POLITECNICO DI TORINO

Master's degree in
Mathematical Engineering

Master's degree thesis

Modeling Credit Default: A Portfolio-Based Comparison of CreditMetrics and CreditRisk+

Supervisor
prof. Patrizia Semeraro

Candidate
Francesca Vogliotti

Academic year 2022-2023

Alla mia Mamma
e al mio Papà

Acknowledgements

Firstly, I would like to sincerely thank Professoressa Patrizia Semeraro, who supervised this thesis work and guided me towards topics that sparked my curiosity, while allowing me to focus on the subjects that resonated with my interests, such as implementation and simulation. Through her consistent support and genuine kindness, she made it possible for me to balance this work with my full-time dream job.

This work marks the final chapter of a journey that commenced years ago, a journey that has facilitated substantial growth in both my personal and academic realms. With immense gratitude, I acknowledge the multitude of individuals who have stood by my side throughout this voyage.

I want to express my gratitude to my family, particularly my mother, Donna, my father, Riccardo, and my brother, Andrea, for their unwavering support and belief in me.

I want to extend my thanks to those who have been with me from the very beginning: Giuseppe P., Francesco, Marco, Alessandra, Andrea, Giuseppe M., Thomas, and Tayisiya. You made my years at Politecnico full of learning, laughter, and support. We have grown so much together over the last five years.

I also wish to convey my gratitude to the friends I made through Alta Scuola Politecnica, individuals who have become lifelong companions: Matteo, Glenda, Francesca, Giovanni, Silvia, Luca, and Mario.

I'm thankful to those who joined me during long study hours, making them more enjoyable: Virginia, Elena, Alessandro, Jacopo, Samuele, Giorgia, Simone, Alessandra, and Lucrezia. You have filled my days with happiness and support, and I will forever cherish our friendship.

An immense thank you goes to the amazing girls I have shared all joys and hardships with in the last three years: Martina and Cecilia.

I want to express my gratitude to those who have always been there for me, my life partners: Rebecca and Elena.

A heartfelt thank you to the people who encouraged me to take risks and explore places I had never thought of, whether it was stepping onto the stage of a startup competition or discovering the 10th arrondissement in Paris: Tommaso, Davide, Andrea, Enrico, and Giuseppe P. You changed my life.

Finally, I want to thank those who made me feel so welcome in Paris: Alice, Elisabeth, Gaspard, Jeremy, Gregoire, and the new amazing batch of Associates.

To all who have left an indelible mark on my journey and have contributed to where I stand today, I am immensely grateful.

Contents

List of Algorithms

Chapter 1

Introduction

Credit risk is the possibility of losing a lender holds, due to a risk of default on a debt that may arise from a borrower failing to make required payments [15]. The modelling of this risk is an indispensable tool utilized by financial institutions worldwide - including banks, insurance companies and investment firms - to effectively measure, control and manage the risk associated with lending. The sphere of credit risk though is by no means confined to just loans or borrowed money; it extends to other areas of financial transactions, including bonds, derivatives, and other financial products. For this reason, this element of risk has triggered a wave of interest from various stakeholders ranging from academia to industry players, regulators and policy makers.

Seemingly abstract, credit risk can potentially put banks at the brink of bankruptcy or stimulate economic crises on a global scale, as the 2008 financial crisis lucidly exhibited. Thus, the ability to properly model and understand credit risk is not only important but utterly crucial to the health of financial institutions and economy as a whole. Organizations seek to comprehend credit risk to ensure sound decision-making, to optimize returns on investment, and to mitigate the possibility of unforeseen financial losses. A comprehensive understanding of credit risk modelling also assists institutions in regulatory compliance. Indeed, regulations such as the ones brought by the Basel III Capital Accord (2010), highlight the significance of credit risk modelling in calculating the Value at Risk (VaR) and the associated economic capital.

This thesis focuses on modelling the risk associated a portfolio of corporate bonds, using and comparing two of the most prominent models: CreditMetrics and CreditRisk+. First, we aim to explain the mathematical foundations of the two frameworks, highlighting the differences and similarities in the two approaches. Second, we explain how to parameterise consistently the inputs of the two framework. Third, we provide step by step implementations of the two models in their various possible implementations, comparing the results yielded in the default and loss distributions relative to two sample portfolios (one with debtors of high quality bonds, one of low). We finally aim to shed light on the possible discrepancies in risk measure calculations and model choice for financial institutions.

The thesis is organised as follows:

- Chapter 2 introduces the foundational concepts for modeling Credit Risk. It presents two widely used mathematical frameworks: Structural and Mixture models, along with their commercial implementations, CreditMetrics and CreditRisk+.
- In Chapter 3, we delve into a practical case study featuring two realistic portfolios. We implement CreditMetrics and CreditRisk+ under varying modeling assumptions, initially assuming independence and later incorporating dependency structures. This chapter reveals the intricacies of maintaining consistency in the implementation, particularly when dealing with dependencies. We conclude by comparing resulting distributions, risk measures, and the efficiency of each approach, yielding consistent results.
- Chapter 4 offers concluding insights into the challenges encountered when modeling credit risk and comparing different models. This thesis work highlights that CreditMetrics is preferable for richer modeling, incorporating credit migrations and coupons. In contrast, CreditRisk+ excels in efficiently managing extensive debtor portfolios, especially in retail settings.

Chapter 2

Credit Risk Modelling

Credit Risk is omnipresent in the portfolio of a typical financial institution, and has been also at the heart of many recent developments on the regulatory side following the 2008 crisis and the Basel III accords. For this reason the modelling of this risk has been an active field of research for academia, financial players and regulators. These models are in fact used to determine the loss distribution of a debt portfolio due to defaults (or credit migrations), and to compute the associated risk measures, in order to comply with the required risk-capital allocations [12].

2.1 Definition of Credit Risk

A comprehensive definition of credit risk can be defined as follows [16]: "the possibility that an unexpected change in the creditworthiness of a counterparty will lead to a corresponding unexpected change in the current value of the related credit exposure." This definition brings to light that credit management not only involves the probability of the counterparty's insolvency but also includes the risk of downgrading, which is the possibility of deterioration of the counterparty's creditworthiness.

It is possible to identify four components of credit risk: Probability of Default (PD), Exposure at Default (EAD), Loss Given Default (LGD) and Maturity (M). The estimation of these variables is expressly required by regulation and is the basis of the internal rating method introduced with (Basel 2).

- The PD indicates the probability that a counterparty will default within a given time horizon. This measure is related to the debtor as such, i.e., it is independent of the type of exposure. The PD distribution can be described by a Bernoulli random variable, which can take on the two different values of "solvency" or "default".
- The EAD is the expected value of the exposure, conditional on the state of default. This variable depends on the technical form of the exposure and can therefore be deterministic or stochastic, when the exposure is not known in the possible states of default.

- The LGD corresponds to the loss rate in case of default, i.e., the expected value (possibly conditioned to adverse scenarios) of the ratio, expressed in percentage terms, between the loss due to default and the EAD. The complement to one of this variable is called Recovery Rate (RR).
- Maturity is the average, for a given exposure, of the remaining contractual durations of the payments.

These measures, when appropriately studied, allow us to estimate Expected Losses (EL) and Unexpected Losses (UL).

- Expected loss can be defined as the mean value of the loss distribution, considered as a random variable, that an entity supposes to suffer on a single investment or a portfolio of investments. However, this loss, precisely because it is "expected", cannot be used to determine the actual degree of risk. It can be calculated as follows:

  $$EL = EAD \cdot PD \cdot LGD \tag{2.1}$$

  Therefore, we can affirm that the expected loss is a function of three elements: exposure at default, probability of default, and loss given default. The expected loss concept is applicable both to single exposure and to the measurement of the entire loan portfolio. To extend the application to the latter case, sum up the individual exposures present in the portfolio. Specifically:

  $$EL = \sum_{i=1}^{n} EL_i = \sum_{i=1}^{n} PD_i \cdot LGD_i \cdot EAD_i \tag{2.2}$$

  We can rewrite the equation in terms of percentages. In this case, we refer to the expected loss rate (ELR). This refers to the fraction of the portfolio value at risk obtained by dividing the expected loss by the total value of the exposures in case of default. Specifically, we have:

  $$ELR = \frac{EL}{\sum_{i=1}^{n} EAD} = \sum_{i=1}^{n} w_i \cdot LGD_i \cdot PD_i \tag{2.3}$$

- The unexpected loss is the average total loss over and above the mean loss. It is calculated as a standard deviation from the mean at a certain confidence level. It is also referred to as Credit VaR. The unexpected loss of a portfolio at a 99% confidence level will be expressed as follows:

  $$UL_{99\%} = D_{99\%} - EL \tag{2.4}$$

  where $D_{99\%}$ represents the 99% VaR quantile, and the VaR is defined as the smallest x such that the cumulative probability of a loss less than or equal to x is at least α, $f_L(l)$ is the loss distribution function:

  $$VaR_\alpha = \inf\{x \in \mathbb{R} : \int_{-\infty}^{x} f_L(l)dl \geq \alpha\}$$

We can consider the realizations of losses as random variables whose distribution can be exemplified in Figure 2.1.

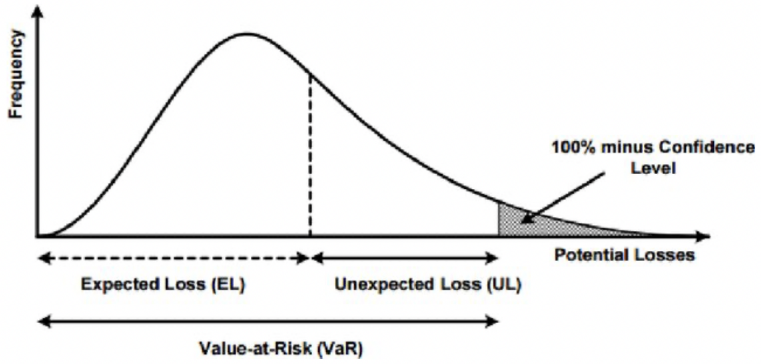


Figure 2.1. The relation between Expected Loss, Unexpected Loss, and Value at Risk, adopted from Bank of International Settlements (BIS), taken from [7]

2.2 Credit Risk Models

Quantitative credit risk models can essentially be structured into two categories [12]: structural (or firm-value) models and reduced-form models.
Merton (1974) stands as the pioneer of all firm-value models. These models suggest that a firm's default is related to the association between its assets and owed liabilities at a specified time end. In structural models, defaults occur when an asset value, represented by a stochastic variable, falls short of the liability threshold, for these reason, particularly at portfolio level, they can be referred as threshold models.
On the contrary, reduced-form models don't explicitly state the mechanism resulting in default, and assume conditioning independence of defaults given shared underlying stochastic factors.
In the case of a portfolio of credit exposures, these models mostly employ Monte Carlo simulation techniques to calculate the distribution of defaults and losses, and the associate risk measures.

The quantitative modelling of credit risk faces some unique challenges [12].
Firstly, there is an insufficiency of publicly accessible credit data. Facts concerning corporate credit quality are often not well-documented. This inhibits corporate lending, given that company management typically has better information about the firm's economic outlook and default risk than potential lenders. These information asymmetries are well-recognized in microeconomical literature. Moreover, the scarcity of data thwarts the effective use of statistical methods in credit risk, which is amplified given that the risk management span is usually a minimum of one year. This data deficiency is the principal barrier to the accurate calibration of credit models.

Secondly, credit losses exhibit a skewed distribution pattern with the upper tail having substantial weight. This means that a typical credit portfolio over time will comprise frequent small profits juxtaposed with infrequent enormous losses. Consequently, a significant proportion of risk capital (or economic capital) is needed to maintain such a portfolio. The economic capital necessary for a loan portfolio (the level of risk capital deemed permissible by the institution's shareholders and board of directors, irrespective of the regulatory framework) often equates to the 99.97% quantile of the loss distribution.

Thirdly, the role of dependence modelling cannot be overlooked. The occurrence of an excessively high number of defaults from different counterparties within a specific time frame poses a significant credit risk to a particular bond or loan portfolio. Importantly, this risk is directly related to the dependency structure of the default events. For a large portfolio, default dependency critically influences the upper tail of a credit loss distribution. This behaviour will be further presented in the case study example.

2.3 Structural Models

The structural model for credit risk, sometimes called contingent claims modeling, was introduced by Merton (1974) and extended by Leland, 1994, Leland and Toft, 1996, Anderson and Sundaresan, 1996, among others. They are credit risk models which aim to provide an explicit relationship between default risk and capital structure [19]. Default occurs whenever a stochastic variable or process, representing an asset value, falls below a threshold representing liabilities. The main advantage of this representation is the fact that they provide an endogenous explanation for default. This approach was pioneered by the works of Black & Scholes (1973) and Merton (1974) which model asset log returns with a geometric brownian motion process and numerous extensions have been proposed throughout literature.
The greatest challenge in structural models lies in the determination of the value of the firm, as corporate debt is not traded. The assumption of asset log returns driven by a diffusion process allows a closed-form formula between the firm's value and its equity via Black & Scholes formula, which can be solved via system of non linear equations.
The Brownian hypothesis has been challenged as it fails to describe historic default data accurately and produces almost zero default probabilities for short maturities [2], phenomenon known in literature as *credit spread puzzle*. For this reason sever enhancements of the model have been proposed throughout the years, better accounting for extreme market events.

2.3.1 The Merton Model

The intuition behind the model is treating a company's equity as a call option on its assets with strike price equal to the book of value of the firm's liabilities, thus allowing for applications of Black-Scholes option pricing methods [1].
The value of the firm is considered to be a tradable asset which obeys a lognormal diffusion process $V_t = S_t + B_t$, where:

- S_t is the value at time t of equity;
- B_t is the book value of the debt at time t, consisting of a single zero coupon bond with face value K and maturity T.

The model assumes friction-less markets (no transactional costs or taxes) and the fact that the firm cannot pay out dividends or issue new debt. At maturity there are two possibilities:

1. if $V_T > K$ there is no default. Debt holders receive K and shareholders receive the residual $S_T = V_T - K$;
2. if $V_T \leq K$ the value of the firm's asset is less than its liabilities, so the firm cannot meet its financial obligations. The firm defaults and debt holders have the first claim on residual asset, gaining $B_T = V_T$, shareholders are left with nothing, so $S_T = 0$.

Therefore, equity value at time T can be written as:

$$S_T = max(V_T - K, 0) = (V_T - K)^+ \tag{2.5}$$

which is the pay-off of a European call option on the firm's assets V_T with strike price equal to the book value of the firm's liabilities. Debt value can be written as:

$$B_T = min(V_T, 0) = K - (K - V_T)^+ \tag{2.6}$$

which is the nominal value of the liabilities minus the pay-off of a European put on V_T, with strike K.
Under the real world probability, Merton assumes that V_t follows a geometric Brownian motion of the type

$$dV_t = \mu_V V_t dt + \sigma_V V_t dW_t. \tag{2.7}$$

As in [6], using Ito's lemma:

$$d(lnV_t) = (\mu - \frac{\sigma^2}{2})dt + \sigma dW_t$$

$$lnV_T \sim N\Big(lnV_t + (\mu - \frac{\sigma^2}{2})(T-t), \sigma^2(T-t)\Big).$$

The probability of default is the probability that the market value of the firm's assets will be less than the book value of the firm's liabilities by the time the debt matures:

$$P(V_T < K) = P(lnV_T < lnK) = \Phi\Big(\frac{K - \mathbb{E}(V_T)}{\sigma(V_T)}\Big)$$

$$PD = \Phi\Big(\frac{ln\frac{K}{V_T} - (\mu - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\Big)$$

Hence it is possible to determine the probability of default knowing V_t, K, μ, σ_V^2. The market value of the firm V_t and its volatility σ are in general not observable, as reliable data on the market value of the firm's debt is in general unavailable. The latter is estimated by reverse engineering the well-known Black-Scholes formulas. Under this framework, a credit default at time T is triggered by the event that shareholders' call option matures out-of-money, with a risk-neutral probability, therefore the value of equity can be written as:

$$S_t = V_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2) \tag{2.8}$$

where

$$d_1 = \frac{ln(V_t/K) + (r + \sigma_V^2/2)(T-t)}{\sigma_V\sqrt{T-t}}, \quad d_2 = d_1 - \sigma_V\sqrt{T-t}.$$

From Ito's lemma the volatility of the market value of the firm can be written as

$$\sigma_V = \left(\frac{V}{S}\right)\Phi(d_1)\sigma_S \tag{2.9}$$

Rearranging the Black-Scholes formula:

$$V_t = \frac{S_t + Ke^{-r(T-t)}\Phi(d_2)}{\Phi(d_1)}. \tag{2.10}$$

As in [5], the first step in implementing the model is to estimate the volatility of the equity from historical data. The second step is to choose a forecasting horizon and a measure of the face value of the firm's debt, in literature the common forecasting horizon is one year and the face value of the firm's debt is taken to be the book value of the firm's total liabilities. After collecting the value of the risk free rate, system (2.8) and (2.9) can be solved to find V and σ_V.

Given the equally spaced time series of N observations of the firm's market capitalization S_t, V_t and σ can by estimated through maximum likelihood methods or through iterative ones, such as in Vassalou and Xing [18].

2.3.2 CreditMetrics

CreditMetrics, introduced in 1997 by J.P. Morgan and RiskMetrics Group, belongs to the structural models since it is based on the model of Merton for the definition of thresholds for the migration of credit. More explicitly, Merton's model is extended by assuming that the assets returns of the firm determine not only its probability of default but also the probability of migrating to any other credit rating. In the model, portfolio risk is assessed due to changes in debt value, which are caused by changes in obligor credit quality [14]. As changes in value of the portfolio are not only caused by possible default, but also by upgrades or downgrade in credit-quality, CreditMetrics falls under the umbrella of *mark-to-market* models. Returns are assumed normally distributed, therefore a change in the credit quality of the firm occurs when its returns fall within certain thresholds in the normal distribution.

In this approach, migrations and consequently defaults depend of the asset value returns of the firms, thus it is not necessary to know the value of debt, actives, and volatilities. CreditMetrics is consequently an "agnostic" model, which takes in input historical values for default and credit migration probabilities, and models the asset value returns of the portfolio with a multivariate normal to incorporate dependencies between borrowers. Each firm is assigned a discrete risk level, identified by the rating class to which it is assigned to by the corresponding rating agency (such as Moody's or Standard & Poor's), which can migrate from a class rating to another, given a certain time horizon. The state of default constitutes an *absorbent* state, meaning that this state, once entered it is not possible to have a transition to any other state. Each state has a different likelihood or probability of occurring, which is derived from historical rating data. Given this premise, it is possible to build a transition matrix, which contains the migration probabilities from all possible states at the end of the time period. Finally, it is possible to calculate the value of the bond given the rating class at the end of the time period.

Risk for a stand-alone exposure

To calculate the distribution of values for a single bond, we present the methodology explained in [14], and summarised in Figure 2.2.

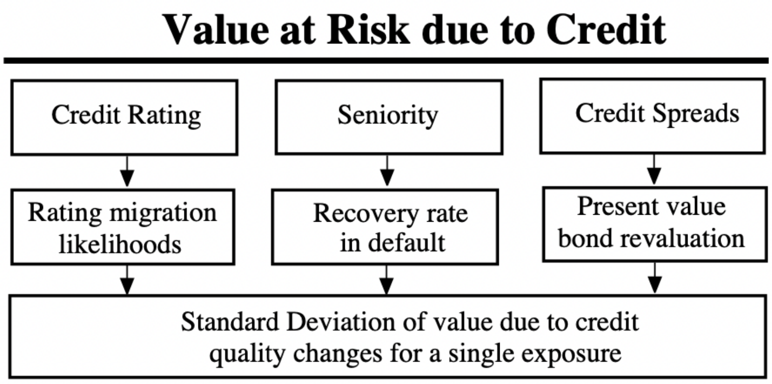


Figure 2.2. Three steps to calculating the credit risk for a "portfolio" of one bond [14]

Given an exposure of a certain rating, for example BB by Standard & Poor's, at the end of the time period either the bond maintained the same credit worthiness, or it migrated to other states (upgraded to BBB, A, AA, AAA, downgraded to B, CCC, or to default). Each outcome above has a different likelihood or probability of occurring, derived from historical rating data, as shown in the example in Figure 2.3.
The next step is calculating the value of the exposure under each possible rating scenario by finding the new present value of the bond's remaining cash flows at its new rating. The discount rate that enters this present value calculation is read from the forward zero curve that extends from the end of the risk horizon to the maturity of the bond. This zero curve is different for each forward rating category. The present value bond revaluation can be calculated as in Formula 3.1.
In the case of default, the amount recovered in is specified by recovery rates, as in Figure

Year-end rating	Probability (%)
AAA	0.02
AA	0.33
A	5.95
BBB	86.93
BB	5.30
B	1.17
CCC	0.12
Default	0.18

Figure 2.3. Probability of credit rating migrations in one year for a BBB bond [14]

2.4.

Recovery rates by seniority class (% of face value, i.e., "par")

Seniority Class	Mean (%)	Standard Deviation (%)
Senior Secured	53.80	26.86
Senior Unsecured	51.13	25.45
Senior Subordinated	38.52	23.81
Subordinated	32.74	20.18
Junior Subordinated	17.09	10.90

Source: Carty & Lieberman [96a] —Moody's Investors Service

Figure 2.4. Recovery rates by seniority class (% of face value, i.e., "par") [14]

The resulting discrete distribution is a set of future values, associated with a corresponding probability of occurring (as shown in the example in Figure 2.5).
Details of this calculation will be presented in the case study in the following chapter.

Obtaining a distribution of values for a portfolio multiple bonds

Firstly, let's consider the idea of a portfolio consisting of two bonds. Similar to our previous approach, we can calculate the distributions of each individual bond and subsequently combine these to ascertain the year-end estimations of the entire portfolio. Given that either of the two bonds may possess any of eight potential values within a one-year span as an outcome of rating shifts, the portfolio may assume 64 (8 x 8) distinct values. The portfolio's worth at the risk horizon is determined for each of these 64 states by simply

Year-end rating	Value ($)	Probability (%)
AAA	109.37	0.02
AA	109.19	0.33
A	108.66	5.95
BBB	107.55	86.93
BB	102.02	5.30
B	98.10	1.17
CCC	83.64	0.12
Default	51.13	0.18

Figure 2.5. Distribution of value of a BBB par bond in one year [14]

summing up the values of the individual bonds and linking them to year-end joint probabilities across the 64 unique states.
This becomes a straightforward process if the ratings results of the two bonds operate independently. In this instance, the joint probability is merely a multiplication of the individual probabilities. However, in reality, the ratings results of the two bonds are interrelated, since they are significantly influenced by similar macro-economic elements. Hence, it's crucially important to consider the correlation between rating migrations when assessing the risk of a portfolio, thus resulting in the framework shown in Figure 2.6. As

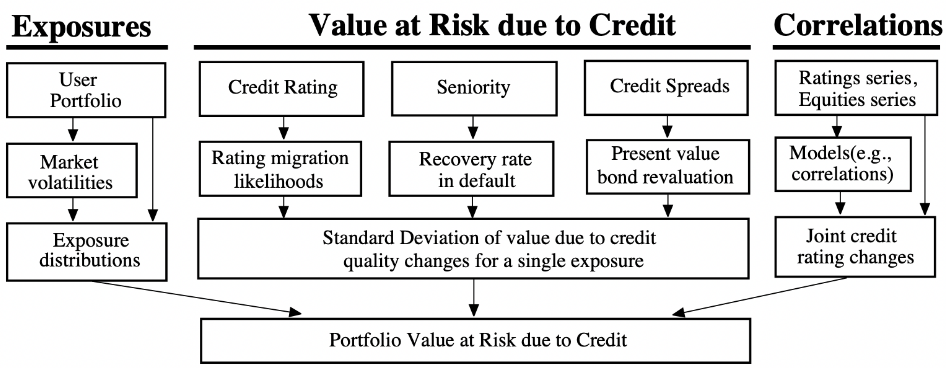


Figure 2.6. CreditMetrics framework for a Portfolio [14]

the portfolio grows larger, the complexity of the distribution grows exponentially and the process described above becomes impossibly costly, as there are 8^N joint rating states. For this reason we use simulation in order to calculate the distribution of defaults, losses and risk measures of a large portfolio.
This can be done using the framework introduced by Merton, and described in section

2.3.1: the underlying firm value is random with some distribution, and if the value of assets should happen to decline so much that the value is less than amount of liabilities outstanding, the firm defaults. This model can be extended to allow for rating migrations, including in addition to the default threshold credit rating up(down)grade thresholds, as in Figure 2.7.

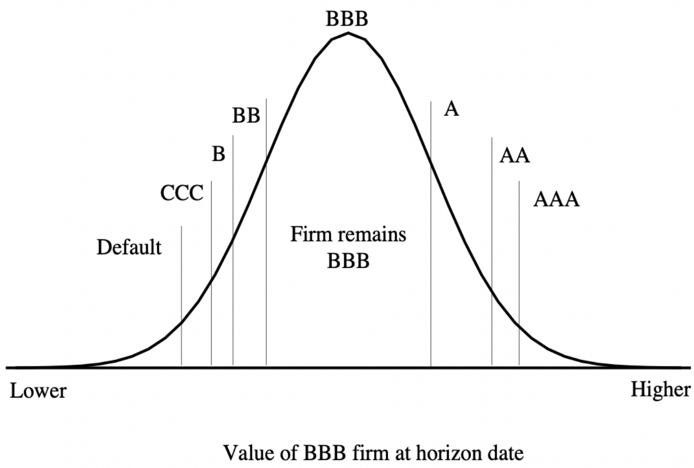


Figure 2.7. Model of firm value and generalized credit quality thresholds, taken from [14]

Essentially, the assumption is that the asset value in one year determines the credit rating (or default) of the company at that time. We assert that the percent changes in asset value are normally distributed, and parameterized by a mean and standard deviation. With this assumption it is possible to calculate the thresholds as:

$$\Pr\{\text{Default}\} = \Pr\{\text{R} < \text{ZDef}\} = \Phi\left(\frac{ZDef}{\sigma}\right) \tag{2.11}$$

$$\Pr\{\text{CCC}\} = \Pr\{\text{ZDef} < \text{R} < \text{ZCCC}\} = \Phi\left(\frac{ZCCC}{\sigma}\right) - \Phi\left(\frac{ZDef}{\sigma}\right) \tag{2.12}$$

and so on. To describe the joint evolution of the portfolio, it is possible to incorporate correlation thanks to the use of the multivariate normal distribution, for example for two obligors:

$$\Pr\{\text{ZB} < \text{R} < \text{ZBB, Z'BBB} < \text{R'} < \text{Z'A}\} = \int_{ZB}^{ZBB}\int_{Z'BBB}^{Z'A} f(r, r'; \Sigma)dr'dr \tag{2.13}$$

The procedure can be split in the following steps:

1. Calculate asset return thresholds for the obligors in the portfolio, as explained above.

2. Generate scenarios of asset returns according to the normal distribution, using commands to such as *mvnrnd* in Matlab and *mvrnorm* in R to generate the random numbers.

3. For each scenario associate the new credit state (AAA, AA, BB, etc or 1,2,3, ..) according to the value of the asset return scenarios.

4. Calculate the value of each exposure in each scenario as above: in the case of migration through zero curves, in the case of default through recovery rates.

5. Summarise the results: given all the scenarios of possible future portfolio values, calculate meaningful risk estimates for the distribution of loss, such as mean loss, volatility (standard deviation), percentile level (VaR).

A realistic implementation is described in the following Case Study chapter.

2.4 Mixture Models

In the mixture model approach, it is assumed that the default risk of an obligor relies on a set of common factors, such as macroeconomic variables, which are modeled stochastically. Once the factors are realized, defaults of individual firms are independent. Dependence occurs due to the individual default probabilities' reliance on these common factors.

2.4.1 Poisson mixture models

As in our case study we will be focusing on implementing the industry model CreditRisk+, let us introduce Poisson mixture models as in [12].

Given some $p < m$ and a p-dimensional random vector $\Psi = (\Psi, ..., \Psi)$, the random vector $\tilde{Y} = (\tilde{Y}_1, ..., \tilde{Y}_m)'$ follows a Poisson mixture model with factors Ψ if there are functions $\lambda_i : R^p \longrightarrow (0, \infty), 1 \leq i \leq m$, such that conditional on $\Psi = \psi$ the random vector $\tilde{Y}$ is a vector of independent Poisson distributed random variables with rate parameter $\lambda_i(\psi)$. We define the $\tilde{M} = \sum_{i=1}^{m} \tilde{Y}_i$ and observe that, for small Poisson parameters λ_i, $\tilde{M}$ is approximately equal to the number of defaulting companies. Given the factors, it is the sum of conditionally independent Poisson variables and therefore its distribution satisfies:

$$P(\tilde{M} = k|\Psi = \psi) = \exp\left(-\sum_{i=1}^{m} \lambda_i(\psi)\right) \frac{\left(\sum_{i=1}^{m} \lambda_i(\psi)\right)^k}{k!}. \tag{2.14}$$

In practice, in the majority of cases it is more common to use one factor models, as they are more tractable.

2.4.2 Probability generating functions and convolution

In order to derive the loss distributions arising from the CreditRisk + framework, we need to introduce the concepts of probability generating functions and convolution, as in [13]. Let us consider a discrete random variable X, and it's the discrete probability distribution $f_x = [f_0, f_1, f_2, ..., f_R]$. It's probability generating function is defined as:

$$P_X(t) = f_0 t^0 + f_1 t^1 + f_2 t^2 + f_3 t^3 + ... + f_R t^R, \tag{2.15}$$

in a more compact writing:

$$P_X^{(n)} = E[t^X] = \sum_{n=0}^{\infty} t^n P[X = n], \tag{2.16}$$

which also corresponds to the expected value of t^X. In the context of credit risk, the variable X could can signify:

- The number of obligors and the probability that they default or not default during a given period of time,
- The exposure in the obligors and the probability that they default or not default during a given period of time.

Now let's consider an example of two obligors, where N and K are independent discrete random variables expressed on non-negative integers. Their sum, symbolized as $J = N + K$, has a probability distribution constituting the convolution of the probability distributions of N and K. This can be mathematically represented as:

$$Pr\{J = j\} = \sum_{n=0}^{j} Pr\{N = n\}Pr\{K = j - n\}, \quad \forall\, j = 0,1,2, ..., n. \tag{2.17}$$

From the perspective of probability generating functions (PGF), the PGF of the sum (N + K) is a product of the PGFs of N and K due to their independence. This is expressed as:

$$P_{N+K}(t) = E[t^{N+K}] = E[t^N t^K] = E[t^N]E[t^K] = P_N(t)P_K(t). \tag{2.18}$$

Again, this can be used both to express discrete probabilities tied to portfolio defaults, and losses.

2.4.3 Convolution by Fast Fourier Transform

An alternative approach to compute the loss distribution, which has the advantage of being more computationally efficient a more stable, is based on the Fast Fourier Transform [13].

The Fourier Transform of a variable X, or Characteristic Function, can be defined as:

$$\phi_X(t) = E[e^{itX}] = P_X(e^{it}), \tag{2.19}$$

where i denotes the imaginary unit with i = sqrt-1. In the multivariate case, the joint characteristic function $(X_1, X_2, ..., X_k)$ is defined as:

$$\phi_{X_1,...,X_k}(t_1, ..., t_k) = E[e^{i(t_1X_1,...,t_kX_k)}] = P_{X_1,...X_k}(e^{it_1}, ..., e^{it_k}). \tag{2.20}$$

More in general:

$$\phi(t) = \int_{-\infty}^{\infty} f(x)e^{itx}dx \tag{2.21}$$

The characteristic function has a key property: for two independent variables, N and K, the characteristic function of their sum (N + K) equals the product of their individual characteristic functions. Leveraging this association, one can perform convolutions using the Fast Fourier Transform algorithm (FFT). Thus, mathematically:

$$\phi_{N+K}(t) = E[e^{it(N+K)}] = E[e^{itN}e^{itK}] = E[e^{itN}]E[e^{itK}] = \phi_N(t)\phi_K(t) \tag{2.22}$$

This is possible as a result of the independence of N and K.

The FFT of the sum of two independent discrete random variables is the product of the FFTs of the individual variables, given that adequate zeros are appended to each probability vector. The FFT, mapping n points to n correlatively, mandates input and output vectors to maintain the same length.

Generally, a longer vector is needed to adequately represent the sum variable compared to the components, as the sum variable tends to embrace larger values with non-zero probability. If not catered for, the tail probabilities for the sum will loop and reveal itself at the vector's commencement, necessitating the addition of adequate zeros to each probability vector's right end.

FFT Algorithm

Given two probability vectors $f = [f_0, f_1, ..., f_{m-1}]$ and $g = [g_0, g_1, ..., g_{k-1}]$, their convolution via Fast Fourier Transform can be computed through the following steps [13]:

1. Pad both vectors f and g with zeros such that each has a length of n $\geq$ m + k.
2. Apply FFT to each of the vectors: $\tilde{f}$ = FFT(f) and $\tilde{g}$ = FFT(g).
3. Compute the product (complex number multiplication), element by element, of the two vectors: $\tilde{h} = \tilde{f}.\tilde{g}$.
4. Apply the inverse function of the FFT (IFFT) to $\tilde{h}$ to retrieve the probability vector as a convolution of f and g.

Algorithm 1: Fast Fourier Transform algorithm for convolution

2.4.4 CreditRisk+

CreditRisk+, developed by Credit Suisse Financial Products (CSFP) in 1997, is an industry example of the Poisson mixture model introduced above. It neither harnesses traditional market data nor employs transition matrices or other standard inputs for models like CreditMetrics. The model asserts that defaults are predominantly random events, and doesn't use an underlying cause for default events, i.e. does not depend on firm´s fundamentals but is modelled as an exogenous variable, which follows a Poisson distribution with stochastic intensity parameter [8].

CreditRisk+ operates as a Poisson mixture model, where the stochastic parameter $\lambda_i(\Psi)$ of the conditional Poisson distribution for firm i equals [12]:

$$\lambda_i(\Psi) = k_i w_i \Psi, \tag{2.23}$$

with a constant $k_i > 0$, non-negative factor weights $w_i = (w_{i_1}, ..., w_{i_p})$ adhering to $\sum_j w_{ij} = 1$, and p-independent Gamma Ga(α_j, β_j)-distributed factors $\Psi_1, ..., \Psi_p$.

The model is interesting, and differs from other industry frameworks, as it allows to build a closed form for the distribution of losses, yielding fast and accurate results, also in the case of very large portfolios. Yet, a key trade-off of not employing simulation is the rudimentary environment it provides to incorporate correlations, which the model executes through a sector analysis. Once more, we note that the model presumes common correlation within a sector, not among individual debtors, these are in fact classified into sub-portfolios, and each sub- portfolio is affected by a specific economic factor.

Default and loss distribution with fixed default rates

Following the implementation described in the CreditRisk + technical document [17], let's consider having N independent debtors in a portfolio. Given independence, probabilities of default for each obligor are fixed, and considered input parameters usually taken from rating agencies. Let's portray the likelihood of debtor A defaulting as p_A and its probability generating function can be represented with an auxiliary variable z as:

$$F(z) = \sum_{n=0}^{\infty} P(n \text{ default}) z^n \tag{2.24}$$

It is straightforward to write the probability generating function for a single obligor explicitly:

$$F_A(z) = 1 - p_A + p_A z = 1 + p_A(z-1) \tag{2.25}$$

Since we are considering default events as independent, the portfolio PGF can be written as the product of individual PGFs, and the logarithm function can be applied to yield:

$$\begin{aligned} F(z) &= \prod_A F_A(z) = \prod_A (1 + p_A(z-1)) \\ log F(z) &= \sum_A log(1 + p_A(z-1)) \end{aligned} \tag{2.26}$$

In the limit, this becomes:

$$F(z) = e^{\sum_A p_A(z-1)} = e^{\mu(z-1)} \tag{2.27}$$

as $log(1 + p_A(z-1)) \approx p_A(z-1)$ when p_A are sufficiently small, and $\mu = \sum_A p_A$, representing the expected number of defaults in one year for the whole portfolio. It is possible to expand F(z) in its Taylor series:

$$F(z) = e^{\mu(z-1)} = e^{-\mu} e^{\mu z} = \sum_{n=0}^{\infty} \frac{e^{-\mu}\mu^n}{n!} z^n. \tag{2.28}$$

Yielding a probability of n defaults for the portfolio in one year equal to, thus following a Poisson distribution:

$$P(n \text{ default}) = \frac{e^{-\mu}\mu^n}{n!} \tag{2.29}$$

.

To lessen computational costs, CreditRisk+ uses exposure bands instead of actual exposures, meaning that levels of integer-valued amounts, known as bands, are introduced into the model. For each exposure, it is rounded to the nearest integer and then to the closest band. Each debtor's credit amount is then conveyed to the nearest band before the loss distribution from these bands is figured out. This roundabout way significantly reduces computational demands though it sacrifices a bit of accuracy. We can divide the portfolio into m exposure bands, symbolized by index j for $1 \leq j \leq m$. According to the original model, ν_j, ϵ_j, and μ_j collectively symbolize the usual exposure, predicted loss and predicted default number in exposure band j.

$$\epsilon_j = \nu_j \times \mu_j \tag{2.30}$$

Therefore, the forecasted loss in terms of default event probability can be stated as follows:

$$\mu_j = \frac{\epsilon_j}{\nu_j} = \sum_{A:\nu_A=\nu_j} \frac{\epsilon_a}{\nu_A} \tag{2.31}$$

Moreover, let μ symbolize the aggregate anticipated count of default incidents in a year, then

$$\mu = \sum_{j=1}^{m} \mu_j = \sum_{j=1}^{m} \frac{\epsilon_j}{\nu_j}. \tag{2.32}$$

After having calculated the distribution of defaults, we can proceed to calculate the distribution of losses due to default events. Let G(z) symbolize the probability generating function of default losses, expressed in multiples of the unit L of exposure:

$$G(z) = \sum_{n=0}^{\infty} P(\text{AggregateLosses} = n \times L) z^n. \tag{2.33}$$

Given that the exposures are presumed to be independent, the exposure bands also are independent. Hence,

$$G(z) = \prod_{i=1}^{m} G_i(z). \tag{2.34}$$

The following equation illustrates the probability generating function for the jth band,

$$G_j = \sum_{n=0}^{\infty} P(n \text{ defaults}) z^{n\nu_j} = \sum_{n=0}^{\infty} \frac{e^{-\mu_j}\mu_j^n}{n!} = e^{-\mu_j+\mu_j z^{\nu_j}}. \tag{2.35}$$

In simple terms, if there's an occurrence of n defaults in the jth band of the portfolio, the loss characteristic function may be explained by this formula. As a result:

$$G_{independent}(z) = \prod_{j=1}^{m} e^{-\mu_j+\mu_j z^{\nu_j}} = e^{-\sum_{j=1}^{m}\mu_j+\sum_{j=1}^{m}\mu_j z^{\nu_j}}. \tag{2.36}$$

The above formula represents the loss probability generating function for the portfolio relative to default losses with a fixed default rate.

Using the FFT algorithm described in the previous section (Algorithm 1), we can outline the overall portfolio's loss distribution for a Poisson mixture model employing the FFT:

1. Decide dimension n of the probability vector f such that it has appropriate length, as discussed above
2. Establish a probability vector for each band j as:

$$P(n\,defaults) = \frac{\mu^n e^{-\mu_j}}{n!}\,\forall\, n = 0,1,2,...,2^r \; and \; j = 1,2,...,m \tag{2.37}$$

3. Compute the portfolio's loss distribution vector via the following formula:

$$\text{IFFT}\left(\prod_{j=1}^{m} \text{FFT}(f_j)\right) \tag{2.38}$$

Algorithm 2: Calculation of the loss distribution using FFT algorithm

This implementation differs from Algorithm 1 as each obligor had a dedicated vector (i.e. f and g). Here (Algorithm 2), we use as many vectors as there are obligors in band j under the recognition that the default probability remains the same for all obligors within that band, denoted as μ_j.

Sector Analysis

As explained in [17], CreditRisk+ models correlation between obligors through the concept of defaults volatility, i.e. incorporating the effects of background factors into the specification of default rates by allowing the default rate μ_j itself to have a probability distribution. Differing from CreditMetrics, CreditRisk+ does not attempt to model correlations explicitly but captures the same concentration effects through the use of default rate volatilities and sector analysis. The premise is that issuers in each sector have some common specifics. The rationale underlying the sector analysis is the fact that default volatilities are driven by some systematic-risk factors, typical for every sector.

- The variability of these default probabilities can be tied to fluctuations in a limited number of background factors, such as economic conditions, influencing debtor outcomes. For instance, a declining economy could increase the likelihood of default for most debtors in a portfolio.
- However, while shifts in the economy or other factors might influence the likelihood,

they do not guarantee debtor defaults. Regardless of economic circumstances, actual defaults should still be regarded as relatively scarce events. Thus, the above analysis concerning rare events proves relevant when properly adjusted. This emphasized here is that uncertainty stems from factors that can have parallel effects on a large number of debtors.

Effects of including default rate volatility can be shown in Figure 2.8.

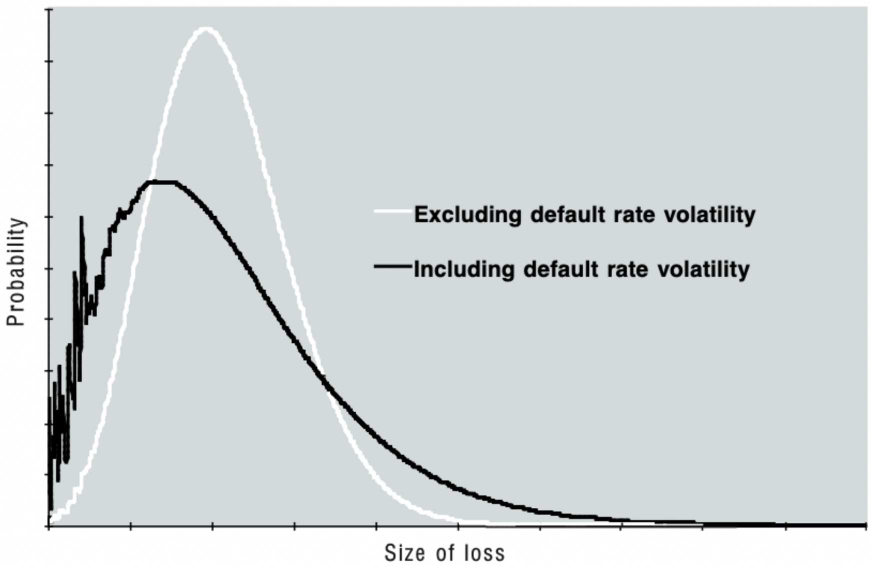


Figure 2.8. Loss distributions of a certain multiple of the chosen unit of exposure calculated without default rate volatility and with default rate volatility, taken from [17]

It is important to notice that, as expected, the two distributions have the same mean loss. However, by incorporating default volatility we allow for a much larger chance of experiencing extreme losses, evident visually through the fatter tail of the distribution. Mathematically, the model views each sector $S_k : 1 \leq k \leq n$ as driven by a singe underlying factor, a random variable x_k described by the mean, μ_k, and its standard deviation, σ_k. For each sector, x_k represents the average default rate. In CreditRisk+, the key assumption is that x_k has Gamma distribution $X_k \sim \Gamma(\alpha_k, \beta_k)$, where:

$$\mu_k = \alpha_k \beta_k \text{ and } \beta_k = \sigma_k^2 / \mu_k.$$

For sector k, we can write the PGF of default events, conditional on $x_k = x$

$$F_k(z)[x_k = x] = e^{x(z-1)} \tag{2.39}$$

Given that x_k has a certain gamma distributed density function $f_k(x)$, the PGF for sector k is the average of the conditional PGFs over all possible values of the mean default rate:

$$F_k(z) = \sum_{n=0}^{\infty} P(\text{n defaults})z^n = \sum_{n=0}^{\infty} z^n \int_{x=0}^{\infty} P(\text{n defaults}|x)f(x)dx = \int_{x=0}^{\infty} e^{x(z-1)} f(x)d(x). \tag{2.40}$$

Plugging in $\Gamma(\alpha_k, \beta_k)$ for $f(x)_k$:

$$F_k(z) = \int_{x=0}^{\infty} e^{x(z-1)} \frac{e^{-x7\beta} x^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_{y=0}^{\infty} \left(\frac{y}{\beta^{-1}+1-z} \right)^{\alpha-1} e^{-y} \frac{dy}{\beta^{-1}+1-z} = \frac{\Gamma(\alpha)}{\beta^\alpha \Gamma(\alpha)(1+\beta^{-1}-z)^\alpha} = \frac{1}{\beta^\alpha (1+\beta^{-1}-z)^\alpha}. \tag{2.41}$$

As $p_k = \frac{\beta_k}{1+\beta_k}$, the PGF for sector k can be written and expanded in its Taylor series:

$$F_k(z) = \left(\frac{1-p_k}{1-p_k z} \right)^{\alpha_k} = (1-p_k)^{\alpha_k} \sum_{n=1}^{\infty} \binom{n+\alpha_k-1}{n} p_k^n z^n. \tag{2.42}$$

The probability of a number n of defaults in each sector is thus:

$$P(\text{n defaults}) = (1-p_k)^{\alpha_k} \binom{n+\alpha_k-1}{n} p_k^n. \tag{2.43}$$

Considering that there is independence between sections, the probability generating function for default events of the portfolio can be written:

$$F(z) = \prod_{k=1}^{n} F_k(z) = \prod_{k=1}^{n} \left(\frac{1-p_k}{1-p_k z} \right)^{\alpha_k}. \tag{2.44}$$

While, after some calculations, the probability generating function for losses has the closed form expression:

$$G_{dependent}(z) = \prod_{k=1}^{n} G_k(z) = \prod_{k=1}^{n} \left(\frac{1-p_k}{1-\frac{p_k}{\mu_k}\sum_{j=1}^{m(k)} \frac{\epsilon_j^{(k)}}{v_j^{(k)}} z^{v_j^{(k)}}} \right)^{\alpha_k}. \tag{2.45}$$

The above formulas are valid for a portfolio that is divided into mutually independent classes, each encompassing a set of obligors, driven by one factor.

In general, it is possible to consider a situation where obligors cannot be grouped in independent classes, but are affected in different ways by systematic factors. This concept falls under the name of General Sector Analysis, which will not be expanded in this thesis work.

Chapter 3

Case Study

CreditRisk+ and CreditMetrics represent two of the most influential benchmark models for studying credit risk, however, direct comparison is not straightforward. While the two serve essentially the same purpose, they are constructed quite differently regarding their mathematical framework, as studied in the previous chapters.
Difficulties in the comparison derive essentially from the different distributional assumptions, different techniques for calibration and CreditMetrics allowing a richer model (mark-to-market approach). Thus, given the same portfolio of credit exposures, the models in general yield different evaluations of credit risk.

Literature regarding direct comparison of the two credit risk models is not very vast, and is mainly concerned with analysing the default component of the portfolio credit risk (the distribution of defaults), thus not considering credit migrations. Previous papers find out that there is a symmetry of factor transformations that allows to consistently parametrise CreditRisk+ and CreditMetrics in the case of portfolios with same credit rating, affected by one risk factor. In particular:

- in Koyluoglu and Hickman [11] theoretical similarities between CreditMetrics and CreditRisk+ are analysed assuming a simplified framework of homogeneous portfolio and Vasicek representation of asset returns. The study focuses on the default distribution rather than the loss one.

- similary, Finger [9] compares CreditRisk+ to restricted CrediMetrics for homogenous portoflios driven by a single economic factor. He arrives to the result that extreme tails of the default distributions generated by the models are very different.

- Diaz and Gemmil [8] extend Koyluoglu and Hickman [11] by developing a three state model for CreditMetrics and comparing the loss distribution between the former and CreditRisk+. They use a portfolio formed by N bonds equally rated, with the same exposure size, the same time to maturity, and affected by a single economic factor.

In the case of restricted CreditMetrics (considering only default), the difficulty in the comparison between the models lies in parametrising the inputs in a constituent way:

- mapping CreditMetrics correlation structure into adequate volatilities for in CreditRisk+. In order to make models comparable, we will introduce a transformation procedure developed by Koyluoglu and Hickman [11];
- recovery rates: as CreditRisk+ does not analyze ”which obligor” defaulted, it is impossible to map the number of defaulted obligors to the corresponding recovery rates of each obligor. Thus, instead of principals, directly LGD is taken as input of the model.

Even more difficulties arise when trying to map CreditMetrics' mark to market implementation, which considers losses arising from default migrations and coupons. The comparison makes sense from a business and solvency perspective, but is hardly mathematically justifiable.

First, the aim of this case study is to show how the CreditMetrics and CreditRisk+ frameworks are implemented in practice, and how different implementation methods and assumptions yield more or less similar results. Secondly, we aim to compare the models in the case of non homogeneous portfolios and different sectors, extending the work done in previous studies through Monte Carlo simulation.

We start through simple sector analysis with no volatilities in default rates (which maps to a diagonal correlation matrix in the CreditMetrics framework). We will compare results given by CreditMetrics in its restricted version (same approach as previous research papers), in its mark-to-market extension, and CreditRisk+, comparing different implementation methods described in technical documents and further papers, including both an analytical approach (based on the probability generation function), its extension through Fast Fourier Transform, and Monte Carlo simulation.

Then, a more advanced approach will follow where the plain fixed default model will be upgraded for default volatilities, accounting for multiple independent sectors.

Both the default distribution and the loss distribution are presented, given different modelling assumptions. Pseudo code will be provided in order to bridge the gap between theory and implementation.

3.1 Portfolio selection

All models are run with a constructed high quality and a low quality portfolio, where all input assumptions are taken from reference literature, in order to make the examples realistic. To construct a distribution of initial credit ratings, data is taken from Gordy [10].

Other necessary inputs are:

1. **probability of default and migration**: neither CreditMetrics not CreditRisk+

Table 3.1. Typical bank portfolio characteristics

Rating	AAA	AA	A	BBB	BB	B	CCC	$\mu(sumPD)$
High quality	4	6	29	36	21	3	1	3.2109
Low quality	1	1	4	15	40	34	5	0.6587

are methodologies to estimate probabilities of default (PDs). Instead, these probabilities, together with migration probabilities, are important input parameters, usually obtained from one of the major rating agencies, which publish updated migration matrices frequently. In our case from the Creditmetrics technical document [14], shown in Table 3.2.

Initial	Rating at year-end (%)							
Rating	AAA	AA	A	BBB	BB	B	CCC	Default
AAA	90.81	8.33	0.68	0.06	0.12	0	0	0
AA	0.70	90.65	7.79	0.64	0.06	0.14	0.02	0
A	0.09	2.27	91.05	5.52	0.74	0.26	0.01	0.06
BBB	0.02	0.33	5.95	86.93	5.30	1.17	0.12	0.18
BB	0.03	0.14	0.67	7.73	80.53	8.84	1.00	1.06
B	0	0.11	0.24	0.43	6.48	83.46	4.07	5.20
CCC	0.22	0	0.22	1.30	2.38	11.24	64.86	19.79

Table 3.2. Transition matrix for credit rating migration

2. **principals**: to let the case study be as general as possible, no assumptions are made on principals, which are chosen randomly between 2M and 15M from each debtor. The total value of the portfolio is 831.64M.

3. **recovery rates**: for simplicity, recovery rates are fixed to 20%, although in a real-life portfolio more case specific assumptions can be very easily incorporated in the models (es. historical recovery rates per industry, recovery rates per bond seniority, etc).

4. **correlation**: first we will assume no correlation to compare CreditMetrics with CreditRisk+ in the most pure way possible, then we will develop a framework to map CM's correlation structure into CR+'s default volatilities. Correlations for the different sectors will be provided in the section below.

3.2 CreditMetrics

CreditMetrics' valuation of the loss distribution due to default is assessed through Monte Carlo simulation. Thanks to simulation, it is possible for financial institutions to incorporate complex portfolio structures and foresee to what degree of exposure the risk is bearable. The implementation follows the directives given in the CreditMetrics technical document [14].
We will focus on two kinds of simulation: default loss simulation and mark-to-market simulation. Even though CreditMetrics is a mark-to-market model, default loss is a very popular topic among regulators and studies that are based on statistical modeling of credit risk, in fact allowing a more consistent comparison with further implementations of CreditRisk+. This same approach is used in Finger [9], Koyluoglu and Hickman [11].

3.2.1 CreditMetrics default loss simulation

The first model we implement is the simplest possible, and the most immediate to directly compare to our following implementations of CreditRisk+. In this restricted version of CreditMetrics, the distribution of losses is driven solely by defaults.

Inputs of the model are:

1. Ratings of obligors: given in Table 3.1 for both the high quality and low quality portfolio;
2. Exposures: uniformly generated between 2M and 15M;
3. Transition Matrix: taken from rating agencies, shown in Table 3.2.
4. Recovery rates: fixed for simplicity at 20%;
5. Correlation Matrix: diagonal, thus assuming independence between borrowers.

Given the inputs, calculation of the loss and default distribution is very simple in Matlab, and detailed in Algorithm 3. In Matlab environment, 100,000 repetitions were carried out for both portfolios, also calculating appropriate risk measures. In Figures 3.1 and 3.2 default and loss distributions for the portfolios, which we can notice to be very different, and a visual indication of the VaR95, VaR99, VaR999.

First of all we notice that the distribution of default is coherent with the total default probabilities, having means in 3.21 and 0.66 (low and high quality portfolio respectively). We also see that although the exposure of the portfolios is the same, the different quality composition yields a very different loss distribution, with much higher risk measures for the low quality portfolio. A more detailed analysis will be carried in the sections below.

Data: Principal, Recovery, default_probability, n_firms, n_scenarios
Result: Loss and default distribution
for *each firm i from 1 to* n_firms **do**
    Calculate the default threshold as norminv(default_probability(i));
end
Generate $n\_scenarios$ for each firm using mvnrnd;
for *each scenario s* **do**
    for *each firm* **do**
        if *scenario(s, firm) > default threshold for the firm* **then**
            loss(s, firm) = Principal(firm) - Recovery(firm);
            default_count(s, firm) = 1;
        end
    end
end

Algorithm 3: Loss and Default Distribution CreditMetrics restricted

Figure 3.1. Restricted CreditMetrics - Low quality portfolio

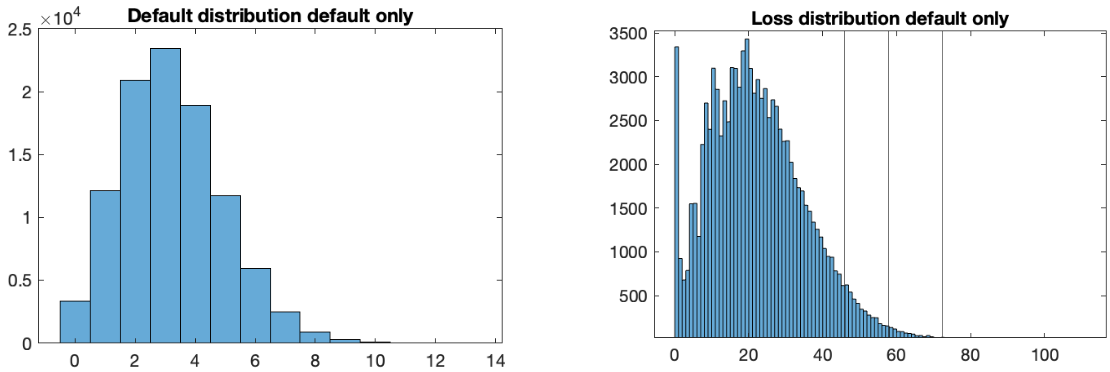
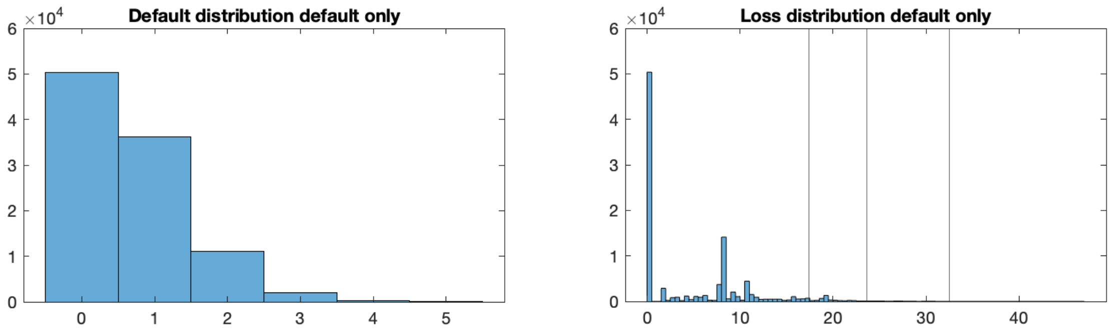


Figure 3.2. Restricted CreditMetrics - High quality portfolio



3.2.2 Mark to Market CreditMetrics

Mark-to-market analysis is a critical technique for evaluating rates or spreads of debtors belonging to various rating classes. This method differs from default loss simulation since it allows us to incorporate the loss attributable to alterations in the creditworthiness of debtors into the yearly portfolio value calculations, thus determining whether the implemented spreads genuinely address the associated risk. In this section, we will clarify how CreditMetrics can be utilized to carry out a mark-to-market simulation, and compare distributions and risk measures to previous implementation. In order to move forward, additional input parameters are needed:

- **coupon rates (annual)** are chosen arbitrarily (Table 3.3) to yield a sensible result with the given spread values. More sophisticated input coupon values could have been chosen, but do not add to the purpose of this analysis. In a real portfolio, this input can be easily derived in a real portfolio in the documentation of the each bond;

Table 3.3. Table of coupon values by rating category

Rating Category	Coupon Value
AAA	2.00
AA	2.75
A	4.25
BBB	4.75
BB	5.25
B	5.75
CCC	6.25

- **maturities**: generated randomly between 1 and 5 years;
- **yields/spreads**: are taken from Atta Mills [3], and shown in Table 3.4.

Table 3.4. One-year forward zero curve for each credit rating category (%)

	Year 1	Year 2	Year 3	Year 4
AAA	3.6	4.17	4.73	5.12
AA	3.65	4.22	4.78	5.17
A	3.72	4.32	4.93	5.32
BBB	4.1	4.67	5.25	5.63
BB	5.55	6.02	6.78	7.27
B	6.05	7.02	8.03	8.52
CCC	15.05	15.02	14.03	13.52

In order to implement CreditMetrics' mark-to-market framework, we can split the simulation process in the following steps [14]:

1. **Generate scenarios**. Each scenario corresponds to a possible "state of the world" at the end of our risk horizon. For our purposes, the "state of the world" is just the credit rating of each of the obligors in our portfolio.

2. **Value portfolio**. For each scenario, we revalue the portfolio to reflect the new credit ratings. This step gives us a large number of possible future portfolio values.

3. **Summarize results**. Given the value scenarios generated in the previous steps, we have an estimate for the distribution of portfolio values. We may then choose to report any number of descriptive statistics for this distribution.

1. Scenario Generation

We are interested in modelling the portfolio value at time T=1 year. In order to do so, we will simulate future states of the world. The steps to scenario generation are as follows:

- Establish asset return thresholds, as in Equation 2.12, for the obligors in the portfolio: we calculate thresholds such that we can map obligors to a certain credit-worthiness given the initial rating, as in Matlab Algorithm 4. Inputs are an object portfolio_data, a vector containing the number of obligors for each rating category, states_rating, a vector containing the rating category states, the transition matrix as detailed in Table 3.2, and the number of firms.
  A Matlab implementation is shown in Algorithm 4, resulting in the asset return thresholds in Table 3.5.

```
Data: portfolio_data, states_rating, Transition_matrix, n_firms
Result: thresholds
initial_states = constructThresholds(portfolio_data, states_rating);
Function construct_initialstates(portfolio_data, states_rating)
    num_firms = sum(portfolio_data);
    initial_states = repelem(states_rating, portfolio_data);
    return initial_states;
Initialize Thresholds matrix as ones(7, n_firms);
for i = 1 to n_firms do
    Compute cumulativeTransition =
     cumsum(Transition_matrix(initial_states(i), 1:7));
    for j = 1 to 7 do
        thresholds(j, i) = norminv(1 - cumulativeTransition(j));
    end
end
```

Algorithm 4: Thresholds Calculation with Initial States

Table 3.5. Threshold Kernel

	AAA	AA	A	BBB	BB	B	CCC
AA	-1.33	2.46	3.12	3.54	3.43		2.85
A	-2.38	-1.36	1.98	2.70	2.93	3.06	2.85
BBB	-2.91	-2.38	-1.51	1.53	2.39	2.70	2.62
BB	-3.04	-2.85	-2.30	-1.49	1.37	2.42	2.11
B		-2.95	-2.72	-2.18	-1.23	1.46	1.74
CCC		-3.54	-3.19	-2.75	-2.04	-1.32	1.02
Default			-3.24	-2.91	-2.30	-1.62	-0.85

- Generate scenarios of asset returns according to the normal distribution using random numbers generator.

- Map the asset return scenarios to credit rating scenarios as in Algorithm 5: for example, for a firm initially rated AAA, all scenarios above -1.33 imply the firm stays in its initial rating category $state = 1$, all values between -1.33 and -2.38 imply a downgrade to rate AA, so $state = 2$ etc. Some migrations have probability 0.

```
Data: n_scenarios, n_firms, thresholds, mu, M
Result: states
scenarios = mvnrnd(mu, M, n_scenarios);
Initialize states matrix as zeros(n_scenarios, n_firms);
for s = 1 to n_scenarios do
    for f = 1 to n_firms do
        for j = 1 to 7 do
            if thresholds(j + 1, f) ≤ scenarios(s, f) ∧
             scenarios(s, f) ≤ thresholds(j, f) then
                states(s, f) = j;
                break;
            end
        end
        if scenarios(s, f) < thresholds(7, f) then
            states(s, f) = 8;
        end
    end
end
```

Algorithm 5: Mapping Asset Return Scenarios to Credit Rating Scenarios

2. Portfolio valuation

With respect to the default simulation, we need to calculate the values of the bonds in all possible states of the world, and then assigning that value for the state emerged by the simulation. This can be done performing a straightforward present value bond revaluation using the zero curves in input.

$$P_{bond} = P_{coupons} + P_{facevalue} = \sum \frac{C}{1+r}^{t} + \frac{F}{1+r}^{T} \tag{3.1}$$

In Matlab this can be implemented as in Algorithm 6.

Data: n_firms, rating_cat, maturities, forward_curves, coupons
Result: values
for $i = 1$ *to* n_firms **do**
  for $r = 1$ *to* $rating\_cat - 1$ **do**
    values(r, i) = bond_value(maturities(i), forward_curves(r,:), coupons(i));
  end
end

Function *bond_value(maturity, forward_curves, coupon)*
  Initialize value as coupon;
  for $i = 1$ *to maturity* **do**
    if $i == 1$ **then**
      value = coupon;
    end
    else if $i == maturity$ **then**
      value = value + (100 + coupon)/(1 + forward_curves(i))$^{(i-1)}$ **end**
      else
        value = value + (coupon)/(1 + forward_curves(i))$^{(i-1)}$ **end**
      end
      return *value*;

Algorithm 6: Bond Value Calculation

In order to calculate the loss, if the bond didn't default we take the difference between the value it would have in its initial state, minus the value at final state. If the bond defaulted, the loss will be the loss given default.
Finally, percentage value changes are multiplied by the individual exposures. The whole process is implemented in Matlab as in Algorithm 7.

```
Data: initial_states, states, values, n_scenarios, n_firms
Result: loss, final_values, default_count
for i = 1 to n_scenarios do
    for j = 1 to n_firms do
        loss(i, j) = values(initial_states(j), j) - values(states(i, j), j);
        final_values(i, j) = values(states(i, j), j);
        if states(i, j) == 8 then
            default_count(i, j) = 1;
        end
    end
end
```

Algorithm 7: Loss and default Distribution CreditMetrics Mark to Market

3. Results

We now compare the results obtained with the Mark to Market implementation of CreditMetrics, and the previous default only implementation.
While the default distribution if the same for the two frameworks, the default one varies significantly, as a result of the incorporation of the migration component of risk.

Low quality portfolio

It is possible to notice that:

- the loss distribution doesn't have a peak on the value 0 (which indicated no defaults in the portfolio), and accounts for a smoother structure without multiple peaks;
- negative values for losses are possible, accounting for positive gains due to favorable credit migrations;
- the loss distribution has a slightly longer tail, reflected in a higher values for VaR measures, and consequently higher estimations of the risk of losses;
- credit migrations do not impact the distribution of defaults.

	Restricted CreditMetrics	Mark to Market CreditMetrics
Mean Loss	22.59	22.99
VaR 99%	57.88	58.84
VaR 99.9%	71.6	73.45
St. deviation	12.07	13.07
Skewness	0.57	0.57
Kurtosis	3.24	3.24

Table 3.6. Descriptive metrics for the loss distribution of the low quality portfolio

Figure 3.3. Mark to Market CM - Low quality portfolio

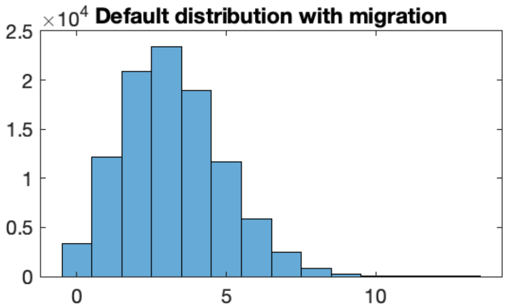


Figure 3.4. Distribution of defaults

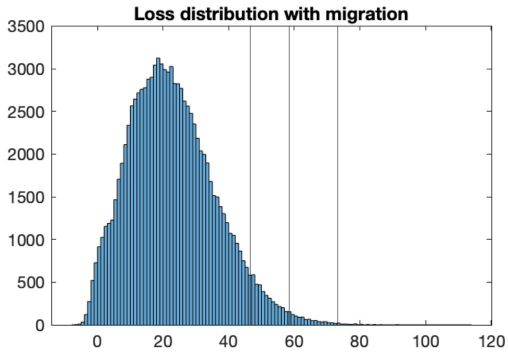


Figure 3.5. Distribution of losses

Figure 3.6. Restricted CM - Low quality portfolio

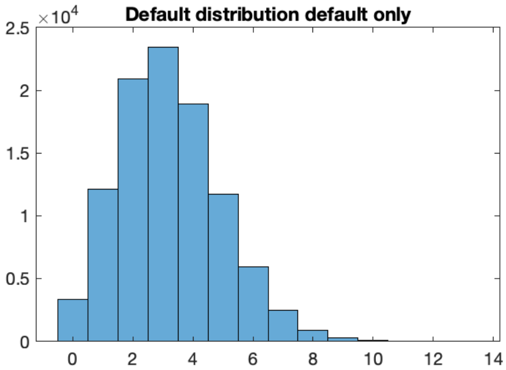


Figure 3.7. Distribution of defaults

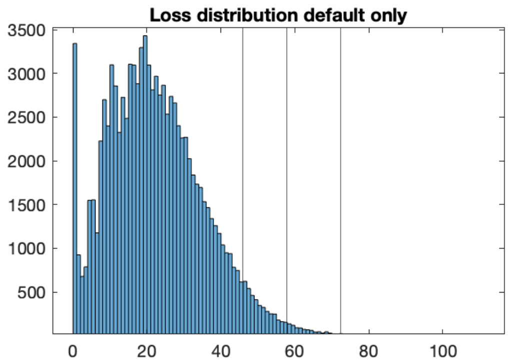


Figure 3.8. Distribution of losses

High quality portfolio

In this case, the loss distribution assumes a very different appearance using the two models:

- it is possible to notice from the default distribution that in about half of runs, no firm defaults. This is evident in the loss distribution for the default only implementation which results highly concentrated on a zero loss;
- in the mark-to-market approach, the distribution results to be much smoother, less concentrated on null values due to migration, and with a clear bi-modal structure;
- as before, positive values are possible do to favourable credit migrations.

	Restricted CreditMetrics	Mark to Market CreditMetrics
Mean Loss	4.90	5.44
VaR 99%	23.71	24.95
VaR 99.9%	32.24	33.78
St. deviation	6.16	6.25
Skewness	1.25	1.24
Kurtosis	4.45	4.56

Table 3.7. Descriptive metrics for the loss distribution of the high quality portfolio

Figure 3.9. Mark to Market CM - High quality portfolio

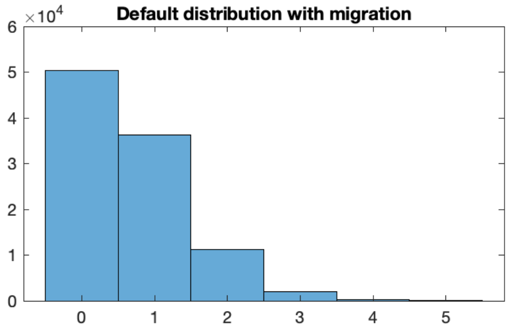


Figure 3.10. Distribution of defaults

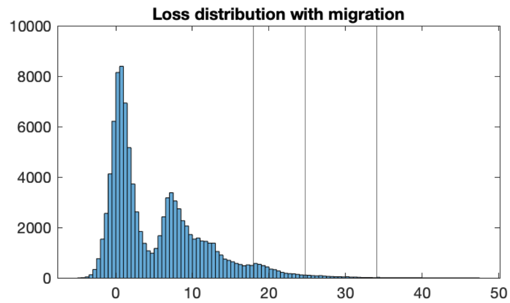


Figure 3.11. Distribution of losses

Figure 3.12. Restricted CM - High quality portfolio

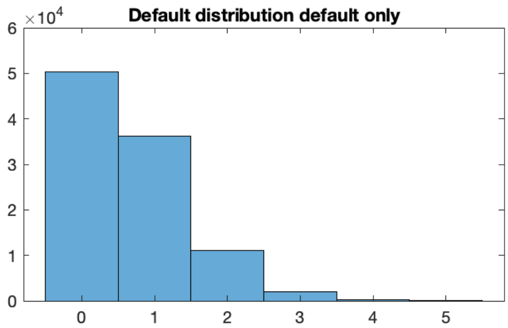


Figure 3.13. Distribution of defaults

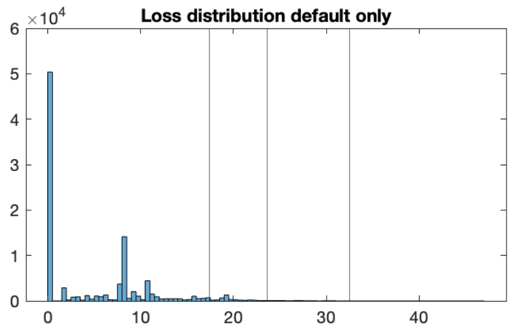


Figure 3.14. Distribution of losses

3.3 CreditRisk+

In this section we are going to apply model CreditRisk+, as described in the Credit Suisse technical document [17] using fixed default rates (thus accounting for independence between borrowers).
CreditRisk+ has the distinguishing features that it can avoid Monte Carlo simulation, using instead a numerical algorithm to calculate aggregate risk, and that it was published as an open methodology document, with the intention of encouraging financial institutions and risk practitioners to develop their own implementations of the model. Since its introduction, CreditRisk+ has consistently attracted the interest of practitioners, financial regulators, and academics, resulting in a significant body of literature on the model, including many enhancements and alternative approaches [20]. We will compare different implementations in order to measure eventual differences and computational efficiency.

3.3.1 Data preparation

What we need to run CreditRisk+, at least in the basic version without default volatility, are probabilities of default and the exposures.

- To follow our consistency-preservation principle, we use the same default probabilities as those used in CreditMetrics.
- It is also necessary to adjust exposures, as in banding approaches the model does not allow to specify which obligor defaulted, thus not making it possible to model recovery rates to defaults. The data is adjused before hand so that the exposures only count for the risk part (only the loss given default). This is done by:

$$exposure = principals - principals * recoveryrates.$$

Additionally, further difficulties arise when comparing CreditMetrics' mark-to-market version, as CreditRisk+ does not incorporate coupons.

3.3.2 Naif Monte Carlo

In the most Naif implementation possible of CreditRisk+ through simulation, the model takes in input:

- default probabilities
- exposures

The model works as shown in Figure 3.15, where the event of default is simulated according to a Poisson distribution with parameter equal to the default probability, and the severity of the losses is given by the exposure, minus the recovery rate. This can be easily implemented as in Algorithm 8.

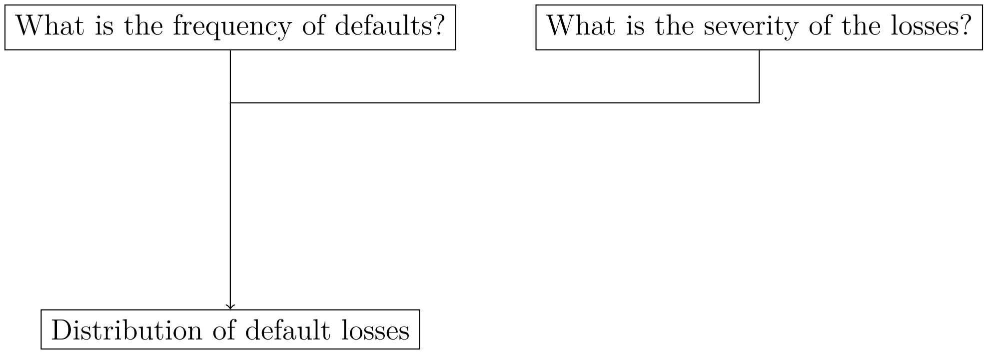


Figure 3.15. CreditRisk+ framework

```
Data: n_scenarios, n_firms, pd, exposure
Result: loss_distribution, def_distribution
for i = 1 to n_scenarios do
    simulated_losses = 0;
    defaults = 0;
    for m = 1 to n_firms do
        num_defaults = poissrnd(pd(m));
        if num_defaults > 0 then
            def_count = def_count + num_defaults;
        end
        num_defaults = min(num_defaults, 1);
        simulated_losses = simulated_losses + exposure(m) * num_defaults;
        defaults = defaults + num_defaults;
    end
    loss_distribution(i) = simulated_losses;
    def_distribution(i) = defaults;
end
```

Algorithm 8: Loss and default distribution with Naif Monte Carlo for CreditRisk+

Resulting in the distributions shown in Figures 3.16 and 3.19 for defaults and losses given 50.000 repetitions.

Figure 3.16. Naif CR+ implementation - Low quality portfolio

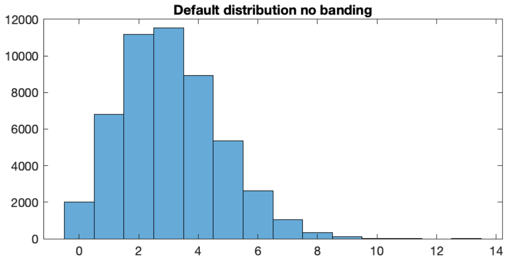


Figure 3.17. Distribution of defaults

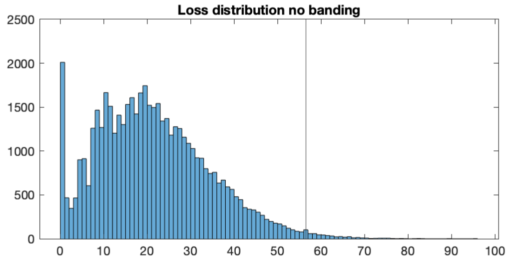


Figure 3.18. Distribution of losses

Figure 3.19. Naif CR+ implementation - High quality portfolio

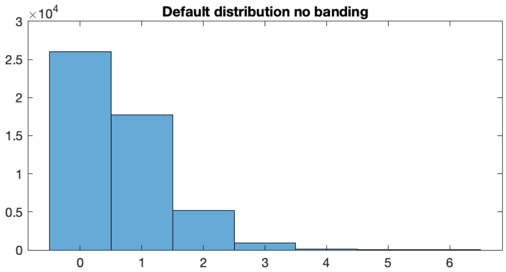


Figure 3.20. Distribution of defaults

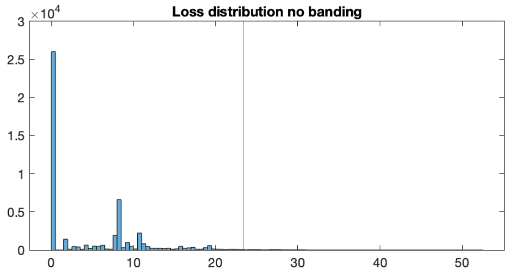


Figure 3.21. Distribution of losses

3.3.3 Monte Carlo with Banding

In the official CreditRisk+ documents, in order to reduce the amount of data to be processed for large portfolios, the exposures, net of the above recovery adjustment, are divided into bands of exposure with the level of exposure in each band being approximated by a common average [17].
The first step in obtaining the distribution of losses from the portfolio in an amenable form is to group the exposures of the portfolio into bands. This has the effect of significantly reducing the amount of data that must be incorporated into the calculation.
While banding reduces the computational complexity, it also introduces an approximation into the calculation. A possible workaround would be choosing a width of the bands that is small compared with the average exposure size characteristic of the portfolio, the approximation is insignificant. This however cannot be applied: suppose we choose the size of band as $B = 0.05M$, which wouldn't mean severe rounding error from actual numbers, we would get a very large numbers of bands m (in fact much higher than the number of debtors in our portfolio) so that the band approach would lose its purpose as there would be no grouping. Therefore, we set $B = 1M$, which can cause a non negligible error.

The portfolio can be now divided into m exposure bands represented by the index j where $1 \leq j \leq m$. We adopt the following notation for each band j:

- ν_j common exposure of the band
- ϵ_j expected loss of each band
- μ_j expected number of defaults of each band

We can then write $\epsilon_j = \nu_j \times \mu_j$, $\mu_j = \frac{\epsilon_j}{\nu_j} = \sum_{A:\nu_A=\nu_j} \frac{\epsilon_A}{\nu_A}$.
Having chosen $B = 1M$, we round the portfolio according to ν_j, and extract the unique exposures. We can follow Algorithm 9 in Matlab to generate exposure bands:

Data: exposure, B, pd
Result: bands
for $i = 1$ *to length(unique_exposures)* **do**
  obligors_in_band = (portfolio(:, 1) == unique_exposures(i));
  bands(i, 1) =unique_exposures(i);
  bands(i, 2) = sum(portfolio(obligors_in_band, 2));
  bands(i, 3) = sum(portfolio(obligors_in_band, 1) · portfolio(obligors_in_band, 2));
  bands(i, 4) = sum(obligors_in_band);
end
bands = sortrows(bands);

Algorithm 9: Exposure Bands Calculation

For the low quality credit portfolio the output for the exposure bands is the one shown in table 3.8 (we ave a similar result for the high quality portfolio):

Table 3.8. Bands in low quality portfolio

ν_j	μ_j	ϵ_j	# of obligors
2.00	0.31	0.61	9.00
3.00	0.27	0.82	13.00
4.00	0.28	1.11	7.00
5.00	0.16	0.80	9.00
6.00	0.34	2.03	12.00
7.00	0.13	0.90	6.00
8.00	0.68	5.42	9.00
9.00	0.34	3.04	15.00
10.00	0.08	0.75	4.00
11.00	0.55	6.10	12.00
12.00	0.08	1.01	4.00

We notice $\sum \mu_j = 3.2109$ ($\sum \mu_j = 0.6576$ in the high quality portfolio), coherently with the original model.

The loss distribution is calculated as before (Algorithm 8), but instead of simulating if each firm defaults or not:

1. first we simulate the number of defaults in each band as a random Poisson with parameter μ_j
2. the loss is calculated multiplying the number of defaults times the exposure in each band ν_j
3. this is repeated for all bands.

Results are consistent with the previous implementation, and can be seen in Figures 3.22 and 3.25.

Figure 3.22. CR+ with banding implementation - Low quality portfolio

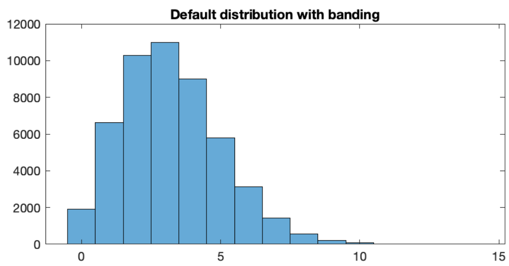


Figure 3.23. Distribution of defaults

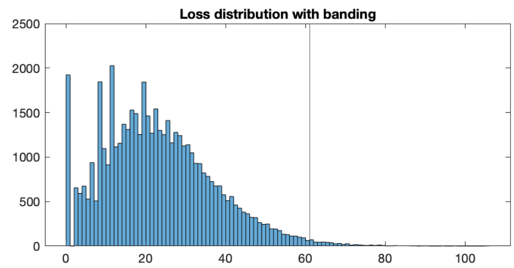


Figure 3.24. Distribution of losses

Figure 3.25. CR+ with banding implementation - High quality portfolio

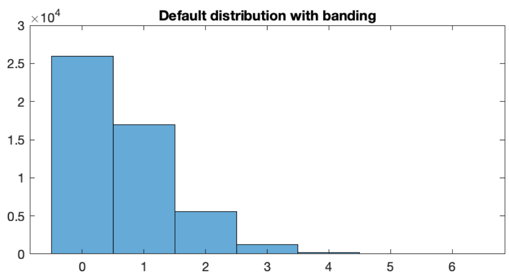


Figure 3.26. Distribution of defaults

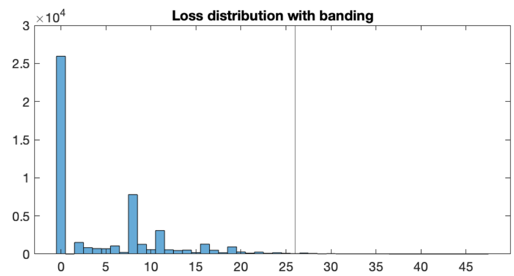


Figure 3.27. Distribution of losses

3.3.4 Probability distribution using Panjer Recursion

While using simulation for the implementation of CreditRisk+ is common, the framework actually has the distinguishing feature that it can avoid Monte Carlo simulation, as both the distribution of default rates and default losses can be described mathematically through probability generating functions.
In fact, the original algorithm proposed by Credit Suisse [17] to implement the CreditRisk+ model is based on a recursive formula known as the Panjer recursion. This allows very fast results in case of large portfolios, but increases the conceptual complexity with respect to simulation.

As previously illustrated in Section 2.4.4, in this simplest case of fixed default rate, the probability generating function for the losses G(z) can be expressed as in the multiplies of unit L of exposure:

$$G(z) = \sum_{n=0}^{\infty} P(\text{AggregateLosses} = n \times L) z^n \tag{3.2}$$

Assuming the exposure default independently, and consequently the exposure bands, we can use nice convolution properties:

$$G(z) = \prod_{i=1}^{m} G_i(z) \tag{3.3}$$

Yielding the probability generating function for the losses of band j:

$$G_j = \sum_{n=0}^{\infty} P(n \text{ defaults}) z^{n\nu_j} = \sum_{n=0}^{\infty} \frac{e^{-\mu_j} \mu_j^n}{n!} = e^{-\mu_j + \mu_j z^{\nu_j}} \tag{3.4}$$

From the PGF, it is actually possible to derive the actual distribution of losses, this is done through Panjer Recursion Technique. For n an integer, let A_n be the probability of a loss of $n \times L$, or n units from the portfolio. In order to compute A_n, it is possible to derive recursive relationships using Taylor series expansions of $G(z)$:

$$p(lossofnL) = \frac{1}{n!} \frac{d^n G(z)}{dz^n} |_{z=0} = A_n \tag{3.5}$$

Utilizing Leibnitz's formula, it is possible to write:

$$\begin{aligned} \left. \frac{d^n G(z)}{n! dz^n} \right|_{z=0} &= \left. \frac{d^{n-1}}{n! dz^{n-1}} \left(G(z) \frac{d}{dz} \sum_{j=1} m \mu_j z^{v_j} \right) \right|_{z=0} = \\ \frac{1}{n!} \sum_{k=0}^{n-1} \binom{n-1}{k} & \left. \frac{d^{n-k-1}}{dz^{n-k-1}} G(z) \frac{d^{k+1}}{dz^{k+1}} \left(\sum_{j=1}^{m} \mu_j z^{v_j} \right) \right|_{z=0} \end{aligned} \tag{3.6}$$

Additionally,

$$\frac{d^{k+1}}{dz^{k+1}}\left(\sum_{j=1}^{m}\mu_j z^{v_j}\right)\Bigg|_{z=0} = \begin{cases}\mu_j(k+1)! & \text{if } k = v_j - 1 \text{ for some j}\\ 0 & \text{otherwise}\end{cases} \tag{3.7}$$

By definition:

$$\frac{d^{n-k-1}}{dz^{n-k-1}}G(z)\Bigg|_{z=0} = (n-k-1)!A_{n-k-1} \tag{3.8}$$

Yielding:

$$A_n = \sum_{\substack{k \leq n-1,\\ k = v_j \text{ for some } j}} \frac{1}{n!}\binom{n-1}{k}(k+1)!(n-k-1)!\mu_j A_{n-k-1} = \sum_{j:v_j\leq n}\frac{\mu_j v_j}{n}A_{n-v_j} \tag{3.9}$$

Relating $\varepsilon_j = \nu_j \times \mu_j$, we gain a recurrence relationship permitting a fast computation of the distribution.

$$A_n = \sum_{j:v_j\leq n}\frac{\epsilon_j}{n}A_{n-v_j}. \tag{3.10}$$

To initiate the computation, we have a formula for the initial term, expressing the probability of zero loss arising from the portfolio:

$$A_0 = G(0) = F(P(0)) = e^{-\nu} = e^{-\sum_{j=1}^{m}\frac{\epsilon_j}{v_j}} \tag{3.11}$$

The computation of the Panjer coefficients is implemented in Matlab as in Algorithm 10. It is important to note, that unlike all previous sections, the output is not a simulation of the distribution of losses, but the exact probability distribution of losses for each discrete loss value. What is particularly powerful is that the calculation depends only on knowledge of ϵ_j and ν_j, which represents a very small amount of data even for a large portfolio consisting of many exposures.

The output is a vector $A(n)$ which contains the discrete probability of a loss equal to the coefficient n+1 (as Matlab indexes start from 1 and not 0).

Results are coherent with the results of simulation, as illustrated in Table 3.9, showing the comparison between the first 20 values yielded by simulation and by Panjer recursion for the low quality portfolio. A direct comparison with the distribution shapes will be discussed in the following sections.

```
Data: bands, principals
Result: A_n
for j = 1 to size(bands, 1) do
    mu_j = bands(j, 2);
    v_j = bands(j, 1);
    eps_j = bands(j, 3);
    sum_A = sum_A - eps_j / v_j;
end
A(1) = exp(sum_A);
for n = 1 to sum(principals) do
    sum_A = 0;
    for j = 1 to size(bands, 1) do
        mu_j = bands(j, 2);
        v_j = bands(j, 1);
        eps_j = bands(j, 3);
        if v_j < n+1 then
            sum_A = sum_A + (eps_j / n) * A((n+1) - v_j);
        end
    end
    A(n+1) = sum_A;
end
```

Algorithm 10: Calculation of Panjer recursion coefficients

3.3.5 Probability distribution using Fast Fourier Transform

In the context of credit risk models, the usefulness of the Panjer recursion is limited by two numerical issues, as computers does not have infinite precision [4]. First, the Panjer recursion cannot deal with arbitrarily large numbers of obligors: as the expected number of defaults in the portfolio increases, the computation of the first term of the recursion becomes increasingly imprecise; above a certain value for the expected number of defaults in the portfolio, the value found for the first term of the recursion becomes meaningless.

Second, the Panjer recursion is numerically unstable in the sense that numerical errors accumulate as more terms in the recursion are computed. This can result in significant errors in the upper tail of the loss distribution and hence in the computation of the portfolio's VaR. Melchiorri [13] developed an algorithm that is numerically more stable and faster than the standard model described in CreditSuisse, based on the Fast Fourier Transform.

As the probability generating function for the loss is discrete, it is possible to view the product of individual loss as a simple convolution that can be computed using the FFT. The calculation of the loss distribution for the whole portfolio, using FFT, can be done following these steps:

- the dimension n of the probability vector f is chosen so that the cumulative loss distribution shall have negligible probability outside the range [0,n]. To this purpose

Loss in millions	Monte Carlo Count	Monte Carlo Percent	Panjer coefficients
0	1923	3.85%	4.03%
1	0	0%	0%
2	655	1.31%	1.23%
3	592	1.18%	1.11%
4	674	1.35%	1.30%
5	528	1.06%	0.98%
6	938	1.88%	1.88%
7	508	1.02%	1.07%
8	1845	3.69%	3.58%
9	1097	2.19%	2.22%
10	912	1.82%	1.93%
11	2027	4.05%	4.04%
12	1115	2.23%	2.33%
13	1154	2.31%	2.37%
14	1369	2.74%	2.79%
15	1310	2.62%	2.56%
16	1529	3.06%	2.92%
17	1489	2.98%	3.20%
18	1253	2.51%	2.53%
19	1843	3.69%	3.61%

Table 3.9. Comparison between results with Monte Carlo and Panjer coefficients

we have chosen $n = 2nextpow2(totalexposure)$

- a probability vector is built for each j bands such that:

$$P(ndefaults) = \frac{\mu_j^n e^{-\mu_j}}{n!}, n = 0,1,2,...,2^r$$

- passing from the probability vector to the loss one requires a good understanding of probability generating function concepts. $P(ndefaults)$ represents the probability of having a number n of defaults, and this information is stored in the index of the vector form. For each band, having n defaults corresponds to a certain loss amount, maintaining the same probability of occurrence. It is thus possible to pass from the probability generating function of defaults, to that of losses by working with the indexes of the vector. This is built by maintaining the same value for the probability, but shifting the indexes so that they match loss values (which can be counter intuitive in the beginning!). In our implementation, for each band j, this yields:

$$l(loss) = \frac{\mu_j^n e^{-\mu_j}}{n!}, n = 0,1,2,...,2^r$$

The Matlab implementation, detailed in Algorithm 11, requires particular attention, as indices in vectors start from value 1. This means that probability values will be shifted by one unit.

Data: bands, n
Result: loss_distribution_fft, def_distribution_fft
for $j = 1$ *to size(bands, 1)* **do**
  l = zeros(1, n);
  p = zeros(1, n);
  μ_j = bands(j, 2);
  v_j = bands(j, 1);
  max_def = bands(j, 4);
  k = 0;
  p(1) = $e^{-\mu_j} \cdot \frac{\mu_j^k}{k!}$;
  l(1) = $e^{-\mu_j} \cdot \frac{\mu_j^k}{k!}$;
  for $k = 1$ *to max_def* $\times$ *3* **do**
    p($k + 1$) = $e^{-\mu_j} \cdot \frac{\mu_j^k}{k!}$;
    l($k \cdot v_j + 1$) = $e^{-\mu_j} \cdot \frac{\mu_j^k}{k!}$;
  end
  loss_distribution = loss_distribution .* fft(l);
  def_distribution = def_distribution .* fft(p);
end
loss_distribution_fft = ifft(loss_distribution);
def_distribution_fft = ifft(def_distribution);

Algorithm 11: Loss and Default Distribution computation with FFT

In Figures 3.28 and 3.31 the resulting density functions, in red the one obtained by Panjer recursion, in blue coefficients obtained by FFT. It is evident that the two methods are equivalent (and yield similar distributions to the ones obtained by Monte Carlo simulation).

Figure 3.28. PGF implementation - Low quality portfolio

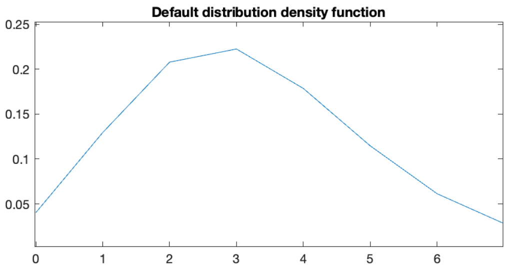


Figure 3.29. Distribution of defaults

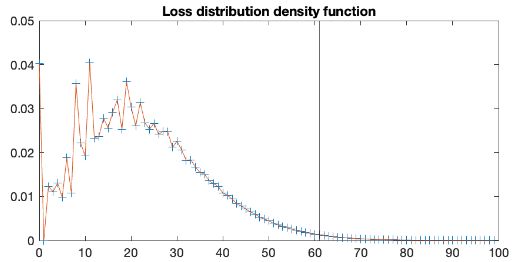


Figure 3.30. Distribution of losses

Figure 3.31. PGF implementation - High quality portfolio

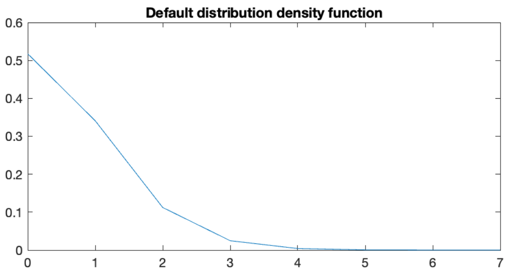


Figure 3.32. Distribution of defaults

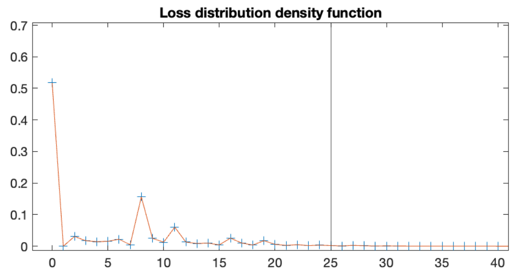


Figure 3.33. Distribution of losses

3.3.6 Comparison of results between implementations

Low quality portfolio

It is possible to notice that:

- from a visual perspective, the methods yield similar distributions, thus confirming the coherency between the different implementations;
- risk measures are similar between CM and CR+ without banding, while introducing the banding implies an approximation that becomes evident in the VaR calculations;
- banding reduces significantly computational time, and methods based on the calculation of the pgf (Panjer and FFT) are the most efficient, as shown in Table 3.10.

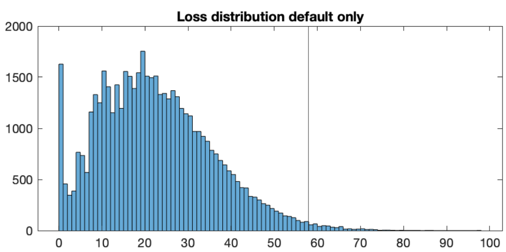


Figure 3.34. CreditMetrics without migration

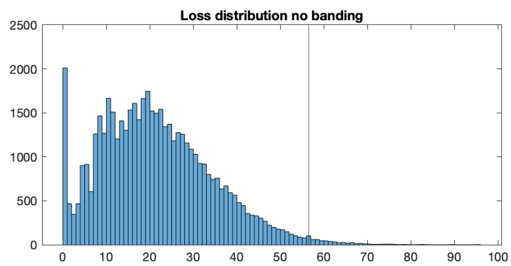


Figure 3.35. CR+ Monte Carlo without banding

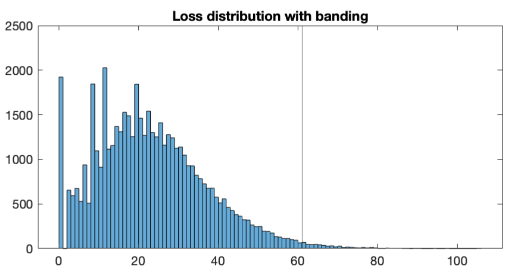


Figure 3.36. CR+ Monte Carlo with banding

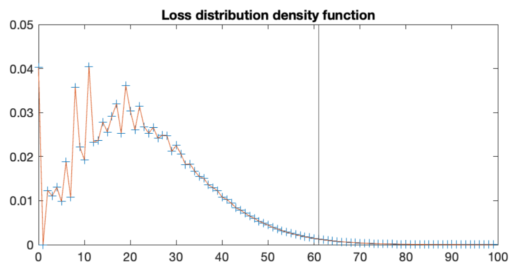


Figure 3.37. CR+ Panjer recursion and FFT

Table 3.10. Comparison for a low quality portolio

Metric	Time	VaR99	VaR999
CM w/o migration	0.23	57.88	71.61
CR+ Monte Carlo w/o banding	40.45	56.5	70.92
CR+ Monte Carlo w/ banding	4.75	61	76
CR+ Panjer recursion	0.023	61	77
CR+ FFT	0.039	61	77

High quality portfolio

It is possible to notice that:

- from a visual perspective, the methods yield very similar distributions, again confirming the coherency between the different implementations;
- again, risk measures are similar between CreditMetrics and CreditRisk+ without banding, while introducing the banding implies an approximation that becomes evident in the VaR calculations, yielding higher values;
- banding reduces significantly computational time, and methods based on the calculation of the pgf (Panjer and FFT) are the most efficient, as shown in Table 3.11.

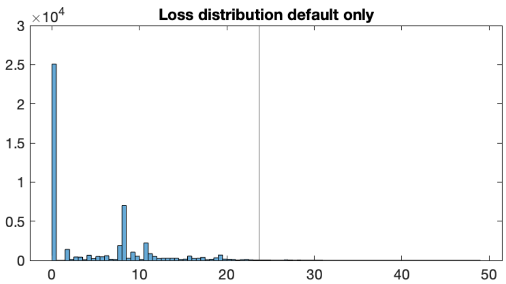


Figure 3.38. Credit Metrics without migration

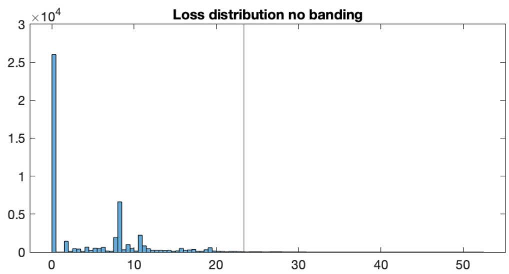


Figure 3.39. CR+ Monte Carlo without banding

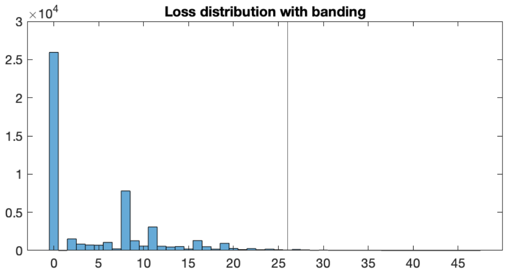


Figure 3.40. CR+ Monte Carlo with banding

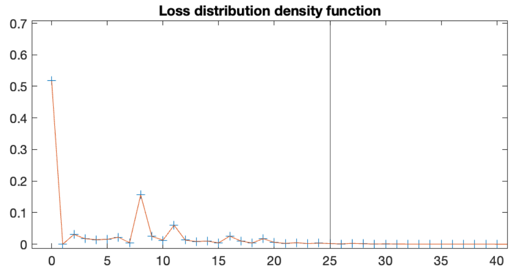


Figure 3.41. CR+ Panger recursion and FFT

Table 3.11. Comparison for a high quality portolio

Metric	Time	VaR99	VaR999
CM w/o migration	0.42	23.7	32.2
CR+ Monte Carlo w/o banding	39.14	23.32	31.56
CR+ Monte Carlo w/ banding	4.4	26	36
CR+ Panjer recursion	0.0076	25	36
CR+ FFT	0.038	25	35

3.4 Sector Analysis

Previously, the default event was modelled as independent, assuming a diagonal structure in CreditMetrics, and using a fixed default probability rate for CreditRisk+.
The purpose of this section is to incorporate a possible correlation structure between borrowers. Using the previous implementations frameworks, this can be respectively accomplished:

- for CreditMetrics, by modifying the correlation matrix M to describe the correlation relationship between the borrowers. This implementation can account in a straight forward way for one-sector, multi-sector independent, and multi-sectors correlated structures.

- for CreditRisk+, the implementation is not as straightforward, as it models the effects of background factors by using default rate volatilities that result in increased defaults, rather than by using default correlations as a direct input.

The fact that CreditRisk+ does not attempt to model correlations explicitly but captures concentration effects through the use of default rate volatilities, makes a sensible mapping between the two models a challenging task.

3.4.1 Setting a common framework for CreditMetrics and CreditRisk+

In this section, we will focusing on building a common framework to compare results from CreditMetrics and CreditRisk+, following the work of Koyluoglu and Hickman [11], using simulation to extend results for the distribution of defaults for an non homogeneous portfolio with multiple independent sectors, in the default only case.
In order to construct a comparable framework for mapping the input parameters, both models will be slightly reformulated in their implementation with respect to the previous sections, without changing the fundamentals.

For borrowers affected by a single economic factor, and similar size exposures and credit ratings, for the realisation of an economic factor X, the conditional probability of default for CreditMetrics can be written:

$$pd_X^{CM} = \phi\left[\frac{c_1 - \sqrt{\rho}X}{\sqrt{1-\rho}}\right] \tag{3.12}$$

where $X \sim N(0{,}1)$, $\phi(c_1) = \bar{pd}$. In CreditRisk+, the economic factor is assumed to have Gamma distribution. To have a coherent mapping, Koyluoglu and Hickman modify the distribution of the economic factor to be normal distributed, and in order the preserve the gamma distribution of the default rate (key assumption of CreditRisk+), they suggest the following transformation:

$$\int_0^{\xi} \Gamma(p_D^{CR};\alpha;\beta)dp = \int_{\chi}^{\infty} \phi(x)dx \tag{3.13}$$

which yields the following conditional default rate

$$pd_X^{CR} = \Psi^{-1}(1-\Phi(X);\alpha;\beta). \tag{3.14}$$

Now the issue is to parameterize consistently the two models. Koyluoglu and Hickman argue that means and standard deviations of default must be the same, finding the following default rate volatility for CreditMetrics, expressed as function of $\bar{p}_D^{CM}$ and ρ.

$$\sigma_{CM}^2 = \int_{-\infty}^{\infty}\left(\Phi\left[\frac{\Phi^{-1}(\bar{p}_D^{CM}) - \sqrt{\rho}X}{\sqrt{1-\rho}}\right] - \bar{p}_D^{CM}\right)^2 \phi(x)dx. \tag{3.15}$$

This way, given the default probabilities $\bar{p}_D^{CM} = \bar{p}_D^{CR}$, we have set a correspondence between the asset correlation ρ, which is taken in input in CreditMetrics, and the default volatility σ_{CR}.

3.4.2 Input parameters

For simplicity of implementation, we sill assume that obligors in the same default rate bracket, are affected in the same way by a background parameter, thus resulting to be assigned to the same sector. This assumption of credit quality correlation, introduced in Chapter 8 of [14], is purely for design purposes, and can be relaxed accounting for different correlation structures. Concretely, this means that we are dividing our portfolios in sub-portfolios with the same probability of default (but different exposures), and then aggregating the results to get the full portfolio view. This implementation is clearly a simplification, but shows in a rather simple way that it is indeed possible to map the two framework's parameters in a sensible ways for simple cases, and constitutes a base for further works.

Following the assumption that the higher PD the higher the volatility, we set the following correlation inputs:

$$\rho = [0.05, 0.075, 0.1, 0.125, 0.15, 0.175, 0.2].$$

3.4.3 CreditMetrics

In order to calculate the loss distribution for CreditMetrics in a way that allows us to consistently model and compare the two frameworks, we proceed as follows:

1. For each sector, calculate:
$$c_1 = \Phi^{-1}(pd)$$
2. For each scenario, simulate the background factor as a standard normal random variable $X \sim N(0,1)$
3. For each realisation, calculate the conditional probability of default
$$pd_X^{CM} = \phi[\frac{c_1 - \sqrt{\rho}X}{\sqrt{1-\rho}}]$$
4. Simulate the default, as a random binomial with probability pd_X^{CM}
5. If the borrower defaults, calculate the corresponding loss
6. Repeat for each scenario, for each sector calculate α_s and β_s for CreditRisk+.

It is possible to calculate the conditional probability of default for each scenario, and each sector, as in Algorithm 12.
Having the conditional default probability for each scenario, we can easily calculate the loss and default distribution as in Algorithm 13.

3.4.4 CreditRisk+

1. using the same scenarios, calculate the conditional default probability
$$pd_X^{CR} = \Psi^{-1}(1 - \Phi(X); \alpha_s; \beta_s)$$
2. Simulate the default as a random Poisson of parameter pd_X^{CR}

Data: n_scenarios, n_sectors, c1, rho
Result: pd_conditional_cm
for $j = 1$ *to n_scenarios* **do**
  x(j) = randn(1);
  for $i = 1$ *to n_sectors* **do**
    pd_conditional_cm(j, i) = normcdf$\left(\frac{c1(i)-\sqrt{rho(i)}\times x(j)}{\sqrt{1-rho(i)}}\right)$;
  end
end

Algorithm 12: Calculation of conditional PDs using Gaussian copula

Data: n_scenarios, n_firms, pd_conditional_cm, sectors, principals, recovery_rates
Result: loss_cm, default_count_cm
for $s = 1$ *to n_scenarios* **do**
  for $f = 1$ *to n_firms* **do**
    p = pd_conditional_cm(s, sectors(f));
    def = binornd(1, p);
    if def *== 1* **then**
      loss_cm(s, f) = principals(f) - principals(f) × recovery_rates(f) / 100;
      default_count_cm(s, f) = 1;
    end
  end
end

Algorithm 13: Calculation of loss and default count for each scenario

3. If the borrower defaults, calculate the corresponding loss

4. Repeat for each sector

Firstly, it is necessary to calculate the respective parameters α_s and β_s so that the model is consistent with the CreditMetrics implementation (described in Matlab Algorithm 14).

Having the parameters of the Gamma distribution, we can proceed to calculate the conditional probability of default for each scenario and sector (Algorithm 15).
Finally, we can calculate default and loss distributions, as in Matlab Algorithm 16.

Data: n_sectors, prob_default_rating, rho
Result: variance_k, volatility_k, alpha_k, beta_k
for $i = 1$ *to n_sectors* **do**
    variance_cm(x) =
    $\left(\text{normcdf}\left(\frac{\text{norminv(prob\_default\_rating}(i))-\sqrt{\text{rho}(i)}\times x}{\sqrt{1-\text{rho}(i)}}\right) - \text{prob\_default\_rating}\,(i)\right)^2 \times$
    normpdf(x);
    variance_k(i) = integral(variance_cm, $-\infty$, ∞);
    volatility_k(i) = $\sqrt{\text{variance\_k}\,(i)}$;
    alpha_k(i) = $\frac{\text{prob\_default\_rating}(i)^2}{\text{variance\_k}(i)}$;
    beta_k(i) = $\frac{\text{variance\_k}(i)}{\text{prob\_default\_rating}(i)}$;
end
Algorithm 14: Calculation of parameters for sector-level default distributions

Data: n_scenarios, n_sectors, x, alpha_k, beta_k
Result: pd_conditional_cr
for $j = 1$ *to n_scenarios* **do**
    for $i = 1$ *to n_sectors* **do**
        pd_conditional_cr$(j,\ i)$ = gaminv$(1 - \text{normcdf}\,(x(j))\,, \alpha_k(i), \beta_k(i))$;
    end
end
Algorithm 15: Calculation of conditional PDs for each scenario and sector

Data: n_scenarios, n_firms, pd_conditional_cr, sectors, principals, recovery_rates
Result: loss_cr, default_count_cr, p_sum
for $s = 1$ *to n_scenarios* **do**
    for $f = 1$ *to n_firms* **do**
        p = pd_conditional_cr(s, sectors(f));
        p_sum(s) = p_sum(s) + p;
        default_count_cr(s, f) = poissrnd(p);
        if *default_count_cr(s, f) > 0* **then**
            loss_cr(s, f) = principals(f) - principals(f) × recovery_rates(f) / 100;
        end
    end
end
Algorithm 16: Calculation of loss and default count for each scenario

3.4.5 Comparison of results

We chose to maintain 50.000 scenarios, although noticing that computational times start becoming significantly higher due to calculations of the inverse which are costly (225.88 seconds).

First, we verify that there is indeed consistency in our parameterization, as default mean and the implied standard deviation are consistent between the CreditMetrics and CreditRisk+ implementations, as visible from Table 3.12.

	Default mean		default st. dev.	
Sector	CM	CR+	CM	CR+
Sector 1	0.0000	0.0000	0.0000	0.0000
Sector 2	0.0000	0.0000	0.0000	0.0000
Sector 3	0.0006	0.0006	0.0009	0.0009
Sector 4	0.0018	0.0018	0.0026	0.0026
Sector 5	0.0106	0.0106	0.0131	0.0132
Sector 6	0.0520	0.0520	0.0497	0.0497
Sector 7	0.1978	0.1978	0.1291	0.1290

Table 3.12. Table of Mean and Standard Deviations for CM and CR

Low quality portfolio

It is possible to notice how adding correlation effects changes the shape of both the default and loss distribution. For the default distribution, shown in Figure 3.42:

- the mean probability of default is maintained, yielding 3.22 average defaults per run as by design;
- we notice a much larger number of scenarios where no borrower defaults;
- we notice scenarios with a much higher number of joint defaults, yielding a distribution that is more skewed and with a significantly fatter tail;
- while the effect of correlation modifies both the CreditMetrics and CreditRisk+ way in a coherent way, the two distributions have slightly different shapes due to the different underlying assumptions of the model.

Figure 3.42. Comparison between the distribution of defaults for the low quality portfolio

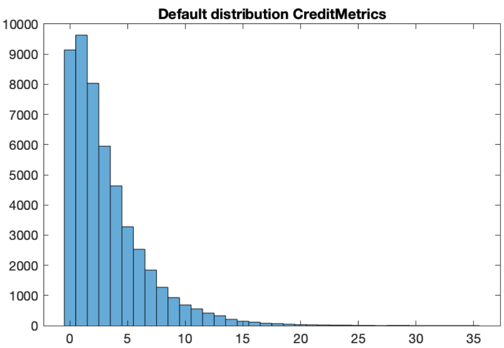


Figure 3.43. CreditMetrics

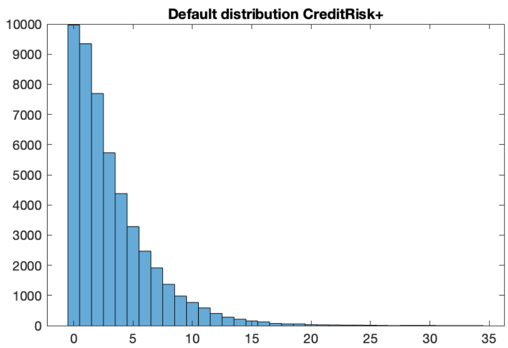


Figure 3.44. CreditRisk+

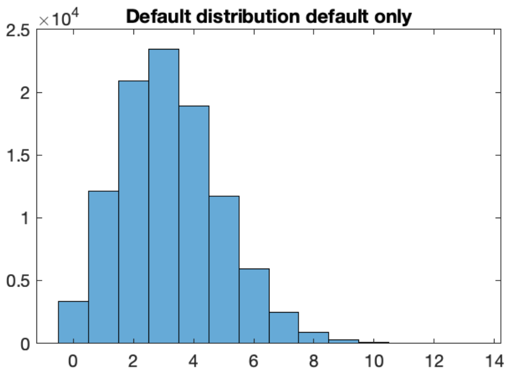


Figure 3.45. CreditMetrics without correlation

Similar considerations are noticeable in the distribution of losses, Figure 3.46:

- from a qualitative point of view, the resulting distributions for the two frameworks are modified by the effect of correlation (volatility) in a similar way;
- scenarios with no defaults yield a much higher number of scenarios without any loss;
- the joint default behaviour yields a distribution that has a much longer and fatter tail, accounting for scenarios where the loss is up to two times higher;
- coherently with results yielded by Gemmil [8], the CreditMetrics loss distribution has higher skewness and kurtosis, yielding higher measures for the Value at Risk, as reported in Table 3.13. This is especially evident for the 0.999 measure, which is about 25% higher with respect to the CreditRisk+.

Figure 3.46. Comparison between the distribution of losses for the low quality portfolio

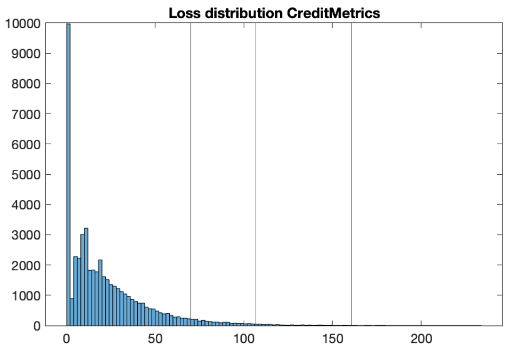


Figure 3.47. CreditMetrics

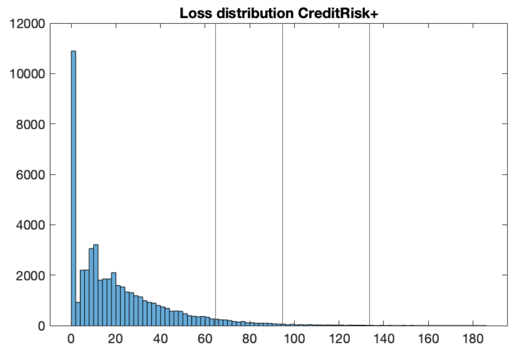


Figure 3.48. CreditRisk+

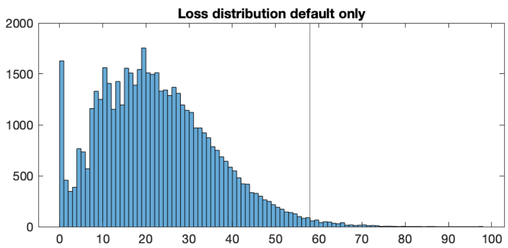


Figure 3.49. CreditMetrics without correlation

	CreditMetrics	CreditRisk+
VaR 95%	70.0809	64.7183
VaR 99%	106.7366	94.7487
VaR 99.9%	160.8724	133.7612
Skewness	1.8597	1.6233
Kurtosis	7.9973	6.5142

Table 3.13. Table of VaR, Skewness, and Kurtosis for losses in the low quality portfolio

High quality portfolio

Although less evident than in the low quality borrower portfolio, also in this case we notice that introducing sector correlation yields default distributions that have longer and heavier tails, Figure 3.50, accounting for higher joint defaults for certain background factors.

Figure 3.50. Comparison between the distribution of defaults for the high quality portfolio

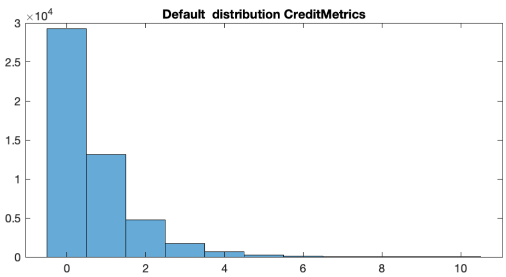


Figure 3.51. CreditMetrics

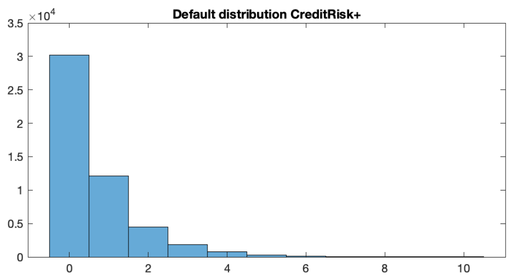


Figure 3.52. CreditRisk+

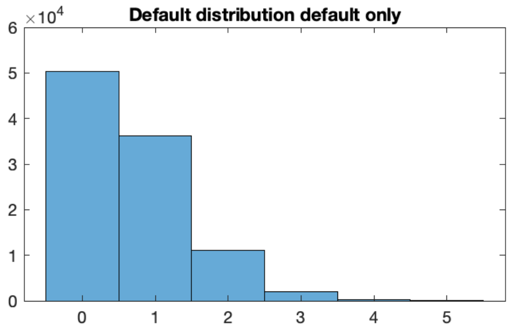


Figure 3.53. CM without correlation

It is possible to notice the same for the loss distribution, Figure 3.54, although the change is again visually less evident than before.

Figure 3.54. Comparison between the distribution of losses for the high quality portfolio

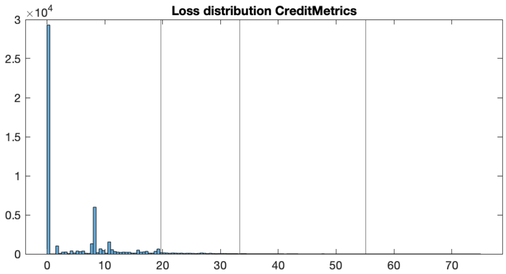


Figure 3.55. CreditMetrics

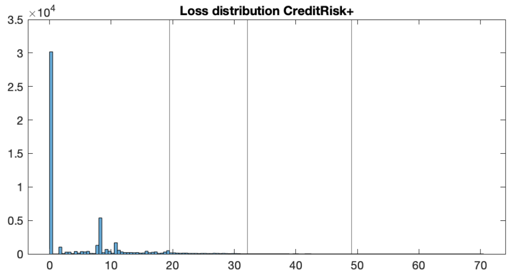


Figure 3.56. CreditRisk+

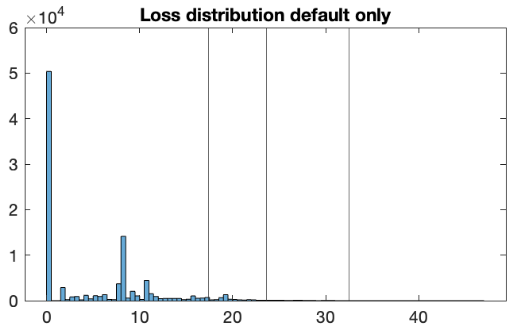


Figure 3.57. CreditMetrics without correlation

Regarding risk measures, displayed in Table 3.14, again CreditMetrics determines higher values for the Value at Risk, although the percentage difference between the two methods of modelling is less evident than for the previous portfolio. Thus, discrepancies between capital requirement for institutions using the two different model are less significant for higher quality portfolios.

	CreditMetrics	CreditRisk+
VaR 95%	19.6955	19.4930
VaR 99%	33.2950	32.1731
VaR 99.9%	55.0732	49.0928
Skewness	2.2117	2.1603
Kurtosis	9.8675	9.0608

Table 3.14. Table of VaR, Skewness, and Kurtosis for CM and CR

Chapter 4

Final Considerations

This thesis work aimed to provide insights on the modelling complexity faced by financial institutions when dealing with credit risk, which, additionally to a portfolio of equity, faces challenges such as:

- scarcity, lack of data;
- skewed distributions;
- underlying correlation structures between borrowers that can lead to extreme events.

We introduced the theory underlying two main approaches when dealing with modelling this risk: Structural and Mixture models, and their respective commercial implementations CreditMetrics and CreditRisk+. We then analysed the resulting default and loss distributions yielded by both models, considering different input and modelling assumptions in two realistic bond portfolios with different exposure sizes, and different credit qualities (extended previous works that mainly focused on homogeneous portfolios).

As general considerations on implementing a sensible comparison of the models:

- from an initial look, the studied credit risk portfolio models seem to be quite different. However, deeper examination reveals that the models can be assimilated to a single general framework, which identifies three critical points of comparison: the default rate distribution, the conditional default distribution, and the convolution / aggregation technique.
- building coherent models with the two approaches isn't a trivial task, but requires adequate preparation of the data, especially regarding exposures, default probabilities and correlation structures. A naive comparison, using parameters estimated from data using different techniques, is quite likely to produce significantly different results for the same portfolio. Furthermore, particularities of CreditMetrics, such as coupons and credit class migrations, are not transferable to CreditRisk+ in a justifiable way if not under mathematical workarounds. For this reason we bench-marked the frameworks using the default-only CreditMetrics model;

- modelling the dependence structure in a coherent way is especially challenging, as CreditMetrics takes in input direct asset-value correlation between borrowers, while CreditRisk+ models dependencies through default volatilities. Mapping the first into the latter is not straightforward, but can be achieved for single economic factor dependencies by modifying the underlying Gamma distribution used in CreditMetrics to be normally distributed, and then introducing an appropriate transformation, as first theorised by [11]. This method ensures coherency in the first (mean) and second moment (standard deviation) of the distribution of defaults. The procedure deeply modifies the standard implementations of the models;

- given an adequate parameterization of the input data, it is actually possible to obtain comparable default and loss distributions, and their associate risk measures, thus enabling the comparison of the results, the efficiency and the implementation methods of the different models;

- with coherent parameterizations, the models still yield some differences in the estimate of risk measures. This is more evident in high confidence intervals (99%, 99.9%).

Diving more in detail into the results:

- CreditMetrics offers a richer model in terms what it is possible to model, incorporating easily losses (and gains) that may arise from credit quality migrations and the concept of coupons. This framework is thus indicated for corporate bond portfolios, where ratings and coupons are easily available data. The effect of including or omitting migrations is more evident in high quality portfolios, where migration risk accounts for a larger part in the overall risk measures;

- regarding efficiency, CreditRisk+ is a faster and less expensive approach for calculating capital requirements, especially when recurring to non simulation implementations, such as calculating Panjer recursion coefficients or Fast Fourier Transform coefficients. Some attention need to be into when using the banding approach, as it can introduce higher (or lower) estimates of the Value at Risk (VaR) due to approximations. In our case this can account to up to 10% for very high confidence intervals in low quality portfolios. Given efficiency properties for very large portfolios, CreditRisk+ may be more suitable for evaluation of different types of debt, such as retail;

- for incorporating dependencies, if we have data regarding Asset value correlations, CreditMetrics allows to easily model pairwise relations between borrowers with the simple use of the multivariate normal distribution. For CreditRisk+ it is relatively simple to implement dependencies in the case of a single sector, thus having a portfolio, or a collection of sub-portfolios, with similar default rates and exposed to the same risk factors. This can be the case of a large retail portfolio exposed to the economy of a certain country. Incorporating more risk factors becomes more complex, and is subject to some arbitrate when assigning the weights to each risk component;

- finally, for the actual estimation of risk measures, discrepancies seem to be almost negligible in cases of lower confidence intervals and high quality portfolios, coherently with what is found by [11]. The driver in the case of differences between the models is the shape of the distribution of default of the models, yielding slightly higher results for CreditMetrics.

In conclusion, the choice of the model to be implemented largely depends on several factors:

1. The type of credit risk within the portfolio (e.g., corporate bonds, retail, other).
2. The size of the portfolio.
3. The objective of the measurement (e.g., the need for data on reserves or provisions, where CreditMetrics provides more information about the size of reserves needed for non-defaulted exposures [8]).
4. The availability of data (e.g., asset default correlations, default rate volatility).

While this thesis provides an initial exploration of the reconcilable differences between two of the most important industry models for credit risk modeling, future studies can explore the capital requirements yielded when dealing with more complex correlation structures.

Bibliography

[1] Zvika Afik, Ohad Arad, and Koresh Galil. Using merton model for default prediction: An empirical assessment of selected alternatives. *Journal of Empirical Finance*, 35:43–67, 2016.

[2] Jean-Philippe Aguilar, Nicolas Pesci, and Victor James. A structural approach to default modelling with pure jump processes. *Applied Mathematical Finance*, 28(1):48–78, 2021.

[3] Ebenezer Fiifi Emire Atta Mills, Bo Yu, and Lanlan Gu. On meeting capital requirements with a chance-constrained optimization model. *SpringerPlus*, 5:1–20, 2016.

[4] Renzo Avesani, Kexue Liu, Alin Mirestean, and Jean Salvati. Review and implementation of credit risk models of the financial sector assessment program (fsap). 2006.

[5] Sreedhar T Bharath and Tyler Shumway. Forecasting default with the merton distance to default model. *The Review of Financial Studies*, 21(3):1339–1369, 2008.

[6] Somnath Chatterjee et al. Modelling credit risk. *Handbooks*, 2015.

[7] Gorton Dan. *Aspects of Modeling Fraud Prevention of Online Financial Services.* PhD thesis, KTH Royal Institute of Technology, 2015.

[8] Diana Diaz and Gordon Gemmill. A systematic comparison of two approaches to measuring credit risk: Creditmetrics versus creditrisk+. *Trans. 27th ICA*, 2002.

[9] C. Finger. Sticks and stones. riskmetrics group. working paper. 1999.

[10] Michael B Gordy. A comparative anatomy of credit risk models. *Journal of Banking & Finance*, 24(1-2):119–149, 2000.

[11] H Ugur Koyluoglu and Andrew Hickman. A generalized framework for credit risk portfolio models. *Manuscript, Oliver Wyman & Company*, 1998.

[12] Alexander J McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: concepts, techniques and tools-revised edition.* Princeton university press, 2015.

[13] Mario R Melchiori. Creditrisk+ by fast fourier transform. *YieldCurve, August*, 2004.

[14] JP Morgan et al. Creditmetrics-technical document. *JP Morgan, New York*, 1997.

[15] Basle Committee on Banking Supervision and Bank for International Settlements. *Principles for the management of credit risk*. Bank for International Settlements, 2000.

[16] Andrea Serata. La gestione del rischio di tasso di interesse del portafoglio bancario: evidenze da un campione di banche italiane. 2021.

[17] Credit Suisse. *CreditRisk+: A credit risk management framework*. Credit Suisse, 1997.

[18] Maria Vassalou and Yuhang Xing. Default risk in equity returns. *The journal of finance*, 59(2):831–868, 2004.

[19] Wei Wang and Wulin Suo. Assessing default probabilities from structural credit risk models. *Available at SSRN 897793*, 2006.

[20] Tom Wilde. Creditrisk+. *Encyclopedia of quantitative finance*, 2010.