risk will solely
be concerned about future losses, and not possible gains. Subsequently, let us define
the random variable $L$, which describes the loss in a portfolio consisting of financial
assets. We call the distribution of $L$ the loss distribution which gives information about the loss probabilities.

One traditional way of measuring risk is the so called Notional-amount approach, where the risk is measured by adding the notional values of the individual assets in the portfolio and weighing it according to its riskiness. This approach has the advantage of being simple, but has a few other shortcomings such as not taking into account the benefits of diversification or netting between short and long positions. Most modern models of measuring risk use the loss distribution to make statistical estimates in order to measure risk. This approach has led to the common risk measures Value-at-Risk (VaR) and the related Expected Shortfall (ES). A more in-depth discussion of these risk measures will be presented in chapter 4. Further, one major problem with analyzing the loss distribution is choosing an appropriate default model that renders the loss. This thesis will discuss several models for doing so, along with an analysis of the benefits and drawbacks of each model. Note that there are also other ways of measuring risk, but the main concern here is the approach of using the loss distribution.

Let us move on to discussing the purpose of managing risk. First of all, we observe that there are different perspectives of risk management. For instance, the view of the stock owners differs from the view of the society as a whole. In other words, the success (or failure) of a large company or financial institution will directly affect the stock owners, but also the general economy and thereby indirectly the society. The significant impacts of the well-being of financial institutions has led to the need for supervision and regulations. Saunders and Cornett (2007, p. 383) describe it as the goal to enhance the net social benefits of banking services to the economy. Some of the most important regulations, which also have close connections to credit risk, are those concerned with capital requirements. Such regulations are primarily set through the Basel Accords, which will be discussed in Section 2.4. Further, note that not all companies and financial institutions are under the same jurisdiction of authoritative supervision. A hedge fund for instance has less restrictions than a commercial bank and is therefore able to operate under greater risks. Regardless of the amount of regulation a financial institution is under, the need for managing risk is equally important. This is why great amounts of money and time is invested in controlling and managing risk.
Another crucial type of systemic risk worth mentioning is the so called contagious risk, which is the risk that the failure of one financial institution can affect others into also failing. The globalization has made financial institutions more exposed to risks in markets around the world, and the crash in one market will ultimately affect most other markets. For instance, dramatically falling housing prices in the U.S. market would have strong effects on the global economy. In recent years, the use of structured products have exploded around the world, fueling the risks for contagious effects on the international markets. The idea behind structured products is to avoid excess risk by grouping financial assets into one financial product and then selling pieces of it to international investors. For instance, an American bank can practically sell some of its risk within the U.S. housing market by packaging its mortgage loans into a structured product and selling it to investors around the world. In the next section, the significance of contagious risk and structured products will be even more evident as we discuss the recent financial crisis and the events leading up to it.

To summarize what has been said about risk management, McNeil, Frey and Embrechts (2005) describe the expectations of the financial system from the view of the society in the following way:

Modern society relies on the smooth functioning of banking and insurance systems and has a collective interest in the stability of such systems. The regulatory process culminating in Basel II has been strongly motivated by the fear of systemic risk, i.e. the danger that problems in a single financial institution may spill over and, in extreme situations disrupt the normal functioning of the entire financial system. (McNeil et al., 2005, p. 15)

Hence, the most important reason for companies and financial institutions to engage in risk management is to induce a stability in their business, and thus contribute to the stability of the economy as a whole. Even though risk management activities are costly, stock owners in general appreciate its benefits since, in the end, it adds value to the company. There are extensive discussions about why risk management can increase the value of a company, though we will only briefly mention some of the conclusions for it. Two of the main reasons are that risk management can reduce tax costs for stock owners, and that it makes bankruptcy (including indirect bankruptcy) less likely. These
points are incentives enough for any company, either required by regulations or not, to invest in risk management.

2.3 Financial Crisis

The recent financial crisis began in 2007 on the U.S. housing market. Even today, early 2011, the aftermath of the financial crisis is still evident on most markets around the world, especially the U.S. markets. The primary reasons for the crisis is still argued between professionals, but we will attempt to summarize some of the key contributing factors while keeping a focus on credit risk related issues. Before the crisis, the U.S. housing market had been blooming for years (The Federal Reserve, 2011). The market was considered extremely stable, prices kept moving in the right direction and few people sensed the bubble that was emerging. The over-confidence within American financial institutions gave rise to the increased use of the so called subprime mortgages. These mortgages had higher interest rates than ordinary mortgage loans, but were constructed to give people with bad credit an opportunity to own their own house. This can be connected to what George Bush said in 2002:

We want everybody in America to own their own home. [...] One of the programs is designed to help deserving families who have bad credit histories to qualify for homeownership loans [...] the low-income home buyer can have just as nice a house as anybody else. (Bush, 2002)

Thus, more people were now able to purchase their own house, which further increased the demand on the housing market and thus pushing up prices even more. The financial institutions saw no problem with these mortgages, since housing prices were still rising, creating a surplus in the loans. However, most borrowers did not sit still in their house, enjoying this surplus. On the contrary, the general consensus was to keep borrowing on the mortgages to gain additional purchasing power. Thus, most houses were highly leveraged. Eventually, the financial institutions recognized their huge exposure to the housing market, and even though they still considered it a stable market, they saw an opportunity to reduce some of its risk. By constructing securitized products in the form of Collateralized Debt Obligations (CDO:s) the financial institutions were able to sell most of the risk to investors. A CDO is essentially a package of many types of loans, creating a product that can be bought and sold. Apart from mortgages these products
can for instance include student loans, car loans and credit card debts. The investors of these products take on the default risk of the loans and in return they make money when payments are made on the loans. Rating agencies had, in general, an over-confidence in these products, giving them a too high rating class. Thus, the CDO:s became very attractive because of their high expected return to risk ratio. Further, the transparency of the CDO:s were low, i.e. investors had low insight into what types of loans were actually included in the CDO:s. Thus, as for their risk control, investors were left to solely rely on the rating agencies. Furthermore, the investors of CDO:s could insure the default risk by purchasing Credit Default Swaps (CDS) from financial institutions. In fact, any person were able to buy CDS:s, and thus bet on the event that the underlying obligors in the CDO:s would default. The insurance company AIG was the financial institution that sold the most CDS:s during the years before the crisis, which also turned out to be devastating for AIG’s financial performance. For a more detailed discussion on the role of CDO:s in the financial crisis see Saunders and Allen (2010, Ch. 1)

In 2007 the U.S. housing market started to turn around and prices dropped. The subprime mortgages were the first to default, and the number of defaults grew dramatically. Before the crash, the financial institutions had taken huge positions in the derivatives market, and up until 2007 they had been making significant profits for doing so. However, once the market turned around, their exposure was unhedged and caused severe damage to the financial institutions. Because of the low transparency in the derivatives market, the financial institutions knew little of their actual exposure, and even less about the exposure that other financial institutions had. In order for the financial industry to function, banks must be able to borrow and lend money between themselves on a short-term basis. However, because of the market instability and uncertainty, banks stopped their inter-bank lending, which practically crippled the financial markets and caused significant liquidity problems. Some large financial institutions soon suffered from severe solvency risk, and in September 15, 2008 Lehman Brothers filed for bankruptcy. For more on the financial crisis see e.g. Saunders and Allen (2010) or The Federal Reserve (2011).

Further, it is clear that credit risk management played a central role throughout the crisis. Many financial institutions relied on their mathematical models to provide them...
with accurate risk measures, but most of these models failed to incorporate the impacts of these extreme events, in particular the default contagion aspects. Ultimately, there were many factors that caused the crisis, even more than those mentioned above. In the aftermath of the crisis, there has been a rising demand for increased transparency in financial products, especially advanced derivatives such as CDO:s. Also, there will be increased regulations with respect to minimum capital requirements and other credit risk related issues (see the next section).

2.4 The Basel Accords

In order to ensure that financial institutions have enough capital on account to meet obligations and to absorb unexpected losses, the Basel Committee on Bank Supervision (BCBS) established the Basel Accords in 1988. The primary purpose of the Basel Accords (Basel I) was to require banks to hold capital in accordance with the perceived credit risk in their portfolio of activities. The Basel Accords serve as recommendations and guidelines for central banks and governments to address the issues of liquidity and insolvency risk within financial institutions. These recommendations do not have legal force in themselves, but member countries of the BCBS adopt these guidelines and implement them in their economy. In early 2011, the BCBS had 27 member countries, including all of the G20-countries (Bank for International Settlements, 2009).

Over the years, Basel I has been criticized for not fulfilling its function. One of the weaknesses was its failure to adequately differentiate between levels of credit risk in assets in general, and commercial and industrial loans in particular. These were some of the reasons to why the Basel Accords were revised and republished as Basel II in 2004. The new set of recommendations and guidelines incorporated credit ratings into the regulatory capital standards and allowed the capital requirements on assets to vary as the credit rating of an obligor changed (Jacques, 2008). Basel II aimed to strengthen international banking requirements as well as to supervise and help implement these requirements.

In the aftermath of the recent financial crisis discussed in Section 2.3, the BCBS began developing a new set of Basel Accords which will be introduced as Basel III. The new changes in the Accords will primarily be concerned with credit risk, and especially the
Value-at-Risk framework. For instance, in order to account for more extreme events in the risk measurement process, Basel III will include a further developed stressed Value-at-Risk. Another change will be the introduction of higher risk weights for the exposure to securitized products, which helps reveal the level of risk inherent in these products. Also, the minimum capital requirements for financial institutions will be increased (Bank for International Settlements, 2010).

2.5 The Importance of Dependence Modelling

One of the most important and difficult task when modelling credit risk is to properly include a realistic default dependence between the obligors in the portfolio. As we have discussed in the previous sections it is not reasonable to assume independence between defaults of companies. On the contrary, globalization has made international companies more susceptible to the economic performance of markets around the world. Structured products and other securitized products are easily sold in the international market, and the trading of such products have exploded in recent years. With a high expected return at a perceived low level of risk these products appeal to many investors. However, the issue with non-transparency implies that the level of risk could be significantly higher than expected, as we have seen in the 2008 financial crisis. Taking all of this into account, we see an increased contagion risk affecting not only international companies but also domestic companies of all sizes in markets around the world. Thus, it is evident that the dependence between obligors in the credit market is substantial, and there is a strong need for credit risk models that do a good job in capturing this dependence structure. However, this task is far from trivial, as David Lando points out in the book Credit Risk Modeling:

Modeling dependence between default events and between credit quality changes is, in practice one of the biggest challenges of credit risk models. The most obvious reason for worrying about dependence is that it affects the distribution of loan portfolio losses and is therefore critical in determining quantiles or other risk measure used for allocating capital for solvency purposes. (Lando, 2004, p. 213)

It is a known fact that financial time series generally have "thick" tails, that is to say that extreme outcomes are more common. In credit risk the default dependence has an
essential role in the upper tail of the loss distribution.

Furthermore, there are several different approaches to model default dependence. One is to introduce a factor variable which is a set of macroeconomic variables that are common for all companies. As will be seen in this thesis, the factor variable will introduce a more realistic default dependence among the obligors in the portfolio.
Chapter 3

Static Credit Risk Models

There is a significant distinction between credit risk models focusing on credit risk management on the one hand, and the pricing of credit risk securities such as CDS:s and CDO:s on the other. In credit risk management so called *static* credit risk models are often used, as opposed to *dynamic* credit risk models which are used in the pricing of credit securities. The major difference between these models is the timing of defaults; an issue which is ignored in static models, while central in dynamic models.

The emphasis in this thesis will be on static portfolio credit risk models. These models are concerned with whether an obligor has defaulted or not up to a given time. The exact timing of the default is not relevant in this context; this is however a concern in dynamic models, which are used when pricing e.g. CDO:s.

In Section 3.1 the fundamental framework in static portfolio credit risk modelling are discussed and these models are divided into two categories; one contains the *Mixture Models*, and the other the *Threshold Models*. However, threshold models are nothing more than special cases of the mixture models. Despite this strong similarity, we maintain this categorization since the two models have different approaches to modelling defaults. Further, common elements and standard notations in credit models are also discussed and introduced in this section. In Section 3.2 one of the simplest models; the binomial model is described. This model is in Section 3.3 extended to allow for dependence among obligors and we enter the frame of the Bernoulli mixture model. The definition and some properties of the Bernoulli mixture model will be stated in a general setting, and even more focus will be put on the one-factor exchangeable Bernoulli
mixture model. The section is continued by describing the behavior in portfolios where the number of obligors is large. It will be concluded that as the portfolio size tends to infinity the loss distribution will be completely determined by the distribution of the mixing variable.

In Section 3.4 we introduce the threshold models. The basic idea in a threshold model is that a default occurs if some value concerning the obligor falls below some threshold. This value can for example be interpreted as the asset value of the obligor, and it can be modelled by a stochastic process. The analysis of threshold models will be done using two different approaches. The analysis in the first approach is similar to the one in the mixture model framework, and draws from the ideas of Merton. This is the concern of Section 3.4. The second approach uses the concept of copulas, which is a useful tool for modelling the dependence structure in credit portfolios. The implication of using copulas in credit risk modelling will be explored in Section 3.5.

The core ideas and choice of notations in this chapter are inspired by McNeil et al. (2005) in *Quantitative Risk Management* and Herbertsson (2009) in his lecture notes from the course *Credit Risk Modelling*.

### 3.1 Fundamentals of Static Credit Risk Modelling

In this first introductive section of this chapter the setting in static portfolio credit risk models will be explained. In these models the interest lies in modelling the default of one company, but above all modelling the defaults of several companies in a portfolio. Regardless of the choice of model there is a general setup of the problem formulation, and a common set of notations. This section serves to present these common characteristics.

The default of company $i$ is modelled by a default indicator, $Y_i$, which is a random variable in $\{0, 1\}$. The default indicator takes the value 1 if the company defaults before time $T$, and 0 otherwise, i.e.

$$Y_i = \begin{cases} 1 & \text{if obligor } i \text{ defaults before time } T \\ 0 & \text{otherwise} \end{cases}$$
It should be noted that we will not make any specific assumption about the time $T$. However, a reasonable time horizon would be to consider a few years. In order to model the credit loss of an obligor, and finally the loss of an entire portfolio of obligors, we need to introduce some further notations and variables. Let $e_i$ be the *exposure* to obligor $i$, i.e. the notional amount owed by obligor $i$. The exposures are known a priori, and thus considered deterministic. Further, let $l_i$ be the *loss rate* for obligor $i$, which represents the percentage lost if obligor $i$ defaults. The loss rate will in many cases be deterministic and the same for all obligors, but as the models will be extended it can also be modelled stochastically. Regarding the loss rate and exposure, note that $l_i \in (0, 1]$ and $e_i > 0$. With these notations, the loss from obligor $i$ is given by:

$$L_i = l_i \cdot e_i \cdot Y_i.$$  

To model the loss in a portfolio of $m$ obligors we define the variable $L^{(m)}$ as the sum of the individual losses.

$$L^{(m)} = \sum_{i=1}^{m} L_i = \sum_{i=1}^{m} l_i \cdot e_i \cdot Y_i.$$  

(3.1.1)

Without loss of generality it can often be assumed that $e_i$ is equal to 1 for all obligors. It will be clear from the context if this is the case.

As mentioned above, loss rates and exposures are often modelled as the same for all obligors. A portfolio where the quantities are the same for all obligors is called a *Homogeneous* portfolio, and this type of portfolio will frequently be assumed throughout this thesis. In this setting it is often convenient to introduce the random variable $N^{(m)}$, which is the number of defaults in the portfolio, i.e. $N^{(m)} = \sum_{i=1}^{m} Y_i$. The reason for introducing $N^{(m)}$ is due to its close relation to $L^{(m)}$. By setting $l_i = l$ and $e_i = 1$ for all $i$, we see that

$$L^{(m)} = \sum_{i=1}^{m} l \cdot 1 \cdot Y_i = l \cdot \sum_{i=1}^{m} Y_i = l \cdot N^{(m)}.$$  

(3.1.2)

We thus observe that $P(L^{(m)} = l \cdot k) = P(l \cdot N^{(m)} = l \cdot k) = P(N^{(m)} = k)$. In other words, to analyze $L^{(m)}$ in a homogeneous portfolio with constant loss rate it is sufficient for us to study the behavior of $N^{(m)}$. In addition to finding the probability distribution of $N^{(m)}$ we will also be interested in its expected value and variance.

We will in this thesis mainly consider so called *Exchangeable* models, which implies
that the individual default probabilities are the same for all companies. The general definition of exchangeability is wider (see McNeil et al. (2005, p. 344)), but this implication suffices for our purpose. In this setting we have $\pi = P(Y_i = 1)$ for all obligors in the portfolio.

It is possible to model the number of defaults in a portfolio in a simple binomial model, however this model is too simplified and will only serve as an introduction to more sophisticated models. By introducing a dependence among the defaults of the companies, the binomial model can be extended to a so called Bernoulli mixture model. The Bernoulli mixture model can be seen as a conditional binomial model, where the condition on a random factor creates default dependence. Further, the Bernoulli mixture model is the main model that will be used in the category of mixture models. The second category consists of the threshold models where the default of a company occurs if a random variable, often interpreted as the asset value of the company, will fall below some deterministic threshold. As argued before, the threshold models can also be classified as a special case of the mixture model.

The role of dependence modelling was discussed in Section 2.5. The threshold model and the mixture model have different approaches to the actual modelling of the dependence between the default indicators. However, to quantify and measure this dependence it is common in both models to use the standard covariance and correlation, as will be explained later.

### 3.2 Binomial Model

The binomial model is a simplistic approach for modelling the defaults of obligors in a portfolio. It is too unrealistic to be useful in practice, but it is a good start for intuitive purposes and it is a stepping stone towards more advanced models.

In the following model, assume we have a homogeneous portfolio with $m$ obligors, where each obligor can either default or not default up to a constant time $T$. Let $Y_i$ be a Bernoulli random variable which takes the value 1 if company $i$ defaults before time $T$, and 0 otherwise. It is assumed that the default indicators, $Y_1, Y_2, \ldots, Y_m$, are inde-
ependent and identically Bernoulli distributed with parameter $p$, i.e. $P(Y_i = 1) = p$ and $P(Y_i = 0) = 1 - p$. Further, $N^{(m)} = \sum_{i=1}^{m} Y_i$, is the number of defaults in a portfolio with $m$ obligors. Since $N^{(m)}$ is a sum of $m$ independent Bernoulli random variables it is binomially distributed with parameters $m$ and $p$, i.e. $N^{(m)} \sim \text{Bin}(m,p)$. Therefore, $P(N^{(m)} = k) = \binom{m}{k} p^k (1 - p)^{m-k}$ and $E[N_m] = mp$.

In order to analyze the credit portfolio loss, we use the notation from Equation (3.1.2). Let $l$ be the constant credit loss rate in case of a default, identical for all obligors. We are mainly interested in analyzing the total credit loss in a portfolio with $m$ obligors, $L^{(m)}$, which is given by (3.1.2). However, recall that in this setting it is enough for us to study $N^{(m)}$. In Figure 3.1 the probability distribution for the number of defaults is plotted in a portfolio with $m = 50$ obligors and an individual default probability of $\pi = 0.05$.

![Figure 3.1: The probability distribution for the number of defaults in the binomial model.](image)

Note from Figure 3.1 that in the binomial model it is unlikely to have many defaults, i.e. the binomial distribution has thin tails. For instance, in our example $P(N^{(m)} \geq 7) \approx 1.2\%$. Even if the default probabilities are increased the tail will still be thin. This
is due to the assumption that the default indicators are independent. Subsequently, we are interested in finding a model which can generate thicker tails and also take into account the dependence between the obligors. Thicker tails essentially means a higher probability of having more extreme loss scenarios, something which is practically impossible to achieve in the binomial model. In the next section the binomial model will be extended into a Bernoulli mixture model, which addresses both of the mentioned problems. The probability distributions of the number of defaults from the two models will be compared, and it will be seen that the mixture model can generate thicker tails.

### 3.3 Bernoulli Mixture Model

It was mentioned in the previous section that the independence assumption of the default indicators is too simplistic and does not capture reality well. The first step towards a more sophisticated static credit portfolio model is the Bernoulli mixture model which introduces a dependence among the defaults of the obligors. In the Bernoulli mixture model, the dependence is modelled by a factor vector, which can be seen as a set of common macroeconomic variables, e.g. interest rates and stock index prices. These macroeconomic variables are also modelled stochastically. Conditional on the factor vector the individual defaults of the obligors are assumed to be independent. Definition 3.3.1 is taken from McNeil et al. (2005) and formally defines a Bernoulli mixture model.

**Definition 3.3.1 (Bernoulli Mixture Model)**

Given some $p < m$ and a $p$-dimensional random vector $\Psi = (\Psi_1, \ldots, \Psi_p)'$, the random vector $Y = (Y_1, \ldots, Y_m)'$ follows a Bernoulli mixture model with factor vector $\Psi$ if there are functions $p_i : \mathbb{R}^p \to [0, 1], 1 \leq i \leq m$, such that conditional on $\Psi$ the components of $Y$ are independent Bernoulli rvs satisfying $P(Y_i = 1 \mid \Psi = \psi) = p_i(\psi)$.

The definition states that in a Bernoulli mixture model the probability that an individual company defaults given the factor vector is a function of the factors. Further, the definition also says that conditional on those factors the elements of the default indicator $Y$ are independent.

Let $y = (y_1, \ldots, y_m)'$, where $y_i \in \{0, 1\}$ for every obligor $i$, be a vector representing
which obligors defaulted and which survived. Since the default indicators are conditionally independent it follows that:

\[ P(Y = y | \Psi = \psi) = \prod_{i=1}^{m} P(Y_i = y_i | \Psi = \psi) = \prod_{i=1}^{m} p_i(\psi)^{y_i} (1 - p_i(\psi))^{1-y_i}. \]  

(3.3.1)

The last equality is due to the fact that given the factors the individual default indicators follow a Bernoulli distribution with parameter \( p_i(\psi) \). By integrating over all possible values of the factors the unconditional distribution, \( P(Y = y) \), of the default indicator is obtained. Below, this will be explained further in a one-factor exchangeable Bernoulli mixture model.

It is also possible to approximate the Bernoulli random variables with Poisson random variables. The Poisson mixture model will not be discussed in this thesis, but the definition is analogous to that of a Bernoulli mixture model; the main difference being that the default vector is Poisson distributed instead of Bernoulli distributed. For more on this see Section 8.4 of McNeil et al. (2005).

### 3.3.1 One-Factor Exchangeable Bernoulli Mixture Model

In this section we let the default indicators for the obligors depend on one single common factor. The reason for this is not only for simplicity of calculation. In many cases it can also be difficult to calibrate a model with more factors because of lack of information. Further, we consider an exchangeable model, which implies that all the individual default probabilities \( p_i \) are the same.

To emphasize that the companies have the same default functions we define \( p(\Psi) := p_i(\Psi) \) for all \( i \). The random variable \( p(\Psi) \) is called a mixing variable. Thus we have that \( P(Y_i = 1 | \Psi) = p(\Psi) \), for all obligors, and this implies that the unconditional probability \( P(Y_i = 1) = \mathbb{E}[p(\Psi)] \). To see this consider the following calculation:

\[
\pi = P(Y_i = 1) = 1 \cdot P(Y_i = 1) + 0 \cdot P(Y_i = 0) = \mathbb{E}[Y_i] = \mathbb{E}[\mathbb{E}[Y_i | \Psi]] \\
= \mathbb{E}[1 \cdot P(Y_i = 1 | \Psi) + 0 \cdot P(Y_i = 0 | \Psi)] = \mathbb{E}[P(Y_i = 1 | \Psi)] = \mathbb{E}[p(\Psi)].
\]  

(3.3.2)

The calculation uses the definition of expected value and the law of iterated expectation.
We now turn to the problem of finding the probability of a certain number of defaults. Recall from last section that $N^{(m)}$ is defined as the number of defaults in a portfolio of $m$ obligors. Thus, we are interested in finding the unconditional probability $P(N^{(m)} = k)$ for $k = 0, 1, \ldots, m$, and in order to find this we start out with the conditional probability $P(N^{(m)} = k \mid \Psi = \psi)$. The argument from Section 3.3 states that the individual default probabilities given the factors are Bernoulli distributed. In this one-factor exchangeable model this implies that the number of defaults $N^{(m)}$ is conditionally binomial distributed with parameters $m$ and $p(\psi)$, since it is the sum of $m$ independent Bernoulli trials with parameter $p(\psi)$, i.e. we have that:

$$P(N^{(m)} = k \mid \Psi = \psi) = \binom{m}{k} p(\psi)^k (1 - p(\psi))^{m-k}.$$ 

In order to find the unconditional probability of the number of defaults we need to integrate over $p(\psi)$, (which is between 0 and 1 since it is a probability) and we get:

$$P(N^{(m)} = k) = \binom{m}{k} \int_0^1 p(\psi)^k (1 - p(\psi))^{m-k} dG(p(\psi)). \quad (3.3.3)$$

where $G$ is the distribution function for the mixing variable $p(\Psi)$, that is $G(x) = P(p(\Psi) \leq x)$. Note here that the mixing variable $p(\Psi)$ plays an important role in order to calculate default probabilities and the probability for the number of defaults. Common distributions for the mixing variables include Beta mixing distribution, Probit-normal mixing distribution and Logit-normal mixing distribution (Frey and McNeil, 2003). These mixing distributions will be explored further in the next chapter.

By using Equation (3.3.3) we can compute the probability distribution for the number of defaults in a Bernoulli mixture model. For example, if we assume that the mixing variable is beta distributed with parameters $a$ and $b$, Equation (3.3.3) boils down to:

$$P(N^{(m)} = k) = \binom{m}{k} \frac{\beta(a + k, b + m - k)}{\beta(a, b)}, \quad a, b > 0$$

where $\beta(a, b) = \int_0^1 z^{a-1}(1 - z)^{b-1} dz, \quad 0 < z < 1$.

Figure 3.2 illustrates the difference between a binomial model and a Bernoulli mixture model with a beta mixing distribution, where both models have the same individual default probability. As can be seen, the mixture model can produce thicker tails than the binomial model.
**Figure 3.2:** The probability distribution of the number of defaults in a binomial model and a Bernoulli mixture model with beta mixing variable. The individual default probability is given by $\pi = 0.05$ in both models, and the portfolio size is $m = 50$ obligors.

**Correlation.** Recall that the reason to go from the simple binomial model to the mixture model was to introduce a dependence among the obligors. The correlation between two default indicators in the binomial model is obviously zero since they are independent. In the mixture model they are only conditionally independent and we claim that the default correlation is nonnegative and determined entirely by the mixing distribution. Following the notation of McNeil et al. (2005), we define the general joint default probability for $k$ firms as:

$$\pi_k = P(Y_{i_1} = 1, \ldots, Y_{i_k} = 1)$$

where $\{i_1, \ldots, i_k\}$ is an arbitrary subset of $\{1, \ldots, m\}$ for $k \in \{2, \ldots, m\}$. In view of Equation (3.3.2) we sometimes write $\pi$ instead of $\pi_1 = P(Y_i = 1)$. With this notation
we can write the correlation between two default indicators as:
\[
\text{Corr}(Y_i, Y_j) = \frac{\text{Cov}(Y_i, Y_j)}{\sqrt{\text{Var}(Y_i)\text{Var}(Y_j)}} = \frac{\mathbb{E}[Y_i Y_j] - \mathbb{E}[Y_i] \mathbb{E}[Y_j]}{\text{Var}(Y_i)} = \frac{\mathbb{E}[Y_i Y_j] - \mathbb{E}[Y_i] \mathbb{E}[Y_j]}{\mathbb{E}[Y_i^2] - \mathbb{E}[Y_i]^2}
\]
\[
= \frac{\pi_2 - \pi^2}{\pi - \pi^2}. \quad (3.3.4)
\]

Equation (3.3.4) is of general form, and only the definition of correlation and our notation is used. In the Bernoulli mixture model an analogous calculation to that of Equation (3.3.2) yields:
\[
\pi_k = P(Y_{i_1} = 1, \ldots, Y_{i_k}) = \mathbb{E}[\mathbb{E}[Y_1 \cdots Y_k | \Psi]] = \mathbb{E}[p(\Psi)^k].
\]

Consequently, in a Bernoulli mixture framework the correlation is thus given by:
\[
\text{Corr}(Y_i, Y_j) = \frac{\pi_2 - \pi^2}{\pi - \pi^2} = \frac{\mathbb{E}[p(\Psi)^2] - \mathbb{E}[p(\Psi)]^2}{\mathbb{E}[p(\Psi)]^2 - \mathbb{E}[p(\Psi)]^2} = \frac{\text{Var}(p(\Psi))}{\mathbb{E}[p(\Psi)] (1 - \mathbb{E}[p(\Psi)])} \geq 0.
\]

Conclusively, we have seen that in the Bernoulli mixture model there is a dependence among the defaults of the obligors, which is completely determined by the choice of mixing variable. In fact, the correlation between \(Y_i\) and \(Y_j\), for any two pairs \(i \neq j\), is completely determined by the first two moments of the mixing variable. In chapter 4 we will come back to the Bernoulli mixture model when applying it in a portfolio credit loss framework.

### 3.3.2 Asymptotic Behavior in Large Portfolios

This section will present some results regarding the asymptotic behavior of large portfolios in a Bernoulli mixture setting. For instance, in one-factor Bernoulli mixture models, both Frey and McNeil (2003) and Herbertsson (2009) suggest that the tail of the mixing distribution practically determines the tail of the loss distribution. This strong result will be investigated further, along with some other findings regarding the properties of Bernoulli mixture models.

At this point we are familiar with the default indicator variable \(Y_i\) from the previous sections. We again implement these indicators into a framework of total portfolio credit loss, which is given by Equation (3.1.1). For simplicity of calculations, assume that \(e_i = 1\), for all obligors. The loss rate on the other hand can be considered random, though doing so will complicate the model significantly. With regard to this, Frey and
McNeil (2003) point out that loss rates given default, $l_i$, and default indicators, $Y_i$, should not be assumed to be independent. This is rather intuitive, since one could expect companies to have greater difficulties recovering in times of a financial crisis than in times of normal growth. Hence, the loss rate, $l_i$, can be modelled so that it is affected by the background economic factors $\Psi$. However, in the model considered here, it is assumed that the loss rates are deterministic and the same for all obligors and given by $l$.

As we have seen in the previous section, Bernoulli mixture models are extensions of the more simple binomial model, the major difference being the introduction of a dependence between defaults. From the binomial setting we will therefore make further use of the variable $N^{(m)} = \sum_{i=1}^{m} Y_i$, which is the number of defaults in the portfolio. Also, in the following analysis we will consider an exchangeable model, i.e. the individual default probabilities are the same for all obligors; that is, $p_i(\Psi) = p(\Psi)$, for all $i$. Recall from Section 3.1 that when studying $L^{(m)}$ under these assumptions, it is enough for us to study $N^{(m)}$. Since the random variable $N^{(m)}$ is created by i.i.d. Bernoulli distributed random variables (independent when conditional on the factor vector) it would be convenient to apply the law of large numbers and thereby analyze the asymptotic behavior. To this end, let us focus on the ratio $N^{(m)}/m$, which can be viewed as the fraction of the number of defaults in the portfolio (also called the default fraction). When $e_i = 1$ for all obligors, the ratio can be also be seen as the average loss for each obligor. To proceed we first notice that conditional on $\Psi$ the random variables $Y_1, \ldots, Y_m$ are i.i.d. with default probability $p(\Psi)$. It can be shown (see Frey and McNeil (2003)) that given an outcome of $\Psi$, the law of large numbers implies:

$$\lim_{m \to \infty} \frac{N^{(m)}}{m} = p(\Psi) \text{ a.s. under the measure } P(\cdot | \Psi)$$

which implies that the event $\lim_{m \to \infty} \frac{N^{(m)}}{m} = p(\Psi)$ given an outcome of $\Psi$ has probability one, i.e.:

$$P \left( \lim_{m \to \infty} \frac{N^{(m)}}{m} = p(\Psi) \Bigm\vert \Psi \right) = 1.$$

More specifically, almost sure convergence also implies convergence in distribution, i.e.
\begin{align}
\lim_{m \to \infty} P \left( \frac{N^{(m)}}{m} \leq x \right) = P \left( p(\Psi) \leq x \right). \quad \text{Inspired by the outline of Lando (2004), this can be shown in the following intuitive way:} \\
\lim_{m \to \infty} P \left( \frac{N^{(m)}}{m} \leq x \ \bigg| \Psi \right) = \begin{cases} 0 & \text{if } p(\Psi) > x \\ 1 & \text{if } p(\Psi) \leq x \end{cases} = \mathbb{I}_{\{p(\Psi) \leq x\}}. \tag{3.3.5}
\end{align}

By using the same reasoning as before, we can derive the unconditional probability,

\[ P \left( \frac{N^{(m)}}{m} \leq x \right) = \mathbb{E} \left[ P \left( \frac{N^{(m)}}{m} \leq x \ \bigg| \Psi \right) \right]. \tag{3.3.6} \]

Finally, by inserting Equation (3.3.5) in Equation (3.3.6) and letting \( m \to \infty \) we see that

\[ \lim_{m \to \infty} P \left( \frac{N^{(m)}}{m} \leq x \right) = \mathbb{E} \left[ \mathbb{I}_{\{p(\Psi) \leq x\}} \right] = P(\Psi) \leq x) = G(x). \tag{3.3.7} \]

Thus, the distribution of the default fraction converges to the mixing distribution as the number of obligors becomes larger. The fact that we in large portfolios can approximate the distribution of the default fraction with the mixing distribution will be used frequently throughout this thesis. For example, in Section 3.4.2 we will derive the widely used large portfolio approximation in a mixture model inspired by the Merton model. Further, Equation (3.3.7) verifies our preconceived notion that the tail of the mixing distribution practically determines the distribution of the default fraction, \( N^{(m)}/m \), and in turn also the tail of the loss distribution, \( L^{(m)} \). In other words, if the mixing variable has a thick tail, allowing for more extreme events, then also the loss distribution will be given a thicker tail and give higher probabilities for extreme events.

Frey and McNeil (2003) present a more general approach to the analysis of asymptotic behavior in large portfolios. In particular, they relax the assumptions about the model being exchangeable and the loss rate being constant. This makes the analysis slightly more technical, but they still arrive at the same conclusions. In the end, Frey and McNeil (2003) also presents a strong result regarding one-factor Bernoulli mixture models, where they are able to link the quantiles of the loss distribution to the quantiles of the mixing distribution.

### 3.4 Threshold Models Using The Mixture Approach

In the Bernoulli mixture model the defaults were modelled in a binomial framework, and a default of an obligor occurred with a given probability without being concerned with
the actual mechanism of default. However, in the threshold models there is a mechanism which describes why there is a default. Common for all threshold models is that a default occurs if some stochastic process falls below a threshold. This stochastic process is often interpreted as the obligor’s asset value. We earlier categorized the threshold models into two approaches, and in this section the first category of threshold models using the mixture approach is discussed. This category builds to a large extent from the framework of the Merton model. It will be seen that even though the mechanism of default differ substantially in the Bernoulli mixture model and this type of threshold model there will be strong similarities in the analysis. This is also the reason behind the title of this section.

3.4.1 The Merton Framework

In the year 1974, the American economist Robert C. Merton published a paper emphasizing the credit risk of companies, i.e. the risk that a company may not be able to meet its financial obligations and pay its debts, see Merton (1974). The so called Merton model proposed a new way of analyzing this type of risk, and the model was soon applied to other areas within finance; the most famous application being to the Black-Scholes model. The core idea behind the Merton model is to make use of the company’s balance sheet. The balance sheet can be divided into three parts; assets, debt and equity. Consider the total assets of a company, and assume that it follows a stochastic process \( V_t, 0 \leq t \leq T \). In order to finance its assets/investments the company takes on debt and issues equity. The Merton model assumes that the total debt of a company can be expressed as a single zero-coupon bond with face value \( \bar{D} \) and maturity \( T \). It is also assumed that companies cannot take on new debt or pay dividends to its stock owners. Let the debt and equity of the company at time \( t \) be denoted by \( D_t \) and \( E_t \) respectively. Finally, by assuming that there are no transaction costs or taxes we must have that the company’s assets is the sum of its debt and equity, i.e. \( V_t = D_t + E_t \), for \( 0 \leq t \leq T \). In this framework, a company will default at time \( T \) if the value of its assets is less than the value of its outstanding debt, i.e. if \( V_T < \bar{D} \). Note that it is assumed that a company can only default exactly on time \( T \), when the debt matures.

\[ \text{Due to his contribution to option pricing theory, Robert C. Merton received the Nobel Memorial Prize in Economic Sciences together with Myron Scholes in 1997. Fischer Black, the third person behind the Black-Scholes model, had tragically died from cancer in 1995.} \]
and not before $T$. Further, note that $D$ is the predetermined notional amount owed which must be paid at time $T$, while $D_t$ is a process which describes the value of the debt (the zero-coupon bond) at time $t$, where $t \leq T$. The zero-coupon price $D_t$ can change over time due to several factors; the most imminent factor being interest rates. For instance, a rise in interest rates will make the bond value go down.

In simplified terms, if the company defaults the debt owners will receive $V_T$, which is less than $\bar{D}$; the amount that was actually owed. The stock owners on the other hand will receive nothing. If the company does not default, then the debt owners will receive their whole claim $\bar{D}$, and the stock owners are left with $V_T - \bar{D}$. These possible payoffs for the debt and stock owners can be described in terms of the payoffs of call and put options.

$$D_T = \min(V_T, \bar{D}) = \bar{D} - (\bar{D} - V_T)^+$$

$$E_T = \max(V_T - \bar{D}, 0) = (V_T - \bar{D})^+$$

In other words, the value of the company’s debt at time $T$ is the nominal amount owed minus the payoff of a European put option on $V_T$ with strike price $\bar{D}$ and time to maturity $T$. The value of the equity on the other hand is the payoff of a European call option on $V_T$. The payoff for the debt and stock owners is illustrated in Figure 3.3.
Before further describing the asset value process $V_t$ we make one more important assumption. Instead of operating under the real-world probability measure $\mathbb{P}$ the following analysis of the Merton framework will be done under the risk-neutral probability measure $\tilde{\mathbb{P}}$. The basic difference between these probability measures has to do with the variable $\mu_V$, which describes the mean rate of change in the asset value process $V_t$. Under $\mathbb{P}$, $\mu_V$ can be estimated from historical data, while under the risk-neutral probability measure $\mu_V$ is set to equal the risk-free interest rate, $r$. (A more technical difference is that under the risk-neutral measure all discounted asset value processes become martingales, which is used in the Merton model when pricing securities which pay-offs depend on the asset value process $V_t$. However, these applications of the Merton model will not be further explored.) For more on this, see for example Shreve (2006).

The Merton model assumes that the asset value process follows a geometric Brownian motion (which is the same assumption made about stock price processes in the Black-Scholes model). In other words, $V_t$ has the following dynamics,

$$dV_t = rV_t dt + \sigma_V V_t dW_t$$

(3.4.1)
where \( r \geq 0, \sigma_V > 0, \) and \( W_t \) is a standard Brownian motion. The parameter \( \sigma_V \) denotes the volatility of the assets, and can either be estimated from historical data or extracted from stock prices. By setting \( t = T \) and solving Equation (3.4.1) for \( V_T \), by using Itô’s Lemma, we get that \( V_T = V_0 \cdot \exp \left( (r - \frac{\sigma_V^2}{2})T + \sigma_V W_T \right) \). Further, note that since \( W_T \) is a standard Brownian motion we have that \( W_T \sim N(0, T) \). By dividing this process by its standard deviation, the process becomes standard normal, i.e. \( \frac{W_T}{\sqrt{T}} \sim N(0, 1) \). To simplify the notation we introduce the random variable \( Z \sim N(0, 1) \), and let \( W_T = \sqrt{T}Z \). This observation will be used in the following computations, where we also recall that the company defaults if \( V_T < \bar{D} \). The probability of default can then be expressed as

\[
P \left( V_T < \bar{D} \right) = P \left( \ln(V_T) < \ln(\bar{D}) \right) = P \left( \ln(V_0) + \left( r - \frac{\sigma_V^2}{2} \right) T + \sigma_V \sqrt{T}Z < \ln(\bar{D}) \right)
\]

\[
= P \left( \sigma_V \sqrt{T}Z < \ln \left( \frac{\bar{D}}{V_0} \right) - \left( r - \frac{\sigma_V^2}{2} \right) T \right)
\]

\[
= P \left( Z < \frac{\ln \left( \frac{\bar{D}}{V_0} \right) - \left( r - \frac{\sigma_V^2}{2} \right) T}{\sigma_V \sqrt{T}} \right)
\]

\[
= \Phi \left( \frac{\ln \left( \frac{\bar{D}}{V_0} \right) - \left( r - \frac{\sigma_V^2}{2} \right) T}{\sigma_V \sqrt{T}} \right)
\]

where \( \Phi \) denotes the standard normal distribution function. It is intuitive that the default probability is increasing in \( \bar{D} \) and this is also clear from the equation above.

Before moving on and applying the Merton model to static portfolio credit risk, some small observations are in order. The Merton framework makes some strong assumptions and simplifications which are needed to make the model mathematically tractable. Every assumption made takes us one step further away from reality, but the real world is obviously much more complex. However, the Merton model has proved to be a stepping stone towards more advanced models as will be seen later on.

### 3.4.2 Mixture Models in the Merton Framework

We build on the Merton framework from the last section and implement the fundamental ideas in a credit portfolio setting. The analysis in this section is based on the lecture notes from Herbertsson (2009). Consider a credit portfolio with \( m \) obligors, and let the
assets of obligor $i$ at time $t$ be denoted by $V_{t,i}$. To introduce a dependence among the obligors the idea is to let the asset values be driven by two stochastic processes; one that is common for all obligors and one process that is individual. Thus, for constants $r$, $\sigma_i$ and $\rho$, we let $V_{t,i}$ have the following dynamics:

$$dV_{t,i} = rV_{t,i}dt + \sigma_i V_{t,i}dB_{t,i}, \quad r \geq 0, \sigma_i > 0 \quad (3.4.2)$$

where $B_{t,i}$ is the stochastic process given by

$$B_{t,i} = \sqrt{\rho}W_{t,0} + \sqrt{1-\rho}W_{t,i}, \quad \rho \geq 0$$

where $W_{t,0}, W_{t,1}, \ldots, W_{t,m}$ are independent standard Brownian motions. With this definition $B_{t,i}$ is also a standard Brownian motion. Further, it can be shown that $\text{Corr} (B_{t,i}, B_{t,j}) = \rho$ for $i \neq j$. Thus, there is a clear dependence among the obligors asset values $V_{t,i}$, which stems from the common stochastic process $W_{t,0}$. We interpret the constant $\rho$ as a correlation parameter. For now this interpretation will suffice, the exact role and implications of $\rho$ will be discussed further on. It is important to remember that $\rho$ is not the correlation between the default indicators, but rather a correlation parameter affecting the correlation between the default indicators.

Furthermore, to make a distinction between the common process and the individual process we let $W_{t,0} = \sqrt{t}\Psi$ and $W_{t,i} = \sqrt{t}Z_i$ for $1 \leq i \leq m$, where $\Psi, Z_i \sim N(0,1)$. Recall that $\Psi$ was the common factor in the last sections about Bernoulli mixture models. As in the previous section $\Psi$ can be interpreted as a random variable modelling the economic climate, which is common for all obligors. To recapitulate, with this notation the dynamics of $V_{t,i}$ in Equation (3.4.2) becomes

$$dV_{t,i} = rV_{t,i}dt + \sigma_i V_{t,i}dB_{t,i}, \quad \text{where } B_{t,i} = \sqrt{\rho t}\Psi + \sqrt{t(1-\rho)}Z_i. \quad (3.4.3)$$

In this model the default of obligor $i$ occurs if its asset value at time $T$ is below some deterministic threshold, $\bar{D}_i$, often interpreted as the debt level for obligor $i$. Thus, $Y_i = 1 \Leftrightarrow V_{T,i} < \bar{D}_i$. Following the discussion from Section 3.4.1 the solution to Equation (3.4.3) is given by:

$$V_{t,i} = V_{0,i} \exp \left( \left( r - \frac{\sigma_i^2}{2} \right) t + \sigma_i \left( \sqrt{\rho t}\Psi + \sqrt{t(1-\rho)}Z_i \right) \right).$$
Finally, the event that obligor \(i\) defaults, i.e. its asset value at time \(T\) is below \(\bar{D}_i\), is equivalent to:

\[
V_{T,i} < \bar{D}_i \iff V_{0,i} \exp \left( (r - \frac{\sigma_i^2}{2})T + \sigma_i \left( \sqrt{\rho T} \Psi + \sqrt{T(1-\rho)} Z_i \right) \right) < \bar{D}_i
\]

\[
\iff \ln (V_{0,i}) + (r - \frac{\sigma_i^2}{2}) T + \sigma_i \sqrt{T} \left( \sqrt{\rho \Psi} + \sqrt{1-\rho} Z_i \right) < \ln \left( \bar{D}_i \right)
\]

\[
\iff \sqrt{\rho \Psi} + \sqrt{1-\rho} Z_i < \frac{\ln \left( \frac{\bar{D}_i}{V_{0,i}} \right) - \left( r - \frac{\sigma_i^2}{2} \right) T}{\sigma_i \sqrt{T}}
\]

\[
\iff Z_i < \frac{C_i - \sqrt{\rho \Psi}}{\sqrt{1-\rho}}
\]

where \(C_i = \frac{\ln \left( \frac{\bar{D}_i}{V_{0,i}} \right) - \left( r - \frac{\sigma_i^2}{2} \right) T}{\sigma_i \sqrt{T}}\) \hspace{1cm} (3.4.4)

Note that all the quantities in \(C_i\) are deterministic and known at time 0. Thus \(C_i\) is a constant, and in fact \(C_i\) can be calculated explicitly, as will be seen below. First we will be concerned with finding the probability of a default in this model. A default occurs if the asset value for obligor \(i\), at time \(T\), is below its debt level \(\bar{D}_i\). From the above calculations we concluded that obligor \(i\) defaults if and only if the inequality (3.4.4) holds. However, the event of a default depends on the common economic variable \(\Psi\), which is standard normal distributed. Conditional on \(\Psi\) the probability that obligor \(i\) defaults is standard normal distributed and given by:

\[
p_i(\Psi) = P(Y_i = 1 \mid \Psi) = P \left( Z_i < \frac{C_i - \sqrt{\rho \Psi}}{\sqrt{1-\rho}} \right) = \Phi \left( \frac{C_i - \sqrt{\rho \Psi}}{\sqrt{1-\rho}} \right), \hspace{1cm} (3.4.5)
\]

where \(\Phi(x)\) is the distribution function of a standard normal random variable. The third equality follows from the fact that \(Z_i\) is independent of \(\Psi\) by definition, for all \(i\). Further we know from Equation (3.3.2) that the unconditional default probability, \(p_i\), for obligor \(i\) is given by:

\[
p_i = P \left( Y_i = 1 \right) = E \left[ P \left( Y_i = 1 \mid \Psi \right) \right].
\]

We now proceed with finding an expression for the quantity \(C_i\). In the beginning of this section we claimed that \(B_{t,i}\), defined as above, is a standard Brownian motion. Thus it has the same distribution as \(\sqrt{t} Z_i\), where \(Z_i \sim N(0,1)\) as usual. With this notation
the probability that obligor $i$ defaults is equivalent to:

\[ p_i = P(Y_i = 1) = P\left(V_{T,i} < \bar{D}_i\right) = P\left(V_{0,i} \exp \left((r - \frac{\sigma_i^2}{2})T + \sigma_i \sqrt{T} Z_i\right) < \bar{D}_i\right) \]

\[ = P\left(Z_i < \frac{\ln \left(\frac{\bar{D}_i}{V_{0,i}}\right) - \left(r - \frac{\sigma_i^2}{2}\right)T}{\sigma_i \sqrt{T}}\right) = P\left(Z_i < C_i\right) = \Phi(C_i). \]

Thus, from this we conclude that $C_i$ is given by $p_i = \Phi(C_i) \Leftrightarrow C_i = \Phi^{-1}(p_i)$.

**Credit Portfolio.** We are now ready to enter the static credit portfolio framework, analogous to previous sections. We start by considering an exchangeable model, i.e. all obligors have the same default probability, $\pi = p_i$ and $p(\psi) = p_i(\psi)$, for all $i$. In order for the default probabilities to be equal, all quantities involved need to be identical for all obligors. To this end, we make use of the notation $C = C_i$. We consider a homogeneous portfolio, and thus we let the loss rate $l$ be constant and the same for all obligors. In this setting we have that the portfolio loss is given by $L^{(m)} = l \cdot N^{(m)}$, where $N^{(m)}$ is the number of defaults in the portfolio. As before it is sufficient to study the behavior of $N^{(m)}$. From Equation (3.4.5) we have that the default probability is given by:

\[ p(\psi) = \Phi\left(\frac{\Phi^{-1}(\pi) - \sqrt{p(\psi)}}{\sqrt{1 - \rho}}\right). \quad (3.4.6) \]

Thus, we have as before that conditional on $\Psi = \psi$ the default indicator $Y_i$ is a Bernoulli random variable with parameter $p(\psi)$. Along with the results of subsection 3.3.1 this implies that the conditional probability of $k$ defaults in the portfolio follows a binomial distribution with parameters $m$ and $p(\psi)$, i.e.:

\[ P\left(N^{(m)} \leq k \mid \Psi = \psi\right) = \sum_{i=0}^{k} P\left(N^{(m)} = i \mid \Psi = \psi\right) = \sum_{i=0}^{k} \binom{m}{i} p(\psi)^i (1 - p(\psi))^{m-i} \]

and thus the unconditional distribution is given by:

\[ P\left(N^{(m)} \leq k\right) = \sum_{i=0}^{k} \binom{m}{i} \int_{-\infty}^{\infty} p(u)^i (1 - p(u))^{m-i} f_\Psi(u) du \quad (3.4.7) \]

where $f_\Psi$ is the density function of a standard normal random variable. Since we know $p(\psi)$, the right hand side of Equation (3.4.7) can be expanded into the following rather
messy expression:

\[
P (N^{(m)} \leq k) = \sum_{i=0}^{k} \binom{m}{i} \int_{-\infty}^{\infty} \Phi \left( \frac{\phi^{-1}(\pi) - \sqrt{\rho}u}{\sqrt{1-\rho}} \right) \\
\cdot \left( 1 - \Phi \left( \frac{\phi^{-1}(\pi) - \sqrt{\rho}u}{\sqrt{1-\rho}} \right) \right)^{m-i} \int_{-\infty}^{\infty} f_{\Psi}(u)du.
\] (3.4.8)

In order to calculate the probability of \( k \) defaults one has to numerically compute the expression on the right hand side of Equation (3.4.8). However, conveniently enough there also exists a large portfolio approximation for the fraction of the number of defaults in the portfolio.

**Large Portfolio Approximation.** We know from subsection 3.3.2 that in the framework of Bernoulli mixture models the distribution of the default fraction converges to the distribution of the mixing variable. This holds analogously in this setting, i.e. in the Merton model the distribution of the default fraction converges to the distribution of \( p(\Psi) \). Thus, if we let \( G \) be the distribution function for \( p(\Psi) \), i.e. \( G(x) = P(p(\Psi) \leq x) \), we have that:

\[
\lim_{m \to \infty} P \left( \frac{N^{(m)}}{m} \leq x \right) = G(x).
\] (3.4.9)

What now remains is to find an explicit expression for \( G(x) \). From Equation (3.4.6) and (3.4.9) we get:

\[
G(x) = P(p(\Psi) \leq x) = P \left( \Phi \left( \frac{\phi^{-1}(\pi) - \sqrt{\rho}\Psi}{\sqrt{1-\rho}} \right) \leq x \right) \\
= P \left( \phi^{-1}(\pi) - \sqrt{\rho} \Psi \leq \phi^{-1}(x) \right) \\
= P \left( -\Psi \leq \frac{1}{\sqrt{\rho}} \left( \sqrt{1-\rho} \phi^{-1}(x) - \phi^{-1}(\pi) \right) \right) \\
= P \left( \Psi \leq \frac{1}{\sqrt{\rho}} \left( \sqrt{1-\rho} \phi^{-1}(x) - \phi^{-1}(\pi) \right) \right) \\
= \phi \left( \frac{1}{\sqrt{\rho}} \left( \sqrt{1-\rho} \phi^{-1}(x) - \phi^{-1}(\pi) \right) \right).
\]

The second last equality follows from the fact that \( -\Psi \) and \( \Psi \) have the same distribution since it is symmetric around 0. Formally, we have that \( P(-\Psi \leq x) = P(\Psi \geq -x) = 1 - P(\Psi \leq -x) = 1 - \Phi(-x) = \Phi(x) \), where the last equality is due to the symmetry
of \( \Phi \). Finally, an approximation for the fraction of the number of defaults is given by:

\[
P \left( \frac{N^{(m)}}{m} \leq x \right) \approx \Phi \left( \frac{1}{\sqrt{\rho}} \left( \sqrt{1 - \rho \Phi^{-1}(x)} - \Phi^{-1}(\pi) \right) \right).
\] (3.4.10)

Equation (3.4.10) is the Large Portfolio Approximation (LPA) in the Merton framework. The LPA in the Merton framework was originally developed by Oldrich Vasicek at KMV, see Vasicek (1991). Today the LPA is widely used in the industry for evaluating large static credit portfolios, see for example Schönbucher (2002).

**Correlation.** We now consider the correlation in the Merton LPA model. It was clear from the definition of \( B_{t,i} \) in (3.4.2) that there is a dependence between the asset values of the obligors, and it was concluded that the correlation between \( B_{t,i} \) and \( B_{t,j} \) was \( \rho \). However, the major concern is not the correlation between the two Brownian motions, but rather the correlation between the default indicators \( Y_i \) and \( Y_j \). As in the Bernoulli mixture model in subsection 3.3.1, the correlation between the default indicators in a homogeneous model can be expressed as in Equation (3.3.4), that is:

\[
\text{Corr}(Y_i, Y_j) = \frac{\pi_2 - \pi^2}{\pi - \pi^2} = \frac{\text{Var}(p(\Psi))}{\mathbb{E}[p(\Psi)] (1 - \mathbb{E}[p(\Psi)])} \geq 0.
\]

In the same way as in the Bernoulli mixture model, the default correlation is determined by the mixing distribution. In the mixture model inspired by Merton we have that the mixing variable in a homogeneous portfolio is given by (3.4.6). However, we run into difficulties when trying to calculate the correlation explicitly. The main reason for this is that it is difficult to analytically express \( \pi_2 \) in terms of \( \pi \) and \( \rho \). Recall that the motivation behind the large portfolio approximation was that it was hard to find a closed form expression for the default distribution. As mentioned above, the situation is the same when we want to calculate the correlation. Thus, we cannot in an easy way express the correlation between defaults in the Merton model. However, we can show that if \( \rho = 0 \) then the default correlation is zero, and if \( \rho > 0 \) then the default correlation is non-zero. To this end, consider a homogeneous portfolio with \( m \) obligors, each with individual default probability \( \pi \). Let the asset value process \( V_{t,i} \) of obligor \( i \) follow the dynamics in Equation (3.4.2), where \( r, \sigma \) and \( \rho \) are nonnegative parameters.

**Proposition 3.4.1 (Correlation in the Merton model)**

Consider the portfolio framework above. For \( i, j = 1, \ldots, m, \ i \neq j \), the correlation
between the default indicators depend on \( \rho \) in the following way:

\[
\rho = 0 \Rightarrow \text{Corr}(Y_i, Y_j) = 0 \\
\rho \neq 0 \Rightarrow \text{Corr}(Y_i, Y_j) \neq 0.
\]

**Proof:**

First, let \( \rho = 0 \). From the dynamics of \( B_{T,i} \) we then have that \( B_{T,i} = W_{T,i} \), where \( W_{T,1}, \ldots, W_{T,m} \) are independent standard Brownian motions. Thus, \( \text{Corr}(B_{T,i}, B_{T,j}) = \text{Corr}(W_{T,i}, W_{T,j}) = 0 \). To prove that the correlation between the default indicators is zero, it is enough to show that their covariance is zero.

\[
\text{Cov}(Y_i, Y_j) = \mathbb{E}[Y_i Y_j] - \mathbb{E}[Y_i] \mathbb{E}[Y_j] = P(Y_i = 1, Y_j = 1) - P(Y_i = 1)P(Y_j = 1)
= P(V_{T,i} < \bar{D}, V_{T,j} < \bar{D}) - \pi^2 = P(B_{T,i} < \sqrt{T}C, B_{T,j} < \sqrt{T}C) - \pi^2
= P(W_{T,i} < \sqrt{T}C, W_{T,j} < \sqrt{T}C) - \pi^2
= P(W_{T,i} < \sqrt{T}C) P(W_{T,j} < \sqrt{T}C) - \pi^2 = \pi^2 - \pi^2 = 0.
\]

This implies that the correlation between the default indicators is zero if \( \rho = 0 \).

Second, let \( \rho \neq 0 \). To show that \( \text{Corr}(Y_i, Y_j) \neq 0 \) it suffices to show that \( \text{Cov}(Y_i, Y_j) \neq 0 \).

We first note that

\[
\text{Cov}(B_{T,i}, B_{T,j}) = E[B_{T,i}B_{T,j}] - E[B_{T,i}] E[B_{T,j}]
= E[\left(\sqrt{\rho}W_{T,0} + \sqrt{1-\rho}W_{T,i}\right) \left(\sqrt{\rho}W_{T,0} + \sqrt{1-\rho}W_{T,j}\right)] - 0
= \rho E\left[W_{T,0}^2\right] = \rho T.
\]

Next, with a similar reasoning as in the first part of this proof, the covariance between the default indicators is:

\[
\text{Cov}(Y_i, Y_j) = P(V_{T,i} < \bar{D}, V_{T,j} < \bar{D}) - \pi^2
\neq P(V_{T,i} < \bar{D}) P(V_{T,j} < \bar{D}) - \pi^2 = \pi^2 - \pi^2 = 0.
\]

The inequality follows from the fact that \( \text{Cov}(B_{T,i}, B_{T,j}) = \rho T \). Thus, if \( \rho \neq 0 \), then \( \text{Corr}(Y_i, Y_j) \neq 0 \). \( \blacksquare \)

From Proposition 3.4.1 it can be concluded that \( \rho \) affects the correlation between defaults. We end this section with illustrating this in a figure. To this end, we will derive
the density of the LPA distribution in the Merton framework. Recall from Equation (3.4.10) that the LPA distribution function in the Merton setting was given by

\[ F(x) = P\left( \frac{N^{(m)}}{m} \leq x \right) \approx \Phi \left( \frac{1}{\sqrt{\rho}} \left( \sqrt{1 - \rho} \Phi^{-1}(x) - \Phi^{-1}(\pi) \right) \right). \]

We thus need to find the derivative of \( F(x) \) with respect to \( x \), i.e. \( \frac{dF}{dx} \). We let \( \phi \) denote the density of a standard normal random variable, hence \( \phi(x) = \Phi'(x) \). We then have

\[
\frac{dF}{dx} = \phi \left( \frac{1}{\sqrt{\rho}} \left( \sqrt{1 - \rho} \Phi^{-1}(x) - \Phi^{-1}(\pi) \right) \right) \cdot \frac{\sqrt{1 - \rho}}{\sqrt{\rho}} \frac{1}{\phi(\Phi^{-1}(x))}
\]

\[
= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{\sqrt{\rho}^2} \left( \sqrt{1 - \rho} \Phi^{-1}(x) - \Phi^{-1}(\pi) \right)^2 \right\} \cdot \sqrt{\frac{1 - \rho}{\rho}} \exp \left\{ -\frac{(\Phi^{-1}(x))^2}{2} \right\}
\]

\[
= \sqrt{\frac{1 - \rho}{\rho}} \exp \left\{ \frac{1}{2} (\Phi^{-1}(x))^2 - \frac{1}{2\rho} \left( \sqrt{1 - \rho} \Phi^{-1}(x) - \Phi^{-1}(\pi) \right)^2 \right\}. \tag{3.4.11}
\]

In the first equality we used the fact that the inverse of a function \( f \) is given by \((f^{-1}(x))' = \frac{1}{f'(f^{-1}(x))}\).

Using the density and distribution of the LPA, Figure 3.4 shows the influence of \( \rho \) on the default fraction in the mixture model inspired by the Merton framework.
Figure 3.4: Illustration of the probability density (left figure) and the probability distribution (right figure) of the default fraction using the large portfolio approximation for four different values of $\rho$. The individual default probability is 5%.

It can be seen that higher values of $\rho$ (and thus higher default correlation) gives rise to thicker tails, i.e. larger probability for extreme events. Further, it can also be seen that lower correlation implies that the default distribution is more centered around its expected value. Note that all distributions, regardless of the value of $\rho$, have the same expected value of 0.05, which is the individual default probability. Let us now instead fix $\rho$ at 5% and let the individual default probabilities vary. The effects of this is shown in Figure 3.5. It can be seen that a higher default probability naturally leads to a higher expected number of defaults and also generates a distribution with thicker tails.
Figure 3.5: Illustration of the probability density (left figure) and the probability distribution (right figure) of the default fraction using the large portfolio approximation for four different default probabilities.

3.5 Threshold Models Using Copulas

The importance of finding a proper way to model dependence between obligors is central throughout this paper. The choice of credit risk model entails a choice of dependence structure, which is crucial for the model as a whole. In mixture models for instance, the practitioner needs to choose an appropriate mixing variable to describe dependence. The same issue is also apparent in threshold models, where there are many ways to model dependence. One way of doing this is to rely on the linear correlation in the Merton type model presented in subsection 3.4.2, where dependence was determined by a common random factor \( \Psi \). Another way to model dependence is to introduce the concept of copulas, which will be further explored in this section. Hult and Lindskog (2007) present two main reasons for introducing copulas:
they are useful for constructing multivariate models with non-Gaussian dependence structure
2) they help us understand dependence beyond linear correlation.

In connection to the second reason, McNeil et al. (2005) point out that linear correlation has several crucial weaknesses, especially in models that are more complex than those that use the regular Gaussian multivariate distribution. Thus, there is a strong need for modelling dependence in a more general way, and this is where copulas become useful. This section will begin with a general discussion of the theory behind the concept of copulas, and then proceed to connect copulas to models of credit risk.

3.5.1 Copulas

The following fundamental theory and notations behind copulas have to a large extent been obtained from McNeil et al. (2005) and Hult and Lindskog (2007).

**Definition 3.5.1 (Copula)**

A d-dimensional copula is a distribution function on $[0,1]^d$ with standard uniform marginal distributions.

Thus, a copula is a distribution function of the form $C : [0,1]^d \rightarrow [0,1]$. In other words, given a vector of random variables $(U_1, \ldots, U_d)$, a copula is the distribution function $P(U_1 \leq u_1, \ldots, U_d \leq u_d)$, where $P(U_i \leq u_i) = u_i$ and $u_i \in [0,1]$ for all $i$. From Definition 3.5.1 we can extract three necessary properties that a copula must satisfy. If a function fulfills these properties then it is a copula.

1. $C(u_1, \ldots, u_d)$ is increasing in each component $u_i$.
2. $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ for all $i \in \{1, \ldots, d\}$, where $u_i \in [0,1]$.
3. For all $(a_1, \ldots, a_d), (b_1, \ldots, b_d) \in [0,1]^d$ with $a_i \leq b_i$ we have

$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\cdots+i_d} C(u_{i_1}, \ldots, u_{i_d}) \geq 0,$$

where $u_{j1} = a_j$ and $u_{j2} = b_j$ for all $j \in \{1, \ldots, d\}$.

The first property follows directly from the fact that a copula is a distribution function.
Further, from the definition of marginal uniform distributions we get the second property. The last property is more technical, but it boils down to the rectangle inequality which ensures that $P(a_1 \leq U_1 \leq b_1, \ldots, a_d \leq U_1 \leq b_d)$ is non-negative.

The core usefulness of copulas is summarized in the following theorem, known as Sklar’s theorem. It essentially connects the joint distribution function to its univariate marginal distribution functions.

**Theorem 3.5.2 (Sklar’s Theorem)**

Let $F$ be a joint distribution function with marginal distribution functions $F_1, \ldots, F_d$. Then there exists a copula $C : [0,1]^d \to [0,1]$ such that, for all $x_1, \ldots, x_d$ in $\mathbb{R} = [-\infty, \infty]$,

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)) \quad (3.5.1)$$

If $F_1, \ldots, F_d$ are continuous, then $C$ is unique. Conversely, if $C$ is a copula and $F_1, \ldots, F_d$ are distribution functions, then $F$ defined by Equation (3.5.1) is a joint distribution function with marginal distribution functions $F_1, \ldots, F_d$.

One of the most fundamental challenges of modelling large portfolios of credit risk is finding an expression for the joint distribution function of the losses. Marginal distributions on the other hand, are often easier to obtain. To this end, Sklar’s theorem is extremely useful as it links the joint distribution function to its univariate marginal distribution functions. Further, observe that if $F$ is continuous then $F_i(F_i^{-1}(u)) = u$. By evaluating Equation (3.5.1) at $x_i = F_i^{-1}(x_i)$ for all $i \in \{1, \ldots, d\}$, we get the following formula

$$C(x_1, \ldots, x_d) = F(F_1^{-1}(x_1), \ldots, F_d^{-1}(x_d)) \quad (3.5.2)$$

Note that Equation 3.5.2 is fundamental since it shows how a copula $C$ can be extracted from a multivariate distribution function $F$ and its marginal distributions $F_1, \ldots, F_d$. Given this, it is natural to define the concept of the copula of a multivariate distribution as follows (see also Definition 5.4 in McNeil et al. (2005)):

**Definition 3.5.3**

If a random vector $X = (X_1, \ldots, X_d)$ has a multivariate distribution $F$ with continuous marginals $F_1, \ldots, F_d$, then the copula of $F$ (or $X$) is the distribution $C$ of $(F_1(X_1), \ldots, F_d(X_d)))$. 

38
Furthermore, the following proposition states a useful property of copulas.

**Proposition 3.5.4**

Suppose \((X_1, \ldots, X_d)\) has continuous marginal distribution functions and copula \(C\) and let \(T_1, \ldots, T_d\) be strictly increasing functions. Then also the random vector 
\[ (T_1(X_1), \ldots, T_d(X_d)) \]
has copula \(C\).

Proposition 3.5.4 states that the copula is invariant under strictly increasing transformations, i.e. it is possible to apply a strictly increasing function to the marginal distribution functions and still preserve the same copula. The results of Sklar’s theorem and Proposition 3.5.4 are central in the discussion of copulas. The proofs of the theorem and proposition respectively can be found in Appendix A.2.

Copulas can be divided into two major classes; *Elliptical* and *Archimedean*. The former contains copulas on elliptical distributions, which is where Sklar’s Theorem is applied. It is assumed that the reader is familiar with elliptical distributions, otherwise we refer to McNeil et al. (2005) for a detailed discussion on the topic. Archimedean copulas differ from elliptical copulas in many ways, and both classes have advantages and disadvantages over the other. The next part of this section will discuss these types of copulas in further detail.

**Elliptical Copulas.** One common elliptical copula is the so called *Gaussian copula*, which according to McNeil et al. (2005) is used in most threshold models in the industry. If we let the vector \(Z\) have a multivariate standard normal distribution, \(Z \sim N_d(0, R)\), where \(R\) is a correlation matrix, then the copula of \(Z\) is the Gaussian copula. By using Sklar’s Theorem together with Equation (3.5.2) and Definition 3.5.3 we can derive the Gaussian copula to be
\[
C^\text{Ga}_R(u) = P(\Phi(X_1) \leq u_1, \ldots, \Phi(X_d) \leq u_d) = \Phi_R(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d))
\]
where \(\Phi_R\) denotes the multivariate standard normal distribution with correlation matrix \(R\).

It turns out that the non-standard multivariate normal distribution also has the Gaussian copula. To see this, consider the random vector \(X\) with distribution \(X \sim N_d(\mu, \Sigma)\), where \(\mu\) is the mean vector and \(\Sigma\) is a covariance matrix. It is possible to transform
by applying a series of strictly increasing transformations on \( X \). Thus, by Proposition 3.5.4, \( X \) has the same copula as \( Z \), i.e. the Gaussian copula.

In a similar way we can let the vector \( X \) have a standard multivariate student-t distribution, and then use Sklar’s Theorem together with Equation (3.5.2) and Definition 3.5.3 to obtain the \( t \) copula as

\[
C^t_{\nu,R}(u) = t_{\nu,R}(t_{\nu}^{-1}(u_1), \ldots, t_{\nu}^{-1}(u_d))
\]

where \( \nu \) is the degrees of freedom. Roughly speaking all elliptical copulas share a common structure. In view of Equation (3.5.2) and Definition 3.5.3 these copulas can be expressed as

\[
C(u) = F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)), \text{ for } u \in [0,1]^d,
\]

where \( F \) is the joint distribution function and \( F_1, \ldots, F_d \) its corresponding marginal distribution functions. One of the advantages of elliptical copulas is that they are quite simple to simulate, as it amounts to simulating from an elliptical distribution such as the normal distribution. Further details about the simulation will be given later in this section. However, elliptical copulas also have some disadvantages. For instance, they are only given in implicit form, i.e. they do not possess closed form expressions. For example the \( d \)-dimensional Gaussian copula is a \( d \)-integral over the density of \( Z \). A further drawback is that by using elliptical copulas we are restricted to a certain dependence structure. As we have argued earlier there is a dependence between more extreme events in credit portfolios, and this sort of dependence cannot be achieved by using elliptical copulas. To this end we introduce the Archimedean copulas.

**Archimedean Copulas.** Archimedean copulas can be used to model a wider class of dependence, and they can also be expressed in closed form. However, the Archimedean copulas are not derived by using Sklar’s Theorem, and it turns out that some technical conditions need to hold in order for the Archimedean ”copulas” to actually be copulas in higher dimensions than two. A full description of Archimedean copulas are beyond the scope of this text, but for more details we refer to Nelsen (2006) or McNeil et al. (2005). Since the focus of this thesis is on credit risk modelling, we will restrict our discussion on Archimedean copulas to the details that are actually needed in our treatment of credit risk models. To this end, we need to introduce more mathematical concepts and notations. We hope that the examples will clarify the motivation behind introducing copulas in general, and Archimedean copulas in particular, to help in the
credit risk context. We begin our analysis in two dimensions, and start by stating a proposition needed to make a formal definition of Archimedean copulas:

**Proposition 3.5.5**
Let \( \phi : [0, 1] \to [0, \infty] \) be a continuous and strictly decreasing function s.t. \( \phi(0) = \infty \) and \( \phi(1) = 0 \). Let \( C : [0, 1]^2 \to [0, 1] \) be given by:
\[
C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2))
\]
(3.5.3)
where \( \phi^{-1} \) is the inverse of \( \phi \). Then \( C \) is a copula if and only if \( \phi \) is convex.

For a proof see Nelsen (2006). By using Proposition 3.5.5 we can formally define Archimedean copulas.

**Definition 3.5.6 (Archimedean Copula and Generator)**
A copula given on the form (3.5.3) is called a bivariate Archimedean copula. The function \( \phi \) defined as in Proposition 3.5.5 is called an Archimedean copula generator.

As can be seen, the Archimedean copulas are given by their generator. One example of a bivariate Archimedean copula is the two-dimensional Clayton copula, which is explained in the following example.

**Example 3.5.7 (Two-dimensional Clayton Copula)**
Let the copula generator \( \phi \) be given by \( \phi(t) = t^{-\theta} - 1 \), and \( \theta > 0 \), which implies that \( \phi^{-1}(t) = (t + 1)^{-\frac{1}{\theta}} \). A generator defined this way yields the Clayton copula, which in two dimensions is given by:
\[
C_{\theta}^{Cl}(u_1, u_2) = \phi^{-1}(u_1^{-\theta} - 1 + u_2^{-\theta} - 1) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}.
\]
In this version of the Clayon copula \( \theta \) is the only parameter to be estimated. There also exists a similar Clayton copula with two parameters, but this version will not be discussed, see e.g. McNeil et al. (2005) for more details.

We will now discuss how the theory for two-dimensional Archimedean copulas can be extended to hold for higher dimensions. We are interested in under which conditions
\[
\phi^{-1}(\phi(u_1) + \cdots + \phi(u_d))
\]
(3.5.4)
is a copula, for \( d \geq 3 \). In order for (3.5.4) to be a copula it is not sufficient that the generator \( \phi \) satisfies the conditions in Proposition 3.5.5. In addition, the inverse of the
generator, $\varphi^{-1}$, needs to be completely monotonic (see Appendix A for a definition).
Thus, it is clear that the generator $\varphi$ needs to fulfill certain conditions in order for (3.5.4) to be a copula. It turns out that a natural candidate for $\varphi$ is the inverse of the Laplace transform to any non-negative random variable. A generator defined this way will satisfy all needed conditions required for (3.5.4) to be a copula. Consequently, we proceed by defining the Laplace transform. This definition follows closely the lecture notes of Herbertsson (2010). By slight abuse of language we sometimes refer to the Laplace transform of the distribution function of a random variable as the Laplace transform of the random variable itself.

**Definition 3.5.8 (Laplace Transform)**

Let $X$ be a non-negative random variable with distribution function $F$. Then, the Laplace transform of $X$ is defined as:

$$
\mathcal{L}_X(s) = \mathbb{E}[e^{-sX}] = \int_0^\infty e^{-sx}dF(x), \quad s \geq 0 \tag{3.5.5}
$$

where the integral is the Lebesgue-Stieltjes integral of the distribution function $F$. Further, if the density $f$ exists we also have that:

$$
\mathcal{L}_X(s) = \int_0^\infty e^{-sx}f(x)dx, \quad s \geq 0.
$$

It is possible to verify that the inverse of the Laplace transform $\mathcal{L}_X$ indeed satisfies the needed conditions to be an Archimedean generator. If we define $\mathcal{L}_X(\infty) = 0$ then $\mathcal{L}_X$ is a continuous, strictly decreasing and completely monotonic function mapping $[0, \infty] \to [0, 1]$. Thus, if we let $\varphi(s) = \mathcal{L}_X^{-1}(s)$, and consequently $\varphi^{-1}(s) = \mathcal{L}_X(s)$ then expression (3.5.4) is a copula in $d$ dimensions. Further, we proceed by stating an example where the Clayton copula is generalized to hold in higher dimensions, and which gives additional motivation for using Laplace transforms. Then we will state a very useful theorem which shows how it is possible to construct a random vector with the distribution given by a multivariate Archimedean copula.

**Example 3.5.9 (Multivariate Clayton Copula)**

As in Example 3.5.7, let the copula generator $\varphi$ be given by $\varphi(t) = t^{-\theta} - 1$, for $\theta > 0$, implying that $\varphi^{-1}(t) = (t + 1)^{-\frac{1}{\theta}}$. This generates the Clayton copula, which in $d$ dimensions take the form:

$$
C^{\text{Cl}}_{\theta}(u_1, \ldots, u_d) = \left(u_1^{-\theta} + \cdots + u_d^{-\theta} - d + 1\right)^{-\frac{1}{\theta}}. \tag{3.5.6}
$$
We now extend this example by showing the link to Laplace transforms. Let $X$ be a gamma distributed random variable with parameters $\frac{1}{\theta}$ and 1, i.e. $X \sim \text{Gam}(\frac{1}{\theta}, 1)$. This implies that $X$ has the density function $f_X(x) = \frac{x^{\frac{1}{\theta} - 1}e^{-x}}{\Gamma(\frac{1}{\theta})}$, where $\Gamma(z)$ is the gamma function given by $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$. By taking the Laplace transform of $X$ we get:

\[
L_X(t) = \mathbb{E}[e^{-tX}] = \int_0^\infty e^{-tx}f(x)dx = \int_0^\infty e^{-tx}\frac{x^{\frac{1}{\theta} - 1}e^{-x}}{\Gamma(\frac{1}{\theta})}dx
\]
\[
= \frac{1}{\Gamma(\frac{1}{\theta})} \int_0^\infty \left(\frac{y}{1+t}\right)^{\frac{1}{\theta} - 1}e^{-y} \frac{1}{1+t}dy = \frac{1}{\Gamma(\frac{1}{\theta})(1+t)^{\frac{1}{\theta} - 1}(1+t)} \int_0^\infty y^{\frac{1}{\theta} - 1}e^{-y}dy
\]
\[
= \frac{\Gamma(\frac{1}{\theta})}{\Gamma(\frac{1}{\theta})(1+t)^{\frac{1}{\theta}}} = (1+t)^{-\frac{1}{\theta}} = \varphi^{-1}(t).
\]

Thus, if we let a random variable be gamma distributed with the parameters above, the Laplace transformation of that random variable will yield the Clayton copula generator and the inverse of the generator. In a more general form it can be said that the Archimedean copulas are determined by an underlying random variable with a given distribution. From this random variable the copula generator and its inverse can be obtained by means of taking the Laplace transform of the random variable.

We are now ready to state a theorem which even further motivates the role of Laplace transforms in Archimedean copulas. We refer to this theorem as the Marshall Olkin theorem, as it was them who originally proved it in Marshall and Olkin (1988). Similar results are also established in McNeil et al. (2005, p. 223). The Marshall Olkin theorem presents a way to construct random vectors that have the multivariate Archimedean copulas as their distribution. Further, it also gives a hint on how to simulate from Archimedean copulas and how to connect the theory of Archimedean copulas to modelling credit risk portfolios. The outline of the theorem and proof follows from Hult and Lindskog (2007) and Herbertsson (2010) in his lecture notes on Archimedean copulas.

**Theorem 3.5.10 (Marshall Olkin)**

Let $\Psi$ be a non-negative random variable with Laplace transform $L_{\Psi}$ given by Equation

\[L_{\Psi}(t) = \mathbb{E}[e^{-t\Psi}] = \int_0^\infty e^{-tx}f(x)dx = \int_0^\infty e^{-tx}\frac{x^{\frac{1}{\theta} - 1}e^{-x}}{\Gamma(\frac{1}{\theta})}dx = \frac{1}{\Gamma(\frac{1}{\theta})} \int_0^\infty \left(\frac{y}{1+t}\right)^{\frac{1}{\theta} - 1}e^{-y} \frac{1}{1+t}dy = \frac{1}{\Gamma(\frac{1}{\theta})(1+t)^{\frac{1}{\theta} - 1}(1+t)} \int_0^\infty y^{\frac{1}{\theta} - 1}e^{-y}dy = \frac{\Gamma(\frac{1}{\theta})}{\Gamma(\frac{1}{\theta})(1+t)^{\frac{1}{\theta}}} = (1+t)^{-\frac{1}{\theta}} = \varphi^{-1}(t).
\]
Further, let \( \varphi(t) = \mathcal{L}_\Psi^{-1}(t) \), and \( \varphi^{-1}(t) = \mathcal{L}_\Psi(t) \). Then, there exists a sequence \( V_1, \ldots, V_d \) of identically distributed random variables defined on \([0, 1]\) that are conditionally independent given \( \Psi \) with conditional distribution given by:

\[
P (V_i \leq v_i \mid \Psi) = \exp(-\Psi \varphi(v_i)).
\] (3.5.7)

Additionally, the distribution function of the random vector \((V_1, \ldots, V_d)\) is the Archimedean copula with generator \( \varphi(t) \), i.e.:

\[
P (V_1 \leq v_1, \ldots, V_d \leq v_d) = \varphi^{-1}(\varphi(v_1) + \cdots + \varphi(v_d)).
\] (3.5.8)

**Proof** Let \( U_1, \ldots, U_d \) be a sequence of i.i.d. standard uniform random variables, which are independent of \( \Psi \). Define \( V_i \) as:

\[
V_i = \varphi^{-1}\left(-\frac{\ln U_i}{\Psi}\right) \quad \text{for all} \quad i \in \{1, \ldots, d\}.
\]

This implies that the distribution of \( V_i \) conditional on \( \Psi \) is given by:

\[
P (V_i \leq v_i \mid \Psi) = P \left( \varphi^{-1}\left(-\frac{\ln U_i}{\Psi}\right) \leq v_i \mid \Psi \right) = P \left( -\frac{\ln U_i}{\Psi} \geq \varphi(v_i) \mid \Psi \right)
\]

\[
= P (\ln U_i \leq -\Psi \varphi(v_i) \mid \Psi) = P (U_i \leq \exp(-\Psi \varphi(v_i)) \mid \Psi)
\]

\[
= \exp(-\Psi \varphi(v_i)).
\]

The second equality uses the fact that \( \varphi \) is strictly decreasing, i.e. if \( x_1 \leq x_2 \) then \( \varphi(x_1) \geq \varphi(x_2) \). This proves (3.5.7).

From the first part of the theorem we have that \( V_1, \ldots, V_d \) are conditionally independent given \( \Psi \) and that the conditional distribution is given by (3.5.7). This implies:

\[
P (V_1 \leq v_1, \ldots, V_d \leq v_d) = \mathbb{E} \left[ P (V_1 \leq v_1, \ldots, V_d \leq v_d \mid \Psi) \right] = \mathbb{E} \left[ \prod_{i=1}^{d} P (V_i \leq v_i \mid \Psi) \right]
\]

\[
= \mathbb{E} \left[ \prod_{i=1}^{d} \exp(-\Psi \varphi(v_i)) \right] = \mathbb{E} \left[ \exp\left( -\sum_{i=1}^{d} \varphi(v_i) \right) \Psi \right]
\]

\[
= \mathcal{L}_\Psi \left( \sum_{i=1}^{d} \varphi(v_i) \right) = \varphi^{-1}\left( \sum_{i=1}^{d} \varphi(v_i) \right).
\]

The last equality follows from that \( \varphi^{-1} = \mathcal{L}_\Psi \) by definition.
We will come back to the Marshall Olkin theorem in subsection 3.5.3 and use the results to construct a mixture model based on an Archimedean copula.

More specifically, in the next two sections we will implement copulas in the analysis of credit risk portfolios. First we proceed by working in a threshold model, and show how loss and default distributions can be estimated by means of simulating from copulas. So far we have established results that reveal how we can simulate from different copulas. These results will be applied when conducting simulations from both elliptical and Archimedean copulas. Furthermore, as argued earlier, the threshold models can be incorporated into the more general class of mixture models. To emphasize this, in Section 3.5.3 we will construct a mixture model by using Archimedean copulas.

3.5.2 Simulations from Copulas in Static Credit Risk Portfolios

We now turn the focus to the connection of copulas in credit risk portfolios. Assume that we are working in the threshold model framework, i.e. an obligor defaults if a random variable, or a stochastic process, is below some threshold at a given time. We let $V_{t,i}$ denote the asset value process for obligor $i$, and we consider a portfolio with $m$ obligors. Further we assume that we have an exchangeable and homogeneous portfolio. As before, an obligor defaults if the asset value at time $T$ is below the deterministic threshold $D$. Note here that in contrast to Section 3.4.2 we do not make any assumption of the dynamics of the process $V_{t,i}$, instead we will make use of copulas. We let $F$ be the joint distribution function and $F_i$ the marginal distribution functions of the asset values at time $T$. We assume that the $F_i$'s are continuous. Thus, since we have an exchangeable portfolio, the individual default probability is

$$P(Y_i = 1) = P(V_{T,i} < D) = F_i(D) = \pi \text{ for all } i.$$ 

As in the section on mixture models, we are interested in calculating the probabilities for the number of defaults, i.e. $P(N^{(m)} = k)$. However, following the discussion from the previous section, we run into problems by trying to calculate this probability explicitly. Fortunately, computing this probability by means of simulations is quite straightforward. First, we define $E_k$ as the event that the first $k$ obligors default. The
probability that $E_k$ occurs is given by:

$$P(E_k) = P(Y_1 = 1, \ldots, Y_k = 1)$$

$$= P(V_{T,1} < D, \ldots, V_{T,k} < D)$$

$$= P(V_{T,1} < D, \ldots, V_{T,k} < D, V_{T,k+1} < \infty, \ldots, V_{T,m} < \infty)$$

$$= C(F_1(D), \ldots, F_k(D), 1, \ldots, 1)$$

$$= C(\pi, \ldots, \pi, 1, \ldots, 1)$$

where $C$ is the copula of the asset values at time $T$. The copula is introduced in the fourth equality by using Sklar’s Theorem. Further, since we have an exchangeable portfolio it holds that:

$$C(u_1, \ldots, u_m) = C(u_{\Pi(1)}, \ldots, u_{\Pi(m)})$$

for an arbitrary permutation $\{\Pi(1), \ldots, \Pi(m)\}$ of $\{1, \ldots, m\}$.

The algorithm for simulating outcomes from a copula differs depending on the choice of copula, as is shown in Appendix A. For instance, the algorithms for simulating from elliptical and Archimedean copulas differ substantially. In elliptical copulas the distribution of the random factors is given from the choice of copula, and simulating these random factors is fairly straightforward. Archimedean copulas on the other hand are determined by its generator, and the distribution to simulate from is not given explicitly (however, we established that the Clayton copula can be obtained from a gamma distributed random variable). It is also possible to use other algorithms, see e.g. McNeil et al. (2005, Ch. 5) or Hult and Lindskog (2007). Simulation algorithms from some of the most common copulas are presented in Appendix A.

We are now interested in how the choice of copula will affect the default distribution. More specifically, we will focus on the difference in dependence structure, which can be investigated by observing the tail of the distribution. For instance, we claimed earlier that by using elliptical copulas we are restricted to a certain dependence structure. To this end, we will use three different types of copulas, both elliptical and Archimedean, to explore the implications of the choice of copula. The three types of copulas will be Gaussian, t and Clayton.
Consider an exchangeable portfolio of 1000 obligors with individual default probability $\pi = 0.049$ and correlation matrix $R$, given by $R(i, j) = 0.0157$ for $i \neq j$ and $R(i, i) = 1$. We will first estimate the default distribution by using the simulation approach from the Gaussian and $t$ copula. Figure 3.6 displays the simulated approximation to $P(N^{(m)} = k)$ when using the above mentioned Gaussian copula and a $t$ copula with 4 degrees of freedom.

It is clear that the choice of copula has a major impact on the default distribution. Even though the input parameters are the same there is a significant difference between the simulated default distributions. Specifically, we observe that the $t$ copula allows for more extreme events.

Further, we will now conduct simulations of the same portfolio by using the Clayton copula. The Clayton parameter that corresponds to this portfolio setting is $\theta = 0.032$. 

Figure 3.6: The simulated default distribution of a portfolio with 1000 obligors, individual default probability $\pi = 0.049$ and correlation matrix $R$. The number of defaults were simulated 10000 times.
Figure 3.7 illustrates the simulated default distribution.

Figure 3.7: The simulated default distribution from a Clayton copula with \( \theta = 0.032 \). The portfolio consists of 1000 obligors with individual default probability of \( \pi = 0.049 \). The number of defaults were simulated 10000 times.

We conclude that the Gaussian copula implies a significantly more centered distribution around its mean, whereas the \( t \) and Clayton copula is more spread with a thicker upper tail. Hence, from both the \( t \) copula and the Clayton copula it can be seen that there is a stronger dependence between defaults. Further, in the Clayton copula the distribution of defaults are from a model perspective given exclusively by the choice of the parameter \( \theta \). That the Clayton copula only has one parameter is a drawback, however as mentioned in Example 3.5.7 there exist generalizations in order to incorporate more parameters. The influence on the default distribution in the Clayton copula by the choice of the parameter \( \theta \) is illustrated in Figure 3.8.
Figure 3.8: The simulated density as a function of $\theta$ and the number of defaults using the Clayton copula. The credit portfolio consists of 100 obligors with individual default probability $\pi = 0.15$. The number of defaults were simulated $10^5$ times.

It can be seen from Figure 3.8 that for low values of $\theta$ the distribution of defaults is more centered around the expected number of defaults, i.e. 15. As $\theta$ increases both tails of the distribution become thicker and the distribution is more spread. Note also that the peak of the probability mass shifts more to the lower tail. When $\theta$ increases even more, this behavior would continue, giving greater probability to events close to zero defaults. Events outside the lower tail would be given almost equal probability. For instance, the probability of having 20 defaults would be almost the same as the probability of having 90 defaults.

Furthermore, a realistic value of $\theta$ in a static credit risk portfolio could range between 0.02 and 0.07. For example, a value of $\theta = 0.032$ was, according to McNeil et al. (2005), obtained by fitting the Clayton copula to a portfolio where the parameters roughly corresponded to an S&P rating B. For more details, see Section 4.2.1.
3.5.3 Bernoulli Mixture Models based on the Archimedean Copula

We are now interested in trying to incorporate the threshold model into a mixture model by using copulas. More specifically, by starting with a threshold model we want to construct a Bernoulli mixture model based on an Archimedean copula. This section follows the ideas of Herbertsson (2010), and also similar outlines can be found in McNeil et al. (2005, p.361). The initial setting is identical to the framework in Section 3.5.2. That is, we assume that we have an homogeneous and exchangeable credit portfolio with $m$ obligors. Further, $V_{T,1}, \ldots, V_{T,m}$ is a sequence of random variables representing obligor $i$'s asset value at time $T$, and an arbitrary obligor defaults if $V_{T,i} < D$. Thus, we let the default indicator $Y_i$ be defined as:

$$Y_i = \begin{cases} 1 & \text{if } V_{T,i} < D \\ 0 & \text{otherwise.} \end{cases}$$

The idea is to use Theorem 3.5.10. in order to construct this mixture model. We give a short recap of the setting:

Let $\Psi$ be a non-negative random variable with Laplace transform $L_\Psi$. Set $\varphi(t) = L^{-1}_\Psi(t)$, so that $\varphi^{-1}(t) = L_\Psi(t)$. Further, let $V_{T,1}, \ldots, V_{T,m}$ be conditionally independent given $\Psi$, with conditional distribution given by:

$$P(V_i \leq v_i \mid \Psi) = \exp(-\Psi \varphi(v_i)).$$

From Theorem 3.5.10 we then know that the vector $(V_{T,1}, \ldots, V_{T,m})$ has a multivariate distribution given by the Archimedean copula with generator $\varphi$. Thus, we have that the conditional default probability is given by:

$$p(\Psi) = P(Y_i = 1 \mid \Psi) = P(V_{T,i} < D \mid \Psi) = \exp(-\Psi \varphi(D))$$

where as before we interpret $\Psi$ as a variable that models the macroeconomic climate. Further, this implies that the unconditional default probability can be calculated as:

$$\pi = P(Y_i = 1) = \mathbb{E}[P(Y_i = 1 \mid \Psi)] = \mathbb{E}[\exp(-\Psi \varphi(D))] = L_\varphi(\varphi(D))$$

$$= \varphi^{-1}(\varphi(D)) = D.$$ 

This implies that the threshold level $D$ in this mixture model is given by the individual default probability. To sum up, we have obtained that the conditional default
probability is given by:

\[ p(\Psi) = \exp(-\Psi \varphi(\pi)) \quad (3.5.9) \]

where \( \pi \) is the default probability and \( \varphi \) is the Archimedean copula generator defined as the inverse of the Laplace transform of \( \Psi \), i.e. \( \varphi(t) = \mathcal{L}_\Psi^{-1}(t) \). In this homogeneous model we are now interested in analyzing the default distribution, and specifically find a large portfolio approximation (LPA) of the normalized loss distribution. From subsection 3.3.2 we know that the fraction of the number of defaults in a portfolio converges almost surely to the mixing variable. Further, this implies that the distribution of the fraction of the number of defaults converges to the distribution of the mixing variable. Hence, we have that:

\[ \lim_{m \to \infty} P\left( \frac{N^{(m)}}{m} \leq x \right) \to G(x), \text{ where } G(x) = P\left( p(\Psi) \leq x \right). \]

Thus, we proceed with finding an expression for the distribution function \( G(x) \).

\[
G(x) = P\left( p(\Psi) \leq x \right) = P\left( \exp(-\Psi \varphi(\pi)) \leq x \right) = P\left( -\Psi \leq \frac{\ln x}{\varphi(\pi)} \right) \\
= P\left( \Psi \geq -\frac{\ln x}{\varphi(\pi)} \right) = 1 - P\left( \Psi \leq -\frac{\ln x}{\varphi(\pi)} \right) = 1 - F\left( -\frac{\ln x}{\varphi(\pi)} \right),
\]

where \( x \in [0, 1] \). To conclude, we have that as the number of obligors in a portfolio grows large, the default fraction in this mixture model based on the Archimedean copula can be approximated by:

\[ P\left( \frac{N^{(m)}}{m} \leq x \right) \approx 1 - F\left( -\frac{\ln x}{\varphi(\pi)} \right), \quad x \in [0, 1] \quad (3.5.10) \]

where \( F \) is the distribution function of the common factor \( \Psi \), and \( \varphi \) is the Archimedean copula generator (defined as \( \varphi = \mathcal{L}_\Psi^{-1} \)).

We end this section with calculating the default distribution in a mixture model based on the Clayton copula. Recall that this implies that \( \Psi \sim \text{Gam}(\frac{1}{\theta}, 1) \) and that the generator is given by \( \varphi(t) = t^{-\theta} - 1 \), for \( \theta \geq 0 \) (See Example 3.5.9). With the Clayton copula the conditional default probability \( p(\Psi) \) in Equation (3.5.9) becomes \( p(\Psi) = \exp \left( -\Psi (\pi^{-\theta} - 1) \right) \). Further, if we let \( F \) be the distribution function of \( \Psi \), then
with the Clayton copula Equation (3.5.10) takes the form:

\[ P\left( \frac{N^{(m)}}{m} \leq x \right) \approx 1 - F\left( -\frac{\ln x}{\pi^{\theta} - 1} \right) = 1 - \frac{\gamma(\frac{1}{\theta}, -\frac{\ln x}{\pi^{\theta} - 1})}{\Gamma(\frac{1}{\theta})}, \]

where

\[ \gamma(s, x) = \int_{0}^{x} t^{s-1} e^{-t} dt \]

\[ \Gamma(s) = \int_{0}^{\infty} t^{s-1} e^{-t} dt. \]

The LPA distribution for the fraction of the number of defaults for different values of \( \theta \) and for different individual default probabilities is illustrated in Figure 3.9.

**Figure 3.9:** The figure to the left shows the LPA distribution for the fraction of the number of defaults for different values of \( \theta \) where the individual default probability is fixed at \( \pi = 0.15 \). The figure to the right shows the LPA distribution for the fraction of the number of defaults for different individual default probabilities, where \( \theta \) is fixed at \( \theta = 0.032 \).

In the left plot of Figure 3.9 it can be seen that larger values of \( \theta \) implies that the distribution function grows more slowly. Thus, there is a greater probability for higher default fractions and we can conclude that larger values of \( \theta \) produce thicker tails.
Up to this point we have studied a few different approaches to modelling the defaults and dependence in primarily homogeneous and exchangeable static credit portfolios. First the Bernoulli mixture model was analyzed, and it was concluded that it in turn is determined by the choice of mixing distribution. Secondly we discussed a threshold model using the mixture approach, where inspiration also came from the Merton model. Lastly we constructed a Bernoulli mixture model from a threshold model using Archimedean copulas. A relevant question at this point is how different the static credit risk models discussed so far are. In other words, if we have a credit portfolio with a given number of obligors, individual default probability and default correlation, how well do the models coincide. This question is one of the concerns of Chapter 4. It will be concluded that these different approaches yield similar results for non-extreme scenarios.
Chapter 4

Risk Measuring in Static Credit Risk Models

In this chapter we use the static credit risk models from chapter 3 to measure and quantify the risk involved in a credit portfolio. The risk of the credit portfolios will primarily be measured by using the two risk measures Value-at-Risk and Expected Shortfall. We will also be concerned with analyzing the similarities of the different static credit risk models discussed so far.

In Section 4.1 the risk measures Value-at-Risk and Expected Shortfall will be defined, and some further properties of these will be discussed. In Section 4.2 the Bernoulli mixture model will be explored more in depth. It was concluded that the loss and default distribution in this model is completely determined by the mixing variable. For this reason, frequently used mixing variables, also involving the mixing variable based on the Clayton copula, will be discussed. Furthermore, we will analyze how the choice of mixing distribution affects the risk in the portfolio and how it affects the tails of the loss distribution. It will be concluded that the mixing distributions will yield similar results, and thus that the choice of mixing distribution is of less concern.

In order to analyze the mixture model inspired by Merton in the same framework as the Bernoulli mixture model some additional information is needed. This is the concern of Section 4.2.2.
4.1 Value-at-Risk and Expected Shortfall

In Section 2.2 the measurement and management of risk was discussed. It was mentioned that two frequently used risk measures are Value-at-Risk (VaR) and Expected Shortfall (ES). Value-at-Risk is the most common way to measure risk in the industry, see e.g. McNeil et al. (2005, p. 37) who writes ”Value-at-Risk (VaR) is probably the most widely used risk measure in financial institutions”. In loose terms the Value-at-Risk gives a value which the loss will not exceed with a high probability. The following definition is similar to the one in Herbertsson (2009) and Hult and Lindskog (2007).

Definition 4.1.1 (Value-at-Risk)
Given a loss \( L \) and a confidence level \( \alpha \in (0, 1) \), the \( \text{VaR}_\alpha(L) \) is the smallest number \( y \) such that the probability that \( L \) exceeds \( y \) is less than or equal to \( 1 - \alpha \);

\[
\text{VaR}_\alpha(L) = \begin{cases} 
\inf \{ y \in \mathbb{R} : P(L \geq y) \leq 1 - \alpha \} \\
\inf \{ y \in \mathbb{R} : 1 - P(L \leq y) \leq 1 - \alpha \} \\
\inf \{ y \in \mathbb{R} : P(L \leq y) \geq \alpha \} \\
\inf \{ y \in \mathbb{R} : F_L(y) \geq \alpha \} 
\end{cases} 
\tag{4.1.1}
\]

Remark: Note that \( \text{VaR}_\alpha(L) \) is simply the \( \alpha \)-quantile of the loss distribution \( F_L(x) \), see e.g. McNeil et al. (2005, p. 39).

Remark: Note that if the distribution function of \( L \) is continuous and strictly increasing then \( \text{VaR}_\alpha(L) = F_L^{-1}(\alpha) \).

Remark: VaR usually comes with a specified time horizon. In credit risk management typically one year is used. So if the loss \( L \) is over one year we talk about the one-year \( \text{VaR}_\alpha(L) \) with confidence level \( \alpha \).

Furthermore, Proposition 4.1.2 states some properties of VaR.

Proposition 4.1.2 (Properties of Value-at-Risk)
Let \( L \), \( L_1 \) and \( L_2 \) be random variables denoting the loss and \( x \) a real number. Then the following properties of VaR holds:

a) Translation Invariance: \( \text{VaR}_\alpha(L + x) = \text{VaR}_\alpha(L) + x \)

b) Positive Homogeneity: \( \text{VaR}_\alpha(xL) = x \text{VaR}_\alpha(L) \)

c) Monotonicity: If \( L_1 \leq L_2 \) almost surely then \( \text{VaR}_\alpha(L_1) \leq \text{VaR}_\alpha(L_2) \)
Proof:

\[ a) \quad \text{VaR}_\alpha(L + x) = \inf \{ y \in \mathbb{R} : P(L + x \leq y) \geq \alpha \} \]
\[ = \inf \{ y \in \mathbb{R} : P(L \leq y - x) \geq \alpha \} = \{ \text{let } y - x = z \} \]
\[ = \inf \{ z + x \in \mathbb{R} : P(L \leq z) \geq \alpha \} = \inf \{ z \in \mathbb{R} : P(L \leq z) \geq \alpha \} + x \]
\[ = \text{VaR}_\alpha(L) + x. \]

\[ b) \quad \text{Analogous to proof of } a). \]

\[ c) \quad \text{If } L_1 \leq L_2 \text{ a.s. then } F_{L_1} \geq F_{L_2}. \text{ This implies:} \]
\[ \text{VaR}_\alpha(L_1) = \inf \{ y \in \mathbb{R} : F_{L_1}(y) \geq \alpha \} \leq \inf \{ y \in \mathbb{R} : F_{L_2}(y) \geq \alpha \} = \text{VaR}_\alpha(L_2). \]

What can be concluded from Proposition 4.1.2 is that Value-at-Risk possesses many of the desired properties that one would wish a risk measure to have. However, there are two drawbacks of VaR. The first is that it is not subadditive which implies that \( \text{VaR}_\alpha(L_1 + L_2) \) is not necessarily less than or equal to \( \text{VaR}_\alpha(L_1) + \text{VaR}_\alpha(L_2) \) for arbitrary \( L_1 \) and \( L_2 \). This is a problem since if two or more assets are combined then one would expect that the total risk would be no greater than the risk of the assets considered separately. Diversification should lower the risk, not the opposite. However, when using VaR as a risk measure the risk of a combined, well-diversified portfolio can be larger than the sum of the individual risks of the members in the portfolio. This fact is known as the non-coherence of VaR as a risk measure. We will not delve deeper into this but for an excellent discussion of coherent risk measures and the drawbacks of VaR we refer to Frey and McNeil (2002).

The second drawback of VaR is that it does not give any information about how severe the losses can be if the VaR is exceeded, which by definition occurs with probability \( 1 - \alpha \).

One risk measure that overcomes both of the above mentioned problems is the expected shortfall. Expected shortfall uses, just as VaR, the distribution of the loss to determine the level of risk. As will be seen, these risk measures are closely related. The definition of expected shortfall is:
Definition 4.1.3 (Expected Shortfall)
Let $L$ be a random variable representing the loss, and assume that $\mathbb{E}[|L|] < \infty$. Then, the expected shortfall at confidence level $\alpha$ is defined as:

$$ES_\alpha(L) = \mathbb{E}[L \mid L \geq \text{VaR}_\alpha(L)].$$

Thus the expected shortfall answers the question: "If the VaR is exceeded, how much do we expect to lose?" By definition, the expected shortfall clearly depends on the Value-at-Risk. However, in contrast to VaR, the expected shortfall is concerned with how the tail of the loss distribution behaves for values greater than the VaR, i.e. it looks further out in the tail. It can also be shown that ES has all the properties of VaR from Proposition 4.1.2, but in addition it possesses the subadditivity property, which implies that ES is a coherent risk measure.

We will now prove a useful identity of expected shortfall which holds if the loss distribution function, $F_L$, is continuous and strictly increasing. However, we first need an auxiliary property, which we state as a lemma.

Lemma 4.1.4
Let $L$ be a random variable with distribution function $F_L$, and further let $U$ be a random variable which have a standard uniform distribution. Then $L$ and $F_L^{-1}(U)$ have the same distribution, i.e.

$$L \overset{d}{=} F_L^{-1}(U).$$ (4.1.2)

Proof  Let $H$ be the distribution function for $F_L^{-1}(U)$. The following shows that $H(x) = F_L(x)$.

$$H(x) = P(F_L^{-1}(U) \leq x) = P(U \leq F_L(x)) = F_U(F_L(x)) = F_L(x).$$

The last equality follows from the fact that the probability that a standard uniform random variable is less than or equal to some number between 0 and 1 is simply that number in return. Note that $F_L(x) \in [0, 1]$ for all $x$ since $F_L$ is a distribution function.

Proposition 4.1.5
Let the notation be the same as in Lemma 4.1.4 and let the distribution of $L$ be contin-
uous and strictly increasing. Then

$$ES_\alpha(L) = \frac{1}{1 - \alpha} \int_\alpha^1 \text{VaR}_\alpha(L) du.$$  \hfill (4.1.3)

**Proof**

$$ES_\alpha(L) = \mathbb{E} [L \mid L \geq \text{VaR}_\alpha(L)] = \frac{\mathbb{E} \left[ L \cdot I_{\{L \geq \text{VaR}_\alpha(L)\}}(L) \right]}{P(L \geq \text{VaR}_\alpha(L))} = \frac{\mathbb{E} \left[ L \cdot I_{\{\text{VaR}_\alpha(L), \infty\}}(L) \right]}{1 - \alpha}. \hfill (4.1.4)$$

Note that in the last equality only the definition of VaR is used; the probability for the loss to exceed the VaR is by definition $1 - \alpha$. Furthermore, we now proceed with using that $L$ has the same distribution as $F_L^{-1}(U)$ and that $\text{VaR}_\alpha(L) = F_L^{-1}(\alpha)$ since $F_L$ is continuous. The numerator in (4.1.4) becomes

$$\mathbb{E} \left[ L \cdot I_{\{\text{VaR}_\alpha(L), \infty\}}(L) \right] = \mathbb{E} \left[ F_L^{-1}(U) \cdot I_{\{F_L^{-1}(\alpha), \infty\}}(F_L^{-1}(U)) \right] = \left\{ \begin{array}{l}
1_{\{F_L^{-1}(\alpha), \infty\}}(F_L^{-1}(U)) = 1 \\
\text{if } F_L^{-1}(\alpha) \leq F_L^{-1}(U) < \infty \\
\text{if } \alpha \leq U < 1
\end{array} \right\} = \mathbb{E} \left[ F_L^{-1}(U) \cdot \eta_{[\alpha,1]}(U) \right] = \int_\alpha^1 F_L^{-1}(u) f_U du = \int_\alpha^1 \text{VaR}_\alpha(L) du.$$

The fact that $\text{VaR}_\alpha(L) = F_L^{-1}(\alpha)$ if $F$ is continuous and strictly increasing was used in the second and the last equality. Note that the density, $f_U$, for a standard uniform random variable is equal to one. Adding this up Equation (4.1.3) has been shown. This identity is useful for calculating the expected shortfall when the distribution function of the loss is continuous and strictly increasing. Figure 4.1 shows an example of VaR and ES in a Bernoulli mixture model with a beta mixing variable. The point here is to illustrate VaR and ES with a given loss distribution. The exact details of the calculations will be explained later.
This section will be concluded with an approximation of VaR and ES when the number of obligors in the portfolio is "large".

**Approximation of VaR and ES in Large Portfolios.** Consider the Bernoulli mixture model from Chapter 3 (which we know also incorporates the threshold models). We are interested in finding an approximation of the VaR and ES in a large portfolio. Recall that we denote the loss distribution of a portfolio with $m$ obligors with $L^{(m)}$ and further we let $G$ be the distribution function of the mixing variable $p(\Psi)$. In the following proposition we will derive an approximation of VaR and ES which holds for homogeneous portfolios in the mixture framework.

**Proposition 4.1.6 (Approximation of VaR and ES in large portfolios)**

Consider a homogeneous portfolio with $m$ obligors, where the loss rate is given by $l$. If
$G$ is continuous and $m$ is large then the following approximations hold:

\[ \text{VaR}_\alpha(L^{(m)}) \approx l \cdot m \cdot G^{-1}(\alpha) \]  

(4.1.5)

\[ \text{ES}_\alpha(L^{(m)}) \approx \frac{l \cdot m}{1 - \alpha} \int_\alpha^1 G^{-1}(u)du. \]  

(4.1.6)

**Proof:**

\[ \text{VaR}_\alpha(L^{(m)}) = \inf \{ y \in \mathbb{R} : P(L^{(m)} \leq y) \geq \alpha \} = \inf \{ y \in \mathbb{R} : P\left(\frac{L^{(m)}}{lm} \leq \frac{y}{lm}\right) \geq \alpha \} \]

\[ = \inf \left\{ y \in \mathbb{R} : P\left(\frac{N^{(m)}}{m} \leq \frac{y}{lm}\right) \geq \alpha \right\} \xrightarrow{m \to \infty} \inf \left\{ y \in \mathbb{R} : G\left(\frac{y}{lm}\right) \geq \alpha \right\} \]

\[ = \left\{ \text{let } \frac{y}{lm} = x \right\} = \inf \{ l \cdot m \cdot x \in \mathbb{R} : G(x) \geq \alpha \} = l \cdot m \cdot G^{-1}(\alpha). \]

**Comment at *:** We proved in Section 3.3.2 that $P\left(\frac{N^{(m)}}{m} \leq \frac{y}{lm}\right) \to G\left(\frac{y}{lm}\right)$ as $m \to \infty$. That is the distribution of the average number of losses converges to the distribution of the the mixing variable as $m$ goes to infinity.

Further, the approximation of expected shortfall follows directly from the VaR-approximation by using the identity from Equation (4.1.3):

\[ \text{ES}_\alpha(L^{(m)}) = \frac{1}{1 - \alpha} \int_\alpha^1 \text{VaR}_\alpha(L^{(m)})du \approx \frac{1}{1 - \alpha} \int_\alpha^1 l \cdot m \cdot G^{-1}(u)du \]

\[ = \frac{l \cdot m}{1 - \alpha} \int_\alpha^1 G^{-1}(u)du. \]

\[ \blacksquare \]

### 4.2 Bernoulli Mixture Model

Mixture models are commonly used in the industry, as they make up the building blocks of, for instance, the model called CreditRisk+. There are many advantages with applying mixture variables to model portfolio credit loss. They are flexible in the sense that there is a wide variety of mixing variables to choose from, and there are several assumptions that can either be relaxed or kept. Thus, when fitting a mixture model to a certain credit portfolio the user will have many options to optimize this fit.
This section begins with a discussion of the essential properties of some of the most common mixing variables. The choice of mixing distribution yields different models, which give rise to *model risk*. As claimed by Frey and McNeil (2003), it will be shown that model risk in a Bernoulli mixture model is not as much of a problem as the issue with determining the individual default probabilities, $\pi$, and the default correlation.

### 4.2.1 Properties of Mixing Distributions

It is intuitive to think that the choice of mixing distribution is essential when concerned with model risk. However, this section will explore this intuition and actually conclude that for a wide range of mixing distributions the difference between the loss distributions are small for non-extreme quantiles. In order to investigate the model risk, the differences between the most common mixing variables will be explored. Therefore, consider an exchangeable one-factor Bernoulli mixture model with mixing variable $p(\Psi)$. One of the most critical properties of these variables is how their tails differ for more extreme events. To this end, let us analyze the probability $P(p(\Psi) > q)$ for the following distributions; Beta, Probit-normal, Logit-normal and Clayton copula.

**Beta mixing distribution.** Let $p(\Psi) = \Psi$ where the random variable $\Psi$ is Beta distributed with parameters $a$ and $b$, i.e. $\Psi \sim \text{Beta}(a, b)$. Then, its density function is given by

$$f(x) = \frac{1}{\beta(a, b)} x^{a-1} (1 - x)^{b-1}, \quad 0 < x < 1, \ a, b > 0,$$

where the beta function is defined as $\beta(a, b) = \int_0^1 x^{a-1} (1 - x)^{b-1} dx, \ a, b > 0$.

The beta distribution is often used within credit risk modelling because of its nice properties. For instance, the output of the beta random variable ranges in the interval $(0, 1)$, which simplifies computations when dealing with probabilities, i.e. there is no need for re-scaling the output of the random variable.

**Probit-normal mixing distribution.** This model is commonly used in the industry, as it is included in both CreditMetrics and KMV (Frey and McNeil, 2003). For this mixing variable it is assumed that $p(\Psi) = \Phi(\mu + \sigma \Psi)$, where $\Psi \sim N(0, 1), \mu \in \mathbb{R}$ and
σ > 0. To find the probability $P(p(Ψ) > q)$ some calculations need to be done, that is:

$$P(p(Ψ) > q) = P(Φ(\mu + σΨ) > q) = P(\mu + σΨ > Φ^{-1}(q)) = P(Ψ > \frac{Φ^{-1}(q) - \mu}{σ})$$

$$= 1 - P(Ψ ≤ \frac{Φ^{-1}(q) - \mu}{σ}) = 1 - Φ\left(\frac{Φ^{-1}(q) - \mu}{σ}\right) = Φ\left(\frac{μ - Φ^{-1}(q)}{σ}\right)$$

where the last equality follows from the fact that $1 - Φ(x) = Φ(-x)$.

**Logit-normal mixing distribution.** For this mixing variable we have that $p(Ψ) = F(\mu + σΨ)$, where $Ψ ∼ N(0, 1)$, $μ ∈ ℝ$, $σ > 0$ and

$$F(y) = \frac{1 + e^{-y}}{1 + e^y} - 1.$$

As in the case above, some simple calculations are needed to extract the probability $P(p(Ψ) > q)$. Thus we have for $q ∈ (0, 1]$ that:

$$P(p(Ψ) > q) = P((1 + e^{-μ+σΨ})^{-1} > q) = P\left(-(μ + σΨ) < \ln\left(\frac{1}{q} - 1\right)\right)$$

$$= P\left(Ψ > -\left(\frac{\ln\left(\frac{1}{q} - 1\right) + \mu}{σ}\right)\right) = P\left(Ψ < \frac{\ln\left(\frac{1}{q} - 1\right) + \mu}{σ}\right)$$

$$= Φ\left(\frac{\ln\left(\frac{1}{q} - 1\right) + \mu}{σ}\right).$$

**Clayton copula.** Consider the threshold model derived from the Clayton copula with generator $φ(t) = t^{−θ} − 1$. The mixing variable in the corresponding Bernoulli mixture model is $p(Ψ) = e^{-Ψ(π^{-θ} - 1)}$, where $θ > 0$ and $Ψ ∼ Gam\left(\frac{1}{θ}, 1\right)$. To compute $P(p(Ψ) > q)$ we see that

$$P(p(Ψ) > q) = P(e^{-Ψ(π^{-θ} - 1)} > q) = P\left(-Ψ(π^{-θ} - 1) > \ln q\right)$$

$$= P\left(Ψ < -\frac{\ln q}{π^{-θ} - 1}\right) = \frac{γ\left(\frac{1}{θ}, -\frac{\ln q}{π^{-θ} - 1}\right)}{Γ\left(\frac{1}{θ}\right)}, \text{ where}$$

$$γ(s, x) = \int_0^x t^{s-1}e^{-t}dt$$

$$Γ(s) = \int_0^∞ t^{s-1}e^{-t}dt.$$
Standard & Poor rating of B would approximately correspond to an individual default probability of $\pi = 0.049$ and $\pi_2 = 0.00313$. These values imply a default correlation of:

$$\rho_Y = \frac{\pi_2 - \pi^2}{\pi - \pi^2} = \frac{0.00313 - 0.049^2}{0.049 - 0.049^2} = 0.0157.$$ 

With all models set to the same $\pi, \pi_2$ it remains to calibrate the parameters of each model. Depending on the model at hand, various techniques are used to determine these parameters. This will be discussed in more detail later on. For now, a summary of the resulting parameter values is presented in Table 4.1.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>S&amp;P rating B</th>
</tr>
</thead>
<tbody>
<tr>
<td>All models</td>
<td>$\pi$</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>$\pi_2$</td>
<td>0.00313</td>
</tr>
<tr>
<td></td>
<td>$\rho_Y$</td>
<td>0.0157</td>
</tr>
<tr>
<td>Beta</td>
<td>$a$</td>
<td>3.08</td>
</tr>
<tr>
<td></td>
<td>$b$</td>
<td>59.8</td>
</tr>
<tr>
<td>Probit-normal</td>
<td>$\mu$</td>
<td>-1.71</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>0.264</td>
</tr>
<tr>
<td>Logit-normal</td>
<td>$\mu$</td>
<td>-3.1</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>0.0.556</td>
</tr>
<tr>
<td>Clayton</td>
<td>$\pi$</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>$\theta$</td>
<td>0.032</td>
</tr>
</tbody>
</table>

Table 4.1: Model parameters that roughly corresponds to a Standard & Poor rating B according to (McNeil et al., 2005, p.364)

Finally, all the information needed to compare the tails of the four mixing distributions has been gathered. In Figure 4.2 the behavior of the tails of the distributions is illustrated. It can be seen that there are no significant differences between the tails up to the 99% quantile.
Recall from Section 3.3.2 that "the tail of the credit loss in large one-factor Bernoulli mixture models is essentially driven by the tail of the mixing variable" (Frey and McNeil, 2003). In the end, this implies that the portfolio loss distribution does not depend as heavily on the choice of mixing distribution, but rather on the estimates of $\pi$, $\pi_2$ and the default correlation. Thus, model risk in a one-factor Bernoulli mixture model for non-extreme quantiles is not as evident as the risk involved with finding good estimates of the individual default probabilities and the default correlation. We will shortly discuss the sensitivity in the Bernoulli mixture model to changes in the individual default probability and correlation. McNeil et al. (2005) states that the tail in Bernoulli mixture model are quite sensitive to changes in these parameters, and further that they can be difficult to estimate.

Further, let us connect our findings with the portfolio risk measurements of Value-at-Risk and Expected Shortfall. For the same mixing variables as above, consider a portfolio with $m = 1000$ obligors and a constant individual loss rate of $l = 60\%$. Since
the distribution functions of the four mixing variables are continuous and strictly increasing, and \( m \) is considered large, we can use the approximations of the VaR and ES from Proposition 4.1.6:

\[
\begin{align*}
\text{VaR}_{\alpha}(L^{(m)}) & \approx l \cdot m \cdot G^{-1}(\alpha) \\
\text{ES}_{\alpha}(L^{(m)}) & \approx \frac{l \cdot m}{1 - \alpha} \int_{\alpha}^{1} G^{-1}(u) du
\end{align*}
\]

where \( L^{(m)} \) is the total portfolio loss, and \( G \) is the distribution function of the mixing variable. Figure 4.3 illustrates the approximated VaR and ES for the four portfolios with \( \alpha \) varying between 97% and close to 100%. Using the approximation above it is straightforward to calculate the VaR, and thus the ES, in the mixture model. If the mixing variable is beta distributed then the inverse of \( G \) is simply given by the inverse of the beta cumulative distribution function. In the other three cases we make use of the above calculated \( P(p(\Psi) > q) \) to get \( G(q) = P(p(\Psi) \leq q) = 1 - P(p(\Psi) > q) \).

When \( G \) is known it is just a matter of calculating its inverse.

In Figure 4.3 the computations of VaR and ES for the four different mixing variables are displayed. In Figure 4.2 it was clear that the Logit-Normal mixing variable had the heaviest tail, and this is also apparent in Figure 4.3 as it is higher than the others.
From Figure 4.3 it can be seen that the VaR and ES increases exponentially as the confidence level $\alpha$ increases. As argued before it can also be seen in this figure that the choice of mixing distribution only implies a small change in the VaR and ES as all VaR and ES curves lie in a rather narrow interval for $\alpha \in [0.97, 1)$. We will now proceed with analyzing the sensitivity of the risk by the influence from individual default probability and correlation. Following the discussion that the choice of mixing distribution is of small importance once the parameters $\pi$ and the default correlation have been estimated, we will proceed with only analyzing the sensitivity for one of these mixing distributions. To this end, we let the mixing variable be beta distributed, which will yield quite simple calculations of the model parameters, i.e. the fitting of the beta parameters $a$ and $b$ to given $\pi$ and default correlation.

When the mixing variable is beta distributed, we have by definition that $p(\Psi) = \Psi$, i.e. $\Psi \sim \text{Beta}(a, b)$. This implies that the parameters $\pi, \pi_2$ and the default correlation are
given by:

\[ \pi = E[p(\Psi)] = E[\Psi] \]

\[ \pi_2 = E[\Psi^2] \]

\[ \text{Corr}(Y_i, Y_j) = \pi_2 - \pi^2 \pi^{-2} \].

Thus, \( \pi \) and \( \pi_2 \) are given by the first and second moment of the beta mixing variable, and the correlation in its turn is also calculated from these moments. In order to estimate the parameters \( a \) and \( b \), we then need to calculate \( E[\Psi] \) and \( E[\Psi^2] \) when \( \Psi \) is beta distributed. This is done in Proposition 4.2.1. However, in order to show these properties one needs to be familiar with some properties of the gamma distribution. These are summarized in Appendix B.1.

**Proposition 4.2.1 (Beta)**

Let \( \Psi \sim \text{Beta}(a, b) \), for \( a, b > 0 \), and let \( f \) be the density function of \( \Psi \). Then:

**a)** \( E[\Psi] = \frac{a}{a + b} \)

**b)** \( E[\Psi^2] = \frac{a(a + 1)}{(a + b)(a + b + 1)} \).

**Proof**

\[ a) \quad E[\Psi] = \int_0^1 xf(x)dx = \int_0^1 \frac{1}{\beta(a, b)} xx^{a-1}(1-x)^{b-1}dx \]

\[ = \int_0^1 \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} x^a(1-x)^{b-1}dx = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a + 1)}{\Gamma(a + b + 1)} \]

\[ = \frac{\Gamma(a + b)}{\Gamma(a)} \cdot \frac{a\Gamma(a)}{(a + b)\Gamma(a + b)} = \frac{a}{a + b} \]

**b)** Follows in the same way as the proof of a).

Hence, in order to fit the parameters \( a \) and \( b \) in the beta distribution to \( \pi \) and \( \pi_2 \) we need to solve the system of equations:

\[
\begin{cases}
\pi = \frac{a}{a + b} \\
\pi_2 = \frac{a(a + 1)}{(a + b)(a + b + 1)}
\end{cases}
\]
and after some calculations we get that

$$\begin{align*}
a &= \frac{(\pi_2 - \pi)\pi}{\pi^2 - \pi_2} \\
b &= \frac{(\pi_2 - \pi)\pi}{\pi^2 - \pi_2} \left( \frac{1}{\pi} - 1 \right)
\end{align*}$$

Given these equations it is now possible to compute the Value-at-Risk and Expected Shortfall for different correlation values. Figure 4.4 illustrates the results when the correlation varies between 0 and 0.5 for four different portfolios.

![Figure 4.4: Value-at-Risk and Expected Shortfall approximated as functions of correlation in portfolios with 1000 obligors and a constant loss rate of 60%. The four portfolios are driven by a beta mixing variable with four different values of the individual default probability.](image)

To illustrate the behavior of ES in relation to VaR, Figure 4.5 shows the difference between these risk measures as the correlation varies.
4.2.2 Mixture Model in the Merton Framework

We begin this section by working within the model discussed in subsection 3.4.2, i.e. the mixture model inspired by Merton. We want to be able to compare this model with the mixing distributions from last section. To do this we need some intermediate steps, since fitting the parameters to data in this model is a bit more complex compared to assuming a beta mixing variable. Given the parameters, $\pi$, $\pi_2$ and the default correlation, we want to be able to see which value of $\rho$ corresponds to these parameters. The parameters $\pi$ and $\pi_2$ are given by:

$$
\pi = \mathbb{E}[p(\Psi)]
$$

$$
\pi_2 = \mathbb{E}[p(\Psi)^2]
$$

where

$$
p(\Psi) = \Phi \left( \frac{\Phi^{-1}(\pi) - \sqrt{\rho} \Psi}{\sqrt{1 - \rho}} \right).
$$
We cannot find a closed form expression of \( \rho \) as a function of \( \pi \) and \( \pi_2 \) as we could with a beta mixing variable. However, we can obtain \( \rho \) by numerical integration. We will not give the details of this computation since the numerical procedures is not the main focus. At this point we are thus able to fit the mixture model inspired by Merton to given values of \( \pi \) and \( \pi_2 \) (or \( \pi \) and default correlation). What is left is then to use these parameters as inputs to calculate the risk, and compare it with using a beta mixing variable. The last step to be able to conduct this comparison is to derive an analytical expression for the VaR and ES in the Merton mixture model. If we can find an expression for the VaR the expression for the ES follows immediately. Furthermore, we will proceed by doing this with the large portfolio approximation. We know from Proposition 4.1.6 that the approximations

\[
\text{VaR}_\alpha(L^{(m)}) \approx l \cdot m \cdot G^{-1}(\alpha) \quad (4.2.1)
\]

\[
\text{ES}_\alpha(L^{(m)}) \approx \frac{l \cdot m}{1 - \alpha} \int_0^1 G^{-1}(u) du. \quad (4.2.2)
\]

hold in a large portfolio in this setting. The distribution function \( G \) is the distribution function of the corresponding mixing variable \( p(\Psi) \), which we in Section 3.4.2 derived to be:

\[
G(x) = \Phi \left( \frac{1}{\sqrt{1 - \rho}} \left( \sqrt{1 - \rho} \Phi^{-1}(x) - \Phi^{-1}(\pi) \right) \right).
\]

Thus, in order to get an approximation of the VaR, we need to find the inverse of \( G \). By simple calculations we get:

\[
G^{-1}(y) = \Phi \left( \frac{\Phi^{-1}(y) \sqrt{1 - \rho} + \Phi^{-1}(\pi)}{\sqrt{1 - \rho}} \right).
\]

Thus the approximations of VaR and ES become:

\[
\text{VaR}_\alpha(L^{(m)}) \approx l \cdot m \cdot \Phi \left( \frac{\Phi^{-1}(\alpha) \sqrt{1 - \rho} + \Phi^{-1}(\pi)}{\sqrt{1 - \rho}} \right) \quad (4.2.3)
\]

\[
\text{ES}_\alpha(L^{(m)}) \approx \frac{l \cdot m}{1 - \alpha} \int_\alpha^1 \Phi \left( \frac{\Phi^{-1}(u) \sqrt{1 - \rho} + \Phi^{-1}(\pi)}{\sqrt{1 - \rho}} \right) du. \quad (4.2.4)
\]

As can be seen the risk in a portfolio, as measured by these two risk measures, depends on a lot of factors. These factors are the loss rate, the number of obligors, the confidence level, the individual default probability and the correlation parameter. We have now obtained everything we need. Figure 4.6 illustrates the difference (or rather similarity) between having a mixing variable inspired by Merton and a beta mixing
variable. The difference is illustrated by showing the impact on the risk as a function of default correlation for two different individual default probabilities with these two mixing variables.

![Graph showing Value-at-Risk and Expected Shortfall with beta mixing variable versus mixing variable inspired by Merton. VaR and ES are computed as functions of the individual default correlation in portfolios of 1000 obligors and a constant loss rate of 60%.

Figure 4.6: Value-at-Risk and Expected Shortfall with a beta mixing variable versus a mixing variable inspired by Merton. VaR and ES are computed as functions of the individual default correlation in portfolios of 1000 obligors and a constant loss rate of 60%.

It is clear from Figure 4.6 that the two mixing variables yield similar results. Thus, we can conclude that the choice of mixing variable is of minor importance when modelling static credit portfolios. The main problem is rather to find good estimates of the parameters $\pi_1$, $\pi_2$ and the default correlation. The most common industry models differ in the approach of finding these parameters.
Chapter 5

Conclusion

We have seen that there is a wide variety of approaches to choose from when modelling static portfolio credit risk. Of crucial importance is the modelling of default dependence, and therein lies also the major challenge. Dependence structures in credit risk portfolios are mathematically complex, and as we have seen there are several ways to address this issue. One concern here is to find a model that can account for a wide range of dependence structures.

The most general model is the Bernoulli mixture model, which makes use of a factor vector of mixing variables to model dependence. This framework is general in the sense that most threshold models can be expressed as a Bernoulli mixture model. The choice of mixing variables is vital, as the distribution of these variables will completely determine the distribution of the loss and also the default correlation. However, when fitting different mixing variables to a given set of data (individual default probabilities and default correlations) we have noted that there is only a minor discrepancy between the distributions of the mixing variables up until the extreme quantiles. Thus, the model risk in Bernoulli mixture models is not as evident as the risk involved with finding accurate estimates of the individual default probabilities and default correlations. Analogously, the portfolio loss distribution in threshold models using the concept of copulas is completely determined by the choice of copula.

Further, it is of interest to study the behavior of a portfolio when the number of obligors is large. It was concluded that in the Bernoulli mixture models the distribution of the
portfolio loss fraction converges to the distribution of the mixing variable. From this asymptotic behavior we have also been able to derive large portfolio approximations for several model frameworks. These large portfolio approximations are extremely useful since the portfolio loss distribution in many models cannot be expressed in closed form.

This thesis can serve as an introduction to static credit risk modelling. We have mainly considered homogeneous and exchangeable portfolios, and as we have seen the task of modelling credit risk is still far from trivial. In homogeneous portfolios it is enough to study default distribution instead of the portfolio loss distribution, which has simplified both calculations and the mathematical modelling. Further extensions from here would be to relax some of the assumptions and make the portfolio more general and heterogeneous. For instance, a first step could be to assume a heterogeneous portfolio and then strategically categorize the obligors into several groups, where each group can be viewed as a homogeneous portfolio. Another direction could be to study the modelling of loss rates more rigorously, especially to model them as stochastic variables. One could also introduce a dependence between loss rates and defaults, where the intuition would be that loss rates tend to be larger when the general economy is in a recession and more defaults occur.
Bibliography


Herbertsson, A. (2009), ‘Credit Risk Modelling course’.


Appendix A

Copulas

A.1 Preliminaries

When dealing with copulas the reader need to be comfortable with two elementary operations of probability and quantile transformations. These are summarized in Proposition A.1.1. The outline of the proposition and proof is similar to the one in McNeil et al. (2005).

Proposition A.1.1

Let \( G \) be a distribution function and let \( G^{-1} \) denote its inverse.

1. Quantile transformation. If \( U \sim U(0, 1) \) has a standard uniform distribution, then \( P(G^{-1}(U) \leq x) = G(x) \).

2. Probability transformation. If \( Y \) has distribution function \( G \), where \( G \) is a continuous univariate distribution function, then \( G(Y) \sim U(0, 1) \).

Proof:

Let \( x \in \mathbb{R} \) and \( u \in (0, 1) \). The first part of the proof follows from:

\[
P(G^{-1}(U) \leq x) = P(U \leq G(x)) = G(x).
\]

To prove the second part we proceed with:

\[
P(G(Y) \leq u) = P(G^{-1}(G(Y)) \leq G^{-1}(u)) = P(Y \leq G^{-1}(u)) = G(G^{-1}(u)) = u,
\]

where the first inequality follows from the fact that \( G \) is continuous.
A.2 Proof of Sklar’s Theorem

Proof (Sklar’s Theorem)
We will prove a simplified version of Sklar’s Theorem, where it is assumed that the marginal distribution functions $F_1, \ldots, F_d$ are continuous. For a more general proof, see Schweizer and Sklar (1983, p. 83). Note first that if $X$ is a random variable and $T$ is an increasing function, then \( \{ X < x \} \subset \{ T(X) \leq T(x) \} \) and \( P(T(X) \leq T(x)) = P(X \leq x) + P(T(X) = T(x), X > x) \). Using this property, for any \( x_1, \ldots, x_d \in \mathbb{R} = [-\infty, \infty] \) we see that if $X$ has distribution function $F$, then

$$F(x_1, \ldots, x_d) = P(F_1(X_1) \leq F_1(x_1), \ldots, F_d(X_d) \leq F_d(x_d)).$$

Since we assume that $F_1, \ldots, F_d$ are continuous, Proposition A.1.1 together with the definition of copula imply that the distribution function of $(F_1(X_1), \ldots, F_d(X_d))$ is a copula. Let this copula be denoted by $C$. Thus, we obtain Equation (3.5.1) in Sklar’s Theorem.

Further, assume that $C$ is a copula and that $F_1, \ldots, F_d$ are univariate distribution functions. It remains to show that these assumptions imply that the function $F$ in Equation (3.5.1) is a joint distribution function with marginal distribution functions $F_1, \ldots, F_d$. To this end, let $U$ be a random vector with distribution function $C$, as in Equation (3.5.1), and define the random vector $X := (F_1^{-1}(U_1), \ldots, F_d^{-1}(U_d))$. Then, we see that

$$P(X_1 \leq x_1, \ldots, X_d \leq x_d) = P(F_1^{-1}(U_1) \leq x_1, \ldots, F_d^{-1}(U_d) \leq x_d)$$

$$= P(U_1 \leq F_1(x_1), \ldots, U_d \leq F_d(x_d))$$

$$= C(F_1(x_1), \ldots, F_d(x_d)).$$

\[\blacksquare\]

A.3 Completely Monotonic

A decreasing function $g(x)$ is called completely monotonic on $(a, b)$ if it satisfies:

$$(-1)^k \frac{d^k}{dx^k} g(x) \geq 0, \text{ for } k \in \mathbb{N}, \text{ and } x \in (a, b). \quad (A.3.1)$$

This definition of a completely monotonic function is from McNeil et al. (2005, p. 222).
A.4 Simulation Algorithms

The following algorithms describe how to simulate from their respective copula. The algorithms are from Hult and Lindskog (2007).

Simulation from Gaussian copula:
Let the correlation matrix be denoted by $R = AA^T$ for some $d \times d$ matrix $A$, where $R$ is strictly positive definite.

- Let $A$ be the Cholesky decomposition of $R$. Then $R = AA^T$.
- Simulate $d$ independent random variables $Z_1, \ldots, Z_d$ from $\Phi(0, 1)$
- Set $X = AZ$
- Set $U_k = \Phi(X_k)$ for $k = 1, \ldots, d$
- $U = (U_1, \ldots, U_d)$ has then distribution function $C_R^{Ga}$

Thus, a default of obligor $i$ occurs if $U_i \leq p_i$.

Simulation from $t$ copula:
Let the correlation matrix be denoted by $R = AA^T$ for some $d \times d$ matrix $A$, where $R$ is strictly positive definite. Further $\chi^2$ denotes the chi-square distribution with $\nu$ degrees of freedom

- Let $A$ be the Cholesky decomposition of $R$. Then $R = AA^T$.
- Simulate $d$ independent random variables $Z_1, \ldots, Z_d$ from $\Phi(0, 1)$
- Simulate a random variable $S$ from $\chi^2$ independent of $Z_i$ for all $i$
- Set $Y = AZ$
- Set $X = \frac{\sqrt{\nu}}{\sqrt{S}} Y$
- Set $U_k = t_\nu(X_k)$ for $k = 1, \ldots, d$
- $U = (U_1, \ldots, U_d)$ has then distribution function $C_{\nu, R}^t$

Thus, a default of obligor $i$ occurs if $U_i \leq p_i$.

Simulation from Clayton copula:
We established in Example 3.5.9 that the Clayton copula can be generated from taking the Laplace transform of a gamma distributed random variable with parameters $\frac{1}{\theta}$ and 1. Furthermore, the algorithm uses the result of the Marshall Olkin theorem, i.e. Theorem 3.5.10. Let $\theta$ be the parameter in the Clayton copula.

- Simulate a random variable $X \sim \Gamma(\frac{1}{\theta}, 1)$
• Simulate $d$ independent random variables $V_1, \ldots, V_d$ from Uni(0, 1)
• If $\Psi(s) = (s + 1) - \frac{1}{s}$, then $U = \left( \Psi\left(-\ln(V_1)\right), \ldots, \Psi\left(-\ln(V_d)\right) \right)$ has d.f. $C^\alpha$.

Thus, a default of obligor $i$ occurs if $U_i \leq p_i$. 
Appendix B

Distributions

B.1 Gamma Distribution

In Proposition 4.2.1 some properties of the gamma distribution, and the gamma function were used. Here these properties are summarized.

Definition B.1.1 A random variable $X$ is gamma distributed with parameters $\lambda, t > 0$, $X \sim \text{Gam}(\lambda, t)$, if its density, $f$, is given by:

$$f(x) = \frac{1}{\Gamma(t)} \lambda^t x^{t-1} e^{-\lambda x}, \quad x \geq 0$$

where $\Gamma(t)$ is the gamma function defined by

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx.$$

One can prove that the gamma function satisfies:

$$\Gamma(t + 1) = t \Gamma(t).$$

Also the following connection between the beta function and gamma function holds:

$$\beta(a, b) = \int_0^1 x^{a-1} (1 - x)^{b-1} dx = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a + b)}.$$