Assessing concentration and diversification in portfolio credit risk models

by

O.C. van den Bergh

to obtain the degree of Master of Science in Applied Mathematics
at the Delft University of Technology,
to be defended publicly on Friday July 31, 2020 at 14:00.

Student number:	4295943	
Project duration:	January 6, 2020 – July 31, 2020	
Thesis committee:	Dr. ir. P. Cirillo,	TU Delft, supervisor
	Prof. dr. ir. C. W. Oosterlee,	TU Delft
	J. Wevers, MSc.	ING Bank, supervisor

Abstract

This thesis explores existing and proposes new methods for assessing concentration risk in default-only credit risk models. Within the existing methods, the analytic Granularity Adjustment is studied in the single factor Gaussian threshold and in the CreditRisk+ framework. These adjustments are tested on a sample portfolio in the presence of recovery risk, and we show that the CreditRisk+ adjustment is more conservative than the Gaussian threshold adjustment. Furthermore, we show that in the presence of recovery risk, the accuracy of the adjustment on exposure level deteriorates. Additionally, the Granularity Adjustment is extended to an independent single factor *t*-threshold model to account for heavier tailed asset returns. Based on the independent single factor *t*-threshold model, we suggest an ASRF equivalent that could serve as an alternative to the current IRB framework. Although much existing literature is focusing on developing analytical methods for measuring concentration risk, recent advances in computational speed make Monte Carlo methods an interesting substitute for measuring concentration risk. Using the methods developed by [35], we propose a split between Monte Carlo based Economic and Regulatory Concentration Risk and show that these measures do not coincide for a given portfolio. This method involves a novel way of assessing idiosyncratic risk in multi factor frameworks. In order to assess sector concentration risk, this thesis proposes a Diversification Factor and a Capital Diversification Index as risk management tools. Finally, this thesis provides a clear account of the effects of concentration, diversification and recovery risk on the portfolio loss distribution for both Gaussian and *t*-threshold models. This thesis was carried out in close cooperation with ING Bank.

Acknowledgements

This thesis is submitted to the Delft Institute of Applied Mathematics in partial fulfillment of the requirements for the degree of Master of Science in Applied Mathematics. I would like to express my gratitude towards Dr. Pasquale Cirillo for his helpful coaching and guidance throughout this work. Furthermore, I would like to thank Prof. dr. ir. Kees Oosterlee for being a member of the Thesis Committee. I would also like to express my gratitude towards ING and towards Arjan de Ridder, Jeroen Wevers and Nicolas Derrien in particular for granting me the opportunity to conduct this research.

Oliver van den Bergh
Amsterdam, July 2020

Contents

I

Introduction

1
Introduction

Within the core tasks of modern banks, supplying the market with various forms of credit is arguably the most important. Naturally, supplying the market with credit brings along risks that should be properly managed. One of these risks is credit risk; the risk that a borrower defaults on its obligations towards the bank. Adequate measurements of credit risk in a banks portfolio is a heavily practiced, regulated and studied topic. The main difficulties involved with modelling credit risk arise from the fact that default events are generally quite rare and they occur unexpectedly. However, when a default does take place, it might lead to significant losses.

In order for the bank to stay solvent and cover unexpected losses due to credit risk, the bank holds capital as a protection against this type of risk. The European Central Bank extensively regulates the amount of capital that has to be held through the Basel Accords. Credit risk in a portfolio arises from two sources; systematic and idiosyncratic risk. Systematic risk is caused by unexpected changes in macroeconomic conditions that affect the global and local economies to which borrowers are exposed. This kind of risk cannot be eliminated through diversification across borrowers. The second source of risk, idiosyncratic risk, arises from shocks to individual borrowers that do not affect other borrowers. This borrower specific risk can be diversified away as the largest exposures in a portfolio account for small shares of the total portfolio exposure.

The Internal Ratings Based approach of the Basel capital framework is well suited in measuring systematic credit risk, but by being portfolio invariant approach, it does not allow for assessing credit risk due to diversification. Essentially, the capital charge in the IRB framework is derived from an asymptotic single risk factor model which inherently implies that idiosyncratic risk is assumed to be diversified in the portfolio. This implies that the IRB framework is unable to explicitly account for concentration risk measurements due to idiosyncratic shocks to the portfolio. Under Basel II and III, banks are obliged to assess concentration risk under Pillar II, but banks and regulators have a large degree of freedom in choosing the exact quantitative tools to measure the additional capital required to cover for concentration risk.

Throughout this thesis, we assess the existing techniques and develop new techniques for quantifying concentration risk in credit portfolios. We restrict ourselves to default only models, therefore we do not take any effects of credit migration into account. In contrast to many existing methods, we both research the effect of recovery risk on concentration risk and the effect of heavier tailed asset returns.

Generally, concentration risk is a topic covering two specific effects: single name concentration risk and sector concentration risk. Single name concentration risk refers to the type of risk that is due to idiosyncratic shocks to large exposures in a portfolio. Sector concentration risk refers to the type of risk that is due to diversifying a portfolio across multiple sectors and regions. In short, concentration risk in credit portfolios arises from either an unequal distribution of loans to single borrowers or industrial and regional sectors. We analyze both variants of concentrations risk and assess techniques to quantify both effects. For major financial institutions, losses due to concentration risk can be extreme. Therefore, adequate measurement of concentration risk in a bank's portfolio of outstanding loans is essential to the financial stability of the banking system. Historically, concentration risk in loan portfolios has been shown to be one of the major causes of banking distress [21].

This thesis has been conducted in close cooperation with ING Bank, a Dutch multinational banking and financial service corporation active in over 40 countries. ING's total assets exceed €890 million in 2019 [11] making it one of the biggest banks worldwide, consistently ranking in the top 30 globally. Additionally, ING is the only Dutch bank to be listed among the Globally Systemic Important Banks.

Contributions of this work

The purpose of this work is twofold. Firstly, we review existing methods for assessing concentration risk in credit portfolios. We explore both Monte Carlo based and analytical methods and compare these methods across aspects such as ease of implementation and accuracy. Secondly, we develop new methods for measuring concentration risk and adapt existing methods to fit different underlying models of credit risk, accounting for effects such as recovery risk. This work is designed to be relevant to practitioners of credit risk management and model validators within all financial institutions. Since this thesis is written in close cooperation with ING's model validation department, this work is focused on exploring and testing existing methods for assessing concentration risk.

More specifically, this work extends existing literature to real-life portfolios and test analytical methods for assessing concentration risk on portfolios many times the size of the portfolios that are under investigation in the existing literature. Furthermore, we test the existing methods without any assumptions on the homogeneity of the portfolio in terms of LGD, PD and EAD. Additionally, we explore the effects of diversification by allowing the obligors to depend on multiple systemic factors, in contrast to a single region or industry in the existing literature. Furthermore, we allow the systematic factors to be governed by a covariance matrix instead of the standard correlation matrix. We also adapt the methods developed in [35] to this new setting. Additionally, we introduce recovery risk across all analytical methods proposed in this work and assess the accuracy of analytical approximation to credit risk in the presence of recovery risk.

Moreover, we respond to the suggestions of existing literature to make different assumptions on the underlying distributions of asset returns by adapting the concentration risk framework to student t-distributed asset returns instead of normally distributed asset returns. This also involves finding a solution to an ASRF and Granularity Adjustment equivalent for a t-threshold model with independent systematic and idiosyncratic risk factors. Next, by adopting a different interpretation of the methods developed by [35], we argue in support of adapting two new forms of concentration risk: Regulatory and Economic Concentration Risk. Lastly, we suggest a measure of diversification by eliminating the effects of diversification through assuming an all-ones matrix and showing that for this matrix, the multi factor threshold model equals an equivalent single factor model.

Outline

In Chapter 2 we start with revisiting the regulatory framework and we introduce the general credit risk model setting. Furthermore, we take a closer look at the definition of concentration risk and shortly discuss ING's current methods for determining Economic Capital. In Chapter 3 we discuss the mathematical tools required to assess concentration risk. Additionally, industry models such as Moody's RiskFrontier and CreditRisk+ are introduced and evaluated shortly. Lastly, we introduce current techniques for measuring concentration risk. Chapter 4 applies the techniques developed to existing risk models and introduces some more recent methods for evaluating concentration risk. Additionally, we shortly explore some computational techniques such as Monte Carlo methods. To verify the performance of the introduced methods, Chapters 5 and 6 are devoted to testing the methods empirically. Chapter 5 introduces the portfolios and correlations structure on which the methods are evaluated, while Chapter 6 focuses on the justification and tests of the methods discussed. Chapter 7 concludes and suggests future research.

2

Concentration risk and the regulatory framework

Throughout this chapter we will introduce the basics of credit risk modelling such as the variables of risk (PD,LGD,EAD) and measures to quantify the riskiness of a credit portfolio, such as the VaR and ES. Additionally, we will evaluate Gaussian threshold models and derive the Asymptotic Single Risk Factor (ASRF) model. Lastly, we will review ING's current method for determining Economic Capital.

2.1. The Regulatory Framework

In the most recent Basel capital framework, the Asymptotic Single-Risk Factor (ASRF) model underpins the Internal Rating Based (IRB) approach. However, this approach does not allow for the explicit measurement of concentration risk. The IRB risk-weight functions of Basel II (on which we will elaborate in Section 2.6.1) are designed to be portfolio invariant by nature, therefore neglecting the effects of concentration on the loss distribution. Essentially, this means that the capital required for a specific loan only depends on the risk profile of this specific loan and must not depend on the portfolio it is part of [30]. A major drawback of this method is its incapability of capturing concentration effects. Therefore, Basel II states that banks should explicitly consider credit risk concentrations in the assessment of capital adequacy under Pillar 2 [31]. Basel II states that "banks should have in place effective internal policies, systems and controls to identify, measure, monitor, and control their credit risk concentrations." These policies cover the following concentrations:

- Significant exposures to an individual counterparty or group of related counterparties.
- Credit exposures to counterparties in the same economic sector or geographic region.
- Credit exposures to counterparties whose financial performance is dependent on the same activity or commodity.
- Indirect credit exposures arising from a bank's CRM activities.

Throughout this work, we will focus on addressing the first two of the aforementioned concentrations. Basel developed a more extensive framework for managing concentration risk in 2013 under the name *Supervisory framework for measuring and controlling large exposures*. In [32] the Committee stresses that the large exposure framework complements the IRB capital standard as the standard IRB approach is not explicitly designed to protect banks from large losses resulting from the sudden default of a single counterparty. Furthermore, the large exposure framework will be part of Basel III when it is implemented on January 1st, 2022. The framework sets prudent limits to large exposures. Large exposures are defined by the sum of all exposures of a bank to a single counterparty that are equal to or exceed 10% of its Tier 1 capital. The limit for large exposures is set at 25% for general banks and at 15% for globally systemically important banks. Some exposures will be exempted from these requirements such as intraday interbank, sovereign and central bank exposures.

2.2. Risk Measurement and Variables of Risk

Credit risk is the risk that the value of a portfolio of loans changes due to unexpected changes in the credit quality of issuers or trading partners. This includes both losses due to defaults as losses caused by changes in credit quality [25]. According to the ECB, credit risk is defined "as the risk of losses due to credit events, i.e. default (an obligor being unwilling or unable to repay its debt) or a change in quality of the credit (rating change)"[28]. Broadly speaking, credit risk can be quantified in default or in migration mode. In default mode, the only relevant risk is the risk of default meaning mark to market losses due to rating migrations are not taken into account. In contrast, migration mode deals with all mark to market gains and losses due to changes in ratings [28]. Default is, in essence, an extreme occasion of rating migration and can therefore be regarded as a particular case of migration mode. Empirically, the calculated credit risk in migration mode is usually higher than in default mode since the probability of a rating downgrade exceeds the probability of an upgrade [28]. In this work, we will focus on default mode only.

Credit risk modelling poses several challenges. The main difficulties when modelling credit risk arise from the fact that default events are rare and occur unexpectedly. Moreover, there is a lack of public information and data regarding the credit quality of corporations. Often, this gives rises to a problem of informational asymmetry in which the management of a firm is better informed about the financial prospects of the firm than the lender, in this case the bank, is. Furthermore, loss distributions are typically heavily skewed and have a relatively heavy upper tail. This translates into a large amount of risk capital to be required to cover potential large losses [25].

The uncertainty of whether an obligor will default within a set time horizon, typically one year, is measured by its probability of default. The probability of default ($PD \in [0,1]$) describes the probability of the default event occurring before the specified time horizon. The exposure at default (EAD) of an obligor denotes the portion of the exposure to the obligor which is lost in case of an occurring default event. The $EAD \in (0,\infty)$ is a deterministic quantity. In case of the default of an obligor, the creditor does not necessarily lose its full exposure. Obligors can partly recover meaning that the creditor can receive a fraction of the notional value of the claim. This setting is measured by the loss given default ($LGD \in [0-\epsilon, 1+\epsilon]$). ϵ can take non zero values as for instance, cost can be made in retrieving the value of the claim. Throughout this work, we will assume $\epsilon = 0$. The LGD is usually modelled as a random variable describing the severity of losses in the default event. Typical values of its expectation range from 40% to 80% [21]. Lastly, a variable of risk to consider is the default correlation. This measure describes the degree to which the default risk of one firm depends on the default risk of another firm. It is characterized by the joint default probabilities of the obligors over the specified time horizon. Default dependence has been shown to have a strong impact on the tail of the loss distribution for a given portfolio.

Generally, credit risk is measured in expected loss (EL) and unexpected loss (UL). The expected loss is easily determined from the aforementioned variables and can easily be managed. Unexpected loss is more difficult to measure and quantify. Economic capital (EC) is held to cover unexpected losses. A more rigorous definition of these risk measures will be provided in the next sections.

2.3. General Model Setting

We specify a probability space, denoted as $(\Omega, \mathcal{F}, \mathbb{P})$, with sample space Ω, σ-algebra $\mathcal{F}$ and probability measure $\mathbb{P}$. This probability space is the domain of all random variables introduced in this paper. Furthermore, we introduce a portfolio consisting of N loans indexed by $i = 1, ..., N$. Moreover, we assume that the exposures have been aggregated in such a way that there is a unique obligor for each position. Therefore, the amount of obligors equals the number of positions. The share of total portfolio exposure for obligor i is then defined by:

$$w_i = \frac{EAD_i}{\sum_{j=1}^{N} EAD_j} \tag{2.1}$$

Clearly, we have $\sum_{i=1}^{N} w_i = 1$. Furthermore, denote D_i as the default indicator of obligor i . At $t = 0$, all obligors are assumed to be in a non-default state. Since at $t = T$, obligor i defaults or does not default, D_i is represented as a Bernoulli random variable taking the values:

$$D_i = \begin{cases} 1 & \text{if counterparty i defaults before time T} \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

with probabilities $\mathbb{P}(D_i = 1) = PD_i$ and $\mathbb{P}(D_i = 0) = 1 - PD_i$. The loss of a single obligor is the random variable L_i:

$$L_i = EAD_i \cdot LGD_i \cdot D_i \tag{2.3}$$

Similarly, the absolute loss of the whole portfolio $L = L_N$ is calculated as the sum over all individual losses in the portfolio:

$$L_N = \sum_{i=1}^{N} EAD_i \cdot LGD_i \cdot D_i \tag{2.4}$$

Throughout this text, the subscript N is regularly omitted when there is no ambiguity about the portfolio size. Furthermore, the following assumption holds:

Assumption 2.3.1. *The exposure at default EAD_i, the loss given default LGD_i and default indicator D_i for all $i = 1, ..., n$ are independent.*

Therefore, the expected loss of loan i is given by:

$$\mathbb{E}[L_i] = \mathbb{E}[EAD_i \cdot LGD_i \cdot D_i] = EAD_i \cdot ELGD_i \cdot \mathbb{E}[D_i] = EAD_i \cdot ELGD_i \cdot PD_i \tag{2.5}$$

where $ELGD := \mathbb{E}[LGD]$. The same analysis can be conducted on the expected loss of the full portfolio resulting in:

$$\mathbb{E}[L_N] = \sum_{i=1}^{N} \mathbb{E}[L_i] = \sum_{i=1}^{N} EAD_i \cdot ELGD_i \cdot PD_i \tag{2.6}$$

2.3.1. Measuring Credit Risk

Essentially, a credit risk model is a function mapping from a set of instrument-level characteristics and market-level parameters to a distribution of portfolio credit losses over some chosen time horizon, often chosen to be 1 year [9]. In order to grasp the risk presented by the portfolio loss distribution, some summary statistic of this distribution is required. Currently, the preferred statistics for this are the Value-at-Risk (VaR) and economic capital (EC).

Value-at-Risk

Value-at-Risk is the most widely used risk measure in financial institutions due to its interpretability and its presence in the Basel II framework for measuring credit risk. The following derivation of the VaR and other risk measures is largely based on [25]. Let $F_L(x) = \mathbb{P}(L \leq x)$ be the distribution of the loss variable, the distribution that is of interest in measuring credit risk. As mentioned before, for large portfolios this distribution is expected to be highly skewed and heavy tailed. The goal is to define a statistic based on F_L that measures the risk of holding the portfolio. Since the support of F_L is unbounded in all models we will consider, the maximum loss is unbounded and therefore not a sufficient statistic. The VaR, however, describes the maximum possible loss which is not exceeded in a given time period at a set confidence level:

Definition 2.3.1. *Consider some confidence level $q \in (0, 1)$. The Value-at-Risk (VaR) of a portfolio at the confidence level q is given by the smallest number x such that the probability that the loss L exceeds x is not larger than $(1 - q)$:*

$$VaR_q(L) = \inf\{x \in \mathbb{R} : \mathbb{P}(L > x) \leq 1 - q\} = \inf\{x \in \mathbb{R} : F_L(x) \geq q\} \tag{2.7}$$

In probabilistic terms, the VaR equals the quantile of the loss distribution. Generally, the VaR can be derived over different holding periods and at all confidence levels. However, typically the holding period equals one year and the confidence level either equals 95%, 99% or 99.9%. Clearly, a higher confidence level leads to a higher VaR. Throughout this text, the following abbreviation for the VaR will be widely used: $VaR_q(L) := \alpha_q(L)$. Considering a continuous and strictly increasing distribution function $F_L(x)$, we have $\alpha_q(L) = F_L^{-1}(q)$, where $F_L^{-1}(q)$ is the quantile function.

The VaR has some major and widely researched drawbacks. Firstly, by its very definition, the VaR at confidence level q does not contain any information about the severity of losses that occur with a probability less than $1-q$. This shortcoming is addressed by the Expected Shortfall (ES). Secondly, the VaR is a non-coherent risk measure as it is not subadditive [1]. This means that the VaR of a merged portfolio of two individual portfolios is not necessarily bounded from above by the sum of the VaR's of the individual portfolios contradicting the intuition of diversification benefits [25].

Expected Shortfall
The Expected Shortfall (ES) is closely related to the VaR since it overcomes the deficiencies of the VaR. The ES averages the VaR over all confidence levels exceeding q, therefore taking the tail of the loss distribution into account. Formally:

Definition 2.3.2. *For a loss L with $\mathbb{E}[|L|], \infty$ and distribution function F_L, the Expected Shortfall (ES) at confidence level $q \in (0,1)$ is defines as*

$$ES_q = \frac{1}{1-q}\int_q^1 VaR_u(L)du \tag{2.8}$$

Clearly, from the definition we have $ES_q \geq VaR_q$. For continuous loss distributions a more intuitive express can be derived:

Proposition 2.3.1. *For an integrable loss variable L with continuous distribution function F_L and any $q \in (0,1)$ we have:*

$$ES_q = \frac{\mathbb{E}[L\mathbb{1}_{L\geq VaR_q(L)}]}{1-q} = \mathbb{E}[L|L \geq VaR_q(L)] \tag{2.9}$$

For the proof of this proposition, see [25] page 45. Proposition 2.3.1 allows for a more intuitive expression of the ES. The ES can be interpreted as the expected loss that is incurred in the event that the VaR is exceeded.

Economic Capital
Economic capital is a measure of unexpected loss for a given portfolio:

Definition 2.3.3. *The Economic Capital (EC) at confidence level $q \in (0,1)$ is defined as the difference between the VaR and the expected loss EL of the portfolio:*

$$EC_q = VaR_q(L) - \mathbb{E}[L] \tag{2.10}$$

For instance, assuming a one year time horizon, at a confidence level of 99.9% the EC can be interpreted as the appropriate capital to cover unexpected losses in 999 out of a 1000 years. Therefore, the EC represents the capital a financial institution should hold to limit the probability of default to a given confidence level. Clearly, with the VaR being portfolio dependent, the EC is portfolio dependent, whereas the EL is independent of the portfolio. Therefore, allocating EC to individual obligors is not as straightforward. For example, the EC charge for a new loan that is added to a well-diversified portfolio is much lower than the EC charge when the same loan is added to a heavily concentrated or very small portfolio. The problem of allocating the VaR and thus the EC to individual obligors is discussed in Section 3.5.1.

2.4. Merton's Structural Model of Default

So far, we have evaluated the general credit risk modeling setting and reviewed the measures of credit risk. To fully define the credit risk model, assumptions have to be made on the default scenario. Throughout this section, we will briefly review Merton's structural model of default. The model proposed by [26] forms the essential basis of all asset value based models and although many models have been developed since 1974, it remains influential and popular in current credit risk modelling practices [21]. The following derivation of the multi-factor model is based on [21] and [25]. Firstly, consider a firm whose asset value is described by a stochastic process V_t. The firm is financed by only two classes of securities, equity S_t and debt B_t. In Merton's model, the firm's debt is given by a zero coupon bond with face value B and maturity T. Furthermore, the model assumes a frictionless market, meaning that there are no taxes or transactions costs involved. The firm defaults if the value of its assets is less than the obliged debt repayment at time T, and default can only occur at maturity T of the bond. The value of the firm is then given by:

$$V_t = S_t + B_t,\ 0 \leq t \leq T \tag{2.11}$$

At maturity, two possible scenarios occur:

1. $V_T > B$ The value of the firms asset's exceeds its debt, so the debtholders receive B and the shareholders receive the residual value $S_T = V_T - B$. No default occurs.
2. $V_T \leq B$ The value of the firm's assets are less than its debt. Hence, the firm cannot meet its obligations and defaults. The debtholders take ownership of the firm $B = V_T$ and the shareholders receive nothing $S_T = 0$.

The Merton model assumes that under the real-word probability measure $\mathbb{P}$ the asset value process $(V_t)_{t\geq 0}$ follows a geometric Brownian motion with drift of the form:

$$dV_t = \mu_V V_t dt + \sigma_V V_t dW_t \tag{2.12}$$

for constants $\mu_V \in \mathbb{R}$, $\sigma_V > 0$ and standard Brownian motion $(W_t)_{t\geq 0}$. The solution at time T of stochastic differential Equation (2.12) is given by:

$$V_T = V_0 e^{((\mu_V - 1/2\sigma_V^2)T + \sigma_V W_T)} \tag{2.13}$$

which implies that $ln(V_T) \sim \mathcal{N}(ln(V_0) + (\mu_V - 1/2\sigma_V^2)T, \sigma_V^2 T)$. In essence, this means that under the dynamics of (2.12), the default probability of the firm is given by:

$$\mathbb{P}(V_T \leq B) = \mathbb{P}(ln(V_T) \leq ln(B)) = \Phi\left(\frac{ln(\frac{B}{V_0}) - (\mu_V - 1/2\sigma_V^2)T}{\sigma_V\sqrt{T}}\right) \tag{2.14}$$

Essentially , Equation (2.14) justifies the methods we will discuss in Section 2.5 in which we explain the more general idea of default occurring for a firm when some critical random variable falls below some deterministic threshold at the end of time period $[0, T]$. In the Merton model, this critical random variable is a lognormally distributed asset value and the default threshold is represented by the firms liabilities.

2.5. Gaussian Factor Threshold Models

In general, factor models are models in which the asset returns (or any other measure of counterparty's well-being) of a counterparty are modelled as a combination of systematic factors and an idiosyncratic factor. A major upside of these models is their economic interpretation. Factor models provide the possibility to interpret the asset returns, and correlations among asset returns, in terms of underlying economic variables. The following derivation of the multi-factor model is based on [21], [13] and [25]. In what follows, focus is on default-only models, meaning that a decision of a counterparty's state is made by comparing the counterparty's asset value to a threshold value. If, at the end of a set period, the counterparty's asset value falls below the threshold, the counterparty defaults. Furthermore, we consider a portfolio of N exposures. Data is aggregated in such a way that each obligor i has one loan with principal EAD_i. Additionally, fix the time horizon $T > 0$ and define r_i to be the asset return of obligor i. We will make the following assumption:

Assumption 2.5.1. *The asset returns r_i depend linearly on K standard normally distributed risk factors $X = (X_1, ..., X_K)$ affecting the counterparty's returns in a systematic way as well as on a standard normally distributed idiosyncratic term ϵ_i. Moreover, the ϵ_i's are independent, uncorrelated and independent from the system factors X_k for every $k \in \{1, ..., K\}$.*

Under this assumption, the asset returns are driven by two different types of risk factors, systematic and idiosyncratic risk factors:

Definition 2.5.1. ***Systematic risk*** *represents the effect of unexpected changes in macroeconomic and financial market conditions on the performance of borrowers. Borrowers may differ in their degree of sensitivity to systematic risk, but few firms are completely indifferent to the wider economic conditions in which they operate. Therefore, the systematic component of portfolio risk is unavoidable and only partly diversifiable [30].*

Definition 2.5.2. ***Idiosyncratic risk*** *represents the effects of risks that are particular to individual borrowers. As a portfolio becomes more fine-grained, in the sense that the largest individual exposures account for a smaller share of total portfolio exposure, idiosyncratic risk is diversified away at the portfolio level. This risk is totally eliminated in an infinitely granular portfolio [30].*

Under assumption 2.5.1, the counterparty's standardized asset returns are described by:

$$r_i = \beta_i Y_i + \sqrt{1-\beta_i^2}\epsilon_i \tag{2.15}$$

in which Y_i is the counterparty's composite factor:

$$Y_i = \sum_{k=1}^{K} \alpha_{i,k} X_k \tag{2.16}$$

Generally, systematic factors X_k represent geographical and industrial effects on the asset returns and factor loading $\alpha_{i,k}$. Furthermore, consider the following decomposition of r_i:

$$\mathbb{V}[r_i] = \beta_i^2 \mathbb{V}[Y_i] + (1-\beta_i^2)\mathbb{V}[\epsilon_i] \tag{2.17}$$

Clearly, Equation (2.17) can be interpreted as splitting the total risk of a counterparty's asset returns into the systematic and idiosyncratic components. β_i^2 expresses the amount of variance of r_i that can be attributed to the systematic risk. Factor $1-\beta_i^2$ captures the contribution of idiosyncratic risk to the total variance. As r_i is standardized, the following condition should be satisfied: $\sum_{n=1}^{K} \alpha_{n,k}^2 = 1$. PD_i denotes the one-year probability of default for counterparty i: $PD_i = \mathbb{P}[r_i < c_i]$ and since we impose that $r_i \sim \mathcal{N}(0,1)$ we obtain:

$$c_i = \Phi^{-1}[PD_i] \tag{2.18}$$

Rewriting the condition $r_i < c_i$ yields:

$$\epsilon_i < \frac{\Phi^{-1}(PD_i) - \beta_i Y_i}{\sqrt{1-\beta_i^2}} \tag{2.19}$$

Therefore, since ϵ_i is normally distributed, the probability of default conditional on the systematic factors can be written as:

$$PD_i(Y_i) = \mathbb{P}[r_i < c_i | Y_i] = \mathbb{P}\left[\epsilon_i < \frac{\Phi^{-1}(PD_i) - \beta_i Y_i}{\sqrt{1-\beta_i^2}} | Y_i\right] = \Phi\left(\frac{\Phi^{-1}(PD_i) - \beta_i Y_i}{\sqrt{1-\beta_i^2}}\right) \tag{2.20}$$

An interesting computation is that of the expected value of the conditional default probability:

$$\mathbb{E}[PD_i(Y_i)] = \mathbb{E}[\mathbb{E}[\mathbb{1}_{r_i<\Phi^{-1}(PD_i)}|Y_i]] = \mathbb{E}[\mathbb{1}_{r_i<\Phi^{-1}(PD_i)}] = \mathbb{P}(r_i < \Phi^{-1}(PD_i)) = \Phi(\Phi^{-1}(PD_i)) = PD_i \tag{2.21}$$

This means, that when we compute the average conditional default probability, the unconditional default probability is recovered. The Gaussian threshold models share this characteristic with other model approaches such as CreditRisk+.

A major upside of this model is the ability to describe the asset correlation between obligors through their common systematic factors. Denote $\mathbf{\Omega} \in \mathbb{R}^{K\times K}$ to be the correlation matrix of the K systematic factors X_k. Then $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Omega})$. Again, the assumption that $\mathbb{V}[r_i] = 1$ means that $\boldsymbol{\alpha}_i'\mathbf{\Omega}\boldsymbol{\alpha}_i = 1 \; \forall \; i \in [1, N]$. Since $\mathbb{V}[X_k] = 1$ and through the independence of X_k and ϵ_i the asset correlations are given by:

$$\rho(r_i, r_j) = cov(r_i, r_j) = \beta_i \beta_j \boldsymbol{\alpha}_i'\mathbf{\Omega}\boldsymbol{\alpha}_j \tag{2.22}$$

In using the model, the factor loadings $\alpha_{i,k}$ and systematic risk component β_i have to be determined for each obligor and correlation matrix $\mathbf{\Omega}$ has to be determined. Usually, these factors are determined through regression models on historical data.

2.6. Measuring Credit Risk with the Single Factor Model

In the previous section, we have derived the Multi Factor Gaussian Threshold model. Multi factor indicates the existence of multiple systematic factors, Gaussian threshold indicates the fact that the asset returns are assumed to be standard Normally distributed. For many practical implications, of which the IRB approach is one, this model is simplified to contain only one single systematic factor, in other words, $K = 1$. Throughout this section, we will explore this model in more detail.

2.6.1. Asymptotic Single Risk Factor Approximation

The Asymptotic Single Risk Factor (ASRF) model, developed by [9], is based on several major assumptions. We will briefly consider these assumptions as they are relevant to the derivation of the granularity adjustment in later sections.

Assumption 2.6.1. *Assume that the variables $U_i \equiv LGD_i D_i$ for $i = 1, ..., N$ are bounded in the interval $[-1, 1]$ and conditional on Y, are mutually independent.*

Assumption 2.6.2. *Let EAD_i be an increasing sequence of postive constants. Assume that*

1. $\sum_{i=1}^{N} EAD_i \uparrow \infty$ *as* $N \to \infty$
2. *There exists a* $\xi > 0$ *such that* $\frac{EAD_N}{\sum_{i=1}^{N} EAD_i} = \mathcal{O}(N^{-(\frac{1}{2}+\xi)})$

Assumption 2.6.2 guarantees that the share of the largest single exposure with respect to the total portfolio exposure vanishes to zero as the number of exposures in the portfolio increases. These assumptions are vital for guaranteeing that the idiosyncratic risk vanishes as more assets are added to the portfolio. In practice, these assumptions are quite weak and are likely to be satisfied by real-world large bank portfolios. Using these assumptions, we can state an important result:

Proposition 2.6.1. *Under assumption 2.6.1 and 2.6.2, conditional on $x = X$, $L_N - \mathbb{E}[L_N|x] \to 0$ almost surely.*

Where L_N denotes the portfolio loss ratio $L_N = \sum_{i=1}^{N} w_i LGD_i D_i$. For the proof of this proposition, relying on the law of large numbers, we refer to appendix A.1. This is the main result on which the ASRF model is built as it enables one to approximate the true loss distribution by the expected value of the conditional loss distribution. Proposition 2.6.1 says that the larger the portfolio is, the more idiosyncratic risk is diversified away and in the limit, the portfolio is driven by systematic risk purely. Regularly in literature, this limiting portfolio is referred to as the infinitely fine-grained portfolio or the asymptotic portfolio [21]. If we now assume the following:

Assumption 2.6.3. *The systematic risk factor X is one dimensional*

Assumption 2.6.4. *There is an open interval B containing $\alpha_q(X)$ and a real number $n_0 < \infty$ such that:*

1. $\forall i \in 1, ..., N$, $\mathbb{E}[U_i|x]$ *in continuous in x on B,*
2. $\mathbb{E}[L_N|x]$ *is monotonously decreasing in x on B for all $N > n_0$,*
3. *for all $N > n_0$,* $\inf_{x \in B} \mathbb{E}[L_N|x] \geq \sup_{x \leq \inf B} \mathbb{E}[L_N|x]$ *and* $\sup_{x \in B} \mathbb{E}[L_N|x] \leq \inf_{x \geq \sup B} \mathbb{E}[L_N|x]$

In short, assumption 2.6.3 imposes a single systematic risk factor as the source of dependence across all obligors in the portfolio. Assumption 2.6.4 essentially guarantees that the neighbourhood of the qth quantile of $\mathbb{E}[L_N|x]$ is associated with neighbourhood of the unique qth quantile of X. This leads us to the following proposition:

Proposition 2.6.2. *Under assumption 2.6.4 and 2.6.3, for $N > n_0$ we have* $\alpha_q(\mathbb{E}[L_N|x]) = \mathbb{E}[L_N|\alpha_{1-q}(X)]$

For the proof of this proposition we refer to appendix 2.8.9 of [13]. Essentially, this relation enables us to calculate the value at risk $\alpha_q(\mathbb{E}[L_N|x])$, which in general is highly complicated, by the exposure-weighted average of the individual assets' conditional expected losses $\mathbb{E}[L_N|\alpha_{1-q}(X)]$. This relation constitutes the core of Basel's regulatory capital formulas. We can now apply the expression derived in Section 2.5. Firstly, we condense the multi-factor model to a single factor model in line with assumption 2.6.3:

$$r_i = \beta_i X + \sqrt{1-\beta_i^2}\epsilon_i \text{ for } i = 1,...,N \tag{2.23}$$

All other assumptions from Section 2.5, such as the normality of X and ϵ_i, also hold in this setting. Therefore, from Equation (2.20) we have:

$$PD_i(\alpha_{1-q}(X)) = \Phi\left(\frac{\Phi^{-1}(PD_i) - \beta_i\Phi^{-1}(1-q)}{\sqrt{1-\beta_i^2}}\right) \tag{2.24}$$

And therefore, when applying proposition 2.6.2 together with the previously derived equation we have in the ASRF setting:

$$VaR_q^{ASRF} = \sum_{i=1}^{N} w_i \cdot LGD_i \cdot \Phi\left(\frac{\Phi^{-1}(PD_i) + \beta_i\Phi^{-1}(q)}{\sqrt{1-\beta_i^2}}\right) \tag{2.25}$$

where we used the symmetry of Φ^{-1}. This result enables one to estimate the value at risk of the loss distribution of a portfolio of N obligors in a single systematic factor setting.

2.6.2. Measuring Credit Risk with the Basel II-III IRB Approach

The Basel IRB approach is set up to be a portfolio invariant framework for assessing credit risk. Therefore, the sum of the risk contributions of the individual facilities in the portfolio defines the measure of risk for the full portfolio. Generally, the risk contribution of an obligor is calculated in the following way within the A-IRB framework:

$$RC_i^{Basel} = EAD_i LGD_i\left[\Phi\left(\frac{\Phi^{-1}(PD_i) + \sqrt{\rho_i}\Phi^{-1}(0.999)}{\sqrt{1-\rho_i}}\right) - PD_i\right] \cdot \frac{1 + b(M_i - 2.5)}{1 - 1.5b} \cdot 1.06 \tag{2.26}$$

with $b = [0.11852 - 0.05478 \cdot ln(PD_i)]^2$ and

$$\rho_i = \Lambda \cdot \left[0.12 \cdot \frac{1 - e^{-50PD_i}}{1 - e^{-50}} + 0.24 \cdot \left(1 - \frac{1 - e^{-50PD_i}}{1 - e^{-50}}\right)\right] \tag{2.27}$$

The total reculatory capital for the full portfolio is then easily determined through $RC = \sum_{i=1}^{N} RC_i^{Basel}$. Taking a closer look at Equation (2.26) and comparing it with Equation (2.24), we can recognize the following structure:

$$RC_i^{Basel} = EAD_i LGD_i\,[PD_i(\alpha_{0.999}(X)) - PD_i] \cdot MA \cdot 1.06 \tag{2.28}$$

With

$$MA = \frac{1 + b(M_i - 2.5)}{1 - 1.5b} \tag{2.29}$$

denoting the Maturity Adjustment. From (2.28) we can notice that, similarly to economic capital, regulatory capital is expressed as a scaled difference between the conditional expected loss on a extreme event and the unconditional expected loss. Additionally, the A-IRB approach shows great similarities with the ASRF model and therefore clearly ignores the effect of idiosyncratic risk. An important detail of the A-IRB approach is the choice of asset correlation (2.27). By construction, the correlation coefficient lies in the interval $\rho \in [0.12, 0.24]$, given $\Lambda = 1$, and is a decreasing function of the unconditional default probability PD_i. Basel's logic behind this choice is that if a counterparty has a large default probability, then much of its inherent default risk is idiosyncratic and it is less highly correlated with other credit counterparties. Conversely, low risk counterparties predominantly face systematic risk and thus exhibit higher levels of asset correlation [2]. The asset correlation attains its maximum value below a default probability of around 3-4% and for default probabilities exceeding 10%, the correlation function essentially equals its lower bound of 0.12. The parameter Λ is a new addition of Basel III and equals 1.25 if firm i is a large financial institution and equals 1 otherwise.

The maturity adjustment (2.29) reflects the notion that long term exposures are riskier than their shorter-term equivalents. This means that regulatory risk capital contributions increase for longer-term maturities. Additionally, through b, the maturity adjustment is a function of default probability. The impact of the maturity adjustment is a decreasing function of its unconditional default probability.

2.6.3. Limitations of the A-IRB approach

The IRB approach to measuring credit risk was developed as a portfolio invariant measure [30]. In essence, this means that the capital required for any set loan depends on the characteristics of that specific loan only, and not on the portfolio the loan is added to. The upside to this approach is that it makes the IRB framework widely applicable across countries and institutions. Furthermore, it allows for a straightforward comparison of riskiness across individual loans [29].

In order to achieve portfolio invariance, the IRB approach builds on two key assumptions:

1. Bank portfolios are infinitely granular
2. There exists only one source of systematic risk

The first assumption implies that all idiosyncratic risk is diversified away. The second assumption implies that there are no diversification possibilities beyond the reduction of idiosyncratic risk. A more informal interpretation of the second assumption is that the banks' portfolio is well diversified across geographical regions and sectors meaning that the only remaining systematic risk is the performance of the global economy [30]. As the IRB approach assumes idiosyncratic risk to be fully diversified away, one only needs to assess the systematic component of risk which results in the IRB formulas for assessing VaR. Being portfolio invariant, the contributing VaR's for each exposure are simply added up to provide the VaR assessment for the full portfolio. Therefore, the IRB approach does not allow for an elaborate correlation structure between individual risks [29]. However, when these assumptions are violated, there is no guarantee that the IRB approach will be accurate. Most likely, the marginal VaR contributions of single exposures to the overall VaR depends on the risk profile of the full portfolio. In general, the IRB-based capital requirements may either under- or overestimate the risk profile of the portfolio depending on the level of diversification and systematic correlation.

2.7. Concentration Risk

Basel defines a risk concentration as "any single exposure or group of exposures with the potential to produce losses large enough (relative to a bank's capital, total assets, or overall risk level) to threaten a bank's health or ability to maintain its core operations"[31]. This definition refers to concentration risk across multiple sorts of risk such as credit risk, market risk, liquidity risk and operational risk. Throughout this work, we will focus on solely on credit risk. Within credit risk, two kinds of concentration risk can be identified: name concentration and sector concentration.

Definition 2.7.1. *Name concentration risk is the residual idiosyncratic risk arising from the deviation from the infinitely fine-grained ideal [2]. Name concentration risk can be split into two types:*

- *Individual name concentration risk refers to the type of concentration risk the results from an exposure to one firm or to a conglomerate of firms that is extremely large compared to the rest of the exposures of the portfolio [13].*
- *Portfolio name concentration risk refers to the risk that occurs if a banks holds a portfolio containing a relatively small amount of firms, each of them with large exposures [13].*

Definition 2.7.2. *Sector concentration risk stems from the existence of multiple systematic factors and arises from the assumption of a single underlying risk factor [2].*

For the first type of name concentration risk, the risk of the portfolio is significantly driven by the idiosyncratic risk of a set of large exposure obligors. The second type of name concentration risk is driven by a lack of diversification. The bank faces high losses if several accidentally occur and are not driven by default correlation of the obligors. Sector concentration can be both geographically or industry-based.

2.8. Internal Capital at ING

This section is omitted on purpose.

II

State of the art

3

Tools for identifying concentration risk

This chapter addresses the mathematical tools needed to assess concentration risk in credit portfolios. General mathematical tools such as the multivariate normal distribution and Pearson's correlation will shortly be studied. Additionally, the concept of recovery risk will be introduced. The problem of capital allocation is treated in this chapter through the introduction of the Euler allocation principle in both an analytical and Monte Carlo setting. Ad hoc measures of concentration, such as the Gini index will be reviewed after which model-based concentration measures, such as the Granularity Adjustment, will be derived.

3.1. Mathematical Tools

3.1.1. Multivariate Statistics

Credit risk models are inherently multivariate as the loss incurred by a credit portfolio depends on a random vector of losses for individual counterparties in the portfolio [25]. In this section, we will briefly touch upon the relevant statistical tools and multivariate distributions that are relevant to credit risk modelling. We assume the reader to be familiar with statistical tools such as the notions of expectation, variance and some basic distributions such as the normal distribution. Furthermore, we assume a basic knowledge of linear algebra. In this section, we follow the standard notation and general concepts from [25].

Multivariate Normal Distribution

A d-dimensional vector of random variables $\mathbf{X} = (X_1, ..., X_d)'$ has a multivariate normal (or Gaussian) distribution if

$$\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \mathbf{AZ} \tag{3.1}$$

where $\mathbf{Z} = (Z_1, ..., Z_k)'$ is a k-dimensional vector of i.i.d. univariate standard normal random variables, and $\mathbf{A} \in \mathbb{R}^{d \times k}$ and $\boldsymbol{\mu} \in \mathbb{R}^d$ are a matrix and vector of constants respectively. The mean of this vector of random variables is $\mathbb{E}[\mathbf{X}] = \boldsymbol{\mu}$ and its covariance matrix is given by $\mathrm{Cov}(\mathbf{X}) = \boldsymbol{\Sigma}$ where $\boldsymbol{\Sigma} = \mathbf{A}^\intercal\mathbf{A}$. The distribution if characterized by its mean vector and covariance matrix, therefore we use the following standard notation throughout this work $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Moreover, if $\boldsymbol{\Sigma}$ has full rank d and is therefore invertible and positive semi-definite, $\mathbf{X}$ has an absolutely continuous distribution function with joint density [25]:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \tag{3.2}$$

where $|\boldsymbol{\Sigma}|$ denotes the determinant of $\boldsymbol{\Sigma}$.

Linear combinations of multivariate normal random vectors, remain multivariate normal. This property will be widely applied throughout this text. Let $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and take $B \in \mathbb{R}^{k \times d}$ and $\mathbf{b} \in \mathbb{R}^k$, then:

$$B\mathbf{X} + \mathbf{b} \sim \mathcal{N}(B\boldsymbol{\mu} + \mathbf{b}, B\boldsymbol{\Sigma}B^\intercal) \tag{3.3}$$

and in the case that $\mathbf{b} \in \mathbb{R}^d$

$$\mathbf{a}^\intercal \mathbf{X} \sim \mathcal{N}(\mathbf{a}^\intercal \mu, \mathbf{a}^\intercal \mathbf{\Sigma} \mathbf{a}) \tag{3.4}$$

Student-*t* Distribution

Besides the normal distribution, another widely used distribution in the practice of portfolio credit risk modelling is the student-*t* distribution, often named the *t*-distribution. Firstly, we introduce the notion of normal variance mixtures:

Definition 3.1.1. *The random vector* $\boldsymbol{X}$ *is said to have a multivariate normal variance mixture distribution if*

$$\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + \sqrt{C}\boldsymbol{A}\boldsymbol{Z} \tag{3.5}$$

where

1. $\boldsymbol{Z} \sim \mathcal{N}(\mathbf{0}, I_k)$
2. $C \geq 0$ *is a non-negative, scalar valued random variable independent of* $\boldsymbol{Z}$
3. $\boldsymbol{A} \in \mathbb{R}^{d \times k}$ *and* $\boldsymbol{\mu} \in \mathbb{R}^d$ *are a matrix and vector of constants respectively.*

Clearly, the resulting mixture **X** is itself not a multivariate normal distribution. However, conditional on C, we can observe the following identity $\mathbf{X}|C = c \sim \mathcal{N}(\boldsymbol{\mu}, c\mathbf{\Sigma})$ with $\mathbf{\Sigma} = AA^\intercal$. [25] suggest interpreting the mixing variable C as a shock that arises from new information and that impacts the volatilities of all assets. Given that C has a finite expectation, the following holds:

$$\mathbb{E}[\mathbf{X}] = \mathbb{E}[\boldsymbol{\mu} + \sqrt{C}\mathbf{A}\mathbf{Z}] = \boldsymbol{\mu} + \mathbb{E}[\sqrt{C}]\mathbf{A}\mathbb{E}[\mathbf{Z}] = \boldsymbol{\mu} \tag{3.6}$$

$$\mathrm{Cov}(\mathbf{X}) = \mathbb{E}[(\sqrt{C}\mathbf{A}\mathbf{Z})(\sqrt{C}\mathbf{A}\mathbf{Z})^\intercal] = \mathbb{E}[C]\mathbf{A}\mathbf{A}^\intercal\mathbb{E}[\mathbf{Z}\mathbf{Z}^\intercal] = \mathbf{\Sigma}\mathbb{E}[C] \tag{3.7}$$

An important note is that $\mathbf{\Sigma}$ is the covariance matrix of $A\mathbf{Z}$ and only the covariance matrix of **X** in the case that $\mathbb{E}[C] = 1$. Generally, $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ are referred to as the location vector and the dispersion matrix of normal mixture distribution. The correlation matrices of **X** and **Z** are the same, given the finite expectation of C.

The multivariate *t* distribution is retrieved from the normal variance mixture distribution if we take C to be a random variable in the form of $\sqrt{\frac{\eta}{W}}$ where $W \sim \chi^2(\eta)$, and η the degrees of freedom of the chi-squared distribution. Then, **X** has a multivariate *t* distribution with η degrees of freedom. Our notation for the multivariate *t* is $\mathbf{X} \sim t_d(\eta, \boldsymbol{\mu}, \mathbf{\Sigma})$. The density of the multivariate *t*-distribution is given by:

$$f(\mathbf{x}) = \frac{\Gamma(\frac{1}{2}(\eta + d))}{\Gamma(\frac{1}{2}\eta)(\eta\pi)^{d/2}|\mathbf{\Sigma}|^{1/2}}\left(1 + \frac{(\mathbf{x}-\boldsymbol{\mu})^\intercal\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{\eta}\right)^{-(\eta+d)/2} \tag{3.8}$$

In Figure 3.1 the univariate *t*-distribution is depicted for varying levels of degree of freedom parameter η. In comparison with the standard normal univariate distribution, the *t*-distribution has heavier tails and has a higher tendency to generate extreme values. Furthermore, the *t*-distribution is clearly bell shaped and symmetrical and as the degrees of freedom increase, the distribution approaches the standard normal distribution.

3.1.2. Monte Carlo Methods

Throughout this work, we will make generous use of the Monte-Carlo methods. MC methods have been of great importance for credit risk modelling and well known industry models such as Moody's KMV and CreditMetrics rely heavily on them [40][19]. This section will briefly address the foundations of the Monte-Carlo method. Throughout this section we will refer to the work of [2] and [7]. Firstly, introduce the integral of an arbitrary function $g(x)$:

$$\gamma = \int_0^1 g(x)dx \tag{3.9}$$

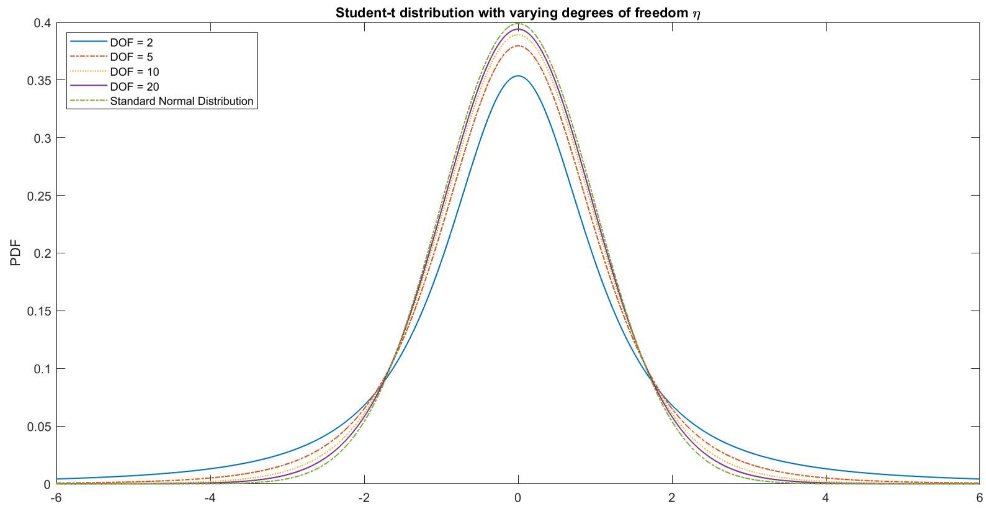

Figure 3.1: Example of the impact of degrees of freedom parameter η on the shape t-distribution. As η increases, the t-distribution approaches the standard normal distribution.

We may represent this integral as an expectation $\mathbb{E}[g(U)]$[1] with U uniformly distributed between 0 and 1. Suppose we draw $U_1, U_2, \ldots$ independently and uniformly from $[0, 1]$. The Monte Carlo estimate $\hat{\gamma}_N$ is given by evaluating the function f at N random points and then averaging the result:

$$\hat{\gamma}_N = \frac{1}{N}\sum_{i=1}^{N} g(U_i) \tag{3.10}$$

If g is integrable over $[0, 1]$, then by the strong law of large numbers:

$$\hat{\gamma}_N \to \gamma \text{ almost surely as } N \to \infty \tag{3.11}$$

The error of the Monte-Carlo estimator goes to zero as $\sqrt{N}$ goes to infinity, otherwise known as $\mathcal{O}(\frac{1}{\sqrt{N}})$ convergence. More generally, consider a random variable X defined on probability space $(\Omega, \mathcal{F}, \mathbb{Q})$. Furthermore, we assume that X is absolutely continuous and has probability density function $f_X(x)$. For some measurable function g, the expectation of $g(X)$ is defined as:

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \tag{3.12}$$

Following the same procedure, we generate N independent realizations of X: $\{X_1, \ldots, X_N\}$. The integral can then be estimated by:

$$\mathbb{E}[\hat{g}(X)] = \frac{1}{N}\sum_{i=1}^{N} g(X_i) \tag{3.13}$$

Similarly, as M tends to infinity, solution (3.13) converges to its true value. However, the question remains how we can apply this method to credit risk problems. In essence, credit risk management is about identifying quantiles of the loss distribution, not necessarily about computing expectations of the form (3.13). The expression for losses is given as:

$$L = \sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot D_i \tag{3.14}$$

[1]Since if U is uniformly distributed on [0,1], $f_U(u) = 1$

And throughout this work, we are interested in the VaR, or the upper quantiles of the loss distribution:

$$\mathbb{P}(L \geq \alpha_q(L)) = \mathbb{P}\left(\sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot D_i \geq \alpha_q(L)\right) \tag{3.15}$$

Using the general fact that $\mathbb{P}(L \geq \alpha_q(L)) = \mathbb{E}[\mathbb{1}_{L \geq \alpha_q(L)}]$ we can transform Equation (3.15) into the following expectation:

$$\mathbb{P}(L \geq \alpha_q(L)) = \mathbb{E}\left[\mathbb{1}_{\{\sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot D_i \geq \alpha_q(L)\}}\right] \tag{3.16}$$

This places the VaR computation in a multidimensional framework of equation 3.12 and this implies we can employ the estimator (3.13). This justifies the use of Monte Carlo simulations in a portfolio credit risk setting. A remaining challenge is that we are considering extreme outcomes for $\alpha_q(L)$ and therefore, the simulation process is both relatively slow and computationally expensive. [8] suggested the use of variance reduction technique called importance sampling to solve this problem. The general idea of importance sampling is to sample in such a way that extreme events are more likely. For more information on this method we refer to [7] and [8].

3.1.3. Correlation

In the practice of portfolio credit risk, we are interested in the joint behaviour of individual exposures. In essence, the overall risk of a portfolio does not only depend on the risk of individual exposures, but also on the correlations between the future values of the counterparties in the portfolios. Therefore, we will shortly treat the notion of correlation in this section. Throughout this work, we will only consider Pearson correlation, which assesses linear relationships between two random variables. Alternatives for the Pearson correlation coefficient are for instance Spearman's rank correlation coefficient or Kendall's rank correlation coefficient. Whereas Pearson's coefficient is only a measure of linear relationships, Spearman's coefficient assesses monotonic relationships in general.

Generally, both the covariance and the Pearson correlation coefficient of two random variables are measures of how the two random variables vary jointly linearly [33]. A positive linear correlation indicates that the two random variables vary jointly in the same direction whereas a negative linear correlation indicates that the two random variables vary jointly in opposite direction. The Pearson correlation coefficient of two random variables X, Y, generally denoted as ρ, is given by:

$$\rho_{XY} = \frac{\text{Cov}(X,Y)}{\sqrt{\sigma_X^2 \sigma_Y^2}} = \frac{\mathbb{E}[(X-\mu_X)(Y-\mu_Y)]}{\sqrt{\sigma_X^2 \sigma_Y^2}} \tag{3.17}$$

Where μ and σ denote the expected value and the standard deviation of the individual random variables, respectively. Using this notation and the fact that asset returns are assumed to be standard Normal random variables, the asset return correlation between any two obligors for the single factor threshold model can easily be determined:

$$\rho_{r_i r_j} = \text{Cov}(r_i, r_j) = \beta_i \beta_j \text{Cov}(X, X) = \beta_i \beta_j \tag{3.18}$$

This immediately clarifies a major upside of factor models. Essentially, instead of defining the pairwise correlation for each counterparty in a portfolio, if suffices to define only the systematic factor loading β_i for each counterparty. It must be noted that the asset correlation is not equal to the default correlation. Default correlation is the degree of correlation between the default indicators for two counterparties, ρ_{D_i,D_j}. From Section 2.3 we already know that D_i is a Bernoulli random variable, using this knowledge we can easily determine ρ_{D_i,D_j}:

$$\rho_{D_i,D_j} = \frac{\text{Cov}(D_i, D_j)}{\sqrt{\mathbb{V}[D_i]\mathbb{V}[D_j]}} = \frac{\mathbb{E}[D_i, D_j] - PD_i PD_j}{\sqrt{(PD_i - PD_i^2)(PD_j - PD_j^2)}} \tag{3.19}$$

This leaves us with determining $\mathbb{E}[D_i, D_j]$:

$$\mathbb{E}[D_i, D_j] = \mathbb{P}(D_i = 1 \wedge D_j = 1) = \Phi_2(\Phi^{-1}(PD_i), \Phi^{-1}(PD_j), \rho_{r_i,r_j}) \tag{3.20}$$

Clearly, the default correlation depends on the joint distribution of r_i and r_j, which in our case is a multivariate Gaussian distribution.

3.2. Recovery Risk

Recovery risk denotes the risk that following a default, contracts of the defaulting obligor cannot be honoured in full, thereby leading to financial loss to the bank. Recovery risk is the complement of LGD Risk. Recovery risk is expressed through the distribution of the LGD parameter. Throughout this work, we will either model the LGD as a deterministic quantity or as a Beta distributed random variable. The Beta distribution is used in two well known industry models; Moody's RiskFrontier and CreditMetrics [6]. The Beta distribution is a viable candidate for LGD modeling as it is defined on the interval $(0,1)$ and therefore can be interpreted as returning a LGD. A random variable X has a Beta distribution, $X \sim Beta(a,b)$, if its density is:

$$f(x) = \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1} \tag{3.21}$$

With $x \in (0,1)$ and $a, b > 0$ and where:

$$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{3.22}$$

Where $\Gamma(\cdot)$ denotes the Gamma function. The parameters a, b are called the shape parameters of the Beta distribution. The mean and variance of the distribution are respectively:

$$\mathbb{E}[X] = \frac{a}{a+b} \tag{3.23}$$

$$\mathbb{V}[X] = \frac{1}{(a+b)^2}\frac{ab}{(a+b+1)} \tag{3.24}$$

However, instead of parameterizing the Beta distribution according to the shape parameters, we will reparameterize the distribution according to its expected value and variance. We model the random loss given default $\widetilde{LGD}$ as a Beta distributed random variable with mean LGD:

$$\mathbb{E}[\widetilde{LGD}] = LGD = \frac{a}{a+b} \tag{3.25}$$

$$\mathbb{V}[\widetilde{LGD}] = \frac{LGD(1-LGD)}{a+b+1} \tag{3.26}$$

We then define a second parameter, $k = a + b + 1$, we can fully describe the shape parameters a, b according to LGD and k:

$$a = (k-1)LGD \tag{3.27}$$

$$b = (k-1)(1-LGD) \tag{3.28}$$

Depending on the choice of parameters, the Beta distribution can be either bell-shaped, U shaped or J shaped. Empirically, historic losses reveal that the LGD occur at either 0% or 100% loss [4]. This indicates the choice of a low k parameters, for instance at a LGD of 0.5 a value of k below 4. This choice produces a LGD distribution with a relatively high density at these extremities. However, if a bank is very confident about their LGD estimates, a higher parameter k is set in order to reduce the variance around the set LGD mean.

3.3. Industry Threshold Models: Moody's RiskFrontier

Throughout this section, we will briefly explore a widely applied industry model; Moody's RiskFrontier. Moody's RiskFrontier, formerly known as Moody's KMV, is the economic capital solution by Moody's Analytics. The model deployed by RiskFrontier is in essence a general multi factor model, although in a somewhat different representation compared to the general multi factor model (2.15) [14]:

$$r_i = \sqrt{RSQ_i}\phi_i + \sqrt{1 - RSQ_i}\epsilon_i \tag{3.29}$$

Where $\phi_i, \epsilon_i \sim \mathcal{N}(0,1)$ and ϕ_i denotes the systematic factor. RSQ_i denotes the measure of an obligor's exposure to systematic risk compared to idiosyncratic risk. By construction, we have that $r_i \sim \mathcal{N}(0,1)$.

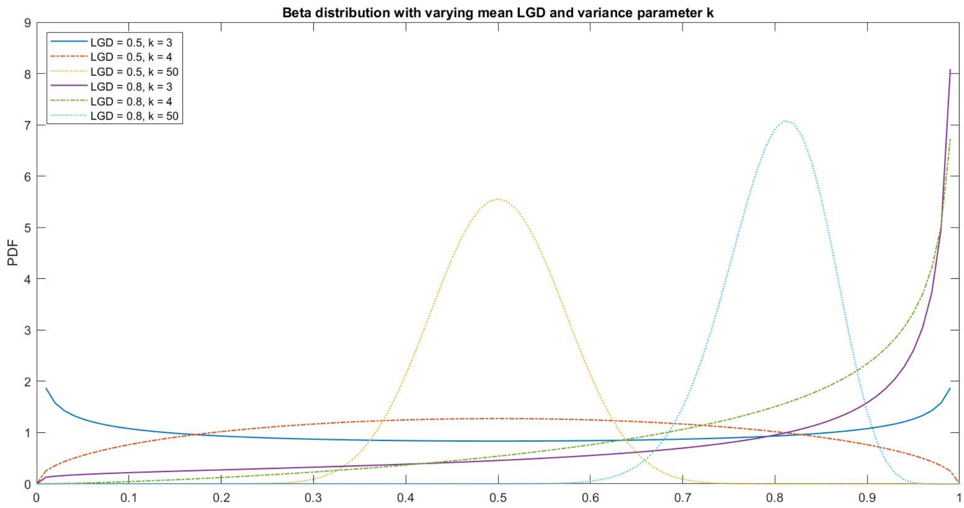

Figure 3.2: Example of the impact of LGD and k on the shape and location of the Beta distribution. The distribution can take on multiple different shapes by adjusting LGD and k.

Clearly, by comparing (3.29) with (2.15) we have that $\beta_i = \sqrt{RSQ_i}$. This observation allows us to easily adapt existing literature on general Gaussian multi factor threshold models to the Moody's RiskFrontier framework. Clearly, by construction, the asset returns are standard normal random variables. The default threshold is defined exactly in line with the aforementioned general multi factor model. Furthermore, the EAD is assumed to be deterministic and the LGD is assumed to be Beta distributed, similarly to the methods described in Section 3.2.

Within the full RiskFrontier framework, the systematic factor ϕ_i is a weighted combination of 245 correlated geographical and sector risk factors where the weights are uniquely determined for each counterparty in the portfolio. In this thesis, we limit ourselves to corporate clients, which reduces the number of correlated risk factors to 110. Of these 110 risk factors, 49 are country-specific and 61 are industry-specific [14]. Clearly, these 110 risk factors are correlated through a 110×110 covariance matrix $\mathbf{\Omega}$. Whereas similar obligors can have an identical systematic factor, the idiosyncratic shocks ϵ_i are obligor specific. Similarly to the general multi factor model, the systematic and idiosyncratic factors are independent.

The values for RSQ_i and covariance matrix $\mathbf{\Omega}$ are a product of Moody's *GCorr Corporate* model. GCorr Corporate has access to historical weekly asset returns of firms around the world, dating back to July 1, 1999 [14]. The asset values are driven by equity prices combined with debt information such as the amount of debt, duration and interest rate level. To compute the RSQ_i and covariance matrix $\mathbf{\Omega}$, market-weighted asset return indexes are computed for each country and industry combination. These combinations are then decomposed into country and industry-specific return indices. Using these specific returns, the covariance matrix of the risk factors can be computed. Having estimated the 110 industry and country risk factors, RSQ_i can be determined by regressing the firm's asset returns on the returns of its systematic factor. When there is a lack of public data, such as for private firms, RSQ_i is modeled based on the size, country and industry weights of the firm. Throughout this work, both RSQ_i and $\mathbf{\Omega}$ are assumed to be given and therefore we do not elaborate on the mathematical methods of estimating these values. For more information regarding GCorr, we refer to [14].

3.4. CreditRisk+

Throughout this section, we will briefly study the CreditRisk+ model, first introduced by Credit Suisse Financial Products (CSFP) and is currently a widely used portfolio credit risk model in the financial service industry. In contrast to threshold models such as the single and multi factor Gaussian threshold model, in these models the conditional default probabilities are modelled directly. Generally, in threshold models, conditional probabilities of default are modelled indirectly through defining a stochastic process for the firms asset value which then indirectly leads to a default event based on the passing of asset values of some kind of pre-defined threshold [21]. The full-blown CreditRisk+ approach is a multi factor implementation, we will however derive the model in the one-factor setting. The derivation and explanation of the CreditRisk+ model is based on the official technical document [40] in combination with [2] and [25]. Generally, the CreditRisk+ model is built around the following assumption:

Assumption 3.4.1. *Assume that the random default probabilities are influenced by a common set of Gamma distributed systematic factors. Therefore, the default events are assumed to be mutually independent conditional on the realizations of the risk factors.*

Firstly, we introduce the notation of the default event for counterparty i:

$$D_i = \begin{cases} 1 & \text{if counterparty i defaults at time T} \\ 0 & \text{otherwise} \end{cases} \tag{3.30}$$

Basically, the CreditRisk+ model is a Poisson-gamma mixture model. Therefore, the default event for counterparty i is given by:

$$\mathbb{1}_{D_i} = \mathbb{1}_{X_i \geq 1} \tag{3.31}$$

Where $X_i \sim \mathcal{P}(p_i(S))$. Furthermore, to fully define the CreditRisk+ model we need an expression for the default probability of counterparty i conditional on the realization of the Gamma distributed systematic factor $S \sim \Gamma(a, b)$:

$$p_i(S) = PD_i(\omega_0 + \omega_1 S) \tag{3.32}$$

Where PD_i denotes the unconditional default probability of counterparty i, similar to its definition in the one and multi factor setting and where $\omega_0, \omega_1 \in \mathbb{R}$. Additionally, the parameters ω are forced to sum to unity, in essence: $\omega_0 + \omega_1 = 1$. Also, the shape and scale parameters a, b that define the Gamma distribution are set in such a way that $\mathbb{E}[S] = 1$. Given that the expected value of a Gamma distributed random variable is given by $\mathbb{E}[S] = ab$ we introduce the new variable ϵ and set $S \sim \Gamma(\epsilon, \frac{1}{\epsilon})$. Revisiting Equation (3.32), we can exploit an economical interpretation of parameters ω_0, ω_1. These parameters have a similar interpretation to the parameter β in the Gaussian threshold model, namely that of relative exposure to the systematic and idiosyncratic risk. ω_1 is essentially a factor loading for the dependence on systematic risk on the default probability of the individual counterparty. Note that the model admits factor loadings that differ across counterparties and using the summing to unity condition for ω_0, ω_1 yields:

$$p_i(S) = PD_i(\omega_0 + \omega_1 S) = PD_i(1 + \omega_i(S - 1)) \tag{3.33}$$

An additional interesting observation we can make is calculating the expected value of the conditional probability of default:

$$\mathbb{E}[p_i(S)] = \mathbb{E}[PD_i(\omega_0 + \omega_1 S)] = [PD_i(\omega_0 + \omega_1 \mathbb{E}[S]) = PD_i \tag{3.34}$$

Which shows the relation between the conditional and unconditional default probability, in a similar way as we have seen with the standard Gaussian threshold model. This indicates that in essence, forcing the Gamma distributed systematic factor to have an expectation of unity is an equivalent trick to forcing the asset returns in the Gaussian threshold model to be standard normally distributed.

3.5. Risk Contributions

In previous sections, multiple models for measuring value at risk have been described and using the value at risk, we have defined economic capital the difference between the portfolio's value at risk and its expected loss. However, the question remains how to allocate the economic capital of the full portfolio to sectors, industries, countries and single-name counterparties within the portfolio. Allocating capital to the various components of the portfolio is vital for management decisions, business planning, performance measurement, pricing and profitability assessments [36].

For instance, the performance of individual loans or business units can be measured using some sort of return on risk-adjusted capital (RORAC) approach. In essence, performance is measured by [25]:

$$RORAC = \frac{Profit}{EC}$$

There is no unique way of allocating capital; each methodology exhibits its own pros and cons depending on the application. The most popular decomposition is the marginal risk contribution, due to its additive properties. Given the definition of the EC, allocating capital based on contributions to the VaR is the most natural choice. However, this brings some difficulties, especially since the VaR refers to one particular value of loss which makes it difficult to obtain accurate risk contributions through Monte Carlo simulation.

Broadly, three capital allocation methodologies can be classified [36]:

- *Stand-alone capital contributions* - sub-portfolios are assigned the amount of economic capital as if it is a stand-alone portfolio. A drawback of this method is that it does not reflect diversification effects. Therefore, the sum of stand-alone economic capital might exceed total EC for the full portfolio.
- *Incremental capital contributions* - capital contributions of sub-portfolios are calculated through calculating the EC for the full portfolio and subtracting the EC for the portfolio without the sub-portfolio. This method captures the amount of capital that would be released if the sub-portfolio were sold. A major disadvantage of this method is that it does not yield an additive risk decomposition.
- *Marginal capital contributions* - By construction, the sum of the marginal capital contributions sum to the total EC of the portfolio. Capital is allocated to each sub-portfolio on a marginal basis. This method will be described in more detail.

3.5.1. Euler Allocation

We define the contributing economic capital (ECC) as:

$$ECC_i = VaRC_q(L_i) - EL_i \tag{3.35}$$

As stated in Equation (2.5) the expected loss per obligor is portfolio invariant. Using this, the problem of calculating the contributing economic capital of individual obligors boils down to determining the contributing VaR (VaRC). This problem is known as capital allocation and a known solution to this problem is Euler's allocation principle which states that when a risk measure positive-homogeneous (such as the VaR and ES) and differentiable with respect to exposure, then the total VaR can be allocated to the individual obligors using a relatively simple formula. Formally, the Euler principle reads [39]:

Theorem 3.5.1. ***Euler's theorem on homogeneous functions***
Let $U \subset \mathbb{R}$ be an open set and let $f : U \to \mathbb{R}$ be a continuously differentiable function. Then f is homogeneous of degree τ if and only if it satisfies:

$$\tau f(u) = \sum_{i=1}^{n} u_i \frac{\partial f(u)}{\partial u} \text{ for } u = (u_1, ..., u_n) \in U \tag{3.36}$$

With the VaR being a homogenous differentiable risk measure of degree 1, the Euler allocation of the VaR simplifies to [18]:

$$VaR_q(L) = \sum_{i=1}^{n} EAD_i \frac{\partial VaR_q(L)}{\partial EAD_i} \tag{3.37}$$

Regularly, we will calculate the VaR as a fraction of total portfolio exposure, which results in the following equation for VaRC [15]:

$$VaR_q(L) = \sum_{i=1}^{n} w_i \frac{\partial VaR_q(L)}{\partial w_i} \tag{3.38}$$

Therefore, we can define the contributing VaR (VaRC) as:

$$VaRC_{q,i} = w_i \frac{\partial VaR_q(L)}{\partial w_i} \tag{3.39}$$

Clearly, Equation (3.38) satisfies the full allocation property, meaning that the sum of the contributing VaRs equals the total portfolio VaR.

3.5.2. Monte Carlo Extension

Equation (3.39) is an analytical approach to determining the contributing VaR of individual obligors. However, since many models apply Monte Carlo sampling in determining the VaR at portfolio level, these models lack an analytical formula for the portfolio VaR and therefore we cannot directly use Equation (3.39) for determining the risk contribution in MC based models. However, [10] determined the derivatives of the VaR and [16] used these results to derive the following expression under appropriate conditions:

$$VaRC_{q,i} = \mathbb{E}[L_i | L = \alpha_q] \tag{3.40}$$

It can easily be shown that Equation (3.40) satisfies the full allocation property by the linearity of the expectation:

$$\begin{aligned} \sum_{i=1}^{N} VaRC_{q,i} &= \sum_{i=1}^{N} \mathbb{E}[L_i | L = \alpha_q] \\ &= \mathbb{E}[\sum_{i=1}^{N} L_i | L = \alpha_q] \\ &= \mathbb{E}[L | L = \alpha_q] = VaR_q(L) \end{aligned} \tag{3.41}$$

Unfortunately, Equation (3.40) poses some computational challenges. Each VaR contribution depends on the probability of an inherently rare event, namely default, conditional on an even more rare extreme loss event. Computationally, this issue translates in a need for an extreme number of MC trials in order to reduce the statistic noise inherent in this approach. Alternatively, we could try to improve the accuracy by not conditioning on the single event $L = \alpha_q$ but on a small range around $L = \alpha_q$: $L \in [(1-\gamma)\alpha_q, (1+\gamma)\alpha_q]$. In Section 4.1 we will go into more depth on this subject.

3.6. Ad Hoc Measures of Concentration

This work focuses on model-based methods for measuring and assessing concentration risk, such as the granularity adjustment. However, we will briefly touch upon on some ad-hoc measures of concentrations such as the Gini coefficient and the Herfindahl-Hirschman index (HHI). These indices will be useful in the empirical analysis as one would expect a positive dependence of any constructed model-based measurement on ad-hoc indices. In principle, both measures can be applied to both name and sector concentrations. We will shortly discuss the aforementioned measures in terms of a set of desirable properties, adapted from [37] and [3]:

1. *Transfer Principle:* The reduction of an exposure and a sequential equal increase of an existing bigger exposure must not decrease the concentration measure.

2. *Uniform distribution principle:* If all exposures are equal, then the concentration measure attains its minimal value.

3. *Lorenz-criterion:* If two portfolios, which are composed of the same number of loans, satisfy that the aggregate size of the k biggest loans of the first portfolio is greater or equal to the size of the k biggest loans in the second portfolio for $1 \leq k \leq n$, then the same inequality must hold between the measures of concentration for the two portfolios.

4. *Superadditivity:* If two or more loans are merged, the measure of concentration must not decrease

5. *Independence of loan quantity:* Consider a portfolio consisting of loans of equal size. The measure of concentration must not increase with an increase in the number of loans.

6. *Irrelevance of small exposures:* Granting an additional loan of relatively low amount must not increase the concentration measure. More formally, if s' denotes a share of a loan and a new loan with a share $\tilde{s} \leq s'$ is granted, then the concentration measure must not increase.

For a more mathematical formal definition of the aforementioned properties, the reader is referred to [3]. Furthermore, [3] has shown that if a concentration measure satisfies properties 1 and 6, all the aforementioned six properties are satisfied.

3.6.1. Gini's Index

As previously mentioned, the most common heuristic approaches to quantifying concentration risk are the Gini coefficient and the Herfindahl-Hirschmann Index [30]. In order to describe the Gini coefficient, we will first introduce the Lorenz curve. The Lorenz curve is a widely used graphical representation of the distribution of a variable z and the degree of inequality of this variable [13].

Definition 3.6.1. *The Lorenz curve is the piece-wise linear function connecting the points* (x_i, y_i) *with*

$$x_i = \frac{i}{n} \text{ and } y_i = \frac{\sum_{j=1}^{i} z_j}{\sum_{j=1}^{n} z_j} \tag{3.42}$$

where z_j *is the ordered set such that* $z_1 \leq z_2 \leq \ldots \leq z_n$

Therefore, x_i denotes the relative amount of included elements and y_i the relative amount of the i smallest elements of z_i. If all elements are of equal size, Equation (3.42) reduces to $y = x$ which is called the *line of perfect equality*. Conversely, the *line of perfect inequality* is the situation where one elements accounts for the total such that $y = 0$ for all $x < 1$ and $y = 1$ for $x = 1$. In terms of concentration risk, the Lorenz curve displays the cumulative share of exposures for each cumulative share of credits.

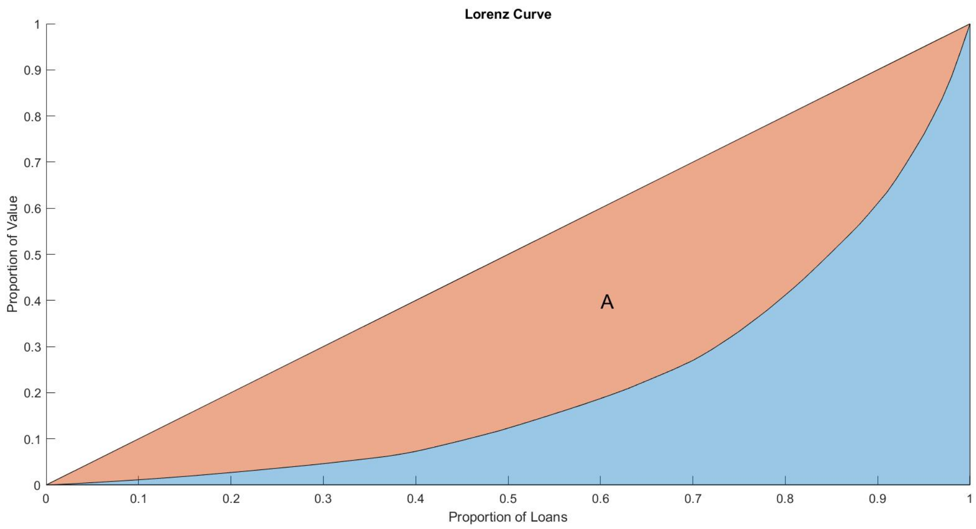

Figure 3.3: Example of the Lorenz curve for a portfolio of credit exposures.

The Gini coefficient is related to the Lorenz curve since it is depicted by the red area (denote as area A) between the line of perfect equality and the Lorenz curve in Figure 3.3. Therefore, it represents a measure for the deviation from a equal distribution as this area would decrease to zero as the Lorenz curve approaches the line of perfect equality. The Gini coefficient G is defined as twice the shaded area A transforming the coefficient from $A \in [0, 0.5]$ to $G \in [0, 1]$. Therefore, the Gini coefficient is given by:

$$G = \frac{\sum_{i=1}^{n}(2i-1)w_i}{N} - 1 \tag{3.43}$$

The major advantage of the Gini coefficient is its ease of implementation and its graphical interpretability [13]. However, its main drawback is that the coefficient does not take the size of a portfolio into account, meaning it represents a measure of inequality instead of concentration. Furthermore, [3] has shown that the Gini coefficient only satisfies properties 1,2,3 and 5.

3.6.2. Herfindahl-Hirschman Index

To overcome the shortcomings of the Gini coefficient, we introduce the Herfindahl-Hirschman Index (HHI).

Definition 3.6.2. *The HHI is defined as the sum of the squared exposure shares (measured as a fraction of the total portfolio) of each facility:*

$$HHI = \sum_{i=1}^{N} w_i^2 = \frac{1}{N^*} \tag{3.44}$$

Where we denote N^* as the number of effective exposures. Comparing the HHI to the Gini Coefficient, the HHI has two major advantages. Firstly, the HHI satisfies all aforementioned six properties of an ideal risk measure. Furthermore, it accounts for the size of the portfolio. By considering the sum squares of the portfolio shares, small exposures affect the level of concentration less than a proportional relationship [3].

Although both the HHI and the Gini Coefficient may seem to be intuitive concentration risk measures, they both lack to incorporate any other information than the EAD of individual exposures and the size of the portfolio whereas we would expect an ideal measure of concentration risk to incorporate EAD, LGD, PD and the correlation structure. Furthermore, there is no intuitive way of applying these measures in constructing add-ons for accounting for concentration risk. Therefore, the indices only provide superficial estimates for concentration risk, both name and sector.

Having mentioned this, these ad hoc measures should only be used for gaining a rough insight into the degree of concentration present in portfolios across sectors and countries and tracking the differences in concentration in time. Thus, there is a need for a more sophisticated model-based approach of measuring and accounting for concentration risk. The Basel Committee has stressed this fact in [30] as model-based measurements *represent a more consistent approach to the measurement and management of all dimensions of credit risk for the portfolio.*

3.7. Model Based Single Name Concentration Risk Measurement

The need for a model-based adjustment to account for undiversified idiosyncratic risk was first introduced in an early draft of Basel II, known as the Second Consultative Paper [22]. The adjustment was called the granularity adjustment and was introduced as a formal component of the minimum required capital rules of the IRB approach. However, the GA was determined differently from the proposed method in this work. The GA was obtained by fitting a functional form between the actual, Monte Carlo obtained, VaR and the ASRF VaR for a set of synthetic portfolios. [41] was the first to introduce an analytically derived formula for the GA based on a linear approximation around the ASRF VaR solution. [23] elaborated on the work by [41] with a more rigorous derivation of Wilde's formula, based on the work by [10]. [12] then extended the method by incorporating higher order terms of the Taylor extension around the ASRF VaR in the GA in an effort to improve accuracy. An approach largely related to [41] is the GA proposed [22] who extend the adjustment to the CreditRisk+ model.

3.7.1. Granularity Adjustment

Before elaborating on the theoretical framework of the granularity adjustment (GA), an intuitive example of how the GA works is presented. This brief example is adopted from the work by [21]. Assume a one-factor structural model of default with systematic factor $X \sim \mathcal{N}(0, \nu^2)$ and that the loss rate conditional on the systematic risk factor $U_i|X \sim \mathcal{N}(0, \sigma^2)$ are both normally distributed with σ and ν being known constants. From this it follows that the unconditional portfolio loss ratio L_N is also a normally distributed random variable. The q[th] quantile of this distribution is given by $\sqrt{\nu^2 + \sigma^2/N} \cdot \Phi^{-1}(q)$. Clearly, as the total number of obligors N increases to infinity, the distribution of L_N converges to that of X, meaning that the qth quantile of the asymptotic distribution is $\nu \cdot \Phi^{-1}(q)$. This expression corresponds to the systematic contribution to the VaR. Therefore, the idiosyncratic contribution can be derived as the difference between the quantile of the true loss ratio L_N and the asymptotic loss L_∞. Applying a Taylor expansion around $\sigma^2/N = 0$ yields:

$$\sqrt{\nu^2 + \sigma^2/N} = \nu + \frac{1}{N}\frac{\sigma^2}{2\nu} + \mathcal{O}(\frac{1}{N^2}) \tag{3.45}$$

Therefore, the idiosyncratic contribution to the VaR can be estimated by $\frac{1}{N}\frac{\sigma^2}{2\nu}\Phi^{-1}(q) + \mathcal{O}(\frac{1}{N^2})$. This add-on contribution to the VaR is what is referred to as the Granularity Adjustment. A more formal derivation and definition is given in the next section.

3.7.2. Name Concentration Granularity Adjustment

The granularity adjustment refers to incorporating the effect of portfolio size in the EC calculation. This section will elaborate on the formal derivation of the granularity adjustment. Firstly, assume an infinitely fine grained portfolio for which the VaR can be estimated under the ASRF model. Furthermore, we assume all dependence across counterparties to be driven by a single systematic factor. For this portfolio, a add-on factor is constructed which takes the finite granularity of the portfolio into account. This factor is determined through a Taylor expansion of the VaR around the ASRF solution. The ASRF model estimates the VaR as the q[th] percentile of the expected loss conditional on the systematic factor $\alpha_q(\mathbb{E}[L|X])$. We are looking for an approximation to the exact adjustment $\alpha_q(L) - \alpha_q(\mathbb{E}[L|X])$ for the effect of undiversified idiosyncratic risk in the portfolio. We start with subdividing the portfolio loss into a systematic and an unsystematic component:

$$L = \mathbb{E}[L|X] + \lambda\{L - \mathbb{E}[L|X]\} =: Y + \lambda Z \tag{3.46}$$

$Y = \mathbb{E}[L|X]$ describes the source of systematic risk in the portfolio loss and the second term $\lambda Z = \lambda\{L - \mathbb{E}[L|X]\}$ describes the the unsystematic idiosyncratic part of the portfolio loss. Clearly, λ describes the fraction of idiosyncratic risk present in the portfolio, which tends to zero if the number of obligors approaches infinitiy. For the GA, we claim that the assumption of infinite granularity is not met, meaning that λ exceeds zero. A Taylor expansion around $\lambda = 0$ yields:

$$\alpha_q(L) = \alpha_q(Y + \lambda Z) = \alpha_q(Y) + \lambda\left[\frac{d\alpha_q(Y + \lambda Z)}{d\lambda}\right]_{\lambda=0} + \frac{\lambda^2}{2}\left[\frac{d^2\alpha_q(Y + \lambda Z)}{d\lambda^2}\right]_{\lambda=0} + \mathcal{O}(\lambda^3) \tag{3.47}$$

Clearly, the first term describes the systematic contribution to the portfolio VaR and the remaining terms add an aditional component of the VaR due to undiversified idiosyncratic components. [10] derived the expression for the first two derivatives of the VaR:

$$\left[\frac{d\alpha_q(Y + \lambda Z)}{d\lambda}\right]_{\lambda=0} = \mathbb{E}[Z|Y = \alpha_q(Y)] \tag{3.48}$$

$$\left[\frac{d^2\alpha_q(Y + \lambda Z)}{d\lambda^2}\right]_{\lambda=0} = -\frac{1}{f_Y(y)}\frac{d}{dy}\left[f_Y(y)\mathbb{V}[Z|Y]\right]_{y=\alpha_q(Y)} \tag{3.49}$$

Where $f_Y(y)$ describes the probability density function of $Y = \mathbb{E}[L|X]$. The first derivative of the VaR equals zero:

$$\mathbb{E}[Z|Y] = \frac{1}{\lambda}\mathbb{E}[L - Y|Y] = \frac{1}{\lambda}\mathbb{E}[L] - \frac{1}{\lambda}\mathbb{E}[\mathbb{E}[L|Y]] = 0 \tag{3.50}$$

Furthermore, we have:

$$\lambda^2\mathbb{V}[Z|Y] = \mathbb{V}[\lambda Z|Y] = \mathbb{V}[L - Y|Y] = \mathbb{V}[L|Y] \tag{3.51}$$

Since the first derivative of the VaR vanishes the second derivative is the only relevant term in calculating the granularity adjustment. Using (3.49) and (3.51) we have:

$$GA = \frac{\lambda^2}{2}\left[\frac{d^2\alpha_q(Y + \lambda Z)}{d\lambda^2}\right]_{\lambda=0} = \frac{\lambda^2}{2}\left(-\frac{1}{f_Y(y)}\frac{d}{dy}\left[f_Y(y)\mathbb{V}[Z|Y]\right]_{y=\alpha_q(Y)}\right) = -\frac{1}{2f_Y(y)}\frac{d}{dy}\left[f_Y(y)\mathbb{V}[L|Y]\right]_{y=\alpha_q(Y)} \tag{3.52}$$

Therefore, Equation (3.55) is the expression for the general GA. This representation of the GA admits any definition of loss. As the conditional expectation $Y = \mathbb{E}[L|X]$ is assumed to be continuous and strictly monotonously decreasing in X, the probability density function $f_Y(y)$ can be transformed using the inverse function theorem (using $g^{-1}(y) = x$):

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right| = \frac{f_X(x)}{|dy/dx|} = -\frac{f_X(x)}{dy/dx} = -\frac{f_X(x)}{\frac{d}{dx}\mathbb{E}[L|X]} \tag{3.53}$$

Furthermore, we have:

$$\begin{aligned} &y = \alpha_q(Y) \\ &\Leftrightarrow \mathbb{E}[L|X] = \alpha_q(\mathbb{E}[L|X]) \\ &\Leftrightarrow \mathbb{E}[L|X] = \mathbb{E}[L|\alpha_{1-q}(X)] \\ &\Leftrightarrow x = \alpha_{1-q}(X) \end{aligned}$$

Therefore, we have:

$$\alpha_q(L) \approx \alpha_q^{ASRF}(L) + GA \tag{3.54}$$

with GA:

$$\begin{aligned} GA &= -\frac{1}{2f_X(x)}\frac{d}{dx}\left[\frac{f_X(x)\mathbb{V}[L|X=x]}{\frac{d}{dx}\mathbb{E}[L|X=x]}\right]_{x=\alpha_{1-q}(X)} \\ &= -\frac{1}{2}\left[\frac{f_X'(x)}{f_X(x)}\frac{\mathbb{V}[L|X=x]}{\frac{d}{dx}\mathbb{E}[L|X=x]} + \frac{\frac{d}{dx}\mathbb{V}[L|X=x]}{\frac{d}{dx}\mathbb{E}[L|X=x]} - \frac{\mathbb{V}[L|X=x]\frac{d^2}{dx^2}\mathbb{E}[L|X=x]}{(\frac{d}{dx}\mathbb{E}[L|X=x])^2}\right]_{x=\alpha_{1-q}(X)} \end{aligned} \tag{3.55}$$

3.7.3. Granularity Adjustment and the HHI

Referring back to the HHI of Section 3.6.2 we expect some kind of positive dependence of the granularity adjustment on the HHI. [13] derived some properties of this relation and showed that the GA is a term of order $\mathcal{O}(\frac{1}{N^*})$ where N^* denotes the number of effective exposures from Equation (3.44). Since conditional on the underlying systematic factors, default events are independent and by construction we have that the individual loss rate is bounded, i.e. $LGD_i Di \in [-1, 1]$ for all $i \in (1, ..., n)$, there exists a finite number $E(x)^* \leq 1$ such that

$$\mathbb{E}[L|X] = \mathbb{E}\left[\sum_{i=1}^{N} w_i \cdot LGD_i \cdot D_i|X\right] = \sum_{i=1}^{N} w_i E(x)^* = E(x)^* \sum_{i=1}^{N} w_i = E(x)^* \tag{3.56}$$

and similarly, there exists a finite number $V(x)^* \leq 1$ such that

$$\mathbb{V}[L|X] = \mathbb{V}\left[\sum_{i=1}^{N} w_i \cdot LGD_i \cdot D_i|X\right] = \sum_{i=1}^{N} w_i^2 \mathbb{V}[LGD_i \cdot D_i|X] = V(x)^* \sum_{i=1}^{N} w_i^2 = HHI \cdot V(x)^* = \frac{1}{N^*}V(x)^* \tag{3.57}$$

Substituting these equations in (3.55) yields

$$GA = -\frac{1}{N^*}\frac{1}{2f_X(x)}\frac{d}{dx}\left[\frac{f_X(x)V(x)^*}{\frac{d}{dx}E(x)^*}\right]_{x=\alpha_{1-q}(X)} \tag{3.58}$$

Clearly, this results states that the granularity adjustment is linear in terms of the HHI, which confirms our suspicion of some positive relationship between the HHI and the GA.

3.7.4. Sector Concentration Risk Measurement

The aforementioned methods are developed to account for name-concentration in economic capital assessments. Some analytic or semi-analytic methods exist that account for sectoral diversification. The seminal paper by [35] describes a rigorous analytical approach that extends the ideas by [23] to a multiple systematic factor setting. Furthermore, they derive this setting in both the VaR as the ES setting. Alternatively, a semi-analytic approach is developed by [5]. The authors construct an adjustment to the single factor model by scaling the economic capital resulting from the ASRF model with a capital diversification factor. The diversification factor is estimated numerically using Monte Carlo methods. More recently, [17] proposed a concentration charge by assessing the impact of different sectors on the portfolio loss curve. For calculating the concentration charge, they propose a method based on Monte Carlo simulations and a method based on the analytical approximation to the VaR by [35].

3.7.5. Pykhtin Multi-Factor Adjustment

In this section, a model is presented that extends the general idea of the GA to the multi factor situation. This model was first described in [35] and has since been widely applied in the literature. For instance, [17] applies the Pykhtin approach to measure sector concentration risk. The multi factor adjustment provides an analytical method for calculating the VaR and EC in a multi factor setting without going through computationally time-consuming Monte Carlo simulations. The idea of [35] is to approximate the portfolio loss L in the multi factor model by adjusting the portfolio loss $\bar{L}$ of the ASRF model. Pykhtin does this by mapping the correlation structure of the multi factor model to a single correlation factor through maximizing the correlation between the new single risk factor and the original sector and geographical factors. Then, similar to the process described in Section 3.7.2, a Taylor series expansion around the constructed single factor model yields the desired adjustment. In this section, we follow the derivations by [35], [13] and [21].

In essence, the distribution of $\bar{L}$, denoting the loss of the adjusted single factor model, can be calculated with the known solution of the ASRF method:

$$\bar{L} = \mu(\bar{X}) = \sum_{i=1}^{N} w_i \cdot LGD_i \cdot \Phi\left(\frac{\Phi^{-1}(PD_i) - c_i\bar{X}}{\sqrt{1-c_i^2}}\right) \tag{3.59}$$

Where c_i is the desired parameter that maximizes the correlation between the existing systematic factors and the new single systematic factor. Since this method shows great similarities with the one-factor granularity adjustment, we start by describing the method for determining the single correlation factor to condense the multi factor model into a single factor model. Denote the original systematic sector factors from the multi factor model as X_k for $k = 1, \ldots, K$ normally distributed systematic factors. To relate the estimated $\tilde{L}$ by the ASRF model to the true portfolio loss L we link the new effective single systematic risk factor $\bar{X}$ to the original systematic risk factors X_k by:

$$\bar{X} = \sum_{k=1}^{K} b_k X_k \tag{3.60}$$

Where $\sum_{k=1}^{K} b_k^2 = 1$ to preserve unit variance of $\bar{X}$. In order to determine $\bar{L}$, c_i and b_k have to be specified. Remember the general multi factor framework:

$$r_i = \beta_i Y_i + \sqrt{1-\beta_i^2}\epsilon_i \tag{3.61}$$

Similar to the granularity adjustment approach, we assume that $\bar{L} = \mathbb{E}[L|\bar{X}]$. To calculate $E[L|\bar{X}]$ we make the following assumption:

$$Y_i = \delta_i \bar{X} + \sqrt{1-\delta_i^2}\eta_i \tag{3.62}$$

Where $\eta_i \sim \mathcal{N}(0,1)$ independent of $\bar{X}$ and:

$$\delta_i = \mathrm{Corr}(Y_i, \bar{X}) = \sum_{k=1}^{K} \alpha_{i,k} b_k \tag{3.63}$$

Using the introduced equations and notations, we can rewrite Equation (3.61) as:

$$r_i = \beta_i \delta_i \bar{X} + \sqrt{1 - \beta_i^2 \delta_i^2} \epsilon_i \tag{3.64}$$

The conditional expectation of L is then clearly given by:

$$\mathbb{E}[L|\bar{X}] = \sum_{i=1}^{N} w_i \cdot LGD_i \cdot \Phi \left(\frac{\Phi^{-1}(PD_i) - \beta_i \delta_i \bar{X}}{\sqrt{1 - \beta_i^2 \delta_i^2}} \right) \tag{3.65}$$

By comparing (3.59) and (3.65) the equality $\bar{L} = \mathbb{E}[L|\bar{X}]$ is obtained if and only if the following restriction for the effective factor loading holds:

$$c_i = \beta_i \delta_i = \beta_i \sum_{k=1}^{K} \alpha_{i,k} b_k \tag{3.66}$$

While the coefficients β_i and $\alpha_{i,k}$ are known, the coefficients b_k are unknown. Unfortunately, determining the coefficients b_k is not obvious nor trivial. Ideally, we want to identify a set $\{b_k\}$ that minimizes the difference between the quantiles of the true loss function and the approximation of the loss function. However, finding such a set is outside the scope of [35]. Therefore, [35] opts for a method such that the coefficients are determined through maximizing the correlation between the single risk factor $\bar{X}$ and the original risk factor Y_i for all i. This leads to the following maximization problem:

$$\max_{b_1,\dots,b_K} \left(\sum_{i=1}^{N} d_i \sum_{k=1}^{K} \alpha_{i,k} b_k \right) \text{ such that } \sum_{k=1}^{K} b_k^2 = 1 \tag{3.67}$$

A solution to this maximization problem is given by an application of the Lagrange multiplier:

$$\Lambda(b_k, \lambda) = \sum_{i=1}^{N} d_i \sum_{k=1}^{K} \alpha_{i,k} b_k - \lambda \left(\sum_{k=1}^{K} b_k^2 - 1 \right) \tag{3.68}$$

Taking partial derivatives leads to a system of $K+1$ equations:

$$\frac{\partial \Lambda(b_k, \lambda)}{\partial b_k} = \sum_{i=1}^{N} d_i \alpha_{i,k} - 2\lambda b_k = 0 \text{ for all } k = 1, \dots, K$$

$$\frac{\partial \Lambda(b_k, \lambda)}{\partial \lambda} = \sum_{k=1}^{K} b_k^2 - 1 = 0$$

The solution for this system of equations is given by the following b_k^{opt}:

$$b_k = \sum_{i=1}^{N} \frac{d_i \alpha_{i,k}}{2\lambda} \tag{3.69}$$

Where the Lagrange multiplier λ is chosen in such a way that $\{b_k\}$ satisfies the second constraint, in essence, such that $\{b_k\}$ has a Euclidian norm of one. Unfortunately, we have not considered the coefficients d_i yet. [35] has empirically determined that the following definition is the best performing choice:

$$d_i = w_i \cdot LGD_i \cdot \Phi \left(\frac{\Phi^{-1}(PD_i) - \beta_i \Phi^{-1}(1-q)}{\sqrt{1 - \beta_i^2}} \right) \tag{3.70}$$

The intuition behind this choice for d_i is that obligors with a high exposure in terms of VaR should have a large weight in the maximization problem whereas obligors with a small VaR should have a minor impact [13]. Summing up, we can now define the comparable one factor model (3.64) through using Equation (4.77) together with the expressions for b_k and d_i we have derived.

Having constructed the loss variable $\bar{L} = \mathbb{E}[L|\bar{X}]$, we can largely adopt the method described in Section 3.7.2. We again perturb the true portfolio loss variable L by $Z := L - \bar{L}$:

$$L = \bar{L} + \lambda Z \tag{3.71}$$

Note that λ as applied here denotes the perturbation coefficient, and is not equal to the Langrange multiplier in previous sections. Applying exactly the same reasoning as in Section 3.7.2 and Section 4.2.1 to retrieve a adjustment of the form:

$$\Delta\alpha_q = \alpha_q(L) - \alpha_q(\bar{L}) \approx -\frac{1}{2}\left[\frac{f_X'(x)}{f_X(x)}\frac{\mathbb{V}[L|X=x]}{\frac{d}{dx}\mathbb{E}[L|X=x]} + \frac{\frac{d}{dx}\mathbb{V}[L|X=x]}{\frac{d}{dx}\mathbb{E}[L|X=x]} - \frac{\mathbb{V}[L|X=x]\frac{d^2}{dx^2}\mathbb{E}[L|X=x]}{(\frac{d}{dx}\mathbb{E}[L|X=x])^2}\right]_{x=\alpha_{1-q}(X)} \tag{3.72}$$

3.7.6. Kurtz's Capital Charge for Concentration Risk

The following section is largely based on the work by [17]. The authors propose a method to compute the (economic) capital charges for concentration risk in a multi factor model setting similar to the model explained in Section 2.5. The sector concentration effect is defined as the impact of the weight of sector losses on the portfolio loss curve. To compute the capital charge, the loss distribution of the portfolio both with and without the sector under consideration is calculated. The sector concentration is then calculated as the difference between the original VaR contribution, and the VaR contribution of the sector after removal. [17] propose two methods to perform the required calculations of the loss distribution, one based on the Monte Carlo approach and one based on the method described in Section 3.7.5.

The authors of [17] argue that sector concentration charges are implicitly present in the analytical EC calculations of [35], but not presented as an isolated component. The Pykhtin framework decomposes the VaR into a linear component, a non-linear multi factor adjustment associated with adjusting the multi factor to a single factor model and a granularity adjustment (see Equation (4.54)). [17] argues that the linear component incorporates a large amount of the sector concentration risk meaning that the multi factor adjustment only measures a small portion of the concentration risk and cannot be used in isolation to compute sector concentration add-ons.

Multi Factor Concentration Charge

The framework for measuring concentration effects from [17] is based on a general multi factor model. We again consider a portfolio of N loans to unique obligors. The obligors can be assigned to M different industries (or geographical regions). Furthermore, define $s(i)$ as a function mapping obligor i to its sector $s(i)$. Additionally, denote N_s as the number of obligors in sector $s = 1, ..., M$. Thus, the total number of obligors equals $N = \sum_{s=1}^{M} N_s$.

In this setting, EC is measured as usual, i.e. as the difference between portfolio VaR and expected loss. Additionally, we apply the Euler allocation principle (see Section 3.5.2) to define the contributing VaR of sub-portfolio $S \subset \{1, ..., N\}$:

$$VaRC_{q,S} = \mathbb{E}[L(S)|L = \alpha_q] \tag{3.73}$$

Furthermore, using this equation we can define the economic capital of sub-portfolio S as:

$$EC_q(S) = VaRC_{q,S} - \mathbb{E}[L(S)] \tag{3.74}$$

Next, we denote the portfolio $P = 1, ..., N$ and the complement of the sub-portfolio as $\bar{S} = P \backslash S$. Furthermore, we assume that an obligor either belongs to S or $\bar{S}$, meaning that obligors cannot be split. According to [17], if the exposure of S is increased linearly (for instance by doubling each exposure), and $\bar{S}$ is kept constant, then $EC(S)$ increases more than linearly. An explanation for this effect is that besides the losses associated with S being higher, scenarios with a high loss rate of S are also more likely to be tail scenarios. This non-linear effect is what [17] refers to as the concentration effect. Formally, the concentration charge is defined through:

Definition 3.7.1. *The **concentration charge** for sub-portfolio S equals*

$$CC_q(S) := EC_q(S) - EC_q^{NC}(S) \tag{3.75}$$

Where $EC_q^{NC}(S)$ denotes the economic capital contribution of portfolio S to the non-concentrated portfolio. $EC_q^{NC}(S)$ is given by:

$$EC_q^{NC}(S) := \lim_{\lambda\to\infty} \mathbb{E}[L(S)|L(S \cup \lambda\bar{S}) = \alpha_q^{S\cup\lambda\bar{S}}] - \mathbb{E}[L(S)] \tag{3.76}$$

With $\lambda \in \mathbb{R}_+$. The sub portfolio $\lambda\bar{S}$ is the original sub portfolio $\bar{S}$ scaled by λ. The intuition behind (3.76) is that all non-linear concentration effects associated with S are eliminated due to the linear scaling of $\bar{S}$. Formally, [17] shows this behaviour using the following argument for $\mu > 0$:

$$\begin{aligned} EC_q^{NC}(\mu S) &= \lim_{\lambda\to\infty} \mathbb{E}[L(\mu S)|L(\mu S \cup \lambda\bar{S}) = \alpha_q^{\mu S\cup\lambda\bar{S}}] - \mathbb{E}[L(\mu S)] \\ &= \lim_{\lambda\to\infty} \mathbb{E}[L(\mu S)|L(\mu S \cup \mu\lambda\bar{S}) = \alpha_q^{\mu S\cup\mu\lambda\bar{S}}] - \mathbb{E}[L(\mu S)] \\ &= \lim_{\lambda\to\infty} \mathbb{E}[L(\mu S)|L(S \cup \lambda\bar{S}) = \alpha_q^{S\cup\lambda\bar{S}}] - \mu\mathbb{E}[L(S)] \\ &= \mu\left(\lim_{\lambda\to\infty} \mathbb{E}[L(S)|L(S \cup \lambda\bar{S}) = \alpha_q^{S\cup\lambda\bar{S}}] - \mathbb{E}[L(S)]\right) = \mu EC_q^{NC}(S) \end{aligned}$$

Where all arguments essentially follow from the positive homogeneity of the VaR and the linearity of the portfolio loss variable $L(\cdot)$. In order to define a method for calculating $EC_q^{NC}(S)$ we use the following proposition:

Proposition 3.7.1. *Let $P = S \cup \bar{S}$ and let the loss distribution of $\bar{S}$ be a strictly increasing function in the vicinity of quantile q. Then*

$$EC_q^{NC}(S) = \mathbb{E}[L(S)|L(\bar{S}) = \alpha_q^{\bar{S}}] - \mathbb{E}[L(S)] \tag{3.77}$$

For the proof of this proposition we refer to Theorem 2.4 in [17]. Using Equations (3.73) through (3.75) we can express the concentration charge as:

$$CC_q(S) = \mathbb{E}[L(S)|L = \alpha_q] - \mathbb{E}[L(S)|L(\bar{S}) = \alpha_q^{\bar{S}}] \tag{3.78}$$

Using this format, we notice that determining the concentration charge essentially boils down to calculating the loss distribution of $\bar{S}$. If we define the sub-portfolio S as all exposures within a specific sector, we need to calculate a new loss distribution $L(\bar{S})$ for each sector. As the numbers of sectors in a portfolio increases, this is a very demanding task as it involves new Monte Carlo simulations for each sector. Alternatively, we could apply the analytic Pykhtin approach formulated in Section 3.7.5 to speed up this task.

4

Model based techniques for assessing concentration risk

Throughout this chapter, the methods derived in Chapter 3 are applied to credit risk models by making assumptions on the distribution of the underlying systematic and idiosyncratic factors. Furthermore, we will briefly touch upon some computational techniques such as MC simulation and sampling correlated random variables. This chapter ends with several extensions on existing methods, such as the introduction of Recovery Risk, applying Euler allocation to both the single factor as the multi factor GA, Economic & Regulatory Concentration Risk and concentration risk in t-threshold models.

4.1. Monte Carlo Approach

All measures of interest (VaR, ES, EC) are obtained from the portfolio loss distribution. The loss distribution is affected by three stochastic sources: systematic factors, idiosyncratic factors and random LGD's. Additionally, the correlation structure of the systematic factors influences the portfolio loss distribution. Because of these complex sources of randomness, loss distributions cannot be determined analytically for realistic portfolios. Therefore, the most widely applied method to calculate the portfolio loss distribution is Monte Carlo simulation. Monte Carlo simulations are used to calculate both portfolio and obligor level risk metrics.

On a high level, the MC sampling algorithm proceeds as follows:

Algorithm 1: Multi-Factor Monte Carlo Method

for *i= 1:trials* **do**

1. Draw a set of K i.i.d. standard normal systematic factor realizations;
2. Correlate systematic factor realizations ;
3. Draw N i.i.d. standard normal idiosyncratic factor realizations;
4. Compute the value of each instrument at horizon using Eq. (2.15) ;
5. Check whether asset returns breach default point ;
    - If no default occurs, set loss to zero ;
    - If default occurs, draw i.i.d. Beta distributed LGD

end

Compute portfolio risk measures from summed portfolio losses ;

Compute obligor level risk measures from obligor specific losses

In order to accurately determine the portfolio loss distribution and risk contributions of individual obligors, the amount of trials is extremely large, up to ten million. Depending on the size of the portfolio, the algorithm takes up to several days to fully run. However, for smaller portfolios a high accuracy can be attained at several hundred thousands of trials. Furthermore, the performance of the algorithm can be improved substantially by removing the for loop and using matrix operations instead. For this work, in which we are dealing with $N = 40.000+$ obligors, a combination of the for loops together with matrix operations is applied to satisfy memory limits of the computing systems.

Determining VaR and contributing VaR using MC
Suppose we perform $M \in \mathbb{N}$ Monte Carlo trials using algorithm 1 resulting in a collection of portfolio losses: $\{L^{(m);m=1,...,M}\}$. The simulation provides us with all the necessary ingredients to estimate the portfolio VaR and the VaR contributions. To estimate the $VaR(L)_q$, we place the collection of portfolio losses in ascending order and denote is as a vector $\hat{L}$. Then the VaR estimator is given by:

$$VaR_q(L) = \hat{L}(\lceil q \cdot M \rceil) \tag{4.1}$$

where $\lceil x \rceil$ denotes the smallest integer greater or equal to x [2]. In order to compute the VaR contributions, we first recall the following fundamental property of conditional expectation:

$$\mathbb{E}[L_i | L = VaR_q(L)] = \frac{\mathbb{E}[L_i \mathbb{1}_{L=VaR_q(L)}]}{\mathbb{P}(L = VaR_q(L))} \tag{4.2}$$

This convenient form leads to the following equation for contributing VaR estimator

$$VaRC_{q,i} = \frac{\sum_{m=1}^{M} L_i^{(m)} \mathbb{1}_{L^{(m)}=VaR_q(L)}}{\sum_{m=1}^{M} \mathbb{1}_{L^{(m)}=VaR_q(L)}} \tag{4.3}$$

for $i = 1, ..., N$. However, in practice, the denominator is often equal to one, since we will typically only observe one single VaR measure at a certain level q. This observation makes the contributing VaR very susceptible to noise. One way to solve this problem is by repeating the entire MC simulation multiple times and then averaging the contributing VaR results over these multiple simulations. However, as mentioned in Section 3.5.2, we opt for a slightly different method. To this extend, we use the following proposition, adapted from [27]:

Proposition 4.1.1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $U : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^n$ be an n-dimensional random variable and $f : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ be a real random variable with finite mean. Then for $\mathbb{P}^U$- almost every $u \in \mathbb{R}^n$ and almost every sequence $(\omega_i)_{i\in\mathbb{N}}$ in Ω, we have*

$$\mathbb{E}[f | U = u] = \lim_{\epsilon \downarrow 0} \lim_{K\to\infty} \frac{\sum_{k=1}^{K} \mathbb{1}_{B_\epsilon(u)}(U(\omega_i)) f(\omega_i)}{\sum_{k=1}^{K} \mathbb{1}_{B_\epsilon(u)}(U(\omega_i))} \tag{4.4}$$

where $\mathbb{P}^U$ denotes the probability distribution of U. By applying this result to compute $VaRC_{q,i}$ using MC methods, we get

$$VaRC_{q,i} \approx \frac{\sum_{m=1}^{M} L_i^{(m)} \mathbb{1}_{L^{(m)} \in [VaR_q(L)-\epsilon, VaR_q(L)+\epsilon]}}{\sum_{m=1}^{M} \mathbb{1}_{L^{(m)} \in [VaR_q(L)-\epsilon, VaR_q(L)+\epsilon]}} \tag{4.5}$$

for a sufficiently large number of MC trials M and a sufficiently small $\epsilon > 0$. Essentially, this means that we evaluate the losses of the full portfolio around the quantile of L_N and average the individual losses L_i at these positions.

4.1.1. Sampling Correlated Random Variables

In step two of the standard MC method, algorithm 1, standard normal random variables have to be correlated according to the covariance matrix $\mathbf{\Omega}$ of the K systematic factors. Suppose we wish to generate a random vector $\mathbf{X} = (X_1, ..., X_k)'$ where $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Omega)$. Let $\mathbf{Z} = (Z_1, ..., Z_k)'$ be a vector of i.i.d. standard normal random variables, i.e. $Z_i \sim \mathcal{N}(0, 1)$ for $i = 1, .., K$. Let $\mathbf{A}$ be a KxK matrix, then it follows that:

$$\mathbf{A}^T\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{A}^T\mathbf{A}) \tag{4.6}$$

Therefore, the problem reduces to finding $\mathbf{A}$ such that $\mathbf{A}^T\mathbf{A} = \mathbf{\Omega}$. A well-established method for solving this problem is the Cholesky decomposition.

Cholesky Decomposition

For any symmetric positive-definite matrix[1] M, the following equation holds:

$$\mathbf{M} = \mathbf{U}^{\intercal}\mathbf{D}\mathbf{U} \tag{4.7}$$

where **U** is an upper triangular matrix and **D** is a diagonal matrix with all diagonal elements being positive. Since by construction, the variance-covariance matrix $\boldsymbol{\Omega}$ is symmetric positive-definite, we have:

$$\boldsymbol{\Omega} = \mathbf{U}^{\intercal}\mathbf{D}\mathbf{U} = (\mathbf{U}^{\intercal}\sqrt{\mathbf{D}})(\sqrt{\mathbf{D}}\mathbf{U}) = (\sqrt{\mathbf{D}}\mathbf{U})^{\intercal}(\sqrt{\mathbf{D}}\mathbf{U}) \tag{4.8}$$

Clearly, matrix $\mathbf{A} = \sqrt{\mathbf{D}}\mathbf{U}$ satisfies $\mathbf{A}^{\intercal}\mathbf{A} = \boldsymbol{\Omega}$. **A** is called the Cholesky Decomposition of $\boldsymbol{\Omega}$. In the actual sampling procedure, we need to take the factor weights of the individual counterparties into account. For each counterparty, denote their vector of factor weights as $\mathbf{w}_i = (w_{i,1}, ..., w_{i,K})$ that satisfies the constraint $\sum_{k=1}^{K} w_{i,k} = 1$. The calculation of the factor loadings proceeds as follows:

Algorithm 2: Weighted Correlated Random Variables

Apply Cholesky decomposition to variance covariance matrix $\Omega_{KxK} = A_{KxK}^{\intercal}A_{KxK}$;
for *i= 1:N* **do**
  Calculate the factor loadings for counterparty i

$$(\alpha_{i,1}, ..., \alpha_{i,K}) = (w_{i,1}, ..., w_{i,K}) \times A_{KxK}$$

  Normalize the factor loadings for counterparty i

$$(\hat{\alpha}_{i,1}, ..., \hat{\alpha}_{i,K}) = \frac{(\alpha_{i,1}, ..., \alpha_{i,K})}{||(\alpha_{i,1}, ..., \alpha_{i,K})||_2}$$

  Set $(\alpha_{i,1}, ..., \alpha_{i,K}) = (\hat{\alpha}_{i,1}, ..., \hat{\alpha}_{i,K})$
end

Essentially, algorithm 2 constitutes step 2 of algorithm 1. Therefore, a combination of the two algorithms constitutes the full MC method for the multi factor threshold model. However, algorithm 2 will prove to be applicable more widely throughout this work.

4.2. Granularity Adjustment applied to Credit Risk Models

Equation (3.55) describes the general form of the GA, without making any assumptions on the distribution of the systematic risk factor or on the variance and expected value of the conditional loss of the portfolio. In order to apply this equation, a model has to be fixed, for instance the one-factor threshold model or the one-factor CreditRisk+ model. In the following sections, we will further evaluate the GA for these specific credit risk models.

4.2.1. First-Order granularity adjustment for the single factor Gaussian threshold model

As we have seen in Equation (2.20) the conditional probability of default on the systematic factor X is given by:

$$PD_i(X) = \Phi\left(\frac{\Phi^{-1}(PD_i) - \beta_i X}{\sqrt{1-\beta_i^2}}\right) \tag{4.9}$$

where the systematic factor X is standard normally distributed. To ease notation, we denote $\mu(x) = \mathbb{E}[L|X=x]$ and $\sigma^2(x) = \mathbb{V}[L|X=x]$. Furthermore, since X is standard normally distributed, it follows from the probabiliy density function of a standard normal random variable that:

$$\frac{f_X'(x)}{f_X(x)} = -x \tag{4.10}$$

[1]**M** is positive definite if for any $\mathbf{a} \in \mathbb{R}^K \backslash \{\mathbf{0}\}$: $\mathbf{a}^{\intercal}\mathbf{M}\mathbf{a} > 0$

Using this notation and Equation (4.10) the GA (3.55) can be easily expressed as:

$$GA = \frac{1}{2}\left[\frac{x \cdot \sigma^2(x) - \sigma^{2'}(x)}{\mu'(x)} + \frac{\sigma^2(x)\mu''(x)}{(\mu'(x))^2}\right]_{x=\Phi^{-1}(1-q)} \tag{4.11}$$

Note that this equation follows from assuming X to be normally distributed alone, no assumptions are made on $\sigma^2(x)$ and $\mu(x)$ yet. In order to determine the first and second derivative of the conditional expectation and the first derivative of the conditional variance a credit risk model has to be set. Therefore, the loss given default are assumed to be independent. Furthermore, denote the expectation of the loss given default to be LGD and its variance as VLGD. The conditional expectation and variance are given by:

$$\mu(x) = \sum_{i=1}^{n} w_i \cdot LGD_i \cdot PD_i(x) \tag{4.12}$$

$$\sigma^2(x) = \sum_{i=1}^{n} w_i^2 \cdot [(LGD_i^2 + VLGD_i) \cdot PD_i(x) - LGD_i^2 \cdot PD_i(x)^2] \tag{4.13}$$

The desired derivatives are given by:

$$\mu'(x) = \sum_{i=1}^{n} w_i \cdot LGD_i \cdot PD_i'(x) \tag{4.14}$$

$$\mu''(x) = \sum_{i=1}^{n} w_i \cdot LGD_i \cdot PD_i''(x) \tag{4.15}$$

$$\sigma^{2'}(x) = \sum_{i=1}^{n} w_i^2 \cdot [(LGD_i^2 + VLGD_i) \cdot PD_i'(x) - LGD_i^2 \cdot (PD_i(x)^2)'] \tag{4.16}$$

Again, to ease notation denote $PD_i(x) = \Phi(z_i)$ with $z_i = \frac{\Phi^{-1}(PD_i) - \beta_i X}{\sqrt{1-\beta_i^2}}$. Using this notation:

$$PD_i'(x) = \frac{d}{dx}\Phi(z_i) = -\frac{\sqrt{\beta_i}}{\sqrt{1-\beta_i}} f_X(z_i) \tag{4.17}$$

$$PD_i''(x) = -\frac{\beta_i}{1-\beta_i} f_X(z_i) \cdot z_i \tag{4.18}$$

$$(PD_i(x)^2)' = \frac{d}{dx}\Phi(z_i)^2 = -2\frac{\sqrt{\beta_i}}{\sqrt{1-\beta_i}} f_X(z_i) \cdot \Phi(z_i) \tag{4.19}$$

Combining Equations (4.12) through (4.19) together with (4.11) yields the following elaborate expression for the GA in case of the single factor Gaussian threshold model:

$$GA = \frac{1}{2}\left[\Phi^{-1}(q)\frac{\sum_{i=1}^{n} w_i^2 \cdot [(LGD_i^2 + VLGD_i)\cdot\Phi(z_i) - LGD_i^2\cdot(\Phi(z_i)^2)]}{\sum_{i=1}^{n} w_i \cdot LGD_i \cdot \frac{\sqrt{\beta_i}}{\sqrt{1-\beta_i}} f_X(z_i)}\right]$$
$$-\frac{1}{2}\left[\frac{\sum_{i=1}^{n} w_i^2 \cdot [(LGD_i^2 + VLGD_i)\cdot \frac{\sqrt{\beta_i}}{\sqrt{1-\beta_i}} f_X(z_i) - 2\cdot LGD_i^2 \cdot \frac{\sqrt{\beta_i}}{\sqrt{1-\beta_i}} f_X(z_i)\cdot\Phi(z_i)]}{\sum_{i=1}^{n} w_i \cdot LGD_i \cdot \frac{\sqrt{\beta_i}}{\sqrt{1-\beta_i}} f_X(z_i)}\right]$$
$$-\frac{1}{2}\left[\sum_{i=1}^{n} w_i^2 \cdot [(LGD_i^2 + VLGD_i)\cdot\Phi(z_i) - LGD_i^2\cdot(\Phi(z_i)^2)]\right]$$
$$\cdot\left[\frac{\sum_{i=1}^{n} w_i \cdot LGD_i \cdot \frac{\beta_i}{1-\beta_i} f_X(z_i)\cdot z_i}{\left(\sum_{i=1}^{n} w_i \cdot LGD_i \cdot \frac{\sqrt{\beta_i}}{\sqrt{1-\beta_i}} f_X(z_i)\right)^2}\right]\Bigg|_{z_i=\frac{\Phi^{-1}(PD_i)-\beta_i\Phi^{-1}(q)}{\sqrt{1-\beta_i^2}}} \quad (4.20)$$

4.2.2. Contributing VaR for the Single Factor Granularity Adjustment

The GA of Equation (3.55) indicates the size of the GA for the full portfolio. However, in managing portfolio credit risk, one is interested in the build up of the total credit risk. In other words, one is looking for the risk contributions of individual counterparties to the total risk. To that extend, we can apply the theory developed in Section 3.5.1. In this section, we limit ourselves to the single factor Gaussian threshold setting of Section 4.2.1, and therefore we have:

$$GA = \frac{1}{2}\left[\frac{x\cdot\sigma^2(x) - \sigma^{2'}(x)}{\mu'(x)} + \frac{\sigma^2(x)\mu''(x)}{(\mu'(x))^2}\right]_{x=\Phi^{-1}(1-q)} \quad (4.21)$$

Furthermore, we introduce a slight abuse of notation where for instance $\mu'(x)$ denotes a partial derivative with respect to x and $\frac{\partial\mu(x)}{\partial w_i}$ denotes the partial derivate with respect to exposure weight w_i. In essence, using Equation (3.39), determining risk contribution reduces to taking the partial derivative of Equation (3.54) with respect to w_i:

$$w_i\frac{\partial\alpha_q(L)}{\partial w_i} \approx w_i\frac{\partial\alpha_q^{ASRF}(L)}{\partial w_i} + w_i\frac{\partial GA}{\partial w_i} \quad (4.22)$$

Where $\frac{\partial\alpha_q^{ASRF}(L)}{\partial w_i}$ allows for a very straightforward solution:

$$w_i\frac{\partial\alpha_q^{ASRF}(L)}{\partial w_i} = w_i\frac{\partial}{\partial w_i}\sum_{i=1}^{N} w_i\cdot LGD_i\cdot\Phi\left(\frac{\Phi^{-1}(PD_i)+\beta_i\Phi^{-1}(q)}{\sqrt{1-\beta_i^2}}\right) = w_i LGD_i\cdot\Phi\left(\frac{\Phi^{-1}(PD_i)+\beta_i\Phi^{-1}(q)}{\sqrt{1-\beta_i^2}}\right) \quad (4.23)$$

Unfortunately, the expression for $w_i\frac{\partial GA}{\partial w_i}$ is slightly more cumbersome to derive. The partial derivative with respect to exposure weight of Equation (4.11) is given by:

$$w_i\frac{\partial}{\partial w_i}GA = w_i\frac{1}{2}\left[x\frac{\partial}{\partial w_i}\left(\frac{\sigma^2(x)}{\mu'(x)}\right) - \frac{\partial}{\partial w_i}\left(\frac{\sigma^{2'}(x)}{\mu'(x)}\right) + \frac{\partial}{\partial w_i}\left(\frac{\sigma^2(x)\mu''(x)}{(\mu'(x))^2}\right)\right]_{x=\Phi^{-1}(1-q)} \quad (4.24)$$

This leaves us with calculating the expressions $\frac{\partial}{\partial w_i}\left(\frac{\sigma^2(x)}{\mu'(x)}\right)$, $\frac{\partial}{\partial w_i}\left(\frac{\sigma^{2'}(x)}{\mu'(x)}\right)$ and $\frac{\partial}{\partial w_i}\left(\frac{\sigma^2(x)\mu''(x)}{(\mu'(x))^2}\right)$. For the exact expressions of these partial derivatives and their application to the single factor Gaussian threshold model, we refer to appendix A.2. Using the exact expressions, the analytical contributing VaR in the single factor setting, is easily calculated using Equation (4.22).

4.2.3. CreditRisk+ Adjustment

The following derivation of the granularity adjustment for assessing economic capital is largely based on the work by [22]. The authors propose a granularity adjustment for portfolio credit VaR that accounts for undiversified idiosyncratic risk in the portfolio based on the CreditRiks+ model. Similarly to the derivation of the GA in the one-factor setting, we start of with the general expression for the GA Equation (3.55). The below described derivation of the GA for the CreditRisk+ model is largely based on the work by [21] in combination with [22].

As we have seen in previous sections, the expressions for $f_X(x)$, $\mathbb{V}[L]$ and $\mathbb{E}[L]$ in Equation (3.55) are model dependent. Instead of basing these expressions on the model that underlies the IRB formulas, [22] propose to base these expressions on the single factor CreditRisk+ model. The CreditRisk+ model is a widely used industry model for portfolio credit risk that was introduced by Credit Suisse Financial Products in 1997 and is further explained in Section 3.4. A major upside to the GA in the CreditRisk+ setting is its analytical tractability and transparency, compared to the intricate expression for the GA in the one factor Merton model setting Equation (4.20). Especially in a regulatory setting, analytic tractability of expressions is a desired property.

Firstly, we define the loss rate as $U_i = LGD_i \cdot D_i$ where D_i denotes the default indicator. A major approximation on which the CreditRisk+ model is build, is the approximation of the default indicator to be Poisson distributed. This assumptions will help us later on in the derivation of the GA. In order to ensure tractability, the following notation is introduced:

$$\mu(x) = \mathbb{E}[L|x] = \sum_{i=1}^{N} w_i \mu_i(x) \tag{4.25}$$

$$\sigma^2(x) = \mathbb{V}[L|x] = \sum_{i=1}^{N} w_i^2 \sigma_i^2(x) \tag{4.26}$$

In the CreditRisk+ setting, $\mu(x)_i$ is given by:

$$\mu_i(x) = LGD_i \cdot PD_i(x) = LGD_i \cdot PD_i \cdot (1 + \omega_i(x-1)) \tag{4.27}$$

For the conditional variance, we have:

$$\sigma_i^2(x) = \mathbb{E}[LGD_i^2 D_i^2|X] - LGD_i^2 \cdot PD_i(x)^2 = \mathbb{E}[LGD_i^2]\mathbb{E}[D_i^2|X] - \mu_i(x)^2 \tag{4.28}$$

Since we assume the default indicator D_i given X to be Poisson distributed, the expected value and variance of a Poisson distributed variable yield $\mathbb{E}[D_i|X] = \mathbb{V}[D_i|X] = PD_i(x)$ which in turn implies by definition of the variance:

$$\mathbb{E}[D_i^2|X] = PD_i(x) + PD_i(x)^2 \tag{4.29}$$

Furthermore, substituting $\mathbb{E}[LGD_i^2] = \mathbb{V}[LGD_i] + LGD_i^2$ and denoting $\mathbb{V}[LGD_i] = VLGD_i$ in Equation (4.28) together with the previous expression yields the following expression for the conditional variance in the CR+ setting:

$$\sigma_i^2(x) = (VLGD_i + LGD_i^2) \cdot (PD_i(x) + PD_i(x)^2) - \mu_i(x)^2 = C_i \mu_i(x) + \mu_i(x)^2 \frac{VLGD_i}{LGD_i^2} \tag{4.30}$$

With $C_i = \frac{VLGD_i + LGD_i^2}{LGD_i}$. Again, similarly to the GA for the one-factor Merton model, an expression for the distribution of the single risk factor X is required. In the CR+ setting, we assume the single risk factor to be Gamma distributed, $X \sim \Gamma(\epsilon, \frac{1}{\epsilon})$.

X being Gamma distributed, the shape-scale version of the Gamma PDF is given by:

$$f_X(x) = \frac{x^{\epsilon-1} e^{-x\epsilon}}{\Gamma(\epsilon)(\frac{1}{\epsilon})^{\epsilon}} \tag{4.31}$$

Where $\Gamma(\epsilon)$ denotes the Gamma function. The first derivative with respect tor x has the following convenient form:

$$\frac{df_X(x)}{dx} = f_X(x)\left(\frac{\epsilon-1}{x} - \epsilon\right) \tag{4.32}$$

Therefore, since the GA requires the ratio of the probability density function to its derivative, we have:

$$\frac{f_X'(x)}{f_X(x)} = \left(\frac{\epsilon - 1}{x} - \epsilon\right) = A \tag{4.33}$$

Furthermore, the linear dependence of $\mu(x)$ on x implies that the second derivative of the conditional expectation $\mu''(x)$ vanishes. Therefore Equation (4.33) together with Equation (3.55) and a simple application of the quotient rule for derivatives yields the following expression for the GA in the CR+ setting:

$$GA = -\tfrac{1}{2}\left(\frac{A\sigma^2(x) - \sigma^{2'}(x)}{\mu'(x)}\right)\Bigg|_{x=\alpha_q(X)} \tag{4.34}$$

Where q is used instead of $1-q$ in the point of evaluation since the probability of default is monotonically increasing function of the state variable in the CR+ setting that is considered [2]. Taking the derivative of Equations (4.25) and (4.26) and plugging these expressions into (4.34) would yield an expression for the granularity adjustment in the CR+ setting. [22] however stress the importance of analytic tractability of their GA. Therefore, they introduce several new terms, of which the first is:

$$R_i = LGD_i PD_i \tag{4.35}$$

The second newly introduced term is the total amount of regulatory capital for obligor i as a proportion of its total exposure in the CR+ setting:

$$K_i = LGD_i PD_i \omega_i (\alpha_q - 1) \tag{4.36}$$

Where $\alpha_q \equiv \alpha_q(X)$. The final new term introduced is an adaption to our before derived A:

$$\delta_q(\epsilon) = -(\alpha_q - 1)\cdot A = (\alpha_q - 1)\left(\frac{1-\epsilon}{\alpha_q} + \epsilon\right) \tag{4.37}$$

In this setting, Equation (4.34) can be rewritten in:

$$GA = \frac{\delta_q(\epsilon)\sigma^2(\alpha_q) - (\alpha_q - 1)\sigma^{2'}(\alpha_q)}{2(\alpha_q - 1)\mu'(\alpha_q)} \tag{4.38}$$

Therefore, in order to finish the derivation of the GA in CR+ scenario, expressions for both the derivative of the conditional expectation and conditional variance with respect to x are required. Using the recently introduced notation, these terms simplify to the following:

$$(\alpha_q - 1)\mu'(\alpha_q) = \sum_{i=1}^{N} w_i K_i = K^* \tag{4.39}$$

$$(\alpha_q - 1)\sigma^{2'}(\alpha_q) = \sum_{i=1}^{N} w_i^2 K_i \left(C_i + 2(R_i + K_i)\frac{VLGD_i}{LGD_i^2}\right) \tag{4.40}$$

Combining Equation (4.38) together with the two previously derived expression, the full expression for the GA in the CR+ setting is given by:

$$GA_{CR+} = \frac{1}{2K^*}\sum_{i=1}^{N} w_i^2 \left(\left(\delta_q(\epsilon)C_i(R_i + Ki) + \delta_q(\epsilon)(R_i + Ki)^2\frac{VLGD_i}{LGD_i^2}\right) - K_i\left(C_i + 2(R_i + K_i)\frac{VLGD_i}{LGD_i^2}\right)\right) \tag{4.41}$$

For many purposes, including part of this thesis, the LGD parameter will be deterministic, $LGD \in [0, 1]$, which clearly yields $\mathbb{E}[LGD] = LGD$ and $VLGD_i = 0$ which in turn greatly simplifies Equation (4.41):

$$GA_{CR+} = \frac{1}{2K^*}\sum_{i=1}^{N} w_i^2 LGD\left(\delta_q(\epsilon)(R_i + Ki) - K_i\right) \tag{4.42}$$

Where we used the fact that for the deterministic case $C_i = \mathbb{E}[LGD] = LGD.$ In order to fully specify the adjustment, a value for shape parameter ϵ is required. [22] propose a value $\epsilon = 0.125.$

Contributing VaR for the CR+ Granularity Adjustment
Similar to Section 4.2.2, to fully assess single name concentration risk, the CR+ GA has to be derived on obligor level. This derivation is fairly straightforward as we can easily apply (3.39) to (4.41). One has to be cautious however, since $K^* = K^*(w_i)$:

$$w_i \frac{\partial GA_{CR+}}{\partial w_i} = w_i \frac{\partial}{\partial w_i}\left(\frac{1}{2K^*}\right)\sum_{n=1}^{N} w_n^2 \left(\left(\delta_q(\epsilon)C_n(R_n+K_n) + \delta_q(\epsilon)(R_n+K_n)^2\frac{VLGD_n}{LGD_n^2}\right)\right.$$
$$\left. -K_n\left(C_n + 2(R_n+K_n)\frac{VLGD_n}{LGD_n^2}\right)\right) + \frac{w_i}{K^*}\left(\left(\delta_q(\epsilon)C_i(R_i+Ki) + \delta_q(\epsilon)(R_i+K_i)^2\frac{VLGD_i}{LGD_i^2}\right)\right. \tag{4.43}$$
$$\left. -K_i\left(C_i + 2(R_i+K_i)\frac{VLGD_i}{LGD_i^2}\right)\right)$$

where we have:

$$\frac{\partial}{\partial w_i}\left(\frac{1}{2K^*}\right) = -\frac{K_i}{2(K^*)^2} \tag{4.44}$$

Referring back to Equation (3.54), one still needs an ASRF equivalent for the CreditRisk+ model. In the CreditRisk+ setting, this expression essentially equals (4.36) on obligor level [22]. This yields an opportunity to match the ASRF model with the CreditRisk+ model, by setting Equations (4.23) and (4.36) equal to each other. From this procedure, and expression can be derived that expresses ω_i as a function of β_i and confidence level q. For details about this procedure, we refer to chapter 6.3.7 of [2]. In this work, we will opt for a different approach in which we set $\omega_i = \beta_i$ and use the usual ASRF solution (3.54) to which the CR+ GA is added.

Remark
Two main limitations are identified for the above-proposed method. Firstly, the GA formula is itself an asymptotic approximation, limiting its applicability on very small portfolios. Secondly, since the GA is based on the CreditRisk+ model, it differs from the underlying methods in determining the IRB formulae. Therefore, there is a model mismatch. This issue is however solved by the GA for the single risk factor model derived previously. Additionally, the GA is derived in a single factor setting. Because of this, the GA only accounts for name concentration risk and is not able to assess sector concentration risk.

4.3. Multi Factor Concentration Adjustment

4.3.1. Pykhtin Multi-Factor Adjustment

In order to derive the explicit adjustment $\Delta\alpha_q$ Equation (3.72) we have to determine the conditional variance and mean of L given $\bar{X} = x$. Luckily, a large part of the work has been done in Section 4.2.1. Since $\mu(x)$ is given by Equation (3.59) we can adapt Equations (4.14) and (4.15) given that the original β_i is replaced by the newly derived correlation c_i. Therefore, it remains to determine conditional variance $\sigma^2(x)$ and its first derivative $\sigma^{2'}(x)$. The conditional variance is given by:

$$\sigma^2(x) := \mathbb{V}[L|\bar{X} = x] \tag{4.45}$$

In contrast to the single-risk factor framework, the defaults are not independent conditional on $x = \bar{X}$. This dependence becomes apparent if we rewrite the asset returns (3.61) in terms of the definitions of Y_i and $\bar{X}$:

$$r_i = c_i\bar{X} + \sum_{k=1}^{K}(\beta_i\alpha_{n,k} - c_i b_k)X_k + \sqrt{1-\beta_i^2}\epsilon_i \tag{4.46}$$

Although the asset returns are not independent conditional on $x = \bar{X}$, they are independent conditional on $\{X_1, ..., X_K\}$. Using the law of total variance, we can decompose the variance to include the conditional independence:

$$\sigma^2(x) = \underbrace{\mathbb{V}[\mathbb{E}[L|\{X_k\}]|\bar{X}]}_{\sigma^2_\infty(x)} + \underbrace{\mathbb{E}[\mathbb{V}[L|\{X_k\}]|\bar{X}]}_{\sigma^2_{GA}(x)} \tag{4.47}$$

The first term on the right-hand side, denoted as $\sigma^2_\infty(x)$, is the variance of the limiting loss distribution $\mathbb{E}[L|\{X_k\}]$ conditional on the effective systematic risk factor $\bar{X}$. It describes the systematic risk adjustment, accounting for the difference between the multi-factor and single factor loss distribution. This is further emphasized by the fact that it disappears if the single systematic factor $\bar{X}$ is set equal to the independent factors $\{X_1, ..., X_K\}$. The second term on the right-hand side, denoted as $\sigma^2_{GA}(x)$, captures the effect of granularity in the portfolio. Therefore, it accounts for the difference between a finite and infinite number of loans in the portfolio [35].

Proposition 4.3.1. *The conditional variance term $\sigma^2_\infty(x)$ is given by:*

$$\sigma^2_\infty(x) = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j LGD_i LGD_j [\Phi_2(\Phi^{-1}(PD_i(x)), \Phi^{-1}(PD_j(x)), \rho^{\bar{X}}_{ij}) - PD_i(x)PD_j(x)] \tag{4.48}$$

where $\Phi_2(\cdot)$ denotes the bivariate normal distribution and $\rho^{\bar{X}}_{ij}$ denotes the asset correlation between asset n and m conditional on $\bar{X}$ given by:

$$\rho^{\bar{X}}_{ij} = \frac{\beta_i\beta_j \sum_{k=1}^{K} \alpha_{n,k}\alpha_{m,k} - c_i c_j}{\sqrt{(1-c_i^2)(1-c_j)}} \tag{4.49}$$

Moreover, the derivative of $\sigma^2_\infty(x)$ is given by:

$$\frac{d}{dx}\sigma^2_\infty(x) = 2\sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j LGD_i LGD_j PD_i'(x)\left[\Phi\left(\frac{\Phi^{-1}(PD_j(x)) - \rho^{\bar{X}}_{ij}\Phi^{-1}(PD_i(x))}{\sqrt{1-(\rho^{\bar{X}}_{ij})^2}}\right) - PD_j(x)\right] \tag{4.50}$$

For the proof of this proposition we refer to the proof of Theorem 10.1.2 of [21]. Similarly, [35] shows:

Proposition 4.3.2. *The conditional variance term $\sigma^2_{GA}(x)$ is given by:*

$$\sigma^2_{GA}(x) = \sum_{i=1}^{N} w_i^2[(LGD_i^2 + VLGD_i)PD_i(x) - LGD_i^2 \cdot \Phi_2(\Phi^{-1}(PD_i(x)), \Phi^{-1}(PD_i(x)), \rho^{\bar{X}}_{nn})] \tag{4.51}$$

Its derivative equals

$$\frac{d}{dx}\sigma^2_{GA}(x) = \sum_{i=1}^{N} w_i^2 PD_i'(x)\left[(LGD_i^2 + VLGD_i) - 2LGD_i^2\Phi\left(\frac{\Phi^{-1}(PD_i(x)) - \rho^{\bar{X}}_{nn}\Phi^{-1}(PD_i(x))}{\sqrt{1-(\rho^{\bar{X}}_{nn})^2}}\right)\right] \tag{4.52}$$

For the proof of this proposition we refer to the proof of Theorem 10.1.3 of [21]. Exploiting the fact that the multi-factor adjustment (3.72) is linear in the conditional variance $\sigma^2(x) = \sigma^2_\infty(x) + \sigma^2_{GA}(x)$ and in its first derivative, the multi-factor adjustment can be rewritten in the following form:

$$\Delta\alpha_q = \Delta\alpha_q^\infty + \Delta\alpha_q^{GA} \tag{4.53}$$

Therefore, the multi-factor adjustment can be split into a systematic risk adjustment component and a granularity adjustment component. To summarize, the approximation of the loss quantile $\alpha_q(L)$ is given by:

$$\alpha_q(L) \approx \alpha_q(\bar{L}) + \Delta\alpha_q = \alpha_q(\bar{L}) + \Delta\alpha_q^\infty + \Delta\alpha_q^{GA} \tag{4.54}$$

In principle, implementation of the Pykhtin model as elaborated on above is fairly straightforward. Unfortunately, the computation of the analytic adjustment poses some challenges. The main problem is that, due to the double sums in Equations (4.48) and (4.50), computation can be extremely time-consuming when applying the adjustment to large portfolios. In addition to the double sums, the calculation of the conditional asset correlations requires $\frac{N^2}{2}$ computations. We will return to this issue in later sections.

Risk Contributions in the Pykhtin Setting
Similarly to the derivation of Section 4.2.2, the aforementioned theory enables us to compute the VaR analytically on portfolio level. We again apply the Euler allocation principle to compute the contributing VaR of each obligor. The individual terms in Equation (4.54) are given by:

$$\alpha_q(\bar{L}) = \mu(x)|_{x=\Phi^{-1}(1-q)} \tag{4.55}$$

$$\Delta\alpha_q^{GA} = \frac{1}{2}\left[\frac{x\cdot\sigma_{GA}^2(x) - \sigma_{GA}^{2'}(x)}{\mu'(x)} + \frac{\sigma_{GA}^2(x)\mu''(x)}{(\mu'(x))^2}\right]_{x=\Phi^{-1}(1-q)} \tag{4.56}$$

$$\Delta\alpha_q^{\infty} = \frac{1}{2}\left[\frac{x\cdot\sigma_{\infty}^2(x) - \sigma_{\infty}^{2'}(x)}{\mu'(x)} + \frac{\sigma_{\infty}^2(x)\mu''(x)}{(\mu'(x))^2}\right]_{x=\Phi^{-1}(1-q)} \tag{4.57}$$

where $\mu(x) = \sum_{i=1}^N w_i LGD_i \Phi(\frac{\Phi^{-1}(PD_i)-c_i X}{\sqrt{1-c_i^2}})$. The expressions for the derivative of $\mu(x)$ are similar to the derivatives presented in 4.2.1 with β_i replaced by the adjusted correlation factor c_i and therefore we will not repeat deriving these equations in the Pykhtin setting. Applying the Euler allocation principle to (4.54) yields:

$$w_i\frac{\partial\alpha_q(L)}{\partial w_i} = w_i\frac{\partial\alpha_q(\bar{L})}{\partial w_i} + w_i\frac{\partial\Delta\alpha_q^{\infty}}{\partial w_i} + w_i\frac{\partial\Delta\alpha_q^{GA}}{\partial w_i} \tag{4.58}$$

The first expression on the right hand side of this equation is easily calculated to equal

$$w_i\frac{\partial\alpha_q(\bar{L})}{\partial w_i} = w_i LGD_i \Phi\left(\frac{\Phi^{-1}(PD_i) - c_i\Phi^{-1}(1-q)}{\sqrt{1-c_i^2}}\right) \tag{4.59}$$

To determine the other two quantities on the right hand side of Equation (4.58) we notice that the adjustments (4.56) and (4.57) are of similar form to (4.21) and therefore the partial derivatives with respect to the weighted exposure are of the form (4.24):

$$w_i\frac{\partial\Delta\alpha_q^{\infty}}{\partial w_i} = w_i\frac{1}{2}\left[x\frac{\partial}{\partial w_i}\left(\frac{\sigma_{\infty}^2(x)}{\mu'(x)}\right) - \frac{\partial}{\partial w_i}\left(\frac{\sigma_{\infty}^{2'}(x)}{\mu'(x)}\right) + \frac{\partial}{\partial w_i}\left(\frac{\sigma_{\infty}^2(x)\mu''(x)}{(\mu'(x))^2}\right)\right]_{x=\Phi^{-1}(1-q)} \tag{4.60}$$

$$w_i\frac{\partial\Delta\alpha_q^{GA}}{\partial w_i} = w_i\frac{1}{2}\left[x\frac{\partial}{\partial w_i}\left(\frac{\sigma_{GA}^2(x)}{\mu'(x)}\right) - \frac{\partial}{\partial w_i}\left(\frac{\sigma_{GA}^{2'}(x)}{\mu'(x)}\right) + \frac{\partial}{\partial w_i}\left(\frac{\sigma_{GA}^2(x)\mu''(x)}{(\mu'(x))^2}\right)\right]_{x=\Phi^{-1}(1-q)} \tag{4.61}$$

Due to its familiar form, we can use previously derived Equations (A.5) through (A.10) with β_i replaced by c_i. However, due to form of the conditional variance terms $\sigma_{\infty}^2(x)$ and $\sigma_{GA}^2(x)$, its partial derivatives take a different form:

$$\frac{\partial\sigma_{\infty}^2(x)}{\partial w_i} = 2LGD_i\sum_{j=1}^n w_j LGD_j\left[\Phi_2(\Phi^{-1}(PD_i(x)),\Phi^{-1}(PD_j(x)),\rho_{ij}^{\bar{X}}) - PD_i(x)PD_j(x)\right] \tag{4.62}$$

$$\frac{\partial\sigma_{GA}^2(x)}{\partial w_i} = 2w_i\left[(LGD_i^2 + VLGD_i)PD_i(x) - LGD_i^2\Phi_2(\Phi^{-1}(PD_i(x)),\Phi^{-1}(PD_i(x)),\rho_{ii}^{\bar{X}})\right] \tag{4.63}$$

$$\begin{aligned}\frac{\partial\sigma_{\infty}^{2'}(x)}{\partial w_i} &= 2LGD_i\sum_{j=1}^N w_j LGD_j PD_j'(x)\left[\Phi\left(\frac{\Phi^{-1}(PD_i(x)) - \rho_{ji}^{\bar{X}}\Phi^{-1}(PD_j(x))}{\sqrt{1-(\rho_{ji}^{\bar{X}})^2}}\right) - PD_i(x)\right]\\ &+2LGD_i\sum_{j=1}^N w_j LGD_j PD_i'(x)\left[\Phi\left(\frac{\Phi^{-1}(PD_j(x)) - \rho_{ji}^{\bar{X}}\Phi^{-1}(PD_i(x))}{\sqrt{1-(\rho_{ji}^{\bar{X}})^2}}\right) - PD_j(x)\right]\end{aligned} \tag{4.64}$$

$$\frac{\partial \sigma_{GA}^{2'}(x)}{\partial w_i} = 2w_i PD_i'(x)\left[(LGD_i^2 + VLGD_i) - 2LGD_i^2\Phi\left(\frac{\Phi^{-1}(PD_i(x)) - \rho_{ii}^{\bar{X}}\Phi^{-1}(PD_i(x))}{\sqrt{1-(\rho_{ii}^{\bar{X}})^2}}\right)\right] \tag{4.65}$$

These equations show great similarities with the equations presented in [17] and [34], however, Equations (4.62) through (4.65) account for random LGD's whereas the expressions derived in the two aforementioned works do not. This implies that the original derivation by [35] were unable to account for the effects of recovery risk, whereas in this work we assume recovery risk through Beta distributed LGD rates.

4.4. An Extension: Economic versus Regulatory Name Concentration Risk

Regulatory Name Concentration Risk

Currently, a fairly widespread and agreed upon measure of single name concentration risk is the granularity adjustment in the single factor case, essentially Equation (3.55). More generally, name concentration risk is measured as the difference in VaR between the true, MC based, VaR resulting from a single factor model and its asymptotic VaR:

$$CR_q(L_N) := VaR_q^{SF}(L_N) - \lim_{N\to\infty} VaR_q^{SF}(L_N) \tag{4.66}$$

Given the assumption of a single factor, we have seen that the asymptotic portfolio VaR $\lim_{N\to\infty} VaR_q^{SF}(L_N)$ can analytically be determined by the ASRF model. Therefore, Equation (4.66) reduces to:

$$CR_q(L) = VaR_q^{SF}(L) - VaR_q^{ASRF}(L) \tag{4.67}$$

For the concentration risk contributions of individual obligors we have:

$$CRC_{q,i}(L) = VaRC_{q,i}^{SF} - VaRC_{q,i}^{ASRF} \tag{4.68}$$

Both $VaR_q^{SF}(L)$ and $VaRC_{q,i}^{SF}$ are a result of numerous Monte Carlo simulations on a single systematic factor and N idiosyncratic factors through the use of the single factor variant of algorithm 1. The expression for $VaR_q^{ASRF}(L)$ is given by Equation (2.25) and the contributing $VaRC_{q,i}^{ASRF}$ for each obligor is easily determined through apply the Euler Allocation principle (3.39) on Equation (2.25):

$$VaRC_{q,i}^{ASRF} = w_i \frac{\partial VaR_q(L)^{ASRF}}{\partial w_i} = w_i \cdot LGD_i \cdot \Phi\left(\frac{\Phi^{-1}(PD_i) + \beta_i\Phi^{-1}(q)}{\sqrt{1-\beta_i^2}}\right) \tag{4.69}$$

Notice that $CR_q(L)$ is equal in terms of VaR and in terms of EC since $VaR_q^{SF}(L) - \mathbb{E}[L] - (VaR_q^{ASRF}(L) - \mathbb{E}[L]) = VaR_q^{SF}(L) - VaR_q^{ASRF}(L)$. Clearly, for a infinitely fine grained portfolio the concentration charge $CR_q(L)$ equals zero since $\lim_{N\to\infty} VaR_q^{SF}(L) = VaR_q^{ASRF}(L)$ by construction of the Asymptotic Single Risk Factor approximation. However, for real life portfolio we expect $CR_q(L)$ to be larger than zero. Furthermore, by construction of the Granularity Adjustment (i.e. Equation (3.54)), we expect $CR_q(L) \approx GA_q(L)$. The concentration charge $CR_q(L)$ will slightly differ from the analytical GA since $CR_q(L)$ is a result of a numerical approximation of the VaR and the GA is a analytical approximation based on a second order Taylor expansion around the VaR. However, since banks often have access to Monte Carlo methods for determining VaR, practitioners may opt for the numerical concentration charge (4.67) instead of the analytical GA. A major upside to the numerical concentration charge is the fact that it is based on existing MC methods and that therefore $CR_q(L)$ always equals the difference between known regulatory capital requirements and the ASRF approximation whereas the GA, being an approximation, does not always equal this difference. However, calculating the numerical concentration charge is more time consuming and computationally intensive compared to the straightforward calculations required for calculating the GA (i.e. (4.20)).

However, we argue that Equation (4.67) only holds in single factor regulatory A-IRB framework. This means using a single global systematic factor, regulatory correlation parameter ρ (i.e. Equation

(2.27)) and using the LGD, EAD and PD determined by the bank's internal risk models. The regulatory correlation parameters are adjusted to the single risk factor model, whereas economic capital is often calculated in a multi factor model to account for diversification effects. In a multi factor model, the correlation parameter β is determined through a regression on multiple systematic factors and therefore differs from ρ. Moreover, the capital requirements from a multi factor model might either exceed or be less than the capital requirements for the single factor model depending on the correlation structure that allows for diversification benefits. This effect is depicted in Figure 4.1 where we compare the loss distribution of a heterogeneous portfolio of a 1000 obligors. One loss distribution is based on a single factor model whereas the other is based on a multi factor model with equal EAD, PD, LGD and β parameters. Clearly, the loss distribution differs greatly between the multi factor and the single factor model. Measuring single name concentration risk using with respect to the multi factor loss distribution by plainly using the ASRF solution yields senseless results because implicitly, one would be measuring the concentration risk with respect to the single factor variant of the multi factor model. Therefore we propose a difference between *Regulatory Concentration Risk* (RCR) and *Economic Concentration Risk* (ECR).

In short, Regulatory Concentration Risk equals the residual idiosyncratic risk present in a single factor regulatory framework, where the asymptotic VaR is estimated by the usual ASRF solution and the true VaR is a result of numerous MC trials. The GA is developed to approximate RCR, but we propose a method for assessing RCR in a Monte Carlo framework. Although being a relevant risk measure, as it effectively equals the add-on risk that the A-IRB method does not take into account under Pillar I, it does not take the effects of diversification into account. To this extend, we introduce Economic Concentration Risk.

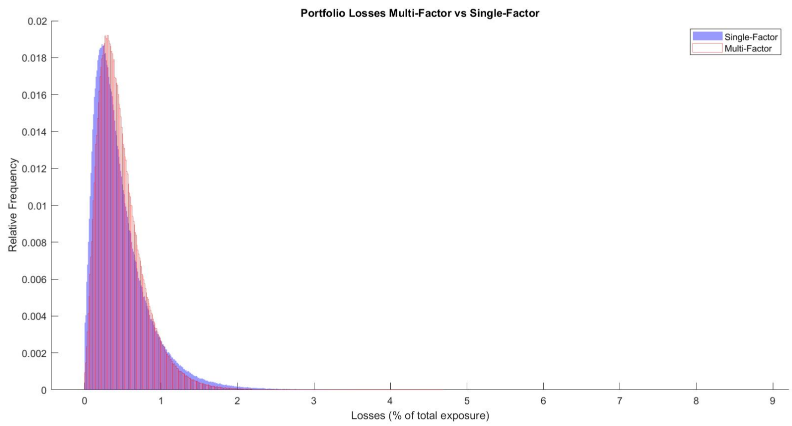

Figure 4.1: Histogram of portfolio losses based on a corporate sub portfolio of ING of 1000 obligors and 1e6 MC trials. One can see clearly that the portfolio loss distribution strongly differs across the used methods with the single factor model seemingly exhibiting a heavier tail, indicating higher capital requirements. The expected loss is equal across both models.

Economic Name Concentration Risk

In this section we assume that a banks measures their EC in a multi factor framework to stress the difference with a single factor regulatory framework. Essentially, we are looking for the following expression for the economic concentration charge:

$$ECR_q(L_N) = VaR_q^{MF}(L_N) - \lim_{N\to\infty} VaR_q^{MF}(L_N) \tag{4.70}$$

Where the superscript MF denotes multi-factor. This expression poses a problem, as we cannot use the tools used in Section 2.6.1 to compute $\lim_{N\to\infty} VaR_q^{MF}(L)$ as assumption 2.6.3 is violated by introducing multiple risk factors. Instead, we propose measuring $ECR_q(L)$ by adjusting a single factor model in such a way it approximates a multi factor model and then compute $ECR_q(L)$ by subtracting the ASRF solution of this adjusted single factor model. The question that remains is how to accurately adjust a multi factor model to a single factor model. This would involve mapping the correlation structure of the systematic factors to a scaling parameter for the single factor model. To this extend, we apply the theory developed by [35] and previously described in Section 3.7.5. In short, we adjust a single factor loss distribution in such a way that $L^{MF} = \tilde{L}^{A\text{-}SF}$. Essentially, we start with the general multi model for asset returns:

$$r_i = \beta_i Y_i + \sqrt{1-\beta_i^2}\epsilon_i \tag{4.71}$$

in which Y_i is the counter-party's composite factor:

$$Y_i = \sum_{k=1}^{K} \alpha_{i,k} X_k \tag{4.72}$$

Where the parameters $\alpha_{i,k}$ are based on a combination of the factor weights of obligor i and the Cholesky decomposition of the covariance matrix of the systematic factors (see Section 4.1.1). As usual, the random variable describing the portfolio loss is then given by:

$$L^{MF} = \sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot \mathbb{1}_{r_i < \Phi^{-1}(PD_i)} \tag{4.73}$$

$\tilde{L}^{A\text{-}SF}$ satisfies a very similar expression:

$$\tilde{L}^{A\text{-}SF} = \sum_{i=1}^{n} EAD_i \cdot LGD_i \cdot \mathbb{1}_{\tilde{r}_i < \Phi^{-1}(PD_i)} \tag{4.74}$$

In order to satisfy the constraint $L^{MF} = \tilde{L}^{A\text{-}SF}$, the following equality has to be satisfied:

$$r_i = \tilde{r}_i \tag{4.75}$$

To satisfy this constraint, the following adjusted single factor model is introduced:

$$\tilde{r}_i = c_i X + \sqrt{1-c_i^2}\epsilon_i \tag{4.76}$$

Parameter c_i is derived completely in line with Section 3.7.5 by maximizing the correlation between X and Y_i which leads to the following expressions:

$$c_i = \beta_i \sum_{k=1}^{K} \alpha_{i,k} b_k \tag{4.77}$$

$$b_k = \sum_{i=1}^{N} \frac{d_i \alpha_{i,k}}{2\lambda} \tag{4.78}$$

$$d_i = w_i \cdot LGD_i \cdot \Phi\left(\frac{\Phi^{-1}(PD_i) - \beta_i \Phi^{-1}(1-q)}{\sqrt{1-\beta_i^2}}\right) \tag{4.79}$$

To empirically show the performance of the adjusted single factor model, compare Figures 4.1 and 4.2 in which the same multi factor loss distribution is depicted but differing single factor models. In Figure 4.1 we naively apply multi factor parameter β_i whereas in Figure 4.2 parameter c_i is applied. Empirically, Figure 4.2 shows that the adjusted single factor model coincides with the original multi factor model for this specific heterogeneous portfolio.

However, Figure 4.2 does not sufficiently depict the tail behaviour of the distribution, which is mainly of interest in terms of credit risk assessments. Figure 4.3 shows that the performance of the model is high across all quantiles of the distribution. Furthermore, from Figure 4.4 it is clear that again, the performance across the frequently used VaR confidence intervals is accurate.

The upside of this method is that known methods, developed for the usual single risk factor model, can be applied to this adjusted form such as the Asymptotic Single Risk Factor approximation. Applying this, similarly to Equation (4.67), Economic Concentration Risk on portfolio level is defined by:

$$ECR_q(L) := VaR_q^{MF}(L) - VaR_q^{A\text{-}ASRF}(L) \tag{4.80}$$

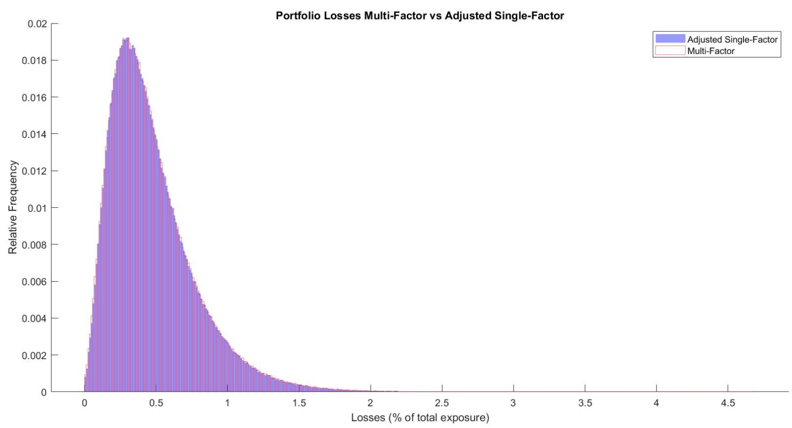

Figure 4.2: Histogram of portfolio losses based on a corporate sub portfolio of ING of 1000 obligors and 1e6 MC trials, similar to the portfolio of Figure 4.1. One can clearly see that the portfolio loss distributions of the two models seemingly overlap.

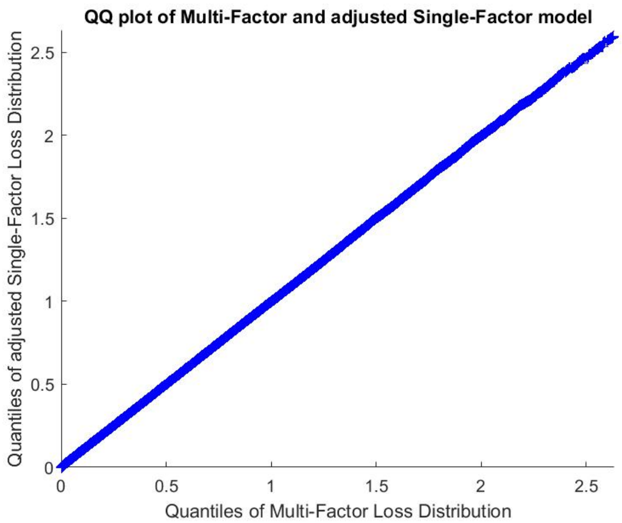

Figure 4.3: QQ plot (quantiles expressed as % of total exposure) of the portfolio loss distributions of the adjusted Single Factor and Multi Factor model of Figure 4.2. The clustering of the data points around the 45° line indicates that the quantiles of the loss distributions of the two different models agree across all quantile levels.

Equivalently, we can define the Economic Concentration Risk Contribution as:

$$ECRC_{q,i}(L) = VaRC_{q,i}^{MF} - VaRC_{q,i}^{A\text{-}ASRF} \tag{4.81}$$

Where the superscript MF denotes Multi Factor and $A\text{-}ASRF$ Adjusted ASRF. Using the framework developed by [9] we can apply assumptions 2.6.1 through 2.6.4 together with proposition 2.6.2 to derive an expression for the $VaR_q^{A\text{-}ASRF}(L)$:

$$VaR_q^{A\text{-}ASRF}(L) = \sum_{i=1}^{N} w_i \cdot LGD_i \cdot \Phi\left(\frac{\Phi^{-1}(PD_i) + c_i\Phi^{-1}(q)}{\sqrt{1-c_i^2}}\right) \tag{4.82}$$

$$VaRC_{q,i}^{A\text{-}ASRF} = w_i\frac{\partial VaR_q(L)^{A\text{-}ASRF}}{\partial w_i} = w_i \cdot LGD_i \cdot \Phi\left(\frac{\Phi^{-1}(PD_i) + c_i\Phi^{-1}(q)}{\sqrt{1-c_i^2}}\right) \tag{4.83}$$

These results are combined in Figure 4.4 which both depicts the effect of undiversified idiosyncratic risk and the excellent performance of the adjusted single risk factor model compared to the full multi factor model in the tail of the loss distribution.

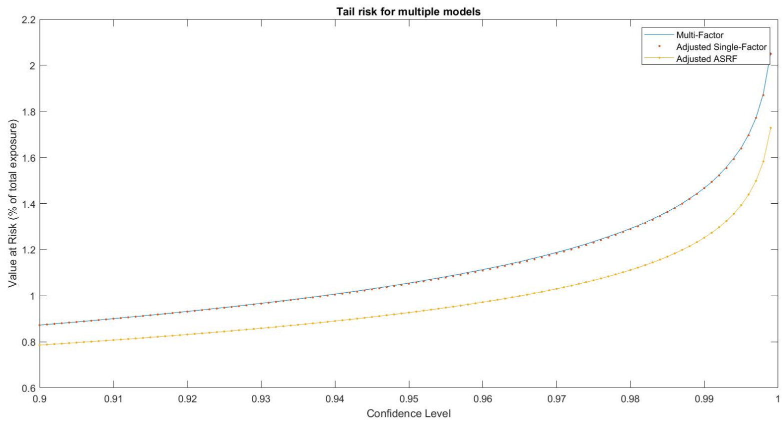

Figure 4.4: This graph depicts the portfolio VaR at quantile levels ranging from 95% upwards based on the portfolio loss distribution of 4.2. Clearly, the VaR of the adjusted Single Factor and Multi Factor model agree to a high degree of accuracy. The gap between the Multi Factor VaR and the adjusted ASRF indicates the amount of economic concentration risk. Parameter c_i is calibrated at $q = 0.999$.

4.5. An Extension: Sector and Geographical Diversification

Definition 2.7.2 states that sector concentration risk stems from the existence of multiple systematic factors and arises from the assumption of a single underlying risk factor. In order to quantify the effect of diversification, the following proposition will prove to be useful:

Proposition 4.5.1. *Assume multi factor model r_i^{MF} with perfectly positive correlated systematic factor, meaning that for any pair of systematic factors (X_i, X_j) we have $Corr(X_i, X_j) = 1$ or in other words $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{J}_{K\times K})$ where $\boldsymbol{J}_{K\times K}$ denotes a matrix of ones. Then r_i^{MF} has an equivalent single factor model $r_i^{E\text{-}SF}$ such that $r_i^{MF} \sim r_i^{E\text{-}SF}$.*

Proof. Firstly, let $r_i^{MF} = \beta_i Y_i + \sqrt{1-\beta_i^2}\epsilon_i$ with $Y_i = \sum_{k=1}^{K} w_{i,k} X_k$ bounded to the constraint $\sum_{k=1}^{K} w_{i,k} = 1$ meaning that the factor weights of counterparty i on systematic factors $X_1, ..., X_K$ sum to one. Since the systematic factors are perfectly correlated, we have $X_1 = X_2 = ... = X_K \equiv X$. Therefore, $Y_i = \sum_{k=1}^{K} w_{i,k} X_k$ reduces to $Y_i = \sum_{k=1}^{K} w_{i,k} X = X \cdot \sum_{k=1}^{K} w_{i,k} = X$. Clearly, this fact yields $r_i^{MF} = \beta_i X + \sqrt{1-\beta_i^2}\epsilon_i$ which is an equivalent single factor model $r_i^{E\text{-}SF}$. □

Essentially, this proposition states that a perfectly correlated multi factor model reduces to a single factor model with identical systematic factor loading β_i. Empirically, we can depict the effect of the proposition by comparing the quantiles of the loss distributions of the two models, simulated in a Monte Carlo setting. This experiment is performed in Figure 4.5. We see that, besides some simulation noise, the quantiles of the two factor models are equivalent.

Proposition 4.5.1 yields a way to measure the effect of systematic diversification for both the full portfolio and individual obligors. The perfectly correlated multi factor model introduced in proposition 4.5.1 has one major implication, namely a lack of diversification benefits. Generally, banks employ multi factor models for more accurate risk assessments but also to profit from the diversification benefits these models possibly imply. The diversification effects stem from the correlation matrix describing the pairwise dependency of the systematic factors. In a single factor model, all counterparties are 100% dependent from the global systematic factor, whereas in a multi factor setting, obligors can have an arbitrary factor loading on a arbitrary amount of systematic factors. Clearly, this greatly increases the effects of diversification. To measure the effect of diversification, we introduce the Diversification Factor:

$$DF = \frac{EC^{MF}}{EC^{E\text{-}SF}}, \; DF \leq 1 \tag{4.84}$$

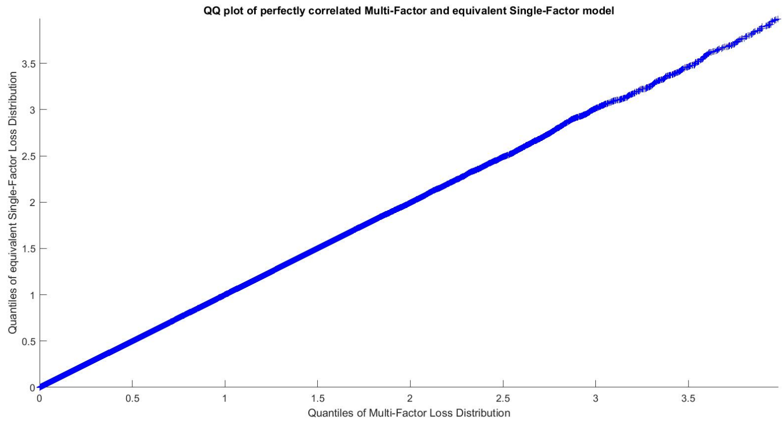

Figure 4.5: QQ plot of the portfolio loss distributions of the perfectly correlated MF and E-SF model, based on a ING corporate sub-portfolio of 1000 obligors and 1e6 MC trials. Clearly, the quantiles of the two models agree which indicates that the loss distributions of the two models are equivalent.

The idea of a diversification factor of this form is not necessarily new, as [5] introduced a definition of similar form. However, [5] introduce it as an adjustment to a single factor model to account for the effects of diversification and numerically approximate it as a function of two parameters that capture the sector concentration and the average cross-sector correlation. We introduce it as a relatively easily observable measure to assess the effects of diversification in the portfolio and thoroughly substantiate the measure through the use of proposition 4.5.1. Through assuming perfect correlation, $\boldsymbol{\Omega} = \boldsymbol{J}$, we assume that all systematic factors vary perfectly in line with one another. This implies that the portfolio risk is insensitive to the sectors, as for instance, an exposure in sector *1* is equally influenced by the systematic factors as an equivalent exposure in sector *2*. The benefits of diversification are therefore best expressed as the discrepancy between the EC of the perfectly correlated and realistically correlated model, which by applying proposition 4.5.1, equals the discrepancy between the multi factor and its equivalent single factor model based EC.

Equation (4.84) produces a measure of diversification, whereas we are also looking to quantify the effect of diversification within subsets of the portfolio. To that extend, consider a portfolio of N loans to unique obligors. The obligors can be assigned to M different industries (or geographical regions). Additionally, denote N_s as the number of obligors in sector $s = 1, \dots, M$. Thus, the total number of obligors in the portfolio equals $N = \sum_{s=1}^{M} N_s$. The diversification factor for sector s is then given by:

$$DF_s = \frac{EC_s^{MF}}{EC_s^{E\text{-}SF}} \text{ , } DF_s \leq 1 \tag{4.85}$$

Where EC_s^{MF} is the economic capital assigned to sector s, based on the Euler allocation principle. We will additionally define the Capital Diversification Index (CDI):

$$CDI = \frac{\sum_{i=1}^{S_s} EC_i^2}{EC_s^{MF}} \tag{4.86}$$

With S_s denoting the amount of exposure with a sector s, essentially $N = \sum_{s=1}^{M} S_s$. The CDI was also first introduced by [5], however, in a slightly different form and again with a different purpose. In [5], the CDI is based on the single factor capital requirements instead of the multi factor capital requirements that take the effects of diversification into account. Basically, the CDI is the HH index, but instead of applying the HHI at exposure level, the HHI is applied at economic capital level. Therefore, the CDI indicates the level EC concentration within a certain sector (either an industry or a region). A low value of the CDI indicates that the EC within the sector is fairly well distributed among the obligors in that sector, whereas a high value of the CDI indicates the contrary.

4.6. An Extension: The *t*-Threshold Model

A major drawback of the general multi factor model described in Section 2.5 is its assumption of standard normally distributed asset returns. However, empirical investigations of asset returns, for example by [24] and [20], reject the normal distribution because of its inability to mimic the empirical fat tails of empirically observed asset returns. [20] suggests using the *t* distribution for modelling asset returns. In this section we will introduce the multi factor t-threshold model and apply aforementioned theory on equivalent and adjusted single factor models to this multi factor t-threshold model. Furthermore, we will derive an ASRF equivalent for an independent *t*-threshold model, which differs from the general *t*-threshold model.

Firstly, we will start by introducing the general t-threshold model:

$$\overline{r}_i = \sqrt{\frac{\eta}{W}}\left[\beta_i Y_i + \sqrt{1-\beta_i^2}\epsilon_i\right] = \beta_i\sqrt{\frac{\eta}{W}}Y_i + \sqrt{1-\beta_i^2}\sqrt{\frac{\eta}{W}}\epsilon_i \tag{4.87}$$

in which Y_i is the counter-party's composite factor:

$$Y_i = \sum_{k=1}^{K}\alpha_{i,k}X_k \tag{4.88}$$

and where $X_k, Y_i, \epsilon_i \sim \mathcal{N}(0,1)$. In addition to the general multi factor model, the scaling factor $\sqrt{\frac{\eta}{W}}$ is introduced in Equation (4.87). In this scaling factor, W denotes the chi-squared distribution, $W \sim \chi^2(\eta)$, and η the degrees of freedom of the chi-squared distribution. Furthermore, we impose that W is independent of X_k, Y_i, ϵ_i. Through scaling factor $\sqrt{\frac{\eta}{W}}$, the standard normally distributed asset returns r_i are transformed into *t* distributed asset returns $\overline{r}_i$ with η degrees of freedom, essentially $\overline{r}_i \sim t(\eta)$. Note that the asset returns are still correlated through the factor loadings $\alpha_{i,k}$. Furthermore, we denote the cumulative distribution function of the *t* distribution as $F(x)$. Analogous to the normally distributed threshold factor model, a default threshold has to be identified.

PD_i denotes the one-year probability of default for counter-party i: $PD_i = \mathbb{P}[\overline{r}_i < c_i]$ and since we assume that $\overline{r}_i \sim t(\eta)$ we obtain:

$$c_i = F^{-1}(PD_i) \tag{4.89}$$

Rewriting the condition $\overline{r}_i < c_i$ yields:

$$\epsilon_i < \frac{F^{-1}(PD_i) - \beta_i Y_i}{\sqrt{1-\beta_i^2}} \tag{4.90}$$

Therefore, since ϵ_i is standard normally distributed, the probability of default conditional on the systematic factors and W can be written as:

$$PD_i(Y_i, W) = \mathbb{P}[\overline{r}_i < c_i] = \mathbb{P}\left[\epsilon_i < \frac{\sqrt{\frac{\eta}{W}}F^{-1}(PD_i) - \beta_i Y_i}{\sqrt{1-\beta_i^2}}\right] = \Phi\left(\epsilon_i < \frac{\sqrt{\frac{\eta}{W}}F^{-1}(PD_i) - \beta_i Y_i}{\sqrt{1-\beta_i^2}}\right) \tag{4.91}$$

Unfortunately, due to its dependence on W, the t-threshold model does not readily admit an asymptotic single risk factor solution. In the original *t*-threshold model of (4.87), the scaling factor $\sqrt{\frac{\eta}{W}}$ introduces a dependence between the systematic and idiosyncratic risk. Whereas Y_i and ϵ_i are independent, $\sqrt{\frac{\eta}{W}}Y_i$ and $\sqrt{\frac{\eta}{W}}\epsilon_i$ are clearly not. However, we can impose independence by assuming the following single factor model:

$$\hat{r}_i = \beta_i\sqrt{\frac{\eta}{W}}X + \sqrt{1-\beta_i^2}\sqrt{\frac{\eta}{W_i}}\epsilon_i \tag{4.92}$$

As both Y and ϵ_i are standard normal random variables, scaling them with $\sqrt{\frac{\eta}{W}}$ and $\sqrt{\frac{\eta}{W_i}}$ respectively yields *t* distributed random variables: $T = \sqrt{\frac{\eta}{W}}X \sim t(\eta)$ and $\zeta_i = \sqrt{\frac{\eta}{W_i}}\epsilon_i \sim t(\eta)$. Using these expressions, we can rewrite Equation (4.92):

$$\hat{r}_i = \beta_i T + \sqrt{1 - \beta_i^2}\zeta_i \tag{4.93}$$

where T and ζ_i are independent. We will refer to this model as the *indepedent single factor t-threshold model*. Equation (4.93) shows great similarities with its Gaussian single factor threshold counterpart which raises the question of the existence of an asymptotic result for the VaR in this setting. The default condition $\hat{r}_i < F^{-1}(PD_i)$ yields

$$\zeta_i < \frac{F^{-1}(PD_i) - \beta_i T_i}{\sqrt{1 - \beta_i^2}} \tag{4.94}$$

With conditional default probability:

$$PD_i(T) = F\left(\frac{F^{-1}(PD_i) - \beta_i T}{\sqrt{1 - \beta_i^2}}\right) \tag{4.95}$$

Referring back to assumptions 2.6.1 through 2.6.4, no assumption is made on the distribution of either the asset returns or the systematic factor. Therefore, this adapted single factor t-threshold model satisfies all assumptions and we can easily deduce the adapted t-threshold ASRF solution by applying proposition 2.6.2 through assuming that the only source of risk is systematic:

$$VaR_q^{t\text{-}ASRF} = \sum_{i=1}^{N} w_i \cdot LGD_i \cdot F\left(\frac{F^{-1}(PD_i) - \beta_i F^{-1}(1-q)}{\sqrt{1 - \beta_i^2}}\right) \tag{4.96}$$

Equation (4.96) exhibits many similarities with its normally distributed single factor counterpart. Furthermore, as we have seen in Section 3.1.1, the *t*-distribution converges to the normal distribution as the degrees of freedom increase (see Figure 4.6). We should however stress that Equation (4.96) is an asymptotic solution to (4.93) and not to (4.87), and therefore applying it as a measure of systematic risk in a portfolio results in a model mismatch. The existence of the ASRF solution in the independent

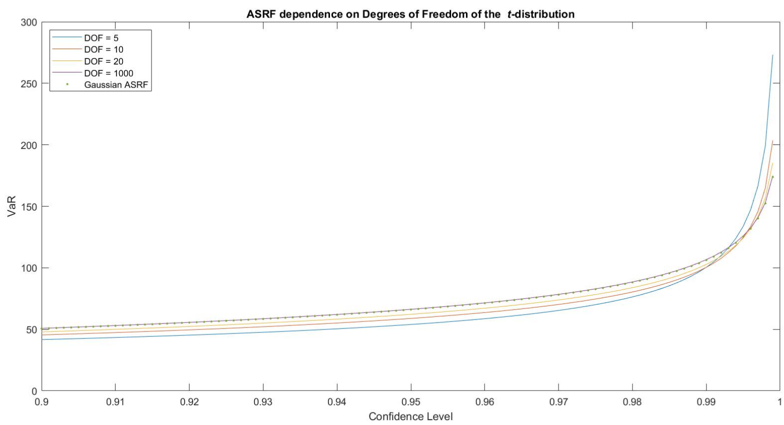

Figure 4.6: Graph depicting the convergence of independent *t*-threshold ASRF solution to the Gaussian threshold ASRF solution as the degrees of freedom of the underlying *t*-distribution increases. The VaR levels displayed are based on a 100 obligor heterogeneous sample portfolio.

single factor t-threshold setting and its similarities with its Gaussian threshold equivalent, raises the question of the existence for a GA in this setting. Intuitively, the GA would be an add on to t-ASRF VaR (4.96) such that the sum of (4.96) and GA approximates the true VaR that results from the independent single factor t-threshold model (4.93). To this extend, we can apply the theory developed in Section 3.7.2 to this new setting. Essentially, since we did not make any assumptions on the distribution of the single systematic factor in the derivation of Equation (3.55), the GA is given by:

$$GA_t = -\frac{1}{2}\left[\frac{f_T'(t)}{f_T(t)}\frac{\mathbb{V}[L|T=t]}{\frac{d}{dt}\mathbb{E}[L|T=t]} + \frac{\frac{d}{dt}\mathbb{V}[L|T=t]}{\frac{d}{dt}\mathbb{E}[L|T=t]} - \frac{\mathbb{V}[L|T=t]\frac{d^2}{dt^2}\mathbb{E}[L|T=t]}{(\frac{d}{dt}\mathbb{E}[L|T=t])^2}\right]_{t=\alpha_{1-q}(T)} \tag{4.97}$$

where T indicates the single t distributed systematic factor, $T = \sqrt{\frac{\eta}{W}}X \sim t(\eta)$. Firstly, we recall the probability density function of a t distributed random variable:

$$f_T(t) = \frac{\Gamma(\frac{\eta+1}{2})}{\sqrt{\pi\eta}\Gamma(\frac{\eta}{2})}\left(1+\frac{t^2}{\eta}\right)^{-\frac{\eta+1}{2}} \tag{4.98}$$

Which results in:

$$f_T'(t) = -t\frac{\eta+1}{\eta}\left(1+\frac{t^2}{\eta}\right)^{-1} f_T(t) \tag{4.99}$$

From which we have

$$\frac{f_T'(t)}{f_T(t)} = -t\frac{\eta+1}{\eta}\left(1+\frac{t^2}{\eta}\right)^{-1} \tag{4.100}$$

Within this new setting, we will define loss exactly in line with Equation (2.4), only the underlying asset returns and default threshold are defined slightly different. Therefore, we can adapt Equations (4.12) through (4.16) by replacing $PD_i(X)$ and its derivatives by $PD_i(T)$ and its derivatives:

$$PD_i'(t) = -\frac{\beta_i}{\sqrt{1-\beta_i^2}} f_T(u_i) \tag{4.101}$$

$$PD_i''(t) = -u_i\frac{\beta_i^2}{\sqrt{1-\beta_i^2}}\frac{\eta+1}{\eta}\left(1+\frac{u^2}{\eta}\right)^{-1} f_T(u_i) \tag{4.102}$$

$$(PD_i(t)^2)' = -2\frac{\beta_i}{\sqrt{1-\beta_i^2}} F_T(u_i) f_T(u_i) \tag{4.103}$$

where $u_i = \frac{F^{-1}(PD_i)-\beta_i t}{\sqrt{1-\beta_i^2}}$. Notice that, contrary to the Gaussian framework, Equations (4.101) through (4.103) are dependent on degrees of freedom η, which introduces an extra parameter to the GA compared to its Gaussian counterpart. The remaining part of the derivation of GA_t and the risk contributions associated with GA_t is analogous to the derivation for the single factor Gaussian threshold model of Sections 4.2.1 and 4.2.2 and will therefore not be repeated in this section.

As we have seen in Figure 4.6, the t-ASRF solution converges to its Gaussian counterpart as the degrees of freedom tend to infinity. For the GA_t, we expect a similar result as η tends to infinity. By the properties of the t-distribution, $f_T(x) \to \phi(x)$ and $F_T(x) \to \Phi(x)$ as $\eta \to \infty$. Therefore, $u_i \to z_i$ with $z_i = \frac{\Phi^{-1}(PD_i)-\beta_i x}{\sqrt{1-\beta_i^2}}$ as $\eta \to \infty$. Additionally, since $\frac{\eta+1}{\eta}\left(1+\frac{u^2}{\eta}\right)^{-1} \to 1$ as $\eta \to \infty$, all the individual components of the GA_t converge to the individual components of the Gaussian GA and therefore, $GA_t \to GA$, as $\eta \to \infty$.

Adjusted *t*-Threshold Model

In the previous section, we derived the general *t*-threshold multi factor model. In order to apply the theory developed in Section 4.4 on economic concentration risk, we need to derive an adjusted single factor *t*-threshold model that mimics the results of the multi factor *t*-threshold model Equation (4.87). Unfortunately, to our knowledge, no theory has been developed for this specific model. However, the approach developed by [35] performs excellently in the Gaussian setting. Therefore, we will try to adopt this method to the *t*-distributed setting.

Since the *t*-threshold model is essentially a scaled version of the Gaussian-threshold model, we can largely adopt the methods from Section 3.7.5. We will derive an adjusted single factor model $\overline{r}_i^{SF}$ such that $\overline{r}_i \sim \overline{r}_i^{SF}$ in which we have

$$\overline{r}_i^{SF} = \sqrt{\frac{\eta}{W}}\left[c_i\bar{X} + \sqrt{1-c_i^2}\epsilon_i\right] \tag{4.104}$$

in which $\bar{X} \sim \mathcal{N}(0,1)$ is the single systematic factor. Analogous to Gaussian case, $\bar{X}$ is linked to the K original systematic factors X_k in the following way:

$$\bar{X} = \sum_{k=1}^{K} b_k X_k \tag{4.105}$$

Where $\sum_{k=1}^{K} b_k^2 = 1$ to preserve unit variance of $\bar{X}$. This reduces our problem to specifying b_k for $k \in (1,\dots,K)$ and c_i for $i \in (1,\dots,N)$. To this extend, we make the following assumption:

$$Y_i = \delta_i\bar{X} + \sqrt{1-\delta_i^2}\zeta_i \tag{4.106}$$

Where again, $\zeta_i \sim \mathcal{N}(0,1)$. Hence, the original systematic factor Y_i is driven by systematic and idiosyncratic risk, weighted by δ_i. δ_i is set to be the correlation between the original systematic factors and the new single systematic factor:

$$\delta_i = \mathrm{Corr}(Y_i,\bar{X}) = \mathrm{Corr}\left(\sum_{k=1}^{K}\alpha_{i,k}X_k, \sum_{k=1}^{K} b_k X_k\right) = \sum_{k=1}^{K}\alpha_{i,k}b_k \tag{4.107}$$

Combining Equations (4.87) and (4.106) yields:

$$\overline{r}_i = \sqrt{\frac{\eta}{W}}\left[\beta_i\delta_i\bar{X} + \sqrt{1-\beta_i^2\delta_i^2}\epsilon_i\right] \tag{4.108}$$

Clearly, by comparing this equation with Equation (4.104), we have $c_i = \beta_i\delta_i = \beta_i\sum_{k=1}^{K}\alpha_{i,k}b_k$. While the coefficients β_i and $\alpha_{i,k}$ are known, the coefficients b_k are unknown. Analogous to [35] method, we determine b_k through maximizing the weighted correlation between the single risk factor $\bar{X}$ and the original risk factor Y_i for all i. This leads to:

$$\max_{b_1,\dots,b_K}\left(\sum_{i=1}^{N} d_i \sum_{k=1}^{K}\alpha_{n,k}b_k\right) \text{ such that } \sum_{k=1}^{K} b_k^2 = 1 \tag{4.109}$$

Where we introduce a weight d_i for each counterparty in the portfolio. The solution to this maximization problem is given by (see Section 3.7.5 for details):

$$b_k = \sum_{i=1}^{N}\frac{d_i\alpha_{n,k}}{2\lambda} \tag{4.110}$$

Where the Lagrange multiplier λ is chosen in such a way that $\{b_k\}$ satisfies the second constraint, in essence, such that $\{b_k\}$ has a Euclidian norm of one. Unfortunately, we have not considered the coefficients d_i yet. Empirically, we will see that, in contrast to the original method, the follow choice works well:

$$d_i = w_i \cdot LGD_i \cdot F\left(\frac{F^{-1}(PD_i) - \beta_i F^{-1}(1-q)}{\sqrt{1-\beta_i^2}}\right) \tag{4.111}$$

We note that Equation (4.111) is essentially equal to the contributing VaR of counterparty i in the independent t-ASRF setting. The intuition behind this choice for d_i is that obligors with a high exposure in terms of VaR should have a large weight in the maximization problem whereas obligors with a small VaR should have a minor impact. Summing up, we can now define the adjusted one factor model (4.104) with $c_i = \beta_i \sum_{k=1}^{K} \alpha_{i,k} b_k$ together with the expressions for b_k and d_i we have derived. Figure 4.7 depicts the performance of the approximation which proves to be accurate. Furthermore, this figure depicts the heavier tail of the t-threshold model compared to its Gaussian counterpart. The aforementioned developed theory makes it possible to apply many of our results on economic concentration risk and sector diversification to the t-threshold model.

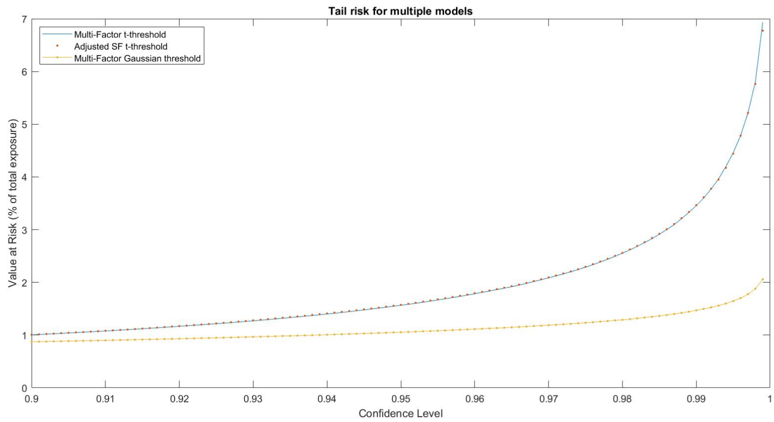

Figure 4.7: This graph depicts the difference in portfolio VaR across the tail of the loss distribution for a single portfolio, but different modelling assumptions. Clearly, the multi factor t-threshold model implies higher capital requirements compared to its Gaussian counterpart. Furthermore, the accuracy of the adjusted single factor model in comparison with the full multi factor t-threshold approach is depicted. Clearly, the loss distributions of the adjusted SF t-threshold and MF t-threshold models largely agree. Results are displayed for a heterogeneous portfolio of 1000 facilities and 1e6 MC trials, for $\eta = 5$.

III

Empirical Analysis

5

Portfolio Data

In the previous chapters, no assumptions have been made on the underlying portfolios in terms of size and heterogeneity. This chapter introduces the portfolios on which the developed methods will be examined. We will also take a closer look at the correlation structure that describes the linear interaction of the 110 systematic factors.

5.1. Portfolios

In order to test both the analytical and Monte Carlo based methods examined and developed in the previous chapters, credit portfolios have to be constructed. Throughout this section, these portfolios will be explored. In order to fully define a portfolio, several parameters are required, which are summarized in Table 5.1.

Abbreviation	Full Name	Typical Values	Source
EAD_i	Exposure at Default	$(0, \infty)$	ING, Matlab, Randomly Generated
PD_i	Probability of Default	$[0, 1]$	ING, Matlab, Randomly Generated
LGD_i	Loss Given Default	$[0, 1]$	ING, Matlab, Randomly Generated
β_i	Correlation Parameter	$[0, 1]$	Moody's, Regulatory Correlation
$\mathbf{\Omega}$	Covariance Matrix	$\mathbb{R}^{K \times K}$	Moody's
$w_{i,k}$	Sectoral Weights	$[0, 1]$	ING, Randomly Generated

Table 5.1: Set of parameters essential for fully defining a portfolio. Depending on the application, the parameters will be gathered from different sources.

5.1.1. Sample Portfolio

To illustrate some of the effects of for instance concentration, diversification, recovery risk and heavier tailed asset returns, we opt for using a fictitious sample portfolio. Using a single sample portfolio across multiple methods allows for a frame of reference in terms of risk measures. To this extend, we use a small heterogeneous portfolio, based on the example portfolio *CreditPortfolioData.mat* that is part of Matlab's Risk Management Toolbox. The portfolio consists of 100 obligors, the total exposure equals 3143 USD and the average exposure equals 31.4 USD. The exposures are assigned to 6 PD classes and 3 LGD classes. The correlation parameter β_i is given by Equation (2.27) with $\Lambda = 1$ and therefore PD dependent. Furthermore, the Gini index equals 0.557 and the HH index equals 0.0239 meaning that from first sight, the exposure is fairly concentrated. The portfolio exists of a broad range of exposures which should make for interesting analysis. Additionally, the portfolio is sufficiently small such that we can focus on individual counterparties and their risk contributions. In order to research the effects of diversification, every individual obligor is assigned to both a country and an industry based on the 110 sectors in table A.1. Since the original Matlab portfolio does not include factor loadings on sectors, the country and industry for each obligor is drawn from a scaled and shifted uniform distribution. Figure 5.1 gives some insight into the distribution of the default probabilities and exposures of our sample portfolio. However, it does not tell the full story. Figure 5.2 yields some insights in the interaction of the PD and

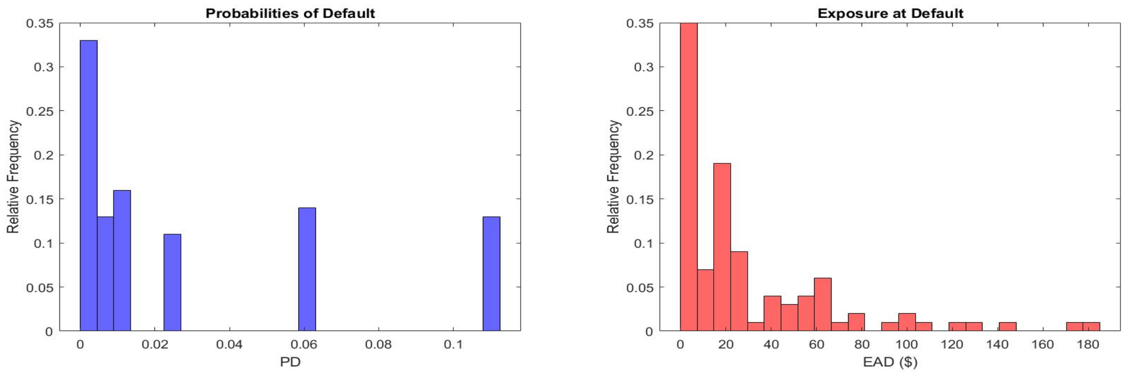

Figure 5.1: Illustration of the individual exposures and unconditional default probabilities associated with the 100 obligor sample portfolio.

exposure. Intuitively, you would expect a bank to take a higher risk on small loans as the potential impact on the business is small and as the exposures grow, one would expect lower probabilities of default. Clearly, this is the case for our sample portfolio which justifies that our sample portfolio broadly matches a real portfolio, although a lot smaller in size.

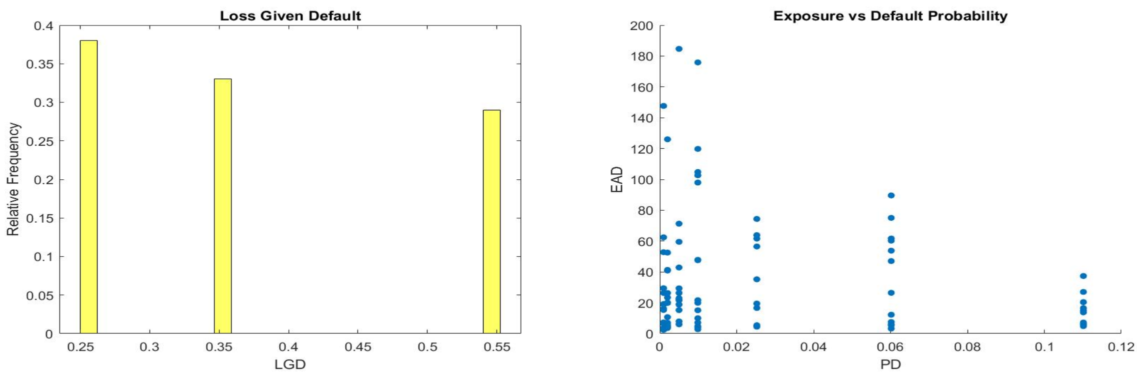

Figure 5.2: LGD and the interaction of EAD and PD for the 100 obligor sample portfolio.

5.1.2. Simulated Portfolios

In order to test some analytical methods across a wider set of portfolios, we suggest simulating portfolio with fairly similar properties to the sample portfolio of Section 5.1.1. Simulating credit portfolios is reasonably straightforward. For our purposes, we adopt the suggestion by [2] to normalize the total exposure to 10.000, split over 100 exposures. This implies that the average exposure per counterparty equals 100 units. The actual exposures are drawn from the Weibull distribution. The values for the PD are drawn from a χ^2 distribution with an average value of 1%. Values for the LGD are drawn from a Beta distribution, centered around an average value of 50%. Both the Weibull and χ^2 distribution are selected for the positive support and skew, which allows for the construction of relatively interesting heterogeneous portfolios. Otherwise, there is no specific reasoning underlying this choice. There are plenty of alternatives to the choices we make, for instance log-normally distributed EAD's and Beta distributed PD's (see, for instance [38]). For illustrative purposes, Figures 5.3 and 5.4 display some of the properties of a single draw of the simulated portfolios. Clearly, for each draw, the simulated portfolio will have slightly different properties.

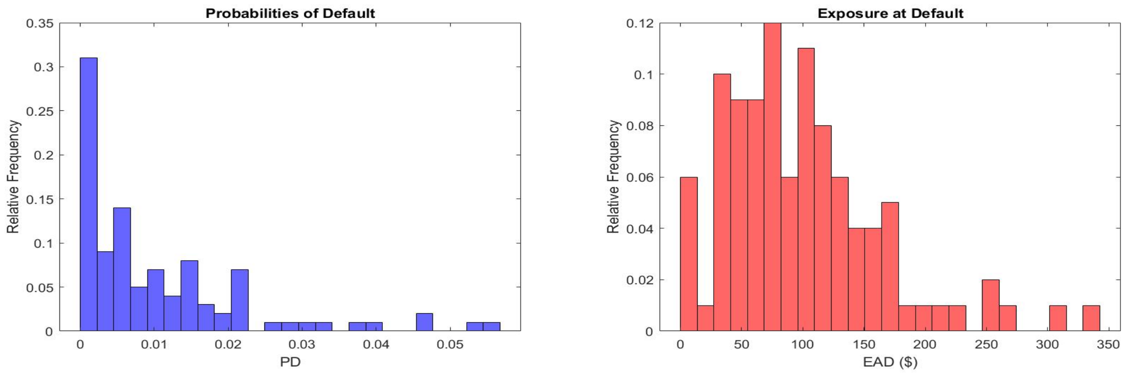

Figure 5.3: Illustration of the individual exposures and unconditional default probabilities associated with the simulated portfolio containing 100 exposures.

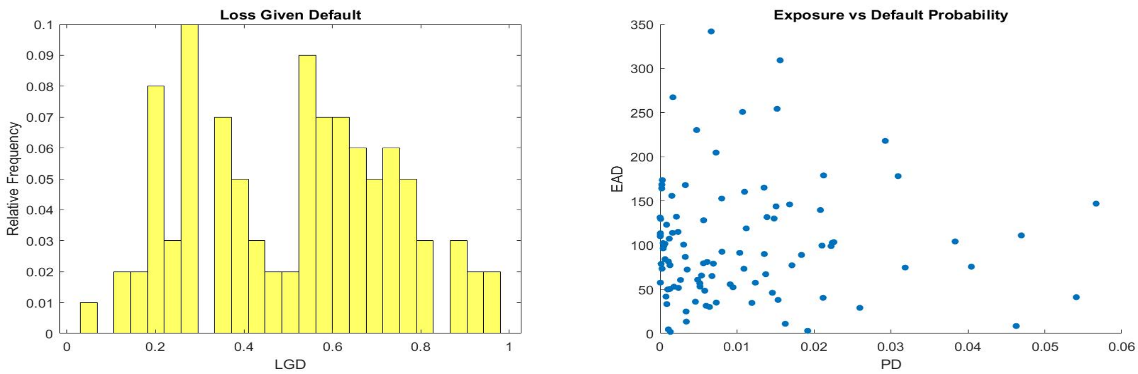

Figure 5.4: LGD and the interaction of EAD and PD for the simulated portfolio containing 100 exposures.

5.2. Correlation Structure

Throughout this work, in order to quantify the effect of systematic diversification, we will impose a covariance matrix to describe the joint behaviour of the systematic factors. The framework for sampling the correlated systematic factors is described in Section 4.1.1. In essence, we are looking for covariance matrix $\mathbf{\Omega} \in \mathbb{R}^{K \times K}$ of the K systematic factors. To this extend, we opt for the $K = 110$ covariance matrix obtained from Moody's Analytics. Due to confidentiality reasons, we cannot publish or display the full covariance matrix in this work. Of the 110 systematic factors, 49 systematic factors match a geographical region such as the *Middle East* or *Eastern Europe* and 61 systematic factors match an industry such as *Agriculture*, *Broadcast Media* or *Automotive*. For the full set of regions and industries, we refer to table A.1 in the appendix. The associated correlation matrix [1] is graphically depicted in Figure 5.5. To further ensure non-replicability, the matrix depicted is a perturbed version of the true correlation matrix. However, the perturbations are relatively modest in order to conserve the main properties of the correlation matrix. The correlation matrix, displayed in Figure 5.5, clearly satisfies the main properties of a correlation matrix by being symmetric and having a diagonal elements $a_{i,i} = 1$ for all $i = 1, ..., 110$. Furthermore, all elements $a_{i,j}$ are bound in the interval $[-1, 1]$, with $a_{i,j} = 1$ indicating perfect positive linear correlation and $a_{i,j} = -1$ indicating perfect negative linear correlation. The first 49 rows and columns indicate the regional correlation, the second 61 rows and columns indicate the industry correlation. This explains the presence of the distinct pattern of two darker shaded square planes and two lighter shaded square planes. The darker shade indicates a higher level of positive linear correlation. Focusing on the first 49 rows of the matrix, we can clearly observe that regions are highly positively correlated with each other but exhibit nearly no or negative linear correlation with the

[1] Knowing the covariance matrix, we can easily calculate the associated correlation matrix since the diagonal elements of the covariance matrix match the variances of the 110 systematic factors

61 industries. Focusing on the second 61 rows, we observe that the different industries are highly positively correlated with each other. This matches our intuition that the impact of changes in a specific country do not necessarily impact a specific industry and vice versa.

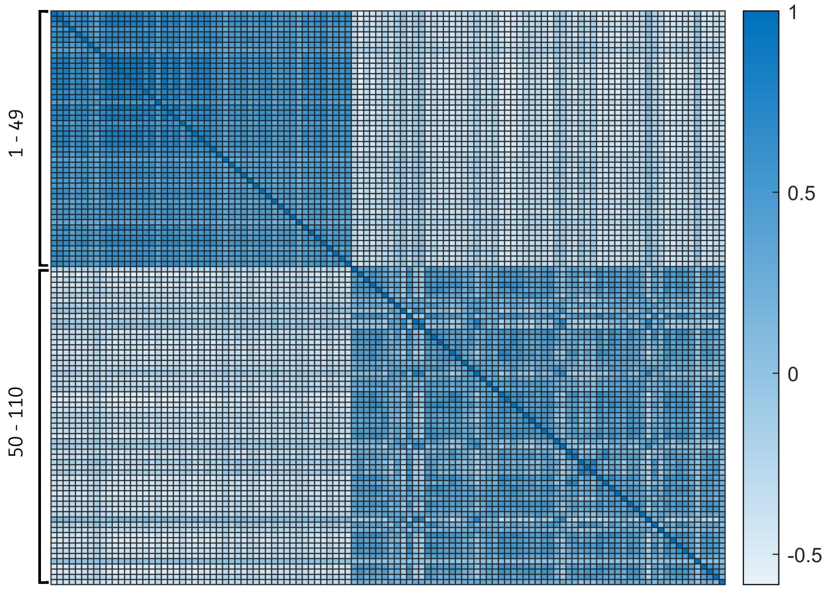

Figure 5.5: Perturbed Correlation Matrix for 110 systematic factors. The darker shaded square patterns can be explained by a high degree of intra-regional and intra-industrial correlation, but a lower degree of correlation between regions and industries.

6 Results

This chapter will commence with exploring the effects of concentration, diversification and random recovery on portfolio credit risk. Additionally, the effects of diversification and concentration will be evaluated for the t-threshold model, for which the effects are shown to be less prevalent. Next, the accuracy of the analytic approximations on portfolio and obligor level to concentration risk are explored in both a deterministic and stochastic LGD setting. Next, we will explore systematic diversification and Economic versus Regulatory Concentration Risk and show that these measures do not coincide. Lastly, we analyse the DF and CDI measures for the sample portfolio.

6.1. Justification: Sample Portfolio

In this section we will briefly discuss some results based on the sample portfolio of Section 5.1.1 to justify some of the claims made on the effects of concentration, diversification and recovery risk. Firstly, by essentially running algorithm 1, the single factor loss distribution for our sample portfolio is depicted in Figure 6.1. The values for VaR_{ASRF}, VaR and ES are all determined at a 99.5% quantile level.

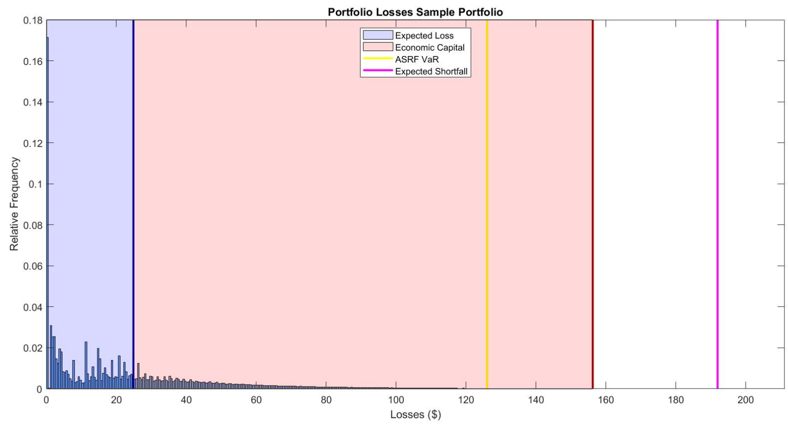

Figure 6.1: Single factor portfolio loss distribution of the sample portfolio based on 3e6 MC trials at $q = 99.5\%$ and deterministic LGD. The portfolio loss distribution indicates the relative frequency (relative to the 3e6 trials) of the losses depicted on the x-axis. Clearly, lower losses are expected to occur more frequently than extreme losses.

This figure yields the first insights into our sample portfolio. Firstly, the default-loss distribution is indeed asymmetric, highly skewed and heavy tailed, which is in line with our general expectations on portfolio loss distributions. Furthermore, one can notice that the ASRF assumptions are not satisfied as the true VaR exceeds its ASRF equivalent. This behaviour is expected since the sample portfolio size of 100 does not agree with the assumption of infinite portfolio size. Figure 6.2 depicts the contributing VaR of the individual obligors to the total VaR. From the first figure, we can conclude that the contributing

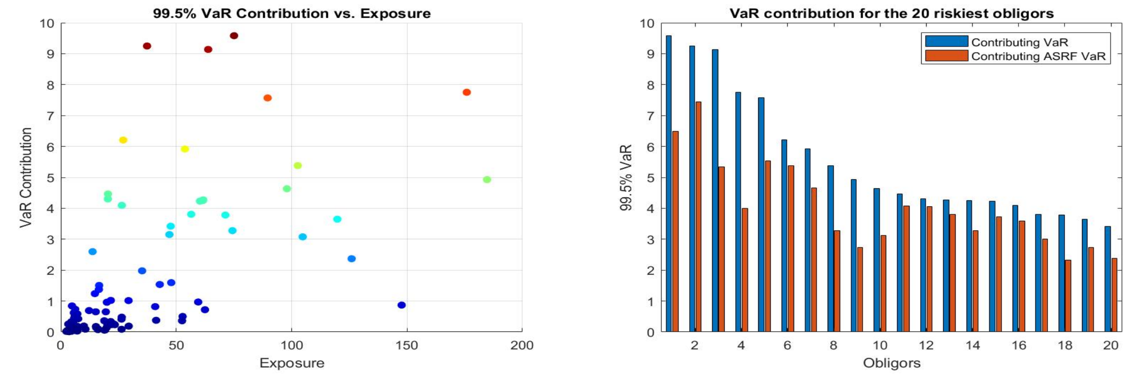

Figure 6.2: *VaR contribution perspective:* These graphs illustrate the contributing VaR of each obligor compared to its exposure and the top 20 riskiest obligors, based on a deterministic LGD and 3e6 MC trials.

VaR differs widely from counterparty to counterparty and that not necessarily, the highest exposures contributes the most to the total risk of the portfolio. In the second figure, the 20 highest contributing obligors are ranked on the basis of contributing VaR and compared with its ASRF contributing VaR equivalent. Firstly, the true contributing VaR seems to exceed its ASRF equivalent, which we would expect from the conclusions based on Figure 6.1. A more interesting observation is that the ranking differs with respect to the true VaR compared to its ASRF equivalent. This means that some loans might seem to be relatively safe in an ASRF setting, whereas in the full Monte Carlo setting, they add more risk than expected to the portfolio. This stresses the need for Pillar II compliance, in which the bank has to operate its own models next to the standard IRB methods.

6.1.1. The Single Name Concentration Effect

Throughout this section we will explore the effects of concentration, or in other words, of residual idiosyncratic risk, on the sample portfolio. We will compare our 100 obligor portfolio with a more diversified portfolio. However, to make a reasonable comparison, the diversified portfolio should resemble the sample portfolio. To this extend, we make a simple adjustment to the sample portfolio by replicating the portfolio ten times but with each exposure 1/10th of its original exposure. The diversified portfolio will be of a size of 1000 obligors but with an equal total exposure to the original sample portfolio. By the portfolio invariance of the ASRF method, the ASRF VaR for both these portfolios is exactly equal. The Gini index for this portfolio is equal to the sample portfolio whereas the HH index is ten times smaller; 0.00239. This observation underlines a major drawback of the Gini index, namely its incapability of taking portfolio size into account. Based on its HHI, the diversified portfolio is ”ten times” less concentrated than the original sample portfolio.

The actual effect of concentration is depicted in Figure 6.3. This figure is based on 1e6 MC trials a single risk factor model for both the diversified and concentrated version of our sample portfolio. Clearly, the concentrated portfolio has a heavier tale and is, therefore, a riskier portfolio. The tail quantiles of the diversified portfolio nearly coincide with the ASRF quantiles, indicating that idiosyncratic risk is nearly diversified away. Increasing the number of exposures in the diversified portfolio further would decrease its HHI and bring its tail quantiles closes to the ASRF equivalent. The difference in VaR between the concentrated and diversified portfolio is entirely due to resididual idiosyncratic risk and is what we call the Concentration Effect.

6.1.2. The Diversification Effect

In Section 6.1.1 we discussed the effects of concentration on a portfolio. In exploring this effect, we limited ourselves to the single factor Gaussian threshold model. Clearly, since this model depends only on a single systematic factor, the effects of diversification are not taken into account. In this section, we will briefly explore the effects of systematic diversification on our sample portfolio. To this extend, we will apply a 110 factor Gaussian threshold model and compare it with the single factor model. As mentioned in Section 5.1.1, regional and industry dependencies are drawn from a uniform distribution. In total, we consider 110 indices, of which 49 (#1-49) are region related and 61 (#50-110) are industry

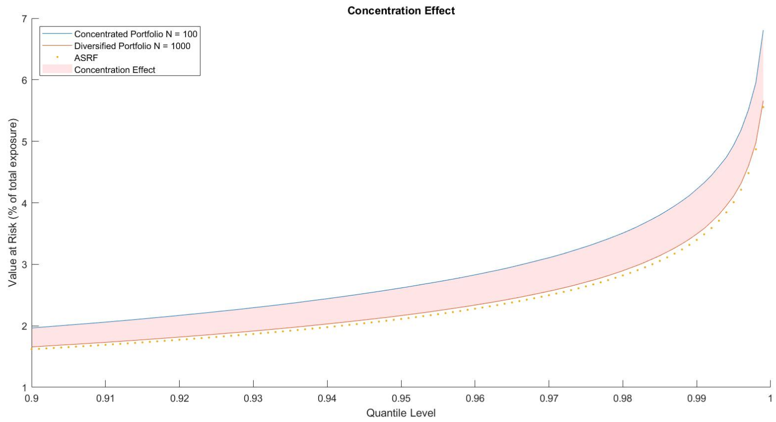

Figure 6.3: The effect of portfolio size on residual idiosyncratic risk for our sample portfolio in the tail of the default loss distribution. The $N = 1000$ portfolio's quantiles approach the ASRF quantiles, indicating that the portfolio is nearly perfectly diversified.

related. Each counterparty is assigned to one country and one industry, as illustrated in Figure 6.4. This is done purely from a practical point of view; one would for instance also be able to assign multiple countries to a specific counterparty if that counterparty is active in multiple regions.

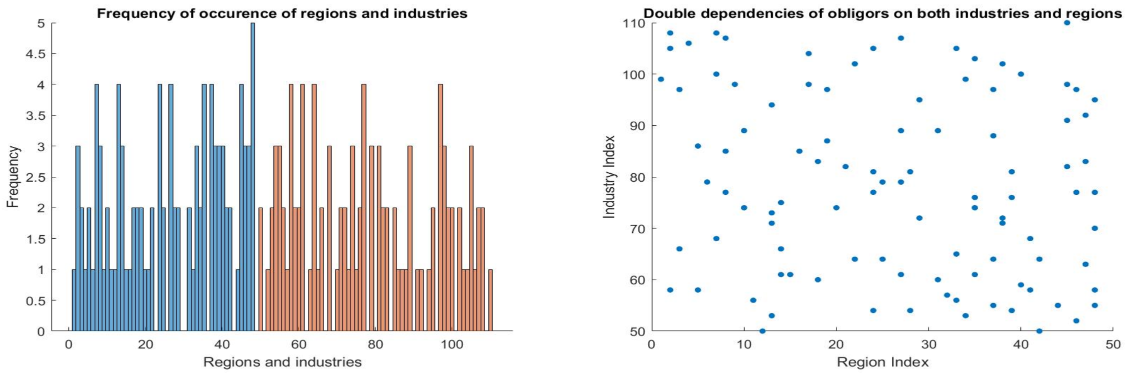

Figure 6.4: *Multi factor perspective:* These graphics depict the dependence on the 49 regional and the 61 industrial systematic factors of the 100 obligors. Our sample portfolio seems to be relatively well diversified across regions and industries.

Figure 6.5 depicts the effect of diversification on the tail of the loss distribution of our sample portfolio. As expected, the single factor model overestimates the VaR compared to the multi factor model since it lacks to take any diversification effects into account. Interestingly, the effect of diversification seems to increase as we move further into the tail of our loss distribution. Additionally, if we compare the ASRF quantiles to the multi factor quantiles, one can notice that the ASRF based VaR might overestimate the risk compared to the multi factor model.

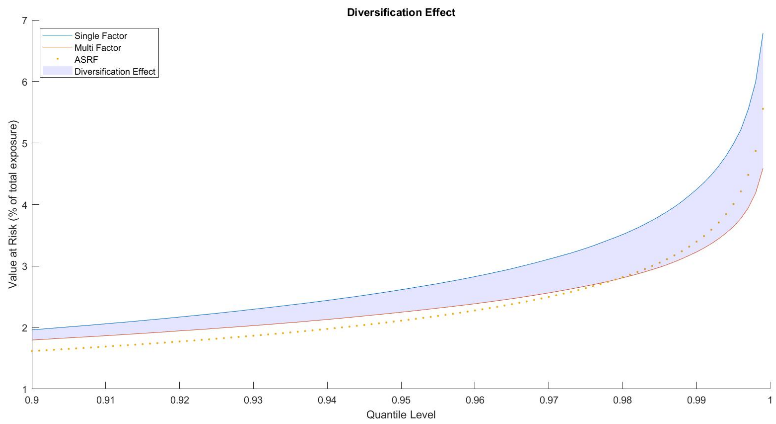

Figure 6.5: The effect of diversification through switching from a single systematic factor to 110 systematic factors, based on the sample portfolio at 1e6 MC trials and deterministic LGD. The effects of diversification are quite apparent, with the ASRF solution even exceeding the true portfolio VaR from quantile levels of approximately 98% upwards.

The results displayed in Figure 6.5 have an interesting link with the theory developed in Section 4.5 as essentially, the single factor results displayed are the EC_q^{E-SF} for our sample portfolio. Summing the EC contributions of the obligors belonging to a certain region yields the results depicted in Figure 6.6. Clearly, across every region, the lack of diversification causes higher risk contributions for the single factor model compared to its multi factor equivalent. More strikingly, the ranking of riskiest regions differs across the two models. For instance, according to the multi factor model, region 19 contributes more to the total EC than region 2, whereas in the single factor framework, this statement does not hold. This clearly suggests that the information in these two risk measures do not coincide. The last five regions do not add any risk to the portfolio risk level as these regions are not present in our sample portfolio. A very similar result follows from splitting the total EC across the 61 industries represented in the portfolio.

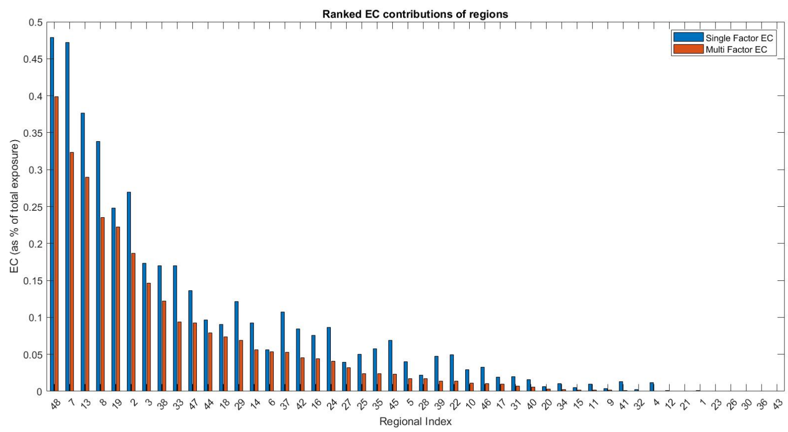

Figure 6.6: EC contribution of the individual regions in the sample portfolio, ranked on the contribution in the Multi Factor framework at a 99.5% quantile level, based on 1e6 MC trials. The EC contribution in the MF case is lower across all regions compared to the SF case. Moreover, the rankings of EC do not match between the two underlying models.

6.1.3. Recovery Risk Effect

In the previous two sections, we explored the effects of concentration and diversification in a deterministic LGD setting. This thesis aims to research a third effect, the effect of recovery risk on concentration and diversification and the ability of analytic approximations, such as the GA, to capture this effect. Therefore, we will explore this effect empirically in this section. Again, we assume are 100 exposures sample portfolio. We apply the single factor Gaussian threshold model to this portfolio, however, we set the LGD to be Beta distributed with recovery parameter k. Intuitively, we expect a higher VaR at a given quantile level for a portfolio with random LGD, as these random recovery rates "add risk" to the portfolio. Our intuition is in line with the results of our sample portfolio, depicted in Figure 6.7. Note that the expected loss is equal, invariant of the Beta distributed LGD. Clearly, the random LGD results in a higher VaR across the tail of the portfolio loss distribution. As the value of k increases, the uncertainty and variance around the LGD estimates decrease and therefore, at high values k, the recovery risk effect diminishes. Furthermore, the results underline the inability of Basel's ASRF method to capture the effect of recovery risk on the portfolio VaR.

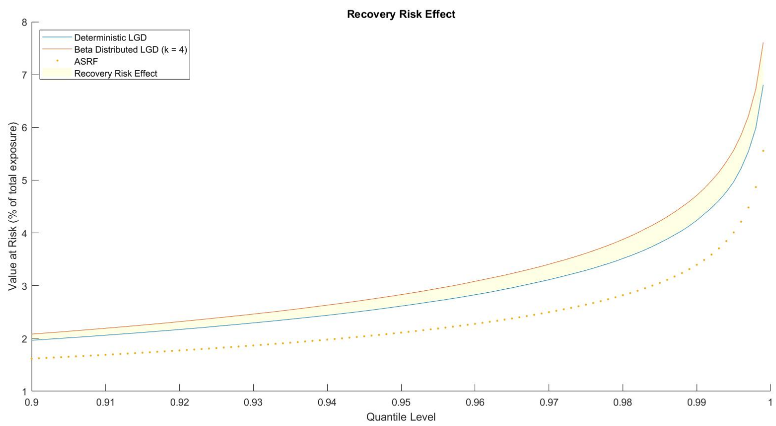

Figure 6.7: Graphic representation of the effects of recovery risk in the single factor setting on our sample portfolio. Introducing Beta distributed LGD rates ($k = 4$) has a clear effect on the tail behaviour of the portfolio loss distribution, increasing the capital requirements at every observed quantile level.

6.1.4. *t*-Threshold Model

Throughout this section, we will shortly address the effects of diversification and concentration for the t-threshold model (4.87). Essentially, we perform a similar analysis to Sections 6.1.1 and 6.1.2 on our sample portfolio, only with different underlying model assumptions. The results of this experiment are displayed in Figure 6.8. Firstly, one will notice that as expected, the VaR is higher across the upper quantiles for the *t*-threshold model compared to its Gaussian counterpart, depicted in Figures 6.3 and 6.5. This result can be attributed to the fatter tails of the *t* distribution and the dependence between systematic and idiosyncratic factors. Secondly, compared to Figures 6.3 and 6.5, both the concentration and diversification is less prevalent. The effects of increasing portfolio size and diversifying the portfolio across sectors is only minimal compared to its Gaussian counterparts. This implies that in the *t*-threshold setting, idiosyncratic risk cannot be diversified away as the number of facilities in a portfolio increases. Furthermore, spreading one's portfolio over multiple sectors and regions has only a limited effect on the risk level of the portfolio. This effect can again be attributed to the fatter tails of the *t* distribution and the dependence between the idiosyncratic and systematic factors. To clarify, we should note that Figure 6.8 displays the results based on $\eta = 5$, meaning the distribution of asset returns substantially differs from the standard normal distribution. For higher degrees of freedom, both the concentration and diversification will be more prevalent and approximate the Gaussian case as the degrees of freedom tend to infinity. From a risk management point of view, these results can be of interest, as the *t*-threshold model is less sensitive to misspecifications in the correlation parameters.

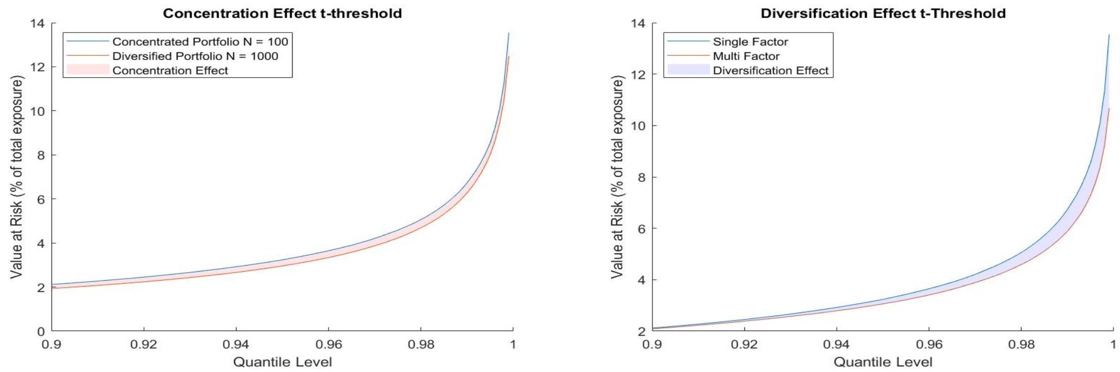

Figure 6.8: *t Threshold perspective:* Depiction of the VaR at varying quantile levels for the t-Threshold model based on 1e6 MC trials at $\eta = 5$. The left figure depicts the effect of increasing the portfolio size on the portfolio VaR. The right figure depicts the effect of switching from a single factor to a multi factor setting. Both the concentration and diversification effect seem to be minimal.

6.2. Analytic Single Name Concentration Risk

Throughout Chapter 4 we developed methods to assess concentration risk in the general Gaussian threshold single risk factor setting (2.23). Throughout this section, we briefly put the single factor GA to practice on our sample portfolio. In essence, we will see whether the GA can 'bridge the gap' of Figure 6.3. We will examine this both in a deterministic LGD and a Beta distributed LGD setting. Furthermore, we will examine whether the risk contributions of individual obligors coincide between the full MC methods and the analytical approximations based on the Euler allocation principle.

6.2.1. Gaussian Threshold GA

First, we focus on the GA of (4.20) and of (4.41) applied to our sample portfolio. In Section 6.1.1 we concluded that there is a fair amount of idiosyncratic risk involved in our sample portfolio as its "true" VaR exceeds its ASRF-VaR across the entire tail of the portfolio loss distribution. The GA was developed to analytically bridge the gap between the ASRF-VaR and the 'true' VaR. We will test the GA on our sample portfolio by applying Equations (4.20) and (4.41) as an add-on to the ASRF VaR of the sample portfolio. This experiment results in Figure 6.9. This figure suggests that both the CR+ and the SF adjustment seem to perform quite well as an add on to the ASRF solution in approximating the true loss distribution of our sample portfolio. In general, both the CR+ and SF GA slightly overestimate the true VaR at the lower quantile levels which means that to this extend, the GA is relatively conservative.

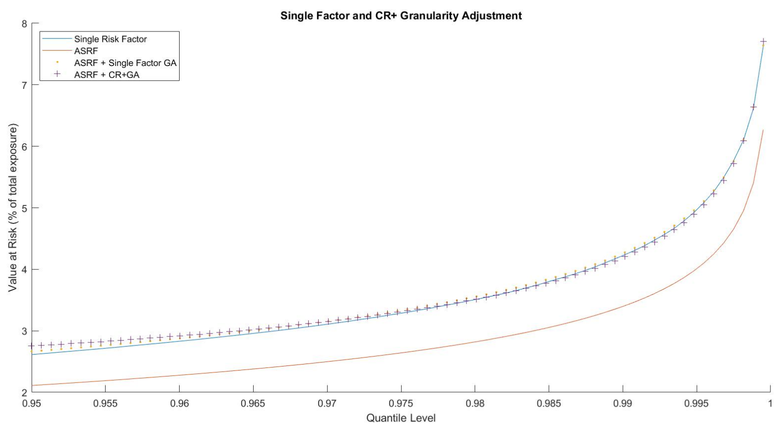

Figure 6.9: Illustration of the performance of the single factor GA based on Equation (4.20) and the CreditRisk+ GA based on Equation (4.41) for the sample portfolio. Both methods slightly overestimate the true portfolio VaR for quantile levels up to approximately 98%, after which both methods provide a reasonable analytical estimation for portfolio VaR. Numerical single risk factor results are based on 1e6 MC trials and a deterministic LGD.

In Table 6.1 we zoom in on some widely used levels of q by comparing the VaR approximations of both the CR+ and the SF Granularity Adjustment with its Monte Carlo based counterpart. Again, across all quantiles, the analytical approximations perform well. However, one can note that in general, the SF GA outperforms the CR+ GA slightly.

q	95%	99%	99.5%	99.99%
MC	2.6140	4.2232	4.9693	9.7030
ASRF + SF GA	2.6649	4.2607	4.9970	9.6359
Δ_{SF-MC}	0.0509	0.0375	0.0276	-0.0671
ASRF + CR_+ GA	2.7530	4.1930	4.9341	9.8537
Δ_{CR_+-MC}	0.1390	-0.0302	-0.0352	0.1508

Table 6.1: Percentage point difference between the true VaR and its analytical approximations across the most frequently used quantile levels. Values are displayed as percentage of total exposure.

Although promising, the aforementioned results do not substantiate the performance of the GA across multiple heterogeneous portfolios. We therefore introduce the simulated portfolios of Section 5.1.2. In essence, we will perform the following experiment:

1. Sample 25 portfolios, each containing 100 obligors and a total exposure of 10.000 each
2. Compute the portfolio loss distribution numerically, using 1e6 MC trials for each portfolio
3. Compute the 99.5% MC-based VaR for each portfolio
4. Compute the 99.5% VaR analytically based on the SF GA
5. Compute the 99.5% VaR analytically based on the CR+ GA
6. Determine the residual sum of squares for the SF and CR+ GA by

$$RSS = \sum_{i=1}^{25} [VaR_i^{MC} - VaR_i^{GA}]^2 \tag{6.1}$$

The results of this experiment are displayed in Figure 6.10. Two interesting conclusions can be drawn from this figure. Firstly, in general, the SF GA seems to slightly outperform the CR+ GA as the VaR values for the SF GA seem to be more centered around the 45° line. This observation is supported by the RSS, which equals 0.11 for the SF GA and 0.70 for the CR+ GA. Secondly, the CR+ GA seems to consistently overestimate the true VaR at the 99.5% quantile. This observation agrees with [22] in which the authors suggest that the GA might have difficulties with smaller portfolios and that it might overstate the effect of granularity.

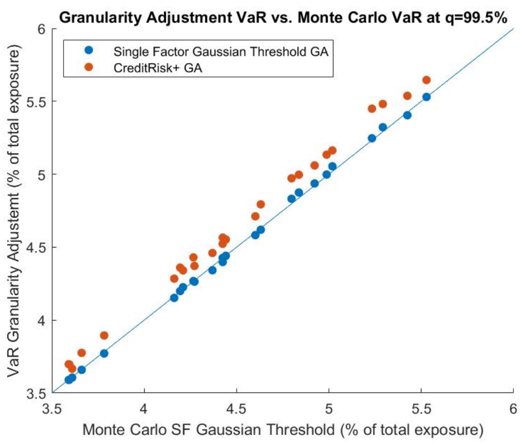

Figure 6.10: Illustration of the performance of the single factor GA based on Equation (4.20) and the CreditRisk+ GA based on Equation (4.41) for 25 simulated portfolios at a single quantile level. The CR+ GA slightly overestimates the portfolio VaR across all simulated portfolios whereas the SF GA performs very well overall. Numerical results are based on 1e6 MC trials and a deterministic LGD.

Risk Contributions
In the previous section we focused our efforts on the accuracy of the GA on portfolio level. However, when assessing concentration risk, we are instead more interested in the contributions of individual obligors to the total risk of the credit portfolio. We therefore introduced methods to determine risk contributions, both in the MC setting and in the analytic setting, using the Euler allocation principle (Sections 3.5.1 and 3.5.2). In this section, we will limit ourselves to the SF GA, as we have seen in previous sections that is seems to slightly outperform the CR+ GA. Ideally, the risk contributions determined through the MC method would equal the risk contributions determined analytically using the GA. Practically, this will outcome will not occur due to two reasons. Firstly, in the previous section we have seen that the GA-VaR does not necessarily exactly equal the MC-VaR. When applying the Euler allocation principle, the risk contributions obey the full allocation property by construction. Therefore, the analytical risk contributions will add up to analytical portfolio VaR whereas the numerical risk contribution will add up to the MC portfolio VaR. Clearly, if the analytical and MC VaR are not equal, their risk contributions cannot be either. Secondly, MC methods always introduce some kind of statistical noise. Using many trials, we can relatively accurately determine the portfolio VaR, whereas the individual risk contributions exhibit more statistical noise. To test the accuracy of the analytic risk contributions, we focus on our sample portfolio. For each of the 100 obligors, we compute the contributing VaR using both MC methods and the analytical approximation Equation (4.24). The results of this experiment are displayed in Figure 6.11. Again, from the scatter plot, we can conclude that the risk contributions largely agree between the analytic and the MC methods. Using this result, the risk contributions for the 20 riskiest obligors based on their ASRF VaR contribution are computed. Where the ASRF contribution can be interpreted as the systematic risk contribution, the GA can be regarded as the idiosyncratic risk contribution. Interestingly, some obligors that add a large amount of systematic risk to the portfolio, do not necessarily add much idiosyncratic risk to the portfolio VaR. This stresses the importance of assessing the effects of idiosyncratic risk on one's portfolio.

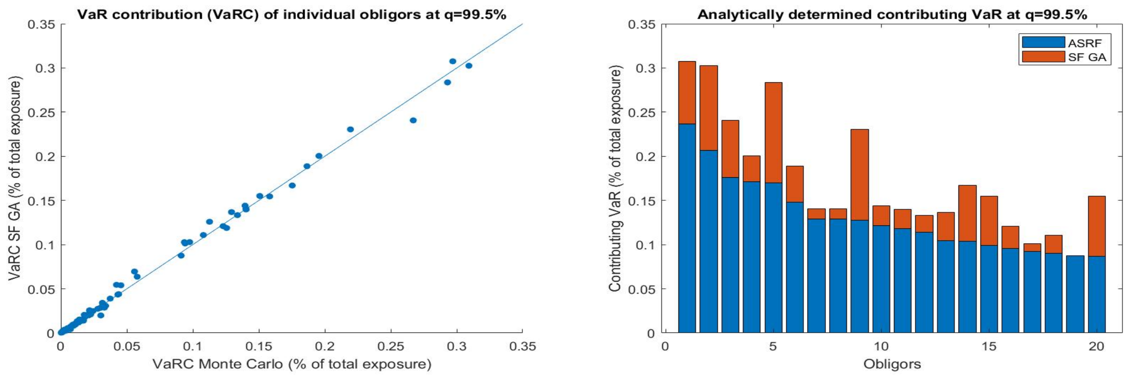

Figure 6.11: *Risk contribution perspective:* Accuracy of the risk contributions of the SF GA compared to its MC counterpart for our sample portfolio. The risk contributions are clustered around the 45° reference line, indicating agreement between the analytical and MC based risk contributions. Furthermore, the GA allows for a clear division in systematic ASRF and idiosyncratic GA risk on obligor level. The impact of the GA on the risk contributions for the highest 20 contributing exposures are substantial. Based on 5e6 MC trials at a 99.5% quantile level.

Overall, we can conclude that, although mathematically rigorous, the GA can be a powerful tool to assess residual idiosyncratic risk in our sample portfolio. However, since the single factor model lacks the ability to account for diversification, the GA as presented above is mainly a useful tool in the regulatory setting in settings where Basel's A-IRB solution underestimates the true risk in a portfolio.

Stochastic LGD
In Section 6.1.3 we have researched the effect of introducing independent Beta distributed LGD's to the single risk factor model. Depending on parameter k, the effect of recovery risk becomes more prevalent compared to its deterministic counterpart. As k increases, the Beta distribution becomes more centered around its expected value and therefore, as k increases, the resulting portfolio VaR approximates its deterministic counterpart. In this section, we will briefly examine the GA's ability to account for recovery risk. In the previous sections, we concluded that in the presence of a deterministic LGD, the GA

performs excellent. However, since a random LGD has a clear impact on the tail of the portfolio loss distribution, we will examine the performance of the GA in the stochastic LGD setting. Both the single factor GA (4.20) and the CreditRisk+ GA (4.41) are derived in a setting that accounts for stochastic LGD's. We again focus our efforts on the sample portfolio and perform an experiment very similar to the deterministic LGD case of Figure 6.9. The results of this experiment are displayed in Figure 6.12. Essentially, Figure 6.12 matches Figure 6.7, but it overlays the analytical results of the stochastic LGD GA. Clearly, the SF GA outperforms the CR+ GA across most quantile levels and especially as we approach the higher quantile levels, the SF GA approximates the true VaR better compared to the CR+ GA. Strikingly, for the lower quantile levels, both the CR+ and SF GA overestimate the true portfolio VaR. In general, the SF GA slightly overestimates the true portfolio VaR across all quantile levels, which coincides with the accuracy of the deterministic LGD versions. While ideally, we would like to exactly match the true VaR analytically, from a regulatory point of view overestimation is preferred over underestimation. We can therefore conclude that the SF GA estimates the true portfolio VaR in the presence of Beta distributed LGD fairly accurately, and consequently, it is a useful tool in analytical regulatory risk approximations. However, taking a closer look at the SF GA's performance on obligor level, by comparing the contributing VaR of the individual exposures to the portfolio VaR, some disagreements occur between the analytical approximation and the numerical solution. Especially for some relatively high contributing exposures, the MC solution does not match the GA approximation. Whereas for the highest contributing exposures one would ideally have the most accurate risk measurement, this is a major drawback to the GA applied to stochastic LGD models.

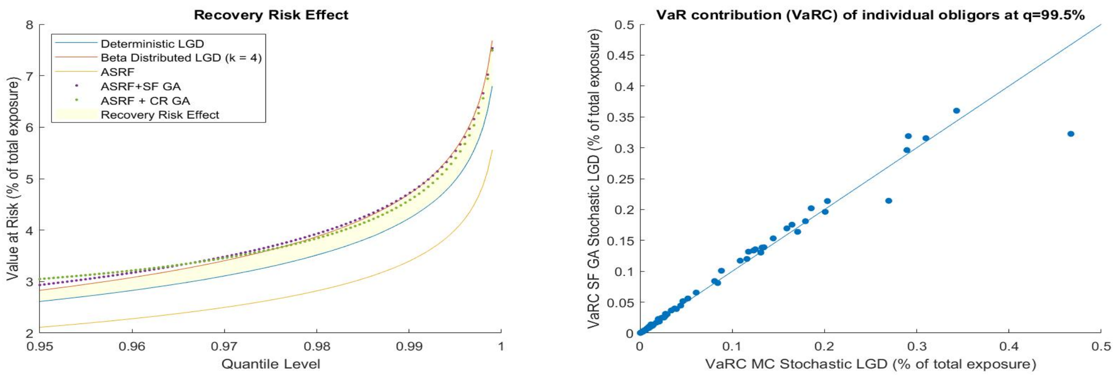

Figure 6.12: Illustration of the performance of the single factor GA based on Equation (4.20) and the CreditRisk+ GA based on Equation (4.41) for the sample portfolio. The risk contributions are based on the SF GA only, as we have seen that it approximates the true VaR more accurately. Compared to the deterministic LGD case, the accuracy of the adjustment on both portfolio as contribution level seems to be worse. Numerical results are based on 1e6 MC trials and stochastic LGD.

6.2.2. Independent *t*-threshold Granularity Adjustment

In Section 4.6, two different *t*-threshold models were introduced. The first model was based on the original normal variance mixture approach, by scaling the original Gaussian threshold factor model with a stochastic scaling factor, resulting in Equation (4.87). In Section 6.1.4, the effects of diversification and concentration in this model have been reviewed, and we concluded that these effects are less prevalent compared with the Gaussian threshold model. Secondly, we introduced the independent *t*-threshold model (4.93), in which the systematic and idiosyncratic factors remain independent, similar to its Gaussian counterpart. For this model, we have derived both an ASRF solution (4.96) and a GA (4.97), and we have shown their convergence to the Gaussian model as the degrees of freedom increase. In this section, we will briefly explore the performance of the *t*-threshold GA on portfolio level on our sample portfolio. In essence, we repeat the same experiment as performed on the Gaussian single factor model by trying to approximate the true portfolio MC based VaR by the sum of the ASRF and the GA analytical approximations. The results of this experiment are displayed in Figure 6.13. For reference purposes, we also included the results for the single factor Gaussian threshold model. The degrees of freedom η are set to 5, in order to ensure a significant difference in portfolio VaR between the two models under review. Taking a closer look at Figure 6.13, we can first note that up until approximately the 99.5% quantile level, the portfolio VaR of the Gaussian threshold model exceeds that of the independent t-threshold model. From this point onwards, a sharp increase in portfolio VaR is apparent for the t-threshold model. Secondly, the t-ASRF solution seems to be a relatively good approximation to the true VaR at very high quantile levels. Thirdly, the t-GA seems to approximate the true solution very well; the accuracy of the analytical approximation seems to be in line with the accuracy of the Gaussian GA of the previous section.

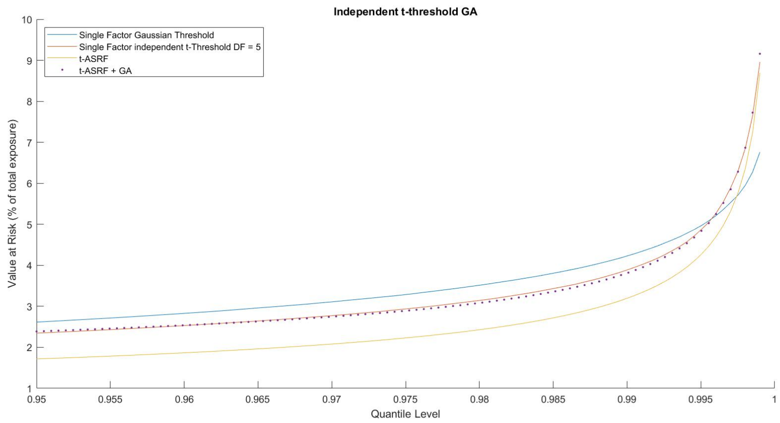

Figure 6.13: Illustration of the performance of the independent t-threshold GA based on Equation (4.97) on estimating the portfolio VaR across the tail of the loss distribution. From a quantile level of 99.5% the independent t-threshold based VaR exceeds its Gaussian counterpart. The GA seems to approximate the true VaR accurately. Numerical results are based on 1e6 MC trials and a deterministic LGD.

From this analysis, we can conclude that the t-GA performs on par with the regular Gaussian GA for their respective models. The independent t-threshold model is particularly interesting from a regulatory perspective. As we have seen, the portfolio EC requirements are higher compared to the current IRB methods, depending on the choice of degrees of freedom and the quantile level. This indicates that the regulator can have more control over capital requirements, by simply adjusting the degrees of freedom in for instance periods of economic downturn. This makes the independent t-threshold model, its ASRF solution and its Granularity Adjustments a powerful tool from a regulatory perspective.

6.3. Analytic Sector Concentration Risk

Deterministic LGD

In the previous sections, we have limited our analysis to single factor models, both the Gaussian and t-threshold variants. Additionally, we have shortly addressed the effects of allowing diversification in the sample portfolio in Section 6.1.2, which essentially describes the effects of switching from a single factor to a multi factor framework. In Sections 3.7.5 and 4.3.1 we explored Pykhtin's multi factor adjustment [35] and derived it for both the portfolio level adjustment and obligor level adjustments. Throughout this section, we will assess the performance of these adjustments for our sample portfolio. Ideally, the adjustments would yield an accurate analytical method to approximate the multi factor VaR solution of Figure 6.5. In order to test the performance of the adjustment, we performed two experiments of which the results are displayed in Figure 6.14. Essentially, we consider two portfolios. Firstly, we consider the standard sample portfolio of 100 exposures with each exposure assigned to both a specific country and specific sector, analogous to Figure 6.4. Secondly, we increased the portfolio in size, using a similar process to the methods described in Section 6.1.1. For both portfolios, the Adjusted ASRF solution is determined according Equation (4.55) and the adjustment is calculated using (4.56) and (4.57). Since the adjusted ASRF model is portfolio invariant, both adjusted ASRF solutions displayed in figure (6.14) equal each other. The MC solution differs for both portfolios since due to the size difference, concentration risk is reduced in the second portfolio compared to the first portfolio. Noticeably, the accuracy of the MF GA analytical approximation differs across the two portfolios. Especially in the small 100 exposure sample portfolio, performance seems to be relatively poor. This indicates that the MF GA has difficulties with approximating the portfolio VaR for very small portfolios with much residual idiosyncratic risk and performance increases as idiosyncratic risk is increasingly diversified away. Although the accuracy is not perfect, the GA adjusted solution is clearly preferable over the adjusted ASRF solution in estimating the true portfolio risk in the Gaussian threshold multi factor setting.

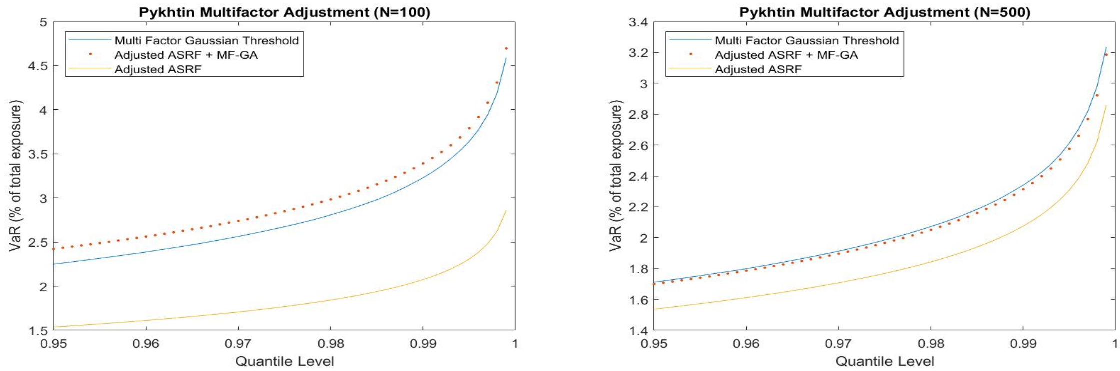

Figure 6.14: Illustration of the performance of the Multi Factor GA based on Equation (4.54) for two sample portfolios of differing size. The overestimation for the $N = 100$ portfolio indicates that the MF GA is inaccurate in the presence of much residual idiosyncratic risk. Numerical results are based on 1e6 MC trials and deterministic LGD.

As mentioned before, a major drawback of the MF GA is its computational complexity as it requires calculating a conditional correlation matrix between any two obligors, in essence Equation (4.49). Furthermore, the MF GA requires calculating the joint conditional default probability between any two obligors, essentially $\Phi_2(\Phi^{-1}(PD_i(x)), \Phi^{-1}(PD_j(x), \rho_{ij}^{\bar{X}})$. This has a major implication on the computing time of this analytical approximation. To illustrate, for the $N = 500$ portfolio, the MC method takes just over a minute to compute the full portfolio loss distribution, whereas the analytical method takes over 21 minutes to compute the 50 discrete values of the portfolio VaR displayed in the second graph of Figure 6.14. However, since we are often only interested in computing the portfolio VaR at a single quantile level, the analytical method outperforms the MC method at 25 seconds. Be that as it may, the memory issues with generating the conditional correlation matrix between any two obligors increase disproportionally compared to the memory needs involved with running MC trials.

Nonetheless, the accuracy of the MF GA on portfolio level seems to be promising. We will now focus our efforts on the risk contributions of the individual obligors and their numerical and analytical approximations, similarly to the results displayed in Figure 6.11. We will apply the theories developed in Section 4.3.1, and Equation (4.58) in particular as it yields a method to allocate the portfolio VaR to the individual exposures analytically. The result of this experiment, performed on the enlarged sample portfolio of 500 exposures, are displayed in Figure 6.15. Overall, the GA and MC approximations are centered around the 45° reference line, indicating that the two approximations broadly agree on the value of the individual risk contributions to the portfolio VaR. Notably, an interesting pattern occurs in which risk contributions are clustered horizontally. This pattern is caused by the enlargement procedure, in which obligors are replicated in the portfolio. In this case, we enlarged the portfolio by a factor of 5. Essentially, this means that that portfolio of 500 obligors exists of 100 unique obligors, in sets of 5 identical obligors. In terms of risk contribution, we would expect the contributions of the identical obligors to contribute an identical amount of risk to the portfolio VaR. Since the risk contribution clusters are lined up horizontally, the analytical solution is equal among the sets of 5 identical obligors, whereas the MC approximation differs among these obligors. This effect stresses a major drawback of MC trials versus an analytical approximation; MC trials inherently exhibit statistical noise.

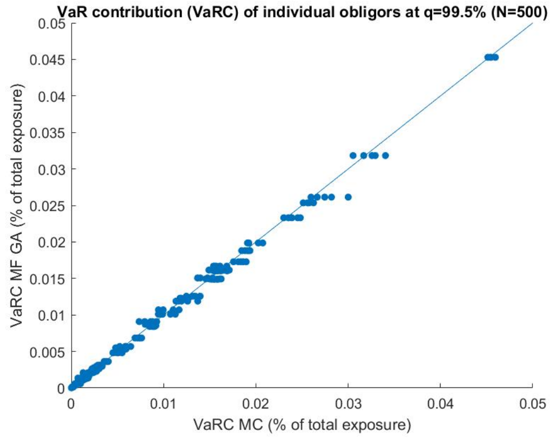

Figure 6.15: Illustration of the performance of the multi factor GA risk contributions based on Equation (4.58) at a single quantile level of 99.5%. Generally, the point are clustered around the 45° reference line, indicating a large degree of agreement between the numerical and analytical methods. Numerical results are based on 2e6 MC trials and a deterministic LGD.

Stochastic LGD

Having analyzed the multi factor granularity adjustment in a deterministic setting, we will now shortly assess its accuracy in the presence of stochastic LGD rates. We continue our analysis on the $N = 500$ portfolio, as we have concluded that the accuracy of the MF GA is relatively poor on the 100 exposure sample portfolio. Essentially, introduce recovery risk to the multi factor Gaussian threshold model using the same methods described in 6.1.3 by setting the LGD parameter k to a value of 4, which represent a high degree of uncertainty in the given LGD rate. For average LGD rates of approximately 50%, a value of $k = 4$ results in a U shaped Beta distribution, with a higher probability of sampling either a 0% or 100% LGD rate. Clearly, we could consider higher values of k, indicating a larger degree of certainty of our LGD estimates. However, a k increases, the results will converge towards a deterministic LGD setting and therefore we opt for low values of k.

We repeat the experiment performed in assessing the accuracy of the multi factor GA in a deterministic LGD setting. The results of this experiment are displayed in Figure 6.16. There are two main takeaways we can conclude from this experiment. Firstly, the introduction of stochastic LGD rates has only a minor effect on the portfolio VaR since the VaR levels in the tail of the loss distribution are nearly equal to those displayed in right graph of Figure 6.14. Secondly, the MF GA seems to accurately approximate both the portfolio VaR as the VaR contributions of the individual obligors analytically for this specific portfolio. Performing the same experiment on the smaller $N = 100$ sample portfolio would yield less accurate results.

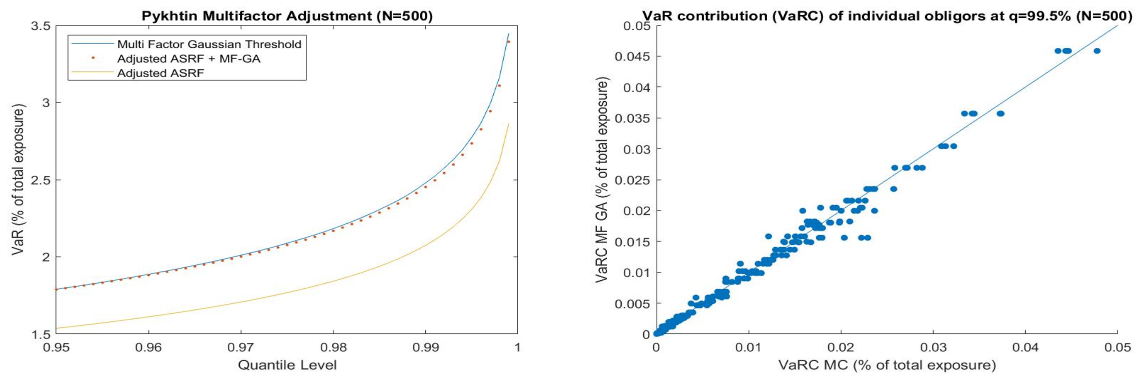

Figure 6.16: Illustration of the performance of the Multi Factor GA based on Equation (4.54). The horizontal clustering of data points in the VaRC plot is due to the method used to reduce the idiosyncratic risk in the portfolio, by replicating exposures 5 times. Numerical results are based on 1.5e6 MC trials and Beta distributed LGD (k=4).

Conclusion
Throughout the previous sections, we have explored the multi factor granularity adjustment. In addition to the original formulation of the adjustment, based on the work by [35], we have adjusted it to both account for stochastic LGD rates and double dependencies of exposures on both a regional and an industry-based systematic factors. By testing the adjustment on the sample portfolio, we deduced that the adjustment systematically overestimates the true portfolio VaR. This indicates that the MF GA overestimates risk in portfolios with a relatively high level of idiosyncratic risk. However, since realistic portfolios tend to contain many more exposures compared to our sample portfolio, we showed that as the portfolio size increases, the MF GA is an accurate approximation to both the portfolio VaR as the VaR contributions in both a deterministic and stochastic LGD framework. Practical implementation of the MF GA is however hindered by its extensive computational needs, making it less of an attractive alternative to the usual MC methods.

6.4. Economic vs Regulatory Concentration Risk

In Section 4.4 we introduced the concepts of Regulatory and Economic Concentration Risk. The difference between these two kinds of risk arises from the different underlying modelling frameworks in which concentration risk can be assessed; single factor versus multi factor models. Regulatory concentration risk is defined to be the difference between the full-blown MC based single factor model and its ASRF equivalent, given by Equation (4.67). Similarly, economic concentration risk is defined as the difference between the full-blown MC based multi factor model and its adjusted ASRF equivalent, given by Equation (4.67). These two forms of concentration risk on our sample portfolio are graphically depicted in Figure 6.17.

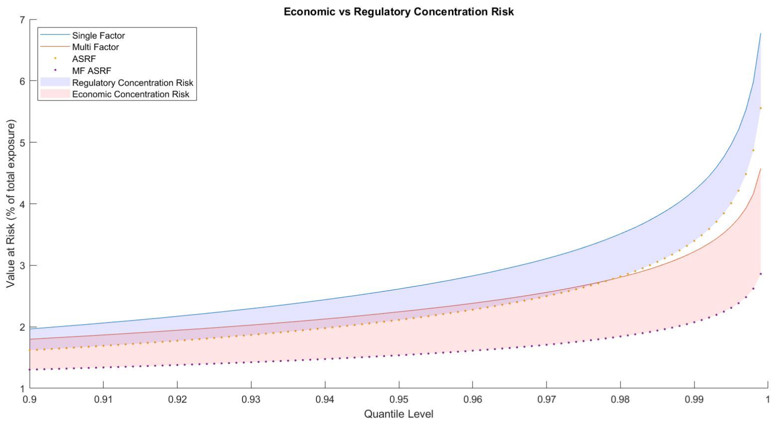

Figure 6.17: Graphical interpretation of the effects of economic concentration risk and regulatory concentration risk on the sample portfolio. Clearly, for high quantile levels the amount of economic concentration risk exceeds the amount of regulatory concentration risk. Numerical results are based on 3e6 MC trials and a deterministic LGD.

As we have concluded before, the portfolio VaR of the single factor framework exceeds that of the multi factor framework, which can be attributed to the diversification effect. Clearly, this does however not imply that regulatory concentration risk also exceeds economic concentration risk. From Figure 6.17 we can conclude that across the higher quantiles, economic concentration risk exceeds regulatory concentration risk. To further stress this observation, we have included Figure 6.18 in which we take a closer look at the highest contributing obligors in terms of EC and concentration risk contributions.

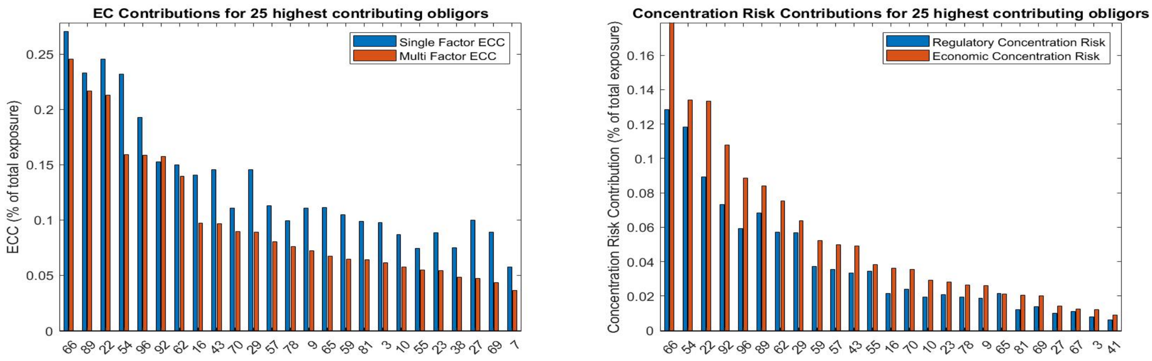

Figure 6.18: Graphical depiction of the highest contributing obligors in terms of EC and concentration risk to total portfolio EC and concentration risk. The contributions are ranked on basis of the highest contributors in the multi factor setting (depicted in red). All results are assessed at the 99.5% quantile, using our sample portfolio.

The left bar graph of Figure 6.18 confirms the observation that the capital requirements are generally higher in the single factor setting compared to the multi factor setting, both for the full portfolio EC as the individual contributions. An important additional observation arises from the fact that the ranking of EC contributions is not preserved, meaning that the single factor model predicts certain exposures to be relatively riskier compared to the multi factor model and vice versa. Generally, the multi factor model is assumed to be a more accurate representation of the world economy by assuming multiple systematic factors. This indicates that the IRB approach, which is based on the single factor model, might assign risk incorrectly.

The right bar graph of Figure 6.18 attributes the concentration risk in the portfolio to the individual exposures, both for regulatory as economic concentration risk. Interestingly, this figure tells a different story compared to the left figure displaying EC contributions. Having already concluded that portfolio economic concentration risk exceeds regulatory concentration risk, we can now conclude that this is also the case at exposure level. So whereas the total risk contribution of exposures (in terms of EC) is higher in the regulatory setting, concentration risk is higher in the economic framework. This conclusion can be attributed to the choice of correlation parameters β_i and c_i. Whereas in the sample portfolio case β_i is equal to ρ_i (2.27), c_i is determined through Equation (4.77). This implies that whereas β_i is only PD dependent, c_i depends on the correlation structure, LGD and EAD. For our sample portfolio this results in higher β_i values compared to the c_i values; the average value for β_i equals 0.1852 whereas the average value for c_i equals 0.0642. A lower value of the correlation parameter (either β_i or c_i) indicates that relatively less of the asset returns can be attributed to systematic factors and more to the idiosyncratic factors, implying more idiosyncratic risk and therefore a higher level of concentration risk.

In a final effort to verify the statements made about RCR and ECR, we perform a similar experiment to that of the justification of the concentration effect. The original sample portfolio is diversified by a factor of 100, with each exposure equalling 1/100$^{\text{th}}$ of its original exposure. In the single factor setting, we have seen that the idiosyncratic risk is diversified for this portfolio and that the true portfolio VaR equals the ASRF approximation of the VaR (compare with Figure 6.3). Therefore, as RCR equals the concentration effect, RCR diminishes by the diversification with a factor 100. When ECR is formulated correctly, we expect a similar effect for the ECR. As the portfolio is diversified, both RCR and ECR should subside. The results of this experiment are displayed in Figure 6.19. The graphical depiction of RCR and ECR verify our understanding, as both have clearly diminished trough the diversification.

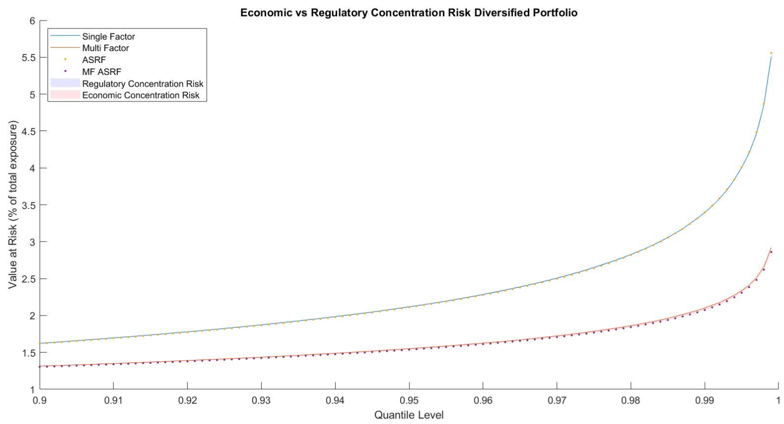

Figure 6.19: Verification of the claims of Economic Concentration Risk by diversifying the portfolio with a factor 100. Both the single factor and multi factor VaR levels are very well approximated by their respective ASRF solutions, indicating the elimination of concentration risk. Numerical results are based on 1e6 MC trials and a deterministic LGD.

In short, the Economic Concentration Risk framework poses a method to asses the levels of idiosyncratic risk in a multi factor framework, a measure that was previously not trivially assessed at ING.

6.5. Sectoral Diversification

In order to assess the benefits diversification, we developed the theory on sectoral diversification in Section 4.5 where we introduced the Diversification Factor (4.85) and the Capital Diversification Index (4.86). Before applying these methods to our sample portfolio, we have to gain a slightly different insight into the distribution of the exposures among the 110 regions and industries of Table A.1. In Figure 6.4 we explored the frequency of occurrence of the different sectors in the sample portfolio, however, this does not yield any insights in the distribution of exposures over among these sectors.

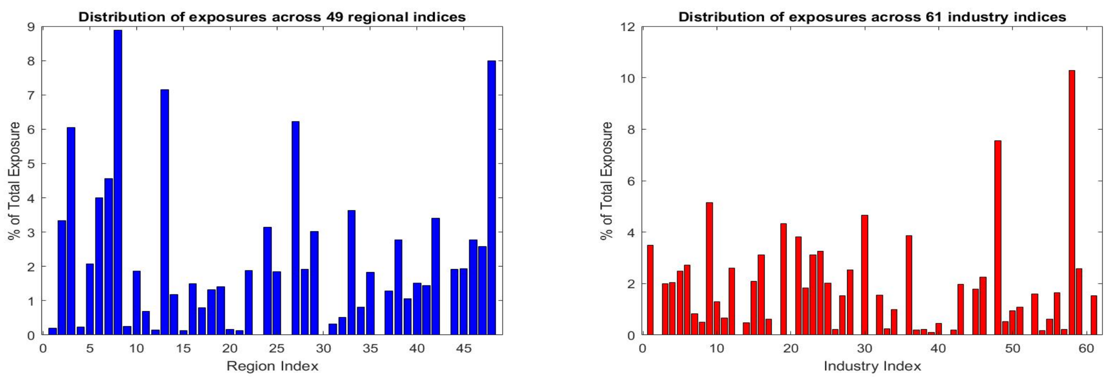

Figure 6.20: Graphical representation of the distribution of exposures among the regions and industries, with indices matching those of A.1.

Whereas from Figure 6.4 it is apparent that the sectors were assigned to the individual obligors uniformly, Figure 6.20 takes the relative exposure of each of the sectors into account with respect to the full portfolio exposure. From this figure, we can conclude that some regions and industries hold a higher share of total portfolio exposure compared to other regions and sectors. This observation is in line with our expectations on realistic portfolios; banks are often focused on certain regions and industries in their lending activities. In order to determine the diversification factor (4.85), the loss distribution of two models has to be determined; the multi factor and the equivalent single factor model. This method essentially matches the procedures performed in Section 6.1.2, since in showing the effects of diversification on portfolio VaR, we essentially compared the multi factor loss distribution with its equivalent single factor loss distribution. Throughout this section, we will focus on the effects of diversification on the individual regions and industries through the use of the diversification factor (4.85). The following experiment will be performed: First, we generate the loss distribution the multi factor and equivalent single factor model using the MC algorithm. Second, for both loss distributions, the portfolio VaR is determined after which the EC risk contributions of the individual exposures are computed. The diversification factor is then easily determined by summing the EC contribution of all exposures assigned to a particular region or industry for both underlying models. Using Equation (4.85), the diversification factor for each region and industry is calculated. The results of this experiment are displayed in Figure 6.21. It should be noted that, by definition, a lower value of the DF indicates a higher degree of diversification benefits. The upside of using the diversification factor is that the underlying models only differ in correlation structure. This means that for both the MF as the E-SF model, the parameters EAD_i, LGD_i, PD_i and β_i are identical, also resulting in an equal expected loss for both models. Therefore, the difference in capital requirements, expressed through the diversification factors, can be attributed to the correlation structure alone. For our sample portfolio, the portfolio diversification factor (4.84) at a quantile level of 99.5% equals 0.6814 which essentially indicates that diversification has an impact of -31.86% on Economic Capital of our sample portfolio. Figure 6.21 assigns the diversification factors of the individual sectors. Clearly, whereas some regions and industries hardly benefit from the effects of diversification (such as region 3 and 19), other sectors benefit to a very high degree from diversification (region 4 for instance). Two effects should be noted when analysing these results. Firstly, because of the relatively small size of our sample portfolio, some sectors are fully defined by a single exposure. This possibly leaves a high degree of idiosyncratic risk impacting the capital requirements for both the MF as the E-SF models. This effect could be offset by instead using the ASRF and A-ASRF models in calculating the DF, which assume all idiosyncratic risk to be diversified away and therefore only addressing the effect of systematic risk. However, for real-life portfolios which are much greater in size, we expect the effects of idiosyncratic risk on the diversification factor to be minimal. Secondly, exposures are assigned to both a region and an industry. Therefore, the diversification benefits are not necessarily fully due to the correlation of the region or industry alone, but by a combination of the two. Ideally, one would like to isolate the effects. A possible suggestion is to run the method twice, once assigning no industries to the exposures, and once by assigning no region to the exposures.

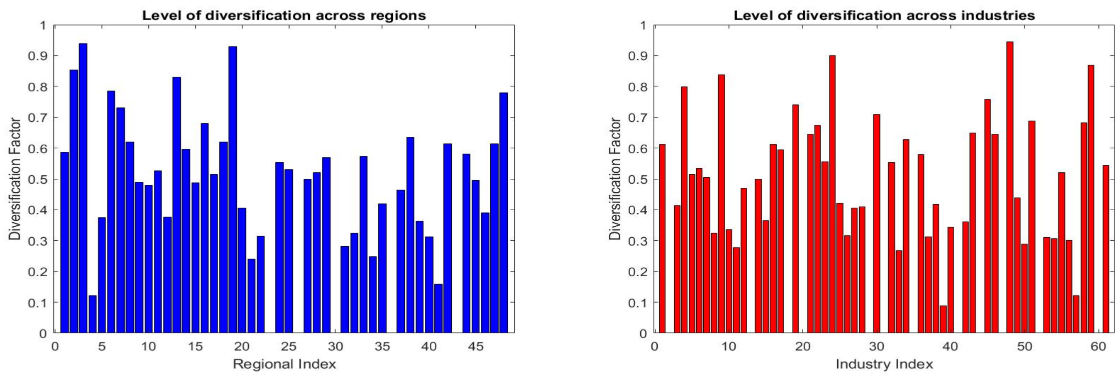

Figure 6.21: Level of diversification among the regions and industries present in our sample portfolio. A lower value of the diversification factor indicates a high level of diversification. When no diversification factor is given, this industry or region is not present in our sample portfolio. Results are based on the 99.5% VaR at 3e6 MC trials and deterministic LGD.

Lastly, we will shortly explore the level of economic capital concentration in the specific sector by applying the CDI (Equation (4.86)) to the sample portfolio at a 99.5% quantile level. The CDI's are displayed in Figure 6.22. As expected, some CDI's equal one which essentially indicates that the specific sector only contains a single obligor. In real life portfolio's, typical values of the CDI are expected to be lower. Low values of the CDI indicated that the economic capital is split fairly among the obligors within the sector, meaning that sector concentration risk is low.

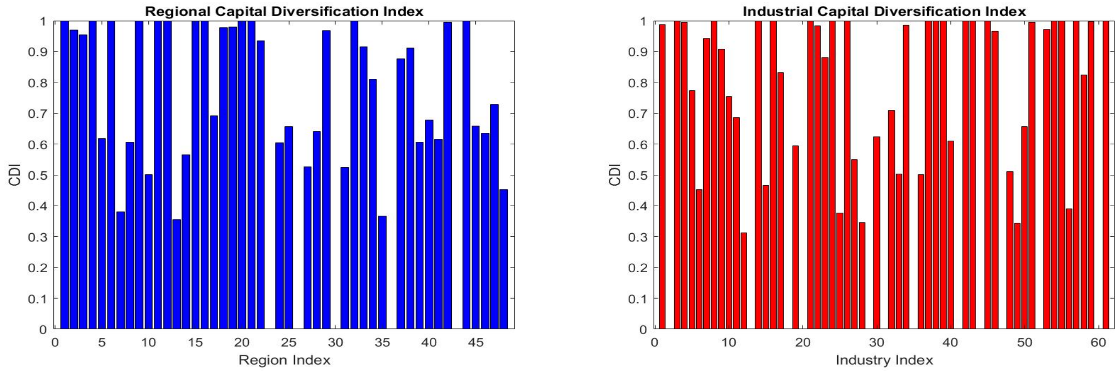

Figure 6.22: Level of capital diversification within the regions and industries present in our sample portfolio. A lower value of the CDI indicates a low level of EC concentration. When no CDI is given, this industry or region is not present in our sample portfolio. Results are based on the 99.5% VaR at 3e6 MC trials and deterministic LGD.

In short, the diversification factor and the capital diversification index can be effective tools for risk management purposes to evaluate the effects of diversification on sector level. It can be applied to investigate which sectors, or even which obligors, benefit to a high degree from diversification and which sectors do not. Additionally, the DF could be used for identifying capital sensitivities, stress testing and possibly for identifying the added benefits or disadvantages of adding certain exposures to existing portfolios.

7 Conclusions

Throughout this work, we have explored a wide range of topics concerning the concept of concentration risk in credit portfolios. The purpose of this thesis is twofold; it should serve both as a study into current methods for assessing concentration risk as an exploration into more advanced methods. This eventually led to considering analytical methods (such as the GA), Monte Carlo based methods and a combination of the two (such as the CDI) for gauging concentration risk. Before considering these methods, this thesis presents a concise overview of the topic of credit risk modelling in general with a strong focus on threshold models, both the Gaussian and *t*-threshold models.

The subject of concentration risk spans two topics; single name concentration risk and sector concentration risk. Single name concentration risk is assessed within the single systematic factor models, analytically through the GA. The GA is derived in two frameworks, the Gaussian threshold and CreditRisk+ (based on the work by [22]) framework. This leads to two different GA's, of which the Gaussian threshold GA outperformed the CR+ GA both in terms of portfolio VaR as for individual risk contributions. Additionally, both methods were tested for their accuracy in accounting for recovery risk. We concluded that although still relatively accurate, the performance of the GA decreases in the presence of recovery risk. The GA has shown to be a powerful tool in regulatory risk management, as an add on to the standard IRB approach built on the ASRF model. It would increase capital requirements to account for idiosyncratic risk, which the standard IRB model is unable to capture due to its inherent assumptions of perfect granularity. However, in practice, multi factor models, relying on numerous correlated systematic factors, are applied which the GA cannot account for. This led us to the topic of sector concentration risk.

To that extend, we introduced a multi factor GA based on the work by [35]. By essentially deriving an adjusted single factor model that mimics the loss distribution of the multi factor model, a multi factor GA (MFGA) is derived. The MF GA has shown to be portfolio dependent in its accuracy, on our small sample portfolio the MF GA overestimated the true VaR substantially. Furthermore, the MF GA in computationally expensive and nearly impossible to solve for larger portfolios without complicated computational techniques and is therefore not deemed to be a suitable alternative to existing MC methods.

In general, analytical methods provide a viable alternative to MC methods, but only on relative small portfolios or in a regulatory setting. Additionally, analytical methods could be used to quickly estimate the add on risk of adding a new loan to an existing portfolio, which would be a demanding task in the MC setting as it involves running the MC algorithm on the full portfolio. However, for more straightforward methods of assessing concentration risk, the MC methods outperform the analytical methods.

Having concluded that in a multi factor setting, the analytical methods do not fully satisfy our needs in terms of accuracy and computational complexity, the methods developed by [35] do present a useful method for assessing concentration risk in a multi factor setting. By applying the methods for developing an adjusted single factor model to mimic the multi factor loss distribution, we have defined the concepts of Regulatory and Economic Concentration Risk, depending in which framework concentration risk is evaluated. Additionally, we derived the diversification factor (DF) and the capital diversification index (CDI) as sector concentration risk indicators and applied both to the sample portfolio. The methods have shown to be present relevant insights into the sample portfolios, such as the disagreement between the economic and regulatory concentration risk values.

Existing literature has repeatedly suggested to evaluate the effects of fatter tailed distributions on concentration risk. In this work, we explored two different models accounting for this effect; the *t*-threshold model and the independent *t*-threshold model. The independent *t*-threshold model allowed for an ASRF and GA solution which converges to its Gaussian counterparts. For the general *t*-threshold model, we explored the possibilities to develop an adjusted single factor *t*-threshold model in the spirit of [35], and we have empirically shown this to be viable. Next, we explored both the effects of diversification and concentration on the *t*-threshold model, and we have shown that these effects are much less apparent within this model compared to its Gaussian counterpart. This makes it an interesting model from a risk management perspective, as the VaR measures are less dependent on uncertainty in parameters such as the covariance matrix describing the linear relationships between the different systematic factors. Lastly, we suggested the independent single factor *t*-threshold model to be a feasible alternative to the current IRB ASRF model, as it allows for an intuitive extra regulatory parameter impacting the capital requirements without any additional computational complexity.

Critical view

All threshold models stem from the early work of Merton [26]. Since then, the advances in terms of modelling have been modest. Throughout this work, we have also mostly evaluated concentration risk within these threshold models, although in multiple forms. This means that all methods presented are inherently limited by the accuracy of the threshold model to compute realistic loss distributions. This shortcoming is not limited to this work but to threshold modelling in general but worth noting.

Throughout this work, we assumed default only models. Basically, this implies that the effects of migration of rating classes have been omitted. However, more advanced methods such as Moody's RiskFrontier and CreditMetrics can take the effects of migration on the portfolio loss distribution into account. Literature on concentration risk in migration models is only limited, but would be worth exploring.

Furthermore, the correlation parameter β_i is assumed to be either given or PD dependent according to IRB correlation Equation (2.27). However, as the correlation parameter is an indirect factor loading to the idiosyncratic risk to which the asset returns are dependent, there is a strong link between this parameter and concentration risk. Through Basel's equations, the correlation parameter is limited to a small range ($[0.12, 0.24]$). In practice, these correlation parameters are not based on the IRB formulas but are modelled, which does not necessarily bound the parameters to the IRB's bounds.

A similar issue arises with the *t*-threshold models studies in this thesis. The IRB correlation parameter is adapted to the IRB ASRF model, which assumes the normally distributed systematic and idiosyncratic factors. In our implementation, we assumed the same equation for the correlation parameters for the Gaussian threshold as the *t*-threshold models, whereas in practice, one would expect different, degrees of freedom dependent, correlation parameters.

Focusing slightly more on the analytical methods, we have concluded that they do not perform accurately across all portfolios and quantile levels. In practice, this means that the analytical method has to be tested against MC methods to verify their accuracy. However, this takes away the upside of the analytical methods being independent on MC methods. This leaves the analytical methods most useful for quickly approximating the risk contributions of added loans to a portfolio.

In this work, plain MC methods are widely applied. However, these methods inherently exhibit some stochastic noise. We have not quantified this level of stochastic noise nor evaluated it. A widely applied method to overcome this shortcoming is the concept of importance sampling. Importance sampling is widely applied in risk management, for instance, in Moody's RiskFrontier simulations. It has also been widely studied, for instance in [7].

Lastly, in this thesis obligors were assumed to depend on a single country and a single industry whereas, in practice, obligors can be active in a set of countries and industries. Whereas many of the developed methods can easily be adapted to this setting, it is not trivial how to assign EC contributions of individual exposures to the sectors on which their asset returns depend. This issue already occurs with the current double dependence setting, as for example the EC contributions of a single country are split over contributions of multiple industries.

Suggestions for further research

Many of the suggested research directions are a direct result of the shortcoming of our approach. This covers, for instance, studying the dependence of concentration risk on the correlation parameter and tweaking the correlation parameters to the *t*-threshold models. Also assigning the risk contribution of exposures among their respective sectors is a relevant topic to pursue.

In addition to the shortcomings, several other topics can be explored. For instance the effect of PD-LGD correlation on concentration risk, or modelling the LGD through a single factor model depending on the same systematic factors as the asset returns. Furthermore, this work limits itself to the VaR as a measure of risk since it has a straightforward link to Economic Capital. However, EC can also be defined with respect to the Expected Shortfall, evaluating the proposed methods for the Expected Shortfall would be of value to this work. Finally, the effects of varying correlation matrices of the systematic factors should be assessed with respect to sectoral diversification and concentration risk.

A

Appendix

A.1. Proof of proposition 2.6.1

In this proof, we will follow the proof of proposition 1 in [9]. In order to prove proposition 2.6.1, we will make use of the following lemmas:

Lemma A.1.1. *Let $\{a_n\}$ be a sequence of positive constraints and $\{Y_n\}$ a sequence of independent random variables. If $a_n \uparrow \infty$ and $\sum_{n=1}^{\infty} \frac{\mathbb{V}[Y_n]}{a_n^2} < \infty$, then for $n \to \infty$*

$$\frac{1}{a_n}\left(\sum_{i=1}^{n} Y_i - \mathbb{E}\left[\sum_{i=1}^{n} Y_i\right]\right) \to 0 \text{ a.s.} \tag{A.1}$$

Lemma A.1.2. *If $\{b_n\}$ is a sequence of positive real numbers such that $\{a_n\}$ is $\mathcal{O}(n^{-\rho})$ for some $\rho > 1$, then*

$$\sum_{n=1}^{\infty} b_n < \infty \tag{A.2}$$

Proof. In order to prove proposition 2.6.1, let $Y_n = EAD_n U_n$ and let $a_n \equiv \sum_{i=1}^{n} EAD_i$. Conditional independence on realization x implies

$$\sum_{n=1}^{\infty} \frac{\mathbb{V}[Y_n|x]}{a_n^2} = \sum_{n=1}^{\infty} \left(\frac{EAD_n}{\sum_{i=1}^{n} EAD_i}\right)^2 \mathbb{V}[U_n|x] \tag{A.3}$$

Clearly, by definition $U_n \in [0,1]$ so $\mathbb{V}[U_n|x] < 1$ for any $X = x$. Therefore, there exists a finite constant V^* such that

$$\sum_{n=1}^{\infty} \frac{\mathbb{V}[Y_n|x]}{a_n^2} \leq V^* \left(\frac{EAD_n}{\sum_{i=1}^{n} EAD_i}\right)^2 \tag{A.4}$$

As we assumed that there exists a $\xi > 0$ such that $\frac{EAD_n}{\sum_{i=1}^{n} EAD_i} = \mathcal{O}(n^{-(\frac{1}{2}+\xi)})$, we have that $\left(\frac{EAD_n}{\sum_{i=1}^{n} EAD_i}\right)^2$ is $\mathcal{O}(n^{-(1+2\xi)})$. By lemma A.1.2, the series sum $\left(\frac{EAD_n}{\sum_{i=1}^{n} EAD_i}\right)^2$ is finite and therefore $\sum_{n=1}^{\infty} \frac{\mathbb{V}[Y_n]}{a_n^2} < \infty$. Furthermore, since we assumed that $\sum_{i=1}^{n} EAD_i \uparrow \infty$ as $n \to \infty$ the conditions of lemma A.1.1 are satisfied. Since the loss ration L_n is equal to $\frac{\sum_{i=1}^{n} Y_i}{a_n}$, proposition 2.6.1 is proved. □

A.2. Analytical expression for the contributing VaR for the single factor Gaussian threshold model

In order to apply equation (4.24), the following expressions have to be derived: $\frac{\partial}{\partial w_i}\left(\frac{\sigma^2(x)}{\mu'(x)}\right)$, $\frac{\partial}{\partial w_i}\left(\frac{\sigma^{2'}(x)}{\mu'(x)}\right)$ and $\frac{\partial}{\partial w_i}\left(\frac{\sigma^2(x)\mu''(x)}{(\mu'(x))^2}\right)$. All derivations follow from straightforward applications of the product rule:

$$\frac{\partial}{\partial w_i}\left(\frac{\sigma^2(x)}{\mu'(x)}\right) = \frac{\mu'(x) - \frac{\partial\sigma^2(x)}{\partial w_i} - \sigma^2(x)\frac{\partial\mu'(x)}{\partial w_i}}{(\mu'(x))^2} \tag{A.5}$$

$$\frac{\partial}{\partial w_i}\left(\frac{\sigma^{2'}(x)}{\mu'(x)}\right) = \frac{\mu'(x) - \frac{\partial\sigma^{2'}(x)}{\partial w_i} - \sigma^{2'}(x)\frac{\partial\mu'(x)}{\partial w_i}}{(\mu'(x))^2} \tag{A.6}$$

$$\frac{\partial}{\partial w_i}\left(\frac{\sigma^2(x)\mu''(x)}{(\mu'(x))^2}\right) = \frac{\partial\sigma^2(x)}{\partial w_i}\frac{\mu''(x)}{(\mu'(x))^2} + \sigma^2(x)\frac{\partial}{\partial w_i}\frac{\mu''(x)}{(\mu'(x))^2} \tag{A.7}$$

Where in the final expression we have:

$$\frac{\partial}{\partial w_i}\frac{\mu''(x)}{(\mu'(x))^2} = \frac{\mu'(x)\frac{\partial\mu''(x)}{\partial w_i} - 2\mu''(x)\frac{\partial\mu'(x)}{\partial w_i}}{(\mu'(x))^3} \tag{A.8}$$

This reduces the problem of deriving the contributing VaR in the single factor setting to determining the following derivatives:

$$\frac{\partial\mu'(x)}{\partial w_i} = LGD_i PD_i'(x) \tag{A.9}$$

$$\frac{\partial\mu''(x)}{\partial w_i} = LGD_i PD_i''(x) \tag{A.10}$$

$$\frac{\partial\sigma^2(x)}{\partial w_i} = 2w_i[(LGD_i^2 + VLGD_i)\cdot PD_i(x) - LGD_i^2\cdot PD_i(x)^2] \tag{A.11}$$

$$\frac{\partial\sigma^{2'}(x)}{\partial w_i} = 2w_i[(LGD_i^2 + VLGD_i)\cdot PD_i'(x) - LGD_i^2\cdot (PD_i(x)^2)'] \tag{A.12}$$

A.3. Geographical location and industry table

Region		Industry	
1	USA/CARIBBEAN	1	AEROSPACE & DEFENSE
2	CANADA	2	AGRICULTURE
3	DENMARK	3	AIR TRANSPORTATION
4	GREECE/SOUTHEAST EUROPE	4	APPAREL & SHOES
5	JAPAN	5	AUTOMOTIVE
6	AUSTRALIA	6	BANKS AND S&LS
7	INDONESIA/SOUTHEAST ASIA	7	BROADCAST MEDIA
8	CHINA	8	BUSINESS PRODUCTS WHSL
9	AUSTRIA	9	BUSINESS SERVICES
10	BELGIUM/LIECHTENSTEIN/LUXEMBOURG	10	CHEMICALS
11	SWITZERLAND	11	COMPUTER HARDWARE
12	GERMANY	12	COMPUTER SOFTWARE
13	SPAIN	13	CONSTRUCTION
14	FRANCE	14	CONSTRUCTION MATERIALS

15	UNITED KINGDOM	15	CONSUMER DURABLES
16	HONG KONG	16	CONSUMER DURABLES RETL/WHSL
17	ITALY	17	CONSUMER PRODUCTS
18	MALAYSIA	18	CONSUMER PRODUCTS RETL/WHSL
19	NETHERLANDS	19	CONSUMER SERVICES
20	SINGAPORE	20	ELECTRICAL EQUIPMENT
21	SWEDEN	21	ELECTRONIC EQUIPMENT
22	THAILAND	22	ENTERTAINMENT & LEISURE
23	SOUTH AFRICA	23	FINANCE COMPANIES
24	FINLAND	24	FINANCE NEC
25	IRELAND	25	FOOD & BEVERAGE
26	NORWAY	26	FOOD & BEVERAGE RETL/WHSL
27	KOREA, REPUBLIC OF	27	FURNITURE & APPLIANCES
28	NEW ZEALAND	28	HOTELS & RESTAURANTS
29	NORTH AFRICA	29	INSURANCE - LIFE
30	CENTRAL AFRICA	30	INSURANCE - PROP/CAS/HEALTH
31	FORMER USSR	31	INVESTMENT MANAGEMENT
32	EAST EUROPE	32	LESSORS
33	PACIFIC OCEAN ISLANDS	33	LUMBER & FORESTRY
34	CENTRAL/SOUTH AMERICA	34	MACHINERY & EQUIPMENT
35	SOUTHERN SOUTH AMERICA	35	MEASURE & TEST EQUIPMENT
36	ISRAEL	36	MEDICAL EQUIPMENT
37	TURKEY	37	MEDICAL SERVICES
38	PHILIPPINES	38	MINING
39	TAIWAN	39	OIL REFINING
40	INDIA	40	OIL, GAS & COAL EXPL/PROD
41	PAKISTAN	41	PAPER
42	PORTUGAL	42	PHARMACEUTICALS
43	SOUTH ASIA	43	PLASTIC & RUBBER
44	POLAND	44	PRINTING
45	MIDDLE EAST	45	PUBLISHING
46	ARGENTINA	46	REAL ESTATE
47	BRAZIL	47	REAL ESTATE INVESTMENT TRUSTS
48	CHILE	48	SECURITY BROKERS & DEALERS
49	MEXICO	49	SEMICONDUCTORS
		50	STEEL & METAL PRODUCTS
		51	TELEPHONE
		52	TEXTILES
		53	TOBACCO
		54	TRANSPORTATION EQUIPMENT
		55	TRANSPORTATION
		56	TRUCKING
		57	MISCELLANEOUS
		58	UTILITIES NEC
		59	UTILITIES, ELECTRIC
		60	UTILITIES, GAS
		61	CABLE TELEVISION

Table A.1: Indices representing the various regions and industries.

Bibliography

[1] Philippe Artzner, Freddy Delbaen, Eber Jean-Marc, and David Heath. Coherent measures of risk. *Mathematical Finance*, 9:203 – 228, 07 1999.

[2] David Bolder. *Credit-Risk Modelling: Theoretical Foundations, Diagnostic Tools, Practical Examples, and Numerical Recipes in Python*. 01 2018. ISBN 978-3-319-94687-0.

[3] Raffaella Calabrese and Francesco Porro. Single-name concentration risk in credit portfolios: a comparison of concentration indices. 01 2012.

[4] Stefano Caselli, Stefano Gatti, and Francesca Querci. The sensitivity of the loss given default rate to systematic risk: New empirical evidence on bank loans. *Journal of Financial Services Research*, 34:1–34, 08 2008.

[5] Juan Carlos Cespedes, Juan Antonio de Juan Herrero, Alex Kreinin, and Dan Rosen. A simple multifactor factor adjustment for the treatment of credit capital diversification. *Journal of Credit Risk*, 2(3):57–85, 2006.

[6] Rongda Chen and Ze Wang. Curve fitting of the corporate recovery rates: The comparison of beta distribution estimation and kernel density estimation. *PloS one*, 8:e68238, 07 2013.

[7] Paul Glasserman. *Monte Carlo methods in financial engineering*. Springer, New York, 2004. ISBN 9781441918222.

[8] Paul Glasserman and Jingyi Li. Importance sampling for portfolio credit risk. *Management Science*, 51:1643–1656, 02 2004.

[9] Michael B. Gordy. A risk-factor model foundation for ratings-based bank capital rules. *Journal of Financial Intermediation*, 12(3):199 – 232, 2003. ISSN 1042-9573.

[10] Christian Gouriéroux, J Laurent, and Olivier Scaillet. Sensitivity analysis of values at risk. *The Journal of Empirical Finance*, 7:225–245, 2000.

[11] ING Group. 2019 annual report ing groep n.v. 2019.

[12] Marc Gürler, Dirk Heithecker, and Martin Hibbeln. Concentration risk under pillar 2: When are credit portfolios infinitely fine grained? *Kredit und Kapital*, 41:79–124, 05 2008.

[13] Martin Hibbeln. *Risk Management in Credit Portfolios*. Physica-Verlag, 2010.

[14] Jimmy Huang, Mariano Lanfranconi, Nihil Patel, and Libor Pospisil. Modeling credit correlations: An overview of moody's analytics gcorr model. *Modeling Methodology from Moody's Analytics*, 12 2012.

[15] Xinzheng Huang, C.W. Oosterlee, and Hans van der Weide. Higher order saddlepoint approximations in the vasicek portfolio credit loss model. *Reports of the Department of Applied Mathematical Analysis*, 11, 01 2006.

[16] Alexandre Kurth and Dirk Tasche. Contributions to credit risk. *Risk*, 16:84–88, 07 2002.

[17] Cornelius Kurtz, Eva Lütkebohmert, and Julian Sester. Calculating capital charges for sector concentration risk. *The Journal of Credit Risk*, 01 2018.

[18] Alvaro Leitao Rodriguez and Luis Ortiz-Gracia. Model-free computation of risk contributions in credit portfolios. *SSRN Electronic Journal*, 01 2018.

[19] Amnon Levy. An overview of modeling credit portfolios. *Modeling Methodology from Moody's Analytics*, 01 2013.

[20] Francois Longin. *The Choice of the Distribution of Asset Returns: How Extreme Value Can Help?: A Handbook of Extreme Value Theory and its Applications*, pages 483–505. 10 2016. ISBN 9781118650196.

[21] Eva Lütkebohmert. *Concentration Risk in Credit Portfolios*. Springer, 2009.

[22] Eva Lütkebohmert and Michael Gordy. Granularity adjustment for regulatory capital assessment. *International Journal of Central Banking*, 9:33–71, 09 2013.

[23] Richard Martin and Tom Wilde. Unsystematic credit risk. *Risk*, 15(11):123–128, 2002.

[24] Alexander McNeil and Rüdiger Frey. Modelling dependent defaults. 06 2001.

[25] Alexander J. McNeil, Rudiger Frey, and Paul Embrechts. *Quantitative Risk Management*. Princeton University Press, 2005.

[26] Robert C. Merton. On the pricing of corporate debt: the risk structure of interest rates. *The Journal of Finance*, 29(2):449–470, 1974.

[27] Agustín Nogales, Paol Perez, and P. Monfort. A monte carlo method to approximate conditional expectations based on a theorem of besicovitch: Application to equivariant estimation of the parameters of the general half-normal distribution. 06 2013.

[28] ECB Task Force of the Market Operation Committee. The use of portfolio credit risk models in central banks. Technical report, 2007. Occasional Paper Series No. 64.

[29] Basel Committee on Banking Supervision. An explanatory note on the basel ii irb risk weight functions. Technical report, 2005. Working paper No. 15.

[30] Basel Committee on Banking Supervision. Studies on credit risk concentration. Technical report, 2006. Working paper No. 15.

[31] Basel Committee on Banking Supervision. International convergence of capital measurement and capital standards. Technical report, 2006.

[32] Basel Committee on Banking Supervision. Supervisory framework for measuring and controlling large exposures. Technical report, 2013.

[33] Kun Park. *Fundamentals of Probability and Stochastic Processes with Applications to Communications*. 01 2018. ISBN 978-3-319-68074-3.

[34] Natalia Puzanova and Klaus Duellmann. Systemic risk contributions: A credit portfolio approach. *Journal of Banking and Finance*, 37, 04 2013.

[35] Michael Pykhtin. Multi-factor adjustment. *Risk*, 17:85–90, 01 2004.

[36] Dan Rosen. Chapter 16 economic credit capital allocation and risk contributions. *Handbooks in Operations Research and Management Science*, 15, 12 2007.

[37] Badreddine Slime. Concentration risk: The comparison of the ad-hoc approach indexes. *Journal of Financial Risk Management*, 05:14, 03 2016.

[38] Badreddine Slime. Credit name concentration risk: Granularity adjustment approximation. *Journal of Financial Risk Management*, 5:246–263, 12 2016.

[39] Dirk Tasche. Capital allocation to business units and sub-portfolios: the euler principle. *Quantitative Finance Papers*, pages 423–453, 01 2008.

[40] Tom Wilde. Creditrisk+: A credit risk management framework. *Credit Suisse First Boston*, 1997.

[41] Tom Wilde. Probing granularity. *Risk*, 14(8):103–106, 2001.